

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 696 773**

51 Int. Cl.:

**C12N 9/02** (2006.01)  
**C12N 9/88** (2006.01)  
**C12P 7/64** (2006.01)  
**C10L 1/02** (2006.01)  
**C12P 5/02** (2006.01)  
**C12P 7/04** (2006.01)  
**C12P 7/24** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **18.05.2009 E 15153942 (6)**

97 Fecha y número de publicación de la concesión europea: **19.09.2018 EP 2927315**

54 Título: **Métodos y composiciones para producir ácidos grasos y alcoholes grasos**

30 Prioridad:

**16.05.2008 US 53955 P**

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

**17.01.2019**

73 Titular/es:

**REG LIFE SCIENCES, LLC (100.0%)  
600 Gateway Boulevard  
South San Francisco, CA 94080, US**

72 Inventor/es:

**SCHIRMER, ANDREAS;  
RUDE, MATHEW y  
BRUBAKER, SHANE**

74 Agente/Representante:

**MILTENYI , Peter**

ES 2 696 773 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

Métodos y composiciones para producir ácidos grasos y alcoholes grasos

**Referencia cruzada a solicitudes relacionadas**

5 Esta solicitud reivindica el beneficio de la solicitud provisional estadounidense n.º 61/053.955, presentada el 16 de mayo de 2008.

**Antecedentes de la invención**

10 El petróleo es un recurso natural y limitado que se encuentra en la Tierra en formas líquida, gaseosa o sólida. El petróleo está compuesto principalmente por hidrocarburos, que están compuestos principalmente por carbono e hidrógeno. También contiene cantidades significativas de otros elementos, tales como nitrógeno, oxígeno o azufre, en diferentes formas.

15 El petróleo es un recurso valioso, pero los productos de petróleo se desarrollan a costes considerables, tanto financieros como ambientales. En primer lugar, deben encontrarse fuentes de petróleo. La exploración de petróleo es una apuesta costosa y arriesgada. El coste de explorar pozos de agua profundos puede superar los 100 millones de dólares. Además, no existe garantía de que estos pozos contengan petróleo. Se estima que sólo el 40% de los pozos perforados conduce a pozos productivos que generan hidrocarburos comerciales. Además del coste económico, la exploración de petróleo conlleva un alto coste ambiental. Por ejemplo, la exploración mar adentro altera los medios marinos circundantes.

20 Después de que se descubra un pozo productivo, el petróleo debe extraerse de la Tierra a un elevado coste. Durante la recuperación primaria, la presión natural subterránea es suficiente para extraer aproximadamente el 20% del petróleo en el pozo. Según cae esta presión natural, se emplean métodos de recuperación secundaria, si fuera económico. Generalmente, la recuperación secundaria implica aumentar la presión del pozo mediante, por ejemplo, inyección de agua, inyección de gas natural o extracción por inyección de gas. Al usar métodos de recuperación secundaria, se recupera de un 5% a un 15% adicional de petróleo. Una vez que se agotan los métodos de recuperación secundaria, pueden usarse métodos de recuperación terciaria, si fuera económico. Los métodos terciarios implican reducir la viscosidad del petróleo para que sea más fácil de extraer. Al usar métodos de recuperación terciaria, se recupera de un 5% a un 15% adicional de petróleo. Por tanto, incluso en las mejores circunstancias, sólo puede extraerse el 50% del petróleo en un pozo. La extracción de petróleo también conlleva un coste ambiental. Por ejemplo, la extracción de petróleo puede dar como resultado grandes filtraciones de petróleo que aflora a la superficie. Además, la perforación mar adentro implica dragar el fondo marino lo que altera o destruye el medio marino circundante.

Puesto que los depósitos de petróleo no se encuentran de manera uniforme en toda la Tierra, el petróleo debe transportarse a lo largo de grandes distancias desde las regiones productoras de petróleo hasta regiones consumidoras de petróleo. Además de los costes de envío, también existe el riesgo ambiental de vertidos de petróleo devastadores.

35 En su forma natural, el petróleo crudo extraído de la Tierra tiene pocos usos comerciales. Es una mezcla de hidrocarburos (por ejemplo, parafinas (o alcanos), olefinas (o alquenos), alquinos, naftenos (o cicloalcanos), compuestos alifáticos, compuestos aromáticos, etc.) de longitud y complejidad variables. Además, el petróleo crudo contiene otros compuestos orgánicos (por ejemplo, compuestos orgánicos que contienen nitrógeno, oxígeno, azufre, etc.) e impurezas (por ejemplo, azufre, sal, ácido, metales, etc.).

40 Por tanto, el petróleo crudo debe refinarse y purificarse antes de que pueda usarse comercialmente. Debido a su elevada densidad energética y a su fácil transportabilidad, la mayor parte del petróleo se refina dando combustibles, tales como combustibles para transporte (por ejemplo, gasolina, diésel, combustible para aviación, etc.), aceite para calefacción, gas de petróleo licuado, etc.

45 El petróleo crudo también es una fuente principal de materiales de partida para producir productos petroquímicos. Las dos clases principales de materiales de partida derivados de petróleo son olefinas de cadena corta (por ejemplo, etileno y propileno) y compuestos aromáticos (por ejemplo, isómeros de benceno y xileno). Estos materiales de partida se derivan de hidrocarburos de cadena más larga en petróleo crudo craqueándolo a un coste considerable usando una variedad de métodos, tales como craqueo catalítico, craqueo a vapor o reformado catalítico. Estos materiales de partida se usan para elaborar productos petroquímicos, que no pueden refinarse directamente a partir de petróleo crudo, tal como monómeros, disolventes, detergentes o adhesivos.

55 Un ejemplo de un material de partida derivado de petróleo crudo es etileno. El etileno se usa para producir productos petroquímicos tales como, polietileno, etanol, óxido de etileno, etilenglicol, poliéster, glicol éter, etoxilato, acetato de vinilo, 1,2-dicloroetano, tricloroetileno, tetracloroetileno, cloruro de vinilo y poli(cloruro de vinilo). Un ejemplo adicional de un material de partida es propileno, que se usa para producir alcohol isopropílico, acrilonitrilo, polipropileno, óxido de propileno, propilenglicol, glicol éteres, butileno, isobutileno, 1,3-butadieno, elastómeros sintéticos, poliolefinas, alfa-olefinas, alcoholes grasos, ácido acrílico, polímeros acrílicos, cloruro de alilo, epiclohidrina y resinas

epoxídicas.

Estos productos petroquímicos pueden usarse entonces para elaborar productos químicos especializados, tales como plásticos, resinas, fibras, elastómeros, productos farmacéuticos, lubricantes o geles. Productos químicos especializados particulares que pueden producirse a partir de materiales de partida petroquímicos son: ácidos grasos, hidrocarburos (por ejemplo, de cadena larga, cadena ramificada, saturados, insaturados, etc.), alcoholes grasos, ésteres, aldehídos grasos, cetonas, lubricantes, etc.

Los productos químicos especializados tienen muchos usos comerciales. Se usan ácidos grasos comercialmente como tensioactivos, por ejemplo, en detergentes y jabones. También pueden usarse como aditivos en combustibles, aceites lubricantes, pinturas, lacas, velas, aceite para ensalada, manteca, cosméticos y emulsionantes. Además, se usan ácidos grasos como activadores del acelerador en productos de caucho. Pueden usarse ácidos grasos como materia prima para producir ésteres metílicos, amidas, aminas, cloruros de ácido, anhídridos, dímeros de cetena, y ácidos peroxi y ésteres.

Los hidrocarburos tienen muchos usos comerciales. Por ejemplo, se usan alcanos de cadena más corta como combustibles. Metano y etano son los constituyentes principales del gas natural. Se usan alcanos de cadena más larga (por ejemplo, desde cinco hasta dieciséis carbonos) como combustibles para transporte (por ejemplo, gasolina, diésel o combustible para aviación). Los alcanos que tienen más de dieciséis átomos de carbono son componentes importantes de aceites combustibles y aceites lubricantes. Incluso alcanos más largos, que son sólidos a temperatura ambiente, pueden usarse, por ejemplo, como cera de parafina. Los alcanos que contienen aproximadamente treinta y cinco carbonos se encuentran en betún, que se usa para revestimiento de carreteras. Además, los alcanos de cadena más larga pueden craquearse para producir hidrocarburos de cadena más corta útiles comercialmente.

Como los alcanos de cadena corta, se usan alquenos de cadena corta en combustibles para transporte. Se usan alquenos de cadena más larga en plásticos, lubricantes y lubricantes sintéticos. Además, se usan alquenos como materia prima para producir alcoholes, ésteres, plastificantes, tensioactivos, aminas terciarias, aceites de recuperación de aceite potenciada, ácidos grasos, tioles, anhídridos alquenilsuccínicos, epóxidos, alcanos clorados, alquenos clorados, ceras, aditivos de combustible y reductores de flujo de arrastre.

Los alcoholes grasos tienen muchos usos comerciales. Los alcoholes grasos de cadena más corta se usan en las industrias cosmética y alimentaria como emulsionantes, emolientes y espesantes. Debido a su naturaleza anfifílica, los alcoholes grasos se comportan como tensioactivos no iónicos, que son útiles como detergentes. Además, se usan alcoholes grasos en ceras, gomas, resinas, pomadas y lociones farmacéuticas, aditivos de aceites lubricantes, agentes textiles antiestáticos y de acabado, plastificantes, cosméticos, disolventes industriales y disolventes para grasas.

Los ésteres tienen muchos usos comerciales. Por ejemplo, el biodiésel, un combustible alternativo, está comprendido por ésteres (por ejemplo, éster metílico de ácido graso, ésteres etílicos de ácido graso, etc.). Algunos ésteres de bajo peso molecular son volátiles con un olor agradable que los hace útiles como fragancias o agentes saborizantes. Además, se usan ésteres como disolventes para lacas, pinturas y barnices. Además, algunas sustancias que se producen de manera natural, tales como ceras, grasas y aceites están comprendidas por ésteres. Los ésteres también se usan como agentes suavizantes en resinas y plásticos, plastificantes, retardantes de llama y aditivos en gasolina y aceite. Además, los ésteres pueden usarse en la fabricación de polímeros, películas, textiles, colorantes y productos farmacéuticos.

Se usan aldehídos para producir muchos productos químicos especializados. Por ejemplo, se usan aldehídos para producir polímeros, resinas (por ejemplo, baquelita), colorantes, saborizantes, plastificantes, perfumes, productos farmacéuticos y otros productos químicos. Algunos se usan como disolventes, conservantes o desinfectantes. Algunos compuestos naturales y sintéticos, tales como vitaminas y hormonas, son aldehídos. Además, muchos azúcares contienen grupos aldehído.

Se usan comercialmente cetonas como disolventes. Por ejemplo, se usa frecuentemente acetona como disolvente, pero también es un material de partida para elaborar polímeros. También se usan cetonas en lacas, pinturas, explosivos, perfumes y procesamiento de textiles. Además, se usan cetonas para producir alcoholes, alquenos, alcanos, iminas y enaminas.

Además, el petróleo crudo es una fuente de lubricantes. Los lubricantes derivados de petróleo están compuestos normalmente por olefinas, particularmente poliolefinas y alfa-olefinas. Los lubricantes pueden o bien refinarse a partir de petróleo crudo o bien fabricarse usando materiales de partida refinados a partir de petróleo crudo.

Obtener estos productos químicos especializados a partir de petróleo crudo requiere una inversión financiera significativa así como una gran cantidad de energía. También es un procedimiento ineficaz porque, con frecuencia, los hidrocarburos de cadena larga en petróleo crudo se craquean para producir monómeros más pequeños. Estos monómeros se usan entonces como material de partida para fabricar los productos químicos especializados más complejos.

Además de los problemas con la exploración, la extracción, el transporte y el refinamiento de petróleo, el petróleo es un recurso limitado y cada vez más escaso. Una estimación del consumo de petróleo mundial es de 30 miles de millones de barriles al año. Según algunas estimaciones, se prevé que a los niveles actuales de producción, las reservas mundiales de petróleo podrían agotarse antes del año 2050.

5 Finalmente, la quema de combustibles a base de petróleo libera gases de efecto invernadero (por ejemplo, dióxido de carbono) y otras formas de contaminación atmosférica (por ejemplo, monóxido de carbono, dióxido de azufre, etc.). Según aumenta la demanda mundial de combustibles, la emisión de gases de efecto invernadero y otras formas de contaminación atmosférica también aumenta. La acumulación de gases de efecto invernadero en la atmósfera conduce a un aumento del calentamiento global. Por tanto, además de dañar el medio ambiente a nivel local (por ejemplo, vertidos de petróleo, dragado de medios marinos, etc.), quemar petróleo también daña el medio ambiente a nivel global.

Debido a los desafíos inherentes planteados por el petróleo, existe la necesidad de una fuente de petróleo renovable que no requiere exploración, extracción y transporte a lo largo de largas distancias, o sustancialmente refinada como el petróleo. También existe la necesidad de una fuente de petróleo renovable que pueda producirse de manera económica y sin crear el tipo de daño ambiental producido por la industria del petróleo y la quema de combustibles a base de petróleo. Por motivos similares, también existe la necesidad de una fuente renovable de productos químicos que se derivan normalmente de petróleo.

La entrada de base de datos UnitProt Q54795 se refiere a SEQ ID NO: 66; la anotación automática lo identifica como una oxidorreductasa no especificada adicionalmente.

20 La entrada de base de datos EMBL U59236 se refiere a SEQ ID NO: 65; la entrada dejó la secuencia sin anotar.

Los documentos WO 2007/136762 A2 y WO 2008/119082 A2 describen la producción de alcoholes grasos mediante expresión recombinante de un gen de tioesterasa y un gen de acil-CoA reductasa en un huésped de producción de *E. coli*.

#### Sumario de la invención

25 La invención se basa, al menos en parte, en la identificación de genes cianobacterianos que codifican para polipéptidos biosintéticos de hidrocarburos. Por consiguiente, la presente invención proporciona lo siguiente:

1. Una célula huésped no humana para su uso en la producción de un aldehído graso o un alcohol graso, dicha célula se modifica por ingeniería genética para expresar un polinucleótido que comprende una secuencia de ácido nucleico que tiene al menos el 70% de identidad de secuencia con SEQ ID NO: 65 ó 69, en la que la secuencia de ácido nucleico codifica para un polipéptido que tiene actividad acil-ACP o acil-CoA reductasa y que es de una cianobacteria.
2. Una célula no humana según el punto 1, en la que el polinucleótido está comprendido por un vector recombinante.
3. Una célula no humana según el punto 1, en la que el polinucleótido se incorpora de manera estable en el ADN genómico de la célula huésped.
4. Una célula no humana según uno cualquiera de los puntos 1-3, en la que la célula huésped es una célula microbiana.
5. Una célula no humana según uno cualquiera de los puntos 1-4, en la que la célula huésped es una célula bacteriana.
- 40 6. Una célula no humana según uno cualquiera de los puntos 1-5, en la que la célula huésped es una célula de *Escherichia* o una célula cianobacteriana.

Además se da a conocer en el presente documento un método de producción de un aldehído, comprendiendo el método producir en una célula huésped un polipéptido que comprende la secuencia de aminoácidos de SEQ ID NO: 66, 68, 70, 72, 74, 76, 78, 80 u 82, o una variante del mismo, y aislar el aldehído de la célula huésped.

45 Además se da a conocer en el presente documento un polipéptido que comprende una secuencia de aminoácidos que tiene al menos aproximadamente el 70%, al menos aproximadamente el 80%, al menos aproximadamente el 85%, al menos aproximadamente el 90%, al menos aproximadamente el 91%, al menos aproximadamente el 92%, al menos aproximadamente el 93%, al menos aproximadamente el 94%, al menos aproximadamente el 95%, al menos aproximadamente el 96%, al menos aproximadamente el 97%, al menos aproximadamente el 98%, o al menos aproximadamente el 99% de identidad con SEQ ID NO: 66, 68, 70, 72, 74, 76, 78, 80 u 82.

Además se da a conocer en el presente documento un polipéptido que comprende la secuencia de aminoácidos de SEQ ID NO: 66, 68, 70, 72, 74, 76, 78, 80 u 82 con una o más sustituciones, adiciones, inserciones o deleciones de aminoácido. En algunas realizaciones, el polipéptido tiene actividad reductasa. Además se da a conocer en el

presente documento un polipéptido que comprende la secuencia de aminoácidos de SEQ ID NO: 66, 68, 70, 72, 74, 76, 78, 80 u 82, con una o más sustituciones de aminoácido conservadoras. Por ejemplo, el polipéptido comprende una o más de las siguientes sustituciones de aminoácido conservadoras: reemplazo de un aminoácido alifático, tal como alanina, valina, leucina e isoleucina, por otro aminoácido alifático; reemplazo de una serina por una treonina; reemplazo de una treonina por una serina; reemplazo de un residuo ácido, tal como ácido aspártico y ácido glutámico, por otro residuo ácido; reemplazo de un residuo que porta un grupo amida, tal como asparagina y glutamina, por otro residuo que porta un grupo amida; intercambio de un residuo básico, tal como lisina y arginina, con otro residuo básico; y reemplazo de un residuo aromático, tal como fenilalanina y tirosina, por otro residuo aromático. El polipéptido puede tener aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100, o más sustituciones, adiciones, inserciones o deleciones de aminoácido. En algunas realizaciones, el polipéptido tiene actividad reductasa.

Además se da a conocer en el presente documento un polipéptido que comprende la secuencia de aminoácidos de SEQ ID NO: 54, 55, 56, 57, 58, 59, 60, 61, 62, 63 ó 64. En determinadas realizaciones, el polipéptido tiene actividad reductasa.

Además se da a conocer en el presente documento un método de producción de un aldehído, comprendiendo el método expresar en una célula huésped un polinucleótido que comprende una secuencia de nucleótidos que tiene al menos aproximadamente el 70%, al menos aproximadamente el 75%, al menos aproximadamente el 80%, al menos aproximadamente el 85%, al menos aproximadamente el 90%, al menos aproximadamente el 91 %, al menos aproximadamente el 92%, al menos aproximadamente el 93%, al menos aproximadamente el 94%, al menos aproximadamente el 95%, al menos aproximadamente el 96%, al menos aproximadamente el 97%, al menos aproximadamente el 98%, o al menos aproximadamente el 99% de identidad de secuencia con SEQ ID NO: 65, 67, 69, 71, 73, 75, 77, 79 u 81. En algunas realizaciones, la secuencia de nucleótidos es SEQ ID NO: 65, 67, 69, 71, 73, 75, 77, 79 u 81. En algunas realizaciones, el método comprende además aislar el aldehído de la célula huésped.

Además se da a conocer en el presente documento una secuencia de nucleótidos que se hibrida con un complemento de SEQ ID NO: 65, 67, 69, 71, 73, 75, 77, 79 u 81, o con un fragmento del mismo, por ejemplo, en condiciones de baja rigurosidad, media rigurosidad, alta rigurosidad o muy alta rigurosidad.

Además se da a conocer en el presente documento una secuencia de nucleótidos que codifica para un polipéptido que comprende: (i) la secuencia de aminoácidos de SEQ ID NO: 66, 68, 70, 72, 74, 76, 78, 80 u 82; o (ii) la secuencia de aminoácidos de SEQ ID NO: 66, 68, 70, 72, 74, 76, 78, 80 u 82 con una o más sustituciones, adiciones, inserciones o deleciones de aminoácido. En algunas realizaciones, el polipéptido comprende la secuencia de aminoácidos de SEQ ID NO: 66, 68, 70, 72, 74, 76, 78, 80 u 82 con una o más sustituciones de aminoácido conservadoras. En algunas realizaciones, el polipéptido tiene actividad reductasa.

Además se da a conocer en el presente documento una secuencia de nucleótidos que codifica para un polipéptido que tiene la misma actividad biológica que un polipéptido que comprende la secuencia de aminoácidos de SEQ ID NO: 66, 68, 70, 72, 74, 76, 78, 80 u 82. En algunas realizaciones, la secuencia de nucleótidos es SEQ ID NO: 65, 67, 69, 71, 73, 75, 77, 79 u 81 o un fragmento de la misma. Además se da a conocer en el presente documento una secuencia de nucleótidos que se hibrida con un complemento de SEQ ID NO: 65, 67, 69, 71, 73, 75, 77, 79 u 81 o con un fragmento del mismo, por ejemplo, en condiciones de baja rigurosidad, media rigurosidad, alta rigurosidad o muy alta rigurosidad. En algunas realizaciones, la actividad biológica es actividad reductasa.

Además se da a conocer en el presente documento un método que comprende transformar una célula huésped con un vector recombinante que comprende una secuencia de nucleótidos que tiene al menos aproximadamente el 70%, al menos aproximadamente el 75%, al menos aproximadamente el 80%, al menos aproximadamente el 85%, al menos aproximadamente el 90%, al menos aproximadamente el 91%, al menos aproximadamente el 92%, al menos aproximadamente el 93%, al menos aproximadamente el 94%, al menos aproximadamente el 95%, al menos aproximadamente el 96%, al menos aproximadamente el 97%, al menos aproximadamente el 98%, o al menos aproximadamente el 99% de identidad de secuencia con SEQ ID NO: 65, 67, 69, 71, 73, 75, 77, 79 u 81. En algunas realizaciones, el vector recombinante comprende además un promotor unido operativamente a la secuencia de nucleótidos. En algunas realizaciones, el promotor es un promotor regulado por desarrollo, específico de orgánulos, específico de tejidos, inducible, constitutivo o específico de células. En realizaciones particulares, el vector recombinante comprende al menos una secuencia seleccionada del grupo que consiste en (a) una secuencia reguladora acoplada operativamente a la secuencia de nucleótidos; (b) un marcador de selección acoplado operativamente a la secuencia de nucleótidos; (c) una secuencia de marcador acoplada operativamente a la secuencia de nucleótidos; (d) un resto de purificación acoplado operativamente a la secuencia de nucleótidos; (e) una secuencia de secreción acoplada operativamente a la secuencia de nucleótidos; y (f) una secuencia de direccionamiento acoplada operativamente a la secuencia de nucleótidos. En determinadas realizaciones, la secuencia de nucleótidos se incorpora de manera estable en el ADN genómico de la célula huésped, y la expresión de la secuencia de nucleótidos está bajo el control de una región de promotor regulada.

En cualquiera de los aspectos descritos en el presente documento, la célula huésped puede seleccionarse del grupo que consiste en una célula de mamífero, célula vegetal, célula de insecto, célula de levadura, célula de hongo, célula de hongos filamentosos y célula bacteriana.

En algunas realizaciones, la célula huésped es una célula bacteriana Gram positiva. En otras realizaciones, la célula huésped es una célula bacteriana Gram negativa.

5 En algunas realizaciones, la célula huésped se selecciona del género *Escherichia*, *Bacillus*, *Lactobacillus*, *Rhodococcus*, *Pseudomonas*, *Aspergillus*, *Trichoderma*, *Neurospora*, *Fusarium*, *Humicola*, *Rhizomucor*, *Kluyveromyces*, *Pichia*, *Mucor*, *Myceliophthora*, *Penicillium*, *Phanerochaete*, *Pleurotus*, *Trametes*, *Chrysosporium*, *Saccharomyces*, *Stenotrophomonas*, *Schizosaccharomyces*, *Yarrowia* o *Streptomyces*.

10 En realizaciones particulares, la célula huésped es una célula de *Bacillus lentus*, una célula de *Bacillus brevis*, una célula de *Bacillus stearothermophilus*, una célula de *Bacillus licheniformis*, una célula de *Bacillus alkalophilus*, una célula de *Bacillus coagulans*, una célula de *Bacillus circulans*, una célula de *Bacillus pumilis*, una célula de *Bacillus thuringiensis*, una célula de *Bacillus clausii*, una célula de *Bacillus megaterium*, una célula de *Bacillus subtilis* o una célula de *Bacillus amyloliquefaciens*.

15 En otras realizaciones, la célula huésped es una célula de *Trichoderma koningii*, una célula de *Trichoderma viride*, una célula de *Trichoderma reesei*, una célula de *Trichoderma longibrachiatum*, una célula de *Aspergillus awamori*, una célula de *Aspergillus fumigates*, una célula de *Aspergillus foetidus*, una célula de *Aspergillus nidulans*, una célula de *Aspergillus niger*, una célula de *Aspergillus oryzae*, una célula de *Humicola insolens*, una célula de *Humicola lanuginosa*, una célula de *Rhodococcus opacus*, una célula de *Rhizomucor miehei* o una célula de *Mucor michei*.

En aún otras realizaciones, la célula huésped es una célula de *Streptomyces lividans* o una célula de *Streptomyces murinus*. En otras realizaciones, la célula huésped es una célula de *Actinomycetes*.

20 En algunas realizaciones, la célula huésped es una célula CHO, una célula COS, una célula VERO, una célula BHK, una célula HeLa, una célula Cv1, una célula MDCK, una célula 293, una célula 3T3 o una célula PC12.

En realizaciones particulares, la célula huésped es una célula *E. coli*, tal como una cepa B, una cepa C, una cepa K o una cepa W de una célula de *E. coli*.

25 En otras realizaciones, la célula huésped es una célula huésped cianobacteriana. En realizaciones particulares, la célula huésped cianobacteriana es una célula enumerada en la tabla 1.

En algunas realizaciones, el aldehído se secreta por la célula huésped.

30 En determinadas realizaciones, la célula huésped sobreexpresa un sustrato descrito en el presente documento. El método puede incluir además transformar la célula huésped con un ácido nucleico que codifica para una enzima descrita en el presente documento, y la célula huésped sobreexpresa un sustrato descrito en el presente documento. El método puede incluir además cultivar la célula huésped en presencia de al menos un sustrato descrito en el presente documento. El sustrato puede ser un derivado de ácido graso, una acil-ACP, un ácido graso, una acil-CoA, un aldehído graso, un alcohol graso o un éster graso.

35 El sustrato de derivado de ácido graso es un sustrato de derivado de ácido graso insaturado, un sustrato de derivado de ácido graso monoinsaturado o un sustrato de derivado de ácido graso saturado. El sustrato de derivado de ácido graso puede ser un sustrato de derivado de ácido graso de cadena lineal, un sustrato de derivado de ácido graso de cadena ramificada o un sustrato de derivado de ácido graso que incluye un resto cíclico.

40 El derivado de ácido graso puede ser un derivado de ácido graso C<sub>3</sub>-C<sub>25</sub>. Por ejemplo, el derivado de ácido graso es un derivado de ácido graso C<sub>3</sub>, C<sub>4</sub>, C<sub>5</sub>, C<sub>6</sub>, C<sub>7</sub>, C<sub>8</sub>, C<sub>9</sub>, C<sub>10</sub>, C<sub>11</sub>, C<sub>12</sub>, C<sub>13</sub>, C<sub>14</sub>, C<sub>15</sub>, C<sub>16</sub>, C<sub>17</sub>, C<sub>18</sub>, C<sub>19</sub>, C<sub>20</sub>, C<sub>21</sub>, C<sub>22</sub>, C<sub>23</sub>, C<sub>24</sub> o C<sub>25</sub>. En particular, el sustrato de derivado de ácido graso es tetradecanoil-ACP, hexadecanoil-ACP, hexadecenoil-ACP u octadecenoil-ACP.

45 En determinados aspectos descritos en el presente documento, el aldehído es un aldehído C<sub>3</sub>-C<sub>25</sub>. Por ejemplo, el aldehído es un aldehído C<sub>3</sub>, C<sub>4</sub>, C<sub>5</sub>, C<sub>6</sub>, C<sub>7</sub>, C<sub>8</sub>, C<sub>9</sub>, C<sub>10</sub>, C<sub>11</sub>, C<sub>12</sub>, C<sub>13</sub>, C<sub>14</sub>, C<sub>15</sub>, C<sub>16</sub>, C<sub>17</sub>, C<sub>18</sub>, C<sub>19</sub>, C<sub>20</sub>, C<sub>21</sub>, C<sub>22</sub>, C<sub>23</sub>, C<sub>24</sub> o C<sub>25</sub>. En algunas realizaciones, el aldehído es tetradecanal, hexadecanal, hexadecenal, octadecanal, octadecenal, metiltetradecanal, metiltetradecenal, metilhexadecanal, metilhexadecenal, metiloctadecanal o metilooctadecenal.

El aldehído puede ser un aldehído de cadena lineal, un aldehído de cadena ramificada o un aldehído cíclico.

El método puede incluir además aislar el aldehído de la célula huésped o del medio de cultivo.

50 Además se da a conocer en el presente documento un microorganismo modificado por ingeniería genética que comprende una secuencia de control exógena incorporada de manera estable en el ADN genómico del microorganismo. En una realización, la secuencia de control se integra aguas arriba de un polinucleótido que comprende una secuencia de nucleótidos que tiene al menos aproximadamente el 70% de identidad de secuencia con SEQ ID NO: 65, 67, 69, 71, 73, 75, 77, 79 u 81. En algunas realizaciones, la secuencia de nucleótidos tiene al menos aproximadamente el 75%, al menos aproximadamente el 80%, al menos aproximadamente el 85%, al menos aproximadamente el 90%, al menos aproximadamente el 91%, al menos aproximadamente el 92%, al menos

aproximadamente el 93%, al menos aproximadamente el 94%, al menos aproximadamente el 95%, al menos aproximadamente el 96%, al menos aproximadamente el 97%, al menos aproximadamente el 98%, o al menos aproximadamente el 99% de identidad de secuencia con SEQ ID NO: 65, 67, 69, 71, 73, 75, 77, 79 u 81. En algunas realizaciones, la secuencia de nucleótidos es SEQ ID NO: 65, 67, 69, 71, 73, 75, 77, 79 u 81.

- 5 El polinucleótido puede ser endógeno al microorganismo. El microorganismo puede expresar un nivel aumentado de un aldehído en relación con un microorganismo silvestre. El microorganismo puede ser una cianobacteria.

Además se da a conocer en el presente documento un método de elaboración de un aldehído, comprendiendo el método cultivar un microorganismo modificado por ingeniería genética descrito en el presente documento en condiciones adecuadas para expresión génica, y aislar el aldehído.

- 10 Además se da a conocer en el presente documento un método de elaboración de un aldehído, que comprende poner en contacto un sustrato con (i) un polipéptido que tiene al menos el 70% de identidad con la secuencia de aminoácidos de SEQ ID NO: 66, 68, 70, 72, 74, 76, 78, 80 u 82, o una variante del mismo; (ii) un polipéptido codificado por una secuencia de nucleótidos que tiene al menos el 70% de identidad con SEQ ID NO: 65, 67, 69, 71, 73, 75, 77, 79 u 81, o una variante del mismo; o (iii) un polipéptido que comprende la secuencia de aminoácidos de SEQ ID NO: 54, 55, 56, 57, 58, 59, 60, 61, 62, 63 ó 64. En algunas realizaciones, el polipéptido tiene actividad reductasa.

- 20 Además se da a conocer en el presente documento un polipéptido que tiene al menos aproximadamente el 80%, al menos aproximadamente el 85%, al menos aproximadamente el 90%, al menos aproximadamente el 91%, al menos aproximadamente el 92%, al menos aproximadamente el 93%, al menos aproximadamente el 94%, al menos aproximadamente el 95%, al menos aproximadamente el 96%, al menos aproximadamente el 97%, al menos aproximadamente el 98%, o al menos aproximadamente el 99% de identidad con SEQ ID NO: 66, 68, 70, 72, 74, 76, 78, 80 u 82. En algunas realizaciones, el polipéptido has la secuencia de aminoácidos de SEQ ID NO: 66, 68, 70, 72, 74, 76, 78, 80 u 82.

- 25 Además se da a conocer en el presente documento un polipéptido que está codificado por una secuencia de nucleótidos que tiene al menos aproximadamente el 75%, al menos aproximadamente el 80%, al menos aproximadamente el 85%, al menos aproximadamente el 90%, al menos aproximadamente el 91%, al menos aproximadamente el 92%, al menos aproximadamente el 93%, al menos aproximadamente el 94%, al menos aproximadamente el 95%, al menos aproximadamente el 96%, al menos aproximadamente el 97%, al menos aproximadamente el 98%, o al menos aproximadamente el 99% de identidad de secuencia con SEQ ID NO: 65, 67, 69, 71, 73, 75, 77, 79 u 81. En algunas realizaciones, el polipéptido está codificado por una secuencia de nucleótidos que tiene SEQ ID NO: 65, 67, 69, 71, 73, 75, 77, 79 u 81.

El sustrato biológico puede ser un derivado de ácido graso, una acil-ACP, un ácido graso, una acil-CoA, un aldehído graso, un alcohol graso o un éster graso.

- 35 El derivado de ácido graso puede ser un derivado de ácido graso C<sub>3</sub>-C<sub>25</sub>. Por ejemplo, el derivado de ácido graso es un derivado de ácido graso C<sub>3</sub>, C<sub>4</sub>, C<sub>5</sub>, C<sub>6</sub>, C<sub>7</sub>, C<sub>8</sub>, C<sub>9</sub>, C<sub>10</sub>, C<sub>11</sub>, C<sub>12</sub>, C<sub>13</sub>, C<sub>14</sub>, C<sub>15</sub>, C<sub>16</sub>, C<sub>17</sub>, C<sub>18</sub>, C<sub>19</sub>, C<sub>20</sub>, C<sub>21</sub>, C<sub>22</sub>, C<sub>23</sub>, C<sub>24</sub> o C<sub>25</sub>. En particular, el sustrato de derivado de ácido graso es tetradecanoil-ACP, hexadecanoil-ACP, hexadecenoil-ACP u octadecenoil-ACP.

- 40 El aldehído puede ser un aldehído C<sub>3</sub>-C<sub>25</sub>. Por ejemplo, el aldehído es un aldehído C<sub>3</sub>, C<sub>4</sub>, C<sub>5</sub>, C<sub>6</sub>, C<sub>7</sub>, C<sub>8</sub>, C<sub>9</sub>, C<sub>10</sub>, C<sub>11</sub>, C<sub>12</sub>, C<sub>13</sub>, C<sub>14</sub>, C<sub>15</sub>, C<sub>16</sub>, C<sub>17</sub>, C<sub>18</sub>, C<sub>19</sub>, C<sub>20</sub>, C<sub>21</sub>, C<sub>22</sub>, C<sub>23</sub>, C<sub>24</sub> o C<sub>25</sub>. Se da a conocer en el presente documento un aldehído que es tetradecanal, hexadecanal, hexadecenal, octadecanal, octadecenal, metiltetradecanal, metiltetradecenal, metilhexadecanal, metilhexadecenal, metiloctadecanal o metiloctadecenal.

El aldehído puede ser un aldehído de cadena lineal, un aldehído de cadena ramificada o un aldehído cíclico.

- 45 Además se da a conocer en el presente documento un aldehído producido mediante cualquiera de los métodos o microorganismos descritos en el presente documento. El aldehído puede tener un  $\delta^{13}\text{C}$  de aproximadamente -15,4 o mayor. Por ejemplo, el aldehído tiene un  $\delta^{13}\text{C}$  de aproximadamente -15,4 a aproximadamente -10,9, por ejemplo, de aproximadamente -13,92 a aproximadamente -13,84. El aldehído también puede tener un  $f_{\text{M}}^{14}\text{C}$  de al menos aproximadamente 1,003. Por ejemplo, el aldehído tiene un  $f_{\text{M}}^{14}\text{C}$  de al menos aproximadamente 1,01 o al menos aproximadamente 1,5. El aldehído puede tener un  $f_{\text{M}}^{14}\text{C}$  de aproximadamente 1,111 a aproximadamente 1,124.

- 50 Además se da a conocer en el presente documento un método de producción de un alcohol graso, comprendiendo el método producir en una célula huésped un polipéptido que comprende la secuencia de aminoácidos de SEQ ID NO: 66, 68, 70, 72, 74, 76, 78, 80 u 82, o una variante del mismo, y aislar el alcohol graso de la célula huésped. El alcohol graso puede secretarse por la célula.

- 55 Además se da a conocer en el presente documento un polipéptido que comprende una secuencia de aminoácidos que tiene al menos aproximadamente el 70%, al menos aproximadamente el 80%, al menos aproximadamente el 85%, al menos aproximadamente el 90%, al menos aproximadamente el 91%, al menos aproximadamente el 92%, al menos aproximadamente el 93%, al menos aproximadamente el 94%, al menos aproximadamente el 95%, al

menos aproximadamente el 96%, al menos aproximadamente el 97%, al menos aproximadamente el 98%, o al menos aproximadamente el 99% de identidad con SEQ ID NO: 66, 68, 70, 72, 74, 76, 78, 80 u 82.

Además se da a conocer en el presente documento un polipéptido que comprende la secuencia de aminoácidos de SEQ ID NO: 66, 68, 70, 72, 74, 76, 78, 80 u 82 con una o más sustituciones, adiciones, inserciones o deleciones de aminoácido. En algunas realizaciones, el polipéptido tiene actividad reductasa. El polipéptido comprende la secuencia de aminoácidos de SEQ ID NO: 66, 68, 70, 72, 74, 76, 78, 80 u 82, con una o más sustituciones de aminoácido conservadoras. Por ejemplo, el polipéptido comprende una o más de las siguientes sustituciones de aminoácido conservadoras: reemplazo de un aminoácido alifático, tal como alanina, valina, leucina e isoleucina, por otro aminoácido alifático; reemplazo de una serina por una treonina; reemplazo de una treonina por una serina; reemplazo de un residuo ácido, tal como ácido aspártico y ácido glutámico, por otro residuo ácido; reemplazo de un residuo que porta un grupo amida, tal como asparagina y glutamina, por otro residuo que porta un grupo amida; intercambio de un residuo básico, tal como lisina y arginina, con otro residuo básico; y reemplazo de un residuo aromático, tal como fenilalanina y tirosina, por otro residuo aromático. Además se da a conocer en el presente documento un polipéptido que tiene aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100, o más sustituciones, adiciones, inserciones o deleciones de aminoácido. En algunas realizaciones, el polipéptido tiene actividad reductasa.

Además se da a conocer en el presente documento un polipéptido que comprende la secuencia de aminoácidos de SEQ ID NO: 54, 55, 56, 57, 58, 59, 60, 61, 62, 63 ó 64. En determinadas realizaciones, el polipéptido tiene actividad reductasa.

Además se da a conocer en el presente documento un método de producción de un alcohol graso, comprendiendo el método expresar en una célula huésped un polinucleótido que comprende una secuencia de nucleótidos que tiene al menos aproximadamente el 70%, al menos aproximadamente el 75%, al menos aproximadamente el 80%, al menos aproximadamente el 85%, al menos aproximadamente el 90%, al menos aproximadamente el 91%, al menos aproximadamente el 92%, al menos aproximadamente el 93%, al menos aproximadamente el 94%, al menos aproximadamente el 95%, al menos aproximadamente el 96%, al menos aproximadamente el 97%, al menos aproximadamente el 98%, o al menos aproximadamente el 99% de identidad de secuencia con SEQ ID NO: 65, 67, 69, 71, 73, 75, 77, 79 u 81. En algunas realizaciones, la secuencia de nucleótidos es SEQ ID NO: 65, 67, 69, 71, 73, 75, 77, 79 u 81. El método puede comprender además aislar el alcohol graso de la célula huésped.

Además se da a conocer en el presente documento una secuencia de nucleótidos que se hibrida con un complemento de SEQ ID NO: 65, 67, 69, 71, 73, 75, 77, 79 u 81, o con un fragmento del mismo, por ejemplo, en condiciones de baja rigurosidad, media rigurosidad, alta rigurosidad o muy alta rigurosidad.

Además se da a conocer en el presente documento una secuencia de nucleótidos que codifica para un polipéptido que comprende: (i) la secuencia de aminoácidos de SEQ ID NO: 66, 68, 70, 72, 74, 76, 78, 80 u 82; o (ii) la secuencia de aminoácidos de SEQ ID NO: 66, 68, 70, 72, 74, 76, 78, 80 u 82 con una o más sustituciones, adiciones, inserciones o deleciones de aminoácido. En algunas realizaciones, el polipéptido comprende la secuencia de aminoácidos de SEQ ID NO: 66, 68, 70, 72, 74, 76, 78, 80 u 82 con una o más sustituciones de aminoácido conservadoras. En algunas realizaciones, el polipéptido tiene actividad reductasa.

Además se da a conocer en el presente documento una secuencia de nucleótidos que codifica para un polipéptido que tiene la misma actividad biológica que un polipéptido que comprende la secuencia de aminoácidos de SEQ ID NO: 66, 68, 70, 72, 74, 76, 78, 80 u 82. En algunas realizaciones, la secuencia de nucleótidos es SEQ ID NO: 65, 67, 69, 71, 73, 75, 77, 79 u 81 o un fragmento de la misma. En algunas realizaciones, la secuencia de nucleótidos se hibrida con un complemento de SEQ ID NO: 65, 67, 69, 71, 73, 75, 77, 79 u 81 o con un fragmento del mismo, por ejemplo, en condiciones de baja rigurosidad, media rigurosidad, alta rigurosidad o muy alta rigurosidad. En algunas realizaciones, la actividad biológica es actividad reductasa.

Además se da a conocer en el presente documento un método que comprende transformar una célula huésped con un vector recombinante que comprende una secuencia de nucleótidos que tiene al menos aproximadamente el 70%, al menos aproximadamente el 75%, al menos aproximadamente el 80%, al menos aproximadamente el 85%, al menos aproximadamente el 90%, al menos aproximadamente el 91 %, al menos aproximadamente el 92%, al menos aproximadamente el 93%, al menos aproximadamente el 94%, al menos aproximadamente el 95%, al menos aproximadamente el 96%, al menos aproximadamente el 97%, al menos aproximadamente el 98%, o al menos aproximadamente el 99% de identidad de secuencia con SEQ ID NO: 65, 67, 69, 71, 73, 75, 77, 79 u 81. Además se da a conocer en el presente documento un vector recombinante que comprende además un promotor unido operativamente a la secuencia de nucleótidos. En algunas realizaciones, el promotor es un promotor regulado por desarrollo, específico de orgánulos, específico de tejidos, inducible, constitutivo o específico de células. En realizaciones particulares, el vector recombinante comprende al menos una secuencia seleccionada del grupo que consiste en (a) una secuencia reguladora acoplada operativamente a la secuencia de nucleótidos; (b) un marcador de selección acoplado operativamente a la secuencia de nucleótidos; (c) una secuencia de marcador acoplada operativamente a la secuencia de nucleótidos; (d) un resto de purificación acoplado operativamente a la secuencia de nucleótidos; (e) una secuencia de secreción acoplada operativamente a la secuencia de nucleótidos; y (f) una secuencia de direccionamiento acoplada operativamente a la secuencia de nucleótidos. En determinadas

realizaciones, la secuencia de nucleótidos se incorpora de manera estable en el ADN genómico de la célula huésped, y la expresión de la secuencia de nucleótidos está bajo el control de una región de promotor regulada.

El método puede incluir además expresar un gen que codifica para una alcohol deshidrogenasa recombinante en la célula huésped.

5 Además se dan a conocer en el presente documento métodos que pueden producir alcoholes grasos que comprenden un alcohol graso C<sub>6</sub>-C<sub>26</sub>. En algunas realizaciones, el alcohol graso comprende un alcohol graso C<sub>6</sub>, C<sub>7</sub>, C<sub>8</sub>, C<sub>9</sub>, C<sub>10</sub>, C<sub>11</sub>, C<sub>12</sub>, C<sub>13</sub>, C<sub>14</sub>, C<sub>15</sub>, C<sub>16</sub>, C<sub>17</sub>, C<sub>18</sub>, C<sub>19</sub>, C<sub>20</sub>, C<sub>21</sub>, C<sub>22</sub>, C<sub>23</sub>, C<sub>24</sub>, C<sub>25</sub> o C<sub>26</sub>. En realizaciones particulares, el alcohol graso es 1-decanol, 1-dodecanol, alcohol 1-miristílico, 1-hexadecanol, octadecanol, tetradecanol o hexadecanol.

10 El alcohol graso puede comprender un alcohol graso de cadena lineal. El alcohol graso puede comprender un alcohol graso de cadena ramificada. El alcohol graso puede comprender un resto cíclico.

El alcohol graso puede ser un alcohol graso insaturado. El alcohol graso puede ser un alcohol graso monoinsaturado. El alcohol graso puede ser un alcohol graso saturado.

15 Además se da a conocer en el presente documento un alcohol graso producido por cualquiera de los métodos o cualquiera de los microorganismos descritos en el presente documento, o un tensioactivo que comprende un alcohol graso producido por cualquiera de los métodos o cualquiera de los microorganismos descritos en el presente documento.

El alcohol graso puede tener un  $\delta^{13}\text{C}$  de aproximadamente -15,4 o más. El alcohol graso puede tener un  $\delta^{13}\text{C}$  de aproximadamente -15,4 a aproximadamente -10,9, o de aproximadamente -13,92 a aproximadamente -13,84.

20 El alcohol graso puede tener un  $f_M^{14}\text{C}$  de al menos aproximadamente 1,003. El alcohol graso puede tener un  $f_M^{14}\text{C}$  de al menos aproximadamente 1,01 o al menos aproximadamente 1,5. El alcohol graso puede tener un  $f_M^{14}\text{C}$  de aproximadamente 1,111 a aproximadamente 1,124.

25 Además se da a conocer en el presente documento un ácido nucleico aislado que consiste en no más de aproximadamente 500 nucleótidos de SEQ ID NO: 65, 67, 69, 71, 73, 75, 77, 79 u 81. En algunas realizaciones, el ácido nucleico consiste en no más de aproximadamente 300 nucleótidos, no más de aproximadamente 350 nucleótidos, no más de aproximadamente 400 nucleótidos, no más de aproximadamente 450 nucleótidos, no más de aproximadamente 550 nucleótidos, no más de aproximadamente 600 nucleótidos, o no más de aproximadamente 650 nucleótidos, de SEQ ID NO: 65, 67, 69, 71, 73, 75, 77, 79 u 81. En algunas realizaciones, el ácido nucleico codifica para un polipéptido que tiene actividad reductasa.

30 Además se da a conocer en el presente documento un ácido nucleico aislado que consiste en no más de aproximadamente el 99%, no más de aproximadamente el 98%, no más de aproximadamente el 97%, no más de aproximadamente el 96%, no más de aproximadamente el 95%, no más de aproximadamente el 94%, no más de aproximadamente el 93%, no más de aproximadamente el 92%, no más de aproximadamente el 91%, no más de aproximadamente el 90%, no más de aproximadamente el 85%, o no más de aproximadamente el 80% de los nucleótidos de SEQ ID NO: 65, 67, 69, 71, 73, 75, 77, 79 u 81. En algunas realizaciones, el ácido nucleico codifica para un polipéptido que tiene actividad reductasa.

35 Además se da a conocer en el presente documento un polipéptido aislado que consiste en no más de aproximadamente 200, no más de aproximadamente 175, no más de aproximadamente 150, o no más de aproximadamente 100 de los aminoácidos de SEQ ID NO: 66, 68, 70, 72, 74, 76, 78, 80 u 82. En algunas realizaciones, el polipéptido tiene actividad reductasa.

40 Además se da a conocer en el presente documento un polipéptido aislado que consiste en no más de aproximadamente el 99%, no más de aproximadamente el 98%, no más de aproximadamente el 97%, no más de aproximadamente el 96%, no más de aproximadamente el 95%, no más de aproximadamente el 94%, no más de aproximadamente el 93%, no más de aproximadamente el 92%, no más de aproximadamente el 91%, no más de aproximadamente el 90%, no más de aproximadamente el 85%, o no más de aproximadamente el 80% de los aminoácidos de SEQ ID NO: 66, 68, 70, 72, 74, 76, 78, 80 u 82. En algunas realizaciones, el polipéptido tiene actividad reductasa.

#### Definiciones

45 En toda la memoria descriptiva, puede hacerse una referencia usando un nombre de gen o nombre de polipéptido abreviado, pero se entiende que un nombre de gen o de polipéptido abreviado de este tipo representa el género de genes o polipéptidos. Tales nombres de genes incluyen todos los genes que codifican para el mismo polipéptido y polipéptidos homólogos que tienen la misma función fisiológica. Los nombres de polipéptido incluyen todos los polipéptidos que tienen la misma actividad (por ejemplo, que catalizan la misma reacción química fundamental).

Los números de registro a los que se hace referencia en el presente documento se derivan de la base de datos del

NCBI (Centro Nacional para Información de Biotecnología) mantenida por el Instituto Nacional de Salud, EE. UU. A menos que se indique lo contrario, los números de registro son tal como se proporcionan en la base de datos de abril de 2009.

5 Los números de EC los establece el Comité de Nomenclatura de la Unión Internacional de Bioquímica y Biología Molecular (NC-IUBMB) (disponible en <http://www.chem.qmul.ac.uk/iubmb/enzyme/>). Los números de EC a los que se hace referencia en el presente documento se derivan de la base de datos KEGG Ligand, mantenida por la Enciclopedia de Genes y Genomas de Kioto, patrocinada en parte por la Universidad de Tokio. A menos que se indique lo contrario, los números de EC son tal como se proporcionan en la base de datos de marzo de 2008.

10 Los artículos “un” y “una” se usan en el presente documento para referirse a uno o a más de uno (es decir, a al menos uno) del objeto gramatical del artículo. A modo de ejemplo, “un elemento” significa un elemento o más de un elemento.

El término “aproximadamente” se usa en el presente documento para significar un valor  $\pm 20\%$  de un valor numérico dado. Por tanto, “aproximadamente el 60%” significa un valor de entre  $60 \pm$  (el 20% de 60) (es decir, entre 48 y 70).

15 Tal como se usa en el presente documento, el término “aldehído” significa un hidrocarburo que tiene la fórmula RCHO caracterizado por un grupo carbonilo insaturado (C=O). En una realización preferida, el aldehído es cualquier aldehído hecho de un ácido graso o derivado de ácido graso. En una realización, el grupo R es de al menos aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 ó 20 carbonos de longitud.

Tal como se usa en el presente documento, un “gen biosintético de aldehído” o un “polinucleótido biosintético de aldehído” es un ácido nucleico que codifica para un polipéptido biosintético de aldehído.

20 Tal como se usa en el presente documento, un “polipéptido biosintético de aldehído” es un polipéptido que es una parte de la ruta biosintética de un aldehído. Tales polipéptidos pueden actuar sobre un sustrato biológico para dar un aldehído. En algunos casos, el polipéptido biosintético de aldehído tiene actividad reductasa.

Tal como se usa en el presente documento, el término “alcano” significa un hidrocarburo que contiene sólo enlaces sencillos carbono-carbono.

25 Tal como se usa en el presente documento, un “gen biosintético de alcano” o un “polinucleótido biosintético de alcano” es un ácido nucleico que codifica para un polipéptido biosintético de alcano.

Tal como se usa en el presente documento, un “polipéptido biosintético de alcano” es un polipéptido que es una parte de la ruta biosintética de un alcano. Tales polipéptidos pueden actuar sobre un sustrato biológico para dar un alcano. En algunos casos, el polipéptido biosintético de alcano tiene actividad descarboxilasa.

30 Tal como se usa en el presente documento, un “gen biosintético de alqueno” o un “polinucleótido biosintético de alqueno” es un ácido nucleico que codifica para un polipéptido biosintético de alqueno.

Tal como se usa en el presente documento, un “polipéptido biosintético de alqueno” es un polipéptido que es una parte de la ruta biosintética de un alqueno. Tales polipéptidos pueden actuar sobre un sustrato biológico para dar un alqueno. En algunos casos, el polipéptido biosintético de alqueno tiene actividad descarboxilasa.

35 Tal como se usa en el presente documento, el término “atenuar” significa debilitar, reducir o disminuir. Por ejemplo, un polipéptido puede atenuarse modificando el polipéptido para reducir su actividad (por ejemplo, modificando una secuencia de nucleótidos que codifica para el polipéptido).

40 Tal como se usa en el presente documento, el término “biodiésel” significa un biocombustible que puede ser un sustituto de diésel, que se deriva de petróleo. Puede usarse biodiésel en motores diésel de combustión interna o bien en forma pura, que se denomina biodiésel “puro”, o bien como una mezcla en cualquier concentración con diésel a base de petróleo. El biodiésel puede incluir ésteres o hidrocarburos, tales como aldehídos y alcanos.

45 Tal como se usa en el presente documento, el término “biocombustible” se refiere a cualquier combustible derivado de biomasa. Los biocombustibles pueden sustituirse por combustibles a base de petróleo. Por ejemplo, los biocombustibles son inclusivos de combustibles para transporte (por ejemplo, gasolina, diésel, carburante, etc.), combustibles de calefacción y combustibles de generación de electricidad. Los biocombustibles son una fuente de energía renovable.

50 Tal como se usa en el presente documento, el término “biomasa” se refiere a una fuente de carbono derivada de material biológico. La biomasa puede convertirse en un biocombustible. Una fuente de biomasa a modo de ejemplo es materia vegetal. Por ejemplo, maíz, caña de azúcar o pasto pueden usarse como biomasa. Otro ejemplo no limitativo de biomasa es materia animal, por ejemplo estiércol de vaca. La biomasa también incluye productos residuales de la industria, agricultura, silvicultura y hogares. Ejemplos de tales productos residuales que pueden usarse como biomasa son restos de fermentación, paja, madera, aguas residuales, basura y restos de comida. La biomasa también incluye fuentes de carbono, tales como hidratos de carbono (por ejemplo, monosacáridos, disacáridos o polisacáridos).

Tal como se usa en el presente documento, la frase “fuente de carbono” se refiere a un sustrato o compuesto adecuado para usarse como fuente de carbono para crecimiento de células procariontas o eucariotas simples. Las fuentes de carbono pueden estar en diversas formas, incluyendo, pero sin limitarse a polímeros, hidratos de carbono, ácidos, alcoholes, aldehídos, cetonas, aminoácidos, péptidos y gases (por ejemplo, CO y CO<sub>2</sub>). Estos incluyen, por ejemplo, diversos monosacáridos, tales como glucosa, fructosa, manosa, y galactosa; oligosacáridos, tales como fructo-oligosacárido y galacto-oligosacárido; polisacáridos tales como xilosa y arabinosa; disacáridos, tales como sacarosa, maltosa, y turanosa; material celulósico, tal como metilcelulosa y carboximetilcelulosa de sodio; ésteres de ácidos grasos saturados o insaturados, tales como succinato, lactato, y acetato; alcoholes, tales como etanol o mezclas del mismo. La fuente de carbono también puede ser un producto de la fotosíntesis, incluyendo, pero sin limitarse a, glucosa. Una fuente de carbono preferida es biomasa. Otra fuente de carbono preferida es glucosa.

Tal como se usa en el presente documento, un “aditivo de reducción del punto de turbidez” es un aditivo añadido a una composición para disminuir o reducir el punto de turbidez de una disolución.

Tal como se usa en el presente documento, la frase “punto de turbidez de un fluido” significa la temperatura a la que los sólidos disueltos ya no son completamente solubles. Por debajo de esta temperatura, los sólidos empiezan a precipitar como fase secundaria dando al fluido un aspecto turbio. En la industria del petróleo, punto de turbidez se refiere a la temperatura por debajo de la que un material solidificado u otro hidrocarburo pesado cristalizan en un aceite crudo, aceite refinado, o combustible para formar un aspecto turbio. La presencia de materiales solidificados influye en el comportamiento de flujo del fluido, la tendencia del fluido a atascar los filtros de combustible, inyectores, etc., la acumulación de materiales solidificados sobre superficies frías (por ejemplo, un ensuciamiento de intercambiador de calor o tubería), y las características de emulsión del fluido con agua.

Una secuencia de nucleótidos es “complementaria” a otra secuencia de nucleótidos si cada una de las bases de las dos secuencias coincide (es decir, puede formar pares de bases de Watson Crick). El término “hebra complementaria” se usa en el presente documento de manera intercambiable con el término “complemento”. El complemento de una hebra de ácido nucleico puede ser el complemento de una hebra codificante o el complemento de una hebra no codificante.

Tal como se usa en el presente documento, el término “condiciones suficientes para permitir la expresión” significa cualquier condición que permite que una célula huésped produzca un producto deseado, tal como un polipéptido, aldehído, o alcano descrito en el presente documento. Las condiciones adecuadas incluyen, por ejemplo, condiciones de fermentación. Las condiciones de fermentación pueden comprender muchos parámetros, tales como intervalos de temperatura, niveles de aireación y composición de los medios. Cada una de estas condiciones, de manera individual y en combinación, permite que la célula huésped crezca. Los medios de cultivo a modo de ejemplo incluyen caldos o geles. Generalmente, el medio incluye una fuente de carbono, tal como glucosa, fructosa, celulosa, o similar, que puede metabolizarse mediante una célula huésped directamente. Además, pueden usarse enzimas en el medio para facilitar la movilización (por ejemplo, la despolimerización de almidón o celulosa a azúcares fermentables) y posterior metabolismo de la fuente de carbono.

Para determinar si las condiciones son suficientes para permitir la expresión, puede cultivarse una célula huésped, por ejemplo, durante aproximadamente 4, 8, 12, 24, 36 ó 48 horas. Durante y/o después de cultivar, pueden obtenerse muestras y analizarse para determinar si las condiciones permiten la expresión. Por ejemplo, las células huésped en la muestra o el medio en el que las células huésped se cultivaron pueden someterse a una prueba para determinar la presencia de un producto deseado. Cuando se somete a prueba la presencia de un producto, pueden usarse ensayos, tales como, pero sin limitarse a, CCF, HPLC, CG/FID, CG/EM, CL/EM, EM.

Se entiende que los polipéptidos descritos en el presente documento pueden tener sustituciones de aminoácidos conservadoras o no esenciales adicionales, que no tienen un efecto sustancial sobre las funciones del polipéptido. Si una sustitución particular será o no tolerada (es decir, no afectará de manera adversa a las propiedades biológicas deseadas, tales como actividad descarboxilasa) puede determinarse tal como se describe en Bowie *et al.*, Science (1990) 247:1306 1310. Una “sustitución de aminoácido conservadora” es una en la que el residuo de aminoácido se sustituye por un residuo de aminoácido que tiene una cadena lateral similar. Las familias de residuos de aminoácido que tienen cadenas laterales similares se han definido en la técnica. Estas familias incluyen aminoácidos con cadenas laterales básicas (por ejemplo, lisina, arginina, histidina), cadenas laterales ácidas (por ejemplo, ácido aspártico, ácido glutámico), cadenas laterales polares no cargadas (por ejemplo, glicina, asparagina, glutamina, serina, treonina, tirosina, cisteína), cadenas laterales no polares (por ejemplo, alanina, valina, leucina, isoleucina, prolina, fenilalanina, metionina, triptófano), cadenas laterales ramificadas beta (por ejemplo, treonina, valina, isoleucina), y cadenas laterales aromáticas (por ejemplo, tirosina, fenilalanina, triptófano, histidina).

Tal como se usa en el presente documento, “elemento de control” significa un elemento de control de la transcripción. Los elementos de control incluyen promotores y potenciadores. El término “elemento de promotor”, “promotor” o “secuencia de promotor” se refiere a una secuencia de ADN que funciona como un interruptor que activa la expresión de un gen. Si el gen se activa, se dice que está transcrito o que participa en la transcripción. Transcripción implica la síntesis de ARNm a partir del gen. Por tanto, un promotor sirve como elemento regulador de la transcripción y también proporciona un sitio para el inicio de la transcripción del gen en ARNm. Los elementos de

control interactúan específicamente con proteínas celulares implicadas en la transcripción (Maniatis *et al.*, Science 236:1237, 1987).

5 Tal como se usa en el presente documento, el término “éster sintasa” significa un péptido que puede producir ésteres grasos. Más específicamente, una éster sintasa es un péptido que convierte un tioéster en un éster graso. En una realización preferida, la éster sintasa convierte un tioéster (por ejemplo, acil-CoA) en un éster graso.

En una realización alternativa, una éster sintasa usa un tioéster y un alcohol como sustratos para producir un éster graso. Las éster sintasas pueden usar tioésteres de cadena corta y larga como sustratos. Además, las éster sintasas pueden usar alcoholes de cadena corta y larga como sustratos.

10 Ejemplos no limitativos de éster sintasas son cera sintasas, cera-éster sintasas, acil-CoA:alcohol transacilasas, aciltransferasas y acil graso-coenzima A:alcohol graso aciltransferasas. Éster sintasas a modo de ejemplo se clasifican en número de clasificación de enzima EC 2.3.1.75. En la figura 40 se proporcionan números de registro de GenBank a modo de ejemplo.

15 Tal como se usa en el presente documento, el término “ácido graso” significa un ácido carboxílico que tiene la fórmula RCOOH. R representa un grupo alifático, preferiblemente un grupo alquilo. R puede comprender entre aproximadamente 4 y aproximadamente 22 átomos de carbono. Los ácidos grasos pueden ser saturados, monoinsaturados o poliinsaturados. En una realización preferida, el ácido graso está compuesto por una ruta biosintética de ácido graso.

20 Tal como se usa en el presente documento, el término “ruta biosintética de ácido graso” significa una ruta biosintética que produce ácidos grasos. La ruta biosintética de ácido graso incluye enzimas de ácido graso que pueden modificarse por ingeniería genética, tal como se describe en el presente documento, para producir ácidos grasos, y en algunas realizaciones puede expresarse con enzimas adicionales para producir ácidos grasos que tienen características de cadena de carbono deseadas.

25 Tal como se usa en el presente documento, el término “derivado de ácido graso” significa productos compuestos en parte a partir de la ruta biosintética de ácido graso del organismo huésped de producción. “Derivado de ácido graso” también incluye productos compuestos en parte a partir de acil-ACP o derivados de acil-ACP. La ruta biosintética de ácido graso incluye enzimas ácido graso sintasa que pueden modificarse por ingeniería genética tal como se describe en el presente documento para producir derivados de ácidos grasos, y en algunos ejemplos pueden expresarse con enzimas adicionales para producir derivados de ácidos grasos que tienen características de cadena de carbono deseadas. Los derivados de ácidos grasos a modo de ejemplo incluyen por ejemplo, ácidos grasos, acil-CoA, aldehído graso, alcoholes de cadena corta y larga, hidrocarburos, alcoholes grasos y ésteres (por ejemplo, ceras, ésteres de ácido graso o ésteres grasos).

35 Tal como se usa en el presente documento, el término “enzimas de derivados de ácidos grasos” significa todas las enzimas que pueden expresarse o sobreexpresarse en la producción de derivados de ácidos grasos. Estas enzimas se denominan conjuntamente en el presente documento enzimas de derivados de ácidos grasos. Estas enzimas pueden ser parte de la ruta biosintética de ácido graso. Los ejemplos no limitativos de enzimas de derivados de ácidos grasos incluyen ácido graso sintasas, tioesterasas, acil-CoA sintasas, acil-CoA reductasas, alcohol deshidrogenasas, alcohol aciltransferasas, acil-CoA reductasa que forma alcohol graso, éster sintasas, polipéptidos biosintéticos de aldehído y polipéptidos biosintéticos de alcano. Enzimas de derivados de ácidos grasos convierten un sustrato en un derivado de ácido graso. En algunos ejemplos, el sustrato puede ser un derivado de ácido graso que la enzima de derivado de ácido graso convierte en un derivado de ácido graso diferente.

45 Tal como se usa en el presente documento, el término “péptido que forma alcohol graso” significa un péptido que puede catalizar la conversión de acil-CoA en alcohol graso, incluyendo acil-CoA reductasa que forma alcohol graso (FAR, EC 1.1.1.\*), acil-CoA reductasa (EC 1.2.1.50) o alcohol deshidrogenasa (EC 1.1.1.1). Además, un experto habitual en la técnica apreciará que algunos de los péptidos que forman alcohol graso catalizarán también otras reacciones. Por ejemplo, algunos péptidos de acil-CoA reductasa aceptarán otros sustratos además de ácidos grasos. Por tanto, también se incluyen tales péptidos no específicos. Las secuencias de ácido nucleico que codifican para péptidos que forman alcohol graso se conocen en la técnica, y tales péptidos están disponibles públicamente. En la figura 40 se proporcionan números de registro de GenBank a modo de ejemplo.

50 Tal como se usa en el presente documento, “enzima de ácido graso” significa cualquier enzima implicada en la biosíntesis de ácido graso. Las enzimas de ácido graso pueden expresarse o sobreexpresarse en células huésped para producir ácidos grasos. Los ejemplos no limitativos de enzimas de ácido graso incluyen ácido graso sintasas y tioesterasas.

55 Tal como se usa en el presente documento, el término “éster graso” significa un éster. En una realización preferida, un éster graso es cualquier éster que está compuesto por un ácido graso, por ejemplo un éster de ácido graso. En una realización, un éster graso contiene un lado A (es decir, la cadena de carbono unida al oxígeno carboxilato) y un lado B (es decir, la cadena de carbono que comprende el carboxilato original). En una realización preferida, cuando el éster graso se deriva a partir de la ruta biosintética de ácido graso, el lado A está contribuido por un alcohol, y el lado B está contribuido por un ácido graso. Cualquier alcohol puede usarse para formar el lado A de los ésteres

grasos. Por ejemplo, el alcohol puede derivarse de la ruta biosintética de ácido graso. Alternativamente, el alcohol puede producirse a través de las rutas biosintéticas de ácidos no grasos. Además, el alcohol puede proporcionarse de manera exógena. Por ejemplo, el alcohol puede suministrarse en el caldo de fermentación en casos en los que el éster graso se produce mediante un organismo. Alternativamente, un ácido carboxílico, tal como un ácido graso o ácido acético, puede suministrarse de manera exógena en casos en los que el éster graso se produce mediante un organismo que también puede producir alcohol.

Las cadenas de carbono que comprenden el lado A o el lado B pueden ser de cualquier longitud. En una realización, el lado A del éster es al menos de aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 10, 12, 14, 16 ó 18 carbonos de longitud. El lado B del éster es de al menos aproximadamente 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24 ó 26 carbonos de longitud. El lado A y/o el lado B pueden ser de cadena lineal o ramificada. Las cadenas ramificadas pueden tener uno o más puntos de ramificación. Además, las cadenas ramificadas pueden incluir ramas cíclicas. Además, el lado A y/o el lado B pueden ser saturados o insaturados. Si son insaturados, el lado A y/o el lado B pueden tener uno o más puntos de insaturación.

En una realización, el éster graso se produce de manera biosintética. En esta realización, en primer lugar se “activa” el ácido graso. Ejemplos no limitativos de ácidos grasos “activados” son acil-CoA, acil-ACP y fosfato de acilo. Acil-CoA puede ser un producto directo de biosíntesis o degradación de ácido graso. Además, acil-CoA puede sintetizarse a partir de un ácido graso libre, una CoA, o un nucleótido trifosfato de adenosina (ATP). Un ejemplo de una enzima que produce acil-CoA es acil-CoA sintasa.

Después de que se active el ácido graso, pueden transferirse fácilmente a un nucleófilo receptor. Nucleófilos a modo de ejemplo son alcoholes, tioles o fosfatos.

En una realización, el éster graso es una cera. La cera puede derivarse de un alcohol de cadena larga y un ácido graso de cadena larga. En otra realización, el éster graso puede derivarse de un acil-tioéster graso y un alcohol. En otra realización, el éster graso es un tioéster de ácido graso, por ejemplo acil graso-coenzima A (CoA). En otras realizaciones, el éster graso es un pantotenato de acilo graso, una proteína portadora de acilo (ACP) o un éster de fosfato graso. Los ésteres grasos tienen muchos usos. Por ejemplo, pueden usarse ésteres grasos como biocombustible.

