

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 449 474**

51 Int. Cl.:

**C12N 9/18** (2006.01)

**C12N 9/10** (2006.01)

**C12N 15/75** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **25.01.2007 E 11184428 (8)**

97 Fecha y número de publicación de la concesión europea: **04.12.2013 EP 2405007**

54 Título: **Producción de una lípido aciltransferasa a partir de células de Bacillus licheniformis**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**19.03.2014**

73 Titular/es:

**DUPONT NUTRITION BIOSCIENCES APS  
(100.0%)  
Langebrogade 1, Postboks 17  
1001 Copenhagen K, DK**

72 Inventor/es:

**KOLKMAN, MARC;  
MIKKELSEN, JØRN DALGAARD y  
LORENTSEN, RIKKE HØEGH**

74 Agente/Representante:

**DE ELZABURU MÁRQUEZ, Alberto**

ES 2 449 474 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**

Producción de una lípido aciltransferasa a partir de células de *Bacillus licheniformis*

**Referencia a las solicitudes relacionadas**

5 Se hace referencia a las siguientes solicitudes relacionadas: Documentos US 2002-0009518, US 2004-0091574, WO2004/064537, WO2004/064987, WO2005/066347, WO2005/066351. Solicitud de Estados Unidos de Número de Serie 60/764.430 presentada el 2 de febrero de 2006, y documento WO2006/008508.

**Campo de la presente invención**

10 La presente invención se refiere a la producción de lípido aciltransferasas. En particular, métodos para la producción de una lípido aciltransferasa mediante la expresión de una lípido aciltransferasa en una célula hospedadora de *Bacillus*, preferiblemente una célula hospedadora de *B. licheniformis*. Además, la presente invención se refiere al uso de *Bacillus* (preferiblemente *B. licheniformis*) para expresar una lípido aciltransferasa y a una célula hospedadora de *Bacillus*, preferiblemente una célula hospedadora de *B. licheniformis*, que comprende en su genoma un gen que codifica una lípido aciltransferasa.

**Antecedentes de la presente invención**

15 Se sabe que las lípido aciltransferasas son ventajosas en las aplicaciones alimentarias. Se ha descubierto que las lípido aciltransferasas tienen una actividad de aciltransferasa significativa en los productos alimenticios. Esta actividad tiene aplicaciones beneficiosas sorprendentes en los métodos de preparación de productos alimenticios.

20 Por ejemplo, el documento WO 2004/064537 describe un método para la producción *in situ* de un emulsionante mediante el uso de una lípido aciltransferasa y las ventajas asociadas a ello. Por lo tanto, existe la necesidad de un método para la producción comercial de lípido aciltransferasas.

No obstante, en general los genes pueden ser difíciles de expresar en hospedadores heterólogos, y la expresión de lípido aciltransferasas en las células hospedadoras puede ser problemática.

25 El documento WO 2004/064537 describe la expresión de dos lípido aciltransferasas de *Aeromonas* en *Bacillus subtilis* y *Escherichia Coli*. No obstante, la expresión en *B. subtilis* es baja, mientras *E. coli* no es un organismo GRAS y, por lo tanto, es inadecuado como hospedador para enzimas que se van a usar en la industria alimentaria.

30 El documento US 6.255.076 describe un método para producir un polipéptido en una célula hospedadora de *Bacillus*. Sin embargo, tal método requiere el uso de un promotor en tándem en el que cada secuencia promotora esté unida de forma operable a una única copia de una secuencia de ácido nucleico que codifica la secuencia polipeptídica. Así, existe la necesidad en la técnica de un método mejorado para la producción de lípido aciltransferasas.

**Aspectos resumen de la presente invención**

Los aspectos de la presente invención se presentan en las reivindicaciones y en el siguiente comentario.

Un aspecto de la presente invención se refiere a un método para la producción de una lípido aciltransferasa que comprende las etapas de:

35 (i) proporcionar una célula hospedadora de *Bacillus* en la que la célula hospedadora de *Bacillus* es una célula de *Bacillus licheniformis*;

(ii) transformar la célula hospedadora de *Bacillus*, en la que la célula hospedadora de *Bacillus* es la célula de *Bacillus licheniformis*, con una secuencia de nucleótidos heteróloga que codifica una lípido aciltransferasa mostrada como SEQ ID N°. 49 y

40 (iii) expresar la lípido aciltransferasa en la célula bajo control de una secuencia promotora.

En la presente memoria se describe una célula hospedadora de *Bacillus* en la que la célula hospedadora de *Bacillus* es distinta de *Bacillus subtilis*, preferiblemente una célula hospedadora de *Bacillus licheniformis*, que comprende una lípido aciltransferasa heteróloga.

45 En la presente memoria también se describe el uso de una célula hospedadora de *Bacillus* en la que la célula hospedadora de *Bacillus* es distinta de *Bacillus subtilis*, preferiblemente una célula hospedadora de *Bacillus licheniformis*, en la producción de una lípido aciltransferasa heteróloga.

De manera adecuada, la expresión en el hospedador de *Bacillus*, en el que el hospedador de *Bacillus* es distinto de *Bacillus subtilis*, y preferiblemente en el que el hospedador de *Bacillus* es *B. licheniformis*, puede dar como resultado una expresión incrementada en comparación con la expresión en *B. subtilis*.

En la presente memoria se describe un vector de expresión que comprende una secuencia de nucleótidos que codifica una lípido aciltransferasa unida de forma operable a una o más secuencia(s) reguladora(s) de forma que la(s) secuencia(s) reguladora(s) es capaz de expresar la secuencia de nucleótidos que codifica una lípido aciltransferasa en un hospedador o célula hospedadora adecuada, preferiblemente en un hospedador (o célula) de *Bacillus* en el que el hospedador (o célula) de *Bacillus* es distinto de *Bacillus subtilis*, preferiblemente en *B. licheniformis* o una célula de *B. licheniformis*.

De manera adecuada, la lípido aciltransferasa puede ser una lípido aciltransferasa recombinante.

#### Aspectos detallados de la presente invención

Según un primer aspecto de la presente invención, se proporciona un método para la producción de una lípido aciltransferasa que comprende las etapas de:

(i) proporcionar una célula hospedadora de *Bacillus* en la que la célula hospedadora de *Bacillus* es una célula de *Bacillus licheniformis*;

(ii) transformar la célula hospedadora de *Bacillus* en la que la célula hospedadora de *Bacillus* es una célula de *Bacillus licheniformis*, con una secuencia de nucleótidos heteróloga que codifica una lípido aciltransferasa mostrada como SEQ ID N°. 49; y

(iii) expresar la lípido aciltransferasa en la célula bajo control de una secuencia promotora.

Además, se puede unir de forma operable una secuencia de nucleótidos que codifica un péptido señal a dicha secuencia de nucleótidos heteróloga que codifica una lípido aciltransferasa.

De manera adecuada, el método de la presente invención puede comprender además la etapa adicional de aislar/recuperar la lípido aciltransferasa.

De manera adecuada, la lípido aciltransferasa puede ser una lípido aciltransferasa recombinante.

De manera adecuada, la secuencia promotora usada de acuerdo con las células hospedadoras, vectores, métodos y/o usos de la presente invención puede ser homóloga respecto de la célula hospedadora. "Homólogo respecto de la célula hospedadora" significa que se origina en el organismo hospedador; es decir, una secuencia promotora que se halla de manera natural en el organismo hospedador. De manera adecuada, la secuencia promotora se puede seleccionar del grupo que consiste en una secuencia de nucleótidos que codifica: un promotor de  $\alpha$ -amilasa, un promotor de proteasa, un promotor de subtilisina, un promotor de proteasa específica de ácido glutámico y un promotor de levansacarasa. De manera adecuada, la secuencia promotora puede ser una secuencia de nucleótidos que codifica: el LAT (p.ej. el promotor de alfa-amilasa de *B. licheniformis*, también conocido como AmyL), AprL (p.ej. promotor de subtilisina Carlsberg), EndoGluC (p.ej. el promotor específico de ácido glutámico de *B. licheniformis*), AmyQ (p.ej. el promotor de alfa amilasa de *B. amyloliquefaciens*) y SacB (p.ej. el promotor de levansacarasa de *B. subtilis*).

En una realización de la presente descripción, la secuencia promotora es la secuencia de -35 a -10 de un promotor de alfa amilasa, preferiblemente la secuencia de -35 a -10 de un promotor de  $\alpha$ -amilasa de *B. licheniformis*. La "secuencia de -35 a -10" describe la posición respecto del sitio de inicio de la transcripción. Tanto "-35" como "-10" son cajas, es decir, varios nucleótidos, cada una comprende 6 nucleótidos, y estas cajas están separadas por 17 nucleótidos. Estos 17 nucleótidos se denominan a menudo "espaciador". Esto se ilustra en la Figura 55, en la que están subrayadas las cajas de -35 y -10. Para evitar dudas, cuando se usa "secuencia de -35 a -10" en la presente memoria se refiere a una secuencia desde el inicio de la caja de -35 hasta el final de la caja de -10, es decir, que incluye la caja de -35, el espaciador de 17 nucleótidos de longitud y la caja de -10.

Una secuencia de nucleótidos que codifica una lípido aciltransferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y/o usos descritos en la presente memoria puede comprender un motivo GDSx y/o un motivo GANDY.

Preferiblemente, la enzima lípido aciltransferasa se caracteriza como una enzima que posee actividad de aciltransferasa y que comprende el motivo de secuencia de aminoácidos GDSX, en el que X es uno o más de los residuos de aminoácidos siguientes L, A, V, I, F, Y, H, Q, T, N, M o S.

De manera adecuada, la secuencia de nucleótidos que codifica una lípido aciltransferasa para el uso en cualquiera de los vectores, células hospedadoras, métodos y/o usos descritos en la presente memoria puede ser obtenible, y preferiblemente se obtiene, de un organismo de uno o más de los géneros siguientes: *Aeromonas*, *Streptomyces*, *Saccharomyces*, *Lactococcus*, *Mycobacterium*, *Streptococcus*, *Lactobacillus*, *Desulfibacterium*, *Bacillus*, *Campylobacter*, *Vibrionaceae*, *Xylella*, *Sulfolobus*, *Aspergillus*, *Schizosaccharomyces*, *Listeria*, *Neisseria*, *Mesorhizobium*, *Ralstonia*, *Xanthomonas* y *Candida*. Preferiblemente, la lípido aciltransferasa es obtenible, y preferiblemente se obtiene, de un organismo del género *Aeromonas*.

Una secuencia de nucleótidos que codifica una lípido aciltransferasa para el uso en cualquiera de los vectores, células hospedadoras, métodos y/o usos descritos en la presente memoria codifica una lípido aciltransferasa que comprende un residuo de ácido aspártico en una posición que corresponde a N-80 en la secuencia de aminoácidos de la lípido aciltransferasa de *Aeromonas hydrophila* mostrada como SEQ ID N° 35.

- 5 Una secuencia de nucleótidos que codifica una lípido aciltransferasa para el uso en cualquiera de los vectores, células hospedadoras, métodos y/o usos descritos en la presente memoria codifica una lípido aciltransferasa que puede comprender la secuencia de aminoácidos mostrada como SEQ ID N° 16, o una secuencia de aminoácidos que tiene un 75% o más de homología con ella. De manera adecuada, la secuencia de nucleótidos que codifica una lípido aciltransferasa codifica una lípido aciltransferasa que puede comprender la secuencia de aminoácidos mostrada como SEQ ID N° 16.

El término "heterólogo", tal como se usa en la presente memoria, significa una secuencia derivada de una fuente genética o especie diferente. Una secuencia heteróloga es una secuencia que no es del hospedador, una secuencia modificada, una secuencia de una cepa de células hospedadoras diferentes, o una secuencia homóloga de una localización cromosómica diferente de la célula hospedadora.

- 15 Una secuencia "homóloga" es una secuencia que se halla en la misma fuente o especie genética, es decir, se da de manera natural en la especie relevante de la célula hospedadora.

La expresión "lípido aciltransferasa recombinante", tal como se usa en la presente memoria, significa que la lípido aciltransferasa se ha producido por medio de recombinación genética. Por ejemplo, la secuencia de nucleótidos que codifica la lípido aciltransferasa se ha insertado en un vector de clonación, lo que da como resultado la célula de *B. licheniformis* caracterizada por la presencia de la lípido aciltransferasa heteróloga.

#### Célula hospedadora

En una realización de la presente invención, la célula hospedadora para el uso en los métodos y/o usos de la presente invención es una célula hospedadora de *Bacillus licheniformis*.

- 25 Se ha descubierto que el uso de una célula hospedadora de *Bacillus licheniformis* da como resultado la expresión incrementada de una lípido aciltransferasa en comparación con otros organismos, tales como *Bacillus subtilis*.

Se ha insertado una lípido aciltransferasa de *Aeromonas salmonicida* en varios vectores de expresión convencionales, diseñados para ser óptimos para la expresión en *Bacillus subtilis*, *Hansenula polymorpha*, *Schizosaccharomyces pombe* y *Aspergillus tubigensis*, respectivamente. Se detectaron solamente niveles muy bajos, sin embargo, en *Hansenula polymorpha*, *Schizosaccharomyces pombe* y *Aspergillus tubigensis*. Los niveles de expresión fueron inferiores a 1 µg/ml, y no fue posible seleccionar las células que produjeron suficiente proteína para iniciar una producción comercial (resultados no mostrados). En contraste, *Bacillus licheniformis* fue capaz de producir niveles de proteína que son atractivos para una producción económicamente viable.

- 30 En particular, se ha descubierto que la expresión en *B. licheniformis* es aproximadamente 100 veces mayor que la expresión en *B. subtilis* bajo control del promotor de aprE, o es aproximadamente 100 veces mayor que la expresión en *S. lividans* bajo control de un promotor de A4 y fusionado a celulosa (resultados no mostrados en la presente memoria).

En otra realización, la célula hospedadora puede ser cualquier *Bacillus licheniformis*.

- 40 La expresión "célula hospedadora", en relación a la presente invención, incluye cualquier célula que comprende una secuencia de nucleótidos que codifica una lípido aciltransferasa como se define en la presente memoria o un vector de expresión como el descrito anteriormente y que se usa en la producción recombinante de una lípido aciltransferasa que tiene las propiedades específicas definidas en la presente memoria.

Así, una realización adicional de la presente descripción proporciona una célula hospedadora que comprende (por ejemplo transformada o transfectada con) una secuencia de nucleótidos de la presente descripción o una secuencia de nucleótidos que expresa un polipéptido que tiene las propiedades específicas definidas en la presente memoria.

- 45 De manera adecuada, en ciertas realizaciones, la célula hospedadora puede ser una cepa deficiente de proteasa o sin proteasa y/o una cepa deficiente de  $\alpha$ -amilasa o sin  $\alpha$ -amilasa.

#### Secuencias Regulatoras

- 50 En ciertas aplicaciones, se puede unir de forma operable una secuencia de lípido aciltransferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y/o usos de la presente descripción a una secuencia reguladora que es capaz de mantener la expresión de la secuencia de nucleótidos, tal como mediante la célula hospedadora elegida (tal como una célula de *B. licheniformis*).

A modo de ejemplo, la presente descripción cubre un vector que comprende la secuencia de nucleótidos de la presente descripción unida de forma operable a dicha secuencia reguladora, es decir, el vector es un vector de expresión.

5 La expresión "unida de forma operable" se refiere a la yuxtaposición en la que los componentes descritos se encuentran en una relación que les permite funcionar en su modo pretendido. Una secuencia reguladora "unida de forma operable" a una secuencia codificante está ligada de tal modo que la expresión de la secuencia codificante se logra en condiciones compatibles con las secuencias de control.

El término "secuencias reguladoras" incluye promotores y potenciadores y otras señales de regulación de la expresión.

10 El término "promotor" se usa en el sentido normal de la técnica, por ejemplo un sitio de unión de polimerasa de ARN.

La expresión incrementada de la secuencia de nucleótidos que codifica la enzima que tiene las propiedades específicas definidas en la presente memoria también se puede conseguir mediante la selección de las regiones reguladoras, p.ej. regiones promotoras, líder de secreción y terminadoras que no son regiones reguladoras para la secuencia de nucleótidos que codifica la enzima en la naturaleza.

15 De manera adecuada, la secuencia de nucleótidos de la presente descripción se puede unir de forma operable a al menos un promotor.

De manera adecuada, la secuencia de nucleótidos que codifica una lípido aciltransferasa se puede unir de forma operable a una secuencia de nucleótidos que codifica una secuencia terminadora. Los ejemplos de secuencias terminadoras adecuadas para el uso en cualquiera de los vectores, células hospedadoras, métodos y/o usos de la presente invención incluyen: una secuencia terminadora de  $\alpha$ -amilasa (por ejemplo, CGGGACTTACCGAAAGAAA-CCATCAATGATGGTTTCTTTTTGTTTCATAAA - SEQ ID N° 64), una secuencia terminadora de proteasa alcalina (por ejemplo, CAAGACTAAAGACCGTTCGCCCGTTTTGCAATAAGCGGGCGAATCTTACATAAAA ATA - SEQ ID N° 65), una secuencia terminadora específica de ácido glutámico (por ejemplo, ACGGCCGTTAGATGTGACAGCCC-GTTCCAAAAGGAAGCGGGCTGTCTTCGTGTAT TATTGT - SEQ ID N° 66), una secuencia terminadora de levanasa (por ejemplo, TCTTTTAAAGGAAAGGCTGGAATGCCCGGCA17CCAGCCACATGATCATCG7TT - SEQ ID N° 67) y una secuencia terminadora de subtilisina E (por ejemplo, GCTGACAAATAAAAAGAAGCAGGTA-TGGAGGAACCTGCTTCTTTTACTATTATTG). De manera adecuada, la secuencia de nucleótidos que codifica una lípido aciltransferasa se puede unir de forma operable a un terminador de  $\alpha$ -amilasa, tal como un terminador de  $\alpha$ -amilasa de *B. licheniformis*.

### 30 **Promotor**

La secuencia promotora a usar de acuerdo con la presente invención puede ser heteróloga u homóloga respecto de la secuencia que codifica una lípido aciltransferasa.

La secuencia promotora puede ser cualquier secuencia promotora capaz de dirigir la expresión de una lípido aciltransferasa en la célula hospedadora de elección.

35 De manera adecuada, la secuencia promotora puede ser homóloga respecto de una especie *Bacillus*, por ejemplo *B. licheniformis*. Preferiblemente, la secuencia promotora es homóloga respecto de la célula hospedadora de elección.

Las secuencias promotoras adecuadas para el uso en la presente invención incluyen: el promotor del gen de alfa-amilasa de *Bacillus licheniformis*, el promotor del gen de subtilisina de *Bacillus licheniformis*, el promotor del gen de subtilisina de *Bacillus subtilis*, el promotor del gen de proteasa alcalina de *Bacillus licheniformis* (gen de subtilisina Carlsberg), el promotor del gen de proteasa específica de ácido glutámico de *B. licheniformis*, el promotor del gen de alfa-amilasa de *B. amyloliquefaciens*; el promotor de levansacarasa de *B. subtilis* y un promotor "consenso" que tiene la secuencia TTGACA para la región de "-35" y TATAAT para la región de "-10" (es decir, el promotor de -35 a -10) del gen de alfa-amilasa.

45 Otros ejemplos de promotores adecuados para dirigir la transcripción de una secuencia de ácido nucleico en los métodos de la presente invención incluyen: el promotor del gen de proteasa alcalina (aprH) de *Bacillus lentus*; el promotor del gen de alfa-amilasa (amyE) de *Bacillus subtilis*; el promotor del gen de amilasa maltogénica (amyM) de *Bacillus stearothermophilus*; el promotor del gen de penicilinas (penP) de *Bacillus licheniformis*; los promotores de los genes xyla y xylB de *Bacillus subtilis*; y/o el promotor del gen CryIIIA de *Bacillus thuringiensis* subsp. *tenebrionis*.

50 En una realización preferida, la secuencia del promotor es un promotor de  $\alpha$ -amilasa (tal como un promotor de  $\alpha$ -amilasa de *Bacillus licheniformis*). Preferiblemente, la secuencia del promotor comprende la secuencia de -35 a -10 del promotor de  $\alpha$ -amilasa de *B. licheniformis*, véanse las Figuras 53 y 55.

**Péptido señal**

La lípido aciltransferasa producida por una célula hospedadora mediante la expresión de la secuencia de nucleótidos que codifica la lípido aciltransferasa se puede secretar o se puede contener intracelularmente dependiendo de la secuencia y/o el vector usados.

5 Se puede usar una secuencia señal para la secreción directa de las secuencias codificantes a través de una membrana celular particular. Las secuencias señal pueden ser naturales o exógenas respecto de la secuencia codificante de la lípido aciltransferasa. Por ejemplo, la secuencia codificante del péptido señal se puede obtener de un gen de amilasa o proteasa de una especie *Bacillus*, preferiblemente de *Bacillus licheniformis*.

10 Las secuencias codificantes adecuadas del péptido señal se pueden obtener de uno o más de los genes siguientes: gen de  $\alpha$ -amilasa maltogénica, gen de subtilisina, gen de beta-lactamasa, gen de proteasa neutra, gen de prsA, y/o gen de aciltransferasa.

15 Preferiblemente, el péptido señal es un péptido señal de  $\alpha$ -amilasa de *B. licheniformis*, aciltransferasa de *Aeromonas* (por ejemplo, mkkwfvclglialtvqa - SEQ ID N° 21), subtilisina de *B. subtilis* (por ejemplo, mrskkiwisilfaltlftmfsnmsaqa - SEQ ID N° 22) o subtilisina de *B. licheniformis* (por ejemplo, mmrkkstwfglmftafmlvftmefsdsasa - SEQ ID N° 23). De manera adecuada, el péptido señal puede ser el péptido señal de  $\alpha$ -amilasa de *B. licheniformis*.

Sin embargo, se puede usar cualquier secuencia codificante del péptido señal capaz de dirigir la lípido aciltransferasa expresada hacia la ruta secretora de una célula hospedadora de *Bacillus* (preferiblemente una célula hospedadora de *B. licheniformis*) de elección.

20 En ciertas realizaciones de la presente descripción, se puede unir de forma operable una secuencia de nucleótidos que codifica un péptido señal a una secuencia de nucleótidos que codifica una lípido aciltransferasa de elección.

La lípido aciltransferasa de elección se puede expresar en una célula hospedadora como se define en la presente memoria en forma de una proteína de fusión.

**Vector de expresión**

El término "vector de expresión" significa una construcción capaz de expresarse *in vivo* o *in vitro*.

25 Preferiblemente, el vector de expresión se incorpora en el genoma del organismo, tal como un hospedador de *B. licheniformis*. El término "incorpora" preferiblemente cubre la incorporación estable en el genoma.

30 La secuencia de nucleótidos que codifica una lípido aciltransferasa como se define en la presente memoria puede estar presente en un vector, en el que la secuencia de nucleótidos está unida de forma operable a las secuencias reguladoras de forma que las secuencias reguladoras son capaces de proporcionar la expresión de la secuencia de nucleótidos en un organismo hospedador adecuado (tal como *B. licheniformis*), es decir, el vector es un vector de expresión.

Los vectores de la presente invención se pueden transformar en una célula hospedadora adecuada como se describió anteriormente para proporcionar la expresión de un polipéptido que tiene actividad de lípido aciltransferasa como se define en la presente memoria.

35 La elección del vector, p.ej. un vector plasmídico, cósmido, virus o fago, inserto genómico, dependerá a menudo de la célula hospedadora en la que se va a introducir. La presente invención puede cubrir otras formas de vectores de expresión que cumplen funciones equivalentes y que son, o pasan a ser, conocidas en la técnica.

Una vez transformado en la célula hospedadora de elección, el vector se puede replicar y funcionar independientemente del genoma de la célula hospedadora, o se puede integrar en el propio genoma.

40 Los vectores pueden contener uno o más marcadores seleccionables -tal como un gen que confiere resistencia a antibióticos, por ejemplo, resistencia a ampicilina, kanamicina, cloranfenicol o tetraciclina. Alternativamente, la selección puede lograrse mediante co-transformación (tal como se describe en el documento WO91/17243).

Se pueden usar los vectores *in vitro*, por ejemplo para la producción de ARN, o se pueden usar para transfectar o transformar una célula hospedadora.

45 Así, en una realización adicional, la descripción proporciona un método para producir secuencias de nucleótidos de la presente invención o secuencias de nucleótidos que codifican polipéptidos que tienen las propiedades específicas definidas en la presente memoria para el uso en cualquiera de los vectores, células hospedadoras, otros métodos y/o usos de la presente invención, introduciendo una secuencia de nucleótidos en un vector replicable, introduciendo el vector en una célula hospedadora compatible, y cultivando la célula hospedadora en condiciones que provocan la replicación del vector.

50

El vector puede comprender además una secuencia de nucleótidos que permita que el vector se replique en la célula hospedadora en cuestión. Los ejemplos de tales secuencias son los orígenes de replicación de los plásmidos pUC19, pACYC177, pUB110, pE194, AMB1 y pIJ702.

### Lípido acil transferasa

5 La secuencia de nucleótidos que codifica una lípido acil transferasa para el uso en cualquiera de los métodos, vectores y/o usos de la presente descripción puede codificar una lípido acil transferasa natural o una lípido acil transferasa variante.

10 Por ejemplo, la secuencia de nucleótidos que codifica una lípido acil transferasa para el uso en la presente descripción puede ser una descrita en los documentos WO2004/064537, WO2004/064987, WO2005/066347, o WO2006/008508.

15 La expresión "lípido acil transferasa", tal como se usa en la presente memoria, significa preferiblemente una enzima que tiene actividad de aciltransferasa (clasificada en general como E.C. 2.3.1.x, por ejemplo 2.3.1.43), por lo que la enzima es capaz de transferir un grupo acilo de un lípido a uno o más sustratos aceptores, tales como uno o más de lo siguiente: un esteroles; un estanol; un carbohidrato; una proteína; una subunidad de proteína; un polialcohol, tal como ácido ascórbico y/o glicerol, preferiblemente glicerol y/o un esteroles, tal como colesterol.

Preferiblemente, la secuencia de nucleótidos que codifica una lípido acil transferasa para el uso en cualquiera de los vectores, células hospedadoras, métodos y/o usos de la presente invención codifica una lípido aciltransferasa que es capaz de transferir un grupo acilo de un fosfolípido (como se define en la presente memoria) a un polialcohol, tal como ácido ascórbico y/o glicerol, lo más preferiblemente glicerol.

20 Para ciertos aspectos, el "aceptor de acilo" según la presente invención puede ser cualquier compuesto que comprende un grupo hidroxilo (-OH), tal como, por ejemplo, alcoholes polivalentes, que incluyen glicerol; esterol; estanoles; carbohidratos; hidroxiaácidos que incluyen ácidos de fruta, ácido cítrico, ácido tartárico, ácido láctico y ácido ascórbico; proteínas o una sub-unidad de las mismas, tal como aminoácidos, hidrolizados proteicos y péptidos (proteína parcialmente hidrolizada) por ejemplo; y mezclas y derivados de los mismos. Preferiblemente, el "aceptor de acilo" según la presente invención no es agua. Preferiblemente, el "aceptor de acilo" según la presente invención es un polialcohol, tal como un polioliol, lo más preferiblemente glicerol. Para el propósito de esta invención, el ácido ascórbico se considera también un polialcohol.

El aceptor de acilo preferiblemente no es un monoglicérido.

El aceptor de acilo preferiblemente no es un diglicérido.

30 En un aspecto, la secuencia de nucleótidos que codifica una lípido aciltransferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y/o usos de la presente invención codifica una lípido aciltransferasa que puede, y también es capaz de transferir un grupo acilo de un lípido a glicerol, y además es capaz de transferir el grupo acilo de un lípido a uno o más de lo siguiente: un carbohidrato, una proteína, una subunidad de proteína, esteroles y/o un estanol, preferiblemente es capaz de transferir a un polialcohol, tal como ácido ascórbico y/o glicerol, lo más preferiblemente un esteroles tal como colesterol, y/o esterol/estanoles vegetales.

Preferiblemente, el sustrato lipídico sobre el que actúa la lípido acil transferasa es uno o más de los lípidos siguientes: un fosfolípido, tal como una lecitina, p.ej. fosfatidilcolina.

40 Dicho sustrato lipídico puede denominarse en la presente memoria "lípido donante de acilo". El término lecitina, tal como se usa en la presente memoria, abarca fosfatidilcolina, fosfatidiletanolamina, fosfatidilinositol, fosfatidilserina y fosfatidilglicerol.

Para ciertos aspectos, preferiblemente la secuencia de nucleótidos que codifica una lípido acil transferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y/o usos de la presente invención codifica una lípido aciltransferasa que es incapaz, o sustancialmente incapaz, de actuar sobre un triglicérido y/o un 1-monoglicérido y/o 2-monoglicérido.

45 Para ciertos aspectos, preferiblemente la secuencia de nucleótidos que codifica una lípido acil transferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y/o usos de la presente invención codifica una lípido aciltransferasa que no exhibe actividad de triacilglicerol lipasa (E.C. 3.1.1.3) o no exhibe actividad significativa de triacilglicerol lipasa (E.C. 3.1.1.3).

50 La capacidad para hidrolizar triglicéridos (actividad E.C. 3.1.1.3) se puede determinar mediante la actividad de lipasa determinada según Food Chemical Codex (3ª Ed., 1981, pág. 492-493) modificado para aceite de girasol y pH 5,5 en lugar de aceite de oliva y pH 6,5. La actividad de lipasa se mide como LUS (unidades de lipasa de girasol, del inglés "lipase units sunflower"), donde 1 LUS se define como la cantidad de enzima que puede liberar 1 [mu]mol de ácidos grasos por minuto a partir de aceite de girasol en las condiciones de ensayo anteriores. Alternativamente, se puede usar el ensayo LUT como se define en el documento WO9845453.

- La secuencia de nucleótidos que codifica una lípido acil transferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y/o usos de la presente invención puede codificar una lípido aciltransferasa que es sustancialmente incapaz de actuar sobre un triglicérido que puede tener una LUS/mg de menos de 1000, por ejemplo menos de 500, tal como menos de 300, preferiblemente menos de 200, más preferiblemente menos de 100, más preferiblemente menos de 50, más preferiblemente menos de 20, más preferiblemente menos de 10, tal como menos de 5, menos de 2, más preferiblemente menos de 1 LUS/mg. Alternativamente, la actividad de LUT/mg es menor que 500, tal como menos de 300, preferiblemente menos de 200, más preferiblemente menos de 100, más preferiblemente menos de 50, más preferiblemente menos de 20, más preferiblemente menos de 10, tal como menos de 5, menos de 2, más preferiblemente menos de 1 LUT/mg.
- La secuencia de nucleótidos que codifica una lípido acil transferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y/o usos de la presente invención puede codificar una lípido aciltransferasa que es sustancialmente incapaz de actuar sobre un monoglicérido se puede determinar mediante el uso de mono-oleato (M7765 1-Oleoil-rac-glicerol, 99%) en lugar de aceite de girasol en el ensayo LUS. 1 MGHU se define como la cantidad de enzima que puede liberar 1 [mu]mol de ácidos grasos por minuto desde un monoglicérido en las condiciones de ensayo.
- La secuencia de nucleótidos que codifica una lípido acil transferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y/o usos de la presente invención codifica una lípido aciltransferasa que es sustancialmente incapaz de actuar sobre un triglicérido que puede tener una MGHU/mg de menos de 5000, por ejemplo menos de 1000, por ejemplo menos de 500, tal como menos de 300, preferiblemente menos de 200, más preferiblemente menos de 100, más preferiblemente menos de 50, más preferiblemente menos de 20, más preferiblemente menos de 10, tal como menos de 5, menos de 2, más preferiblemente menos de 1 MGHU/mg.
- De manera adecuada, la secuencia de nucleótidos que codifica una lípido aciltransferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y/o usos de la presente invención codifica una lípido aciltransferasa que puede exhibir una o más de las siguientes actividades de fosfolipasa: actividad de fosfolipasa A2 (E.C. 3.1.1.4) y/o actividad de fosfolipasa A1 (E.C. 3.1.1.32). La lípido acil transferasa puede tener también actividad de fosfolipasa B (E.C 3.1.1.5).
- De manera adecuada, para ciertos aspectos la lípido aciltransferasa puede ser capaz de transferir un grupo acilo de un fosfolípido a un polialcohol, preferiblemente glicerol y/o ácido ascórbico.
- Para ciertos aspectos, preferiblemente la secuencia de nucleótidos que codifica una lípido aciltransferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y/o usos de la presente invención codifica una lípido aciltransferasa que es capaz de transferir un grupo acilo de un fosfolípido a un esteroil y/o un estanol para formar al menos un éster de esteroil y/o un éster de estanol.
- La lípido aciltransferasa puede ser capaz de transferir un grupo acilo de un lípido a un polioliol tal como glicerol, y/o un esteroil tal como colesterol o esteroides/estanoles vegetales. Así, en una realización, el "aceptor de acilo" según la presente invención puede ser glicerol y/o colesterol o esteroides/estanoles vegetales.
- Preferiblemente, la enzima lípido aciltransferasa puede caracterizarse usando los siguientes criterios:
- la enzima posee una actividad de acil transferasa que se puede definir como una actividad de transferencia de éster, en la que la parte acilo de un enlace éster original de un lípido donante de acilo se transfiere a un aceptor de acilo, preferiblemente glicerol o colesterol, para formar un éster nuevo; y
  - la enzima comprende el motivo de secuencia de aminoácidos GDSX, donde X es uno o más de los siguientes residuos de aminoácido L, A, V, I, F, Y, H, Q, T, N, M o S.
- Preferiblemente, X del motivo GDSX es L o Y. Más preferiblemente, X del motivo GDSX es L. Así, preferiblemente la enzima según la presente invención comprende el motivo de secuencia de aminoácidos GDSL.
- El motivo GDSX consta de cuatro aminoácidos conservados. Preferiblemente, la serina del motivo es una serina catalítica de la enzima lípido acil transferasa. De manera adecuada, la serina del motivo GDSX puede estar en una posición que corresponde a Ser-16 en la enzima lípido aciltransferasa de *Aeromonas hydrophila* mostrada en Brumlik & Buckley (Journal of Bacteriology, abril de 1996, Vol. 178, N° 7, págs. 2060-2064).
- Para determinar si una proteína tiene el motivo GDSX según la presente invención, la secuencia se compara preferiblemente con los perfiles del modelo de Markov ocultos (perfiles HMM) de la base de datos pfam de acuerdo con los procedimientos mostrados en los documentos WO2004/064537 o WO20041064987.
- Preferiblemente, la enzima lípido acil transferasa se puede alinear mediante el uso de la secuencia consenso Pfam00657 (para una explicación completa, véanse los documentos WO2004/064537 o WO2004/064987).
- Preferiblemente, una coincidencia positiva con el perfil del modelo de Markov oculto (perfil HMM) de la familia de dominio pfam00657 indica la presencia del dominio GDSL o GDSX según la presente invención.



Preferiblemente, cuando se alinea con la secuencia consenso Pfam00657 la lípido aciltransferasa para el uso en los métodos o usos de la invención puede tener al menos uno, preferiblemente más de uno, preferiblemente más de dos, de lo siguiente, un bloque GDSx, un bloque GANDY, un bloque HPT. De forma adecuada, la lípido aciltransferasa puede tener un bloque GDSx y un bloque GANDY. Alternativamente, la enzima puede tener un bloque GDSx y un bloque HPT. Preferiblemente, la enzima comprende al menos un bloque GDSx. Véanse los documentos WO2004/064537 o WO2004/064987 para detalles adicionales.

Preferiblemente, los residuos del motivo GANDY se seleccionan entre GANDY, GGND, GGNDL, más preferiblemente GANDY.

Preferiblemente, cuando se alinea con la secuencia consenso Pfam00657, la enzima para el uso en los métodos o usos de la invención tiene al menos uno, preferiblemente más de uno, preferiblemente más de dos, preferiblemente más de tres, preferiblemente más de cuatro, preferiblemente más de cinco, preferiblemente más de seis, preferiblemente más de siete, preferiblemente más de ocho, preferiblemente más de nueve, preferiblemente más de diez, preferiblemente más de once, preferiblemente más de doce, preferiblemente más de trece, preferiblemente más de catorce de los siguientes residuos de aminoácidos en comparación con la secuencia polipeptídica de referencia de *A. hydrophila*, es decir la SEQ ID N° 1: 28hid, 29hid, 30hid, 31hid, 32gly, 33Asp, 34Ser, 35hid, 130hid, 131Gly, 132Hid, 133Asn, 134Asp, 135hid, 309His.

El dominio GDSX de pfam00657 es un identificador único que distingue proteínas que poseen dicho dominio de otras enzimas.

La secuencia consenso pfam00657 se presenta en la Figura 3 como SEQ ID N° 2. Ésta se deriva de la identificación de la familia pfam 00657, base de datos versión 6, que también puede referirse en la presente memoria como pfam00657.6.

La secuencia consenso se puede actualizar mediante el uso de publicaciones adicionales de la base de datos pfam (por ejemplo, véanse los documentos WO2004/064537 o WO2004/064987).

En una realización, la secuencia de nucleótidos que codifica una enzima lípido acil transferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y/o usos de la presente invención codifica una lípido aciltransferasa que se puede caracterizar mediante el uso de los criterios siguientes:

(i) la enzima posee actividad de acil transferasa que se puede definir como una actividad de transferencia de éster, en la que la parte acilo de un enlace éster original de un lípido donante de acilo se transfiere a un aceptor de acilo, preferiblemente glicerol o colesterol, para formar un éster nuevo, preferiblemente éster de monoglicérido o de colesterol, respectivamente;

(ii) la enzima comprende el motivo de secuencia de aminoácidos GDSX, donde X es uno o más de los siguientes residuos de aminoácido L, A, V, I, F, Y, H, Q, T, N, M o S;

(iii) la enzima comprende His-309 o comprende un residuo de histidina en una posición que corresponde a His-309 en la enzima lípido aciltransferasa de *Aeromonas hydrophila* mostrada en las Figuras 2 y 4 (SEQ ID N° 1 o SEQ ID N° 3).

Preferiblemente, el residuo de aminoácido del motivo GDSX es L.

En la SEQ ID N° 3 o la SEQ ID N° 1 los primeros 18 residuos de aminoácido forman una secuencia señal. La His-309 de la secuencia de longitud completa, que es la proteína que incluye la secuencia señal, es igual a la His-291 de la parte madura de la proteína, es decir, la secuencia sin la secuencia señal.

La secuencia de nucleótidos que codifica una enzima lípido acil transferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y usos descritos en la presente memoria codifica una lípido aciltransferasa que comprende la siguiente tríada catalítica: Ser-34, Asp-306 y His-309, o comprende un residuo de serina, un residuo de ácido aspártico y un residuo de histidina, respectivamente, en las posiciones que corresponden a Ser-34, Asp-306 y His-309 en la enzima lípido acil transferasa de *Aeromonas hydrophila* mostrada en la Figura 4 (SEQ ID N° 3) o la Figura 2 (SEQ ID N° 1). Como se ha indicado anteriormente, en la secuencia mostrada en la SEQ ID N° 3 o la SEQ ID N° 1 los primeros 18 residuos de aminoácido forman una secuencia señal. Ser-34, Asp-306 e His-309 de la secuencia de longitud completa, que es la proteína que incluye la secuencia señal, son iguales a Ser-16, Asp-288 e His-291 de la parte madura de la proteína, es decir, la secuencia sin la secuencia señal. En la secuencia consenso pfam00657, tal como se da en la Figura 3 (SEQ ID N° 2), los residuos del centro activo corresponden a Ser-7, Asp-345 e His-348.

La secuencia de nucleótidos que codifica una enzima lípido acil transferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y usos descritos en la presente memoria codifica una lípido aciltransferasa que se puede caracterizar mediante el uso de los criterios siguientes:

la enzima posee actividad de acil transferasa que puede definirse como actividad de transferencia de éster en la que la parte acilo de un enlace éster original de un primer lípido donante de acilo se transfiere a un aceptor de acilo para formar un nuevo éster; y

5 la enzima comprende al menos Gly-32, Asp-33, Ser-34, Asp-134 y His-309 o comprende residuos de glicina, ácido aspártico, serina, ácido aspártico e histidina en posiciones que corresponden a Gly-32, Asp-33, Ser-34, Asp-306 e His-309, respectivamente, en la enzima lípido aciltransferasa de *Aeromonas hydrophila* mostrada en SEQ ID N° 3 o SEQ ID N° 1.

10 La secuencia de nucleótidos que codifica una enzima lípido aciltransferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y usos descritos en la presente memoria puede ser una de las secuencias de nucleótidos siguientes:

- (a) la secuencia de nucleótidos mostrada como SEQ ID N° 36 (véase la Figura 29);
- (b) la secuencia de nucleótidos mostrada como SEQ ID N° 38 (véase la Figura 31);
- (c) la secuencia de nucleótidos mostrada como SEQ ID N° 39 (véase la Figura 32);
- (d) la secuencia de nucleótidos mostrada como SEQ ID N° 42 (véase la Figura 35);
- 15 (e) la secuencia de nucleótidos mostrada como SEQ ID N° 44 (véase la Figura 37);
- (f) la secuencia de nucleótidos mostrada como SEQ ID N° 46 (véase la Figura 39);
- (g) la secuencia de nucleótidos mostrada como SEQ ID N° 48 (véase la Figura 41);
- (h) la secuencia de nucleótidos mostrada como SEQ ID N° 49 (véase la Figura 57);
- (i) la secuencia de nucleótidos mostrada como SEQ ID N° 50 (véase la Figura 58);
- 20 (j) la secuencia de nucleótidos mostrada como SEQ ID N° 51 (véase la Figura 59);
- (k) la secuencia de nucleótidos mostrada como SEQ ID N° 52 (véase la Figura 60);
- (l) la secuencia de nucleótidos mostrada como SEQ ID N° 53 (véase la Figura 61);
- (m) la secuencia de nucleótidos mostrada como SEQ ID N° 54 (véase la Figura 62);
- (n) la secuencia de nucleótidos mostrada como SEQ ID N° 55 (véase la Figura 63);
- 25 (o) la secuencia de nucleótidos mostrada como SEQ ID N° 56 (véase la Figura 64);
- (p) la secuencia de nucleótidos mostrada como SEQ ID N° 57 (véase la Figura 65);
- (q) la secuencia de nucleótidos mostrada como SEQ ID N° 58 (véase la Figura 66);
- (r) la secuencia de nucleótidos mostrada como SEQ ID N° 59 (véase la Figura 67);
- (s) la secuencia de nucleótidos mostrada como SEQ ID N° 60 (véase la Figura 68);
- 30 (t) la secuencia de nucleótidos mostrada como SEQ ID N° 61 (véase la Figura 69);
- (u) la secuencia de nucleótidos mostrada como SEQ ID N° 62 (véase la Figura 70);
- (v) la secuencia de nucleótidos mostrada como SEQ ID N° 63 (véase la Figura 71);
- (w) o

35 una secuencia de nucleótidos que tiene un 70% o más, preferiblemente un 75% o más, de identidad con cualquiera de las secuencias mostradas como SEQ ID N° 36, SEQ ID N° 38, SEQ ID N° 39, SEQ ID N° 42, SEQ ID N° 44, SEQ ID N° 46, SEQ ID N° 48, SEQ ID N° 49, SEQ ID N° 50, SEQ ID N° 51, SEQ ID N° 52, SEQ ID N° 53, SEQ ID N° 54, SEQ ID N° 55, SEQ ID N° 56, SEQ ID N° 57, SEQ ID N° 58, SEQ ID N° 59, SEQ ID N° 60, SEQ ID N° 61, SEQ ID N° 62 o SEQ ID N° 63.

40 De manera adecuada, la secuencia de nucleótidos puede tener un 80% o más, preferiblemente un 85% o más, más preferiblemente un 90% o más y aún más preferiblemente un 95% o más de identidad con cualquiera de las secuencias mostradas como SEQ ID N° 36, SEQ ID N° 38, SEQ ID N° 39, SEQ ID N° 42, SEQ ID N° 44, SEQ ID N° 46, SEQ ID N° 48, SEQ ID N° 49, SEQ ID N° 50, SEQ ID N° 51, SEQ ID N° 52, SEQ ID N° 53, SEQ ID N° 54, SEQ ID N° 55, SEQ ID N° 56, SEQ ID N° 57, SEQ ID N° 58, SEQ ID N° 59, SEQ ID N° 60, SEQ ID N° 61, SEQ ID N° 62 o SEQ ID N° 63.

La secuencia de nucleótidos que codifica una enzima lípido aciltransferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y usos descritos en la presente memoria es una secuencia de nucleótidos que tiene un 70% o más, preferiblemente un 75% o más, de identidad con cualquiera de las secuencias mostradas como: SEQ ID N° 49, SEQ ID N° 50, SEQ ID N° 51, SEQ ID N° 62 y SEQ ID N° 63. De forma adecuada, la secuencia de nucleótidos puede tener un 80% o más, preferiblemente un 85% o más, más preferiblemente un 90% o más e incluso más preferiblemente un 95% o más, de identidad con respecto a una cualquiera de las secuencias mostradas como: SEQ ID N° 49, SEQ ID N° 50, SEQ ID N° 51, SEQ ID N° 62 y SEQ ID N° 63.

En una realización, la secuencia de nucleótidos que codifica una enzima lípido aciltransferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y usos de la presente descripción es una secuencia de nucleótidos que tiene un 70% o más, un 75% o más, 80% o más, preferiblemente un 85% o más, más preferiblemente un 90% o más y aún más preferiblemente un 95% o más de identidad con la secuencia mostrada como SEQ ID N° 49.

La secuencia de nucleótidos que codifica una enzima lípido acil transferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y usos descritos en la presente memoria puede codificar una lípido aciltransferasa que comprende una o más de las secuencias de aminoácidos siguientes:

- (i) la secuencia de aminoácidos mostrada como SEQ ID N° 3
- (ii) la secuencia de aminoácidos mostrada como SEQ ID N° 4
- (iii) la secuencia de aminoácidos mostrada como SEQ ID N° 5
- (iv) la secuencia de aminoácidos mostrada como SEQ ID N° 6
- (v) la secuencia de aminoácidos mostrada como SEQ ID N° 7
- (vi) la secuencia de aminoácidos mostrada como SEQ ID N° 8
- (vii) la secuencia de aminoácidos mostrada como SEQ ID N° 9
- (viii) la secuencia de aminoácidos mostrada como SEQ ID N° 10
- (ix) la secuencia de aminoácidos mostrada como SEQ ID N° 11
- (x) la secuencia de aminoácidos mostrada como SEQ ID N° 12
- (xi) la secuencia de aminoácidos mostrada como SEQ ID N° 13
- (xii) la secuencia de aminoácidos mostrada como SEQ ID N° 14
- (xiii) la secuencia de aminoácidos mostrada como SEQ ID N° 1
- (xiv) la secuencia de aminoácidos mostrada como SEQ ID N° 15 o

una secuencia de aminoácidos que tiene un 75%, 80%, 85%, 90%, 95%, 98% o más de identidad con cualquiera de las secuencias mostradas como SEQ ID N° 1, SEQ ID N° 3, SEQ ID N° 4, SEQ ID N° 5, SEQ ID N° 6, SEQ ID N° 7, SEQ ID N° 8, SEQ ID N° 9, SEQ ID N° 10, SEQ ID N° 11, SEQ ID N° 12, SEQ ID N° 13, SEQ ID N° 14, o SEQ ID N° 15.

Una secuencia de nucleótidos que codifica una enzima lípido acil transferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y usos descritos en la presente memoria puede codificar una lípido aciltransferasa que comprende la secuencia de aminoácidos mostrada como SEQ ID N° 3 o como SEQ ID N° 4 o SEQ ID N° 1 o SEQ ID N° 15 o comprende una secuencia de aminoácidos que tiene un 75% o más, preferiblemente un 80% o más, preferiblemente un 85% o más, preferiblemente un 90% o más, preferiblemente un 95% o más, de identidad con la secuencia de aminoácidos mostrada como SEQ ID N° 3 o la secuencia de aminoácidos mostrada como SEQ ID N° 4 o la secuencia de aminoácidos mostrada como SEQ ID N° 1 o la secuencia de aminoácidos mostrada como SEQ ID N° 15.

La secuencia de nucleótidos que codifica una enzima lípido acil transferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y usos descritos en la presente memoria puede codificar una lípido aciltransferasa que comprende una secuencia de aminoácidos que tiene un 80% o más, preferiblemente un 85% o más, más preferiblemente un 90% o más y aún más preferiblemente un 95% o más de identidad con cualquiera de las secuencias mostradas como SEQ ID N° 3, SEQ ID N° 4, SEQ ID N° 5, SEQ ID N° 6, SEQ ID N° 7, SEQ ID N° 8, SEQ ID N° 9, SEQ ID N° 10, SEQ ID N° 11, SEQ ID N° 12, SEQ ID N° 13, SEQ ID N° 14, SEQ ID N° 1, o SEQ ID N° 15.

