

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 993 463**

51 Int. Cl.:

**C12N 9/02** (2006.01)

**C12P 5/02** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **19.05.2017 PCT/EP2017/062061**

87 Fecha y número de publicación internacional: **23.11.2017 WO17198802**

96 Fecha de presentación y número de la solicitud europea: **19.05.2017 E 17726576 (6)**

97 Fecha y número de publicación de la concesión europea: **06.11.2024 EP 3458577**

54 Título: **Nueva ácido graso descarboxilasa y sus usos**

30 Prioridad:

**20.05.2016 EP 16305583**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**30.12.2024**

73 Titular/es:

**COMMISSARIAT À L'ENERGIE ATOMIQUE ET  
AUX ENERGIES ALTERNATIVES (33.33%)  
25 Rue Leblanc Bat Le Ponant  
75015 Paris, FR;  
CENTRE NATIONAL DE LA RECHERCHE  
SCIENTIFIQUE (33.33%) y  
UNIVERSITE D'AIX-MARSEILLE (33.33%)**

72 Inventor/es:

**BEISSON, FRÉDÉRIC;  
SORIGUE, DAMIEN;  
LEGERET, BERTRAND;  
CUINE, STÉPHAN;  
BLANGY, STÉPHANIE y  
PELTIER, GILLES**

74 Agente/Representante:

**VALLEJO LÓPEZ, Juan Pedro**

Observaciones:

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

ES 2 993 463 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

Nueva ácido graso descarboxilasa y sus usos

5 **Campo de la invención**

La presente invención se refiere al campo de la enzimología, en particular a la ácido graso descarboxilasa.

10 **Antecedentes de la invención**

Los alcanos y los alquenos son los componentes principales de los combustibles fósiles, pero también se producen naturalmente en plantas, insectos, cianobacterias y algunas bacterias. Las diatomeas y las colonias de clorofíceas *Botryococcus braunii* también son conocidas por producir alca(que)nos con más de 21 átomos de carbono.

15 La mayoría de las rutas sintéticas de los alca(que)nos implican la conversión de ácidos grasos activados (acil-ACP o acil-CoA) en un compuesto intermedio de aldehído y la descarboxilación final a alca(que)no. Estas reacciones están catalizadas por diversas enzimas. En las plantas, los alcanos de cadena muy larga de 27 a 33 átomos de carbono se producen a partir de ácidos grasos mediante la acción de dos oxidoreductasas posibles homólogas ECERIFERUM1 (CER1) y ECERIFERUM3 (CER3) (Bernard et al., 2012, *The Plant Cell*, 24(7), 3106-18; Bourdenx et al., 2011, *Plant Physiology*, 156(1), 29-45). En los insectos, la síntesis de alca(que)nos cuticulares a partir de 21 a 37 átomos de carbono implica una acil éster reductasa y un citocromo P450 (Qiu et al., 2012, *PNAS*, 109, 14858-14863). En las cianobacterias, la ruta se compone de una acil-ACP reductasa y una aldehído deformilante oxigenasa (Li et al., 2012, *Biochemistry*, 51(40), 7908-16; Rude et al., 2011, *Applied and Environmental Microbiology*, 77(5), 1718-27). Solo las bacterias *Jeotgalicoccus spp* y *Pseudomonas sp* son capaces de producir hidrocarburos mediante la descarboxilación directa de ácidos grasos libres (Grant et al 2015, *Journal of the American Chemical Society*, 137(15), 4940-3; Rui et al., 2014, *PNAS*, 111, 18237-18242; Rude et al., 2011, *Applied and Environmental Microbiology*, 77(5), 1718-27). Sin embargo, estas ácido graso descarboxilasas bacterianas pueden producir alquenos (que tienen insaturaciones terminales) pero no alcanos. En las microalgas, no se conoce ninguna enzima que catalice la síntesis de hidrocarburos. Se han hecho intentos de purificar el sistema de síntesis de alcanos de la microalga *Botryococcus braunii* y se ha sugerido que la síntesis transcurre mediante la descarboxilación de un compuesto intermedio de ácido graso mediante una enzima de cobalto-porfirina (Dennis y Kolattukudy, 1992, *PNAS*, 89(12), 5306-10). Sin embargo, la proteína nunca se ha identificado.

35 Por lo tanto, existe una fuerte necesidad de enzimas adecuadas para producir alcanos/alquenos a partir de ácidos grasos.

**Sumario de la invención**

40 La presente invención se refiere a la identificación de una nueva clase de enzimas para la síntesis de alcanos y alquenos, que se ha identificado en microalgas. Es la primera vez que una alcano sintasa se ha identificado en estos microorganismos. La enzima cataliza la conversión de ácidos grasos libres en alcanos y/o alquenos mediante descarboxilación. La enzima pertenece a una superfamilia de proteínas dependientes de FAD presentes en procariontas y eucariotas (GMC (Glucosa-Metanol-Colina) oxidoreductasas) y que incluye varias enzimas (principalmente, alcohol oxidasas). De este modo, ninguna GMC oxidoreductasa procedente de microalgas se había caracterizado desde un punto de vista bioquímico, y ningún otro miembro de la superfamilia de las GMC oxidoreductasas se había identificado como una alcano sintasa. Por lo tanto, se trata de una nueva enzima para la síntesis de alcanos y alquenos con aplicación en biocombustibles, química verde, diagnóstico y nutrición.

50 Por consiguiente, la presente invención se refiere al uso de un polipéptido o de una célula que expresa dicho polipéptido para producir alcanos/alquenos a partir de ácidos grasos, en el que el polipéptido tiene actividad ácido graso descarboxilasa y comprende una secuencia que tiene al menos un 40 % con la SEQ ID No 1.

55 Preferentemente, el polipéptido comprende la secuencia de consenso G-X-L(X)4-C-[D/E]-X-G-[A/G]-F-X-[K/R] (SEQ ID No 4), siendo X un aminoácido. De forma alternativa o adicional, uno, dos, tres, cuatro, cinco o seis aminoácidos en las posiciones C372, R391, Y406, Q426, H512 y N515 de la SEQ ID No 1 están conservados. De forma alternativa o adicional, al menos un 40 % de los restos de aminoácido situados entre las posiciones 388-428 son restos hidrófobos seleccionados entre los grupos que consisten en V, I, L, M, F, W, C, A, e Y.

60 En una realización más específica, el polipéptido se caracteriza además por que su secuencia de aminoácidos se selecciona del grupo que consiste en las SEQ ID No 1-3 y 5-14 y secuencias de aminoácidos que tienen al menos un 80 % de identidad con una de ellas.

Preferentemente, el polipéptido procede de algas, preferentemente microalgas o cianobacterias.

65 Preferentemente, los ácidos grasos y los correspondientes alcanos y alquenos descarboxilados comprenden de 8 a 24 átomos de carbono, preferentemente de 12 a 22 átomos de carbono.

Opcionalmente, los ácidos grasos y los correspondientes alcanos y alquenos descarboxilados están sustituidos y/o interrumpidos por uno o varios grupos funcionales. Preferentemente, los ácidos grasos y los correspondientes alcanos y alquenos descarboxilados están sustituidos por uno o varios grupos tales como un hidroxilo, un alcohol C<sub>1</sub>-C<sub>3</sub>, un acilo C<sub>1</sub>-C<sub>3</sub>, un éster C<sub>1</sub>-C<sub>3</sub>, una amina C<sub>1</sub>-C<sub>3</sub>, un grupo amino, una amida C<sub>1</sub>-C<sub>3</sub>, un carboxilo, un aldehído, un epoxi, un halógeno, un alcoxi C<sub>1</sub>-C<sub>3</sub>, un tioalquilo C<sub>1</sub>-C<sub>3</sub>, una imina C<sub>1</sub>-C<sub>3</sub>, un nitrilo, un grupo sulfuro tal como una sulfona C<sub>1</sub>-C<sub>3</sub> o sulfóxido C<sub>1</sub>-C<sub>3</sub>, un tiol, un nitro, un ciano, un alquil C<sub>1</sub>-C<sub>3</sub> halógeno, o puede estar interrumpido por un heteroátomo tal como O, N o S, un grupo acetilénico, un éter tal como un grupo divinil éter, o un grupo oxo.

La presente invención también se refiere a un método para producir alcanos y alquenos a partir de ácidos grasos, en el que un polipéptido tal como se ha definido anteriormente se pone en contacto con ácidos grasos y se somete a luz en presencia del cofactor FAD del polipéptido. Preferentemente, la luz tiene una longitud de onda comprendida entre 300 y 540 nm, más preferentemente entre 400 y 520 nm. Por ejemplo, la luz es una luz blanca o cualquier luz que contenga fotones de color azul o UV (por ejemplo, de 400 a 520 nm).

La presente invención también se refiere a una célula hospedadora recombinante diseñada por ingeniería genética para producir alcanos/alquenos que comprende una secuencia de ácido nucleico heterólogo que codifica un polipéptido heterólogo que tiene una actividad ácido graso descarboxilasa y que comprende una secuencia que tiene al menos un 95 % de identidad con la secuencia de aminoácidos seleccionada de las SEQ ID No, siendo la célula hospedadora preferentemente una bacteria, una microalga, un hongo filamentoso o una levadura.

En una primera realización, el polipéptido heterólogo se expresa simultáneamente con una tioesterasa, preferentemente en una microalga, o una cianobacteria. En otra realización, el polipéptido heterólogo se expresa simultáneamente con una lipasa, preferentemente en una bacteria.

La presente invención se refiere a un método para producir alcanos y alquenos a partir de ácidos grasos, en el que se cultiva una célula hospedadora recombinante de acuerdo con la presente invención, y se recuperan los alcanos y alquenos.

Además, la presente invención se refiere al uso de un polipéptido como se ha definido anteriormente para la dosificación de ácidos grasos en una muestra. Más en particular, el método para dosificar ácidos grasos en una muestra comprende (a) poner en contacto la muestra con una ácido graso descarboxilasa de la presente invención en condiciones adecuadas para convertir los ácidos grasos en alcanos/alquenos, (b) recuperar los alcanos/alquenos y/o CO<sub>2</sub>; y (c) cuantificar los alcanos/alquenos y/o CO<sub>2</sub>. Como alternativa, el método para dosificar ácidos grasos en una muestra comprende (a) poner en contacto la muestra con una ácido graso descarboxilasa de la presente invención en condiciones adecuadas para convertir los ácidos grasos en alcanos/alquenos, y (b) medir la fluorescencia emitida por la ácido graso descarboxilasa. Opcionalmente, el método comprende una etapa anterior de poner en contacto la muestra con una lipasa, en particular en condiciones adecuadas para convertir los ácidos grasos en ácidos grasos libres.

La presente invención, también se refiere al uso de un polipéptido tal como se ha definido anteriormente en la descarboxilación de ácidos grasos.

#### Descripción detallada de la invención

Los inventores han identificado una nueva clase de enzimas capaz de conversión de ácidos grasos libres en alcanos y alquenos mediante una descarboxilación del ácido graso.

#### Definiciones

*Aproximadamente*: Cuando se usa en el presente documento, "aproximadamente" significa más o menos el 10 %, preferentemente más o menos el 5 %. Por ejemplo, aproximadamente 100 significa entre 90 y 110, preferentemente entre 95 y 105.

*Secuencia de codificación*: La expresión "secuencia de codificación" significa un polinucleótido, que especifica directamente la secuencia de aminoácidos del polipéptido. Los límites de la secuencia de codificación están determinados generalmente por un marco de lectura abierto, que comienza con un codón de inicio tal como ATG, GTG, o TTG y finaliza con un codón de terminación tal como TAA, TAG, o TGA. La secuencia de codificación puede ser ADN genómico, ADNc, ADN sintético, o una combinación de los mismos. El código genético se puede optimizar para la célula hospedadora.

*Secuencias de control*: La expresión "secuencias de control" significa secuencias de ácidos nucleicos necesarias para la expresión de un polinucleótido que codifica una enzima de la presente invención. Las secuencias de control pueden ser naturales (es decir, del mismo género) o heterólogas (es decir, de un gen diferente y/o de una especie diferente) a la del polinucleótido que codifica la enzima. Preferentemente, las secuencias de control son heterólogas. Se preferirán las secuencias de control bien conocidas y actualmente utilizadas por el experto en la materia. Dichas

secuencias de control incluyen, aunque no de forma limitativa, un líder, secuencia de poliadenilación, secuencia de propéptido, promotor, secuencia péptido señal, y terminador de la transcripción. Como mínimo, las secuencias de control incluyen un promotor, y señales de inicio y finalización de la transcripción. Las secuencias de control se pueden proporcionar con enlazadores con el fin de introducir sitios de restricción específicos que faciliten la ligadura de las secuencias de control a la región de codificación del polinucleótido que codifica la enzima. La combinación funcional entre las secuencias de control y las secuencias de codificación se puede denominar como *casete de expresión*.

*Expresión:* El término "expresión" incluye cualquier etapa implicada en la producción de un polipéptido incluidas, aunque no de forma limitativa, transcripción, modificaciones posteriores a la transcripción, traducción, modificación posterior a la traducción, y secreción.

*Vector de expresión:* El término "vector de expresión" significa una molécula de ADN lineal o circular que comprende un polinucleótido que codifica la enzima de la invención y está operativamente unido a las secuencias de control que proporcionan su expresión. A continuación, el vector de expresión comprende un casete de expresión adecuado para expresar la enzima de la invención.

*Aislado:* El término "aislado" significa una sustancia en una forma o entorno que no se produce en la naturaleza. Los ejemplos no limitantes de sustancias aisladas incluyen (1) cualquier sustancia de origen no natural, (2) cualquier sustancia que incluya, aunque no de forma limitativa, cualquier enzima, variante, ácido nucleico, proteína, péptido o cofactor, al que se haya eliminado al menos parcialmente uno o más o todos de los componentes naturales con los que está asociado en la naturaleza; (3) cualquier sustancia modificada por la mano del hombre, con respecto a dicha sustancia cuando se encuentra en la naturaleza; o (4) cualquier sustancia modificada mediante el aumento de la sustancia con respecto a otros componentes con los que está asociado de forma natural (por ejemplo, múltiples copias de un gen que codifica la sustancia; uso de un promotor más fuerte que el promotor asociado de forma natural con el gen que codifica la sustancia).

*Recombinante:* Recombinante se refiere a una construcción de ácido nucleico, un vector y una proteína producida mediante ingeniería genética.

*Heterólogo:* en el contexto de una célula hospedadora, un vector o construcción de ácido nucleico, designa una secuencia de codificación de la enzima introducida en la célula hospedadora, el vector o la construcción de ácido nucleico mediante ingeniería genética. En el contexto de una célula hospedadora, significa que la secuencia de codificación de la enzima se origina de una fuente diferente de la célula en la que se ha introducido o bien que la secuencia de codificación de la enzima procede de la misma especie que la célula hospedadora en la que se introduce pero se considera heteróloga debido a que su entorno es no natural, por ejemplo, porque está bajo el control de un promotor que no es su promotor natural, o se ha introducido en una ubicación que difiere de su localización natural.

*Construcción de ácidos nucleicos:* La expresión "construcción de ácidos nucleicos" significa una molécula de ácido nucleico, tanto monocatenaria como bicatenaria, que se ha modificado para contener segmentos de ácidos nucleicos de una forma que no existiría en la naturaleza o que sea sintética, que comprende una o más secuencias de control.

*Unido de manera operativa:* La expresión "unido de manera operativa" significa una configuración en la que una secuencia de control está situada en una posición relativa adecuada con respecto a una secuencia de codificación, de tal forma que la secuencia de control dirige la expresión de la secuencia de codificación.

*Expresar en exceso:* La expresión expresar en exceso significa expresar o producir la expresión de un ácido nucleico o polipéptido en una célula en mayor concentración de la que se expresa normalmente en una correspondiente célula de tipo natural.

*Identidad de secuencia:* La identidad de secuencia entre dos secuencias de aminoácidos se describe por el parámetro "identidad de secuencia". Para los fines de la presente invención, el "porcentaje de identidad" entre dos secuencias de aminoácidos (A) y (B) se determina por comparación de dos secuencias alineadas de forma óptima, mediante una ventana de comparación. Dicho alineamiento de secuencias se puede llevar a cabo por métodos bien conocidos, por ejemplo, usando el algoritmo de alineamiento global de Needleman-Wunsch. El programa informático de análisis de proteína empareja secuencias similares utilizando medidas de similitud asignadas a diversas sustituciones, deleciones y otras modificaciones, que incluyen sustituciones de aminoácidos conservativas. Una vez que se ha obtenido el alineamiento total, el porcentaje de identidad se puede obtener dividiendo el número completo de restos de aminoácidos idénticos alineados por el número completo de restos contenidos en la secuencia más larga entre la secuencia (A) y (B).

La identidad de secuencia se determina normalmente usando un programa informático de análisis de secuencias. Para comparar dos secuencias de aminoácidos, se puede usar, por ejemplo, la herramienta "Emboss needle" para alineamiento de secuencias emparejadas que proporciona EMBL-EBI y está disponible en: [www.ebi.ac.uk/Tools/services/web/toolform.ebi?tool=emboss\\_needle&context=protein](http://www.ebi.ac.uk/Tools/services/web/toolform.ebi?tool=emboss_needle&context=protein), utilizando la configuración por defecto: (i) Matriz: BLOSUM62, (ii) Apertura de hueco: 10, (iii) extensión de hueco: 0,5, (iv) formato de salida: pareja, (v) penalización por finalización de hueco: falso, (v) finalización de hueco abierto: 10, (vii) finalización de extensión de

hueco: 0,5.

Como alternativa, la Identidad de secuencia también se puede determinar de forma típica usando el programa informático de análisis de secuencia Clustal Omega con el algoritmo HAlign y su configuración predeterminada como su motor de alineación principal. El algoritmo se describe en Söding, J. (2005) 'Protein homology detection by HMM-HMM comparison'. *Bioinformatics* 21, 951-960, con la configuración predeterminada. ([www.ebi.ac.uk/Tools/msa/clustalo/](http://www.ebi.ac.uk/Tools/msa/clustalo/))

*Aminoácidos*: Las secuencias de aminoácidos definidas en el presente documento están representadas por un símbolo de una sola letra, como se muestra a continuación: A, Ala, (alanina); R, Arg, (arginina); N, Asn, (asparagina); D, Asp, (ácido aspártico); C, Cys, (cisteína); Q, Gln, (glutamina); E, Glu, (ácido glutámico); G, Gly, (glicina); H, His, (histidina); I, Ile, (isoleucina); L, Leu, (leucina); K, Lys, (lisina); M, Met, (metionina); F, Phe, (fenilalanina); P, Pro, (prolina); S, Ser, (serina); T, Thr, (treonina); W, Trp, (triptófano); Y, Tyr, (tirosina); y V, Val, (valina).

Por "consiste esencialmente de" se pretende indicar que el polipéptido tiene la SEQ ID No indicada y puede comprender además una alteración, es decir, una sustitución, inserción, y/o delección, de no más de 20 aminoácidos, preferentemente de no más de 10 aminoácidos. En particular, el polipéptido puede tener alteraciones en no más de 11, 10, 9, 8, 7, 6, 5, 4, 3, 2 o 1 aminoácido, por ejemplo, puede tener una sustitución, inserción, y/o delección de 1, 2, 3, 4, 5, 6, 7, 8, 9 o 10 aminoácidos. Una sustitución significa la sustitución del aminoácido que ocupa una posición por un aminoácido diferente; una delección significa la eliminación del aminoácido que ocupa una posición; y una inserción significa añadir un aminoácido adyacente a, e inmediatamente después, del aminoácido que ocupa una posición. La sustitución puede ser una sustitución conservativa. Los ejemplos de sustituciones conservativas se realizan entre los grupos de aminoácidos básicos (arginina, lisina e histidina), aminoácidos ácidos (ácido glutámico y ácido aspártico), aminoácidos polares (glutamina y asparagina), aminoácidos hidrófobos (leucina, isoleucina y valina), aminoácidos aromáticos (fenilalanina, triptófano, y tirosina) y aminoácidos pequeños (glicina, alanina, serina, treonina y metionina). Las sustituciones de aminoácidos que no alteran de forma general la actividad específica son conocidos en la técnica y se describen, por ejemplo, en H. Neurath y R.L. Hill (1979, *En, The Proteins*, Academic Press, Nueva York). Las sustituciones habituales son las siguientes Ala/Ser, Val/Ile, Asp/Glu, Thr/Ser, Ala/Gly, Ala/Thr, Ser/Asn, Ala/Val, Ser/Gly, Tyr/Phe, Ala/Pro, Lys/Arg, Asp/Asn, Leu/Ile, Leu/Ala, Ala/Glu, y Asp/Gly.

Como alternativa, los cambios de aminoácidos son tal tipo que las propiedades fisicoquímicas de los polipéptidos quedan alteradas. Por ejemplo, los cambios de aminoácidos pueden mejorar la estabilidad térmica del polipéptido, alterar la especificidad de sustrato, cambiar el pH óptimo, y similares. Los aminoácidos esenciales de un polipéptido se pueden identificar según procedimientos conocidos en la materia, tales como mutagénesis dirigida al emplazamiento, o mutagénesis de barrido de alanina (Cunningham y Wells, 1989, *Science* 244:1081-1085). En la última técnica, se introducen mutaciones únicas de alanina en cada resto de la molécula, y las moléculas mutantes resultantes se someten a ensayo para determinar la capacidad de producir 4-HBA a partir de L-Tirosina para identificar los restos de aminoácidos que son fundamentales para la actividad de la molécula. Véanse también, Hilton et al., 1996, *J. Biol. Chem.* 271: 4699-4708. El sitio activo de la enzima u otra interacción biológica también se puede determinar mediante el análisis físico de la estructura, según se determina por técnicas tales como resonancia magnética nuclear, cristalografía, difracción de electrones, o marcado de fotoafinidad, junto con la mutación de los aminoácidos del posible sitio de contacto. Véase, por ejemplo, de Vos et al., 1992, *Science* 255:306-312; Smith et al., 1992, *J. Mol. Biol.* 224:899-904; Wlodaver et al., 1992, *FEBS Lett.* 309: 59-64. La identidad de aminoácidos esenciales también se puede inferir a partir de un alineamiento con un polipéptido relacionado.

Se pueden llevar a cabo sustituciones, delecciones, y/o inserción de uno o varios aminoácidos, y someterse a ensayo usando métodos conocidos de mutagénesis, recombinación, y/o intercambio de cadenas, seguido por un procedimiento de cribado relevante, tales como los divulgados por Reidhaar-Olson y Sauer, 1988, *Science* 241:53-57; Bowie y Sauer, 1989, *Proc. Natl. Acad. Sci. USA* 86:2152-2156; documento WO 95/17413; o documento WO 95/22625. Otros métodos que se pueden usar incluyen la PCR propensa a error, expresión en fagos (por ejemplo, Lowman et al., 1991, *Biochemistry* 30: 10832-10837; patente de Estados Unidos n.º 5.223.409; documento WO 92/06204), y mutagénesis dirigida a la región (Derbyshire et al., 1986, *Gene* 46:145; Ner et al., 1988, *DNA* 7:127). Los métodos de mutagénesis/intercambio de cadenas se pueden combinar con métodos de cribado automatizados de alto rendimiento para detectar la actividad de polipéptidos clonados mutagenizados expresados por células hospedadoras (Ness et al., 1999, *Nature Biotechnology* 17:893-896). Las moléculas de ADN mutagenizado que codifican polipéptidos activos se pueden recuperar de las células hospedadoras y secuenciarse rápidamente usando métodos convencionales en la técnica. Estos métodos permiten la determinación rápida de la importancia de restos de aminoácidos individuales en un polipéptido.

*Conservado*: Por aminoácido conservado se pretende indicar que una secuencia definida está alineada con la secuencia de referencia y que los residuos de la secuencia definida correspondiente a la posición indicada en la secuencia de referencia son idénticos a los restos de aminoácidos de la secuencia de referencia. El alineamiento se puede llevar a cabo por cualquier método disponible, y en particular, por el método divulgado anteriormente para la determinación de la identidad, más preferentemente mediante Clustal Omega. La posición del residuo se indica en la secuencia de referencia.

*Purificar* o Purificado: Tal como se usa en el presente documento, los términos "purificar", "purificado". o "purificación" significan la eliminación o el aislamiento de una molécula de su entorno mediante, por ejemplo, aislamiento o separación. Las moléculas "prácticamente purificadas" están exentas en al menos aproximadamente un 60 %, preferentemente exentas en al menos aproximadamente un 75 %, y más preferentemente exentas en al menos aproximadamente un 90 % de otros componentes con los que están asociadas. Tal como se usa en el presente documento, estos términos se refieren también a la eliminación de los contaminantes de una muestra. Por ejemplo, la eliminación de los contaminantes puede dar como resultado un aumento en el porcentaje de alcanos/alquenos de una muestra. Por ejemplo, cuando los alcanos/alquenos se producen en una célula hospedadora, los alcanos/alquenos se pueden purificar por eliminación de las proteínas de la célula hospedadora. Tras la purificación, el porcentaje de alcanos/alquenos en la muestra aumenta. Los términos "purificar", "purificado". y "purificación" no exigen pureza absoluta. Son términos relativos. De este modo, por ejemplo, cuando se producen olefinas en células hospedadoras, un alcano/alqueno purificado es uno que está prácticamente separado de otros componentes celulares (por ejemplo, ácidos nucleicos, polipéptidos, lípidos, carbohidratos, u otros hidrocarburos). En otro ejemplo, una preparación purificada de alcanos/alquenos es una en la que los alcanos/alquenos están prácticamente exentos de contaminantes, tales como los que podrían estar presentes después de una fermentación. En algunas realizaciones, los alcanos/alquenos están purificados cuando al menos aproximadamente un 50 % en peso de una muestra está compuesta por los alcanos/alquenos. En otras realizaciones, una olefina está purificada cuando al menos aproximadamente un 60 %, 70 %, 80 %, 85 %, 90 %, 92 %, 95 %, 98 %, o 99 % o más en peso de una muestra está compuesta por los alcanos/alquenos.

Actividad ácido graso descarboxilasa:

Por "actividad ácido graso descarboxilasa" se denomina la eliminación del grupo ácido carboxílico del ácido graso, en particular, la eliminación directa sin un compuesto intermedio de aldehído y sin introducir una insaturación en el extremo. La actividad ácido graso descarboxilasa se puede medir por cualquier método disponible para la persona experta. Más preferentemente, la actividad se puede medir según el siguiente método.

En los ensayos enzimáticos *in vitro*, las reacciones se llevaron a cabo en viales precintados que contenían 100 nmoles de sustrato (de forma típica, ácido graso C16:0 libre) disuelto en el disolvente adecuado (etanol o dimetil sulfóxido), 2 nmoles de alcano C16 como patrón interno, y de 5 a 10 µg de enzima purificada (que contiene su cofactor FAD) con 500 µl del tampón de actividad (Tris-HCl 50 mM pH 8,2 con NaCl 100 mM) o un extracto de proteína total. Los viales se incubaron a 25°C en presencia de luz blanco (o cualquier luz que contenga fotones azules) a 2000 µmol.fotón.m<sup>-2</sup>.s<sup>-1</sup> en un agitador rotatorio a 250 rpm. A continuación, se inyectó hidróxido de sodio (10 µl a 10 M) en el vial para detener la reacción y las muestras se enfriaron sobre hielo. Los hidrocarburos se extrajeron con hexano y se cuantificaron mediante cromatografía de gases acoplada a un detector de ionización por llama y a espectrometría de masas (GC-FID-MS).