Tal como se usa en el presente documento, “fracción de carbono moderno” o “ $f_M$ ” tiene el mismo significado que el definido por los Materiales de referencia convencionales (SRM) 4990B y 4990C del Instituto Nacional de Normas y Tecnología (NIST), conocidos como patrones de ácidos oxálicos HOxI y HOxII, respectivamente. La definición fundamental se refiere a 0,95 veces la razón de isótopo  $^{14}\text{C}/^{12}\text{C}$  HOxI (denominado AD 1950). Esto es aproximadamente equivalente a madera previa a la Revolución Industrial con decaimiento corregido. Para la biosfera viva actual (material vegetal), la  $f_M$  es de aproximadamente 1.1.

Los cálculos de “homología” entre dos secuencias pueden realizarse tal como sigue. Las secuencias se alinean con fines de comparación óptima (por ejemplo, pueden introducirse huecos en uno o ambos de un primer y un segundo aminoácido o secuencia de ácido nucleico para alineación óptima y pueden descartarse secuencias no homólogas con fines de comparación). En una realización preferida, la longitud de una secuencia de referencia que se alinea con fines de comparación es de al menos aproximadamente el 30%, preferiblemente de al menos aproximadamente el 40%, más preferiblemente de al menos aproximadamente el 50%, incluso más preferiblemente de al menos aproximadamente el 60%, e incluso más preferiblemente de al menos aproximadamente el 70%, al menos aproximadamente el 80%, al menos aproximadamente el 90%, o aproximadamente el 100% de la longitud de la secuencia de referencia. Se comparan entonces los residuos de aminoácido o nucleótidos en las correspondientes posiciones de aminoácido o posiciones de nucleótido. Cuando una posición en la primera secuencia se ocupa por el mismo residuo de aminoácido o nucleótido que la posición correspondiente en la segunda secuencia, entonces las moléculas son idénticas en esa posición (tal como se usa en el presente documento, “identidad” de aminoácido o ácido nucleico es equivalente a “homología” de aminoácido o ácido nucleico). El porcentaje de identidad entre las dos secuencias es una función del número de posiciones idénticas compartidas por las secuencias, considerando el número de huecos y la longitud de cada hueco, que han de introducirse para alineación óptima de las dos secuencias.

La comparación de secuencias y la determinación de porcentaje de homología entre dos secuencias pueden lograrse usando un algoritmo matemático. En una realización preferida, el porcentaje de homología entre dos secuencias de aminoácidos se determina usando el algoritmo de Needleman y Wunsch (1970), J. Mol. Biol. 48:444-453, algoritmo que se ha incorporado en el programa GAP en el paquete de software GCG, usando o bien una matriz Blossum 62 o una matriz PAM250, y un peso de hueco de 16, 14, 12, 10, 8, 6 ó 4 y un peso de longitud de 1, 2, 3, 4, 5 ó 6. En aún otra realización preferida, el porcentaje de homología entre dos secuencias de nucleótidos se determina usando el programa GAP en el paquete de software GCG, usando una matriz NWSgapdna.CMP y un peso de hueco de 40, 50, 60, 70 u 80 y un peso de longitud de 1, 2, 3, 4, 5 ó 6. Un conjunto particularmente preferido de parámetros (y el que debe usarse si el practicante está inseguro sobre qué parámetros deben aplicarse para determinar si una molécula está dentro de una limitación de homología de las reivindicaciones) son una matriz de puntuación Blossum 62 con una penalidad de hueco de 12, una penalidad de extensión de hueco de 4 y una penalidad de hueco de desplazamiento del marco de lectura de 5.

Tal como se usa en el presente documento, un “célula huésped” es una célula usada para producir un producto descrito en el presente documento (por ejemplo, un aldehído o alcano descrito en el presente documento). Una célula huésped puede modificarse para expresar o sobreexpresar genes seleccionados o para tener expresión atenuada de genes seleccionados. Los ejemplos no limitativos de células huésped incluyen células vegetales, de animales, humanas, bacterianas, de levadura o de hongos filamentosos.

Tal como se usa en el presente documento, el término “se hibrida en condiciones de baja rigurosidad, media rigurosidad, alta rigurosidad o muy alta rigurosidad” describe condiciones para la hibridación y el lavado. Puede encontrarse una guía para realizar las reacciones de hibridación en Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1 – 6.3.6. Se describen métodos acuosos y no acuosos en esa referencia y puede usarse cualquier método. Las condiciones de hibridación específicas a las que se hace referencia en el presente documento son tal como sigue: 1) condiciones de hibridación de baja rigurosidad en cloruro de sodio/citrato de sodio 6X (SSC) a aproximadamente 45°C, seguido por dos lavados en SSC 0,2X, SDS al 0,1% al menos a 50°C (la temperatura de los lavados puede aumentarse hasta 55°C para condiciones de rigurosidad baja); 2) condiciones de hibridación de rigurosidad media en SSC 6X a aproximadamente 45°C, seguido por uno o más lavados en SSC 0,2X, SDS al 0,1% a 60°C; 3) condiciones de hibridación de rigurosidad alta en SSC 6X a aproximadamente 45°C, seguido por uno o más lavados en SSC 0,2X, SDS al 0,1% a 65°C; y preferiblemente 4) condiciones de hibridación de rigurosidad muy altas son fosfato de sodio 0,5 M, SDS al 7% a 65°C, seguido por uno o más lavados a SSC 0,2X, SDS al 1% a 65°C. Las condiciones de rigurosidad muy alta (4) son las condiciones preferidas a menos que se especifique lo contrario.

El término “aislado” tal como se usa en el presente documento con respecto a ácidos nucleicos, tales como ADN o ARN, se refiere a moléculas separadas de otros ADN u ARN, respectivamente, que están presentes en la fuente natural del ácido nucleico. Además, un “ácido nucleico aislado” incluye fragmentos de ácido nucleico, tales como fragmentos que no se producen de manera natural. El término “aislado” también se usa en el presente documento para referirse a polipéptidos, que se aíslan a partir de otras proteínas celulares, y abarca tanto polipéptidos endógenos purificados como polipéptidos recombinantes. El término “aislado” tal como se usa en el presente documento también se refiere a un ácido nucleico o polipéptido que está sustancialmente libre de material celular, material viral, o medio de cultivo cuando se produce mediante técnicas de ADN recombinante. El término “aislado” tal como se usa en el presente documento también se refiere a un ácido nucleico o polipéptido que está sustancialmente libre de precursores químicos u otros productos químicos cuando se sintetiza químicamente.

Tal como se usa en el presente documento, el “nivel de expresión de un gen en una célula” se refiere al nivel de ARNm, transcrito(s) nacientes de pre-ARNm, transcrito que procesa productos intermedios, ARNm maduro(s) y/o productos de degradación codificados por el gen en la célula.

Tal como se usa en el presente documento, el término “microorganismo” significa especies microbianas procariotas y eucariotas de los dominios Archaea, Bacteria y Eucarya, incluyendo el último levadura y hongos filamentosos, protozoos, algas, o Protista superiores. El término “célula microbiana”, tal como se usa en el presente documento, significa una célula de un microorganismo.

Tal como se usa en el presente documento, el término “ácido nucleico” se refiere a polinucleótidos, tales como ácido desoxirribonucleico (ADN), y, cuando sea apropiado, ácido ribonucleico (ARN). El término también incluye análogos de o bien ARN o bien ADN compuestos por análogos de nucleótido, y, tal como es aplicable a la realización que está describiéndose, polinucleótidos monocatenarios y bicatenarios (sentido o antisentido), EST, cromosomas, ADNc, ARNm y ARNr.

Tal como se usa en el presente documento, el término “unido operativamente” significa que una secuencia de nucleótidos seleccionada (por ejemplo, que codifica para un polipéptido descrito en el presente documento) está en proximidad con un promotor para permitir que el promotor regule la expresión de la secuencia de nucleótidos seleccionada. Además, el promotor se sitúa aguas arriba de la secuencia de nucleótidos seleccionada en cuanto a la dirección de la transcripción y traducción. Por “unido operativamente” quiere decirse que una secuencia de nucleótidos y una(s) secuencia(s) reguladora(s) se conectan de tal manera para permitir la expresión génica cuando las moléculas apropiadas (por ejemplo, proteínas de activador de la transcripción) se unen a la(s) secuencia(s) reguladora(s).

El término “o” se usa en el presente documento para significar, y se usa de manera intercambiable con, el término “y/o”, a menos que el contexto indique claramente lo contrario.

Tal como se usa en el presente documento, “sobreexpresar” significa expresar o provocar que se exprese un ácido nucleico, polipéptido o hidrocarburo en una célula a una concentración mayor que la que se expresa normalmente en una célula silvestre correspondiente. Por ejemplo, un polipéptido puede “sobreexpresarse” en una célula huésped recombinante cuando el polipéptido está presente en una mayor concentración en la célula huésped recombinante en comparación con su concentración en una célula huésped no recombinante de la misma especie.

Tal como se usa en el presente documento, “coeficiente de reparto”, o “P”, se define como la concentración en equilibrio de un compuesto en una fase orgánica dividida entre la concentración en equilibrio en una fase acuosa

(por ejemplo, caldo de fermentación). En una realización de un sistema bifásico descrito en el presente documento, la fase orgánica se forma mediante el aldehído o alcano durante el proceso de producción. Sin embargo, en algunos ejemplos, puede proporcionarse una fase orgánica, tal como proporcionando una capa de octano, para facilitar la separación del producto. Cuando se describe un sistema bifásico, las características de reparto de un compuesto pueden describirse como logP. Por ejemplo, un compuesto con un logP de 1 repartiría 10:1 con respecto a la fase orgánica. Un compuesto con un logP de -1 repartiría 1:10 con respecto a la fase orgánica. Al elegir un caldo de fermentación y fase orgánica apropiados, un aldehído o alcano con un alto valor de logP puede separarse para dar la fase orgánica incluso a concentraciones muy bajas en el recipiente de fermentación.

Tal como se usa en el presente documento, el término “purificar”, “purificado” o “purificación” significa la retirada o aislamiento de una molécula de su entorno mediante, por ejemplo, aislamiento o separación. Moléculas “sustancialmente purificadas” están al menos aproximadamente el 60% libres, preferiblemente al menos aproximadamente el 75% libres, y más preferiblemente al menos aproximadamente el 90% libres de otros componentes con los que se asocian. Tal como se usa en el presente documento, estos términos también se refieren a la retirada de contaminantes de una muestra. Por ejemplo, la retirada de contaminantes puede dar como resultado un aumento del porcentaje de aldehídos o alcanos en una muestra. Por ejemplo, cuando se producen aldehídos o alcanos en una célula huésped, los aldehídos o alcanos pueden purificarse mediante la retirada de proteínas de la célula huésped. Tras la purificación, el porcentaje de aldehídos o alcanos en la muestra se aumenta.

Los términos “purificar”, “purificado” y “purificación” no requieren pureza absoluta. Son términos relativos. Por tanto, por ejemplo, cuando se producen aldehídos o alcanos en células huésped, un aldehído purificado o alcano purificado es uno que se separa sustancialmente de otros componentes celulares (por ejemplo, ácidos nucleicos, polipéptidos, lípidos, hidratos de carbono u otros hidrocarburos). En otro ejemplo, una preparación de aldehído purificado o alcano purificado es una en la que el aldehído o alcano está sustancialmente libre de contaminantes, tales como aquellos que pueden estar presentes tras la fermentación. En algunas realizaciones, un aldehído o un alcano se purifica cuando al menos aproximadamente el 50% en peso de una muestra está compuesto por el aldehído o alcano. En otras realizaciones, un aldehído o un alcano se purifican cuando al menos aproximadamente el 60%, el 70%, el 80%, el 85%, el 90%, el 92%, el 95%, el 98% o el 99% o más en peso de una muestra está compuesto por el aldehído o alcano.

Tal como se usa en el presente documento, el término “polipéptido recombinante” se refiere a un polipéptido que se produce mediante técnicas de ADN recombinante, en la que generalmente ADN que codifica para el polipéptido expresado o ARN se inserta en un vector de expresión adecuado y que a su vez se usa para transformar una célula huésped para producir el polipéptido o ARN.

Tal como se usa en el presente documento, el término “sustancialmente idéntico” (o “sustancialmente homólogo”) se usa para referirse a una primera secuencia de aminoácidos o nucleótidos que contiene un número suficiente de residuos de aminoácido o nucleótidos idénticos o equivalentes (por ejemplo, con una cadena lateral similar, por ejemplo, sustituciones de aminoácidos conservadas) con respecto a una segunda secuencia de aminoácidos o nucleótidos de modo que las secuencias de aminoácidos o nucleótidos primera y segunda tienen actividades similares.

Tal como se usa en el presente documento, el término “sintasa” significa una enzima que cataliza un procedimiento de síntesis. Tal como se usa en el presente documento, el término sintasa incluye sintasas, sintetasas y ligasas.

Tal como se usa en el presente documento, el término “transfección” significa la introducción de un ácido nucleico (por ejemplo, por medio de un vector de expresión) en una célula receptora mediante transferencia de genes mediada por ácido nucleico.

Tal como se usa en el presente documento, “transformación” se refiere a un procedimiento en el que el genotipo de una célula se cambia como resultado de la captación celular de ácido nucleico exógeno. Esto puede dar como resultado que la célula transformada exprese una forma recombinante de un ARN o polipéptido. En el caso de expresión antisentido del gen transferido, se altera la expresión de una forma que se produce de manera natural del polipéptido.

Tal como se usa en el presente documento, una “proteína de transporte” es un polipéptido que facilita el movimiento de uno o más compuestos en y/o fuera de un orgánulo celular y/o una célula.

Tal como se usa en el presente documento, una “variante” de polipéptido X se refiere a un polipéptido que tiene la secuencia de aminoácidos de polipéptido X en la que uno o más residuos de aminoácido se alteran. La variante puede tener cambios conservadores o cambios no conservadores. Puede encontrarse una guía para determinar qué residuos de aminoácido pueden sustituirse, insertarse o deleccionarse sin afectar a la actividad biológica usando programas informáticos bien conocidos en la técnica, por ejemplo, software LASERGENE (DNASTAR).

El término “variante”, cuando se usa en el contexto de una secuencia de polinucleótido, puede abarcar una secuencia de polinucleótido relacionada con la de un gen o la secuencia codificante del mismo. Esta definición también puede incluir, por ejemplo, variantes “alélicas”, “de corte y empalme”, “de especies” o “polimórficas”. Una variante de corte y empalme puede tener una identidad significativa con respecto a un polinucleótido de referencia,

pero tendrá generalmente un número mayor o menor de polinucleótidos debido a corte y empalme alternativo de exones durante el procesamiento de ARNm. El polipéptido correspondiente puede poseer dominios funcionales adicionales o una ausencia de dominios. Las variantes de especies son secuencias de polinucleótido que varían de una especie a otra. Los polipéptidos resultantes tendrán generalmente una identidad de aminoácido significativa en relación entre sí. Una variante polimórfica es una variación en la secuencia de polinucleótido de un gen particular entre individuos de una especie dada.

Tal como se usa en el presente documento, el término "vector" se refiere a una molécula de ácido nucleico que puede transportar otro ácido nucleico al que se ha unido. Un tipo de vector útil es un episoma (es decir, un ácido nucleico que puede realizar replicación extracromosómica). Vectores útiles son aquellos que pueden realizar replicación autónoma y/o expresión de ácidos nucleicos a los que se unen. Los vectores que pueden dirigir la expresión de genes a los que se unen operativamente se denominan en el presente documento "vectores de expresión". En general, los vectores de expresión de utilidad en técnicas de ADN recombinante están a menudo en forma de "plásmidos", que se refieren generalmente a bucles de ADN bicatenario circular que, en su forma de vector, no están unidos al cromosoma. En la presente memoria descriptiva, "plásmido" y "vector" se usan de manera intercambiable, puesto que el plásmido es la forma más comúnmente usada de vector. Sin embargo, también se incluyen otras formas de vectores de expresión que sirven funciones equivalentes y que se vuelven conocidas en la técnica posteriormente aquí presente.

A menos que se defina lo contrario, todos los términos técnicos y científicos usados en el presente documento tienen el mismo significado que entiende comúnmente un experto habitual en la técnica a la que pertenece esta invención. Aunque pueden usarse métodos y materiales similares o equivalentes a los descritos en el presente documento en la práctica o pruebas de la presente invención, se describen a continuación métodos y materiales adecuados. En caso de conflicto, la presente memoria descriptiva, incluyendo definiciones, prevalecerá. Además, los materiales, métodos y ejemplos son sólo ilustrativos y no se pretende que sean limitativos.

Otras características y ventajas de la invención serán evidentes a partir de la siguiente descripción detallada y a partir de las reivindicaciones.

#### Breve descripción de los dibujos

La figura 1A es un trazo de CG/EM de hidrocarburos producidos por células de *Prochlorococcus marinus* CCMP1986. La figura 1B es un patrón de fragmentación de masa del pico a 7,55 min de la figura 1A.

La figura 2A es un trazo de CG/EM de hidrocarburos producidos por células de *Nostoc punctiforme* PCC73102. La figura 2B es un patrón de fragmentación de masa del pico a 8,73 min de la figura 2A.

La figura 3A es un trazo de CG/EM de hidrocarburos producidos por células de *Gloeobaceter violaceus* ATCC29082. La figura 3B es un patrón de fragmentación de masa del pico a 8,72 min de la figura 3A.

La figura 4A es un trazo de CG/EM de hidrocarburos producidos por células de *Synechocystis* sp. PCC6803. La figura 4B es un patrón de fragmentación de masa del pico a 7,36 min de la figura 4A.

La figura 5A es un trazo de CG/EM de hidrocarburos producidos por células silvestres de *Synechocystis* sp. PCC6803. La figura 5B es un trazo de CG/EM de hidrocarburos producidos por células de *Synechocystis* sp. PCC6803 con una delección de los genes *slI0208* y *slI0209*.

La figura 6A es un trazo de CG/EM de hidrocarburos producidos por células silvestres de *E. coli* MG1655. La figura 6B es un trazo de CG/EM de hidrocarburos producidos por células de *E. coli* MG1655 que expresan *Synechococcus elongatus* PCC7942 YP\_400611 (Synpcc7942\_1594) (SEQ ID NO: 65).

La figura 7 es un trazo de CG/EM de hidrocarburos producidos por células de *E. coli* que expresan *Cyanothece* sp. ATCC51142 *cce\_1430* (YP\_001802846) (SEQ ID NO: 69).

La figura 8A es un trazo de CG/EM de hidrocarburos producidos por células de *E. coli* que expresan *Synechococcus elongatus* PCC7942 YP\_400611 (Synpcc7942\_1594) (SEQ ID NO: 65) y *Synechococcus elongatus* PCC7942 YP\_400610 (Synpcc7942\_1593) (SEQ ID NO: 1). La figura 8B representa patrones de fragmentación de masa del pico a 6,98 min de la figura 8A y de pentadecano. La figura 8C representa patrones de fragmentación de masa del pico a 8,12 min de la figura 8A y de 8-heptadeceno.

La figura 9 es un trazo de CG/EM de hidrocarburos producidos por células de *E. coli* MG1655 que expresan *Synechococcus elongatus* PCC7942 YP\_400611 (Synpcc7942\_1594) (SEQ ID NO: 65) y *Nostoc punctiforme* PCC73102 Npun02004178 (ZP\_00108838) (SEQ ID NO: 5).

La figura 10 es un trazo de CG/EM de hidrocarburos producidos por células de *E. coli* MG1655 que expresan *Synechococcus elongatus* PCC7942 YP\_400611 (Synpcc7942\_1594) (SEQ ID NO: 65) y *Synechocystis* sp. PCC6803 *slI0208* (NP\_442147) (SEQ ID NO: 3).

La figura 11 es un trazo de CG/EM de hidrocarburos producidos por células de *E. coli* MG1655 que expresan

*Synechococcus elongatus* PCC7942 YP\_400611 (Synpcc7942\_1594) (SEQ ID NO: 65) y *Nostoc* sp. PCC7210 alr5283 (NP\_489323) (SEQ ID NO: 7).

5 La figura 12 es un trazo de CG/EM de hidrocarburos producidos por células de *E. coli* MG1655 que expresan *Synechococcus elongatus* PCC7942 YP\_400611 (Synpcc7942\_1594) (SEQ ID NO: 65) y *Acaryochloris marina* optimizado con codón MBIC11017 AM1\_4041 (YP\_001518340) (SEQ ID NO: 46).

La figura 13 es un trazo de CG/EM de hidrocarburos producidos por células de *E. coli* MG1655 que expresan *Synechococcus elongatus* PCC7942 YP\_400611 (Synpcc7942\_1594) (SEQ ID NO: 65) y *Thermosynechococcus elongatus* optimizado con codón BP-1 tll1313 (NP\_682103) (SEQ ID NO: 47).

10 La figura 14 es un trazo de CG/EM de hidrocarburos producidos por células de *E. coli* MG1655 que expresan *Synechococcus elongatus* PCC7942 YP\_400611 (Synpcc7942\_1594) (SEQ ID NO: 65) y *Synechococcus* sp. optimizado con codón JA-3-3Ab CYA\_0415 (YP\_473897) (SEQ ID NO: 48).

La figura 15 es un trazo de CG/EM de hidrocarburos producidos por células de *E. coli* MG1655 que expresan *Synechococcus elongatus* PCC7942 YP\_400611 (Synpcc7942\_1594) (SEQ ID NO: 65) y *Gloeobacter violaceus* PCC7421 gll3146 (NP\_926092) (SEQ ID NO: 15).

15 La figura 16 es un trazo de CG/EM de hidrocarburos producidos por células de *E. coli* MG1655 que expresan *Synechococcus elongatus* PCC7942 YP\_400611 (Synpcc7942\_1594) (SEQ ID NO: 65) y *Prochlorococcus marinus* optimizado con codón MIT9313 PMT1231 (NP\_895059) (SEQ ID NO: 49).

20 La figura 17 es un trazo de CG/EM de hidrocarburos producidos por células de *E. coli* MG1655 que expresan *Synechococcus elongatus* PCC7942 YP\_400611 (Synpcc7942\_1594) (SEQ ID NO: 65) y *Prochlorococcus marinus* CCMP1986 PMM0532 (NP\_892650) (SEQ ID NO: 19).

La figura 18 es un trazo de CG/EM de hidrocarburos producidos por células de *E. coli* MG1655 que expresan *Synechococcus elongatus* PCC7942 YP\_400611 (Synpcc7942\_1594) (SEQ ID NO: 65) y *Prochlorococcus marinus* optimizado con codón NATL2A PMN2A\_1863 (YP\_293054) (SEQ ID NO: 51).

25 La figura 19 es un trazo de CG/EM de hidrocarburos producidos por células de *E. coli* MG1655 que expresan *Synechococcus elongatus* PCC7942 YP\_400611 (Synpcc7942\_1594) (SEQ ID NO: 65) y *Synechococcus* sp. optimizado con codón RS9917 RS9917\_09941 (ZP\_01079772) (SEQ ID NO: 52).

La figura 20 es un trazo de CG/EM de hidrocarburos producidos por células de *E. coli* MG1655 que expresan *Synechococcus elongatus* PCC7942 YP\_400611 (Synpcc7942\_1594) (SEQ ID NO: 65) y *Synechococcus* sp. optimizado con codón RS9917 RS9917\_12945 (ZP\_01080370) (SEQ ID NO: 53).

30 La figura 21 es un trazo de CG/EM de hidrocarburos producidos por células de *E. coli* MG1655 que expresan *Synechococcus elongatus* PCC7942 YP\_400611 (Synpcc7942\_1594) (SEQ ID NO: 65) y *Cyanothece* sp. ATCC51142 cce\_0778 (YP\_001802195) (SEQ ID NO: 27).

35 La figura 22 es un trazo de CG/EM de hidrocarburos producidos por células de *E. coli* MG1655 que expresan *Synechococcus elongatus* PCC7942 YP\_400611 (Synpcc7942\_1594) (SEQ ID NO: 65) y *Cyanothece* sp. PCC7425 Cyan7425\_0398 (YP\_002481151) (SEQ ID NO: 29).

La figura 23 es un trazo de CG/EM de hidrocarburos producidos por células de *E. coli* MG1655 que expresan *Synechococcus elongatus* PCC7942 YP\_400611 (Synpcc7942\_1594) (SEQ ID NO: 65) y *Cyanothece* sp. PCC7425 Cyan7425\_2986 (YP\_002483683) (SEQ ID NO: 31).

40 La figura 24A es un trazo de CG/EM de hidrocarburos producidos por células de *E. coli* MG1655 que expresan *Prochlorococcus marinus* CCMP1986 PMM0533 (NP\_892651) (SEQ ID NO: 71). La figura 24B es un trazo de CG/EM de hidrocarburos producidos por células de *E. coli* MG1655 que expresan *Prochlorococcus marinus* CCMP1986 PMM0533 (NP\_892651) (SEQ ID NO: 71) y *Prochlorococcus marinus* CCMP1986 PMM0532 (NP\_892650) (SEQ ID NO: 19).

45 La figura 25A es un trazo de CG/EM de hidrocarburos producidos por células de *E. coli* MG1655  $\Delta$ fadE lacZ::P<sub>trc</sub> 'tesA-fadD. La figura 25B es un trazo de CG/EM de hidrocarburos producidos por células de *E. coli* MG1655  $\Delta$ fadE lacZ::P<sub>trc</sub> 'tesA-fadD que expresan *Synechococcus elongatus* PCC7942 YP\_400611 (Synpcc7942\_1594) (SEQ ID NO: 65) y *Acaryochloris marina* MBIC11017 AM1\_4041 (YP\_001518340) (SEQ ID NO: 9).

50 La figura 26A es un trazo de CG/EM de hidrocarburos producidos por células de *E. coli* MG1655  $\Delta$ fadE lacZ::P<sub>trc</sub> 'tesA-fadD que expresan *Synechocystis* sp. PCC6803 sll0209 (NP\_442146) (SEQ ID NO: 67). La figura 26B es un trazo de CG/EM de hidrocarburos producidos por células de *E. coli* MG1655  $\Delta$ fadE lacZ::P<sub>trc</sub> 'tesA-fadD que expresan *Synechocystis* sp. PCC6803 sll0209 (NP\_442146) (SEQ ID NO: 67) y *Synechocystis* sp. PCC6803 sll0208 (NP\_442147) (SEQ ID NO: 3).

La figura 27A es un trazo de CG/EM de hidrocarburos producidos por células de *E. coli* MG1655  $\Delta$ fadD lacZ::P<sub>trc</sub>-tesA que expresan la cepa de *M. smegmatis* MC2 155 MSMEG\_5739 (YP\_889972) (SEQ ID NO: 85). La figura 27B es un trazo de CG/EM de hidrocarburos producidos por células de *E. coli* MG1655  $\Delta$ fadD lacZ::P<sub>trc</sub>-tesA que expresan la cepa de *M. smegmatis* MC2 155 MSMEG\_5739 (YP\_889972) (SEQ ID NO: 85) y *Nostoc punctiforme* PCC73102 Npun02004178 (ZP\_00108838) (SEQ ID NO: 5).

La figura 28 es una representación gráfica de hidrocarburos producidos por células de *E. coli* MG1655  $\Delta$ fadD lacZ::P<sub>trc</sub>-tesA que expresan la cepa de *M. smegmatis* MC2 155 MSMEG\_5739 (YP\_889972) (SEQ ID NO: 85) o bien sola o bien en combinación con *Nostoc* sp. PCC7120 alr5283 (SEQ ID NO: 7), *Nostoc punctiforme* PCC73102 Npun02004178 (SEQ ID NO: 5), *P. marinus* CCMP1986 PMM0532 (SEQ ID NO: 19), *G. violaceus* PCC7421 gll3146 (SEQ ID NO: 15), *Synechococcus* sp. RS9917\_09941 (SEQ ID NO: 23), *Synechococcus* sp. RS9917\_12945 (SEQ ID NO: 25), o *A. marina* MBIC11017 AM1\_4041 (SEQ ID NO: 9).

La figura 29A es una representación de la estructura tridimensional de una proteína de la subunidad  $\beta$  de ribonucleasa reductasa clase I, RNR $\beta$ . La figura 29B es una representación de la estructura tridimensional de *Prochlorococcus marinus* MIT9313 PMT1231 (NP\_895059) (SEQ ID NO: 17). La figura 29C es una representación de la estructura tridimensional del sitio activo de *Prochlorococcus marinus* MIT9313 PMT1231 (NP\_895059) (SEQ ID NO: 17).

La figura 30A es un trazo de CG/EM de hidrocarburos producidos por células de *E. coli* MG1655 que expresan *Nostoc punctiforme* PCC73102 Npun02004178 (ZP\_00108838) (SEQ ID NO: 5). La figura 30B es un trazo de CG/EM de hidrocarburos producidos por células de *E. coli* MG1655 que expresan *Nostoc punctiforme* PCC73102 Npun02004178 (ZP\_00108838), variante Y123F. La figura 30C es un trazo de CG/EM de hidrocarburos producidos por células de *E. coli* MG1655 que expresan *Nostoc punctiforme* PCC73102 Npun02004178 (ZP\_00108838), variante Y126F.

La figura 31 representa trazas de CG/EM de hidrocarburos producidos *in vitro* usando *Nostoc punctiforme* PCC73102 Npun02004178 (ZP\_00108838) (SEQ ID NO: 6) y octadecanal (A); Npun02004178 (ZP\_00108838) (SEQ ID NO: 6), octadecanal, ferredoxina reductasa de espinaca, y NADPH (B); octadecanal, ferredoxina de espinaca, ferredoxina reductasa de espinaca, y NADPH (C); o Npun02004178 (ZP\_00108838) (SEQ ID NO: 6), ferredoxina de espinaca, y ferredoxina de espinaca (D).

La figura 32 representa trazas de CG/EM de hidrocarburos producidos *in vitro* usando *Nostoc punctiforme* PCC73102 Npun02004178 (ZP\_00108838) (SEQ ID NO: 6), NADPH, octadecanal, y o bien (A) ferredoxina de espinaca y ferredoxina reductasa de espinaca; (B) *N. punctiforme* PCC73102 Npun02003626 (ZP\_00109192) (SEQ ID NO: 88) y *N. punctiforme* PCC73102 Npun02001001 (ZP\_00111633) (SEQ ID NO: 90); (C) Npun02003626 (ZP\_00109192) (SEQ ID NO: 88) y *N. punctiforme* PCC73102 Npun02003530 (ZP\_00109422) (SEQ ID NO: 92); o bien (D) Npun02003626 (ZP\_00109192) (SEQ ID NO: 88) y *N. punctiforme* PCC73102 Npun02003123 (ZP\_00109501) (SEQ ID NO: 94).

La figura 33A es un trazo de CG/EM de hidrocarburos producidos *in vitro* usando octadecanoil-CoA, *Synechococcus elongatus* PCC7942 YP\_400611 (Synpcc7942\_1594) (SEQ ID NO: 66), NADH y Mg<sup>2+</sup>. La figura 33B es un trazo de CG/EM de hidrocarburos producidos *in vitro* usando octadecanoil-CoA, *Synechococcus elongatus* PCC7942 YP\_400611 (Synpcc7942\_1594) (SEQ ID NO: 66), NADPH y Mg<sup>2+</sup>. La figura 33C es un trazo de CG/EM de hidrocarburos producidos *in vitro* usando octadecanoil-CoA, *Synechococcus elongatus* PCC7942 YP\_400611 (Synpcc7942\_1594) (SEQ ID NO: 66) y NADPH.

La figura 34A es un trazo de CG/EM de hidrocarburos producidos *in vitro* usando octadecanoil-CoA, NADPH marcado, *Synechococcus elongatus* PCC7942 YP\_400611 (Synpcc7942\_1594) (SEQ ID NO: 66), y NADPH no marcado. La figura 34B es un trazo de CG/EM de hidrocarburos producidos *in vitro* usando octadecanoil-CoA, NADPH marcado, *Synechococcus elongatus* PCC7942 YP\_400611 (Synpcc7942\_1594) (SEQ ID NO: 66), y S-(4-<sup>2</sup>H)NADPH. La figura 34C es un trazo de CG/EM de hidrocarburos producidos *in vitro* usando octadecanoil-CoA, NADPH marcado, *Synechococcus elongatus* PCC7942 YP\_400611 (Synpcc7942\_1594) (SEQ ID NO: 66), y R-(4-<sup>2</sup>H)NADPH.

La figura 35 es un trazo de CG/EM de hidrocarburos en el sobrenadante libre de células producido por células de *E. coli* MG1655  $\Delta$ fadE en medios Che-9 que expresan *Synechococcus elongatus* PCC7942 YP\_400611 (Synpcc7942\_1594) (SEQ ID NO: 65).

La figura 36 es un trazo de CG/EM de hidrocarburos en el sobrenadante libre de células producido por células de *E. coli* MG1655  $\Delta$ fadE en medios Che-9 que expresan *Synechococcus elongatus* PCC7942 YP\_400611 (Synpcc7942\_1594) (SEQ ID NO: 65) y *Nostoc punctiforme* PCC73102 Npun02004178 (ZP\_00108838) (SEQ ID NO: 5).

La figura 37 es un trazo de CG/EM de hidrocarburos producidos por células de *E. coli* MG1655 que expresan *Nostoc* sp. PCC7120 alr5283 (NP\_489323) (SEQ ID NO: 7) y *Nostoc* sp. PCC7120 alr5284 (NP\_489324) (SEQ ID NO: 81).

La figura 38 es una lista de ejemplos de homólogos de *Synechococcus elongatus* PCC7942 YP\_400610 (Synpcc7942\_1593) (SEQ ID NO: 1) a partir de una base de datos metagenómica.

La figura 39 es una lista de ejemplos de homólogos de *Synechococcus elongatus* PCC7942 YP\_400611 (Synpcc7942\_1594) (SEQ ID NO: 65) a partir de una base de datos metagenómica.

- 5 La figura 40 es una tabla que identifica diversos genes que pueden expresarse, sobreexpresarse o atenuarse para aumentar la producción de sustratos particulares.

### Descripción detallada

10 Se dan a conocer en el presente documento composiciones y métodos de producción de aldehídos, alcoholes grasos e hidrocarburos (tales como alcanos, alquenos y alquinos) a partir de sustratos, por ejemplo, una acil-ACP, un ácido graso, una acil-CoA, un aldehído graso o un sustrato de alcohol graso (por ejemplo, tal como se describe en el documento PCT/US08/058788). Tales aldehídos, alcanos y alquenos son útiles como biocombustibles (por ejemplo, sustitutos para gasolina, diésel, carburante, etc.), productos químicos especializados (por ejemplo, lubricantes, aditivo de combustible, etc.), o materia prima para conversión química adicional (por ejemplo, combustibles, polímeros, plásticos, textiles, disolventes, adhesivos, etc.). La invención se basa, en parte, en la  
15 identificación de genes que están implicados en la biosíntesis de aldehído, alcano y alqueno.

Tales genes biosintéticos de alcano y alqueno incluyen, por ejemplo, *Synechococcus elongatus* PCC7942 Synpcc7942\_1593 (SEQ ID NO: 1), *Synechocystis* sp. PCC6803 sll0208 (SEQ ID NO: 3), *Nostoc punctiforme* PCC 73102 Npun02004178 (SEQ ID NO: 5), *Nostoc* sp. PCC 7120 alr5283 (SEQ ID NO: 7), *Acaryochloris marina* MBIC11017 AM1\_4041 (SEQ ID NO: 9), *Thermosynechococcus elongatus* BP-1 tll1313 (SEQ ID NO: 11),  
20 *Synechococcus* sp. JA-3-3A CYA\_0415 (SEQ ID NO: 13), *Gloeobacter violaceus* PCC 7421 gll3146 (SEQ ID NO: 15), *Prochlorococcus marinus* MIT9313 PM123 (SEQ ID NO: 17), *Prochlorococcus marinus* subsp. *pastoris* cepa CCMP1986 PMM0532 (SEQ ID NO: 19), *Prochlorococcus marinus* cepa NATL2A PMN2A\_1863 (SEQ ID NO: 21), *Synechococcus* sp. RS9917 RS9917\_09941 (SEQ ID NO: 23), *Synechococcus* sp. RS9917 RS9917\_12945 (SEQ ID NO: 25), *Cyanothece* sp ATCC51142 cce\_0778 (SEQ ID NO: 27), *Cyanothece* sp PCC7245 Cyan7425DRAFT\_1220 (SEQ ID NO: 29), *Cyanothece* sp. PCC7245 cce\_0778 (SEQ ID NO: 31), *Anabaena variabilis* ATCC29413 YP\_323043 (Ava\_2533) (SEQ ID NO: 33), y *Synechococcus elongatus* PCC6301 YP\_170760 (syc0050\_d) (SEQ ID NO: 35). Otros genes biosintéticos de alcano y alqueno se enumeran en la tabla 1 y la figura 38.

Los genes biosintéticos de aldehído incluyen, por ejemplo, *Synechococcus elongatus* PCC7942 Synpcc7942\_1594 (SEQ ID NO: 65), *Synechocystis* sp. PCC6803 sll0209 (SEQ ID NO: 67), *Cyanothece* sp ATCC51142 cce\_1430 (SEQ ID NO: 69), *Prochlorococcus marinus* subsp. *pastoris* cepa CCMP1986 PMM0533 (SEQ ID NO: 71), *Gloeobacter violaceus* PCC7421 NP\_96091 (gll3145) (SEQ ID NO: 73), *Nostoc punctiforme* PCC73102 ZP\_00108837 (Npun02004176) (SEQ ID NO: 75), *Anabaena variabilis* ATCC29413 YP\_323044 (Ava\_2534) (SEQ ID NO: 77), *Synechococcus elongatus* PCC6301 YP\_170761 (syc0051\_d) (SEQ ID NO: 79), y *Nostoc* sp. PCC 7120 alr5284 (SEQ ID NO: 81). Otros genes biosintéticos de aldehído se enumeran en la tabla 1 y la figura 39.

35 Usando los métodos descritos en el presente documento, pueden prepararse aldehídos, alcoholes grasos, alcanos y alquenos usando uno o más genes biosintéticos de aldehído, alcano y/o alqueno o polipéptidos descritos en el presente documento, o variantes de los mismos, usando células huésped o métodos libres de células.

Tabla 1: Homólogos de genes biosintéticos de aldehído y alcano en genomas cianobacterianos

Cianobacteria	Gen biosintético de alcano		Gen biosintético de aldehído	
	número de registro	% de ID	número de registro	% de ID
<i>Synechococcus elongatus</i> PCC 7942	YP_400610	100	YP_400611	100
<i>Synechococcus elongatus</i> PCC 6301	YP_170760	100	YP_170761	100
<i>Microcoleus chthonoplastes</i> PCC 7420	EDX75019	77	EDX74978	70
<i>Arthrospira maxima</i> CS-328	EDZ94963	78	EDZ94968	68
<i>Lyngbya</i> sp. PCC 8106	ZP_01619575	77	ZP_01619574	69
<i>Nodularia spumigena</i> CCY9414	ZP_01628096	77	ZP_01628095	70
<i>Trichodesmium erythraeum</i> IMS101	YP_721979	76	YP_721978	69
<i>Microcystis aeruginosa</i> NIES-843	YP_001660323	75	YP_001660322	68
<i>Microcystis aeruginosa</i> PCC 7806	CAO90780	74	CAO90781	67
<i>Nostoc</i> sp. PCC 7120	NP_489323	74	NP_489324	72
<i>Nostoc azollae</i> 0708	EEG05692	73	EEG05693	70
<i>Anabaena variabilis</i> ATCC 29413	YP_323043	74	YP_323044	73
<i>Crocospaera watsonii</i> WH 8501	ZP_00514700	74	ZP_00516920	67
<i>Synechocystis</i> sp. PCC 6803	NP_442147	72	NP_442146	68
<i>Synechococcus</i> sp. PCC 7335	EDX86803	73	EDX87870	67
<i>Cyanothece</i> sp ATCC 51142	YP_001802195	73	YP_001802846	67
<i>Cyanothece</i> sp CCY0110	ZP_01738578	72	ZP_01728620	68

ES 2 696 773 T3

<i>Nostoc punctiforme</i> PCC 73102	ZP_00108838	72	ZP_00108837	71
<i>Acaryochloris marina</i> MBIC11017	YP_001518340	71	YP_001518341	66
<i>Cyanothece</i> sp PCC 7425	YP_002481151	71	YP_002481152	70
<i>Cyanothece</i> sp PCC 8801	ZP_02941459	70	ZP_02942716	69
<i>Thermosynechococcus elongatus</i> BP-1	NP_682103	70	NP_682102	70
<i>Synechococcus</i> sp. JA-2-3B'a(2-13)	YP_478639	68	YP_478638	63
<i>Synechococcus</i> sp. RCC307	YP_001227842	67	YP_001227841	64
<i>Synechococcus</i> sp. WH 7803	YP_001224377	68	YP_001224378	65
<i>Synechococcus</i> sp. WH 8102	NP_897829	70	NP_897828	65
<i>Synechococcus</i> sp. WH 7805	ZP_01123214	68	ZP_01123215	65
<i>Synechococcus</i> marino no cultivado tipo A GOM 3O12	ABD96376	70	ABD96375	65
<i>Synechococcus</i> sp. JA-3-3Ab	YP_473897	68	YP_473896	62
<i>Synechococcus</i> marino no cultivado tipo A GOM 3O6	ABD96328	70	ABD96327	65
<i>Synechococcus</i> marino no cultivado tipo A GOM 3M9	ABD96275	68	ABD96274	65
<i>Synechococcus</i> sp. CC9311	YP_731193	63	YP_731192	63
<i>Synechococcus</i> marino no cultivado tipo A 5B2	ABB92250	69	ABB92249	64
<i>Synechococcus</i> sp. WH 5701	ZP_01085338	66	ZP_01085337	67
<i>Gloeobacter violaceus</i> PCC 7421	NP_926092	63	NP_926091	67
<i>Synechococcus</i> sp. RS9916	ZP_01472594	69	ZP_01472595	66
<i>Synechococcus</i> sp. RS9917	ZP_01079772	68	ZP_01079773	65
<i>Synechococcus</i> sp. CC9605	YP_381055	66	YP_381056	66
<i>Cianobium</i> sp. PCC 7001	EDY39806	64	EDY38361	64
<i>Prochlorococcus marinus</i> cepa MIT 9303	YP_001016795	63	YP_001016797	66
<i>Prochlorococcus marinus</i> cepa MIT9313	NP_895059	63	NP_895058	65
<i>Synechococcus</i> sp. CC9902	YP_377637	66	YP_377636	65
<i>Prochlorococcus marinus</i> cepa MIT 9301	YP_001090782	62	YP_001090783	62
<i>Synechococcus</i> sp. BL107	ZP_01469468	65	ZP_01469469	65
<i>Prochlorococcus marinus</i> cepa AS9601	YP_001008981	62	YP_001008982	61
<i>Prochlorococcus marinus</i> cepa MIT9312	YP_397029	62	YP_397030	61
<i>Prochlorococcus marinus</i> subsp. <i>pastoris</i> cepa CCMP1986	NP_892650	60	NP_892651	63
<i>Prochlorococcus marinus</i> cepa MIT 9211	YP_001550420	61	YP_001550421	63
<i>Cyanothece</i> sp PCC 7425	YP_002483683	59	-	
<i>Prochlorococcus marinus</i> cepa NATL2A	YP_293054	59	YP_293055	62
<i>Prochlorococcus marinus</i> cepa NATL1A	YP_001014415	59	YP_001014416	62
<i>Prochlorococcus marinus</i> subsp. <i>marinus</i> cepa CCMP1375	NP_874925	59	NP_874926	64
<i>Prochlorococcus marinus</i> cepa MIT 9515_05961	YP_001010912	57	YP_001010913	63
<i>Prochlorococcus marinus</i> cepa MIT 9215_06131	YP_001483814	59	YP_001483815	62
<i>Synechococcus</i> sp. RS9917	ZP_01080370	43	-	
<i>Synechococcus</i> marino no cultivado tipo A GOM 5D20			ABD96480	65

Variantes y genes biosintéticos de aldehído, alcano y alqueno

Los métodos y composiciones descritos en el presente documento incluyen, por ejemplo, genes biosintéticos de alcano o alqueno que tienen la secuencia de nucleótidos de SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33 ó 35, así como variantes de polinucleótido de las mismas. En algunos casos, el gen biosintético de alcano o alqueno codifica para uno o más de los motivos de aminoácido descritos en el presente documento. Por

ejemplo, el gen biosintético de alcanos o alquenos puede codificar para un polipéptido que comprende SEQ ID NO: 37, 38, 39, 41, 42, 43 ó 44. El gen biosintético de alcanos o alquenos también puede incluir un polipéptido que comprende SEQ ID NO: 40 y también uno cualquiera de SEQ ID NO: 37, 38 ó 39.

5 Los métodos y composiciones descritos en el presente documento también incluyen, por ejemplo, genes biosintéticos de aldehído que tienen la secuencia de nucleótidos de SEQ ID NO: 65, 67, 69, 71, 73, 75, 77, 79 u 81, así como variantes de polinucleótido de las mismas. En algunos casos, el gen biosintético de aldehído codifica para uno o más de los motivos de aminoácido descritos en el presente documento. Por ejemplo, el gen biosintético de aldehído puede codificar para un polipéptido que comprende SEQ ID NO: 54, 55, 56, 57, 58, 59, 60, 61, 62, 63 ó 64.

10 Las variantes pueden producirse de manera natural o crearse *in vitro*. En particular, tales variantes pueden crearse usando técnicas de modificación por ingeniería genética, tales como mutagénesis dirigida al sitio, mutagénesis química aleatoria, procedimientos de delección de exonucleasa III y técnicas de clonación convencionales. Alternativamente, tales variantes, fragmentos, análogos o derivados pueden crearse usando síntesis químicas o procedimientos de modificación.

15 En la técnica se conocen bien métodos de elaborar variantes. Estos incluyen procedimientos en los que secuencias de ácido nucleico obtenidas a partir de aislados naturales se modifican para generar ácidos nucleicos que codifican para polipéptidos que tienen características que potencian su valor en aplicaciones industriales o de laboratorio. En tales procedimientos, se genera y se caracteriza un gran número de secuencias de variante que tienen una o más diferencias de nucleótido con respecto a la secuencia obtenida a partir del aislado natural. Normalmente, estas diferencias de nucleótido dan como resultado cambios de aminoácido con respecto a los polipéptidos codificados por los ácidos nucleicos de los aislados naturales.

20 Por ejemplo, pueden crearse variantes usando PCR propensa a error (véase, por ejemplo, Leung *et al.*, Technique 1:11-15, 1989; y Caldwell *et al.*, PCR Methods Applic. 2:28-33, 1992). En PCR propensa a error, se realiza PCR en condiciones en las que la fidelidad de copiado de la ADN polimerasa es baja, de modo que se obtiene una elevada tasa de mutaciones puntuales a lo largo de toda la longitud del producto de PCR. Brevemente, en tales procedimientos, los ácidos nucleicos que van a mutagenizarse (por ejemplo, una secuencia de polinucleótido biosintético de aldehído o alcanos), se mezclan con cebadores de PCR, tampón de reacción, MgCl<sub>2</sub>, MnCl<sub>2</sub>, Taq polimerasa, y una concentración apropiada de dNTP para lograr una elevada tasa de mutación puntual a lo largo de toda la longitud del producto de PCR. Por ejemplo, la reacción puede realizarse usando 20 fmoles de ácido nucleico que va a mutagenizarse (por ejemplo, una secuencia de polinucleótido biosintético de aldehído o alcanos), 30 pmoles de cada cebador de PCR, un tampón de reacción que comprende KCl 50 mM, Tris HCl 10 mM (pH 8,3) y gelatina al 0,01%, MgCl<sub>2</sub> 7 mM, MnCl<sub>2</sub> 0,5 mM, 5 unidades de Taq polimerasa, dGTP 0,2 mM, dATP 0,2 mM, dCTP 1 mM y dTTP1 mM. Puede realizarse PCR durante 30 ciclos de 94°C durante 1 min, 45°C durante 1 min y 72°C durante 1 min. Sin embargo, se apreciará que estos parámetros pueden variarse según sea apropiado. Los ácidos nucleicos mutagenizados se clonan entonces en un vector apropiado y las actividades de los polipéptidos codificados por los ácidos nucleicos mutagenizados se evalúan.

35 También pueden crearse variantes usando mutagénesis dirigida a oligonucleótidos para generar mutaciones específicas del sitio en cualquier ADN clonado de interés. La mutagénesis de oligonucleótidos se describe en, por ejemplo, Reidhaar-Olson *et al.*, Science 241:53-57, 1988. Brevemente, en tales procedimientos una pluralidad de oligonucleótidos bicatenarios que portan una o más mutaciones que van a introducirse en el ADN clonado se sintetizan y se insertan en el ADN clonado que va a mutagenizarse (por ejemplo, una secuencia de polinucleótido biosintético de aldehído o alcanos). Los clones que contienen el ADN mutagenizado se recuperan, y las actividades de los polipéptidos que codifican se evalúan.

40 Otro método para generar variantes es PCR de ensamblaje. PCR de ensamblaje implica el ensamblaje de un producto de PCR a partir de una mezcla de pequeños fragmentos de ADN. Se produce un gran número de diferentes reacciones de PCR en paralelo en el mismo vial, cebando los productos de una reacción los productos de otra reacción. Se describe PCR de ensamblaje en, por ejemplo, en la patente estadounidense n.º 5.965.408.

45 Todavía otro método de generación de variantes es mutagénesis de PCR sexual. En la mutagénesis de PCR sexual, se produce recombinación homóloga forzada entre moléculas de ADN de secuencia de ADN diferente, pero muy relacionada, *in vitro* como resultado de fragmentación aleatoria de la molécula de ADN basándose en la homología de secuencia. Esto se sigue por fijación del cruzamiento mediante extensión de cebador en una reacción de PCR. Se describe mutagénesis de PCR sexual en, por ejemplo, Stemmer, PNAS, USA 91:10747-10751, 1994.

50 También pueden crearse variantes mediante mutagénesis *in vivo*. En algunas realizaciones, se generan mutaciones aleatorias en una secuencia de ácido nucleico propagando la secuencia en una cepa bacteriana, tal como una cepa de *E. coli*, que lleva mutaciones en una o más de las rutas de reparación de ADN. Tales cepas "mutágenas" tienen una tasa de mutación aleatoria mayor que la de una cepa silvestre. Propagar una secuencia de ADN (por ejemplo, una secuencia de polinucleótido biosintético de aldehído o alcanos) en una de estas cepas generará eventualmente mutaciones aleatorias dentro del ADN. Las cepas mutágenas adecuadas para su uso para mutagénesis *in vivo* se describen en, por ejemplo, la publicación PCT n.º WO 91/16427.

También pueden generarse variantes usando mutagénesis de casete. En la mutagénesis de casete, una pequeña región de una molécula de ADN bicatenario se sustituye por un "casete" de oligonucleótido sintético que difiere de la secuencia nativa. El oligonucleótido contiene a menudo una secuencia nativa completa y/o parcialmente aleatorizada.

5 También puede usarse mutagénesis conjunta recurrente para generar variantes. La mutagénesis conjunta recurrente es un algoritmo para la modificación por ingeniería genética de proteínas (es decir, mutagénesis de proteínas) desarrollada para producir poblaciones diversas de mutantes relacionados fenotípicamente cuyos miembros difieren en la secuencia de aminoácidos. Este método usa un mecanismo de retroalimentación para controlar rondas sucesivas de mutagénesis de casete combinatoria. Se describe mutagénesis conjunta recurrente en, por ejemplo, Arkin *et al.*, PNAS, USA 89:7811-7815, 1992.

10 En algunas realizaciones, se crean variantes usando mutagénesis de conjunto exponencial. La mutagénesis de conjunto exponencial es un procedimiento para generar bibliotecas combinatorias con un elevado porcentaje de mutantes únicos y funcionales, en la que pequeños grupos de residuos se aleatorizan en paralelo para identificar, en cada posición alterada, aminoácidos que conducen a proteínas funcionales. Se describe mutagénesis de conjunto exponencial en, por ejemplo, Delegrave *et al.*, Biotech. Res. 11:1548-1552, 1993. Se describen mutagénesis aleatoria y dirigida al sitio en, por ejemplo, Arnold, Curr. Opin. Biotech. 4:450-455, 1993.

15 En algunas realizaciones, se crean variantes usando procedimientos de intercambio en los que porciones de una pluralidad de ácidos nucleicos que codifican para distintos polipéptidos se fusionan juntas para crear secuencias de ácido nucleico quiméricas que codifican para polipéptidos quiméricos tal como se describe en, por ejemplo, las patentes estadounidenses n.ºs 5.965.408 y 5.939.250.

20 Las variantes de polinucleótido también incluyen análogos de ácido nucleico. Los análogos de ácido nucleico pueden modificarse en el residuo de base, residuo de azúcar, o esqueleto de fosfato para mejorar, por ejemplo, la estabilidad, hibridación o solubilidad del ácido nucleico. Las modificaciones en el resto de base incluyen desoxiuridina para desoxitimidina y 5-metil-2'-desoxicitidina o 5-bromo-2'-doxicitidina para desoxicitidina. Las modificaciones del residuo de azúcar incluyen modificación del 2'-hidroxilo del azúcar de ribosa para formar azúcares 2'-O-metilo o 2'-O-alilo. El esqueleto de fosfato de desoxirribosa puede modificarse para producir ácidos nucleicos de morfolino, en el que cada resto de base se une a un anillo morfolino de 6 miembros, o ácidos nucleicos peptídicos, en los que el esqueleto de desoxifosfato se reemplaza por un esqueleto pseudopeptídico y las cuatro bases se conservan. (Véase, por ejemplo, Summerton *et al.*, Antisense Nucleic Acid Drug Dev. (1997) 7:187-195; y Hyrup *et al.*, Bioorgan. Med. Chem. (1996) 4:5-23). Además, el esqueleto de desoxifosfato puede sustituirse por, por ejemplo, un esqueleto de fosforotioato o fosforoditioato, un esqueleto de fosforoamidita o fosfotriéster de alquilo.

25 Los polipéptidos biosintéticos de aldehído y alcanos Synpcc7942\_1594 (SEQ ID NO: 66) y Synpcc7942\_1593 (SEQ ID NO: 2) tienen homólogos en otras cianobacterias (se representan ejemplos no limitativos en la tabla 1). Por tanto, cualquier secuencia de polinucleótido que codifica para un homólogo enumerado en la tabla 1, o una variante del mismo, puede usarse como polinucleótido biosintético de aldehído o alcanos en los métodos descritos en el presente documento. Cada cianobacteria enumerada en la tabla 1 tiene copias de ambos genes. El nivel de identidad de secuencia de los productos génicos oscila entre el 61% y el 73% para Synpcc7942\_1594 (SEQ ID NO: 66) y entre el 43% y el 78% para Synpcc7942\_1593 (SEQ ID NO: 2).

30 Se enumeran homólogos adicionales del polipéptido biosintético de aldehído Synpcc7942\_1594 (SEQ ID NO: 66) en la figura 39, y puede usarse cualquier secuencia de polinucleótido que codifica para un homólogo enumerado en la figura 39, o una variante del mismo, como polinucleótido biosintético de aldehído en los métodos descritos en el presente documento. Se enumeran homólogos adicionales del polipéptido biosintético de alcanos Synpcc7942\_1593 (SEQ ID NO: 2) en la figura 38, y puede usarse cualquier secuencia de polinucleótido que codifica para un homólogo enumerado en la figura 38, o una variante del mismo, como polinucleótido biosintético de alcanos en los métodos descritos en el presente documento.

35 En determinados casos, un gen biosintético de aldehído, alcanos y/o alqueno se optimiza por codón para la expresión en una célula huésped particular. Por ejemplo, para la expresión en *E. coli*, pueden optimizarse uno o más codones tal como se describe en, por ejemplo, Grosjean *et al.*, Gene 18:199-209 (1982).

#### Variantes y polipéptidos biosintéticos de aldehído, alcanos y alqueno

40 Los métodos y composiciones descritos en el presente documento también incluyen polipéptidos biosintéticos de alcanos o alqueno que tienen la secuencia de aminoácidos de SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34 ó 36, así como variantes de polipéptido de la misma. En algunos casos, un polipéptido biosintético de alcanos o alqueno es uno que incluye uno o más de los motivos de aminoácido descritos en el presente documento. Por ejemplo, el polipéptido biosintético de alcanos o alqueno puede incluir la secuencia de aminoácidos de SEQ ID NO: 37, 38, 39, 41, 42, 43 ó 44. El polipéptido biosintético de alcanos o alqueno también puede incluir la secuencia de aminoácidos de SEQ ID NO: 40 y también una cualquiera de SEQ ID NO: 37, 38 ó 39.

45 Los métodos y composiciones descritos en el presente documento también incluyen polipéptidos biosintéticos de aldehído que tienen la secuencia de aminoácidos de SEQ ID NO: 66, 68, 70, 72, 74, 76, 78, 80 u 82, así como

variantes de polipéptido de la misma. En algunos casos, un polipéptido biosintético de aldehído es uno que incluye uno o más de los motivos de aminoácido descritos en el presente documento. Por ejemplo, el polipéptido biosintético de aldehído puede incluir la secuencia de aminoácidos de SEQ ID NO: 54, 55, 56, 57, 58, 59, 60, 61, 62, 63 ó 64.

5 Variantes de polipéptido biosintético de aldehído, alcano y alqueno pueden ser variantes en las que uno o más residuos de aminoácido se sustituyen por un residuo de aminoácido conservado o no conservado (preferiblemente un residuo de aminoácido conservado). Tal residuo de aminoácido sustituido puede o no ser uno codificado por el código genético.

10 Sustituciones conservadoras son aquellas que sustituyen un aminoácido dado en un polipéptido por otro aminoácido de características similares. Sustituciones conservadoras típicas son los siguientes reemplazos: reemplazo de un aminoácido alifático, tal como alanina, valina, leucina e isoleucina, por otro aminoácido alifático; reemplazo de una serina por una treonina o viceversa; reemplazo de un residuo ácido, tal como ácido aspártico y ácido glutámico, por otro residuo ácido; reemplazo de un residuo que porta un grupo amida, tal como asparagina y glutamina, por otro residuo que porta un grupo amida; intercambio de un residuo básico, tal como lisina y arginina, con otro residuo básico; y reemplazo de un residuo aromático, tal como fenilalanina y tirosina, por otro residuo aromático.

15 Otras variantes de polipéptido son aquellas en las que uno o más residuos de aminoácido incluyen un grupo sustituyente. Todavía otras variantes de polipéptido son aquellas en las que el polipéptido se asocia con otro compuesto, tal como un compuesto para aumentar la semivida del polipéptido (por ejemplo, polietilenglicol).

20 Variantes de polipéptido adicionales son aquellas en las que aminoácidos adicionales se fusionan con el polipéptido, tal como una secuencia líder, una secuencia secretora, una secuencia de proproteína, o una secuencia que facilita la purificación, el enriquecimiento o la estabilización del polipéptido.

En algunos casos, una variante de polipéptido biosintético de alcano o alqueno conserva la misma función biológica que un polipéptido que tiene la secuencia de aminoácidos de SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34 ó 36 (por ejemplo, conserva actividad biosintética de alcano o alqueno) y tiene una secuencia de aminoácidos sustancialmente idéntica a la misma.

25 En otros casos, las variantes de polipéptido biosintético de alcano o alqueno tienen al menos aproximadamente el 50%, al menos aproximadamente el 55%, al menos aproximadamente el 60%, al menos aproximadamente el 65%, al menos aproximadamente el 70%, al menos aproximadamente el 75%, al menos aproximadamente el 80%, al menos aproximadamente el 85%, al menos aproximadamente el 90%, al menos aproximadamente el 95%, o más de aproximadamente el 95% de homología con la secuencia de aminoácidos de SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34 ó 36. En otra realización, las variantes de polipéptido incluyen un fragmento que comprende al menos aproximadamente 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100 ó 150 aminoácidos consecutivos del mismo.

30 En algunos casos, una variante de polipéptido biosintético de aldehído conserva la misma función biológica que un polipéptido que tiene la secuencia de aminoácidos de SEQ ID NO: 66, 68, 70, 72, 74, 76, 78, 80 u 82 (por ejemplo, conserva actividad biosintética de aldehído) y tiene una secuencia de aminoácidos sustancialmente idéntica a la misma.

35 En aún otros casos, las variantes de polipéptido biosintético de aldehído tienen al menos aproximadamente el 50%, al menos aproximadamente el 55%, al menos aproximadamente el 60%, al menos aproximadamente el 65%, al menos aproximadamente el 70%, al menos aproximadamente el 75%, al menos aproximadamente el 80%, al menos aproximadamente el 85%, al menos aproximadamente el 90%, al menos aproximadamente el 95%, o más de aproximadamente el 95% de homología con la secuencia de aminoácidos de SEQ ID NO: 66, 68, 70, 72, 74, 76, 78, 80 u 82. En otra realización, las variantes de polipéptido incluyen un fragmento que comprende al menos aproximadamente 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100 ó 150 aminoácidos consecutivos del mismo.

40 Las variantes de polipéptido o fragmentos de las mismas pueden obtenerse aislando ácidos nucleicos que las codifican usando técnicas descritas en el presente documento o expresando ácidos nucleicos sintéticos que las codifican. Alternativamente, pueden obtenerse variantes de polipéptido o fragmentos de las mismas mediante procedimientos de enriquecimiento o purificación bioquímicos. La secuencia de variantes de polipéptido o fragmentos puede determinarse mediante digestión proteolítica, electroforesis en gel y/o microsecuenciación. La secuencia de las variantes de polipéptido biosintético de alcano o alqueno o fragmentos pueden entonces compararse con la secuencia de aminoácidos de SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34 ó 36 usando cualquiera de los programas descritos en el presente documento. La secuencia de las variantes de polipéptido biosintético de aldehído o fragmentos puede compararse con la secuencia de aminoácidos de SEQ ID NO: 66, 68, 70, 72, 74, 76, 78, 80 u 82 usando cualquiera de los programas descritos en el presente documento.

45 Las variantes de polipéptido y fragmentos de las mismas pueden someterse a ensayo para determinar actividad de producción de aldehído, alcohol graso, alcano y/o alqueno usando métodos de rutina. Por ejemplo, las variantes de polipéptido o fragmento pueden ponerse en contacto con un sustrato (por ejemplo, un sustrato de derivado de ácido graso u otro sustrato descrito en el presente documento) en condiciones que permiten que la variante de polipéptido funcione. Una disminución en el nivel del sustrato o un aumento en el nivel de un aldehído, alcano o alqueno pueden

medirse para determinar la actividad de producción de aldehído, alcohol graso, alcano o alqueno, respectivamente.

#### Anticuerpos de polipéptido biosintético antialdehído, antialcohol graso, antialcano y antialqueno

Los polipéptidos biosintéticos de aldehído, alcohol graso, alcano y alqueno descritos en el presente documento también pueden usarse para producir anticuerpos dirigidos contra polipéptidos biosintéticos de aldehído, alcohol graso, alcano y alqueno. Tales anticuerpos pueden usarse, por ejemplo, para detectar la expresión de un polipéptido biosintético de aldehído, alcohol graso, alcano o alqueno usando métodos conocidos en la técnica. El anticuerpo puede ser, por ejemplo, un anticuerpo policlonal; un anticuerpo monoclonal o fragmento de unión a antígeno del mismo; un anticuerpo modificado tal como un anticuerpo quimérico, anticuerpo reconfigurado, anticuerpo humanizado, o fragmento del mismo (por ejemplo, Fab', Fab, F(ab')<sub>2</sub>); o un anticuerpo biosintético, por ejemplo, un anticuerpo de cadena sencilla, anticuerpo de dominio único (DAB), Fv, Fv de cadena sencilla (scFv), o similares.

Se describen métodos de elaboración y uso de anticuerpos policlonales y monoclonales, por ejemplo, en Harlow *et al.*, *Using Antibodies: A Laboratory Manual: Portable Protocol I*. Cold Spring Harbor Laboratory (1 de diciembre de 1998). En la técnica se conocen métodos para elaborar anticuerpos modificados y fragmentos de anticuerpo (por ejemplo, anticuerpos quiméricos, anticuerpos reconfigurados, anticuerpos humanizados, o fragmentos de los mismos, por ejemplo, fragmentos Fab', Fab, F(ab')<sub>2</sub>); o anticuerpos biosintéticos (por ejemplo, anticuerpos de cadena sencilla, anticuerpos de dominio único (DAB), Fv, Fv de cadena sencilla (scFv), y similares), y pueden encontrarse, por ejemplo, en Zola, *Monoclonal Antibodies: Preparation and Use of Monoclonal Antibodies and Engineered Antibody Derivatives*, Springer Verlag (15 de diciembre de 2000; 1ª edición).

#### Sustratos

Las composiciones y métodos descritos en el presente documento pueden usarse para producir aldehídos, alcoholes grasos, alcanos y/o alquenos a partir de un sustrato apropiado. Sin querer limitarse por ninguna teoría particular, se cree que los polipéptidos biosintéticos de alcano o alqueno descritos en el presente documento producen alcanos o alquenos a partir de sustratos por medio de un mecanismo de descarbonilación. En algunos casos, el sustrato es un derivado de ácido graso, por ejemplo, un aldehído graso, y puede producirse un alcano que tiene patrones de ramificación particulares y longitud de cadena de carbono a partir de un derivado de ácido graso, por ejemplo, un aldehído graso, que tiene esas características particulares. En otros casos, el sustrato es un derivado de ácido graso insaturado, por ejemplo, un aldehído graso insaturado, y puede producirse un alqueno que tiene patrones de ramificación particulares y longitud de cadena de carbono a partir de un derivado de ácido graso insaturado, por ejemplo, un aldehído graso insaturado, que tiene esas características particulares.

Sin querer limitarse por ninguna teoría particular, se cree que los polipéptidos biosintéticos de aldehído descritos en el presente documento producen aldehídos a partir de sustratos por medio de un mecanismo de reducción. En determinados casos, el sustrato es una acil-ACP.

Sin querer limitarse por ninguna teoría particular, se cree que los alcoholes grasos descritos en el presente documento se producen a partir de sustratos por medio de un mecanismo de reducción. En determinados casos, el sustrato es un aldehído graso.

Por consiguiente, cada etapa dentro de una ruta biosintética que conduce a la producción de estos sustratos puede modificarse para producir o sobreproducir el sustrato de interés. Por ejemplo, genes conocidos implicados en la ruta biosintética de ácido graso, la ruta de aldehído graso y la ruta de alcohol graso pueden expresarse, sobreexpresarse o atenuarse en células huésped para producir un sustrato deseado (véase, por ejemplo, el documento PCT/US08/058788). Se proporcionan genes a modo de ejemplo en la figura 40.

#### Síntesis de sustratos

La ácido graso sintasa (FAS) es un grupo de polipéptidos que catalizan el inicio y la elongación de cadenas de acilo (Marrakchi *et al.*, *Biochemical Society*, 30:1050-1055, 2002). La proteína portadora de acilo (ACP) junto con las enzimas en la ruta de FAS controla la longitud, el grado de saturación y la ramificación de los derivados de ácidos grasos producidos. La ruta biosintética de ácido graso implica los precursores acetil-CoA y malonil-CoA. Las etapas en esta ruta se catalizan mediante enzimas de las familias génicas de acetil-CoA carboxilasa (acc) y biosíntesis de ácido graso (fab) (véase, por ejemplo, Heath *et al.*, *Prog. Lipid Res.* 40(6):467-97 (2001)).