La secuencia de nucleótidos que codifica una enzima lípido acil transferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y usos descritos en la presente memoria puede codificar una lípido aciltransferasa que comprende una o más de las secuencias de aminoácidos siguientes:

- (a) una secuencia de aminoácidos mostrada como los residuos de aminoácido 1-100 de la SEQ ID N° 3 o la SEQ ID N° 1;
- (b) una secuencia de aminoácidos mostrada como los residuos de aminoácido 101-200 de la SEQ ID N° 3 o la SEQ ID N° 1;
- 5 (c) una secuencia de aminoácidos mostrada como los residuos de aminoácido 201-300 de la SEQ ID N° 3 o la SEQ ID N° 1; o
- (d) una secuencia de aminoácidos que tiene un 75% o más, preferiblemente un 85% o más, más preferiblemente un 90% o más, incluso más preferiblemente un 95% o más, de identidad con respecto a una cualquiera de las secuencias de aminoácidos definidas en los anteriores apartados (a)-(c).
- 10 La enzima lípido acil transferasa para el uso en los métodos y usos descritos en la presente memoria puede comprender una o más de las secuencias de aminoácidos siguientes:
- (a) una secuencia de aminoácidos mostrada como los residuos de aminoácido 28-39 de SEQ ID N° 3 o SEQ ID N° 1;
- 15 (b) una secuencia de aminoácidos mostrada como los residuos de aminoácido 77-88 de SEQ ID N° 3 o SEQ ID N° 1;
- (c) una secuencia de aminoácidos mostrada como los residuos de aminoácido 126-136 de SEQ ID N° 3 o SEQ ID N° 1;
- (d) una secuencia de aminoácidos mostrada como los residuos de aminoácido 163-175 de SEQ ID N° 3 o SEQ ID N° 1;
- 20 (e) una secuencia de aminoácidos mostrada como los residuos de aminoácido 304-311 de SEQ ID N° 3 o SEQ ID N° 1; o
- (f) una secuencia de aminoácidos que tiene un 75% o más, preferiblemente un 85% o más, más preferiblemente un 90% o más, incluso más preferiblemente un 95% o más, de identidad con respecto a una cualquiera de las secuencias de aminoácidos definidas en los anteriores apartados (a)-(e).
- 25 La secuencia de nucleótidos que codifica una enzima lípido acil transferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y usos descritos en la presente memoria codifica una lípido aciltransferasa que puede ser la lípido acil transferasa de *Candida parapsilosis* como se muestra en el documento EP 1 275 711. Así, la lípido acil transferasa para el uso en el método y usos descritos en la presente memoria puede ser una lípido acil transferasa que comprende una de las secuencias de aminoácidos mostrada en SEQ ID N° 17 o SEQ ID N° 18.
- 30 Como preferencia, la secuencia de nucleótidos que codifica una enzima lípido acil transferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y usos de la presente invención codifica una lípido aciltransferasa que puede ser una lípido acil transferasa (lípido aciltransferasa) que comprende la secuencia de aminoácidos mostrada como SEQ ID N° 16, o una secuencia de aminoácidos que tiene un 75% o más, preferiblemente un 85% o más, más preferiblemente un 90% o más, aún más preferiblemente un 95% o más, aún
- 35 más preferiblemente un 98% o más, o aún más preferiblemente un 99% o más de identidad respecto de SEQ ID N° 16. Esta enzima se podría considerar una enzima variante.
- En un aspecto, la secuencia de nucleótidos que codifica una enzima lípido aciltransferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y usos de la presente invención codifica una lípido aciltransferasa que puede ser una lecitina:colesterol aciltransferasa (LCAT) o una variante de la misma (por ejemplo, una variante hecha mediante evolución molecular).
- 40 En la técnica se conocen LCAT adecuadas y se pueden obtener a partir de uno o más de los siguientes organismos, por ejemplo: mamíferos, rata, arroz, pollos, *Drosophila melanogaster*, plantas, que incluyen *Arabidopsis* y *Oryza sativa*, nemátodos, hongos y levadura.
- En una realización, la secuencia de nucleótidos que codifica una enzima lípido aciltransferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y usos de la presente descripción codifica una lípido aciltransferasa que puede ser la lípido aciltransferasa obtenible, preferiblemente obtenida, de las cepas de *E. coli* TOP 10 que albergan pPet12aAhydro y pPet12aASalmo depositadas por Danisco A/S de Langebrogade 1, DK-1001 Copenhague K, Dinamarca bajo el Tratado de Budapest sobre el Reconocimiento Internacional del Depósito de Microorganismos a los Fines del Procedimiento en Materia de Patentes en la National Collection of Industrial, Marine and Food Bacteria (NCIMB) 23 St. Machar Street, Aberdeen Escocia, GB el 22 de diciembre de 2003 con los
- 50 números de acceso NCIMB 41204 y NCIMB 41205, respectivamente.
- Una secuencia de nucleótidos que codifica una enzima lípido aciltransferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y usos descritos en la presente memoria puede codificar una fosfolípido glicerol

- 5 acil transferasa. Las fosfolípido glicerol acil transferasas incluyen las aisladas de *Aeromonas spp.*, preferiblemente *Aeromonas hydrophila* o *A. salmonicida*, lo más preferible *A. salmonicida* o variantes de la misma. Las lípido acil transferasas descritas en la presente memoria están codificadas por SEQ ID N°s 1, 3, 4, 15 y 16. El especialista reconocerá que es preferible que los péptidos señal de la acil transferasa sean sometidos a ruptura durante la expresión de la transferasa. El péptido señal de SEQ ID 1, 3, 4, 15 y 16 son los aminoácidos 1-18. Por lo tanto, las regiones más preferidas son los aminoácidos 19-335 para SEQ ID N° 1 y SEQ ID N° 3 (*A. hydrophila*) y los aminoácidos 19-336 para SEQ ID N° 4, SEQ ID N° 15 y SEQ ID N° 16. (*A. salmonicida*). Cuando se usan para determinar la homología de identidad de las secuencias de aminoácidos, es preferible que las alineaciones tal como se describen en la presente memoria usen la secuencia madura.
- 10 Por lo tanto, las regiones más preferidas para determinar la homología (identidad) son los aminoácidos 19-335 para SEQ ID N° 1 y 3 (*A. hydrophila*) y los aminoácidos 19-336 para SEQ ID N°s 4, 15 y 16. (*A. salmonicida*). Las SEQ ID N° 34 y 35 son secuencias de proteínas maduras de una lípido acil transferasa de *A. hydrophila* y *A. salmonicida*, respectivamente.
- 15 Una secuencia de nucleótidos que codifica una enzima lípido aciltransferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y usos descritos en la presente memoria codifica una lípido aciltransferasa que se puede aislar también de *Thermobifida*, preferiblemente *T. fusca*, lo más preferiblemente la codificada por SEQ ID N° 28.
- 20 Una secuencia de nucleótidos que codifica una enzima lípido aciltransferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y usos descritos en la presente memoria codifica una lípido aciltransferasa que se puede aislar también de *Streptomyces*, preferiblemente *S. avermitis*, lo más preferiblemente la codificada por SEQ ID N° 32. Otras enzimas posibles de *Streptomyces* descritas en la presente memoria incluyen las codificadas por SEQ ID N°s 5, 6, 9, 10, 11, 12, 13, 14, 31, y 33.
- Una enzima para los usos descritos en la presente memoria también se puede aislar de *Corynebacterium*, preferiblemente *C. efficiens*, lo más preferiblemente la codificada por SEQ ID N° 29.
- 25 La secuencia de nucleótidos que codifica una enzima lípido aciltransferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y usos descritos en la presente memoria puede codificar una lípido aciltransferasa que comprende cualquiera de las secuencias de aminoácidos mostradas como SEQ ID N°s 37, 38, 40, 41, 43, 45, o 47 o una secuencia de aminoácidos que tiene al menos un 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97% o 98% de identidad con ellas, o codificadas por cualquiera de las secuencias de nucleótidos mostradas como SEQ ID N°s 36, 39, 42, 44, 46, o 48 o una secuencia de nucleótidos que tiene al menos un 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97% o 98% de identidad con ellas.
- 30 La secuencia nucleica que codifica una enzima lípido aciltransferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y usos descritos en la presente memoria se selecciona del grupo que consiste en:
- a) un ácido nucleico que comprende una secuencia de nucleótidos mostrada en la SEQ ID N° 36;
- 35 b) un ácido nucleico que está relacionado con la secuencia de nucleótidos de la SEQ ID N° 36 mediante degeneración del código genético; y
- c) un ácido nucleico que comprende una secuencia de nucleótidos que tiene al menos un 70% de identidad con la secuencia de nucleótidos mostrada en SEQ ID N° 36.
- 40 En una realización, una secuencia de nucleótidos que codifica una enzima lípido aciltransferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y usos descritos en la presente memoria codifica una lípido aciltransferasa que comprende una secuencia de aminoácidos como se muestra en SEQ ID N° 37 o una secuencia de aminoácidos que tiene al menos un 60% de identidad con ella.
- 45 La secuencia de nucleótidos que codifica una enzima lípido aciltransferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y usos descritos en la presente memoria puede codificar una lípido aciltransferasa que comprende cualquiera de las secuencias de aminoácidos mostradas como SEQ ID N° 37, 38, 40, 41, 43, 45 ó 47 o una secuencia de aminoácidos que tiene al menos un 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97% o 98% de identidad con ellas, o codificadas por cualquiera de las secuencias de nucleótidos mostradas como SEQ ID N° 39, 42, 44, 46 ó 48 o una secuencia de nucleótidos que tiene al menos un 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97% o 98% de identidad con ellas.
- 50 La secuencia de nucleótidos que codifica una enzima lípido aciltransferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y usos descritos en la presente memoria puede codificar una lípido aciltransferasa que comprende cualquiera de las secuencias de aminoácidos mostradas como SEQ ID N° 38, 40, 41, 45 ó 47 o una secuencia de aminoácidos que tiene al menos un 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97% o 98% de identidad con ellas para los usos descritos en la presente memoria.

- 5 La secuencia de nucleótidos que codifica una lípido aciltransferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y usos descritos en la presente memoria puede codificar una lípido aciltransferasa que comprende cualquiera de las secuencias de aminoácidos mostradas como SEQ ID N° 38, 40, o 47 o una secuencia de aminoácidos que tiene al menos un 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97% o 98% de identidad con ellas para los usos descritos en la presente memoria.
- 10 La secuencia de nucleótidos que codifica una lípido aciltransferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y usos descritos en la presente memoria puede codificar una lípido aciltransferasa que comprende la secuencia de aminoácidos mostrada como SEQ ID N° 47 o una secuencia de aminoácidos que tiene al menos un 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97% o 98% de identidad con ella.
- 15 La secuencia de nucleótidos que codifica una lípido aciltransferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y usos descritos en la presente memoria puede codificar una lípido aciltransferasa que comprende la secuencia de aminoácidos mostrada como SEQ ID N° 41 o una secuencia de aminoácidos que tiene al menos un 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97% o 98% de identidad con ella.
- 20 La secuencia de nucleótidos que codifica una lípido aciltransferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y usos descritos en la presente memoria se selecciona del grupo que consiste en:
- a) un ácido nucleico que comprende una secuencia de nucleótidos mostrada en la SEQ ID N° 36;
  - b) un ácido nucleico que está relacionado con la secuencia de nucleótidos de la SEQ ID N° 36 mediante degeneración del código genético; y
  - c) un ácido nucleico que comprende una secuencia de nucleótidos que tiene al menos un 70% de identidad con la secuencia de nucleótidos mostrada en SEQ ID N° 36.
- 25 La lípido aciltransferasa descrita en la presente memoria puede ser una lípido aciltransferasa obtenible, preferiblemente obtenida, a partir de las cepas de *Streptomyces* L130 o L131 depositadas por Danisco A/S de Langebrogade 1, DK-1001 Copenhagen K, Dinamarca, bajo el Tratado de Budapest sobre el Reconocimiento Internacional del Depósito de Microorganismos a los Fines del Procedimiento en Materia de Patentes en la National Collection of Industrial, Marine and Food Bacteria (NCIMB) 23 St. Machar Street, Aberdeen Escocia, GB el 25 de junio de 2004, con los números de acceso del NCIMB 41226 y NCIMB 41227, respectivamente.
- 30 Las secuencias de nucleótidos que codifican una lípido aciltransferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y usos descritos en la presente memoria pueden codificar un polinucleótido que codifica una lípido aciltransferasa (SEQ ID N° 16); o pueden codificar una secuencia de aminoácidos de una lípido aciltransferasa (SEQ ID N° 17).
- 35 Una secuencia de nucleótidos que codifica una lípido aciltransferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y usos descritos en la presente memoria puede codificar una secuencia de aminoácidos que se puede identificar mediante alineación respecto de la secuencia de L131 (SEQ ID N° 37) con el uso de Align X, el algoritmo de alineación de pares Clustal W de VectorNTI mediante el uso de los ajustes por defecto.
- 40 Una alineación de la L131 y homólogos de *S. avermitilis* y *T. fusca* ilustra la conservación del motivo GDSx (GDSY en L131 y *S. avermitilis* y *T. fusca*), la caja GANDY, que es GGND A o GGNDL, y el bloque HPT (que se considera que es la histidina catalítica conservada). Estos tres bloques conservados están resaltados en la Figura 42.
- 45 Cuando se alinea respecto de la secuencia consenso pfam Pfam00657 (como se describió en el documento WO04/064987) y/o la secuencia de L131 descrita en la presente memoria (SEQ ID N° 37) es posible identificar tres regiones conservadas, el bloque GDSx, el bloque GANDY y el bloque HTP (véase el documento WO04/064987 para detalles adicionales).
- 50 Cuando se alinea respecto de la secuencia consenso pfam Pfam00657 (como se describió en el documento WO04/064987) y/o la secuencia de L131 descrita en la presente memoria (SEQ ID N° 37)
- i) La secuencia de nucleótidos que codifica una lípido aciltransferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y usos descritos en la presente memoria puede codificar una lípido aciltransferasa que tiene un motivo GDS<sub>x</sub>, más preferiblemente un motivo GDS<sub>x</sub> seleccionado del motivo GDSL o GDSY.
- y/o

ii) La secuencia de nucleótidos que codifica una lípido aciltransferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y usos descritos en la presente memoria puede codificar una lípido aciltransferasa que tiene un bloque GANDY, más preferiblemente un bloque GANDY que comprende amino GGNDx, más preferiblemente GGNDa o GGNDL.

5 y/o

iii) La secuencia de nucleótidos que codifica una lípido aciltransferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y usos descritos en la presente memoria codifica una lípido aciltransferasa que tiene preferiblemente un bloque HTP.

y preferiblemente

10 iv) La secuencia de nucleótidos que codifica una lípido aciltransferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y usos descritos en la presente memoria puede codificar una lípido aciltransferasa que tiene preferiblemente un motivo GDSx o GDSY, y un bloque GANDY que comprende amino GGNDx, preferiblemente GGNDa o GGNDL, y un bloque HTP (histidina conservada).

Lípido acil transferasa variante

15 En una realización preferida, la secuencia de nucleótidos que codifica una lípido aciltransferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y usos de la presente descripción puede codificar una lípido aciltransferasa que es una lípido acil transferasa variante.

20 Se pueden usar las variantes que tienen una actividad incrementada sobre los fosfolípidos, tal como actividad hidrolítica incrementada y/o actividad transferasa incrementada, preferiblemente actividad transferasa incrementada sobre los fosfolípidos.

Preferiblemente, la lípido aciltransferasa variante se prepara mediante una o más modificaciones de aminoácidos de las lípido acil transferasas como se definieron anteriormente en la presente memoria.

25 De manera adecuada, la secuencia de nucleótidos que codifica una lípido aciltransferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y usos de la presente descripción puede codificar una lípido aciltransferasa que puede ser una lípido aciltransferasa variante, en cuyo caso la enzima se puede caracterizar porque la enzima comprende el motivo de secuencia de aminoácidos GDSX, en el que X es uno o más de los residuos de aminoácido siguientes L, A, V, I, F, Y, H, Q, T, N, M o S, y en la que la enzima variante comprende una o más modificaciones de aminoácidos en comparación con una secuencia originaria en uno o más de los residuos de aminoácido definidos en el conjunto 2 o conjunto 4 o conjunto 6 o conjunto 7 (como se define en el documento WO2005/066347 y más adelante en la presente memoria).

30 Por ejemplo, la lípido aciltransferasa variante se puede caracterizar porque la enzima comprende el motivo de secuencia de aminoácidos GDSX, en el que X es uno o más de los residuos de aminoácido siguientes L, A, V, I, F, Y, H, Q, T, N, M o S, y en la que la enzima variante comprende una o más modificaciones de aminoácidos en comparación con una secuencia originaria en uno o más de los residuos de aminoácido detallados en el conjunto 2 o conjunto 4 o conjunto 6 o conjunto 7 (como se define en el documento WO2005/066347 y más adelante en la presente memoria) identificadas alineando estructuralmente dicha secuencia originaria con el modelo estructural de P10480 definido en la presente memoria, que se obtiene preferiblemente mediante alineación estructural de las coordenadas de la estructura cristalina de P10480 con 1IVN.PDB y/o 1DEO.PDB como se define en el documento WO2005/066347 y más adelante en la presente memoria.

35 Una secuencia de nucleótidos que codifica una lípido aciltransferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y usos descritos en la presente memoria puede codificar una lípido aciltransferasa variante que se puede caracterizar porque la enzima comprende el motivo de secuencia de aminoácidos GDSX, en el que X es uno o más de los residuos de aminoácido siguientes L, A, V, I, F, Y, H, Q, T, N, M o S, y en la que la enzima variante comprende una o más modificaciones de aminoácidos en comparación con una secuencia originaria en uno o más de los residuos de aminoácidos mostrados en el conjunto 2 identificados cuando dicha secuencia originaria se alinea con la secuencia consenso pfam (SEQ ID N° 2 - Figura 3) y se modifica según un modelo estructural de P10480 para asegurar un solapamiento con el mejor ajuste como se define en el documento WO2005/066347 y más adelante en la presente memoria.

40 La secuencia de nucleótidos que codifica una lípido aciltransferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y usos descritos en la presente memoria puede codificar una enzima lípido aciltransferasa variante que puede comprender una secuencia de aminoácidos, cuya secuencia de aminoácidos se muestra como SEQ ID N° 34, SEQ ID N° 3, SEQ ID N° 4, SEQ ID N° 5, SEQ ID N° 6, SEQ ID N° 7, SEQ ID N° 8, SEQ ID N° 19, SEQ ID N° 10, SEQ ID N° 11, SEQ ID N° 12, SEQ ID N° 13, SEQ ID N° 14, SEQ ID N° 1, SEQ ID N° 15, SEQ ID N° 25, SEQ ID N° 26, SEQ ID N° 27, SEQ ID N° 28, SEQ ID N° 29, SEQ ID N° 30, SEQ ID N° 32, o SEQ ID N° 33 excepto por una o más modificaciones de aminoácidos en uno o más de los residuos de aminoácidos

definidos en el conjunto 2 o conjunto 4 o conjunto 6 o conjunto 7 (como se define en el documento WO2005/066347 y más adelante en la presente memoria) identificados mediante alineación de secuencias con SEQ ID N° 34.

De manera alternativa, la secuencia de nucleótidos que codifica una lípido aciltransferasa puede codificar una enzima lípido aciltransferasa variante que comprende una secuencia de aminoácidos, cuya secuencia de aminoácidos se muestra como SEQ ID N° 34, SEQ ID N° 3, SEQ ID N° 4, SEQ ID N° 5, SEQ ID N° 6, SEQ ID N° 7, SEQ ID N° 8, SEQ ID N° 19, SEQ ID N° 10, SEQ ID N° 11, SEQ ID N° 12, SEQ ID N° 13, SEQ ID N° 14, SEQ ID N° 1, SEQ ID N° 15, SEQ ID N° 25, SEQ ID N° 26, SEQ ID N° 27, SEQ ID N° 28, SEQ ID N° 29, SEQ ID N° 30, SEQ ID N° 32, o SEQ ID N° 33 excepto por una o más modificaciones de aminoácidos en uno cualquiera o más de los residuos de aminoácidos definidos en el conjunto 2 o conjunto 4 o conjunto 6 o conjunto 7 como se definieron en el documento WO2005/066347 y más adelante en la presente memoria, identificadas alineando estructuralmente dicha secuencia originaria con el modelo estructural de P10480 definido en la presente memoria, que se obtiene preferiblemente mediante alineación estructural de las coordenadas de la estructura cristalina de P10480 con 1IVN.PDB y/o 1 DEO.PDB como se muestra en el documento WO2005/066347 y más adelante en la presente memoria.

De manera alternativa, la secuencia de nucleótidos que codifica una lípido aciltransferasa puede codificar una enzima lípido aciltransferasa variante que comprende una secuencia de aminoácidos, cuya secuencia de aminoácidos se muestra como SEQ ID N° 34, SEQ ID N° 3, SEQ ID N° 4, SEQ ID N° 5, SEQ ID N° 6, SEQ ID N° 7, SEQ ID N° 8, SEQ ID N° 19, SEQ ID N° 10, SEQ ID N° 11, SEQ ID N° 12, SEQ ID N° 13, SEQ ID N° 14, SEQ ID N° 1, SEQ ID N° 15, SEQ ID N° 25, SEQ ID N° 26, SEQ ID N° 27, SEQ ID N° 28, SEQ ID N° 29, SEQ ID N° 30, SEQ ID N° 32, o SEQ ID N° 33 excepto por una o más modificaciones de aminoácidos en uno o más de los residuos de aminoácido mostrados en el conjunto 2 identificados cuando dicha secuencia originaria se alinea con la secuencia consenso pfam (SEQ ID N° 2) y se modifica según un modelo estructural de P10480 para asegurar el solapamiento con el mejor ajuste como se muestra en el documento WO2005/066347 y más adelante en la presente memoria.

Preferiblemente, la enzima originaria es una enzima que comprende, o que es homóloga a, la secuencia de aminoácidos mostrada como SEQ ID N° 34 y/o SEQ ID N° 15 y/o SEQ ID N° 35.

Preferiblemente, la secuencia de nucleótidos que codifica una lípido aciltransferasa puede codificar una enzima variante que comprende una secuencia de aminoácidos, cuya secuencia de aminoácidos se muestra como SEQ ID N° 34 o SEQ ID N° 35 excepto por una o más modificaciones de aminoácidos en uno o más de los residuos de aminoácido definidos en el conjunto 2 o conjunto 4 o conjunto 6 o conjunto 7 como se define en el documento WO2005/066347 y más adelante en la presente memoria.

### Definición de conjuntos

Conjunto de aminoácidos 1:

Conjunto de aminoácidos 1 (obsérvese que éstos son los aminoácidos de 1IVN - Figura 53 y Figura 54) Gly8, Asp9, Ser10, Leu11, Ser12, Tyr15, Gly44, Asp45, Thr46, Glu69, Leu70, Gy71, Gly72, Asn73, Asp74, Gly75, Lieu76, Gln106, Ile107, Arg108, Leu109, Pro110, Tyr113, Phe121, Phe139, Phe140, Met141, Tyr145, Met151, Asp154, His157, Gly155, Ile156, Pro158

Los motivos altamente conservados, tales como GDS<sub>x</sub> y residuos catalíticos, fueron deseleccionados del conjunto 1 (residuos subrayados). Para evitar dudas, el conjunto 1 define los residuos de aminoácido a menos de 10 Å del átomo de carbono central de un glicerol en el centro activo del modelo 1IVN.

Conjunto de aminoácidos 2:

Conjunto de aminoácidos 2 (obsérvese que la numeración de los aminoácidos se refiere a los aminoácidos en la secuencia madura P10480)

Leu17, Lys22, Met23, Gly40, Asn80, Pro81, Lys82, Asn87, Asn88, Trp111, Val112, Ala114, Tyr117, Leu118, Pro156, Gly159, Glen160, Asn161, Pro162, Ser163, Ala164, Arg165, Ser166, Gln167, Lys168, Val169, Val170, Glu171, Ala172, Tyr179, His180, Asn181, Met209, Leu210, Arg211, Asn215, Lys284, Met285, Gln289 y Val290.

Tabla de residuos seleccionados en el Conjunto 1 comparados con el Conjunto 2:

Modelo IVN		P10480	
IVN	Homólogo A.hy	Número de Residuo en la Secuencia Madura	
	PFAM	Estructura	
Gly8	Gly32		
Asp9	Asp33		
Ser10	Ser34		
Leu11	Leu35	Leu17	



ES 2 449 474 T3

Modelo IVN			P10480
IVN	Homólogo A.hyd		Número de Residuo en la Secuencia Madura
	PFAM	Estructura	
Ser12	Ser36		Ser18
			Lys22
			Met23
Tyr15	Gly58		Gly40
Gly44	Asn98		Asn80
Asp45	Pro99		Pro81
Thr46	Lys100		Lys82
			Asn87
			Asn88
Glu69	Trp129		Trp111
Leu70	Val130		Val112
Gly71	Gly131		
Gly72	Ala132		Ala114
Asn73	Asn133		
Asp74	Asp134		
Gly75	Tyr135		Tyr117
Leu76	Leu136		Leu118
Gln106		Pro174	Pro156
Ile107		Gly177	Gly159
Arg108		Gln178	Gln160
Leu109		Asn179	Asn161
Pro110		180 a 190	Pro162
Tyr113			Ser163
			Ala164
			Arg165
			Ser166
			Gln167
			Lys168
			Val169
			Val170
			Glu171
			Ala172
Phe121	His198	Tyr197	Tyr179
		His198	His180
		Asn199	Asn181
Phe139	Met227		Met209
Phe140	Leu228		Leu210
Met141	Arg229		Arg211
Tyr145	Asn233		Asn215
			Lys284
Met151	Met303		Met285
Asp154	Asp306		
Gly155	Gln307		Gln289
Ile156	Val308		Val290
His157	His309		
Pro158	Pro310		

Conjunto de aminoácidos 3:

5 El conjunto de aminoácidos 3 es idéntico al conjunto 2 pero se refiere a la secuencia codificante de *Aeromonas salmonicida* (SEQ ID N° 4), es decir, los números de los residuos de aminoácido son 18 unidades mayores en el conjunto 3, ya que esto refleja la diferencia entre la numeración de aminoácidos en la proteína madura (SEQ ID N° 34) en comparación con la proteína que incluye una secuencia señal (SEQ ID N° 25).

10 Las proteínas maduras de GDSX de *Aeromonas salmonicida* (SEQ ID N° 4) y de GDSX de *Aeromonas hydrophila* (SEQ ID N° 34) difieren en cinco aminoácidos. Estos son Thr3Ser, Gln182Lys, Glu309Ala, Ser310Asn, y Gly318-, en los que el residuo de *salmonicida* se enumera primero y el residuo de *hydrophila* se enumera después. La proteína de *hydrophila* tiene una longitud de solamente 317 aminoácidos y carece de un residuo en la posición 318. El GDSX de *Aeromonas salmonicida* tiene una actividad considerablemente mayor sobre los lípidos polares, tales como sustratos galactolipídicos, que la proteína de *Aeromonas hydrophila*. Se llevó a cabo un escaneo de sitio para las cinco posiciones de aminoácido.

Conjunto de aminoácidos 4:

El conjunto de aminoácidos 4 es S3, Q182, E309, S310, y -318.

15 Conjunto de aminoácidos 5:

F13S, D15N, S18G, S18V, Y30F, D116N, D116E, D157 N, Y226F, D228N, Y230F.

Conjunto de aminoácidos 6:

20 El conjunto de aminoácidos 6 es Ser3, Leu17, Lys22, Met23, Gly40, Asn80, Pro81, Lys82, Asn 87, Asn88, Trp111, Val112, Ala114, Tyr117, Leu118, Pro156, Gly159, Gln160, Asn161, Pro162, Ser163, Ala164, Arg165, Ser166, Gln167, Lys168, Val169, Val170, Glu171, Ala172, Tyr179, His180, Asn181, Gln182, Met209, Leu210, Arg211, Asn215, Lys284, Met285, Gln289, Val290, Glu309, Ser310, -318.

La numeración de los aminoácidos del conjunto 6 se refiere a los residuos de aminoácido de P10480 (SEQ ID N° 25) - se pueden determinar los aminoácidos correspondientes en otras cadenas principales de secuencia mediante alineación de homología y/o alineación estructural con P10480 y/o 1IVN.

25 Conjunto de aminoácidos 7:

30 El conjunto de aminoácidos 7 es Ser3, Leu17, Lys22, Met23, Gly40, Asn80, Pro81, Lys82, Asn 87, Asn88, Trp111, Val112, Ala114, Tyr117, Leu118, Pro156, Gly159, Gln160, Asn161, Pro162, Ser163, Ala164, Arg165, Ser166, Gln167, Lys168, Val169, Val170, Glu171, Ala172, Tyr179, His180, Asn181, Gln182, Met209, Leu210, Arg211, Asn215, Lys284, Met285, Gln289, Val290, Glu309, Ser310, -318, Y30X (en el que X se selecciona de A, C, D, E, G, H, I, K, L, M, N, P, Q, R, S, T, V, o W), Y226X (en el que X se selecciona de A, C, D, E, G, H, I, K, L, M, N, P, Q, R, S, T, V, o W), Y230X (en el que X se selecciona de A, C, D, E, G, H, I, K, L, M, N, P, Q, R, S, T, V, o W), S18X (en el que X se selecciona de A, C, D, E, F, H, I, K, L, M, N, P, Q, R, T, W o Y), D157X (en el que X se selecciona de A, C, E, F, G, H, I, K, L, M, P, Q, R, S, T, V, W o Y).

35 La numeración de los aminoácidos del conjunto 7 se refiere a los residuos de aminoácido de P10480 (SEQ ID N° 25) - se pueden determinar los aminoácidos correspondientes en otras cadenas principales de secuencia mediante alineación de homología y/o alineación estructural con P10480 y/o 1IVN.

La enzima variante descrita en la presente memoria puede comprender una o más de las modificaciones de aminoácidos siguientes en comparación con la enzima originaria:

S3E, A, G, K, M, Y, R, P, N, T o G

40 E309Q, R o A, preferiblemente Q o R

-318Y, H, S o Y, preferiblemente Y.

Preferiblemente, X del motivo GDSX es L. Por tanto, preferiblemente la enzima originaria comprende el motivo de aminoácidos GDSL.

45 La primera lípido aciltransferasa originaria puede comprender cualquiera de las secuencias de aminoácidos siguientes: SEQ ID N° 34, SEQ ID N° 3, SEQ ID N° 4, SEQ ID N° 5, SEQ ID N° 6, SEQ ID N° 7, SEQ ID N° 8, SEQ ID N° 19, SEQ ID N° 10, SEQ ID N° 11, SEQ ID N° 12, SEQ ID N° 13, SEQ ID N° 14, SEQ ID N° 1, SEQ ID N° 15, SEQ ID N° 25, SEQ ID N° 26, SEQ ID N° 27, SEQ ID N° 28, SEQ ID N° 29, SEQ ID N° 30, SEQ ID N° 32 o SEQ ID N° 33.

50 De forma adecuada, dicha segunda lípido aciltransferasa relacionada puede comprender una cualquiera de las siguientes secuencias de aminoácidos: SEQ ID N° 3, SEQ ID N° 34, SEQ ID N° 4, SEQ ID N° 5, SEQ ID N° 6, SEQ

## ES 2 449 474 T3

ID N° 7, SEQ ID N° 8, SEQ ID N° 19, SEQ ID N° 10, SEQ ID N° 11, SEQ ID N° 12, SEQ ID N° 13, SEQ ID N° 14, SEQ ID N° 1, SEQ ID N° 15, SEQ ID N° 25, SEQ ID N° 26, SEQ ID N° 27, SEQ ID N° 28, SEQ ID N° 29, SEQ ID N° 30, SEQ ID N° 32 o SEQ ID N° 33.

5 La enzima variante comprende al menos una modificación de aminoácido con respecto a la enzima originaria. La enzima variante descrita en la presente memoria puede comprender al menos 2, preferiblemente al menos 3, preferiblemente al menos 4, preferiblemente al menos 5, preferiblemente al menos 6, preferiblemente al menos 7, preferiblemente al menos 8, preferiblemente al menos 9, preferiblemente al menos 10, modificaciones de aminoácidos en comparación con la enzima originaria.

10 Cuando se hace referencia a residuos de aminoácidos específicos en la presente memoria, la numeración es la obtenida de la alineación de la secuencia variante con la secuencia de referencia mostrada como SEQ ID N° 34 o SEQ ID N° 35.

La enzima variante descrita en la presente memoria puede comprender una o más de las sustituciones de aminoácidos siguientes:

- S3A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, T, V, W o Y; y/o
- 15 L17A, C, D, E, F, G, H, I, K, M, N, P, Q, R, S, T, V, W o Y; y/o  
S18A, C, D, E, F, H, I, K, L, M, N, P, Q, R, T, W o Y; y/o  
K22A, C, D, E, F, G, H, I, L, M, N, P, Q, R, S, T, V, W, o Y; y/o  
M23A, C, D, E, F, G, H, I, K, L, N, P, Q, R, S, T, V, W, o Y; y/o  
Y30A, C, D, E, G, H, I, K, L, M, N, P, Q, R, S, T, V o W; y/o
- 20 G40A, C, D, E, F, H, I, K, L, M, N, P, Q, R, S, T, V, W o Y; y/o  
N80A, C, D, E, F, G, H, I, K, L, M, P, Q, R, S, T, V, W o Y; y/o  
P81A, C, D, E, F, G, H, I, K, L, M, N, Q, R, S, T, V, W o Y; y/o  
K82A, C, D, E, F, G, H, I, L, M, N, P, Q, R, S, T, V, W o Y; y/o  
N87A, C, D, E, F, G, H, I, K, L, M, P, Q, R, S, T, V, W o Y; y/o
- 25 N88A, C, D, E, F, G, H, I, K, L, M, P, Q, R, S, T, V, W o Y; y/o  
W111A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W o Y; y/o  
V112A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, W o Y; y/o  
A114C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W o Y; y/o  
Y117A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V o W; y/o
- 30 L118A, C, D, E, F, G, H, I, K, M, N, P, Q, R, S, T, V, W o Y; y/o  
P156A, C, D, E, F, G, H, I, K, L, M, N, Q, R, S, T, V, W, o Y; y/o  
D157A, C, E, F, G, H, I, K, L, M, P, Q, R, S, T, V, W o Y; y/o  
G159A, C, D, E, F, H, I, K, L, M, N, P, Q, R, S, T, V, W o Y; y/o  
Q160A, C, D, E, F, G, H, I, K, L, M, N, P, R, S, T, V, W o Y; y/o
- 35 N161A, C, D, E, F, G, H, I, K, L, M P, Q, R, S, T, V, W o Y; y/o  
P162A, C, D, E, F, G, H, I, K, L, M, N, Q, R, S, T, V, W, o Y; y/o  
S163A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, T, V, W o Y; y/o  
A164C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W o Y; y/o  
R165A, C, D, E, F, G, H, 1, K, L, M, N, P, Q, S, T, V, W, o Y; y/o
- 40 S166A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, T, V, W o Y; y/o  
Q167A, C, D, E, F, G, H, I, K, L, M, N, P, R, S, T, V, W o Y; y/o

- K168A, C, D, E, F, G, H, I, L, M, N, P, Q, R, S, T, V, W o Y; y/o  
V169A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, W o Y; y/o  
V170A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, W, o Y; y/o  
E171A, C, D, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W o Y; y/o  
5 A172C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W o Y; y/o  
Y179A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V o W; y/o  
H180A, C, D, E, F, G, I, K, L, M, P, Q, R, S, T, V, W, o Y; y/o  
N181A, C, D, E, F, G, H, I, K, L, M, P, Q, R, S, T, V, W o Y; y/o  
G182A, C, D, E, F, G, H, I, K, L, M, N, P, R, S, T, V, W, o Y, preferiblemente K; y/o  
10 M209A, C, D, E, F, G, H, I, K, L, N, P, Q, R, S, T, V, W o Y; y/o  
L210 A, C, D, E, F, G, H, I, K, M, N, P, Q, R, S, T, V, W o Y; y/o  
R211 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, o Y; y/o  
N215 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W o Y; y/o  
Y226A, C, D, E, G, H, I, K, L, M, N, P, Q, R, S, T, V o W; y/o  
15 Y230A, C, D, E, G, H, I, K, L, M, N, P, Q, R, S, T, V o W; y/o  
K284A, C, D, E, F, G, H, I, L, M, N, P, Q, R, S, T, V, W o Y; y/o  
M285A, C, D, E, F, G, H, I, K, L, N, P, Q, R, S, T, V, W o Y; y/o  
Q289A, C, D, E, F, G, H, I, K, L, M, N, P, R, S, T, V, W o Y; y/o  
V290A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, W o Y; y/o  
20 E309A, C, D, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W o Y; y/o  
S310A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, T, V, W o Y.

Además de, o alternativamente, puede haber una o más extensiones C-terminales. Preferiblemente, la extensión C-terminal adicional está compuesta de uno o más aminoácidos alifáticos, preferiblemente un aminoácido apolar, más preferiblemente de I, L, V o G. Así, la enzima variante descrita en la presente memoria puede comprender una o más de las siguientes extensiones C-terminales: 3181, 318L, 318V, 318G.

Las enzimas variantes preferidas pueden presentar una actividad hidrolítica disminuida frente a fosfolípidos, tal como fosfatidilcolina (PC), también pueden presentar una actividad de transferasa incrementada a partir de un fosfolípido.

Las enzimas variantes preferidas pueden presentar una actividad de transferasa incrementada a partir de un fosfolípido, tal como fosfatidilcolina (PC), también pueden presentar una actividad hidrolítica incrementada frente a un fosfolípido.

La modificación de uno o más de los siguientes residuos puede dar como resultado una enzima variante que tiene una actividad de transferasa absoluta incrementada frente a fosfolípido:

S3, D157, S310, E309, Y179, N215, K22, Q289, M23, H180, M209, L210, R211, P81, V112, N80, L82, N88, N87

Las modificaciones específicas preferidas que pueden proporcionar una enzima variante que tiene una actividad de transferasa mejorada a partir de un fosfolípido pueden seleccionarse entre una o más de las siguientes:

S3A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, T, V, W o Y; preferiblemente N, E, K, R, A, P o M, lo más preferiblemente S3A

D157A, C, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W o Y; preferiblemente D157S, R, E, N, G, T, V, Q, K o C

S310A, C, D, E, F, G, H, i, K, L, M, N, P, Q, R, T, V, W o Y; preferiblemente S310T -318 E

40 E309A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, T, V, W o Y; preferiblemente E309 R, E, L, R o A

## ES 2 449 474 T3

- Y179A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V o W; preferiblemente Y179 D, T, E, R, N, V, K, Q o S, más preferiblemente E, R, N, V, K o Q
- N215A, C, D, E, F, G, H, I, K, L, M, P, Q, R, S, T, V, W o Y; preferiblemente N215 S, L, R o Y
- K22A, C, D, E, F, G, H, I, L, M, N, P, Q, R, S, T, V, W o Y; preferiblemente K22 E, R, C o A
- 5 Q289A, C, D, E, F, G, H, I, K, L, M, N, P, R, S, T, V, W o Y; preferiblemente Q289 R, E, G, P o N
- M23A, C, D, E, F, G, H, I, K, L, N, P, Q, R, S, T, V, W o Y; preferiblemente M23 K, Q, L, G, T o S
- H180A, C, D, E, F, G, I, K, L, M, P, Q, R, S, T, V, W o Y; preferiblemente H180 Q, R o K M209 A, C, D, E, F, G, H, I, K, L, N, P, Q, R, S, T, V, W o Y; preferiblemente M209 Q, S, R, A, N, Y, E, V o L
- L210A, C, D, E, F, G, H, I, K, M, N, P, Q, R, S, T, V, W o Y; preferiblemente L210 R, A, V, S, T, I, W o M
- 10 R211A, C, D, E, F, G, H, I, K, L, M, N, P, Q, S, T, V, W o Y; preferiblemente R211T
- P81A, C, D, E, F, G, H, I, K, L, M, N, Q, R, S, T, V, W o Y; preferiblemente P81G
- V112A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, W o Y; preferiblemente V112C
- N80A, C, D, E, F, G, H, I, K, L, M, P, Q, R, S, T, V, W o Y; preferiblemente N80 R, G, N, D, P, T, E, V, A o G
- 15 L82A, C, D, E, F, G, H, I, M, N, P, Q, R, S, T, V, W o Y; preferiblemente L82N, S o E N88A, C, D, E, F, G, H, I, K, L, M, P, Q, R, S, T, V, W o Y; preferiblemente N88C
- N87A, C, D, E, F, G, H, I, K, L, M, P, Q, R, S, T, V, W o Y; preferiblemente N87M o G

La modificación preferida de uno o más de los siguientes residuos da como resultado una enzima variante que tiene una actividad de transferasa absoluta incrementada frente a fosfolípido:

- S3 N, R, A, G
- 20 M23 K, Q, L, G, T, S
- H 180 R
- L82 G
- Y179 E, R, N, V, K o Q
- E309 R, S, L o A
- 25 Una modificación preferida es N80D. Este es el caso especialmente cuando se usa la secuencia de referencia SEQ ID N° 35 como cadena principal. Así, la secuencia de referencia puede ser SEQ ID N° 16. Esta modificación se puede dar en combinación con una o más modificaciones adicionales. Por lo tanto, la secuencia de nucleótidos que codifica una lípido aciltransferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y usos descritos en la presente memoria puede codificar una lípido aciltransferasa que comprende SEQ ID N° 35.
- 30 Como se indicó anteriormente, cuando se hace referencia a residuos de aminoácidos específicos en la presente memoria, la numeración es la obtenida de la alineación de la secuencia variante con la secuencia de referencia mostrada como SEQ ID N° 34 o SEQ ID N° 35.
- 35 Como preferencia, la secuencia de nucleótidos que codifica una lípido aciltransferasa para el uso en cualquiera de las células hospedadoras, vectores, métodos y usos descritos en la presente memoria puede codificar un lípido que comprende la secuencia de aminoácidos mostrada como SEQ ID N° 16, o una secuencia de aminoácidos que tiene un 75% o más, preferiblemente un 85% o más, más preferiblemente un 90% o más, aún más preferiblemente un 95% o más, aún más preferiblemente un 98% o más, o aún más preferiblemente un 99% o más de identidad respecto de SEQ ID N° 16. Esta enzima puede considerarse una enzima variante.
- 40 Para los fines descritos en la presente memoria, el grado de identidad se basa en el número de elementos de secuencia que son iguales. El grado de identidad para las secuencias de aminoácidos puede determinarse de forma adecuada por medio de programas de ordenador conocidos en la técnica, tal como Vector NTI 10 (Invitrogen Corp.). Para la alineación por pares la puntuación usada preferiblemente es BLOSUM62 con una penalización por apertura de hueco de 10,0 y una penalización por extensión de hueco de 0,1.
- 45 De forma adecuada, el grado de identidad con respecto a una secuencia de aminoácidos se determina sobre al menos 20 aminoácidos contiguos, preferiblemente sobre al menos 30 aminoácidos contiguos, preferiblemente sobre

al menos 40 aminoácidos contiguos, preferiblemente sobre al menos 50 aminoácidos contiguos, preferiblemente sobre al menos 60 aminoácidos contiguos.

De forma adecuada, el grado de identidad con respecto a la secuencia de aminoácidos puede determinarse sobre la secuencia completa.

- 5 De manera adecuada, la secuencia de nucleótidos que codifica una enzima lípido aciltransferasa/lípido acil transferasa descrita en la presente memoria puede ser obtenible, y preferiblemente se obtiene, a partir de organismos de uno o más de los géneros siguientes: *Aeromonas*, *Streptomyces*, *Saccharomyces*, *Lactococcus*, *Mycobacterium*, *Streptococcus*, *Lactobacillus*, *Desulfibacterium*, *Bacillus*, *Campylobacter*, *Vibrionaceae*, *Xylella*, *Sulfolobus*, *Aspergillus*, *Schizosaccharomyces*, *Listeria*, *Neisseria*, *Mesorhizobium*, *Ralstonia*, *Xanthomonas*,  
10 *Candida*, *Thermobifida* y *Corynebacterium*.

- De manera adecuada, la secuencia de nucleótidos que codifica una enzima lípido aciltransferasa/lípido acil transferasa descrita en la presente memoria puede ser obtenible, y preferiblemente se obtiene, a partir de uno o más de los organismos siguientes: *Aeromonas hydrophila*, *Aeromonas salmonicida*, *Streptomyces coelicolor*, *Streptomyces rimosus*, *Mycobacterium*, *Streptococcus pyogenes*, *Lactococcus lactis*, *Streptococcus pyogenes*,  
15 *Streptococcus thermophilus*, *Streptomyces thermosacchari*, *Streptomyces avermitilis*, *Lactobacillus helveticus*, *Desulfibacterium dehalogenans*, *Bacillus sp.*, *Campylobacter jejuni*, *Vibrionaceae*, *Xylella fastidiosa*, *Sulfolobus solfataricus*, *Saccharomyces cerevisiae*, *Aspergillus terreus*, *Schizosaccharomyces pombe*, *Listeria innocua*, *Listeria monocytogenes*, *Neisseria meningitidis*, *Mesorhizobium loti*, *Ralstonia solanacearum*, *Xanthomonas campestris*, *Xanthomonas axonopodis*, *Candida parapsilosis*, *Thermobifida fusca* y *Corynebacterium efficiens*.

- 20 En un aspecto, preferiblemente la secuencia de nucleótidos que codifica una lípido aciltransferasa para el uso en cualquier método de la presente invención codifica una enzima lípido acil transferasa según la presente invención obtenible, preferiblemente obtenida o derivada, de *Aeromonas spp.*, o *Aeromonas salmonicida*.

- Las enzimas que funcionan como lípido aciltransferasas de acuerdo con la presente invención se pueden identificar de manera rutinaria mediante el uso del ensayo mostrado en el Ejemplo 12 del documento WO2004/064537.  
25 Mediante el uso de este ensayo, en el que hay un contenido de agua muy elevado, aproximadamente del 95%, las lípido aciltransferasas/lípido acil transferasa de acuerdo con la presente invención son aquellas que tienen al menos un 2% de actividad de aciltransferasa (actividad de transferasa relativa), preferiblemente al menos un 5% de actividad de transferasa relativa, preferiblemente al menos un 10% de actividad de transferasa relativa, preferiblemente al menos un 15%, 20%, 25% 26%, 28%, 30%, 40% 50%, 60% o 75% de actividad de transferasa  
30 relativa.

Las fosfolipasas pueden actuar como enzimas acil-transferasas en medios con bajo contenido de agua. Por lo tanto, se considera que en lugar de o además de la enzima fosfolípido aciltransferasa se puede usar una enzima fosfolipasa cuando el proceso para la modificación del aceite o grasa comestible tiene lugar en un medio con bajo contenido de agua.

- 35 La expresión "contenido elevado de agua", tal como se usa en la presente memoria, significa cualquier sustrato o producto alimenticio con más de un 3% de contenido de agua, preferiblemente más de un 4%, más de un 5%, más de un 6%, más de un 7%, más de un 8%, más de un 9%, más de un 10%, más de un 20%, más de un 30%, más de un 40%, más de un 50%, más de un 60%, más de un 70%, más de un 80% o más de un 90%.

- La expresión "contenido bajo de agua", tal como se usa en la presente memoria, significa cualquier sustrato o  
40 producto alimenticio con menos de un 3% de contenido de agua, preferiblemente menos de un 2%, menos de un 1% o menos de un 0,5%, menos de un 0,3%, menos de un 0,2, menos de un 0,1, menos de un 0,05, o menos de un 0,01%.

Para evitar dudas, la leche es un medio con contenido elevado de agua mientras que la mantequilla es un medio con contenido bajo de agua.

- 45 Las fosfolipasas adecuadas para los usos descritos en la presente memoria incluyen la fosfolipasa A1, fosfolipasa A2, o fosfolipasa B. La fosfolipasa A1, fosfolipasa A2, o fosfolipasa B se pueden usar también coordinadamente con la actividad de lípido acil transferasa. La fosfolipasa C y/o D se pueden usar también coordinadamente con la actividad de lípido acil transferasa/actividad de fosfolipasa A1, A2 y/o B análogamente al documento WO2005/089562. Las fosfolipasas preferidas pueden incluir la fosfolipasa A2, tal como Lecitase<sup>TM</sup> o la fosfolipasa  
50 de *Fusarium venenatum* y *Tuber albidum* descritas en el documento WO2004/97012 (Novozymes/Chr. Hansen). La fosfolipasa de *Fusarium venenatum* la comercializa Novozymes como MAX YIELD<sup>TM</sup>.

#### Aislada

En un aspecto, el método de la presente invención comprende la etapa adicional de recuperar/aislar la lípido aciltransferasa. Así, la lípido aciltransferasa producida puede estar en forma aislada.

En otro aspecto, la secuencia de nucleótidos que codifica una lípido aciltransferasa para el uso en la presente invención puede estar en forma aislada.

5 El término "aislada" significa que la secuencia o proteína está al menos sustancialmente libre de al menos otro componente con el que la secuencia o proteína está asociada de manera natural en la naturaleza y como se halla en la naturaleza.

**Purificada**

En un aspecto, el método de la presente invención comprende la etapa adicional de purificar la lípido aciltransferasa.

En otro aspecto, la secuencia de nucleótidos que codifica una lípido aciltransferasa para el uso en la presente invención puede estar en forma purificada.

10 El término "purificada" significa que la secuencia se encuentra en un estado relativamente puro -por ejemplo, al menos aproximadamente pura en un 51%, o al menos aproximadamente un 75%, o al menos aproximadamente un 80%, o al menos pura en aproximadamente un 90%, o al menos pura en aproximadamente un 95% o al menos pura en aproximadamente un 98%.