Polipéptido de ácido graso descarboxilasa:

La presente invención se refiere a un polipéptido que tiene una actividad ácido graso descarboxilasa, denominado en el presente documento "ácido graso descarboxilasa", y que comprende, consiste esencialmente o consiste en una secuencia que tiene al menos un 40 % de identidad con la SEQ ID No 1, SEQ ID No 5 o SEQ ID No 7, preferentemente la SEQ ID No 1.

Por ejemplo, la ácido graso descarboxilasa comprende, consiste esencialmente o consiste en una secuencia que tiene al menos un 42 %, 45 %, 50 % o 55 % de identidad con la SEQ ID No 1, SEQ ID No 5 o SEQ ID No 7, preferentemente la SEQ ID No 1.

En otro aspecto, la ácido graso descarboxilasa comprende, consiste esencialmente o consiste en una secuencia que tiene al menos un 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 % o 99 % de identidad con cualquiera de las SEQ ID No 1-3 y 5-14.

En un aspecto adicional, la ácido graso descarboxilasa comprende, consiste esencialmente o consiste en la secuencia de cualquiera de las SEQ ID No 1-3 y 5-14. En una realización preferida, la ácido graso descarboxilasa comprende, consiste esencialmente o consiste en la secuencia de cualquiera de las SEQ ID No 1, 5, 7-14, más preferentemente cualquiera de las SEQ ID No 1, 5, y 7.

En una realización, la ácido graso descarboxilasa comprende la secuencia de consenso G-X<sub>1</sub>-L-(X)<sub>4</sub>-C-[D/E]-X<sub>2</sub>-G-[A/G]-F-X<sub>3</sub>-[K/R/S/E] (SEQ ID No 4), siendo X un aminoácido (SEQ ID No 26). En una realización preferida, la ácido graso descarboxilasa comprende la secuencia de consenso G-X<sub>1</sub>-L-(X)<sub>4</sub>-C-[D/E]-X<sub>2</sub>-G-[A/G]-F-X<sub>3</sub>-[K/R] (SEQ ID No 4), siendo X un aminoácido.

Preferentemente, X<sub>1</sub> se puede seleccionar entre el grupo que consiste en P, L y G. Preferentemente, (X)<sub>4</sub> puede ser más específicamente [T/A]-[T/S/C]-[P/T/A]-[G/A]. Preferentemente, X<sub>2</sub> se puede seleccionar entre el grupo que consiste en H, N y R. Preferentemente, X<sub>3</sub> puede ser un aminoácido hidrófobo, especialmente seleccionado entre el grupo que consiste en L, V, A y F.

En una realización específica, la ácido graso descarboxilasa comprende la secuencia de consenso de G-X<sub>1</sub>-L-(X)<sub>4</sub>-C-[D/E]-X<sub>2</sub>-G-[A/G]-F-X<sub>3</sub>-[K/R/S/E] (SEQ ID No 26), en la que

- 5 - X<sub>1</sub> se puede seleccionar entre el grupo que consiste en P, L y G;  
 - (X)<sub>4</sub> puede ser más específicamente [T/A]-[T/S/C]-[P/T/A]-[G/A];  
 - X<sub>2</sub> se puede seleccionar entre el grupo que consiste en H, N y R; y  
 - X<sub>3</sub> puede ser un aminoácido hidrófobo, especialmente seleccionado entre el grupo que consiste en L, V, A y F.

10 En una realización muy específica, la ácido graso descarboxilasa comprende la secuencia consenso G-X<sub>1</sub>-L-(X)<sub>4</sub>-C-[D/E]-X<sub>2</sub>-G-[A/G]-F-X<sub>3</sub>-[K/R] (SEQ ID No 4), en la que

- X<sub>1</sub> se puede seleccionar entre el grupo que consiste en P, L y G;  
 - (X)<sub>4</sub> puede ser más específicamente [T/A]-[T/S/C]-[P/T/A]-[G/A];  
 15 - X<sub>2</sub> se puede seleccionar entre el grupo que consiste en H, N y R; y  
 - X<sub>3</sub> puede ser un aminoácido hidrófobo, especialmente seleccionado entre el grupo que consiste en L, V, A y F.

En otra realización adicional o preferida, la ácido graso descarboxilasa comprende aminoácidos conservados con referencia a la SEQ ID No 1. Basándose en la secuencia de referencia de la SEQ ID No 1, uno, dos, tres, cuatro, cinco  
 20 o todos los restos C372, R391, Y406, Q426, H512 y N515 de la SEQ ID No 1 están conservados. Lo mismo puede ser cierto para el correspondiente resto en las posiciones de otra ácido graso descarboxilasa correspondiente a dichas posiciones en la SEQ ID No 1.

En otra realización adicional o preferida, la ácido graso descarboxilasa comprende una región que forma un túnel  
 25 hidrófobo en el que el sustrato, el ácido graso libre, pueda entrar. Esta región está situada entre los restos en las posiciones 391 a 426 de la SEQ ID No 1. Por consiguiente, al menos un 40 % de los restos de aminoácido situados entre las posiciones 388-428 son restos hidrófobos, preferentemente seleccionados entre los grupos que consisten en V, I, L, M, F, W, C, A, e Y.

30 La ácido graso descarboxilasa tiene cualquier origen tal como de bacterias o algas.

Preferentemente, la ácido graso descarboxilasa procede de algas, preferentemente microalgas o cianobacterias. Por ejemplo, la ácido graso descarboxilasa es una GMC oxidorreductasa que tiene una actividad ácido graso  
 35 descarboxilasa procedente de *Chlorella*, *Chlamydomonas*, *Phaeodactylum*, *Coccomyxa*, *Volvox*, *Ectocarpus*, *Emiliania*, *Aureococcus*, *Chondrus*, *Galdieria* o *Nannochloropsis*. Se conocen muchas especies de microalgas y se pueden encontrar, por ejemplo, en la base de datos AlgaeBase ([www.algaebase.org/](http://www.algaebase.org/)).

En una realización particular, la ácido graso descarboxilasa es de *Chlorella variabilis*, en particular *Chlorella variabilis*  
 40 NC64A. En otra realización particular, la ácido graso descarboxilasa es de *Chlamydomonas reinhardtii*. En otra realización particular, la ácido graso descarboxilasa es de *Phaeodactylum tricornutum*. En otra realización particular, la ácido graso descarboxilasa es de *Coccomyxa subellipsoidea*, en particular *Coccomyxa subellipsoidea* C-169. En otra realización particular, la ácido graso descarboxilasa es de *Volvox carteri*. En otra realización particular, la ácido graso descarboxilasa es de *Ectocarpus siliculosus*. En otra realización particular, la ácido graso descarboxilasa es de *Emiliania huxleyi*. En otra realización particular, la ácido graso descarboxilasa es de *Aureococcus anophagefferens*.  
 45 En otra realización particular, la ácido graso descarboxilasa es de *Nannochloropsis gaditana*.

Preferentemente, la ácido graso descarboxilasa no tiene más de aproximadamente 750, 700, 650, o 600 aminoácidos de longitud.

50 La presente invención también se refiere a un polipéptido híbrido o polipéptido de fusión en el que la secuencia de aminoácidos de la ácido graso descarboxilasa como se ha definido anteriormente se fusiona con el extremo N o el extremo C de una región de otro polipéptido. La actividad ácido graso descarboxilasa permanece en el polipéptido híbrido o polipéptido de fusión. Preferentemente, la región de otro polipéptido se fusiona en el extremo N de la ácido graso descarboxilasa. Como alternativa, la región de otro polipéptido se fusiona en el extremo C de la ácido graso  
 55 descarboxilasa. Las técnicas para producir polipéptidos de fusión son conocidos en la materia, e incluyen ligar las secuencias de codificación que codifican la enzima y la región adicional de otro polipéptido de forma que estén en marco y dicha expresión del polipéptido de fusión está bajo el control de los mismos promotores y terminador. Los polipéptidos de fusión también se pueden construir usando tecnología de inteínas en las que los polipéptidos de fusión se crean después de la traducción (Cooper et al., 1993, EMBO J. 12: 2575-2583; Dawson et al., 1994, Science  
 60 266:776-779).

La región de adición del polipéptido de fusión se puede seleccionar para potenciar la estabilidad de la enzima de acuerdo con la presente divulgación, para fomentar la secreción (tal como un péptido señal hidrófobo en el extremo N) o la proteína de fusión de una célula (tal como una célula bacteriana o una célula de levadura) o para ayudar en la  
 65 purificación de la proteína de fusión. Más en particular, la repetición adicional puede ser una etiqueta útil para la purificación o la inmovilización de la enzima. Dicha etiqueta es bien conocida del experto en la materia, por ejemplo,

como una etiqueta His (His<sub>6</sub>), una etiqueta FLAG, una etiqueta HA (epítipo derivado de la proteína hemaglutinina de la gripe), una proteína de unión a maltosa (MPB), una etiqueta MYC (epítipo derivado de la protooncoproteína humana MYC), una etiqueta STREP o una etiqueta GST (glutatión S-transferasa pequeña). La repetición adicional puede ser tiorredoxina.

5 Un polipéptido de fusión puede comprender además un sitio de escisión entre la enzima y la región de adición. Tras la secreción o purificación de la proteína de fusión, el sitio se escinde liberando los dos polipéptidos. Los ejemplos de sitios de escisión incluyen, aunque no de forma limitativa, los sitios divulgados en Martin et al., 2003, J. Ind. Microbiol. Biotechnol. 3:568-576; Svetina et al., 2000, J. Biotechnol. 76:245-251; Rasmussen-Wilson et al., 1997, Appl. Environ. Microbiol. 63:3488-3493; Ward et al., 1995, Biotechnology 13:498-503; y Contreras et al., 1991, Biotechnology 9:378-381; Eaton et al., 1986, Biochemistry 25: 505-512; Collins-Racie et al., 1995, Biotechnology 13:982-987; Carter et al., 1989, Proteins: Structure, Function, and Genetics 6: 240-248; y Stevens, 2003, Drug Discovery World 4: 35-48. Por ejemplo, el sitio de escisión puede ser un sitio de escisión TEV (*virus del grabado del tabaco*). Otros sitios de escisión son bien conocidos del experto en la técnica.

15 En una realización particular, la presente invención se refiere a un polipéptido que comprende una etiqueta de histidina, una tiorredoxina, un sitio de escisión y la ácido graso descarboxilasa que se ha definido anteriormente (por ejemplo, véase la SEQ. ID No 3 para dicha construcción con la ácido graso descarboxilasa de *Chlorella*).

20 La presente invención se refiere además a una construcción o vector de ácido nucleico recombinante que comprende una secuencia de ácido nucleico que codifica la ácido graso descarboxilasa que se ha definido anteriormente. Más en particular, la construcción o vector de ácido nucleico es adecuada para expresar dicha ácido graso descarboxilasa. Además, se proporciona una célula hospedadora recombinante que comprende un ácido nucleico, una construcción de ácido nucleico recombinante o un vector recombinante que comprende una secuencia de ácido nucleico que codifica la ácido graso descarboxilasa que se ha definido anteriormente.

Ácidos nucleicos que codifican la ácido graso descarboxilasa y construcciones de ácido nucleico

30 La presente invención se refiere a un polinucleótido que codifica una ácido graso descarboxilasa de la presente invención. El ácido nucleico puede ser ADN (ADNc o ADNg), ARN, o una mezcla de los dos. Puede estar en forma monocatenaria o en forma de duplete o una mezcla de los dos. Puede comprender nucleótidos modificados, que comprenden por ejemplo un enlace modificado, una base de purina o pirimidina modificada, o un azúcar modificado. Se puede preparar por cualquier método conocido en la materia, incluidas síntesis química, recombinación, y mutagénesis.

35 Opcionalmente, la secuencia codificante se puede optimizar para su expresión en la célula hospedadora. En particular, como las ácido graso descarboxilasas de la presente invención proceden de algas, en particular microalgas, la secuencia de ácido nucleico que codifica la enzima se puede optimizar para un hospedador bacteriano, tal como *E. coli*, una hospedadora de levadura, o incluso una microalga que sea diferente de aquella donde se origina la enzima.

40 La presente invención, también se refiere a construcciones de ácidos nucleicos que comprenden un polinucleótido que codifica una ácido graso descarboxilasa de acuerdo con la presente divulgación unido de manera operativa a una o más secuencias de control que dirigen la expresión de la secuencia de codificación en una célula hospedadora adecuada en condiciones compatibles con las secuencias de control. Un polinucleótido se puede manipular de diferentes formas para proporcionar la expresión de la ácido graso descarboxilasa. La manipulación del polinucleótido antes de su inserción en un vector puede ser deseable o necesaria dependiendo del vector de expresión. Las técnicas para modificador polinucleótidos usando métodos de ADN recombinante son bien conocidas en la materia.

45 La secuencia de control puede incluir un promotor que se reconoce una célula hospedadora o un sistema de expresión *in vitro* para la expresión de un polinucleótido que codifica una ácido graso descarboxilasa de la presente invención. El promotor contiene secuencias de control de la transcripción que median en la expresión de la ácido graso descarboxilasa. El promotor puede ser cualquier polinucleótido que muestra actividad transcripcional en la célula hospedadora, incluidos promotores mutantes, truncados e híbridos, y se puede obtener a partir de genes que codifican polipéptidos extracelulares o intracelulares tanto homólogos como heterólogos respecto de la célula hospedadora. Opcionalmente, el promotor puede ser inducible. Opcionalmente, el promotor es un promotor fuerte que permite la expresión en exceso de la ácido graso descarboxilasa. Opcionalmente, el promotor es un promotor inducible fuerte.

50 Los ejemplos de promotores adecuados en una célula hospedadora bacteriana son los promotores obtenidos a partir del gen de la alfa amilasa de *Bacillus amyloliquefaciens* (amyQ), gen de la alfa amilasa de *Bacillus licheniformis* (amyL), gen de la alfa penicilinasas de *Bacillus licheniformis* (penP), gen de la amilasa maltogénica de *Bacillus stearothermophilus* (amyM), gen de la levansucrasa de *Bacillus subtilis* (sacB), genes xylA y xylB de *Bacillus subtilis*, gen cryIIA de *Bacillus thuringiensis* (Agaisse y Lereclus, 1994, Molecular Microbiology 13: 97-107), operón lac de *E. coli*, promotor tcr de *E. coli* (Egon et al., 1988, Gene 69: 301-315), gen de la agarasa de *Streptomyces coelicolor* (dagA), y gen de la beta-lactamasa procariota (Villa-Kamaroff et al., 1978, Proc. Natl. Acad. Sci. USA 75: 3727-3731), así como el promotor tac (DeBoer et al., 1983, Proc. Natl. Acad. Sci. USA 80:21-25. Otros promotores se describen en "Useful proteins from recombinant bacteria" en Gilbert et al., 1980, Scientific American 242:74-94; y en Sambrook

et al., 1989. Los ejemplos de promotores en tándem se divulgan el documento WO 99/43835.

Los ejemplos de promotores adecuados para células hospedadoras de hongos filamentosos son los promotores obtenidos a partir de los genes de acetamidasa de *Aspergillus nidulans*, alfa-amilasa neutra de *Aspergillus niger*, alfa-amilasa estable en medio ácido de *Aspergillus niger*, glucoamilasa (glaA) de *Aspergillus niger* o *Aspergillus awamori*, amilasa TAKA de *Aspergillus oryzae*, proteasa alcalina de *Aspergillus oryzae*, triosa fosfato isomerasa de *Aspergillus oryzae*, proteasa análoga a tripsina de *Fusarium oxysporum* (documento WO96/00787), amiloglucosidasa de *Fusarium venenatum* (documento WO 00/56900), *Fusarium venenatum* Daria (WO 00/56900), *Fusarium venenatum* Quinn (WO 00/56900), lipasa de *Rhizomucor miehei*, aspártico proteinasa de *Rhizomucor miehei*, beta-glucosidasa de *Trichoderma reesei*, celobiohidrolasa I de *Trichoderma reesei*, celobiohidrolasa II de *Trichoderma reesei*, endoglucanasa I de *Trichoderma reesei*, endoglucanasa II de *Trichoderma reesei*, endoglucanasa III de *Trichoderma reesei*, endoglucanasa IV de *Trichoderma reesei*, endoglucanasa V de *Trichoderma reesei*, xilanasa I de *Trichoderma reesei*, xilanasa II de *Trichoderma reesei*, beta-xilosidasa de *Trichoderma reesei*, así como el promotor NA2-tpi (un promotor modificado a partir de un gen de la alfa-amilasa neutra de *Aspergillus* en el que el líder no traducido se ha sustituido por un líder no traducido de un gen de la triosa fosfato isomerasa de *Aspergillus*; los ejemplos no limitantes incluyen promotores modificados a partir de un gen de la alfa-amilasa neutra de *Aspergillus niger* en el que el líder no traducido se ha sustituido por un líder no traducido de un gen de la triosa fosfato isomerasa de *Aspergillus nidulans* o *Aspergillus oryzae*; y promotores mutantes, truncados, e híbridos de los mismos.

En un hospedador de levadura, los promotores útiles se obtienen a partir de los genes de la enolasa de *Saccharomyces cerevisiae* (ENO-1), galactoquinasa de *Saccharomyces cerevisiae* (GAL1), alcohol deshidrogenasa/gliceraldehído-3-fosfato deshidrogenasa de *Saccharomyces cerevisiae* (ADH1, ADH2/GAP), triosa fosfato isomerasa de *Saccharomyces cerevisiae* (TPI), metalotioneína de *Saccharomyces cerevisiae* (CUP1), y 3-fosfoglicerato quinasa de *Saccharomyces cerevisiae*. Otros promotores de utilidad para las células hospedadoras de levadura se describen en Romanos et al., 1992, Yeast 8: 423-488.

Un promotor inducible puede ser sensible a, por ejemplo, la intensidad de la luz o a alta o baja temperatura, y/o puede ser sensible a compuestos específicos. El promotor inducible puede ser, por ejemplo, un promotor sensible a hormonas (por ejemplo, un promotor sensible a ecdisona, tal como se describe en el documento EP 6.379.945), un promotor de la metalotioneína (describe en el documento US 6.410.828), un promotor relacionado con la patogénesis (PR) que puede ser sensible a una sustancia química tal como, por ejemplo, ácido salicílico, etileno, timina, y/o BTH (documento US 5.689.044), o similares, o alguna combinación de los mismos. Un promotor inducible también puede ser sensible a la luz o la oscuridad (documento US 8.318.482; US 5.750.385; US 5.639.952), metales (Eukaryotic Cell 2:995-1002 (2003)) o temperatura (US 5.447.858; Abe et al. Plant Cell Physiol. 49: 625-632 (2008); Shroda et al. Plant J. 21: 121-131 (2000)). Los ejemplos anteriores son no limitantes de los tipos de promotores o promotores específicos que se pueden utilizar. La secuencia promotora puede ser de cualquier organismo, siempre que sea funcional en el organismo hospedador. En determinadas realizaciones, los promotores inducibles están formados por la fusión de una o más porciones o dominios de un promotor inducible conocido con al menos una porción de un promotor diferente que puede operar en la célula hospedadora, por ejemplo, para transmitir inducibilidad a un promotor que opera en la especie hospedadora.

Un experto en la materia apreciará rápidamente que varias secuencias promotoras conocidas se pueden desplegar de forma útil para especies de microalgas. Por ejemplo, los promotores habitualmente utilizados para impulsar la expresión de transgenes en microalgas incluyen varias versiones del promotor 35S del virus del mosaico de la coliflor (CaMV35S), que se ha usado tanto en dinoflagelados como en clorofitas (Chow et al, Plant Cell Rep., 18:778-780, 1999; Jarvis y Brown, Curr. Genet., 317-321, 1991; Lohuis y Miller, Plant J., 13:427-435, 1998). El promotor SV40 del virus de simio también se ha notificado como activo en algunas algas (Gan et al., J. Appl. Phycol., 151, 345-349, 2003; Qin et al, Hydrobiologia 398-399, 469-472, 1999). Los promotores de RBCS2 (ribulosa bifosfato carboxilasa, subunidad pequeña) (Fuhrmann et al, PlantJ., 19:353-361,1999) y Psad (proteína abundante del complejo del fotosistema I; Fischer y Rochaix, FEBS Lett. 581:5555-5560, 2001) de *Chlamydomonas* también pueden ser de utilidad. Los promotores de fusión de HSP70A/RBCS2 y HSP70A/p2TUB (tubulina) (Schroda et al, Plant J., 21:121-131, 2000) también pueden ser de utilidad para una expresión mejorada de los transgenes, donde el promotor HSP70A puede servir como activador transcripcional cuando se coloca antes de otros promotores. La expresión de alto nivel de un gen de interés también se puede conseguir en, por ejemplo, especies de diatomeas, bajo el control de un promotor de un gen fcp que codifica una proteína de unión a fucoxantina-clorofila a/b de diatomea (Falciatore et al, Mar. Biotechnol, 1:239-251, 1999; Zaslavskaja et al, J. Phycol. 36:379-386, 2000) o el gen vcp que codifica una proteína de unión a eustigmatofito violaxantina-clorofila a/b (véase la Patente de Estados Unidos n.º 8.318.482). Si así se desea, los promotores inducibles pueden proporcionar una expresión rápida y bien controlada de genes en microalgas transgénicas. Por ejemplo, las regiones promotoras de los genes NR que codifican la nitrato reductasa se pueden usar como dichos promotores inducibles. La actividad del promotor NR se suprime, de forma típica, mediante amonio, y se induce cuando el amonio se sustituye por nitrato (Poulsen y Kroger, FEBS Lett 272:3413-3423, 2005), por tanto, la expresión génica se puede activar o desactivar cuando las células de microalgas se hacen crecer en presencia de amonio/nitrato. Se conocen varios promotores de algas, y se pueden usar, incluidos los divulgados en los documentos US 2013/0023035; US 2013-0323780, US 2014-0154806; y US 2014/-0363892.

La secuencia de control puede ser también un terminador de la transcripción, que se reconoce por una célula

hospedadora para finalizar la transcripción. El terminador está unido de manera operativa con el extremo 3' del polinucleótido que codifica el polipéptido. En la presente invención se puede utilizar cualquier terminador que sea funcional en la célula hospedadora.

- 5 Los terminadores preferidos para células hospedadoras bacterianas se obtiene de los genes de la proteasa alcalina de *Bacillus clausii* (aprH), alfa amilasa de *Bacillus licheniformis* (amyL), ARN ribosómico de *Escherichia coli* (rrnB).

Los terminadores preferidos para células hospedadoras de hongos filamentosos se obtienen de los genes de la antranilato sintasa de *Aspergillus nidulans*, glucoamilasa de *Aspergillus niger*, alfa-glucosidasa de *Aspergillus niger*,  
10 amilasa TAKA de *Aspergillus oryzae*, y proteasa de tipo tripsina de *Fusarium oxysporum*.

Los terminadores preferidos para células hospedadoras de levadura se obtienen de los genes de la enolasa de *Saccharomyces cerevisiae*, citocromo C de *Saccharomyces cerevisiae* (CYC1), y gliceraldehído-3-fosfato deshidrogenasa de *Saccharomyces cerevisiae*. Otros terminadores de utilidad para las células hospedadoras de  
15 levadura se describen en Romanos et al., 1992, anteriormente citado.

La secuencia de control también puede ser una región estabilizante del ARNm después de un promotor y antes de la secuencia de codificación de un gen que aumenta la expresión del gen.

- 20 Los ejemplos de regiones estabilizantes del ARNm adecuadas se obtienen a partir de un gen cryIIIA de *Bacillus thuringiensis* (documento WO 94/25612) y un gen SP82 de *Bacillus subtilis* (Hue et al, 1995, Journal of Bacteriology 177: 3465-3471).

La secuencia de control puede ser también una secuencia líder, una región no traducida de un ARNm que es importante para la traducción por la célula hospedadora. La secuencia líder está unida de manera operativa con el extremo 5' del polinucleótido que codifica la ácido graso descarboxilasa. Se puede utilizar cualquier líder que sea funcional en la célula hospedadora.

- 30 Los líderes preferidos para las células hospedadoras de hongos filamentosos se obtienen ge los genes de la amilasa TAKA de *Aspergillus oryzae* y la triosa fosfato isomerasa de *Aspergillus nidulans*.

Los líderes adecuados para células hospedadoras de levaduras se obtienen de los genes de la enolasa de *Saccharomyces cerevisiae* (ENO-1), 3-fosfoglicerato quinasa de *Saccharomyces cerevisiae*, factor alfa de *Saccharomyces cerevisiae*, y alcohol deshidrogenasa/gliceraldehído-3-fosfato deshidrogenasa de *Saccharomyces cerevisiae* (ADH2/GAP).  
35

La secuencia de control puede ser también una secuencia de poliadenilación, una secuencia unida de manera operativa con el extremo 3' del polinucleótido que codifica la ácido graso descarboxilasa y, cuando se transcribe, se reconoce por la célula hospedadora como una señal para añadir restos de poliadenosina al ARNm transcrito. Se puede utilizar cualquier secuencia de poliadenilación que sea funcional en la célula hospedadora.  
40

Las secuencias de poliadenilación preferidas para células hospedadoras de hongos filamentosos se obtienen de los genes de la antranilato sintasa de *Aspergillus nidulans*, glucoamilasa de *Aspergillus niger*, alfa-glucosidasa de *Aspergillus niger*, amilasa TAKA de *Aspergillus oryzae*, y proteasa de tipo tripsina de *Fusarium oxysporum*.  
45

Las secuencias de poliadenilación útiles para las células hospedadoras de levaduras se describen en Guo y Sherman, 1995, Mol. Cellular Biol. 15: 5983-5990.

La secuencia de control también puede ser un péptido señal de la región de codificación que codifica un péptido señal unido al extremo N de la ácido graso descarboxilasa y dirige la ácido graso descarboxilasa a la ruta secretora de la célula. El extremo 5' de la secuencia de codificación del polinucleótido puede contener inherentemente una secuencia de codificación de un péptido señal unida de forma natural en el marco de lectura de traducción con el segmento de la región codificante que codifica la ácido graso descarboxilasa. Como alternativa, el extremo 5' de la secuencia de codificación puede contener una secuencia de codificación de un péptido señal que sea extraña para la secuencia de codificación.  
50  
55 Una secuencia de codificación de un péptido señal extraña puede ser necesaria cuando la secuencia de codificación no contiene de forma natural una secuencia de codificación del péptido señal. Como alternativa, una secuencia de codificación de un péptido señal extraña puede sustituir simplemente la secuencia de codificación de un péptido señal natural para potenciar la secreción del polipéptido. Sin embargo, se pude utilizar cualquier secuencia de codificación de un péptido señal que dirija el polipéptido expresado a la ruta secretora de la célula hospedadora. El péptido señal también puede ser un péptido transitorio de cloroplasto, el péptido transitorio de cloroplasto de la ácido graso descarboxilasa o cualquier otro péptido transitorio de cloroplasto.  
60

Las secuencias de codificación de un péptido señal eficaces para células hospedadoras bacterianas son las secuencias de codificación de un péptido señal obtenidas de los genes de la amilasa maltogénica de *Bacillus* NCIB 1 1837, subtilisina de *Bacillus licheniformis*, beta-lactamasa de *Bacillus licheniformis*, alfa-amilasa de *Bacillus stearothermophilus*, proteasas neutras de *Bacillus stearothermophilus* (nprT, nprS, nprM), y prsA de *Bacillus subtilis*.  
65

Otros péptidos señal se describen en Simonen y Palva, 1993, *Microbiological Reviews* 57:109-137.