Pueden modificarse por ingeniería genética células huésped para expresar sustratos de derivado de ácido graso expresando o sobreexpresando de manera recombinante genes de acetil-CoA y/o malonil-CoA sintasa. Por ejemplo, para aumentar la producción de acetil-CoA, uno o más de los siguientes genes pueden expresarse en una célula huésped: pdh, panK, aceEF (que codifica para el componente deshidrogenasa E1p y el componente dihidrolipoamida aciltransferasa E2p de los complejos de piruvato y 2-oxoglutarato deshidrogenasa), fabH, fabD, fabG, acpP y fabF. Los números de registro de GenBank a modo de ejemplo para estos genes son: pdh (BAB34380, AAC73227, AAC73226), panK (también conocido como coaA, AAC76952), aceEF (AAC73227, AAC73226), fabH (AAC74175), fabD (AAC74176), fabG (AAC74177), acpP (AAC74178), fabF (AAC74179). Además, los niveles de expresión de fadE, gpsA, ldhA, pflb, adhE, pta, poxB, ackA y/o ackB pueden atenuarse o inactivarse en una célula huésped modificada por ingeniería genética mediante transformación con plásmidos replicativos o no replicativos

condicionalmente que contienen mutaciones anuladoras o de delección de los genes correspondientes o sustituyendo secuencias de promotor o de potenciador. Números de registro de GenBank a modo de ejemplo para estos genes son: *fadE* (AAC73325), *gspA* (AAC76632), *ldhA* (AAC74462), *pflB* (AAC73989), *adhE* (AAC74323), *pta* (AAC75357), *poxB* (AAC73958), *ackA* (AAC75356) y *ackB* (BAB81430). Las células huésped resultantes tendrán niveles de producción de acetil-CoA aumentados cuando se hacen crecer en un ambiente apropiado.

5 La sobreexpresión de malonil-CoA puede efectuarse introduciendo *accABCD* (por ejemplo, número de registro AAC73296, EC 6.4.1.2) en una célula huésped. Los ácidos grasos pueden sobreexpresarse además en células huésped introduciendo en la célula huésped una secuencia de ADN que codifica para una lipasa (por ejemplo, números de registro CAA89087, CAA98876).

10 Además, la inhibición de *PlsB* puede conducir a un aumento en los niveles de acil-ACP de cadena larga, que inhibirá etapas tempranas en la ruta (por ejemplo, *accABCD*, *fabH* y *fabI*). La mutación de *plsB* (por ejemplo, número de registro AAC77011) D311E puede usarse para aumentar la cantidad de acil-CoA disponible.

15 Además, una célula huésped puede modificarse por ingeniería genética para sobreexpresar un gen *sfa* (supresor de *fabA*, por ejemplo, número de registro AAN79592) para aumentar la producción de ácidos grasos monoinsaturados (Rock *et al.*, J. Bacteriology 178:5382-5387, 1996).

En algunos casos, pueden modificarse por ingeniería genética células huésped para expresar, sobreexpresar o atenuar la expresión de una tioesterasa para aumentar la producción de sustrato de ácido graso. La longitud de cadena de un sustrato de ácido graso se controla mediante tioesterasa. En algunos casos, un gen *tes* o *fat* puede sobreexpresarse. En otros casos, pueden producirse ácidos grasos C<sub>10</sub> atenuando tioesterasa C<sub>18</sub> (por ejemplo, números de registro AAC73596 y P0ADA1), que usa C<sub>18:1</sub>-ACP, y que expresa tioesterasa C<sub>10</sub> (por ejemplo, número de registro Q39513), que usa C<sub>10</sub>-ACP. Esto da como resultado una población relativamente homogénea de ácidos grasos que tienen una longitud de cadena de carbono de 10. En aún otros casos, pueden producirse ácidos grasos C<sub>14</sub> atenuando tioesterasas endógenas que producen ácidos no grasos C<sub>14</sub> y que expresan las tioesterasas, que usan C<sub>14</sub>-ACP (por ejemplo, número de registro Q39473). En algunas situaciones, pueden producirse ácidos grasos C<sub>12</sub> expresando tioesterasas que usan C<sub>12</sub>-ACP (por ejemplo, número de registro Q41635) y atenuando tioesterasas que producen ácidos no grasos C<sub>12</sub>. La sobreproducción de acetil-CoA, malonil-CoA y ácido graso puede verificarse usando métodos conocidos en la técnica, por ejemplo, usando precursores radiactivos, HPLC, y CG-EM posterior a la lisis celular. En la tabla 2 se enumeran ejemplos no limitativos de tioesterasas que pueden usarse en los métodos descritos en el presente documento.

30 Tabla 2: Tioesterasas

Número de registro	Organismo fuente	Gen	Producto preferencial producido
AAC73596	<i>E. coli</i>	<i>tesA</i> sin secuencia líder	C <sub>18:1</sub>
AAC73555	<i>E. coli</i>	<i>tesB</i>	
Q41635, AAA34215	<i>Umbellularia californica</i>	<i>fatB</i>	C <sub>12:0</sub>
Q39513; AAC49269	<i>Cuphea hookeriana</i>	<i>fatB2</i>	C <sub>8:0</sub> - C <sub>10:0</sub>
AAC49269; AAC72881	<i>Cuphea hookeriana</i>	<i>fatB3</i>	C <sub>14:0</sub> - C <sub>16:0</sub>
Q39473, AAC49151	<i>Cinnamomum camphorum</i>	<i>fatB</i>	C <sub>14:0</sub>
CAA85388	<i>Arabidopsis thaliana</i>	<i>fatB</i> [M141T]*	C <sub>16:1</sub>
NP 189147; NP 193041	<i>Arabidopsis thaliana</i>	<i>fatA</i>	C <sub>18:1</sub>
CAC39106	<i>Bradyrhizobium japonicum</i>	<i>fatA</i>	C <sub>18:1</sub>
AAC72883	<i>Cuphea hookeriana</i>	<i>fatA</i>	C <sub>18:1</sub>
AAL79361	<i>Helianthus annuus</i>	<i>fatA1</i>	

\* Mayer *et al.*, BMC Plant Biology 7:1-11, 2007

Formación de aldehídos, alcoholes grasos, alcanos y alquenos ramificados

35 Pueden producirse aldehídos, alcoholes grasos, alcanos y alquenos que contienen puntos de ramificación usando derivados ramificados de ácidos grasos como sustratos. Por ejemplo, aunque *E. coli* produce de manera natural derivados de cadena lineal de ácidos grasos (sFA), *E. coli* también puede modificarse por ingeniería genética para producir derivados de cadena ramificada de ácidos grasos (brFA) introduciendo y expresando o sobreexpresando genes que proporcionan precursores ramificados en *E. coli* (por ejemplo, familias de genes *bkd*, *ilv*, *icm* y *fab*). Además, una célula huésped puede modificarse por ingeniería genética para expresar o sobreexpresar genes que codifican para proteínas para la elongación de brFA (por ejemplo, ACP, FabF, etc.) y/o para delecionar o atenuar los genes de célula huésped correspondientes que conducen normalmente a sFA.

40 La primera etapa en la formación de brFA es la producción de los ácido  $\alpha$ -ceto correspondientes mediante una aminoácido aminotransferasa de cadena ramificada. Las células huésped pueden incluir de manera endógena genes que codifican para tales enzimas o tales genes pueden introducirse de manera recombinante. *E. coli*, por ejemplo,

expresa de manera endógena una enzima de este tipo, IlvE (EC 2.6.1.42; registro de GenBank YP\_026247). En algunas células huésped, puede no expresarse una aminoácido aminotransferasa de cadena ramificada heteróloga. Sin embargo, puede introducirse y expresarse de manera recombinante, si no es endógena, *E. coli* IlvE o cualquier otra aminoácido aminotransferasa de cadena ramificada (por ejemplo, IlvE de *Lactococcus lactis* (registro de GenBank AAF34406), IlvE de *Pseudomonas putida* (registro de GenBank NP\_745648), o IlvE de *Streptomyces coelicolor* (registro de GenBank NP\_629657)).

La segunda etapa es la descarboxilación oxidativa de los ácidos  $\alpha$ -ceto con la acil-CoA de cadena ramificada correspondiente. Esta reacción puede catalizarse mediante un complejo de ácido  $\alpha$ -ceto deshidrogenasa de cadena ramificada (bkd; EC 1.2.4.4.) (Denoya *et al.*, J. Bacteriol. 177:3504, 1995), que consiste en subunidades de E1 $\alpha$ / $\beta$  (descarboxilasa), E2 (dihidrolipoil transacilasa) y E3 (dihidrolipoil deshidrogenasa). Estos complejos de ácido  $\alpha$ -ceto deshidrogenasa de cadena ramificada son similares a los complejos de piruvato y  $\alpha$ -cetoglutarato deshidrogenasa. Cualquier microorganismo que posee brFA y/o crece sobre aminoácidos de cadena ramificada puede usarse como fuente para aislar genes bkd para la expresión en células huésped, por ejemplo, *E. coli*. Además, *E. coli* tiene el componente E3 como parte de su complejo de piruvato deshidrogenasa (lpd, EC 1.8.1.4, registro de GenBank NP\_414658). Por tanto, puede ser suficiente expresar sólo los genes E1 $\alpha$ / $\beta$  y E2 bkd. La tabla 3 enumera ejemplos no limitativos de genes bkd de varios microorganismos que pueden introducirse y expresarse de manera recombinante en una célula huésped para proporcionar precursores de acil-CoA de cadena ramificada.

Tabla 3: Genes bkd de microorganismos seleccionados

Organismo	Gen	n.º de registro de GenBank
<i>Streptomyces coelicolor</i>	bkdA1 (E1 $\alpha$ )	NP_628006
	bkdB1 (E1 $\beta$ )	NP_628005
	bkdC1 (E2)	NP_638004
<i>Streptomyces coelicolor</i>	bkdA2 (E1 $\alpha$ )	NP_733618
	bkdB2 (E1 $\beta$ )	NP_628019
	bkdC2 (E2)	NP_628018
<i>Streptomyces avermitilis</i>	bkdA (E1a)	BAC72074
	bkdB (E1b)	BAC72075
	bkdC (E2)	BAC72076
<i>Streptomyces avermitilis</i>	bkdF (E1 $\alpha$ )	BAC72088
	bkdG (E1 $\beta$ )	BAC72089
	bkdH (E2)	BAC72090
<i>Bacillus subtilis</i>	bkdAA (E1 $\alpha$ )	NP_390288
	bkdAB (E1 $\beta$ )	NP_390288
	bkdB (E2)	NP_390288
<i>Pseudomonas putida</i>	bkdA1 (E1 $\alpha$ )	AAA65614
	bkdA2 (E1 $\beta$ )	AAA65615
	bkdC (E2)	AAA65617

En otro ejemplo, puede elaborarse isobutiril-CoA en una célula huésped, por ejemplo en *E. coli*, mediante la coexpresión de una crotonil-CoA reductasa (Ccr, EC 1.6.5.5, 1.1.1.1) e isobutiril-CoA mutasa (subunidad grande IcmA, EC 5.4.99.2; subunidad pequeña IcmB, EC 5.4.99.2) (Han y Reynolds, J. Bacteriol. 179:5157, 1997). Crotonil-CoA es un producto intermedio en la biosíntesis de ácido graso en *E. coli* y otros microorganismos. En la tabla 4 se enumeran ejemplos no limitativos de genes ccr e icm de microorganismos seleccionados.

Tabla 4: Genes ccr e icm de microorganismos seleccionados

Organismo	Gen	n.º de registro de GenBank
<i>Streptomyces coelicolor</i>	Ccr	NP_630556
	icmA	NP_629554
	icmB	NP_630904
<i>Streptomyces cinnamonensis</i>	ccr	AAD53915
	icmA	AAC08713
	icmB	AJ246005

Además de la expresión de los genes bkd, el inicio de la biosíntesis de brFA usa proteína portadora de  $\beta$ -cetoacil-acil sintasa III (FabH, EC 2.3.1.41) con especificidad por acil-CoA de cadena ramificada (Li *et al.*, J. Bacteriol. 187:3795-3799, 2005). En la tabla se enumeran 5 ejemplos no limitativos de tales enzimas FabH. Pueden expresarse genes fabH que están implicados en la biosíntesis de ácido graso de cualquier microorganismo que contiene brFA en una célula huésped. Las enzimas Bkd y FabH de células huésped que no producen de manera natural brFA pueden no soportar la producción de brFA. Por tanto, bkd y fabH pueden expresarse de manera recombinante. Pueden

5 insertarse vectores que contienen los genes *bkd* y *fabH* en una célula huésped de este tipo. De manera similar, el nivel endógeno de producción de Bkd y FabH puede no ser suficiente para producir brFA. En este caso, pueden sobreexpresarse. Además, otros componentes de la ruta de biosíntesis de ácido graso pueden expresarse o sobreexpresarse, tales como proteínas portadoras de acilo (ACP) y proteína portadora  $\beta$ -cetoacil-acil sintasa II (*fabF*, EC 2.3.1.41) (en la tabla 5 se enumeran ejemplos no limitativos de candidatos). Además de expresar estos genes, algunos genes en la biosíntesis endógena de la ruta de ácido graso pueden atenuarse en la célula huésped (por ejemplo, los genes de *E. coli* *fabH* (n.º de registro de GenBank NP\_415609) y/o *fabF* (n.º de registro de GenBank NP\_415613)).

Tabla 5: Genes FabH, ACP y *fabF* de microorganismos seleccionados con brFA

Organismo	Gen	n.º de registro de GenBank
<i>Streptomyces coelicolor</i>	<i>fabH1</i> ACP <i>fabF</i>	NP_626634 NP_626635 NP_626636
<i>Streptomyces avermitilis</i>	<i>fabH3</i> <i>fabC3</i> (ACP) <i>fabF</i>	NP_823466 NP_823467 NP_823468
<i>Bacillus subtilis</i>	<i>fabH_A</i> <i>fabH_B</i> ACP <i>fabF</i>	NP_389015 NP_388898 NP_389474 NP_389016
<i>Stenotrophomonas maltophilia</i>	SmaIDRAFT_0818 (FabH) SmaIDRAFT_0821 (ACP) SmaIDRAFT_0822 (FabF)	ZP_01643059 ZP_01643063 ZP_01643064
<i>Legionella pneumophila</i>	FabH ACP <i>fabF</i>	YP_123672 YP_123675 YP_123676

10

#### Formación de aldehídos, alcoholes grasos, alcanos y alquenos cíclicos

Pueden producirse aldehídos, alcoholes grasos, alcanos y alquenos cíclicos usando derivados de ácidos grasos cíclicos como sustratos. Para producir sustratos de derivado de ácido graso cíclicos, pueden introducirse genes que proporcionan precursores cíclicos (por ejemplo, las familias de genes *ans*, *chc* y *plm*) en la célula huésped y expresarse para permitir el inicio de biosíntesis de ácido graso de precursores cíclicos. Por ejemplo, para convertir una célula huésped, tal como *E. coli*, en una que puede sintetizar derivados de ácidos grasos  $\omega$ -cíclicos (*cyFA*), un gen que proporciona el precursor cíclico ciclohexilcarbonil-CoA (CHC-CoA) (Cropp *et al.*, Nature Biotech. 18:980-983, 2000) puede introducirse y expresarse en la célula huésped. Los ejemplos no limitativos de genes que proporcionan CHC-CoA en *E. coli* incluyen: *ansJ*, *ansK*, *ansL*, *chcA* y *ansM* de la agrupación de genes *ansatrienina* de *Streptomyces collinus* (Chen *et al.*, Eur. J. Biochem. 261: 98-107, 1999) o *plmJ*, *plmK*, *plmL*, *chcA* y *plmM* de la agrupación de genes de foslactomicina B de *Streptomyces* sp. HK803 (Palaniappan *et al.*, J Biol. Chem. 278:35552-35557, 2003) junto con el gen *chcB* (Patton *et al.*, Biochem. 39:7595-7604, 2000) de *S. collinus*, *S. avermitilis* o *S. coelicolor* (véase la tabla 6). Los genes enumerados en la tabla 5 pueden expresarse entonces para permitir el inicio y la elongación de ácidos grasos  $\omega$ -cíclicos. Alternativamente, los genes homólogos pueden aislarse a partir de microorganismos que producen *cyFA* y expresados en una célula huésped (por ejemplo, *E. coli*).

25

Tabla 6: Genes para la síntesis de CHC-CoA

Organismo	Gen	n.º de registro de GenBank
<i>Streptomyces collinus</i>	<i>ansJK</i> <i>ansL</i> <i>chcA</i> <i>ansM</i> <i>chcB</i>	U72144*  AF268489
<i>Streptomyces</i> sp. HK803	<i>pmlJK</i> <i>pmlL</i> <i>chcA</i> <i>pmlM</i>	AAQ84158 AAQ84159 AAQ84160 AAQ84161
<i>Streptomyces coelicolor</i>	<i>chcB/caiD</i>	NP_629292
<i>Streptomyces avermitilis</i>	<i>chcB/caiD</i>	NP_629292

\*Sólo se anota *chcA* en la entrada de GenBank U72144, *ansJKLM* son según Chen *et al.* (Eur. J. Biochem. 261:98-107, 1999).

Los genes enumerados en la tabla 5 (fabH, ACP y fabF) permiten el inicio y la elongación de derivados de ácidos grasos  $\omega$ -cíclicos porque tienen amplia especificidad de sustrato. Si la coexpresión de cualquiera de estos genes con los genes enumerados en la tabla 6 no produce cyFA, entonces pueden aislarse homólogos de fabH, ACP y/ o fabF de microorganismos que producen cyFA (por ejemplo, los enumerados en la tabla 7) (por ejemplo, usando cebadores de PCR de degeneración o sondas de secuencia de ADN heterólogas) y coexpresarse.

Tabla 7: Ejemplos no limitativos de microorganismos que contienen ácidos grasos  $\omega$ -cíclicos

Organismo	Referencia
<i>Curtobacterium pusillum</i>	ATCC19096
<i>Alicyclobacillus acidoterrestris</i>	ATCC49025
<i>Alicyclobacillus acidocaldarius</i>	ATCC27009
<i>Alicyclobacillus cycloheptanicus</i> *	Moore, J. Org. Chem. 62: págs. 2173, 1997.

\*Usa cicloheptilcarbonil-CoA y no ciclohexilcarbonil-CoA como precursor para la biosíntesis de cyFA.

#### Niveles de saturación de aldehído, alcohol graso y alqueno

El grado de saturación en derivados de ácidos grasos puede controlarse regulando el grado de saturación de productos intermedios de derivado de ácido graso. Las familias de genes sfa, gns y fab pueden expresarse o sobreexpresarse para controlar la saturación de ácidos grasos. La figura 40 enumera ejemplos no limitativos de genes en estas familias de genes que pueden usarse en los métodos y células huésped descritos en el presente documento.

Las células huésped pueden modificarse por ingeniería genética para producir ácidos grasos insaturados modificando por ingeniería genética la célula huésped para sobreexpresar fabB o haciendo crecer la célula huésped a bajas temperaturas (por ejemplo, menos de 37°C). FabB tiene preferencia sobre cis- $\delta$ 3deconoil-ACP y da como resultado producción de ácido graso insaturado en *E. coli*. La sobreexpresión de fabB da como resultado la producción de un porcentaje significativo de ácidos grasos insaturados (de Mendoza *et al.*, J. Biol. Chem. 258:2098-2101, 1983). El gen fabB puede insertarse en y expresarse en células huésped que no tienen de manera natural el gen. Estos derivados de ácidos grasos insaturados pueden usarse entonces como productos intermedios en células huésped que se modifican por ingeniería genética para producir derivados de ácidos grasos, tales como aldehídos grasos, alcoholes grasos o alquenos.

En otros casos, un represor de la biosíntesis de ácido graso, por ejemplo, fabR (registro de GenBank NP\_418398), puede delecionarse, lo que dará como resultado una producción aumentada de ácido graso insaturado en *E. coli* (Zhang *et al.*, J. Biol. Chem. 277:15558, 2002). Pueden realizarse delecciones similares en otras células huésped. Puede lograrse un aumento adicional de derivados de ácidos grasos insaturados, por ejemplo, sobreexpresando fabM (trans-2, cis-3-deconoil-ACP isomerasa, registro de GenBank DAA05501) y expresión controlada de fabK (trans-2-enoil-ACP reductasa II, registro de GenBank NP\_357969) de *Streptococcus pneumoniae* (Marrakchi *et al.*, J. Biol. Chem. 277: 44809, 2002), mientras se deleciona *E. coli* fabI (trans-2-enoil-ACP reductasa, registro de GenBank NP\_415804). En algunos ejemplos, el gen fabF endógeno puede atenuarse, aumentado por tanto el porcentaje de palmitoleato (C16:1) producido.

#### Otros sustratos

Otros sustratos que pueden usarse para producir aldehídos, alcoholes grasos, alcanos y alquenos en los métodos descritos en el presente documento son acil-ACP, acil-CoA, un aldehído graso o un alcohol graso, que se describen en, por ejemplo, el documento PCT/US08/058788. En la figura 40 se enumeran genes a modo de ejemplo que pueden alterarse para expresar o sobreexpresar estos sustratos en células huésped. Otros genes a modo de ejemplo se describen en el documento PCT/US08/058788.

#### Modificación por ingeniería genética de células huésped para producir aldehídos, alcoholes grasos, alcanos y alquenos

Pueden usarse diversas células huésped para producir aldehídos, alcoholes grasos, alcanos y/o alquenos, tal como se describe en el presente documento. Una célula huésped puede ser cualquier célula procariota o eucariota. Por ejemplo, un polipéptido descrito en el presente documento puede expresarse en células bacterianas (tales como *E. coli*), células de insecto, células de levadura o mamífero (tales como células de ovario de hámster chino (CHO), células COS, células VERO, células BHK, células HeLa, células Cv1, células MDCK, células 293, células 3T3 o células PC12). Otras células huésped a modo de ejemplo incluyen células de los miembros del género *Escherichia*, *Bacillus*, *Lactobacillus*, *Rhodococcus*, *Pseudomonas*, *Aspergillus*, *Trichoderma*, *Neurospora*, *Fusarium*, *Humicola*, *Rhizomucor*, *Kluyveromyces*, *Pichia*, *Mucor*, *Myceliophthora*, *Penicillium*, *Phanerochaete*, *Pleurotus*, *Trametes*, *Chrysosporium*, *Saccharomyces*, *Schizosaccharomyces*, *Yarrowia* o *Streptomyces*. Aún otras células huésped a modo de ejemplo pueden ser una célula de *Bacillus lentus*, una célula de *Bacillus brevis*, una célula de *Bacillus stearothermophilus*, una célula de *Bacillus licheniformis*, una célula de *Bacillus alkalophilus*, una célula de *Bacillus coagulans*, una célula de *Bacillus circulans*, una célula de *Bacillus pumilis*, una célula de *Bacillus thuringiensis*, una

célula de *Bacillus clausii*, una célula de *Bacillus megaterium*, una célula de *Bacillus subtilis*, una célula de *Bacillus amyloliquefaciens*, una célula de *Trichoderma koningii*, una célula de *Trichoderma viride*, una célula de *Trichoderma reesei*, una célula de *Trichoderma longibrachiatum*, una célula de *Aspergillus awamori*, una célula de *Aspergillus fumigates*, una célula de *Aspergillus foetidus*, una célula de *Aspergillus nidulans*, una célula de *Aspergillus niger*, una célula de *Aspergillus oryzae*, una célula de *Humicola insolens*, una célula de *Humicola lanuginosa*, una célula de *Rhizomucor miehei*, una célula de *Mucor michei*, una célula de *Streptomyces lividans*, una célula de *Streptomyces murinus* o una célula de *Actinomycetes*.

Otros ejemplos no limitativos de células huésped son los enumerados en la tabla 1.

En una realización preferida, la célula huésped es una célula de *E. coli*. En una realización más preferida, la célula huésped es de las cepas de *E. coli* B, C, K o W. Otras células huésped adecuadas las conocen los expertos en la técnica.

Pueden usarse diversos métodos bien conocidos en la técnica para modificar por ingeniería genética células huésped para producir aldehídos, alcoholes grasos, alcanos y/o alquenos. Los métodos incluyen el uso de vectores, preferiblemente vectores de expresión, que contienen un ácido nucleico que codifica para un polipéptido biosintético de aldehído, alcohol graso, alcano y/o alqueno descrito en el presente documento, o una variante de polipéptido o fragmento del mismo. Tal como se usa en el presente documento, el término "vector" se refiere a una molécula de ácido nucleico que puede transportar otro ácido nucleico al que se ha unido. Un tipo de vector es un "plásmido", que se refiere a un bucle de ADN bicatenario circular en el que pueden ligarse segmentos de ADN adicionales. Otro tipo de vector es un vector viral, en el que puede ligarse segmentos de ADN adicionales en el genoma viral. Determinados vectores pueden realizar replicación autónoma en una célula huésped en la que se introducen (por ejemplo, vectores bacterianos que tienen un origen bacteriano de replicación y vectores de mamíferos episomales). Otros vectores (por ejemplo, vectores de mamíferos no episomales) se integran en el genoma de una célula huésped tras la introducción en la célula huésped y se replican de ese modo junto con el genoma huésped. Además, determinados vectores, tales como vectores de expresión, pueden dirigir la expresión de genes a los que se unen de manera operativa. En general, los vectores de expresión usados en técnicas de ADN recombinante están a menudo en forma de plásmidos. Sin embargo, también pueden usarse otras formas de vectores de expresión, tales como vectores virales (por ejemplo, retrovirus de replicación defectuosa, adenovirus y virus adenoasociados).

Los vectores de expresión recombinante descritos en el presente documento incluyen un ácido nucleico descrito en el presente documento en una forma adecuada para la expresión del ácido nucleico en una célula huésped. Los vectores de expresión recombinantes pueden incluir una o más secuencias de control, seleccionadas basándose en la célula huésped que va a usarse para la expresión. La secuencia de control está operativamente unida a la secuencia de ácido nucleico que va a expresarse. Tales secuencias de control se describen, por ejemplo, en Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, Calif. (1990). Las secuencias de control incluyen aquellas que dirigen la expresión constitutiva de una secuencia de nucleótidos en muchos tipos de células huésped y aquellas que dirigen la expresión de la secuencia de nucleótidos sólo en determinadas células huésped (por ejemplo, secuencias reguladoras específicas de tejido). Los expertos en la técnica apreciarán que el diseño del vector de expresión puede depender de factores tales como la elección de la célula huésped que va a transformarse, el nivel de expresión de proteína deseado, etc. Los vectores de expresión descritos en el presente documento pueden introducirse en células huésped para producir polipéptidos, incluyendo polipéptidos de fusión, codificados por los ácidos nucleicos tal como se describe en el presente documento.

Pueden diseñarse vectores de expresión recombinante para la expresión de un polipéptido biosintético de aldehído, alcohol graso, alcano y/o alqueno o variante en células procariotas o eucariotas (por ejemplo, células bacterianas, tales como *E. coli*, células de insecto (usando vectores de expresión de baculovirus), células de levadura o células de mamífero). Se comentan adicionalmente células huésped adecuadas en Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). Alternativamente, el vector de expresión recombinante puede transcribirse y traducirse *in vitro*, por ejemplo, usando secuencias reguladoras de promotor de T7 y polimerasa de T7.

La expresión de polipéptidos en procariotas, por ejemplo, *E. coli*, se lleva a cabo con la mayor frecuencia con vectores que contienen promotores constitutivos o inducibles que dirigen la expresión de polipéptidos o bien de fusión o bien no de fusión. Los vectores de fusión añaden varios aminoácidos a un polipéptido codificado en los mismos, habitualmente en el extremo amino-terminal del polipéptido recombinante. Tales vectores de fusión sirven normalmente para tres fines: (1) aumentar la expresión del polipéptido recombinante; (2) aumentar la solubilidad del polipéptido recombinante; y (3) ayudar en la purificación del polipéptido recombinante actuando como ligando en la purificación por afinidad. Con frecuencia, en vectores de expresión de fusión, se introduce un sitio de escisión proteolítica en la unión del resto de fusión y el polipéptido recombinante. Esto permite la separación del polipéptido recombinante del resto de fusión tras la purificación del polipéptido de fusión. Los ejemplos de tales enzimas, y sus secuencias de reconocimiento relacionadas, incluyen factor Xa, trombina y enterocinasas. Los vectores de expresión de fusión a modo de ejemplo incluyen pGEX (Pharmacia Biotech Inc; Smith *et al.*, *Gene* 67:31-40 (1988)), pMAL (New England Biolabs, Beverly, Mass.) y pRITS (Pharmacia, Piscataway, N.J.), que fusionan glutatión S-transferasa (GST), proteína de unión a maltosa E o proteína A, respectivamente, al polipéptido recombinante diana.

Los ejemplos de vectores de expresión de *E. coli* inducibles, no de fusión, incluyen pTrc (Amann *et al.*, Gene (1988) 69:301-315) y pET 11d (Studier *et al.*, Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, Calif. (1990) 60-89). La expresión génica diana a partir del vector pTrc se basa en la transcripción de ARN polimerasa del huésped a partir de un promotor de fusión de trp-lac híbrido. La expresión génica diana a partir del vector pET 11d se basa en la transcripción de un promotor de fusión de gn10-lac de T7 mediada por una ARN polimerasa viral coexpresada (gn1 de T7). Esta polimerasa viral se suministra por cepas huésped BL21(DE3) o HMS174(DE3) a partir de un profago  $\lambda$  residente que aloja un gen gn1 de T7 bajo el control transcripcional del promotor lacUV 5.

Una estrategia para maximizar la expresión de polipéptido recombinante es expresar el polipéptido en una célula huésped con una capacidad alterada para escindir proteolíticamente el polipéptido recombinante (véase Gottesman, Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, Calif. (1990) 119-128). Otra estrategia es alterar la secuencia de ácido nucleico que va a insertarse en un vector de expresión de modo que los codones individuales para cada aminoácido son los usados preferiblemente en la célula huésped (Wada *et al.*, Nucleic Acids Res. (1992) 20:2111-2118). Tal alteración de secuencias de ácido nucleico puede llevarse a cabo mediante técnicas de síntesis de ADN convencionales.

En otra realización, la célula huésped es una célula de levadura. En esta realización, el vector de expresión es un vector de expresión de levadura. Los ejemplos de vectores para la expresión en levadura *S. cerevisiae* incluyen pYepSec1 (Baldari *et al.*, EMBO J. (1987) 6:229-234), pMFa (Kurjan *et al.*, Cell (1982) 30:933-943), pJRY88 (Schultz *et al.*, Gene (1987) 54:113-123), pYES2 (Invitrogen Corporation, San Diego, Calif.) y picZ (Invitrogen Corp, San Diego, Calif.).

Alternativamente, puede expresarse un polipéptido descrito en el presente documento en células de insecto usando vectores de expresión de baculovirus. Los vectores de baculovirus disponibles para la expresión de proteínas en células de insecto en cultivo (por ejemplo, células Sf9) incluyen, por ejemplo, la serie pAc (Smith *et al.*, Mol. Cell Biol. (1983) 3:2156-2165) y la serie pVL (Lucklow *et al.*, Virology (1989) 170:31-39).

En aún otra realización, los ácidos nucleicos descritos en el presente documento pueden expresarse en células de mamífero usando un vector de expresión de mamífero. Los ejemplos de vectores de expresión de mamífero incluyen pCDM8 (Seed, Nature (1987) 329:840) y pMT2PC (Kaufman *et al.*, EMBO J. (1987) 6:187-195). Cuando se usan en células de mamífero, las funciones de control del vector de expresión pueden proporcionarse por elementos reguladores virales. Por ejemplo, los promotores habitualmente usados se derivan de polioma, adenovirus 2, citomegalovirus y virus del simio 40. Otros sistemas de expresión adecuados para células tanto procariontas como eucariotas se describen en los capítulos 16 y 17 de Sambrook *et al.*, eds., Molecular Cloning: A Laboratory Manual. 2ª, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989.

Pueden introducirse vectores en células procariontas o eucariotas mediante técnicas de transformación o transfección convencionales. Tal como se usan en el presente documento, los términos "transformación" y "transfección" se refieren a una variedad de técnicas reconocidas en la técnica para introducir ácido nucleico foráneo (por ejemplo, ADN) en una célula huésped, incluyendo coprecipitación con fosfato de calcio o cloruro de calcio, transfección mediada por DEAE-dextrano, lipofección o electroporación. Pueden encontrarse métodos adecuados para transformar o transfectar células huésped, por ejemplo, en Sambrook *et al.* (citado anteriormente).

Para la transformación estable de células bacterianas, se sabe que, dependiendo del vector de expresión y la técnica de transformación usados, sólo una pequeña fracción de células captarán y replicarán el vector de expresión. Con el fin de identificar y seleccionar estos transformantes, puede introducirse un gen que codifica para un marcador seleccionable (por ejemplo, resistencia a antibióticos) en las células huésped junto con el gen de interés. Los marcadores seleccionables incluyen los que confieren resistencia a fármacos, tales como ampicilina, canamicina, cloranfenicol o tetraciclina. Pueden introducirse ácidos nucleicos que codifican para un marcador seleccionable en una célula huésped en el mismo vector que el que codifica para un polipéptido descrito en el presente documento o pueden introducirse en un vector separado. Pueden identificarse células transfectadas de manera estable con el ácido nucleico introducido mediante selección con fármaco (por ejemplo, las células que han incorporado el gen de marcador seleccionable sobrevivirán, mientras que las otras células morirán).

Para una transfección estable de células de mamífero, se sabe que, dependiendo del vector de expresión y la técnica de transfección usados, sólo una pequeña fracción de células puede integrar el ADN foráneo en su genoma. Con el fin de identificar y seleccionar estos integrantes, puede introducirse un gen que codifica para un marcador seleccionable (por ejemplo, resistencia a antibióticos) en las células huésped junto con el gen de interés. Los marcadores seleccionables preferidos incluyen aquéllos que confieren resistencia a fármacos, tales como G418, higromicina y metotrexato. Pueden introducirse ácidos nucleicos que codifican para un marcador seleccionable en una célula huésped en el mismo vector que el que codifica para un polipéptido descrito en el presente documento o pueden introducirse en un vector separado. Pueden identificarse las células transfectadas de manera estable con el ácido nucleico introducido mediante selección con fármaco (por ejemplo, las células que han incorporado el gen de marcador seleccionable sobrevivirán, mientras que las otras células morirán).

En determinados métodos, se coexpresan un polipéptido biosintético de aldehído y un polipéptido biosintético de

alcano o alqueno en una única célula huésped. En métodos alternativos, se coexpresan un polipéptido biosintético de aldehído y un polipéptido de alcohol deshidrogenasa en una única célula huésped.

#### Proteínas de transporte

5 Las proteínas de transporte pueden exportar polipéptidos e hidrocarburos (por ejemplo, aldehídos, alcanos y/o alquenos) fuera de una célula huésped. Muchas proteínas de transporte y eflujo sirven para excretar una amplia variedad de compuestos y pueden modificarse de manera natural para ser selectivas para tipos particulares de hidrocarburos.

10 Ejemplos no limitativos de proteínas de transporte adecuadas son proteínas de transporte de casete de unión a ATP (ABC), proteínas de eflujo y proteínas transportadoras de ácidos grasos (FATP). Los ejemplos no limitativos adicionales de proteínas de transporte adecuadas incluyen las proteínas de transporte de ABC de organismos tales como *Caenorhabditis elegans*, *Arabidopsis thaliana*, *Alkaligenes eutrophus* y *Rhodococcus erythropolis*. Las proteínas de transporte de ABC a modo de ejemplo que pueden usarse se enumeran en la figura 40 (por ejemplo, CER5, AtMRP5, AmiS2 y AtPGP1). También pueden elegirse células huésped por su capacidad endógena para secretar hidrocarburos. La eficacia de la producción de hidrocarburos y secreción en el entorno de la célula huésped (por ejemplo, medio de cultivo, caldo de fermentación) puede expresarse como una razón de producto intracelular con respecto a producto extracelular. En algunos ejemplos, la razón puede ser de aproximadamente 5:1, 4:1, 3:1, 2:1, 1:1, 1:2, 1:3, 1:4 ó 1:5.

#### Fermentación

20 La producción y el aislamiento de aldehídos, alcoholes grasos, alcanos y/o alquenos puede potenciarse empleando técnicas de fermentación beneficiosas. Un método para maximizar la producción al tiempo que se reducen los costes es aumentar el porcentaje de la fuente de carbono que se convierte en productos de hidrocarburo.

25 Durante ciclos de vida celulares normales, se usa carbono en funciones celulares, tales como producción de lípidos, sacáridos, proteínas, ácidos orgánicos y ácidos nucleicos. Reducir la cantidad de carbono necesaria para actividades relacionadas con el crecimiento puede aumentar la eficacia de conversión de la fuente de carbono en producto. Esto puede lograrse, por ejemplo, haciendo crecer en primer lugar células huésped hasta una densidad deseada (por ejemplo, una densidad que se alcanza al máximo de la fase de crecimiento logarítmica). En tal punto, pueden emplearse genes de punto de control de la replicación para detener el crecimiento de células. Específicamente, pueden usarse mecanismos de percepción de quórum (revisado en Camilli et al., Science 311:1113, 2006; Venturi FEMS Microbio. Rev. 30:274-291, 2006; y Reading et al., FEMS Microbiol. Lett. 254:1-11, 2006) para activar genes de punto de control, tales como p53, p21 u otros genes de punto de control.

30 Los genes que pueden activarse para detener la replicación y el crecimiento celular en *E. coli* incluyen genes *umuDC*. La sobreexpresión de genes *umuDC* detiene la progresión desde la fase estacionaria hasta el crecimiento exponencial (Murlí et al., J. of Bact. 182: 1127, 2000). UmuC es una ADN polimerasa que puede llevar a cabo la síntesis translesión sobre lesiones no codificantes (la base mecánica de la mayoría de las mutagénesis por UV y químicas). Los productos del gen *umuDC* están implicados en el proceso de síntesis translesión y también sirven como un punto de control de daño de la secuencia de ADN. Los productos del gen *umuDC* incluyen UmuC, UmuD, umuD', UmuD<sub>2</sub>C, UmuD<sub>2</sub>' y UmuD<sub>2</sub>. Simultáneamente, pueden activarse genes que producen productos, minimizando así la necesidad de usar rutas de replicación y mantenimiento mientras se prepara un aldehído, alcano y/o alqueno. También pueden modificarse por ingeniería células huésped para expresar *umuC* y *umuD* de *E. coli* en pBAD24 bajo el sistema del promotor *prpBCDE* mediante síntesis *de novo* de este gen con los genes de producción de producto final apropiados.

45 El porcentaje de carbonos introducidos convertidos en aldehídos, alcoholes grasos, alcanos y/o alquenos puede ser un impulsor de coste. Cuanto más eficaz es el proceso (es decir, mayor es el porcentaje de carbonos introducidos convertidos en aldehídos, alcoholes grasos, alcanos y/o alquenos), menos caro será el proceso. Para fuentes de carbono que contienen oxígeno (por ejemplo, fuentes a base de glucosa y otros hidratos de carbono), el oxígeno debe liberarse en forma de dióxido de carbono. Por cada 2 átomos de oxígeno liberados, también se libera un átomo de carbono conduciendo a una eficacia metabólica teórica máxima de aproximadamente el 34% (p/p) (para productos derivados de ácidos grasos). Sin embargo, esta cifra cambia para otros productos de hidrocarburo y fuentes de carbono. Las eficacias típicas en la bibliografía son aproximadamente inferiores al 5%. Las células huésped modificadas por ingeniería genética para producir aldehídos, alcoholes grasos, alcanos y/o alquenos pueden tener una eficacia superior a aproximadamente el 1, 3, 5, 10, 15, 20, 25 y 30%. En un ejemplo, las células huésped pueden mostrar una eficacia de aproximadamente el 10% a aproximadamente el 25%. En otros ejemplos, tales células huésped pueden mostrar una eficacia de aproximadamente el 25% a aproximadamente el 30%. En otros ejemplos, las células huésped pueden mostrar una eficacia superior al 30%.

55 La célula huésped puede modificarse por ingeniería genética adicionalmente para expresar celulosomas recombinantes, tales como los descritos en la solicitud de PCT número PCT/US2007/003736. Estos celulosomas pueden permitir que la célula huésped use material celulósico como fuente de carbono. Por ejemplo, la célula huésped puede modificarse por ingeniería genética adicionalmente para expresar invertasas (EC 3.2.1.26) de modo

que puede usarse sacarosa como fuente de carbono. De manera similar, la célula huésped puede modificarse por ingeniería genética usando las enseñanzas descritas en las patentes estadounidenses n.ºs 5.000.000; 5.028.539; 5.424.202; 5.482.846 y 5.602.030; de modo que la célula huésped puede asimilar carbono eficazmente y usar materiales celulósicos como fuentes de carbono.

- 5 En un ejemplo, la cámara de fermentación puede encerrar una fermentación que está experimentando una reducción continua. En este caso, puede crearse un entorno reductor estable. El equilibrio electrónico puede mantenerse mediante la liberación de dióxido de carbono (en forma gaseosa). Esfuerzos para aumentar el equilibrio de NAD/H y NADP/H también pueden facilitar la estabilización del equilibrio electrónico. También puede potenciarse la disponibilidad de NADPH intracelular modificando por ingeniería genética la célula huésped para expresar una NADH:NADPH transhidrogenasa. La expresión de una o más NADH:NADPH transhidrogenasas convierte la NADH producida en la glicólisis en NADPH, lo que puede potenciar la producción de aldehídos, alcanos y/o alquenos.

10 Para la producción a pequeña escala, las células huésped modificadas por ingeniería genética pueden hacerse crecer en lotes de, por ejemplo, aproximadamente 100 ml, 500 ml, 1 l, 2 l, 5 l, o 10 l; fermentarse; e inducirse para expresar aldehídos, alcoholes grasos, alcanos y/o alquenos deseados basándose en los genes específicos codificados en los plásmidos apropiados. Por ejemplo, pueden incubarse células de *E. coli* BL21 (DE3) que llevan pBAD24 (con resistencia a ampicilina y la ruta de síntesis de aldehído, alcohol graso, alcano o alqueno) así como pUMVC1 (con resistencia a kanamicina y el sistema de sobreexpresión de acetil-CoA/malonil-CoA) durante la noche en matraces de 2 l a 37°C agitados a > 200 rpm en 500 ml de medio LB complementado con ampicilina 75 µg/ml y kanamicina 50 µg/ml hasta que los cultivos alcancen una DO<sub>600</sub> de > 0,8. Tras lograr una DO<sub>600</sub> de > 0,8, las células pueden complementarse con propionato de sodio 25 mM (pH 8,0) para activar los sistemas de genes modificados por ingeniería genética para la producción y para detener la proliferación celular activando las proteínas UmuC y UmuD. Puede realizarse inducción durante 6 h a 30°C. Tras la incubación, los medios pueden examinarse para determinar aldehídos, alcoholes grasos, alcanos y/o alquenos usando CG-EM.

25 Para producción a gran escala, las células huésped modificadas por ingeniería genética pueden hacerse crecer en lotes de 10 l, 100 l, 1000 l, o mayores; fermentarse; e inducirse para expresar aldehídos, alcoholes grasos, alcanos y/o alquenos deseados basándose en los genes específicos codificados en los plásmidos apropiados. Por ejemplo, pueden incubarse células de *E. coli* BL21 (DE3) que llevan pBAD24 (con resistencia a ampicilina y la ruta de síntesis de aldehído y/o alcano) así como pUMVC1 (con resistencia a kanamicina y el sistema de sobreexpresión de acetil-CoA/malonil-CoA) a partir de un cultivo de siembra de 500 ml para fermentaciones de 10 l (5 l para fermentaciones de 100 l, etc.) en medios LB (libres de glicerol) con kanamicina 50 µg/ml y ampicilina 75 µg/ml a 37°C, y agitarse a > 200 rpm hasta que los cultivos alcancen una DO<sub>600</sub> de > 0,8 (normalmente 16 h). Pueden complementarse los medios continuamente para mantener propionato de sodio 25 mM (pH 8,0) para activar los sistemas de genes modificados por ingeniería genética para la producción y para detener la proliferación celular activando las proteínas umuC y umuD. Pueden complementarse los medios continuamente con glucosa para mantener una concentración de 25 g/100 ml.

Tras la primera hora de inducción, pueden retirarse alícuotas de no más del 10% del volumen celular total cada hora y permitirse que se asienten sin agitación para permitir que los aldehídos, alcanos y/o alquenos afloren a la superficie y experimenten una separación de fase espontánea. Entonces puede recogerse el componente de aldehído, alcoholes grasos, alcano y/o alqueno, y devolverse la fase acuosa a la cámara de reacción. La cámara de reacción puede operarse de manera continua. Cuando la DO<sub>600</sub> cae por debajo de 0,6, las células pueden sustituirse por un nuevo lote hecho crecer a partir de un cultivo de siembra.

#### Producción de aldehídos, alcoholes grasos, alcanos y alquenos usando métodos libres de células

45 En algunos métodos descritos en el presente documento, puede producirse un aldehído, alcoholes grasos, alcano y/o alqueno usando un polipéptido purificado descrito en el presente documento y un sustrato descrito en el presente documento. Por ejemplo, una célula huésped puede modificarse por ingeniería genética para expresar polipéptido biosintético de aldehído, alcoholes grasos, alcano y/o alqueno o variante tal como se describe en el presente documento. La célula huésped puede cultivarse en condiciones adecuadas para permitir la expresión del polipéptido. Pueden generarse entonces extractos libres de células usando métodos conocidos. Por ejemplo, las células huésped pueden lisarse usando detergentes o mediante sonicación. Los polipéptidos expresados pueden purificarse usando métodos conocidos. Tras obtener los extractos libres de células, pueden añadirse los sustratos descritos en el presente documento a los extractos libres de células y mantenerse en condiciones que permitan la conversión de los sustratos en aldehídos, alcoholes grasos, alcanos y/o alquenos. Los aldehídos, alcoholes grasos, alcanos y/o alquenos pueden entonces separarse y purificarse usando técnicas conocidas.

#### Procesamiento tras la producción

55 Los aldehídos, alcoholes grasos, alcanos y/o alquenos producidos durante la fermentación pueden separarse de los medios de fermentación. Puede usarse cualquier técnica conocida para separar aldehídos, alcoholes grasos, alcanos y/o alquenos de medios acuosos. Un procedimiento de separación a modo de ejemplo es un procedimiento de separación de dos bases (bifásico). Este procedimiento implica fermentar las células huésped modificadas por ingeniería genética en condiciones suficientes para producir un aldehído, alcoholes grasos, alcano y/o alqueno,

permitiendo que el aldehído, alcoholes grasos, alcano y/o alqueno se recojan en una fase orgánica, y separando la fase orgánica del caldo de fermentación acuoso. Este método puede practicarse en un entorno de fermentación tanto discontinua como continua.

5 La separación bifásica usa la inmiscibilidad relativa de aldehídos, alcoholes grasos, alcanos y/o alquenos para facilitar la separación. Inmiscible se refiere a la incapacidad relativa de un compuesto para disolverse en agua y se define por el coeficiente de reparto del compuesto. Un experto habitual en la técnica apreciará que al elegir un caldo de fermentación y fase orgánica, de modo que el aldehído, alcano y/o alqueno que está produciéndose tenga un alto valor de logP, el aldehído, alcano y/o alqueno pueden separarse en la fase orgánica, incluso a concentraciones muy bajas, en el recipiente de fermentación.

10 Los aldehídos, alcoholes grasos, alcanos y/o alquenos producidos por los métodos descritos en el presente documento pueden ser relativamente inmiscibles en el caldo de fermentación, así como en el citoplasma. Por tanto, el aldehído, alcoholes grasos, alcano y/o alqueno pueden recogerse en una fase orgánica o bien de manera intracelular o bien extracelular. La recogida de los productos en la fase orgánica puede reducir el impacto del aldehído, alcoholes grasos, alcano y/o alqueno sobre la función celular y puede permitir que la célula huésped produzca más producto.

15 Los métodos descritos en el presente documento pueden dar como resultado la producción de compuestos homogéneos en los que al menos aproximadamente el 60%, el 70%, el 80%, el 90% o el 95% de los aldehídos, alcoholes grasos, alcanos y/o alquenos producidos tendrán longitudes de la cadena de carbono que varían en menos de aproximadamente 6 carbonos, menos de aproximadamente 4 carbonos, o menos de aproximadamente 2 carbonos. Estos compuestos también pueden producirse con un grado relativamente uniforme de saturación. Estos compuestos pueden usarse directamente como combustibles, aditivos de combustible, productos químicos especializados, materiales de partida para la producción de otros compuestos químicos (por ejemplo, polímeros, tensoactivos, plásticos, textiles, disolventes, adhesivos, etc.), o aditivos de productos de cuidado personal. Estos compuestos también pueden usarse como materia prima para reacciones posteriores, por ejemplo, hidrogenación, craqueo catalítico (por medio de hidrogenación, pirólisis, o ambas), para producir otros productos.

20 En algunas realizaciones, los aldehídos, alcoholes grasos, alcanos y/o alquenos producidos usando los métodos descritos en el presente documento pueden contener entre aproximadamente el 50% y aproximadamente el 90% de carbono; o entre aproximadamente el 5% y aproximadamente el 25% de hidrógeno. En otras realizaciones, los aldehídos, alcoholes grasos, alcanos y/o alquenos producidos usando los métodos descritos en el presente documento pueden contener entre aproximadamente el 65% y aproximadamente el 85% de carbono; o entre aproximadamente el 10% y aproximadamente el 15% de hidrógeno.

#### Composiciones de combustible y composiciones de productos químicos especializados

35 Los aldehídos, alcoholes grasos, alcanos y/o alquenos descritos en el presente documento pueden usarse como o convertirse en un combustible o como un producto químico especializado. Un experto habitual en la técnica apreciará que, según el fin previsto del combustible o producto químico especializado, pueden producirse y usarse diferentes aldehídos, alcoholes grasos, alcanos y/o alquenos. Por ejemplo, un aldehído, alcohol graso, alcano y/o alqueno ramificado puede ser deseable para combustible de automóviles que pretende usarse en climas fríos. Además, cuando los aldehídos, alcoholes grasos, alcanos y/o alquenos descritos en el presente documento se usan como materia prima para la producción de combustible o producto químico especializado, un experto habitual en la técnica apreciará que las características de la materia prima de aldehído, alcohol graso, alcano y/o alqueno afectarán a las características del combustible o producto químico especializado producido. Por tanto, las características del producto de combustible o producto químico especializado pueden seleccionarse produciendo aldehídos, alcoholes grasos, alcanos y/o alquenos particulares para su uso como materia prima.

40 Usando los métodos descritos en el presente documento, pueden producirse biocombustibles que tienen calidades de combustible deseadas a partir de aldehídos, alcoholes grasos, alcanos y/o alquenos. Aldehídos, alcoholes grasos, alcanos y/o alquenos producidos de manera biológica representan una nueva fuente de biocombustibles, que pueden usarse como carburante, diésel o gasolina. Algunos biocombustibles elaborados usando aldehídos, alcoholes grasos, alcanos y/o alquenos no se han producido a partir de fuentes renovables y son nuevas composiciones de materia. Estos nuevos combustibles o productos químicos especializados pueden distinguirse de combustibles o productos químicos especializados derivados de carbono petroquímico basándose en la huella de carbono isotópico doble. Además, la fuente específica de carbono de fuentes biológicas (por ejemplo, glucosa frente a glicerol) puede determinarse mediante la huella de carbono isotópico doble (véase, por ejemplo, la patente estadounidense n.º 7.169.588, en particular col. 4, línea 31, a col. 6, línea 8).

45 Los aldehídos, alcoholes grasos, alcanos y/o alquenos y los biocombustibles, productos químicos especializados y mezclas asociados pueden distinguirse de sus homólogos derivados de productos petroquímicos basándose en  $^{14}\text{C}$  ( $f_M$ ) y huella de carbono isotópico doble. En algunos ejemplos, el aldehído, alcohol graso, alcano y/o alqueno en la composición de biocombustible pueden tener una fracción de carbono moderno ( $f_M^{14}\text{C}$ ) de, por ejemplo, al menos aproximadamente 1,003, 1,010 ó 1,5.

En algunos ejemplos, puede elaborarse una composición de biocombustible que incluye aldehídos, alcoholes grasos, alcanos y/o alquenos que tienen  $\delta^{13}\text{C}$  de desde aproximadamente -15,4 hasta aproximadamente -10,9, donde los aldehídos, alcoholes grasos, alcanos y/o alquenos representan al menos aproximadamente el 85% de material de fuentes biológicas (es decir, derivado a partir de un recurso renovable, tal como biomasa, materiales celulósicos y azúcares) en la composición.

La capacidad para distinguir estos productos derivados de manera biológica es beneficiosa para monitorizar estos materiales en el mercado. Por ejemplo, combustibles o productos químicos especializados que comprenden tanto perfiles de isótopo de carbono derivados de manera biológica como a base de petróleo pueden distinguirse de combustibles y productos químicos especializados elaborados únicamente por materiales a base de carbono. Por tanto, los aldehídos, alcoholes grasos, alcanos y/o alquenos descritos en el presente documento pueden seguirse en el mercado o identificarse en el mercado basándose en su perfil único. Además, pueden identificarse otros materiales competidores como derivados de manera biológica o derivados de una fuente petroquímica.

Se usan aditivos de combustible para potenciar el rendimiento de un combustible o motor. Por ejemplo, pueden usarse aditivos de combustible para alterar el punto de congelación/gelificación, punto de turbidez, poder de lubricación, viscosidad, estabilidad oxidativa, calidad de ignición, nivel de octano y/o punto de inflamación. En los Estados Unidos, todos los aditivos de combustible deben registrarse en la Agencia de Protección Ambiental. Los nombres de los aditivos de combustible y las empresas que comercializan los aditivos de combustible están disponibles públicamente contactando con la EPA o viendo la página web de la agencia. Un experto habitual en la técnica apreciará que los biocombustibles a base de aldehído y/o alcano descritos en el presente documento pueden mezclarse con uno o más aditivos de combustible para conferir una calidad deseada.

Los biocombustibles a base de aldehído, alcoholes grasos, alcano y/o alqueno descritos en el presente documento pueden mezclarse con otros combustibles, tales como diversos alcoholes, tales como etanol y butanol, y productos derivados de petróleo, tales como gasolina, diésel o carburante.

En algunos ejemplos, la mezcla puede incluir al menos aproximadamente el 10%, el 15%, el 20%, el 30%, el 40%, el 50% o el 60% en peso del aldehído, alcoholes grasos, alcano, o alqueno. En otros ejemplos, puede elaborarse una composición de biocombustible que incluye al menos aproximadamente el 5%, el 10%, el 20%, el 30%, el 40%, el 50%, el 60%, el 70%, el 80%, el 85%, el 90% o el 95% de un aldehído, alcoholes grasos, alcano o alqueno que incluyen una cadena de carbono que es de 8, 10, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 ó 22 carbonos de longitud. Tales composiciones de biocombustible puede incluir además al menos un aditivo seleccionado de un aditivo de reducción del punto de turbidez que puede reducir el punto de turbidez hasta menos de aproximadamente 5°C, o 0°C; un tensioactivo; una microemulsión; al menos aproximadamente el 5%, el 10%, el 15%, el 20%, el 30%, el 40%, el 50%, el 60%, el 70%, el 80%, el 85%, el 90% o el 95% de combustible diésel a partir de triglicéridos; gasolina derivada de petróleo; o combustible diésel a partir de petróleo.

### Ejemplos

La invención se describe adicionalmente en los siguientes ejemplos, que no limitan el alcance de la invención descrita en las reivindicaciones.

#### Ejemplo 1. Detección y verificación de la biosíntesis de alcanos en cianobacterias seleccionadas

Se seleccionaron siete cianobacterias, cuyas secuencias de genoma completo están disponibles públicamente, para la verificación y/o detección de la biosíntesis de alcanos: *Synechococcus elongatus* PCC7942, *Synechococcus elongatus* PCC6301, *Anabaena variabilis* ATCC29413, *Synechocystis* sp. PCC6803, *Nostoc punctiforme* PCC73102, *Gloeobacter violaceus* ATCC 29082 y *Prochlorococcus marinus* CCMP1986. Sólo las tres primeras cepas cianobacterianas de esta lista se había notificado anteriormente que contenían alcanos (Han *et al.*, J. Am. Chem. Soc. 91:5156-5159 (1969); Fehler *et al.*, Biochem. 9:418-422 (1970)). Las cepas se hicieron crecer de manera fotoautotrófica en matraces de agitación en 100 ml de los medios apropiados (enumerados en la tabla 8) durante 3-7 días a 30°C a una intensidad de luz de aproximadamente 3.500 lux. Se extrajeron las células para la detección de alcanos tal como sigue: se centrifugaron células a partir de un volumen de cultivo de 1 ml durante 1 min a 13.000 rpm, se resuspendieron los sedimentos celulares en metanol, se agitaron con vórtex durante 1 min y luego se sonicaron durante 30 min. Tras la centrifugación durante 3 min a 13.000 rpm, se transfirieron los sobrenadantes a viales nuevos y se analizaron mediante CG-EM. Se analizaron las muestras o bien en una columna capilar DP-5 30 m (diámetro interno de 0,25 mm) o bien en una columna DP-5 30 m de alta temperatura (diámetro interno de 0,25 mm) usando el siguiente método.

Tras una inyección sin división de 1  $\mu\text{l}$  (temperatura de entrada mantenida a 300°C) sobre la columna de CG/EM, se mantuvo el horno a 100°C durante 3 min. Se elevó la temperatura hasta 320°C a una velocidad de 20°C/min. Se mantuvo el horno a 320°C durante 5 min adicionales. La velocidad de flujo del gas portador helio fue de 1,3 ml/min. El cuadrupolo de EM realizó un barrido desde 50 hasta 550 m/z. Se compararon los tiempos de retención y patrones de fragmentación de picos de producto con referencias auténticas para confirmar la identidad de pico.

De las siete cepas, seis producían principalmente heptadecano y una producía pentadecano (*P. marinus* CCMP1986); una de estas cepas producía metil-heptadecano además de heptadecano (*A. variabilis* ATCC29413)

(véase la tabla 8). Por tanto, se verificó la biosíntesis de alcanos en tres cianobacterias notificadas previamente, y se detectó biosíntesis de alcanos en cuatro cianobacterias que no se sabía previamente que producían alcanos: *P. marinus* CCMP1986 (véase la figura 1), *N. punctiforme* PCC73102 (véase la figura 2), *G. violaceus* ATCC 29082 (véase la figura 3) y *Synechocystis sp.* PCC6803 (véase la figura 4).

5 La figura 1A representa el trazo de CG/EM de células de *Prochlorococcus marinus* CCMP1986 extraídas con metanol. El pico a 7,55 min tenía el mismo tiempo de retención que pentadecano (Sigma). En la figura 1B, se muestra el patrón de fragmentación de masas del pico de pentadecano. El pico 212 corresponde al peso molecular de pentadecano.

10 La figura 2A representa el trazo de CG/EM de células de *Nostoc punctiforme* PCC73102 extraídas con metanol. El pico a 8,73 min tiene el mismo tiempo de retención que heptadecano (Sigma). En la figura 2B, se muestra el patrón de fragmentación de masas del pico de heptadecano. El pico 240 corresponde al peso molecular de heptadecano.

15 La figura 3A representa el trazo de CG/EM de células de *Gloeobacter violaceus* ATCC29082 extraídas con metanol. El pico a 8,72 min tiene el mismo tiempo de retención que heptadecano (Sigma). En la figura 3B, se muestra el patrón de fragmentación de masas del pico de heptadecano. El pico 240 corresponde al peso molecular de heptadecano.

La figura 4A representa el trazo de CG/EM de células de *Synechocystis sp.* PCC6803 extraídas con metanol. El pico a 7,36 min tiene el mismo tiempo de retención que heptadecano (Sigma). En la figura 4B, se muestra el patrón de fragmentación de masas del pico de heptadecano. El pico 240 corresponde al peso molecular de heptadecano.

Tabla 8: Hidrocarburos detectados en cianobacterias seleccionadas

Cianobacteria	N.º de la ATCC	Genoma	Medio	Alcanos	
				Notificado	Verificado <sup>2</sup>
<i>Synechococcus elongatus</i> PCC7942	27144	2,7 Mb	BG-11	C17:0	<b>C17:0</b> , C15:0
<i>Synechococcus elongatus</i> PCC6301	33912	2,7 Mb	BC-11	C17:0	<b>C17:0</b> , C15:0
<i>Anabaena variabilis</i>	29413	6,4 Mb	BG-11	C17:0, 7- u 8- Me-C17:0	<b>C17:0</b> , Me- C17:0
<i>Synechocystis sp.</i> PCC6803	27184	3,5 Mb	BG-11	-	<b>C17:0</b> , C15:0
<i>Prochlorococcus marinus</i> CCMP1986 <sup>1</sup>	-	1,7 Mb	-	-	<b>C15:0</b>
<i>Nostoc punctiforme</i> PCC73102	29133	9,0 Mb	ATCC819	-	<b>C17:0</b>
<i>Gloeobacter violaceus</i>	29082	4,6 Mb	BG11	-	<b>C17:0</b>
<sup>1</sup> Las células para la extracción fueron un regalo de Jacob Waldbauer (MIT)					
<sup>2</sup> el hidrocarburo principal está en negrita					

20 El análisis genómico produjo dos genes que estaban presentes en las cepas productoras de alcanos. Los homólogos de *Synechococcus elongatus* PCC7942 de estos genes se representan en la tabla 9 y son Synpcc7942\_1593 (SEQ ID NO: 1) y Synpcc7942\_1594 (SEQ ID NO: 65).

Tabla 9: Genes cianobacterianos productores de alcanos

ID de objeto génico	Etiq. de locus	Registro de Genbank	Nombre del gen	Longitud	COG	Pfam	InterPro	Notas
637800026	Synpcc 7942_1593	YP_400610	Proteína hipotética	231 aa	-	pfam02915	IPR009078 IPR003251	Rubreritrina de tipo ferritina/ribonucleótido reductasa
637800027	Synpcc 7942_1594	YP_400611	Proteína hipotética	341 aa	COG5322	pfam00106	IPR000408 IPR016040 IPR002198	Deshidrogenasa de cadena corta de unión a deshidrogenasa NAD(P) predicha

25 Ejemplo 2. La delección de los genes sll0208 y sll0209 en *Synechocystis sp.* PCC6803 conduce a pérdida de biosíntesis de alcanos

30 Los genes que codifican para la supuesta descarboxilasa (sll0208; NP\_442147) (SEQ ID NO: 3) y enzima generadora de aldehído (sll0209; NP\_442146) (SEQ ID NO: 67) de *Synechocystis sp.* PCC6803 se deleccionaron tal como sigue. Se amplificó aproximadamente 1 kb de ADN flanqueante en el sentido de 5' y en el sentido de 3' usando el cebador sll0208/9-KO1 (CGCGGATCCCTTGATTCTACTGCGGCGAGT) con cebador sll0208/9-KO2 (CACGCACCTAGGTTCCACTCCCATGGTATAACAGGGGCGTTGGACTCC TGTG) y cebador sll0208/9-KO3 (GTTATACCATGGGAGTGTGAACCTAGGTGCGTGGCCGACAGGATAGGG-CGTGT) con cebador sll0208/9-KO4

(CGCGGATCCAACGCATCCTCACTAGTCGGG), respectivamente. Se usaron los productos de PCR en una PCR cruzada con cebadores sll0208/9-KO1 y sll0208/9-KO4 para amplificar el casete de delección de sll0208/sll0209 de aproximadamente 2 kb, que se clonó en el sitio BamHI del vector de clonación pUC 19. Entonces se amplificó un casete de resistencia a kanamicina (aph, KanR) a partir del plásmido pRL27 (Larsen *et al.*, Arch. Microbiol. 178:193 (2002)) usando los cebadores Kan-aph-F (CATGCCATGGAAAGCCACGTTGTGTCTCAAATCTCTG) y Kan-aph-R (CTAGTCTAGAGCGCTGAGGTCTGCCTCGTGAA), que entonces se cortó con NcoI y XbaI y se clonó en los sitios NcoI y AvrII del casete de delección sll0208/sll0209, creando un casete de inserción de KanR con delección de sll0208/sll0209 en pUC 19. El vector que contiene el casete, que no se replica en cianobacterias, se transformó en *Synechocystis sp.* PCC6803 (Zang *et al.*, 2007, J. Microbiol., vol. 45, págs. 241) y se seleccionaron transformantes (por ejemplo, integrantes cromosómicos mediante recombinación homóloga doble) en placas de agar BG-11 que contenían kanamicina 100 µg/ml en un incubador equipado con luz a 30°C. Volvieron a sembrarse en estría colonias resistentes a kanamicina una vez y luego se sometieron a análisis genotípico usando PCR con cebadores de diagnóstico.

Se cultivaron mutantes de inserción con delección confirmada en 12 ml de medio BG11 con kanamicina 50 µg/ml durante 4 días a 30°C en un incubador-agitador equipado con luz. Entonces se centrifugó 1 ml de caldo (1 min a 13.000 g) y se extrajeron los sedimentos celulares con 0,1 ml de metanol. Tras la extracción, se centrifugaron de nuevo las muestras y se sometieron los sobrenadantes a análisis de CG-EM tal como se describió en el ejemplo 1.

Tal como se muestra en la figura 5, las cepas de *Synechocystis sp.* PCC6803 en las que se deleccionaron los genes sll0208 y sll0209 perdieron su capacidad para producir heptadeceno y octadecenal. Este resultado demuestra que los genes sll0208 y sll0209 en *Synechocystis sp.* PCC6803 y los genes ortólogos en otras cianobacterias (véase la tabla 1) son responsables de la biosíntesis de alcanos y aldehídos grasos en estos organismos.

#### Ejemplo 3. Producción de aldehídos grasos y alcoholes grasos en *E. coli* a través de expresión heteróloga de orf1594 de *Synechococcus elongatus* PCC7942

Se amplificó el ADN genómico que codifica para orf1594 de *Synechococcus elongatus* PCC7942 (YP\_400611; supuesta enzima generadora de aldehído) (SEQ ID NO: 65) y se clonó en los sitios NcoI y EcoRI del vector OP-80 (derivado de pCL1920) bajo el control del promotor P<sub>trc</sub>. Se transformó el constructo resultante ("OP80-PCC7942\_1594") en *E. coli* MG1655 y se hicieron crecer las células a 37°C en medio mínimo M9 con glucosa al 1% (p/v) como fuente de carbono y complementado con espectinomicina 100 µg/ml. Cuando el cultivo alcanzó una DO<sub>600</sub> de 0,8-1,0, se indujo con IPTG 1 mM y se hicieron crecer las células durante 18-20 h adicionales a 37°C. Se extrajeron células a partir de 0,5 ml de cultivo con 0,5 ml de acetato de etilo. Tras la sonicación durante 60 min, se centrifugó la muestra a 15.000 rpm durante 5 min. Se analizó la capa de disolvente mediante CG-EM tal como se describió en el ejemplo 1.

Tal como se muestra en la figura 6, células de *E. coli* transformadas con el vector que lleva orf1594 de *Synechococcus elongatus* PCC7942 produjeron los siguientes aldehídos grasos y alcoholes grasos: hexadecanal, octadecenal, tetradecenol, hexadecenol, hexadecanol y octadecenol. Este resultado indica que orf1594 de PCC7942 (i) genera aldehídos *in vivo* como posibles sustratos para la descarbonilación y (ii) puede reducir acil-ACP como sustratos, que son la forma más abundante de ácidos grasos activados en células de *E. coli* silvestres. Por tanto, la enzima se denominó acil-ACP reductasa. *In vivo*, los aldehídos grasos aparentemente se reducen adicionalmente a los correspondientes alcoholes grasos mediante una actividad aldehído reductasa de *E. coli* endógena.