**Clonación de una secuencia de nucleótidos que codifica un polipéptido según la presente invención**

15 Una secuencia de nucleótidos que codifica un polipéptido que tiene las propiedades específicas definidas en la presente memoria, o un polipéptido que es adecuado para modificación, puede ser aislado a partir de cualquier célula u organismo que produzca dicho polipéptido. En la técnica se conocen varios métodos para el aislamiento de secuencias de nucleótidos.

20 Por ejemplo, se puede construir una biblioteca de ADN y/o de ADNc usando ADN cromosómico o ARN mensajero del organismo que produce el polipéptido. Si se conoce la secuencia de aminoácidos del polipéptido, se pueden sintetizar sondas de oligonucleótidos marcadas y usarlas para identificar los clones que codifican el polipéptido a partir de la biblioteca genómica preparada a partir del organismo. De forma alternativa, se podría usar una sonda de oligonucleótidos marcada que contiene secuencias homólogas a otro gen de polipéptido conocido para identificar los clones que codifican el polipéptido. En el último caso, se usan condiciones de hibridación y lavado de baja severidad.

25 Alternativamente, se podrían identificar clones que codifican polipéptidos insertando fragmentos de ADN genómico en un vector de expresión, tal como un plásmido, transformando bacterias negativas en enzimas con la biblioteca de ADN genómica resultante, y a continuación llevando a placa las bacterias transformadas sobre agar que contiene una enzima inhibida por el polipéptido, permitiendo con ello que los clones expresen el polipéptido que va a ser identificado.

30 En otra alternativa adicional, la secuencia de nucleótidos que codifica el polipéptido puede prepararse sintéticamente mediante métodos estándar, por ejemplo el método de fosforoamidita descrito por Beucage S.L. et al (1981) Tetrahedron Letters 22, pág. 1859-1869, o el método descrito por Matthes et al (1984) EMBO J. 3, pág. 801-805. En el método de fosforoamidita, se sintetizan oligonucleótidos, por ejemplo en un sintetizador de ADN automático, se purifican, se maduran, se ligan y se clonan en vectores apropiados.

35 La secuencia de nucleótidos puede ser origen mixto genómico y sintético, de origen mixto sintético y de ADNc, o de origen mixto genómico y de ADNc, preparada ligando fragmentos de origen sintético, genómico o de ADNc (según sea apropiado) de acuerdo con las técnicas estándar. Cada fragmento ligado corresponde a varias partes de la secuencia de nucleótidos completa. La secuencia de ADN también puede prepararse mediante reacción en cadena de la polimerasa (PCR) usando cebadores específicos, por ejemplo como se describe en el documento US 4.683.202 o en Saiki R K et al (Science (1988) 239, págs. 487-4991).

**Secuencias de nucleótidos**

45 La presente invención también abarca secuencias de nucleótidos que codifican polipéptidos que tienen las propiedades específicas definidas en la presente memoria. El término "secuencia de nucleótidos" tal como se usa en la presente memoria se refiere a una secuencia de oligonucleótido o a una secuencia de polinucleótido, y a sus derivados, fragmentos, variantes y homólogos (tales como sus porciones). La secuencia de nucleótidos puede ser de origen genómico o sintético o recombinante, y puede ser bicatenaria o monocatenaria, que representa la cadena directa o la inversa.

50 La expresión "secuencia de nucleótidos" con respecto a la presente invención incluye el ADN genómico, ADNc, ADN sintético y ARN. Preferiblemente significa ADN, más preferiblemente ADNc para la secuencia codificante.

En una realización preferida, la secuencia de nucleótidos *per se* que codifica un polipéptido que tiene las propiedades específicas definidas en la presente memoria no cubre la secuencia de nucleótidos nativa en su medio natural cuando está unida a su(s) secuencia(s) asociada(s) de forma natural que también está(n) en su entorno natural. A fin de facilitar la referencia, denominaremos a esta realización preferida "secuencia de nucleótidos no

nativa". A este respecto, la expresión "secuencia de nucleótidos nativa" significa una secuencia de nucleótidos completa que está en su medio nativo y cuando está unida de forma operable a un promotor completo con el que está asociada de manera natural, cuyo promotor también está en su medio nativo. Así, el polipéptido de la presente invención se puede expresar mediante una secuencia de nucleótidos en su organismo nativo pero en el que la secuencia de nucleótidos no está bajo el control del promotor con el que está asociada de manera natural dentro de ese organismo.

Preferiblemente el polipéptido no es un polipéptido nativo. A este respecto, el término "polipéptido nativo" significa un polipéptido completo que se encuentra en su entorno nativo y cuando ha sido expresado por su secuencia de nucleótidos nativa.

En general, la secuencia de nucleótidos que codifica polipéptidos que tienen las propiedades específicas definidas en la presente memoria se prepara mediante el uso de técnicas de ADN recombinante (es decir, ADN recombinante). Sin embargo, en una realización alternativa de la invención, la secuencia de nucleótidos se podría sintetizar, por completo o en parte, mediante el uso de métodos químicos muy conocidos en la técnica (véase Caruthers MH et al (1980) Nuc Acids Res Symp Ser 215-23 y Horn T et al (1980) Nuc Acids Res Symp Ser 225-232).

## 15 Evolución molecular

Una vez que se ha aislado la secuencia de nucleótidos que codifica una enzima, o se ha identificado una secuencia de nucleótidos putativa que codifica una enzima, puede ser deseable modificar la secuencia de nucleótidos seleccionada, por ejemplo puede ser deseable mutar la secuencia para preparar una enzima de acuerdo con la presente invención.

Se pueden introducir mutaciones usando oligonucleótidos sintéticos. Estos oligonucleótidos contienen secuencias de nucleótidos que flanquean los sitios de mutación deseados.

Un método adecuado se describe en Morinaga et al (Biotechnology (1984) 2, p. 646-649). Otro método para introducir mutaciones en secuencias de nucleótidos que codifican enzima se describe en Nelson y Long (Analytical Biochemistry (1989), 180, p. 147-151).

En lugar de mutagénesis dirigida, tal como se ha descrito previamente, se pueden introducir mutaciones aleatoriamente, por ejemplo usando un equipo comercial como el equipo de mutagénesis GeneMorph PCR de Stratagene, o el equipo de mutagénesis aleatoria Diversify PCR de Clontech. El documento EP 0 583 265 se refiere a métodos de optimización de mutagénesis basada en PCR, que también puede combinarse con el uso de análogos de ADN mutagénicos tales como los descritos en el documento EP 0 866 796. Las tecnologías de PCR propensas a error son adecuadas para la producción de variantes de lípido acil transferasas con características preferidas. El documento WO0206457 se refiere a la evolución molecular de lipasas.

Un tercer método para obtener nuevas secuencias es fragmentar secuencias de nucleótidos no idénticos, usando cualquier número de enzimas de restricción o una enzima tal como DNasa I, y reconstruyendo secuencias de nucleótidos completas que codifican para proteínas funcionales. Alternativamente, se puede usar una o múltiples secuencias de nucleótidos no idénticos e introducir mutaciones durante la reconstrucción de la secuencia de nucleótidos completa. El barajado de ADN y las tecnologías de barajado de familia son adecuadas para la producción de variantes de lípido acil transferasas con características preferidas. Los métodos adecuados para llevar a cabo el "barajado" se pueden encontrar en los documentos EP0 752 008, EP1 138 763, EP1 103 606. El barajado también puede combinarse con otras formas de mutagénesis de ADN, tal como se describe en los documentos US 6.180.406 y WO 01/34835.

Por tanto, es posible producir numerosas mutaciones, dirigidas o aleatorias, en una secuencia de nucleótidos, tanto *in vivo* como *in vitro*, y posteriormente cribar en busca de una funcionalidad mejorada del polipéptido codificado mediante diversos medios. Usando métodos de recombinación mediados *in silico* y *in vivo* (véanse los documentos WO 00/58517, US 6.344.328, US 6.361.974), por ejemplo, se puede llevar a cabo una evolución molecular en la que la variante producida mantiene una homología muy baja con respecto a enzimas o proteínas conocidas. Las variantes obtenidas de este modo pueden presentar una analogía estructural significativa con respecto a enzimas de transferasa conocidas, pero tienen una homología de secuencia de aminoácidos muy baja.

Como ejemplo no limitante, adicionalmente, las mutaciones o variantes naturales de una secuencia de polinucleótido pueden recombinarse con mutaciones naturales o de otro tipo o con variantes naturales para producir nuevas variantes. Dichas nuevas variantes también pueden ser escrutadas para determinar mejoras de la funcionalidad en el polipéptido codificado.

La aplicación de los métodos de evolución molecular mencionados anteriormente y otros similares permite la identificación y la selección de variantes de las enzimas de la presente invención que tienen características preferidas sin ningún conocimiento previo de la estructura o la función de la proteína, y permite la producción de mutaciones o variantes no predecibles pero beneficiosas. Existen numerosos ejemplos de la aplicación de evolución molecular en la técnica para la optimización o la alteración de la actividad enzimática, ejemplos que incluyen, aunque sin limitación, uno o más de los siguientes: una expresión y/o actividad optimizadas en una célula



hospedadora o in vitro, una actividad enzimática incrementada, una especificidad de sustrato y/o producto alterada, una estabilidad enzimática o estructural aumentada o disminuida, una actividad/especificidad enzimática alterada en condiciones ambientales preferidas, por ejemplo, temperatura, pH, sustrato.

5 Como resultará evidente para el especialista en la técnica, usando herramientas de evolución molecular se puede alterar una enzima para mejorar la funcionalidad de la enzima.

10 De manera adecuada, la secuencia de nucleótidos que codifica una lípido aciltransferasa usada en la descripción puede codificar una lípido aciltransferasa variante, es decir, la lípido aciltransferasa puede contener al menos una sustitución, deleción o adición de aminoácidos, en comparación con una enzima originaria. Las enzimas variantes retienen al menos el 1%, 2%, 3%, 5%, 10%, 15%, 20%, 30%, 40%, 50 %, 60%, 70%, 80%, 90%, 95%, 97%, 99% de homología con respecto a la enzima originaria. Las enzimas originarias adecuadas pueden incluir cualquier enzima con actividad de esterasa o lipasa. Preferiblemente, la enzima originaria se alinea con la secuencia consenso pfam00657.

15 En una realización preferida, una enzima lípido aciltransferasa variante retiene o incorpora al menos uno o más de los residuos de aminoácido de la secuencia consenso pfam00657 encontrados en los bloques GDSx, GANDY y HPT.

Las enzimas, tal como las lipasas con poca o nula actividad de lípido aciltransferasa en un entorno acuoso se pueden mutar usando herramientas de evolución molecular para introducir o potenciar la actividad de transferasa, produciendo de este modo una enzima lípido aciltransferasa con una actividad significativa de transferasa adecuada para el uso en las composiciones y métodos de la presente invención.

20 De manera adecuada, la secuencia de nucleótidos que codifica una lípido aciltransferasa para el uso en cualquiera de los vectores, células hospedadoras, métodos y/o usos de la presente descripción puede codificar una lípido aciltransferasa que puede ser una variante con actividad enzimática incrementada sobre los lípidos polares, preferiblemente fosfolípidos y/o glicolípidos en comparación con la enzima originaria. Preferiblemente, dichas variantes también presentan baja o nula actividad con lípidos liso polares. La actividad mejorada con lípidos polares, 25 fosfolípidos y/o glicolípidos, puede ser el resultado de hidrólisis y/o actividad de transferasa o de una combinación de ambos.

Las lípido aciltransferasas variantes pueden presentar una actividad disminuida con triglicéridos, y/o monoglicéridos y/o diglicéridos, en comparación con la enzima originaria.

30 De forma adecuada, la enzima variante puede tener actividad nula con triglicéridos y/o monoglicéridos y/o diglicéridos.

De forma alternativa, la enzima variante puede presentar una actividad incrementada con triglicéridos, y/o también puede presentar una actividad incrementada con uno o más de los siguientes, lípidos polares, fosfolípidos, lecitina, fosfatidilcolina, glicolípidos, monoglicérido de digalactosilo, monoglicérido de monogalactosilo.

35 Se conocen variantes de lípido aciltransferasas, y una o más de dichas variantes pueden ser adecuadas para el uso en los métodos y usos según la presente descripción y/o en las composiciones enzimáticas según la presente descripción. A modo de ejemplo solamente, las variantes de lípido aciltransferasas que se describen en las referencias siguientes se pueden usar de acuerdo con la presente descripción: Hilton & Buckley J Biol. Chem. 1991 enero 15: 266 (2): 997-1000; Robertson et al J. Biol. Chem. 21 de enero de 1994; 269(3): 2146-50; Brumlik et al J. Bacteriol, abril de 1996; 178 (7): 2060-4; Peelman et al Protein Sci. 1998 marzo; 7(3): 587-99.

#### 40 **Secuencias de aminoácidos**

La presente invención también abarca las secuencias de aminoácidos codificadas por una secuencia de nucleótidos que codifica una lípido aciltransferasa para el uso en cualquiera de los vectores, células hospedadoras, métodos y/o usos de la presente invención.

45 Tal como se usa en la presente memoria, la expresión "secuencia de aminoácidos" es sinónima del término "polipéptido" y/o del término "proteína". En algunos casos, la expresión "secuencia de aminoácidos" es sinónima del término "péptido".

La secuencia de aminoácidos puede prepararse/aislarse a partir de una fuente adecuada, o puede fabricarse sintéticamente o puede prepararse mediante el uso de técnicas de ADN recombinante.

50 De forma adecuada, las secuencias de aminoácidos pueden obtenerse a partir de los polipéptidos aislados mostrados en la presente memoria mediante técnicas estándar.

Un método adecuado para determinar secuencias de aminoácidos a partir de polipéptidos aislados es el siguiente:

El polipéptido purificado se puede liofilizar y se pueden disolver 100 µg del material liofilizado en 50 µL de una mezcla de urea 8 M y bicarbonato amónico 0,4 M, pH 8,4. La proteína disuelta puede ser desnaturalizada y reducida

durante 15 minutos a 50°C después de ventilación con nitrógeno y adición de 5 µL de ditiotreitól 45 mM. Tras enfriar hasta temperatura ambiente, se pueden añadir 5 µL de yodoacetamida 100 mM para derivatizar los residuos de cisteína durante 15 minutos a temperatura ambiente a oscuras y en atmósfera de nitrógeno.

5 Se pueden añadir 135 µL de agua y 5 µg de endoproteinasa Lys-C en 5 µL de agua a la anterior mezcla de reacción y la digestión se puede llevar a cabo a 37°C en atmósfera de nitrógeno durante 24 horas.

10 Los péptidos resultantes pueden separarse mediante HPLC de fase inversa en una columna VYDAC C18 (0,46x15 cm; 10 µm; The Separation Group, California, EE.UU.) usando un disolvente A: 0,1% de TFA en agua y un disolvente B: 0,1% de TFA en acetonitrilo. Los péptidos seleccionados se pueden volver a someter a cromatografía en una columna Develosil C18 usando el mismo sistema de disolventes, antes de la secuenciación N-terminal. La secuenciación se puede realizar usando un secuenciador Applied Biosystems 476A empleando ciclos rápidos de líquido pulsado siguiendo las instrucciones del fabricante (Applied Biosystems, California, EE.UU.).

### Identidad de secuencia u homología de secuencia

15 En la presente memoria, el término "homólogo" significa una entidad que tiene una determinada homología con las secuencias de aminoácidos objetivo de la invención y con las secuencias de nucleótidos objetivo de la invención. En la presente memoria, el término "homología" puede igualarse a "identidad".

La secuencia de aminoácidos homóloga y/o la secuencia de nucleótidos homóloga deberían proporcionar y/o codificar un polipéptido que retiene la actividad funcional y/o potencia la actividad de la enzima.

20 En el presente contexto, se considera que una secuencia homóloga incluye una secuencia de aminoácidos que puede ser idéntica a la secuencia objetivo de la invención en al menos un 75, un 85 o un 90%, preferiblemente idéntica a la secuencia objetivo en al menos un 95 o un 98%. Habitualmente, los homólogos comprenderán los mismos sitios activos, etc. que la secuencia de aminoácidos objetivo. Aunque la homología también puede considerarse en términos de similitud (es decir, residuos de aminoácido que tienen propiedades/funciones químicas similares), en el contexto de la presente invención se prefiere expresar la homología en términos de identidad de secuencia.

25 En el presente contexto, se considera que una secuencia homóloga incluye una secuencia de nucleótidos que puede ser idéntica en al menos un 75, 85 ó 90%, preferiblemente idéntica en al menos un 95 o un 98%, a una secuencia de nucleótidos que codifica un polipéptido de la presente invención (la secuencia objetivo). Habitualmente los homólogos comprenderán las mismas secuencias que codifican para los sitios activos, etc., que la secuencia objetivo. Aunque la homología también puede considerarse en términos de similitud (es decir, residuos de aminoácido que tienen propiedades/funciones químicas similares), en el contexto de la presente invención se prefiere expresar la homología en términos de identidad de secuencia.

30 Las comparaciones de homología se pueden llevar a cabo visualmente, o de forma más habitual con la ayuda de programas de comparación de secuencias fácilmente disponibles. Estos programas de ordenador disponibles comercialmente pueden calcular el % de homología entre dos o más secuencias.

35 El % de homología se puede calcular para secuencias contiguas, es decir una secuencia se alinea con la otra secuencia y cada aminoácido de una secuencia es comparado directamente con el correspondiente aminoácido de la otra secuencia, residuo a residuo. Esto se denomina alineación "sin huecos". Habitualmente, dichas alineaciones sin huecos se llevan a cabo sólo con un número relativamente corto de residuos.

40 Aunque este es un método muy simple y consistente, no tiene en consideración que, por ejemplo, en un par de secuencias por lo demás idénticas, una inserción o eliminación provocará que los siguientes residuos de aminoácido queden fuera de la alineación, dando potencialmente como resultado una gran reducción del % de homología cuando se realiza una alineación global. Consecuentemente, la mayoría de los métodos de comparación están diseñados para producir alineaciones óptimas que tienen en cuenta las posibles inserciones y eliminaciones sin penalizar de forma indebida la puntuación global de homología. Esto se logra mediante la inserción de "huecos" en la alineación de secuencia para intentar maximizar la homología local. Sin embargo, estos métodos más complejos asignan "penalizaciones por hueco" a cada hueco que se produce en la alineación de tal modo que, para el mismo número de aminoácidos idénticos, una alineación de secuencia con el mínimo número posible de huecos - lo que refleja la relación entre las dos secuencias comparadas - conseguirá un puntuación mayor que uno con muchos huecos. Habitualmente se usan "costes de hueco afines" que aplican un coste relativamente alto para la existencia de un hueco, y una penalización menor para cada residuo subsiguiente en el hueco. Este el sistema de puntuación de huecos usado más habitualmente. Por supuesto, una elevada penalización por hueco producirá alineaciones optimizadas con menores huecos. La mayoría de los programas de alineación permiten modificar las penalizaciones por hueco. Sin embargo, es preferible usar los valores por defecto cuando se usa dicho software para comparaciones de secuencia.

55 Por lo tanto, el cálculo del % máximo de homología en primer lugar requiere la producción de una alineación óptima, considerando las penalizaciones por hueco. Un programa informático adecuado para llevar a cabo tal alineación es el Vector NTI (Invitrogen Corp.). Los ejemplos de otros programas informáticos que pueden llevar a cabo

comparaciones de secuencias incluyen, pero sin limitación, el paquete BLAST (véase Ausubel et al 1999 Short Protocols in Molecular Biology, 4ª Ed - Capítulo 18), y FASTA (Altschul et al 1990 J. Mol. Biol. 403-410). Tanto el BLAST como el FASTA se encuentran disponibles para búsquedas con y sin conexión a Internet (véase Ausubel et al 1999, páginas 7-58 a 7-60). Sin embargo, para ciertas aplicaciones, se prefiere usar el programa Vector NTI. También está disponible una nueva herramienta, denominada BLAST 2 Sequences, para comparar secuencias de proteína y nucleótidos (véase FEMS Microbiol Lett 1999 174 (2): 247-50; FEMS Microbiol Lett 1999 177 (1): 187-8 y tatiana@nobi.nlm.nih.gov).

Aunque el % de homología final se puede medir en términos de identidad, el propio proceso de alineación habitualmente no se basa en una comparación de pares todo-o-nada. En lugar de eso, generalmente se usa una matriz de puntuación de escala de similitud que asigna puntuaciones a cada comparación por pares en base a la similitud química o la distancia de evolución. Un ejemplo de dicha matriz usado habitualmente es la matriz BLOSUM62 - la matriz por defecto para el tipo de programas BLAST. Los programas Vector NTI usan en general los valores públicos por defecto o una tabla de comparación de símbolos a medida si se suministra (véase el manual del usuario para detalles adicionales). Para ciertas aplicaciones, se prefiere usar los valores por defecto para el paquete Vector NTI.

Alternativamente, los porcentajes de homología se pueden calcular usando la característica de alineación múltiple de Vector NTI (Invitrogen Corp.), basado en un algoritmo, análogo a CLUSTAL (Higgins DG & Sharp PM (1988), Gene 73(1), 237-244).

Una vez que el software ha producido una alineación óptima, es posible calcular el % de homología, preferiblemente el % de identidad de secuencia. El software habitualmente realiza esto como parte de la comparación de secuencias y genera un resultado numérico.

En caso de que usen penalizaciones por hueco al determinar la identidad de secuencia, se usarán preferiblemente los siguientes parámetros para la alineación por pares:

PARA BLAST	
APERTURA DE HUECO	0
EXTENSIÓN DE HUECO	0

PARA CLUSTAL	ADN	PROTEÍNA	
TAMAÑO DE PALABRA	2	1	K triple
PENALIZACIÓN POR HUECO	15	10	
EXTENSIÓN DE HUECO	6,66	0,1	

En una realización, preferiblemente la identidad de secuencia para las secuencias de nucleótidos se determina mediante el uso de CLUSTAL con los ajustes de penalización por hueco y extensión de hueco como se definieron anteriormente.

De forma adecuada, el grado de identidad con respecto a una secuencia de nucleótidos se determina sobre al menos 20 nucleótidos contiguos, preferiblemente sobre al menos 30 nucleótidos contiguos, preferiblemente sobre al menos 40 nucleótidos contiguos, preferiblemente sobre al menos 50 nucleótidos contiguos, preferiblemente sobre al menos 60 nucleótidos contiguos, preferiblemente sobre al menos 100 nucleótidos contiguos.

De forma adecuada, el grado de identidad con respecto a la secuencia de nucleótidos puede determinarse sobre la secuencia completa.

En una realización, el grado de identidad de secuencia de aminoácidos de acuerdo con la presente invención se puede determinar de manera adecuada por medio de programas informáticos conocidos en la técnica, tales como Vector NTI 10 (Invitrogen Corp.). Para la alineación por pares la matriz usada preferiblemente es BLOSUM62 con una penalización por apertura de hueco de 10,0 y una penalización por extensión de hueco de 0,1.

De forma adecuada, el grado de identidad con respecto a una secuencia de aminoácidos se determina sobre al menos 20 aminoácidos contiguos, preferiblemente sobre al menos 30 aminoácidos contiguos, preferiblemente sobre al menos 40 aminoácidos contiguos, preferiblemente sobre al menos 50 aminoácidos contiguos, preferiblemente sobre al menos 60 aminoácidos contiguos.

De forma adecuada, el grado de identidad con respecto a la secuencia de aminoácidos puede determinarse sobre la secuencia completa.

Las secuencias también pueden tener eliminaciones, inserciones o sustituciones de residuos de aminoácidos que producen un cambio silencioso y dan como resultado una sustancia funcionalmente equivalente. Se pueden realizar

sustituciones de aminoácidos deliberadas en base a la similitud en polaridad, carga, solubilidad, hidrofobicidad, hidrofiliidad y/o naturaleza anfipática de los residuos, siempre que se mantenga la actividad de unión secundaria de la sustancia. Por ejemplo, los aminoácidos cargados negativamente incluyen ácido aspártico y ácido glutámico; los aminoácidos cargados positivamente incluyen lisina y arginina; y los aminoácidos con grupos de cabeza polares no cargados que tienen valores de hidrofiliidad similares incluyen leucina, isoleucina, valina, glicina, alanina, asparagina, glutamina, serina, treonina, fenilalanina y tirosina.

Se pueden realizar sustituciones conservativas, por ejemplo según la Tabla mostrada continuación. Los aminoácidos del mismo bloque de la segunda columna y preferiblemente de la misma línea de la tercera columna pueden ser sustituidos entre sí:

ALIFÁTICO	No polar	G A P
		I L V
	Polar - no cargado	C S T M
		N Q
	Polar - cargado	D E
		K R
AROMÁTICO		H F W Y

La presente descripción también abarca la sustitución homóloga (sustitución y reemplazamiento se usan ambos en la presente memoria para indicar el intercambio de un residuo de aminoácido existente por un residuo alternativo) que puede producirse, es decir sustitución de tipo-por-tipo tal como básico por básico, ácido por ácido, polar por polar, etc. También se puede producir la sustitución no homóloga, es decir a partir de una clase de residuo por otra, o alternativamente que implica la inclusión de aminoácidos no naturales tales como ornitina (denominada a partir de aquí Z), ornitina de ácido diaminobutírico (denominada a partir de aquí B), ornitina de norleucina (denominada a partir de aquí O), pirilalanina, tienilalanina, naftilalalina y fenilglicina.

También se pueden realizar reemplazamientos con aminoácidos no naturales.

Las secuencias de aminoácidos variantes pueden incluir grupos espaciadores que pueden insertarse entre dos residuos de aminoácido cualesquiera de la secuencia, que incluyen grupos alquilo tales como grupos metilo, etilo o propilo, además de espaciadores de aminoácidos tales como residuos de glicina o β-alanina. Otra forma adicional de variación, que implica la presencia de uno o más residuos de aminoácido en forma peptoide, será bien entendida por los especialistas en la técnica. Para evitar dudas, "la forma peptoide" se usa para referirse a residuos de aminoácido variantes en los que el grupo sustituyente de carbono α está sobre el átomo de nitrógeno del residuo, en lugar de sobre el carbono α. Los procesos para preparar péptidos en la forma peptoide se conocen en la técnica, por ejemplo Simon RJ et al., PNAS (1992) 89(20), 9367-9371 y Horwell DC, Trends Biotechnol. (1995) 13(4), 132-134.

Las secuencias de nucleótidos para el uso en la presente invención o que codifican un polipéptido que tiene las propiedades específicas definidas en la presente memoria pueden incluir dentro de ellas nucleótidos sintéticos o modificados. En la técnica se conoce una serie de tipos diferentes de modificación de oligonucleótidos. Éstas incluyen cadenas principales de metilfosfonato y fosforotioato y/o la adición de cadenas de acridina o polilisina en los extremos 3' y/o 5' de la molécula. Para los propósitos de la presente invención, debe entenderse que las secuencias de nucleótidos descritas en la presente memoria pueden modificarse mediante cualquier método disponible en la técnica. Dichas modificaciones pueden llevarse a cabo con el objetivo de potenciar la actividad *in vivo* o la duración de las secuencias de nucleótidos.

La presente descripción también abarca el uso de secuencias de nucleótidos que sean complementarias a las secuencias discutidas en la presente memoria, o cualquier derivado o fragmento de las mismas. Si la secuencia es complementaria a un fragmento de la misma, entonces dicha secuencia puede usarse como sonda para identificar secuencias codificantes similares en otros organismos, etc.

Los polinucleótidos que no son un 100% homólogos respecto de las secuencias de la presente invención, pero que entran dentro del alcance de la invención, se pueden obtener de varias maneras. Otras variantes de las secuencias descritas en la presente memoria se pueden obtener, por ejemplo, empleando sondas en bibliotecas de ADN preparadas a partir de una serie de individuos, por ejemplo de individuos procedentes de diferentes poblaciones. Adicionalmente, se pueden obtener otros homólogos víricos/bacterianos, o celulares, particularmente homólogos celulares encontrados en células de mamífero (por ejemplo, células de rata, ratón, bovinas y de primate), y dichos homólogos y sus fragmentos en general serán capaces de hibridarse de forma selectiva con las secuencias mostradas en el listado de secuencias incluido en la presente memoria. Dichas secuencias pueden obtenerse mediante el empleo de sondas en bibliotecas de ADNc preparadas o de bibliotecas de ADN genómicas de otras especies animales, y sondeando dichas bibliotecas con sondas que comprenden la totalidad o una parte de una cualquiera de las secuencias del listado de secuencias anexo, en condiciones de media a alta severidad. Se aplican consideraciones similares a la obtención de homólogos de especie y variantes alélicas de las secuencias de polipéptidos o nucleótidos de la invención.

Las variantes y los homólogos de cepa/especie también pueden obtenerse usando PCR degenerada, que usará cebadores diseñados para seleccionar secuencias dentro de las variantes y homólogos que codifican secuencias de aminoácidos conservadas dentro de las secuencias de la presente invención. Se pueden predecir las secuencias conservadas, por ejemplo, alineando las secuencias de aminoácidos de varios homólogos/variantes. Las alineaciones de secuencia se pueden llevar a cabo usando un software de ordenador conocido en la técnica. Por ejemplo, se usa ampliamente el programa GCG Wisconsin PileUp.

Los cebadores usados en la PCR degenerativa contendrán una o más posiciones degeneradas y se usarán en condiciones de severidad menores a los usados para clonar secuencias con cebadores de secuencia individuales contra secuencias conocidas.

Alternativamente, dichos polipéptidos pueden obtenerse mediante mutagénesis dirigida de secuencias caracterizadas. Esto puede ser útil, por ejemplo, cuando se requieren cambios de secuencia de codón silenciosos para optimizar las preferencias de codón para una célula hospedadora concreta en la que se van a expresar las secuencias de polinucleótidos. Pueden ser deseables otros cambios de secuencia con el objetivo de introducir sitios de reconocimiento de polipéptido de restricción, o para alterar la propiedad o la función de los polipéptidos codificados por los polinucleótidos.

Los polinucleótidos (secuencias de nucleótidos) de la invención pueden usarse para producir un cebador, p.ej. un cebador de PCR, un cebador para una reacción de amplificación alternativa, una sonda, p.ej. marcada con una etiqueta reveladora por medios convencionales usando etiquetas radiactivas o no radiactivas, o los polinucleótidos pueden clonarse en vectores. Dichos cebadores, sondas y otros fragmentos tendrán una longitud de al menos 15 nucleótidos, preferiblemente de al menos 20 nucleótidos, por ejemplo al menos 25, 30 ó 40 nucleótidos, y también quedan abarcados por el término "polinucleótidos" de la descripción tal como se usa en la presente memoria.

Los polinucleótidos tales como los polinucleótidos de ADN y las sondas según la invención se pueden producir de manera recombinante, sintética, o mediante cualquier medio disponible para los expertos en la técnica. También pueden clonarse mediante técnicas estándar.

En general, los cebadores se producirán por medios sintéticos, lo que implica la fabricación por etapas de la secuencia de ácido nucleico deseada, nucleótido a nucleótido. Las técnicas para llevar a cabo esto empleando técnicas automatizadas se encuentran fácilmente disponibles en la técnica.

Los polinucleótidos de mayor tamaño generalmente se producirán usando medios recombinantes, por ejemplo usando una técnica de clonación PCR (reacción en cadena de polimerasa). Esto implicará preparar un par de cebadores (por ejemplo, de aproximadamente 15 a 30 nucleótidos) que escolten una región de la secuencia dirigida a lípido que se desea clonar, poniendo los cebadores en contacto con ARNm o ADNc obtenido de una célula animal o humana, llevando a cabo la reacción en cadena de polimerasa en condiciones que conlleven la amplificación de la región deseada, aislando el fragmento aislado (por ejemplo, purificando la mezcla de reacción sobre gel de agarosa) y recuperando el ADN amplificado. Los cebadores pueden diseñarse para contener sitios de reconocimiento de enzima de restricción, de tal modo que el ADN amplificado puede ser clonado en un vector de clonación adecuado.

### Hibridación

La presente descripción también contempla secuencias que son complementarias a las secuencias de la presente invención o secuencias que son capaces de hibridarse con las secuencias de la presente invención o con secuencias que son complementarias a las mismas.

El término "hibridación" tal como se usa en la presente memoria incluirá "el proceso mediante el cual una cadena de ácido nucleico se une a una cadena complementaria a través de emparejamiento de bases", así como el proceso de amplificación llevado a cabo en las tecnologías de reacción en cadena de polimerasa (PCR).

La presente descripción también abarca el uso de secuencias de nucleótidos que son capaces de hibridarse con las secuencias que son complementarias a las secuencias objetivo discutidas en la presente memoria, o cualquier derivado o fragmento de las mismas.

La presente descripción también abarca secuencias que son complementarias a las secuencias que son capaces de hibridarse con las secuencias de nucleótidos discutidas en la presente memoria.

Las condiciones de hibridación se basan en la temperatura de fusión ( $T_f$ ) del complejo de unión al nucleótido, como se muestra en Berger y Kimmel (1987, Guide to Molecular Cloning Techniques, Methods in Enzymology, Vol. 152, Academic Press, San Diego CA), y confieren una "severidad" definida como se explica a continuación.

La máxima severidad se produce habitualmente a aproximadamente  $T_f - 5^\circ\text{C}$  ( $5^\circ\text{C}$  por debajo de la  $T_f$  de la sonda); elevada severidad a aproximadamente entre  $5^\circ\text{C}$  y  $10^\circ\text{C}$  por debajo de la  $T_f$ ; severidad intermedia a aproximadamente entre  $10^\circ\text{C}$  y  $20^\circ\text{C}$  por debajo de la  $T_f$ ; y baja severidad a aproximadamente entre  $20^\circ\text{C}$  y  $25^\circ\text{C}$  por debajo de la  $T_f$ ; Como entenderán los especialistas en la técnica, se puede usar una hibridación de severidad

máxima para identificar o detectar secuencias de nucleótidos idénticas, y se puede usar una hibridación de severidad intermedia (o baja) para identificar o detectar secuencias de polinucleótidos similares o relacionadas.

5 Preferiblemente, la presente descripción abarca secuencias que son complementarias a secuencias que son capaces de hibridarse en condiciones de alta severidad o en condiciones de severidad intermedia con secuencias de nucleótidos que codifican polipéptidos que tienen las propiedades específicas definidas en la presente memoria.

Más preferiblemente, la presente descripción abarca secuencias que son complementarias a secuencias que son capaces de hibridarse en condiciones de alta severidad (p.ej., 65°C y 0,1xSSC {1xSSC = NaCl 0,15 M, citrato de Na 0,015 M, pH 7,0}) a secuencias de nucleótidos que codifican polipéptidos que tienen las propiedades específicas definidas en la presente memoria.

10 La presente descripción también se refiere a secuencias de nucleótidos que pueden hibridarse a las secuencias de nucleótidos discutidas en la presente memoria (que incluyen secuencias complementarias a las discutidas en la presente memoria).

15 La presente descripción también se refiere a secuencias de nucleótidos que son complementarias a secuencias que pueden hibridarse a las secuencias de nucleótidos discutidas en la presente memoria (que incluyen secuencias complementarias a las discutidas en la presente memoria).

También se incluyen dentro del alcance de la presente descripción las secuencias de polinucleótido que son capaces de hibridarse con las secuencias de nucleótidos discutidas en la presente memoria en condiciones de severidad entre intermedia y máxima.

20 En un aspecto preferido, la presente descripción cubre secuencias de nucleótidos que pueden hibridarse con las secuencias de nucleótidos discutidas en la presente memoria, o su complemento, en condiciones de severidad (por ejemplo, 50°C y 0,2 x SSC).

En un aspecto más preferido, la presente descripción cubre secuencias de nucleótidos que pueden hibridarse con las secuencias de nucleótidos discutidas en la presente memoria, o su complemento, en condiciones de alta severidad (por ejemplo, 65°C y 0,1 x SSC).

## 25 **Expresión de polipéptidos**

Una secuencia de nucleótidos para el uso en la presente invención o para codificar un polipéptido que tiene las propiedades específicas definidas en la presente memoria se puede incorporar en un vector replicable recombinante. El vector puede usarse para replicar y expresar la secuencia de nucleótidos, en forma de polipéptido, en y/o a partir de una célula hospedadora compatible. La expresión se puede controlar usando secuencias de control que incluyen promotores/potenciadores y otras señales de regulación de la expresión. Se pueden usar promotores procarionóticos y promotores funcionales en células eucarióticas. Se pueden usar promotores específicos de tejido o específicos de estímulo. También se pueden usar promotores quiméricos que comprenden elementos de secuencia de dos o más promotores diferentes descritos anteriormente.

35 El polipéptido producido por una célula recombinante hospedadora mediante expresión de la secuencia de nucleótidos puede ser secretado o puede estar contenido intracelularmente dependiendo de la secuencia y/o del vector usado. Pueden diseñarse secuencias codificantes con secuencias señal que dirijan la secreción de las secuencias codificantes de la sustancia a través de una membrana concreta de célula procariótica o eucariótica.

## **Construcciones**

40 El término "construcción" - que es sinónimo a términos tales como "conjugado", "casete" e "híbrido" - incluye una secuencia de nucleótidos que codifica un polipéptido que tiene las propiedades específicas definidas en la presente memoria para el uso según la presente invención unida directamente o indirectamente a un promotor. Un ejemplo de una unión indirecta es la provisión de un grupo espaciador adecuado tal como una secuencia intrónica, tal como el intrón Sh1 o el intrón ADH, entre el promotor y la secuencia de nucleótidos de la presente invención. Lo mismo es válido para el término "fusionado" con respecto a la presente invención, que incluye la unión directa o indirecta. En algunos casos, los términos no cubren la combinación natural de la secuencia de nucleótidos que codifican para la proteína asociada de forma ordinaria al promotor génico natural y cuando ambos están en su entorno natural.

La construcción incluso pueden contener o expresar un marcador que permite la selección de la construcción genética.

50 Para algunas aplicaciones, la construcción preferiblemente comprende al menos una secuencia de nucleótidos de la presente invención o una secuencia de nucleótidos que codifica un polipéptido que tiene las propiedades específicas definidas en la presente memoria ligada de forma operable al promotor.

**Organismo**

El término "organismo" en relación a la presente invención incluye cualquier organismo que pueda comprender una secuencia de nucleótidos según la presente invención o una secuencia de nucleótidos que codifique un polipéptido que tenga las propiedades específicas definidas en la presente memoria y/o productos obtenidos a partir del mismo.

5 La expresión "organismo transgénico" en relación con la presente invención incluye cualquier organismo que comprende una secuencia de nucleótidos que codifica un polipéptido que tiene las propiedades específicas definidas en la presente memoria y/o los productos obtenidos a partir del mismo, y/o en el que un promotor puede permitir la expresión de la secuencia de nucleótidos que codifica un polipéptido que tiene las propiedades específicas definidas en la presente memoria en el organismo. Preferiblemente, la secuencia de nucleótidos se incorpora al genoma del organismo.

El término "organismo transgénico" no cubre secuencias codificantes de nucleótidos nativos en su entorno natural cuando están bajo el control de su promotor nativo, que también está en su entorno natural.

15 Por lo tanto, el organismo transgénico de la presente descripción incluye un organismo que comprende una cualquiera, o combinaciones, de las secuencias de nucleótidos que codifican para un polipéptido que tiene las propiedades específicas definidas en la presente memoria, construcciones como las definidas en la presente memoria, vectores como los definidos en la presente memoria, plásmidos como los definidos en la presente memoria, células como las definidas en la presente memoria, o los productos de los mismos. Por ejemplo, el organismo transgénico puede comprender también una secuencia de nucleótidos que codifica un polipéptido que tiene las propiedades específicas definidas en la presente memoria bajo control de un promotor que no está asociado con una secuencia que codifica una lípido aciltransferasa en la naturaleza.

**Transformación de células/organismos hospedadores**

25 Los conocimientos sobre la transformación de hospedadores procarióticos están bien documentados en la técnica, por ejemplo véase Sambrook et al (Molecular Cloning: A Laboratory Manual, 2ª edición, 1989, Cold Spring Harbor Laboratory Press). Si se usa un hospedador procariótico la secuencia de nucleótidos puede necesitar ser modificada de forma adecuada antes de la transformación - por ejemplo mediante la eliminación de intrones.

Se conocen diversos métodos para la transformación de la especie *Bacillus*.

**Secreción**

30 A menudo es deseable que el polipéptido sea secretado desde el hospedador de expresión al medio de cultivo, desde donde la enzima puede recuperarse más fácilmente. Según la presente invención, la secuencia líder de secreción se puede seleccionar basándose en el hospedador de expresión deseado. También se pueden usar secuencias señal híbridas con el contexto de la presente invención.

35 Los ejemplos típicos de secuencias líder de secreción que no están asociadas con una secuencia de nucleótidos que codifica una lípido aciltransferasa en la naturaleza son las que se originan del gen de amiloglucosidasa (AG) fúngica (*glaA* - las versiones de 18 y 24 aminoácidos, p.ej. de *Aspergillus*), el gen del factor-a (de levadura, p.ej. *Saccharomyces*, *Kluyveromyces* y *Hansenula*) o el gen de  $\alpha$ -amilasa (*Bacillus*).

**Detección**

En la técnica se conoce una variedad de protocolos para detectar y medir la expresión de la secuencia de aminoácidos. Los ejemplos incluyen el ensayo inmunosorbente ligado a enzima (ELISA), el radioinmunoensayo (RIA) y la clasificación de células activada por fluorescencia (FACTS).

40 Los especialistas en la técnica conocen una amplia variedad de etiquetas y técnicas de conjugación y pueden usarse en diversos ensayos con ácido nucleico y aminoácidos.

Una serie de compañías tales como Pharmacia Biotech (Piscataway, NJ), Promega (Madison, WI) y US Biochemical Corp. (Cleveland, OH) suministran kits y protocolos comerciales para dichos procedimientos.

45 Las moléculas informadoras o etiquetas adecuadas incluyen radionucleidos, enzimas, agentes fluorescentes, quimioluminiscentes o cromogénicos, así como sustratos, cofactores, inhibidores, partículas magnéticas, y similares. Las patentes que describen el uso de dichas etiquetas incluyen los documentos US-A-3.817.837; US-A-3.850.752; US-A-3.939.350; US-A-3.996.345; US-A-4.277.437; US-A-4.275.149 y US-A-4.366.241.

Asimismo, se pueden producir inmunoglobulinas recombinantes como se indica en el documento US-A-4.816.567.

**Proteínas de fusión**

50 En el método de la presente descripción, la lípido aciltransferasa se puede producir en forma de una proteína de fusión, por ejemplo para ayudar en la extracción y purificación de la misma. Los ejemplos de compañeros de

proteína de fusión incluyen glutation-S-transferasa (GST), 6xHis, GAL4 (dominios de unión de ADN y/o de activación transcripcional) y  $\beta$ -galactosidasa. También puede ser conveniente incluir un sitio de ruptura proteolítica entre el compañero de la proteína de fusión y la secuencia proteica de interés para permitir la eliminación de secuencias de proteína de fusión. Preferiblemente, la proteína de fusión no dificultará la actividad de la secuencia de proteína.

5 Los sistemas de expresión de fusión de genes en *E. coli* se han revisado en Curr. Opin. Biotechnol. (1995) 6(5): 501-6.

10 En otra realización de la descripción, la secuencia de aminoácidos de un polipéptido que tiene las propiedades específicas definidas en la presente memoria se puede ligar a una secuencia no nativa para codificar una proteína de fusión. Por ejemplo, para cribar bibliotecas de péptidos en busca de agentes capaces de afectar a la actividad de una sustancia, puede ser útil codificar una sustancia quimérica que exprese un epítipo no nativo que sea reconocido por un anticuerpo disponible comercialmente.

La invención de describirá a continuación, a modo de ejemplo solamente, con referencia a las siguientes Figuras y Ejemplos.