Las secuencias de codificación del péptido señal eficaces para células hospedadoras de hongos filamentosos son las secuencias de codificación del péptido señal obtenidas de los genes de la amilasa neutra de *Aspergillus niger*, glucoamilasa de *Aspergillus niger*, amilasa TAKA de *Aspergillus oryzae*, celulasa de *Humicola insolens*, endoglucanasa V de *Humicola insolens*, lipasa de *Humicola lanuginosa*, y proteinasa aspártica de *Rhizomucor miehei*.

Los péptidos señal útiles para las células hospedadoras de levaduras se obtienen de los genes del factor alfa de *Saccharomyces cerevisiae* y la invertasa de *Saccharomyces cerevisiae*. Otras secuencias de codificación del péptido señal adecuadas se describen en Romanos et al., 1992, *anteriormente citado*.

También puede ser deseable añadir secuencias reguladoras que regulen la expresión del polipéptido con respecto al crecimiento de la célula hospedadora. Los ejemplos de sistemas reguladores son aquellos que producen la activación o desactivación de la expresión génica en respuesta a un estímulo químico o físico, incluida la presencia de un compuesto regulador. Los sistemas reguladores en sistemas procariotas incluyen los sistemas de operadores lac, tac, y trp. En levaduras, se puede usar el sistema ADH2 o el sistema GAL1. En hongos filamentosos, el promotor de la glucoamilasa de *Aspergillus niger*, el promotor de la alfa amilasa TAKA de *Aspergillus oryzae*, y el promotor de la glucoamilasa *Aspergillus oryzae* son de utilidad. Otros ejemplos de secuencias reguladoras son las que permiten la amplificación génica. En sistemas eucariotas, estas secuencias reguladoras incluyen el gen de la dihidrofolato reductasa que se amplifica en presencia de metotrexato, y los genes de la metalotioneína que se amplifican con metales pesados. En estos casos, el polinucleótido que codifica el polipéptido estaría unido de manera operativa con la secuencia reguladora.

### 25 Vectores de expresión

La presente invención también se refiere a vectores de expresión recombinante que comprenden una construcción de ácidos nucleicos como se ha divulgado anteriormente, o un polinucleótido que codifica un ácido graso descarboxilasa de la presente invención, un promotor, y señales de inicio y finalización de la transcripción. Las diferentes secuencias de nucleótidos y de control se pueden unir entre sí para producir un vector de expresión recombinante que puede incluir uno o más sitios de restricción convenientes para permitir la inserción o sustitución del polinucleótido que codifica la ácido graso descarboxilasa en dichos sitios. Como alternativa, el polinucleótido se puede expresar insertando el polinucleótido o una construcción de ácidos nucleicos que comprende el polinucleótido en un vector de expresión adecuado.

Para crear el vector de expresión, la secuencia de codificación se ubica en el vector de forma que la secuencia de codificación esté unida de manera operativa a las secuencias de control adecuadas para la expresión.

El vector de expresión recombinante puede ser cualquier vector (por ejemplo, un plásmido o virus) que se puede someter convenientemente a procedimientos de ADN recombinante y puede llevar a cabo la expresión del polinucleótido. La elección del vector dependerá de forma típica de la compatibilidad del vector con la célula hospedadora en la que se introduce el vector. El vector puede ser un plásmido lineal o circular cerrado.

El vector puede ser un vector de replicación autónoma, es decir, un vector que existe como entidad extracromosómica, cuya replicación es independiente de la replicación cromosómica, por ejemplo, un plásmido, un elemento extracromosómico, un minicromosoma, o un cromosoma artificial. El vector puede contener cualquier medio para garantizar la autorreplicación. Como alternativa, el vector puede ser uno que, cuando se introduce en la célula hospedadora, se integra en el genoma y se replica junto con el uno o varios cromosomas donde se ha integrado. Adicionalmente, un único vector o plásmido o dos o más vectores o plásmidos que conjuntamente contienen el ADN total se puede introducir en el genoma de la célula hospedadora, o un se puede usar un transposón.

El vector contiene preferentemente uno o más marcadores seleccionables que permiten la selección sencilla de las células transformadas, transfectadas, transducidas, o similares. Un marcador seleccionable es un gen cuyo producto produce resistencia a biocidas o virus, resistencia a metales pesados, prototrofia a auxotrofia, y similares.

Los ejemplos de marcadores bacterianos seleccionables son genes o marcadores de *Bacillus licheniformis* o *Bacillus subtilis* que transmiten resistencia a antibióticos tales como ampicilina, cloranfenicol, kanamicina, neomicina, espectinomycin, tetraciclina. Los marcadores adecuados para células hospedadoras de levaduras incluyen, aunque no de forma limitativa, ADE2, HIS3, LEU2, LYS2, MET3, TRP1, y URA3. Los marcadores seleccionables para su uso en una célula hospedadora de hongo filamentoso incluyen, aunque no de forma limitativa, amdS (acetamidasa), argB (ornitina carbamoiltransferasa), bar (fosfinotricina acetiltransferasa), hph (higromicina fosfotransferasa), niaD (nitrato reductasa), pyrG (orotidina-5'-fosfato descarboxilasa), sC (sulfato adeniltransferasa), y trpC (antranilato sintasa), así como equivalentes de las mismas. Preferidos para su uso en una célula de *Aspergillus* son los genes amdS y pyrG de *Aspergillus nidulans* o *Aspergillus oryzae* y un gen de *Streptomyces hygroscopicus*.

El vector contiene preferentemente uno o más elementos que permiten la integración del vector en el genoma de la célula hospedadora o la replicación autónoma del vector en la célula independiente del genoma.

5 Cuando se produce la integración en el genoma de la célula hospedadora, la integración de las secuencias en el genoma puede basarse en la recombinación homóloga o no homóloga. Como alternativa, el vector puede contener polinucleótidos adicionales para dirigir la integración mediante recombinación homóloga en el genoma de la célula hospedadora en una o varias localizaciones precisas en uno o varios cromosomas. Para aumentar la posibilidad de integración en una localización precisa, los elementos de integración deberán contener un número suficiente de ácidos nucleicos, tales como de 100 a 10.000 pares de bases, de 400 a 10.000 pares de bases, y de 800 a 10.000 pares de bases, que tienen un elevado grado de identidad de secuencia con la correspondiente secuencia diana para potenciar la probabilidad de la recombinación homóloga. Los elementos de integración puede ser cualquier secuencia que es homóloga con la secuencia diana en el genoma de la célula hospedadora. Adicionalmente, los elementos de integración pueden ser polinucleótidos no codificantes o codificantes. Por otro lado, el vector se puede integrar en el genoma de la célula hospedadora mediante recombinación no homóloga.

15 Para la replicación autónoma, el vector puede comprender además un origen de replicación que permite al vector replicarse de forma autónoma en la célula hospedadora en cuestión. El origen de replicación puede ser cualquier replicador plásmido que media la replicación autónoma que funciona en una célula. La expresión "origen de replicación" o "replicador plásmido" significa un polinucleótido que permite a un plásmido o vector para replicarse in vivo. Los ejemplos de orígenes de replicación bacterianos son los orígenes de replicación de los plásmidos pBR322, pUC19, pACYC177, y pACYC184 que permiten la replicación de *E. coli*, y pUB1 10, pE194, pTA1060, y pAMβ1 que permiten la replicación en *Bacillus*. Los ejemplos de orígenes de replicación para su uso en una célula hospedadora de levaduras son el origen de replicación de 2 micrómetros, ARS1, ARS4, la combinación de ARS1 y CEN3, y la combinación de ARS4 y CEN6. Los ejemplos de orígenes de replicación útiles en una célula de hongo filamentoso son AMA1 y ANS1 (Gems et al., 1991, Gene 98: 61-67; Cullen et al., 1987, Nucleic Acids Res. 15:9163-9175; documento WO 00/24883). El aislamiento del gen AMA1 y la construcción de plásmidos o vectores que comprenden el gen se puede llevar a cabo según los métodos descritos en el documento WO 00/24883.

30 Se puede introducir más de una copia de un polipéptido de la presente invención en una célula hospedadora para aumentar la producción de un polipéptido. Un aumento en el número de copias del polinucleótido se puede obtener mediante la integración de al menos una copia adicional de la secuencia en el genoma de la célula hospedadora o mediante la inclusión de un gen marcador seleccionable amplificable con el polinucleótido donde las células contienen copias amplificadas del gen marcador seleccionable y, por tanto, copias adicionales del polinucleótido, se puede seleccionar mediante el cultivo de las células en presencia del agente seleccionable adecuado.

35 Los procedimientos utilizados para ligar los elementos descritos anteriormente para construir los vectores de expresión recombinante de la presente invención son bien conocidos del experto en la materia (véase, por ejemplo, Sambrook et al., 1989, supra).

### Células hospedadoras

40 Se divulga una célula hospedadora recombinante que expresa una ácido graso descarboxilasa de la presente invención, más específicamente una célula hospedadora recombinante diseñada mediante ingeniería genética para producir alcanos/alquenos.

45 Se divulga una célula hospedadora recombinante que puede expresar una ácido graso descarboxilasa endógena pero con una expresión en exceso de la misma (por ejemplo, mediante el control de la expresión mediante un promotor heterólogo fuerte y/o mediante el aumento en el número de copias que codifica la ácido graso descarboxilasa en la célula).

50 Como alternativa, también se divulga una célula hospedadora recombinante que puede expresar una ácido graso descarboxilasa heteróloga de la presente invención.

55 En un aspecto particular, la invención se refiere a una célula recombinante diseñada mediante ingeniería genética para producir alcanos/alquenos y comprende una secuencia de ácido nucleico que codifica un polipéptido heterólogo que tiene una actividad ácido graso descarboxilasa y que comprende una secuencia que tiene al menos un 95 % de identidad con una secuencia de aminoácidos seleccionada entre las SEQ ID No 1-3 y 5-14.

60 Además, o como otra alternativa, la célula hospedadora recombinante se ha diseñado mediante ingeniería genética para aumentar la producción de ácidos grasos o favorecer la producción de ácidos grasos preferenciales. Por ejemplo, la célula hospedadora puede ser una bacteria (incluida una bacteria fotosintética), un hongo filamentoso, una levadura o una microalga de una especie diferente a la especie en la que se origina la ácido graso descarboxilasa.

65 Por consiguiente, también se divulgan células hospedadoras recombinantes que comprenden un polinucleótido que codifica una ácido graso descarboxilasa de acuerdo con la presente divulgación unido de manera operativa a una o más secuencias de control que dirigen la producción de la ácido graso descarboxilasa de la presente invención. Una construcción o vector que comprende un polinucleótido que codifica una ácido graso descarboxilasa de acuerdo con la presente divulgación se introduce en una célula hospedadora de tal forma que la construcción o el vector se

5 mantiene como integrante cromosómico o un vector extracromosómico autorreplicante como se ha descrito anteriormente. En el caso de células hospedadoras recombinantes de acuerdo con la invención, dicha construcción o vector comprende un polinucleótido que codifica una ácido graso descarboxilasa que comprende una secuencia que tiene al menos un 95 % de identidad con una secuencia de aminoácidos seleccionada entre las SEQ ID No 1-3 y 5-14.

10 La expresión "célula hospedadora" abarca cualquier progenie de una célula precursora que no sea idéntica a la célula precursora debido a mutaciones que se producen durante la replicación. La elección de una célula hospedadora, en gran medida, dependerá del gen que codifica el polipéptido y su fuente.

La célula hospedadora puede ser cualquier célula útil en la producción recombinante de una ácido graso descarboxilasa de la presente invención, por ejemplo, una procariota o una eucariota.

15 La célula hospedadora procariota puede ser cualquier bacteria Gram-positiva o Gram-negativa. Las bacterias Gram-positivas incluyen, aunque no de forma limitativa, *Bacillus*, *Clostridium*, *Enterococcus*, *Geobacillus*, *Lactobacillus*, *Lactococcus*, *Oceanobacillus*, *Staphylococcus*, *Streptococcus*, y *Streptomyces*. Las bacterias Gram-negativas incluyen, aunque no de forma limitativa, *Campylobacter*, *E. coli*, *Flavobacterium*, *Fusobacterium*, *Helicobacter*, *Ilyobacter*, *Neisseria*, *Pseudomonas*, *Salmonella*, y *Ureaplasma*. La célula bacteriana hospedadora puede ser cualquier célula de *Bacillus* incluidas, aunque no de forma limitativa, *Bacillus alkalophilus*, *Bacillus amyloliquefaciens*, *Bacillus brevis*, *Bacillus circulans*, *Bacillus clausii*, *Bacillus coagulans*, *Bacillus firmus*, *Bacillus lautus*, *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus megaterium*, *Bacillus pumilus*, *Bacillus stearothermophilus*, *Bacillus subtilis*, y *Bacillus thuringiensis*. La célula bacteriana hospedadora puede ser también cualquier célula de *Streptococcus* incluida, aunque no de forma limitativa, células de *Streptococcus equisimilis*, *Streptococcus pyogenes* *Streptococcus uberis*, *Streptococcus equi* y *Streptococcus zooepidemicus*. La célula bacteriana hospedadora puede ser adicionalmente cualquier célula de *Streptomyces* incluidas, aunque no de forma limitativa, células de *Streptomyces achromogenes*, *Streptomyces avermitilis*, *Streptomyces coelicolor*, *Streptomyces griseus*, y *Streptomyces lividans*. En una realización particular, la célula hospedadora es una cianobacteria. Las cianobacterias incluyen, aunque no de forma limitativa, especies de los géneros *Synechococcus*, *Synechocystis*, *Anabaena*, *Spirulina*.

30 La introducción de ADN en una célula de *Bacillus* se puede ver afectada mediante transformación de protoplastos (véase, por ejemplo, Chang y Cohen, 1979, Mol. Gen. Genet. 168: 111-115), transformación de células competentes (véase, por ejemplo, Young y Spizizen, 1961, J. Bacteriol. 81: 823-829, o Dubnau y Davidoff-Abelson, 1971, J. Mol. Biol. 56: 209-221), electroporación (véase, por ejemplo, Shigekawa y Dower, 1988, Biotechniques 6: 742-751), o conjugación (véase, por ejemplo, Koehler y Thorne, 1987, J. Bacteriol. 169: 5271-5278). La introducción de ADN en una célula de *E. coli* se puede realizar mediante transformación de protoplastos (véase, por ejemplo, Hanahan, 1983, J. Mol. Biol. 166: 557-580), o electroporación (véase, por ejemplo, Dower et al, 1988, Nucleic Acids Res. 16: 6127-6145). La introducción de ADN en una célula de *Streptomyces* se puede llevar a cabo mediante transformación de protoplastos, electroporación (véase, por ejemplo, Gong et al., 2004, Folia Microbiol. (Praha) 49: 399-405), conjugación (véase, por ejemplo, Mazodier et al., 1989, J. Bacteriol. 171: 3583-3585), o transducción (véase, por ejemplo, Burke et al., 2001, Proc. Natl. Acad. Sci. USA 98:6289-6294). La introducción de ADN en una célula de *Pseudomonas* se puede llevar a cabo mediante electroporación (véase, por ejemplo, Choi et al., 2006, J. Microbiol. Methods 64: 391 - 397) o conjugación (véase, por ejemplo, Pinedo y Smets, 2005, Appl. Environ. Microbiol. 71: 51-57). La introducción de ADN en una célula de *Streptococcus* se puede llevar a cabo mediante competencia natural (véase, por ejemplo, Perry y Kuramitsu, 1981, Infect. Immun. 32:1295-1297), transformación de protoplastos (véase, por ejemplo, Catt y Jollick, 1991, Microbios 68:189-207), electroporación (véase, por ejemplo, Buckley et al., 1999, Appl. Environ. Microbiol. 65:3800-3804), o conjugación (véase, por ejemplo, Clewell, 1981, Microbiol. Rev. 45:409-436). Sin embargo, se puede utilizar cualquier método conocido en la materia para introducir el ADN en una célula hospedadora.

50 La célula hospedadora también puede ser una célula eucariota, tal como una célula de mamífero, insecto, planta u hongo. La célula hospedadora puede ser una célula fúngica. "Hongos" tal como se usa en el presente documento, son los filos *Ascomycota*, *Basidiomycota*, *Chytridiomycota*, y *Zygomycota*, así como *Oomycota* y todos los hongos mitospóricos (como se define en Hawksworth et al., En, Ainsworth and Bisby's Dictionary of The Fungi, 8ª edición, 1995, CAB International, University Press, Cambridge, Reino Unido). La célula hospedadora fúngica puede ser una célula de levadura. "Levadura" tal como se usa en el presente documento incluye levaduras ascospóricas (Endomicetales), levaduras basidiosporógenas, y levaduras que pertenecen a los *Fungi imperfecti* (Blastomicetos). Puesto que la clasificación de las levaduras puede variar en el futuro, para los fines de la presente invención, las levaduras deberán definirse como se describe en Biology and Activities of Yeast (Skinner, Passmore, y Davenport, editores, Soc. App. Bacteriol. Symposium Series N.º 9, 1980). La célula hospedadora de levaduras puede ser una célula de *Candida*, *Hansenula*, *Kluyveromyces*, *Pichia*, *Saccharomyces*, *Schizosaccharomyces*, o *Yarrowia*, tales como una célula de *Kluyveromyces lactis*, *Saccharomyces carlsbergensis*, *Saccharomyces cerevisiae*, *Saccharomyces diastaticus*, *Saccharomyces douglasii*, *Saccharomyces kluyveri*, *Saccharomyces norbensis*, *Saccharomyces oviformis*, o *Yarrowia lipolytica*. La célula hospedadora fúngica puede ser una célula de hongo filamentoso. "Hongos filamentosos" incluyen todas las formas filamentosas de la subdivisión *Eumycota* y *Oomycota* (tal como se define en Hawksworth et al., 1995, supra). Los hongos filamentosos se caracterizan generalmente por una pared micelular compuesta de quitina, celulosa, glucano, quitosán, manano, y otros polisacáridos complejos. La célula hospedadora de hongos filamentosos puede ser una célula de *Acremonium*, *Aspergillus*, *Aureobasidium*,

*Bjerkandera*, *Ceriporiopsis*, *Chrysosporium*, *Coprinus*, *Coriolus*, *Cryptococcus*, *Filibasidium*, *Fusarium*, *Humicola*, *Magnaporthe*, *Mucor*, *Myceliophthora*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, *Penicillium*, *Phanerochaete*, *Phlebia*, *Piromyces*, *Pleurotus*, *Schizophyllum*, *Talaromyces*, *Thermoascus*, *Thielavia*, *Tolypocladium*, *Trametes*, o *Trichoderma*. Por ejemplo, la célula hospedadora de hongos filamentosos puede ser una célula de *Aspergillus awamori*, *Aspergillus foetidus*, *Aspergillus fumigatus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger*, *Aspergillus oryzae*, *Bjerkandera adusta*, *Ceriporiopsis aneirina*, *Ceriporiopsis caregiea*, *Ceriporiopsis gilvescens*, *Ceriporiopsis pannocinta*, *Ceriporiopsis rivulosa*, *Ceriporiopsis subrufa*, *Ceriporiopsis subvermispora*, *Chrysosporium inops*, *Chrysosporium keratinophilum*, *Chrysosporium lucknowense*, *Chrysosporium merdarium*, *Chrysosporium pannicola*, *Chrysosporium queenslandicum*, *Chrysosporium tropicum*, *Chrysosporium zonatum*, *Coprinus cinereus*, *Coriolus hirsutus*, *Fusarium bactridioides*, *Fusarium cerealis*, *Fusarium crookwellense*, *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium graminum*, *Fusarium heterosporum*, *Fusarium negundi*, *Fusarium oxysporum*, *Fusarium reticulatum*, *Fusarium roseum*, *Fusarium sambucinum*, *Fusarium sarcochromum*, *Fusarium sporotrichioides*, *Fusarium sulphureum*, *Fusarium torulosum*, *Fusarium trichothecioides*, *Fusarium venenatum*, *Humicola insolens*, *Humicola lanuginosa*, *Mucor miehei*, *Myceliophthora thermophila*, *Neurospora crassa*, *Penicillium purpurogenum*, *Phanerochaete chrysosporium*, *Phlebia radiata*, *Pleurotus eryngii*, *Thielavia terrestris*, *Trametes villosa*, *Trametes versicolor*, *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei*, o *Trichoderma viride*.

Las células fúngicas se pueden transformar mediante un proceso que implica la formación de protoplastos, transformación de los protoplastos, y regeneración de la pared celular de una manera conocida por sí misma. Los procedimientos adecuados para transformar células hospedadoras de *Aspergillus* y *Trichoderma* se describen en el documento EP 238023, Yelton et al., 1984, Proc. Natl. Acad. Sci. USA 81:1470-1474, y Christensen et al., 1988, Bio/Technology 6:1419-1422. Los métodos adecuados para transformar especies de *Fusarium* se describen en Malardier et al., 1989, Gene 78: 147-156, y el documento WO 96/00787. Las levaduras se pueden transformar usando los procedimientos descritos en Becker y Guarente, En Abelson, J.N. y Simon, M.I., editores, Guide to Yeast Genetics and Molecular Biology, Methods in Enzymology, Volumen 194, pp 182-187, Academic Press, Inc., Nueva York; Ito et al., 1983, J. Bacteriol. 153:163; e Hinnen et al., 1978, Proc. Natl. Acad. Sci. USA 75:1920.

La célula puede ser una célula de mamífero, por ejemplo, COS, CHO (documentos US 4.889.803; US 5.047.335). En una realización particular, la célula es no humana y no embrionica. Además, la ácido graso descarboxilasa de la invención se puede producir mediante un animal transgénico no humano, por ejemplo, en la leche producida por el animal.

Las especies de algas adecuadas para el método de la invención incluyen microalgas tales como, por ejemplo, especies de los géneros *Achnanthes*, *Amphiprora*, *Amphora*, *Ankistrodesmus*, *Asteromonas*, *Aurantiochytrium*, *Boekelovia*, *Bolidomonas*, *Borodinella*, *Botrydium*, *Botryococcus*, *Bracteococcus*, *Chaetoceros*, *Carteria*, *Chlamydomonas*, *Chlorococcum*, *Chlorogonium*, *Chlorella*, *Chroomonas*, *Chrysosphaera*, *Cricosphaera*, *Crypthecodinium*, *Cryptomonas*, *Cyclotella*, *Desmodesmus*, *Dunaliella*, *Elipsoidon*, *Emiliania*, *Eudorina*, *Eremosphaera*, *Ernodesmius*, *Euglena*, *Eustigmatos*, *Franceia*, *Fragilaria*, *Fragilaropsis*, *Gloeothamnion*, *Gonium*, *Haematococcus*, *Hantzschia*, *Heterosigma*, *Hymenomonas*, *Isochrysis*, *Lepocinclis*, *Lobosphaera*, *Micractinium*, *Micrasterias*, *Monodus*, *Monoraphidium*, *Nannochloris*, *Nannochloropsis*, *Navicula*, *Neochloris*, *Nephrochloris*, *Nephroselmis*, *Nitzschia*, *Ochromonas*, *Oedogonium*, *Oocystis*, *Ostreococcus*, *Pandorina*, *Parachlorella*, *Parietochloris*, *Parietichytrium*, *Pascheria*, *Pavlova*, *Pelagomonas*, *Phaiodactylum*, *Phagus*, *Picochlorum*, *Platymonas*, *Pleurochrysis*, *Pleurococcus*, *Prototheca*, *Pseudochlorella*, *Pseudochoricystis*, *Pseudoneochloris*, *Pseudostaurastrum*, *Pyramimonas*, *Pyrobotrys*, *Scenedesmus*, *Schizochlamydeella*, *Schizochytrium*, *Skeletonema*, *Spyrogyra*, *Stichococcus*, *Tetrachlorella*, *Tetraselmis*, *Thalassiosira*, *Thrautochytrium*, *Tribonema*, *Ulva*, *Vaucheria*, *Viridiella*, *Vischeria*, y *Volvox*. Los ejemplos no limitantes de especies especialmente adecuadas incluyen, por ejemplo, diatomeas tales como, por ejemplo, una especie de cualquiera de los géneros *Amphora*, *Chaetoceros*, *Cyclotella*, *Cylindrotheca*, *Fistulifera*, *Fragilaria*, *Fragilaropsis*, *Navicula*, *Nitzschia*, *Phaeodactylum*, *Pseudonitzia*, o *Thalassiosira*, o eustigmatofitos, por ejemplo, *Eustigmatos*, *Monodus*, *Nannochloropsis*, o *Vischeria*.