#### Ejemplo 4. Producción de aldehídos grasos y alcoholes grasos en *E. coli* a través de expresión heteróloga de cce\_1430 de *Cyanothece sp.* ATCC51142

Se amplificó el ADN genómico que codifica para cce\_1430 de *Cyanothece sp.* ATCC51142 (YP\_001802846; supuesta enzima generadora de aldehído) (SEQ ID NO: 69) y se clonó en los sitios NcoI y EcoRI del vector OP-80 (derivado de pCL1920) bajo el control del promotor P<sub>trc</sub>. Se transformó el constructo resultante en *E. coli* MG1655 y se hicieron crecer las células a 37°C en medio mínimo M9 con glucosa al 1% (p/v) como fuente de carbono y complementado con espectinomicina 100 µg/ml. Se cultivaron las células y se extrajeron como en el ejemplo 3 y se analizaron mediante CG-EM tal como se describió en el ejemplo 26.

Tal como se muestra en la figura 7, células de *E. coli* transformadas con el vector que lleva cce\_1430 de *Cyanothece sp.* ATCC51142 produjeron los siguientes aldehídos grasos y alcoholes grasos: hexadecanal, octadecenal, tetradecenol, hexadecenol, hexadecanol y octadecenol. Este resultado indica que cce\_1430 de ATCC51142 (i) genera aldehídos *in vivo* como posibles sustratos para la descarbonilación y (ii) puede reducir acil-ACP como sustratos, que son la forma más abundante de ácidos grasos activados en células de *E. coli* silvestres. Por tanto, esta enzima es también una acil-ACP reductasa.

#### Ejemplo 5. Producción de alcanos y alquenos en *E. coli* a través de expresión heteróloga de orf1594 de *Synechococcus elongatus* PCC7942 y orf1593 de *Synechococcus elongatus* PCC7942

Se amplificó el ADN genómico que codifica para orf1593 de *Synechococcus elongatus* PCC7942 (YP\_400610; supuesta descarbonilasa) (SEQ ID NO: 1) y se clonó en los sitios NdeI y XhoI del vector OP-183 (derivado de

pACYC) bajo el control del promotor P<sub>trc</sub>. Se cotransformó el constructo resultante con OP80-PCC7942\_1594 en *E. coli* MG1655 y se hicieron crecer las células a 37°C en medio mínimo M9 complementado con espectinomomicina 100 µg/ml y carbenicilina 100 µg/ml. Se cultivaron las células y se extrajeron como en el ejemplo 3 y se analizaron mediante CG-EM tal como se describió en el ejemplo 1.

5 Tal como se muestra en la figura 8, células de *E. coli* cotransformadas con los vectores que llevan PCC7942\_1594 de *S. elongatus* y PCC7942\_1593 de *S. elongatus* produjeron los mismos aldehídos grasos y alcoholes grasos que en el ejemplo 3, pero también pentadecano y heptadeceno. Este resultado indica que PCC7942\_1593 en *E. coli* convierte hexadecanal y octadecenal en pentadecano y heptadeceno, respectivamente, y por tanto es una aldehído graso descarboxilasa activa.

10 Ejemplo 6. Producción de alcanos y alquenos en *E. coli* a través de expresión heteróloga de orf1594 de *Synechococcus elongatus* PCC7942 y Npun02004178 de *Nostoc punctiforme* PCC73102

Se amplificó el ADN genómico que codifica para Npun02004178 de *Nostoc punctiforme* PCC73102 (ZP\_00108838; supuesta descarboxilasa) (SEQ ID NO: 5) y se clonó en los sitios NdeI y XhoI del vector OP-183 (derivado de pACYC) bajo el control del promotor P<sub>trc</sub>. Se cotransformó el constructo resultante con OP80-PCC7942\_1594 en *E. coli* MG1655 y se hicieron crecer las células a 37°C en medio mínimo M9 complementado con espectinomomicina 100 µg/ml y carbenicilina 100 µg/ml. Se cultivaron las células y se extrajeron como en el ejemplo 3 y se analizaron mediante CG-EM tal como se describió en el ejemplo 1.

20 Tal como se muestra en la figura 9, células de *E. coli* cotransformadas con los vectores que llevan PCC7942\_1594 de *S. elongatus* y Npun02004178 de *N. punctiforme* PCC73102 produjeron los mismos aldehídos grasos y alcoholes grasos que en el ejemplo 3, pero también tridecano, pentadeceno, pentadecano y heptadeceno. Este resultado indica que Npun02004178 en *E. coli* convierte tetradecanal, hexadecenal, hexadecanal y octadecenal en tridecano, pentadeceno, pentadecano y heptadeceno, respectivamente, y por tanto es una aldehído graso descarboxilasa activa.

25 Ejemplo 7. Producción de alcanos y alquenos en *E. coli* a través de expresión heteróloga de orf1594 de *Synechococcus elongatus* PCC7942 y slI0208 de *Synechocystis* sp. PCC6803

Se amplificó el ADN genómico que codifica para slI0208 de *Synechocystis* sp. PCC6803 (NP\_442147; supuesta descarboxilasa) (SEQ ID NO: 3) y se clonó en los sitios NdeI y XhoI del vector OP-183 (derivado de pACYC) bajo el control del promotor P<sub>trc</sub>. Se cotransformó el constructo resultante con OP80-PCC7942\_1594 en *E. coli* MG1655 y se hicieron crecer las células a 37°C en medio mínimo M9 complementado con espectinomomicina 100 µg/ml y carbenicilina 100 µg/ml. Se cultivaron las células y se extrajeron como en el ejemplo 3 y se analizaron mediante CG-EM tal como se describió en el ejemplo 1.

35 Tal como se muestra en la figura 10, células de *E. coli* cotransformadas con los vectores que llevan PCC7942\_1594 de *S. elongatus* y slI0208 de *Synechocystis* sp. PCC6803 produjeron los mismos aldehídos grasos y alcoholes grasos que en el ejemplo 3, pero también pentadecano y heptadeceno. Este resultado indica que Npun02004178 en *E. coli* convierte hexadecanal y octadecenal en pentadecano y heptadeceno, respectivamente, y por tanto es una aldehído graso descarboxilasa activa.

Ejemplo 8. Producción de alcanos y alquenos en *E. coli* a través de expresión heteróloga de orf1594 de *Synechococcus elongatus* PCC7942 y alr5283 de *Nostoc* sp. PCC7210

40 Se amplificó el ADN genómico que codifica para alr5283 de *Nostoc* sp. PCC7210 (NP\_489323; supuesta descarboxilasa) (SEQ ID NO: 7) y se clonó en los sitios NdeI y XhoI del vector OP-183 (derivado de pACYC) bajo el control del promotor P<sub>trc</sub>. Se cotransformó el constructo resultante con OP80-PCC7942\_1594 en *E. coli* MG1655 y se cultivaron las células a 37°C en medio mínimo M9 complementado con espectinomomicina 100 µg/ml y carbenicilina 100 µg/ml. Se cultivaron las células y se extrajeron como en el ejemplo 3 y se analizaron mediante CG-EM tal como se describió en el ejemplo 1.

45 Tal como se muestra en la figura 11, células de *E. coli* cotransformadas con los vectores que llevan PCC7942\_1594 de *S. elongatus* y alr5283 *Nostoc* sp. de PCC7210 produjeron los mismos aldehídos grasos y alcoholes grasos que en el ejemplo 3, pero también pentadecano y heptadeceno. Este resultado indica que alr5283 en *E. coli* convierte hexadecanal y octadecenal en pentadecano y heptadeceno, respectivamente, y por tanto es una aldehído graso descarboxilasa activa.

50 Ejemplo 9. Producción de alcanos y alquenos en *E. coli* a través de expresión heteróloga de orf1594 de *Synechococcus elongatus* PCC7942 y AM1\_4041 de *Acaryochloris marina* MBIC11017

Se optimizaron los codones del ADN genómico que codifica para AM1\_4041 de *Acaryochloris marina* MBIC11017 (YP\_001518340; supuesta descarboxilasa) (SEQ ID NO: 9) para su expresión en *E. coli* (SEQ ID NO: 46), se sintetizó y se clonó en los sitios NdeI y XhoI del vector OP-183 (derivado de pACYC) bajo el control del promotor P<sub>trc</sub>. Se cotransformó el constructo resultante con OP80-PCC7942\_1594 en *E. coli* MG1655 y se hicieron crecer las células a 37°C en medio mínimo M9 complementado con espectinomomicina 100 µg/ml y carbenicilina 100 µg/ml. Se

cultivaron las células y se extrajeron como en el ejemplo 3 y se analizaron mediante CG-EM tal como se describió en el ejemplo 26.

5 Tal como se muestra en la figura 12, células de *E. coli* cotransformadas con los vectores que llevan PCC7942\_1594 de *S. elongatus* y AM1\_4041 de *A. marina* MBIC11017 produjeron los mismos aldehídos grasos y alcoholes grasos que en el ejemplo 3, pero también tridecano, pentadecano, heptadecano y octadecano. Este resultado indica que AM1\_4041 en *E. coli* convierte tetradecanal, hexadecanal, hexadecanal y octadecanal en tridecano, pentadecano, heptadecano y octadecano, respectivamente, y por tanto es una aldehído graso descarboxilasa activa.

Ejemplo 10. Producción de alcanos y alquenos en *E. coli* a través de expresión heteróloga de orf1594 de *Synechococcus elongatus* PCC7942 y BP-1 tll1313 de *Thermosynechococcus elongatus*

10 Se optimizaron los codones del ADN genómico que codifica para BP-1 tll1313 de *Thermosynechococcus elongatus* (NP\_682103; supuesta descarboxilasa) (SEQ ID NO: 11) para su expresión en *E. coli* (SEQ ID NO: 47), se sintetizó y se clonó en los sitios NdeI y XhoI del vector OP-183 (derivado de pACYC) bajo el control del promotor P<sub>trc</sub>. Se cotransformó el constructo resultante con OP80-PCC7942\_1594 en *E. coli* MG1655 y se hicieron crecer las células a 37°C en medio mínimo M9 complementado con espectinomicina 100 µg/ml y carbenicilina 100 µg/ml. Se cultivaron las células y se extrajeron como en el ejemplo 3 y se analizaron mediante CG-EM tal como se describió en el ejemplo 26.

20 Tal como se muestra en la figura 13, células de *E. coli* cotransformadas con los vectores que llevan PCC7942\_1594 de *S. elongatus* y BP-1 tll1313 de *T. elongatus* produjeron los mismos aldehídos grasos y alcoholes grasos que en el ejemplo 3, pero también pentadecano y heptadecano. Este resultado indica que tll1313 en *E. coli* convierte hexadecanal y octadecanal en pentadecano y heptadecano, respectivamente, y por tanto es una aldehído graso descarboxilasa activa.

Ejemplo 11. Producción de alcanos y alquenos en *E. coli* a través de expresión heteróloga de orf1594 de *Synechococcus elongatus* PCC7942 y CYA\_0415 de *Synechococcus* sp. JA-3-3Ab

25 Se optimizaron los codones del ADN genómico que codifica para CYA\_0415 de *Synechococcus* sp. JA-3-3Ab (YP\_473897; supuesta descarboxilasa) (SEQ ID NO: 13) para la expresión en *E. coli* (SEQ ID NO: 48), se sintetizó y se clonó en los sitios NdeI y XhoI del vector OP-183 (derivado de pACYC) bajo el control del promotor P<sub>trc</sub>. Se cotransformó el constructo resultante con OP80-PCC7942\_1594 en *E. coli* MG1655 y se hicieron crecer las células a 37°C en medio mínimo M9 complementado con espectinomicina 100 µg/ml y carbenicilina 100 µg/ml. Se cultivaron las células y se extrajeron como en el ejemplo 3 y se analizaron mediante CG-EM tal como se describió en el ejemplo 26.

35 Tal como se muestra en la figura 14, células de *E. coli* cotransformadas con los vectores que llevan PCC7942\_1594 de *S. elongatus* y CYA\_0415 de *Synechococcus* sp. JA-3-3Ab produjeron los mismos aldehídos grasos y alcoholes grasos que en el ejemplo 3, pero también pentadecano y heptadecano. Este resultado indica que Npun02004178 en *E. coli* convierte hexadecanal y octadecanal en pentadecano y heptadecano, respectivamente, y por tanto es una aldehído graso descarboxilasa activa.

Ejemplo 12. Producción de alcanos y alquenos en *E. coli* a través de expresión heteróloga de orf1594 de *Synechococcus elongatus* PCC7942 y gll3146 de *Gloeobacter violaceus* PCC7421

40 Se amplificó el ADN genómico que codifica para gll3146 de *Gloeobacter violaceus* PCC7421 (NP\_926092; supuesta descarboxilasa) (SEQ ID NO: 15) y se clonó en los sitios NdeI y XhoI del vector OP-183 (derivado de pACYC) bajo el control del promotor P<sub>trc</sub>. Se cotransformó el constructo resultante con OP80-PCC7942\_1594 en *E. coli* MG1655 y se hicieron crecer las células a 37°C en medio mínimo M9 complementado con espectinomicina 100 µg/ml y carbenicilina 100 µg/ml. Se cultivaron las células y se extrajeron como en el ejemplo 3 y se analizaron mediante CG-EM tal como se describió en el ejemplo 1.

45 Tal como se muestra en la figura 15, células de *E. coli* cotransformadas con los vectores que llevan PCC7942\_1594 de *S. elongatus* y gll3146 de *G. violaceus* PCC7421 produjeron los mismos aldehídos grasos y alcoholes grasos que en el ejemplo 3, pero también pentadecano y heptadecano. Este resultado indica que gll3146 en *E. coli* convierte hexadecanal y octadecanal en pentadecano y heptadecano, respectivamente, y por tanto es una aldehído graso descarboxilasa activa.

Ejemplo 13. Producción de alcanos y alquenos en *E. coli* a través de expresión heteróloga de orf1594 de *Synechococcus elongatus* PCC7942 y PMT1231 de *Prochlorococcus marinus* MIT9313

50 Se optimizaron los codones del ADN genómico que codifica para PMT1231 de *Prochlorococcus marinus* MIT9313 (NP\_895059; supuesta descarboxilasa) (SEQ ID NO: 17) para su expresión en *E. coli* (SEQ ID NO: 49), se sintetizó y se clonó en los sitios NdeI y XhoI del vector OP-183 (derivado de pACYC) bajo el control del promotor P<sub>trc</sub>. Se cotransformó el constructo resultante con OP80-PCC7942\_1594 en *E. coli* MG1655 y se hicieron crecer las células a 37°C en medio mínimo M9 complementado con espectinomicina 100 µg/ml y carbenicilina 100 µg/ml. Se cultivaron las células y se extrajeron como en el ejemplo 3 y se analizaron mediante CG-EM tal como se describió en el

ejemplo 26.

Tal como se muestra en la figura 16, células de *E. coli* cotransformadas con los vectores que llevan PCC7942\_1594 de *S. elongatus* y PMT1231 de *P. marinus* MIT9313 produjeron los mismos aldehídos grasos y alcoholes grasos que en el ejemplo 3, pero también pentadecano y heptadeceno. Este resultado indica que PMT 1231 en *E. coli* convierte hexadecanal y octadecenal en pentadecano y heptadeceno, respectivamente, y por tanto es una aldehído graso descarboxilasa activa.

Ejemplo 14. Producción de alcanos y alquenos en *E. coli* a través de expresión heteróloga de orf1594 de *Synechococcus elongatus* PCC7942 y PMM0532 de *Prochlorococcus marinus* CCMP1986

Se amplificó el ADN genómico que codifica para PMM0532 de *Prochlorococcus marinus* CCMP1986 (NP\_892650; supuesta descarboxilasa) (SEQ ID NO: 19) y se clonó en los sitios NdeI y XhoI del vector OP-183 (derivado de pACYC) bajo el control del promotor P<sub>trc</sub>. Se cotransformó el constructo resultante con OP80-PCC7942\_1594 en *E. coli* MG1655 y se hicieron crecer las células a 37°C en medio mínimo M9 complementado con espectinomicina 100 µg/ml y carbenicilina 100 µg/ml. Se cultivaron las células y se extrajeron como en el ejemplo 3 y se analizaron mediante CG-EM tal como se describió en el ejemplo 1.

Tal como se muestra en la figura 17, células de *E. coli* cotransformadas con los vectores que llevan PCC7942\_1594 *S. elongatus* y PMM0532 de *P. marinus* CCMP1986 produjeron los mismos aldehídos grasos y alcoholes grasos que en el ejemplo 3, pero también pentadecano y heptadeceno. Este resultado indica que PMM0532 en *E. coli* convierte hexadecanal y octadecenal en pentadecano y heptadeceno, respectivamente, y por tanto es una aldehído graso descarboxilasa activa.

Ejemplo 15. Producción de alcanos y alquenos en *E. coli* a través de expresión heteróloga de orf1594 de *Synechococcus elongatus* PCC7942 y PMN2A 1863 de *Prochlorococcus marinus* NATL2A

Se optimizaron los codones del ADN genómico que codifica para PMN2A\_1863 de *Prochlorococcus marinus* NATL2A (YP\_293054; supuesta descarboxilasa) (SEQ ID NO: 21) para su expresión en *E. coli* (SEQ ID NO: 51), se sintetizó y se clonó en los sitios NdeI y XhoI del vector OP-183 (derivado de pACYC) bajo el control del promotor P<sub>trc</sub>. Se cotransformó el constructo resultante con OP80-PCC7942\_1594 en *E. coli* MG1655 y se hicieron crecer las células a 37°C en medio mínimo M9 complementado con espectinomicina 100 µg/ml y carbenicilina 100 µg/ml. Se cultivaron las células y se extrajeron como en el ejemplo 3 y se analizaron mediante CG-EM tal como se describió en el ejemplo 26.

Tal como se muestra en la figura 18, células de *E. coli* cotransformadas con los vectores que llevan PCC7942\_1594 de *S. elongatus* y PMN2A\_1863 de *P. marinus* NATL2A produjeron los mismos aldehídos grasos y alcoholes grasos que en el ejemplo 3, pero también pentadecano y heptadeceno. Este resultado indica que PMN2A\_1863 en *E. coli* convierte hexadecanal y octadecenal en pentadecano y heptadeceno, respectivamente, y por tanto es una aldehído graso descarboxilasa activa.

Ejemplo 16. Producción de alcanos y alquenos en *E. coli* a través de expresión heteróloga de orf1594 de *Synechococcus elongatus* PCC7942 y RS9917 09941 de *Synechococcus* sp. RS9917

Se optimizaron los codones del ADN genómico que codifica para RS9917\_09941 de *Synechococcus* sp. RS9917 (ZP\_01079772; supuesta descarboxilasa) (SEQ ID NO: 23) para su expresión en *E. coli* (SEQ ID NO: 52), se sintetizó y se clonó en los sitios NdeI y XhoI del vector OP-183 (derivado de pACYC) bajo el control del promotor P<sub>trc</sub>. Se cotransformó el constructo resultante con OP80-PCC7942\_1594 en *E. coli* MG1655 y se hicieron crecer las células a 37°C en medio mínimo M9 complementado con espectinomicina 100 µg/ml y carbenicilina 100 µg/ml. Se cultivaron las células y se extrajeron como en el ejemplo 3 y se analizaron mediante CG-EM tal como se describió en el ejemplo 26.

Tal como se muestra en la figura 19, células de *E. coli* cotransformadas con los vectores que llevan PCC7942\_1594 de *S. elongatus* y RS9917\_09941 de *Synechococcus* sp. RS9917 produjeron los mismos aldehídos grasos y alcoholes grasos que en el ejemplo 3, pero también pentadecano y heptadeceno. Este resultado indica que RS9917\_09941 en *E. coli* convierte hexadecanal y octadecenal en pentadecano y heptadeceno, respectivamente, y por tanto es una aldehído graso descarboxilasa activa.

Ejemplo 17. Producción de alcanos y alquenos en *E. coli* a través de expresión heteróloga de orf1594 de *Synechococcus elongatus* PCC7942 y RS9917 12945 de *Synechococcus* sp. RS9917

Se optimizaron los codones del ADN genómico que codifica para RS9917\_12945 *Synechococcus* sp. RS9917 (ZP\_01080370; supuesta descarboxilasa) (SEQ ID NO: 25) para su expresión en *E. coli* (SEQ ID NO: 53), se sintetizó y se clonó en los sitios NdeI y XhoI del vector OP-183 (derivado de pACYC) bajo el control del promotor P<sub>trc</sub>. Se cotransformó el constructo resultante con OP80-PCC7942\_1594 en *E. coli* MG1655 y se hicieron crecer las células a 37°C en medio mínimo M9 complementado con espectinomicina 100 µg/ml y carbenicilina 100 µg/ml. Se cultivaron las células y se extrajeron como en el ejemplo 3 y se analizaron mediante CG-EM tal como se describió en el ejemplo 26.

Tal como se muestra en la figura 20, células de *E. coli* cotransformadas con los vectores que llevan PCC7942\_1594 de *S. elongatus* y RS9917\_12945 de *Synechococcus* sp. RS9917 produjeron los mismos aldehídos grasos y alcoholes grasos que en el ejemplo 3, pero también pentadecano y heptadeceno. Este resultado indica que RS9917\_12945 en *E. coli* convierte hexadecanal y octadecenal en pentadecano y heptadeceno, respectivamente, y por tanto es una aldehído graso descarboxilasa activa.

Ejemplo 18. Producción de alcanos y alquenos en *E. coli* a través de expresión heteróloga de orf1594 de *Synechococcus elongatus* PCC7942 y cce\_0778 de *Cyanothece* sp. ATCC51142

Se sintetizó el ADN genómico que codifica para cce\_0778 de *Cyanothece* sp ATCC51142 (YP\_001802195; supuesta descarboxilasa) (SEQ ID NO: 27) y se clonó en los sitios NdeI y XhoI del vector OP-183 (derivado de pACYC) bajo el control del promotor P<sub>trc</sub>. Se cotransformó el constructo resultante con OP80-PCC7942\_1594 en *E. coli* MG1655 y se hicieron crecer las células a 37°C en medio mínimo M9 complementado con espectinomicina 100 µg/ml y carbenicilina 100 µg/ml. Se cultivaron las células y se extrajeron como en el ejemplo 3 y se analizaron mediante CG-EM tal como se describió en el ejemplo 26.

Tal como se muestra en la figura 21, células de *E. coli* cotransformadas con los vectores que llevan PCC7942\_1594 de *S. elongatus* y cce\_0778 de *Cyanothece* sp ATCC51142 produjeron los mismos aldehídos grasos y alcoholes grasos que en el ejemplo 3, pero también tridecano, pentadeceno, pentadecano y heptadeceno. Este resultado indica que cce\_0778 de ATCC51142 en *E. coli* convierte tetradecanal, hexadecenal, hexadecanal y octadecenal en tridecano, pentadeceno, pentadecano y heptadeceno, respectivamente, y por tanto es una aldehído graso descarboxilasa activa.

Ejemplo 19. Producción de alcanos y alquenos en *E. coli* a través de expresión heteróloga de orf1594 de *Synechococcus elongatus* PCC7942 y Cyan 7425\_0398 de *Cyanothece* sp. PCC7425

Se sintetizó el ADN genómico que codifica para Cyan7425\_0398 de *Cyanothece* sp PCC7425 (YP\_002481151; supuesta descarboxilasa) (SEQ ID NO: 29) y se clonó en los sitios NdeI y XhoI del vector OP-183 (derivado de pACYC) bajo el control del promotor P<sub>trc</sub>. Se cotransformó el constructo resultante con OP80-PCC7942\_1594 en *E. coli* MG1655 y se hicieron crecer las células a 37°C en medio mínimo M9 complementado con espectinomicina 100 µg/ml y carbenicilina 100 µg/ml. Se cultivaron las células y se extrajeron como en el ejemplo 3 y se analizaron mediante CG-EM tal como se describió en el ejemplo 26.

Tal como se muestra en la figura 22, células de *E. coli* cotransformadas con los vectores que llevan PCC7942\_1594 de *S. elongatus* y Cyan7425\_0398 de *Cyanothece* sp PCC7425 produjeron los mismos aldehídos grasos y alcoholes grasos que en el ejemplo 3, pero también tridecano, pentadeceno, pentadecano y heptadeceno. Este resultado indica que Cyan7425\_0398 en *E. coli* convierte tetradecanal, hexadecenal, hexadecanal y octadecenal en tridecano, pentadeceno, pentadecano y heptadeceno, respectivamente, y por tanto es una aldehído graso descarboxilasa activa.

Ejemplo 20. Producción de alcanos y alquenos en *E. coli* a través de expresión heteróloga de orf1594 de *Synechococcus elongatus* PCC7942 y Cyan7425\_2986 de *Cyanothece* sp. PCC7425

Se sintetizó el ADN genómico que codifica para Cyan7425\_2986 de *Cyanothece* sp PCC7425 (YP\_002483683; supuesta descarboxilasa) (SEQ ID NO: 31) y se clonó en los sitios NdeI y XhoI del vector OP-183 (derivado de pACYC) bajo el control del promotor P<sub>trc</sub>. Se cotransformó el constructo resultante con OP80-PCC7942\_1594 en *E. coli* MG1655 y se hicieron crecer las células a 37°C en medio mínimo M9 complementado con espectinomicina 100 µg/ml y carbenicilina 100 µg/ml. Se cultivaron las células y se extrajeron como en el ejemplo 3 y se analizaron mediante CG-EM tal como se describió en el ejemplo 26.

Tal como se muestra en la figura 23, células de *E. coli* cotransformadas con los vectores que llevan PCC7942\_1594 de *S. elongatus* y Cyan7425\_2986 de *Cyanothece* sp PCC7425 produjeron los mismos aldehídos grasos y alcoholes grasos que en el ejemplo 3, pero también tridecano, pentadeceno, pentadecano y heptadeceno. Este resultado indica que Cyan7425\_2986 en *E. coli* convierte tetradecanal, hexadecenal, hexadecanal y octadecenal en tridecano, pentadeceno, pentadecano y heptadeceno, respectivamente, y por tanto es una aldehído graso descarboxilasa activa.

Ejemplo 21. Producción de alcanos y alquenos en *E. coli* a través de expresión heteróloga de PMM0533 de *Prochlorococcus marinus* CCMP1986 y PMM0532 de *Prochlorococcus marinus* CCMP1986

Se amplificó el ADN genómico que codifica para PMM0533 de *P. marinus* CCMP1986 (NP\_892651; supuesta enzima generadora de aldehído) (SEQ ID NO: 71) y PMM0532 de *Prochlorococcus marinus* CCMP1986 (NP\_892650; supuesta descarboxilasa) (SEQ ID NO: 19) y se clonó en los sitios NcoI y EcoRI del vector OP-80 y los sitios NdeI y XhoI del vector OP-183, respectivamente. Se transformaron por separado los constructos resultantes y se cotransformaron en *E. coli* MG1655 y se hicieron crecer las células a 37°C en medio mínimo M9 complementado con espectinomicina 100 µg/ml y carbenicilina 100 µg/ml. Se cultivaron las células y se extrajeron como en el ejemplo 3 y se analizaron mediante CG-EM tal como se describió en el ejemplo 26.

Tal como se muestra en la figura 24A, células de *E. coli* células transformadas con sólo el vector que lleva PMM0533 de *P. marinus* CCMP1986 no produjeron ningún aldehído graso o alcohol graso. Sin embargo, células de *E. coli* cotransformadas con vectores que llevan PMM0533 y PMM0532 produjeron hexadecanol, pentadecano y heptadeceno (figura 24B). Este resultado indica que PMM0533 sólo proporciona sustratos de aldehído graso para la reacción de descarbonilación cuando interacciona con una descarbonilasa, tal como PMM0532.

Ejemplo 22. Producción de alcanos y alquenos en una cepa de *E. coli* productora de acil graso-CoA a través de expresión heteróloga de orf1594 de *Synechococcus elongatus* PCC7942 y AM1\_4041 de *Acaryochloris marina* MBIC11017

Se sintetizó el ADN genómico que codifica para AM1\_4041 de *Acaryochloris marina* MBIC1 1017 (YP\_001518340; supuesta aldehído graso descarbonilasa) (SEQ ID NO: 9) y se clonó en los sitios NdeI y XhoI del vector OP-183 (derivado de pACYC) bajo el control del promotor P<sub>trc</sub>. Se cotransformó el constructo resultante con OP80-PCC7942\_1594 en *E. coli* MG1655 ΔfadE lacZ::P<sub>trc</sub>'tesA-fadD. Esta cepa expresa una versión citoplasmática de la tioesterasa de *E. coli*, 'TesA, y la acil-CoA sintetasa de *E. coli*, FadD, bajo el control del promotor P<sub>trc</sub>, y por tanto produce acil graso-CoA. Se cultivaron las células a 37°C en medio mínimo M9 complementado con espectinomicina 100 μg/ml y carbenicilina 100 μg/ml. Se cultivaron las células y se extrajeron como en el ejemplo 3 y se analizaron mediante CG-EM tal como se describió en el ejemplo 1.

Tal como se muestra en la figura 25, estas células de *E. coli* cotransformadas con PCC7942\_1594 de *S. elongatus* y AM1\_4041 de *A. marina* MBIC11017 también produjeron alcanos y alcoholes grasos. Este resultado indica que PCC7942\_1594 de *S. elongatus* puede usar acil-CoA como sustrato para producir hexadecenal, hexadecanal y octadecenal, que entonces se convierte en pentadeceno, pentadecano y heptadeceno, respectivamente, mediante AM1\_4041 de *A. marina* MBIC11017.

Ejemplo 23. Producción de alcanos y alquenos en una cepa de *E. coli* productora de acil graso-CoA a través de expresión heteróloga de slI0209 de *Synechocystis* sp. PCC6803 y slI0208 de *Synechocystis* sp. PCC6803

Se sintetizó el ADN genómico que codifica para slI0208 de *Synechocystis* sp. PCC6803 (NP\_442147; supuesta aldehído graso descarbonilasa) (SEQ ID NO: 3) y se clonó en los sitios NdeI y XhoI del vector OP-183 (derivado de pACYC) bajo el control del promotor P<sub>trc</sub>. Se sintetizó el ADN genómico que codifica para *Synechocystis* sp. PCC6803 slI0209 (NP\_442146; acil-ACP reductasa) (SEQ ID NO: 67) y se clonó en los sitios NcoI y EcoRI del vector OP-183 (derivado de pACYC) bajo el control del promotor P<sub>trc</sub>. Se cotransformaron los constructos resultantes en *E. coli* MG1655 ΔfadE lacZ::P<sub>trc</sub>'tesA-fadD. Esta cepa expresa una versión citoplasmática de la tioesterasa de *E. coli*, 'TesA, y la acil-CoA sintetasa de *E. coli*, FadD, bajo el control del promotor P<sub>trc</sub>, y por tanto produce acil graso-CoA. Se cultivaron las células a 37°C en medio mínimo M9 complementado con espectinomicina 100 μg/ml y carbenicilina 100 μg/ml. Se cultivaron las células y se extrajeron como en el ejemplo 3 y se analizaron mediante CG-EM tal como se describió en el ejemplo 26.

Tal como se muestra en la figura 26, estas células de *E. coli* transformadas con slI0209 de *Synechocystis* sp. PCC6803 no produjeron ningún aldehído graso o alcohol graso. Sin embargo, cuando se cotransformaron con slI0208 y slI0209 de *Synechocystis* sp. PCC6803, produjeron alcanos, aldehídos grasos y alcoholes grasos. Este resultado indica que slI0209 de *Synechocystis* sp. PCC6803 puede usarse acil-CoA como sustrato para producir aldehídos grasos tales como tetradecanal, hexadecanal y octadecenal, pero sólo cuando se coexpresa con una aldehído graso descarbonilasa. Los aldehídos grasos aparentemente se reducen adicionalmente a los correspondientes alcoholes grasos, tetradecanol, hexadecanol y octadecanol, mediante una actividad aldehído reductasa de *E. coli* endógena. En este experimento, se convirtió octadecenal en heptadeceno mediante slI0208 de *Synechocystis* sp. PCC6803.

Ejemplo 24. Producción de alcanos y alquenos en una cepa de *E. coli* productora de aldehído graso a través de expresión heteróloga de Npun02004178 de *Nostoc punctiforme* PCC73102 y varios de sus homólogos

Se amplificó el ADN genómico que codifica para Npun02004178 de *Nostoc punctiforme* PCC73102 (ZP\_00108838; supuesta aldehído graso descarbonilasa) (SEQ ID NO: 5) y se clonó en los sitios NdeI y XhoI del vector OP-183 (derivado de pACYC) bajo el control del promotor P<sub>trc</sub>. Se amplificó el ADN genómico que codifica para el orf MSMEG\_5739 de la cepa de *Mycobacterium smegmatis* MC2 155 (YP\_889972, supuesta ácido carboxílico reductasa) (SEQ ID NO: 85) y se clonó en los sitios NcoI y EcoRI del vector OP-180 (derivado de pCL1920) bajo el control del promotor P<sub>trc</sub>. Se cotransformaron los dos constructos resultantes en *E. coli* MG1655 ΔfadD lacZ::P<sub>trc</sub>-'tesA. En esta cepa, MSMEG\_5739 proporcionó aldehídos grasos, que reduce ácidos grasos libres (formados por la acción de 'TesA) a aldehídos grasos. Se hicieron crecer las células a 37°C en medio mínimo M9 complementado con espectinomicina 100 μg/ml y carbenicilina 100 μg/ml. Se cultivaron las células y se extrajeron como en el ejemplo 3 y se analizaron mediante CG-EM tal como se describió en el ejemplo 1.

Tal como se muestra en la figura 27, estas células de *E. coli* cotransformadas con los vectores que llevan Npun02004178 de *N. punctiforme* PCC73102 y MSMEG\_5739 de la cepa de *M. smegmatis* MC2 155 produjeron tridecano, pentadeceno y pentadecano. Este resultado indica que Npun02004178 en *E. coli* convierte tetradecanal, hexadecenal y hexadecanal proporcionados por la ácido carboxílico reductasa MSMEG\_5739 en tridecano,



25 ml de diclorometano. A continuación, se añadieron 200 mg de clorocromato de piridinio (TCI America) a la disolución y se agitó durante la noche. Se secó la mezcla de reacción a vacío para eliminar el diclorometano. Se resuspendieron los productos restantes en hexano y se filtraron a través de papel de filtro Whatman. Entonces se secó el filtrado a vacío y se resuspendió en 5 ml de hexano y se purificó mediante cromatografía ultrarrápida en sílice. Se cargó la mezcla sobre la columna alimentada por gravedad en hexano y luego se lavó con dos volúmenes de columna de hexano. Entonces se eluyó el octadecanal con una mezcla 8:1 de hexano y acetato de etilo. Se agruparon las fracciones que contenían octadecanal y se analizaron usando los métodos de CG/EM descritos a continuación. El producto final era el 95% puro tal como se determinó mediante este método.

Para someter a prueba la proteína Npun02004178 para determinar la actividad de descarbonilación, se establecieron los siguientes ensayos enzimáticos. Se establecieron reacciones de 200  $\mu$ l en tampón fosfato de sodio 100 mM a pH 7,2 con los siguientes componentes a sus concentraciones finales respectivas: 30  $\mu$ M de proteína Npun02004178 purificada, octadecanal 200  $\mu$ M, ferredoxina de espinaca 0,11  $\mu$ g/ml (Sigma), ferredoxina reductasa de espinaca 0,05 unidades/ml (Sigma) y NADPH 1 mM (Sigma). Los controles negativos incluían la reacción anterior sin Npun02004178, la reacción anterior sin octadecanal y la reacción anterior sin ferredoxina de espinaca, ferredoxina reductasa y NADPH. Se incubó cada reacción a 37°C durante 2 h antes de extraerse con 100  $\mu$ l de acetato de etilo. Se analizaron las muestras mediante CG/EM usando los siguientes parámetros: tiempo de ejecución: 13,13 min; columna: HP-5-EM n.º de parte 19091S-433E (longitud de 30 metros; D.I.: orificio estrecho de 0,25 mm; película: 0,25 $\mu$ M); inyección: entrada 1 il Agilent 6850; entrada: 300 C sin división; gas portador: helio; flujo: 1,3 ml/min; temp. del horno: 75°C mantenidos 5 min, 320 a 40°C/min, 320 mantenidos 2 min; det.: Agilent 5975B VL MSD; temp. de det.: 330°C; barrido: 50-550 M/Z. Se usó heptadecano de Sigma como referencia auténtica para determinar el tiempo de retención de los compuestos y el patrón de fragmentación.

Tal como se muestra en la figura 31, se observó conversión *in vitro* de octadecanal en heptadecano en presencia de Npun02004178. La descarbonilación enzimática de octadecanal mediante Npun02004178 era dependiente de la adición de ferredoxina de espinaca reductasa, ferredoxina y NADPH.

A continuación, se determinó si ferredoxinas y ferredoxina reductasas cianobacterianas pueden reemplazar a las proteínas de espinaca en el ensayo de aldehído graso descarbonilasa *in vitro*. Se clonaron por separado los siguientes cuatro genes en los sitios NdeI y XhoI de pET-15b: Npun02003626 de *N. punctiforme* PCC73102 (ZP\_00109192, ferredoxina oxidoreductasa petH sin el dominio de ligador de aloficocianina n-terminal) (SEQ ID NO: 87), Npun02001001 de *N. punctiforme* PCC73102 (ZP\_00111633, ferredoxina 1) (SEQ ID NO: 89), Npun02003530 de *N. punctiforme* PCC73102 (ZP\_00109422, ferredoxina 2) (SEQ ID NO: 91) y Npun02003123 de *N. punctiforme* PCC73102 (ZP\_00109501, ferredoxina 3) (SEQ ID NO: 93). Se expresaron las cuatro proteínas y se purificaron tal como se describió anteriormente. Se obtuvieron 1 mg/ml de cada ferredoxina y 4 mg/ml de la ferredoxina oxidoreductasa. Se sometieron a prueba las tres ferredoxinas cianobacterianas con la ferredoxina oxidoreductasa cianobacterianas usando la configuración enzimática descrita anteriormente con los siguientes cambios. La concentración final de la ferredoxina reductasa era de 60  $\mu$ g/ml y las ferredoxinas estaban a 50  $\mu$ g/ml. Las reacciones enzimáticas extraídas fueron mediante CG/EM usando los siguientes parámetros: tiempo de ejecución: 6,33 min; columna: J&W 122-5711 DB-5ht (longitud de 15 metros; D.I.: orificio estrecho de 0,25 mm; película: 0,10  $\mu$ M); inyección: entrada de 1  $\mu$ l Agilent 6850; entrada: 300°C sin división; gas portador: helio; flujo: 1,3 ml/min; temp. del horno: 100°C mantenidos 0,5 min, 260 a 30°C/min, 260 mantenidos 0,5 min; det.: Agilent 5975B VL MSD; temp. de det.: 230°C; barrido: 50-550 M/Z.

Tal como se muestra en la figura 32, se observó conversión *in vitro* dependiente de Npun02004178 de octadecanal en heptadecano en presencia de NADPH y la ferredoxina oxidoreductasa cianobacterianas y cualquiera de las tres ferredoxinas cianobacterianas.

#### Ejemplo 27. Caracterización bioquímica de orf1594 de *Synechococcus elongatus* PCC7942

Se clonó el ADN genómico que codifica para orf1594 de *S. elongatus* PCC7492 (SEQ ID NO: 65) en los sitios NcoI y XhoI del vector pET-28b bajo el control del promotor de T7. La proteína PCC7942\_orf1594 resultante contenía una etiqueta de His C-terminal. Se transformó una cepa de *E. coli* BL21 (DE3) (Invitrogen) con el plásmido y se expresó la proteína PCC7942\_orf1594 y se purificó tal como se describió en el ejemplo 22. Se almacenó la disolución de proteína en el siguiente tampón: fosfato de sodio 50 mM, pH 7,5, NaCl 100 mM, THP 1 mM, glicerol al 10%. Se determinó la concentración de proteína usando el ensayo Bradford (Biorad). Se obtuvieron 2 mg/ml de proteína PCC7942\_orf1594.

Para someter a prueba la proteína PCC7942\_orf1594 para determinar la actividad acil-ACP o acil-CoA reductasa, se establecieron los siguientes ensayos enzimáticos. Se establecieron reacciones de 100  $\mu$ l en tampón Tris-HCl 50 mM a pH 7,5 con los siguientes componentes a sus concentraciones finales respectivas: 10  $\mu$ M de proteína PCC7942\_orf1594 purificada, acil-CoA o acil-ACP 0,01-1 mM, MgCl<sub>2</sub> 2 mM, NADPH 0,2-2 mM. Se incubaron las reacciones durante 1 h a 37°C y se detuvieron añadiendo 100  $\mu$ l de acetato de etilo (que contenía 1-octadeceno 5 mg/l como patrón interno). Se agitaron con vórtex las muestras durante 15 min y se centrifugaron a velocidad máx. durante 3 min para la separación de fases. Se transfirieron 80  $\mu$ l de la capa superior a viales de vidrio de CG y se analizaron mediante CG/EM tal como se describió en el ejemplo 26. Se calculó la cantidad de aldehído formada

basándose en el patrón interno.

Tal como se muestra en la figura 33, PCC7942\_orf1594 podía reducir octadecanoil-CoA a octadecanal. La actividad reductasa requería cationes divalentes tales como  $Mg^{2+}$ ,  $Mn^{2+}$  o  $Fe^{2+}$  y NADPH como donador de electrones. NADH no soportaba actividad reductasa. PCC7942\_orf1594 también podía reducir octadecanoil-CoA y octadecenoil-ACP a octadecenal. Se determinaron los valores de  $K_m$  para la reducción de octadecanoil-CoA, octadecenoil-CoA y octadecenoil-ACP en presencia de NADPH 2 mM como  $45 \pm 20 \mu M$ ,  $82 \pm 22 \mu M$  y  $7,8 \pm 2 \mu M$ , respectivamente. Estos resultados demuestran que PCC7942\_orf1594, *in vitro*, reduce tanto acil-CoA como acil-ACP y que la enzima tiene aparentemente una mayor afinidad por acil-ACP en comparación con acil-CoA. El valor de  $K_m$  para NADPH en presencia de octadecanoil-CoA 0,5 mM para PCC7942\_orf1594 se determinó como  $400 \pm 80 \mu M$ .

A continuación, se examinó la transferencia de hidruro estereoespecífica desde NADPH hasta un aldehído graso catalizada por PCC7942\_orf1594. Se preparó deuterio-NADPH según el siguiente protocolo. Se añadieron 5 mg de  $NADP^+$  y 3,6 mg de D-glucosa-1-d a 2,5 ml de tampón fosfato de sodio 50 mM (pH 7,0). Se inició la producción enzimática de NADPH marcado mediante la adición de 5 unidades de glucosa deshidrogenasa de o bien *Bacillus megaterium* (USB Corporation) para la producción de R-(4- $^2H$ )NADPH o bien *Thermoplasma acidophilum* (Sigma) para la producción de S-(4- $^2H$ )NADPH. Se incubó la reacción durante 15 min a 37°C, se filtró por centrifugación usando un filtro de centrifuga Amicon Ultra de MWCO de 10 KDa (Millipore), se congeló inmediatamente sobre hielo seco y se almacenó a -80°C.

La reacción de ensayo *in vitro* contenía Tris-HCl 50 mM (pH 7,5), 10  $\mu M$  de proteína PCC7942\_orf1594 purificada, octadecanoil-CoA 1 mM,  $MgCl_2$  2 mM y 50  $\mu l$  de deuterio-NADPH (preparado tal como se describió anteriormente) en un volumen total de 100  $\mu l$ . Tras una incubación de 1 h, se extrajo el producto de la reacción enzimática y se analizó tal como se describió anteriormente. El aldehído graso resultante detectado mediante CG/EM era octadecanal (véase la figura 34). Debido a que la transferencia de hidruro desde NADPH es estereoespecífica, se sintetizaron tanto R-(4- $^2H$ )NADPH como S-(4- $^2H$ )NADPH. Se observó octadecanal con una masa unitaria de uno positivo usando sólo el S-(4- $^2H$ )NADPH. El hecho de que el aldehído graso estaba marcado indica que el hidrógeno deuterado se ha transferido desde el NADPH marcado hasta el aldehído graso marcado. Esto demuestra que se usa NADPH en esta reacción enzimática y que la transferencia de hidruro catalizada por PCC7942\_orf1594 es estereoespecífica.

#### Ejemplo 28. Producción intracelular y extracelular de aldehídos grasos y alcoholes grasos en *E. coli* a través de expresión heteróloga de orf1594 de *Synechococcus elongatus* PCC7942

Se amplificó el ADN genómico que codifica para orf1594 de *Synechococcus elongatus* PCC7942 (YP\_400611; acil-ACP reductasa) (SEQ ID NO: 65) y se clonó en los sitios NcoI y EcoRI del vector OP-80 (derivado de pCL1920) bajo el control del promotor  $P_{trc}$ . Se cotransformó el constructo resultante en *E. coli* MG1655  $\Delta fadE$  y se hicieron crecer las células a 37°C en 15 ml de medio mínimo Che-9 con glucosa al 3% (p/v) como fuente de carbono y complementado con espectinomicina y carbenicilina 100  $\mu g/ml$ , respectivamente. Cuando el cultivo alcanzó una  $DO_{600}$  de 0,8-1,0, se indujo con IPTG 1 mM y se hicieron crecer las células durante 24-48 h adicionales a 37°C. El medio mínimo Chc-9 se define como:  $Na_2HPO_4$  6 g/l,  $KH_2PO_4$  3 g/l, NaCl 0,5 g/l,  $NH_4Cl$  2 g/l,  $MgSO_4$  0,25 g/l x 7  $H_2O$ ,  $CaCl_2$  11 mg/l,  $Fe_3Cl$  27 mg/l x 6  $H_2O$ ,  $ZnCl_2$  2 mg/l x 4  $H_2O$ ,  $Na_2MoO_4$  2 mg/l x 2  $H_2O$ ,  $CuSO_4$  1,9 mg/l x 5  $H_2O$ ,  $H_3BO_3$  0,5 mg/l, tiamina 1 mg/l, Bis-Tris 200 mM (pH 7,25) y Triton-X100 al 0,1% (v/v). Cuando el cultivo alcanzó una  $DO_{600}$  de 1,0-1,2, se indujo con IPTG 1 mM y se permitió que las células crecieran durante 40 h adicionales a 37°C. Se extrajeron células a partir de 0,5 ml de cultivo con 0,5 ml de acetato de etilo para determinar la producción de hidrocarburos totales tal como se describió en el ejemplo 26. Adicionalmente, se separaron las células y el sobrenadante mediante centrifugación (4.000 g a TA durante 10 min) y se extrajeron por separado.

El cultivo produjo 620 mg/l de aldehídos grasos (tetradecanal, heptadecenal, heptadecanal y octadecenal) y 1670 mg/l de alcoholes grasos (dodecanol, tetradecanol, tetradecanol, heptadecanol, heptadecanol y octadecanol). La figura 35 muestra el cromatograma del sobrenadante extraído. Se determinó que el 73% de los aldehídos grasos y alcoholes grasos estaban en el sobrenadante libre de células.

#### Ejemplo 29. Producción intracelular y extracelular de alcanos y alquenos en *E. coli* a través de expresión heteróloga de orf1594 de *Synechococcus elongatus* PCC7942 y Npun02004178 de *Nostoc punctiforme* PCC73102

Se amplificó el ADN genómico que codifica para orf1594 de *Synechococcus elongatus* PCC7942 (YP\_400611; acil-ACP reductasa) (SEQ ID NO: 65) y se clonó en los sitios NcoI y EcoRI del vector OP-80 (derivado de pCL1920) bajo el control del promotor  $P_{trc}$ . Se amplificó el ADN genómico que codifica para Npun02004178 de *Nostoc punctiforme* PCC73102 (ZP\_00108838; aldehído graso descarboxilasa) (SEQ ID NO: 5) y se clonó en los sitios NdeI y XhoI del vector OP-183 (derivado de pACYC) bajo el control del promotor  $P_{trc}$ . Se cotransformaron los constructos resultantes en *E. coli* MG1655  $\Delta fadE$  y se hicieron crecer las células a 37°C en 15 ml de medio mínimo Che9 con glucosa al 3% (p/v) como fuente de carbono y complementado con espectinomicina y carbenicilina 100  $\mu g/ml$ , respectivamente. Se hicieron crecer las células, se separaron del caldo, se extrajeron y se analizaron tal como se describió en el ejemplo 28.

El cultivo produjo 323 mg/l de alcanos y alquenos (tridecano, pentadeceno, pentadecano y heptadeceno), 367 mg/l de aldehídos grasos (tetradecanal, heptadecenal, heptadecanal y octadecenal) y 819 mg/l de alcoholes grasos

(tetradecanol, heptadecanol, heptadecanol y octadecanol). La figura 36 muestra el cromatograma del sobrenadante extraído. Se determinó que el 86% de los alcanos, alquenos, aldehídos grasos y alcoholes grasos estaban en el sobrenadante libre de células.

Ejemplo 30. Producción de alcanos y alquenos en *E. coli* a través de expresión heteróloga de alr5284 de *Nostoc sp.* PCC7210 y alr5283 de *Nostoc sp.* PCC7210

5

Se amplificó el ADN genómico que codifica para alr5284 de *Nostoc sp.* PCC7210 (NP\_489324; supuesta enzima generadora de aldehído) (SEQ ID NO: 81) y se clonó en los sitios NcoI y EcoRI del vector OP-80 (derivado de pCL1920) bajo el control del promotor P<sub>trc</sub>. Se amplificó el ADN genómico que codifica para alr5283 de *Nostoc sp.* PCC7210 (NP\_489323; supuesta descarboxilasa) (SEQ ID NO: 7) y se clonó en los sitios NdeI y XhoI del vector OP-183 (derivado de pACYC) bajo el control del promotor P<sub>trc</sub>. Se cotransformaron los constructos resultantes en *E. coli* MG1655 y se hicieron crecer las células a 37°C en 15 ml de medio mínimo Che9 con glucosa al 3% (p/v) como fuente de carbono y complementado con espectinomicina y carbenicilina 100 µg/ml, respectivamente (tal como se describió en el ejemplo 28). Se extrajeron células a partir de 0,5 ml de cultivo y se analizaron tal como se describió en el ejemplo 3 y se analizaron mediante CG-EM tal como se describió en el ejemplo 26.

10

15

Tal como se muestra en la figura 37, células de *E. coli* cotransformadas con los vectores que llevaban alr5284 de *Nostoc sp.* PCC7210 y alr5283 de *Nostoc sp.* PCC7210 produjeron tridecano, pentadeceno, pentadecano, tetradecanol y hexadecanol. Este resultado indica que la coexpresión de alr5284 y alr5283 de *Nostoc sp.* PCC7210 es suficiente para que *E. coli* produzca alcoholes grasos, alcanos y alquenos.

20

25

30

35

40

**Lista de secuencias**

- <110> LS9, Inc
- <120> MÉTODOS Y COMPOSICIONES PARA PRODUCIR HIDROCARBUROS
- <130> EP73108IFZRPpau
- 5 <140> aún sin asignar
- <141> 18-05-2009
- <150> 09747777.2
- <151> 18-05-2009
- <150> Documento PCT/US2009/044409
- 10 <151> 18-05-2009
- <150> 61/053.955
- <151> 16-05-2008
- <160> 100
- <170> Patent In versión 3.5
- 15 <210> 1
- <211> 696
- <212> ADN
- <213> *Synechococcus elongatus*
- <400> 1
- atgcgcgagc ttgaagccag ccttgaactg gactttcaaa gcgagtccca caaagacgct 60
- tacagccgca tcaacgcgat cgtgattgaa ggcaacaag aggcgttcga caactacaat 120
- cgccttgctg agatgctgcc cgaccagcgg gatgagcttc acaagctagc caagatggaa 180
- cagcgccaca tgaaaggcctt tatggcctgt ggcaaaaatc tctccgtcac tcttgacatg 240
- ggttttgccc agaaatTTTT cgagcgttg cagcagaact tcaaagcggc ggctgaggaa 300
- ggcaaggctg tcacctgcct actgattcaa togctaataca togagtgctt tgcgatcgcg 360
- gcttacaaca tctacatccc agtggcggat gcttttgccc gcaaaatcac ggagggggtc 420
- gtgcgcgacg aatacctgca ccgcaacttc ggtgaagagt ggctgaaggc gaatTTTgat 480
- gcttccaaag ccgaactgga agaagccaat cgtcagaacc tgcccttggg ttggctaatag 540
- ctcaacgaag tggccgatga tgctcgcgaa ctcgggatgg agcgtgagtc gctcgtcgag 600
- gactttatga ttgcctacgg tgaagctctg gaaaacatcg gcttcacaac gcgcgaaatc 660
- 20 atgcgtatgt ccgcctatgg ccttgccggcc gtttga 696
- <210> 2
- <211> 231
- <212> PRT
- <213> *Synechococcus elongatus*
- 25 <400> 2

ES 2 696 773 T3

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

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

Gln Glu Ala Phe Asp Asn Tyr Asn Arg Leu Ala Glu Met Leu Pro Asp  
 35 40 45

Gln Arg Asp Glu Leu His Lys Leu Ala Lys Met Glu Gln Arg His Met  
 50 55 60

Lys Gly Phe Met Ala Cys Gly Lys Asn Leu Ser Val Thr Pro Asp Met  
 65 70 75 80

Gly Phe Ala Gln Lys Phe Phe Glu Arg Leu His Glu Asn Phe Lys Ala  
 85 90 95

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

Ile Ile Glu Cys Phe Ala Ile Ala Ala Tyr Asn Ile Tyr Ile Pro Val  
 115 120 125

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

Tyr Leu His Arg Asn Phe Gly Glu Glu Trp Leu Lys Ala Asn Phe Asp  
 145 150 155 160

Ala Ser Lys Ala Glu Leu Glu Glu Ala Asn Arg Gln Asn Leu Pro Leu  
 165 170 175

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

Met Glu Arg Glu Ser Leu Val Glu Asp Phe Met Ile Ala Tyr Gly Glu  
 195 200 205

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

Ala Tyr Gly Leu Ala Ala Val  
 225 230

<210> 3

<211> 696

ES 2 696 773 T3

<212> ADN

<213> *Synechocystis sp.*

<400> 3

atgcccgagc ttgctgtccg caccgaattt gactattcca gcgaaattta caaagacgcc	60
tatagccgca tcaacgccat tgtcattgaa ggccaacagg aagcctacag caactacctc	120
cagatggcgg aactcttgcc ggaagacaaa gaagagttga cccgcttggc caaaatggaa	180
aaccgccata aaaaaggttt ccaagcctgt ggcaacaacc tccaagtgaa ccctgatatg	240
ccctatgccc aggaatTTTT cgcgggtctc catggcaatt tccagcaagc ttttagcgaa	300
gggaaagttg ttacctgttt attgatccag gctttgatta tcgaagcttt tgcgatcgcc	360
gcctataaca tatatatccc tgtggcggac gactttgctc ggaaaatcac tgagggcgta	420
gtcaaggacg aatacaccca cctcaactac ggggaagaat ggctaaaggc caactttgcc	480
accgctaagg aagaactgga gcaggccaac aaagaaaacc tacccttagt gtggaaaatg	540
ctcaaccaag tgcaggggga cgccaaggta ttgggcatgg aaaaagaagc cctagtggaa	600
gattttatga tcagctacgg cgaagccctc agtaacatcg gcttcagcac cagggaaatt	660
atgcgtatgt cttcctacgg tttggccgga gtctag	696

5 <210> 4

<211> 231

<212> PRT

<213> *Synechocystis sp.*

<400> 4

ES 2 696 773 T3

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

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

Gln Glu Ala Tyr Ser Asn Tyr Leu Gln Met Ala Glu Leu Leu Pro Glu  
 35 40 45

Asp Lys Glu Glu Leu Thr Arg Leu Ala Lys Met Glu Asn Arg His Lys  
 50 55 60

Lys Gly Phe Gln Ala Cys Gly Asn Asn Leu Gln Val Asn Pro Asp Met  
 65 70 75 80

Pro Tyr Ala Gln Glu Phe Phe Ala Gly Leu His Gly Asn Phe Gln His  
 85 90 95

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

Ile Ile Glu Ala Phe Ala Ile Ala Ala Tyr Asn Ile Tyr Ile Pro Val  
 115 120 125

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

Tyr Thr His Leu Asn Tyr Gly Glu Glu Trp Leu Lys Ala Asn Phe Ala  
 145 150 155 160

Thr Ala Lys Glu Glu Leu Glu Gln Ala Asn Lys Glu Asn Leu Pro Leu  
 165 170 175

Val Trp Lys Met Leu Asn Gln Val Gln Gly Asp Ala Lys Val Leu Gly  
 180 185 190

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

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

Ser Tyr Gly Leu Ala Gly Val  
 225 230

<210> 5

<211> 699

5 <212> ADN

<213> *Nostoc punctiforme*

ES 2 696 773 T3

<400> 5

```
atgcagcagc ttacagacca atctaaagaa ttagatttca agagcgaac atacaaagat      60
gcttatagcc ggattaatgc gatcgtgatt gaaggggaac aagaagccca tgaaaattac      120
atcacactag cccaactgct gccagaatct catgatgaat tgattcgctt atccaagatg      180
gaaagccgcc ataagaaagg atttgaagct tgtgggcgca atttagctgt taccocagat      240
ttgcaatttg ccaaagagtt tttctccggc ctacaccaa attttcaaac agctgccgca      300
gaagggaaag tggttacttg tctgttgatt cagtctttaa ttattgaatg ttttgcgatc      360
gcagcatata acatttacat ccccgttgcc gacgatttgc cccgtaaaat tactgaagga      420
gtagttaaag aagaatacag ccacctcaat tttggagaag tttggttga agaacacttt      480
gcagaatcca aagctgaact tgaacttgca aatcgccaga acctaccat cgtctggaaa      540
atgctcaacc aagtagaagg tgatgccac acaatggcaa tggaaaaaga tgctttggta      600
gaagacttca tgattcagta tggtagaagca ttgagtaaca ttggtttttc gactcgcgat      660
attatgcgct tgtcagccta cggactcata ggtgcttaa                               699
```

<210> 6

<211> 232

5 <212> PRT

<213> *Nostoc punctiforme*

<400> 6

ES 2 696 773 T3

Met Gln Gln Leu Thr Asp Gln Ser Lys Glu Leu Asp Phe Lys Ser Glu  
 1 5 10 15

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

Glu Gln Glu Ala His Glu Asn Tyr Ile Thr Leu Ala Gln Leu Leu Pro  
 35 40 45

Glu Ser His Asp Glu Leu Ile Arg Leu Ser Lys Met Glu Ser Arg His  
 50 55 60

Lys Lys Gly Phe Glu Ala Cys Gly Arg Asn Leu Ala Val Thr Pro Asp  
 65 70 75 80

Leu Gln Phe Ala Lys Glu Phe Phe Ser Gly Leu His Gln Asn Phe Gln  
 85 90 95

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

Leu Ile Ile Glu Cys Phe Ala Ile Ala Ala Tyr Asn Ile Tyr Ile Pro  
 115 120 125

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

Glu Tyr Ser His Leu Asn Phe Gly Glu Val Trp Leu Lys Glu His Phe  
 145 150 155 160

Ala Glu Ser Lys Ala Glu Leu Glu Leu Ala Asn Arg Gln Asn Leu Pro  
 165 170 175

Ile Val Trp Lys Met Leu Asn Gln Val Glu Gly Asp Ala His Thr Met  
 180 185 190

Ala Met Glu Lys Asp Ala Leu Val Glu Asp Phe Met Ile Gln Tyr Gly  
 195 200 205

Glu Ala Leu Ser Asn Ile Gly Phe Ser Thr Arg Asp Ile Met Arg Leu  
 210 215 220

Ser Ala Tyr Gly Leu Ile Gly Ala  
 225 230

<210> 7

ES 2 696 773 T3

<211> 696

<212> ADN

<213> *Nostoc sp.*

<400> 7

```

atgcagcagg ttgcagccga tttagaaatt gatttcaaga gcgaaaaata taaagatgcc      60
tatagtcgca taaatgcatg cgtgattgaa ggggaacaag aagcatatga gaattacatt      120
caactatccc aactgctgcc agacgataaa gaagacctaa ttgcctctc gaaaatggaa      180
agccgtcaca aaaaaggatt tgaagcttgt ggacggaacc tacaagtatc accagatatg      240
gagtttgcca aagaattctt tgctggacta cacggtaact tccaaaaagc ggcggctgaa      300
ggtaaaatcg ttacctgtct attgattcag tccctgatta ttgaatgttt tgcgatcgcc      360
gcatacaata tctacattcc cgttgcgtgac gattttgctc gtaaaatcac tgagggtgta      420
gtcaaagatg aatacagcca cctcaacttc ggcgaagttt ggttacagaa aaattttgcc      480
caatccaaag cagaattaga agaagctaata cgtcataatc ttcccatagt ttggaaaatg      540
ctcaatcaag tcgcggatga tgccgcagtc ttagctatgg aaaaagaagc cctagtcgaa      600
gattttatga ttcagtacgg cgaagcgtaa agtaatattg gcttcacaac cagagatatt      660
atgCGGgatgt cagcctacgg acttacagca gcttaa      696
  
```

5

<210> 8

<211> 231

<212> PRT

<213> *Nostoc sp.*

10 <400> 8

```

Met  Gln  Gln  Val  Ala  Ala  Asp  Leu  Glu  Ile  Asp  Phe  Lys  Ser  Glu  Lys
 1                                     5                10                15

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

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

Asp  Lys  Glu  Asp  Leu  Ile  Arg  Leu  Ser  Lys  Met  Glu  Ser  Arg  His  Lys
 50                55                60

Lys  Gly  Phe  Glu  Ala  Cys  Gly  Arg  Asn  Leu  Gln  Val  Ser  Pro  Asp  Met
65                70                75

Glu  Phe  Ala  Lys  Glu  Phe  Phe  Ala  Gly  Leu  His  Gly  Asn  Phe  Gln  Lys
                85                90                95

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

ES 2 696 773 T3

Ile Ile Glu Cys Phe Ala Ile Ala Ala Tyr Asn Ile Tyr Ile Pro Val  
 115 120 125

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

Tyr Ser His Leu Asn Phe Gly Glu Val Trp Leu Gln Lys Asn Phe Ala  
 145 150 155 160

Gln Ser Lys Ala Glu Leu Glu Glu Ala Asn Arg His Asn Leu Pro Ile  
 165 170 175

Val Trp Lys Met Leu Asn Gln Val Ala Asp Asp Ala Ala Val Leu Ala  
 180 185 190

Met Glu Lys Glu Ala Leu Val Glu Asp Phe Met Ile Gln Tyr Gly Glu  
 195 200 205

Ala Leu Ser Asn Ile Gly Phe Thr Thr Arg Asp Ile Met Arg Met Ser  
 210 215 220

Ala Tyr Gly Leu Thr Ala Ala  
 225 230

<210> 9

<211> 696

<212> ADN

5 <213> *Acaryochloris marina*

<400> 9

ES 2 696 773 T3

atgccccaaa ctcaggctat ttcagaaatt gacttctata gtgacaccta caaagatgct 60  
 tacagtcgta ttgacggcat tgtgatcgaa ggtgagcaag aagcgcgatga aaactatatt 120  
 cgtcttggcg aaatgctgcc tgagcaccaa gacgacttta tccgcctgtc caagatggaa 180  
 gcccgtcata agaaaggggtt tgaagcctgc ggtcgcaact taaaagtaac ctgcatccta 240  
 gactttgccc ggcgtttctt ttccgactta cacaagaatt ttcaagatgc tgcagctgag 300  
 gataaagtgc caacttgctt agtgattcag tccttgatca ttgagtgttt tgcgatcgca 360  
 gcttacaaca tctatatccc cgtcgctgat gactttgccc gtaagattac agagtctgtg 420  
 gttaaggatg agtatcaaca cctcaattat ggtgaagagt ggcttaaagc tcaactcgat 480  
 gatgtgaaag cagaaatcca agaagcta atcgcaaaaacc tccccatcgt ttggagaatg 540  
 ctgaacgaag tggacaagga tgcggccggt ttaggaatgg aaaaagaagc cctggttgaa 600  
 gacttcatga tccagtatgg tgaagccctt agcaatattg gtttctctac aggcgaaatt 660  
 atgcggatgt ctgcctatgg tcttgtggct gcgtaa 696

<210> 10

<211> 231

<212> PRT

5 <213> *Acaryochloris marina*

<400> 10

ES 2 696 773 T3

Met Pro Gln Thr Gln Ala Ile Ser Glu Ile Asp Phe Tyr Ser Asp Thr  
1 5 10 15

Tyr Lys Asp Ala Tyr Ser Arg Ile Asp Gly Ile Val Ile Glu Gly Glu  
20 25 30

Gln Glu Ala His Glu Asn Tyr Ile Arg Leu Gly Glu Met Leu Pro Glu  
35 40 45

His Gln Asp Asp Phe Ile Arg Leu Ser Lys Met Glu Ala Arg His Lys  
50 55 60

Lys Gly Phe Glu Ala Cys Gly Arg Asn Leu Lys Val Thr Cys Asp Leu  
65 70 75 80

Asp Phe Ala Arg Arg Phe Phe Ser Asp Leu His Lys Asn Phe Gln Asp  
85 90 95

Ala Ala Ala Glu Asp Lys Val Pro Thr Cys Leu Val Ile Gln Ser Leu  
100 105 110

Ile Ile Glu Cys Phe Ala Ile Ala Ala Tyr Asn Ile Tyr Ile Pro Val  
115 120 125

Ala Asp Asp Phe Ala Arg Lys Ile Thr Glu Ser Val Val Lys Asp Glu  
130 135 140

Tyr Gln His Leu Asn Tyr Gly Glu Glu Trp Leu Lys Ala His Phe Asp  
145 150 155 160

Asp Val Lys Ala Glu Ile Gln Glu Ala Asn Arg Lys Asn Leu Pro Ile  
165 170 175

Val Trp Arg Met Leu Asn Glu Val Asp Lys Asp Ala Ala Val Leu Gly  
180 185 190

Met Glu Lys Glu Ala Leu Val Glu Asp Phe Met Ile Gln Tyr Gly Glu  
195 200 205

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

Ala Tyr Gly Leu Val Ala Ala  
 225 230

<210> 11

<211> 696

<212> ADN

5 <213> *Thermosynechococcus elongatus*

<400> 11

```

atgacaacgg ctaccgctac acctgttttg gactaccata gcgatcgcta caaggatgcc      60
tacagccgca ttaacgccat tgtcattgaa ggtgaacagg aagctcacga taactatatc      120
gatttagcca agctgctgcc acaacaccaa gaggaactca cccgccttgc caagatggaa      180
gctcgccaca aaaagggggt tgaggcctgt ggtcgcaacc tgagcgtaac gccagatatg      240
gaatttgcca aagccttctt tgaaaaactg cgcgctaact ttcagagggc tctggcggag      300
ggaaaaactg cgacttgtct tctgattcaa gctttgatca tcgaatcctt tgcgatcgcg      360
gcctacaaca tctacatccc aatggcggat cctttcgccc gtaaaattac tgagagtgtt      420
gttaaggacg aatacagcca cctcaacttt ggcgaaatct ggctcaagga aactttgaa      480
agcgtcaaag gagagctcga agaagccaat cgcgccaatt tacccttggg ctggaaaatg      540
ctcaaccaag tggaagcaga tgccaaagtg ctcggcgatg aaaaagatgc ctttgtggaa      600
gacttcatga ttcagtacag tggtgcccta gaaaatatcg gctttaccac ccgcgaaatt      660
atgaagatgt cagtttatgg cctcactggg gcataa      696
    