15 La Figura 1 muestra la secuencia de aminoácidos de una lípido aciltransferasa madura mutante de *Aeromonas salmonicida* (GCAT) con una mutación de Asn80Asp (en particular, el aminoácido 80 está en la secuencia madura) (SEQ ID 16);

La Figura 2 muestra una secuencia de aminoácidos (SEQ ID N° 1) de una lípido acil transferasa de *Aeromonas hydrophila* (ATCC #7965);

La Figura 3 muestra una secuencia consenso pfam00657 de la base de datos versión 6 (SEQ ID N° 2);

20 La Figura 4 muestra una secuencia de aminoácidos (SEQ ID N° 3) obtenida del organismo *Aeromonas hydrophila* (P10480; GI:121051);

La Figura 5 muestra una secuencia de aminoácidos (SEQ ID N° 4) obtenida del organismo *Aeromonas salmonicida* (AAG098404; GI:9964017);

25 La Figura 6 muestra una secuencia de aminoácidos (SEQ ID N° 5) obtenida del organismo *Streptomyces coelicolor* A3(2) (número de acceso de Genbank NP\_631558);

La Figura 7 muestra una secuencia de aminoácidos (SEQ ID N° 6) obtenida del organismo *Streptomyces coelicolor* A3(2) (número de acceso de Genbank: CAC42140);

La Figura 8 muestra una secuencia de aminoácidos (SEQ ID N° 7) obtenida a partir del organismo *Saccharomyces cerevisiae* (número de acceso de Genbank P41734);

30 La Figura 9 muestra una secuencia de aminoácidos (SEQ ID N° 8) obtenida a partir del organismo *Ralstonia* (número de acceso de Genbank: AL646052);

La Figura 10 muestra la SEQ ID N° 9. Proteína hipotética conservada Scoe1 con código de acceso de proteína del NCBI CAB39707.1 GI:4539178 [Streptomyces coelicolor A3(2)];

35 La Figura 11 muestra una secuencia de aminoácidos mostrada como SEQ ID N° 10. Proteína hipotética conservada Scoe2 con código de acceso de proteína NCBI CAC01477.1 GI:9716139 [Streptomyces coelicolor A3(2)];

La Figura 12 muestra una secuencia de aminoácidos (SEQ ID N° 11) de proteína secretada putativa Scoe3 con código de acceso de proteína NCBI CABS8833.1 GI:7635996. [Streptomyces coelicolor A3(2)];

La Figura 13 muestra una secuencia de aminoácidos (SEQ ID N° 12) de una proteína secretada putativa Scoe4 con código de acceso de proteína NCBI CAB89450.1 GI:7672261. [Streptomyces coelicolor A3(2)];

40 La Figura 14 muestra una secuencia de aminoácidos (SEQ ID N° 13) de lipoproteína putativa Scoe5 con código de acceso de proteína NCBI CAB62724.1 GI:6562793 [Streptomyces coelicolor A3(2)];

La Figura 15 muestra una secuencia de aminoácidos (SEQ ID N° 14) de GDSL-lipasa Srim1 con código de acceso de proteína NCBI AAK84028.1 GI:15082088 [Streptomyces rimosus];

45 La Figura 16 muestra una secuencia de aminoácidos (SEQ ID N° 15) de una lípido aciltransferasa de *Aeromonas salmonicida* subespecie *Salmonicida* (ATCC# 14174);

La Figura 17 muestra la SEQ ID N° 19. Proteína hipotética conservada Scoe1 con código de acceso de proteína NCBI CAB39707.1 GI:4539178 [Streptomyces coelicolor A3(2)];



- La Figura 18 muestra una secuencia de aminoácidos (SEQ ID N° 25) de la construcción de fusión usada para la mutagénesis del gen de lípido aciltransferasa de *Aeromonas hydrophila*. Los aminoácidos subrayados son un péptido señal de xilanasas;
- 5 La Figura 19 muestra una secuencia de polipéptido de una enzima lípido aciltransferasa de *Streptomyces* (SEQ ID N° 26);
- La Figura 20 muestra una secuencia de polipéptido de una enzima lípido aciltransferasa de *Thermobifida* (SEQ ID N° 27);
- La Figura 21 muestra una secuencia de polipéptido de una enzima lípido aciltransferasa de *Thermobifida* (SEQ ID N° 28);
- 10 La Figura 22 muestra un polipéptido de una enzima lípido aciltransferasa de *Corynebacterium efficiens* con GDSx de 300 aminoácidos (SEQ ID N° 29);
- La Figura 23 muestra un polipéptido de una enzima lípido aciltransferasa de *Novosphingobium aromaticivorans* con GDSx de 284 aminoácidos (SEQ ID N° 30);
- 15 La Figura 24 muestra un polipéptido de una enzima lípido aciltransferasa de *Streptomyces coelicolor* con GDSx de 269 aa (SEQ ID N° 31);
- La Figura 25 muestra un polipéptido de una enzima lípido aciltransferasa de *Streptomyces avermitilis* \ GDSx de 269 aminoácidos (SEQ ID N° 32);
- La Figura 26 muestra un polipéptido de una enzima lípido aciltransferasa de *Streptomyces* (SEQ ID N° 33);
- 20 La Figura 27 muestra una secuencia de aminoácidos (SEQ ID N° 34) obtenida del organismo *Aeromonas hydrophila* (P10480; GI:121051) (a destacar, esta es la secuencia madura);
- La Figura 28 muestra la secuencia de aminoácidos (SEQ ID N° 35) de una lípido aciltransferasa madura mutante de *Aeromonas salmonicida* (GCAT) (a destacar, esta es la secuencia madura);
- La Figura 29 muestra una secuencia de nucleótidos (SEQ ID N° 36) de *Streptomyces thermosacchari*;
- La Figura 30 muestra una secuencia de aminoácidos (SEQ ID N° 37) de *Streptomyces thermosacchari*;
- 25 La Figura 31 muestra una secuencia de aminoácidos (SEQ ID N° 38) de *Thermobifida fusca*/GDSx de 548 aminoácidos;
- La Figura 32 muestra una secuencia de nucleótidos (SEQ ID N° 39) de *Thermobifida fusca*;
- La Figura 33 muestra una secuencia de aminoácidos (SEQ ID N° 40) de *Thermobifida fusca*/GDSx;
- 30 La Figura 34 muestra una secuencia de aminoácidos (SEQ ID N° 41) de *Corynebacterium efficiens*/GDSx de 300 aminoácidos;
- La Figura 35 muestra una secuencia de nucleótidos (SEQ ID N° 42) de *Corynebacterium efficiens*;
- La Figura 36 muestra una secuencia de aminoácidos (SEQ ID N° 43) de *S. coelicolor*/ GDSx de 268 aminoácidos;
- La Figura 37 muestra una secuencia de nucleótidos (SEQ ID N° 44) de *S. coelicolor*;
- La Figura 38 muestra una secuencia de aminoácidos (SEQ ID N° 45) de *S. avermitilis*;
- 35 La Figura 39 muestra una secuencia de nucleótidos (SEQ ID N° 46) de *S. avermitilis*;
- La Figura 40 muestra una secuencia de aminoácidos (SEQ ID N° 47) de *Thermobifida fusea*/GDSx;
- La Figura 41 muestra una secuencia de nucleótidos (SEQ ID N° 48) de *Thermobifida fusca*/GDSx;
- La Figura 42 muestra una alineación de L131 y los homólogos de *S. avermitilis* y *T. fusca* que ilustra la conservación del motivo GDSx (GDSY en L131 y *S. avermitilis* y *T. fusca*), la caja GANDY, que es GGND A o GGND L, y el bloque HPT (que se considera que es la histidina catalítica conservada). Estos tres bloques conservados están resaltados;
- 40 La Figura 43 muestra la SEQ ID N° 17 que es la secuencia de aminoácidos de una lípido aciltransferasa de *Candida parapsilosis*;
- La Figura 44 muestra la SEQ ID N° 18 que es la secuencia de aminoácidos de una lípido aciltransferasa de *Candida parapsilosis*;

- La Figura 45 muestra una representación de cinta de la estructura cristalina de 1IVN.PDB que tiene glicerol en el centro activo. La Figura se realizó usando el visor Deep View Swiss-PDB;
- La Figura 46 muestra la Estructura Cristalina de 1IVN.PDB - vista lateral usando el visor Deep View Swiss-PDB, con glicerol en el centro activo - los residuos a menos de 10 Å del centro activo de glicerol están en color negro;
- 5 La Figura 47 muestra la Estructura Cristalina de 1IVN.PDB - vista superior usando el visor Deep View Swiss-PDB, con glicerol en el centro activo - los residuos a menos de 10 Å del centro activo de glicerol están en color negro;
- La Figura 48 muestra la alineación 1;
- La Figura 49 muestra la alineación 2;
- 10 Las Figuras 50 y 51 muestran una alineación de 1IVN con P10480 (P10480 es la secuencia de la base de datos correspondiente a la enzima de *A. hydrophila*), esta alineación se obtuvo de la base de datos PFAM y se usó en el proceso de construcción del modelo; y
- La Figura 52 muestra una alineación en la que P10480 es la secuencia de la base de datos de *Aeromonas hydrophila*. Esta secuencia se usa para la construcción del modelo y para la selección de sitio. Nótese que se muestra la proteína completa (SEQ ID N° 25), la proteína madura (equivalente a SEQ ID N° 34) comienza en el residuo 19. *A. sal* es lipasa GDSX de *Aeromonas salmonicida* (SEQ ID N° 4), *A. hyd* es lipasa GDSX de *Aeromonas hydrophila* (SEQ ID N° 34). La secuencia consenso contiene un \* en la posición de una diferencia entre las secuencias presentadas.
- 15 La Figura 53 muestra una construcción génica usada en el Ejemplo 1;
- Figura 54: muestra una construcción génica de codón optimizado (n° 052907) usada en el Ejemplo 1; y
- 20 La Figura 55 muestra la secuencia del inserto XhoI que contiene el gen precursor LAT-KLM3', las cajas de -35 y -10 están subrayadas;
- La Figura 56 muestra BML780-KLM3'CAP50 (que comprende la SEQ ID No. 16 - colonia superior) y BML780 (la cepa hospedadora vacía - colonia inferior) tras 48 h de crecimiento a 37 °C en agar de tributirina al 1%.
- 25 La Figura 57 muestra una secuencia de nucleótidos de *Aeromonas salmonicida* (SEQ ID N° 49) que incluye la secuencia señal (preLAT - posiciones 1 a 87);
- La Figura 58 muestra una secuencia de nucleótidos (SEQ ID N° 50) que codifica una lípido acil transferasa según la presente invención obtenida del organismo *Aeromonas hydrophila*;
- La Figura 59 muestra una secuencia de nucleótidos (SEQ ID N° 51) que codifica una lípido acil transferasa según la presente invención obtenida del organismo *Aeromonas salmonicida*;
- 30 La Figura 60 muestra una secuencia de nucleótidos (SEQ ID N° 52) que codifica una lípido acil transferasa según la presente invención obtenida del organismo *Streptomyces coelicolor* A3(2) (número de acceso de Genbank NC\_003888.1:8327480..8328367);
- La Figura 61 muestra una secuencia de nucleótidos (SEQ ID N° 53) que codifica una lípido acil transferasa según la presente invención obtenida del organismo *Streptomyces coelicolor* A3(2) (número de acceso de Genbank AL939131.1:265480..266367);
- 35 La Figura 62 muestra una secuencia de nucleótidos (SEQ ID N° 54) que codifica una lípido acil transferasa según la presente invención obtenida del organismo *Saccharomyces cerevisiae* (número de acceso de Genbank Z75034);
- La Figura 63 muestra una secuencia de nucleótidos (SEQ ID N° 55) que codifica una lípido acil transferasa según la presente invención obtenida del organismo *Ralstonia*;
- 40 La Figura 64 muestra una secuencia de nucleótidos mostrada como SEQ ID N° 56 que codifica la proteína hipotética conservada con código de acceso de proteína del NCBI CAB39707.1 GI:4539178 [*Streptomyces coelicolor* A3(2)];
- La Figura 65 muestra una secuencia de nucleótidos mostrada como SEQ ID N° 57 que codifica una proteína hipotética conservada Scoe2 con código de acceso de proteína del NCBI CAC01477.1 GI:9716139 [*Streptomyces coelicolor* A3(2)];
- 45 La Figura 66 muestra una secuencia de nucleótidos mostrada como SEQ ID N° 58 que codifica una proteína secretada putativa Scoe3 con código de acceso de proteína del NCBI CAB88833.1 GI:7635996. [*Streptomyces coelicolor* A3(2)];

La Figura 67 muestra una secuencia de nucleótidos mostrada como SEQ ID N° 59 que codifica una proteína secretada putativa Scoe4 con código de acceso de proteína del NCBI CAB89450.1 GI:7672261. [*Streptomyces coelicolor* A3(2)];

5 La Figura 68 muestra una secuencia de nucleótidos mostrada como SEQ ID N° 60, que codifica una lipoproteína putativa Scoe5 con código de acceso de proteína del NCBI CAB62724.1 GI:6562793 [*Streptomyces coelicolor* A3(2)];

La Figura 69 muestra una secuencia de nucleótidos mostrada como SEQ ID N° 61 que codifica una GDLSL-lipasa Srim1 con código de acceso de proteína del NCBI AAK84028.1 G1:15082088 [*Streptomyces rimosus*];

10 La Figura 70 muestra una secuencia de nucleótidos (SEQ ID N° 62) que codifica una lípido aciltransferasa de *Aeromonas hydrophila* (ATCC #7965);

La Figura 71 muestra una secuencia de nucleótidos (SEQ ID N° 63) que codifica una lípido aciltransferasa de *Aeromonas salmonicida* subsp. *Salmonicida* (ATCC# 14174); y

La Figura 72 muestra una secuencia de nucleótidos (SEQ ID N° 24) que codifica una enzima de *Aeromonas hydrophila* que incluye un péptido señal de xilanasasa.

15 Ejemplo 1

Expresión de KLM3' en *Bacillus licheniformis*

20 Se expresó una secuencia de nucleótidos (SEQ ID N° 49) que codifica una lípido aciltransferasa (SEQ. ID N° 16, más adelante en la presente memoria KLM3') en *Bacillus licheniformis* como una proteína de fusión con el péptido señal de *B. licheniformis* [alfa]-amilasa (LAT) (véanse las Figs. 53 y 54). Para la expresión óptima en *Bacillus*, se encargó una construcción génica de codón optimizado (n° 052907) a Geneart (Geneart AG, Regensburg, Alemania).

25 La construcción n° 052907 contiene un promotor LAT incompleto (solo la secuencia de -10) delante del gen precursor LAT-KLM3' y la transcripción LAT (Tlat) después del gen precursor LAT-KLM3' (véanse las Figs. 53 y 55). Para crear un fragmento *Xho*I que contenga el gen precursor LAT-KLM3' flanqueado por el promotor LAT completo en el extremo 5' y el terminador LAT en el extremo 3', se llevó a cabo una amplificación mediante PCR (reacción en cadena de la polimerasa) con los cebadores Plat5XhoI\_FW y EBS2XhoI\_RV y la construcción génica 052907 como molde.

Plat5XhoI\_FW:

**ccccgctcgaggcttttcttttgggaagaaaatatagggaaaaiggtactgttataaaatfc  
ggaatatttatacaatatcatatggttcacattgaaagggg**

EBS2XhoI\_RV: tggaatctcaaggtttatcctttacctgtctcc

30 La PCR se llevó a cabo en un termociclador con polimerasa de ADN Phusion High Fidelity (Finnzymes OY, Espoo, Finlandia) según las instrucciones del fabricante (temperatura de maduración de 55°C).

El fragmento de PCR resultante se digirió con enzima de restricción *Xho*I y se ligó con T4 ADN ligasa en pCatH digerido con *Xho*I según las instrucciones del proveedor (Invitrogen, Carlsbad, Calif., EE.UU.).

35 La mezcla de ligadura se transformó en la cepa SC6.1 de *B. subtilis* tal como se describe en la Solicitud de Patente de EE.UU. US20020182734 (Publicación Internacional WO 02/14490). La secuencia del inserto de *Xho*I que contiene el gen precursor de LAT-KLM3' se confirmó mediante secuenciación de ADN (BaseClear, Leiden, Holanda) y uno de los clones plasmídicos correctos se denominó pCatH-KLM3'(ori1) (Figura 53). Se transformó el pCatH-KLM3'(ori1) en la cepa BML780 de *B. licheniformis* (un derivado de BRA7 y BML612, véase el documento WO2005111203) a la temperatura permisiva (37 °C).

40 Se seleccionó un transformante resistente a neomicina (neoR) y resistente a cloranfenicol (CmR) y se denominó BML780(pCatH-KLM3'(ori1)). El plásmido de BML780(pCatH-KLM3'(ori1)) se integró en la región catH del genoma de *B. licheniformis* cultivando la cepa a una temperatura no permisiva (50 °C) en medio con 5 µg/mL de cloranfenicol. Se seleccionó un clon resistente a CmR y se denominó BML780-pCatH-KLM3'(ori1). Se cultivó de nuevo el BML780-pCatH-KLM3'(ori1) a la temperatura permisiva durante varias generaciones sin antibióticos para  
45 abrir las secuencias del vector y a continuación se seleccionó un clon CmR sensible a neomicina (neoS). En este clon se escinden las secuencias del vector de pCatH en el cromosoma (que incluyen el gen de resistencia a neomicina) y solo queda el casete catH-LATKLM3'. A continuación, se amplificó el casete catH - LATKLM3' sobre el cromosoma cultivando la cepa en/sobre un medio con concentraciones crecientes de cloranfenicol. Después de  
50 varias rondas de amplificación, se seleccionó un clon (resistente a 50 µg/mL de cloranfenicol) y se designó BML780-KLM3'CAP50. Para verificar la expresión de KLM3', se cultivó BML780-KLM3'CAP50 y BML780 (la cepa hospedadora vacía) durante 48h a 37°C en una placa de agar Heart Infusion (Bacto) con un 1% de tributirina. Una

zona clara, indicativa de la actividad de lípido aciltransferasa, fue claramente visible alrededor de la colonia de BML780-KLM3'CAP50 pero no alrededor de la cepa hospedadora BML780 (véase la Figura 56). Este resultado demuestra que se expresa una cantidad sustancial de KLM3' en la cepa BML780-KLM3'CAP50 de *B. licheniformis* y que dichas moléculas de KLM3' son funcionales.

5 Ejemplo comparativo 1

Construcción de vector

10 La construcción de plásmido es pCS32new N80D, que es un derivado de pCCmini que porta la secuencia que codifica la forma madura de la Glicerofosfolípido-colesterol aciltransferasa nativa de *Aeromonas salmonicida* con una sustitución de Asn a Asp en la posición 80 (KLM3'), bajo control del promotor p32 y con una secuencia señal de CGTasa.

La cepa hospedadora usada para la expresión es la cepa OS21ΔAprE de *bacillus subtilis*.

El nivel de expresión se mide como la actividad de transferasa, expresada como el % de colesterol esterificado, calculado a partir de la diferencia de colesterol libre en la muestra de referencia y de colesterol libre en la muestra de enzima en reacciones con PC ( $T_{PC}$ ) como molécula donante y colesterol como molécula aceptora.

15 Condiciones de cultivo

20 5 ml de caldo LB (digestión enzimática de caseína, 10 g/l; extracto de levadura bajo en sodio, 5 g/l; Cloruro Sódico, 5 g/l; excipientes inertes para comprimidos, 2 g/l) complementado con 50 mg/l de kanamicina, se inoculó con una única colonia y se incubó a 30 °C durante 6 horas a 205 rpm. Se usaron 0,7 ml de este cultivo para inocular 50 ml de medio SAS ( $K_2HPO_4$ , 10 g/l; MOPS (ácido 3-morfolinopropano sulfónico), 40 g/l; Cloruro Sódico, 5 g/l; Antiespumante (Sin 260), 5 gotas/l; harina de soja desgrasada, 20 g/l; Biospringer 106 (YE 100% de peso seco), 20 g/l) complementado con 50 mg/l de kanamicina y una disolución de hidrolizados de almidón con contenido elevado de maltosa (60 g/l). La incubación continuó durante 40 horas a 30 °C y 180 rpm antes de separar el sobrenadante del cultivo mediante centrifugación a 19000 rpm durante 30 min. El sobrenadante se transfirió a un tubo limpio y se usó directamente para medir la actividad de transferasa.

25 Preparación de sustratos y reacción enzimática

Se combinaron PC (Avanti Polar Lipids #441601) y colesterol (Sigma C8503) en la proporción 9:1, se disolvieron en cloroformo, y se evaporaron hasta sequedad.

El sustrato se preparó mediante dispersión de un 3% de PC:Colesterol 9:1 en tampón Hepes 50 mM de pH 7.

30 Se transfirieron 0,250 ml de disolución de sustrato a un tubo de vidrio de 3 ml con tapa de rosca. Se añadieron 0,025 ml de sobrenadante de cultivo y la mezcla se incubó a 40 °C durante 2 horas. También se preparó una muestra de referencia con agua en vez de enzima. El calentamiento de la mezcla de reacción en un baño de agua en ebullición durante 10 minutos detuvo la reacción enzimática. Se añadieron 2 ml de un 99% de etanol a la mezcla de reacción antes de someterla a un análisis de colesterol.

Ensayo de colesterol

35 100 µl de sustrato que contenía 1,4 U/ml de Colesterol oxidasa (SERVA Electrophoresis GmbH N° de cat. 17109), 0,4 mg/ml de ABTS (Sigma A-1888), 6 U/ml de Peroxidasa (Sigma 6782) en Tris-HCl 0,1 M, pH 6,6 y un 0,5 % de Triton X-100 (Sigma X-100) se incubaron a 37 °C durante 5 minutos antes de añadir 5 µl de muestra de reacción enzimática y mezclar. La mezcla de reacción se incubó durante otros 5 minutos y se midió la  $DO_{405}$ . El contenido de colesterol se calculó a partir de los análisis de disoluciones patrón de colesterol que contenían 0,4 mg/ml, 0,3 mg/ml, 0,20 mg/ml, 0,1 mg/ml, 0,05 mg/ml, y 0 mg/ml de colesterol en un 99% de EtOH.

Resultados

La tabla muestra la media de 8 cultivos de expresión distintos.

Cepa	$T_{PC}^a$
OS21ΔAprE[pCS32new]	74,2±10,1 <sup>b</sup>

<sup>a</sup>  $T_{PC}$  es la actividad de transferasa, expresada como el % de colesterol esterificado, calculada a partir de la diferencia de colesterol libre en la muestra de referencia y el colesterol libre en la muestra de enzima en reacciones con PC como molécula donante y colesterol como molécula aceptora.

<sup>b</sup> Media de 8 cultivos de expresión distintos

La invención se describe adicionalmente en los siguientes párrafos numerados;

1. Un método para la producción de una lípido aciltransferasa que comprende las etapas de:
  - (i) proporcionar una célula de *Bacillus licheniformis*;
  - 5 (ii) transformar la célula de *Bacillus licheniformis* con una secuencia de nucleótidos heteróloga que codifica una lípido aciltransferasa; y
  - (iii) expresar la lípido aciltransferasa en la célula bajo control de una secuencia promotora.
2. Un método según el párrafo 1 en el que la secuencia promotora no está asociada de forma nativa con la secuencia de nucleótidos que codifica una lípido aciltransferasa.
3. Un método según el párrafo 2, en el que se une de forma operable una secuencia de nucleótidos que codifica un péptido señal a dicha secuencia de nucleótidos heteróloga que codifica una lípido aciltransferasa.
4. Un método según cualquiera de los párrafos precedentes, en el que dicho método comprende la etapa adicional de aislar/recuperar la lípido aciltransferasa.
5. Un método según cualquiera de los párrafos precedentes en el que dicha secuencia promotora es homóloga respecto de la célula hospedadora.
6. Un método según cualquiera de los párrafos precedentes en el que dicha secuencia promotora se selecciona del grupo que consiste en una secuencia promotora de  $\alpha$ -amilasa, una secuencia promotora de proteasa, una secuencia promotora de subtilisina, una secuencia promotora de proteasa específica de ácido glutámico y una secuencia promotora de levansacarasa.
7. Un método según cualquiera de los párrafos precedentes en el que dicha secuencia promotora es la secuencia promotora de  $\alpha$ -amilasa.
8. Un método según cualquiera de los párrafos precedentes en el que la secuencia de nucleótidos que codifica una lípido acil transferasa codifica una lípido aciltransferasa que comprende un motivo GDSx y/o un motivo GANDY.
9. Un método según cualquiera de los párrafos precedentes en el que la secuencia de nucleótidos que codifica una lípido aciltransferasa codifica una lípido aciltransferasa que se caracteriza como una enzima que posee actividad de acil transferasa y que comprende el motivo de secuencia de aminoácidos GDSX, en el que X es uno o más de los residuos de aminoácidos siguientes L, A, V, I, F, Y, H, Q, T, N, M o S.
10. Un método según cualquiera de los párrafos precedentes en el que la secuencia de nucleótidos que codifica una lípido aciltransferasa codifica una lípido aciltransferasa que es obtenible de un organismo de uno o más de los géneros siguientes: *Aeromonas*, *Streptomyces*, *Saccharomyces*, *Lactococcus*, *Mycobacterium*, *Streptococcus*, *Lactobacillus*, *Desulfitobacterium*, *Bacillus*, *Campylobacter*, *Vibrionaceae*, *Xylella*, *Sulfolobus*, *Aspergillus*, *Schizosaccharomyces*, *Listeria*, *Neisseria*, *Mesorhizobium*, *Ralstonia*, *Xanthomonas* y *Candida*.
11. Un método según cualquiera de los párrafos precedentes en el que la secuencia de nucleótidos que codifica una lípido aciltransferasa codifica una lípido aciltransferasa que es obtenible de un organismo del género *Aeromonas*.
12. Un método según cualquiera de los párrafos precedentes en el que la secuencia de nucleótidos que codifica una lípido aciltransferasa codifica una lípido aciltransferasa que comprende un residuo de ácido aspártico en una posición que corresponde a N-80 en la secuencia de aminoácidos de la lípido aciltransferasa de *Aeromonas hydrophila* mostrada como SEQ ID N° 35.
13. Un método según cualquiera de los párrafos precedentes, en el que la secuencia de nucleótidos que codifica una lípido aciltransferasa codifica una lípido aciltransferasa que comprende la secuencia de aminoácidos mostrada como SEQ ID N° 16, o una secuencia de aminoácidos que tiene al menos un 75% de homología con ella.
14. Un método según cualquiera de los párrafos precedentes en el que la secuencia de nucleótidos que codifica una lípido aciltransferasa codifica una lípido aciltransferasa que comprende la secuencia de aminoácidos mostrada como SEQ ID N° 16.
15. Una célula hospedadora de *Bacillus licheniformis* que comprende una secuencia de nucleótidos que codifica una lípido aciltransferasa heteróloga.
16. Una célula hospedadora según el párrafo 14 en la que la secuencia de nucleótidos que codifica una lípido aciltransferasa codifica una lípido aciltransferasa que comprende un motivo GDSx y/o un motivo GANDY.
17. Una célula hospedadora según el párrafo 15 o el párrafo 16 en la que la secuencia de nucleótidos que codifica una lípido aciltransferasa codifica una lípido aciltransferasa que se caracteriza como una enzima que posee

actividad de aciltransferasa y que comprende el motivo de secuencia de aminoácidos GDSX, en el que X es uno o más de los residuos de aminoácidos siguientes L, A, V, I, F, Y, H, Q, T, N, M o S.

- 5 18. Una célula hospedadora según cualquiera de los párrafos 15 a 17 en la que la lípido aciltransferasa es obtenible de un organismo de uno o más de los géneros siguientes: *Aeromonas*, *Streptomyces*, *Saccharomyces*, *Lactococcus*, *Mycobacterium*, *Streptococcus*, *Lactobacillus*, *Desulfitobacterium*, *Bacillus*, *Campylobacter*, *Vibrionaceae*, *Xylella*, *Sulfolobus*, *Aspergillus*, *Schizosaccharomyces*, *Listeria*, *Neisseria*, *Mesorhizobium*, *Ralstonia*, *Xanthomonas* y *Candida*.
- 10 19. Una célula hospedadora según cualquiera de los párrafos 15 a 17 en la que la secuencia de nucleótidos que codifica una lípido aciltransferasa codifica una lípido aciltransferasa que es obtenible de un organismo del género *Aeromonas*.
- 20 20. Una célula hospedadora según cualquiera de los párrafos 15 a 19 en la que la secuencia de nucleótidos que codifica una lípido aciltransferasa codifica una lípido aciltransferasa que comprende un residuo de ácido aspártico en una posición que corresponde a N-80 en la secuencia de aminoácidos de la lípido aciltransferasa de *Aeromonas hydrophila* mostrada como SEQ ID N° 35.
- 15 21. Una célula hospedadora según cualquiera de los párrafos 15 a 20 en la que la secuencia de nucleótidos que codifica una lípido aciltransferasa codifica una lípido aciltransferasa que comprende la secuencia de aminoácidos mostrada como SEQ ID N° 16, o una secuencia de aminoácidos que tiene un 75% o más de homología con ella.
- 20 22. Una célula hospedadora según cualquiera de los párrafos 15 a 21 en la que la secuencia de nucleótidos que codifica una lípido aciltransferasa codifica una lípido aciltransferasa que comprende la secuencia de aminoácidos mostrada como SEQ ID N° 16.
23. El uso de una célula hospedadora de *Bacillus licheniformis* en la producción de una lípido aciltransferasa heteróloga.
24. El uso según el párrafo 23 en el que existe una expresión incrementada en comparación con la expresión en una célula hospedadora de *B. subtilis*.
- 25 25. Un vector de expresión que comprende una secuencia de nucleótidos que codifica una lípido aciltransferasa unida de forma operable a una secuencia promotora homóloga respecto de *B. licheniformis*.
26. Un vector de expresión según el párrafo 25, en el que el promotor no está asociado de forma nativa con la secuencia de nucleótidos que codifica una lípido aciltransferasa.
- 30 27. Un vector de expresión según el párrafo 25 o el párrafo 26 en el que una secuencia de nucleótidos que codifica un péptido señal está unido de forma operable a dicha secuencia de nucleótidos heteróloga que codifica una lípido aciltransferasa.
- 35 28. Un vector de expresión según cualquiera de los párrafos precedentes en el que dicha secuencia promotora se selecciona del grupo que consiste en una secuencia promotora de  $\alpha$ -amilasa, una secuencia promotora de proteasa, una secuencia promotora de subtilisina, una secuencia promotora de proteasa específica de ácido glutámico y una secuencia promotora de levansacarasa.
29. Un vector de expresión según cualquiera de los párrafos precedentes en el que dicha secuencia promotora es la secuencia promotora de  $\alpha$ -amilasa.
30. Un método para la producción de una lípido aciltransferasa que comprende las etapas de:
- (i) proporcionar una célula de *Bacillus*, en la que la célula de *Bacillus* es distinta de *B. subtilis*;
- 40 (ii) transformar la célula de *Bacillus*, en la que la célula de *Bacillus* es distinta de *B. subtilis* con una secuencia de nucleótidos heteróloga que codifica una lípido aciltransferasa; y
- (iii) expresar la lípido aciltransferasa en la célula bajo control de una secuencia promotora.
31. Una célula hospedadora de *Bacillus* en la que la célula de *Bacillus* es distinta de *B. subtilis* que comprende una secuencia de nucleótidos que codifica una lípido aciltransferasa heteróloga.
- 45 32. El uso de una célula hospedadora de *Bacillus* en la que la célula hospedadora de *Bacillus* es distinta de *B. subtilis* en la producción de una lípido aciltransferasa heteróloga.
33. Un uso como se describió sustancialmente antes en la presente memoria con referencia a los Ejemplos y Figuras.
- 50 34. Un vector de expresión como se describió sustancialmente antes en la presente memoria con referencia a los Ejemplos y Figuras.

35. Una célula hospedadora de *Bacillus licheniformis* como se describió sustancialmente antes en la presente memoria con referencia a los Ejemplos y Figuras.

Listado de secuencias

- 5 <110> DANISCO A/S
- <120> PROCESO
- <130> P025001WO
- 10 <140> PCT/IB2007/000558
- <141> 25-01-2007
- <160> 118
- 15 <170> PatentIn versión 3.5
- <210> 1
- <211> 335
- 20 <212> PRT
- <213> Aeromonas hydrophila
- <400> 1

```

Met Lys Lys Trp Phe Val Cys Leu Leu Gly Leu Val Ala Leu Thr Val
 1          5          10          15

Gln Ala Ala Asp Ser Arg Pro Ala Phe Ser Arg Ile Val Met Phe Gly
 20          25          30

Asp Ser Leu Ser Asp Thr Gly Lys Met Tyr Ser Lys Met Arg Gly Tyr
 35          40          45

Leu Pro Ser Ser Pro Pro Tyr Tyr Glu Gly Arg Phe Ser Asn Gly Pro
 50          55          60

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

Asn Glu Ala Glu Gly Gly Ala Thr Ala Val Ala Tyr Asn Lys Ile Ser
 85          90          95

Trp Asn Pro Lys Tyr Gln Val Ile Asn Asn Leu Asp Tyr Glu Val Thr
 100         105         110

Gln Phe Leu Gln Lys Asp Ser Phe Lys Pro Asp Asp Leu Val Ile Leu
 115         120         125

Trp Val Gly Ala Asn Asp Tyr Leu Ala Tyr Gly Trp Asn Thr Glu Gln
 130         135         140

Asp Ala Lys Arg Val Arg Asp Ala Ile Ser Asp Ala Ala Asn Arg Met
 145         150         155         160

Val Leu Asn Gly Ala Lys Gln Ile Leu Leu Phe Asn Leu Pro Asp Leu
 165         170         175
    
```

25

ES 2 449 474 T3

Gly Gln Asn Pro Ser Ala Arg Ser Gln Lys Val Val Glu Ala Val Ser  
 180 185 190

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

Leu Ala Pro Thr Gly Met Val Lys Leu Phe Glu Ile Asp Lys Gln Phe  
 210 215 220

Ala Glu Met Leu Arg Asp Pro Gln Asn Phe Gly Leu Ser Asp Val Glu  
 225 230 235 240

Asn Pro Cys Tyr Asp Gly Gly Tyr Val Trp Lys Pro Phe Ala Thr Arg  
 245 250 255

Ser Val Ser Thr Asp Arg Gln Leu Ser Ala Phe Ser Pro Gln Glu Arg  
 260 265 270

Leu Ala Ile Ala Gly Asn Pro Leu Leu Ala Gln Ala Val Ala Ser Pro  
 275 280 285

Met Ala Arg Arg Ser Ala Ser Pro Leu Asn Cys Glu Gly Lys Met Phe  
 290 295 300

Trp Asp Gln Val His Pro Thr Thr Val Val His Ala Ala Leu Ser Glu  
 305 310 315 320

Arg Ala Ala Thr Phe Ile Ala Asn Gln Tyr Glu Phe Leu Ala His  
 325 330 335

- 5 <210> 2
- <211> 361
- <212> PRT
- <213> Artificial

- 10 <220>
- <223> Secuencia de consenso

<400> 2

Ile Val Ala Phe Gly Asp Ser Leu Thr Asp Gly Glu Ala Tyr Tyr Gly  
 1 5 10 15

Asp Ser Asp Gly Gly Gly Trp Gly Ala Gly Leu Ala Asp Arg Leu Thr  
 20 25 30

Ala Leu Leu Arg Leu Arg Ala Arg Pro Arg Gly Val Asp Val Phe Asn  
 35 40 45

Arg Gly Ile Ser Gly Arg Thr Ser Asp Gly Arg Leu Ile Val Asp Ala  
 50 55 60

15



ES 2 449 474 T3

Leu Val Ala Leu Leu Phe Leu Ala Gln Ser Leu Gly Leu Pro Asn Leu  
 65 70 75 80  
 Pro Pro Tyr Leu Ser Gly Asp Phe Leu Arg Gly Ala Asn Phe Ala Ser  
 85 90 95  
 Ala Gly Ala Thr Ile Leu Pro Thr Ser Gly Pro Phe Leu Ile Gln Val  
 100 105 110  
 Gln Phe Lys Asp Phe Lys Ser Gln Val Leu Glu Leu Arg Gln Ala Leu  
 115 120 125  
 Gly Leu Leu Gln Glu Leu Leu Arg Leu Leu Pro Val Leu Asp Ala Lys  
 130 135 140  
 Ser Pro Asp Leu Val Thr Ile Met Ile Gly Thr Asn Asp Leu Ile Thr  
 145 150 155 160  
 Ser Ala Phe Phe Gly Pro Lys Ser Thr Glu Ser Asp Arg Asn Val Ser  
 165 170 175  
 Val Pro Glu Phe Lys Asp Asn Leu Arg Gln Leu Ile Lys Arg Leu Arg  
 180 185 190  
 Ser Asn Asn Gly Ala Arg Ile Ile Val Leu Ile Thr Leu Val Ile Leu  
 195 200 205  
 Asn Leu Gly Pro Leu Gly Cys Leu Pro Leu Lys Leu Ala Leu Ala Leu  
 210 215 220  
 Ala Ser Ser Lys Asn Val Asp Ala Ser Gly Cys Leu Glu Arg Leu Asn  
 225 230 235 240  
 Glu Ala Val Ala Asp Phe Asn Glu Ala Leu Arg Glu Leu Ala Ile Ser  
 245 250 255  
 Lys Leu Glu Asp Gln Leu Arg Lys Asp Gly Leu Pro Asp Val Lys Gly  
 260 265 270  
 Ala Asp Val Pro Tyr Val Asp Leu Tyr Ser Ile Phe Gln Asp Leu Asp  
 275 280 285  
 Gly Ile Gln Asn Pro Ser Ala Tyr Val Tyr Gly Phe Glu Thr Thr Lys  
 290 295 300  
 Ala Cys Cys Gly Tyr Gly Gly Arg Tyr Asn Tyr Asn Arg Val Cys Gly  
 305 310 315 320  
 Asn Ala Gly Leu Cys Asn Val Thr Ala Lys Ala Cys Asn Pro Ser Ser  
 325 330 335  
 Tyr Leu Leu Ser Phe Leu Phe Trp Asp Gly Phe His Pro Ser Glu Lys  
 340 345 350  
 Gly Tyr Lys Ala Val Ala Glu Ala Leu  
 355 360

ES 2 449 474 T3

<210> 3  
 <211> 335  
 <212> PRT  
 <213> Aeromonas hydrophila

5

<400> 3

Met Lys Lys Trp Phe Val Cys Leu Leu Gly Leu Val Ala Leu Thr Val  
 1 5 10 15

Gln Ala Ala Asp Ser Arg Pro Ala Phe Ser Arg Ile Val Met Phe Gly  
 20 25 30

Asp Ser Leu Ser Asp Thr Gly Lys Met Tyr Ser Lys Met Arg Gly Tyr  
 35 40 45

Leu Pro Ser Ser Pro Pro Tyr Tyr Glu Gly Arg Phe Ser Asn Gly Pro  
 50 55 60

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

Asn Glu Ala Glu Gly Gly Pro Thr Ala Val Ala Tyr Asn Lys Ile Ser  
 85 90 95

Trp Asn Pro Lys Tyr Gln Val Ile Asn Asn Leu Asp Tyr Glu Val Thr  
 100 105 110

Gln Phe Leu Gln Lys Asp Ser Phe Lys Pro Asp Asp Leu Val Ile Leu  
 115 120 125

Trp Val Gly Ala Asn Asp Tyr Leu Ala Tyr Gly Trp Asn Thr Glu Gln  
 130 135 140

Asp Ala Lys Arg Val Arg Asp Ala Ile Ser Asp Ala Ala Asn Arg Met  
 145 150 155 160

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

Gly Gln Asn Pro Ser Ala Arg Ser Gln Lys Val Val Glu Ala Ala Ser

ES 2 449 474 T3

180

185

190

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

Leu Ala Pro Thr Gly Met Val Lys Leu Phe Glu Ile Asp Lys Gln Phe  
210 215 220

Ala Glu Met Leu Arg Asp Pro Gln Asn Phe Gly Leu Ser Asp Gln Arg  
225 230 235 240

Asn Ala Cys Tyr Gly Gly Ser Tyr Val Trp Lys Pro Phe Ala Ser Arg  
245 250 255

Ser Ala Ser Thr Asp Ser Gln Leu Ser Ala Phe Asn Pro Gln Glu Arg  
260 265 270

Leu Ala Ile Ala Gly Asn Pro Leu Leu Ala Gln Ala Val Ala Ser Pro  
275 280 285

Met Ala Ala Arg Ser Ala Ser Thr Leu Asn Cys Glu Gly Lys Met Phe  
290 295 300

Trp Asp Gln Val His Pro Thr Thr Val Val His Ala Ala Leu Ser Glu  
305 310 315 320

Pro Ala Ala Thr Phe Ile Glu Ser Gln Tyr Glu Phe Leu Ala His  
325 330 335

<210> 4

<211> 336

<212> PRT

<213> Aeromonas salmonicida

<400> 4

Met Lys Lys Trp Phe Val Cys Leu Leu Gly Leu Ile Ala Leu Thr Val  
1 5 10 15

Gln Ala Ala Asp Thr Arg Pro Ala Phe Ser Arg Ile Val Met Phe Gly  
20 25 30

Asp Ser Leu Ser Asp Thr Gly Lys Met Tyr Ser Lys Met Arg Gly Tyr  
35 40 45

Leu Pro Ser Ser Pro Pro Tyr Tyr Glu Gly Arg Phe Ser Asn Gly Pro  
50 55 60

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

5

10

ES 2 449 474 T3

Asn Glu Ala Glu Gly Gly Ala Thr Ala Val Ala Tyr Asn Lys Ile Ser  
85 90 95

Trp Asn Pro Lys Tyr Gln Val Tyr Asn Asn Leu Asp Tyr Glu Val Thr  
100 105 110

Gln Phe Leu Gln Lys Asp Ser Phe Lys Pro Asp Asp Leu Val Ile Leu  
115 120 125

Trp Val Gly Ala Asn Asp Tyr Leu Ala Tyr Gly Trp Asn Thr Glu Gln  
130 135 140

Asp Ala Lys Arg Val Arg Asp Ala Ile Ser Asp Ala Ala Asn Arg Met  
145 150 155 160

Val Leu Asn Gly Ala Lys Gln Ile Leu Leu Phe Asn Leu Pro Asp Leu  
165 170 175

Gly Gln Asn Pro Ser Ala Arg Ser Gln Lys Val Val Glu Ala Val Ser  
180 185 190

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

Leu Ala Pro Thr Gly Met Val Lys Leu Phe Glu Ile Asp Lys Gln Phe  
210 215 220

Ala Glu Met Leu Arg Asp Pro Gln Asn Phe Gly Leu Ser Asp Val Glu  
225 230 235 240

Asn Pro Cys Tyr Asp Gly Gly Tyr Val Trp Lys Pro Phe Ala Thr Arg  
245 250 255

Ser Val Ser Thr Asp Arg Gln Leu Ser Ala Phe Ser Pro Gln Glu Arg  
260 265 270

Leu Ala Ile Ala Gly Asn Pro Leu Leu Ala Gln Ala Val Ala Ser Pro  
275 280 285

Met Ala Arg Arg Ser Ala Ser Pro Leu Asn Cys Glu Gly Lys Met Phe  
290 295 300

Trp Asp Gln Val His Pro Thr Thr Val Val His Ala Ala Leu Ser Glu  
305 310 315 320

Arg Ala Ala Thr Phe Ile Glu Thr Gln Tyr Glu Phe Leu Ala His Gly  
325 330 335

<210> 5

<211> 295

5 <212> PRT

<213> Streptomyces coelicolor

<400> 5

ES 2 449 474 T3

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

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

Pro Ala Gln Ala Thr Pro Thr Leu Asp Tyr Val Ala Leu Gly Asp Ser  
 35 40 45

Tyr Ser Ala Gly Ser Gly Val Leu Pro Val Asp Pro Ala Asn Leu Leu  
 50 55 60

Cys Leu Arg Ser Thr Ala Asn Tyr Pro His Val Ile Ala Asp Thr Thr  
 65 70 75 80

Gly Ala Arg Leu Thr Asp Val Thr Cys Gly Ala Ala Gln Thr Ala Asp  
 85 90 95

Phe Thr Arg Ala Gln Tyr Pro Gly Val Ala Pro Gln Leu Asp Ala Leu  
 100 105 110

Gly Thr Gly Thr Asp Leu Val Thr Leu Thr Ile Gly Gly Asn Asp Asn  
 115 120 125

Ser Thr Phe Ile Asn Ala Ile Thr Ala Cys Gly Thr Ala Gly Val Leu  
 130 135 140

Ser Gly Gly Lys Gly Ser Pro Cys Lys Asp Arg His Gly Thr Ser Phe  
 145 150 155 160

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

Leu Gly Val Arg Ala Arg Ala Pro His Ala Arg Val Ala Ala Leu Gly  
 180 185 190

Tyr Pro Trp Ile Thr Pro Ala Thr Ala Asp Pro Ser Cys Phe Leu Lys  
 195 200 205

Leu Pro Leu Ala Ala Gly Asp Val Pro Tyr Leu Arg Ala Ile Gln Ala  
 210 215 220

His Leu Asn Asp Ala Val Arg Arg Ala Ala Glu Glu Thr Gly Ala Thr  
 225 230 235 240  
 Tyr Val Asp Phe Ser Gly Val Ser Asp Gly His Asp Ala Cys Glu Ala  
 245 250 255

Pro Gly Thr Arg Trp Ile Glu Pro Leu Leu Phe Gly His Ser Leu Val  
 260 265 270

Pro Val His Pro Asn Ala Leu Gly Glu Arg Arg Met Ala Glu His Thr  
 275 280 285

Met Asp Val Leu Gly Leu Asp  
 290 295

ES 2 449 474 T3

<210> 6

<211> 295

<212> PRT

5 <213> Streptomyces coelicolor

<400> 6

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

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

Pro Ala Gln Ala Thr Pro Thr Leu Asp Tyr Val Ala Leu Gly Asp Ser  
35 40 45

Tyr Ser Ala Gly Ser Gly Val Leu Pro Val Asp Pro Ala Asn Leu Leu  
50 55 60

Cys Leu Arg Ser Thr Ala Asn Tyr Pro His Val Ile Ala Asp Thr Thr  
65 70 75 80

Gly Ala Arg Leu Thr Asp Val Thr Cys Gly Ala Ala Gln Thr Ala Asp  
85 90 95

Phe Thr Arg Ala Gln Tyr Pro Gly Val Ala Pro Gln Leu Asp Ala Leu  
100 105 110

Gly Thr Gly Thr Asp Leu Val Thr Leu Thr Ile Gly Gly Asn Asp Asn  
115 120 125

Ser Thr Phe Ile Asn Ala Ile Thr Ala Cys Gly Thr Ala Gly Val Leu  
130 135 140

Ser Gly Gly Lys Gly Ser Pro Cys Lys Asp Arg His Gly Thr Ser Phe  
145 150 155 160

10

ES 2 449 474 T3

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

Leu Gly Val Arg Ala Arg Ala Pro His Ala Arg Val Ala Ala Leu Gly  
 180 185 190

Tyr Pro Trp Ile Thr Pro Ala Thr Ala Asp Pro Ser Cys Phe Leu Lys  
 195 200 205

Leu Pro Leu Ala Ala Gly Asp Val Pro Tyr Leu Arg Ala Ile Gln Ala  
 210 215 220

His Leu Asn Asp Ala Val Arg Arg Ala Ala Glu Glu Thr Gly Ala Thr  
 225 230 235 240

Tyr Val Asp Phe Ser Gly Val Ser Asp Gly His Asp Ala Cys Glu Ala  
 245 250 255

Pro Gly Thr Arg Trp Ile Glu Pro Leu Leu Phe Gly His Ser Leu Val  
 260 265 270

Pro Val His Pro Asn Ala Leu Gly Glu Arg Arg Met Ala Glu His Thr  
 275 280 285

Met Asp Val Leu Gly Leu Asp  
 290 295

<210> 7

<211> 238

5 <212> PRT

<213> *Saccharomyces cerevisiae*

<400> 7

Met Asp Tyr Glu Lys Phe Leu Leu Phe Gly Asp Ser Ile Thr Glu Phe  
 1 5 10 15

Ala Phe Asn Thr Arg Pro Ile Glu Asp Gly Lys Asp Gln Tyr Ala Leu  
 20 25 30

Gly Ala Ala Leu Val Asn Glu Tyr Thr Arg Lys Met Asp Ile Leu Gln  
 35 40 45

Arg Gly Phe Lys Gly Tyr Thr Ser Arg Trp Ala Leu Lys Ile Leu Pro  
 50 55 60

Glu Ile Leu Lys His Glu Ser Asn Ile Val Met Ala Thr Ile Phe Leu  
 65 70 75 80

Gly Ala Asn Asp Ala Cys Ser Ala Gly Pro Gln Ser Val Pro Leu Pro  
 85 90 95

10

ES 2 449 474 T3

Glu Phe Ile Asp Asn Ile Arg Gln Met Val Ser Leu Met Lys Ser Tyr  
 100 105 110

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

Trp Glu Lys Glu Lys Ser Glu Glu Ile Ala Leu Gly Tyr Phe Arg Thr  
 130 135 140

Asn Glu Asn Phe Ala Ile Tyr Ser Asp Ala Leu Ala Lys Leu Ala Asn  
 145 150 155 160

Glu Glu Lys Val Pro Phe Val Ala Leu Asn Lys Ala Phe Gln Gln Glu  
 165 170 175

Gly Gly Asp Ala Trp Gln Gln Leu Leu Thr Asp Gly Leu His Phe Ser  
 180 185 190

Gly Lys Gly Tyr Lys Ile Phe His Asp Glu Leu Leu Lys Val Ile Glu  
 195 200 205

Thr Phe Tyr Pro Gln Tyr His Pro Lys Asn Met Gln Tyr Lys Leu Lys  
 210 215 220

Asp Trp Arg Asp Val Leu Asp Asp Gly Ser Asn Ile Met Ser  
 225 230 235

<210> 8  
 <211> 347  
 5 <212> PRT  
 <213> Ralstonia sp.