Más concretamente, las microalgas que se pueden utilizar incluyen, aunque no de forma limitativa, *Achnanthes orientalis*, *Agmenellum*, *Amphiprora hyaline*, *Amphora cojfeiformis*, *Amphora cojfeiformis lineata*, *Amphora coffeijbrms punctata*, *Amphora cojfeiformis taylori*, *Amphora cojfeiformis tenuis*, *Amphora delicatissima*, *Amphora delicatissima capitaia*, *Amphora sp.*, *Anahaena*, *Anabaena variabilis*, *Ankistrodesmus*, *Ankistrodesmus jalcatu*, *Boekelovia hooglandii*, *Borodinella sp.*, *Botryococciis braunii*, *Botryococcus sudeiicus*, *Bracteococcus minor*, *Bracteococcus medionucleatus*, *Carteria*, *Chaetoceros gracilis*, *Chaetoceros muelleri*, *Chaetoceros suhsalsum*, *Chaetoceros sp.*, *Chlamydomonas reinhardtii*, *Chlamydomonase moewusi*, *Chlamydomonas nivalis*, *Chlamydomonas caudate*, *Chlorella anitrata*, *Chlorella antarctica*, *Chlorella aureoviridis*, *Chlorella Candida*, *Chlorella capsulate*, *Chlorella desiccata*, *Chlorella ellipsoidea*, *Chlorella emersonii*, *Chlorella fusca*, *Chlorella fusca var. vacuolata*, *Chlorella glucotropha*, *Chlorella infusionum*, *Chlorella infusionum var. actophila*, *Chlorella in fusionum var. auxenophila*, *Chlorella kessleri*, *Chlorella lobophora* (cepa SAG 37.88), *Chlorella luteoviridis*, *Chlorella luteoviridis var. aureoviridis*, *Chlorella luteoviridis var. lutescens*, *Chlorella miniata*, *Chlorella minutissima*, *Chlorella mutabilis*, *Chlorella nocturna*, *Chlorella ova lis*, *Chlorella parva*, *Chlorella photophila*, *Chlorella pringsheimii*, *Chlorella protothecoides var. acidicola*, *Chlorella regularis*, *Chlorella regularis var. minima*, *Chlorella regularis var. umhricata*, *Chlorella reisiigii*, *Chlorella saccharophila*, *Chlorella saccharophila var. ellipsoidea*, *Chlorella salina*, *Chlorella simplex*, *Chlorella sorokiniana*, *Chlorella sp.*,

*Chlorella sphaerica*, *Chlorella stigmatophora*, *Chlorella variabilis*, *Chlorella vanniellii*, *Chlorella vulgaris*, *Chlorella vulgaris f. tenia*, *Chlorella vulgaris var. autotrophica*, *Chlorella vulgaris var. viridis*, *Chlorella vulgaris var. vulgaris*, *Chlorella vulgaris var. vulgaris f. tertia*, *Chlorella vulgaris var. vulgaris f. viridis*, *Chlorella xanthella*, *Chlorella zojingensis*, *Chlorella trehouxioides*, *Chlorella vulgaris*, *Chlorococcum infusionum*, *Chlorococcum sp.*, *Chlorogonium*, *Chroomonas sp.*, *Chrysosphaera sp.*, *Coccomyxa suhellioides C-169*, *Cricosphaera sp.*, *Cryptohelocodium cohnii*, *Cryptomonas sp.*, *Cyclotella cryptica*, *Cyclotella meneghiniana*, *Cyclotella sp.*, *Dunaliella sp.*, *Dimaliella bardawil*, *Dunaliella hioculata*, *Dunaliella granulata*, *Dunaliella maritime*, *Dunaliella minuta*, *Dunaliella parva*, *Dunaliella peircei*, *Dunaliella primolecta*, *Dunaliella saiina*, *Dimaliella terricola*, *Dunaliella tertiolecta*, *Dimaliella viridis*, *Dunaliella tertiolecta*, *Eremosphaera viridis*, *Eremosphaera sp.*, *Ellipsoidon sp.*, *Euglena*, *Franceia sp.*, *Fragilaria crotonensis*, *Fragilaria sp.*, *Gleocapsa sp.*, *Gloeothamnion sp.*, *Haematococcus plwialis*, *Hymenomonas sp.*, *Isochrysis affgalhana*, *Isochrysis galbana*, *Lepocinclis*, *Micractinium*, *Micractinium*, *Micromonas*, *Micromonas pusilla*, *Monoraphidium minuium*, *Monoraphidium sp.*, *Nannochloris sp.*, *Nannochloropsis saiina*, *Nannochloropsis sp.*, *Navicula acceptata*, *Navicula hiskanterae*, *Navicula pseudotenelloides*, *Navicula pelliculosa*, *Navicula saprophila*, *Navicula sp.*, *Nephrochloris sp.*, *Nephroselrnis sp.*, *Nitzschia communis*, *Nitzschia alexandrina*, *Nitzschia communis*, *Nitzschia dissipata*, *Nitzschia frustulum*, *Nitzschia hantzschiana*, *Nitzschia inconspicua*, *Nitzschia intermedia*, *Nitzschia microcephala*, *Nitzschia pusilla*, *Nitzschia pusilla elliptica*, *Nitzschia pusilla monoensis*, *Nitzschia quadrangular*, *Nitzschia sp.*, *Nostoc sp.*, *Nostoc Pimctiforme*, *Ochromonas sp.*, *Oocystis parva*, *Oocystis pusilla*, *Oocystis sp.*, *Oscillatoria limnetica*, *Oscillatoria sp.*, *Oscillatoria suborevis*, *Osterococcus*, *Osterococcus lucimarinis*, *Osterococcus tauri*, *Parachlorella kessleri*, *Pascheria acidophila*, *Pavlova sp.*, *Phagus*, *Phaeodactylum tricorutum*, *Pnornidium*, *Piaimonas sp.*, *Pleurochrysis carterae*, *Pleurochrysis dentate*, *Pleurochrysis sp.*, *Prochlorococcus marinus*, *Prototheca wickerhamii*, *Prototheca stagnant*, *Prototheca portoricensis*, *Prototheca moiriformis*, *Prototheca zopfii*, *Pseudochlorella aquatica*, *Pyramimonas sp.*, *Pyroboirys*, *Rhodococcus opacus*, *Sarcinoid chrysohyte*, *Scenedesmus armatus*, *Scynechocystis sp.*, *Scynechococcus*, *Schizochytrium*, *Spirogyra*, *Spirulina platensis*, *Stichococcus sp.*, *Synechococcus sp.*, *Tetraedron*, *Thalassiosira pseudonana*, *Tetraselmis sp.*, *Tetraselmis suecica*, *Thalassiosira weissflogii*, *Viridiella fridericana*, y *Volvox carteri*. En algunas realizaciones, los miembros del género *Nannochloropsis* se seleccionan entre *N. gaditana*, *N. granulata*, *N. limnetica*, *N. oceanica*, *N. oculata*, y *N. salina*.

Las especies de algas adecuadas para el método de la invención incluyen también, aunque no de forma limitativa, especies de algas de los géneros de Rhodophyta tales como *Cyanidioschyzon*, *Gracilaria*, *Kappaphycus*, *Porphyridium*, y *Porphyra*; Charophyta tales como *Closterium* y *Penium*; Chlorophyta tales como *Eudorina*, *Gonium*, *Haematococcus*, *Lobosphaera*, *Micrasterias*, *Ostreococcus*, *Pandorina*, *Parachlorella*, *Platymonas*, *Pseudochoricystis*, *Scenedesmus*, *Ulva*, y *Volvox*; Phaeophyta tales como *Laminaria*.

La célula hospedadora también se puede seleccionar entre especies de los géneros de Dinophyta tales como *Amphidinium* y *Symbiodinium*; Chlorarachniophyta tales como *Lotharella*; Euglenozoa tales como *Euglena*.

De acuerdo con la presente invención, alcanos/alquenos que tienen patrones de ramificaciones concretos, niveles de saturación, y longitud de la cadena de carbono, se puede producir a partir de sustratos de ácidos grasos que tienen dichas características particulares con la ácido graso descarboxilasa de la invención. Por consiguiente, cada etapa dentro de una ruta biosintética de ácidos grasos se puede modificar para producir o producir en exceso un sustrato de ácido graso de interés. Por ejemplo, los genes conocidos implicados en la ruta biosintética de ácido graso se pueden expresar, expresar en exceso, o atenuarse en células hospedadoras para producir un sustrato de ácido graso deseado (véase, por ejemplo, los documentos WO 08/119082, US 8.183.028). Los genes ilustrativos se proporcionan en la Fig. 1 del documento WO08/119082 o en la Tabla I y la col 25-30 del documento US 8.183.028.

Después, la célula hospedadora se puede modificar aumentando o dirigiendo la producción de lípidos y/o la producción de ácidos grasos. La célula hospedadora modificada se puede seleccionar por su capacidad de producir ácidos grasos libres o un perfil concreto de ácidos grasos libres, por ejemplo, después de mutagénesis aleatoria, o se puede preparar mediante ingeniería genética por introducción de nuevos genes o aumento en la expresión de algunos genes (genes implicados en el anabolismo de los ácidos grasos) y/o por delección de otros genes o reducción de su expresión (genes implicados en el catabolismo de los ácidos grasos libres). Además, la célula hospedadora se puede modificar para aumentar la producción de ácidos grasos preferentes y la disminución de la producción de ácidos grasos no deseados.

Por ejemplo, para la producción de lípidos, una célula hospedadora (tal como aunque no de forma limitativa, una célula hospedadora de alga o heteroconta) pueden opcionalmente incluir uno o más genes no naturales que codifican polipéptidos que funcionan en la biosíntesis de lípidos, que incluyen, aunque no de forma limitativa, polipéptidos que codifican enzimas para la producción de ácidos grasos, derivados de ácidos grasos, y/o glicerolípidos que incluyen, aunque no de forma limitativa, gen de diacilglicerol aciltransferasa (DGAT), gen de la glicerolfosfato aciltransferasa (GPAT), un gen de ácido lisofosfatídico aciltransferasa (deshidrogenasa) (LPAAT), un gen de ácido fosfatídico fosfatasa (PAP), y/o gen de monoacilglicerol aciltransferasa (MGAT).

En una realización preferida, una célula hospedadora se ha modificado para expresar o expresar en exceso una lipasa. Una lipasa hidroliza un acil-lípido (tal como un triglicérido o un glicerofosfolípido o un glicerogalactolípido) para producir ácidos grasos libres, y glicerol. Preferentemente, la lipasa será una enzima lipolítica con una amplia actividad de triacilglicerol, fosfolipasa y galactolipasa. Los ejemplos incluyen la proteína relacionada con la lipasa pancreática de tipo 2 de cobaya, la cutinasa de *Fusarium* o la lipasa de *Staphylococcus hyicus* (Liu et al. 2011, PNAS 108:6905). En

una realización preferida, la célula hospedadora es bacteriana y expresa, preferentemente expresa en exceso, una lipasa. Por consiguiente, un objeto de la presente invención es una célula hospedadora recombinante que expresa, preferentemente expresa en exceso, una ácido graso descarboxilasa de la presente invención y una lipasa, siendo una de ellas heteróloga respecto a la célula hospedadora. En una realización preferida, la célula hospedadora es bacteriana.

En una realización preferida, la célula hospedadora se ha diseñado mediante ingeniería genética para expresar, expresar en exceso o atenuar la expresión de una tioesterasa para aumentar la producción de ácidos grasos o favorecer la producción de ácidos grasos preferentes. Preferentemente, la célula hospedadora se ha modificado para expresar o expresar en exceso una tioesterasa. Por consiguiente, un objeto de la presente invención es una célula hospedadora recombinante que expresa, preferentemente expresa en exceso, una ácido graso descarboxilasa de la presente invención y una tioesterasa, siendo una de ellas heteróloga respecto a la célula hospedadora. En una realización preferida, la célula hospedadora es una microalga o una cianobacteria.

Al terminar la biosíntesis de ácidos grasos, la proteína transportadora acilo (ACP) tioesterasa determina funcionalmente la longitud y la identidad del producto de ácido graso final (Salas et al., (2002) Archives of Biochemistry and Biophysics 403: 25-34). Basándose en los alineamientos de secuencias de aminoácidos, se ha demostrado que las tioesterasas vegetales se agrupan en dos familias, los FatA, que muestran una notable preferencia por 18:1-ACP con actividad poco importante sobre 18:0-ACP y 16:0-ACP; y los FatB, que hidrolizan los acil-ACP primarios con longitudes de cadenas comprendidas entre 8-16 átomos de carbono (Voelker, en Genetic Engineering Volumen 18. Editado por: Setlow JK. Nueva York, Plenum Press; 1996:111-133; Ginalski, et al., Nucl Acids Res (2003) 31:3291-3292; y Jones, et al, (1995) Plant Cell 7: 359-371).

La tioesterasa pertenece a la enzima de EC 3.1.2, más particularmente de EC 3.1.2.2 (Palmitoil-CoA hidrolasa), EC 3.1.2.14 (Oleoil-[acil-proteína transportadora] hidrolasa), EC 3.1.2.18 (acil-CoA hidrolasa de cadena corta dependiente de ADP), EC 3.1.2.19 (acil-CoA hidrolasa de cadena media dependiente de ADP), EC 3.1.2.20, EC 3.1.2.21 (Dodecanoil-[acil-proteína transportadora] hidrolasa), EC 3.1.2.22 (Palmitoil-proteína hidrolasa).

Los ejemplos de tioesterasa se han descrito en los documentos WO14120829 y WO16044779. También incluyen una tioesterasa de *Escherichia coli* codificada mediante tesB (véase el n.º de registro GenBank AAA24665.1), una tioesterasa de *Lactobacillus brevis* (n.º de registro GenBank ABJ63754.1), y una esterasa de *Lactobacillus plantarum* (n.º de registro GenBank CCC78182.1). Otro ejemplo de tioesterasa de plantas adecuada para preparar ácidos grasos de cadena media como se describe en Radakovits et al, 2011, Metabolic Engineering 13(1):89-95.

En una realización preferida, la tioesterasa se selecciona para favorecer ácidos grasos de cadena corta y media.

En un aspecto particular de la presente invención, también se proporciona un método para producir una ácido graso descarboxilasa de acuerdo con la presente invención, que comprende cultivar la célula hospedadora tal como se ha definido anteriormente, en condiciones que conducen a la producción de la ácido graso descarboxilasa, y recuperar y/o purificar la ácido graso descarboxilasa. Como alternativa, también se proporciona un método para producir una ácido graso descarboxilasa de acuerdo con la presente invención, que comprende la expresión *in vitro* de la ácido graso descarboxilasa con un ácido nucleico que codifica la ácido graso descarboxilasa como se ha definido anteriormente. Opcionalmente, el método también comprende una etapa de inmovilizar la ácido graso descarboxilasa sobre un soporte sólido.

La enzima se puede recuperar utilizando métodos conocidos en la técnica. Por ejemplo, la enzima se puede recuperar del medio nutriente mediante procedimientos convencionales que incluyen, aunque no de forma limitativa, recogida, centrifugación, filtración, extracción, secado mediante pulverización, evaporación, o precipitación.

La enzima se puede purificarse mediante varios procedimientos conocidos en la materia incluidos, aunque no de forma limitativa, cromatografía (por ejemplo, intercambio iónico, afinidad, hidrofobia, cromatofocalización, y exclusión molecular), procedimientos electroforéticos (por ejemplo, focalización isoeléctrica preparativa), diferencias de solubilidad (por ejemplo, precipitación con sulfato de amonio, SDS-PAGE, o extracción (véase, por ejemplo, Protein Purification, Janson y Ryden, editores, VCH Publishers, Nueva York, 1989) para obtener polipéptidos prácticamente puros. En un aspecto alternativo, la enzima no se recupera, sino en su lugar, una célula hospedadora de la presente invención que expresa la enzima como fuente de la enzima.

#### Uso en la producción de alcanos/alquenos

La presente invención también se refiere al uso de una ácido graso descarboxilasa como se ha definido anteriormente, o soporte sólido que comprende la ácido graso descarboxilasa, o una célula hospedadora recombinante que comprende un ácido nucleico, una construcción de ácido nucleico recombinante o un vector recombinante que comprende una secuencia de ácido nucleico que codifica la ácido graso descarboxilasa que se ha definido anteriormente, para producir alcanos/alquenos a partir de ácidos grasos.

#### Sustratos de ácido graso:

Los alcanos/alquenos se producen mediante la enzima de acuerdo con la presente invención a partir de ácidos grasos. Especialmente, los ácidos grasos son ácidos grasos libres que comprenden una cadena alifática y un grupo de ácido carboxilo. Pueden estar saturados o insaturados. Pueden ser monoinsaturados o poliinsaturados. Pueden comprender de 2 a 36 átomos de carbono. En general, los ácidos grasos se clasifican en cuatro grupos según la longitud de la cadena alifática: (1) Los ácidos grasos de cadena corta (SCFA) son ácidos grasos con colas alifáticas de menos de seis átomos de carbono (por ejemplo, ácido butírico); (2) Los ácidos grasos de cadena media (MCFA) son ácidos grasos con colas alifáticas de 6-12 átomos de carbono, que se pueden encontrar en triglicéridos de cadena media; (3) Los ácidos grasos de cadena larga (LCFA) son ácidos grasos con colas alifáticas de 13 a 21; (4) Los ácidos grasos de cadena muy larga (VLCFA) son ácidos grasos con colas alifáticas más largas de 22 átomos de carbono. La presente invención es muy relevante para los MCFA y LCFA, más específicamente para los LCFA. En una realización preferida, los ácidos grasos comprenden de 8 o 10 a 24 átomos de carbono, preferentemente de 12 a 22 átomos de carbono. En algunas realizaciones, el sustrato de ácido graso es un ácido graso 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> o C<sub>22</sub>, preferentemente un ácido graso 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> o C<sub>22</sub>, aún más preferentemente un ácido graso 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> o C<sub>22</sub>. Más preferentemente, el ácido graso comprende de 8 a 20 átomos de carbono, preferentemente de 12 a 18 átomos de carbono, aún más preferentemente de 14 a 18 átomos de carbono, e incluso más preferentemente de 16 o 17 átomos de carbono. En un aspecto específico, el sustrato de ácido graso es un ácido graso 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>, o C<sub>21</sub>, preferentemente un ácido graso 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>, o C<sub>19</sub>, aún más preferentemente un ácido graso C<sub>12</sub>, C<sub>13</sub>, C<sub>14</sub>, C<sub>15</sub>, C<sub>16</sub>, C<sub>17</sub>, o C<sub>18</sub>. La cadena alifática puede ser lineal, ramificada, o incluso incluir un resto cíclico.

En una realización particular, el ácido graso comprende una cadena ramificada. La cadena ramificada puede incluir una cadena principal con sustituciones del grupo alquilo C<sub>1</sub>-C<sub>3</sub>, preferentemente por metilo. En particular, el ácido graso puede ser un ácido graso terpenoide tales como el ácido fitánico y el ácido pristánico.

Además, la cadena alifática del ácido graso puede estar sustituida por uno o varios grupos tales como un hidroxilo, un alcohol C<sub>1</sub>-C<sub>3</sub>, un acilo C<sub>1</sub>-C<sub>3</sub>, un éster C<sub>1</sub>-C<sub>3</sub>, una amina C<sub>1</sub>-C<sub>3</sub>, un grupo amino, una amida C<sub>1</sub>-C<sub>3</sub>, un carboxilo, un aldehído, un epoxi, un halógeno, un alcoxi C<sub>1</sub>-C<sub>3</sub>, un tioalquilo C<sub>1</sub>-C<sub>3</sub>, una imina C<sub>1</sub>-C<sub>3</sub>, un nitrilo, un grupo sulfuro tal como una sulfona C<sub>1</sub>-C<sub>3</sub> o un sulfóxido C<sub>1</sub>-C<sub>3</sub>, un tiol, un nitro, un ciano, un haloalquilo C<sub>1</sub>-C<sub>3</sub>, o puede estar interrumpido por un heteroátomo tal como O, N o S, un grupo acetilénico, un éter tal como un grupo divinil éter, o un grupo oxo. Preferentemente, la cadena alifática del ácido graso puede estar sustituida por uno o varios grupos tales como un hidroxilo, un carboxilo, un epoxi, un halógeno, un tiol o un nitrilo, o puede estar interrumpida por un átomo de S, un grupo acetilénico, o un éter tal como un grupo divinil éter. Preferentemente, la cadena alifática está sustituida por un grupo, en particular en el extremo de la cadena alifática. En una realización muy significativa, la cadena alifática del ácido graso está sustituida por un hidroxilo, especialmente en el extremo de la cadena alifática.

En otra realización, la cadena alifática del ácido graso está no sustituida.

Los sustratos de ácido graso pueden ser un ácido graso aislado o purificado o bien una mezcla de ácidos grasos como se ha mencionado anteriormente.

Por "ácido graso libre" se pretende indicar una forma en la que el ácido graso está en su forma que tiene un grupo carboxilo, es decir, no en una forma esterificada o como otros derivados de ácidos grasos.

#### **Alcanos/alquenos:**

Los alcanos/alquenos de la presente invención se obtienen mediante descarboxilación de los correspondientes ácidos grasos. Por consiguiente, los alcanos se obtienen a partir de ácidos grasos saturados y los alquenos a partir de ácidos grasos insaturados. La enzima de la presente invención no introduce una insaturación en el extremo. Por lo tanto, en una realización preferida, los alquenos obtenidos no presentan una insaturación en el extremo.

Los alcanos/alquenos obtenidos mediante los usos y métodos de la presente invención comprenden de 8 o 10 a 24 átomos de carbono, preferentemente de 12 a 22 átomos de carbono. En algunas realizaciones, el alcano/alqueno es un alcano/alqueno 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> o C<sub>22</sub>, preferentemente un alcano/alqueno 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> o C<sub>22</sub>, aún más preferentemente un alcano/alqueno 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> o C<sub>22</sub>. Más preferentemente, el alcano/alqueno comprende de 8 a 20 átomos de carbono preferentemente de 12 a 18 átomos de carbono, aún más preferentemente de 14 a 18 átomos de carbono, e incluso más preferentemente de 16 a 17 átomos de carbono. En un aspecto específico, el alcano/alqueno es un alcano/alqueno 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>, o C<sub>21</sub>, preferentemente un alcano/alqueno 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>, o C<sub>19</sub>, aún más preferentemente un alcano/alqueno C<sub>12</sub>, C<sub>13</sub>, C<sub>14</sub>, C<sub>15</sub>, o C<sub>16</sub>. La cadena alifática/de hidrocarburo puede ser lineal, ramificada, o incluso incluir un resto cíclico.

En una realización particular, el alcano/alqueno comprende una cadena ramificada. La cadena ramificada puede incluir una cadena principal con sustituciones del grupo alquilo C<sub>1</sub>-C<sub>3</sub>, preferentemente por metilo. En particular, los ácidos grasos pueden ser alcanos terpenoides tales como fitano y pristano.

Además, como la cadena alifática de los ácidos grasos puede estar sustituida o interrumpida por uno o varios grupos como se ha detallado anteriormente, el correspondiente alcano/alqueno después de la descarboxilación también comprende estar sustituido o interrumpido por uno o varios grupos. Por consiguiente, los alcanos/alquenos pueden estar sustituidos por uno o varios grupos tales como un hidroxilo, un alcohol C<sub>1</sub>-C<sub>3</sub>, un acilo C<sub>1</sub>-C<sub>3</sub>, un éster C<sub>1</sub>-C<sub>3</sub>, una amina C<sub>1</sub>-C<sub>3</sub>, un grupo amino, una amida C<sub>1</sub>-C<sub>3</sub>, un carboxilo, un aldehído, un epoxi, un halógeno, un alcoxi C<sub>1</sub>-C<sub>3</sub>, un tioalquilo C<sub>1</sub>-C<sub>3</sub>, una imina C<sub>1</sub>-C<sub>3</sub>, un nitrilo, un grupo sulfuro tal como una sulfona C<sub>1</sub>-C<sub>3</sub> o un sulfóxido C<sub>1</sub>-C<sub>3</sub>, un tiol, un nitro, un ciano, un haloalquilo C<sub>1</sub>-C<sub>3</sub>, o puede estar interrumpido por un heteroátomo tal como O, N o S, un grupo acetilénico, un éter tal como un grupo divinil éter, o un grupo oxo. Preferentemente, los alcanos/alquenos pueden estar sustituidos por uno o varios grupos tales como un hidroxilo, un carboxilo, un epoxi, un halógeno, un tiol o un nitrilo, o puede estar interrumpida por un átomo de S, un grupo acetilénico, o un éter tal como un grupo divinil éter. Preferentemente, el alcano/alqueno pueden estar sustituidos por un grupo, en particular en el extremo de la cadena alifática. En una realización muy significativa, el alcano/alqueno están sustituidos por un hidroxilo, especialmente en el extremo de la cadena alifática y es un alcohol.

Los alcanos/alquenos pueden ser un alcano/alqueno aislado o purificado o una mezcla de alcanos/alquenos como se ha mencionado anteriormente. En una realización muy significativa, la enzima puede producir alcanos y alquenos C<sub>13</sub>-C<sub>17</sub>, especialmente cuando se expresa en bacterias tales como *E. coli*.

## Producción de alcanos/alquenos en un sistema exento de células

En un primer aspecto, la presente invención se refiere al uso de una ácido graso descarboxilasa tal como se ha definido anteriormente para la producción de alcanos/alquenos a partir de ácidos grasos, especialmente ácidos grasos libres. Por consiguiente, la presente invención se refiere a un método para producir alcanos/alquenos en el que la ácido graso descarboxilasa que se ha definido anteriormente se pone en contacto con un ácido graso o mezcla de ácidos grasos y se expone a la luz, convirtiendo de esta forma el ácido graso o mezcla de ácidos grasos en los correspondientes alcanos/alquenos. El ácido graso o mezcla de ácidos grasos puede estar en una forma purificada o puede estar presente en una composición/producto en bruto. La ácido graso descarboxilasa puede estar purificada, aislada o presente en un extracto de proteínas, en particular un extracto de proteínas total. En general, el cofactor FAD ya está presente junto con la ácido graso descarboxilasa formando un complejo. Opcionalmente, el cofactor FAD se puede añadir junto con la ácido graso descarboxilasa.

La enzima necesita exposición a luz que contiene fotones azules para su actividad. En una realización preferida, la luz azul tiene una longitud de onda de 400 a 520 nm, preferentemente de 450 nm a 495 nm, especialmente aproximadamente 450 nm. La cantidad de luz proporcionada puede ser por ejemplo de 10 a 3000  $\mu\text{mol.fotón.m}^{-2}.\text{s}^{-1}$ , preferentemente aproximadamente 2000  $\mu\text{mol.fotón.m}^{-2}.\text{s}^{-1}$ .

Sin embargo, la luz puede tener un intervalo de longitud de onda más amplio comprendido de 300 a 540 nm. De hecho, los inventores observaron que la FAD, cofactor de la enzima, es capaz de absorber luz en el intervalo de longitudes de onda de 300 a 540 nm y la enzima debería poder descarboxilar los ácidos grasos en este intervalo. De hecho, el espectro de absorción de FAD está comprendido entre las longitudes de onda de 300 a 540 nm.

La enzima también necesita FAD (flavina adenina dinucleótido) como cofactor, que está habitualmente unido a la enzima purificada de extractos de microalgas o a partir de la expresión heteróloga.

El método puede comprender también una etapa adicional de recuperar los alcanos/alquenos. Los alcanos/alquenos se pueden extraer o purificar por cualquier método disponible para el experto en la materia. Los alcanos/alquenos se pueden recuperar en la fase orgánica. Por ejemplo, se pueden extraer con un disolvente orgánico, por ejemplo, con hexano. Pero los alcanos/alquenos también deberían recuperarse de la fase gaseosa del cultivo por condensación.

## Producción de alcanos/alquenos en un sistema de célula hospedadora recombinante

La presente invención se refiere al uso de una célula hospedadora recombinante tal como se ha como se ha descrito anteriormente para la producción de alcanos/alquenos a partir de ácidos grasos, especialmente ácidos grasos libres. También se refiere a un método para producir alcanos/alquenos a partir de ácidos grasos, en el que una célula hospedadora recombinante como se ha descrito anteriormente. se cultiva en condiciones eficaces para expresar la ácido graso descarboxilasa de la presente invención.

Por consiguiente, la presente invención se refiere a un cultivo celular que comprende una célula hospedadora recombinante como se ha descrito anteriormente cultivada en condiciones eficaces para expresar la ácido graso descarboxilasa de la presente invención.

Opcionalmente, la célula hospedadora recombinante se puede cultivar en un medio que comprende ácidos grasos. De forma alternativa o adicional, la célula hospedadora recombinante produce o comprende ácidos grasos.

En el método, la célula hospedadora recombinante se cultiva preferentemente en condiciones para obtener una

biomasa, preferentemente una biomasa rica en lípidos, especialmente ácidos grasos. Después, en una segunda etapa, se inicia la conversión de los ácidos grasos en alcanos/alquenos. Por ejemplo, si la ácido graso descarboxilasa ya está expresada en la célula hospedadora recombinante, la segunda etapa se puede iniciar aplicando una luz que contiene fotones de color azul a la célula hospedadora recombinante para permitir la actividad ácido graso  
 5 descarboxilasa. De hecho, los inventores han mostrado que la producción de alcano se puede controlar mediante luz, especialmente su longitud de onda. La luz azul (>400 nm y <530 nm) permite la producción de alcanos mientras que la luz roja (>600 nm y <700 nm) no lo hace. Después, en el método de la presente invención, la producción de alcanos se puede modular *in vivo* mediante luz. Como alternativa, si la expresión de la ácido graso descarboxilasa se controla mediante un promotor inducible, la segunda etapa se puede iniciar induciendo la expresión de la ácido graso  
 10 descarboxilasa y sometiendo las células a fotones (300-540 nm). Por otra parte, la intensidad de la luz (flujo de fotones) modula la actividad enzimática. Un flujo de fotones elevado aumenta la actividad enzimática, mientras que un flujo de fotones bajo disminuye la actividad enzimática.

Como alternativa, el método puede comprender el cultivo de la célula hospedadora recombinante en condiciones que fomentan conjuntamente la producción de una biomasa, preferentemente rica en lípidos, especialmente ácidos grasos, y la conversión de los ácidos grasos en alcanos/alquenos.

Por "rica" se entiende que la biomasa comprende al menos un 20,30,40,50, 60 o 70 % en peso de lípidos, especialmente ácidos grasos, por biomasa seca.