```

<210> 12

<211> 231

10 <212> PRT

<213> *Thermosynechococcus elongatus*

<400> 12

ES 2 696 773 T3

Met Thr Thr Ala Thr Ala Thr Pro Val Leu Asp Tyr His Ser Asp Arg  
1 5 10 15

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

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

His Gln Glu Glu Leu Thr Arg Leu Ala Lys Met Glu Ala Arg His Lys  
50 55 60

Lys Gly Phe Glu Ala Cys Gly Arg Asn Leu Ser Val Thr Pro Asp Met  
65 70 75 80

Glu Phe Ala Lys Ala Phe Phe Glu Lys Leu Arg Ala Asn Phe Gln Arg

ES 2 696 773 T3

				85					90					95		
Ala	Leu	Ala	Glu	Gly	Lys	Thr	Ala	Thr	Cys	Leu	Leu	Ile	Gln	Ala	Leu	
			100					105					110			
Ile	Ile	Glu	Ser	Phe	Ala	Ile	Ala	Ala	Tyr	Asn	Ile	Tyr	Ile	Pro	Met	
		115					120					125				
Ala	Asp	Pro	Phe	Ala	Arg	Lys	Ile	Thr	Glu	Ser	Val	Val	Lys	Asp	Glu	
	130					135					140					
Tyr	Ser	His	Leu	Asn	Phe	Gly	Glu	Ile	Trp	Leu	Lys	Glu	His	Phe	Glu	
145					150					155					160	
Ser	Val	Lys	Gly	Glu	Leu	Glu	Glu	Ala	Asn	Arg	Ala	Asn	Leu	Pro	Leu	
				165					170					175		
Val	Trp	Lys	Met	Leu	Asn	Gln	Val	Glu	Ala	Asp	Ala	Lys	Val	Leu	Gly	
			180					185					190			
Met	Glu	Lys	Asp	Ala	Leu	Val	Glu	Asp	Phe	Met	Ile	Gln	Tyr	Ser	Gly	
		195					200					205				
Ala	Leu	Glu	Asn	Ile	Gly	Phe	Thr	Thr	Arg	Glu	Ile	Met	Lys	Met	Ser	
	210					215					220					
Val	Tyr	Gly	Leu	Thr	Gly	Ala										
225					230											

<210> 13

<211> 732

<212> ADN

5 <213> *Synechococcus* sp.

<400> 13

ES 2 696 773 T3

atggccccag cgaacgtcct gcccaacacc cccccgtccc ccactgatgg gggcggcact	60
gccctagact acagcagccc aaggtatcgg caggcctact cccgcatcaa cggatttgtt	120
atcgaaggcg aacaagaagc ccacgacaac tacctcaagc tggccgaaat gctgccgaa	180
gctgcagagg agctgcgcaa gctggccaag atggaattgc gccacatgaa aggcttcag	240
gcctgcggca aaaacctgca ggtggaacct gatgtggagt ttgcccgcg ctttttcgcg	300
cccttgcggg acaatttcca aagcgccgca gcggcagggg atctggtctc ctgttttgtc	360
attcagtctt tgatcatcga gtgctttgcc attgccgcct acaacatcta catcccggtt	420
gccgatgact ttgccgcaa gatcaccgag ggggtagtta aggacgagta tctgcacctc	480
aattttgggg agcgctggct gggcgagcac tttgccgagg ttaaagccca gatcgaagca	540
gccaacgccc aaaatctgcc tctagttcgg cagatgctgc agcaggtaga ggcggatgtg	600
gaagccattt acatggatcg cgaggccatt gtagaagact tcatgatcgc ctacggcgag	660
gccctggcca gcatcggctt caacacccgc gaggtaatgc gcctctcggc ccagggctctg	720
cgggccgcct ga	732

<210> 14

<211> 243

5 <212> PRT

<213> *Synechococcus* sp.

<400> 14

ES 2 696 773 T3

Met Ala Pro Ala Asn Val Leu Pro Asn Thr Pro Pro Ser Pro Thr Asp  
1 5 10 15

Gly Gly Gly Thr Ala Leu Asp Tyr Ser Ser Pro Arg Tyr Arg Gln Ala  
20 25 30

Tyr Ser Arg Ile Asn Gly Ile Val Ile Glu Gly Glu Gln Glu Ala His  
35 40 45

Asp Asn Tyr Leu Lys Leu Ala Glu Met Leu Pro Glu Ala Ala Glu Glu  
50 55 60

Leu Arg Lys Leu Ala Lys Met Glu Leu Arg His Met Lys Gly Phe Gln  
65 70 75 80

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

Ala Phe Phe Ala Pro Leu Arg Asp Asn Phe Gln Ser Ala Ala Ala Ala  
100 105 110

Gly Asp Leu Val Ser Cys Phe Val Ile Gln Ser Leu Ile Ile Glu Cys  
115 120 125

Phe Ala Ile Ala Ala Tyr Asn Ile Tyr Ile Pro Val Ala Asp Asp Phe  
130 135 140

Ala Arg Lys Ile Thr Glu Gly Val Val Lys Asp Glu Tyr Leu His Leu  
145 150 155 160

Asn Phe Gly Glu Arg Trp Leu Gly Glu His Phe Ala Glu Val Lys Ala  
165 170 175

Gln Ile Glu Ala Ala Asn Ala Gln Asn Leu Pro Leu Val Arg Gln Met  
180 185 190

ES 2 696 773 T3

Leu Gln Gln Val Glu Ala Asp Val Glu Ala Ile Tyr Met Asp Arg Glu  
 195 200 205

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

Ile Gly Phe Asn Thr Arg Glu Val Met Arg Leu Ser Ala Gln Gly Leu  
 225 230 235 240

**Arg Ala Ala**

<210> 15

<211> 708

<212> ADN

5 <213> *Gloeobacter violaceus*

<400> 15

gtgaaccgaa cgcaccgtc cagcgccgcg cttgattacc gctccgacac ctaccgcat 60  
 gcgtactccc gcatcaatgc catcgtcctt gaaggcgagc ggggaagccca cgccaactac 120  
 cttaccctcg ctgagatgct gccggaccat gccgaggcgc tcaaaaaact ggccgcatg 180  
 gaaaatcgcc acttcaaagg cttccagtcc tgcgcccgca acctcgaagt cacgccggac 240  
 gaccggttg caagggccta cttcgaacag ctcgacggca actttcagca ggcggcggca 300  
 gaaggtgacc ttaccacctg catggtcatc caggcactga tcatcgagtg cttcgcaatt 360  
 gcggcctaca acgtctacat tccggtggcc gacgcgtttg cccgcaaggt gaccgagggc 420  
 gtcgtcaagg acgagtacac ccacctcaac tttgggcagc agtggctcaa agagcgcttc 480  
 gtgaccgtgc gcgagggcat cgagcgcgcc aacgcccaga atctgcccac cgtctggcgg 540  
 atgctcaacg ccgtcgaagc ggacaccgaa gtgctgcaga tggataaaga agcgatecgc 600  
 gaagacttta tgatcgcta cggatgaagc ttgggcgaca tccggtttttc gatgcgcgac 660  
 gtgatgaaga tgtccgcccg cggccttgcc tctgcccccc gccagtga 708

<210> 16

<211> 235

10 <212> PRT

<213> *Gloeobacter violaceus*

<400> 16

ES 2 696 773 T3

Met Asn Arg Thr Ala Pro Ser Ser Ala Ala Leu Asp Tyr Arg Ser Asp  
1                                   5                                   10                                   15

Thr Tyr Arg Asp Ala Tyr Ser Arg Ile Asn Ala Ile Val Leu Glu Gly  
                                  20                                   25                                   30

Glu Arg Glu Ala His Ala Asn Tyr Leu Thr Leu Ala Glu Met Leu Pro  
                  35                                   40                                   45

ES 2 696 773 T3

Asp His Ala Glu Ala Leu Lys Lys Leu Ala Ala Met Glu Asn Arg His  
 50 55 60

Phe Lys Gly Phe Gln Ser Cys Ala Arg Asn Leu Glu Val Thr Pro Asp  
 65 70 75 80

Asp Pro Phe Ala Arg Ala Tyr Phe Glu Gln Leu Asp Gly Asn Phe Gln  
 85 90 95

Gln Ala Ala Ala Glu Gly Asp Leu Thr Thr Cys Met Val Ile Gln Ala  
 100 105 110

Leu Ile Ile Glu Cys Phe Ala Ile Ala Ala Tyr Asn Val Tyr Ile Pro  
 115 120 125

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

Glu Tyr Thr His Leu Asn Phe Gly Gln Gln Trp Leu Lys Glu Arg Phe  
 145 150 155 160

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

Ile Val Trp Arg Met Leu Asn Ala Val Glu Ala Asp Thr Glu Val Leu  
 180 185 190

Gln Met Asp Lys Glu Ala Ile Val Glu Asp Phe Met Ile Ala Tyr Gly  
 195 200 205

Glu Ala Leu Gly Asp Ile Gly Phe Ser Met Arg Asp Val Met Lys Met  
 210 215 220

Ser Ala Arg Gly Leu Ala Ser Ala Pro Arg Gln  
 225 230 235

<210> 17

<211> 732

<212> ADN

ES 2 696 773 T3

<213> *Prochlorococcus marinus*

<400> 17

atgcctacgc ttgagatgcc tgtggcagct gttcttgaca gcactgttgg atcttcagaa	60
gccctgccag acttcacttc agatagatat aaggatgcat acagcagaat caacgcaata	120
gtcattgagg gcgaacagga agcccatgac aattacatcg cgattggcac gctgcttccc	180
gatcatgtcg aagagctcaa gcggccttgcc aagatggaga tgaggcacia gaagggcttt	240
acagcttgcg gcaagaacct tggcgttgag gctgacatgg acttcgcaag ggagtttttt	300
gctcctttgc gtgacaactt ccagacagct ttagggcagg ggaaaacacc tacatgcttg	360
ctgatccagg cgctcttgat tgaagccttt gctatttcgg ctatcacac ctatatccct	420
gtttctgacc cctttgctcg caagattact gaaggtgtcg tgaaggacga gtacacacac	480
ctcaattatg gcgaggcttg gctcaaggcc aatctggaga gttgccgtga ggagttgctt	540
gaggccaatc gcgagaacct gcctctgatt cgccggatgc ttgatcaggt agcaggtgat	600
gctgccgtgc tgcagatgga taaggaagat ctgattgagg atttcttaat cgcctaccag	660
gaatctctca ctgagattgg cttaacact cgtgaaatta cccgtatggc agcggcagct	720
cttgtgagct ga	732

5 <210> 18

<211> 243

<212> PRT

<213> *Prochlorococcus marinus*

<400> 18

ES 2 696 773 T3

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

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

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

His Asp Asn Tyr Ile Ala Ile Gly Thr Leu Leu Pro Asp His Val Glu  
 50 55 60

Glu Leu Lys Arg Leu Ala Lys Met Glu Met Arg His Lys Lys Gly Phe  
 65 70 75 80

Thr Ala Cys Gly Lys Asn Leu Gly Val Glu Ala Asp Met Asp Phe Ala  
 85 90 95

Arg Glu Phe Phe Ala Pro Leu Arg Asp Asn Phe Gln Thr Ala Leu Gly  
 100 105 110

Gln Gly Lys Thr Pro Thr Cys Leu Leu Ile Gln Ala Leu Leu Ile Glu  
 115 120 125

Ala Phe Ala Ile Ser Ala Tyr His Thr Tyr Ile Pro Val Ser Asp Pro  
 130 135 140

Phe Ala Arg Lys Ile Thr Glu Gly Val Val Lys Asp Glu Tyr Thr His



ES 2 696 773 T3

atgcaaacac tcgaatctaa taaaaaaact aatctagaaa attctattga tttacccgat	60
tttactactg attcttacia agacgcttat agcaggataa atgcaatagt tattgaagg	120
gaacaagagg ctcatgataa ttacatttcc ttagcaacat taattcctaa cgaattagaa	180
gagttaacta aattagcgaa aatggagctt aagcacaaaa gaggcctttac tgcattgtga	240
agaaatctag gtgttcaagc tgacattgatt tttgctaaag aattcttttc caaattacat	300
ggtaattttc aggttgcggt atctaattggc aagacaacta catgcctatt aatacaggca	360
attttaattg aagcttttgc tatatccgcg tatcacgttt acataagagt tgctgatcct	420
ttcgcgaaaa aaattaccca aggtgttggt aaagattgat atcttcattt aaattatgga	480
caagaatggc taaaagaaaa tttagcgact tgtaaagatt agctaattgga agcaaataag	540
gttaaccttc cattaatcaa gaagattgta gatcaagtct cggaagattc ttcagtacta	600
gctatggata gggagaatt aatggaagaa ttcattgatt cctatcagga cactctcctt	660
gaaatagggt tagataatag agaaattgca agaattggca tggctgctat agtttaa	717

<210> 20

<211> 238

<212> PRT

5 <213> *Prochlorococcus marinus*

<400> 20

ES 2 696 773 T3

Met Gln Thr Leu Glu Ser Asn Lys Lys Thr Asn Leu Glu Asn Ser Ile  
1 5 10 15

Asp Leu Pro Asp Phe Thr Thr Asp Ser Tyr Lys Asp Ala Tyr Ser Arg  
20 25 30

Ile Asn Ala Ile Val Ile Glu Gly Glu Gln Glu Ala His Asp Asn Tyr  
35 40 45

Ile Ser Leu Ala Thr Leu Ile Pro Asn Glu Leu Glu Glu Leu Thr Lys  
50 55 60

Leu Ala Lys Met Glu Leu Lys His Lys Arg Gly Phe Thr Ala Cys Gly  
65 70 75 80

Arg Asn Leu Gly Val Gln Ala Asp Met Ile Phe Ala Lys Glu Phe Phe  
85 90 95

Ser Lys Leu His Gly Asn Phe Gln Val Ala Leu Ser Asn Gly Lys Thr  
100 105 110

Thr Thr Cys Leu Leu Ile Gln Ala Ile Leu Ile Glu Ala Phe Ala Ile  
115 120 125

Ser Ala Tyr His Val Tyr Ile Arg Val Ala Asp Pro Phe Ala Lys Lys  
130 135 140

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

Gln Glu Trp Leu Lys Glu Asn Leu Ala Thr Cys Lys Asp Glu Leu Met  
165 170 175

Glu Ala Asn Lys Val Asn Leu Pro Leu Ile Lys Lys Met Leu Asp Gln  
180 185 190

Val Ser Glu Asp Ala Ser Val Leu Ala Met Asp Arg Glu Glu Leu Met  
195 200 205

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

Asp Asn Arg Glu Ile Ala Arg Met Ala Met Ala Ala Ile Val  
225 230 235

ES 2 696 773 T3

<210> 21

<211> 726

<212> ADN

<213> *Prochlorococcus marinus*

5 <400> 21

atgcaagctt ttgcatccaa caatttaacc gtagaaaaag aagagctaag ttctaactct	60
cttccagatt tcacctcaga atcttatacaa gatgcttaca gcagaatcaa tgcagttgta	120
attgaagggg agcaagaagc ttattctaata tttcttgatc tcgctaaatt gattcctgaa	180
catgcagatg agcttgtgag gctagggag atggagaaaa agcatatgaa tggtttttgt	240
gcttgccggga gaaatcttgc tgtaaagcct gatatgcctt ttgcaaagac ctttttctca	300
aaactccata ataatttttt agaggctttc aaagtaggag atacgactac ctgtctccta	360
attcaatgca tcttgattga atcttttgca atatccgcat atcacgttta tatacgtgtt	420
gctgatccat tcgccaaaag aatcacagag ggtggtgtcc aagatgaata cttgcatttg	480
aactatggtc aagaatggct taaggccaat ctagagacag ttaagaaaga tcttatgagg	540
gctaataagg aaaacttgcc tcttataaag tccatgctcg atgaagtttc aaacgacgcc	600
gaagtccttc atatggataa agaagagtta atggaggaat ttatgattgc ttatcaagat	660
tcccttcttg aatatggctt tgataataga gaaattgcaa gaatggctct tgcagcgggtg	720
atataa	726

<210> 22

<211> 241

<212> PRT

10 <213> *Prochlorococcus marinus*

<400> 22

ES 2 696 773 T3

Met	Gln	Ala	Phe	Ala	Ser	Asn	Asn	Leu	Thr	Val	Glu	Lys	Glu	Glu	Leu
1				5					10					15	
Ser	Ser	Asn	Ser	Leu	Pro	Asp	Phe	Thr	Ser	Glu	Ser	Tyr	Lys	Asp	Ala
			20					25					30		
Tyr	Ser	Arg	Ile	Asn	Ala	Val	Val	Ile	Glu	Gly	Glu	Gln	Glu	Ala	Tyr
		35					40					45			
Ser	Asn	Phe	Leu	Asp	Leu	Ala	Lys	Leu	Ile	Pro	Glu	His	Ala	Asp	Glu
	50					55					60				
Leu	Val	Arg	Leu	Gly	Lys	Met	Glu	Lys	Lys	His	Met	Asn	Gly	Phe	Cys
65					70					75					80
Ala	Cys	Gly	Arg	Asn	Leu	Ala	Val	Lys	Pro	Asp	Met	Pro	Phe	Ala	Lys
				85					90					95	
Thr	Phe	Phe	Ser	Lys	Leu	His	Asn	Asn	Phe	Leu	Glu	Ala	Phe	Lys	Val
			100					105					110		

ES 2 696 773 T3

Gly Asp Thr Thr Thr Cys Leu Leu Ile Gln Cys Ile Leu Ile Glu Ser  
 115 120 125

Phe Ala Ile Ser Ala Tyr His Val Tyr Ile Arg Val Ala Asp Pro Phe  
 130 135 140

Ala Lys Arg Ile Thr Glu Gly Val Val Gln Asp Glu Tyr Leu His Leu  
 145 150 155 160

Asn Tyr Gly Gln Glu Trp Leu Lys Ala Asn Leu Glu Thr Val Lys Lys  
 165 170 175

Asp Leu Met Arg Ala Asn Lys Glu Asn Leu Pro Leu Ile Lys Ser Met  
 180 185 190

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

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

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

Ile

<210> 23

<211> 732

<212> ADN

5 <213> *Synechococcus* sp.

<400> 23

ES 2 696 773 T3

atgccgaccc ttgagacgtc tgaggctgcc gttcttgaag actcgatggc ttcaggctcc	60
cggctgcctg atttcaccag cgaggcttac aaggacgcct acagccgcat caatgcgac	120
gtgatcgagg gtgagcagga agcgcacgac aactacatcg ccctcggcac gctgatcccc	180
gagcagaagg atgagctggc ccgtctcgcc cgcattgaga tgaagcacat gaaggggttc	240
acctcctgtg gccgcaatct cggcgtggag gcagacctc cctttgctaa ggaattcttc	300
gccccctgc acgggaactt ccaggcagct ctccaggagg gcaaggtggt gacctgcctg	360
ttgattcagg cgctgctgat tgaagcgttc gccatttccg cctatcacat ctacatcccg	420
gtggcggatc ccttcgctcg caagatcact gaaggtgtgg tgaaggatga gtacaccac	480
ctcaattacg gccaggaatg gctgaaggcc aattttgagg ccagcaagga tgagctgatg	540
gaggccaaca aggccaatct gcctctgac cgcctcgatgc tggagcaggt ggcagccgac	600
gccgccgtgc tgcagatgga aaaggaagat ctgatcgaag atttcctgat cgcttaccag	660
gaggccctct gcgagatcgg tttcagctcc cgtgacattg ctcgcatggc cgccgctgcc	720
ctcgcggtct ga	732

<210> 24

<211> 243

5 <212> PRT

<213> *Synechococcus sp.*

<400> 24

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

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

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

His Asp Asn Tyr Ile Ala Leu Gly Thr Leu Ile Pro Glu Gln Lys Asp  
50 55 60

Glu Leu Ala Arg Leu Ala Arg Met Glu Met Lys His Met Lys Gly Phe  
65 70 75 80

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

Lys Glu Phe Phe Ala Pro Leu His Gly Asn Phe Gln Ala Ala Leu Gln  
100 105 110

Glu Gly Lys Val Val Thr Cys Leu Leu Ile Gln Ala Leu Leu Ile Glu  
115 120 125

Ala Phe Ala Ile Ser Ala Tyr His Ile Tyr Ile Pro Val Ala Asp Pro  
130 135 140

Phe Ala Arg Lys Ile Thr Glu Gly Val Val Lys Asp Glu Tyr Thr His  
145 150 155 160

Leu Asn Tyr Gly Gln Glu Trp Leu Lys Ala Asn Phe Glu Ala Ser Lys  
165 170 175

Asp Glu Leu Met Glu Ala Asn Lys Ala Asn Leu Pro Leu Ile Arg Ser  
180 185 190

Met Leu Glu Gln Val Ala Ala Asp Ala Ala Val Leu Gln Met Glu Lys

ES 2 696 773 T3

195		200		205
Glu Asp Leu Ile Glu Asp Phe Leu Ile Ala Tyr Gln Glu Ala Leu Cys				
210		215		220
Glu Ile Gly Phe Ser Ser Arg Asp Ile Ala Arg Met Ala Ala Ala Ala				
225		230		235 240

Leu Ala Val

<210> 25

<211> 681

<212> ADN

5 <213> *Synechococcus* sp.

<400> 25

atgaccacgc tgcactttgc cagtgcggcc taccgcgagg cctacagccg gatcaacggc	60
gttgtgattg tgggcgaagg tctcgccaat cgccatttcc agatggtggc gcggcgcatt	120
cccgctgacg gcgacgagct gcagcggctc ggacgcatgg agggagacca tgccagcggc	180
tttgtgggct gtggctgcaa cctcgggtgtg gtggccgata tgcccctggc ccggcgcctg	240
tttcagcccc tccatgatct gttcaaacgc cagcaccacg acggcaatcg ggcccgaatgc	300
ctggtgatcc aggggttgat cgtggaatgt ttcgccgtgg cggcttaccg ccactacctg	360
ccggtggccg atgcctacgc ccggccgata accgcagcgg tgatgaacga tgaatcggaa	420
cacctcgact acgctgagac ctggctgcag cgccatttcc atcaggtgaa ggcccgggtc	480
agcgcggtgg tgggtggaggc gttgccgctc accctggcga tgttgcaatc gcttgetgca	540
gacatgcgac agatcggcat ggatccggtg gagaccctgg ccagcttcag tgaactgttt	600
cggaagcgt tggaatcggc ggggtttgag gctgtggagg ccagggcact gctgatgcga	660
gcggccgccc ggatggtctg a	681

<210> 26

<211> 226

10 <212> PRT

<213> *Synechococcus* sp.

<400> 26

ES 2 696 773 T3

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

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

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

ES 2 696 773 T3

Arg Leu Gly Arg Met Glu Gly Asp His Ala Ser Ala Phe Val Gly Cys  
50 55 60

Gly Arg Asn Leu Gly Val Val Ala Asp Leu Pro Leu Ala Arg Arg Leu  
65 70 75 80

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

Arg Ala Glu Cys Leu Val Ile Gln Gly Leu Ile Val Glu Cys Phe Ala  
100 105 110

Val Ala Ala Tyr Arg His Tyr Leu Pro Val Ala Asp Ala Tyr Ala Arg  
115 120 125

Pro Ile Thr Ala Ala Val Met Asn Asp Glu Ser Glu His Leu Asp Tyr  
130 135 140

Ala Glu Thr Trp Leu Gln Arg His Phe Asp Gln Val Lys Ala Arg Val  
145 150 155 160

Ser Ala Val Val Val Glu Ala Leu Pro Leu Thr Leu Ala Met Leu Gln  
165 170 175

Ser Leu Ala Ala Asp Met Arg Gln Ile Gly Met Asp Pro Val Glu Thr  
180 185 190

Leu Ala Ser Phe Ser Glu Leu Phe Arg Glu Ala Leu Glu Ser Val Gly  
195 200 205

Phe Glu Ala Val Glu Ala Arg Arg Leu Leu Met Arg Ala Ala Ala Arg  
210 215 220

Met Val  
225

<210> 27

<211> 696

<212> ADN

ES 2 696 773 T3

<213> *Cyanothece* sp.

<400> 27

atgcaagagc ttgctttacg ctcagagcctt gattttaaca gcgaaaccta taaagatgct	60
tacagtcgca tcaatgctat tgtcattgaa ggggaacaag aagcctatca aaattatctt	120
gatatggcgc aacttctccc agaagacgag gctgagttaa ttcgtctctc caagatggaa	180
aaccgtcaca aaaaaggctt tcaagcctgt ggcaagaatt tgaatgtgac cccagatatg	240
gactacgctc aacaatTTTT tgctgaactt catggcaact tccaaaaggc aaaagccgaa	300
ggcaaaattg tcacttgctt attaattcaa tctttgatca tcgaagcctt tgcgatcgcc	360
gcttataata tttatattcc tgtggcagat ccctttgctc gtaaaatcac cgaaggggta	420
gttaaggatg aatataccca cctcaatTTT ggggaagtct ggtaaaaga gcattttgaa	480
gcctctaaag cagaattaga agacgcaaat aaagaaaatt tacccttTgt ttggcaaatg	540
ctcaaccaag ttgaaaaaga tgccgaagtg ttagggatgg agaaagaagc cttagtggaa	600
gatttcatga ttagttatgg agaagcttta agtaatattg gtttctctac ccgtgagatc	660
atgaaaatgt ctgcttacgg gctacgggct gcttaa	696

5 <210> 28

<211> 231

<212> PRT

<213> *Cyanothece* sp.

<400> 28

ES 2 696 773 T3

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

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

Gln Glu Ala Tyr Gln Asn Tyr Leu Asp Met Ala Gln Leu Leu Pro Glu  
35 40 45

Asp Glu Ala Glu Leu Ile Arg Leu Ser Lys Met Glu Asn Arg His Lys  
50 55 60

Lys Gly Phe Gln Ala Cys Gly Lys Asn Leu Asn Val Thr Pro Asp Met  
65 70 75 80

Asp Tyr Ala Gln Gln Phe Phe Ala Glu Leu His Gly Asn Phe Gln Lys  
85 90 95

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

Ile Ile Glu Ala Phe Ala Ile Ala Ala Tyr Asn Ile Tyr Ile Pro Val  
115 120 125

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

Tyr Thr His Leu Asn Phe Gly Glu Val Trp Leu Lys Glu His Phe Glu  
145 150 155 160

ES 2 696 773 T3

Ala Ser Lys Ala Glu Leu Glu Asp Ala Asn Lys Glu Asn Leu Pro Leu  
 165 170 175

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

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

Ala Leu Ser Asn Ile Gly Phe Ser Thr Arg Glu Ile Met Lys Met Ser  
 210 215 220

Ala Tyr Gly Leu Arg Ala Ala  
 225 230

<210> 29

<211> 696

<212> ADN

5 <213> *Cyanotheca* sp.

<400> 29

```

atgcctcaag tgcagtcccc atcggctata gacttctaca gtgagaccta ccaggatgct      60
tacagccgca ttgatgcgat cgtgatcgag ggagaacagg aagcccacga caattacctg      120
aagctgacgg aactgctgcc ggattgtcaa gaagatctgg tccggctggc caaaatggaa      180
gcccgtcaca aaaaagggtt tgaagcttgt ggccgcaatc tcaaggtcac acccgatatg      240
gagtttgctc aacagttctt tgctgacctg cacaacaatt tccagaaagc tgctgctggcc      300
aacaaaattg ccacctgtct ggtgatccag gccctgatta ttgagtgctt tgccatcgcc      360
gcttataaca tctatattcc tgtcgctgat gactttgccc gcaaaattac cgaaaacgtg      420
gtcaaagacg aatacaccoa cctcaacttt ggtgaagagt ggctcaaagc taactttgat      480
agccagcggg aagaagtgga agcggccaac cgggaaaacc tgccgatcgt ctggcggatg      540
ctcaatcagg tagagactga tgctcacggt ttaggtatgg aaaaagaggc tttagtggaa      600
agcttcatga tccaatatgg tgaagccctg gaaaatattg gtttctcgac ccgtgagatc      660
atgcgcatgt ccgtttacgg cctctctgcg gcataa      696
    
```

<210> 30

<211> 231

10 <212> PRT

ES 2 696 773 T3

<213> *Cyanothece sp.*

<400> 30

Met Pro Gln Val Gln Ser Pro Ser Ala Ile Asp Phe Tyr Ser Glu Thr  
1 5 10 15

Tyr Gln Asp Ala Tyr Ser Arg Ile Asp Ala Ile Val Ile Glu Gly Glu  
20 25 30

ES 2 696 773 T3

Gln Glu Ala His Asp Asn Tyr Leu Lys Leu Thr Glu Leu Leu Pro Asp  
35 40 45

Cys Gln Glu Asp Leu Val Arg Leu Ala Lys Met Glu Ala Arg His Lys  
50 55 60

Lys Gly Phe Glu Ala Cys Gly Arg Asn Leu Lys Val Thr Pro Asp Met  
65 70 75 80

Glu Phe Ala Gln Gln Phe Phe Ala Asp Leu His Asn Asn Phe Gln Lys  
85 90 95

Ala Ala Ala Ala Asn Lys Ile Ala Thr Cys Leu Val Ile Gln Ala Leu  
100 105 110

Ile Ile Glu Cys Phe Ala Ile Ala Ala Tyr Asn Ile Tyr Ile Pro Val  
115 120 125

Ala Asp Asp Phe Ala Arg Lys Ile Thr Glu Asn Val Val Lys Asp Glu  
130 135 140

Tyr Thr His Leu Asn Phe Gly Glu Glu Trp Leu Lys Ala Asn Phe Asp  
145 150 155 160

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

Val Trp Arg Met Leu Asn Gln Val Glu Thr Asp Ala His Val Leu Gly  
180 185 190

Met Glu Lys Glu Ala Leu Val Glu Ser Phe Met Ile Gln Tyr Gly Glu  
195 200 205

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

Val Tyr Gly Leu Ser Ala Ala  
225 230

ES 2 696 773 T3

<211> 702

<212> ADN

<213> *Cyanothece sp.*

<400> 31

5 atgtctgatt gcgccacgaa cccagccctc gactattaca gtgaaaccta ccgcaatgct 60  
 taccggcggg tgaacgggat tgtgattgaa ggcgagaagc aagcctacga caactttatc 120  
 cgcttagctg agctgctccc agagtatcaa gcggaattaa cccgtctggc taaaatggaa 180  
 gcccgccacc agaagagctt tgttgccctg gcccaaatc tcaaggttag cccggactta 240  
 gactttgcgg cacagttttt tgctgaactg catcaaattt ttgcatctgc agcaaatgcy 300  
 ggccagggtg ctacctgtct ggttggtgcaa gccctgatca ttgaatgctt tgcgatcgcc 360  
 gcctacaata cctatattgcc agtagcggat gaatttgccc gtaaagtac cgcacccgtt 420  
 gttcaggacg agtacagcca cctaaacttt ggtgaagtct ggctgcagaa tgcgtttgag 480  
 cagtgtaaag acgaaattat cacagctaac cgtcttgctc tgccgctgat ctggaaaatg 540  
 ctcaaccagg tgacaggcga attgocgatt ctgggcatgg acaaagcttc tctggtagaa 600  
 gactttagca ctcgctatgg agaggccctg ggccagattg gtttcaaact atctgaaatt 660  
 ctctccctgt ccgttcaggg tttacaggcg gttacgcctt ag 702

<210> 32

<211> 233

<212> PRT

10 <213> *Cyanothece sp.*

<400> 32

ES 2 696 773 T3

Met Ser Asp Cys Ala Thr Asn Pro Ala Leu Asp Tyr Tyr Ser Glu Thr  
 1 5 10 15

Tyr Arg Asn Ala Tyr Arg Arg Val Asn Gly Ile Val Ile Glu Gly Glu  
 20 25 30

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

Tyr Gln Ala Glu Leu Thr Arg Leu Ala Lys Met Glu Ala Arg His Gln  
 50 55 60

Lys Ser Phe Val Ala Cys Gly Gln Asn Leu Lys Val Ser Pro Asp Leu  
 65 70 75 80

Asp Phe Ala Ala Gln Phe Phe Ala Glu Leu His Gln Ile Phe Ala Ser  
 85 90 95

Ala Ala Asn Ala Gly Gln Val Ala Thr Cys Leu Val Val Gln Ala Leu  
 100 105 110

Ile Ile Glu Cys Phe Ala Ile Ala Ala Tyr Asn Thr Tyr Leu Pro Val  
 115 120 125

Ala Asp Glu Phe Ala Arg Lys Val Thr Ala Ser Val Val Gln Asp Glu  
 130 135 140

Tyr Ser His Leu Asn Phe Gly Glu Val Trp Leu Gln Asn Ala Phe Glu  
 145 150 155 160

Gln Cys Lys Asp Glu Ile Ile Thr Ala Asn Arg Leu Ala Leu Pro Leu  
 165 170 175

Ile Trp Lys Met Leu Asn Gln Val Thr Gly Glu Leu Arg Ile Leu Gly  
 180 185 190

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

Ala Leu Gly Gln Ile Gly Phe Lys Leu Ser Glu Ile Leu Ser Leu Ser  
 210 215 220

Val Gln Gly Leu Gln Ala Val Thr Pro  
 225 230

ES 2 696 773 T3

<210> 33

<211> 696

<212> ADN

<213> *Anabaena variabilis*

5 <400> 33

```

atgcagcagg ttgcagccga tttagaaatc gatttcaaga gcgaaaaata taaagatgcc      60
tatagtcgca taaatgcgat cgtgattgaa ggggaacaag aagcatatga gaattacatt      120
caactatccc aactgctgcc agacgataaa gaagacctaa ttcgcctctc gaaaatggaa      180
agtcgccaca aaaaaggatt tgaagcttgt ggacggaacc tgcaagtatc cccagacata      240
gagttcgcta aagaattctt tgccgggcta cacggttaatt tccaaaaagc ggcagctgaa      300
ggtaaagttg tcaactgcct attgattcaa tccctgatta ttgaatgttt tgcgatcgcc      360
gcatacaata tctacatccc cgtggctgac gatttcgccc gtaaaatcac tgagggtgta      420
gttaaagatg aatacagtca cctcaacttc ggcgaaagttt ggttacagaa aaatttcgct      480
caatcaaaag cagaactaga agaagctaat cgtcataatc ttcccatagt ctggaaaatg      540
ctcaatcaag ttgccgatga tgcggcagtc ttagctatgg aaaaagaagc cctagtggaa      600
gattttatga ttcagtacgg cgaagcacta agtaatattg gcttcacaac cagagatatt      660
atgcggatgt cagcctacgg actcacagca gcttaa      696

```

<210> 34

<211> 231

<212> PRT

10 <213> *Anabaena variabilis*

<400> 34

**Met Gln Gln Val Ala Ala Asp Leu Glu Ile Asp Phe Lys Ser Glu Lys**



ES 2 696 773 T3

<211> 765

<212> ADN

<213> *Synechococcus elongatus*

<400> 35

gtgcgtagcc	cctgggatcc	accaaatacc	acattctccc	tctcatccgt	gtcaggagac	60
cgcagactca	tgccgcagct	tgaagccagc	cttgaactgg	actttcaaag	cgagtcctac	120
aaagacgctt	acagccgcat	caacgcgatc	gtgattgaag	gcgaacaaga	ggcgttcgac	180
aactacaatc	gccttgctga	gatgctgccc	gaccagcggg	atgagcttca	caagctagcc	240
aagatggaac	agcgccacat	gaaaggcttt	atggcctgtg	gcaaaaatct	ctccgtcact	300
cctgacatgg	gttttgccca	gaaatthttc	gagcgcttgc	acgagaactt	caaagcggcg	360
gctgcggaag	gcaaggctgt	cacctgccta	ctgattcaat	cgctaatacat	cgagtgtttt	420
gcgatcgcgg	cttacaacat	ctacatccca	gtggcgggatg	cttttgcccg	caaaatcacg	480
gaggggggtcg	tgccgcgacga	atacctgcac	cgcaacttcg	gtgaagagtg	gctgaaggcg	540
aatthttgatg	cttccaaagc	cgaactggaa	gaagccaatc	gtcagaacct	gcccttggtt	600
tggctaatagc	tcaacgaagt	ggccgatgat	gctcgcgaac	tcgggatgga	gcgtgagtcg	660
ctcgtcgagg	actthtatgat	tgctacgggt	gaagctctgg	aaaacatcgg	cttcacaacg	720
cgcgaaatca	tgcgtagtgc	cgctatggc	cttgcggccg	tttga		765

5

<210> 36

<211> 254

<212> PRT

<213> *Synechococcus elongatus*

10 <400> 36

ES 2 696 773 T3

Met Arg Thr Pro Trp Asp Pro Pro Asn Pro Thr Phe Ser Leu Ser Ser  
 1 5 10 15

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

Leu Asp Phe Gln Ser Glu Ser Tyr Lys Asp Ala Tyr Ser Arg Ile Asn  
 35 40 45

Ala Ile Val Ile Glu Gly Glu Gln Glu Ala Phe Asp Asn Tyr Asn Arg  
 50 55 60

Leu Ala Glu Met Leu Pro Asp Gln Arg Asp Glu Leu His Lys Leu Ala  
 65 70 75 80

Lys Met Glu Gln Arg His Met Lys Gly Phe Met Ala Cys Gly Lys Asn  
 85 90 95

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

ES 2 696 773 T3

Leu His Glu Asn Phe Lys Ala Ala Ala Ala Glu Gly Lys Val Val Thr  
 115 120 125

Cys Leu Leu Ile Gln Ser Leu Ile Ile Glu Cys Phe Ala Ile Ala Ala  
 130 135 140

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

Glu Gly Val Val Arg Asp Glu Tyr Leu His Arg Asn Phe Gly Glu Glu  
 165 170 175

Trp Leu Lys Ala Asn Phe Asp Ala Ser Lys Ala Glu Leu Glu Glu Ala  
 180 185 190

Asn Arg Gln Asn Leu Pro Leu Val Trp Leu Met Leu Asn Glu Val Ala  
 195 200 205

Asp Asp Ala Arg Glu Leu Gly Met Glu Arg Glu Ser Leu Val Glu Asp  
 210 215 220

Phe Met Ile Ala Tyr Gly Glu Ala Leu Glu Asn Ile Gly Phe Thr Thr  
 225 230 235 240

Arg Glu Ile Met Arg Met Ser Ala Tyr Gly Leu Ala Ala Val  
 245 250

<210> 37

<211> 19

<212> PRT

5 <213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: péptido sintético

<220>

<221> MOD\_RES

10 <222> (2) .. (3)

<223> Cualquier aminoácido

<220>

<221> MOD\_RES

<222> (6) .. (6)

15 <223> Cualquier aminoácido

<220>

<221> MOD\_RES

<222> (8) .. (11)

<223> Cualquier aminoácido

<220>

5 <221> MOD\_RES

<222> (13) .. (14)

<223> Cualquier aminoácido

<220>

<221> MOD\_RES

10 <222> (17) .. (18)

<223> Cualquier aminoácido

<400> 37

**Tyr Xaa Xaa Ala Tyr Xaa Arg Xaa Xaa Xaa Xaa Val Xaa Xaa Gly Glu**  
**1 5 10 15**

**Xaa Xaa Ala**

<210> 38

15 <211> 15

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: péptido sintético

20 <220>

<221> MOD\_RES

<222> (2) .. (3)

<223> Cualquier aminoácido

<220>

25 <221> MOD\_RES

<222> (6) .. (7)

<223> Cualquier aminoácido

<220>

<221> MOD\_RES

30 <222> (9) .. (11)

<223> Cualquier aminoácido

<220>

<221> MOD\_RES

<222> (13) .. (14)

ES 2 696 773 T3

<223> Cualquier aminoácido

<400> 38

**Leu Xaa Xaa Met Glu Xaa Xaa His Xaa Xaa Xaa Phe Xaa Xaa Cys**  
**1 5 10 15**

<210> 39

5 <211> 17

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: péptido sintético

10 <220>

<221> MOD\_RES

<222> (2) .. (4)

<223> Cualquier aminoácido

<220>

15 <221> MOD\_RES

<222> (6) .. (9)

<223> Cualquier aminoácido

<220>

<221> MOD\_RES

20 <222> (11) .. (11)

<223> Cualquier aminoácido

<220>

<221> MOD\_RES

<222> (14) .. (15)

25 <223> Cualquier aminoácido

<400> 39

**Cys Xaa Xaa Xaa Gln Xaa Xaa Xaa Xaa Glu Xaa Phe Ala Xaa Xaa Ala**  
**1 5 10 15**

<210> 40

<211> 19

30 <212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: péptido sintético

<220>

35 <221> MOD\_RES

- <222> (2) .. (3)  
 <223> Cualquier aminoácido  
 <220>  
 <221> MOD\_RES
- 5 <222> (5) .. (7)  
 <223> Cualquier aminoácido  
 <220>  
 <221> MOD\_RES  
 <222> (9) .. (10)
- 10 <223> Cualquier aminoácido  
 <220>  
 <221> MOD\_RES  
 <222> (12) .. (17)  
 <223> Cualquier aminoácido
- 15 <400> 40
- |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Thr | Xaa | Xaa | Val | Xaa | Xaa | Xaa | Glu | Xaa | Xaa | His | Xaa | Xaa | Xaa | Xaa | Xaa |
| 1   |     |     | 5   |     |     |     |     | 10  |     |     |     |     |     | 15  |     |
- Xaa Trp Leu**
- <210> 41  
 <211> 23  
 <212> PRT
- 20 <213> Secuencia artificial  
 <220>  
 <223> Descripción de la secuencia artificial: péptido sintético  
 <220>  
 <221> MOD\_RES
- 25 <222> (2) .. (2)  
 <223> Lys, Arg o Gln  
 <220>  
 <221> MOD\_RES  
 <222> (3) .. (3)
- 30 <223> Cualquier aminoácido  
 <220>  
 <221> MOD\_RES  
 <222> (6) .. (6)  
 <223> Ser o Arg

<220>  
 <221> MOD\_RES  
 <222> (8) .. (8)  
 <223> Ile o Val  
 5 <220>  
 <221> MOD\_RES  
 <222> (9) .. (9)  
 <223> Asp o Asn  
 <220>  
 10 <221> MOD\_RES  
 <222> (10) .. (10)  
 <223> Gly o Ala  
 <220>  
 <221> MOD\_RES  
 15 <222> (11) .. (11)  
 <223> Val o Ile  
 <220>  
 <221> MOD\_RES  
 <222> (13) .. (13)  
 20 <223> Ile o Leu  
 <220>  
 <221> MOD\_RES  
 <222> (14) .. (14)  
 <223> Qualquer aminoácido  
 25 <220>  
 <221> MOD\_RES  
 <222> (17) .. (18)  
 <223> Qualquer aminoácido  
 <220>  
 30 <221> MOD\_RES  
 <222> (20) .. (21)  
 <223> Qualquer aminoácido  
 <220>  
 <221> MOD\_RES  
 35 <222> (22) .. (22)  
 <223> Asn o His  
 <220>

<221> MOD\_RES

<222> (23) .. (23)

<223> Tyr o Phe

<400> 41

Tyr Xaa Xaa Ala Tyr Xaa Arg Xaa Xaa Xaa Xaa Val Xaa Xaa Gly Glu  
 1 5 10 15

Xaa Xaa Ala Xaa Xaa Xaa Xaa  
 5 20

<210> 42

<211> 21

<212> PRT

<213> Secuencia artificial

10 <220>

<223> Descripción de la secuencia artificial: péptido sintético

<220>

<221> MOD\_RES

<222> (2) .. (2)

15 <223> Ala, Ser o Gly

<220>

<221> MOD\_RES

<222> (3) .. (3)

<223> Cualquier aminoácido

20 <220>

<221> MOD\_RES

<222> (6) .. (6)

<223> Cualquier aminoácido

<220>

25 <221> MOD\_RES

<222> (7) .. (7)

<223> Arg, Asp o Lys

<220>

<221> MOD\_RES

30 <222> (9) .. (10)

<223> Cualquier aminoácido

<220>

<221> MOD\_RES

<222> (11) .. (11)  
 <223> Ala, Ser o Gly  
 <220>  
 <221> MOD\_RES  
 5 <222> (13) .. (13)  
 <223> Cualquier aminoácido  
 <220>  
 <221> MOD\_RES  
 <222> (14) .. (14)  
 10 <223> Ala, Ser o Gly  
 <220>  
 <221> MOD\_RES  
 <222> (16) .. (16)  
 <223> Ala o Gly  
 15 <220>  
 <221> MOD\_RES  
 <222> (17) .. (17)  
 <223> Cualquier aminoácido  
 <220>  
 20 <221> MOD\_RES  
 <222> (20) .. (20)  
 <223> Cualquier aminoácido  
 <220>  
 <221> MOD\_RES  
 25 <222> (21) .. (21)  
 <223> Val o Ile  
 <400> 42  
**Leu Xaa Xaa Met Glu Xaa Xaa His Xaa Xaa Xaa Phe Xaa Xaa Cys Xaa**  
**1 5 10 15**  
  
**Xaa Asn Leu Xaa Xaa**  
**20**  
 <210> 43  
 30 <211> 21  
 <212> PRT  
 <213> Secuencia artificial  
 <220>

<223> Descripción de la secuencia artificial: péptido sintético

<220>

<221> MOD\_RES

<222> (2) .. (2)

5 <223> Leu, Met o Phe

<220>

<221> MOD\_RES

<222> (3) .. (3)

<223> Val o Leu

10 <220>

<221> MOD\_RES

<222> (4) .. (4)

<223> Ile o Val

<220>

15 <221> MOD\_RES

<222> (6) .. (6)

<223> Cualquier aminoácido

<220>

<221> MOD\_RES

20 <222> (7) .. (7)

<223> Leu o Ile

<220>

<221> MOD\_RES

<222> (8) .. (8)

25 <223> Ile, Leu o Met

<220>

<221> MOD\_RES

<222> (9) .. (9)

<223> Ile o Val

30 <220>

<221> MOD\_RES

<222> (11) .. (11)

<223> Cualquier aminoácido

<220>

35 <221> MOD\_RES

<222> (14) .. (14)

<223> Ile o Val

- <220>  
 <221> MOD\_RES  
 <222> (15) .. (15)  
 <223> Ala o Ser
- 5 <220>  
 <221> MOD\_RES  
 <222> (18) .. (18)  
 <223> Asn, His o Arg
- <220>
- 10 <221> MOD\_RES  
 <222> (19) .. (19)  
 <223> Cualquier aminoácido
- <220>
- <221> MOD\_RES
- 15 <222> (21) .. (21)  
 <223> Ile o Leu
- <400> 43
- Cys Xaa Xaa Xaa Gln Xaa Xaa Xaa Xaa Glu Xaa Phe Ala Xaa Xaa Ala**  
**1 5 10 15**
- Tyr Xaa Xaa Tyr Xaa**  
**20**
- <210> 44
- 20 <211> 26  
 <212> PRT  
 <213> Secuencia artificial
- <220>  
 <223> Descripción de la secuencia artificial: péptido sintético
- 25 <220>  
 <221> MOD\_RES  
 <222> (2) .. (2)  
 <223> Cualquier aminoácido
- <220>
- 30 <221> MOD\_RES  
 <222> (3) .. (3)  
 <223> Phe o Tyr
- <220>

- <221> MOD\_RES  
 <222> (5) .. (5)  
 <223> Arg o Lys  
 <220>
- 5 <221> MOD\_RES  
 <222> (6) .. (6)  
 <223> Cualquier aminoácido  
 <220>
- <221> MOD\_RES  
 10 <222> (7) .. (7)  
 <223> Ile o Val  
 <220>
- <221> MOD\_RES  
 <222> (9) .. (10)  
 15 <223> Cualquier aminoácido  
 <220>
- <221> MOD\_RES  
 <222> (12) .. (12)  
 <223> Val o Met  
 20 <220>
- <221> MOD\_RES  
 <222> (13) .. (13)  
 <223> Cualquier aminoácido  
 <220>
- 25 <221> MOD\_RES  
 <222> (14) .. (14)  
 <223> Asp o Glu  
 <220>
- <221> MOD\_RES  
 30 <222> (16) .. (16)  
 <223> Tyr o Ser  
 <220>
- <221> MOD\_RES  
 <222> (17) .. (17)  
 35 <223> Cualquier aminoácido  
 <220>
- <221> MOD\_RES

<222> (19) .. (19)  
 <223> Leu o Arg  
 <220>  
 <221> MOD\_RES  
 5 <222> (20) .. (20)  
 <223> Asn o Asp  
 <220>  
 <221> MOD\_RES  
 <222> (21) .. (21)  
 10 <223> Tyr o Phe  
 <220>  
 <221> MOD\_RES  
 <222> (22) .. (22)  
 <223> Ala o Gly  
 15 <220>  
 <221> MOD\_RES  
 <222> (23) .. (23)  
 <223> Glu o Gln  
 <220>  
 20 <221> MOD\_RES  
 <222> (24) .. (24)  
 <223> Cualquier aminoácido  
 <400> 44  
**Asp Xaa Xaa Ala Xaa Xaa Xaa Thr Xaa Xaa Val Xaa Xaa Xaa Glu Xaa**  
**1 5 10 15**  
  
**Xaa His Xaa Xaa Xaa Xaa Xaa Xaa Trp Leu**  
**20 25**  
 25 <210> 45  
 <211> 699  
 <212> ADN  
 <213> Secuencia artificial  
 <220>  
 30 <223> Descripción de la secuencia artificial: polinucleótido sintético  
 <400> 45

ES 2 696 773 T3

atgcagcaac tgacggatca gagcaaagaa ctggacttca aaagcgaac ctacaaggac 60  
 gcgtattctc gtatcaacgc tatcgttatc gagggtgaac aagaagcgca cgagaattac 120  
 attacctgg cgcagctgct gcctgaatcc cacgatgaac tgattcgtct gagcaaaatg 180  
 gagtgcgctc acaaaaaggg ttttgaggcc tgcggtcgta acctggcggc cactccggac 240  
 ctgcagttcg ctaaggagtt cttcagcggc ctgcatcaaa actttcagac ggcagcggcg 300  
 gaaghtaagg ttgtcacctg cctgctgatt caaagcctga tcattgagtg tttcgctatc 360  
 gcagcctata acatttacat cccggtggcg gacgattttg cacgcaagat cactgagggc 420  
 gtggttaaag aagaatacag ccacctgaac ttcggtgagg tctggttgaa ggagcacttt 480  
 gcggaaagca aggcggagct ggaattggca aatcgtcaaa acctgccgat cgtgtggaaa 540  
 atgctgaatc aagtggaggg tgatgcacac acgatggcta tggaaaaaga cgctctggtg 600  
 gaggacttca tgatccagta cggcggaggc ctgagcaaca ttggctttag caccctgac 660  
 attatgcgcc tgagcgcgta tggcctgatc ggtgcgtaa 699

<210> 46

<211> 696

5 <212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 46

atgccgcaaa cgcaagctat tagcgaatt gatttctatt ctgacaccta taaggacgct 60  
 tactctcgta tcgatggat cgtgatcgag ggtgagcaag aggcgcatga gaactacatt 120  
 cgtctgggtg aaatggtgcc tgagcatcaa gacgacttta tccgtttgag caagatggag 180  
 gcccgtcaca agaaggcctt tgaggcttgt ggtcgtaact tgaagtgac ttgcatctg 240  
 gacttcgcgc gtcgcttctt ctcgacactg cacaagaact tccaagatgc tgcggccgag 300  
 gataaagttc cgacctgctt ggttattcag tcctgatca tcgaatgctt cgcgattgca 360  
 gcgtataaca ttacatccc ggttgccgat gatttcgctc gtaagattac cgagagcgtc 420  
 gtcaaggacg aataccagca tctgaaactat ggcgaggagt ggctgaaggc ccatttcgac 480  
 gacgtgaagg ccgagatcca ggaagcaaat cgcaagaatc tgccgatcgt ttggcgtatg 540  
 ctgaacgagg ttgacaagga cgcagcagtg ctgggcatgg agaaggaagc gttggttgaa 600  
 gacttcatga ttcaatacgg tgaggccctg tccaacattg gcttttctac cggcgagatc 660  
 atgcgtatgt ctgcgtacgg tctggtggca gcctaa 696

10

ES 2 696 773 T3

<210> 47

<211> 696

<212> ADN

<213> Secuencia artificial

5 <220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 47

```
atgaccaccg cgaccgcaac gccggtgctg gactatcaca gcgaccgcta caaggacgca      60
tacagccgca tcaacgcgat tgtcatcgaa ggtgaacaag aggcccacga caattacatt      120
gatctggcta aactgctgcc tcaacaccaa gaagagctga cccgtctggc gaagatggag      180
gcccgccaca agaagggttt tgaagcgtgc ggtcgcaatc tgtccgttac cccggatatg      240
gagttcgcga aagcgttctt tgagaagctg cgcgcgaact ttcagcgtgc cctggcggag      300
ggtaagaccg caacctgtct gctgatccag gcgttgatca ttgaatcctt cgcaattgcc      360
gcgtacaaca ttacatccc tatggccgat ccgtttgccg gcaagattac cgaaagcgtc      420
gtcaaggatg aatactctca cttgaacttt ggcgaaatct ggttgaagga acatttcgag      480
agcgtcaagg gcgagttgga ggaagctaac cgtgcgaatc tgccgctggt ttggaagatg      540
ttgaatcagg tcgaggcaga cgcaaaggtc ctgggcatgg agaaggatgc tctggtggaa      600
gactttatga tccagtactc cggtgcgctg gagaacatcg gctttaccac ccgtgaaatc      660
atgaaaatgt ctgtgtatgg cctgaccggc gcgtaa      696
```

10 <210> 48

<211> 732

<212> ADN

<213> Secuencia artificial

<220>

15 <223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 48

ES 2 696 773 T3

atggcgcctg caaacgtgct gccaaatacg ccgccgagcc cgaccgatgg tggtggtacg	60
gccctggact acagctctcc gcgttaccgt caggcgtaca gccgtatcaa tggcattggt	120
atcgaaggcg agcaggaagc gcacgataac tacctgaagt tggcggagat gctgcctgag	180
gctgccgagg aactgcgtaa gctggcaaag atggaattgc gtcacatgaa gggctttcag	240
gcttgccgca agaacttgca ggtggagcct gacgtcgagt ttgcccgcgc tttcttcgcg	300
ccgctgcgcg acaacttcca atccgcagca gcggccggtg atctggtttc ctgtttcgtc	360
atccaaagcc tgatcatcga gtgttttgcg atcgtcgcgt ataacattta catcccggtt	420
gcagacgact tcgcccgtaa gatcacggag ggcgtgggta aggacgagta tctgcatctg	480
aatttcggcg agcgttggtt gggngaacac ttcgcagagg ttaaagcaca gatcgaggca	540
gccaatgccc agaacctgcc gctggtgcmc caaatgctgc agcaagttga ggcggacgtc	600
gaggcaatct atatggaccg tgaggcgatc gttgaggatt tcatgattgc ttatggcgaa	660
gcgctggcaa gcattggctt caacacgcgc gaagtgatgc gtctgagcmc acagggcttg	720
cgtgcagcat aa	732

<210> 49

<211> 732

<212> ADN

5 <213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 49

atgccgacgt tggagatgcc ggtcgcctgc gtcctggaca gcacggtcgg tagctctgag	60
gcgctgccgg actttaccag cgaccgctac aaagacgctt attcgcgtat caacgcgatt	120
gtgatcgagg gtgaacaaga agcccacgac aactacatcg caattggcac cctgttgccg	180

ES 2 696 773 T3

gaccatgtgg aagaactgaa acgtctggcg aaaatggaaa tgcgtcacia gaaaggtttc 240  
 accgcgtgcg gtaagaactt ggggtgtgaa gccgatatgg acttcgcccg tgagttcttt 300  
 gccccgttgc gcgacaactt tcaaaccgcg ctgggtcaag gcaagacccc tacgtgtctg 360  
 ttgatccaag cgctgctgat tgaagcgttc gcgatctcgg cctaccacac ttacattccg 420  
 gttagcgatc cgttcgcacg taagatcact gaaggtgtcg ttaaggacga atacacccat 480  
 ctgaactacg gtgaggcatg gctgaaggcg aatctggaga gctgccgcga ggaactgctg 540  
 gaagcgaacc gtgagaatct gccgctgatc cgccgcatgc tggatcaggt cgcgggcgac 600  
 gcggcagtc tgcagatgga taaggaagac ctgatcgaag acttcctgat tgcttaccaa 660  
 gagagcttga ctgagatcgg ctttaacacg cgtgaaatca cccgtatggc cgcagcggcg 720  
 ctggtcagct aa 732

<210> 50

<211> 717

<212> ADN

5 <213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 50

atgcaaacc tggagagcaa caagaaaacc aacctggaaa acagcattga cctgccagat 60  
 ttcacgacgg acagctacia ggatgcgtat tcccgtatca atgctatcgt cattgaaggt 120  
 gaacaggaag cccatgacia ctatatcagc ctggccaccc tgatcccga tgaactggag 180  
 gaattgacca aactggccaa gatggagctg aaacaciaac gtggctttac ggcattgcgt 240  
 cgcaatctgg gtgttcaggc cgatatgatc tttgcgaaag agtttttctc taagctgcac 300  
 ggcaacttcc aagttgcgct gagcaacggt aagacgacca cctgcttgct gatccaggcc 360  
 atcttgattg aagccttcgc gatttccgcg taccacgtgt acattcgtgt cgcggacccg 420  
 tttgcgaaaa agattactca aggtgtggtg aaggatgagt acctgcacct gaactatggt 480  
 caggaatggt tgaaggagaa tctggcaacc tgtaaggacg aactgatgga agcaacaaa 540  
 gttaatctgc cgctgattaa gaaaatgctg gatcaggtga gcgaggatgc ctctgtgttg 600  
 gctatggatc gtgaggagct gatggaggag ttcattgatc cgtatcagga caccctgttg 660  
 gaaatcggtc tggacaatcg tgaattgcg cgtatggcaa tggctgcgat tgtgtaa 717

10 <210> 51

<211> 726

<212> ADN

ES 2 696 773 T3

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 51

atgcaggcct	tcgcaagcaa	taacctgacg	gtcgaaaagg	aagaactgag	ctccaatagc	60
ctgccggatt	tcaccagcga	gagctataag	gatgcatact	ctcgtatcaa	tgccgtggtt	120
atcgaagggtg	aacaagaggc	ttattctaac	tttctggacc	tggccaagct	gatcccggag	180
cacgccgacg	agctgggtgcg	cttgggtaag	atggaaaaga	aacacatgaa	cggcttctgc	240
gcgtgtggtc	gtaacttggc	agttaaacca	gacatgccgt	tcgccaagac	gttcttttagc	300
aagctgcaca	acaatttcct	ggaggcgttt	aaggtgggcg	atacgacgac	ctgtttgttg	360
atccaatgca	tcttgatcga	gtcctttgcc	atcagcgcgt	accacgtgta	cattcgcgtg	420
gcagatccgt	ttgccaagcg	tatcacggaa	ggtgttgttc	aagacgagta	cctgcatttg	480
aattacggtc	aagagtggct	gaaagcgaac	ctggagactg	tgaagaaaga	cctgatgcgc	540
gcgaacaaag	agaatctgcc	attgattaag	tctatgctgg	acgaagtctc	caacgacgct	600
gaagtgctgc	acatggataa	agaagagctg	atggaagagt	ttatgattgc	atatcaggac	660
agcctgctgg	aaattggcct	ggacaaccgc	gagatcgcac	gcatggcgct	ggcagcggtt	720
atttaa						726

5

<210> 52

<211> 732

<212> ADN

<213> Secuencia artificial

10

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 52

ES 2 696 773 T3

atgccgaccc tggaaactag cgaggtggca gttctggaag actcgatggc cagcggtagc	60
cgctgccgg actttaccag cgaggcctat aaggacgcgt atagccgtat caatgcgatc	120
gtgattgaag gcgagcaaga agcgcgatgac aactacattg cactggggcac gctgatccca	180
gaacagaagg acgagctggc tcgcctggct cgtatggaaa tgaaacacat gaagggcttt	240
accagctgtg gtcgtaacct ggggtgtgaa gcggatctgc cgttcgcgaa ggagttcttc	300
gcaccgctgc atggtaactt tcaggcggcg ctgcaggaag gtaaggtggg gacctgtctg	360
ctgattcagg cactgctgat tgaggcgttc gccattagcg cttatcacat ttacattccg	420
gttgctgacc cgtttgcacg caagattacc gaaggtggtg tgaaagacga gtatacccat	480
ctgaactacg gtcaagagtg gttgaaggcg aatttcgaag cctccaaaga cgaactgatg	540
gaagccaaca aggcgaatct gccgctgatc cgttctatgc tggaacaagt cgctgctgat	600
gcggccgtgc tgcaaatgga gaaagaggac ctgattgaag acttcctgat cgcatatcaa	660
gaagctctgt gtgagattgg cttctcgtcc cgtgatatcg cccgcatggc ggcagccgca	720
ctggcggttt aa	732

<210> 53

<211> 681

5 <212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polinucleótido sintético

<400> 53

ES 2 696 773 T3

atgaccaat tggactttgc atctgctgca taccgtgagg catacagccg tatcaatggt 60  
 gtcgttattg ttggcgaggg cctggcgaat cgtcacttcc aaatgctggc gcgtcgcatt 120  
 ccggcagacc gtgacgaatt gcaacgtttg ggccgcatgg agggtgacca cgcaagcgcc 180  
 tttgttggtt gcggtcgsaa tctgggtgtg gtcgctgata tgccgctggc acgccgcctg 240  
 ttccagccgc tgcatactct gttcaagcgt cacgaccacg acggtaaccg tgctgaatgc 300  
 ctggtgatcc agggctctgat tgttgagtgc tttgctggtt ccgctgatac tcattacctg 360  
 ccggtggcag acgctgatac ccgtccgata accgctgctg ttatgaatga cgagagcgaa 420  
 cacctggact acgcagaaac ctggctgcag cgcacttctg accaagttaa agcccgcgtg 480  
 agcgtctgtg ttgtggaggc gctgccgctg acgctggcga tgttgcaaag cctggctgca 540  
 gatatgcgcc aaatcggcat ggacccggtg gaaacgctgg cgagcttcag cgagctgttt 600  
 cgtgaagcgc tggaaagcgt tggttttgaa gcggtcgaag cgcgccgttt gctgatgcgt 660  
 gctgcagctc gtatggttta a 681

<210> 54

<211> 28

<212> PRT

5 <213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: péptido sintético

<220>

<221> MOD\_RES

10 <222> (3) .. (3)

<223> Cualquier aminoácido

<220>

<221> MOD\_RES

<222> (9) .. (12)

15 <223> Cualquier aminoácido

<220>

<221> MOD\_RES

<222> (14) .. (26)

<223> Cualquier aminoácido y esta región puede abarcar 12 ó 13 residuos

20 <400> 54

ES 2 696 773 T3

Gly Ala Xaa Gly Asp Ile Gly Ser Xaa Xaa Xaa Xaa Trp Xaa Xaa Xaa  
1 5 10 15

Xaa Ala Arg  
20 25

<210> 55

<211> 34

<212> PRT

5 <213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: polipéptido sintético

<220>

<221> MOD\_RES

10 <222> (5) .. (5)

<223> Val o Ile

<220>

<221> MOD\_RES

<222> (6) .. (6)

15 <223> Cualquier aminoácido

<220>

<221> MOD\_RES

<222> (17) .. (17)

<223> Cys o Thr

20 <220>

<221> MOD\_RES

<222> (21) .. (21)

<223> Asp o Ile

<220>

25 <221> MOD\_RES

<222> (22) .. (22)

<223> Cualquier aminoácido

<220>

<221> MOD\_RES

30 <222> (24) .. (27)

<223> Cualquier aminoácido

<220>

<221> MOD\_RES

<222> (28) .. (28)

<223> Asp o Glu

<220>

<221> MOD\_RES

5 <222> (30) .. (30)

<223> Ile, Leu o Phe

<220>

<221> MOD\_RES

<222> (32) .. (32)

10 <223> Thr, Ile o Val

<400> 55

Ala Thr Val Ala Xaa Xaa Gly Ala Thr Gly Asp Ile Gly Ser Ala Val  
 1 5 10 15

Xaa Arg Trp Leu Xaa Xaa Lys Xaa Xaa Xaa Xaa Xaa Leu Xaa Leu Xaa  
 20 25 30

Ala Arg

<210> 56

<211> 11

15 <212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: péptido sintético

<220>

20 <221> MOD\_RES

<222> (1) .. (1)

<223> Cualquier aminoácido

<220>

<221> MOD\_RES

25 <222> (3) .. (3)

<223> Cualquier aminoácido

<220>

<221> MOD\_RES

<222> (4) .. (4)

30 <223> Phe, Leu

<220>



<220>  
 <221> MOD\_RES  
 <222> (6) .. (6)  
 <223> Cualquier aminoácido  
 5 <220>  
 <221> MOD\_RES  
 <222> (10) .. (10)  
 <223> Gln o Val  
 <220>  
 10 <221> MOD\_RES  
 <222> (11) .. (11)  
 <223> Cualquier aminoácido  
 <220>  
 <221> MOD\_RES  
 15 <222> (12) .. (12)  
 <223> Leu o Val  
 <220>  
 <221> MOD\_RES  
 <222> (13) .. (13)  
 20 <223> Asp o Glu  
 <220>  
 <221> MOD\_RES  
 <222> (14) .. (16)  
 <223> Cualquier aminoácido  
 25 <220>  
 <221> MOD\_RES  
 <222> (17) .. (17)  
 <223> Val o Ile  
 <220>  
 30 <221> MOD\_RES  
 <222> (18) .. (18)  
 <223> Cualquier aminoácido  
 <400> 58

**Leu Xaa Xaa Trp Xaa Xaa Ala Pro Pro Xaa Xaa Xaa Xaa Xaa Xaa Xaa**  
**1 5 10 15**

35 <210> 59  
 <211> 21

<212> PRT

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: péptido sintético

5 <220>

<221> MOD\_RES

<222> (2) .. (3)

<223> Cualquier aminoácido

<220>

10 <221> MOD\_RES

<222> (5) .. (6)

<223> Cualquier aminoácido

<220>

<221> MOD\_RES

15 <222> (8) .. (8)

<223> Cualquier aminoácido

<220>

<221> MOD\_RES

<222> (10) .. (10)

20 <223> Cualquier aminoácido

<220>

<221> MOD\_RES

<222> (12) .. (13)

<223> Cualquier aminoácido

25 <220>

<221> MOD\_RES

<222> (15) .. (15)

<223> Cualquier aminoácido

<220>

30 <221> MOD\_RES

<222> (17) .. (17)

<223> Cualquier aminoácido

<400> 59

**Ser Xaa Xaa Gly Xaa Xaa Ile Xaa Gly Xaa Tyr Xaa Xaa Ser Xaa Phe**

1

5

10

15

**Xaa Pro Glu Met Leu**  
20

<210> 60

<211> 27

<212> PRT

5 <213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: péptido sintético

<220>

<221> MOD\_RES

10 <222> (2) .. (2)

<223> Cualquier aminoácido

<220>

<221> MOD\_RES

<222> (4) .. (4)

15 <223> Cualquier aminoácido

<220>

<221> MOD\_RES

<222> (7) .. (9)

<223> Cualquier aminoácido

20 <220>

<221> MOD\_RES

<222> (12) .. (14)

<223> Cualquier aminoácido

<220>

25 <221> MOD\_RES

<222> (16) .. (20)

<223> Cualquier aminoácido

<220>

<221> MOD\_RES

30 <222> (22) .. (23)

<223> Cualquier aminoácido

<400> 60

ES 2 696 773 T3

Lys Xaa Ala Xaa Arg Lys Xaa Xaa Xaa Ala Met Xaa Xaa Xaa Gln Xaa  
 1 5 10 15

Xaa Xaa Xaa Xaa Ile Xaa Xaa Leu Gly Gly Phe  
 20 25

<210> 61

<211> 14

<212> PRT

5 <213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: péptido sintético

<220>

<221> MOD\_RES

10 <222> (2) .. (2)

<223> Leu, Val o Met

<220>

<221> MOD\_RES

<222> (3) .. (4)

15 <223> Cualquier aminoácido

<220>

<221> MOD\_RES

<222> (5) .. (5)

<223> Ala o Ser

20 <220>

<221> MOD\_RES

<222> (6) .. (6)

<223> Asp o Asn

<220>

25 <221> MOD\_RES

<222> (7) .. (7)

<223> Cualquier aminoácido

<220>

<221> MOD\_RES

30 <222> (8) .. (9)

<223> Val o Ile

<220>

<221> MOD\_RES

<222> (14) .. (14)

<223> Leu, Met o Ile

<400> 61

**Ala Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Trp Val Ala Ser Xaa**  
**1 5 10**

5 <210> 62

<211> 12

<212> PRT

<213> Secuencia artificial

<220>

10 <223> Descripción de la secuencia artificial: péptido sintético

<220>

<221> MOD\_RES

<222> (2) .. (5)

<223> Cualquier aminoácido

15 <400> 62

**Pro Xaa Xaa Xaa Xaa Asp Gly Gly Tyr Pro Lys Asn**  
**1 5 10**

<210> 63

<211> 25

<212> PRT

20 <213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: péptido sintético

<220>

<221> MOD\_RES

25 <222> (8) .. (8)

<223> Cualquier aminoácido

<220>

<221> MOD\_RES

<222> (10) .. (16)

30 <223> Cualquier aminoácido

<220>

<221> MOD\_RES

<222> (19) .. (20)

<223> Cualquier aminoácido

35 <220>

<221> MOD\_RES

<222> (22) .. (23)

<223> Cualquier aminoácido

<400> 63

**Asn Phe Ser Trp Gly Arg Asn Xaa Ile Xaa Xaa Xaa Xaa Xaa Xaa Xaa**  
**1 5 10 15**

**Ile Gly Xaa Xaa Ser Xaa Xaa His Gly**  
**5 20 25**

<210> 64

<211> 9

<212> PRT

<213> Secuencia artificial

10 <220>

<223> Descripción de la secuencia artificial: péptido sintético

<400> 64

**Phe Thr Thr Gly Asn Thr His Thr Ala**  
**1 5**

<210> 65

15 <211> 1026

<212> ADN

<213> *Synechococcus elongatus*

<400> 65

**atgttcggtc ttatcgggtca tctcaccagt ttggagcagg cccgcgacgt ttctcgcagg 60**  
**atgggctacg acgaatacgc cgatcaagga ttggagtttt ggagtagcgc tcctcctcaa 120**  
**atcgttgatg aatcacagc caccagtgcc acaggcaagg tgattcacgg tcgctacatc 180**  
**gaatcgtgtt tcttgccgga aatgctggcg gcgcgccgct tcaaaacagc cacgcgcaaa 240**  
**gttctcaatg ccatgtocca tgcccaaaaa cacggcatcg acatctcggc cttgggggggc 300**

ES 2 696 773 T3

tttacctcga ttatthttcga gaatttcgat ttggccagtt tgcggcaagt gcgcgacact	360
accttggagt ttgaacgggt caccaccggc aatactcaca cggcctacgt aatctgtaga	420
caggtggaag ccgctgctaa aacgctgggc atcgacatta cccaagcgac agtagcggtt	480
gtcggcgcga ctggcgatat cggtagcgct gtctgccgct ggctcgacct caaactgggt	540
gtcggtgatt tgatcctgac ggcgcgcaat caggagcggt tggataacct gcaggctgaa	600
ctcggccggg gcaagattct gcccttgaa gccgctctgc cggagctga ctttatcgtg	660
tgggtcgcca gtatgcctca gggcgtagt atcgaccag caaccctgaa gcaaccctgc	720
gtcctaatacg acgggggcta ccccaaaaac ttgggcagca aagtccaagg tgagggcatc	780
tatgtcctca atggcgggggt agttgaacat tgcttcgaca tcgactggca gatcatgtcc	840
gctgcagaga tggcgcggcc cgagcgccag atgtttgcct gctttgccga ggcgatgctc	900
ttggaatttg aaggctggca tactaacttc tcctggggcc gcaaccaaat cacgatcgag	960
aagatggaag cgatcgggta ggcacgggt cggcacggct tccaaccctt ggcattggca	1020
atttga	1026

<210> 66

<211> 341

<212> PRT

5 <213> *Synechococcus elongatus*

<400> 66

ES 2 696 773 T3

Met Phe Gly Leu Ile Gly His Leu Thr Ser Leu Glu Gln Ala Arg Asp  
 1 5 10 15

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

Phe Trp Ser Ser Ala Pro Pro Gln Ile Val Asp Glu Ile Thr Val Thr  
 35 40 45

Ser Ala Thr Gly Lys Val Ile His Gly Arg Tyr Ile Glu Ser Cys Phe  
 50 55 60

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

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

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

Ser Leu Arg Gln Val Arg Asp Thr Thr Leu Glu Phe Glu Arg Phe Thr

ES 2 696 773 T3

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

<210> 67

ES 2 696 773 T3

<211> 1023

<212> ADN

<213> *Synechocystis sp.*

<400> 67

atgtttggtc ttattggtca tctcacgagt ttagaacacg cccaagcggg tgctgaagat	60
ttaggctatc ctgagtacgc caaccaaggc ctggatTTTT ggtggtcggc tcctccccaa	120
gtggttgata attttcaggt gaaaagtgtg acggggcagg tgattgaagg caaatatgtg	180
gagtcttgct ttttgccgga aatgttaacc caacggcgga tcaaagcggc cattcgtaa	240
atcctcaatg ctatggccct ggcccaaaag gtgggcttgg atattacggc cctgggaggc	300
ttttcttcaa tcgtatttga agaatttaac ctcaagcaa ataatcaagt ccgcaatgtg	360
gaactagatt ttcagcgggt caccactggt aatacccaca ccgcttatgt gatctgccgt	420
caggtcgagt ctggagctaa acagttgggt attgatctaa gtcaggcaac ggtagcgggt	480
tgtggcgcca cgggagatat tggtagcgcc gtatgtcggt ggtagatag caaacatcaa	540
gttaaggaat tattgctaatt tgcccgtaac cgccaaagat tggaaaatct ccaagaggaa	600
ttgggtcggg gcaaaaattat ggatttgaa acagccctgc ccaggcaga tattattggt	660
tgggtggcta gtatgcccga ggggtagaa attgccccggg aaatgctgaa aaagccctgt	720
ttgattgtgg atgggggcta tcccagaat ttagacacca gggtgaaagc ggatggggtg	780
catattctca agggggggat tgtagaacat tcccttgata ttacctgga aattatgaag	840
attgtggaga tggatattcc ctcccggca atgttcgcct gttttgcgga ggccattttg	900
ctagagtttg agggctggcg cactaatttt tcctggggcc gcaaccaa atccgtaaat	960
aaaatggagg cgattggtga agcttctgtc aagcatggct tttgccttt agtagctctt	1020
5 tag	1023

<210> 68

<211> 340

<212> PRT

<213> *Synechocystis sp.*

10 <400> 68

ES 2 696 773 T3

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

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

Phe Trp Cys Ser Ala Pro Pro Gln Val Val Asp Asn Phe Gln Val Lys  
35 40 45

Ser Val Thr Gly Gln Val Ile Glu Gly Lys Tyr Val Glu Ser Cys Phe  
50 55 60

Leu Pro Glu Met Leu Thr Gln Arg Arg Ile Lys Ala Ala Ile Arg Lys  
65 70 75 80

ES 2 696 773 T3

Ile Leu Asn Ala Met Ala Leu Ala Gln Lys Val Gly Leu Asp Ile Thr  
85 90 95

Ala Leu Gly Gly Phe Ser Ser Ile Val Phe Glu Glu Phe Asn Leu Lys  
100 105 110

Gln Asn Asn Gln Val Arg Asn Val Glu Leu Asp Phe Gln Arg Phe Thr  
115 120 125

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

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

Cys Gly Ala Thr Gly Asp Ile Gly Ser Ala Val Cys Arg Trp Leu Asp  
165 170 175

Ser Lys His Gln Val Lys Glu Leu Leu Leu Ile Ala Arg Asn Arg Gln  
180 185 190

Arg Leu Glu Asn Leu Gln Glu Glu Leu Gly Arg Gly Lys Ile Met Asp  
195 200 205

Leu Glu Thr Ala Leu Pro Gln Ala Asp Ile Ile Val Trp Val Ala Ser  
210 215 220

Met Pro Lys Gly Val Glu Ile Ala Gly Glu Met Leu Lys Lys Pro Cys  
225 230 235 240

Leu Ile Val Asp Gly Gly Tyr Pro Lys Asn Leu Asp Thr Arg Val Lys  
245 250 255

Ala Asp Gly Val His Ile Leu Lys Gly Gly Ile Val Glu His Ser Leu  
260 265 270

Asp Ile Thr Trp Glu Ile Met Lys Ile Val Glu Met Asp Ile Pro Ser  
275 280 285

Arg Gln Met Phe Ala Cys Phe Ala Glu Ala Ile Leu Leu Glu Phe Glu  
290 295 300

Gly Trp Arg Thr Asn Phe Ser Trp Gly Arg Asn Gln Ile Ser Val Asn  
305 310 315 320

Lys Met Glu Ala Ile Gly Glu Ala Ser Val Lys His Gly Phe Cys Pro

ES 2 696 773 T3

325

330

335

Leu Val Ala Leu  
340

<210> 69

<211> 1023

<212> ADN

5 <213> *Cyanothece sp.*

<400> 69