<400> 8

Met Asn Leu Arg Gln Trp Met Gly Ala Ala Thr Ala Ala Leu Ala Leu  
 1 5 10 15

Gly Leu Ala Ala Cys Gly Gly Gly Gly Thr Asp Gln Ser Gly Asn Pro  
 20 25 30

Asn Val Ala Lys Val Gln Arg Met Val Val Phe Gly Asp Ser Leu Ser  
 35 40 45

Asp Ile Gly Thr Tyr Thr Pro Val Ala Gln Ala Val Gly Gly Gly Lys  
 50 55 60

10 Phe Thr Thr Asn Pro Gly Pro Ile Trp Ala Glu Thr Val Ala Ala Gln  
 65 70 75 80



ES 2 449 474 T3

Leu Gly Val Thr Leu Thr Pro Ala Val Met Gly Tyr Ala Thr Ser Val  
85 90 95

Gln Asn Cys Pro Lys Ala Gly Cys Phe Asp Tyr Ala Gln Gly Gly Ser  
100 105 110

Arg Val Thr Asp Pro Asn Gly Ile Gly His Asn Gly Gly Ala Gly Ala  
115 120 125

Leu Thr Tyr Pro Val Gln Gln Gln Leu Ala Asn Phe Tyr Ala Ala Ser  
130 135 140

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

Ser Asn Asp Ile Phe Phe Trp Thr Thr Ala Ala Ala Thr Ser Gly Ser  
165 170 175

Gly Val Thr Pro Ala Ile Ala Thr Ala Gln Val Gln Gln Ala Ala Thr  
180 185 190

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

Val Tyr Val Phe Asn Leu Pro Asp Ser Ser Leu Thr Pro Asp Gly Val  
210 215 220

Ala Ser Gly Thr Thr Gly Gln Ala Leu Leu His Ala Leu Val Gly Thr  
225 230 235 240

Phe Asn Thr Thr Leu Gln Ser Gly Leu Ala Gly Thr Ser Ala Arg Ile  
245 250 255

Ile Asp Phe Asn Ala Gln Leu Thr Ala Ala Ile Gln Asn Gly Ala Ser  
260 265 270

Phe Gly Phe Ala Asn Thr Ser Ala Arg Ala Cys Asp Ala Thr Lys Ile  
275 280 285

Asn Ala Leu Val Pro Ser Ala Gly Gly Ser Ser Leu Phe Cys Ser Ala  
290 295 300

Asn Thr Leu Val Ala Ser Gly Ala Asp Gln Ser Tyr Leu Phe Ala Asp  
305 310 315 320

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

Leu Ala Arg Leu Leu Ala Asp Asn Val Ala His  
340 345

- 5 <210> 9
- <211> 261
- <212> PRT
- <213> Streptomyces coelicolor

- 10 <400> 9

ES 2 449 474 T3

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

Gly Asp Pro Gly Pro Asp Gly Ala Phe Val Gly Trp Ala Asp Arg Leu  
 20 25 30

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

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

Val Pro Arg Val Val Gly Leu Ala Pro Asp Leu Val Ser Phe Ala Ala  
 65 70 75 80

Gly Gly Asn Asp Ile Ile Arg Pro Gly Thr Asp Pro Asp Glu Val Ala  
 85 90 95

Glu Arg Phe Glu Leu Ala Val Ala Ala Leu Thr Ala Ala Ala Gly Thr  
 100 105 110

Val Leu Val Thr Thr Gly Phe Asp Thr Arg Gly Val Pro Val Leu Lys  
 115 120 125

His Leu Arg Gly Lys Ile Ala Thr Tyr Asn Gly His Val Arg Ala Ile  
 130 135 140

Ala Asp Arg Tyr Gly Cys Pro Val Leu Asp Leu Trp Ser Leu Arg Ser  
 145 150 155 160

Val Gln Asp Arg Arg Ala Trp Asp Ala Asp Arg Leu His Leu Ser Pro  
 165 170 175

Glu Gly His Thr Arg Val Ala Leu Arg Ala Gly Gln Ala Leu Gly Leu  
 180 185 190

Arg Val Pro Ala Asp Pro Asp Gln Pro Trp Pro Pro Leu Pro Pro Arg  
 195 200 205

Gly Thr Leu Asp Val Arg Arg Asp Asp Val His Trp Ala Arg Glu Tyr  
 210 215 220

Leu Val Pro Trp Ile Gly Arg Arg Leu Arg Gly Glu Ser Ser Gly Asp  
 225 230 235 240

His Val Thr Ala Lys Gly Thr Leu Ser Pro Asp Ala Ile Lys Thr Arg  
 245 250 255

Ile Ala Ala Val Ala  
 260

- 5 <210> 10
- <211> 260
- <212> PRT
- <213> Streptomyces coelicolor

ES 2 449 474 T3

<400> 10

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

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

Trp Ala Asp Leu Leu Ala Thr Arg Met Ala Ala Arg Ser Pro Gly Phe  
35 40 45

Arg Tyr Ala Asn Leu Ala Val Arg Gly Lys Leu Ile Gly Gln Ile Val  
50 55 60

Asp Glu Gln Val Asp Val Ala Ala Ala Met Gly Ala Asp Val Ile Thr  
65 70 75 80

Leu Val Gly Gly Leu Asn Asp Thr Leu Arg Pro Lys Cys Asp Met Ala  
85 90 95

Arg Val Arg Asp Leu Leu Thr Gln Ala Val Glu Arg Leu Ala Pro His  
100 105 110

Cys Glu Gln Leu Val Leu Met Arg Ser Pro Gly Arg Gln Gly Pro Val  
115 120 125

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

Asp Leu Ala Gly Arg His Gly Ala Val Val Val Asp Leu Tyr Gly Ala  
145 150 155 160

Gln Ser Leu Ala Asp Pro Arg Met Trp Asp Val Asp Arg Leu His Leu  
165 170 175

Thr Ala Glu Gly His Arg Arg Val Ala Glu Ala Val Trp Gln Ser Leu  
180 185 190

Gly His Glu Pro Glu Asp Pro Glu Trp His Ala Pro Ile Pro Ala Thr  
195 200 205

Pro Pro Pro Gly Trp Val Thr Arg Arg Thr Ala Asp Val Arg Phe Ala  
210 215 220

Arg Gln His Leu Leu Pro Trp Ile Gly Arg Arg Leu Thr Gly Arg Ser  
225 230 235 240

Ser Gly Asp Gly Leu Pro Ala Lys Arg Pro Asp Leu Leu Pro Tyr Glu  
245 250 255

Asp Pro Ala Arg  
260

5

ES 2 449 474 T3

<210> 11  
 <211> 454  
 <212> PRT  
 <213> Streptomyces coelicolor

5

<400> 11

```

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

Arg Ala Leu Leu Ala Ala Ile Val Thr Leu Ile Val Ala Ile Ser Ala
 20          25          30

Ala Ile Tyr Ala Gly Ala Ser Ala Asp Asp Gly Ser Arg Asp His Ala
 35          40          45

Leu Gln Ala Gly Gly Arg Leu Pro Arg Gly Asp Ala Ala Pro Ala Ser
 50          55          60

Thr Gly Ala Trp Val Gly Ala Trp Ala Thr Ala Pro Ala Ala Ala Glu
 65          70          75          80

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

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

Leu Tyr Gly Gln Ser Pro Leu Thr Val Thr His Ala Ser Ile Ala Leu
 115         120         125

Ala Ala Gly Pro Asp Thr Ala Ala Ala Ile Ala Asp Thr Met Arg Arg
 130         135         140
    
```

ES 2 449 474 T3

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

ES 2 449 474 T3

Val Val Asp Phe Asp Lys Ala Leu Arg Asp Pro Tyr Asp Pro Arg Arg  
 405 410 415

Met Arg Ser Asp Tyr Asp Ser Gly Asp His Leu His Pro Gly Asp Lys  
 420 425 430

Gly Tyr Ala Arg Met Gly Ala Val Ile Asp Leu Ala Ala Leu Lys Gly  
 435 440 445

Ala Ala Pro Val Lys Ala  
 450

<210> 12

<211> 340

<212> PRT

<213> Streptomyces coelicolor

5

<400> 12

Met Thr Ser Met Ser Arg Ala Arg Val Ala Arg Arg Ile Ala Ala Gly  
 1 5 10 15

Ala Ala Tyr Gly Gly Gly Gly Ile Gly Leu Ala Gly Ala Ala Ala Val  
 20 25 30

Gly Leu Val Val Ala Glu Val Gln Leu Ala Arg Arg Arg Val Gly Val  
 35 40 45

Gly Thr Pro Thr Arg Val Pro Asn Ala Gln Gly Leu Tyr Gly Gly Thr  
 50 55 60

Leu Pro Thr Ala Gly Asp Pro Pro Leu Arg Leu Met Met Leu Gly Asp  
 65 70 75 80

Ser Thr Ala Ala Gly Gln Gly Val His Arg Ala Gly Gln Thr Pro Gly  
 85 90 95

Ala Leu Leu Ala Ser Gly Leu Ala Ala Val Ala Glu Arg Pro Val Arg  
 100 105 110

Leu Gly Ser Val Ala Gln Pro Gly Ala Cys Ser Asp Asp Leu Asp Arg  
 115 120 125

Gln Val Ala Leu Val Leu Ala Glu Pro Asp Arg Val Pro Asp Ile Cys  
 130 135 140

Val Ile Met Val Gly Ala Asn Asp Val Thr His Arg Met Pro Ala Thr  
 145 150 155 160

Arg Ser Val Arg His Leu Ser Ser Ala Val Arg Arg Leu Arg Thr Ala  
 165 170 175

10

ES 2 449 474 T3

Gly Ala Glu Val Val Val Gly Thr Cys Pro Asp Leu Gly Thr Ile Glu  
180 185 190

Arg Val Arg Gln Pro Leu Arg Trp Leu Ala Arg Arg Ala Ser Arg Gln  
195 200 205

Leu Ala Ala Ala Gln Thr Ile Gly Ala Val Glu Gln Gly Gly Arg Thr  
210 215 220

Val Ser Leu Gly Asp Leu Leu Gly Pro Glu Phe Ala Gln Asn Pro Arg  
225 230 235 240

Glu Leu Phe Gly Pro Asp Asn Tyr His Pro Ser Ala Glu Gly Tyr Ala  
245 250 255

Thr Ala Ala Met Ala Val Leu Pro Ser Val Cys Ala Ala Leu Gly Leu  
260 265 270

Trp Pro Ala Asp Glu Glu His Pro Asp Ala Leu Arg Arg Glu Gly Phe  
275 280 285

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

Glu Val Ala Ala Ala Met Pro Thr Gly Pro Arg Gly Pro Trp Ala Leu  
305 310 315 320

Leu Lys Arg Arg Arg Arg Arg Arg Val Ser Glu Ala Glu Pro Ser Ser  
325 330 335

Pro Ser Gly Val  
340

<210> 13

<211> 305

5 <212> PRT

<213> Streptomyces coelicolor

<400> 13

Met Gly Arg Gly Thr Asp Gln Arg Thr Arg Tyr Gly Arg Arg Arg Ala  
1 5 10 15

Arg Val Ala Leu Ala Ala Leu Thr Ala Ala Val Leu Gly Val Gly Val  
20 25 30

Ala Gly Cys Asp Ser Val Gly Gly Asp Ser Pro Ala Pro Ser Gly Ser  
35 40 45

10

ES 2 449 474 T3

Pro Ser Lys Arg Thr Arg Thr Ala Pro Ala Trp Asp Thr Ser Pro Ala  
 50 55 60

Ser Val Ala Ala Val Gly Asp Ser Ile Thr Arg Gly Phe Asp Ala Cys  
 65 70 75 80

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

Ala Lys Val Asp Ser Leu Ala Val Arg Leu Leu Gly Lys Ala Asp Ala  
 100 105 110

Ala Glu His Ser Trp Asn Tyr Ala Val Thr Gly Ala Arg Met Ala Asp  
 115 120 125

Leu Thr Ala Gln Val Thr Arg Ala Ala Gln Arg Glu Pro Glu Leu Val  
 130 135 140

Ala Val Met Ala Gly Ala Asn Asp Ala Cys Arg Ser Thr Thr Ser Ala  
 145 150 155 160

Met Thr Pro Val Ala Asp Phe Arg Ala Gln Phe Glu Glu Ala Met Ala  
 165 170 175

Thr Leu Arg Lys Lys Leu Pro Lys Ala Gln Val Tyr Val Ser Ser Ile  
 180 185 190

Pro Asp Leu Lys Arg Leu Trp Ser Gln Gly Arg Thr Asn Pro Leu Gly  
 195 200 205

Lys Gln Val Trp Lys Leu Gly Leu Cys Pro Ser Met Leu Gly Asp Ala  
 210 215 220

Asp Ser Leu Asp Ser Ala Ala Thr Leu Arg Arg Asn Thr Val Arg Asp  
 225 230 235 240

Arg Val Ala Asp Tyr Asn Glu Val Leu Arg Glu Val Cys Ala Lys Asp  
 245 250 255

Arg Arg Cys Arg Ser Asp Asp Gly Ala Val His Glu Phe Arg Phe Gly  
 260 265 270

Thr Asp Gln Leu Ser His Trp Asp Trp Phe His Pro Ser Val Asp Gly  
 275 280 285

Gln Ala Arg Leu Ala Glu Ile Ala Tyr Arg Ala Val Thr Ala Lys Asn  
 290 295 300

Pro

305

- 5 <210> 14
- <211> 268
- <212> PRT
- <213> Streptomyces rimosus

- 10 <400> 14



ES 2 449 474 T3

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

Pro Ala Leu Ala Leu Phe Gly Ala Ser Ala Ala Val Ser Ala Pro Arg  
 20 25 30

Ile Gln Ala Thr Asp Tyr Val Ala Leu Gly Asp Ser Tyr Ser Ser Gly  
 35 40 45

Val Gly Ala Gly Ser Tyr Asp Ser Ser Ser Gly Ser Cys Lys Arg Ser  
 50 55 60

Thr Lys Ser Tyr Pro Ala Leu Trp Ala Ala Ser His Thr Gly Thr Arg  
 65 70 75 80

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

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

Ile Gly Gly Asn Asp Ala Gly Phe Ala Asp Thr Met Thr Thr Cys Asn  
 115 120 125

Leu Gln Gly Glu Ser Ala Cys Leu Ala Arg Ile Ala Lys Ala Arg Ala  
 130 135 140

Tyr Ile Gln Gln Thr Leu Pro Ala Gln Leu Asp Gln Val Tyr Asp Ala  
 145 150 155 160

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

Arg Phe Tyr Lys Leu Gly Gly Ser Cys Ala Val Gly Leu Ser Glu Lys  
 180 185 190

Ser Arg Ala Ala Ile Asn Ala Ala Ala Asp Asp Ile Asn Ala Val Thr  
 195 200 205

Ala Lys Arg Ala Ala Asp His Gly Phe Ala Phe Gly Asp Val Asn Thr  
 210 215 220

Thr Phe Ala Gly His Glu Leu Cys Ser Gly Ala Pro Trp Leu His Ser  
 225 230 235 240

Val Thr Leu Pro Val Glu Asn Ser Tyr His Pro Thr Ala Asn Gly Gln  
 245 250 255

Ser Lys Gly Tyr Leu Pro Val Leu Asn Ser Ala Thr  
 260 265

- 5 <210> 15
- <211> 336
- <212> PRT
- <213> Aeromonas salmonicida

ES 2 449 474 T3

<400> 15

Met Lys Lys Trp Phe Val Cys Leu Leu Gly Leu Ile Ala Leu Thr Val  
 1 5 10 15

Gln Ala Ala Asp Thr Arg Pro Ala Phe Ser Arg Ile Val Met Phe Gly  
 20 25 30

Asp Ser Leu Ser Asp Thr Gly Lys Met Tyr Ser Lys Met Arg Gly Tyr  
 35 40 45

Leu Pro Ser Ser Pro Pro Tyr Tyr Glu Gly Arg Phe Ser Asn Gly Pro  
 50 55 60

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

Asn Glu Ala Glu Gly Gly Ala Thr Ala Val Ala Tyr Asn Lys Ile Ser  
 85 90 95

Trp Asn Pro Lys Tyr Gln Val Ile Asn Asn Leu Asp Tyr Glu Val Thr  
 100 105 110

Gln Phe Leu Gln Lys Asp Ser Phe Lys Pro Asp Asp Leu Val Ile Leu  
 115 120 125

Trp Val Gly Ala Asn Asp Tyr Leu Ala Tyr Gly Trp Asn Thr Glu Gln  
 130 135 140

Asp Ala Lys Arg Val Arg Asp Ala Ile Ser Asp Ala Ala Asn Arg Met  
 145 150 155 160

Val Leu Asn Gly Ala Lys Gln Ile Leu Leu Phe Asn Leu Pro Asp Leu  
 165 170 175

Gly Gln Asn Pro Ser Ala Arg Ser Gln Lys Val Val Glu Ala Val Ser

ES 2 449 474 T3

180

185

190

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

Leu Ala Pro Thr Gly Met Val Lys Leu Phe Glu Ile Asp Lys Gln Phe  
210 215 220

Ala Glu Met Leu Arg Asp Pro Gln Asn Phe Gly Leu Ser Asp Val Glu  
225 230 235 240

Asn Pro Cys Tyr Asp Gly Gly Tyr Val Trp Lys Pro Phe Ala Thr Arg  
245 250 255

Ser Val Ser Thr Asp Arg Gln Leu Ser Ala Phe Ser Pro Gln Glu Arg  
260 265 270

Leu Ala Ile Ala Gly Asn Pro Leu Leu Ala Gln Ala Val Ala Ser Pro  
275 280 285

Met Ala Arg Arg Ser Ala Ser Pro Leu Asn Cys Glu Gly Lys Met Phe  
290 295 300

Trp Asp Gln Val His Pro Thr Thr Val Val His Ala Ala Leu Ser Glu  
305 310 315 320

Arg Ala Ala Thr Phe Ile Glu Thr Gln Tyr Glu Phe Leu Ala His Gly  
325 330 335

<210> 16

<211> 318

5 <212> PRT

<213> Aeromonas salmonicida

<400> 16

Ala Asp Thr Arg Pro Ala Phe Ser Arg Ile Val Met Phe Gly Asp Ser  
1 5 10 15

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

Ser Ser Pro Pro Tyr Tyr Glu Gly Arg Phe Ser Asn Gly Pro Val Trp  
35 40 45

Leu Glu Gln Leu Thr Lys Gln Phe Pro Gly Leu Thr Ile Ala Asn Glu  
50 55 60

10 Ala Glu Gly Gly Ala Thr Ala Val Ala Tyr Asn Lys Ile Ser Trp Asp  
65 70 75 80

ES 2 449 474 T3

Pro Lys Tyr Gln Val Ile Asn Asn Leu Asp Tyr Glu Val Thr Gln Phe  
85 90 95

Leu Gln Lys Asp Ser Phe Lys Pro Asp Asp Leu Val Ile Leu Trp Val  
100 105 110

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

Lys Arg Val Arg Asp Ala Ile Ser Asp Ala Ala Asn Arg Met Val Leu  
130 135 140

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

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

Ser Ala Tyr His Asn Lys Leu Leu Leu Asn Leu Ala Arg Gln Leu Ala  
180 185 190

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

Met Leu Arg Asp Pro Gln Asn Phe Gly Leu Ser Asp Val Glu Asn Pro  
210 215 220

Cys Tyr Asp Gly Gly Tyr Val Trp Lys Pro Phe Ala Thr Arg Ser Val  
225 230 235 240

Ser Thr Asp Arg Gln Leu Ser Ala Phe Ser Pro Gln Glu Arg Leu Ala  
245 250 255

Ile Ala Gly Asn Pro Leu Leu Ala Gln Ala Val Ala Ser Pro Met Ala  
260 265 270

Arg Arg Ser Ala Ser Pro Leu Asn Cys Glu Gly Lys Met Phe Trp Asp  
275 280 285

Gln Val His Pro Thr Thr Val Val His Ala Ala Leu Ser Glu Arg Ala  
290 295 300

Ala Thr Phe Ile Glu Thr Gln Tyr Glu Phe Leu Ala His Gly  
305 310 315

- <210> 17
- <211> 465
- <212> PRT
- <213> Candida parapsilosis
- <400> 17

5

ES 2 449 474 T3

Met Arg Tyr Phe Ala Ile Ala Phe Leu Leu Ile Asn Thr Ile Ser Ala  
 1 5 10 15

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

Pro Gln Gly Tyr Glu Ala Gln Pro Leu Gly Ser Ile Leu Lys Thr Arg  
 35 40 45

Asn Val Pro Asn Pro Leu Thr Asn Val Phe Thr Pro Val Lys Val Gln  
 50 55 60

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

Asn Ala Ile Val Thr Thr Ile Ile Gln Pro Phe Asn Ala Lys Lys Asp  
 85 90 95

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

Ala Pro Ser Tyr Ala Ile Gln Tyr Gly Ser Asp Ile Ser Thr Leu Thr  
 115 120 125

Thr Gln Gly Glu Met Tyr Tyr Ile Ser Ala Leu Leu Asp Gln Gly Tyr  
 130 135 140

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

Gly Leu Gln Ser Gly Arg Ala Thr Leu Asn Ser Leu Arg Ala Thr Leu  
 165 170 175

Lys Ser Gly Asn Leu Thr Gly Val Ser Ser Asp Ala Glu Thr Leu Leu  
 180 185 190

Trp Gly Tyr Ser Gly Gly Ser Leu Ala Ser Gly Trp Ala Ala Ala Ile  
 195 200 205

Gln Lys Glu Tyr Ala Pro Glu Leu Ser Lys Asn Leu Leu Gly Ala Ala  
 210 215 220

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

Ser Gly Pro Phe Ala Gly Ile Ile Ser Asn Ala Leu Ala Gly Ile Gly  
 245 250 255

ES 2 449 474 T3

Asn Glu Tyr Pro Asp Phe Lys Asn Tyr Leu Leu Lys Lys Val Ser Pro  
 260 265 270

Leu Leu Ser Ile Thr Tyr Arg Leu Gly Asn Thr His Cys Leu Leu Asp  
 275 280 285

Gly Gly Ile Ala Tyr Phe Gly Lys Ser Phe Phe Ser Arg Ile Ile Arg  
 290 295 300

Tyr Phe Pro Asp Gly Trp Asp Leu Val Asn Gln Glu Pro Ile Lys Thr  
 305 310 315 320

Ile Leu Gln Asp Asn Gly Leu Val Tyr Gln Pro Lys Asp Leu Thr Pro  
 325 330 335

Gln Ile Pro Leu Phe Ile Tyr His Gly Thr Leu Asp Ala Ile Val Pro  
 340 345 350

Ile Val Asn Ser Arg Lys Thr Phe Gln Gln Trp Cys Asp Trp Gly Leu  
 355 360 365

Lys Ser Gly Glu Tyr Asn Glu Asp Leu Thr Asn Gly His Ile Thr Glu  
 370 375 380

Ser Ile Val Gly Ala Pro Ala Ala Leu Thr Trp Ile Ile Asn Arg Phe  
 385 390 395 400

Asn Gly Gln Pro Pro Val Asp Gly Cys Gln His Asn Val Arg Ala Ser  
 405 410 415

Asn Leu Glu Tyr Pro Gly Thr Pro Gln Ser Ile Lys Asn Tyr Phe Glu  
 420 425 430

Ala Ala Leu His Ala Ile Leu Gly Phe Asp Leu Gly Pro Asp Val Lys  
 435 440 445

Arg Asp Lys Val Thr Leu Gly Gly Leu Leu Lys Leu Glu Arg Phe Ala  
 450 455 460

Phe  
 465

<210> 18

<211> 471

5 <212> PRT

<213> Candida parapsilosis

<400> 18

ES 2 449 474 T3

Met Arg Tyr Phe Ala Ile Ala Phe Leu Leu Ile Asn Thr Ile Ser Ala  
 1 5 10 15

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

Pro Gln Gly Tyr Glu Ala Gln Pro Leu Gly Ser Ile Leu Lys Thr Arg  
 35 40 45

Asn Val Pro Asn Pro Leu Thr Asn Val Phe Thr Pro Val Lys Val Gln  
 50 55 60

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

Asn Ala Ile Val Thr Thr Ile Ile Gln Pro Phe Asn Ala Lys Lys Asp  
 85 90 95

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

Ala Pro Ser Tyr Ala Ile Gln Tyr Gly Ser Asp Ile Ser Thr Leu Thr  
 115 120 125

Thr Gln Gly Glu Met Tyr Tyr Ile Ser Ala Leu Leu Asp Gln Gly Tyr  
 130 135 140

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

Gly Leu Gln Ser Gly Arg Ala Thr Leu Asn Ser Leu Arg Ala Thr Leu  
 165 170 175

Lys Ser Gly Asn Leu Thr Gly Val Ser Ser Asp Ala Glu Thr Leu Leu  
 180 185 190

Trp Gly Tyr Ser Gly Gly Ser Leu Ala Ser Gly Trp Ala Ala Ala Ile  
 195 200 205

Gln Lys Glu Tyr Ala Pro Glu Leu Ser Lys Asn Leu Leu Gly Ala Ala  
 210 215 220

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

Ser Gly Pro Phe Ala Gly Ile Ile Ser Asn Ala Leu Ala Gly Ile Gly  
 245 250 255

Asn Glu Tyr Pro Asp Phe Lys Asn Tyr Leu Leu Lys Lys Val Ser Pro

ES 2 449 474 T3

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

<210> 19

<211> 261

5 <212> PRT

<213> Streptomyces coelicolor

<400> 19

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



ES 2 449 474 T3

Gly Asp Pro Gly Pro Asp Gly Ala Phe Val Gly Trp Ala Asp Arg Leu  
 20 25 30

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

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

Val Pro Arg Val Val Gly Leu Ala Pro Asp Leu Val Ser Phe Ala Ala  
 65 70 75 80

Gly Gly Asn Asp Ile Ile Arg Pro Gly Thr Asp Pro Asp Glu Val Ala  
 85 90 95

Glu Arg Phe Glu Leu Ala Val Ala Ala Leu Thr Ala Ala Ala Gly Thr  
 100 105 110

Val Leu Val Thr Thr Gly Phe Asp Thr Arg Gly Val Pro Val Leu Lys  
 115 120 125

His Leu Arg Gly Lys Ile Ala Thr Tyr Asn Gly His Val Arg Ala Ile  
 130 135 140

Ala Asp Arg Tyr Gly Cys Pro Val Leu Asp Leu Trp Ser Leu Arg Ser  
 145 150 155 160

Val Gln Asp Arg Arg Ala Trp Asp Ala Asp Arg Leu His Leu Ser Pro  
 165 170 175

Glu Gly His Thr Arg Val Ala Leu Arg Ala Gly Gln Ala Leu Gly Leu  
 180 185 190

Arg Val Pro Ala Asp Pro Asp Gln Pro Trp Pro Pro Leu Pro Pro Arg  
 195 200 205

Gly Thr Leu Asp Val Arg Arg Asp Asp Val His Trp Ala Arg Glu Tyr  
 210 215 220

Leu Val Pro Trp Ile Gly Arg Arg Leu Arg Gly Glu Ser Ser Gly Asp  
 225 230 235 240

His Val Thr Ala Lys Gly Thr Leu Ser Pro Asp Ala Ile Lys Thr Arg  
 245 250 255

Ile Ala Ala Val Ala  
 260

<210> 20  
 <211> 230  
 <212> PRT  
 <213> Aspergillus aculeatus  
 <400> 20

5

ES 2 449 474 T3

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

Gly Ser Gly Thr Asn Gly Trp Gly Glu Tyr Leu Ala Ser Tyr Leu Ser  
20 25 30

Ala Thr Val Val Asn Asp Ala Val Ala Gly Arg Ser Ala Arg Ser Tyr  
35 40 45

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

Asp Tyr Val Ile Val Glu Phe Gly His Asn Asp Gly Gly Ser Leu Ser  
65 70 75 80

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

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

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

Ile Leu Ser Ser Gln Thr Pro Asn Asn Pro Trp Glu Thr Gly Thr Phe  
130 135 140

Val Asn Ser Pro Thr Arg Phe Val Glu Tyr Ala Glu Leu Ala Ala Glu  
145 150 155 160

Val Ala Gly Val Glu Tyr Val Asp His Trp Ser Tyr Val Asp Ser Ile  
165 170 175

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

His Thr His Thr Ser Pro Ala Gly Ala Glu Val Val Ala Glu Ala Phe  
195 200 205

Leu Lys Ala Val Val Cys Thr Gly Thr Ser Leu Lys Ser Val Leu Thr  
210 215 220

Thr Thr Ser Phe Glu Gly  
225 230

<210> 21

<211> 18

5 <212> PRT

<213> Aeromonas sp.

<400> 21

Met Lys Lys Trp Phe Val Cys Leu Leu Gly Leu Ile Ala Leu Thr Val  
1 5 10 15

10 Gln Ala

ES 2 449 474 T3

<210> 22  
 <211> 29  
 <212> PRT  
 <213> Bacillus subtilis

5

<400> 22

Met Arg Ser Lys Lys Leu Trp Ile Ser Leu Leu Phe Ala Leu Thr Leu  
 1 5 10 15

Ile Phe Thr Met Ala Phe Ser Asn Met Ser Ala Gln Ala  
 20 25

10

<210> 23  
 <211> 29  
 <212> PRT  
 <213> Bacillus licheniformis

15

<400> 23

Met Met Arg Lys Lys Ser Phe Trp Phe Gly Met Leu Thr Ala Phe Met  
 1 5 10 15

Leu Val Phe Thr Met Glu Phe Ser Asp Ser Ala Ser Ala  
 20 25

20

<210> 24  
 <211> 1047  
 <212> ADN  
 <213> Aeromonas hydrophila

25

<400> 24

atgtttaagt ttaaaaagaa tttcttagtt ggattatcgg cagctttaat gagtattagc 60  
 ttgttttcgg caaccgcctc tgcagctagc gccgacagcc gtcccgcctt ttcccgatc 120  
 gtgatgttcg gcgacagcct ctccgatacc ggcaaaatgt acagcaagat gcgcggttac 180  
 ctcccctcca gcccgcccta ctatgagggc cgtttctcca acggaccctg ctggctggag 240  
 cagctgacca aacagttccc gggctgacc atcgccaacg aagcggagg cggtgccact 300  
 gccgtggctt acaacaagat ctctggaat cccaagtatc aggtcatcaa caacctggac 360

ES 2 449 474 T3

tacgaggtca cccagttctt gcagaaagac agcttcaagc cggacgatct ggtgatcctc 420  
 tgggtcgggtg ccaatgacta tctggcctat ggctggaaca cggagcagga tgccaagcgg 480  
 gttcgcgatg ccatcagcga tgccggcaac cgcgatgtac tgaacgggtgc caagcagata 540  
 ctgctgttca acctgccgga tctgggccag aaccctcag ctgcagctca gaaggtggtc 600  
 gaggcgggtca gccatgtctc cgcctatcac aaccagctgc tgctgaacct ggcacgccag 660  
 ctggcccca cggcatggt aaagctgttc gagatcgaca agcaatttgc cgagatgctg 720  
 cgtgatccgc agaacttcgg cctgagcgc gtcgagaacc cctgctacga cggcggctat 780  
 gtgtggaagc cgtttgccac ccgcagcgtc agcaccgacc gccagctctc cgccttcagt 840  
 ccgcaggaac gcctcggcat cgcggcaac ccgctgctgg cacaggcctg tgccagctct 900  
 atggcccggc gcagcggcag cccctcaac tgtgagggca agatgttctg ggatcaggtg 960  
 cacccgacca ctgtcgtgca cgcagccctg agcgagcggc ccgccacctt catcgcgaac 1020  
 cagtacgagt tcctcgcca ctgatga 1047

<210> 25  
 <211> 347  
 <212> PRT  
 <213> Artificial

<220>  
 <223> Constructo de fusión

<400> 25

Met Phe Lys Phe Lys Lys Asn Phe Leu Val Gly Leu Ser Ala Ala Leu  
 1 5 10 15

Met Ser Ile Ser Leu Phe Ser Ala Thr Ala Ser Ala Ala Ser Ala Asp  
 20 25 30

Ser Arg Pro Ala Phe Ser Arg Ile Val Met Phe Gly Asp Ser Leu Ser  
 35 40 45

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

Pro Pro Tyr Tyr Glu Gly Arg Phe Ser Asn Gly Pro Val Trp Leu Glu  
 65 70 75 80

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

Gly Gly Ala Thr Ala Val Ala Tyr Asn Lys Ile Ser Trp Asn Pro Lys  
 100 105 110

Tyr Gln Val Ile Asn Asn Leu Asp Tyr Glu Val Thr Gln Phe Leu Gln  
 115 120 125

ES 2 449 474 T3

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

Asn Asp Tyr Leu Ala Tyr Gly Trp Asn Thr Glu Gln Asp Ala Lys Arg  
145 150 155 160

Val Arg Asp Ala Ile Ser Asp Ala Ala Asn Arg Met Val Leu Asn Gly  
165 170 175

Ala Lys Gln Ile Leu Leu Phe Asn Leu Pro Asp Leu Gly Gln Asn Pro  
180 185 190

Ser Ala Arg Ser Gln Lys Val Val Glu Ala Val Ser His Val Ser Ala  
195 200 205

Tyr His Asn Gln Leu Leu Leu Asn Leu Ala Arg Gln Leu Ala Pro Thr  
210 215 220

Gly Met Val Lys Leu Phe Glu Ile Asp Lys Gln Phe Ala Glu Met Leu  
225 230 235 240

Arg Asp Pro Gln Asn Phe Gly Leu Ser Asp Val Glu Asn Pro Cys Tyr  
245 250 255

Asp Gly Gly Tyr Val Trp Lys Pro Phe Ala Thr Arg Ser Val Ser Thr  
260 265 270

Asp Arg Gln Leu Ser Ala Phe Ser Pro Gln Glu Arg Leu Ala Ile Ala  
275 280 285

Gly Asn Pro Leu Leu Ala Gln Ala Val Ala Ser Pro Met Ala Arg Arg  
290 295 300

Ser Ala Ser Pro Leu Asn Cys Glu Gly Lys Met Phe Trp Asp Gln Val  
305 310 315 320

His Pro Thr Thr Val Val His Ala Ala Leu Ser Glu Arg Ala Ala Thr  
325 330 335

Phe Ile Ala Asn Gln Tyr Glu Phe Leu Ala His  
340 345

<210> 26  
<211> 267  
<212> PRT  
<213> Streptomyces sp.

5

<400> 26

ES 2 449 474 T3

Met Arg Leu Thr Arg Ser Leu Ser Ala Ala Ser Val Ile Val Phe Ala  
 1 5 10 15

Leu Leu Leu Ala Leu Leu Gly Ile Ser Pro Ala Gln Ala Ala Gly Pro  
 20 25 30

Ala Tyr Val Ala Leu Gly Asp Ser Tyr Ser Ser Gly Asn Gly Ala Gly  
 35 40 45

Ser Tyr Ile Asp Ser Ser Gly Asp Cys His Arg Ser Asn Asn Ala Tyr  
 50 55 60

Pro Ala Arg Trp Ala Ala Ala Asn Ala Pro Ser Ser Phe Thr Phe Ala  
 65 70 75 80

Ala Cys Ser Gly Ala Val Thr Thr Asp Val Ile Asn Asn Gln Leu Gly  
 85 90 95

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

Asp Ala Gly Phe Ala Asp Ala Met Thr Thr Cys Val Thr Ser Ser Asp  
 115 120 125

Ser Thr Cys Leu Asn Arg Leu Ala Thr Ala Thr Asn Tyr Ile Asn Thr  
 130 135 140

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

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

Ala Ser Asn Pro Trp Tyr Cys Leu Gly Leu Ser Asn Thr Lys Arg Ala  
 180 185 190

Ala Ile Asn Thr Thr Ala Asp Thr Leu Asn Ser Val Ile Ser Ser Arg  
 195 200 205

Ala Thr Ala His Gly Phe Arg Phe Gly Asp Val Arg Pro Thr Phe Asn  
 210 215 220

Asn His Glu Leu Phe Phe Gly Asn Asp Trp Leu His Ser Leu Thr Leu  
 225 230 235 240

Pro Val Trp Glu Ser Tyr His Pro Thr Ser Thr Gly His Gln Ser Gly  
 245 250 255

Tyr Leu Pro Val Leu Asn Ala Asn Ser Ser Thr  
 260 265

5 <210> 27  
 <211> 548  
 <212> PRT  
 <213> Thermobifida sp.

10 <400> 27

ES 2 449 474 T3

Met Leu Pro His Pro Ala Gly Glu Arg Gly Glu Val Gly Ala Phe Phe  
 1 5 10 15

Ala Leu Leu Val Gly Thr Pro Gln Asp Arg Arg Leu Arg Leu Glu Cys  
 20 25 30

His Glu Thr Arg Pro Leu Arg Gly Arg Cys Gly Cys Gly Glu Arg Arg  
 35 40 45

Val Pro Pro Leu Thr Leu Pro Gly Asp Gly Val Leu Cys Thr Thr Ser  
 50 55 60

Ser Thr Arg Asp Ala Glu Thr Val Trp Arg Lys His Leu Gln Pro Arg  
 65 70 75 80

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

Gly Gln Gly Ser Pro Gly Val Leu Trp Cys Gly Arg Glu Gly Cys Arg  
 100 105 110

Phe Glu Val Cys Arg Arg Asp Thr Pro Gly Leu Ser Arg Thr Arg Asn  
 115 120 125

Gly Asp Ser Ser Pro Pro Phe Arg Ala Gly Trp Ser Leu Pro Pro Lys  
 130 135 140

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

Tyr Ser Leu Leu Arg Thr Asp Arg Pro Asp Gly Pro Arg Gly Arg Phe  
 165 170 175

Val Gly Ser Gly Pro Arg Ala Ala Thr Arg Arg Arg Leu Phe Leu Gly  
 180 185 190

Ile Pro Ala Leu Val Leu Val Thr Ala Leu Thr Leu Val Leu Ala Val  
 195 200 205

Pro Thr Gly Arg Glu Thr Leu Trp Arg Met Trp Cys Glu Ala Thr Gln  
 210 215 220

ES 2 449 474 T3

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



ES 2 449 474 T3

Pro Trp Val Asn Gly Val Gln Leu Arg Asp Leu Ala Thr Gly Val Thr  
 485 490 495

Val Asp Arg Ser Thr Phe His Pro Asn Ala Ala Gly His Arg Ala Val  
 500 505 510

Gly Glu Arg Val Ile Glu Gln Ile Glu Thr Gly Pro Gly Arg Pro Leu  
 515 520 525

Tyr Ala Thr Phe Ala Val Val Ala Gly Ala Thr Val Asp Thr Leu Ala  
 530 535 540

Gly Glu Val Gly  
 545

<210> 28

<211> 372

5 <212> PRT

<213> Thermobifida sp.

<400> 28

Met Gly Ser Gly Pro Arg Ala Ala Thr Arg Arg Arg Leu Phe Leu Gly  
 1 5 10 15

Ile Pro Ala Leu Val Leu Val Thr Ala Leu Thr Leu Val Leu Ala Val  
 20 25 30

Pro Thr Gly Arg Glu Thr Leu Trp Arg Met Trp Cys Glu Ala Thr Gln  
 35 40 45

Asp Trp Cys Leu Gly Val Pro Val Asp Ser Arg Gly Gln Pro Ala Glu  
 50 55 60

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

Asn Tyr Tyr Ala Leu Gly Asp Ser Tyr Ser Ser Gly Asp Gly Ala Arg  
 85 90 95

Asp Tyr Tyr Pro Gly Thr Ala Val Lys Gly Gly Cys Trp Arg Ser Ala  
 100 105 110

Asn Ala Tyr Pro Glu Leu Val Ala Glu Ala Tyr Asp Phe Ala Gly His  
 115 120 125

Leu Ser Phe Leu Ala Cys Ser Gly Gln Arg Gly Tyr Ala Met Leu Asp  
 130 135 140

10 Ala Ile Asp Glu Val Gly Ser Gln Leu Asp Trp Asn Ser Pro His Thr  
 145 150 155 160

ES 2 449 474 T3

Ser Leu Val Thr Ile Gly Ile Gly Gly Asn Asp Leu Gly Phe Ser Thr  
 165 170 175

Val Leu Lys Thr Cys Met Val Arg Val Pro Leu Leu Asp Ser Lys Ala  
 180 185 190

Cys Thr Asp Gln Glu Asp Ala Ile Arg Lys Arg Met Ala Lys Phe Glu  
 195 200 205

Thr Thr Phe Glu Glu Leu Ile Ser Glu Val Arg Thr Arg Ala Pro Asp  
 210 215 220

Ala Arg Ile Leu Val Val Gly Tyr Pro Arg Ile Phe Pro Glu Glu Pro  
 225 230 235 240

Thr Gly Ala Tyr Tyr Thr Leu Thr Ala Ser Asn Gln Arg Trp Leu Asn  
 245 250 255

Glu Thr Ile Gln Glu Phe Asn Gln Gln Leu Ala Glu Ala Val Ala Val  
 260 265 270

His Asp Glu Glu Ile Ala Ala Ser Gly Gly Val Gly Ser Val Glu Phe  
 275 280 285

Val Asp Val Tyr His Ala Leu Asp Gly His Glu Ile Gly Ser Asp Glu  
 290 295 300

Pro Trp Val Asn Gly Val Gln Leu Arg Asp Leu Ala Thr Gly Val Thr  
 305 310 315 320

Val Asp Arg Ser Thr Phe His Pro Asn Ala Ala Gly His Arg Ala Val  
 325 330 335

Gly Glu Arg Val Ile Glu Gln Ile Glu Thr Gly Pro Gly Arg Pro Leu  
 340 345 350

Tyr Ala Thr Phe Ala Val Val Ala Gly Ala Thr Val Asp Thr Leu Ala  
 355 360 365

Gly Glu Val Gly  
 370

<210> 29  
 <211> 300  
 <212> PRT  
 <213> Corynebacterium efficiens

<400> 29

5

ES 2 449 474 T3

Met Arg Thr Thr Val Ile Ala Ala Ser Ala Leu Leu Leu Leu Ala Gly  
 1 5 10 15

Cys Ala Asp Gly Ala Arg Glu Glu Thr Ala Gly Ala Pro Pro Gly Glu  
 20 25 30

Ser Ser Gly Gly Ile Arg Glu Glu Gly Ala Glu Ala Ser Thr Ser Ile  
 35 40 45

Thr Asp Val Tyr Ile Ala Leu Gly Asp Ser Tyr Ala Ala Met Gly Gly  
 50 55 60

Arg Asp Gln Pro Leu Arg Gly Glu Pro Phe Cys Leu Arg Ser Ser Gly  
 65 70 75 80

Asn Tyr Pro Glu Leu Leu His Ala Glu Val Thr Asp Leu Thr Cys Gln  
 85 90 95

Gly Ala Val Thr Gly Asp Leu Leu Glu Pro Arg Thr Leu Gly Glu Arg  
 100 105 110

Thr Leu Pro Ala Gln Val Asp Ala Leu Thr Glu Asp Thr Thr Leu Val  
 115 120 125

Thr Leu Ser Ile Gly Gly Asn Asp Leu Gly Phe Gly Glu Val Ala Gly  
 130 135 140

Cys Ile Arg Glu Arg Ile Ala Gly Glu Asn Ala Asp Asp Cys Val Asp  
 145 150 155 160

Leu Leu Gly Glu Thr Ile Gly Glu Gln Leu Asp Gln Leu Pro Pro Gln  
 165 170 175

Leu Asp Arg Val His Glu Ala Ile Arg Asp Arg Ala Gly Asp Ala Gln  
 180 185 190

Val Val Val Thr Gly Tyr Leu Pro Leu Val Ser Ala Gly Asp Cys Pro  
 195 200 205

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

Thr Gly Gln Ile Asn Glu Thr Val Arg Glu Ala Ala Glu Arg His Asp  
 225 230 235 240

Ala Leu Phe Val Leu Pro Asp Asp Ala Asp Glu His Thr Ser Cys Ala  
 245 250 255

Pro Pro Gln Gln Arg Trp Ala Asp Ile Gln Gly Gln Gln Thr Asp Ala  
 260 265 270

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

Val Arg Asp Ala Leu Gly Leu Glu Pro Val Gln Pro  
 290 295 300

ES 2 449 474 T3

<210> 30  
 <211> 284  
 <212> PRT  
 <213> Novosphingobium aromaticivorans

5

<400> 30

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

Ala Leu Ala Gly Leu Ala Pro Ala Ala Thr Val Ala Arg Glu Ala Pro  
 20 25 30

Leu Ala Glu Gly Ala Arg Tyr Val Ala Leu Gly Ser Ser Phe Ala Ala  
 35 40 45

Gly Pro Gly Val Gly Pro Asn Ala Pro Gly Ser Pro Glu Arg Cys Gly  
 50 55 60

Arg Gly Thr Leu Asn Tyr Pro His Leu Leu Ala Glu Ala Leu Lys Leu  
 65 70 75 80

Asp Leu Val Asp Ala Thr Cys Ser Gly Ala Thr Thr His His Val Leu  
 85 90 95

Gly Pro Trp Asn Glu Val Pro Pro Gln Ile Asp Ser Val Asn Gly Asp  
 100 105 110

Thr Arg Leu Val Thr Leu Thr Ile Gly Gly Asn Asp Val Ser Phe Val  
 115 120 125

Gly Asn Ile Phe Ala Ala Ala Cys Glu Lys Met Ala Ser Pro Asp Pro  
 130 135 140

Arg Cys Gly Lys Trp Arg Glu Ile Thr Glu Glu Glu Trp Gln Ala Asp  
 145 150 155 160

Glu Glu Arg Met Arg Ser Ile Val Arg Gln Ile His Ala Arg Ala Pro  
 165 170 175

Leu Ala Arg Val Val Val Val Asp Tyr Ile Thr Val Leu Pro Pro Ser  
 180 185 190

ES 2 449 474 T3

Gly Thr Cys Ala Ala Met Ala Ile Ser Pro Asp Arg Leu Ala Gln Ser  
 195 200 205

Arg Ser Ala Ala Lys Arg Leu Ala Arg Ile Thr Ala Arg Val Ala Arg  
 210 215 220

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

His Pro Cys Ser Ala Lys Pro Trp Ser Asn Gly Leu Ser Ala Pro Ala  
 245 250 255

Asp Asp Gly Ile Pro Val His Pro Asn Arg Leu Gly His Ala Glu Ala  
 260 265 270

Ala Ala Ala Leu Val Lys Leu Val Lys Leu Met Lys  
 275 280

<210> 31

<211> 268

5 <212> PRT

<213> Streptomyces coelicolor

<400> 31

Met Arg Arg Phe Arg Leu Val Gly Phe Leu Ser Ser Leu Val Leu Ala  
 1 5 10 15

Ala Gly Ala Ala Leu Thr Gly Ala Ala Thr Ala Gln Ala Ala Gln Pro  
 20 25 30

Ala Ala Ala Asp Gly Tyr Val Ala Leu Gly Asp Ser Tyr Ser Ser Gly  
 35 40 45

Val Gly Ala Gly Ser Tyr Ile Ser Ser Ser Gly Asp Cys Lys Arg Ser  
 50 55 60

Thr Lys Ala His Pro Tyr Leu Trp Ala Ala Ala His Ser Pro Ser Thr  
 65 70 75 80

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

Gly Gln Leu Gly Pro Leu Ser Ser Gly Thr Gly Leu Val Ser Ile Ser  
 100 105 110

Ile Gly Gly Asn Asp Ala Gly Phe Ala Asp Thr Met Thr Thr Cys Val  
 115 120 125

10 Leu Gln Ser Glu Ser Ser Cys Leu Ser Arg Ile Ala Thr Ala Glu Ala

ES 2 449 474 T3

130

135

140

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

Ile Ser Asp Lys Ala Pro Asn Ala His Val Val Val Ile Gly Tyr Pro  
165 170 175

Arg Phe Tyr Lys Leu Gly Thr Thr Cys Ile Gly Leu Ser Glu Thr Lys  
180 185 190

Arg Thr Ala Ile Asn Lys Ala Ser Asp His Leu Asn Thr Val Leu Ala  
195 200 205

Gln Arg Ala Ala Ala His Gly Phe Thr Phe Gly Asp Val Arg Thr Thr  
210 215 220

Phe Thr Gly His Glu Leu Cys Ser Gly Ser Pro Trp Leu His Ser Val  
225 230 235 240

Asn Trp Leu Asn Ile Gly Glu Ser Tyr His Pro Thr Ala Ala Gly Gln  
245 250 255

Ser Gly Gly Tyr Leu Pro Val Leu Asn Gly Ala Ala  
260 265

<210> 32

<211> 269

5 <212> PRT

<213> Streptomyces avermitilis

<400> 32

Met Arg Arg Ser Arg Ile Thr Ala Tyr Val Thr Ser Leu Leu Leu Ala  
1 5 10 15

Val Gly Cys Ala Leu Thr Gly Ala Ala Thr Ala Gln Ala Ser Pro Ala  
20 25 30

Ala Ala Ala Thr Gly Tyr Val Ala Leu Gly Asp Ser Tyr Ser Ser Gly  
35 40 45

Val Gly Ala Gly Ser Tyr Leu Ser Ser Ser Gly Asp Cys Lys Arg Ser  
50 55 60

Ser Lys Ala Tyr Pro Tyr Leu Trp Gln Ala Ala His Ser Pro Ser Ser  
65 70 75 80

Phe Ser Phe Met Ala Cys Ser Gly Ala Arg Thr Gly Asp Val Leu Ala  
85 90 95

10

ES 2 449 474 T3

Asn Gln Leu Gly Thr Leu Asn Ser Ser Thr Gly Leu Val Ser Leu Thr  
 100 105 110

Ile Gly Gly Asn Asp Ala Gly Phe Ser Asp Val Met Thr Thr Cys Val  
 115 120 125

Leu Gln Ser Asp Ser Ala Cys Leu Ser Arg Ile Asn Thr Ala Lys Ala  
 130 135 140

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

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

Arg Phe Tyr Lys Leu Gly Gly Ser Cys Leu Ala Gly Leu Ser Glu Thr  
 180 185 190

Lys Arg Ser Ala Ile Asn Asp Ala Ala Asp Tyr Leu Asn Ser Ala Ile  
 195 200 205

Ala Lys Arg Ala Ala Asp His Gly Phe Thr Phe Gly Asp Val Lys Ser  
 210 215 220

Thr Phe Thr Gly His Glu Ile Cys Ser Ser Ser Thr Trp Leu His Ser  
 225 230 235 240

Leu Asp Leu Leu Asn Ile Gly Gln Ser Tyr His Pro Thr Ala Ala Gly  
 245 250 255

Gln Ser Gly Gly Tyr Leu Pro Val Met Asn Ser Val Ala  
 260 265

<210> 33

<211> 267

5 <212> PRT

<213> Streptomyces sp.