El método puede comprender una etapa adicional de aislar o recuperar los alcanos/alquenos de la célula hospedadora recombinante o del medio de cultivo.

Los alcanos/alquenos se pueden extraer o purificar por cualquier método disponible para el experto en la materia. Por ejemplo, se pueden extraer con un disolvente orgánico, por ejemplo, con hexano. Los alcanos/alquenos se pueden recuperar en la fase orgánica.

Un proceso de separación ilustrativo es un proceso de separación en dos fases (bifásico). Este proceso implica fermentar las células hospedadoras diseñadas mediante ingeniería genética en condiciones suficientes para producir alcanos/alquenos, dejar que los alcanos/alquenos se recojan en una fase orgánica, y separar la fase orgánica del caldo de fermentación acuoso. Este método se puede llevar a la práctica en un escenario de fermentación tanto continuo o como discontinuo.

Los métodos y usos de acuerdo con la presente invención para producir alcanos/alquenos se pueden utilizar para preparar biocombustibles, materiales útiles en química, especialmente en cosmética y otros campos tales como plásticos, resinas, fibras, elastómeros, agentes farmacéuticos, lubricantes o geles.

Los alquenos descritos en el presente documento se pueden usar tal cual o convertirse en un combustible. El experto en la materia apreciará que, dependiendo del uso previsto del combustible, se pueden producir y utilizar diferentes alquenos. Por ejemplo, los alquenos ramificados pueden ser deseables para combustible de automoción de uso previsto en climas fríos. Además, cuando los alquenos descritos en el presente documento se usan como materia prima para la producción de combustibles, un experto en la materia apreciará que las características de la materia prima de olefina afectarán las características del combustible producido. Por tanto, las características del combustible producido se pueden seleccionar mediante la producción de alquenos en concreto para usar como materia prima.

Usando los métodos descritos en el presente documento, se pueden producir biocombustibles que tengan las calidades de combustible deseadas a partir de alquenos. Los alquenos producidos de forma biológica representan una nueva fuente de biocombustibles, que se pueden utilizar como combustible de motores a reacción, diésel, o gasolina. Algunos biocombustibles fabricados usando alquenos no se han producido a partir de fuentes renovables y son nuevas composiciones de materia.

### Dosificación de ácido graso

La ácido graso descarboxilasa de la invención también es útil para la cuantificación de ácidos grasos libres, y se puede incluir en un kit para la cuantificación de ácidos grasos libres. En particular, la presente invención se refiere a un método para cuantificar ácidos grasos en una muestra, que comprende poner en contacto la muestra con una ácido graso descarboxilasa de la presente invención en condiciones adecuadas para convertir los ácidos grasos en alcanos/alquenos, recuperar los alcanos/alquenos y cuantificar los alcanos/alquenos. La muestra puede ser una muestra de alimento o una muestra biológica, en particular, una muestra de fluido biológico, tal como sangre, suero, plasma, orina y similares. Los alcanos/alquenos producidos se pueden extraer con un disolvente orgánico y cuantificarse, en particular mediante cromatografía de gases acoplada a un detector de ionización por llama y a espectrometría de masas (GC-FID-MS).

En una realización particular, la ácido graso descarboxilasa de la invención también es útil para la cuantificación de ácidos grasos totales, y se puede incluir en un kit para la cuantificación de ácidos grasos totales. En particular, la presente invención se refiere a un método para cuantificar ácidos grasos en una muestra, que comprende poner en

contacto la muestra con una lipasa en condiciones adecuadas para convertir los ácidos grasos en ácidos grasos libres y cuantificar los ácidos grasos libres como se ha detallado anteriormente. Por consiguiente, el método puede comprender poner en contacto la muestra con una lipasa en condiciones adecuadas para convertir los ácidos grasos en ácidos grasos libres, poner en contacto una ácido graso descarboxilasa de la presente invención en condiciones adecuadas para convertir los ácidos grasos en alcanos/alquenos, recuperar los alcanos/alquenos y cuantificar los alcanos/alquenos. Los alcanos/alquenos producidos se pueden extraer con un disolvente orgánico y cuantificarse.

Cuando la ácido graso descarboxilasa convierte los ácidos grasos en alcanos/alquenos, produce CO<sub>2</sub> como subproducto. Después, en un método alternativo, la presente invención se refiere a un método para cuantificar ácidos grasos en una muestra, que comprende poner en contacto la muestra con una ácido graso descarboxilasa de la presente invención en condiciones adecuadas para convertir los ácidos grasos en alcanos/alquenos, recuperar el CO<sub>2</sub> y cuantificar el CO<sub>2</sub>. La cantidad de CO<sub>2</sub> se puede medir por cualquier método conocido del experto en la materia, por ejemplo, cromatografía de gases-espectrometría de masas (CG-EM).

En un aspecto adicional, los inventores observaron que la fluorescencia de la ácido graso descarboxilasa varía en función de la concentración de sustrato. Por lo tanto, la variación de la fluorescencia proporciona un medio adicional para cuantificar los ácidos grasos de una muestra. Después, en un método adicional alternativo, la presente invención se refiere a un método para cuantificar ácidos grasos en una muestra, que comprende poner en contacto la muestra con una ácido graso descarboxilasa de la presente invención en condiciones adecuadas para convertir los ácidos grasos en alcanos/alquenos, y medir la fluorescencia emitida por la ácido graso descarboxilasa. Preferentemente, la fluorescencia se mide a una longitud de onda en el intervalo de 500 a 700, la cinética se realiza preferentemente a aproximadamente 540 nm.

Por supuesto, el método para cuantificar los ácidos grasos de una muestra puede combinar varios métodos como se ha detallado anteriormente, por ejemplo, cuantificar los alcanos/alquenos y CO<sub>2</sub>, cuantificar los alcanos/alquenos y medir la fluorescencia de la enzima, cuantificar CO<sub>2</sub> y medir la fluorescencia de la enzima, o la combinación de los tres parámetros.

La dosificación de ácido graso es útil para diagnosticar enfermedades, por ejemplo, enfermedades hepáticas (documento WO 15089102), diabetes y preeclampsia (documento WO13170369).

### Eliminación de ácidos grasos

La ácido graso descarboxilasa de la invención también puede ser útil para eliminar los ácidos grasos libres de una composición, por ejemplo, de aceites alimentarios, en particular de aceites alimentarios refinados. De hecho, el uso de la ácido graso descarboxilasa de la invención puede permitir la descarboxilación de ácidos grasos sin ningún proceso químico. Por lo tanto, la presente invención se refiere al uso de un polipéptido tal como se ha definido anteriormente en la descarboxilación de ácidos grasos, eliminando de esta forma los ácidos grasos, o un método para eliminar los ácidos grasos libres de una composición que comprende poner en contacto la ácido graso descarboxilasa de la invención con la composición en condiciones adecuadas para descarboxilar los ácidos grasos libres de la composición.

Otros aspectos y ventajas de la presente invención se describen en la siguiente sección experimental, que deberá considerarse como ilustrativa y no limitante del ámbito de la presente solicitud.

### Breve descripción de las figuras

**Figura 1. Detección de una actividad alcano sintasa en un homogenado de *Chlorella*.** Un homogenado celular de *Chlorella variabilis* NC64A se incubó durante la noche en un vial precintado con ácido palmítico perdeuterado (D31). Los hidrocarburos de la fase gaseosa se extrajeron mediante microextracción en fase sólida (SPME) y se analizaron mediante cromatografía de gases acoplada a espectrometría de masas (GC-MS). Se extrajeron los iones correspondientes a  $m/z = 243,5 \pm 0,5$ .

Panel superior: porción del cromatograma correspondiente al producto de pentadecano marcado; control: homogenado precalentado a 95°C durante 30 minutos. Panel inferior: espectro de masas del pentadecano marcado.

**Figura 2. Resumen del procedimiento de purificación parcial de la alcano sintasa y análisis de los perfiles de proteínas en varias etapas de la purificación.** Los ensayos de la actividad alcano sintasa se llevaron a cabo en fracciones de celular y las fracciones se eluyeron como se describe en la Figura 1 para el homogenado celular. Se llevó a cabo electroforesis de proteínas en un gel de acrilamida al 10 % en condiciones de desnaturalización. 1: después de la filtración en gel; 2: después de un caudal Q rápido; 3 y 4: después de Mono Q (fracciones enviadas para análisis proteómico).

**Figura 3.** Número de proteínas detectadas mediante análisis proteómico después de cada una de tres purificaciones y listado de las 10 proteínas comunes. Se realizaron tres purificaciones independientes (A, B, C).

**Figura 4. La alcano sintasa de *Chlorella* es una GMC oxidorreductasa prevista en cloroplastos.** Los datos se recuperaron de PFAM y ProtParam. Aminoácidos: 654; Peso molecular: 69070 Da pl teórico: 9,075. Se predice su localización en el cloroplasto por Predalگو.

**Figura 5. Análisis de proteínas en células de *E. coli* que expresan la GMC oxidorreductasa de *Chlorella*.** La enzima de *Chlorella* se expresó como una proteína marcada con His en el extremo C. Izquierda: perfil de la proteína total como se observa mediante electroforesis en gel de poliacrilamida. Derecha: Transferencia Western usando anticuerpos dirigidos contra His. C: cepa de control con el vector vacío; GMC: cepa que expresa la GMC oxidorreductasa de *Chlorella*.

**Figura 6. Análisis del contenido de hidrocarburo en células de *E. coli* que expresan la GMC oxidorreductasa de *Chlorella*.** Las células se saponificaron y el contenido de hidrocarburos se analizó mediante SPME y GCMS. Se deberá destacar que no se añadió sustrato de ácido graso. Panel superior: región de los cromatogramas correspondientes a los hidrocarburos. Paneles inferiores: espectros de masas de los alcanos detectados. Control: cepa con el vector vacío; GMC: cepa que expresa la GMC oxidorreductasa de *Chlorella*. Los datos cuantificados muestran media  $\pm$  s.d. (n = 3).

**Figura 7. Análisis de hidrocarburos en microalgas transgénicas que expresan la GMC oxidorreductasa.** Los hidrocarburos se analizaron mediante GC-MS en una cepa de la microalga *Chlamydomonas reinhardtii* en la que el genoma cloroplástico se ha transformado con un gen que codifica la GMC oxidorreductasa. Panel superior: parte del cromatograma GC que muestra los picos del hidrocarburo. Panel inferior: cuantificación de los hidrocarburos. WT: cepa natural. GMC OE: cepa que expresa la GMC oxidorreductasa de *Chlorella*.

**Figura 8. Alineamiento múltiple (A) y árbol filogenético (B) de la familia de la GMC oxidorreductasa.** La alineación múltiple se construyó con Clustal O (1.2.1) usando secuencias de *Chlorella* (SEQ ID No 1); *Chlamydomonas* (SEQ ID No 5); *Coccomyxa* (SEQ ID No 9); *Volvox* (SEQ ID No 10); *Ectocarpus* (SEQ ID No 11); *Emiliania* (SEQ ID No 12); *Aureococcus* (SEQ ID No 13); *Phaeodactylum* (SEQ ID No 7); *Nannochloropsis* (SEQ ID No 14). El árbol filogenético no enraizado se construyó usando un conjunto de 56 secuencias de proteínas GMC oxidorreductasa procedentes de varios orígenes (método de unión de vecinos). La GMC oxidorreductasa de *Chlorella* está recuadrada.

**Figura 9. Fracciones purificadas de las GMC oxidorreductasas de *Chlorella* recombinantes.** La enzima se expresó en *E. coli* como una proteína marcada con His en el extremo y se purificó en una columna Ni. F1 a F4 son fracciones de elución. L: escalera de pesos moleculares.

**Figura 10. Actividad relativa de la GMC oxidorreductasa de *Chlorella* en diferentes ácidos grasos.** La enzima recombinante purificada se incubó con ácidos grasos de diferentes longitudes de cadena con luz blanca y los productos se analizaron mediante GC-MS. Los datos cuantificados muestran media  $\pm$  s.d. (n = 3)

**Figura 11: CO<sub>2</sub> como subproducto de descarboxilación catalizado por la alcano sintasa.** Cuantificación relativa del <sup>13</sup>CO<sub>2</sub> liberado tras incubación de l-<sup>13</sup>C-palmitato con la enzima recombinante purificada. Los datos cuantificados muestran media  $\pm$  s.d. (n = 3).

**Figura 12: La GMC oxidorreductasa de *Chlorella* convierte ácidos terpenoicos en metil alca(que)nos.** La enzima recombinante purificada se incubó con ácido fitánico con luz blanca y los productos se analizaron mediante GC-MS.

**Figura 13: La GMC oxidorreductasa de *Chlorella* convierte ácidos hidroxiácidos en alcoholes.** La enzima recombinante purificada se incubó con hidroxipalmitato con luz blanca y los productos se analizaron mediante GC-MS.

**Figura 14. Producción *in vitro* de alcanos a partir de ácidos grasos usando GMC oxidorreductasa junto con una lipasa.** GMC reductasa de *Chlorella* y lipasa de *Rhizopus* se incubaron con triheptadecanoato de glicerilo y los productos se analizaron mediante GC-MS. Los datos cuantificados muestran media  $\pm$  s.d. (n = 3).

**Figura 15: Espectro de absorción de la GMC oxidorreductasa de *Chlorella*.**

**Figura 16: Dependencia de la luz de las GMC oxidorreductasas de algas.** a, Espectro de absorbancia y espectro de acción de la GMC oxidorreductasa de *Chlorella*. b, Seguimiento de la liberación de <sup>13</sup>CO<sub>2</sub> por espectrometría de masas con entrada de membrana tras la incubación de l-<sup>13</sup>C-palmitato con la enzima de *Chlorella*. La mezcla de reacción se expuso sucesivamente a condiciones de luz (azul o roja) a 850  $\mu$ moles.fotones<sup>-2</sup>s<sup>-1</sup>, c, Dependencia de la actividad de la GMC oxidorreductasa de *Chlorella* con la intensidad de la luz, d, Variación de hidrocarburos totales en células de *Chlamydomonas* durante un cultivo con luz azul y después luz roja a 30  $\mu$ moles.fotones<sup>-2</sup>s<sup>-1</sup>. Los datos cuantificados muestran media  $\pm$  s.d. (n = 3).

**Figura 17: Cuantificación de hidrocarburos en células de *E. coli* que expresan la alcano sintasa de *Chlorella* y**

cultivadas en la oscuridad o en luz (1000  $\mu\text{moles}\cdot\text{fotones}^{-2}\cdot\text{s}^{-1}$  de luz blanca que contenía fotones de 400 a 800 nm). Los datos cuantificados muestran media  $\pm$  s.d. (n = 3). Nd, no detectado.

5 **Figura 18. Cinética de la fluorescencia de la GMC oxidorreductasa en presencia de sustrato.** La excitación fue a 450 nm  $\pm$  5 y la emisión a 540 nm. Se usaron varias concentraciones de ácido palmítico.

10 **Figura 19: Cuantificación de hidrocarburos en células de *E. coli* que expresan la GMC oxidorreductasa de *Chlamydomonas* o *Phaeodactylum*.** Las células se transmitieron y el contenido de hidrocarburos se analizó mediante extracción con disolvente y GCMS. Se deberá destacar que no se añadió sustrato de ácido graso. No se detectaron alcanos en las células de *E. coli* transformadas con un vector vacío. Tr, trazas. Los datos cuantificados muestran media  $\pm$  s.d. (n = 3).

15 **Figura 20: Cuantificación de hidrocarburos en células de *E. coli* que expresan simultáneamente una GMC oxidorreductasa de alga y una lipasa.** Las células se transmitieron y el contenido de hidrocarburos se analizó mediante extracción con disolvente y GCMS. Se deberá destacar que no se añadió sustrato de ácido graso. Control negativo: Células de *E. coli* transformadas con un vector vacío. Nd, no detectado. Cr: *Chlamydomonas reinhardtii*; Cv: *Chlorella variabilis*. La lipasa es de la bacteria *Staphylococcus hyicus* (Uniprot P04635). Los datos cuantificados muestran media  $\pm$  s.d. (n = 3).

20 **Breve descripción del listado de secuencias**

SEQ No	ID	Descripción
1		Secuencia de aminoácidos de la proteína GMC derivada de <i>Chlorella variabilis</i> NC64A sin el péptido transitorio de cloroplasto posible
2		Secuencia de aminoácidos de la proteína GMC derivada de <i>Chlorella variabilis</i> NC64A con el péptido transitorio de cloroplasto posible
3		Secuencia de aminoácidos de una proteína GMC derivada de <i>Chlorella variabilis</i> NC64A sin el péptido transitorio de cloroplasto posible, pero con una etiqueta his, tiorredoxina y un sitio de escisión TEV ( <i>virus del grabado del tabaco</i> ) en el extremo N
4		secuencia consenso
5		Secuencia de aminoácidos de la proteína GMC derivada de <i>Chlamydomonas reinhardtii</i> sin el péptido transitorio de cloroplasto posible
6		Secuencia de aminoácidos de la proteína GMC derivada de <i>Chlamydomonas reinhardtii</i> con el péptido transitorio de cloroplasto posible
7		Secuencia de aminoácidos de la proteína GMC derivada de <i>Phaeodactylum tricornutum</i> sin el péptido transitorio de cloroplasto posible
8		Secuencia de aminoácidos de la proteína GMC derivada de <i>Phaeodactylum tricornutum</i> con el péptido transitorio de cloroplasto posible
9		Secuencia de aminoácidos de la proteína GMC derivada de <i>Coccomyxa subellipsoidea</i> sin el péptido transitorio de cloroplasto posible
10		Secuencia de aminoácidos de la proteína GMC derivada de <i>Volvox carteri</i> sin el péptido transitorio de cloroplasto posible
11		Secuencia de aminoácidos de la proteína GMC derivada de <i>Ectocarpus siliculosus</i> sin el péptido transitorio de cloroplasto posible
12		Secuencia de aminoácidos de la proteína GMC derivada de <i>Emiliania huxleyi</i> sin el péptido transitorio de cloroplasto posible

(continuación)

SEQ No	ID	Descripción
13		Secuencia de aminoácidos de la proteína GMC derivada de <i>Aureococcus anophagefferens</i> sin el péptido transitorio de cloroplasto posible
14		Secuencia de aminoácidos de la proteína GMC derivada de <i>Nannochloropsis gaditana</i> sin el péptido transitorio de cloroplasto posible
15		Secuencias de ácido nucleico que codifica la SEQ ID No 1
16		Secuencias de ácido nucleico que codifica la SEQ ID No 5
17		Secuencias de ácido nucleico que codifica la SEQ ID No 7
18		Secuencias de ácido nucleico que codifica la SEQ ID No 9
19		Secuencias de ácido nucleico que codifica la SEQ ID No 10
20		Secuencias de ácido nucleico que codifica la SEQ ID No 11
21		Secuencias de ácido nucleico que codifica la SEQ ID No 12
22		Secuencias de ácido nucleico que codifica la SEQ ID No 13
23		Secuencias de ácido nucleico que codifica la SEQ ID No 14

## EJEMPLOS

5 En el presente documento, los inventores identificaron en la microalga modelo *Chlorella variabilis* NC64A una enzima que cataliza la síntesis de alca(que)nos. La enzima se purificó parcialmente usando ácido palmítico marcado con deuterio como sustrato y microextracción en fase sólida para capturar el producto de pentadecano. Se identificó una proteína candidato que pertenecía a la familia de la Glucosa-Metanol-Colina oxidorreductasa mediante el análisis proteómico de tres purificaciones parciales independientes. La expresión heteróloga de este gen candidato de *Chlorella* en *Escherichia coli* dio como resultado la producción de hidrocarburos lineales de 13 a 17 átomos de carbono, mostrando que una sola enzima es suficiente para producir alca(que)nos de tipo combustible. Esta alcano sintasa de *Chlorella* es una proteína de 69 kDa prevista en cloroplastos que utiliza FAD como cofactor. Los ensayos *in vitro* muestran que puede usar ácidos grasos C12 a C22 para formar alca(que)nos. Se ha descubierto que la actividad de esta enzima era estrictamente dependiente de los fotones 400 a 540 nm pero que también podría trabajar de 300 a 10 400 nm. Estos resultados amplían, por tanto, el conocimiento actual del repertorio catalítico de la familia de la Glucosa-Metanol-Colina oxidorreductasa y abre un nuevo camino para la producción renovable e impulsada por luz de alca(que)nos en microorganismos.

**Resultados**

20 **Purificación parcial de una actividad alcano sintasa a partir de *Chlorella variabilis* NC64A**

Los inventores han mostrado que varias microalgas, incluida *Chlorella variabilis* NC64A, tienen la capacidad de sintetizar alcanos y alquenos C15-C17. En el mismo trabajo, también han comprobado que los ácidos palmíticos deuterados añadidos de forma exógena a cultivos de *Chlorella* se puede convertir en alcanos y alquenos. Para identificar la ruta enzimática de la síntesis de alcanos en microalgas, los inventores seleccionaron un enfoque de purificación tradicional basado en el uso de ácido palmítico marcado con deuterio como sustrato.

La primera etapa fue confirmar que se podría medir una actividad enzimática en un homogenado de células de *Chlorella*. El ácido palmítico deuterado se añadió a un homogenado celular y se incubaron durante la noche en un vial precintado. El producto de pentadecano esperado se extrajo mediante microextracción en fase sólida (SPME) y se analizaron mediante cromatografía de gases acoplada a espectrometría de masas (GC-MS). Se pudo detectar un pico a 12,03 minutos correspondiente al pentadecano marcado sobre células intactas, pero que estaba ausente en el homogenado de control precalentado (**Figura 1**).

Este experimento mostró, por tanto, que un homogenado de *Chlorella* tenía una actividad enzimática de síntesis de alcanos. Puesto que la ruta de síntesis de alcanos identificada en la mayoría de microorganismos tiene un compuesto intermedio de aldehído, los inventores usaron el mismo experimento usando aldehído C16 marcado, pero no se pudo detectar el pentadecano marcado.

Por tanto, el ácido palmítico marcado se usó para estudiar la actividad durante la totalidad del procedimiento de purificación (Figura 2). Cuando las células se centrifugaron a 50.000 g, la mayoría de la actividad se descubrió en la fracción sobrenadante. Pero después de una segunda centrifugación a 105.000 g, la actividad se descubrió principalmente en el aglomerado (fracción microsómica). Se estudiaron diferentes detergentes para solubilizar la actividad, y se descubrió que el más eficaz era Triton X100.

A continuación, se llevaron a cabo algunos ensayos preliminares antes de purificar adicionalmente la actividad. algunos cofactores, tales como ferredoxina, ferredoxina reductasa, NADP, NADPH y ATP se añadieron en diferentes combinaciones de la fracción microsómica solubilizada. Para ninguno de ellos se descubrió un aumento en la actividad y no se añadieron a los ensayos realizados sobre las fracciones purificadas. Los inventores también observaron que,

en tres días, la actividad de la fracción microsómica almacenada a 4 °C disminuyó en un 90 %, lo que indica que la totalidad del proceso de purificación debía realizarse en pocos días.

5 La purificación parcial de la actividad solubilizada implicó una primera etapa de filtración en gel con una columna preparativa Superdex 200 y a continuación dos columnas de intercambio aniónico, una Q de flujo rápido y una mono Q final más resolutive. La mayoría de las fracciones se analizaron para determinar la actividad alcano sintasa usando el ensayo anteriormente descrito. El contenido de proteína de las fracciones más activas se analizó por electroforesis en gel de acrilamida en condiciones de desnaturalización (**Figura 2**).

10 Se llevaron a cabo tres purificaciones parciales. Las fracciones con la mayor actividad después de la etapa de purificación final se enviaron para análisis proteómico. Tomando un punto de corte de 2 cuentas de péptidos como mínimo, solamente diez proteínas fueron comunes entre las tres purificaciones (**Figura 3**). Nueve de estas proteínas eran homólogos evidentes de enzimas bien caracterizadas. El único candidato que pertenecía a un grupo de enzimas con un intervalo diverso de actividades era una posible enzima de la familia de la Glucosa-Metanol-Colina (GMC) oxidorreductasa. Este candidato era también el único con el recuento peptídico más elevado. Por tanto, se seleccionó para su expresión heteróloga.

#### **La alcano sintasa de *Chlorella* es un miembro de la familia de la GMC oxidorreductasa**

20 El gen que codifica la GMC oxidorreductasa de *Chlorella* no estaba completamente incluido en los EST públicamente disponibles. Un ADNc de aproximadamente 2 kb se clonó un extracto de ARN total procedente de *Chlorella*. Codificaba una proteína de 69 kDa (**Figura 4**) y se predijo que estaría localizada en los cloroplastos mediante el programa informático adaptado a las secuencias de microalgas Predalgo. Se predijo que la señal de direccionamiento al cloroplasto del extremo N tendría 50 restos de longitud. El programa informático TMHMM no predijo dominio transmembrana. Este ADNc se expresó en *Escherichia coli* como una proteína marcada con His en su extremo C (**Figura 5**). La presencia de la proteína recombinante se comprobó mediante transferencia Western. El análisis mediante SPME y GC-MS de los productos volátiles de las células de *E. coli* que expresaban la GMC oxidorreductasa mostraron la presencia de alcanos de cadena larga de 13 a 17 átomos de carbono, que estaban ausentes de las células de *E. coli* transformadas con un vector vacío (**Figura 6**). Estos resultados demuestran, por tanto, que la expresión de la GMC oxidorreductasa de *Chlorella* fue suficiente para producir alcanos y alquenos en *E. coli*. Además, en una cepa de *Chlamydomonas* cuyo genoma cloroplástico se transformó con el ADNc que codificaba la enzima de *Chlorella*, se observó un aumento en el contenido de heptadeceno, y también apareció heptadecano (**Figura 7**). Este resultado, por tanto, indica que la enzima de *Chlorella* es funcional dentro de un cloroplasto.

35 Las búsquedas BlastP utilizando la GMC oxidorreductasa de *Chlorella* u otras GMC oxidorreductasas bioquímicamente caracterizadas de otras especies se llevaron a cabo en bases de datos públicas para recuperar varias secuencias de proteínas GMC oxidorreductasas. Alineamientos múltiples de secuencias de algas indicaron que algunos restos de *Chlorella* tales como C372, R391, Y406, Q426, H512 y N515 estaban muy conservados en otras algas (**Figura 8A**) y se trazó un árbol filogenético no enraizado (**Figura 8B**). Las búsquedas Blast indicaron que, en la mayoría de los organismos, las GMC oxidorreductasas eran familias de un solo gen. En *Chlorella variabilis* NC64A, la alcano sintasa de *Chlorella* es la única GMC oxidorreductasa. El árbol filogenético mostró que la familia de las GMC oxidorreductasas consiste en al menos 7 grupos, uno de los cuales sigue sin tener un miembro bioquímicamente caracterizado. De manera interesante, secuencias de macroalgas pardas (*Ectocarpus*), de macroalgas rojas (*Chondrus*) y de microalgas de varios orígenes (la Trebouxiiofícea *Chlorella*, la Clorofícea *Chlamydomonas* y Volvox, la Coccomyxaceae *Coccomyxa*, la Eustigmatofícea *Nannochloropsis*, la Coccolithofícea *Emiliana*, la diatomea *Phaeodactylum*) se agruparon entre sí. La enzima de *Chlorella* es la primera que está funcionalmente caracterizada en este grupo de GMC oxidorreductasas procedentes de algas.