```

atgtttgggt taattgggtca ttttacaagt ttagaacacg cccactccgt tgctgatgcc      60
tttggctatg gcccatagcg cactcagga cttgatttgt ggtggtctgc tccaccccaa      120
ttcgtcgagc attttcatgt tactagcatc acaggacaaa ccatcgaagg aaagtatata      180
gaatccgctt tcttaccaga aatgctgata aagcgacgga ttaaagcagc aattcgcaaa      240
atactgaatg cgatggcctt tgctcagaaa aataacctta acatcacagc attagggggc      300
ttttcttoga ttatttttga agaatttaat ctcaaagaga atagacaagt tcgtaatgtc      360
tctttagagt ttgatcgctt caccaccgga aacaccata ctgcttatat catttgtcgt      420
caagttgaac aggcacccgc taaactaggg attgacttat cccaagcaac ggttgctatt      480
tgccggggcaa ccggagatat tggcagtgca gtgtgtcgtt ggtagatag aaaaaccgat      540
accaggaac tattcttaat tgctcgcaat aaagaacgat tacaacgact gcaagatgag      600
ttgggacggg gtaaaattat gggattggag gaggctttac ccgaagcaga tattatcgtt      660
tgggtggcga gtatgcccaa aggagtggaa attaatgccg aaactctcaa aaaaccctgt      720
ttaattatcg atggtgggta tcttaagaat ttagacacaa aattaaaca tctgatgtc      780
catatcctga aagggggaat tgtagaacat tctctagata ttgactggaa gattatggaa      840
actgtcaata tggatgttcc ttctcgtcaa atgtttgctt gttttgccga agccatttta      900
ttagagtttg aacaatggca cactaatttt tcttggggac gcaatcaaat tacagtgact      960
aaaatggaac aataggaga agcttctgtc aaacatgggt tacaaccggt gttgagttgg     1020
taa                                                                                   1023

```

<210> 70

<211> 340

10 <212> PRT

<213> *Cyanothece sp.*

<400> 70

ES 2 696 773 T3

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

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

ES 2 696 773 T3

Leu Trp Cys Ser Ala Pro Pro Gln Phe Val Glu His Phe His Val Thr  
 35 40 45  
 Ser Ile Thr Gly Gln Thr Ile Glu Gly Lys Tyr Ile Glu Ser Ala Phe  
 50 55 60  
 Leu Pro Glu Met Leu Ile Lys Arg Arg Ile Lys Ala Ala Ile Arg Lys  
 65 70 75 80  
 Ile Leu Asn Ala Met Ala Phe Ala Gln Lys Asn Asn Leu Asn Ile Thr  
 85 90 95  
 Ala Leu Gly Gly Phe Ser Ser Ile Ile Phe Glu Glu Phe Asn Leu Lys  
 100 105 110  
 Glu Asn Arg Gln Val Arg Asn Val Ser Leu Glu Phe Asp Arg Phe Thr  
 115 120 125  
 Thr Gly Asn Thr His Thr Ala Tyr Ile Ile Cys Arg Gln Val Glu Gln  
 130 135 140  
 Ala Ser Ala Lys Leu Gly Ile Asp Leu Ser Gln Ala Thr Val Ala Ile  
 145 150 155 160  
 Cys Gly Ala Thr Gly Asp Ile Gly Ser Ala Val Cys Arg Trp Leu Asp  
 165 170 175  
 Arg Lys Thr Asp Thr Gln Glu Leu Phe Leu Ile Ala Arg Asn Lys Glu  
 180 185 190  
 Arg Leu Gln Arg Leu Gln Asp Glu Leu Gly Arg Gly Lys Ile Met Gly  
 195 200 205  
 Leu Glu Glu Ala Leu Pro Glu Ala Asp Ile Ile Val Trp Val Ala Ser  
 210 215 220  
 Met Pro Lys Gly Val Glu Ile Asn Ala Glu Thr Leu Lys Lys Pro Cys  
 225 230 235 240  
 Leu Ile Ile Asp Gly Gly Tyr Pro Lys Asn Leu Asp Thr Lys Ile Lys  
 245 250 255  
 His Pro Asp Val His Ile Leu Lys Gly Gly Ile Val Glu His Ser Leu  
 260 265 270  
 Asp Ile Asp Trp Lys Ile Met Glu Thr Val Asn Met Asp Val Pro Ser  
 275 280 285

ES 2 696 773 T3

Arg Gln Met Phe Ala Cys Phe Ala Glu Ala Ile Leu Leu Glu Phe Glu  
 290 295 300

Gln Trp His Thr Asn Phe Ser Trp Gly Arg Asn Gln Ile Thr Val Thr  
 305 310 315 320

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

Leu Leu Ser Trp  
 340

<210> 71

<211> 1041

<212> ADN

5 <213> *Prochlorococcus marinus*

<400> 71

```

atgtttgggc ttataggtca ttcaactagt tttgaagatg caaaaagaaa ggcttcatta      60
ttgggctttg atcatattgc ggatggtgat ttagatgttt ggtgcacagc tccacctcaa      120
ctagttgaaa atgtagaggt taaaagtgct ataggtatat caattgaagg ttcttatatt      180
gattcatggt tcgttcctga aatgctttca agatttaaaa cggcaagaag aaaagtatta      240
aatgcaatgg aattagctca aaaaaaagggt attaatatta ccgctttggg ggggttcact      300
tctatcatct ttgaaaattt taatctcctt caacataagc agattagaaa cacttcacta      360
gagtgggaaa ggtttacaac tggttaact cactactgct gggttatttg caggcaatta      420
gagatgaatg ctcttaaaat aggtattgat cttaaaagcg caacagttgc tgtagttggt      480
gctactggag atataggcag tgctgtttgt cgatggttaa tcaataaaac aggtattggg      540
gaacttcttt tggtagctag gcaaaaggaa cccttggatt ctttgcaaaa ggaattagat      600
gggtggaacta tcaaaaatct agatgaagca ttgcctgaag cagatattgt tgtatgggta      660
gcaagtatgc caagacaat ggaaatcgat gctaataatc ttaaacaacc atgtttaatg      720
attgatggag gttatccaaa gaatctagat gaaaaatttc aaggaaataa tatacatggt      780
gtaaaaggag gtatagtaag attcttcaat gatataggtt ggaatatgat ggaactagct      840
gaaatgcaaa atccccagag agaaatgttt gcatgctttg cagaagcaat gattttagaa      900
tttgaaaaat gtcatacaaa ctttagctgg ggaagaaata atatatctct cgagaaaatg      960
gagtttattg gagctgcttc tgtaaagcat ggcttctctg caattggcct agataagcat     1020
ccaaaagtac tagcagtttg a                                             1041
    
```

<210> 72

<211> 346

<212> PRT

<213> *Prochlorococcus marinus*

5 <400> 72

ES 2 696 773 T3

Met Phe Gly Leu Ile Gly His Ser Thr Ser Phe Glu Asp Ala Lys Arg  
1 5 10 15

Lys Ala Ser Leu Leu Gly Phe Asp His Ile Ala Asp Gly Asp Leu Asp  
20 25 30

Val Trp Cys Thr Ala Pro Pro Gln Leu Val Glu Asn Val Glu Val Lys  
35 40 45

Ser Ala Ile Gly Ile Ser Ile Glu Gly Ser Tyr Ile Asp Ser Cys Phe  
50 55 60

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

Asn Ala Met Glu Leu Ala Gln Lys Lys Gly Ile Asn Ile Thr Ala Leu  
85 90 95

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

Lys Gln Ile Arg Asn Thr Ser Leu Glu Trp Glu Arg Phe Thr Thr Gly  
115 120 125

Asn Thr His Thr Ala Trp Val Ile Cys Arg Gln Leu Glu Met Asn Ala  
130 135 140

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

Ala Thr Gly Asp Ile Gly Ser Ala Val Cys Arg Trp Leu Ile Asn Lys  
165 170 175

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

Asp Ser Leu Gln Lys Glu Leu Asp Gly Gly Thr Ile Lys Asn Leu Asp  
195 200 205

Glu Ala Leu Pro Glu Ala Asp Ile Val Val Trp Val Ala Ser Met Pro  
210 215 220

Lys Thr Met Glu Ile Asp Ala Asn Asn Leu Lys Gln Pro Cys Leu Met  
225 230 235 240

ES 2 696 773 T3

Ile Asp Gly Gly Tyr Pro Lys Asn Leu Asp Glu Lys Phe Gln Gly Asn  
245 250 255

Asn Ile His Val Val Lys Gly Gly Ile Val Arg Phe Phe Asn Asp Ile  
260 265 270

Gly Trp Asn Met Met Glu Leu Ala Glu Met Gln Asn Pro Gln Arg Glu  
275 280 285

Met Phe Ala Cys Phe Ala Glu Ala Met Ile Leu Glu Phe Glu Lys Cys  
290 295 300

His Thr Asn Phe Ser Trp Gly Arg Asn Asn Ile Ser Leu Glu Lys Met  
305 310 315 320

Glu Phe Ile Gly Ala Ala Ser Val Lys His Gly Phe Ser Ala Ile Gly  
325 330 335

Leu Asp Lys His Pro Lys Val Leu Ala Val  
340 345

<210> 73

<211> 1053

<212> ADN

5 <213> *Gloeobacter violaceus*

<400> 73

ES 2 696 773 T3

atgtttggcc tgatcggaca cttgaccaat ctttcccatg cccagcgggt cgcccgcgac	60
ctgggctacg acgagtatgc aagccacgac ctcgaattct ggtgcatggc ccctccccag	120
gcggtcgatg aaatcacgat caccagcgtc accggtcagg tgatccacgg tcagtacgtc	180
gaatcgtgct ttctgccgga gatgctcgcc cagggccgct tcaagaccgc catgcgcaag	240
atcctcaatg ccatggccct ggtccagaag cgcggcatcg acattacggc cctgggaggc	300
ttctcgtcga tcatcttoga gaatttcagc ctcgataaat tgctcaacgt ccgcgacatc	360
accctcgaca tccagcgctt caccaccggc aacaccacaca cggcctacat cctttgtcag	420
caggtcgagc agggcgcggc acgctacggc atcgatccgg ccaaagcgac cgtggcggta	480
gtcggggcca ccggcgacat cggtagcgcc gtctgccgat ggctcaccga ccgcgccggc	540
atccacgaac tcttgctggt ggcccgcgac gccgaaaggc tcgaccggct gcagcaggaa	600
ctcggcaccg gtcggatcct gccggtcgaa gaagcacttc ccaaagccga catcgtcgtc	660
tgggcgcct cgatgaacca gggcatggcc atcgaccccg ccggcctgcg caccctctgc	720
ctgctcatcg acggcggcta cccaagaac atggccggca ccctgcagcg cccgggcatc	780
catatcctcg acggcggcat ggtcgagcac tcgctcgaca tcgactggca gatcatgtcg	840
tttctaaatg tgcccaacct cgcccgcag ttcttcgcct gcttcgccga gtcgatgctg	900
ctggaattcg aagggttca cttcaatctt tcttggggcc gcaaccacat caccgtcgag	960
aagatggccc agatcggctc gctgtctaaa aaacatggct ttcgtcccct gcttgaacct	1020
agtcagcgca gcggcgaact cgtacacgga taa	1053

<210> 74

<211> 350

5 <212> PRT

<213> *Gloeobacter violaceus*

<400> 74

ES 2 696 773 T3

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

Val Ala Arg Asp Leu Gly Tyr Asp Glu Tyr Ala Ser His Asp Leu Glu  
 20 25 30

Phe Trp Cys Met Ala Pro Pro Gln Ala Val Asp Glu Ile Thr Ile Thr  
 35 40 45

Ser Val Thr Gly Gln Val Ile His Gly Gln Tyr Val Glu Ser Cys Phe  
 50 55 60

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

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

Ala Leu Gly Gly Phe Ser Ser Ile Ile Phe Glu Asn Phe Ser Leu Asp  
 100 105 110

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

Thr Gly Asn Thr His Thr Ala Tyr Ile Leu Cys Gln Gln Val Glu Gln  
 130 135 140

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

Val Gly Ala Thr Gly Asp Ile Gly Ser Ala Val Cys Arg Trp Leu Thr  
 165 170 175

Asp Arg Ala Gly Ile His Glu Leu Leu Leu Val Ala Arg Asp Ala Glu  
 180 185 190

ES 2 696 773 T3

Arg Leu Asp Arg Leu Gln Gln Glu Leu Gly Thr Gly Arg Ile Leu Pro  
 195 200 205

Val Glu Glu Ala Leu Pro Lys Ala Asp Ile Val Val Trp Val Ala Ser  
 210 215 220

Met Asn Gln Gly Met Ala Ile Asp Pro Ala Gly Leu Arg Thr Pro Cys  
 225 230 235 240

Leu Leu Ile Asp Gly Gly Tyr Pro Lys Asn Met Ala Gly Thr Leu Gln  
 245 250 255

Arg Pro Gly Ile His Ile Leu Asp Gly Gly Met Val Glu His Ser Leu  
 260 265 270

Asp Ile Asp Trp Gln Ile Met Ser Phe Leu Asn Val Pro Asn Pro Ala  
 275 280 285

Arg Gln Phe Phe Ala Cys Phe Ala Glu Ser Met Leu Leu Glu Phe Glu  
 290 295 300

Gly Leu His Phe Asn Phe Ser Trp Gly Arg Asn His Ile Thr Val Glu  
 305 310 315 320

Lys Met Ala Gln Ile Gly Ser Leu Ser Lys Lys His Gly Phe Arg Pro  
 325 330 335

Leu Leu Glu Pro Ser Gln Arg Ser Gly Glu Leu Val His Gly  
 340 345 350

<210> 75

<211> 1020

<212> ADN

5 <213> *Nostoc punctiforme*

<400> 75

ES 2 696 773 T3

atgtttggtc taattggaca tctgactagt ttagaacacg ctcaagccgt agcccaagaa 60  
 ttgggataacc cagaatatgc cgatcaaggg ctagactttt ggtgcagcgc cccgccgcaa 120  
 attgtcgata gtattattgt caccagtgtt actgggcaac aaattgaagg acgatatgta 180  
 gaatcttgct ttttgccgga aatgctagct agtcgccgca tcaaagccgc aacacggaaa 240  
 atcctcaacg ctatggccca tgcacagaag cacggcatta acatcacagc tttaggcgga 300  
 ttttcctcga ttatTTTTga aaactttaag ttagagcagt ttagccaagt ccgaaatata 360  
 aagctagagt ttgaacgctt caccacagga aacacgcata ctgcctacat tatttgtaag 420  
 caggtggaag aagcatccaa acaactggga attaacttat caaacgcgac tgttgccgta 480  
 tgtggagcaa ctggggatat tggtagtgcc gttacacgct ggctagatgc gagaacagat 540  
 gtccaagaac tcctgctaata cgcccgcgat caagaacgct tcaaagagtt gcaaggcgaa 600  
 ctggggcggg ggaaaatcat gggtttgaca gaagcactac cccaagccga tgttgtagtt 660  
 tgggttgcta gtatgccag aggcgtggaa attgaccca ccactttgaa acaaccctgt 720  
 ttgttgattg atggtggcta tcctaaaaac ttagcaacia aaattcaata tcctggcgta 780  
 cacgtgttaa atggtgggat tgtagagcat tcctggata ttgactggaa aattatgaaa 840  
 atagtcaata tggacgtgcc agcccgtcag ttgtttgcct gttttgccga atcaatgcta 900  
 ctggaatttg agaagttata cacgaacttt tcgtggggac ggaatcagat taccgtagat 960  
 aaaatggagc agattggccg ggtgtcagta aaacatggat ttagaccggt gttggtttag 1020

<210> 76

<211> 339

5 <212> PRT

<213> *Nostoc punctiforme*

<400> 76

ES 2 696 773 T3

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

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

Phe Trp Cys Ser Ala Pro Pro Gln Ile Val Asp Ser Ile Ile Val Thr  
 35 40 45

Ser Val Thr Gly Gln Gln Ile Glu Gly Arg Tyr Val Glu Ser Cys Phe  
 50 55 60

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

Ile Leu Asn Ala Met Ala His Ala Gln Lys His Gly Ile Asn Ile Thr  
 85 90 95

Ala Leu Gly Gly Phe Ser Ser Ile Ile Phe Glu Asn Phe Lys Leu Glu  
 100 105 110

Gln Phe Ser Gln Val Arg Asn Ile Lys Leu Glu Phe Glu Arg Phe Thr  
 115 120 125

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

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

ES 2 696 773 T3

Cys Gly Ala Thr Gly Asp Ile Gly Ser Ala Val Thr Arg Trp Leu Asp  
 165 170 175

Ala Arg Thr Asp Val Gln Glu Leu Leu Leu Ile Ala Arg Asp Gln Glu  
 180 185 190

Arg Leu Lys Glu Leu Gln Gly Glu Leu Gly Arg Gly Lys Ile Met Gly  
 195 200 205

Leu Thr Glu Ala Leu Pro Gln Ala Asp Val Val Val Trp Val Ala Ser  
 210 215 220

Met Pro Arg Gly Val Glu Ile Asp Pro Thr Thr Leu Lys Gln Pro Cys  
 225 230 235 240

Leu Leu Ile Asp Gly Gly Tyr Pro Lys Asn Leu Ala Thr Lys Ile Gln  
 245 250 255

Tyr Pro Gly Val His Val Leu Asn Gly Gly Ile Val Glu His Ser Leu  
 260 265 270

Asp Ile Asp Trp Lys Ile Met Lys Ile Val Asn Met Asp Val Pro Ala  
 275 280 285

Arg Gln Leu Phe Ala Cys Phe Ala Glu Ser Met Leu Leu Glu Phe Glu  
 290 295 300

Lys Leu Tyr Thr Asn Phe Ser Trp Gly Arg Asn Gln Ile Thr Val Asp  
 305 310 315 320

Lys Met Glu Gln Ile Gly Arg Val Ser Val Lys His Gly Phe Arg Pro  
 325 330 335

Leu Leu Val

<210> 77

<211> 1020

<212> ADN

5 <213> *Anabaena variabilis*

<400> 77

ES 2 696 773 T3

atgtttggtc taattggaca tctgacaagt ttagaacacg ctcaagcggg agctcaagaa 60  
 ctgggatacc cagaatacgc cgaccaaggg ctagatTTTT ggtgcagcgc tccaccgcaa 120  
 atagttgacc acattaaagt tactagcatt actggtgaaa taattgaagg gaggtatgta 180  
 gaatcttgct ttttaccaga aatgctagcc agccgtagga ttaaagccgc aaccgcaaa 240  
 gtcctcaatg ctatggctca tgctcaaaaa catggcattg acatcaccgc tttgggtggt 300  
 ttctcctcca ttatTTTTga aaacttcaa ttggaacagt ttagccaagt tcgtaatgtc 360  
 aactagagt ttgaacgctt cactacaggc aactacaca cagcttatat catttgcgg 420  
 caggtagaac aagcatcaca acaactcggc attgaactct cccaagcaac agtagctata 480  
 tgtggggcta ctggtgacat tggtagtgca gttactcgtc ggctggatgc caaacagac 540  
 gtaaaagaat tactgttaat cgcccgtaat caagaacgtc tccaagagtt gcaaagcgag 600  
 ttgggacgcg gtaaaatcat gagcctagat gaagcattgc ctcaagctga tattgtagtt 660  
 tgggtagcta gtatgcctaa aggcgtggaa attaatcctc aagttttgaa acaaccctgt 720  
 ttattgattg atggtggtta tccgaaaaac ttgggtacaa aagttcagta tcctggtggt 780  
 tatgtactga acggaggtat cgtcgaacat tccctagata ttgactggaa aatcatgaaa 840  
 atagtcaata tggatgtacc tgcacgcaa ttatTTTgctt gttttgcgga atctatgctc 900  
 ttggaatttg agaagttgta cacgaacttt tcttgggggc gcaatcagat taccgtagac 960  
 aaaatggagc agattggtca agcatcagtg aaacatgggt ttagaccact gctggtttag 1020

<210> 78

<211> 339

5 <212> PRT

<213> *Anabaena variabilis*

<400> 78

ES 2 696 773 T3

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

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

Phe Trp Cys Ser Ala Pro Pro Gln Ile Val Asp His Ile Lys Val Thr  
 35 40 45

Ser Ile Thr Gly Glu Ile Ile Glu Gly Arg Tyr Val Glu Ser Cys Phe  
 50 55 60

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

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

Ala Leu Gly Gly Phe Ser Ser Ile Ile Phe Glu Asn Phe Lys Leu Glu  
 100 105 110

Gln Phe Ser Gln Val Arg Asn Val Thr Leu Glu Phe Glu Arg Phe Thr  
 115 120 125

ES 2 696 773 T3

Thr Gly Asn Thr His Thr Ala Tyr Ile Ile Cys Arg Gln Val Glu Gln  
 130 135 140

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

Cys Gly Ala Thr Gly Asp Ile Gly Ser Ala Val Thr Arg Trp Leu Asp  
 165 170 175

Ala Lys Thr Asp Val Lys Glu Leu Leu Leu Ile Ala Arg Asn Gln Glu  
 180 185 190

Arg Leu Gln Glu Leu Gln Ser Glu Leu Gly Arg Gly Lys Ile Met Ser  
 195 200 205

Leu Asp Glu Ala Leu Pro Gln Ala Asp Ile Val Val Trp Val Ala Ser  
 210 215 220

Met Pro Lys Gly Val Glu Ile Asn Pro Gln Val Leu Lys Gln Pro Cys  
 225 230 235 240

Leu Leu Ile Asp Gly Gly Tyr Pro Lys Asn Leu Gly Thr Lys Val Gln  
 245 250 255

Tyr Pro Gly Val Tyr Val Leu Asn Gly Gly Ile Val Glu His Ser Leu  
 260 265 270

Asp Ile Asp Trp Lys Ile Met Lys Ile Val Asn Met Asp Val Pro Ala  
 275 280 285

Arg Gln Leu Phe Ala Cys Phe Ala Glu Ser Met Leu Leu Glu Phe Glu  
 290 295 300

Lys Leu Tyr Thr Asn Phe Ser Trp Gly Arg Asn Gln Ile Thr Val Asp  
 305 310 315 320

Lys Met Glu Gln Ile Gly Gln Ala Ser Val Lys His Gly Phe Arg Pro  
 325 330 335

Leu Leu Val

<210> 79

<211> 1026

<212> ADN

ES 2 696 773 T3

<213> *Synechococcus elongatus*

<400> 79

atgttcggtc ttatcgggtca tctcaccagt ttggagcagg cccgcgacgt ttctcgcagg	60
atgggctacg acgaatacgc cgatcaagga ttggagtttt ggagtagcgc tcctcctcaa	120
atcgttgatg aatcacacgt caccagtgcc acaggcaagg tgattcacgg tcgctacatc	180
gaatcgtggt tcttgccgga aatgctggcg gcgcgccgct tcaaaacagc cacgcgcaa	240
gttctcaatg ccatgtccca tgcccaaaaa cacggcatcg acatctcggc cttggggggc	300
tttacctcga ttattttcga gaatttcgat ttggccagtt tgccgcaagt gcgcgacact	360
accttgaggt ttgaacgggt caccaccggc aatactcaca cggcctacgt aatctgtaga	420
caggtggaag ccgctgctaa aacgctgggc atcgacatta cccaagcgac agtagcgggt	480
gtcggcgcga ctggcgatat cggtagcgt gtctgccgct ggctcgacct caaactgggt	540
gtcggtgatt tgatcctgac ggcgcgcaat caggagcgtt tgataacct gcaggctgaa	600
ctcggccggg gcaagattct gcccttgaa gccgctctgc cgggaagctga ctttatcgtg	660
tgggtcgcca gtatgcctca gggcgtagt atcgaccag caaccctgaa gcaaccctgc	720
gtcctaatcg acgggggcta ccccaaaaac ttgggcagca aagtccaagg tgagggcatc	780
tatgtcctca atggcggggt agttgaacat tgcttcgaca togactggca gatcatgtcc	840
gctgcagaga tggcgcggcc cgagcgccag atgtttgcct gctttgccga ggcgatgctc	900
ttggaatttg aaggctggca tactaacttc tcctggggcc gcaaccaaat cacgatcgag	960
aagatggaag cgatcgggtga ggcacggtg cgccacggct tccaaccctt ggcatggca	1020
atttga	1026

<210> 80

5 <211> 340

<212> PRT

<213> *Synechococcus elongatus*

<400> 80

ES 2 696 773 T3

Met Phe Gly Leu Ile Gly His Leu Thr Ser Leu Glu Gln Ala Arg Asp  
1 5 10 15

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

Phe Trp Ser Ser Ala Pro Pro Gln Ile Val Asp Glu Ile Thr Val Thr  
35 40 45

Ser Ala Thr Gly Lys Val Ile His Gly Arg Tyr Ile Glu Ser Cys Phe  
50 55 60

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

ES 2 696 773 T3

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

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

Ser Leu Arg Gln Val Arg Asp Thr Thr Leu Glu Phe Glu Arg Phe Thr  
115 120 125

Thr Gly Asn Thr His Thr Ala Tyr Val Ile Cys Arg Gln Val Glu Ala  
130 135 140

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

Val Gly Ala Thr Gly Asp Ile Gly Ser Ala Val Cys Arg Trp Leu Asp  
165 170 175

Leu Lys Leu Gly Val Gly Asp Leu Ile Leu Thr Ala Arg Asn Gln Glu  
180 185 190

Arg Leu Asp Asn Leu Gln Ala Glu Leu Gly Arg Gly Lys Ile Leu Pro  
195 200 205

Leu Glu Ala Ala Leu Pro Glu Ala Asp Phe Ile Val Trp Val Ala Ser  
210 215 220

Met Pro Gln Gly Val Val Ile Asp Pro Ala Thr Leu Lys Gln Pro Cys  
225 230 235 240

Val Leu Ile Asp Gly Gly Tyr Pro Lys Asn Leu Gly Ser Lys Val Gln  
245 250 255

Gly Glu Gly Ile Tyr Val Leu Asn Gly Gly Val Val Glu His Cys Phe  
260 265 270

Asp Ile Asp Trp Gln Ile Met Ser Ala Ala Glu Met Ala Arg Pro Glu  
275 280 285

Arg Gln Met Phe Ala Cys Phe Ala Glu Ala Met Leu Leu Glu Phe Glu  
290 295 300

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

Lys Met Glu Ala Ile Gly Glu Ala Ser Val Arg His Gly Phe Gln Pro  
325 330 335

Leu Ala Leu Ala  
340

<210> 81

<211> 1020

<212> ADN

5 <213> *Nostoc sp.*

<400> 81

```

atgtttggtc taattggaca tctgacaagt ttagaacacg ctcaagcggg agctcaagaa      60
ctgggatacc cagaatacgc cgaccaaggg ctagatTTTT ggtgtagcgc tccaccgcaa      120
atagttgacc acattaaagt tactagtatt actggtgaaa taattgaagg gaggtatgta      180
gaatcttgct ttttaccgga gatgctagcc agtcgtcgga ttaaagccgc aaccgcaaaa      240
gtcctcaatg ctatggctca tgctcaaaag aatggcattg atatcacagc tttgggtggt      300
ttctcctcca ttatTTTTga aaactttaa ttggagcagt ttagccaagt tcgtaatgtg      360
acactagagt ttgaacgctt cactacaggc aacactcaca cagcatatat tatttgcgg      420
caggtagaac aagcatcaca acaactcggc attgaactct cccaagcaac agtagctata      480
tgtggggcta ctggtgatat tggtagtgca gttactcgct ggctggatgc taaaacagac      540
gtgaaagaat tgctgttaat cgcccgtaat caagaacgtc tccaagagtt gcaaagcgag      600
ctgggacgcg gtaaaatcat gagccttgat gaagcactgc cccaagctga tatcgtagtt      660
tgggtagcca gtatgcctaa aggtgtggaa attaatcctc aagttttgaa gcaaccctgt      720
ttgctgattg atggggggtta tccgaaaaac ttgggtacaa aagttcagta tcctgggtgtt      780
tatgtactga acggcgggat cgtcgaacat tcgctggata ttgactggaa aatcatgaaa      840
atagtcaata tggatgtacc tgcacgcaa ttatTTgctt gttttgcgga atctatgctc      900
ttggaatttg agaagttgta cacgaacttt tcttgggggc gcaatcagat taccgtagac      960
aaaatggagc agattgggtca agcatcagtg aaacatgggt ttagaccact gctggtttag     1020

```

<210> 82

<211> 339

10 <212> PRT

<213> *Nostoc sp.*

<400> 82

ES 2 696 773 T3

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

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

Phe Trp Cys Ser Ala Pro Pro Gln Ile Val Asp His Ile Lys Val Thr  
35 40 45

ES 2 696 773 T3

Ser Ile Thr Gly Glu Ile Ile Glu Gly Arg Tyr Val Glu Ser Cys Phe  
50 55 60

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

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

Ala Leu Gly Gly Phe Ser Ser Ile Ile Phe Glu Asn Phe Lys Leu Glu  
100 105 110

Gln Phe Ser Gln Val Arg Asn Val Thr Leu Glu Phe Glu Arg Phe Thr  
115 120 125

Thr Gly Asn Thr His Thr Ala Tyr Ile Ile Cys Arg Gln Val Glu Gln  
130 135 140

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

Cys Gly Ala Thr Gly Asp Ile Gly Ser Ala Val Thr Arg Trp Leu Asp  
165 170 175

Ala Lys Thr Asp Val Lys Glu Leu Leu Leu Ile Ala Arg Asn Gln Glu  
180 185 190

Arg Leu Gln Glu Leu Gln Ser Glu Leu Gly Arg Gly Lys Ile Met Ser  
195 200 205

Leu Asp Glu Ala Leu Pro Gln Ala Asp Ile Val Val Trp Val Ala Ser  
210 215 220

Met Pro Lys Gly Val Glu Ile Asn Pro Gln Val Leu Lys Gln Pro Cys  
225 230 235 240

Leu Leu Ile Asp Gly Gly Tyr Pro Lys Asn Leu Gly Thr Lys Val Gln  
245 250 255

Tyr Pro Gly Val Tyr Val Leu Asn Gly Gly Ile Val Glu His Ser Leu  
260 265 270

Asp Ile Asp Trp Lys Ile Met Lys Ile Val Asn Met Asp Val Pro Ala  
275 280 285

Arg Gln Leu Phe Ala Cys Phe Ala Glu Ser Met Leu Leu Glu Phe Glu

ES 2 696 773 T3

290

295

300

Lys Leu Tyr Thr Asn Phe Ser Trp Gly Arg Asn Gln Ile Thr Val Asp  
305 310 315 320

Lys Met Glu Gln Ile Gly Gln Ala Ser Val Lys His Gly Phe Arg Pro  
325 330 335

Leu Leu Val

<210> 83

<211> 1026

<212> ADN

5 <213> *Synechococcus elongatus*

<400> 83

```

atgtttggtc tgattgggtca cctgaccagc ttggaacaag cgcgtgacgt cagccgccgt      60
atgggttatg atgaatacgc tgatcaaggc ctggagtttt ggagcagcgc gccaccgcag      120
atcgtc gatg agatcaccgt gacctccgca accggttaagg tcatccacgg ccgctacatt      180
gagtctgct tctgctga gatgctggca gctcgcgctt tcaaaacggc cactcgtaag      240
gttctgaatg cgatgtccca tgcgcaaaag catggcattg acattagcgc cttggggcgt      300
tttacgtcga ttatcttcga gaacttcgat ctggcctctt tgcgccaggt gcgtgacacg      360
accttgaggt ttgagcgttt taccacgggt aatacgcaca ccgcttacgt tatctgtcgc      420
caagtcgaag cagcagccaa aaccctgggt attgatatca cccaggccac cgtcgcctgt      480
gtgggtgcta ccggtgatat tggttccgcg gtttgccggt ggctggatct gaaactgggt      540
gttggcgatc tgatcctgac ggcgcgtaat caggagcgtc tggacaacct gcaagccgag      600
ttgggtcgcg gtaagatcct gccggttgag gcagcgttgc cggaggcaga cttcatcgtc      660
tgggttgctg ctatgccgca ggggtgtgtt atcgaccggg cgacctgaa acagccgtgc      720
gtgctgattg atggcggcta tccgaaaaac ctgggcagca aggtccaagg cgaggggtatc      780
tatgtcctga atggcgggtg ggttgagcat tgcttcgaca ttgactggca gatcatgagc      840
gcagcagaaa tggcgcgtcc ggagcgcctc atgtttgcct gttttgcaga agccatgctg      900
ctggagttcg aaggctggca tacgaatttc agctggggtc gtaatcagat taccattgaa      960
aagatggaag cgattgggtga agcaagcgtg cgtcatgggt ttcagccact ggcgctggct     1020
atttaa                                           1026

```

<210> 84

<211> 1041

ES 2 696 773 T3

<212> ADN

<213> *Prochlorococcus marinus*

<400> 84

```

atgtttggtc tgattggcca cagcacgagc tttgaggacg caaagcgtaa ggcgagcctg      60
ctgggctttg atcatattgc tgatggcgac ctggacgtct ggtgcacggc acctccgcaa      120
ctggttgaga atgtcgaggt gaaatcggcg attggcattt ccatcgaagg ctccctacatc      180
gacagctggt tcgtgccgga gatgttgagc cgtttcaaaa ccgcacgtcg caaagttctg      240
aatgcaatgg agctggcaca aaagaagggc atcaacatca cggcgctggg tggtttcacc      300
agcattatct ttgagaactt caatctgttg cagcataaac agatccgtaa taccagcctg      360
gagtgggaac gctttaccac gggtaacacc cacaccgctg gggatgatctg ccgccagctg      420
gagatgaatg cgccgaaaat cggattgac ctgaaaagcg cgacgggtggc agttggtggc      480
gcaactggcg acattggttc ggccgtttgt cgctggctga ttaacaagac cggtatcggc      540
gaattggttc tggtcgctcg ccagaaggag cctctggaca gcctgcaaaa agagctggac      600
ggtggtacga tcaagaacct ggatgaagcg ctgccagaag cggacatcgt cgtctgggtc      660
gcatctatgc cgaaaactat ggaaatcgat gccacaatc tgaaacaacc gtgcctgatg      720
atcgatggcg gctacccgaa gaacttggat gagaagtttc aaggcaataa catccacgtt      780
gtgaagggtg gtattgtccg tttcttcaat gatatcggtt ggaacatgat ggaactggct      840
gaaatgcaga acccgcaacg tgagatgttc gcttgttttg cggaggccat gattctggag      900
ttcgagaaat gccataccaa tttcagctgg ggtcgcaaca acattagcct ggagaaaatg      960
gagttcatcg gcgctgcgag cgtaagcac ggcttcagcg cgattggttt ggataaacat     1020
ccgaaggtcc tggcagttta a                                             1041

```

5 <210> 85

<211> 3522

<212> ADN

<213> *Mycobacterium smegmatis*

<400> 85

ES 2 696 773 T3

atgaccagcg atgttcacga cgccacagac ggcgtcaccg aaaccgcact cgacgacgag	60
cagtcgaccc gccgcatcgc cgagctgtac gccaccgatc ccgagttcgc cgccgcccga	120
ccgttgcccc ccgtggtcga cgcggcgcac aaaccggggc tgcggctggc agagatcctg	180
cagaccctgt tcaccggcta cggtgaccgc ccggcgctgg gataccgcgc ccgtgaactg	240
gccaccgacg agggcggggc caccgtgacg cgtctgctgc cgcggttcga caccctcacc	300
tacgcccagg tgtggtcgcg cgtgcaagcg gtcgccgcgg ccctgcgcca caacttcgcg	360
cagccgatct accccggcga cgccgtcgcg acgatcggtt tcggagatcc cgattacctg	420
acgctggatc tcgtatgcgc ctacctgggc ctcgtgagtg ttccgctgca gcacaacgca	480
ccggtcagcc ggctcgcccc gatcctggcc gaggtcgaac cgcggatcct caccgtgagc	540
gccgaatacc tcgacctcgc agtcgaatcc gtgcgggacg tcaactcggg gtcgcagctc	600

ES 2 696 773 T3

gtggtggtcg	accatcacc	cgaggtcgac	gaccaccgcg	acgcactggc	ccgcgcgcgt	660
gaacaactcg	ccggcaaggg	catcgccgtc	accaccctgg	acgcgatcgc	cgacgagggc	720
gccgggctgc	cggccgaacc	gatctacacc	gccgaccatg	atcagcgcct	cgcgatgac	780
ctgtacacct	cgggttccac	cggcgcaccc	aagggtgcga	tgtacaccga	ggcgatggtg	840
gcgcggctgt	ggaccatgtc	gttcatcacg	ggtgaccca	cgccggtcat	caacgtcaac	900
ttcatgccgc	tcaaccacct	gggcggggcg	atccccattt	ccaccgccgt	gcagaacggg	960
ggaaccagtt	acttcgtacc	ggaatccgac	atgtccacgc	tgttcgagga	tctcgcgctg	1020
gtgcgcccga	ccgaactcgg	cctggttccg	cgcgtcgccg	acatgctcta	ccagcaccac	1080
ctcgccaccg	tcgaccgcct	ggtcacgcag	ggcgccgacg	aactgaccgc	cgagaagcag	1140
gccggtgccg	aactgcgtga	gcaggtgctc	ggcggacgcg	tgatcaccgg	attcgtcagc	1200
accgcaccgc	tggccgcgga	gatgagggcg	ttcctcgaca	tcaccctggg	cgcacacatc	1260
gtcgacggct	acgggctcac	cgagaccggc	gccgtgacac	gcgacgggtg	gatcgtgcgg	1320
ccaccggtga	tcgactacaa	gctgatcgac	gttcccgaac	tcggctactt	cagcaccgac	1380
aagccctacc	cgcgtaggca	actgctggtc	aggtagcaaa	cgctgactcc	cggttactac	1440
aagcgcgccg	aggtcaccgc	gagcgtcttc	gaccgggacg	gctactacca	caccggcgac	1500
gtcatggccg	agaccgcacc	cgaccacctg	gtgtacgtgg	accgtcgcaa	caacgtcctc	1560
aaactcgcgc	agggcgagtt	cgtggcggtc	gccaacctgg	aggcggtggt	ctccggcgcg	1620
gcgctggtgc	gccagatctt	cgtgtacggc	aacagcgagc	gcagtttctt	tctggccgtg	1680
gtggtcccga	cgccggaggc	gctcgagcag	tacgatccgg	ccgcgctcaa	ggccgcgctg	1740
gccgactcgc	tcgagcgcac	cgcacgcgac	gccgaactgc	aatcctacga	ggtgccggcc	1800
gatttcatcg	tcgagaccga	gccgttcagc	gccgccaacg	ggctgctgtc	gggtgtcgga	1860
aaactgctgc	ggcccaacct	caaagaccgc	tacgggcagc	gcctggagca	gatgtacgcc	1920
gatatcgcg	ccagcgaggc	caaccagttg	cgcgaaactgc	ggcgcgcggc	cgccacacaa	1980
ccggtgatcg	acaccctcac	ccaggccgct	gccacgatcc	tcggcaccgg	gagcgaggtg	2040
gcatccgacg	cccacttcac	cgacctgggc	ggggattccc	tgtagggcgt	gacactttcg	2100
aacctgctga	gcgatttctt	cggtttcgaa	gttcccgtcg	gcaccatcgt	gaaccgggcc	2160
accaacctcg	cccaactcgc	ccagcacatc	gaggcgcagc	gcaccgcggg	tgaccgcagg	2220
ccgagtttca	ccaccgtgca	cggcgcggac	gccaccgaga	tccgggcgag	tgagctgacc	2280
ctggacaagt	tcatcgacgc	cgaaaacgctc	cgggccgcac	cggtctgtcc	caaggtcacc	2340
accgagccac	ggacgggtgt	gctctcgggc	gccaacggct	ggctgggccc	gttctctcag	2400
ttgcagtggc	tggaacgcct	ggcacctgtc	ggcggcaccc	tcatcacgat	cgtgcggggc	2460

ES 2 696 773 T3

cgcgacgacg ccgcggcccg cgcacggctg acccaggcct acgacaccga tcccgagttg 2520  
 tcccgcgcgt tcgccgagct ggccgaccgc cacctgcggg tggtcgccgg tgacatcggc 2580  
 gacccgaatc tgggcctcac acccgagatc tggcaccggc tcgccgccga ggtcgacctg 2640  
 gtggtgcatc cggcagcgt ggtcaaccac gtgctcccct accggcagct gttcggcccc 2700  
 aacgtcgtgg gcacggccga ggtgatcaag ctggccctca ccgaacggat caagcccgtc 2760  
 acgtacctgt ccaccgtgtc ggtggccatg gggatccccg acttcgagga ggacggcgac 2820  
 atccggaccg tgagcccggg gcgcccgtc gacggcggat acgccaacgg ctacggcaac 2880  
 agcaagtggg ccggcgaggt gctgctgctg gaggcccacg atctgtgctg gctgcccgtg 2940  
 gcgacgttcc gctcggacat gatcctggcg catccgcgct accgcggtca ggtcaacgtg 3000  
 ccagacatgt tcacgcgact cctggtgagc ctcttgatca ccggcgtcgc gcgcggtcgc 3060  
 ttctacatcg gagacggtga gcgcccggg gcgcaactacc ccggcctgac ggtcgatttc 3120  
 gtggccgagg cggtcacgac gctcggcgcg cagcagcgcg agggatacgt gtctacgac 3180  
 gtgatgaacc cgcacgacga cgggatctcc ctggatgtgt tcgtggactg gctgatccgg 3240  
 gcgggcccac cgatcgaccg ggtcgacgac tacgacgact gggcgcgtcg gttcgagacc 3300  
 gcggtgaccg cgcttcccga gaagcgcgcg gcacagaccg tactgccgct gctgcacgcg 3360  
 ttccgcgctc cgcaggcacc gttgcggggc gcacccgaac ccacggaggt gttccacgcc 3420  
 gcggtgcgca ccgcgaaggt gggcccggga gacatcccgc acctcgacga ggcgctgatc 3480  
 gacaagtaca tacgcatct gcgtgagttc ggtctgatct ga 3522

<210> 86

<211> 1173

<212> PRT

5 <213> *Mycobacterium smegmatis*

<400> 86

ES 2 696 773 T3

Met Thr Ser Asp Val His Asp Ala Thr Asp Gly Val Thr Glu Thr Ala  
1 5 10 15

Leu Asp Asp Glu Gln Ser Thr Arg Arg Ile Ala Glu Leu Tyr Ala Thr  
20 25 30

Asp Pro Glu Phe Ala Ala Ala Ala Pro Leu Pro Ala Val Val Asp Ala  
35 40 45

Ala His Lys Pro Gly Leu Arg Leu Ala Glu Ile Leu Gln Thr Leu Phe  
50 55 60

Thr Gly Tyr Gly Asp Arg Pro Ala Leu Gly Tyr Arg Ala Arg Glu Leu  
65 70 75 80

ES 2 696 773 T3

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

Asp Thr Leu Thr Tyr Ala Gln Val Trp Ser Arg Val Gln Ala Val Ala  
100 105 110

Ala Ala Leu Arg His Asn Phe Ala Gln Pro Ile Tyr Pro Gly Asp Ala  
115 120 125

Val Ala Thr Ile Gly Phe Ala Ser Pro Asp Tyr Leu Thr Leu Asp Leu  
130 135 140

Val Cys Ala Tyr Leu Gly Leu Val Ser Val Pro Leu Gln His Asn Ala  
145 150 155 160

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

Leu Thr Val Ser Ala Glu Tyr Leu Asp Leu Ala Val Glu Ser Val Arg  
180 185 190

Asp Val Asn Ser Val Ser Gln Leu Val Val Phe Asp His His Pro Glu  
195 200 205

Val Asp Asp His Arg Asp Ala Leu Ala Arg Ala Arg Glu Gln Leu Ala  
210 215 220

Gly Lys Gly Ile Ala Val Thr Thr Leu Asp Ala Ile Ala Asp Glu Gly  
225 230 235 240

Ala Gly Leu Pro Ala Glu Pro Ile Tyr Thr Ala Asp His Asp Gln Arg  
245 250 255

Leu Ala Met Ile Leu Tyr Thr Ser Gly Ser Thr Gly Ala Pro Lys Gly  
260 265 270

Ala Met Tyr Thr Glu Ala Met Val Ala Arg Leu Trp Thr Met Ser Phe  
275 280 285

Ile Thr Gly Asp Pro Thr Pro Val Ile Asn Val Asn Phe Met Pro Leu  
290 295 300

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

Gly Thr Ser Tyr Phe Val Pro Glu Ser Asp Met Ser Thr Leu Phe Glu  
325 330 335

ES 2 696 773 T3

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

Ala Asp Met Leu Tyr Gln His His Leu Ala Thr Val Asp Arg Leu Val  
 355 360 365

Thr Gln Gly Ala Asp Glu Leu Thr Ala Glu Lys Gln Ala Gly Ala Glu  
 370 375 380

Leu Arg Glu Gln Val Leu Gly Gly Arg Val Ile Thr Gly Phe Val Ser  
 385 390 395 400

Thr Ala Pro Leu Ala Ala Glu Met Arg Ala Phe Leu Asp Ile Thr Leu  
 405 410 415

Gly Ala His Ile Val Asp Gly Tyr Gly Leu Thr Glu Thr Gly Ala Val  
 420 425 430

Thr Arg Asp Gly Val Ile Val Arg Pro Pro Val Ile Asp Tyr Lys Leu  
 435 440 445

Ile Asp Val Pro Glu Leu Gly Tyr Phe Ser Thr Asp Lys Pro Tyr Pro  
 450 455 460

Arg Gly Glu Leu Leu Val Arg Ser Gln Thr Leu Thr Pro Gly Tyr Tyr  
 465 470 475 480

Lys Arg Pro Glu Val Thr Ala Ser Val Phe Asp Arg Asp Gly Tyr Tyr  
 485 490 495

His Thr Gly Asp Val Met Ala Glu Thr Ala Pro Asp His Leu Val Tyr  
 500 505 510

Val Asp Arg Arg Asn Asn Val Leu Lys Leu Ala Gln Gly Glu Phe Val  
 515 520 525

Ala Val Ala Asn Leu Glu Ala Val Phe Ser Gly Ala Ala Leu Val Arg  
 530 535 540

Gln Ile Phe Val Tyr Gly Asn Ser Glu Arg Ser Phe Leu Leu Ala Val  
 545 550 555 560

Val Val Pro Thr Pro Glu Ala Leu Glu Gln Tyr Asp Pro Ala Ala Leu  
 565 570 575

Lys Ala Ala Leu Ala Asp Ser Leu Gln Arg Thr Ala Arg Asp Ala Glu  
 580 585 590

ES 2 696 773 T3

Leu Gln Ser Tyr Glu Val Pro Ala Asp Phe Ile Val Glu Thr Glu Pro  
 595 600 605  
 Phe Ser Ala Ala Asn Gly Leu Leu Ser Gly Val Gly Lys Leu Leu Arg  
 610 615 620  
 Pro Asn Leu Lys Asp Arg Tyr Gly Gln Arg Leu Glu Gln Met Tyr Ala  
 625 630 635 640  
 Asp Ile Ala Ala Thr Gln Ala Asn Gln Leu Arg Glu Leu Arg Arg Ala  
 645 650 655  
 Ala Ala Thr Gln Pro Val Ile Asp Thr Leu Thr Gln Ala Ala Ala Thr  
 660 665 670  
 Ile Leu Gly Thr Gly Ser Glu Val Ala Ser Asp Ala His Phe Thr Asp  
 675 680 685  
 Leu Gly Gly Asp Ser Leu Ser Ala Leu Thr Leu Ser Asn Leu Leu Ser  
 690 695 700  
 Asp Phe Phe Gly Phe Glu Val Pro Val Gly Thr Ile Val Asn Pro Ala  
 705 710 715 720  
 Thr Asn Leu Ala Gln Leu Ala Gln His Ile Glu Ala Gln Arg Thr Ala  
 725 730 735  
 Gly Asp Arg Arg Pro Ser Phe Thr Thr Val His Gly Ala Asp Ala Thr  
 740 745 750  
 Glu Ile Arg Ala Ser Glu Leu Thr Leu Asp Lys Phe Ile Asp Ala Glu  
 755 760 765  
 Thr Leu Arg Ala Ala Pro Gly Leu Pro Lys Val Thr Thr Glu Pro Arg  
 770 775 780  
 Thr Val Leu Leu Ser Gly Ala Asn Gly Trp Leu Gly Arg Phe Leu Thr  
 785 790 795 800  
 Leu Gln Trp Leu Glu Arg Leu Ala Pro Val Gly Gly Thr Leu Ile Thr  
 805 810 815  
 Ile Val Arg Gly Arg Asp Asp Ala Ala Ala Arg Ala Arg Leu Thr Gln  
 820 825 830  
 Ala Tyr Asp Thr Asp Pro Glu Leu Ser Arg Arg Phe Ala Glu Leu Ala



ES 2 696 773 T3

Pro Ile Asp Arg Val Asp Asp Tyr Asp Asp Trp Val Arg Arg Phe  
 1085 1090 1095

Glu Thr Ala Leu Thr Ala Leu Pro Glu Lys Arg Arg Ala Gln Thr  
 1100 1105 1110

Val Leu Pro Leu Leu His Ala Phe Arg Ala Pro Gln Ala Pro Leu  
 1115 1120 1125

Arg Gly Ala Pro Glu Pro Thr Glu Val Phe His Ala Ala Val Arg  
 1130 1135 1140

Thr Ala Lys Val Gly Pro Gly Asp Ile Pro His Leu Asp Glu Ala  
 1145 1150 1155

Leu Ile Asp Lys Tyr Ile Arg Asp Leu Arg Glu Phe Gly Leu Ile  
 1160 1165 1170

<210> 87

<211> 921

<212> ADN

5 <213> *Nostoc punctiforme*

<400> 87

ES 2 696 773 T3

atgactcaag cgaaagccaa aaaagaccac ggtgacgttc ctgttaacac ttaccgtccc	60
aatgctccat ttattggcaa ggtaatatct aatgaacat tagtcaaaga aggtggtatt	120
ggtattgttc aacaccttaa atttgaccta tctgggtgggg atttgaagta tatagaaggt	180
caaagtattg gcattattcc gccaggttta gacaagaacg gcaagcctga aaaactcaga	240
ctatattcca tcgcctcaac tcgtcatggt gatgatgtag atgataagac agtatcactg	300
tgcgtccgcc agttggagta caagcaccca gaaactggcg aaacagtcta cgggtgttgc	360
tctacgcacc tgtgtttcct caagccaggg gaagaggtaa aaattacagg gcctgtgggt	420
aaggaaatgt tgttacccaa tgaccctgat gctaattgta tcatgatggc tactggaaca	480
ggtattgcbc cgatgcgggc ttacttgtgg cgtcagttta aagatgcgga aagagcggct	540
aaccacagaat accaatttaa aggattctct tggctaatat ttggcgtacc tacaactcca	600
aaccttttat ataaggaaga actggaagag attcaacaaa aatatcctga gaacttccgc	660
ctaactgctg ccatcagccg cgaacagaaa aatccccaaag gcggtagaat gtatattcaa	720
gaccgcgtag cagaacatgc tgatgaattg tggcagttga ttaaaaatga aaaaaccac	780
acttacattt gcggtttgcg cggtatggaa gaaggtattg atgcagcctt aactgctgct	840
gctgctaagg aaggcgtaac ctggagtgat taccagaagc aactcaagaa agccggtcgc	900
tggcacgtag aaacttacta a	921

<210> 88

<211> 437

<212> PRT

5 <213> *Nostoc punctiforme*

<400> 88

ES 2 696 773 T3

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

Ser Arg Ile Phe Val Tyr Glu Val Val Gly Leu Arg Gln Gly Glu Glu  
 20 25 30

Thr Asp Gln Thr Asn Tyr Pro Ile Arg Lys Ser Gly Ser Val Phe Ile  
 35 40 45

Arg Val Pro Tyr Asn Arg Met Asn Gln Glu Met Arg Arg Ile Thr Arg  
 50 55 60

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

Val Asn Gly Lys Ala Ser Phe Gly Asn Ala Thr Ser Val Val Ser Glu  
 85 90 95

Leu Ala Lys Ser Gly Glu Thr Ala Asn Ser Glu Gly Asn Gly Lys Ala  
 100 105 110

Thr Pro Val Asn Ala His Ser Ala Glu Glu Gln Asn Lys Asp Lys Lys  
 115 120 125

Gly Asn Thr Met Thr Gln Ala Lys Ala Lys Lys Asp His Gly Asp Val  
 130 135 140

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

Ser Asn Glu Pro Leu Val Lys Glu Gly Gly Ile Gly Ile Val Gln His  
 165 170 175

Leu Lys Phe Asp Leu Ser Gly Gly Asp Leu Lys Tyr Ile Glu Gly Gln  
 180 185 190

Ser Ile Gly Ile Ile Pro Pro Gly Leu Asp Lys Asn Gly Lys Pro Glu  
 195 200 205

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

Asp Asp Lys Thr Val Ser Leu Cys Val Arg Gln Leu Glu Tyr Lys His



ES 2 696 773 T3

<213> *Nostoc punctiforme*

<400> 89

```

atgccaaactt ataaagtgac actaattaac gaggctgaag ggctgaacac aacccttgat      60
gttgaggacg atacctatat tctagacgca gctgaagaag ctggtattga cctgccctac      120
tcttgccgcg ctggtgcttg ctctacttgt gcaggtaaac tcgtatcagg taccgtcgat      180
caaggcgatc aatcattctt agatgacgat caaatagaag ctggatatgt actgacctgt      240
gttgcttacc caacttctaa tgtcacgatc gaaactcaca aagaagaaga actctattaa      300
  
```

5 <210> 90

<211> 99

<212> PRT

<213> *Nostoc punctiforme*

<400> 90

```

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

Thr Thr Leu Asp Val Glu Asp Asp Thr Tyr Ile Leu Asp Ala Ala Glu
 20                25                30

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

Thr Cys Ala Gly Lys Leu Val Ser Gly Thr Val Asp Gln Gly Asp Gln
 50                55                60

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

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

10 Glu Leu Tyr

<210> 91

<211> 369

<212> ADN

<213> *Nostoc punctiforme*

15 <400> 91

ES 2 696 773 T3

atgtcccgta catacacaat taaagttcgc gatcgcgcca ctggcaaac acacacccta 60  
 aaagtgccag aagaccgta tatcctgcac actgccgaaa aacaaggtgt ggaactaccg 120  
 ttttctgtc gcaacggagc ttgcaccgct tgtgctgtga gggattgtc aggagaaatt 180  
 tatcaaccag aggcgatcgg attgtcacca gatttacgct agcaaggta tgccctgttg 240  
 tgtgtgagtt atccccgttc tgacttgaa gtagagacac aagacgaaga tgaagtctac 300  
 gaactccagt ttgggcgcta ttttgctaag gggaaagta aagcgggtt accgttagat 360  
 gaggaataa 369

<210> 92

<211> 122

<212> PRT

5 <213> *Nostoc punctiforme*

<400> 92

Met Ser Arg Thr Tyr Thr Ile Lys Val Arg Asp Arg Ala Thr Gly Lys  
 1 5 10 15  
 Thr His Thr Leu Lys Val Pro Glu Asp Arg Tyr Ile Leu His Thr Ala  
 20 25 30  
 Glu Lys Gln Gly Val Glu Leu Pro Phe Ser Cys Arg Asn Gly Ala Cys  
 35 40 45  
 Thr Ala Cys Ala Val Arg Val Leu Ser Gly Glu Ile Tyr Gln Pro Glu  
 50 55 60  
 Ala Ile Gly Leu Ser Pro Asp Leu Arg Gln Gln Gly Tyr Ala Leu Leu  
 65 70 75 80  
 Cys Val Ser Tyr Pro Arg Ser Asp Leu Glu Val Glu Thr Gln Asp Glu  
 85 90 95  
 Asp Glu Val Tyr Glu Leu Gln Phe Gly Arg Tyr Phe Ala Lys Gly Lys  
 100 105 110  
 Val Lys Ala Gly Leu Pro Leu Asp Glu Glu  
 115 120

<210> 93

<211> 321

10 <212> ADN

<213> *Nostoc punctiforme*

<400> 93

```

atgcccaaaa cttacaccgt agaaatcgat catcaaggca aaattcatac cttgcaagtt      60
cctgaaaatg aaacgatctt atcagttgcc gatgctgctg gtttggaact gccgagttct      120
tgtaatgcag gtgtttgcac aacttgcgcc ggtcaaataa gccaggggaac tgtggatcaa      180
actgatggca tgggcgttag tccagattta caaaagcaag gttacgtatt gctttgtgtt      240
gcgaaacccc tttctgattt gaaacttgaa acagaaaagg aagacatagt ttatcagtta      300
caatttgcca aagacaaata a                                          321
    
```

<210> 94

<211> 106

5 <212> PRT

<213> *Nostoc punctiforme*

<400> 94

```

Met Pro Lys Thr Tyr Thr Val Glu Ile Asp His Gln Gly Lys Ile His
1          5          10          15

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

Ala Gly Leu Glu Leu Pro Ser Ser Cys Asn Ala Gly Val Cys Thr Thr
          35          40          45

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

Gly Val Ser Pro Asp Leu Gln Lys Gln Gly Tyr Val Leu Leu Cys Val
65          70          75          80

Ala Lys Pro Leu Ser Asp Leu Lys Leu Glu Thr Glu Lys Glu Asp Ile
          85          90          95

Val Tyr Gln Leu Gln Phe Gly Lys Asp Lys
          100          105
    
```

<210> 95

10 <211> 30

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: cebador sintético

15 <400> 95

	<b>cgcggatccc ttgattctac tgcggcgagt</b>	<b>30</b>
	<210> 96	
	<211> 53	
	<212> ADN	
5	<213> Secuencia artificial	
	<220>	
	<223> Descripción de la secuencia artificial: cebador sintético	
	<400> 96	
	<b>cacgcaccta ggttcacact cccatggtat aacaggggcg ttggactcct gtg</b>	<b>53</b>
10	<210> 97	
	<211> 53	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
15	<223> Descripción de la secuencia artificial: cebador sintético	
	<400> 97	
	<b>gttataccat gggagtgtga acctaggtgc gtggccgaca ggatagggcg tgt</b>	<b>53</b>
	<210> 98	
	<211> 30	
20	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Descripción de la secuencia artificial: cebador sintético	
	<400> 98	
25	<b>cgcggatcca acgcatcctc actagtcggg</b>	<b>30</b>
	<210> 99	
	<211> 38	
	<212> ADN	
	<213> Secuencia artificial	
30	<220>	
	<223> Descripción de la secuencia artificial: cebador sintético	
	<400> 99	
	<b>catgccaatgg aaagccacgt tgtgtctcaa aatctctg</b>	<b>38</b>
	<210> 100	
35	<211> 32	

<212> ADN

<213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: cebador sintético

5 <400> 100

**ctagtctaga gcgctgaggt ctgcctcgtg aa**

**32**

10

**REIVINDICACIONES**

1. Célula huésped no humana para su uso en la producción de un aldehído graso o un alcohol graso, dicha célula modificada por ingeniería genética para expresar un polinucleótido que comprende una secuencia de ácido nucleico que tiene al menos el 70% de identidad de secuencia con SEQ ID NO: 65 o 69, en la que la secuencia de ácido nucleico codifica para un polipéptido que tiene actividad acil-ACP o acil-CoA reductasa y que es de una cianobacteria.
2. Célula huésped no humana según la reivindicación 1, en la que el polinucleótido está comprendido por un vector recombinante.
3. Célula huésped no humana según la reivindicación 1, en la que el polinucleótido se incorpora de manera estable en el ADN genómico de la célula huésped.
4. Célula huésped no humana según una cualquiera de las reivindicaciones 1-3, en la que la célula huésped es una célula microbiana.
5. Célula huésped no humana según una cualquiera de las reivindicaciones 1-4, en la que la célula huésped es una célula bacteriana.
6. Célula huésped no humana según una cualquiera de las reivindicaciones 1-5, en la que la célula huésped es una célula de *Escherichia* o una célula cianobacteriana.