<400> 33

Met Arg Leu Thr Arg Ser Leu Ser Ala Ala Ser Val Ile Val Phe Ala  
 1 5 10 15

Leu Leu Leu Ala Leu Leu Gly Ile Ser Pro Ala Gln Ala Ala Gly Pro  
 20 25 30

Ala Tyr Val Ala Leu Gly Asp Ser Tyr Ser Ser Gly Asn Gly Ala Gly  
 35 40 45

10 Ser Tyr Ile Asp Ser Ser Gly Asp Cys His Arg Ser Asn Asn Ala Tyr

ES 2 449 474 T3

50

55

60

Pro Ala Arg Trp Ala Ala Ala Asn Ala Pro Ser Ser Phe Thr Phe Ala  
65 70 75 80

Ala Cys Ser Gly Ala Val Thr Thr Asp Val Ile Asn Asn Gln Leu Gly  
85 90 95

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

Asp Ala Gly Phe Ala Asp Ala Met Thr Thr Cys Val Thr Ser Ser Asp  
115 120 125

Ser Thr Cys Leu Asn Arg Leu Ala Thr Ala Thr Asn Tyr Ile Asn Thr  
130 135 140

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

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

Ala Ser Asn Pro Trp Tyr Cys Leu Gly Leu Ser Asn Thr Lys Arg Ala  
180 185 190

Ala Ile Asn Thr Thr Ala Asp Thr Leu Asn Ser Val Ile Ser Ser Arg  
195 200 205

Ala Thr Ala His Gly Phe Arg Phe Gly Asp Val Arg Pro Thr Phe Asn  
210 215 220

Asn His Glu Leu Phe Phe Gly Asn Asp Trp Leu His Ser Leu Thr Leu  
225 230 235 240

Pro Val Trp Glu Ser Tyr His Pro Thr Ser Thr Gly His Gln Ser Gly  
245 250 255

Tyr Leu Pro Val Leu Asn Ala Asn Ser Ser Thr  
260 265

<210> 34

<211> 317

5 <212> PRT

<213> Aeromonas hydrophila

<400> 34

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



ES 2 449 474 T3

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

ES 2 449 474 T3

Ala Arg Ser Ala Ser Thr Leu Asn Cys Glu Gly Lys Met Phe Trp Asp  
 275 280 285

Gln Val His Pro Thr Thr Val Val His Ala Ala Leu Ser Glu Pro Ala  
 290 295 300

Ala Thr Phe Ile Glu Ser Gln Tyr Glu Phe Leu Ala His  
 305 310 315

<210> 35

<211> 318

<212> PRT

<213> Aeromonas salmonicida

<400> 35

Ala Asp Thr Arg Pro Ala Phe Ser Arg Ile Val Met Phe Gly Asp Ser  
 1 5 10 15

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

Ser Ser Pro Pro Tyr Tyr Glu Gly Arg Phe Ser Asn Gly Pro Val Trp  
 35 40 45

Leu Glu Gln Leu Thr Lys Gln Phe Pro Gly Leu Thr Ile Ala Asn Glu  
 50 55 60

Ala Glu Gly Gly Ala Thr Ala Val Ala Tyr Asn Lys Ile Ser Trp Asn  
 65 70 75 80

Pro Lys Tyr Gln Val Ile Asn Asn Leu Asp Tyr Glu Val Thr Gln Phe  
 85 90 95

Leu Gln Lys Asp Ser Phe Lys Pro Asp Asp Leu Val Ile Leu Trp Val  
 100 105 110

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

Lys Arg Val Arg Asp Ala Ile Ser Asp Ala Ala Asn Arg Met Val Leu  
 130 135 140

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

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

Ser Ala Tyr His Asn Lys Leu Leu Leu Asn Leu Ala Arg Gln Leu Ala  
 180 185 190

5

10

ES 2 449 474 T3

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

Met Leu Arg Asp Pro Gln Asn Phe Gly Leu Ser Asp Val Glu Asn Pro  
 210 215 220

Cys Tyr Asp Gly Gly Tyr Val Trp Lys Pro Phe Ala Thr Arg Ser Val  
 225 230 235 240

Ser Thr Asp Arg Gln Leu Ser Ala Phe Ser Pro Gln Glu Arg Leu Ala  
 245 250 255

Ile Ala Gly Asn Pro Leu Leu Ala Gln Ala Val Ala Ser Pro Met Ala  
 260 265 270

Arg Arg Ser Ala Ser Pro Leu Asn Cys Glu Gly Lys Met Phe Trp Asp  
 275 280 285

Gln Val His Pro Thr Thr Val Val His Ala Ala Leu Ser Glu Arg Ala  
 290 295 300

Ala Thr Phe Ile Glu Thr Gln Tyr Glu Phe Leu Ala His Gly  
 305 310 315

- <210> 36
- <211> 1371
- <212> ADN
- <213> Streptomyces thermosacchari

<400> 36

acagggccgat gcacggaacc gtacctttcc gcagtgaagc gctctcccc catcgttcgc	60
cgggacttca tccgcgattt tggcatgaac acttccttca acgcgcgtag cttgctacaa	120
gtgcggcagc agaccgcctc gttggaggct cagtgagatt gaccgatcc ctgtcggccc	180
catccgtcat cgtcttcgcc ctgctgctcg cgctgctggg catcagcccg gccagggcag	240
ccggcccggc ctatgtggcc ctgggggatt cctattcctc gggcaacggc gccggaagt	300
acatcgattc gagcgggtgac tgtcaccgca gcaacaacgc gtaccccgcc cgctgggcgg	360
cggccaacgc accgtcctcc ttcaccttcg cggcctgctc gggagcggtg accacggatg	420
tgatcaacaa tcagctgggc gccctcaacg cgtccaccgg cctggtgagc atcaccatcg	480
gcggcaatga cgcgggcttc gcggacgcga tgaccacctg cgtcaccagc tcggacagca	540
cctgcctcaa ccggctggcc accgccacca actacatcaa caccacctg ctgcgccggc	600
tcgacgcggt ctacagccag atcaaggccc gtgcccccaa cgccgcgctg gtcgtcctcg	660
gctaccccgcg catgtacctg gcctcgaacc cctggtactg cctgggcctg agcaacacca	720
agcgcgcggc catcaacacc accgccgaca ccctcaactc ggtgatctcc tcccgggcca	780

5

10

ES 2 449 474 T3

ccgcccacgg attccgattc ggcgatgtcc gcccgacctt caacaaccac gaactgttct 840  
 tcggcaacga ctggctgcac tcaactaccc tgccggtgtg ggagtcgtac caccaccca 900  
 gcacgggcca tcagagcggc tatctgccgg tcctcaacgc caacagctcg acctgatcaa 960  
 cgcacggccg tgcccgcccc gcgcgtcacg ctcggcggcg gcgcccgcgc gcgttgatca 1020  
 gcccacagtg ccggtgacgg tcccaccgtc acggtcgagg gtgtacgtca cggtggcgcc 1080  
 gctccagaag tggaacgtca gcaggaccgt ggagccgtcc ctgacctcgt cgaagaactc 1140  
 cggggtcagc gtgatcacc ccccccgta gccggggggc aaggcggcgc cgaactcctt 1200  
 gtaggacgtc cagtcgtgcg gcccgcggtt gccaccgtcc gcgtagaccg cttccatggt 1260  
 cgccagccgg tccccgcgga actcgggtgg gatgtccgtg cccaaggtgg tcccgggtgt 1320  
 gtccgagagc accgggggct cgtaccggat gatgtgcaga tccaagaat t 1371

<210> 37

<211> 267

<212> PRT

<213> Streptomyces thermosacchari

<400> 37

Met Arg Leu Thr Arg Ser Leu Ser Ala Ala Ser Val Ile Val Phe Ala  
 1 5 10 15

Leu Leu Leu Ala Leu Leu Gly Ile Ser Pro Ala Gln Ala Ala Gly Pro  
 20 25 30

Ala Tyr Val Ala Leu Gly Asp Ser Tyr Ser Ser Gly Asn Gly Ala Gly  
 35 40 45

Ser Tyr Ile Asp Ser Ser Gly Asp Cys His Arg Ser Asn Asn Ala Tyr  
 50 55 60

Pro Ala Arg Trp Ala Ala Ala Asn Ala Pro Ser Ser Phe Thr Phe Ala  
 65 70 75 80

Ala Cys Ser Gly Ala Val Thr Thr Asp Val Ile Asn Asn Gln Leu Gly  
 85 90 95

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

Asp Ala Gly Phe Ala Asp Ala Met Thr Thr Cys Val Thr Ser Ser Asp  
 115 120 125

Ser Thr Cys Leu Asn Arg Leu Ala Thr Ala Thr Asn Tyr Ile Asn Thr  
 130 135 140

5

10

ES 2 449 474 T3

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

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

Ala Ser Asn Pro Trp Tyr Cys Leu Gly Leu Ser Asn Thr Lys Arg Ala  
180 185 190

Ala Ile Asn Thr Thr Ala Asp Thr Leu Asn Ser Val Ile Ser Ser Arg  
195 200 205

Ala Thr Ala His Gly Phe Arg Phe Gly Asp Val Arg Pro Thr Phe Asn  
210 215 220

Asn His Glu Leu Phe Phe Gly Asn Asp Trp Leu His Ser Leu Thr Leu  
225 230 235 240

Pro Val Trp Glu Ser Tyr His Pro Thr Ser Thr Gly His Gln Ser Gly  
245 250 255

Tyr Leu Pro Val Leu Asn Ala Asn Ser Ser Thr  
260 265

<210> 38

<211> 548

5 <212> PRT

<213> Thermobifida fusca

<400> 38

Met Leu Pro His Pro Ala Gly Glu Arg Gly Glu Val Gly Ala Phe Phe  
1 5 10 15

Ala Leu Leu Val Gly Thr Pro Gln Asp Arg Arg Leu Arg Leu Glu Cys  
20 25 30

His Glu Thr Arg Pro Leu Arg Gly Arg Cys Gly Cys Gly Glu Arg Arg  
35 40 45

Val Pro Pro Leu Thr Leu Pro Gly Asp Gly Val Leu Cys Thr Thr Ser  
50 55 60

Ser Thr Arg Asp Ala Glu Thr Val Trp Arg Lys His Leu Gln Pro Arg  
65 70 75 80

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

Gly Gln Gly Ser Pro Gly Val Leu Trp Cys Gly Arg Glu Gly Cys Arg  
100 105 110

10

ES 2 449 474 T3

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

ES 2 449 474 T3

Cys Thr Asp Gln Glu Asp Ala Ile Arg Lys Arg Met Ala Lys Phe Glu  
 370 375 380

Thr Thr Phe Glu Glu Leu Ile Ser Glu Val Arg Thr Arg Ala Pro Asp  
 385 390 395 400

Ala Arg Ile Leu Val Val Gly Tyr Pro Arg Ile Phe Pro Glu Glu Pro  
 405 410 415

Thr Gly Ala Tyr Tyr Thr Leu Thr Ala Ser Asn Gln Arg Trp Leu Asn  
 420 425 430

Glu Thr Ile Gln Glu Phe Asn Gln Gln Leu Ala Glu Ala Val Ala Val  
 435 440 445

His Asp Glu Glu Ile Ala Ala Ser Gly Gly Val Gly Ser Val Glu Phe  
 450 455 460

Val Asp Val Tyr His Ala Leu Asp Gly His Glu Ile Gly Ser Asp Glu  
 465 470 475 480

Pro Trp Val Asn Gly Val Gln Leu Arg Asp Leu Ala Thr Gly Val Thr  
 485 490 495

Val Asp Arg Ser Thr Phe His Pro Asn Ala Ala Gly His Arg Ala Val  
 500 505 510

Gly Glu Arg Val Ile Glu Gln Ile Glu Thr Gly Pro Gly Arg Pro Leu  
 515 520 525

Tyr Ala Thr Phe Ala Val Val Ala Gly Ala Thr Val Asp Thr Leu Ala  
 530 535 540

Gly Glu Val Gly  
 545

- <210> 39
- <211> 3000
- 5 <212> ADN
- <213> Thermobifida fusca

<400> 39

ggtagtgtaac cagaacaccc gtagtgtagc gtgggtagtcc aggtgcaggt gcaggttctt	60
caactgctcc agcaggatgc cgcctgggcc gtgcacgatg gccttgggca ggcctgtggt	120
ccccgacgag tacagcaccc atagcggatg gtcgaacggc agcggggtga actccagttc	180
cgcgcttcg cccgcggctt cgaactccgc ccaggacagg gtgtcggcga cagggccgca	240
10 gccaggtac ggcaggacga cggtagtgctg caggctgggc atgccgtcgc gcagggcttt	300

ES 2 449 474 T3

gagcacgtca cggcggtcga agtcottacc gccgtagcgg tagccgtcca cggccagcag 360  
 cactttcggg tcgatctcgc cgaaccggtc gaggacgctg cgcaccccga agtcggggga 420  
 acaggacgac caggtcgcac cgatcgcggc gcaggcgagg aatgcgggccg tcgcctcggc 480  
 gatgttcggc aggtaggcca cgaccgggtc gccggggccc accccgaggc tgcggagggc 540  
 cgcagcgatc gcgcggtgc gggtcgcgag ttctcccag gtccactcgg tcaacggccg 600  
 gagttcggac gcgtgccgga tcgccacggc tgatgggtca cggtcgcgga agatgtgctc 660  
 ggcgtagttg agggggcgc cggggaacca gacggcgccg ggcatggcgt cggaggcgag 720  
 cactgtggtg tacggggtg cgcgcgcac ccggtagtag tcccagatcg cggaccagaa 780  
 tccttcgagg tcggttaccg accagcgcca cagtgcctcg tagtccggtg cgtccacacc 840  
 gcggtgctcc cgcaccagc ggggtaacgc ggtgaggtg gcgcttctt tgcgctcctc 900  
 gtcgggactc cacaggatcg gcggtgcgg cttagtgctc atgaaacgcg acccctcgt 960  
 ggacgggtgc gatcggtga gcgtcgggtg cctcccctaa cgctccccg tgacggagtg 1020  
 ttgtgcacca catctagcac gcgggacgcg gaaaccgtat ggagaaaaca cctacaacc 1080  
 cggccggacg gtgggttcg gccacacta ggggtcgggt gcctgcttc cgggcagggc 1140  
 agtcccggg tgctgtggtg cgggggggag ggctgtcgtc tcgagggtg cggcgggac 1200  
 actcccggc tcagccgtac ccgcaacggg gacagttctc ctcccttcg ggtggtgag 1260  
 tcccttccc cgaatgcgg cgagatctc cagtcagccc ggaaaacacc cgctgtgccc 1320  
 aggtactctt tgcttgaac agacaggccg gacggccac gggggaggtt tgtgggcagc 1380  
 ggaccacgtg cggcgaccag acgacggtg ttctcggta tccccgctct tgtactgtg 1440  
 acagcgctca cgctggtctt ggctgtccc acggggcgcg agacgctgtg gcgcatgtg 1500  
 tgtgaggcca cccaggactg gtgcctggg gtgccggtcg actcccgcg acagcctgcg 1560  
 gaggacggc agtttctgct gctttctccg gtccaggcag cgacctggg gaactattac 1620  
 gcgctcggg attcgtactc ttcgggggac ggggcccgcg actactatcc cggcaccgcg 1680  
 gtgaagggcg gttgctggc gtccgctaac gcctatccg agctggctgc cgaagcctac 1740  
 gacttcggc gacacttgc gttcctggc tcgagcgcc agcgcgcta cgccatgctt 1800  
 gacgctatcg acgaggtcgg ctgcgagctg gactggaact cccctcacac gtcgctggtg 1860  
 acgatcggga tcggcgcaa cgatctggg ttctccacg ttttgaagac ctgcatggtg 1920  
 cgggtgccg tcgtggacag caaggcgtgc acggaccagg aggacgctat ccgcaagcgg 1980  
 atggcgaaat tcgagacgac gtttgaagag ctcatcagc aagtgcgcac ccgcgcccg 2040  
 gacgcccgga tccttgcgt gggctacccc cggattttc cggaggaacc gaccgccc 2100  
 tactacacgc tgaccgag caaccagcgg tggtcaacg aaaccattca ggagttcaac 2160  
 cagcagctcg ccgaggctgt cgcggtccac gacgaggaga ttgccgcgtc gggcggggtg 2220  
 ggcagcgtgg agttcgtgga cgtctaccac gcgtggagc gccacgagat cggctcggac 2280



ES 2 449 474 T3

gagccgtggg tgaacggggt gcagttgcgg gacctcgcca ccggggtgac tgtggaccgc 2340  
 agtaccttcc accccaacgc cgctgggcac cgggcggtcg gtgagcgggt catcgagcag 2400  
 atcgaaaccg gcccgggccg tccgctctat gccactttcg cggtggtggc gggggcgacc 2460  
 gtggacactc tcgcgggcga ggtggggtga cccggcttac cgtccggccc gcaggtctgc 2520  
 gagcactgcg gcgatctggt ccaactgccc gtgcagttcg tcttcggtga tgaccagcgg 2580  
 cggggagagc cggatcgttg agccgtgcgt gtctttgacg agcacacccc gctgcaggag 2640  
 ccgttcgcac agttctcttc cggtgggccag agtcgggtcg acgtcgatcc cagcccacag 2700  
 gccgatgctg cgggcgcgca ccacgccgtt gccgaccagt tggtcgaggc gggcgcgcag 2760  
 cacgggggag agggcgcgga catggtccag gtaaggccg tcgcgacga ggctcaccac 2820  
 ggcagtgccg accgcgcagg cgagggcggt gccgccgaag gtgctgccgt gctggccggg 2880  
 gcggatcacg tcgaagactt ccgcgtcgcc taccgccgcc gccacgggca ggatgccgcc 2940  
 gccacgcgt ttgccgaaca ggtagatata ggcgtcgact ccgctgtggt cgcaggcccg 3000

<210> 40  
 <211> 372  
 <212> PRT  
 <213> Thermobifida fusca

5

<400> 40

Val Gly Ser Gly Pro Arg Ala Ala Thr Arg Arg Arg Leu Phe Leu Gly  
 1 5 10 15  
 Ile Pro Ala Leu Val Leu Val Thr Ala Leu Thr Leu Val Leu Ala Val  
 20 25 30  
 Pro Thr Gly Arg Glu Thr Leu Trp Arg Met Trp Cys Glu Ala Thr Gln  
 35 40 45  
 Asp Trp Cys Leu Gly Val Pro Val Asp Ser Arg Gly Gln Pro Ala Glu  
 50 55 60  
 Asp Gly Glu Phe Leu Leu Leu Ser Pro Val Gln Ala Ala Thr Trp Gly  
 65 70 75 80  
 Asn Tyr Tyr Ala Leu Gly Asp Ser Tyr Ser Ser Gly Asp Gly Ala Arg  
 85 90 95  
 Asp Tyr Tyr Pro Gly Thr Ala Val Lys Gly Gly Cys Trp Arg Ser Ala  
 100 105 110  
 Asn Ala Tyr Pro Glu Leu Val Ala Glu Ala Tyr Asp Phe Ala Gly His  
 115 120 125

10

ES 2 449 474 T3

Leu Ser Phe Leu Ala Cys Ser Gly Gln Arg Gly Tyr Ala Met Leu Asp  
 130 135 140

Ala Ile Asp Glu Val Gly Ser Gln Leu Asp Trp Asn Ser Pro His Thr  
 145 150 155 160

Ser Leu Val Thr Ile Gly Ile Gly Gly Asn Asp Leu Gly Phe Ser Thr  
 165 170 175

Val Leu Lys Thr Cys Met Val Arg Val Pro Leu Leu Asp Ser Lys Ala  
 180 185 190

Cys Thr Asp Gln Glu Asp Ala Ile Arg Lys Arg Met Ala Lys Phe Glu  
 195 200 205

Thr Thr Phe Glu Glu Leu Ile Ser Glu Val Arg Thr Arg Ala Pro Asp  
 210 215 220

Ala Arg Ile Leu Val Val Gly Tyr Pro Arg Ile Phe Pro Glu Glu Pro  
 225 230 235 240

Thr Gly Ala Tyr Tyr Thr Leu Thr Ala Ser Asn Gln Arg Trp Leu Asn  
 245 250 255

Glu Thr Ile Gln Glu Phe Asn Gln Gln Leu Ala Glu Ala Val Ala Val  
 260 265 270

His Asp Glu Glu Ile Ala Ala Ser Gly Gly Val Gly Ser Val Glu Phe  
 275 280 285

Val Asp Val Tyr His Ala Leu Asp Gly His Glu Ile Gly Ser Asp Glu  
 290 295 300

Pro Trp Val Asn Gly Val Gln Leu Arg Asp Leu Ala Thr Gly Val Thr  
 305 310 315 320

Val Asp Arg Ser Thr Phe His Pro Asn Ala Ala Gly His Arg Ala Val  
 325 330 335

Gly Glu Arg Val Ile Glu Gln Ile Glu Thr Gly Pro Gly Arg Pro Leu  
 340 345 350

Tyr Ala Thr Phe Ala Val Val Ala Gly Ala Thr Val Asp Thr Leu Ala  
 355 360 365

Gly Glu Val Gly  
 370

- <210> 41
- <211> 300
- <212> PRT
- <213> Corynebacterium efficiens

5

<400> 41

ES 2 449 474 T3

Met Arg Thr Thr Val Ile Ala Ala Ser Ala Leu Leu Leu Leu Ala Gly  
 1 5 10 15

Cys Ala Asp Gly Ala Arg Glu Glu Thr Ala Gly Ala Pro Pro Gly Glu  
 20 25 30

Ser Ser Gly Gly Ile Arg Glu Glu Gly Ala Glu Ala Ser Thr Ser Ile  
 35 40 45

Thr Asp Val Tyr Ile Ala Leu Gly Asp Ser Tyr Ala Ala Met Gly Gly  
 50 55 60

Arg Asp Gln Pro Leu Arg Gly Glu Pro Phe Cys Leu Arg Ser Ser Gly  
 65 70 75 80

Asn Tyr Pro Glu Leu Leu His Ala Glu Val Thr Asp Leu Thr Cys Gln  
 85 90 95

Gly Ala Val Thr Gly Asp Leu Leu Glu Pro Arg Thr Leu Gly Glu Arg  
 100 105 110

Thr Leu Pro Ala Gln Val Asp Ala Leu Thr Glu Asp Thr Thr Leu Val  
 115 120 125

Thr Leu Ser Ile Gly Gly Asn Asp Leu Gly Phe Gly Glu Val Ala Gly  
 130 135 140

Cys Ile Arg Glu Arg Ile Ala Gly Glu Asn Ala Asp Asp Cys Val Asp  
 145 150 155 160

Leu Leu Gly Glu Thr Ile Gly Glu Gln Leu Asp Gln Leu Pro Pro Gln  
 165 170 175

Leu Asp Arg Val His Glu Ala Ile Arg Asp Arg Ala Gly Asp Ala Gln  
 180 185 190

Val Val Val Thr Gly Tyr Leu Pro Leu Val Ser Ala Gly Asp Cys Pro  
 195 200 205

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

Thr Gly Gln Ile Asn Glu Thr Val Arg Glu Ala Ala Glu Arg His Asp  
 225 230 235 240

ES 2 449 474 T3

Ala Leu Phe Val Leu Pro Asp Asp Ala Asp Glu His Thr Ser Cys Ala  
 245 250 255

Pro Pro Gln Gln Arg Trp Ala Asp Ile Gln Gly Gln Gln Thr Asp Ala  
 260 265 270

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

Val Arg Asp Ala Leu Gly Leu Glu Pro Val Gln Pro  
 290 295 300

<210> 42  
 <211> 3000  
 <212> ADN  
 <213> Corynebacterium efficiens

<400> 42

5

ttctgggggtg ttatgggggtt gttatcggct cgtcctgggt ggatcccgcc aggtggggta 60  
 ttcacggggg acttttgtgt ccaacagccg agaatgagtg ccctgagcgg tgggaatgag 120  
 gtgggcgggg ctgtgtcgcc atgagggggc ggcgggctct gtggtgcccc gcgacccccg 180  
 gccccgtga gcggtgaatg aaatccggct gtaatcagca tcccgtgcc accccgtcgg 240  
 ggagtcagc gcccgagtg tctacgcagt cggatcctct cggactcggc catgctgtcg 300  
 gcagcatcg cgtcccgggt cttggcgctc ctcggctgtt ctgcctgctg tccttggaa 360  
 gcgaaatgat caccggggag tgatacaccg gtggtctcat cccggatgcc cacttcggcg 420  
 ccatccggca attcgggcag ctccgggtgg aagtaggtgg catccgatgc gtcggtgacg 480  
 ccatagtggg cgaagatctc atcctgctcg aggggtctca ggccactctc cggatcgata 540  
 tcggggcggt ccttgatggc gtccttgctg aaaccgaggt gcagcttgtg ggcttccaat 600  
 ttcgaccac ggagcgggac gaggctgaa tgacggccga agagcccgtg gtggacctca 660  
 acgaaggtgg gtagtcccgt gtcattcattg aggaacacgc cctccaccgc acccagcttg 720  
 tggccggagt tgtcgtaggc gctggcatcc agaaggaaa cgatctcata tttgtcggtg 780  
 tgctcagaca tgatcttctt ttgctgtcgg tgtctggtag taccacggta gggctgaatg 840  
 caactgttat ttttctgtta ttttaggaat tgggtccatat cccacaggct ggctgtggtc 900  
 aaatcgatcat caagtaatcc ctgtcacaca aaatgggtgg tgggagccct ggtcgcgggt 960  
 ccgtgggagg cgcctgcccc cgcaggatcg tcggcatcgg cggatctggc cggtagcccc 1020  
 cggatgaataa aatcattctg taaccttcat cacggttgggt ttaggtatc cggccctttc 1080  
 gtcctgacct cgtccccggc gcgcgggagc ccgcggttg cggtagacag gggagacgtg 1140  
 gacaccatga ggacaacggc catcgcagca agcgcattac tccttctcgc cggatgcgcg 1200  
 gatggggccc gggaggagac cgccgggtgca ccgcccgggtg agtcctccgg gggcatccgg 1260

10

ES 2 449 474 T3

gaggagggg cggaggcgtc gacaagcatc accgacgtct acatcgccct cggggattcc 1320  
 tatgcggcga tgggcgggcg ggatcagccg ttacgggggtg agccgttctg cctgcgctcg 1380  
 tccgtaatt acccgaact cctccacgca gaggtcaccg atctcacctg ccagggggcg 1440  
 gtgaccggg atctgctcga acccaggacg ctgggggagc gcacgctgcc ggcgaggtg 1500  
 gatgcgctga cggaggacac caccctggtc accctctcca tcgggggcaa tgacctcgga 1560  
 ttcggggagg tggcgggatg catccgggaa cggatcgccg gggagaacgc tgatgattgc 1620  
 gtggacctgc tgggggaaac catcggggag cagctcgatc agcttcccc gcagctggac 1680  
 cgcgtgcacg aggctatccg ggaccgccc ggggacgccc aggttgtggt caccggttac 1740  
 ctgccgctcg tgtctgccg ggactgcccc gaactggggg atgtctccga ggcggatcgt 1800  
 cgttgggcgg ttgagctgac cgggcagatc aacgagaccg tgcgcgaggc ggccgaacga 1860  
 cacgatgccc tctttgtcct gcccgacgat gccgatgagc acaccagttg tgcaccccca 1920  
 cagcagcgtt gggcggatat ccagggcca cagaccgatg cctatccgct gcacccgacc 1980  
 tccgccggcc atgagggcat ggcgcgcc gtccgggagc cgctgggcct ggaaccggtc 2040  
 cagccgtagc gccgggcg cgcttgtcga cgaccaacc atgccaggct gcagtcacat 2100  
 ccgcacatag cgcgcgccc cgatggagta cgcaccatag aggatgagcc cgatgccgac 2160  
 gatgatgagc agcacactgc cgaagggtt ttccccgagg gtgcgcagag ccgagtccag 2220  
 acctgcggcc tgctccggat catgggcca accggcgatg acgatcaaca cccccaggat 2280  
 cccgaaggcg ataccacggc cgacataacc ggctgttccg gtgatgatga tcgcggtccc 2340  
 gacctgccct gaccccgcac ccgcctccag atcctcccgg aaatcccggg tggccccctt 2400  
 ccagagggtg tagacacccc cccccagtac caccagccc gcgaccacaa ccagcaccac 2460  
 accccagggt tgggatagga cgggtggcgt gacatcggtg gcggtctccc catcggaggt 2520  
 gctgccccc cggcgaagg tggaggtggt caccgccagg gagaagtaga ccatggccat 2580  
 gaccgcccc ttggcccttt ccttgaggtc ctgcccggc agcagctggc tcaattgcca 2640  
 gagtcccagg gccgccagg cgatgacggc aaccacagc aggaactgcc caccggagc 2700  
 ctccgcgatg gtggccagg cacctgaatt cgaggcctca tcaccgaac cgccgatcc 2760  
 agtggcgatg cgcaccgca tccaccgat gaggatgtgc agtatgcca ggacaatgaa 2820  
 accacctctg gccaggtg tcaagcggg gtggtcctcg gcctggtcgg cagcccgttc 2880  
 gatcgtccgt ttcgcgatc tgggtgcgcc cttatccata gctccattg aaccgcttg 2940  
 aggggtgggc ggccactgct agggcggatt gtgatctgaa ctgtgatgtt ccatcaacce 3000

<210> 43

<211> 268

5 <212> PRT

<213> Streptomyces coelicolor

<400> 43

ES 2 449 474 T3

Met Arg Arg Phe Arg Leu Val Gly Phe Leu Ser Ser Leu Val Leu Ala  
 1 5 10 15

Ala Gly Ala Ala Leu Thr Gly Ala Ala Thr Ala Gln Ala Ala Gln Pro  
 20 25 30

Ala Ala Ala Asp Gly Tyr Val Ala Leu Gly Asp Ser Tyr Ser Ser Gly  
 35 40 45

Val Gly Ala Gly Ser Tyr Ile Ser Ser Ser Gly Asp Cys Lys Arg Ser  
 50 55 60

Thr Lys Ala His Pro Tyr Leu Trp Ala Ala Ala His Ser Pro Ser Thr  
 65 70 75 80

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

Gly Gln Leu Gly Pro Leu Ser Ser Gly Thr Gly Leu Val Ser Ile Ser  
 100 105 110

Ile Gly Gly Asn Asp Ala Gly Phe Ala Asp Thr Met Thr Thr Cys Val  
 115 120 125

Leu Gln Ser Glu Ser Ser Cys Leu Ser Arg Ile Ala Thr Ala Glu Ala  
 130 135 140

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

Ile Ser Asp Lys Ala Pro Asn Ala His Val Val Val Ile Gly Tyr Pro  
 165 170 175

Arg Phe Tyr Lys Leu Gly Thr Thr Cys Ile Gly Leu Ser Glu Thr Lys  
 180 185 190

Arg Thr Ala Ile Asn Lys Ala Ser Asp His Leu Asn Thr Val Leu Ala  
 195 200 205

Gln Arg Ala Ala Ala His Gly Phe Thr Phe Gly Asp Val Arg Thr Thr  
 210 215 220

Phe Thr Gly His Glu Leu Cys Ser Gly Ser Pro Trp Leu His Ser Val  
 225 230 235 240

Asn Trp Leu Asn Ile Gly Glu Ser Tyr His Pro Thr Ala Ala Gly Gln  
 245 250 255

Ser Gly Gly Tyr Leu Pro Val Leu Asn Gly Ala Ala  
 260 265

5 <210> 44  
 <211> 2000  
 <212> ADN  
 <213> Streptomyces coelicolor

10 <400> 44

ES 2 449 474 T3

cccggcggcc cgtgcaggag cagcagccgg cccgcgatgt cctcgggct cgtcttcac 60  
 aggccgtcca tcgcgtcggc gaccggcgcc gtgtagttgg cccggacctc gtcccaggtg 120  
 cccgcggcga tctggcgggt ggtgcggtgc gggccgcgcc gaggggagac gtaccagaag 180  
 cccatcgtca cgttctccgg ctgcggttcg ggctcgtccg ccgctccgtc cgtcgcctcg 240  
 ccgagcacct tctcggcgag gtccggcgtg gtcgccgtca ccgtgacgtc ggcgccccgg 300  
 ctccagcgg agatcagcag cgtccagccg tcgccctccg ccagcgtcgc gctgcggtcg 360  
 tcgtcgcggg cgatccgcag cacgcgcgcg cccggcgcca gcagcgtggc gccggaccgt 420  
 acgcggtcga tgttcgccgc gtgcgagtac ggctgctcac ccgtggcga acggccgag 480  
 aacagcgcgt cgacgacgtc ggacggggag tcgctgtcgt ccacgttgag ccggatcggc 540  
 agggcttcgt gcgggttcac ggacatgtcg ccatgatcgg gcacccggcc gccgcgtgca 600  
 cccgctttcc cgggcacgca cgacaggggc tttctcggcg tcttccgtcc gaacttgaac 660  
 gagtgtcagc catttcttgg catggacact tccagtcaac gcgcgtagct gctaccacgg 720  
 ttgtggcagc aatcctgcta agggaggttc catgagacgt ttccgacttg tcggcttcct 780  
 gaggttcgtc gtcctcggcg cgggcgccgc cctcaccggg gcagcgaccg cccaggcggc 840  
 ccaaccgcgc gccccgacg gctatgtggc cctcggcgac tcctactcct cgggggtcgg 900  
 agcgggcagc tacatcagct cgagcggcga ctgcaagcgc agcacgaag cccatcccta 960  
 cctgtggggc gcccccact cgccctccac gttcgacttc accgcctgtt cggcgccccg 1020  
 tacgggtgat gttctctccg gacagctcgg cccgctcagc tcgggcaccg gcctcgtctc 1080  
 gatcagcatc ggcggcaacg acgcccgttt cgccgacacc atgacgacct gtgtgctcca 1140  
 gtcggagagc tcctgcctgt cgcggatcgc caccgccgag gcgtacgtcg actcgacgt 1200  
 gcccgcaag ctcgacggcg tctactcggc aatcagcgac aaggcgccga acgcccacgt 1260  
 cgtcgtcacc ggctaccgc gcttctacaa gctcggcacc acctgcatcg gcctgtccga 1320  
 gaccaagcgg acggcgatca acaaggcctc cgaccacctc aacaccgtcc tcgcccagcg 1380  
 cgccgcggcc cacggcttca ccttcggcga cgtacgcacc accttcaccg gccacgagct 1440  
 gtgctccggc agcccctggc tgcacagcgt caactggctg aacatcggcg agtcgtacca 1500  
 ccccaccgcg gccggccagt ccggtggcta cctgccggtc ctcaacggcg ccgcctgacc 1560  
 tcaggcggaa ggagaagaag aaggagcggg gggagacgag gaggggagg ccccgcccga 1620  
 cggggctccc gtccccgtct ccgtctccgt cccggtcccg caagtcaccg agaacgccac 1680  
 cgcgtcggac gtggcccga ccggactccg cãcctccacg cgcacggcac tctcgaacgc 1740  
 gccggtgtcg tcgtgcgtcg tcaccaccac gccgtcctgg cgcgagcgt cgccgcccga 1800  
 cgggaaggac agcgtccgcc accccgatc ggagaccgac ccgtccgcgg tcaccaccg 1860  
 gtagccgacc tccgcgggca gcccccgac cgtgaacgtc gccgtgaacg cgggtgcccg 1920  
 gtcgtcggc gccggacagg cccccgagta gtgggtgcgc gagcccacca cggtcacctc 1980  
 caccgactgc gctcgggggc 2000

5 <210> 45  
 <211> 269  
 <212> PRT

ES 2 449 474 T3

<213> Streptomyces avermitilis

<400> 45

Met Arg Arg Ser Arg Ile Thr Ala Tyr Val Thr Ser Leu Leu Leu Ala  
1 5 10 15

Val Gly Cys Ala Leu Thr Gly Ala Ala Thr Ala Gln Ala Ser Pro Ala  
20 25 30

Ala Ala Ala Thr Gly Tyr Val Ala Leu Gly Asp Ser Tyr Ser Ser Gly  
35 40 45

Val Gly Ala Gly Ser Tyr Leu Ser Ser Ser Gly Asp Cys Lys Arg Ser  
50 55 60

Ser Lys Ala Tyr Pro Tyr Leu Trp Gln Ala Ala His Ser Pro Ser Ser  
65 70 75 80

Phe Ser Phe Met Ala Cys Ser Gly Ala Arg Thr Gly Asp Val Leu Ala  
85 90 95

Asn Gln Leu Gly Thr Leu Asn Ser Ser Thr Gly Leu Val Ser Leu Thr  
100 105 110

Ile Gly Gly Asn Asp Ala Gly Phe Ser Asp Val Met Thr Thr Cys Val  
115 120 125

Leu Gln Ser Asp Ser Ala Cys Leu Ser Arg Ile Asn Thr Ala Lys Ala  
130 135 140

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

5

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

Arg Phe Tyr Lys Leu Gly Gly Ser Cys Leu Ala Gly Leu Ser Glu Thr  
180 185 190

Lys Arg Ser Ala Ile Asn Asp Ala Ala Asp Tyr Leu Asn Ser Ala Ile  
195 200 205

Ala Lys Arg Ala Ala Asp His Gly Phe Thr Phe Gly Asp Val Lys Ser  
210 215 220

Thr Phe Thr Gly His Glu Ile Cys Ser Ser Ser Thr Trp Leu His Ser  
225 230 235 240

Leu Asp Leu Leu Asn Ile Gly Gln Ser Tyr His Pro Thr Ala Ala Gly  
245 250 255

Gln Ser Gly Gly Tyr Leu Pro Val Met Asn Ser Val Ala  
260 265

<210> 46

10 <211> 1980



ES 2 449 474 T3

<212> ADN  
 <213> Streptomyces avermitilis

<400> 46

5  
 ccaccgccgg gtcggcggcg agtctcctgg cctcggtcgc ggagagggtg gccgtgtagc 60  
 cgttcagcgc ggcgcogaac gtcttcttca ccgtgccgcc gtactcgttg atcaggccct 120  
 tgcccttgct cgacgcggcc ttgaagccgg tgcccttctt gagegtgacg atgtagctgc 180  
 ccttgatcgc ggtgggggag ccggcggcga gcaccgtgcc ctcgccggg gtggcctggg 240  
 cgggcagtgc ggtgaatccg cccacagagg cgccggtcgc cacggcgggt atcgcgcgca 300  
 tccggatcct cttgctacgc agctgtgcc taagaggag tccctcctcg ggcagcggcg 360  
 cgccctgggt gggcgcacgg ctgtgggggg tgcgcgcgtc atcacgcaca cggccctgga 420  
 gcgtcgtgtt ccgccctggg ttgagtaaag cctcggccat ctacgggggt ggctcaaggg 480  
 agttgagacc ctgtcatgag tctgacatga gcacgcaatc aacggggccg tgagcacccc 540  
 ggggcgaccc cggaaagtgc cgagaagtct tggcatggac acttctctgc aacacgcgta 600  
 gctggtagca cggttacggc agagatcctg ctaaaggagg gttccatgag acgttcccga 660  
 attacgcat acgtgacctc actcctcctc gccgtcggct gcgccctcac cggggcagcg 720  
 acggcgcagg cgtccccagc cgccgcggcc acgggctatg tggccctcgg cgactcgtac 780  
 tcgtccggtg tcggcgcggc cagctacctc agctccagcg gcgactgcaa gcgcagtctg 840  
 aaggcctatc cgtacctctg gcaggccggc cattcacctc cgtcgttcag tttcatggct 900  
 tgctcggggc ctctgacggg tgatgtcctg gccaatcagc tcggcaccct gaactcgtcc 960  
 accggcctgg tctccctcac catcggaggc aacgacgcgg gcttctccga cgtcatgacg 1020  
 acctgtgtgc tccagtccga cagcgcctgc ctctcccga tcaacacggc gaaggcgtac 1080  
 gtcgactcca ccctgcccg ccaactcgac agcgtgtaca cggcgatcag cacgaaggcc 1140  
 ccgtcggccc atgtggcgtg gctgggctac ccccgcttct acaaactggg cggctcctgc 1200  
 ctcgccggcc tctcggagac caagcggctc gccatcaaag aocgggccga ctatctgaac 1260  
 agcggcatcg ccaagcgcgc cgccgaccac ggcttcacct tcggcgacgt caagagcacc 1320  
 ttcaccggcc atgagatctg ctccagcagc acctggctgc acagtctcga cctgctgaac 1380  
 atcggccagt cctaccaccc gaccgcggcc ggccagtcgg gcggctatct gccggtcagt 1440  
 aacagcgtgg cctgagctcc caocggcctga attttaagg cctgaatttt taaggcgaag 1500  
 gtgaaccgga agcggaggcc ccgtccgtcg ggtctcctcg cgcacaggtc accgagaacg 1560  
 gcacggagtt ggacgtcgtg cgcaaccggg cgcgcacctc gacggcgatc tcgttcgaga 1620  
 tcgttccgct cgtgtcgtac gtggtgacga acacctgctt ctgctgggtc tttccgccgc 1680  
 tcgccgggaa ggacagcgtc ttocagccc gatccgggac ctcgcccttc ttggtcacc 1740  
 agcggtagct cacctcgacc ggcacccggc ccaccgtgaa ggtcgccgtg aacgtgggcg 1800  
 cctgggcggg gggcggcggg caggcaccgg agtagtcggg gtgcacgccc gtgaccgtca 1860  
 ccttcacgga ctgggcccgc ggggtcgtcg taccgccgcc gccacggccc cctcccggag 1920  
 tggagcccga gctgtggtcg cccccccgt cggcgttgtc gtcctcgggg gttttcgaac 1980

10 <210> 47  
 <211> 372  
 <212> PRT  
 <213> Thermobifida fusca

15 <400> 47

ES 2 449 474 T3

Met Gly Ser Gly Pro Arg Ala Ala Thr Arg Arg Arg Leu Phe Leu Gly  
 1 5 10 15

Ile Pro Ala Leu Val Leu Val Thr Ala Leu Thr Leu Val Leu Ala Val  
 20 25 30

Pro Thr Gly Arg Glu Thr Leu Trp Arg Met Trp Cys Glu Ala Thr Gln  
 35 40 45

Asp Trp Cys Leu Gly Val Pro Val Asp Ser Arg Gly Gln Pro Ala Glu  
 50 55 60

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

Asn Tyr Tyr Ala Leu Gly Asp Ser Tyr Ser Ser Gly Asp Gly Ala Arg  
 85 90 95

Asp Tyr Tyr Pro Gly Thr Ala Val Lys Gly Gly Cys Trp Arg Ser Ala  
 100 105 110

ES 2 449 474 T3

Asn Ala Tyr Pro Glu Leu Val Ala Glu Ala Tyr Asp Phe Ala Gly His  
 115 120 125

Leu Ser Phe Leu Ala Cys Ser Gly Gln Arg Gly Tyr Ala Met Leu Asp  
 130 135 140

Ala Ile Asp Glu Val Gly Ser Gln Leu Asp Trp Asn Ser Pro His Thr  
 145 150 155 160

Ser Leu Val Thr Ile Gly Ile Gly Gly Asn Asp Leu Gly Phe Ser Thr  
 165 170 175

Val Leu Lys Thr Cys Met Val Arg Val Pro Leu Leu Asp Ser Lys Ala  
 180 185 190

Cys Thr Asp Gln Glu Asp Ala Ile Arg Lys Arg Met Ala Lys Phe Glu  
 195 200 205

Thr Thr Phe Glu Glu Leu Ile Ser Glu Val Arg Thr Arg Ala Pro Asp  
 210 215 220

Ala Arg Ile Leu Val Val Gly Tyr Pro Arg Ile Phe Pro Glu Glu Pro  
 225 230 235 240

Thr Gly Ala Tyr Tyr Thr Leu Thr Ala Ser Asn Gln Arg Trp Leu Asn  
 245 250 255

Glu Thr Ile Gln Glu Phe Asn Gln Gln Leu Ala Glu Ala Val Ala Val  
 260 265 270

His Asp Glu Glu Ile Ala Ala Ser Gly Gly Val Gly Ser Val Glu Phe  
 275 280 285

Val Asp Val Tyr His Ala Leu Asp Gly His Glu Ile Gly Ser Asp Glu  
 290 295 300

Pro Trp Val Asn Gly Val Gln Leu Arg Asp Leu Ala Thr Gly Val Thr  
 305 310 315 320

Val Asp Arg Ser Thr Phe His Pro Asn Ala Ala Gly His Arg Ala Val  
 325 330 335

Gly Glu Arg Val Ile Glu Gln Ile Glu Thr Gly Pro Gly Arg Pro Leu  
 340 345 350

Tyr Ala Thr Phe Ala Val Val Ala Gly Ala Thr Val Asp Thr Leu Ala  
 355 360 365

Gly Glu Val Gly  
 370

- 5 <210> 48
- <211> 968
- <212> ADN
- <213> Thermobifida fusca

ES 2 449 474 T3

<400> 48

ctgcagacac ccgccccgcc ttctcccgga tcgtcatggt cggcgactcc ctcagcgaca	60
ccggcaagat gtactccaag atgcgcggt acctgccgtc ctccccgccg tactacgagg	120
gccgttctc gaacggcccc gtctggctgg agcagctgac gaagcagttc cccggcctga	180
cgatcgccaa cgaggccgag gggggcgcga ccgcagtgc ctacaacaag atctcctgga	240
acccgaagta ccaggtcatt aacaacctcg actacgaggt caccagttc ttgcagaagg	300
actcgttcaa gcccgacgac ctggatcatcc tgtgggtggg cgccaacgac tacctggcct	360
acggttgaa cacggagcag gacgccaagc gggcgcgga cgccatctcg gacgcgcaa	420
accgatggt cctgaacggc gcgaagcaga tcctgtgtt caacctgcc gacctgggcc	480
agaacctgct cccccgctcc cagaaggtcg tcgaggcctg ctgcacgtg tccgcctacc	540
acaacaagct gtcctcaac ctcgccccgc agctcgccc gacgggcatg gtcaagctgt	600
tcgagatcga caagcagttc gcggagatgc tgcgcgacct ccagaacttc ggctgagcg	660
acgtggagaa cccgtgctac gacggcggct acgtgtgaa gccgttcgcc acccgtccg	720
tctcgaccga ccggcagctg tcggccttct cgccccagga gcgcctggcg atcgctggca	780
acccgctcct ggacacggcg gtagcttcgc cgatggccc ccgctcggcc tcgcccctca	840
actgcgaggg caagatgttc tgggaccagg tccacccac caccgtggtc caccgcccc	900
tctcggagcg cgccgccacc ttcacgcaga ccagctacga gttcctcgcc cactagtcta	960
gaggatcc	968

- 5 <210> 49
- <211> 1044
- <212> ADN
- <213> Aeromonas salmonicida

10 <400> 49

atgaaacaac aaaaacggct ttacgccgga ttgctgacgc tgttatttgc gtcacatctc	60
ttgctgcctc attctgcagc ttcagcagca gatacaagac cggcgtttag ccgcatcgtc	120
atggttgag atagcctgag cgatacgggc aaaatgtata gcaaatgag aggctatctt	180
ccgtcaagcc cgccgtatta tgaaggccgc tttagcaatg gaccggtctg gctggaacaa	240
ctgacgaaac aatctccggg actgacgatc gctaataag cagaaggagg agcaacagcg	300
gtcgcctata acaaatcag ctgggacctg aaatatcagg tcatcaacaa cctggactat	360
gaagtacac agtttcttca gaaagacagc tttaaaccgg atgatctggt catcctttgg	420

ES 2 449 474 T3

gtcggcgcca atgattatct ggcgtatggc tggaacacag aacaagatgc caaaagagtc 480  
 agagatgcca tcagcgatgc cgtaataga atggctctga acggcgccaa acaaactctg 540  
 ctgtttaacc tgccggatct gggacaaaat ccgagcgcca gaagccaaaa agtcgctgaa 600  
 gcagtcagcc atgtcagcgc ctatcataac aaactgctgc tgaacctggc aagacaattg 660  
 gcaccgacgg gaatggtaa attgtttgaa attgacaaac agtttgccga aatgctgaga 720  
 gatccgcaaa attttggcct gagcgatgtc gaaaaccctg gctatgatgg cggatatgtc 780  
 tggaaaccgt ttgccacaag aagcgtcagc acggatagac aactgtcagc gtttagcccc 840  
 caagaaagac tggcaatcgc cggaaatccg cttttggcac aagcagttgc ttcaccgatg 900  
 gcaagaagat cagcaagccc gctgaattgc gaaggcaaaa tgttttggga tcaggtccat 960  
 ccgacaacag ttgtccatgc tgccctttca gaaagagcgg cgacgtttat cgaaacacag 1020  
 tatgaatttc tggcccatgg ctga 1044

<210> 50  
 <211> 1005  
 <212> ADN  
 <213> Aeromonas hydrophila

5

<400> 50

atgaaaaaat ggtttgtgtg tttattggga ttggtcgcgc tgacagttca ggcagccgac 60  
 agccgtcccg ccttctcccg gatcgtgatg tttggcgaca gcctctccga taccggcaag 120  
 atgtacagca agatgocgcg ttacctcccc tccagcccc cctactatga gggccgcttc 180  
 tccaacgggc ccgtctggct ggagcagctg accaacgagt tcccgggcct gaccatagcc 240  
 aacgaggcgg aaggcggacc gaccgcccgt gcttacaaca agatctcctg gaatcccaag 300  
 tatcaggtca tcaacaacct ggactacgag gtcacccagt tcctgcaaaa agacagcttc 360  
 aagccggacg atctgggtgat cctctgggtc ggcgccaacg actatctggc ctatggctgg 420  
 aacacagagc aggatgccaa gcgggtgcgc gacgcatca gcgatgcggc caaccgcatg 480  
 gtgctgaacg gcgccaagga gatactgctg ttcaacctgc cggatctggg ccagaacccc 540  
 tcggcccgca gccagaaggt ggtcgaggcg gccagccatg tctccgcta ccacaaccag 600  
 ctgctgctga acctggcacg ccagctggct cccaccggca tggatgaagct gttcgagatc 660  
 gacaagcagt ttgccgagat gctgctgat ccgcagaact tcggcctgag cgaccagagg 720  
 aacgcctgct acggtggcag ctatgtatgg aagccgtttg cctcccgcag cgccagcacc 780  
 gacagccagc tctccgcctt caaccgcgag gagcgcctcg ccatcgccgg caaccgcgctg 840  
 ctggcccagg ccgtcgccag ccccatggct gcccgcagcg ccagcaccct caactgtgag 900  
 ggcaagatgt tctgggatca ggtccacccc accactgtcg tgcacgcccg cctgagcggag 960  
 10 cccgcccgca ccttcatcga gagccagtac gagttcctcg cccac 1005

<210> 51  
 <211> 1011  
 <212> ADN  
 <213> Aeromonas salmonicida

15

<400> 51

ES 2 449 474 T3

atgaaaaaat ggtttgttg tttattggg ttgatcgcg tgacagttca ggcagccgac 60  
actcgccccg ccttctcccc gatcgtgatg ttcggcgaca gcctctccga taccggcaaa 120  
atgtacagca agatgcgcgg ttacctcccc tccagccccg cctactatga gggccgtttc 180  
tccaacggac ccgtctggct ggagcagctg accaagcagt tcccgggtct gaccatcgcc 240  
aacgaagcgg aaggcgggtg cactgccgtg gcttacaaca agatctcctg gaatcccaag 300  
tatcaggtct acaacaacct ggactacgag gtcacccagt tcttgacaga agacagcttc 360  
aagccggacg atctggtgat cctctgggtc ggtgccaatg actatctggc atatggtggt 420  
aatacggagc aggatgccaa gcgagttcgc gatgccatca gcgatgcggc caaccgcatg 480  
gtactgaacg gtgccaaagca gatactgctg ttcaacctgc cggatctggg ccagaaccgg 540  
tcagccccga gtcagaaggt ggtcgaggcg gtcagccatg tctccgccta tcacaacaag 600  
ctgctgtga acctggcagc ccagctggcc cccaccggca tggtaaagct gttcgagatc 660  
gacaagcaat ttgccgagat gctgctgat cgcgagaact tccgctgag cgacgtcgag 720  
aaccctgct acgacggcgg ctatgtgtgg aagccgtttg ccaccgcgag cgtcagcacc 780  
gaccgccagc tctccgcctt cagtccgag gaacgcctcg ccacgcggc caaccgctg 840  
ctggcacagg ccgttgccag tctatggcc cgcgcgagc ccagccccct caactgtgag 900  
ggcaagatgt tctgggatca ggtacaccgc accactgtcg tgcacgcagc cctgagcgag 960  
cgcgcggcca cctcatcga gaccagtac gagttcctcg cccacggatg a 1011

<210> 52

<211> 888

<212> ADN

<213> Streptomyces coelicolor

<400> 52

atgccgaagc ctgcccttcg ccgtgtcatg accgcgacag tcgccgccgt cggcacgctc 60  
gccctcggcc tcaccgagc caccgcccac gccgcgccg cccaggccac tccgaccctg 120  
gactacgtcg ccctcgcgga cagctacagc gccggctccg gcgtcctgcc cgtcgacccc 180  
gccaacctgc tctgtctgcg ctcgacggcc aactaccccc acgtcatcgc ggacacgagc 240  
ggcggccgcc tcacggacgt cacctgcggc gccgcgcaga ccgccgactt cacgcgggcc 300  
cagtaccggg gcgtcgcacc ccagttggac gcgctcggca ccggcacgga cctggtcacg 360  
ctcaccatcg gcggcaacga caacagcacc ttcaccaacg ccatcacggc ctgcggcacg 420  
gcgggtgtcc tcagcggcgg caagggcagc ccctgcaagg acaggcacgg cacctccttc 480  
gacgacgaga tcgaggccaa cacgtacccc gcgctcaagg aggcgctgct cggcgtccgc 540  
gccagggtc cccacgccag ggtggcggct ctgggtacc cgtggatcac cccggccacc 600  
gccgaccgt cctgttctct gaagctcccc ctgcgccggc gtgacgtgcc ctacctggg 660  
gccatccagg cacacctcaa cgacgcggtc cggcgggccg ccgaggagac cggagccacc 720  
tacgtggact tctccggggg gtccgacggc cagcagcct gcgaggcccc cggcaccgcg 780  
tggatcgaac cgctgctctt cgggcacagc ctggttccc tccacccaa cgcctggggc 840  
gagcggcgca tggccgagca cacgatggac gtocctcgcc tggactga 888

ES 2 449 474 T3

<210> 53  
 <211> 888  
 <212> ADN  
 <213> Streptomyces coelicolor

5

<400> 53

tcagtcacagg ccgaggacgt ccatcgtgtg ctccgcatg cgccgctcgc ccagggcggt	60
gggggtggacg ggaacgaggc tgtgccccgaa gagcagcggg tcgatccagc ggggtgccggg	120
ggcctcgcag gcgtcgtggc cgtcggacac cccggagaag tccacgtagc tggctccggg	180
ctcctcggcg gcccgccgga ccgcgtcgtt gaggtgtgcc tggatggccc gcaggtaggg	240
cacgtcaccg gcggcgaggg ggagcttcag gaagcaggac gggtcggcgg tggccggggg	300
gatccacggg tagccgagag ccgccaccct ggcgtgggga gccctggcgc ggacgccgag	360
cagcgcctcc ttgagcgcgg ggtacgtgtt ggcctcgatc tcgtcgtcga aggagtgcc	420
gtgcctgtcc ttgcaggggc tgccttgcc gccctgagg acaccgccg tgcgcagggc	480
cgtgatggcg ttgatgaagg tgctgttgc gttgccgccg atggtgagcg tgaccaggtc	540
cgtgccggtg ccgagcgcgt ccaactgggg tgcgacgccg gggactggg cccgcgtgaa	600
gtcggcggtc tgcgcggcgc cgcaggtgac gtccgtgagg cgggcgcccg tcgtgtccgc	660
gatgacgtgg gggtagttgg ccgtcgagcg cagacagagc aggttgccgg ggtcgacggg	720
caggacgccg gagccggcgc tgtagctgtc gccgagggcg acgtagtcca gggtcggagt	780
ggcctgggcg ggcgcgcggt gggcggtggc gtcggtgagg ccgagggcga gcgtgccgac	840
ggcggcgact gtcgcgggtca tgacacggcg aagggcaggg ttcggcat	888

10

<210> 54  
 <211> 717  
 <212> ADN  
 <213> Saccharomyces cerevisiae

15

<400> 54

atggattacg agaagtttct gttatttggg gattccatta ctgaatttgc ttttaatact	60
aggcccattg aagatggcaa agatcagtat gctottggag ccgcattagt caacgaatat	120
acgagaaaaa tggatattct tcaaagaggg ttcaaagggg acacttctag atgggcgttg	180
aaaatacttc ctgagatfff aaagcatgaa tccaatattg tcatggccac aatatttttg	240
gggtgccaacg atgcatgctc agcaggtccc caaagtgtcc ccctccccga atttatcgat	300
aatattcgtc aatgggtatc tttgatgaag tcttaccata tccgtoctat tataatagga	360
ccggggctag tagatagaga gaagtgggaa aaagaaaaat ctgaagaaat agctctcgga	420
tacttccgta ccaacgagaa ctttgccatt tattccgatg ccttagcaaa actagccaat	480
gaggaaaaag ttcccttcgt ggctttgaat aaggcgtttc aacaggaagg tggatgatgct	540
tggcaacaac tgtaacaga tggactgcac ttttccggaa aagggtaaa aatttttcat	600
gacgaattat tgaaggtcat tgagacattc taccaccaat atcatcccaa aaacatgcag	660
tacaaactga aagattggag agatgtgcta gatgatggat ctaacataat gtcttga	717

20

ES 2 449 474 T3

<210> 55  
 <211> 1044  
 <212> ADN  
 <213> Ralstonia sp.