#### **La alcano sintasa de *Chlorella* es una fotoenzima activada por luz que actúa sobre varios ácidos grasos**

50 Para caracterizar adicionalmente su actividad, la GMC oxidorreductasa de *Chlorella* se expresó en *Escherichia coli* como una proteína marcada con His y purificada en una columna de níquel (**Figura 9**). Se descubrió que la proteína recombinante purificada era activa sobre varios ácidos grasos libres saturados comprendidos de C12 a C22 átomos de carbono (**Figura 10**), lo que explica el amplio perfil de ácidos grasos formados en *E. coli* (**Figura 6**).

55 Para determinar el coproducto de la reacción, los inventores usaron ácido palmítico marcado con <sup>13</sup>C en el grupo carboxílico. Observaron la producción de <sup>13</sup>CO<sub>2</sub> lo que demuestra que la enzima libera CO<sub>2</sub> como coproducto, es decir, es una ácido graso descarboxilasa. (Figura 11)

60 Para caracterizar adicionalmente su actividad alcano sintasa, se usaron como sustrato ácidos grasos sustituidos. Se descubrió que el ácido fitánico (ácido 3,7,11,15-tetrametilhexadecanoico) se convertía en alcanos ramificados, lo que indica que la alcano sintasa es activa sobre ácidos terpenoicos (Figura 12). El uso de ácido palmítico 16-OH como sustrato llevó a la producción de pentadecano mostrando que la GMC oxidorreductasa de *Chlorella* se podía usar para producir alcoholes grasos (Figura 13).

65 *In vitro*, la alcano sintasa no pudo usar TAG (triacilglicerol) directamente como sustrato (Figura 14). Sin embargo, en

presencia de TAG (triacilglicerol) lipasa que libera ácidos grasos, se observó producción de hidrocarburos. Esta observación muestra que la GMC oxidorreductasa se podría utilizar para producir hidrocarburos a partir de aceites y residuos ricos en lípidos.

- 5 La Figura 15 muestra que la FAD puede absorber fotones de 320 a 540 nm. Esto fue coherente con el espectro de acción de la enzima (Figura 16A).

Para determinar si la alcano sintasa era una enzima impulsada por luz (fotoenzima) o una enzima activada por luz, se controló la producción de CO<sub>2</sub> durante la reacción. La actividad de la enzima está impulsada por la presencia de fotones y se detiene inmediatamente cuando la luz se apaga. (Figura 16B). La producción de pentadecano aumenta linealmente con la intensidad de la luz, lo que confirma el hecho de que la GMC oxidorreductasa de *Chlorella* es una fotoenzima (Figura 16C). La dependencia de la luz de la alcano sintasa de alga se podría utilizar para modular la producción de hidrocarburos en células de algas usando la calidad de la luz (azul o roja) tal como se observó para *Chlamydomonas reinhardtii* (Figura 16D). Las microalgas producen hidrocarburos en presencia de fotones azules, mientras que para el rojo, la producción se detiene y el contenido de alcano disminuye. En células de *E. coli* que expresan la alcano fotosintasa de hidrocarburos, la luz también se puede usar como inductor para seleccionar el tiempo para una producción óptima de alcanos y alquenos. (Figura 17).

#### 20 **La concentración de sustrato modifica la fluorescencia de la enzima.**

Basándose en la dependencia de la luz de la enzima, los inventores sospechaban que la fluorescencia de la enzima podría cambiar durante la actividad de la enzima con o sin sustrato. Con este fin, los inventores obtuvieron el espectro de fluorescencia y la cinética (Figura 18). Los inventores observaron que la fluorescencia y la cinética cambiaban notablemente con la concentración de sustrato. Puede ser útil utilizar parámetros de fluorescencia para determinar las concentraciones de ácidos grasos.

#### 25 **Las GMC oxidorreductasas de *Chlamydomonas* y *Phaeodactylum* son también alcano sintasas**

Para investigar la posibilidad de que otras GMC oxidorreductasas de algas tengan también actividad ácido graso descarboxilasa y por tanto sean alcano sintasas, las GMC oxidorreductasas de otras clorofíceas (*Chlamydomonas reinhardtii*) y una diatomea (*Phaeodactylum tricornutum*) se expresaron en *Escherichia coli* y los ácidos grasos totales y los hidrocarburos de las células bacterianas se analizaron mediante GC-MS-FID. La enzima de *Chlamydomonas* produjo la formación en células de *E. coli* de pentadecano y heptadecano, así como de sus análogos monoinsaturados (Figura 19). Los mismos compuestos se formaron con la enzima de *Phaeodactylum*, aunque en menores cantidades.

#### 30 **Expresión simultánea de una lipasa con la GMC oxidorreductasa impulsa la síntesis de alcanos en *E. coli***

Las combinaciones de ácidos grasos libres suelen ser pequeñas en las células vivas porque los ácidos grasos son perjudiciales para la estructura de la membrana. Para observar si la producción de alcanos y alquenos se podría impulsar en *E. coli* aumentando la cantidad de ácidos grasos libres disponibles, una lipasa bacteriana (Uniprot P04635) se expresó simultáneamente con la alcano sintasa de *Chlorella* o *Chlamydomonas* (Figura 20). Esto dio como resultado un aumento de dos a tres veces en la mayoría de los hidrocarburos, en comparación con las cepas de *E. coli* que expresan las GMC oxidorreductasas en solitario.

#### 45 **Discusión**

Los alcanos y los alquenos son compuestos interesantes para la producción de combustibles y especialmente interesantes para la industria química. En este trabajo, usando purificación parcial y análisis proteómico, los inventores fueron capaces de identificar una alcano sintasa de microalga derivada de *Chlorella*. Se trata de un miembro de la familia de la GMC oxidorreductasa. Cuando se expresa en *E. coli*, esta proteína sola es capaz de producir alcanos y alquenos. El principal interés de esta enzima es su aparente capacidad para catalizar una descarboxilación formal de los ácidos grasos libres para formar hidrocarburos saturados y el hecho de que sea una fotoenzima. El mecanismo y los posibles requerimientos de cofactores se esperan que sean diferentes de la citocromo P450 alcano sintasa bacteriana. Las GMC oxidorreductasas de *Chlorella* amplían de esta forma la combinación de enzimas sintetizadoras de alcanos y ofrecen nuevas oportunidades de aplicaciones biotecnológicas.

#### 50 ***El grupo de las GMC oxidorreductasa de algas***

La familia de la GMC oxidorreductasa se describió por primera vez en 1992. La comparación con la secuencia de la proteína de la glucosa deshidrogenasa, colina deshidrogenasa, glucosa oxidasa y metanol oxidasa de diferentes organismos (respectivamente: *Drosophila melanogaster*, *Escherichia coli*, *Aspergillus niger*, y *Hansenula polymorpha*) mostró baja similitud, pero motivos conservados. Estas enzimas contienen un sitio de flavoenzima y un pliegue  $\beta\alpha\beta$  canónico de unión a ADP cerca de su extremo amino. Los estudios estructurales confirman que estas proteínas están compuestas por un dominio de unión FAD en el extremo N, y un dominio de unión a sustrato en el extremo C. El dominio de unión a FAD forma el pliegue alfa-beta típico de las proteínas de unión a dinucleótidos, mientras que el dominio de unión a sustrato consiste en una beta lámina rodeada de alfa hélices. La topología general de estas

proteínas está conservada, aunque elementos estructurales insertados aparecen tanto en la colina deshidrogenasa como en la alcohol deshidrogenasa.

5 Los miembros de la familia de la GMC oxidorreductasa catalizan diversas reacciones, principalmente la oxidación de alcoholes a aldehídos. Esta familia incluye glucosa y metanol oxidasas, alcohol graso oxidasa, colina deshidrogenasa. Pero la familia también incluye una liasa de almendra que actúa sobre hidroximandelonitrilo, que muestra que la familia tiene mecanismos catalíticos muy diversos. La dependencia estricta de la actividad de la luz es probable que esté mediada por el cofactor FAD que se encuentra asociado a la enzima recombinante. La presencia de FAD era consistente con el hecho de que la enzima de *Chlorella* tiene un dominio de unión a FAD como todas las  
10 GMC oxidorreductasas (**Figura 15**).

De manera interesante, todas las especies de microalgas que han mostrado producir alca(que)nos de cadena larga o muy larga tienen un homólogo de la GMC oxidorreductasa de *Chlorella*, pero la única especie que no tiene alca(que)nos detectables (*Ostreococcus tauri*) no tiene homólogo de GMC. Parece por tanto bastante probable que los miembros del grupo de las GMC oxidorreductasas de algas sean todas alcano sintasas. Esta idea está respaldada por la demostración de que las GMC oxidasas de *Chlamydomonas reinhardtii* y *Phaeodactylum tricornutum* tienen actividad ácido graso descarboxilasa (**Figura 19**).

### 20 **Posibles aplicaciones biotecnológicas de la alcano sintasa de microalgas**

El descubrimiento de una ruta en microalgas para la síntesis de alcanos tiene interés biotecnológico porque las microalgas son una plataforma prometedora para la producción de lípidos, aunque la recogida de biomasa, extracción de aceite y conversión en biodiésel es muy costosa. La producción en microalgas de alcanos volátiles de tipo combustible que se pudieran recuperar fácilmente del medio de cultivo debería paliar estos problemas.

25 *In vitro*, la enzima de *Chlorella* es capaz de actuar sobre varios ácidos grasos, incluidos de cadena media (**Figura 10**). El hecho de que las enzimas actúan sobre los ácidos grasos libres es una clara ventaja en comparación con las enzimas de plantas, cianobacterias o insectos que actúan sobre los aldehídos grasos porque estos compuestos intermedios se pueden degradar mediante muchas reductasas endógenas (Rodríguez y Atsumi, 2014, *Metabolic Engineering*, 25, 227-37). Asimismo, los ácidos grasos libres se pueden generar mediante varios tipos de lipasas, algunas de las cuales son muy activas y pueden usar varios sustratos lípidos (Andersson, 1996, *Biochim Biophys Acta*, 1302, 236-240). El aumento de la disponibilidad de los ácidos grasos libres mediante la expresión simultánea de una lipasa impulsa la producción de alcanos (**Figura 20**). En *E. coli*, se produce algo de tridecano cuando la GMC oxidorreductasa de *Chlorella* se expresa (**Figura 6**). Este alcano es más corto que el observado en algas, y se debe a la presencia en la *E. coli* de un nuevo sustrato (ácido mirístico) para la enzima. Por consiguiente, las GMC oxidorreductasas de algas se podrían utilizar para generar grandes cantidades de alcanos de cadena media a corta, en particular cuando se usa junto con una lipasa en un microorganismo que acumula ácidos grasos de cadena corta a media.

40 Como la alcano sintasa es una fotoenzima, se puede usar luz para modular finamente la producción de alcanos *in vitro* e *in vivo*. En primer lugar, se puede usar la presencia o ausencia de fotones a 320-540 para seleccionar el momento de la producción de alcanos. En segundo lugar, se puede usar la intensidad de la luz para aumentar o disminuir la velocidad de síntesis de alcanos. En el presente documento se realizaron experimentos con luz continua, pero es posible que otras condiciones, como destellos, pudieran ser interesantes para la síntesis de hidrocarburos.

45 Finalmente, los productos (alcanos y alquenos), subproducto (CO<sub>2</sub>), y fluorescencia de la enzima se pueden usar para estimar la concentración de ácidos grasos libres en una muestra (o de ácidos grasos totales si se usa junto con una lipasa).

### 50 **Materiales y métodos**

#### **Cepas y condiciones de cultivo**

55 Se usaron las cepas de *Chlamydomonas reinhardtii* de tipo natural CC124 (nitrógeno; mt-) y CC125 (nitrógeno; 415 mt+). *Chlorella variabilis* NC64A tuvo como procedencia el laboratorio de J.L. Van Etten (Universidad de Nebraska). Todas las cepas se hicieron crecer de forma rutinaria en matraces cónicos en agitadores de incubación a 25 °C (Infors HT) con aire enriquecido con un 2 % (v/v) de CO<sub>2</sub>, con agitación a 140 rpm e intensidad de la luz a 120 μmol fotones m<sup>-2</sup>s<sup>-1</sup> para *Chlamydomonas* y 70 para *Chlorella*. *Chlamydomonas* y *Chlorella* se cultivaron en medio Tris-Acetato-Fosfato (TAP) y medio mínimo. Las células se contaron de forma rutinaria con un Multisizer<sup>TM</sup> 3 (Coulter).

#### 60 **Purificación de la alcano sintasa natural**

Se usó un sistema rápido para cromatografía líquida de proteínas (AKTApurifier900, GE Healthcare). El ensayo de actividad de la alcano sintasa se describe en la siguiente sección. Las células de *Chlorella* (200.10<sup>9</sup>) se centrifugaron durante una hora a 6.000 g y los aglomerados celulares se congelaron en nitrógeno líquido y se almacenaron durante una hora a menos 80 °C. Las células se resuspendieron en tampón de lisis que contenía Tris 20 mM (pH 8,0), NaCl

100 mM y EDTA 1 mM (tampón A) y se rompieron con un Cell Disruption (Constant) a una presión de 2 kbar. El homogenado se centrifugó dos veces durante 40 min a 50.000 g. El sobrenadante se recogió y se centrifugó durante 90 min a 105.000 g. Los aglomerados microsómicos resultantes se resuspendieron durante la noche a 4 °C con agitación en un tampón A al que se había añadido Triton X100 2,7 mM. Se llevó a cabo la ultracentrifugación del sobrenadante a 105.000 g durante 90 min y el sobrenadante se cargó en una columna de filtración en gel. Las fracciones más activas se combinaron, se concentraron usando un filtro de ultracentrífuga de 30 kDa Amicon® y el tampón se cambió por dilución a Tris 20 mM (pH 8,0), NaCl 50 mM, EDTA 1 mM, tampón Triton X100 al 0,05 % (p/v) (tampón B). La segunda etapa de purificación implicó una columna de intercambio aniónico (HiTrap Q. FF, GE Healthcare). las proteínas se eluyeron con un gradiente (0-100 %) de un tampón Tris 20 mM (pH 8,0), NaCl 1 M, EDTA 1 mM, y tampón Triton X100 al 0,05 % (p/v) (tampón C). Las fracciones más activas se combinaron, se concentraron usando un filtro de ultracentrífuga de 30 kDa Amicon® y el tampón se cambió por dilución al tampón B. La tercera etapa de purificación implicó una columna de intercambio aniónico fuerte (Mono Q. GI, GE Healthcare). Las proteínas se eluyeron usando un gradiente de tampón C. Las fracciones más activas se guardaron para análisis proteómico.

#### 15 **Ensayo de actividad para la purificación de proteínas**

Los ensayos enzimáticos se llevaron a cabo en viales de vidrio transparentes sellados usando tapones provistos de septo. Las mezclas de reacción contenían 500 µl de cada fracción de purificación, 200 µM de ácido palmítico perdeuterado (solución madre 10 mM en etanol) y 45 nmol de hexadecano como patrón interno (solución madre 4,5 mM en cloroformo). Los viales se agitaron a 120 rpm durante la noche a 25 °C con luz blanca (intensidad 120 µmol fotones m<sup>-2</sup>s<sup>-1</sup>). La reacción se detuvo por adición de 10 µl de NaOH 10 M a través del septo usando una jeringa. Los hidrocarburos producidos se analizaron incubando el espacio superior del vial en una fibra para microextracción en fase sólida (SPME) (sílice fundida DVB PDMS, 65 µm doble-polar, Supelco) montada en un soporte. Después de 15 min de incubación a temperatura ambiente, la fibra de SPME se introdujo inmediatamente en el inyector del GC-MS y se desorbió a 250 °C. El análisis mediante GC-MS se llevó a cabo como se describe a continuación.

#### **Análisis proteómico**

La preparación de la proteína, la digestión en gel y los análisis por nano-LC-MS/MS se llevaron a cabo como se ha descrito anteriormente. En resumen, las proteínas solubilizadas en tampón Laemmli se apilaron sobre un gel NuPAGE al 4-12 % (p/v) (Invitrogen) y se tiñeron con azul de Coomassie R-250 (BioRad). Las bandas del gel se recortaron, y las proteínas del gel se digirieron con tripsina (Promega). Los péptidos resultantes se analizaron mediante cromatografía de nanolíquidos acoplada a espectrometría de masas en tándem (Ultimate 3000 acoplado a LTQ-Orbitrap Velos Pro, Thermo Scientific) usando un gradiente de 120 min. Los péptidos y las proteínas se identificaron mediante búsquedas contaminantes en las bases de datos Uniprot (*Chlorella variabilis* taxonomy, versión de septiembre de 2016), contaminantes clásicos (propia) y las correspondientes bases de datos invertidas usando Mascot (versión 2.5.1). Se usó el programa informático Proline para filtrar los resultados (conservación de péptidos de rango 1, identificación de péptidos FDR < 1 % calculado por puntuaciones de péptidos usando la estrategia de la base de datos invertida, longitud del péptido > 7, y un mínimo de 1 péptido específico por grupo de proteína identificado) antes de realizar una compilación, agrupación y comparación de los grupos de proteínas procedentes de diferentes muestras. Solamente las proteínas identificadas con un recuento espectral mínimo de 2 se tuvieron en cuenta para comparación adicional.

#### **Análisis de proteínas y transferencias Western**

A los extractos de proteínas se añadió colorante de carga LDS NuPAGE IX final, se ebulló durante 10 min a 95°C, se resolvió usando SDS-PAGE reductor al 10 % (p/v) con tampón de análisis MOPS y se tiñó con nitrato de plata. Para la detección de las proteínas marcadas con His, los polipéptidos resueltos por SDS-PAGE se transfirieron a una membrana de nitrocelulosa usando un sistema de transferencia semiseco, y las etiquetas His se revelaron usando anticuerpos de conejo dirigidos contra Ig de conejo marcada con His conjugado con peroxidasa de rábano picante y sustrato ECL (Amersham Biosciences).

#### **Clonación de ADNc de alcano sintasa y purificación de alcano sintasa recombinante**

Los ARN totales se extrajeron de células de *Chlorella* por un método de fenol-cloroformo y los ADNc se sintetizaron usando transcriptasa inversa Superscript® III. El ADNc que codifica la GMC oxidorreductasa se amplificó usando cebadores diseñados en las posibles UTR 5' y 3' Cebador directo: ATGGCGTCAATTACATCGCG (SEQ ID No 24); Cebador inverso: TCATGCTGCCACTGTGCG (SEQ. ID No 25), se clonó en un plásmido TOPO XL y se secuenció. La secuencia correspondiente a los restos 62-654 de la alcano sintasa se amplificó a partir de un gen sintético optimizado por codones para la expresión en *E. coli* usando un cebador directo 5'-CTG TAC TTC CAA TCA GCC AGC GCA GTT GAA GAT ATT C-3' (SEQ ID No 27) y un cebador inverso: 5'-TAT CCA CCT TTA CTG TTA TCA TGC AAC GGTTGC CGG TG-3' (SEQ ID No 28). y se clonó en el vector pLIC07, que introdujo después del codón de inicio ATG un casete que codificaba una tiorredoxina con 6 etiquetas His y un sitio de escisión de proteasa del virus del grabado del tabaco (TEV). La alcano sintasa recombinante se produjo en células BL21-CodonPlus (DE3)-RIL *E. coli* cultivadas en medio TB a 37°C hasta DO 0,9. En esta etapa, la temperatura se disminuyó hasta 17 °C y las células se hicieron crecer durante 18 h más. Las células se recogieron mediante centrifugación (4.000 g durante 10 min) y el aglomerado

se congeló. El aglomerado de células se resuspendió en tampón de lisis durante 30 min a 4 °C (10 ml de tampón de lisis para un litro de células a DO=1). El tampón de lisis contenía NaCl 300 mM, Tris 50 mM pH 8,0, imidazol 10 mM, glicerol al 5 % (p/v), 0,25 mg ml<sup>-1</sup> de lisozima, MgSO<sub>4</sub> 20 mM, 10 µg ml<sup>-1</sup> de DNasa, y antiproteasas. Tras la resuspensión, las células se lisaron mediante sonicación, y se centrifugaron durante 30 min a 8.000 g. El sobrenadante se recogió y la enzima se purificó mediante FPLC. La primera purificación se llevó a cabo en una columna de níquel y la proteína se eluyó mediante un gradiente en etapas usando 50 % (v/v) de un segundo tampón que contenía NaCl 300 mM, Tris 50 mM pH 8,0, imidazol 500 mM, glicerol al 5 % (p/v). La proteasa del virus del grabado del tabaco (a 1 mg por 10 mg de proteína total) se usó para cortar la etiqueta His y la tiorredoxina. Se llevó a cabo una diálisis durante la noche en presencia del TEV para cambiar el tampón a un tampón que contenía NaCl 300 mM, Tris 50 mM pH 8,0, imidazol 10 mM, glicerol al 5 % (p/v). Se llevó a cabo una segunda cromatografía FPLC en una columna de níquel para separar la proteína de la tiorredoxina marcada con His. La última etapa de purificación fue una columna de filtración en gel (Superdex200 26/600mm GE Healthcare). El tampón usado en esta etapa contenía NaCl 150 mM, Tris 10 mM pH 8,0, glicerol al 5 % (p/v). La proteína se concentró usando filtros de ultracentrifugación de 50 kDa Amicon® y se almacenaron a -80 °C después de añadir glicerol al 20 % (p/v).

#### **Expresión de GMC oxidorreductasa de *Chlorella variabilis*, *Chlamydomonas reinhardtii* y *Phaeodactylum tricornutum* en *E. coli*.**

Alcano sintasa de *Chlamydomonas reinhardtii* y *Phaeodactylum tricornutum* se amplificó a partir de un gen sintético optimizado por codones para *E. coli* usando para *Chlamydomonas reinhardtii* un cebador directo 5'-TAC TTC CAA TCA ATG CTG GGT CCG AAA ACC -3' (SEQ ID No 29) y un cebador inverso, 5'-TAT CCA CCT TTA CTG TTC TAC TAA ACT GCA ACC GGC TGA CG -3' (SEQ ID No 30). Para *Phaeodactylum tricornutum*, el cebador directo es 5'-TAC TTC CAA TCA ATG AAA AGC CTG CGT AGC-3' (SEQ ID No 31), el cebador inverso es 5'-TAT CCA CCT TTA CTG TTC TAC TAT GCG CTT GCG GTG-3' (SEQ ID No 32). Los genes se clonaron en el vector pLIC07, que introdujo después del codón de inicio ATG un casete que codificaba una tiorredoxina con 6 etiquetas His y un sitio de escisión de proteasa del virus del grabado del tabaco (TEV). Las cepas de *E. coli* que expresan la GMC oxidasa de *Chlorella variabilis* NC64A, *Chlamydomonas reinhardtii* o *Phaeodactylum tricornutum* se hicieron crecer a 37°C con agitación a 180 rpm y luz a 100 µmol.fotón.m<sup>-2</sup>.s<sup>-1</sup>. Cuando la DO alcanzó 0,6, se añadieron 500 µM de isopropil-β-D-1-tiogalactopiranosido. A continuación, las células se hicieron crecer durante 24 horas a 37°C, se recogieron, se transmetilaron con metanol al que se había añadido ácido sulfúrico al 5 %, y los hidrocarburos se extrajeron con hexano y se analizaron mediante GC-MS como se ha descrito anteriormente.

#### **Expresión simultánea en *E. coli* de una GMX oxidorreductasa y una lipasa**

Las cepas de *E. coli* transformadas con el vector que expresa la GMC oxidasa de *Chlorella variabilis* NC64A (o *Chlamydomonas reinhardtii*) y/o un vector que expresa la lipasa de la bacteria *Staphylococcus hyicus*, se hicieron crecer en medio TB a 37°C. La expresión se indujo con 1 mM de isopropil-β-D-1-tiogalactopiranosido (a que se había añadido un 0,2 % de arabinosa para la expresión simultánea). Las células se hicieron crecer durante la noche a 25°C a 100 µmol.fotón.m<sup>-2</sup>.s<sup>-1</sup> y 6 h a 2.000 µmol.fotón.m<sup>-2</sup>.s<sup>-1</sup>. Las células se recogieron (4 ml a DO=5), se transmetilaron usando metanol al que se había añadido ácido sulfúrico al 5 % y los hidrocarburos se analizaron mediante SPME y GC-MS como se ha descrito anteriormente.

#### **Ensayo enzimático con enzima purificada**

Todos los ensayos se llevaron a cabo en viales de vidrio transparentes sellados usando tapones provistos de septo. El pH óptimo se determinó con un tampón universal Teorell Stenhagen (ácido cítrico monohidrato 33 mM, ácido fosfórico 33 mM, NaOH 100 mM, ácido bórico 16,7 mM, pH 8,5 ajustado con HCl 1 N). Otros ensayos se realizaron en Tris HCl 100 mM pH 8,5, NaCl 100 mM. Las mezclas de reacción (500 µl) contenían de forma típica 160 nM enzima purificada (solución madre 2,5 mg ml<sup>-1</sup>) y 400 µM de sustrato (solución madre 10 mM en etanol). En algunos ensayos, se usó una lipasa junto con la alcano sintasa. En este caso, el sustrato fue un triacilglicerol. En general, las muestras se agitaron a 200 rpm durante 15 min bajo luz blanca hecha con LED a 2.000 µmol.fotón.m<sup>-2</sup>.s<sup>-1</sup>. Tras las incubaciones, las muestras se calentaron a 95 °C durante 15 min para detener la reacción enzimática. Las muestras se enfriaron y se añadió patrón interno (hexadecano) (45 nmol de una solución madre 4,5 mM en cloroformo). A continuación, se añadió NaOH a la mezcla de reacción (10 µl de una solución madre de 10 M) y las muestras se vortizaron durante 5 min. A continuación, se añadieron 250 µl de hexano, y las muestras se vortizaron durante 5 min para extraer los alcanos y los alquenos. La fase de hexano se recogió mediante centrifugación y se analizó mediante GC-MS-FID. El análisis se llevó a cabo por inyección directa de 100 µl del espacio vacío en una GC-MS. En la Fig. 11 la actividad en l-<sup>13</sup>C-palmitato se controló por liberación de <sup>13</sup>C<sub>2</sub> usando espectrometría de masas con entrada de membrana. Las iluminaciones se proporcionaron con LED CBT-120 (Luminous, Billerica) tanto azul (máximo a 460 nm, 25 nm FWHM), o rojo (máximo a 635 nm, 15 nm FWHM) con una intensidad de 850 µmol.fotón.m<sup>-2</sup>.s<sup>-1</sup>.

#### **Fluorescencia de la alcano sintasa**

La enzima se analizó por espectroscopia UV-Vis (espectrofotómetro Uvikon XS de Secomam). El espectro de absorbancia se midió sobre la enzima purificada en un tampón que contenía Tris 100 mM pH 8,5, y NaCl 100 mM. El espectro de fluorescencia (500 a 700 nm) se midió en un Varian Cary Eclipse usando un flujo de excitación a 450 nm

con ranura de 10 nm para cinética, la fluorescencia se midió a 540 nm usando un flujo de excitación a 450 nm con ranura de 10 nm.