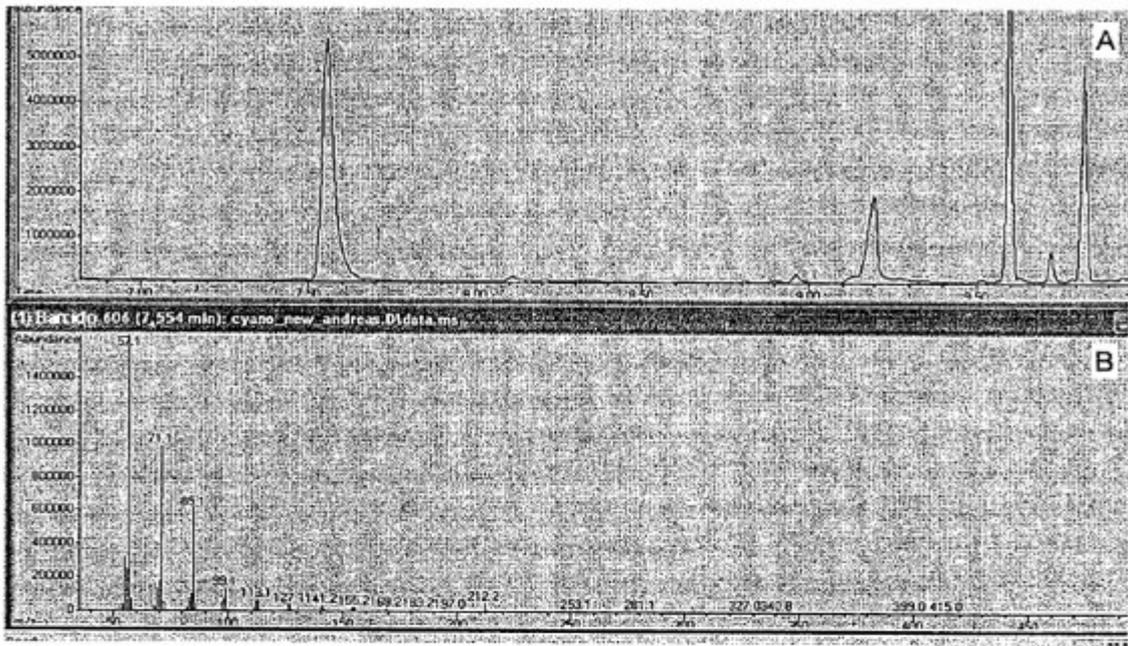


FIG. 1

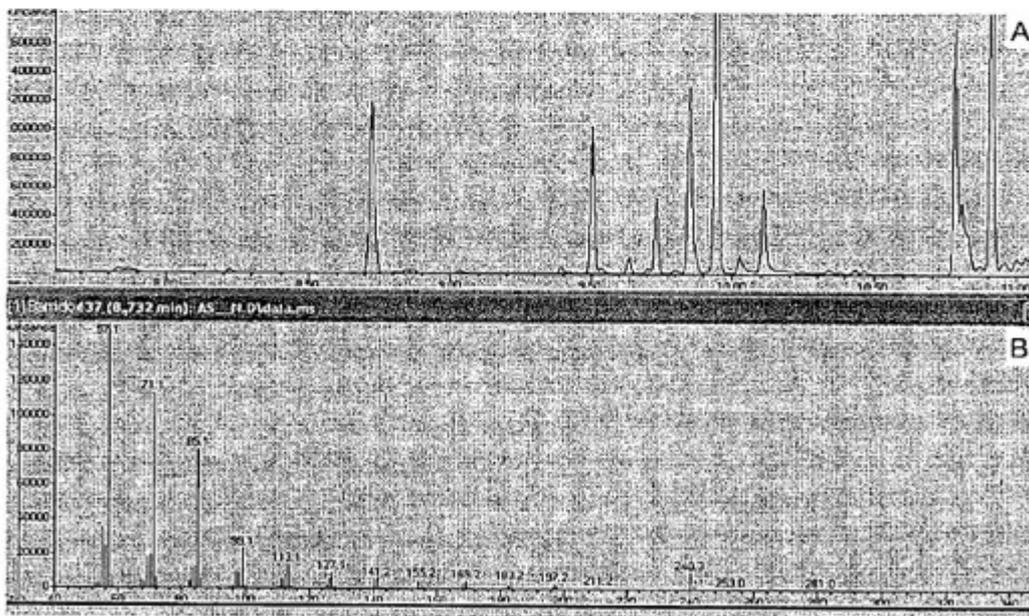


FIG. 2

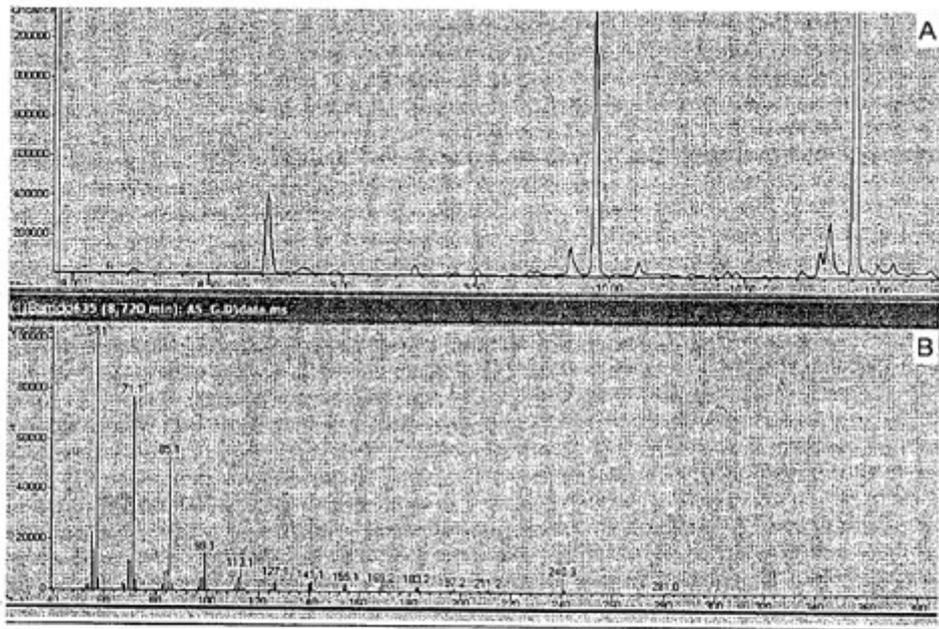


FIG. 3

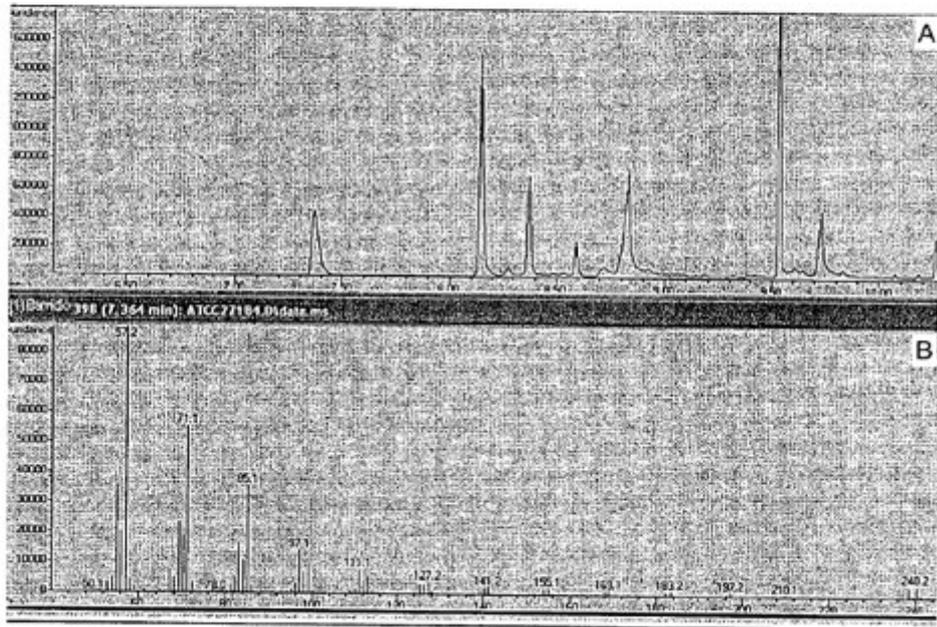


FIG. 4

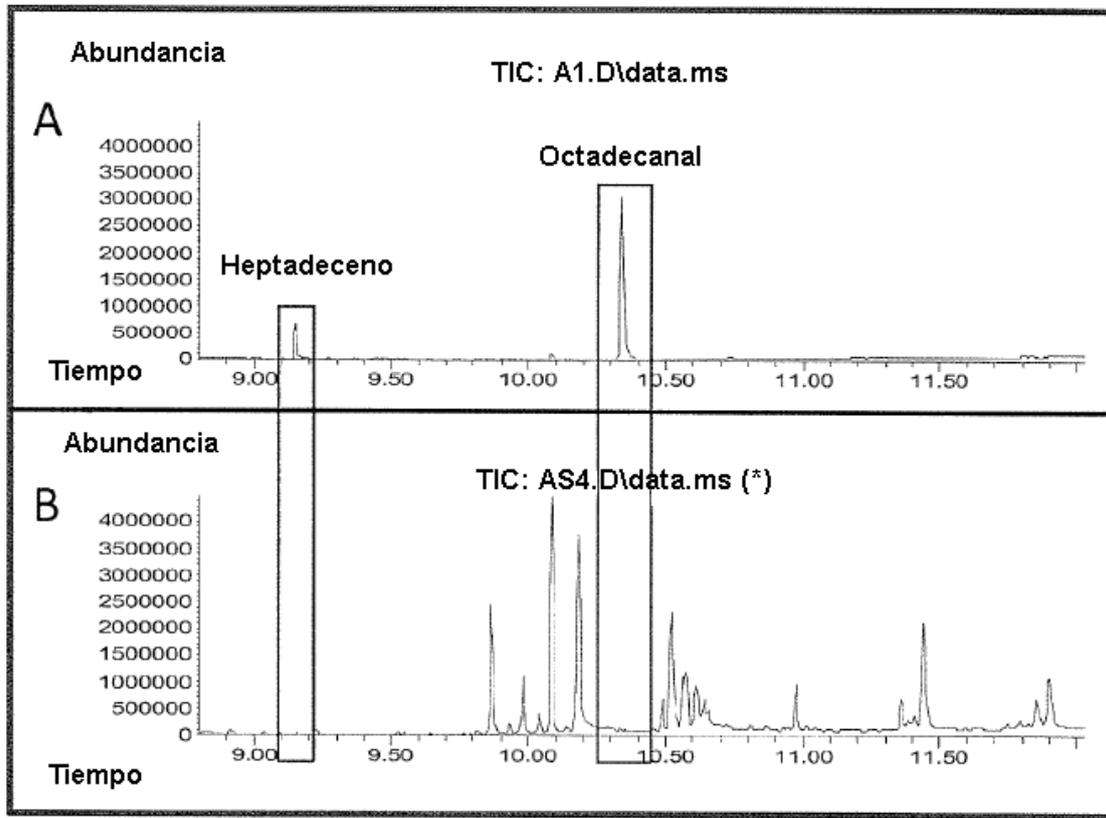


FIG. 5

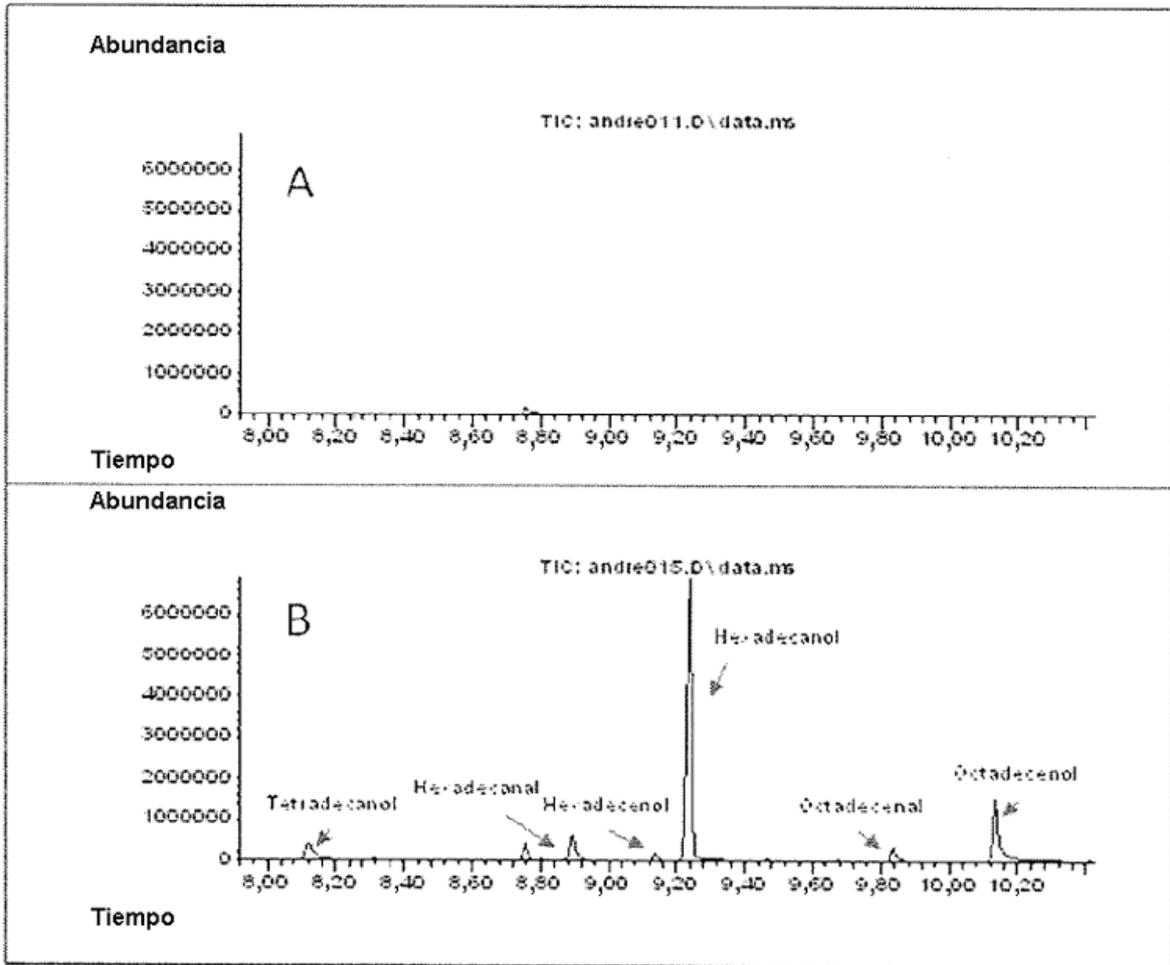


FIG. 6

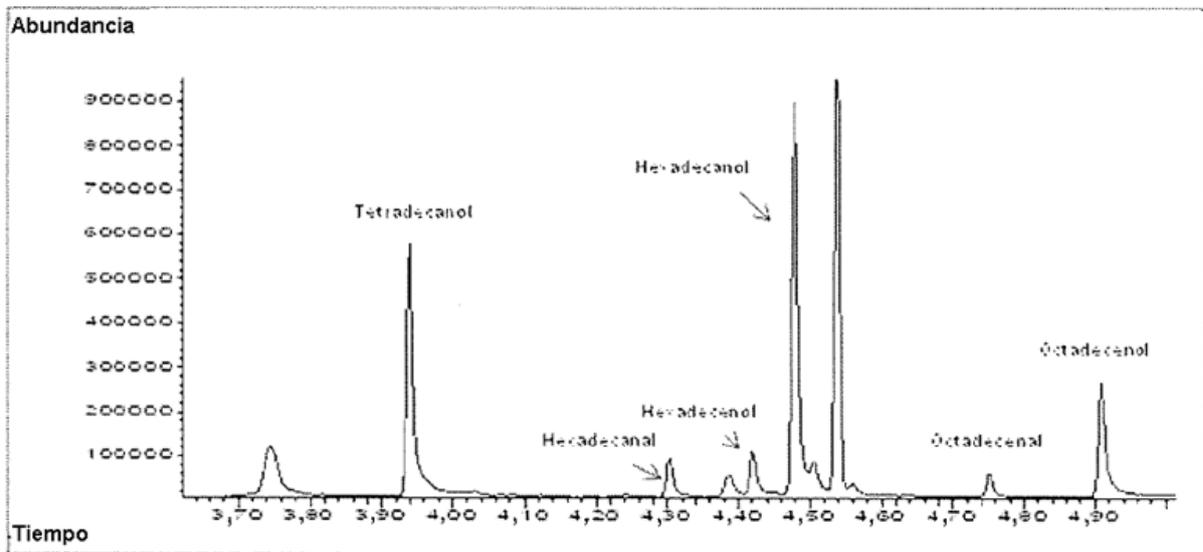


FIG. 7

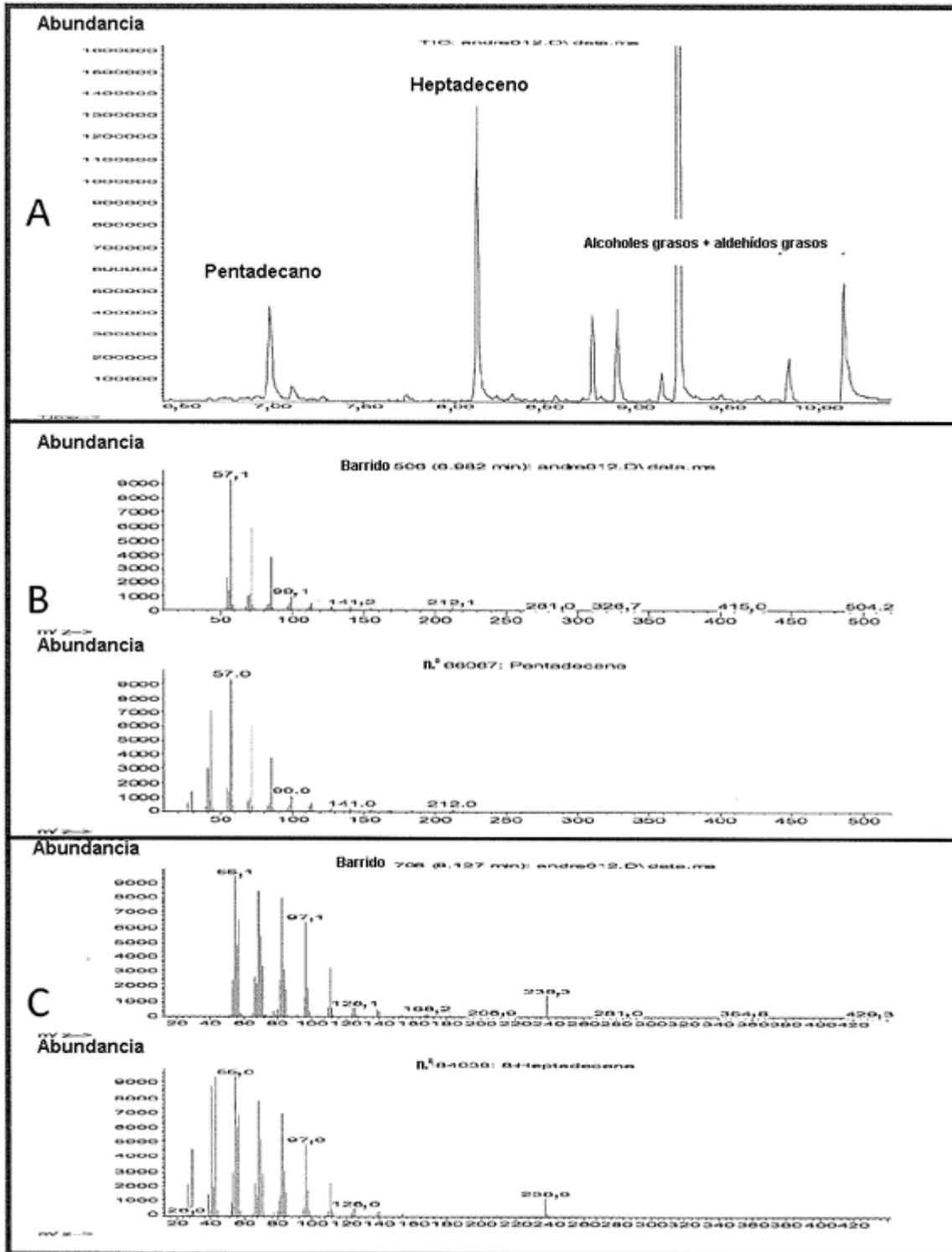


FIG. 8

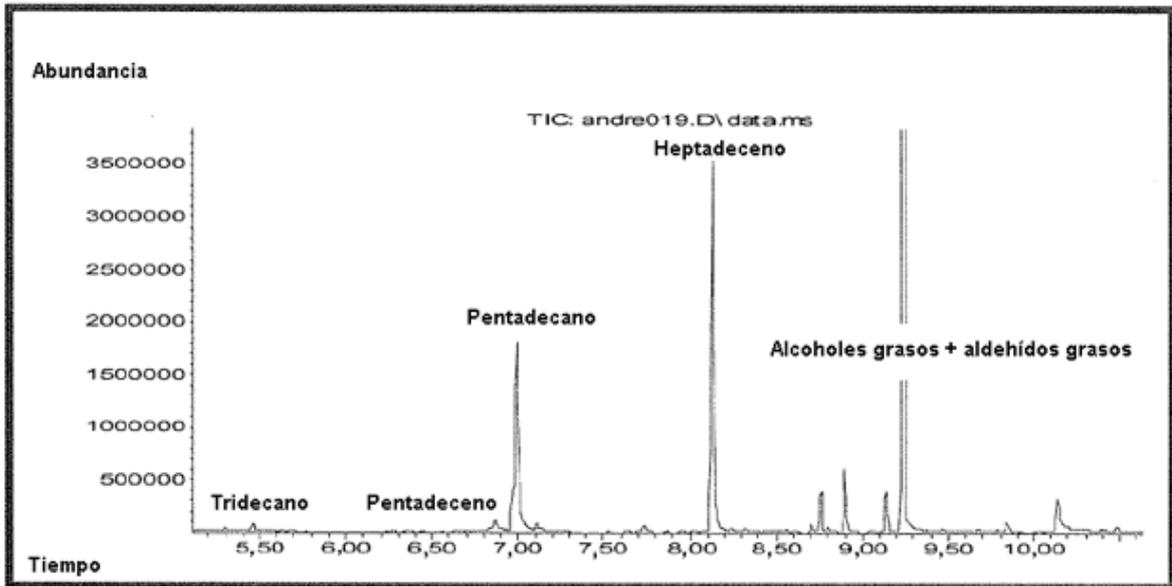


FIG. 9

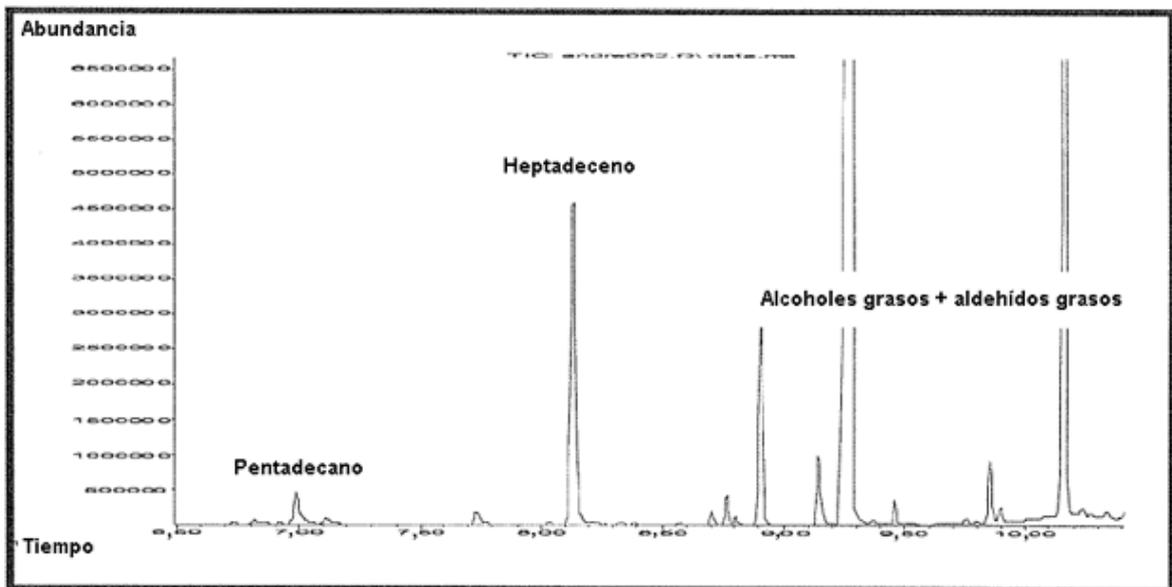


FIG. 10

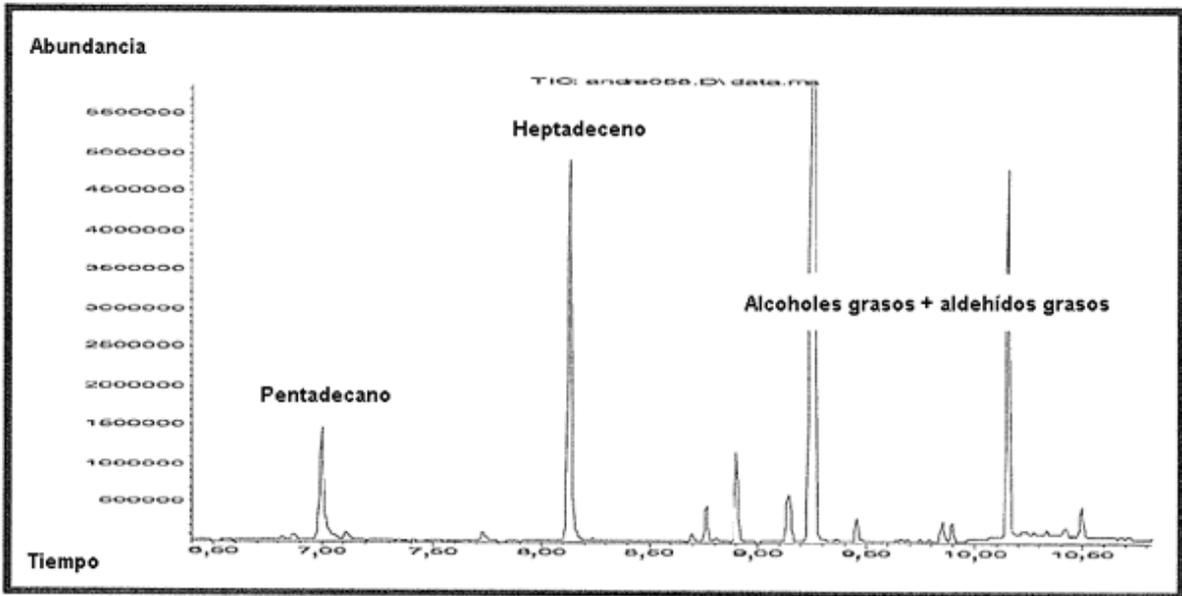


FIG. 11

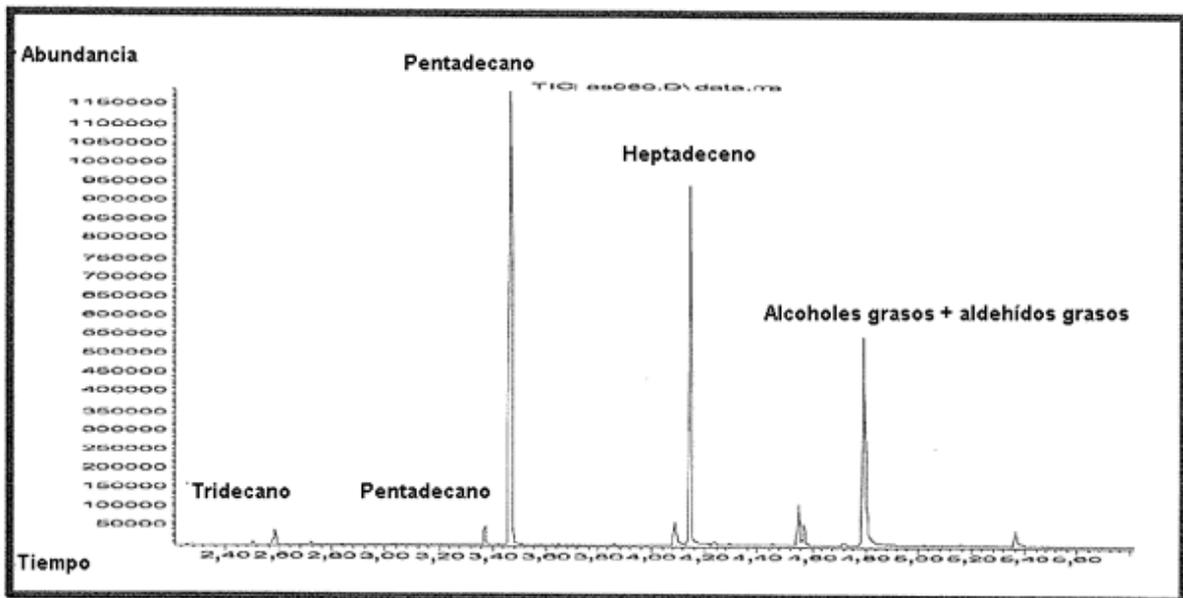


FIG. 12

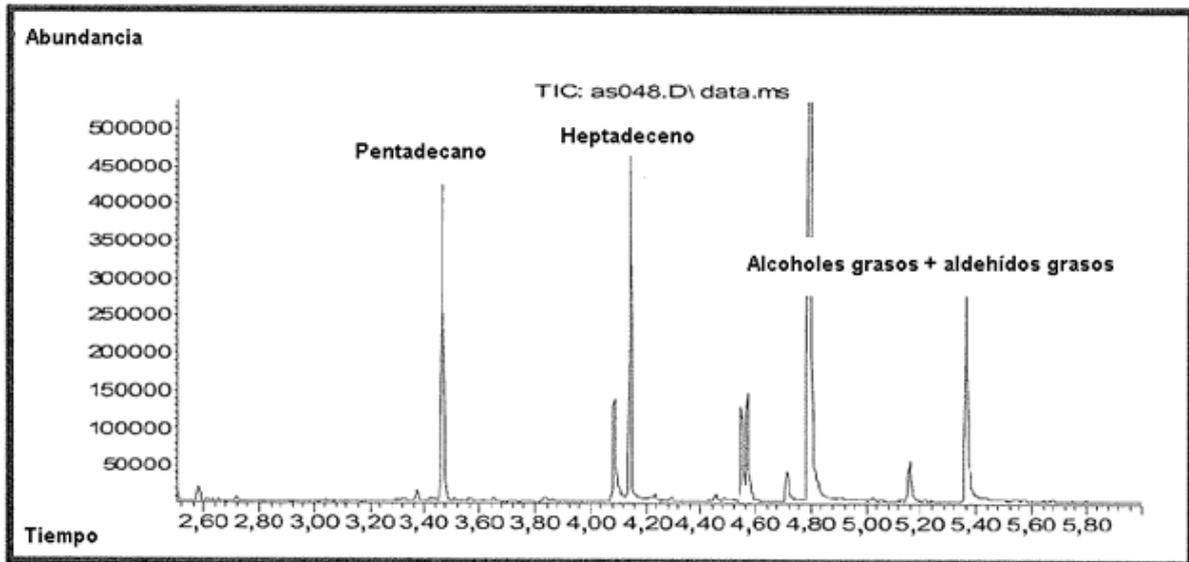


FIG. 13

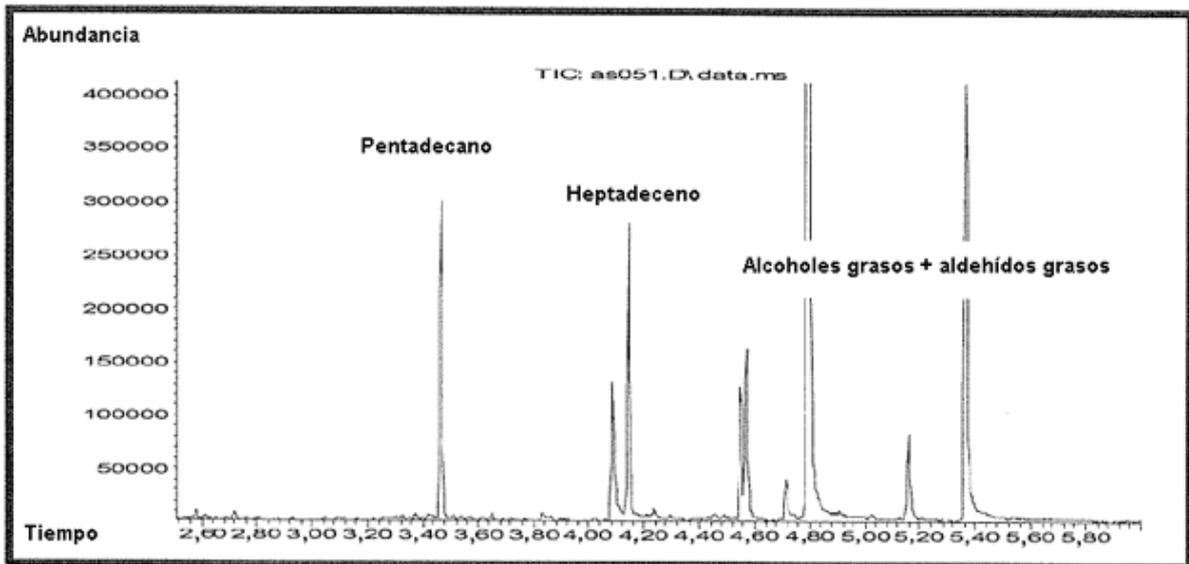


FIG. 14

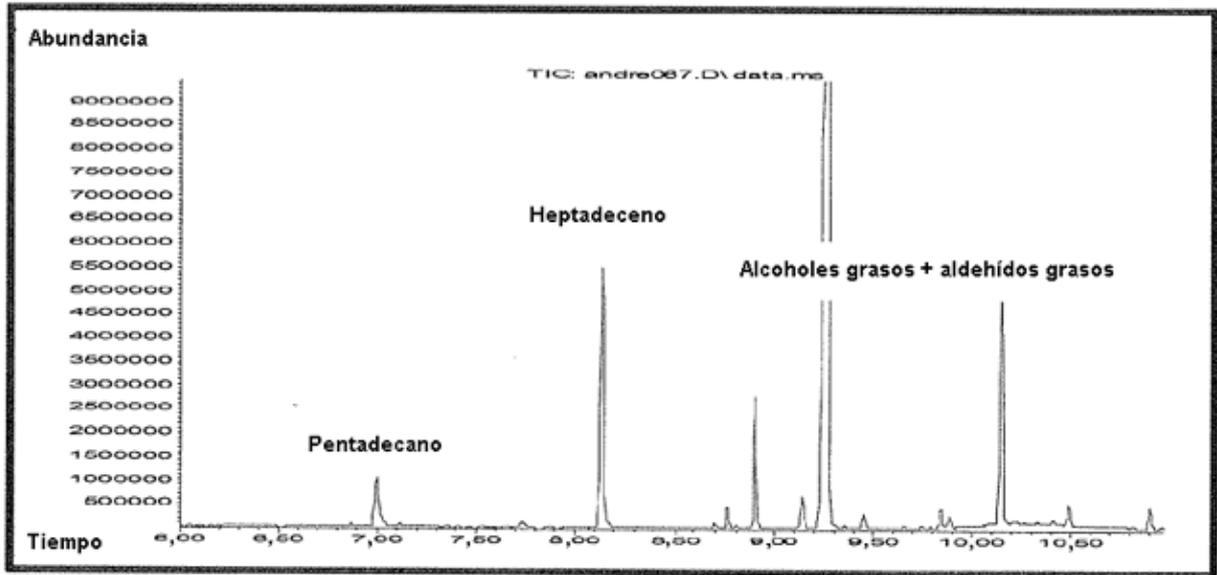


FIG. 15

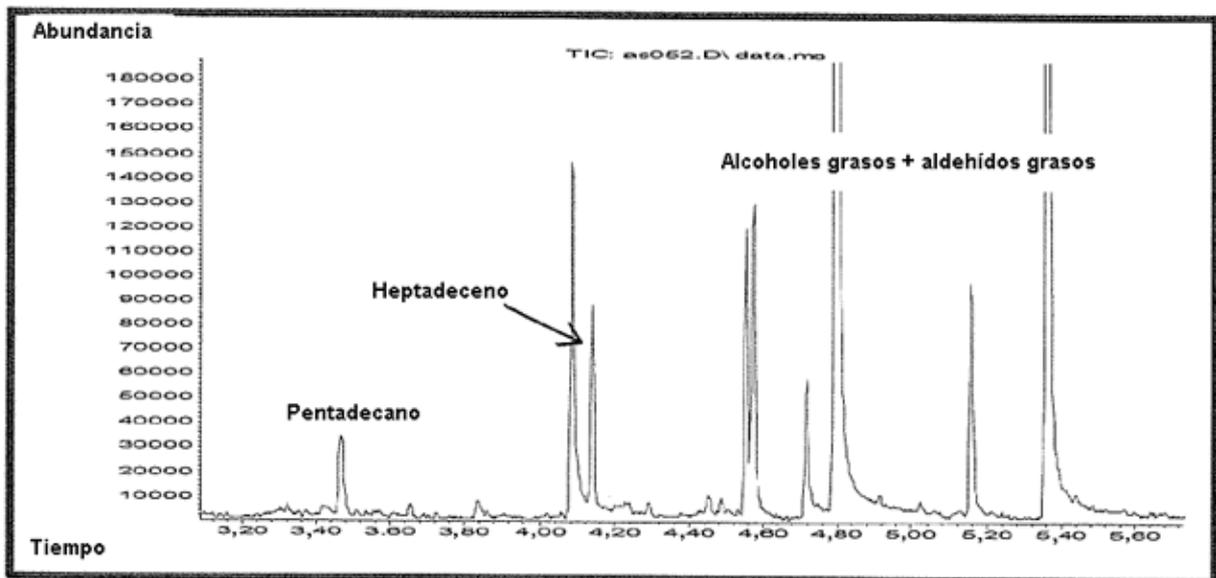


FIG. 16

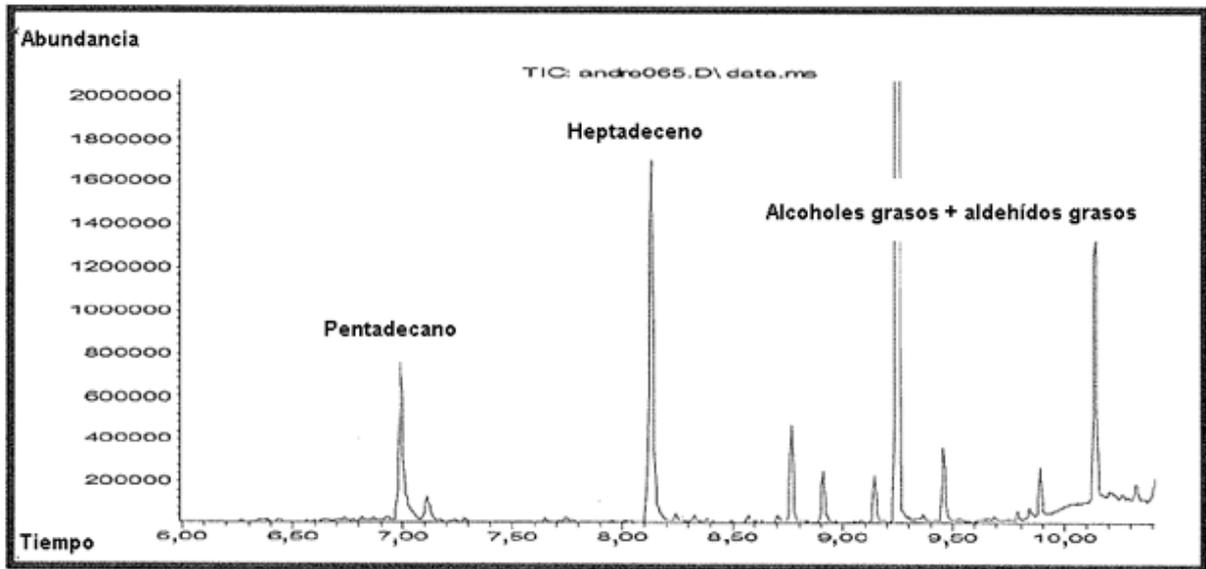


FIG. 17

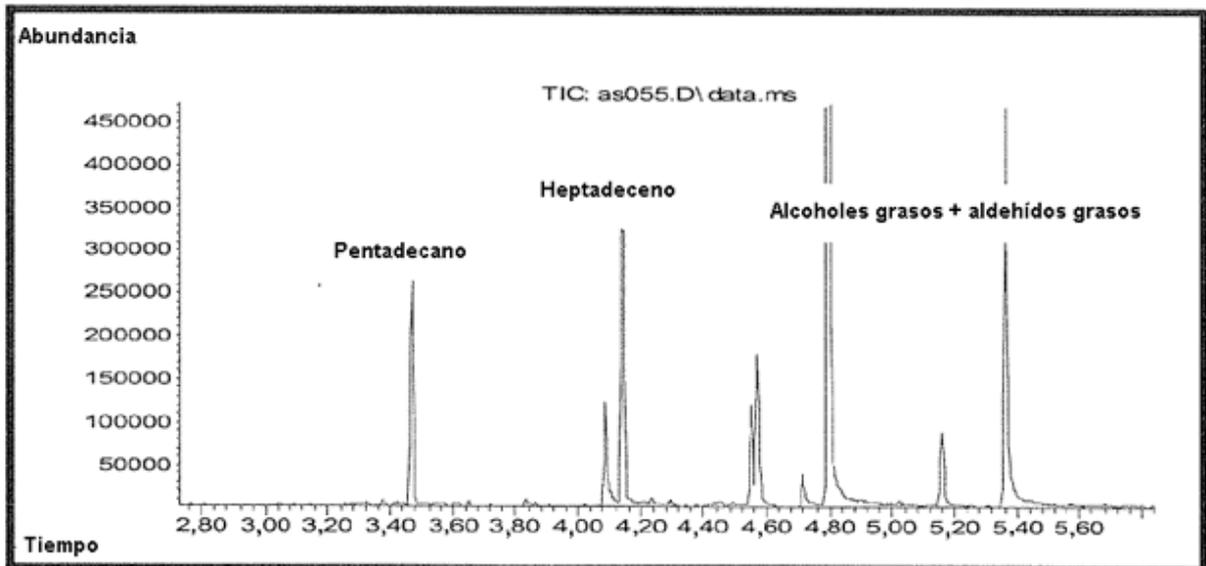


FIG. 18

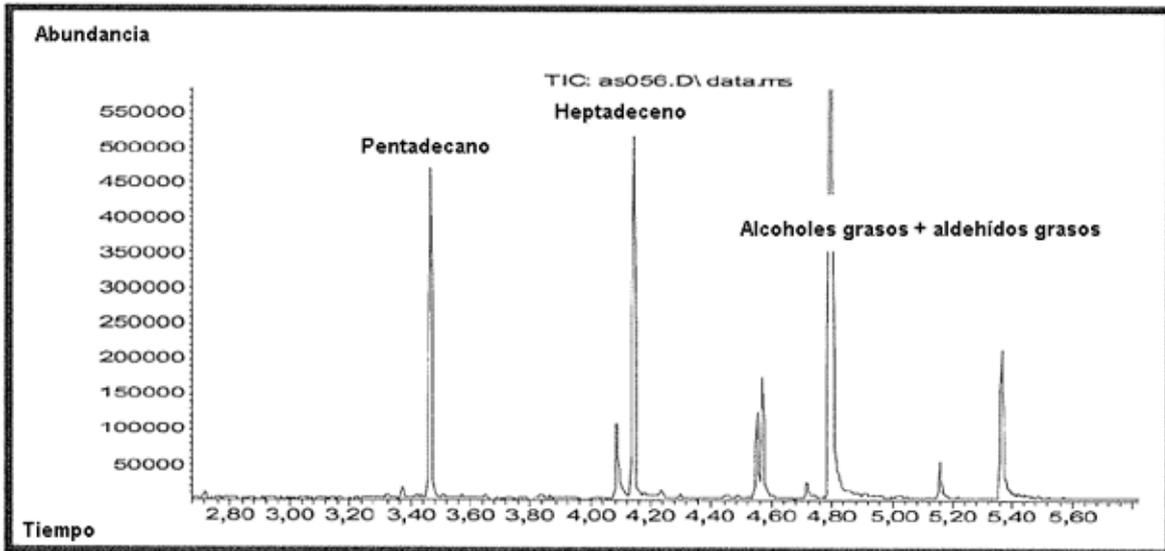


FIG. 19

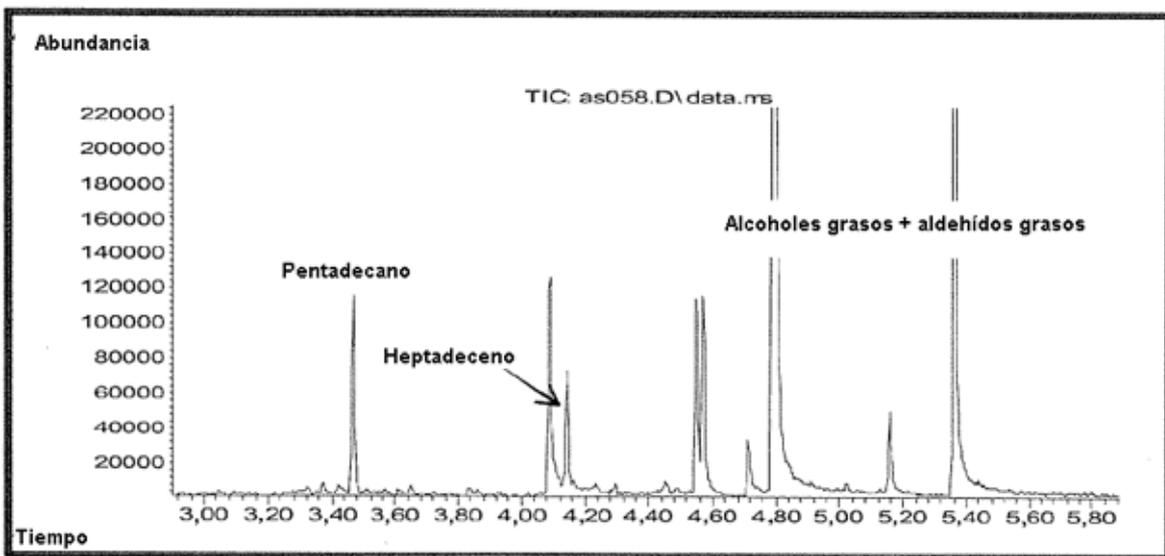


FIG. 20

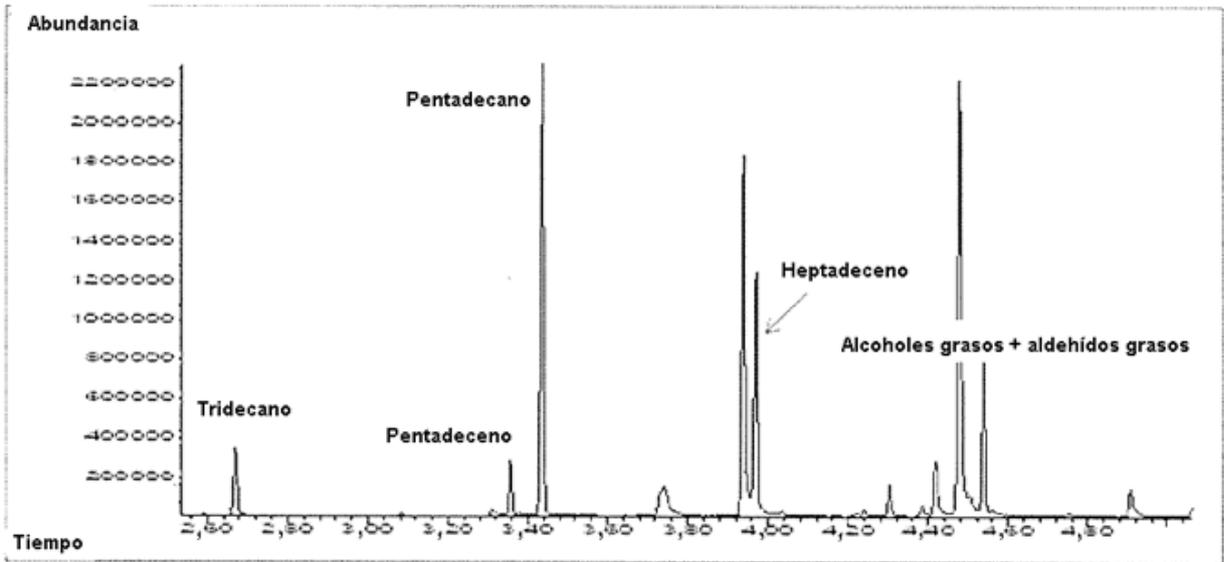


FIG. 21

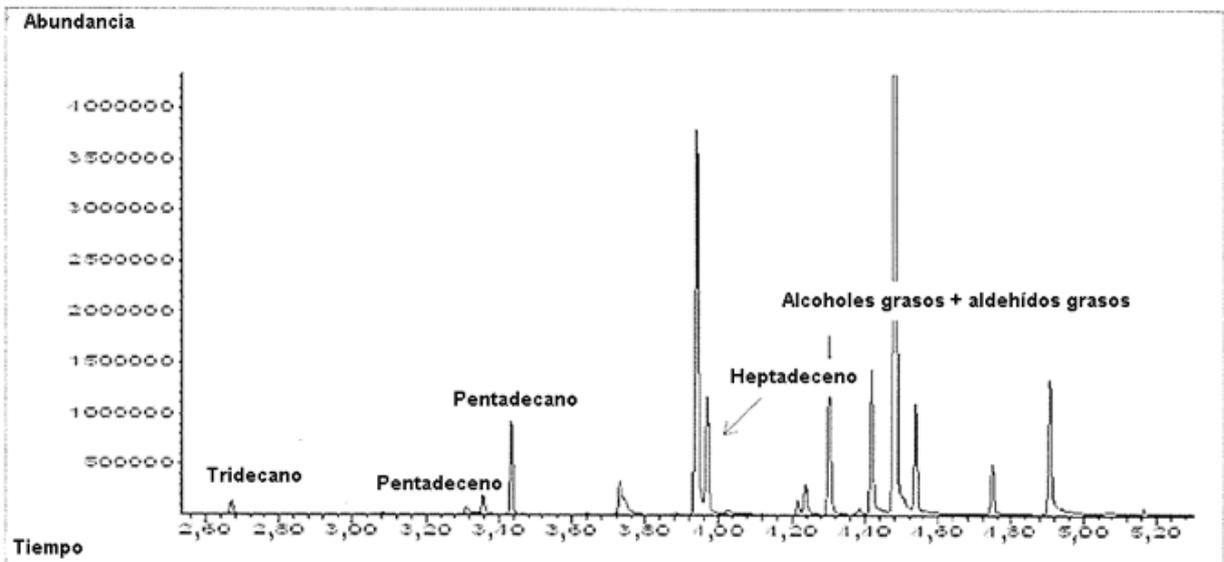


FIG. 22

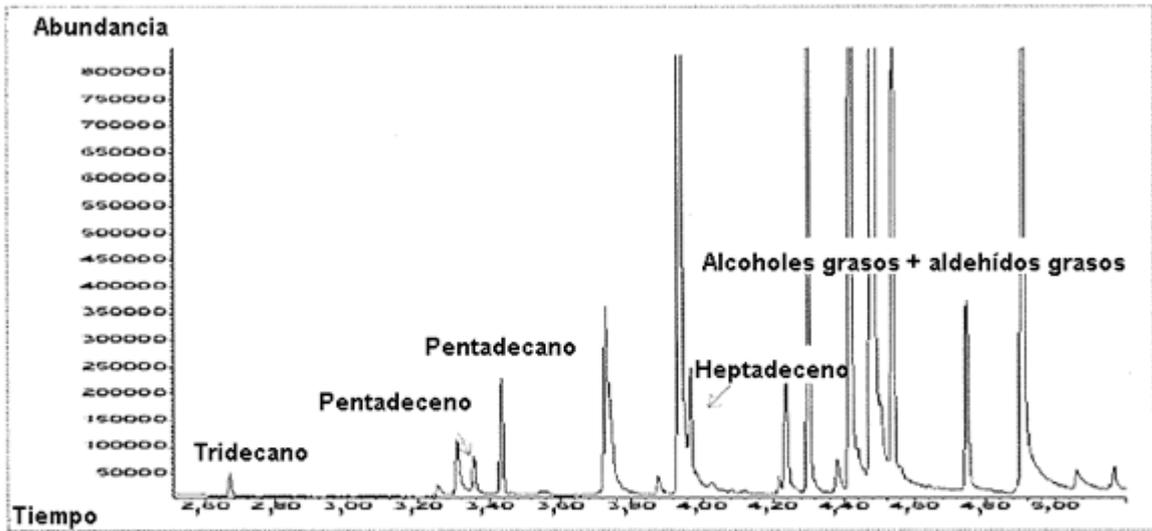


FIG. 23

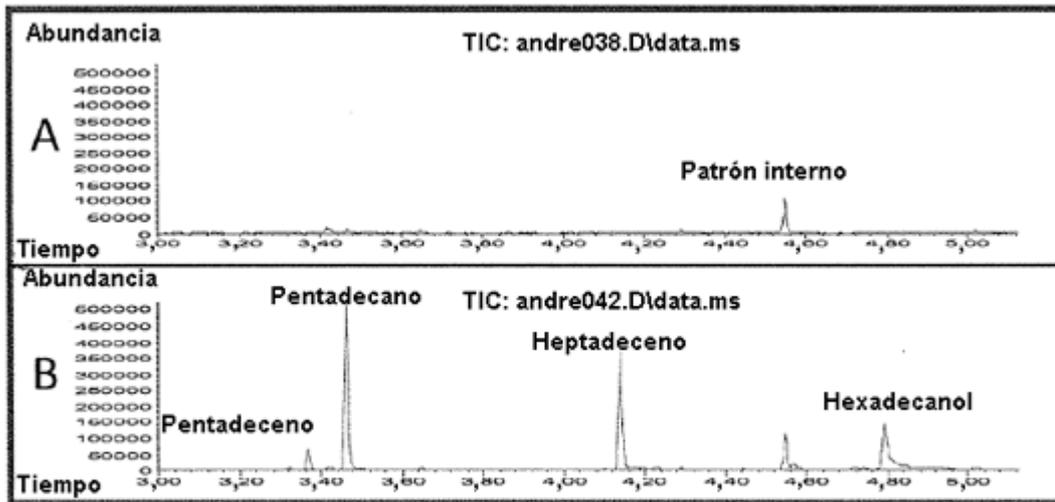


FIG. 24

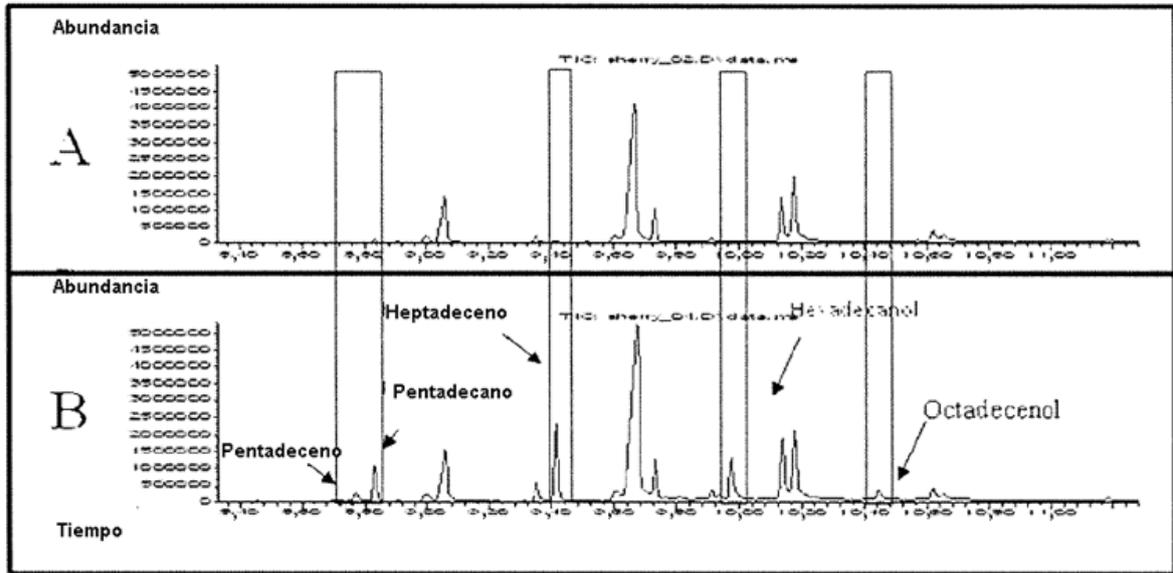


FIG. 25

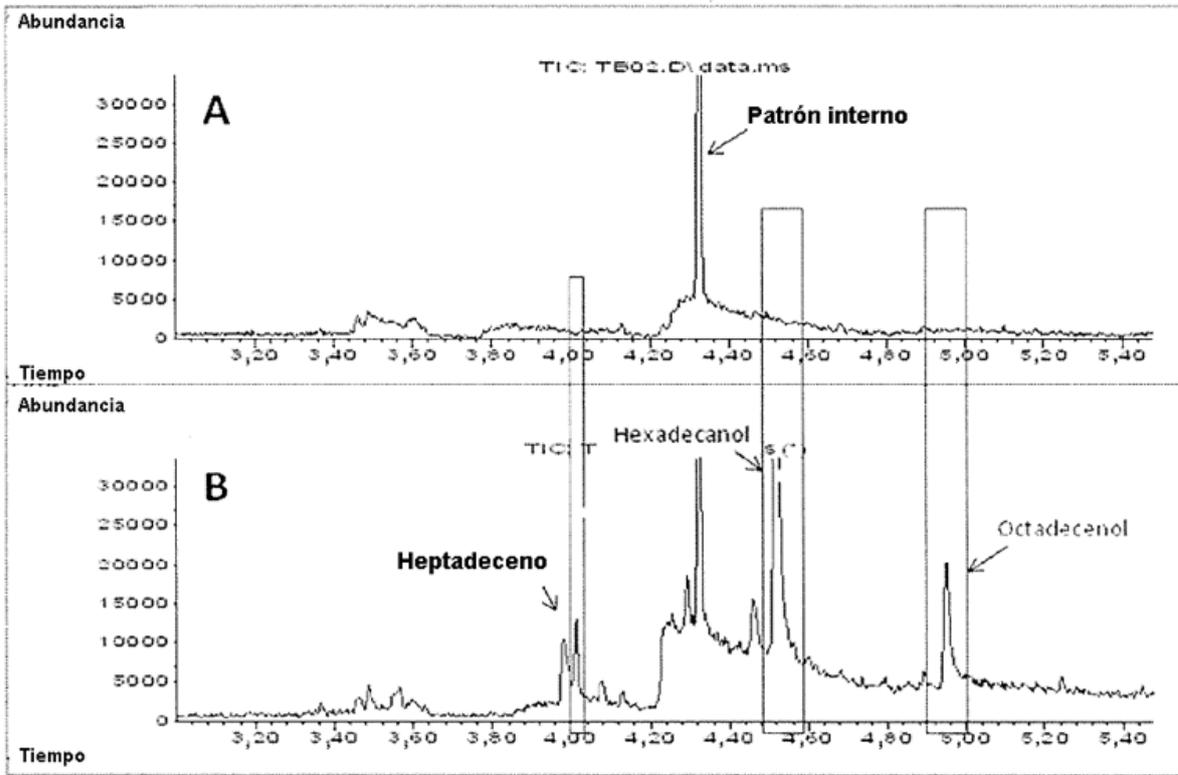


FIG. 26

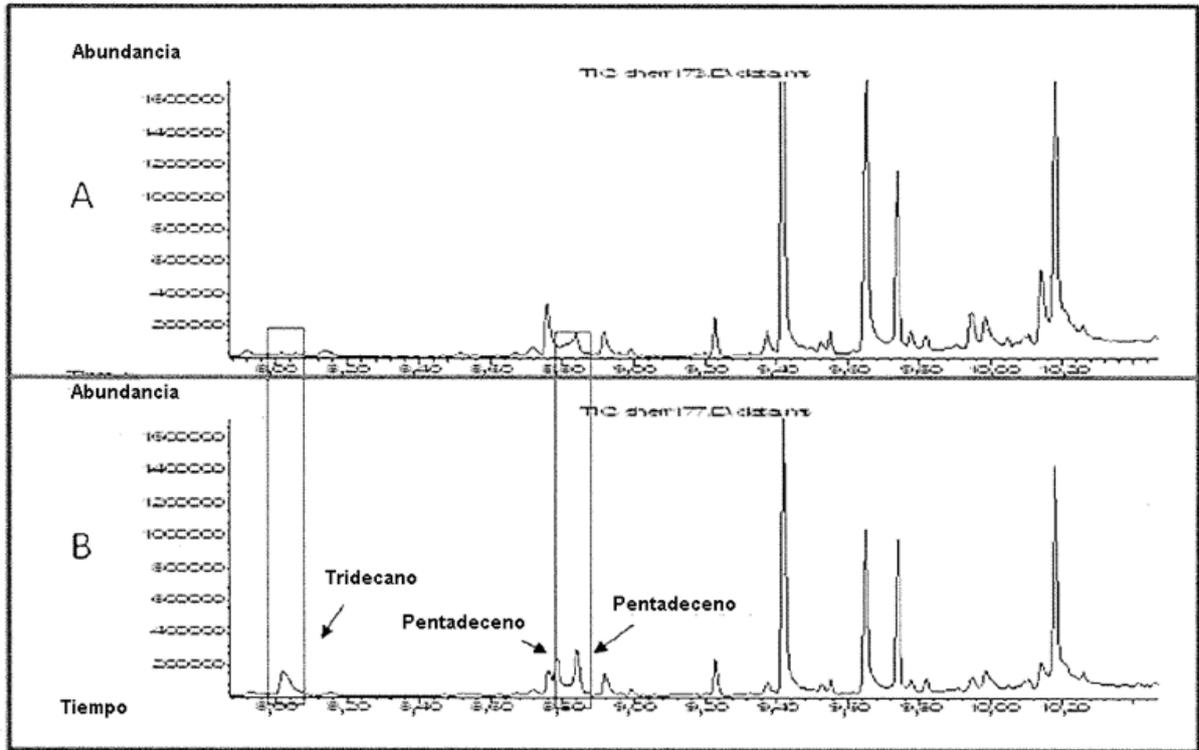


FIG. 27

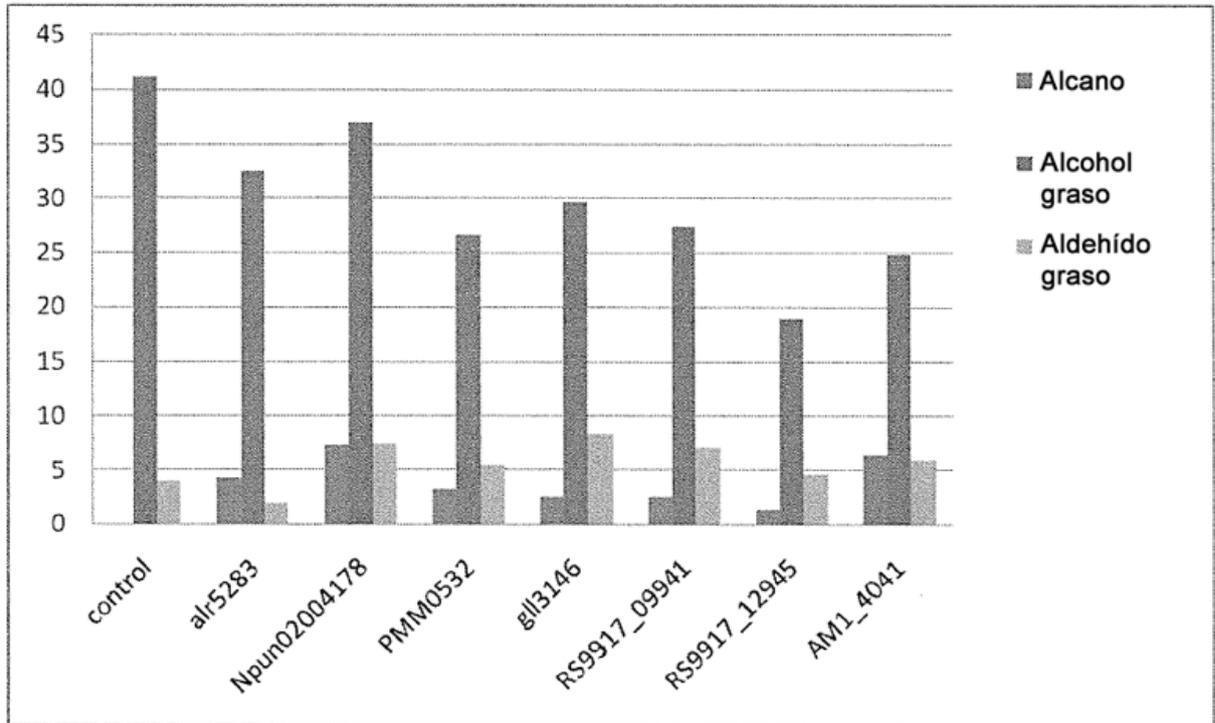


FIG. 28

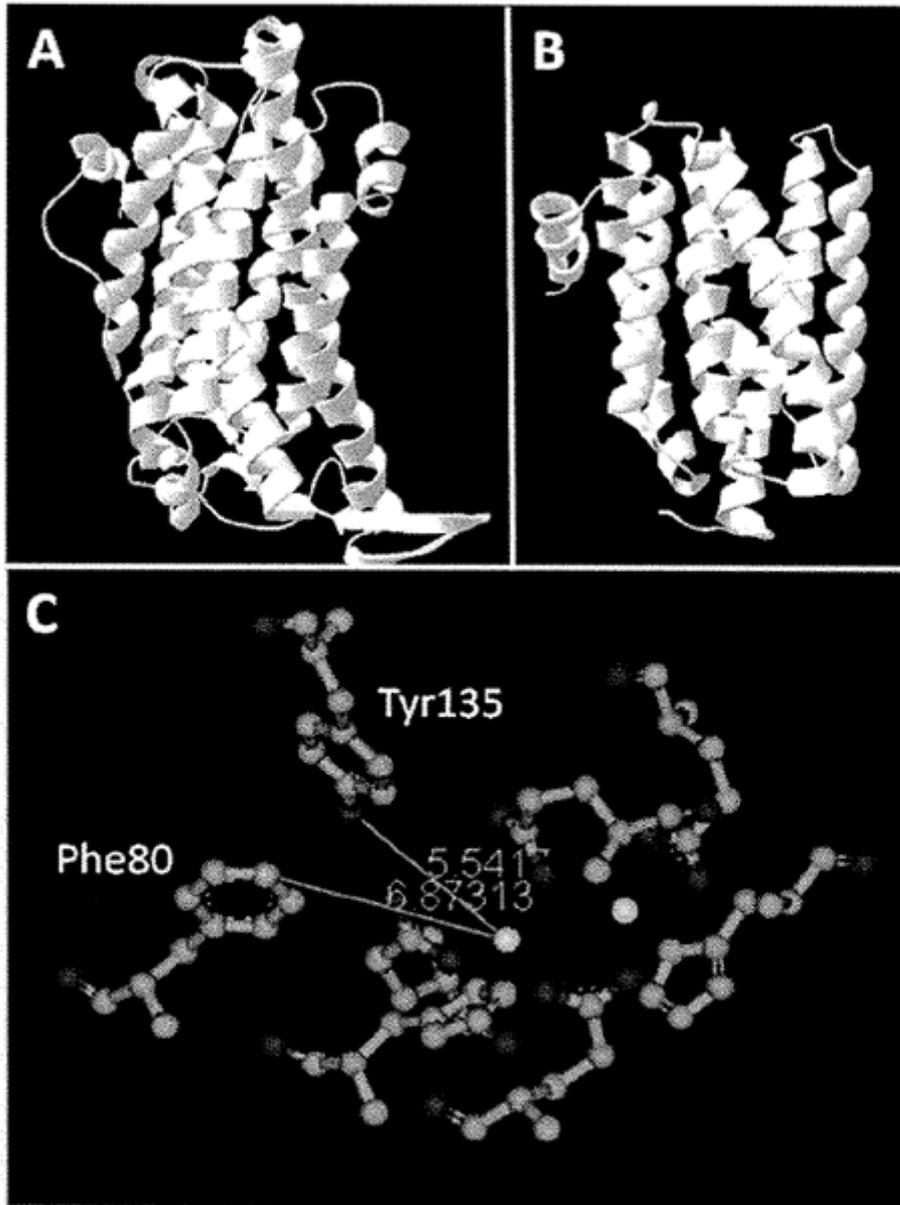


FIG. 29

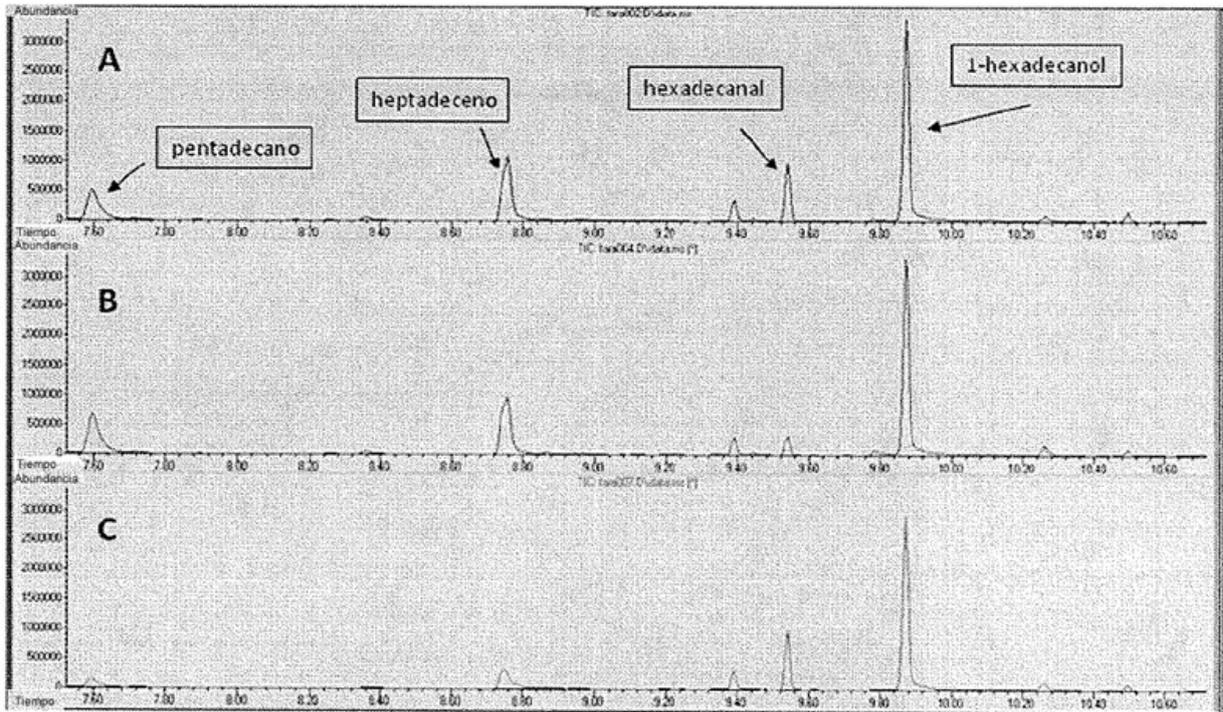


FIG. 30

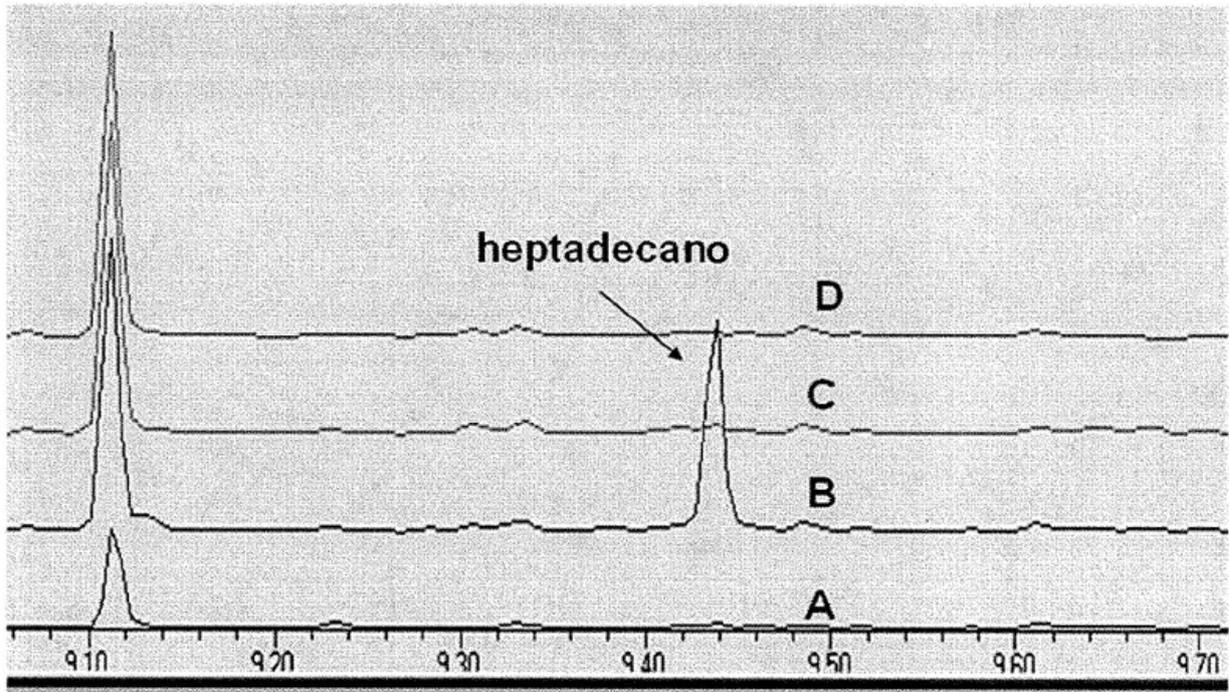


FIG. 31

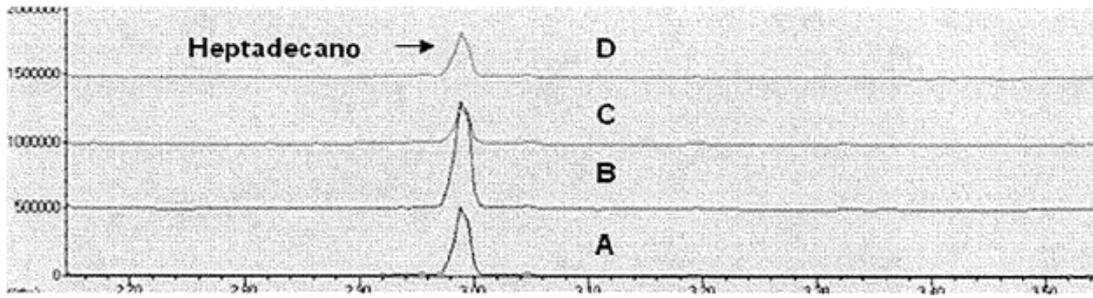


FIG. 32

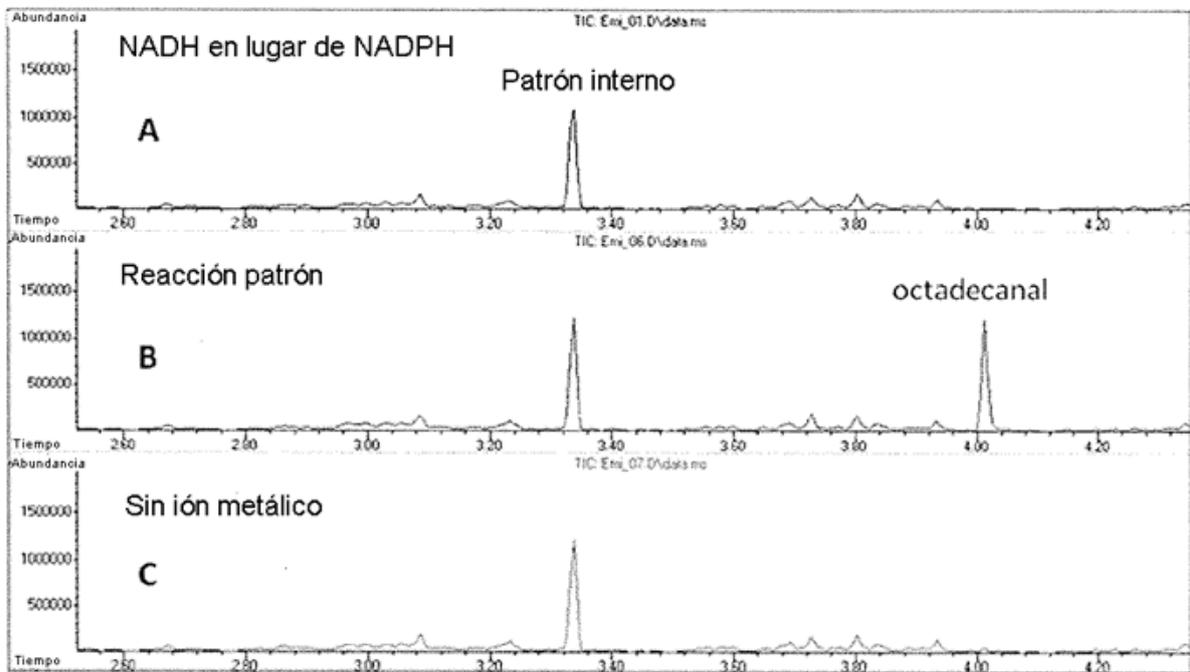


FIG. 33

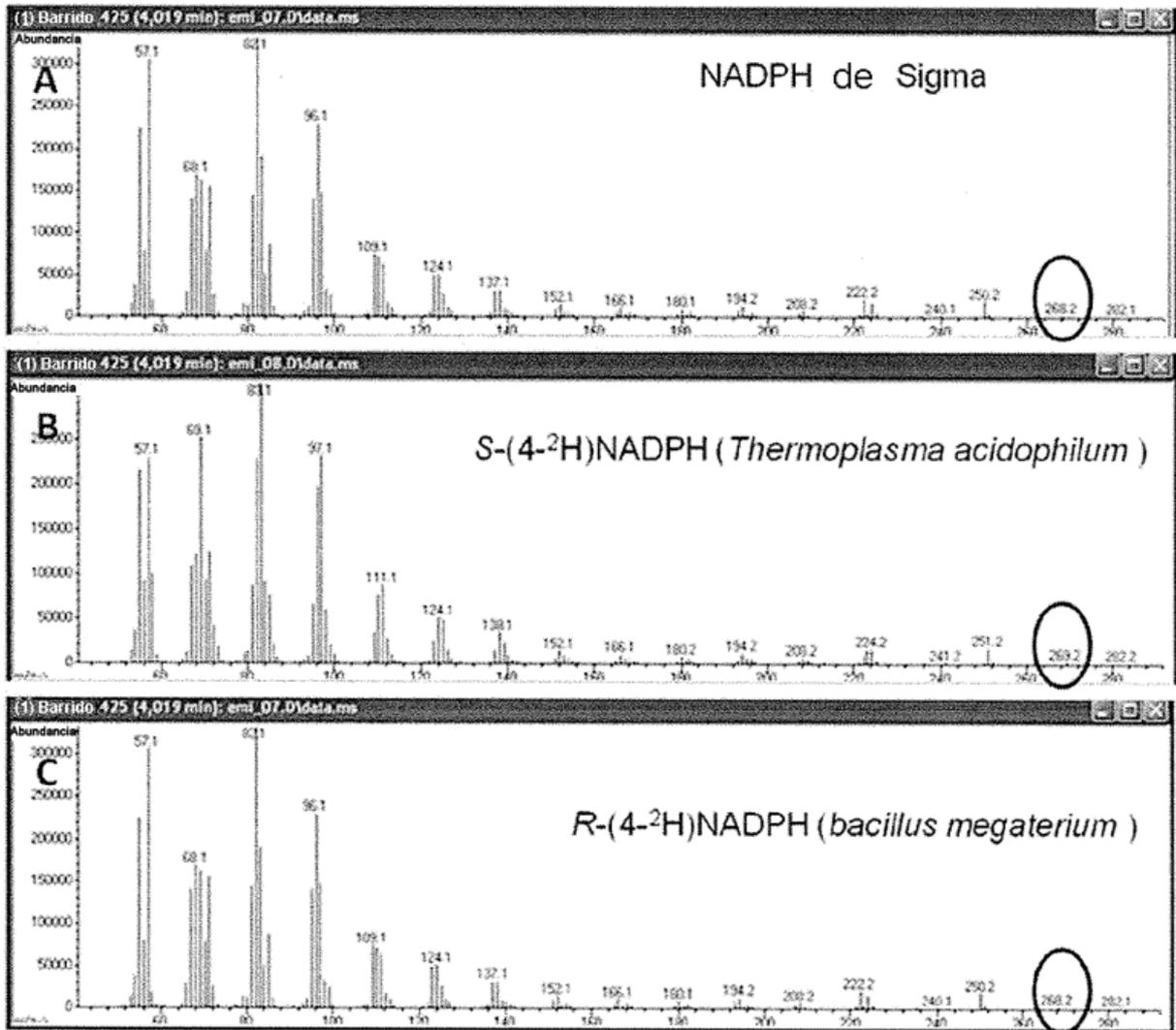


FIG. 34

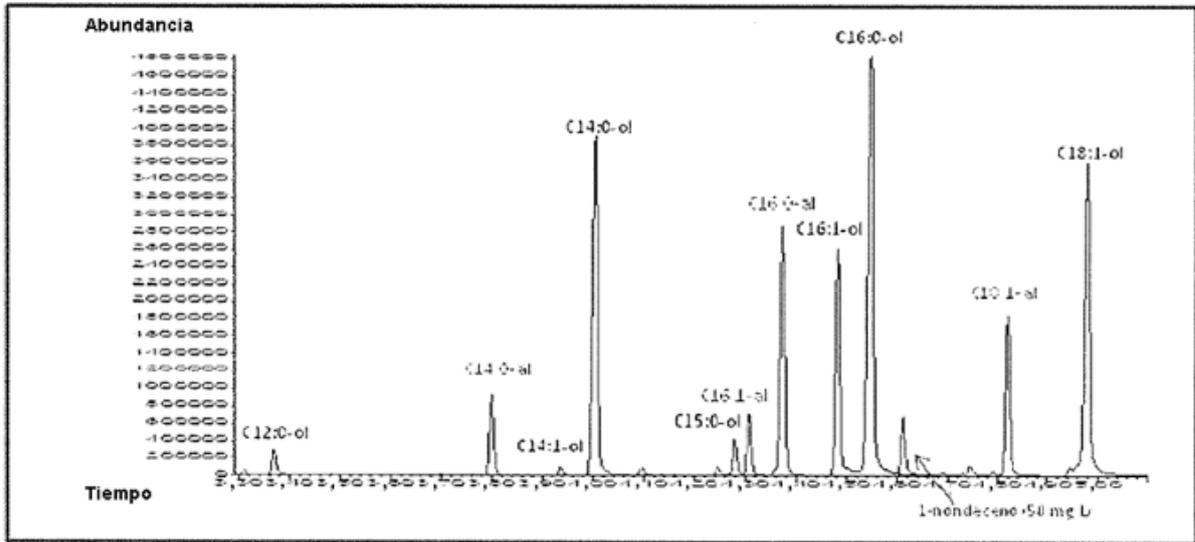


FIG. 35

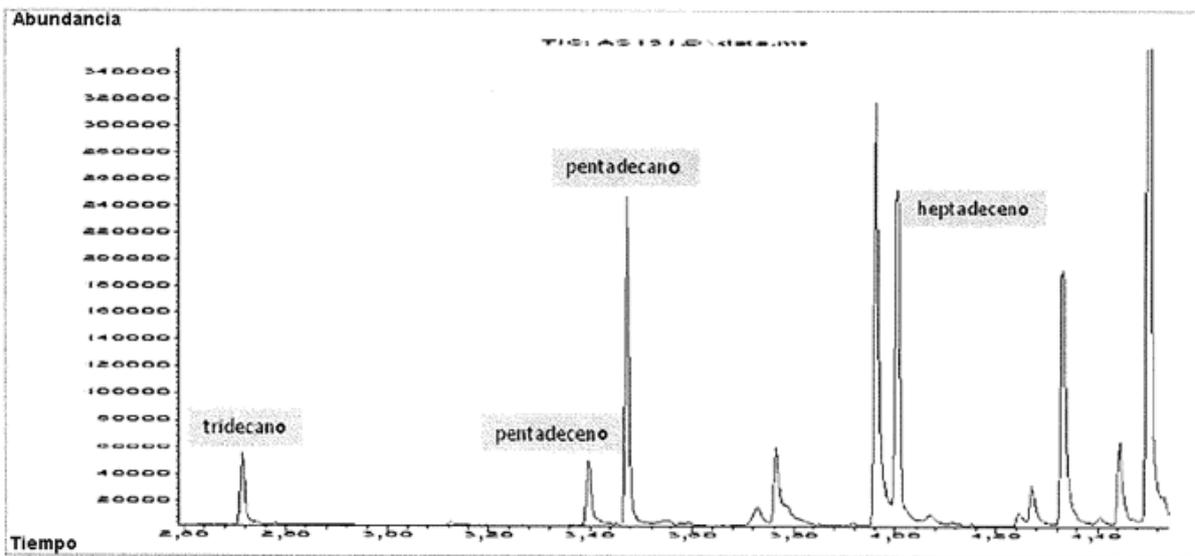


FIG. 36

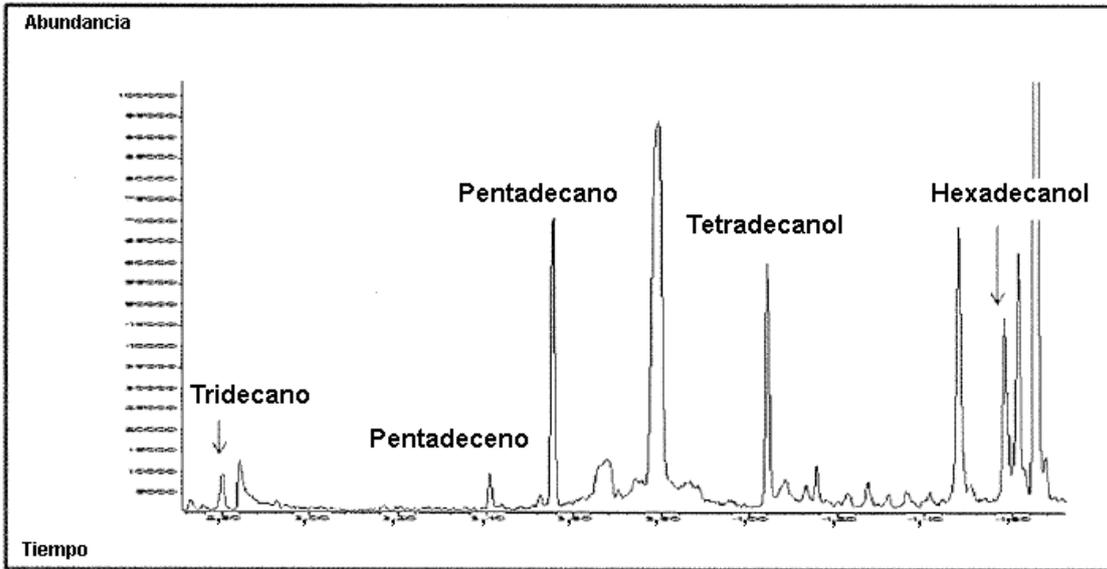


FIG. 37

FIG. 38

Números de registro a partir del 10 de abril de 2009

Número de registro	% de identidad	% de similitud	Longitud de alineación
gi 135970898 gb EBL05614.1	59,5	74,3	237
gi 134964254 gb EBE59803.1	57,6	74,2	229
gi 142528845 gb ECY73505.1	60,4	77,9	222
gi 135713677 gb EBJ38387.1	61,1	78,7	221
gi 141225813 gb ECQ49060.1	59,7	77,8	221
gi 144115151 gb EDI97334.1	67,7	80,9	220
gi 142133005 gb ECV83152.1	67,3	80,9	220
gi 137965371 gb EBX01252.1	67,3	80,9	220
gi 134786157 gb EBD42319.1	67,3	80,9	220
gi 136216894 gb EBM66672.1	62,3	77,7	220
gi 143271262 gb EDE04654.1	63,0	78,5	219
gi 135973786 gb EBL07573.1	63,0	78,1	219
gi 140222739 gb ECK35865.1	63,0	77,6	219
gi 139710482 gb ECG93903.1	62,6	78,1	219
gi 140109767 gb ECJ60294.1	62,6	78,5	219
gi 137939755 gb EBW86789.1	62,6	78,1	219
gi 140086977 gb ECJ44922.1	62,6	78,1	219
gi 143729007 gb EDG48416.1	62,6	78,1	219
gi 143217179 gb EDD66368.1	62,6	78,1	219
gi 138728422 gb ECB60007.1	62,6	78,1	219
gi 143540790 gb EDF53461.1	62,6	78,1	219
gi 143580323 gb EDF73830.1	62,6	78,1	219
gi 137317024 gb EBT41871.1	62,6	78,1	219
gi 143567212 gb EDF67415.1	62,6	78,1	219
gi 140726723 gb ECN09681.1	62,6	78,1	219
gi 136249402 gb EBM88688.1	62,6	78,1	219
gi 141156650 gb ECQ02018.1	62,6	78,1	219
gi 143550472 gb EDF58473.1	62,6	78,1	219
gi 139581046 gb ECG04925.1	62,6	78,1	219
gi 141904835 gb ECU06854.1	62,6	78,1	219
gi 143596624 gb EDF78559.1	62,1	78,1	219
gi 142808717 gb EDA73967.1	62,1	78,1	219
gi 137639134 gb EBV19468.1	62,1	78,1	219
gi 140970945 gb ECO75236.1	62,1	77,6	219

ES 2 696 773 T3

gi 137724335 gb EBV66164.1	62,1	78,1	219
gi 143095952 gb EDC78454.1	66,1	80,3	218
gi 135919851 gb EBK71172.1	66,1	80,3	218
gi 143083445 gb EDC69308.1	66,1	80,3	218
gi 137949431 gb EBW92260.1	63,1	77,6	214
gi 136336883 gb EBN48108.1	62,2	78,0	214
gi 136008042 gb EBL28916.1	62,2	78,0	214
gi 134796061 gb EBD49256.1	60,8	76,6	214
gi 143142117 gb EDD12240.1	60,8	76,6	214
gi 141842906 gb ECT63492.1	60,3	77,6	214
gi 137436109 gb EBU09046.1	60,3	77,1	214
gi 142796298 gb EDA64685.1	71,0	83,3	210
gi 138143480 gb EBX98705.1	70,5	83,3	210
gi 136004890 gb EBL27104.1	68,6	82,4	210
gi 142206956 gb ECW39307.1	67,1	82,4	210
gi 135813588 gb EBK00762.1	62,4	78,6	210
gi 137008728 gb EBR69250.1	62,7	79,0	209
gi 141443295 gb ECR86509.1	63,0	78,9	208
gi 135999440 gb EBL23890.1	63,1	79,1	206
gi 136001501 gb EBL25083.1	62,6	77,7	206
gi 137008499 gb EBR69124.1	62,3	77,5	204
gi 135813080 gb EBK00444.1	60,8	77,5	204
gi 139947991 gb ECI56786.1	63,6	79,8	198
gi 136255251 gb EBM92608.1	61,1	77,3	198
gi 141717528 gb ECS91623.1	58,6	76,8	198
gi 141153056 gb ECP99448.1	59,3	75,3	194
gi 140654921 gb ECM59517.1	67,9	81,9	193
gi 140993407 gb ECO91072.1	63,9	79,6	191
gi 143171676 gb EDD33295.1	63,9	79,6	191
gi 143659341 gb EDG12240.1	59,5	75,8	190
gi 138539256 gb ECA29247.1	68,1	81,9	188
gi 141964470 gb ECU48335.1	64,5	80,3	183
gi 139227663 gb ECE28885.1	63,9	79,2	183
gi 135926500 gb EBK75672.1	58,2	76,4	182
gi 140708786 gb ECM97174.1	68,4	81,4	177
gi 139523141 gb ECF65392.1	62,6	79,3	174
gi 137874181 gb EBW49523.1	61,5	76,4	174
gi 143221750 gb EDD69688.1	60,9	75,9	174
gi 140086962 gb ECJ44914.1	63,4	80,2	172
gi 142781070 gb EDA53384.1	61,8	77,1	170
gi 139775004 gb ECH37282.1	61,0	76,3	169

gi 136260534 gb EBM96169.1	61.0	75.7	169
gi 137441185 gb EBU11854.1	58.1	75.5	167
gi 136330550 gb EBN43814.1	65.1	80.7	166
gi 139650149 gb ECG51660.1	67.3	80.6	165
gi 143638003 gb EDF99869.1	63.6	80.0	165
gi 137949739 gb EBW92432.1	61.7	75.9	162
gi 143382653 gb EDE68551.1	63.9	81.0	158
gi 138989189 gb ECC70595.1	63.9	80.4	158
gi 138408887 gb EBZ46853.1	64.1	78.2	156
gi 137230040 gb EBS93199.1	61.3	76.1	155
gi 141605381 gb ECS53894.1	65.6	81.8	154
gi 137858747 gb EBW40699.1	64.9	81.8	154
gi 140209383 gb ECK27191.1	66.0	79.1	153
gi 142753984 gb EDA33411.1	60.7	75.3	150
gi 137242084 gb EBS99775.1	64.9	82.4	148
gi 136229422 gb EBM75188.1	64.6	82.3	147
gi 140311369 gb ECK89744.1	73.3	85.6	146
gi 140866197 gb ECO03647.1	65.1	82.2	146
gi 139229558 gb ECE29833.1	61.4	76.6	145
gi 141659030 gb ECS68172.1	61.4	76.6	145
gi 139580852 gb ECG04786.1	65.3	81.9	144
gi 138338712 gb EBZ05758.1	65.3	81.9	144
gi 136204827 gb EBM58548.1	61.3	76.8	142
gi 139095530 gb ECD38154.1	64.0	81.3	139
gi 136351648 gb EBN58190.1	60.1	76.1	138
gi 138155154 gb EBY06350.1	67.9	80.3	137
gi 137644530 gb EBV22059.1	63.1	80.8	130
gi 143775710 gb EDG72409.1	61.5	79.2	130
gi 143500330 gb EDF32920.1	63.1	77.9	122
gi 139709584 gb ECG93249.1	71.1	85.1	121
gi 142537519 gb ECY79816.1	62.8	81.0	121
gi 137944410 gb EBW89433.1	64.2	78.3	120
gi 137387955 gb EBT81682.1	63.9	78.2	119
gi 139955976 gb ECI62054.1	71.2	84.8	118
gi 137251843 gb EBT05348.1	62.1	77.6	116
gi 138442523 gb EBZ70326.1	62.1	76.7	116
gi 141590592 gb ECS49420.1	60.7	78.6	112
gi 143187997 gb EDD45026.1	60.7	77.7	112
gi 143655969 gb EDG10472.1	61.5	78.0	109
gi 139459255 gb ECF24788.1	62.9	79.1	105
gi 141976584 gb ECU56751.1	62.9	79.1	105

ES 2 696 773 T3

gi 139233104 gb ECE31190.1	69,6	83,3	102
gi 139233107 gb ECE31193.1	70,3	82,2	101
gi 138582711 gb ECA59616.1	61,4	77,2	101
gi 138442855 gb EBZ70514.1	61,4	75,3	101
gi 137662676 gb EBV31757.1	60,4	76,2	101
gi 139846062 gb ECH87248.1	62,6	74,8	99
gi 136935327 gb EBR27657.1	60,6	74,5	94
gi 137466252 gb EBU25703.1	67,0	80,2	91
gi 137820604 gb EBW18665.1	63,7	80,2	91
gi 138539295 gb ECA29276.1	62,6	79,1	91
gi 136239262 gb EBM81844.1	64,7	81,2	85
gi 136294902 gb EBN19768.1	64,7	81,2	85
gi 137413136 gb EBT96003.1	61,2	76,5	85
gi 137641041 gb EBV20354.1	60,0	72,9	85
gi 142508710 gb ECY58869.1	64,3	82,1	84
gi 140096399 gb ECJ51008.1	62,2	81,7	82
gi 137938664 gb EBW86178.1	64,0	81,3	75
gi 137275448 gb EBT18729.1	58,1	70,3	74
gi 141955842 gb ECU42610.1	63,0	80,8	73
gi 139221707 gb ECE24659.1	66,2	83,1	71
gi 142508709 gb ECY58868.1	52,9	68,6	70
gi 137523719 gb EBU55323.1	65,2	82,6	69
gi 140781524 gb ECN46583.1	65,6	82,8	64
gi 137627577 gb EBV13553.1	60,0	73,3	60
gi 141951833 gb ECU39722.1	59,7	73,7	57
gi 137232510 gb EBS94613.1	59,7	73,7	57

Punto de corte usado: >50% de identidad con y >25% de longitud de synpcc7942\_1593

FIG. 39

Números de registro a partir del 10 de abril de 2009

Número de registro	% de identidad	% de similitud	Longitud de alineación
gi 143288250 gb EDE13503.1	71,3	80,5	87
gi 142342310 gb ECX39602.1	71,3	80,5	87
gi 137949588 gb EBW92346.1	70,7	85,9	92
gi 139984340 gb ECI81897.1	70,6	78,8	85
gi 140249046 gb ECK54318.1	69,7	82,0	267
gi 142111437 gb ECV67406.1	69,2	81,3	182
gi 142994709 gb EDC04737.1	68,8	81,2	138
gi 143066602 gb EDC56955.1	67,7	81,5	248
gi 138840827 gb ECC11022.1	67,1	78,8	146
gi 137829071 gb EBW23606.1	66,7	77,1	96
gi 142133008 gb ECV83155.1	66,5	79,5	337
gi 143095956 gb EDC78458.1	66,4	79,1	339
gi 144115152 gb EDI97335.1	66,4	79,4	339
gi 140732156 gb ECN13587.1	66,1	79,5	254
gi 136241230 gb EBM83170.1	66,0	78,6	103
gi 140001769 gb ECI93451.1	65,9	76,9	91
gi 139305662 gb ECE48752.1	65,9	80,0	205
gi 137634503 gb EBV17219.1	65,8	80,7	114
gi 138584841 gb ECA61142.1	65,5	81,9	116

ES 2 696 773 T3

gi 135919849 gb EBK71170.1	65.4	76.6	107
gi 141153057 gb ECP99449.1	65.4	79.6	280
gi 141976585 gb ECU56752.1	65.3	77.6	98
gi 142206955 gb ECW39306.1	65.3	78.5	340
gi 141804802 gb ECT36785.1	65.3	76.8	95
gi 138931154 gb ECC47219.1	65.1	76.7	86
gi 138408888 gb EBZ46854.1	65.1	78.0	255
gi 134743188 gb EBD14908.1	64.8	78.8	307
gi 138168794 gb EBY16028.1	64.8	77.6	304
gi 135749749 gb EBJ60721.1	64.8	79.1	105
gi 138338711 gb EBZ05757.1	64.6	80.0	175
gi 142827948 gb EDA88477.1	64.5	77.7	121
gi 138361576 gb EBZ15968.1	64.4	78.7	267
gi 140517919 gb ECM08416.1	64.4	77.4	115
gi 135813081 gb EBK00445.1	64.3	79.7	143
gi 137627576 gb EBV13552.1	64.3	80.1	171
gi 141161845 gb ECQ05757.1	64.2	77.1	109
gi 140992134 gb ECO90156.1	64.2	76.8	95
gi 137796334 gb EBW04596.1	64.2	78.9	279
gi 137619413 gb EBV08950.1	64.1	77.6	223
gi 140517917 gb ECM08414.1	64.1	79.6	181
gi 137232509 gb EBS94612.1	64.1	79.0	181
gi 135811491 gb EBJ99446.1	64.0	77.7	314

ES 2 696 773 T3

gi 141167040 gb ECQ09480 .1	64,0	78,8	250
gi 143441820 gb EDE97777. 1	64,0	78,2	261
gi 140311368 gb ECK89743. 1	64,0	79,3	261
gi 140970943 gb ECO75234 .1	64,0	76,6	111
gi 136817739 gb EBQ60666 .1	63,8	78,5	340
gi 141717529 gb ECS91624. 1	63,7	78,8	146
gi 137632337 gb EBV16047. 1	63,7	77,7	256
gi 137662677 gb EBV31758. 1	63,6	79,7	143
gi 140091056 gb ECJ47190. 1	63,6	78,2	280
gi 143217178 gb EDD66367 .1	63,5	78,8	137
gi 139984339 gb ECI81896. 1	63,5	76,3	156
gi 139382506 gb ECE73591. 1	63,4	78,9	194
gi 140096397 gb ECJ51006. 1	63,4	79,7	153
gi 134606350 gb EBC34611. 1	63,4	77,9	131
gi 140705175 gb ECM95033 .1	63,4	77,0	191
gi 139846064 gb ECH87250 .1	63,1	76,0	179
gi 137953535 gb EBW9457 2.1	63,1	78,0	241
gi 143738737 gb EDG53066 .1	63,1	76,6	111
gi 141951832 gb ECU39721 .1	63,1	76,6	111
gi 143271261 gb EDE04653. 1	63,1	76,6	111
gi 139846065 gb ECH87251 .1	63,0	78,8	146
gi 137251844 gb EBT05349. 1	63,0	78,8	146
gi 136249401 gb EBM88687 .1	63,0	78,8	146

ES 2 696 773 T3

gi 134628580 gb EBC48074. 1	63.0	77.0	100
gi 136312048 gb EBN31461 .1	62.9	76.2	143
gi 143221751 gb EDD69689 .1	62.9	78.9	194
gi 141955844 gb ECU42612 .1	62.9	75.7	140
gi 143395654 gb EDE73119. 1	62.8	77.5	218
gi 142781071 gb EDA53385 .1	62.8	79.3	164
gi 136303394 gb EBN25555 .1	62.7	77.7	319
gi 143557688 gb EDF62238. 1	62.6	77.9	131
gi 136008043 gb EBL28917. 1	62.6	77.6	294
gi 143596625 gb EDF78560. 1	62.6	77.0	318
gi 137641042 gb EBV20355. 1	62.6	79.1	163
gi 136231267 gb EBM76426 .1	62.5	76.7	339
gi 143175604 gb EDD36054 .1	62.5	76.7	339
gi 142508708 gb ECY58867. 1	62.5	76.7	339
gi 135926501 gb EBK75673. 1	62.5	76.6	320
gi 141955884 gb ECU42641 .1	62.4	77.8	189
gi 142821119 gb EDA83282 .1	62.4	78.7	202
gi 134609411 gb EBC36492. 1	62.4	75.8	194
gi 142885864 gb EDB27722 .1	62.3	75.4	207
gi 136204828 gb EBM58549 .1	62.3	77.0	318
gi 143580324 gb EDF73831. 1	62.2	76.7	339
gi 143766375 gb EDG67769 .1	62.2	77.0	339
gi 143500332 gb EDF32922. 1	62.2	76.4	339

ES 2 696 773 T3

gi 139233105 gb ECE31191.1	62,1	76,5	132
gi 143738779 gb EDG53089.1	62,1	76,8	314
gi 134964255 gb EBE59804.1	62,1	77,1	153
gi 140863545 gb ECO01751.1	62,1	77,0	269
gi 137944409 gb EBW8943.2.1	62,0	78,5	158
gi 143411619 gb EDE81261.1	62,0	76,7	339
gi 142753988 gb EDA33415.1	62,0	77,0	339
gi 139580853 gb ECG04787.1	61,9	75,7	202
gi 141227933 gb ECQ50606.1	61,9	75,1	173
gi 143659340 gb EDG12239.1	61,8	78,3	157
gi 136935328 gb EBR27658.1	61,8	77,7	157
gi 137275449 gb EBT18730.1	61,8	77,5	204
gi 138585243 gb ECA61437.1	61,8	76,5	136
gi 139195947 gb ECE06889.1	61,7	76,2	269
gi 139424973 gb ECF02640.1	61,7	76,3	274
gi 141380828 gb ECR42772.1	61,7	75,8	227
gi 136351647 gb EBN58189.1	61,7	77,0	339
gi 136304410 gb EBN26254.1	61,6	75,7	185
gi 139948037 gb ECI56814.1	61,6	76,8	224
gi 135970899 gb EBL05615.1	61,6	76,8	211
gi 138627165 gb ECA90647.1	61,5	75,5	143
gi 137395720 gb EBT86160.1	61,5	74,8	143
gi 140086960 gb ECJ44912.1	61,5	77,1	249

ES 2 696 773 T3

gi 141024916 gb ECP11582. 1	61,4	76,7	223
gi 139095531 gb ECD38155. 1	61,4	76,7	210
gi 141659029 gb ECS68171. 1	61,4	77,2	127
gi 139969430 gb ECI71470. 1	61,4	76,7	215
gi 136986729 gb EBR56775. 1	61,4	75,7	202
gi 143634197 gb EDF97600. 1	61,3	76,7	313
gi 135973785 gb EBL07572. 1	61,3	76,3	279
gi 143200944 gb EDD54508 .1	61,3	74,6	173
gi 137787263 gb EBV99371. 1	61,2	76,3	232
gi 139204136 gb ECE12313. 1	61,2	76,7	219
gi 136001500 gb EBL25082. 1	61,2	76,3	219
gi 141874476 gb ECT85572. 1	61,2	76,4	237
gi 137905325 gb EBW6737 5.1	61,1	74,9	175
gi 140089341 gb ECJ46519. 1	61,1	75,4	203
gi 140855194 gb ECN95754 .1	61,1	75,9	203
gi 134965622 gb EBE60718. 1	61,1	75,8	339
gi 141527125 gb ECS15588. 1	61,1	76,0	208
gi 136218988 gb EBM68086 .1	61,0	73,4	154
gi 142364499 gb ECX54765. 1	60,9	76,1	330
gi 141603393 gb ECS53340. 1	60,8	74,8	143
gi 136216893 gb EBM66671 .1	60,8	76,3	245
gi 143743653 gb EDG56305 .1	60,8	73,7	148
gi 140222741 gb ECK35867. 1	60,7	73,3	163

gi 142389887 gb ECX71634.1	60,6	74,8	198
gi 139315697 gb ECE51398.1	60,6	76,4	241
gi 136255250 gb EBM92607.1	60,5	76,0	291
gi 137387954 gb EBT81681.1	60,5	76,6	124
gi 139229559 gb ECE29834.1	60,4	75,3	202
gi 140866196 gb ECO03646.1	60,1	74,6	303
gi 139229561 gb ECE29836.1	60,0	74,0	100
gi 140957440 gb ECO66006.1	59,9	75,9	274
gi 143567213 gb EDF67416.1	59,8	74,8	286
gi 139955973 gb ECI62051.1	59,8	72,2	97
gi 140726724 gb ECN09682.1	59,7	73,4	154
gi 139775003 gb ECH37281.1	59,6	73,7	99
gi 137949740 gb EBW9243.3.1	58,7	71,7	92
gi 139650150 gb ECG51661.1	57,7	73,2	97
gi 142528844 gb ECY73504.1	56,7	70,0	90

Punto de corte usado: >50% de identidad con y >25% de longitud de synpcc7942\_1594

Fig. 40  
 Los números de registro son de NCBI, GenBank, publicación 159.0 del 15 de abril de 2007  
 Los números EC son de KEGG, publicación 42.0 de abril de 2007 (más actualizaciones diarias hasta marzo de 2008).

CATEGORÍA	GEN	NOMBRE	REGISTRO	NÚMERO EC	MODIFICACIÓN	USO	ORGANISMO
1. Aumento de la producción de ácidos grasos / aumento de la producción de productos							
		aumento de acil-CoA					
		reducción del catabolismo de derivados y productos intermedios					
		reducción de la inhibición por retroalimentación					
		atenuación de otras rutas que consumen ácidos grasos					
	accA	Acetil-CoA carboxilasa, subunidad A (carboxiltransferasa alfa)	AAC73296, NP_414727	6.4.1.2	Sobreexpresión	aumentar la producción de malonil-CoA	<i>Escherichia coli</i> , <i>Lactococci</i>
	accB	Acetil-CoA carboxilasa, subunidad B (BCCP : proteína portadora de carboxilo de biotina)	NP_417721	6.4.1.2	Sobreexpresión	aumentar la producción de malonil-CoA	<i>Escherichia coli</i> , <i>Lactococci</i>
	accC	Acetil-CoA carboxilasa, subunidad C (biotina carboxilasa)	NP_417722	6.4.1.2, 6.3.4.14	Sobreexpresión	aumentar la producción de malonil-CoA	<i>Escherichia coli</i> , <i>Lactococci</i>
	accD	Acetil-CoA carboxilasa, subunidad D (carboxiltransferasa beta)	NP_416819	6.4.1.2	Sobreexpresión	aumentar la producción de malonil-CoA	<i>Escherichia coli</i> , <i>Lactococci</i>
	aceE	piruvato deshidrogenasa, subunidad E1	NP_414656, AAC73226	1.2.4.1	Sobreexpresión	aumentar la producción de acetil-CoA	<i>Escherichia coli</i>

aceF	piruvato deshidrogenasa, subunidad E2	NP_414657	2.3.1.12	Sobreexpresión	aumentar la producción de acetil-CoA	<i>Escherichia coli</i>
ackA	acetato cinasa	AAC75356, NP_416799	2.7.2.1	Delección o reducción	aumentar la producción de acetil-CoA	<i>Escherichia coli</i>
ackB	acetato cinasa AckB	BAB81430	2.7.2.1	Delección o reducción	aumentar la producción de acetil-CoA	<i>Escherichia coli</i>
acpP	proteína portadora de acilo	AAC74178	NINGUNO	Sobreexpresión	aumentar la producción de acetil-CoA	<i>Escherichia coli</i>
fadD	acil-CoA sintasa	AP_002424	2.3.1.86, 6.2.1.3	Sobreexpresión	aumentar la producción de ácidos grasos	<i>Escherichia coli</i> W3110
adhE	alcohol deshidrogenasa	CAA47743	1.1.1.1, 1.2.1.10	Delección o reducción	aumentar la producción de acetil-CoA	<i>Escherichia coli</i> W3111
cer1	Aldehído descarbonilasa	BAA11024	4.1.99.5	Sobreexpresión	aumentar la producción de acetil-CoA	<i>Arabidopsis thaliana</i>
fabA	beta-hidroxidecanoil tioéster deshidrasa	NP_415474	4.2.1.60	expresión	producción de acil graso-CoA	<i>E. coli</i> K12