5

<400> 55

atgaacctgc	gtcaatggat	ggcgccgcc	acggctgcc	ttgccttggg	cttggccgcg	60
tgcgggggcg	gtgggaccga	ccagagcggc	aatcccaatg	tcgccaaggt	gcagcgcag	120
gtggtgttcg	gcgacagcct	gagcgatatc	ggcacctaca	cccccgtcgc	gcaggcggtg	180
ggcggcggca	agttcaccac	caaccgggc	ccgatctggg	ccgagaccgt	ggccgcgcaa	240
ctgggcgtga	cgctcacgcc	ggcgggatg	ggctacgcca	cctccgtgca	gaattgcccc	300
aaggccggct	gcttcgacta	tgcgcagggc	ggctcgcgcg	tgaccgatcc	gaacggcatc	360
ggccacaacg	gcggcgcggg	ggcgtgacc	taccgggttc	agcagcagct	cgccaacttc	420
tacgcggcca	gcaacaacac	attcaacggc	aataacgatg	tcgtcttcgt	gctggccggc	480
agcaacgaca	ttttcttctg	gaccactcgc	gcgccacca	gcggctccgg	cgtgacgccc	540
gccattgcca	cggcccaggt	gcagcaggcc	gcgacggacc	tggtcggcta	tgtcaaggac	600
atgatcgcca	agggtgcgac	gcaggtctac	gtgttcaacc	tgcccgacag	cagcctgacg	660
ccggacggcg	tggcaagcgg	cacgaccggc	caggcgtctg	tgcacgcgct	ggtgggcacg	720
ttcaacacga	cgctgcaaag	gggctggcc	ggcacctcgg	cgcgcatcat	cgacttcaac	780
gcacaactga	ccgcggcgat	ccagaatggc	gcctcgttcg	gcttcgcca	caccagcggc	840
cgggcctcgc	acgccaccaa	gatcaatgcc	ctggtgccga	gcgccggcgg	cagctcgtcg	900
ttctgctcgg	ccaacacgct	ggtggcttcc	ggtgcggacc	agagctacct	gttcgccgac	960
ggcgtgcacc	cgaccacggc	cggccatcgc	ctgatcgcca	gcaacgtgct	ggcgcgcctg	1020
ctggcggata	acgtcgcgca	ctga				1044

10 <210> 56  
 <211> 786  
 <212> ADN  
 <213> Streptomyces coelicolor

15 <400> 56



ES 2 449 474 T3

gtgatcgggt cgtacgtggc ggtgggggac agcttcaccg agggcgctcg cgacccccgc 60  
 cccgacgggg cgctcgtcgg ctggggccgac cggctcgcgg tactgctcgc ggaccggcgc 120  
 cccgagggcg acttcacgta cacgaacctc gccgtgcgcg gcaggctcct cgaccagatc 180  
 gtggcggaac aggtcccggc ggtcgtcggc ctgcgcggcg acctcgtctc gttcgcggcg 240  
 ggcggaacg acatcatccg gcccggcacc gatcccgcag aggtcgcgga gcggttcgag 300  
 ctggcggtgg ccgcgctgac cgccggggcc ggaaccgtcc tggtgaccac cgggttcgac 360  
 acccgggggg tgcccgtcct caagcacctg cgcggaaga tcgccacgta caacgggcac 420  
 gtccgcgcca tcgccgaccg ctacggctgc ccggtgctcg acctgtggtc gctgcggagc 480  
 gtccaggacc gcaggcggcg ggacggcgac cggctgcacc tgtcgcggga ggggcacacc 540  
 cgggtggcgc tcgcgcgggg gcaggccctg ggctgcgcg tcccggccga ccctgaccag 600  
 ccctggccgc ccctgccgcc gcgcggcacg ctgcagctcc ggcgcgacga cgtgactgg 660  
 gcgcgcgagt acctggtgcc gtggatcggg cgccggctgc ggggcgagtc gtcgggcgac 720  
 cacgtgacgg ccaaggggac gctgtcggcg gacgccatca agacgcggat cgcccggtg 780  
 gcctga 786

<210> 57

<211> 783

5 <212> ADN

<213> Streptomyces coelicolor

<400> 57

atgcagacga accccgcgta caccagtctc gtgcgcgtcg gcgactcctt caccgagggc 60  
 atgtcggacc tgctgccga cggctcctac cgtggctggg ccgacctcct cgccaccgg 120  
 atggcgggcc gctccccgg cttccggtac gccaacctgg cgggtgcgcg gaagctgatc 180  
 ggacagatcg tcgacgagca ggtggacgtg gccgcccga tgggagccga cgtgatcacg 240  
 ctggtcggcg ggtcaacga cacgctgcgg cccaagtgcg acatggcccg ggtgcgggac 300  
 ctgctgacct aggccgtgga acggctcggc ccgactgcg agcagctggt gctgatgcgc 360  
 agtcccggtc gccagggtcc ggtgctggag cgcttccggc ccgcatgga ggccctgttc 420  
 gccgtgatcg acgacctggc cgggcccgcac ggcgcccgtg tcgtcgacct gtacggggcc 480  
 cagtcgctgg ccgacctcg gatgtgggac gtggaccggc tgcacctgac cgccgagggc 540  
 caccgccggg tcgcggaggc ggtgtggcag tcgctcggcc acgagcccga ggacccccgag 600  
 tggcacgcgc cgatcccggc gacgcccgg ccgggggtgg tgacgcgcag gaccgaggac 660  
 gtccggttcg ccgcgacga cctgctgccc tggataggcc gcaggctgac cgggcgctcg 720  
 tccggggacg gcctgccggc caagcggccc gacctgctgc cctacgagga ccccgcacgg 780  
 tga 783

10

<210> 58

<211> 1365

<212> ADN

15 <213> Streptomyces coelicolor

<400> 58

ES 2 449 474 T3

atgacccggg gtcgtgacgg ggggtgcgggg gcgccccca ccaagcaccg tgccctgctc 60  
 gcggcgatcg tcaccctgat agtggcgatc tccgcgcca tatacgcggg agcgtccggg 120  
 gacgacggca gcagggacca cgcgctgcag gccggaggcc gtctcccacg aggagacgcc 180  
 gccccgcgt ccaccggtgc ctgggtgggc gcctgggcca ccgcaccggc cgcggccgag 240  
 ccgggcaccg agacgaccgg cctggcgggc cgctccgtgc gcaacgtcgt gcacacctcg 300  
 gtcggcgcca ccggcgcgcg gatcaccctc tcgaacctgt acgggcagtc gccgctgacc 360  
 gtcacacacg cctcgatcgc cctggccgcc gggcccgaca ccgccgccgc gatcgcggac 420  
 accatgcgcc ggctcacctt cggcggcagc gcccggtga tcatcccggc gggcgccag 480  
 gtgatgagcg acaccgcccg cctcgccatc ccctacgggg cgaacgtcct ggtcaccacg 540  
 tactccccca tcccgtccgg gccggtgacc taccatccgc aggcccggca gaccagctac 600  
 ctggccgacg gcgaccgcac ggcggacgtc accgccgtcg cgtacaccac ccccacgcc 660  
 tactggcgct acctgaccgc cctcgacgtg ctgagccacg aggccgacgg cacggtcgtg 720  
 gcgttcggcg actccatcac cgacggcgcc cgctcgcaga gcgacgcaa ccaccgctgg 780  
 accgacgtcc tcgccgcacg cctgcacgag gcggcgggcg accggccggga cacgccccgc 840  
 tacagcgtcg tcaacgaggg catcagcggc aaccggctcc tgaccagcag gccggggcgg 900  
 ccggccgaca acccgagcgg actgagccgg ttccagcggg acgtgctgga acgcaccaac 960  
 gtcaaggccg tcgtcgtcgt cctcggcgtc aacgacgtcc tgaacagccc ggaactcgcc 1020  
 gaccgcgacg ccacccctgac cggcctgcgc accctcgtcg accgggcgca ccccgggga 1080  
 ctgcgggtcg tcggcgccac gatcacgccg ttccggcggt accggcgcta caccgaggcc 1140  
 cgcgagacga tgcggcagga ggtcaacgag gagatccgct ccggccgggt cttcgacacg 1200  
 gtcgtcgact tcgacaaggc cctgcgcgac ccgtacgacc cgcgccggat gcgctccgac 1260  
 tacgacagcg gcgaccacct gacccccgc gacaaggggt acgcgcgcat gggcgcggtc 1320  
 atcgacctgg ccgcgctgaa gggcgcggcg ccggtcaagg cgtag 1365

<210> 59  
 <211> 1023  
 <212> ADN  
 <213> Streptomyces coelicolor

<400> 59

atgacgagca tgtcgagggc gagggtgcg cggcggatcg cggccggcgc ggcgtacggc 60  
 ggcggcggca tcggcctggc gggagcggcg gcggtcggtc tgggtggcgc cgaggtgcag 120  
 ctggccagac gcaggggtgg ggtgggcacg ccgaccggg tgccgaacgc gcagggactg 180  
 tacggcggca ccctgcccac ggcggcgac ccgccgctgc ggctgatgat gctgggcgac 240

ES 2 449 474 T3

tccacggccg ccgggcaggg cgtgcaccgg gccgggcaga cgcggggcgc gctgctggcg 300  
 tccgggctcg cggcgggtggc ggagcggccg gtgcggtggc ggtcggtcgc ccagccgggg 360  
 gcgtgctcgg acgacctgga ccggcaggtg gcgctggtgc tcgccgagcc ggaccgggtg 420  
 cccgacatct gcgtgatcat ggtcggcgcc aacgacgtca cccaccggat gccggcgacc 480  
 cgctcgggtgc ggcacctgtc ctccggcgga cggcggctgc gcacggccgg tgcggaggtg 540  
 gtggtcggca cctgtccgga cctgggcacg atcgagcggg tgcggcagcc gctgcgctgg 600  
 ctggcccggc gggcctcacg gcagctcgcg gcggcacaga ccatcggcgc cgtcagagcag 660  
 ggcgggcgca cgggtgctcgt gggcgacctg ctgggtccgg agttcgcgca gaaccgcgg 720  
 gagctcttcg gccccgacaa ctaccacccc tccgccgagg ggtaccacac ggcccgcatg 780  
 gcggtactgc cctcgggtgtg cgcgcgctc ggcctgtggc cggccgacga ggagcaccgc 840  
 gacgcgctgc gccgcgaggg ctctctgccg gtggcgcgcg cggcggcgga ggcggcgtcc 900  
 gaggcgggta cggaggtcgc cgcgccatg cctacggggc ctccggggcc ctgggcgctg 960  
 ctgaagcggc ggagacggcg tcgggtgtcg gaggcggaac cgtccagccc gtccggcggt 1020  
 tga 1023

<210> 60  
 <211> 918  
 <212> ADN  
 <213> Streptomyces coelicolor

5

<400> 60

atgggtcag ggacggacca gcggacggcg tacggccgtc gccgggcgcg tgtcgcgctc 60  
 gccgccctga ccgccccgt cctgggcgtg ggcgtggcgg gctgcgactc cgtgggcggc 120  
 gactcaccgc ctcttccgg cagcccgtcg aagcggacga ggacggcgcc cgcctgggac 180  
 accagcccgg cgtccgtcgc cgcctgggc gactccatca cgcgcggctt cgacgcctgt 240  
 gcggtgctgt cggactgcc ccggaggtgtc tgggcgaccg gcagcagcgc gaaggtcgac 300  
 tcgctggccg tacggtgct ggggaaggcg gacgcggccg agcacagctg gaactacggc 360  
 gtcaccgggg cccgatggc ggacctgacc gctcaggtga cgcgggcggc gcagcgcgag 420  
 ccggagctgg tggcgggtgat ggcggggcg aacgacgcgt gccggccac gacctcggcg 480  
 atgacgccgg tggcggactt ccgggcgag ttcgaggagg cgatggccac cctgcgcaag 540  
 aagctcccca aggcgcaggt gtacgtgtcg agcatcccgg acctcaagcg gctctggtcc 600  
 caggcccgca ccaaccgct gggcaagcag gtgtggaagc tcggcctgtg cccgtcgatg 660  
 ctgggcgacg cggactccct ggactcggcg gcgaccctgc ggcgcaacac ggtgcgcgac 720  
 cgggtggcgg actacaacga ggtgctgcgg gaggtctgcg cgaaggaccg gcggtgccgc 780  
 agcgacgacg gcgcggtgca cgagttccgg ttcggcacgg accagttgag cactgggac 840  
 tggttccacc cgagtgtgga cggccaggcc cggctggcgg agatcgcta ccgcgcggtc 900  
 accgcaaga atccctga 918

10

<210> 61  
 <211> 1068  
 <212> ADN

15

ES 2 449 474 T3

<213> Streptomyces rimosus

<400> 61

	ttcatcacia cgatgtcaca acaccggcca tccgggtcat ccctgatcgt gggaatgggt	60
	gacaagcctt cccgtgacga aagggctctg ctacatcaga aatgacagaa atcctgctca	120
	gggaggttcc atgagactgt cccgacgcgc ggccacggcg tccgcgctcc tctcaccccc	180
	ggcgctcgcg ctcttcggcg cgagcgccgc cgtgtccgcg ccgcgaatcc aggccaccga	240
	ctacgtggcc ctcgcgact cctactcctc gggggtcggc gcgggcagct acgacagcag	300
	cagtggctcc tgtaagcgca gcaccaagt cctaccggcc ctgtgggccc cctcgcacac	360
	cggtacgcgg ttcaacttca ccgcctgttc gggcgccgcg acaggagacg tgctggccaa	420
	gcagctgacc ccggtcaact ccggcaccga cctggtcagc attaccatcg gcggcaacga	480
	cgcggtcttc gccgacacca tgaccactg caacctccag ggcgagagcg cgtgcctggc	540
	gcggtatcgc aagcgcgcg cctacatcca gcagacgctg cccgcccagc tggaccaggt	600
	ctacgacgcc atcgacagcc gggcccccgc agcccaggtc gtcgtcctgg gctaccgcg	660
	cttctacaag ctggggcgca gctgcccgt cggctctctg gagaagtccc gcgcggccat	720
	caacgccgcc gccgacgaca tcaacgccgt caccgccaag cgcgccgccc accacggctt	780
	cgcttcggg gacgtcaaca cgaccttgc cgggacagag ctgtgctccg gcgccccctg	840
	gctgcacagc gtcacccttc ccgtggagaa ctctaccac cccacggcca acggacagtc	900
	caagggttac ctgccctcc tgaactccgc cacctgatct cgcggctact ccgcccctga	960
	cgaagtccc cccccggcg gggcttcgcc gtaggtgccc gtaccgccgt cgcccctcgc	1020
5	gccggtggcc ccgccgtac tgccgccgcc cccggacgcg gtcggttc	1068

<210> 62

<211> 1008

<212> ADN

10 <213> Aeromonas hydrophila

<400> 62

	atgaaaaaat ggtttgtgtg tttattggga ttggtcgcgc tgacagttca ggcagccgac	60
	agtcgccccg ccttttcccg gatcgtgatg ttcggcgaca gcctctccga taccggcaaa	120
	atgtacagca agatgcgcgg ttacctccc tccagcccgc cctactatga gggccgtttc	180
	tccaacggac ccgtctggct ggagcagctg accaaacagt tcccgggtct gaccatcgcc	240
	aacgaagcgg aagcgggtgc cactgccgtg gcttacaaca agatctcctg gaatcccaag	300
15	tatcaggtca tcaacaacct ggactacgag gtcaccagc tcttgacagaa agacagcttc	360

ES 2 449 474 T3

aagccggacg atctggtgat cctctgggtc ggtgccaatg actatctggc ctatggctgg 420  
aacacggagc aggatgccaa gcgggttcgc gatgccatca gcgatgcggc caaccgcatg 480  
gtactgaacg gtgccaagca gatactgctg ttcaacctgc cggatctggg ccagaacccg 540  
tcagctcgca gtcagaaggt ggtcgaggcg gtcagccatg tctccgccta tcacaaccag 600  
ctgctgctga acctggcacg ccagctggcc cccaccggca tggtaaagct gttcgagatc 660  
gacaagcaat ttgccgagat gctgctgat ccgcagaact tcggcctgag cgacgtcgag 720  
aaccctgct acgacggcgg ctatgtgtgg aagccgtttg ccaccgcag cgtcagcacc 780  
gaccgccagc tctccgcctt cagtccgcag gaacgcctcg ccatcgccgg caaccgcctg 840  
ctggcacagg ccgttgccag tcctatggcc cgccgcagcg ccagcccct caactgtgag 900  
ggcaagatgt tctgggatca ggtacaccg accactgtcg tgcacgcagc cctgagcgag 960  
cgcgccgcca ccttcacgc gaaccagtac gagttcctcg cccactga 1008

<210> 63  
<211> 1011  
5 <212> ADN  
<213> Aeromonas salmonicida

<400> 63

atgaaaaaat ggtttgtttg ttatttgggg ttgatcgcgc tgacagttca ggcagccgac 60  
actcgcgccg ccttctcccg gatcgtgatg ttcggcgaca gcctctccga taccggcaaa 120  
atgtacagca agatgcgcgg ttacctcccc tcagcccgcc cctactatga gggccgtttc 180  
tccaacggac ccgtctggct ggagcagctg accaagcagt tcccgggtct gaccatcgcc 240  
aacgaagcgg aagggcgtgc cactgccgtg gttacaaca agatctcctg gaatccaag 300  
tatcaggtca tcaacaacct ggactacgag gtcaccagc tcttgcaaaa agacagcttc 360  
aagccggacg atctggtgat cctctgggtc ggtgccaatg actatctggc atatggctgg 420  
aatacggagc aggatgccaa gcgagttcgc gatgccatca gcgatgcggc caaccgcatg 480  
gtactgaacg gtgccaagca gatactgctg ttcaacctgc cggatctggg ccagaacccg 540  
tcagcccga gtcagaaggt ggtcgaggcg gtcagccatg tctccgccta tcacaacaag 600  
ctgctgctga acctggcacg ccagctggcc cccaccggca tggtaaagct gttcgagatc 660  
gacaagcaat ttgccgagat gctgctgat ccgcagaact tcggcctgag cgacgtcgag 720  
aaccctgct acgacggcgg ctatgtgtgg aagccgtttg ccaccgcag cgtcagcacc 780  
gaccgccagc tctccgcctt cagtccgcag gaacgcctcg ccatcgccgg caaccgcctg 840  
ctggcacagg ccgttgccag tcctatggcc cgccgcagcg ccagcccct caactgtgag 900  
ggcaagatgt tctgggatca ggtacaccg accactgtcg tgcacgcagc cctgagcgag 960  
10 cgcgccgcca ccttcacga gaccagtac gagttcctcg cccacggatg a 1011

<210> 64  
<211> 51  
15 <212> ADN  
<213> Artificial

<220>  
<223> Secuencia terminador

# ES 2 449 474 T3

<400> 64  
 5 cgggacttac cgaaagaac catcaatgat ggttcttt ttgtcataa a 51  
 <210> 65  
 <211> 59  
 <212> ADN  
 <213> Artificial  
 10 <220>  
 <223> Secuencia terminador  
 <400> 65  
 15 caagactaaa gaccgttcgc ccgttttgc aataagcggg cgaatcttac ataaaaata 59  
 <210> 66  
 <211> 61  
 20 <212> ADN  
 <213> Artificial  
 <220>  
 <223> Secuencia terminador  
 25 <400> 66  
 acggccgtta gatgtgacag cccgttccaa aaggaagcgg gctgtcttcg tgtattattg 60  
 t 61  
 30 <210> 67  
 <211> 54  
 <212> ADN  
 <213> Artificial  
 35 <220>  
 <223> Secuencia terminador  
 <400> 67  
 40 tcttttaaag gaaaggctgg aatgcccggc attccagcca catgatcatc gttt 54  
 <210> 68  
 <211> 56  
 <212> ADN  
 45 <213> Artificial  
 <220>  
 <223> Secuencia terminador  
 50 <400> 68  
 gctgacaaat aaaaagaagc aqgtatggag gaacctgctt cttttacta ttattg 56  
 <210> 69  
 55 <211> 102  
 <212> ADN  
 <213> Artificial  
 <220>  
 60 <223> Cebador oligonucleótido  
 <400> 69

ES 2 449 474 T3

```

ccccgctcga ggcttttctt ttggaagaaa atatagggaa aatgggtactt gttaaaaatt      60

cggaatattt atacaatatc atatgtttca cattgaaagg gg                            102

<210> 70
<211> 35
5  <212> ADN
   <213> Artificial

<220>
<223> Cebador oligonucleótido
10 <400> 70

tggaatctcg aggtttatc ctttacctg tctcc                                     35

15 <210> 71
   <211> 8
   <212> PRT
   <213> Artificial

20 <220>
   <223> Secuencia de consenso

   <400> 71

  Met Arg Arg Ser Arg Phe Leu Ala
25  1           5

   <210> 72
   <211> 8
   <212> PRT
30 <213> Artificial

   <220>
   <223> Secuencia de consenso

35 <400> 72

  Ala Leu Ile Leu Leu Thr Leu Ala
   1           5

   <210> 73
   <211> 5
   <212> PRT
40 <213> Artificial

   <220>
45 <223> Secuencia de consenso

   <400> 73

  Ala Arg Ala Ala Pro
   1           5

50 <210> 74
   <211> 11
   <212> PRT
   <213> Artificial

55 <220>
   <223> Secuencia de consenso

   <400> 74

60 Tyr Val Ala Leu Gly Asp Ser Tyr Ser Ser Gly
   1           5           10

```

<210> 75  
 <211> 5  
 <212> PRT  
 5 <213> Artificial  
  
 <220>  
 <223> Secuencia de consenso  
  
 10 <400> 75  
  
**Gly Ala Gly Ser Tyr**  
**1 5**  
  
 <210> 76  
 15 <211> 4  
 <212> PRT  
 <213> Artificial  
  
 <220>  
 20 <223> Secuencia de consenso  
  
 <400> 76  
  
**Ser Ser Gly Asp**  
**1**  
 25  
 <210> 77  
 <211> 15  
 <212> PRT  
 <213> Artificial  
 30  
 <220>  
 <223> Secuencia de consenso  
  
 <400> 77  
 35  
**Arg Ser Thr Lys Ala Tyr Pro Ala Leu Trp Ala Ala Ala His Ala**  
**1 5 10 15**  
  
 <210> 78  
 <211> 5  
 40 <212> PRT  
 <213> Artificial  
  
 <220>  
 <223> Secuencia de consenso  
 45  
 <400> 78  
  
**Ser Ser Phe Ser Phe**  
**1 5**  
 50  
 <210> 79  
 <211> 12  
 <212> PRT  
 <213> Artificial  
  
 55 <220>  
 <223> Secuencia de consenso  
  
 <400> 79  
  
**Ala Cys Ser Gly Ala Arg Thr Tyr Asp Val Leu Ala**  
 60 **1 5 10**  
  
 <210> 80



<211> 15  
 <212> PRT  
 <213> Artificial

5 <220>  
 <223> Secuencia de consenso

<400> 80

**Leu Val Ser Ile Thr Ile Gly Gly Asn Asp Ala Gly Phe Ala Asp**  
 10 1 5 10 15

<210> 81  
 <211> 6  
 <212> PRT  
 <213> Artificial

15 <220>  
 <223> Secuencia de consenso

20 <400> 81

**Met Thr Thr Cys Val Leu**  
 1 5

<210> 82  
 <211> 6  
 <212> PRT  
 <213> Artificial

25 <220>  
 <223> Secuencia de consenso

30 <400> 82

**Ser Asp Ser Ala Cys Leu**  
 1 5

35 <210> 83  
 <211> 4  
 <212> PRT  
 <213> Artificial

40 <220>  
 <223> Secuencia de consenso

45 <400> 83

**Thr Leu Pro Ala**  
 1

<210> 84  
 <211> 9  
 <212> PRT  
 <213> Artificial

50 <220>  
 <223> Secuencia de consenso

55 <400> 84

**Arg Leu Asp Ser Val Tyr Ser Ala Ile**  
 1 5

60 <210> 85  
 <211> 4  
 <212> PRT

<213> Artificial

<220>  
<223> Secuencia de consenso

5 <400> 85

**Thr Arg Ala Pro**  
1

10 <210> 86  
<211> 12  
<212> PRT  
<213> Artificial

15 <220>  
<223> Secuencia de consenso

<400> 86

**Ala Arg Val Val Val Leu Gly Tyr Pro Arg Ile Tyr**  
20 1 5 10

<210> 87  
<211> 4  
<212> PRT

25 <213> Artificial

<220>  
<223> Secuencia de consenso

30 <400> 87

**Leu Gly Leu Ser**  
1

35 <210> 88  
<211> 11  
<212> PRT  
<213> Artificial

<220>  
40 <223> Secuencia de consenso

<400> 88

**Thr Lys Arg Ala Ala Ile Asn Asp Ala Ala Asp**  
45 1 5 10

<210> 89  
<211> 12  
<212> PRT  
<213> Artificial

50 <220>  
<223> Secuencia de consenso

<400> 89

55 **Leu Asn Ser Val Ile Ala Lys Arg Ala Ala Asp His**  
1 5 10

<210> 90  
<211> 7  
60 <212> PRT  
<213> Artificial

<220>  
 <223> Secuencia de consenso  
  
 <400> 90  
 5     **Gly Phe Thr Phe Gly Asp Val**  
       1                           5  
  
 <210> 91  
 <211> 7  
 10    <212> PRT  
       <213> Artificial  
  
 <220>  
 <223> Secuencia de consenso  
 15  
 <400> 91  
  
       **Gly His Glu Leu Cys Ser Ala**  
       1                           5  
 20    <210> 92  
       <211> 9  
       <212> PRT  
       <213> Artificial  
  
 25    <220>  
       <223> Secuencia de consenso  
  
       <400> 92  
  
       **Pro Trp Leu His Ser Leu Thr Leu Pro**  
 30    1                           5  
  
       <210> 93  
       <211> 6  
       <212> PRT  
 35    <213> Artificial  
  
       <220>  
       <223> Secuencia de consenso  
  
 40    <400> 93  
  
       **Ser Tyr His Pro Thr Ala**  
       1                           5  
  
 45    <210> 94  
       <211> 13  
       <212> PRT  
       <213> Artificial  
  
       <220>  
 50    <223> Secuencia de consenso  
  
       <400> 94  
  
       **Gly His Ala Ala Gly Tyr Leu Pro Val Leu Asn Ser Ile**  
       1                           5                           10  
 55    <210> 95  
       <211> 232  
       <212> PRT  
       <213> Aspergillus aculeatus  
 60    <400> 95

ES 2 449 474 T3

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

Gly Ser Gly Thr Asn Gly Trp Gly Glu Tyr Leu Ala Ser Tyr Leu Ser  
 20 25 30

Ala Thr Val Val Asn Asp Ala Val Ala Gly Arg Ser Ala Arg Ser Tyr  
 35 40 45

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

Asp Tyr Val Ile Val Glu Phe Gly His Asn Asp Gly Gly Ser Leu Ser  
 65 70 75 80

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

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

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

Ile Leu Ser Ser Gln Thr Pro Asn Asn Pro Trp Glu Thr Gly Thr Phe  
 130 135 140

Val Asn Ser Pro Thr Arg Phe Val Glu Tyr Ala Glu Leu Ala Ala Glu  
 145 150 155 160

Val Ala Gly Val Glu Tyr Val Asp His Trp Ser Tyr Val Asp Ser Ile  
 165 170 175

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

His Thr His Thr Ser Pro Ala Gly Ala Glu Val Val Ala Glu Ala Phe  
 195 200 205

Leu Lys Ala Val Val Cys Thr Gly Thr Ser Leu Lys Ser Val Leu Thr  
 210 215 220

Thr Thr Ser Phe Glu Gly Thr Cys  
 225 230

<210> 96  
 <211> 184  
 <212> PRT  
 <213> Escherichia coli

<400> 96

5

10

ES 2 449 474 T3

Ala Asp Thr Leu Leu Ile Leu Gly Asp Ser Leu Ser Ala Gly Tyr Arg  
 1 5 10 15

Met Ser Ala Ser Ala Ala Trp Pro Ala Leu Leu Asn Asp Lys Trp Gln  
 20 25 30

Ser Lys Thr Ser Val Val Asn Ala Ser Ile Ser Gly Asp Thr Ser Gln  
 35 40 45

Gln Gly Leu Ala Arg Leu Pro Ala Leu Leu Lys Gln His Gln Pro Arg  
 50 55 60

Trp Val Leu Val Glu Leu Gly Gly Asn Asp Gly Leu Arg Gly Phe Gln  
 65 70 75 80

Pro Gln Gln Thr Glu Gln Thr Leu Arg Gln Ile Leu Gln Asp Val Lys  
 85 90 95

Ala Ala Asn Ala Glu Pro Leu Leu Met Gln Ile Arg Leu Pro Ala Asn  
 100 105 110

Tyr Gly Arg Arg Tyr Asn Glu Ala Phe Ser Ala Ile Tyr Pro Lys Leu  
 115 120 125

Ala Lys Glu Phe Asp Val Pro Leu Leu Pro Phe Phe Met Glu Glu Val  
 130 135 140

Tyr Leu Lys Pro Gln Trp Met Gln Asp Asp Gly Ile His Pro Asn Arg  
 145 150 155 160

Asp Ala Gln Pro Phe Ile Ala Asp Trp Met Ala Lys Gln Leu Gln Pro  
 165 170 175

Leu Val Asn His Asp Ser Leu Glu  
 180

- 5 <210> 97
- <211> 308
- <212> PRT
- <213> Aeromonas hydrophila

- 10 <400> 97

ES 2 449 474 T3

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

Lys Met Arg Gly Tyr Leu Pro Ser Ser Pro Pro Tyr Tyr Glu Gly Arg  
 20 25 30

Phe Ser Asn Gly Pro Val Trp Leu Glu Gln Leu Thr Asn Glu Phe Pro  
 35 40 45

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

Tyr Asn Lys Ile Ser Trp Asn Pro Lys Tyr Gln Val Ile Asn Asn Leu  
 65 70 75 80

Asp Tyr Glu Val Thr Gln Phe Leu Gln Lys Asp Ser Phe Lys Pro Asp  
 85 90 95

Asp Leu Val Ile Leu Trp Val Gly Ala Asn Asp Tyr Leu Ala Tyr Gly  
 100 105 110

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

Ala Ala Asn Arg Met Val Leu Asn Gly Ala Lys Glu Ile Leu Leu Phe

ES 2 449 474 T3

130

135

140

Asn Leu Pro Asp Leu Gly Gln Asn Pro Ser Ala Arg Ser Gln Lys Val  
145 150 155 160

Val Glu Ala Ala Ser His Val Ser Ala Tyr His Asn Gln Leu Leu Leu  
165 170 175

Asn Leu Ala Arg Gln Leu Ala Pro Thr Gly Met Val Lys Leu Phe Glu  
180 185 190

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

Leu Ser Asp Gln Arg Asn Ala Cys Tyr Gly Gly Ser Tyr Val Trp Lys  
210 215 220

Pro Phe Ala Ser Arg Ser Ala Ser Thr Asp Ser Gln Leu Ser Ala Phe  
225 230 235 240

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

Ala Val Ala Ser Pro Met Ala Ala Arg Ser Ala Ser Thr Leu Asn Cys  
260 265 270

Glu Gly Lys Met Phe Trp Asp Gln Val His Pro Thr Thr Val Val His  
275 280 285

Ala Ala Leu Ser Glu Pro Ala Ala Thr Phe Ile Glu Ser Gln Tyr Glu  
290 295 300

Phe Leu Ala His  
305

<210> 98

<211> 167

5 <212> PRT

<213> Escherichia coli

<400> 98

Leu Leu Ile Leu Gly Asp Ser Leu Ser Ala Gly Tyr Arg Met Ser Ala  
1 5 10 15

Ser Ala Ala Trp Pro Ala Leu Leu Asn Asp Lys Trp Gln Ser Lys Thr  
20 25 30

10

Ser Val Val Asn Ala Ser Ile Ser Gly Asp Thr Ser Gln Gln Gly Leu  
35 40 45

ES 2 449 474 T3

Ala Arg Leu Pro Ala Leu Leu Lys Gln His Gln Pro Arg Trp Val Leu  
50 55 60

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

Thr Glu Gln Thr Leu Arg Gln Ile Leu Gln Asp Val Lys Ala Ala Asn  
85 90 95

Ala Glu Pro Leu Leu Met Gln Ile Arg Leu Pro Ala Asn Tyr Gly Arg  
100 105 110

Arg Tyr Asn Glu Ala Phe Ser Ala Ile Tyr Pro Lys Leu Ala Lys Glu  
115 120 125

Phe Asp Val Pro Leu Leu Pro Phe Phe Met Glu Glu Val Tyr Leu Lys  
130 135 140

Pro Gln Trp Met Gln Asp Asp Gly Ile His Pro Asn Arg Asp Ala Gln  
145 150 155 160

Pro Phe Ile Ala Asp Trp Met  
165

<210> 99

<211> 295

5 <212> PRT

<213> Aeromonas hydrophila

<400> 99

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

Lys Met Arg Gly Tyr Leu Pro Ser Ser Pro Pro Tyr Tyr Glu Gly Arg  
20 25 30

Phe Ser Asn Gly Pro Val Trp Leu Glu Gln Leu Thr Asn Glu Phe Pro  
35 40 45

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

Tyr Asn Lys Ile Ser Trp Asn Pro Lys Tyr Gln Val Ile Asn Asn Leu  
65 70 75 80

Asp Tyr Glu Val Thr Gln Phe Leu Gln Lys Asp Ser Phe Lys Pro Asp  
85 90 95

10 Asp Leu Val Ile Leu Trp Val Gly Ala Asn Asp Tyr Leu Ala Tyr Gly



ES 2 449 474 T3

100 105 110

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

Ala Ala Asn Arg Met Val Leu Asn Gly Ala Lys Glu Ile Leu Leu Phe  
130 135 140

Asn Leu Pro Asp Leu Gly Gln Asn Pro Ser Ala Arg Ser Gln Lys Val  
145 150 155 160

Val Glu Ala Ala Ser His Val Ser Ala Tyr His Asn Gln Leu Leu Leu  
165 170 175

Asn Leu Ala Arg Gln Leu Ala Pro Thr Gly Met Val Lys Leu Phe Glu  
180 185 190

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

Leu Ser Asp Gln Arg Asn Ala Cys Tyr Gly Gly Ser Tyr Val Trp Lys  
210 215 220

Pro Phe Ala Ser Arg Ser Ala Ser Thr Asp Ser Gln Leu Ser Ala Phe  
225 230 235 240

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

Ala Val Ala Ser Pro Met Ala Ala Arg Ser Ala Ser Thr Leu Asn Cys  
260 265 270

Glu Gly Lys Met Phe Trp Asp Gln Val His Pro Thr Thr Val Val His  
275 280 285

Ala Ala Leu Ser Glu Pro Ala  
290 295

<210> 100  
 <211> 335  
 <212> PRT  
 <213> Aeromonas hydrophila

<400> 100

Met Lys Lys Trp Phe Val Cys Leu Leu Gly Leu Val Ala Leu Thr Val  
1 5 10 15

Gln Ala Ala Asp Ser Arg Pro Ala Phe Ser Arg Ile Val Met Phe Gly  
20 25 30

5

10

ES 2 449 474 T3

Asp Ser Leu Ser Asp Thr Gly Lys Met Tyr Ser Lys Met Arg Gly Tyr  
 35 40 45  
 Leu Pro Ser Ser Pro Pro Tyr Tyr Glu Gly Arg Phe Ser Asn Gly Pro  
 50 55 60  
 Val Trp Leu Glu Gln Leu Thr Asn Glu Phe Pro Gly Leu Thr Ile Ala  
 65 70 75 80  
 Asn Glu Ala Glu Gly Gly Pro Thr Ala Val Ala Tyr Asn Lys Ile Ser  
 85 90 95  
 Trp Asn Pro Lys Tyr Gln Val Ile Asn Asn Leu Asp Tyr Glu Val Thr  
 100 105 110  
 Gln Phe Leu Gln Lys Asp Ser Phe Lys Pro Asp Asp Leu Val Ile Leu  
 115 120 125  
 Trp Val Gly Ala Asn Asp Tyr Leu Ala Tyr Gly Trp Asn Thr Glu Gln  
 130 135 140  
 Asp Ala Lys Arg Val Arg Asp Ala Ile Ser Asp Ala Ala Asn Arg Met  
 145 150 155 160  
 Val Leu Asn Gly Ala Lys Glu Ile Leu Leu Phe Asn Leu Pro Asp Leu  
 165 170 175  
 Gly Gln Asn Pro Ser Ala Arg Ser Gln Lys Val Val Glu Ala Ala Ser  
 180 185 190  
 His Val Ser Ala Tyr His Asn Gln Leu Leu Leu Asn Leu Ala Arg Gln  
 195 200 205  
 Leu Ala Pro Thr Gly Met Val Lys Leu Phe Glu Ile Asp Lys Gln Phe  
 210 215 220  
 Ala Glu Met Leu Arg Asp Pro Gln Asn Phe Gly Leu Ser Asp Gln Arg  
 225 230 235 240  
 Asn Ala Cys Tyr Gly Gly Ser Tyr Val Trp Lys Pro Phe Ala Ser Arg  
 245 250 255  
 Ser Ala Ser Thr Asp Ser Gln Leu Ser Ala Phe Asn Pro Gln Glu Arg  
 260 265 270  
 Leu Ala Ile Ala Gly Asn Pro Leu Leu Ala Gln Ala Val Ala Ser Pro  
 275 280 285  
 Met Ala Ala Arg Ser Ala Ser Thr Leu Asn Cys Glu Gly Lys Met Phe  
 290 295 300  
 Trp Asp Gln Val His Pro Thr Thr Val Val His Ala Ala Leu Ser Glu  
 305 310 315 320  
 Pro Ala Ala Thr Phe Ile Glu Ser Gln Tyr Glu Phe Leu Ala His  
 325 330 335

ES 2 449 474 T3

<210> 101  
 <211> 318  
 <212> PRT  
 <213> Aeromonas salmonicida

5

<400> 101

Ala Asp Thr Arg Pro Ala Phe Ser Arg Ile Val Met Phe Gly Asp Ser  
 1 5 10 15  
 Leu Ser Asp Thr Gly Lys Met Tyr Ser Lys Met Arg Gly Tyr Leu Pro  
 20 25 30  
 Ser Ser Pro Pro Tyr Tyr Glu Gly Arg Phe Ser Asn Gly Pro Val Trp  
 35 40 45  
 Leu Glu Gln Leu Thr Lys Gln Phe Pro Gly Leu Thr Ile Ala Asn Glu  
 50 55 60  
 Ala Glu Gly Gly Ala Thr Ala Val Ala Tyr Asn Lys Ile Ser Trp Asn  
 65 70 75 80  
 Pro Lys Tyr Gln Val Tyr Asn Asn Leu Asp Tyr Glu Val Thr Gln Phe  
 85 90 95  
 Leu Gln Lys Asp Ser Phe Lys Pro Asp Asp Leu Val Ile Leu Trp Val  
 100 105 110  
 Gly Ala Asn Asp Tyr Leu Ala Tyr Gly Trp Asn Thr Glu Gln Asp Ala  
 115 120 125  
 Lys Arg Val Arg Asp Ala Ile Ser Asp Ala Ala Asn Arg Met Val Leu  
 130 135 140  
 Asn Gly Ala Lys Gln Ile Leu Leu Phe Asn Leu Pro Asp Leu Gly Gln  
 145 150 155 160  
 Asn Pro Ser Ala Arg Ser Gln Lys Val Val Glu Ala Val Ser His Val  
 165 170 175  
 Ser Ala Tyr His Asn Lys Leu Leu Leu Asn Leu Ala Arg Gln Leu Ala  
 180 185 190

ES 2 449 474 T3

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

Met Leu Arg Asp Pro Gln Asn Phe Gly Leu Ser Asp Val Glu Asn Pro  
 210 215 220

Cys Tyr Asp Gly Gly Tyr Val Trp Lys Pro Phe Ala Thr Arg Ser Val  
 225 230 235 240

Ser Thr Asp Arg Gln Leu Ser Ala Phe Ser Pro Gln Glu Arg Leu Ala  
 245 250 255

Ile Ala Gly Asn Pro Leu Leu Ala Gln Ala Val Ala Ser Pro Met Ala  
 260 265 270

Arg Arg Ser Ala Ser Pro Leu Asn Cys Glu Gly Lys Met Phe Trp Asp  
 275 280 285

Gln Val His Pro Thr Thr Val Val His Ala Ala Leu Ser Glu Arg Ala  
 290 295 300

Ala Thr Phe Ile Glu Thr Gln Tyr Glu Phe Leu Ala His Gly  
 305 310 315

<210> 102  
 <211> 50  
 <212> PRT  
 <213> Artificial

<220>  
 <223> Secuencia de consenso

<400> 102

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

Thr Gly Lys Met Tyr Ser Lys Met Arg Gly Tyr Leu Pro Ser Ser Pro  
 20 25 30

Pro Tyr Tyr Glu Gly Arg Phe Ser Asn Gly Pro Val Trp Leu Glu Gln  
 35 40 45

Leu Thr  
 50

<210> 103  
 <211> 13  
 <212> PRT  
 <213> Artificial

<220>  
 <223> Secuencia de consenso

<400> 103

Phe Pro Gly Leu Thr Ile Ala Asn Glu Ala Glu Gly Gly  
 1 5 10

<210> 104

<211> 79  
 <212> PRT  
 <213> Artificial

5 <220>  
 <223> Secuencia de consenso

<400> 104

Thr Ala Val Ala Tyr Asn Lys Ile Ser Trp Asn Pro Lys Tyr Gln Val  
 1 5 10 15

Ile Asn Asn Leu Asp Tyr Glu Val Thr Gln Phe Leu Gln Lys Asp Ser  
 20 25 30

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

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

10 Ala Ile Ser Asp Ala Ala Asn Arg Met Val Leu Asn Gly Ala Lys  
 65 70 75

<210> 105  
 <211> 23  
 <212> PRT  
 <213> Artificial

15 <220>  
 <223> Secuencia de consenso

20 <400> 105

Ile Leu Leu Phe Asn Leu Pro Asp Leu Gly Gln Asn Pro Ser Ala Arg  
 1 5 10 15

Ser Gln Lys Val Val Glu Ala  
 20

25 <210> 106  
 <211> 8  
 <212> PRT  
 <213> Artificial

30 <220>  
 <223> Secuencia de consenso

<400> 106

Ser His Val Ser Ala Tyr His Asn  
 1 5

35 <210> 107  
 <211> 38  
 <212> PRT  
 <213> Artificial

40 <220>  
 <223> Secuencia de consenso

<400> 107

45

ES 2 449 474 T3

Leu Leu Leu Asn Leu Ala Arg Gln Leu Ala Pro Thr Gly Met Val Lys  
 1 5 10 15

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

Asn Phe Gly Leu Ser Asp  
 35

5 <210> 108  
 <211> 7  
 <212> PRT  
 <213> Artificial

10 <220>  
 <223> Secuencia de consenso  
 <400> 108

Tyr Val Trp Lys Pro Phe Ala  
 1 5

15 <210> 109  
 <211> 5  
 <212> PRT  
 <213> Artificial

20 <220>  
 <223> Secuencia de consenso  
 <400> 109

25 Gln Leu Ser Ala Phe  
 1 5

30 <210> 110  
 <211> 22  
 <212> PRT  
 <213> Artificial

<220>  
 <223> Secuencia de consenso  
 35 <400> 110

Pro Gln Glu Arg Leu Ala Ile Ala Gly Asn Pro Leu Leu Ala Gln Ala  
 1 5 10 15

Val Ala Ser Pro Met Ala  
 20

40 <210> 111  
 <211> 4  
 <212> PRT  
 <213> Artificial

45 <220>  
 <223> Secuencia de consenso  
 <400> 111

50 Arg Ser Ala Ser  
 1

<210> 112

ES 2 449 474 T3

<211> 24  
<212> PRT  
<213> Artificial

5 <220>  
<223> Secuencia de consenso

<400> 112

**Leu Asn Cys Glu Gly Lys Met Phe Trp Asp Gln Val His Pro Thr Thr**  
**1 5 10 15**

**Val Val His Ala Ala Leu Ser Glu**  
**20**

10

<210> 113  
<211> 5  
<212> PRT  
15 <213> Artificial

<220>  
<223> Secuencia de consenso

20 <400> 113

**Ala Ala Thr Phe Ile**  
**1 5**

25 <210> 114  
<211> 7  
<212> PRT  
<213> Artificial

<220>  
30 <223> Secuencia de consenso

<400> 114

**Gln Tyr Glu Phe Leu Ala His**  
**1 5**

35 <210> 115  
<211> 1225  
<212> ADN  
<213> Artificial

40 <220>  
<223> Inserto XhoI que contiene el gen precursor LAT-KLM3'