### **Espectrometría de masas con entrada de membrana (MIMS)**

Las mediciones en línea de  $^{12}\text{CO}_2$  ( $m/z = 44$ ) y  $^{13}\text{CO}_2$  ( $m/z = 45$ ) se controlaron usando espectrometría de masas (modelo Prima B, Thermo Scientific). El sistema de entrada de membrana consistió en una cámara de electrodo de oxígeno termorregulada (Hansa Tech), que estaba conectada con la línea de vacío del espectrómetro de masas mediante una membrana fina de teflón permeable a gases (espesor 0,03 mm (0,001 pulgadas), YSI Inc.), que precinta la parte inferior de la cámara. Para los análisis, 20  $\mu\text{l}$  de enzima purificada a 2,5  $\text{mg ml}^{-1}$  y 30  $\mu\text{l}$  de sustrato a 10 mM en dimetil sulfóxido ( $^{13}\text{C}$ -ácido palmítico) se añadieron a 1,45 ml de tampón Tris/Acetato/Borato 100 mM, pH 6,5 que contenía NaCl 100 mM, se introdujo en la cámara de medición, termorregulada a 25 °C, y se agitó continuamente. Los gases disueltos en el medio se difundieron a través de la membrana de teflón hasta la fuente de iones del espectrómetro de masas.

### **Cultivos en fotobiorreactores**

*Chlamydomonas reinhardtii* CC124 (nit1 nit2; mt-) y *Chlamydomonas reinhardtii* que expresaba en exceso el gen de la alcano sintasa se cultivaron en medio mínimo (Harris, 1989) en fotobiorreactores de un litro (BIOSTAT Aplus, Sartorius Stedim Biotech) operados como turbidostatos. Se midió  $A_{880}$  de forma continua usando una sonda de biomasa (Excellprobe, Exner) y los cultivos se mantuvieron a una  $A_{880}$  constante por inyección en medio nuevo. El pH se mantuvo a un valor constante de 7 por inyección de KOH (0,2 N) o HCl (0,2 N). Los cultivos se agitaron usando una hélice metálica (250 rpm). El caudal de gas se ajustó a 0,5  $\text{l min}^{-1}$ . Aire enriquecido con  $\text{CO}_2$  al 2 % (v/v) se generó usando caudalímetros máxicos (EL flow, Bronkhorst). La luz blanca se proporcionó mediante ocho tubos fluorescentes (Osram Dulux L 18 W) situados radialmente alrededor del fotobiorreactor. Los inventores usaron un filtro azul (363 special medium blue, Lee filters, USA) y un filtro rojo (113 magenta, Lee filters, USA) para proporcionar respectivamente luz azul y roja. Ambas luces tenían la misma intensidad (35  $\mu\text{mol.fotón.m}^{-2}.\text{s}^{-1}$ ).

### **Transmetilación**

Para cuantificar los hidrocarburos junto con los ácidos grasos, se usó la transmetilación de células completas. En resumen, aglomerados celulares (cien millos de células para *Chlamydomonas*, doscientos millones de células para *Chlorella variabilis* NC64A y 20 ml de unidades de  $\text{DO}^{-1}$  de *E. coli*) recibieron 2 ml de una solución que contenía metanol con ácido sulfúrico al 5 % (v/v) y 25  $\mu\text{g}$  de triheptadecanoato (de una solución madre de 2,5  $\text{mg ml}^{-1}$  en cloroformo) y 5  $\mu\text{g}$  de 16:0-alcano (solución madre de 1  $\text{mg ml}^{-1}$  en cloroformo) se incluyeron como patrones internos. Las muestras se incubaron a 85 °C durante 90 min en tubos de vidrio precintados. Después de enfriar, FAME y los hidrocarburos se extrajeron por adición de 250  $\mu\text{l}$  de hexano y 500  $\mu\text{l}$  de NaCl al 0,9 % (p/v). Las muestras se vortizaron durante 10 min y la fase orgánica se separó de la fase acuosa por centrifugación a 3.000 g durante 2 min. La fase de hexano se recuperó, y 1  $\mu\text{l}$  se inyectó en el GC-MS/FID.

### **Análisis por GC-MS**

Los análisis mediante cromatografía de gases acoplada a espectrometría de masas (GC-MS), que se realizaron después de la microextracción en fase sólida (SPME), se llevaron a cabo con la siguiente configuración. Un cromatógrafo de gases de Thermo-Fischer serie Focus acoplado a un espectrómetro de masas Thermo-Fischer DSQII (cuadripolo simple) se usó con una columna capilar polar DB-5HT (Agilent) (longitud 30 m, diámetro interno 0,25 mm, espesor de película 0,1  $\mu\text{m}$ ). El transportador de helio se ajustó a 1  $\text{ml min}^{-1}$ . La temperatura del horno se programó con un tiempo de espera inicial de 2 min a 50°C, luego una rampa de 50°C a 300°C a 10°C  $\text{min}^{-1}$ , y un tiempo de espera final de 3 min a 300°C. Las muestras se inyectaron en modo sin división (2 min) a 250 °C. El MS se realizó en modo barrido completo para 40-500 amu (ionización por impacto de electrones, 70 eV) y los picos se cuantificaron en función de la corriente total de iones usando patrones internos. Para la determinación del cosustrato, se usó una columna HP-PLOT Q (0,32 mm diámetro X 30 m) y el  $\text{CO}_2$ ,  $^{13}\text{CO}_2$  y argón se analizaron usando una temperatura de horno de 40 °C y monitorización de ion simple ( $m/z$  40, 44, 45).

### **Análisis por GC-MS/FID**

Los análisis mediante cromatografía de gases acoplada a espectrometría de masas y detector de ionización por llama (GC-MS/FID) se llevaron a cabo solamente después de las reacciones de transmetilación para cuantificar los ácidos grasos y los hidrocarburos conjuntamente. Los análisis se llevaron a cabo en un cromatógrafo de gases Agilent 7890A acoplado a un espectrómetro de masas Agilent 5975C (cuadripolo simple). Se usó una columna capilar polar Zebron 7HG-G007-11 (Phenomenex) (longitud 30 m, diámetro interno 0,25 mm, espesor de película 0,25  $\mu\text{m}$ ). El transportador de hidrógeno se ajustó a 1  $\text{ml min}^{-1}$ . La temperatura del horno se programó con un tiempo de espera inicial de 2 min a 60 °C, una primera rampa de 60°C a 150°C a 20°C  $\text{min}^{-1}$  con un tiempo de espera de 5 min a 150°C, después una segunda rampa de 150 °C a 240 °C a 6 °C  $\text{min}^{-1}$  y un tiempo de espera final de 3 min a 240 °C. Las muestras se inyectaron en modo sin división (1 min) a 250 °C. El MS se analizó en modo de barrido completo 40-350 amu (ionización por impacto de electrones a 70 eV) y los picos se cuantificaron según la señal FID usando los patrones

internos.

**Filogenia**

- 5 Para construir el árbol filogenético, 56 secuencias de aminoácidos de GMC oxidorreductasas derivadas de procariotas y eucariotas se recuperaron de Cyanobase (<http://genome.kazusa.or.jp/cyanobase/>), NCBI (<http://www.ncbi.nlm.nih.gov/>), Phytozome (<https://phytozome.jgi.doe.gov>) o Cyanidioschyzon merolae (<http://merolae.biol.s.u-tokyo.ac.jp/>). Las secuencias se alinearon con el programa MAFFT versión 7. El alineamiento resultante se refinó manualmente con SeaView versión 4 y las regiones donde la homología era dudosa se eliminaron para análisis posterior. Un total de 266 posiciones de aminoácidos se mantuvieron para el análisis filogenético. El árbol se obtuvo mediante Neighbor-Joining (NJ), enfoque del Phylogenetic Inference Package Phylip versión 3.69. El programa PROTDIST se usó para crear matrices de distancia. El programa NEIGHBOR se usó para el análisis NJ, y el orden de introducción de secuencias se aleatorizó (20 jumbles). Los programas SEQBOOT y CONSENSE se usaron para impulsar los cálculos de los valores sobre 100 réplicas y reconstrucciones del árbol de consenso, respectivamente.
- 10
- 15 Los árboles filogenéticos se trazaron con Dendroscope versión 3.

## REIVINDICACIONES

1. Uso de un polipéptido o de una célula que expresa el polipéptido para producir alcanos/alquenos a partir de ácidos grasos, en donde el polipéptido tiene actividad ácido graso descarboxilasa y comprende una secuencia que tiene al menos un 40 % de identidad con la SEQ ID No 1.
2. El uso de acuerdo con la reivindicación 1, en donde el polipéptido comprende la secuencia de consenso de G-X-L-(X)<sub>4</sub>-C-[D/E]-X-G-[A/G]-F-X-[K/R] (SEQ ID No 4), siendo X un aminoácido.
3. El uso de acuerdo con las reivindicaciones 1 o 2, en donde uno, dos, tres, cuatro, cinco o seis aminoácidos en las posiciones C372, R391, Y406, Q426, H512 y N515 de la SEQ ID No 1 están conservados.
4. El uso de acuerdo con una cualquiera de las reivindicaciones 1-3, en donde al menos un 40 % de los restos de aminoácido situados entre las posiciones 388-428 son restos hidrófobos seleccionados de los grupos que consisten en V, I, L, M, F, W, C, A, e Y.
5. El uso de acuerdo con una cualquiera de las reivindicaciones 1-4, en donde el polipéptido tiene una secuencia de aminoácidos seleccionada del grupo que consiste en:
- las SEQ ID No 2-3 y 5-14 y
  - las secuencias de aminoácidos que tienen al menos un 80 % de identidad con una de las SEQ ID No 2-3 y 5-14 y que comprende una secuencia que tiene al menos un 40 % de identidad con la SEQ ID No 1.
6. El uso de acuerdo con una cualquiera de las reivindicaciones 1-4, en donde el polipéptido tiene una secuencia de aminoácidos que tiene al menos un 80 % de identidad con la SEQ ID No 1.
7. El uso de acuerdo con una cualquiera de las reivindicaciones 1-6, en donde el polipéptido es de algas, preferentemente microalgas o cianobacterias.
8. El uso de acuerdo con una cualquiera de las reivindicaciones 1-7, en donde los ácidos grasos y los correspondientes alcanos/alquenos descarboxilados comprenden de 8 a 24 átomos de carbono preferentemente de 12 a 22 átomos de carbono.
9. El uso de acuerdo con una cualquiera de las reivindicaciones 1-8, en donde los ácidos grasos y los correspondientes alcanos/alquenos descarboxilados están sustituidos por uno o varios grupos funcionales tales como un hidroxilo, un alcohol C<sub>1</sub>-C<sub>3</sub>, un acilo C<sub>1</sub>-C<sub>3</sub>, un éster C<sub>1</sub>-C<sub>3</sub>, una amina C<sub>1</sub>-C<sub>3</sub>, un grupo amino, una amida C<sub>1</sub>-C<sub>3</sub>, un carboxilo, un aldehído, un epoxi, un halógeno, un alcoxi C<sub>1</sub>-C<sub>3</sub>, un tioalquilo C<sub>1</sub>-C<sub>3</sub>, una imina C<sub>1</sub>-C<sub>3</sub>, un nitrilo, un grupo sulfuro tal como sulfona C<sub>1</sub>-C<sub>3</sub> o sulfóxido C<sub>1</sub>-C<sub>3</sub>, un tiol, un nitro, un ciano, un haloalquilo C<sub>1</sub>-C<sub>3</sub>, y/o están interrumpidos por uno o varios grupos funcionales tales como O, N o S, un grupo acetilénico, un éter tal como un grupo divinil éter o un grupo oxo.
10. Un método para producir alcanos/alquenos a partir de ácidos grasos, en donde un polipéptido como se define en una cualquiera de las reivindicaciones 1-7 se pone en contacto con ácidos grasos y se somete a luz con una longitud de onda comprendida entre 300 y 540 nm, preferentemente entre 400 y 520 nm.
11. Una célula hospedadora recombinante diseñada mediante ingeniería genética para producir alcanos/alquenos que comprende una secuencia de ácido nucleico que codifica un polipéptido heterólogo que tiene una actividad ácido graso descarboxilasa y que comprende una secuencia que tiene al menos un 95 % de identidad con una secuencia de aminoácidos seleccionada entre las SEQ ID No 1-3 y 5-14, siendo la célula hospedadora preferentemente una bacteria, una microalga, un hongo filamentoso o una levadura.
12. La célula hospedadora recombinante de acuerdo con la reivindicación 11, en donde el polipéptido heterólogo se expresa simultáneamente con una tioesterasa, preferentemente en una microalga o una cianobacteria y/o con una lipasa, preferentemente en una bacteria.
13. Un método para producir alcanos/alquenos a partir de ácidos grasos, en donde se cultiva una célula hospedadora recombinante de acuerdo con una cualquiera de las reivindicaciones 11-12 y se recuperan los alcanos/alquenos.
14. Uso de un polipéptido como se define en una cualquiera de las reivindicaciones 1-7 para la dosificación de ácidos grasos libres en una muestra.
15. Un método para dosificar ácidos grasos libres en una muestra, que comprende
- a. poner en contacto la muestra con una ácido graso descarboxilasa como se define en cualquiera de las reivindicaciones 1-6 en condiciones adecuadas para convertir los ácidos grasos en alcanos/alquenos,
  - b. recuperar los alcanos/alquenos y/o el CO<sub>2</sub>; y

c. cuantificar los alcanos/alquenos y/o el CO<sub>2</sub>.

16. Un método para dosificar ácidos grasos libres en una muestra, que comprende

- 5       a. poner en contacto la muestra con una ácido graso descarboxilasa como se define en una cualquiera de las reivindicaciones 1-7 en condiciones adecuadas para convertir los ácidos grasos en alcanos/alquenos,  
      b. medir la fluorescencia emitida por la ácido graso descarboxilasa.

10      17. El método de las reivindicaciones 15 o 16, en donde el método comprende una etapa previa de poner en contacto la muestra con una lipasa.

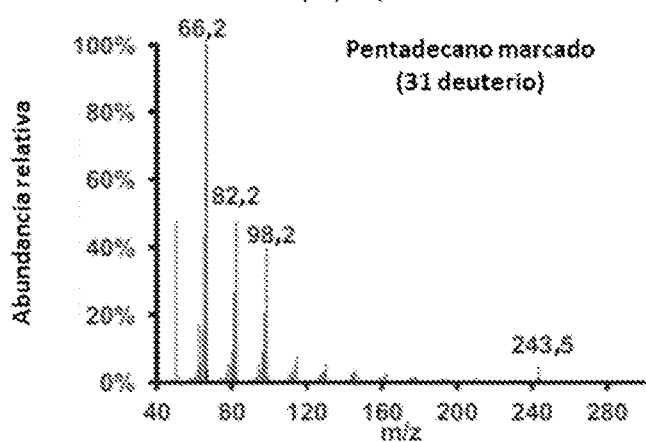
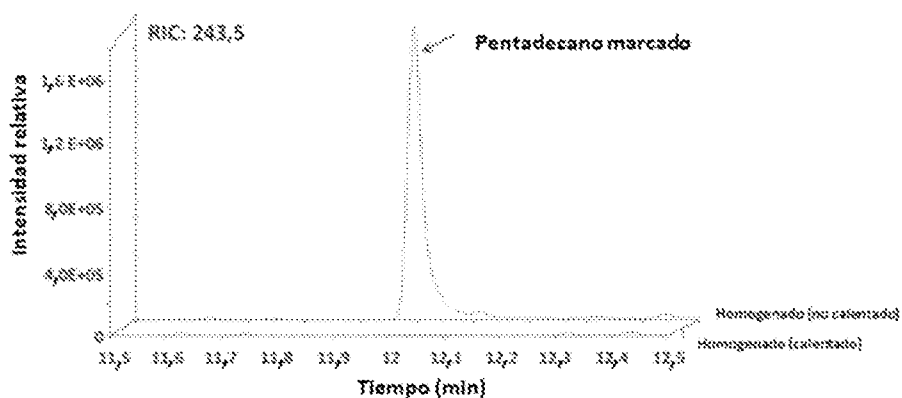


FIGURA 1

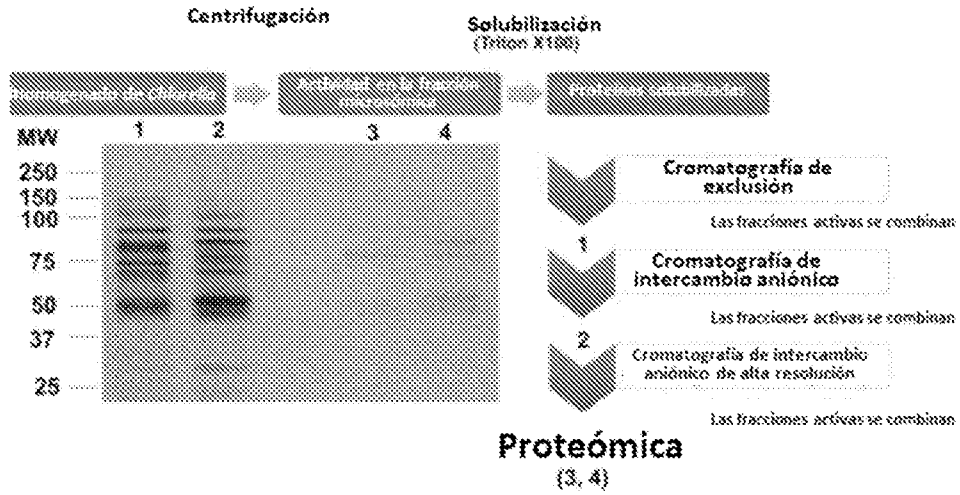


FIGURA 2

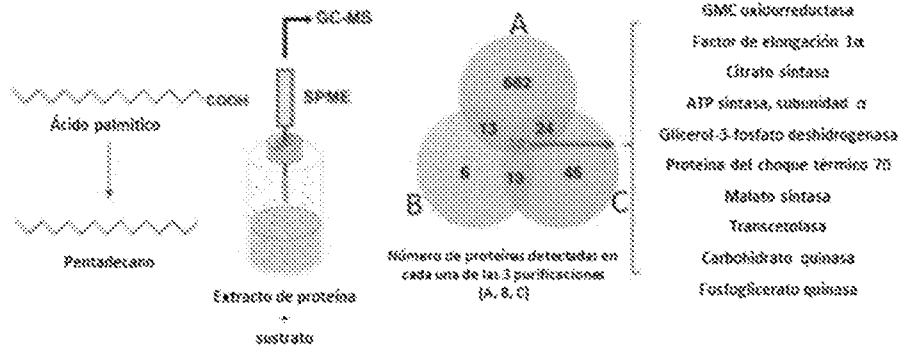
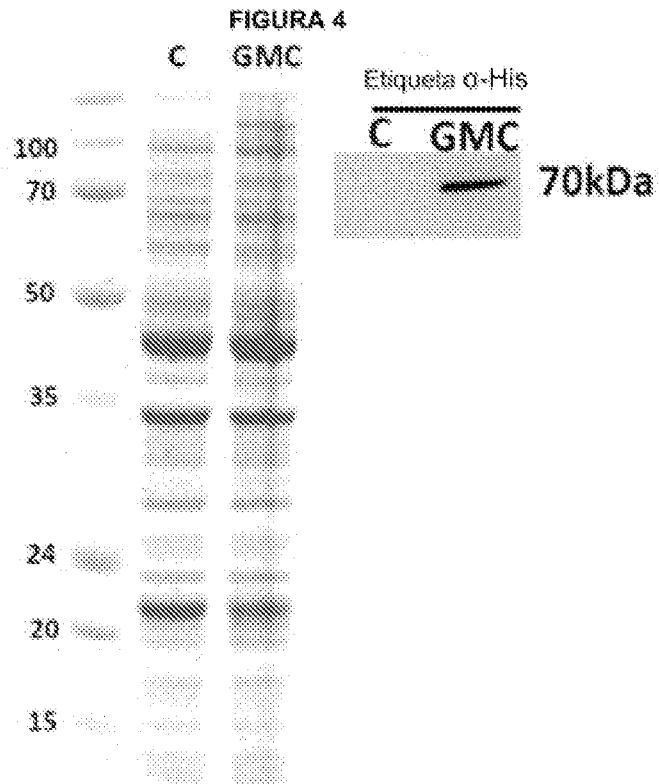


FIGURA 3



**FIGURA 5**

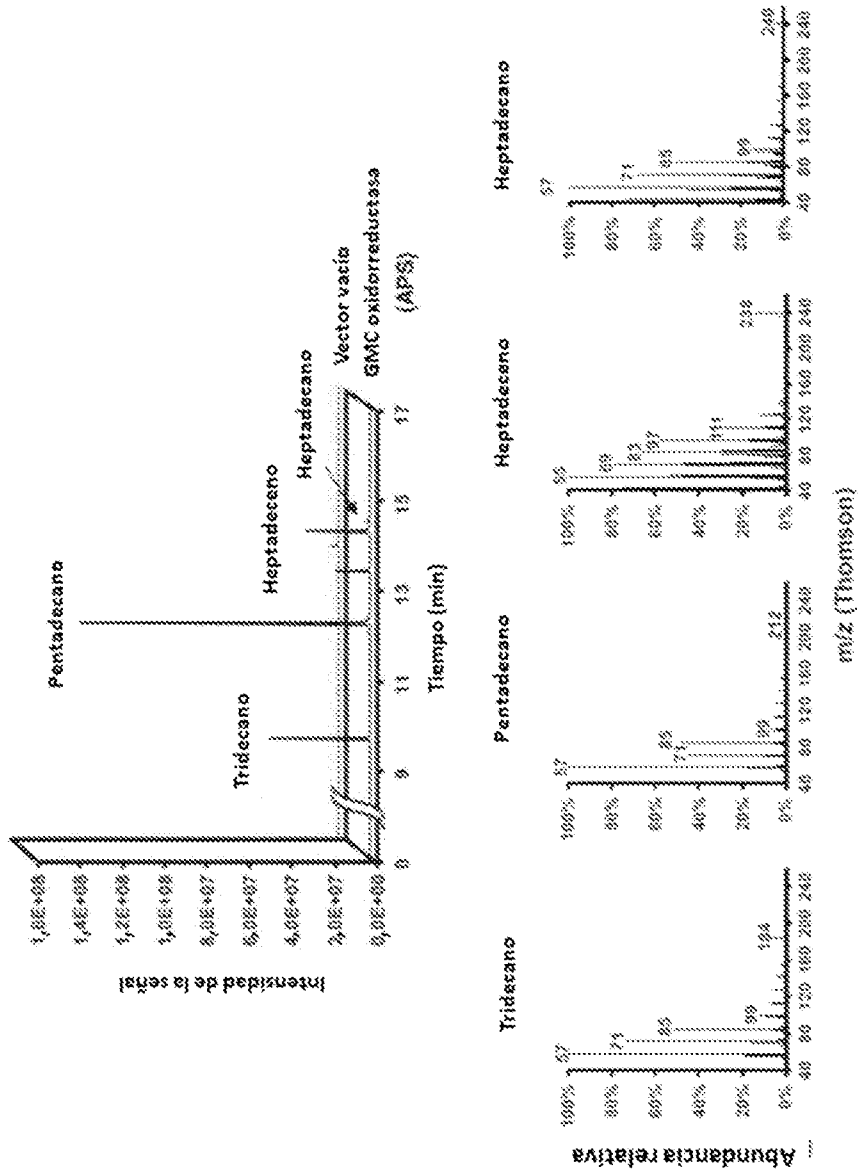


FIGURA 5

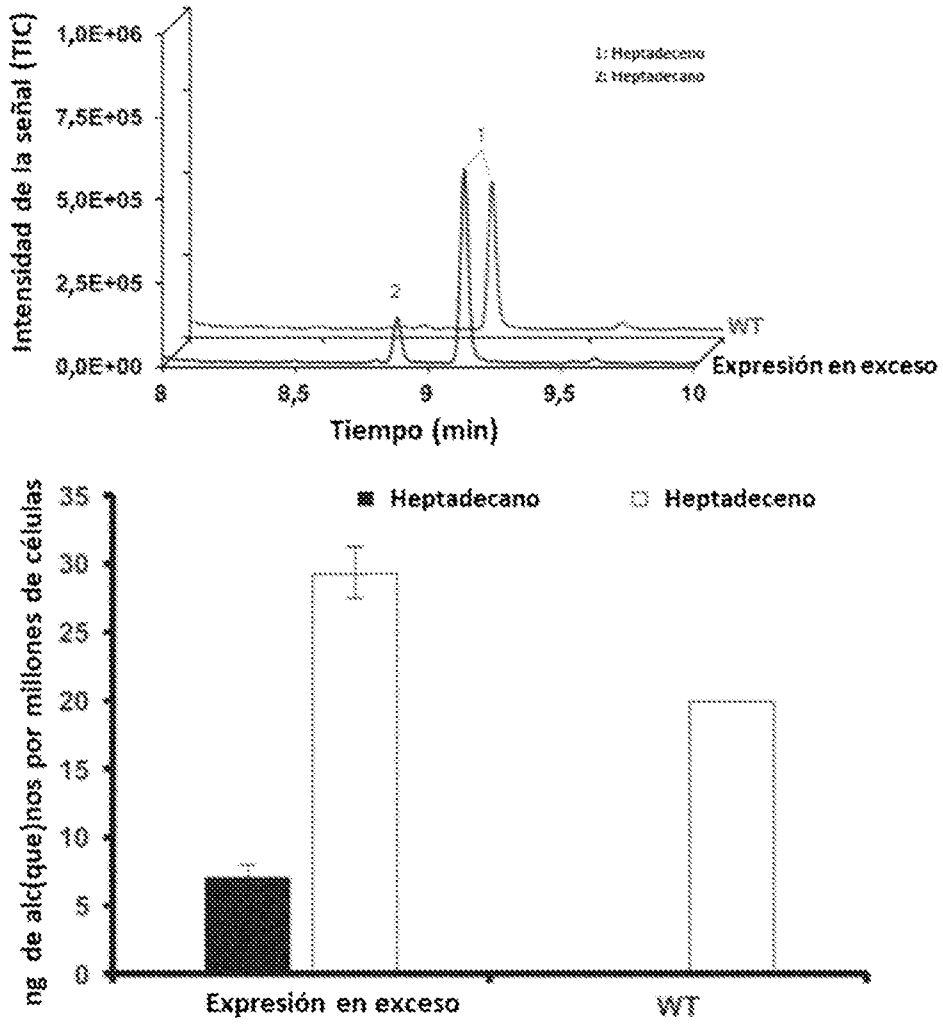


FIGURA 7





1 INKQKALMIRKQVLEKVVLEK--KMLQVLEKAVTQVNIQQA---AKKQKALOVEFQYDQHPD 284  
 5 EKKQKADPTTPTTLKPAKAE--GNLKVVIKARATVVNIKQ----SSQARTTQVVEYDQK-QPS 288  
 9 KKHKPADCTFMKLEFVMEK--KMLVVIKIKKTLKHEKED---GKKTVEKQVTFQVQKQDQ 292  
 10 KKHKADPTTPTTLKPAKAE--GNLKVVIKARATVVNIKQ---STQFARKQVVEYDQK-QPS 296  
 11 KKHKPADCTFMKLEFVMEK--KMLVVIKIKKTLKHEKED---GKKTVEKQVTFQVQKQDQ 298  
 12 KAKKRCYAKQKYLAKAKAE--ANLVVVTGKQVTLALDKAYDQKQVLEKQVVEYDQK-QPS 299  
 13 KKHKPADCTFMKLEFVMEK--KMLVVIKIKKTLKHEKED---GKKTVEKQVTFQVQKQDQ 303  
 7 VNKKADPTTPTTLKPAKAE--GNLKVVIKARATVVNIKQ---STQFARKQVVEYDQK-QPS 299  
 14 KKHKPADCTFMKLEFVMEK--KMLVVIKIKKTLKHEKED---GKKTVEKQVTFQVQKQDQ 306  
 \* \* \* \* \*