fabD	[proteína portadora de acilo] S-maloniltransferasa	AAC74176	2.3.1.39	Sobreexpresión	aumentar la producción de acetil-CoA	<i>E. coli</i> K12
fabF	3-oxoacil-[proteína portadora de acilo] sintasa II	AAC74179	2.3.1.179	Delección o sobreexpresión	aumentar la producción de acetil-CoA	<i>E. coli</i> K12

	fabG	3-oxoacil-[proteína portadora de acilo] reductasa	AAC74177	1.1.1.100	Sobreexpresión	umentar la producción de acetil-CoA	<i>E. coli</i> K12
	fabH	3-oxoacil-[proteína portadora de acilo] sintasa III	AAC74175	2.3.1.180	Sobreexpresión	umentar la producción de acetil-CoA	<i>E. coli</i> K12, <i>lactococci</i>
	fabI	enoil-[proteína portadora de acilo] reductasa, dependiente de NADH	NP_415804	1.3.1.9	expresión	producción de acil graso-CoA	<i>E. coli</i> K12, <i>lactococci</i>
	fabR	Represor transcripcional (3R)-hidroximiristol proteína portadora de acilo deshidratasa	NP_418398	NINGUNO	Delección o reducción	modular la producción de ácidos grasos insaturados	<i>E. coli</i> K12
	fabZ		NP_414722	4.2.1.-			<i>E. coli</i> K12
	fadE	acil-CoA deshidrogenasa	AAC73325	1.3.99.3, 1.3.99.-	Delección o reducción	umentar la producción de acetil-CoA	
	acr1	Acil graso-CoA reductasa	YP_047869, AAC45217	1.2.1.42	Sobreexpresión	para la producción de alcoholes grasos	<i>Acinetobacter sp., es decir, calcoaceticus</i>
	GST, gshB	Glutación sintasa	P04425	6.3.2.3	Delección o reducción	umentar acil-CoA	<i>E. coli</i> K12
	gpsA	sn-glicerol 3-fosfato deshidrogenasa biosintética	AAC76632, NP_418065	EC: 1.1.1.94	Delección o reducción	umentar la producción de acetil-CoA	<i>E. coli</i> K12
	ldhA	lactato deshidrogenasa	AAC74462, NP_415898	EC: 1.1.1.27, 1.1.1.28	Delección o reducción	umentar la producción de acetil-CoA	<i>E. coli</i> K12
	Lipasa	Triglicérido lipasa	CAA89087, CAA98876	3.1.1.3	expresión	umentar la producción de ácidos grasos	<i>Saccharomyces cerevisiae</i>
		Malonil-CoA descarboxilasa	AAA26500	4.1.1.9, 4.1.1.41	Sobreexpresión		<i>Saccharopolyspora erythraea</i>
	panD	aspartato 1-descarboxilasa	BAB96708	4.1.1.11	Sobreexpresión	umentar acil-CoA	<i>Escherichia coli</i> W3110

panK a.k.a. coaA	pantotenato cinasa	AAC76952	2.7.1.33	Sobreexpresión	aumentar la producción de acetil-CoA	<i>E. coli</i>
panK a.k.a. coaA, R106K	pantotenato cinasa	AAC76952 BAB34380, AAC73226, NP_415392	2.7.1.33	Expresión, sobreexpresión, mutación R 106K	aumentar la producción de acetil-CoA	<i>E. coli</i>
pdh	Piruvato deshidrogenasa	AAC73989, P09373	1.2.4.1	Sobreexpresión	aumentar la producción de acetil-CoA	
pf1B	formiato acetiltransferasa (piruvato formiato liasa)	AAC77011	EC: 2.3.1.54	Delección o reducción	aumentar la producción de acetil-CoA	
plsB	aciltransferasa	AAC73958, NP_415392	2.3.1.15	Mutación D311E	reducir los límites en la reserva de acil-CoA	<i>E. coli</i> K12
poxB	piruvato oxidasa	AAC75357, NP_416800	1.2.2.2	Delección o reducción	aumentar la producción de acetil-CoA	
pta	fosfoacetiltransferasa	CAA46822	2.3.1.8	Delección o reducción	aumentar la producción de acetil-CoA	
udhA	nucleótido de piridina transhidrogenasa		1.6.1.1	Sobreexpresión	conversión de NADH en NADPH o viceversa	
fadB	3-hidroxiacetil-CoA epimerasa/delta(3)-cis-delta(2)-trans-enoil-CoA isomerasa/enoil-CoA hidratasa y 3-hidroxiacetil-CoA deshidrogenasa fusionadas	AP_003956	4.2.1.17, 5.1.2.3, 5.3.3.8, 1.1.1.35	Delección o reducción	Bloquear la degradación de ácidos grasos	<i>E. coli</i>



	umbelluria)									
	fatB3	tioesterasa	AAC72881	3.1.2.14	expresión o sobreexpresión	C14:0 - C16:0	Cuphea hookeriana			
	fatB (cinnam onum)	tioesterasa	Q39473	3.1.2.14	expresión o sobreexpresión	C14:0	Cinnamomum camphora			
	fatB[M 141T]*	tioesterasa	CAA85388	3.1.2.14	expresión o sobreexpresión	C16:1	Arabidopsis thaliana			
	fatA1 (Heliant hus)	tioesterasa	AAL79361	3.1.2.14	expresión o sobreexpresión	C18:1	Helianthus annuus			
	atfata (FATA ACIL-ACP TIOESTE- RASA DE ARABI- DOPSIS	tioesterasa	NP_189147, NP_1936041	3.1.2.14	expresión o sobreexpresión	C18:1	Arabidopsis thaliana			
	fatA	tioesterasa	CAC39106	3.1.2.14	expresión o sobreexpresión	C18:1	Brassica juncea			
	fatA (cuphea )	tioesterasa	AAC72883	3.1.2.14	expresión o sobreexpresión	C18:1	Cuphea hookeriana			
	2B. Control de ramificación									
	atenuación de FabH									



				EC 1.2.4.4					
				EC 1.2.4.4	NP_628006	a-cetoácido de cadena ramificada deshidrogenasa subunidad a (E1a)	expresión o sobreexpresión	producir precursores de acil de cadena ramificada-CoA	<i>Streptomyces coelicolor</i>
				EC 1.2.4.4	NP_628005	a-cetoácido de cadena ramificada deshidrogenasa subunidad a (E1b)	expresión o sobreexpresión	producir precursores de acil de cadena ramificada-CoA	<i>Streptomyces coelicolor</i>
				EC 2.3.1.168	NP_628004	dihidrolipoil transacetilasa (E2)	expresión o sobreexpresión	producir precursores de acil de cadena ramificada-CoA	<i>Streptomyces coelicolor</i>
				EC 1.2.4.4	NP_733618	a-cetoácido de cadena ramificada deshidrogenasa subunidad a (E1a)	expresión o sobreexpresión	producir precursores de acil de cadena ramificada-CoA	<i>Streptomyces coelicolor</i>
				EC 1.2.4.4	NP_628019	a-cetoácido de cadena ramificada deshidrogenasa subunidad b (E1b)	expresión o sobreexpresión	producir precursores de acil de cadena ramificada-CoA	<i>Streptomyces coelicolor</i>
				EC 2.3.1.168	NP_628018	dihidrolipoil transacetilasa (E2)	expresión o sobreexpresión	producir precursores de acil de cadena ramificada-CoA	<i>Streptomyces coelicolor</i>
subunidad bkd - E2									
- dihidrolipoil transacetilasa									
bkdA1									
bkdB1									
bkdC1									
bkdA2									
bkdB2									
bkdC2									

bkdA	a-cetoácido de cadena ramificada deshidrogenasa subunidad a (E1a)	<u>BAC72074</u>	EC 1.2.4.4	expresión o sobreexpresión	producir precursores de acil de cadena ramificada-CoA	<i>Streptomyces avermitilis</i>
bkdB	a-cetoácido de cadena ramificada deshidrogenasa subunidad b (E1b)	<u>BAC72075</u>	EC 1.2.4.4	expresión o sobreexpresión	producir precursores de acil de cadena ramificada-CoA	<i>Streptomyces avermitilis</i>
bkdC	dihidrolipoil transacetilasa (E2)	<u>BAC72076</u>	EC 2.3.1.168	expresión o sobreexpresión	producir precursores de acil de cadena ramificada-CoA	<i>Streptomyces avermitilis</i>
bkdF	a-cetoácido de cadena ramificada deshidrogenasa subunidad a (E1a)	<u>BAC72088</u>	EC 1.2.4.4	expresión o sobreexpresión	producir precursores de acil de cadena ramificada-CoA	<i>Streptomyces avermitilis</i>
bkdG	a-cetoácido de cadena ramificada deshidrogenasa subunidad b (E1b)	<u>BAC72089</u>	EC 1.2.4.4	expresión o sobreexpresión	producir precursores de acil de cadena ramificada-CoA	<i>Streptomyces avermitilis</i>
bkdH	dihidrolipoil transacetilasa (E2)	<u>BAC72090</u>	EC 2.3.1.168	expresión o sobreexpresión	producir precursores de acil de cadena ramificada-CoA	<i>Streptomyces avermitilis</i>
bkdAA	a-cetoácido de cadena ramificada deshidrogenasa subunidad a (E1a)	NP_390285	EC 1.2.4.4	expresión o sobreexpresión	producir precursores de acil de cadena ramificada-CoA	<i>Bacillus subtilis</i>
bkdAB	a-cetoácido de cadena ramificada deshidrogenasa subunidad b (E1b)	NP_390284	EC 1.2.4.4	expresión o sobreexpresión	producir precursores de acil de cadena ramificada	<i>Bacillus subtilis</i>



IlVE	aminoácido de cadena ramificada aminotransferasa	NP_629657	2.6.1.42	expresión o sobreexpresión	producir a-cetoácidos ramificados	<i>Streptomyces coelicolor</i>
ccr	crotonil-CoA reductasa	NP_630556	1.6.5.5.1.1.1.1	expresión o sobreexpresión	convertir crotonil-CoA en butiril-CoA	<i>Streptomyces coelicolor</i>
ccr	crotonil-CoA reductasa	AAD53915	1.6.5.5.1.1.1.1	expresión o sobreexpresión	convertir crotonil-CoA en butiril-CoA	<i>Streptomyces cinnamonensis</i>
lcmA, isobutiril-CoA mutasa	isobutiril-CoA mutasa, subunidad A	NP_629554	5.4.99.2	expresión o sobreexpresión	convertir butiril-CoA en isobutiril-CoA	<i>Streptomyces coelicolor</i>
lcmA, isobutiril-CoA mutasa	isobutiril-CoA mutasa, subunidad A	AAC08713	5.4.99.2	expresión o sobreexpresión	convertir butiril-CoA en isobutiril-CoA	<i>Streptomyces cinnamonensis</i>
lcmB, isobutiril-CoA mutasa	isobutiril-CoA mutasa, subunidad B	NP_630904	5.4.99.2	expresión o sobreexpresión	convertir butiril-CoA en isobutiril-CoA	<i>Streptomyces coelicolor</i>
lcmB, isobutiril-CoA mutasa	isobutiril-CoA mutasa, subunidad B	CAB59633	5.4.99.2	expresión o sobreexpresión	convertir butiril-CoA en isobutiril-CoA	<i>Streptomyces cinnamonensis</i>

genes de FabH, ACP y fabF con especificidad por acil de cadena ramificada-CoA									
IIVe	aminoácido de cadena ramificada aminotransferasa	CAC12788	EC2.6.1.42	sobreexpresión	aminoácido de cadena ramificada aminotransferasa	Staphylococcus carnosus			
FabH1	beta-cetoacil-ACP sintasa III	NP_62634	2.3.1.180	expresión o sobreexpresión	biosíntesis de ácidos grasos de cadena ramificada	Streptomyces coelicolor			
ACP	proteína portadora de acilo	NP_626635	NINGUNO	expresión o sobreexpresión	Iniciación y elongación de la biosíntesis de ácidos grasos de cadena ramificada	Streptomyces coelicolor			
FabF	beta-cetoacil-ACP sintasa II	NP_626636	2.3.1.179	expresión o sobreexpresión	Elongación de la biosíntesis de ácidos grasos de cadena ramificada	Streptomyces coelicolor			

FabH3	beta-cetoacil-ACP sintasa III	NP_823466	2.3.1.180	expresión o sobreexpresión	iniciación de la biosíntesis de ácidos grasos de cadena ramificada	<i>Streptomyces avermitilis</i>
FabC3 (ACP)	proteína portadora de acilo	NP_823467	NINGUNO	expresión o sobreexpresión	iniciación y elongación de la biosíntesis de ácidos grasos de cadena ramificada	<i>Streptomyces avermitilis</i>
FabF	beta-cetoacil-ACP sintasa II	NP_823468	2.3.1.179	expresión o sobreexpresión	elongación de la biosíntesis de ácidos grasos de cadena ramificada	<i>Streptomyces avermitilis</i>
FabH_A	beta-cetoacil-ACP sintasa III	NP_389015	2.3.1.180	expresión o sobreexpresión	iniciación de la biosíntesis de ácidos grasos de cadena ramificada	<i>Bacillus subtilis</i>
FabH_B	beta-cetoacil-ACP sintasa III	NP_388898	2.3.1.180	expresión o sobreexpresión	iniciación de la biosíntesis de ácidos grasos de cadena ramificada	<i>Bacillus subtilis</i>

ACP	proteína portadora de acilo	NP_389474	NINGUNO	expresión o sobreexpresión	iniciación y elongación de la biosíntesis de ácidos grasos de cadena ramificada	<i>Bacillus subtilis</i>
FabF	beta-cetoacil-ACP sintasa II	NP_389016	2.3.1.179	expresión o sobreexpresión	elongación de la biosíntesis de ácidos grasos de cadena ramificada	<i>Bacillus subtilis</i>
SmaID RAFT_0818	beta-cetoacil-ACP sintasa III	ZP_01643059	2.3.1.180	expresión o sobreexpresión	iniciación de la biosíntesis de ácidos grasos de cadena ramificada	<i>Stenotrophomonas maltophilia</i>
SmaID RAFT_0821	proteína portadora de acilo	ZP_01643063	NINGUNO	expresión o sobreexpresión	iniciación y elongación de la biosíntesis de ácidos grasos de cadena ramificada	<i>Stenotrophomonas maltophilia</i>
SmaID RAFT_0822	beta-cetoacil-ACP sintasa II	ZP_01643064	2.3.1.179	expresión o sobreexpresión	elongación de la biosíntesis de ácidos grasos de cadena ramificada	<i>Stenotrophomonas maltophilia</i>



AnsJ	deshidratasa (supuesta)	no disponible	no disponible	expresión o sobreexpresión	biosíntesis de ciclohexilcarbonil-CoA	<i>Streptomyces collinus</i>
AnsK	CoA ligasa (supuesta)	no disponible	no disponible	expresión o sobreexpresión	biosíntesis de ciclohexilcarbonil-CoA	<i>Streptomyces collinus</i>
AnsL	deshidrogenasa (supuesta)	no disponible	no disponible	expresión o sobreexpresión	biosíntesis de ciclohexilcarbonil-CoA	<i>Streptomyces collinus</i>
ChcA	enoiil-CoA reductasa	U72144	EC 1.3.1.34	expresión o sobreexpresión	biosíntesis de ciclohexilcarbonil-CoA	<i>Streptomyces collinus</i>
AnsM	oxidorreductasa (supuesta)	no disponible	no disponible	expresión o sobreexpresión	biosíntesis de ciclohexilcarbonil-CoA	<i>Streptomyces collinus</i>
PlmJ	deshidratasa (supuesta)	AAQ84158	no disponible	expresión o sobreexpresión	biosíntesis de ciclohexilcarbonil-CoA	<i>Streptomyces sp. HK803</i>
PlmK	CoA ligasa (supuesta)	AAQ84158	no disponible	expresión o sobreexpresión	biosíntesis de ciclohexilcarbonil-CoA	<i>Streptomyces sp. HK803</i>
PlmL	deshidrogenasa (supuesta)	AAQ84159	no disponible	expresión o sobreexpresión	biosíntesis de ciclohexilcarbonil-CoA	<i>Streptomyces sp. HK803</i>



GnsA	supresores de la mutación secG null	ABD18647.1	NINGUNO	Sobreexpresión	aumentar ésteres de ácidos grasos insaturados	<i>E. coli</i>
GnsB Véase también la sección 2A- artículos con :0 son insaturados (sin dobles enlaces) y con :1 son saturados (1 doble enlace)	supresores de la mutación secG null	AAC74076.1	NINGUNO	Sobreexpresión	aumentar ésteres de ácidos grasos insaturados	<i>E. coli</i>
fabB	3-oxoacil-[proteína portadora de acilo] sintasa I	BAA16180	EC:2.3.1.41	sobreexpresión	modular la producción de ácidos grasos insaturados	<i>Escherichia coli</i>
fabK	trans-2-enoil-ACP reductasa II	AAF98273	1.3.1.9	expresión	modular la producción de ácidos grasos insaturados	<i>Streptococcus pneumoniae</i>
fabL	enoiil-(proteína portadora de acilo) reductasa	AAU39821	1.3.1.9	expresión	modular la producción de ácidos grasos insaturados	<i>Bacillus licheniformis</i> DSM 13

	fabM	trans-2, cis-3-decenil-ACP isomerasa	DAA05501	4.2.1.17	Sobreexpresión	modificar la producción de ácidos grasos insaturados	<i>Streptococcus mutans</i>
3. Salida de producto final.							
3A. Salida de cera							
	AT3G51970	alcohol de cadena larga O-graso aciltransferasa	NP_190765	2.3.1.26	expresión	producción de cera	<i>Arabidopsis thaliana</i>
		fosfoesterasa (véase sección de control de longitud de cadena)			expresión	aumentar la producción de ácidos grasos	
		acil-CoA reductasa que forma alcohol graso		1.1.1.*	expresión	convertir acil-coa en alcohol graso	
	acr1	acil-CoA reductasa (ACR1)	YP_047869	1.2.1.42	expresión	convertir acil-coa en alcohol graso	<i>Acinetobacter sp. ADPI</i>
	yqtA	alcohol deshidrogenasa	AP_003562	1.1.-.-	expresión	aumento	<i>E. coli W3110</i>
	ELO1	Ácido graso elongasa	BAD98251	2.3.1.-	expresión	producir ácidos grasos de longitud de cadena muy larga	<i>Pichia angusta</i>
	plsC	aciltransferasa	AAA16514	2.3.1.51	expresión		<i>Saccharomyces cerevisiae</i>
	DAGA T/DGA T	diacilglicerol aciltransferasa	AAF19262	2.3.1.20	expresión	producción de cera	<i>Arabidopsis thaliana</i>
	hws	acil-CoA cera alcohol aciltransferasa	AAX48018	2.3.1.20	expresión	producción de cera	<i>Homo sapiens</i>



	ALDH	coenzima A-acilante aldehído deshidrogenasa	AAT66436	1.2.1.3	expresión	producir	<i>Clostridium beijerinckii</i>
	AdhE	aldehído-alcohol deshidrogenasa	AAN80172	1.1.1.1 1.2.1.10	expresión	producir	<i>Escherichia coli CF1073</i>
3C. Salida de éster de ácido graso							
	tiesterasa	véase la sección de control de la longitud de cadena			expresión	producir	<i>Acinetobacter sp. ADP1</i>
	acr1	acil-CoA reductasa	YP_047869	1.2.1.42	expresión	producir	<i>E. Coli K12</i>
	yqhD	alcohol deshidrogenasa	AP_003562	1.1.-.-	expresión	producir	<i>Fragaria x ananassa</i>
	AAT	alcohol O-acetiltransferasa	AAG13130	2.3.1.84	expresión	producir	
4. Exportación							
	exportador de éster de cera (familia FATP, proteína de trans- porte de ácido graso (cadena larga)					exportar cera	<i>Drosophila melanogaster</i>
	proteína de trans- porte ABC	supuesto transportador de alcano	AANT3268	NINGUNO	expresión	exportar productos	<i>Rhodococcus erythropolis</i>
	CERS	transportador de cera	At1g51500, AY734542, At3g21090, At1g51460	NINGUNO	expresión	exportar productos	<i>Arabidopsis thaliana</i>

	AtMRP 5	asociada a resistencia a múltiples fármacos de <i>Arabidopsis thaliana</i>	NP_171908	NINGUNO	expresión	exportar productos	<i>Arabidopsis thaliana</i>
	AmiS2	transportador ABC AmiS2	JC5491	NINGUNO	expresión	exportar productos	<i>Rhodococcus sp.</i>
	AtPGP1	P GLICOPROTEINA 1 DE <i>ARABIDOPSIS THALIANA</i>	NP_181228	NINGUNO	expresión	exportar productos	<i>Arabidopsis thaliana</i>
	AcrA	supuesta proteína de transporte de flujo de salida de múltiples fármacos, acrA	CAF23274	NINGUNO	expresión	exportar productos	<i>Candidatus Prochloromydia amoebophila UWE25</i>
	AcrB	probable proteína de transporte de flujo de salida de múltiples fármacos, acrB	CAF23275	NINGUNO	expresión	exportar productos	<i>Candidatus Prochloromydia amoebophila UWE25</i>
	ToIC	proteína de membrana externa [biogénesis de envuelta celular, la proteína transmembrana afecta a la formación del septo y la permeabilidad de la membrana celular	ABD59001	NINGUNO	expresión	exportar productos	<i>Francisella tularensis subsp. novicida</i>
	AcrE	Proteína F de resistencia a acriflavina	YP_312213	NINGUNO	expresión	exportar productos	<i>Shigella sonnei Ss046</i>
	AcrF	transportador de flujo de salida de múltiples fármacos	P24181	NINGUNO	expresión	exportar productos	<i>Escherichia coli</i>
	tll1618	transportador de flujo de salida de múltiples fármacos	NP_682408.1	NINGUNO	expresión	exportar productos	<i>Thermosynechococcus elongatus BP-1J</i>
	tll1619	transportador de flujo de salida de múltiples fármacos	NP_682409.1	NINGUNO	expresión	exportar productos	<i>Thermosynechococcus elongatus BP-1J</i>
	tll1639	transportador de flujo de salida de múltiples fármacos	NP_680930.1	NINGUNO	expresión	exportar productos	<i>Thermosynechococcus elongatus BP-1J</i>
5. Fermentación							