1 ENLQKALMIRKQVLEKVVLEK--KMLQVLEKAVTQVNIQQA---AKKQKALOVEFQYDQHPD 304  
 5 EKKQKADPTTPTTLKPAKAE--GNLKVVIKARATVVNIKQ----SSQARTTQVVEYDQK-QPS 308  
 9 KKHKPADCTFMKLEFVMEK--KMLVVIKIKKTLKHEKED---GKKTVEKQVTFQVQKQDQ 312  
 10 KKHKADPTTPTTLKPAKAE--GNLKVVIKARATVVNIKQ---STQFARKQVVEYDQK-QPS 316  
 11 KKHKPADCTFMKLEFVMEK--KMLVVIKIKKTLKHEKED---GKKTVEKQVTFQVQKQDQ 318  
 12 KAKKRCYAKQKYLAKAKAE--ANLVVVTGKQVTLALDKAYDQKQVLEKQVVEYDQK-QPS 319  
 13 KKHKPADCTFMKLEFVMEK--KMLVVIKIKKTLKHEKED---GKKTVEKQVTFQVQKQDQ 323  
 7 VNKKADPTTPTTLKPAKAE--GNLKVVIKARATVVNIKQ---STQFARKQVVEYDQK-QPS 319  
 14 KKHKPADCTFMKLEFVMEK--KMLVVIKIKKTLKHEKED---GKKTVEKQVTFQVQKQDQ 326  
 \*\* \* \* \* \* \*

1 CLTRAPVVEKQVTFQVQKQDQ---KQKQKALOVEFQYDQHPD---AKKQKALOVEFQYDQHPD 331  
 5 EKKQKADPTTPTTLKPAKAE--GNLKVVIKARATVVNIKQ----SSQARTTQVVEYDQK-QPS 335  
 9 KKHKPADCTFMKLEFVMEK--KMLVVIKIKKTLKHEKED---GKKTVEKQVTFQVQKQDQ 339  
 10 KKHKADPTTPTTLKPAKAE--GNLKVVIKARATVVNIKQ---STQFARKQVVEYDQK-QPS 343  
 11 VTNKADPTTPTTLKPAKAE--GNLKVVIKARATVVNIKQ---STQFARKQVVEYDQK-QPS 347  
 12 VVVVYS--KQKQ---KQKQKALOVEFQYDQHPD---AKKQKALOVEFQYDQHPD 350  
 13 CLTRAPVVEKQVTFQVQKQDQ---KQKQKALOVEFQYDQHPD---AKKQKALOVEFQYDQHPD 353  
 7 AVVVEKQVTFQVQKQDQ---KQKQKALOVEFQYDQHPD---AKKQKALOVEFQYDQHPD 354  
 14 VTIANNI--TRPI---KQKQKALOVEFQYDQHPD---AKKQKALOVEFQYDQHPD 360  
 \* \* \* \* \*

1 GQKQKALOVEFQYDQHPD---AKKQKALOVEFQYDQHPD---AKKQKALOVEFQYDQHPD 400  
 5 GQKQKALOVEFQYDQHPD---AKKQKALOVEFQYDQHPD---AKKQKALOVEFQYDQHPD 404  
 9 GQKQKALOVEFQYDQHPD---AKKQKALOVEFQYDQHPD---AKKQKALOVEFQYDQHPD 408  
 10 GQKQKALOVEFQYDQHPD---AKKQKALOVEFQYDQHPD---AKKQKALOVEFQYDQHPD 412  
 11 GQKQKALOVEFQYDQHPD---AKKQKALOVEFQYDQHPD---AKKQKALOVEFQYDQHPD 416  
 12 GQKQKALOVEFQYDQHPD---AKKQKALOVEFQYDQHPD---AKKQKALOVEFQYDQHPD 420  
 13 GQKQKALOVEFQYDQHPD---AKKQKALOVEFQYDQHPD---AKKQKALOVEFQYDQHPD 424  
 7 GQKQKALOVEFQYDQHPD---AKKQKALOVEFQYDQHPD---AKKQKALOVEFQYDQHPD 428  
 14 GQKQKALOVEFQYDQHPD---AKKQKALOVEFQYDQHPD---AKKQKALOVEFQYDQHPD 434  
 \* \* \* \* \*

1 LKQKALOVEFQYDQHPD---AKKQKALOVEFQYDQHPD---AKKQKALOVEFQYDQHPD 439  
 5 LKQKALOVEFQYDQHPD---AKKQKALOVEFQYDQHPD---AKKQKALOVEFQYDQHPD 443  
 9 LKQKALOVEFQYDQHPD---AKKQKALOVEFQYDQHPD---AKKQKALOVEFQYDQHPD 447  
 10 LKQKALOVEFQYDQHPD---AKKQKALOVEFQYDQHPD---AKKQKALOVEFQYDQHPD 451  
 11 LKQKALOVEFQYDQHPD---AKKQKALOVEFQYDQHPD---AKKQKALOVEFQYDQHPD 455  
 12 LKQKALOVEFQYDQHPD---AKKQKALOVEFQYDQHPD---AKKQKALOVEFQYDQHPD 459  
 13 LKQKALOVEFQYDQHPD---AKKQKALOVEFQYDQHPD---AKKQKALOVEFQYDQHPD 463  
 7 LKQKALOVEFQYDQHPD---AKKQKALOVEFQYDQHPD---AKKQKALOVEFQYDQHPD 467  
 14 LKQKALOVEFQYDQHPD---AKKQKALOVEFQYDQHPD---AKKQKALOVEFQYDQHPD 473  
 \* \* \* \* \*

FIGURA 9 8 {continuación}

```

1  DDQIDENIRHAIHQRAITDTCYHMAHDSSESV--DQQLVHGVNHELVVDASVVPKIPG 556
5  DHELEYTRDTVHGHALVPTAMHAGSAGASVSESELEVVGVVHLEVVDAQVLPKIPG 547
9  DERIDGFIKDTVHDGNANVPTCEHGVNQH--EIV--DFGLVFNHIDGLEVDASVVPKIPG 547
10 DFDIEEYVHTVVDGNALVPTAMGPTTPEGGSEVDSADLVVFNVDGLEVVDAQVLPKIPG 547
11 NFRICSEYIKSTVHDANALVPTCYHGEREDHMSV--NSALEVVMVAGLEVTIDGQVMPKIPG 537
12 DDQIDSEYIKSEVVDANALTEECYHGLDFDFAAVL--DQHLVFNVDGLEVDASVVPKIPG 557
13 EQQLAAYVHTANTANAVVPTCYHGEDSDALEVV--DQHLVFNVDGLEVVDAQVLPKIPG 512
7  DEEIDENIRHSLTANALVPTCYHGTG--EGRSV--GPLEVVICVHGVVGGDSVVPKIPG 523
14 DEELDAYIKHSLTANALVPTCYHGVVEDHNSV--DFECPVIGVSELEVVDAQVVPKIPG 572
      ( ( ( ( ( ) *c.** . . . ) ** ) * ( ( ** *c** * ( *

```

```

1  GQTGAATVWVREAAALLEGKATIQAASAAKSTV-----AA----- 594
5  GQTGAATVWVREAAALLEGKATIQAASAAKSTV----- 580
9  GQTGAATVWVREAAALLEGKATIQAASAAKSTV-----PALA----- 585
10 GQTGAATVWVREAAALLEGKATIQAASAAKSTV----- 580
11 GQTGAATVWVREAAALLEGKATIQAASAAKSTV----- 585
12 GQTGAATVWVREAAALLEGKATIQAASAAKSTV----- 580
13 GQTGAATVWVREAAALLEGKATIQAASAAKSTV----- 585
7  GQTGAATVWVREAAALLEGKATIQAASAAKSTV-----IIRKSGSTEDGATTACA----- 569
14 GQTGAATVWVREAAALLEGKATIQAASAAKSTV----- 580
      *** 1 2 3 4 5 6 7 8 9 10 11 12 13 14

```

FIGURA 3B (continuación)

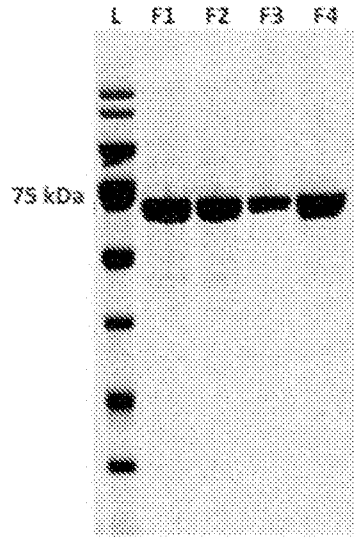


FIGURA 9

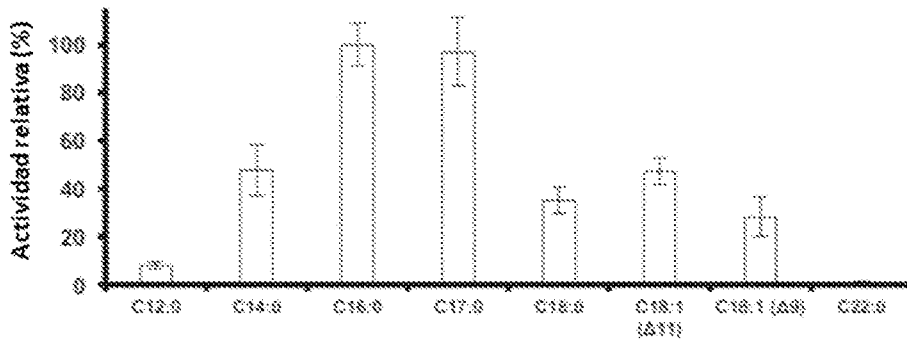


FIGURA 10

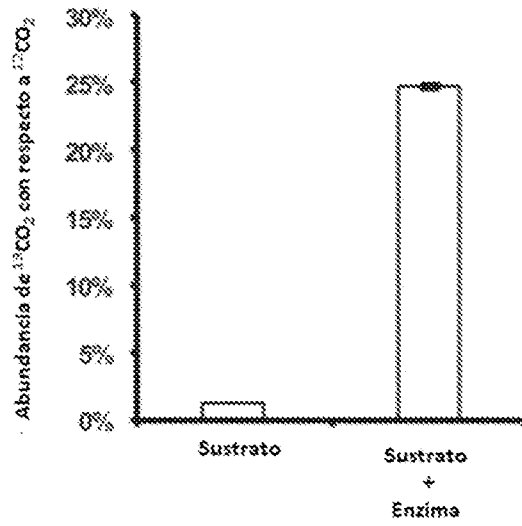


FIGURA 11

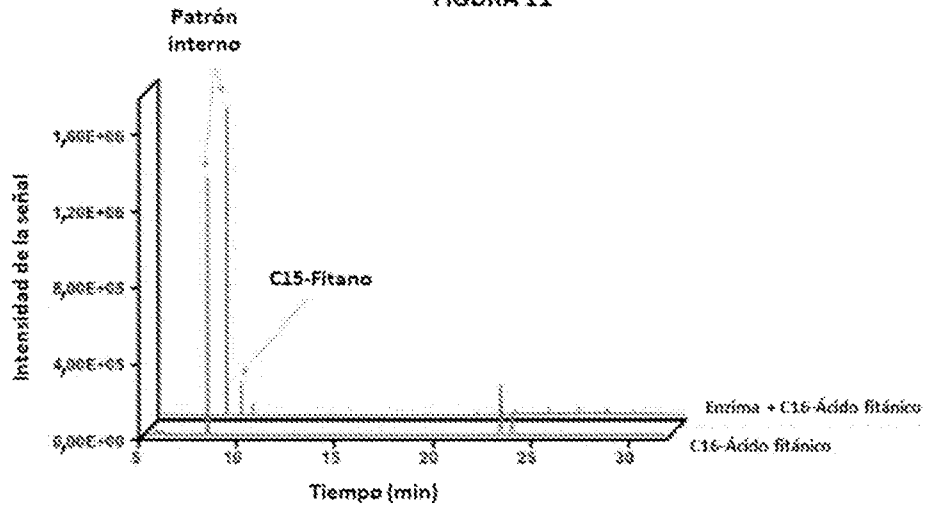


FIGURA 12

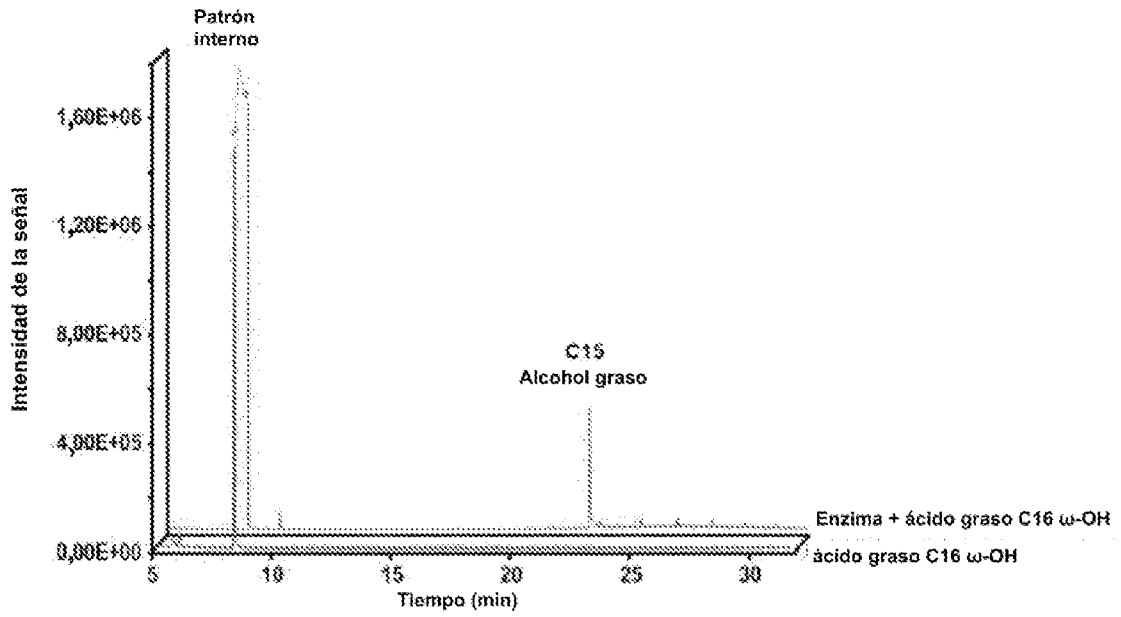


FIGURA 13

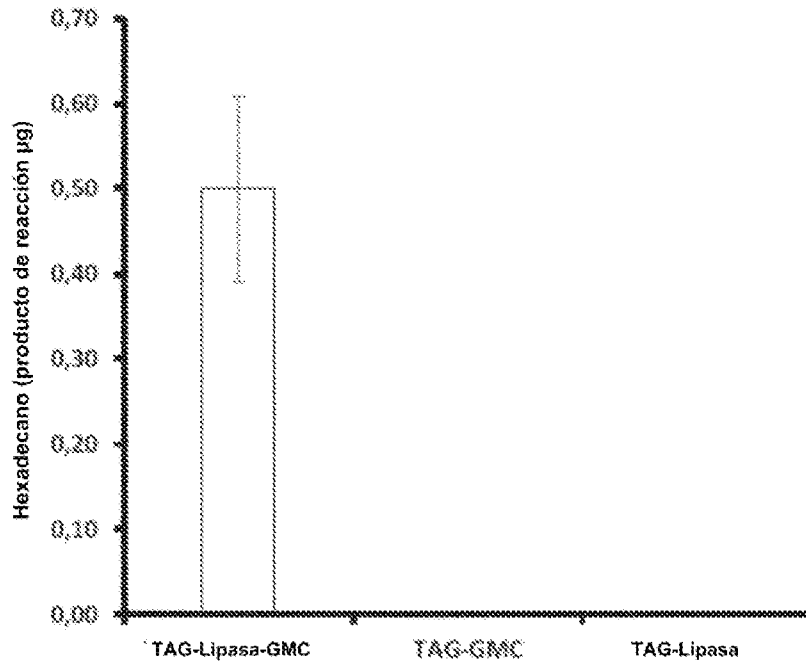


FIGURA 14

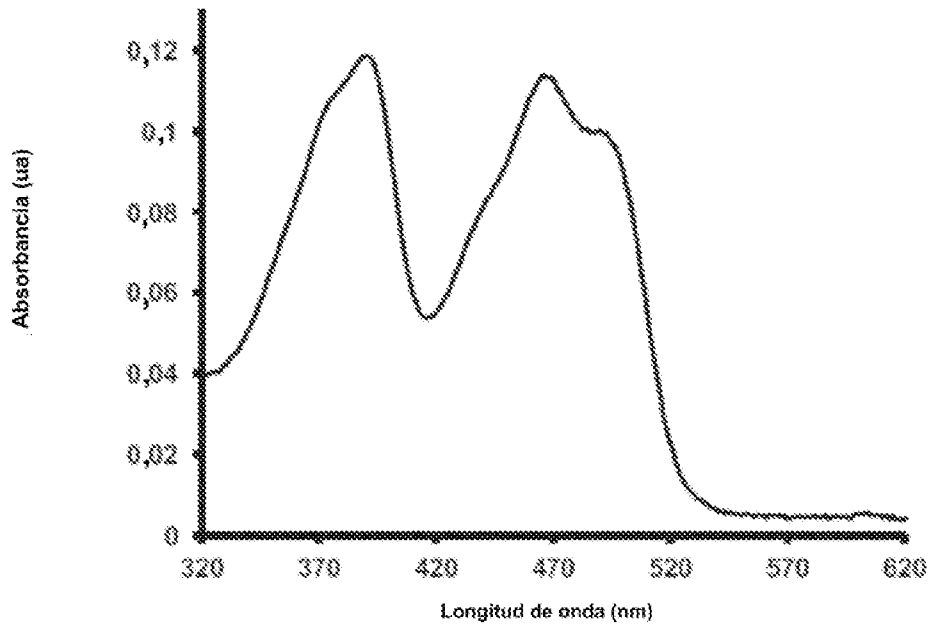


FIGURA 15

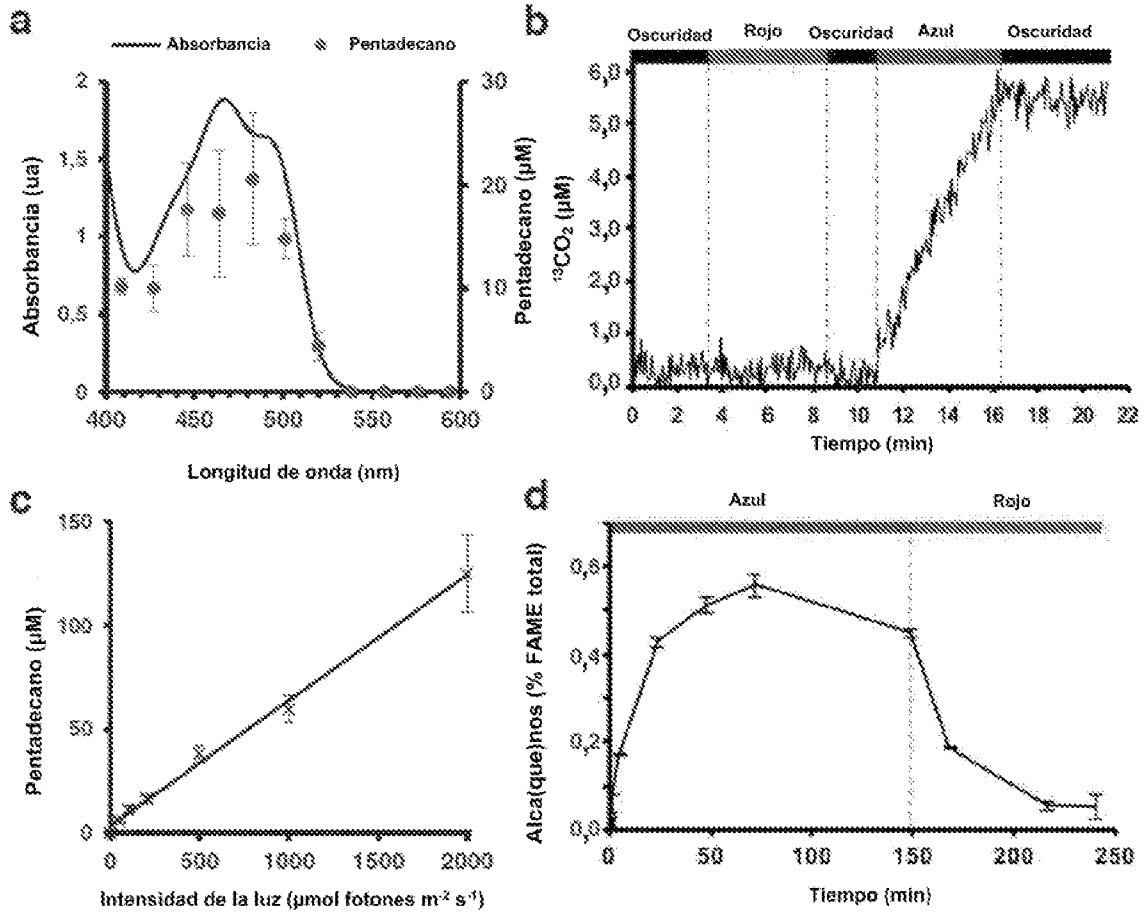


FIGURA 16

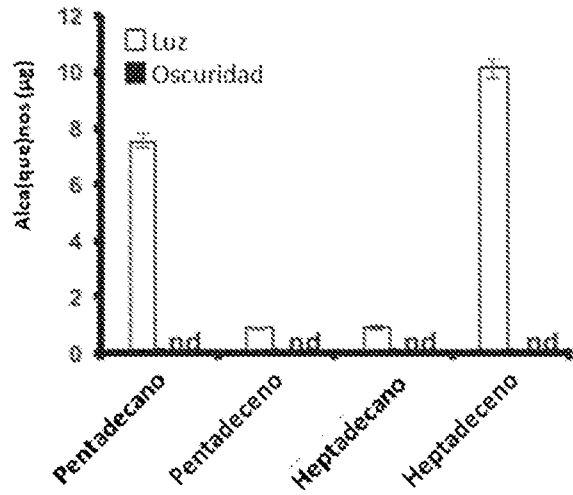


FIGURA 17

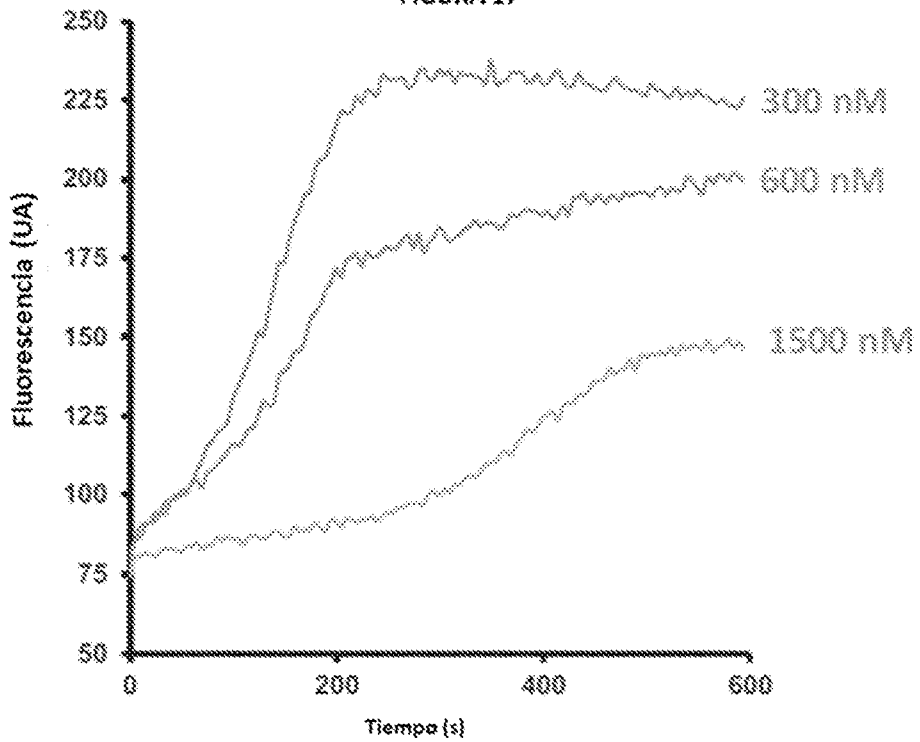


FIGURA 18

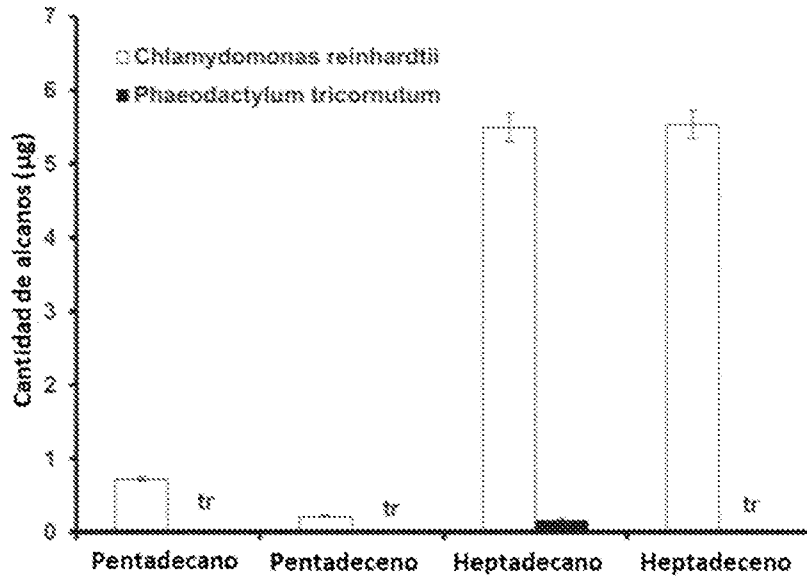


FIGURA 19

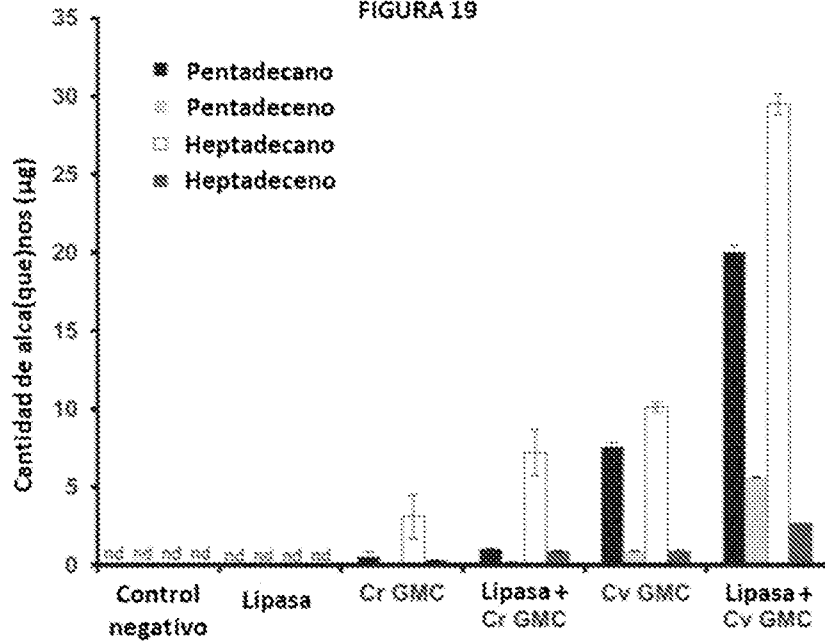


FIGURA 20