<220>  
45 <221> CDS  
<222> (101)..(1144)

<400> 115

ES 2 449 474 T3

gcttttcttt tggaagaaaa tatagggaaa atggtacttg ttaaaaattc ggaatattta	60
tacaatatca tatgtttcac attgaaaggg gaggagaatc atg aaa caa caa aaa	115
Met Lys Gln Gln Lys	
1 5	
cgg ctt tac gcc cga ttg ctg acg ctg tta ttt gcg ctc atc ttc ttg	163
Arg Leu Tyr Ala Arg Leu Leu Thr Leu Leu Phe Ala Leu Ile Phe Leu	
10 15 20	
ctg cct cat tct gca gct tca gca gca gat aca aga ccg gcg ttt agc	211
Leu Pro His Ser Ala Ala Ser Ala Ala Asp Thr Arg Pro Ala Phe Ser	
25 30 35	
cgg atc gtc atg ttt gga gat agc ctg agc gat acg ggc aaa atg tat	259
Arg Ile Val Met Phe Gly Asp Ser Leu Ser Asp Thr Gly Lys Met Tyr	
40 45 50	
agc aaa atg aga ggc tat ctt ccg tca agc ccg ccg tat tat gaa ggc	307
Ser Lys Met Arg Gly Tyr Leu Pro Ser Ser Pro Pro Tyr Tyr Glu Gly	
55 60 65	
cgc ttt agc aat gga ccg gtc tgg ctg gaa caa ctg acg aaa caa ttt	355
Arg Phe Ser Asn Gly Pro Val Trp Leu Glu Gln Leu Thr Lys Gln Phe	
70 75 80 85	
ccg gga ctg acg atc gct aat gaa gca gaa gga gga gca aca gcg gtc	403
Pro Gly Leu Thr Ile Ala Asn Glu Ala Glu Gly Gly Ala Thr Ala Val	
90 95 100	
gcc tat aac aaa atc agc tgg gac ccg aaa tat cag gtc atc aac aac	451
Ala Tyr Asn Lys Ile Ser Trp Asp Pro Lys Tyr Gln Val Ile Asn Asn	
105 110 115	
ctg gac tat gaa gtc aca cag ttt ctt cag aaa gac agc ttt aaa ccg	499
Leu Asp Tyr Glu Val Thr Gln Phe Leu Gln Lys Asp Ser Phe Lys Pro	
120 125 130	
gat gat ctg gtc atc ctt tgg gtc ggc gcc aat gat tat ctg gcg tat	547
Asp Asp Leu Val Ile Leu Trp Val Gly Ala Asn Asp Tyr Leu Ala Tyr	
135 140 145	
ggc tgg aac aca gaa caa gat gcc aaa aga gtc aga gat gcc atc agc	595
Gly Trp Asn Thr Glu Gln Asp Ala Lys Arg Val Arg Asp Ala Ile Ser	
150 155 160 165	
gat gcc gct aat aga atg gtc ctg aac ggc gcc aaa caa atc ctg ctg	643



ES 2 449 474 T3

Asp Ala Ala Asn Arg Met Val Leu Asn Gly Ala Lys Gln Ile Leu Leu  
 170 175 180

ttt aac ctg ccg gat ctg gga caa aat ccg agc gcc aga agc caa aaa 691  
 Phe Asn Leu Pro Asp Leu Gly Gln Asn Pro Ser Ala Arg Ser Gln Lys  
 185 190 195

gtc gtc gaa gca gtc agc cat gtc agc gcc tat cat aac aaa ctg ctg 739  
 Val Val Glu Ala Val Ser His Val Ser Ala Tyr His Asn Lys Leu Leu  
 200 205 210

ctg aac ctg gca aga caa ttg gca ccg acg gga atg gtt aaa ttg ttt 787  
 Leu Asn Leu Ala Arg Gln Leu Ala Pro Thr Gly Met Val Lys Leu Phe  
 215 220 225

gaa att gac aaa cag ttt gcc gaa atg ctg aga gat ccg caa aat ttt 835  
 Glu Ile Asp Lys Gln Phe Ala Glu Met Leu Arg Asp Pro Gln Asn Phe  
 230 235 240 245

ggc ctg agc gat gtc gaa aac ccg tgc tat gat ggc gga tat gtc tgg 883  
 Gly Leu Ser Asp Val Glu Asn Pro Cys Tyr Asp Gly Gly Tyr Val Trp  
 250 255 260

aaa ccg ttt gcc aca aga agc gtc agc acg gat aga caa ctg tca gcg 931  
 Lys Pro Phe Ala Thr Arg Ser Val Ser Thr Asp Arg Gln Leu Ser Ala  
 265 270 275

ttt agc ccg caa gaa aga ctg gca atc gcc gga aat ccg ctt ttg gca 979  
 Phe Ser Pro Gln Glu Arg Leu Ala Ile Ala Gly Asn Pro Leu Leu Ala  
 280 285 290

caa gca gtt gct tca ccg atg gca aga aga tca gca agc ccg ctg aat 1027  
 Gln Ala Val Ala Ser Pro Met Ala Arg Arg Ser Ala Ser Pro Leu Asn  
 295 300 305

tgc gaa ggc aaa atg ttt tgg gat cag gtc cat ccg aca aca gtt gtc 1075  
 Cys Glu Gly Lys Met Phe Trp Asp Gln Val His Pro Thr Thr Val Val  
 310 315 320 325

cat gct gcc ctt tca gaa aga gcg gcg acg ttt atc gaa aca cag tat 1123  
 His Ala Ala Leu Ser Glu Arg Ala Ala Thr Phe Ile Glu Thr Gln Tyr  
 330 335 340

gaa ttt ctg gcc cat ggc tga gttaacagag gacggatttc ctgaaggaaa 1174  
 Glu Phe Leu Ala His Gly  
 345

tccgtttttt tattttaagc ttggagacaa ggtaaaggat aaaacctcga g 1225

<210> 116  
 <211> 347  
 <212> PRT  
 <213> Artificial

5

<220>  
 <223> Constructo sintético

10

<400> 116

Met Lys Gln Gln Lys Arg Leu Tyr Ala Arg Leu Leu Thr Leu Leu Phe  
 1 5 10 15

Ala Leu Ile Phe Leu Leu Pro His Ser Ala Ala Ser Ala Ala Asp Thr

ES 2 449 474 T3

20 25 30

Arg Pro Ala Phe Ser Arg Ile Val Met Phe Gly Asp Ser Leu Ser Asp  
 35 40 45

Thr Gly Lys Met Tyr Ser Lys Met Arg Gly Tyr Leu Pro Ser Ser Pro  
 50 55 60

Pro Tyr Tyr Glu Gly Arg Phe Ser Asn Gly Pro Val Trp Leu Glu Gln  
 65 70 75 80

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

Gly Ala Thr Ala Val Ala Tyr Asn Lys Ile Ser Trp Asp Pro Lys Tyr  
 100 105 110

Gln Val Ile Asn Asn Leu Asp Tyr Glu Val Thr Gln Phe Leu Gln Lys  
 115 120 125

Asp Ser Phe Lys Pro Asp Asp Leu Val Ile Leu Trp Val Gly Ala Asn  
 130 135 140

Asp Tyr Leu Ala Tyr Gly Trp Asn Thr Glu Gln Asp Ala Lys Arg Val  
 145 150 155 160

Arg Asp Ala Ile Ser Asp Ala Ala Asn Arg Met Val Leu Asn Gly Ala  
 165 170 175

Lys Gln Ile Leu Leu Phe Asn Leu Pro Asp Leu Gly Gln Asn Pro Ser  
 180 185 190

Ala Arg Ser Gln Lys Val Val Glu Ala Val Ser His Val Ser Ala Tyr  
 195 200 205

His Asn Lys Leu Leu Leu Asn Leu Ala Arg Gln Leu Ala Pro Thr Gly  
 210 215 220

Met Val Lys Leu Phe Glu Ile Asp Lys Gln Phe Ala Glu Met Leu Arg  
 225 230 235 240

Asp Pro Gln Asn Phe Gly Leu Ser Asp Val Glu Asn Pro Cys Tyr Asp  
 245 250 255

Gly Gly Tyr Val Trp Lys Pro Phe Ala Thr Arg Ser Val Ser Thr Asp  
 260 265 270

Arg Gln Leu Ser Ala Phe Ser Pro Gln Glu Arg Leu Ala Ile Ala Gly  
 275 280 285

ES 2 449 474 T3

Asn Pro Leu Leu Ala Gln Ala Val Ala Ser Pro Met Ala Arg Arg Ser  
 290 295 300

Ala Ser Pro Leu Asn Cys Glu Gly Lys Met Phe Trp Asp Gln Val His  
 305 310 315 320

Pro Thr Thr Val Val His Ala Ala Leu Ser Glu Arg Ala Ala Thr Phe  
 325 330 335

Ile Glu Thr Gln Tyr Glu Phe Leu Ala His Gly  
 340 345

- 5 <210> 117
- <211> 4
- <212> PRT
- <213> Artificial
- 10 <220>
- <223> Resto secuencia
- 15 <220>
- <221> SITIO
- <222> (4) .. (4)
- <223> Xaa puede ser uno o mas de los siguientes restos de aminoácidos L, A, V, I, F, Y, H, Q, T, N, M o S.
- <400> 117
- Gly Asp Ser Xaa**
- 1**
- 20 <210> 118
- <211> 5
- <212> PRT
- <213> Artificial
- 25 <220>
- <223> Resto secuencia
- 30 <220>
- <221> SITIO
- <222> (2) .. (2)
- <223> Puede ser Gly
- 35 <220>
- <221> SITIO
- <222> (5)..(5)
- <223> Puede ser Ala o Leu
- <400> 118
- Gly Ala Asn Asp Tyr**
- 1 5**

**REIVINDICACIONES**

1. Un método para la producción de una lipido aciltransferasa que comprende las etapas de:
  - (i) proporcionar una célula de *Bacillus*, en la que la célula de *Bacillus* es *B. licheniformis*;
  - 5 (ii) transformar la célula de *Bacillus*, en la que la célula de *Bacillus* es *B. licheniformis* con una secuencia de nucleótidos heteróloga que codifica una lipido aciltransferasa mostrada como SEQ ID N° 49; y
  - (iii) expresar la lipido aciltransferasa en la célula bajo control de una secuencia promotora.
2. Un método según la reivindicación 1 en el que la secuencia promotora no está asociada de forma nativa con la secuencia de nucleótidos que codifica una lipido aciltransferasa.
- 10 3. Un método según cualquiera de las reivindicaciones precedentes, en el que dicho método comprende la etapa adicional de aislar/recuperar la lipido aciltransferasa.
4. Un método según cualquiera de las reivindicaciones precedentes en el que dicha secuencia promotora es homóloga respecto de la célula hospedadora.
- 15 5. Un método según cualquiera de las reivindicaciones precedentes en el que dicha secuencia promotora se selecciona del grupo que consiste en una secuencia promotora de  $\alpha$ -amilasa, una secuencia promotora de proteasa, una secuencia promotora de subtilisina, una secuencia promotora de proteasa específica de ácido glutámico y una secuencia promotora de levansacarasa.
6. Una lipido aciltransferasa obtenible mediante el método de cualquiera de las reivindicaciones precedentes y codificada por SEQ ID N° 49.

## FIGURA 1 (SEQ ID N° 16)

```

1  ADTREAFSRI VMFGDSLSDT GKMYSKMRGY LPSSPPYYEG RFSNGPVWLE QLTQFPGLT
61  IANEAREGGAT AVAYNKLSWD PKYQVINNLD YEVTQFLQKD SEKPDDLVLV WVGANDYLAY
121  GWNTEQDAKR VRDAISDAAN RMVINGAKQI LLENLPDLGQ NPSARSQKVV EAVSHVSAYH
181  NKLLLNLARQ LAPTEGMVKLF EIDKQFAEML RDPQNFGLSD VENPCYDGGY VWKPFATRSV
241  STDRQLSAFS PQRERLAIAGN PLLAQAVASP MARRSASPLN CEGKMFWDQV RPTTVVHAAL
301  SERAATFIET QYEFLAHG

```

## FIGURA 2 (SEQ ID N° 1)

```

1  MKKWFVCLLG LVALTVQAAD SRPAFSRIVM FGDSLSDTGK MYSKMRGYLP
51  SSPYYEGRF SNGFPVWLEQL TKQFPGLTIA NEAEGCATAV AYNKISWNP
101  YQVINNLDYE VTQFLQKDSF KPDDLVLVW GANDYLAYGW NTEQDAKRVR
151  DAISDAANRM VINGAKQILL FNLPDLGQNP SARSQKVVEA VSHVSAYHNQ
201  LLNINLARQIA PTGMVKLFEI DKQFAEMLRD PQNFGLSDVE NPCYDGGYVW
251  KPFATRSVST DRQLSAFSPQ ERLAIAGNPL LAQAVASPMA RRSASPLNCE
301  GKMFWDQVHP TTVVHAALSE RAATFIANQY EFLAH*

```

FIGURA 3 (SEQ ID N° 2)

```

1 ivafGD5LTd geayygsdsdg ggwgagladr Ltallrlrar prgvdvfnrg isGrtsdGrl
61 ivDalvallF laqslglpnL pPYLsgdflr GANFAsagAt Ilptsgpfli QvqFkdksq
121 vlelrqalgl lqellrlpv ldakspdlvt imiGtNDlit saffgpkste sdrnsvpaf
181 kdnlrqlikr lrsnngarii vlitlvilnl gplGClPikl alalassknv dasgclerln
241 eavadfneal relaiskled qlrkdgldpv kgadvpyvDl ysifqldgi qnpsayvyGF
301 ettkaCCGyG gryNynrvCG naglcnvtak aCnpssylls flfw5gff5ps ekGykavAea
361 1

```

FIGURA 4 (SEQ ID N° 3)

```

1 mkkwfvcllg lvaltvqaad srpafsrivm fgdsldtgk mysknrgylp sspyyegrf
61 sngpvwleql tnefpgltia neaeggptav aynkiswnpk yqvinnldye vtqflqkdsf
121 kpddlvilwv gandylaygw nteqdarvr daisdaanzm vlngakeill fnlpldgqp
181 sarsqkvvea ashvsayhnq lllnlarqla ptgmvklfei dkqfaemlrd pqrfglsdqr
241 nacyggsyvw kpfarsast dsqlsafnpq erlaiagnpl laqavaspma arsastlnce
301 gkmfwdgvhp ttvvhalse paatfiesqy eflah

```

FIGURA 5 (SEQ ID N° 4)

```

1 mkkwfvcllg lialtvqaad trpafsrivm fgdsldstgk myskmrgylp sspyyegrf
61 sngpvwleql tkqfpgltia neaeggatav ayukiswnpk yqvynldye vtgflqkdsf
121 kpddlvilwv gandylaygw ntegdakrvr daisdaanrm vlngakqill fnlpdlgqnp
181 sarsqkvvea vshvsayhmk lllnlargla ptgmvklfei dkqfaemlrd pqnfglsdve
241 npcydgggyvw kpfatrsvst drqlsafspg erlaiagnpl laqavaspmma rrsasplnce
301 gkmfwdqwhp tivvhaalse raatfietqy eflahg

```

FIGURA 6 (SEQ ID N° 5)

```

1 mpkpalrrvm tatvaavgtl algltdatah aapaqatptl dyvalgdsys agsgvlpvdp
61 anllclrsta nyphviadtt garltdvtcg aqtdadftra qyvgvapqld algtgtdlvt
121 ltiggndnst finaitacgt agvlsqgkgs pckdrhgtsf ddeieantyp alkeallgvr
181 arapharvaa lgypwitpat adpscflklp laagdvpylr aiqahlnlav rraaeetgat
241 yvdfsgvsdg hdaceapgtr wiepllfghs lvpvhpnalg errmaehtmd vlgld

```

FIGURA 7 (SEQ ID N° 6)

```

1 mpkpalrrvm tatvaavgtl algltdatah aapaqatptl dyvalgdsys agsgvlpvdp
61 anllclrsta nyphviadtt garltdvtcg aqtdadftra qyvgvapqld algtgtdlvt
121 ltiggndnst finaitacgt agvlsqgkgs pckdrhgtsf ddeieantyp alkeallgvr
181 arapharvaa lgypwitpat adpscflklp laagdvpylr aiqahlnlav rraaeetgat
241 yvdfsgvsdg hdaceapgtr wiepllfghs lvpvhpnalg errmaehtmd vlgld

```

FIGURA 8 (SEQ ID N° 7)

```

1 mdyekfllfg dsitefafnt rpiedgkdqy algaalvney trkmdilcrg fkgysrwal
61 kilpeilkhe snivmatifl gandacsagp qsvplpefid nirqmvslnk syhirpliig
121 pglvdrekwe kekseeialg yfrtnenfai ysdalaklan eekvpfvaln kafqgegda
181 wqqltdgih fsgkgykifh dellkvietf ypyhpknmq yklkdwrdvl ddgsnims

```

FIGURA 9 (SEQ ID N° 8)

10	20	30	40	50	60
MNLRQWAGAA	TAALALGLAA	CGGGGTDQSG	NPNVAKVQRM	VVFGDSLSDI	GTYTPVAQAV
70	80	90	100	110	120
GGGKFTTNP	PIWAETVAAQ	LCVTLTPAVM	GYATSVQNC	KAGCFDYAQS	GSRVTDPNGI
130	140	150	160	170	180
GHNGGAGALT	YFVQQQLANF	YAASNTFNG	NNDVVFVLAG	SNDIFFWTTA	AATSGSGVTP
190	200	210	220	230	240
AIATAQVQQA	ATDLVGYVKD	MIAKGATQVY	VFNLPDSSLT	PDGVASGTTG	QALLHALVGT
250	260	270	280	290	300
FNTTLOSGLA	GTSARIIDFN	AQLTAAIQNG	ASFGFANTSA	RACDATKINA	LVPSAGESSL
310	320	330	340		
FCSANFLVAS	GADQSYLFAD	GVHPTTAGHR	LIASNVLARL	LADNVAH	



FIGURA 10 (SEQ ID N° 9)

```

1 migsvyavgd sftegvqdpd pdgafvgwad rlavlladrr pegdftytnl avrgrlldqi
61 vaeqvprvvg lapdlvsfaa ggndiirpgt dpdevaerfe lavaalataa gtvltvtgfd
121 trgvpvkhl rgkiatyngv vraiadrygc pvldlwslrs vqdrrawdada rlbhspeght
181 rvalraggal glrvpadpdq pwpplpprgt ldvrrddvhw areylvpwig rrlrgessgd
241 hvtakgtlsp daiktriaav a

```

FIGURA 11 (SEQ ID N° 10)

```

1 mqtncpaytsl vavgdsftg msdllpdgsy rgwadllatr maarspgfry anlavrgkli
61 gqivdeqvvd aaamgadvit lvgglndtir pkcōmarvrd lltqaverla phceqlvlmr
121 spgrgppvle rfrpmealf aviddlagrh gavvvdlyga qsladpzmwd vdrlhltagv
181 hrrvaeavwg slghepedpe whapipatpp pgwvtrrtad vrfarqhlpp wigrrltgrs
241 sgdglpakrp dllypedpar

```

FIGURA 12 (SEQ ID N° 11)

```

1 mtrgrdggag apptkhrall aaivtlivai saaiyagasa ddgsrdhalq aggrlprgda
61 apastgawvg awatapaaae pgtettglag rsvrnvvhts vggtagaritl snlyggsplt
121 vthasialaa gpdtaaaiaa tmrrltfjgs arviipaggq vmsdtaarai pyganvlvt
181 yspipsqpvvt yhpqarqtsy ladgdrtadv tavayttptp ywryltaldv lsheadgtvv
241 afgdsitdga rsgsdanhrw tdvlaarlhe aagdrdtpr ysvvnegisg nrlltsrpgx
301 padnpglsr fqrđvlertn vkavvvvlgv ndvlnspela drdailtqlr tlvdraharg
361 lrvvgatitp fggygytea retmrqevne eirsgrvfdt vvdfdkalrd pydprzmsd
421 ydsgdhllpg dkgymrgav idlaalkgea pvka

```

FIGURA 13 (SEQ ID N° 12)

```

1 mtsmsrarva rriaagaayg gggiglagaa avglvvaevq larrrvvgvt ptrvpnaqgl
61 yggtlptagd pplrlmmlgd staagqgvhr agqtpgalla sglaavaerp vrlgsvaqqp
121 acsddldrqv alvlaepdrv pdicvimvga ndvthmpat rsvrhlssav rrlrtagaev
181 vvgtcpdlgt iervrqplrw larrasrqla aaqtigaveq ggrrvslgdl lgpefaqnpr
241 elfgpōnyhp saegyataam avlpsvcaal glwpadeehp dalrregflp varaaeaaas
301 eagtevaaam ptgprgpwal lkrrrrrrrs eaepsspsgv

```

FIGURA 14 (SEQ ID N° 13)

```

1 mgrgtđqrtr yrrrrarval aaltaavlgv gvagcdsvgg dspapsgps krtrtapawd
61 tspasvaavg dsitrgfdac avlsdcpevs watgssakvd slavrllgka daaehswnya
121 vtgarmadlt aqvtraaqre pelvavmaga ndacrsttsa mtpvadfraa feeamatlrk
181 klpkavvyvs sipdlkrlws qgrtnplgkq vwklglcpm lgdadslđsa atlrntvrd
241 rvadynevlr evcakdrrcr sddgavhefr fgtdqlshwd wfhpsvdgga rlaeaiayrav
301 taknp

```

FIGURA 15 (SEQ ID N° 14)

```

1 mrlsrraata sallltpala lfgasaavsa priqatdyva lgdsyssgvg agsydsssqs
61 ckrstksypa lwaashtgtr inftacsgr tgdlvakqit pvnsqtdlvs itiggndagf
121 adtmittcnlq gesaclaria karayiqqtl paqldqvyda idsrapaaqv vvlgyprfyk
181 lggscavgl s eksraaina addinavtak raadhgfafg dvnttfaghe lcsgapwlhs
241 vtlpvensyh ptangqskgy lpvlnsat
    
```

FIGURA 16 (SEQ ID N° 15)

```

1 MKKWFVCLLG LIALTVQAAD TRPAFSRIVM FGDSLSDTGK MYSKMRGYLP
51 SSPPYEGRE SNGPVWLEQL TKQFPGLTIA NEAEGGATAV AYNKISWNPK
101 YQVINNDYE VTQFLQKDSF KPDDLVLWV GANDYLAYGW NTEQDAKRVR
151 DAISDAANRM VLNGAKQILL FNLPDLQNP SARSQKVEA VSHVSAYHMK
201 LLINLARQLA PTGMVKLEFI DKQFAEMLRD PQNFGLSDVE NPCYDGGYVW
251 KPFAFRSVST DRQLSAFSPQ ERLAIAGNPL LAQAVASPMA RRSASPLNCE
301 GKMFWQVHP TTVVHAALSE RAATFIETQY EFLANG*
    
```

FIGURA 17 (SEQ ID N° 19)

```

1 migsyvavgd sftegvgdpg pdgafvgwad rlavlladdr pegdftytnl avrgrlldqi
61 vaeqvprvvg lapdlvsfaa ggndiirpqt dpdevaerfe lavaalaaa gtvlvttgfd
121 trgvplkhl rgkياتyngh vraiadrygc pvldlwslrs vqdrrawd ad rhlspgght
181 rvalragqal glrvpadpcq pwpplpprgt ldvrrddvhw areylvpwig rrlrgessgd
241 hvtakgtlsp daiktriaav a
    
```

FIGURA 18 (SEQ ID N° 25)

```

1  MFKFKKNFLV  GLSAAALMSIS  LFSATASAAS  ADSREAFSRI  VMFGDSLSDT
51  GKMYSKMRGY  LPSSPFYYEG  RFSNGFPVWLE  QLTQKQFPGIT  IANEAEGGAT
101 AVAYNKISWN  PKYQVINNLD  YEVTQFLQKD  SFKPDDLVLIL  WVGANDYLAY
151 GWNTEQDAKR  VRDAISDAAN  RMVLNGAKQI  LLENLPLDGGQ  NPSARSQKVV
201 EAVSHVSAYH  NQLLLNLRQ  LAPTCMVKLF  EIDKQFAEML  RDPQNFGLSD
251 VENPCYDGGY  VWKPFATRSV  STDRQLSAFS  PQERLAIAGN  PLLAQAVASP
301 MARRSASPLN  CEGKMFWDQV  HPTTVVHAAL  SERAATFIAN  QYEFIAH**

```

FIGURA 19 (SEQ ID N° 26)

```

MRLTRLSAASVIVFALLLLALLGISPAQAAGPAYVALGDSYSSGNGAGSYIDS SGDCHRSN
NAYPARWAAAANAPSSFTFAACSGAVTTDVINNQLGALNASTGLVSTITIGNDAGFADAMTT
CVTSSDSTCLNRLATATNYINTLLARLDAVYSQIKARAPNARVVVLGYPRMYLASNPWYC
LGLSNTKRAAINTTADTLNSVISSRATAHGFRFGDVRPTFNNHELFFGNDWLHSLTLPVWE
SYHPTSTGHQSGYL PVLNANSST

```

FIGURA 20 (SEQ ID N° 27)

ZP 00058717

```

1 mlphpagerg evgaffallv gtpqdrllrl echetrplrg rcgcgerrvp pltlpgdgvI
61 cttsstrdae tvwrkhlqpr pdggfrrphlg vcollagqgs pgvlwcgreg crfevcrirdt
121 pglstrngd ssppfragws lppkogeisq sarktpavpr ysllrtdrpd gprgrfvgsq
181 praatrrrlf lgipalvlvt altlviavpt gretlwrnwc eatqdwclgv pvdsrgqpae
241 dgeflllspv qaatwgnyya lgdsyssgdg ardyypgtav kggcwsana ypelvaeayd
301 faghlsflac sgqrgyamld aidevgsqld wnsphsylv igiggndlgf stvlktcmvr
361 vplldskact dqsdairkrn akfettfeel isevrtrapd arilvgypr ifpeeptgay
421 ytltasnqrw lnetiqefnq qlaeavavhd eeiaasggvg svefvdvyha ldgheigsde
481 pwnvgvqlrd latgvtvdrs tfhpnaaghr avgervieqi etgpprplya tfavvagatv
541 dtlagevg
    
```

FIGURA 21 (SEQ ID N° 28)

```

1 mgsgpraatr rrlflgipal vlvtaltlvi avptgretlw mwceatqdw clgvpvdsrg
61 qpaedgefll lspvqaatwg nyyalgdsys sgdgardyyp gtavkkgcwr sanaypelva
121 eaydfaghls flacsgqrgy amldaidevg sqldwnspht slvtigiggn dlqfstvlkt
181 cmvrvplls kactdqedai rkrmakfett feelisevrt rapdarilv gyprifpeep
241 tgayyiltas nqrwlnetiq efnqqlaeav avhdeeiaas ggvgsvfvd vyhaldghei
301 gsdepwngv qlrdlatgvt vdrstfhpna aghravgerv ieqietgppr plyatfavva
361 gatvdtlage vg
    
```

**FIGURA 22 (SEQ ID N° 29)**

```

1 mrttviaasa llllagcadg areetagapp gessggiree gaeastsitd vyialgdsya
61 amggrdqpir gepfclrsg nypellhaev tdltcggavt gdlleprtlg ertlpaqvda
121 ltedttlvtl siggndlgfg evagcireri agenaddcvd ligetigeql dqlppqldrv
181 heairdragd aqvvtgylp lvsagdcpel gdvseadrw aveltggine tvreaaerhd
241 alfvlpddad ehtscappqg rwadicggqt dayplhptsa gneamaaavr dalglepvqp

```

**FIGURA 23 (SEQ ID N° 30)**

ZP\_00094165

```

1 mgqvklfarr capvllalag lapaatvare aplaegaryv algssfaagp gvgpnapgsp
61 ercgrgtlly phllaealkl dlvdatsga tthhvlgpwn evppqidsvn gdtrlvtlti
121 ggndvsfvgn ifaaacekma spdprcckwr eiteewqad eermrsivrq iharaplarv
181 vvdyyitvlp psgtcaamai spdrlaqsrs aakrlarita rvareegasl lkfshisrrh
241 hpcsakpwsn glsapaddgi pvhpnrlgha eaaaalvklv klnk //

```

**FIGURA 24 (SEQ ID N° 31)**

NP\_625998.

```

1 mrrfrlvgfl sslvlaagaa ltgaataqaa qpaaadgyva lgdsyssgvg agsyisssgd
61 ckrstkahpy lwaaahpst fdftacsgar tgdvlsqqlg plssgtglvs isiggn dagf
121 admttcvlq sessclsria taeayvdstl pgkldgvysa isdkapnahv vvigyprfyk
181 lgttciglse tkrtainkas dhlntvlaqr aaahgftfgd vrttftghel csgspwlhsv
241 nwlalnigesyh ptaagqsggy lpvlnгаа

```

//

**FIGURA 25 (SEQ ID N° 32)**

NP\_827753.

```

1 mrrsritayv tslllavgca ltgaataqas paaaatgyva lgdsyssgvg agsylsssgd
61 ckrskaypy lwqaahspss fsfmacsgar tgdvlnqlg tlnsstglvs ltiggn dagf
121 sdvmttcvlq sdsaclsrin takayvdstl pgqldsvyta istkapsahv avlgyprfyk
181 lggscлагls etkrsainda adylnsaiak raadhgftfg dvkstftghe icssstwlhs
241 ldllnigqsy hptaagqsgg ylpvnmnsva

```

//

**FIGURA 26 (SEQ ID N° 33)**

```

MRLTRSLSAASVIVFALLLALLGISAQAAGPAYVALGDSYSSGNGAGSYIDS SSGDCHRSN
NAYPARWAAANAPSSFTFAACSGAVTTDVINNQLGALNASTGLVSIITIGGNDAGFADAMTT
CVTSSDSTCLNRLATATNYINTLLARLDAVYSQIKARAPNARVVVLGYPRMYLASNPWYC
LGLSNTKRAAINTTADTLNSVISSRATAHGFRFGDVRPTFNHLEFFGNDWLHSLTLPVWE
SYHPTSTGHQSGYLEVLNANSSST

```

FIGURA 27 (SEQ ID N° 34)

ADSRPAFSRIVMFGDSLSDTGKMYSKMRGYLPSSPPYYEGRFSNGPVWLEQLTNEFPGLTIANEAEGGPT  
 AVAYNKISWNEPKYQVINNLDYEVTOFLQKDSFKPDDLVLWVGANDYLAYGWNTAQDAKVRDAISDAAN  
 RMVLNGAKEILLFNLPDLGQNPSARSQKVVEAASHVSAVHNQLLLNLARQLAPTMVKLFEIDKQFAEML  
 RDPQNFGLSDQRNACYGGSYVWKPFASSASTDSQLSAFNPQERLAIAGNPLLAQAVASPMARSASTLN  
 CE  
 GKMFWDQVHPTTVVHAALSEPAATFIESQYEFIAH

FIGURA 28 (SEQ ID N° 35)

1	ADTRPAFSRI	VMFGDSLSDT	GKMYSKMRGY	LPSSPPYYEG	RFSNGPVWLE	QLTKQFPGLT
61	IANEAEGGAT	AVAYNKISWN	PKYQVINNL	YEVTOFLQKD	SFKPDDLVL	WVGANDYLAY
121	GWNTAQDAK	VRDAISDAAN	RMVLNGAKQI	LLFNLPDLGQ	NPSARSQKV	EAVSHVSAYH
181	NKLLLNLARQ	LAPTMVKLFE	EIDKQFAEML	RDPQNFGLSD	VENPCYDGGY	VWKPFAATRSV
241	STDRQLSAFS	PQERLAIAGN	PLLAQAVASP	MARRSASPLN	CEGKMFWDQV	HPTTVVHAAL
301	SERAATFIET	QYEFIAH				

FIGURA 29 (SEQ ID N° 36)

ACAGGCCGATGCACGGPACCGTACCTTTCCGCAGTGAAGCGCTCTCCCCCATCGTTCCG  
 CGGGACTTCATCCGCGATTTTGGCATGAACACTTCCTTCAACGCGCGTAGCTTGCTACAA  
 GTGCGGCAGCAGACCCGCTCGTTGGAGGCTCAGTGAGATTGACCCGATCCCTGTCGGCCG  
 CATCCGTCATCGTCTTCGCCCTGCTGCTCGCGCTGCTGGGCATCAGCCCGGCCAGGCAG  
 CCGGCCCGCCTATGTGGCCCTGGGGGATTCCCTATTCTCGGGCAACGGCGCCGGAAAGTT  
 ACATCGATTGAGCGGTGACTGTACCCGAGCAACAACCGTACCCCGCCCGCTGGGCGG  
 CGGCCAACGCACCGTCCCTCCTTACCTTCGCGGCTGCTCGGGAGCGGTGACCACGGATG  
 TGATCAACAATCAGCTGGGCGCCCTCAACGCGTCCACCGCCCTGGTGGATCACCATCG  
 GCGGCAATGACGCGGGCTTCGCGGACGCGATGACCACCTGCGTCACCAGCTCGGACAGCA  
 CCTGCCTCAACCGGCTGGCCACCGCCACCACTACATCAACACCACCTGCTCGCCCGGC  
 TCGACGCGGTCTACAGCCAGATCAAGGCCGTGCCCCCAACGCCCGCTGGTCTGCTCG  
 GCTACCCGCGCATGTACCTGGCCTCGAACCCCTGGTACTGCTGGGCTGAGCAACACCA  
 AGCGCGCGGCCATCAACACCACCGCGACACCCTCAACTCGGTGATCTCTCCCGGGCCA  
 CCGCCACGGATTCGATTCGGCGATGTCCGCCGACCTTCAACAACCACGAACTGTTCT  
 TCGGCAACGACTGGCTGCACTCACTCACCTGCGGCTGGGAGTGTACCAACCCACCA  
 GCACGGGCCATCAGAGCGGCTATCTGCCGCTCTCAACGCCAACAGCTCGACCTGATCAA  
 CGCACGGCCGTGCCCGCCCGCGCGTCACTCGCGCGGGCGCCGAGCGCGTGTATCA  
 GCCACAGTGCCGGTGACGGTCCCACCGTCAAGGTGAGGGTGTACGTACGGTGGCGCC  
 GCTCCAGAACTGGAACGTCAGCAGGACCGTGGAGCCGTCCCTGACCTCGTGAAGAACTC  
 CGGGGTGAGCGTATCACCCCTCCCCGTAGCCGGGGGGAAGGGCGGCCGAACCTCCTT  
 GTAGGACGTCAGTCTGCGGCCCGCGTTGCCACCGTCCGCGTAGACCGCTCCATGGT  
 CGCCAGCCGGTCCCCGCGAACTCGGTGGGGATGTCCGTGCCAAGGTGGTCCCGGTGGT  
 GTCCGAGAGCACCGGGGCTCGTACCGGATGATGTGCAGATCCAAAGAATT

FIGURA 30 (SEQ ID N° 37)

MRLTRLSAASVIVFALLLALLGISPAQAAGPAYVALGDSYSSGNGAGSYIDSSGDCHRSN  
 NAYPARWAAANAPSSFTFAACSGAVTTPVINNLGALNASTGLVSIITIGGNDAGFADAMTT  
 CVTSSDSTCLNRLATATNYINTLLARLDAVY S QIKARAPNARVVVLGYPRMYLASNPWYC  
 LGLSNTKRAAINTTADTLNSVISRATAHGFRFGDVRPTFNNHELFFGNDWLHSLTLPVWE  
 SYHPTSTGHQSGYLEFVLNANSST



FIGURA 31 (SEQ ID N° 38)

```
1 mlphpagerg evgaffallv gtpqdrirl echetrplrg rcgcgerivp pltlpgdgvl
61 cttsstrdae twwrkhlqpr pdggfrphig vgcclagqgs pglwlcgreg crfevcrddt
121 pglstrtrngd ssppfragws lppkcgeisq sarktpavpr ysllrtdrpd gprgrfvgsq
181 praatrrrlf lgipalvlvt altlavlavpt gretlwrwc eatqdwclgv pvdsrgqpae
241 dgefillspv qaatwgnyya lgdsyssgdg ardyypgtav kggcwsana ypelvaeayd
301 faghlsflac sggrgyamld aidevgsqld wnsphtslvt igiggnldgf stvlktcmvr
361 vplldskact dgedairkrm akfettfeel isevrtrapd arilvgypr ifpeeptgay
421 yltasnqrw lnetiqefnq qlaeavavhd eeiaasggvg svefvdvyha ldgheigsde
481 pwvngvqlrd latgvtvdrs tfhpnaaghr avgervieqi etgpprplya tfavvagatv
541 dtlagevg
```

FIGURA 32 (SEQ ID N° 39)

```

1 ggtggtgaac cagaacaccc ggtcgtcggc gtgggcgtcc aggtgcagg tgcaggttctt
61 caactgctcc agcaggatgc cgcctgtggc gtgcacgatg gccttgggca ggctgtggg
121 ccccgacgag tacagcacc atagcggatg gtcgaacggc agcgggggga actccagttc
181 cgcgccttcg cccgcggctt cgaactcgc ccaggacagg gtgtcggcga cagggccgca
241 gcccaggtag ggcaggacga cgggtgtgctg caggctgggc atgcccgtgc gcagggcttt
301 gagcacgtca cggcgggtcg agtccctacc gccgtagcgg tagccgtcca cggccagcag
361 cactttcggg tcgatctcgc cgaaccggtc gaggacgctg cgcaccccga agtcggggga
421 acaggacgac caggtcgcac cgtcgcggc gcaggcagg aatgcccggc tcgctcggc
481 gatgttcggc aggtaggcca cgaaccggtc gccggggccc accccgaggc tcgggagggc
541 cgcagcgtac gccggcgtgc gggtcggcag ttctcccag gtccactcgg tcaaccggcc
601 gagttcggc gcgtgccgga tcgccacggc tgatgggtca cggtcggcga agatgtgctc
661 ggcgtagttg agggtagggc cggggaacca gacggcggc ggcatgggt cggaggcga
721 cactgtggtg tacggggtgg cggcgcgcac ccggtagtag tcccagctc cggaccagaa
781 tctctcagg tcggttacc accagcgcga cagtgcctc tagtcgggt cgtccacacc
841 ggggtgctcc cgcacccagc gggtagaacg ggtgaggtg gcgcttctt tgcgctctc
901 gtcgggactc cacagtagc gggcgtcgg cttgagtgtc atgaaacgc accctcgt
961 ggacgggtgc gatgcgggtg gcgtcgggtg cctcccctaa cgtcccggc tgacggagt
1021 ttgtgcacca catctagcac gccggacgc gaaaccgtat ggagaaaaa cctacaacc
1081 cggcggacg gtgggtttc gccacacta ggggtcgggt gcctgcttc cgggcagggc
1141 agtcccggg tgctgtggtg cggcggggag ggtgtcgtc tcgaggtgt cggcgggac
1201 actcccggc tcagccgtac ccgcaaccgg gacagtctc cctccctcc ggctggatg
1261 tcccttccc cgaatgagg cgagatctc cagtacgcc ggaaaaacc cgtgtgccc
1321 aggtactctt tgcttcgaac agacaggccg gacggtcac gggggagggt tgtgggagc
1381 ggaccacgtg cggcagacc acgacgggtt ttctcggta tccccgctc tgtacttgt
1441 acagcgtcca cgtcgtctt ggtgtcggc agcggcggc agacgctgt gccatgtg
1501 tgtgagcca cccaggactg gtgcctggg gtgcccgtc actcccggc acagcctcg
1561 gaggaaggc agtttctgt cgttctccg gtcaggcag gcacctggg gaactattc
1621 gcgctcggg attcgtactc ttccggggac gggcccggc actactatc cggcaccgc
1681 gtagaaggcg gttgtggyg gtcgcgtaac gcctatcgg agctggtcg cgaagcctc
1741 gacttcggc gacactgtc gttcctggc tcagcgggc agcggccta gccoatgct
1801 gacgctatc acgaggtcgg ctccagctg gactggaaact cccctcacac gtcgctggt
1861 acgatcggga tcggcggcaa cgtctgggg ttctccacgg ttttgaaag ctgcatggt
1921 cgggtgcgc tgctggacag caaggcgtgc acggaccag aggaagctat ccgcaaggc
1981 atggcgaat tcgagacgac gtttgaagag ctcatcagc aagtgcgac ccgcgcggc
2041 gacgcccgga tccctgtcgt gggctaccoc cggattttc cggaggaaac gaccggcgc
2101 tactacacgc tgaccgcgag caaccagcgg tggctcaacg aaaccattc ggagttcac
2161 cagcagctc cggaggctgt cgcggtcac gacgaggaga ttgcccgtc gggcggggg
2221 ggcagcgtg agttcgtgga cgtctaccac gcgttgacg gccacagat cggctcggc
2281 gagccgtgg tgaccgggtg gcagttgctg gacctgcga cgggggtgac tgtggaccg
2341 agtaccttc accccaacgc cgtggggcag cggcgggtc gtgagcggg catcgagca
2401 atcgaaacg gcccgggccg tccgctctat gccactttc cgggtgtggc gggggcgac
2461 gtggacactc tcgcccggca ggtgggggtg cccggcttac cgtcccggc cgaggtctc
2521 gagcactcgc gcgatctggt ccactgcca gtcagttcg tcttcggta tgaccagcg
2581 cggggagagc cggatrgttg agccgtgctg gctttgacg agcacaccc gctgagagc
2641 ccgcttcgac agttctcttc cggtgccag agtcgggtc acgtcgatc cagcccacg
2701 gccgatgctg cggcggcga ccacgccgtt gccgaccagt tggtcgaggc gggcggcag
2761 cacggggcgc agggcggca catggtccag gtaaggccg tcgcccagc ggctcaccac
2821 ggcagtgcc accgcgagg cgagggcgtt gccgcccag gtgctgccc gctggccgg
2881 cgggatcacg tcgaagactt ccgctcggc tacggcccg gccacggca ggtatgccc
2941 gccagcgcct ttgccgaaca ggtagatata ggcgtcgaact ccgctgtggt cgcaggccc

```

FIGURA 33 (SEQ ID N° 40)

```

1  vgsqpraatr  rrlflgipal  vlvtaltlvt  avptgretlw  rmwceatqdw  clgvpvdsrg
61  qpaedgefl1  lspvqaatwg  nyyalgdsys  sgdgardyyp  gtavkggcwr  sanaypelva
121 eaydfaghls  flacsggrgy  amldaidevg  sqldwnspht  slvtigiggn  dlqfstvlkt
181 cmvrvpllds  kactdqedai  rkrmakfett  feelisevrt  rapdarilvv  gyprifpeep
241 tgayytiltas  nqrwlnetiq  efnqqlaeav  avhdeeiaas  ggvgsvervd  vyhalqghei
301 gsdepwvngv  qlrdlatgvt  vdrstfhpna  aghravgerv  ieqietgpgr  plyatfavva
361 gatvdtlage  vg

```

FIGURA 34 (SEQ ID N° 41)

```

1  mrttviaasa  llllagcadg  areetagapp  gessggiree  gaeastsitd  vyialgdsya
61  amggrdqplr  gepfclrssg  nypellhaev  tdltcggavt  gdllprtlig  ertlpaqvda
121 ltadtltlvtl  siggndlgfg  evagcireri  agenaddcvd  llgetigeql  dqlppqldrv
181 heairdragd  aqvvtgylp  lvsagdcpel  gdvseadrrw  aveltgqine  tvreaaerhd
241 alfvlpddad  ehtscappqq  rwadiqqqqt  dayplhptsa  gheamaaavr  dalglepvqp

```

FIGURA 35 (SEQ ID N° 42)

```

1 ttctgggggtg ttatgggggtt gttatcggct cgtcctgggt ggatcccgcc agtgggggta
61 ttacgggggg acttttgtgt ccaacagccg agaatgagtg ccctgagcgg tgggaatgag
121 gtgggcgggg ctgtgtcgcc atgagggggc ggcgggctct gtggtgcccc gcgacccccg
181 gccccgggtga gcggtgaatg aaatccggct gtaatcagca tcccgtgcc accccgtcgg
241 ggaggtcagc gcccgagtg tctacgcagt cggatcctct cggactcggc catgctgtcg
301 gcagcatcgc gctcccggtt cttggcgtcc ctccgctgtt ctgctgtctg tccctgggag
361 gcgaaatgat caccggggag tgatacaccc gtggtctcat cccggatgcc cacttcggcg
421 ccacccggca attcggggcag ctccgggtgg aagtaggtgg catccgatgc gtcggtgacg
481 ccatagtggg cgaagatctc atcctgctcg aggtgctca ggccactctc cggatcgata
541 tggggggcgt ccttgatggc gtccttgcctg aaacggaggt gcagcttctg ggcttccaat
601 ttccgaccac ggagcgggac gaggctggaa tgacggccga agagccccgtg gtggacctca
661 acgaagggtg gtagtcccgt gtcatcattg aggaacaagc cctccaccgc acccagcttg
721 tggccggagt tgcgttaggc gctggcatcc agaaggaaa cगतctcata ttigtccgtg
781 tgctcagaca tgatcttccct ttgctgtcgg tgctcgttac taccacggta gggctgaatg
841 caactgttat tttctgttta ttttaggaat tggctccatc cccacaggct ggctgtggtc
901 aaatcgtcat caagtaelcc ctgtcacaca aaatgggtgg tgggagccct ggtcggggtt
961 ccgtgggagg cgcctgtccc cgcaggatcg tccgcatcgg cggatctggc cggtaaccoc
1021 cgggtgaataa aatcattctg taaccttcat cacggttggg tttagggtac cgcoccttcc
1081 gtccctgacc cgtccccggc ggcggggagc cgcggggtg cggtagacag gggagacgtg
1141 gacaccatga ggacaacggt catcgcagca agcgcattac tccctctcgc cggatgcccg
1201 gatggggccc gggaggagac cgcgggtgca cgcgggggtg agtccctcgg gggcatccgg
1261 gaggaggggg cggaggcgtc gcaagcatc accgagctct acatcgcctc cggggattcc
1321 tatgcggcga tgggcgggcg ggatcagccc ttacggggtg agccttctg cctgcgctcg
1381 tccggtaat acccggact cctccacgca gaggtcaccg atctcacctg ccaggggcgg
1441 gtgaccgggg atctgctcga acccaggacg ctgggggagc gcacgctgcc gggcagggtg
1501 gatgcgctga cggaggacac caccctgtgc accctctcca tccggggcaa tgacctcga
1561 ttccgggagg tggcgggatg catccgggaa cggatcggcg gggagaaccg ggatgctgac
1621 gtggacctgc tgggggaaac catcggggag cagctogatc agcttcccc gcagctggac
1681 cgcgtgcacg aggtatccg ggaccgcgcc ggggacgcgc aggttgggtt caccggttac
1741 ctgcccgtcg tgtctgcccg ggaactgccc gaaatggggg atgtctccga ggcggatcgt
1801 cgttggggcg ttgagctgac cgggcagatc aacgagaccg tggcgaggcc ggcgaaccga
1861 cacgatgccc tctttgtcct gcccgacgat gccgatgagc acaccagttg tgcacccccca
1921 cagcagcgcct gggcgatat ccaggggcaa cagaccgatg cctatcgcct gcacccgacc
1981 tccgcccggc atgaggcgat ggcggcccgc gtcgggagc cgttgggctt ggaacgggtc
2041 cagccgtagc gccggggcgc cgttctcga cgaccaacc atgccaggct gcagtccatc
2101 ccgcacatag cgcgcgcccg cgatggagta cgcaccatag aggatgagcc cगतgccgac
2161 gatgatgagc agcacactgc cgaagggttg ttccccgagg gtgocgagag cगतtccag
2221 acctgcggcc tgctccggat catgggccc aaccgagatg acgatcaaca cccccaggat
2281 cccgaaggcg ataccacggg cgacataacc ggctgttccg gtgatgatga tccgggtccc
2341 gacctgcctt gaccccgac ccgcctccag atcctcccgg aaatccccgg tggccccctt
2401 ccagaggttg tagacaccgc cccccagtac caccagccc gcgaccacaa ccagacacc
2461 accccaggtt tgggataggc cgggtggcggg gacatcgttg gcggtctccc catcggaggt
2521 gctgcccgcc cgggcgaaag tggaggtggt caccgcccag gagaagtaga ccattggccat
2581 gaccgcccc ttggcccttt cctttaggtc ctgcccgc agcagctggc tcaattgcca
2641 agtcccagg gcccccagg cगतgaccgc aaaccaagg aggaaactgc caccggagc
2701 ctccgcgatg gtggccaggc cacctgaatt cगतcctca tcaccgaaac cgcggatcc
2761 agtggcgatg cgcaccgcga tccaccgat gaggatgtgc agtatgccc ggacaatgaa
2821 accacctctg gccaggtggc tcagcgggg gtggctcctg gcctggctcg cagcccgttc
2881 gatcgtccgt tccggggtc tgggtgcgcc cttatccata gctcccattg aaccgcttg
2941 aggggtgggc ggccactgtc agggcggtt gtgatctgaa ctgtgatgtt ccatcaacc

```

FIGURA 36 (SEQ ID N° 43)

```

1 mrrfrlvglf sslvlaagaa ltgaataqaa qpaaadgyva lgdsyssgvg agsyisssgd
61 ckrstkahpy lwaahspst fdftacs gar tgdvlsqqlg plssgtglvs isiggn dagf
121 adtmttcvlq sessclsria taedaydstl pgkldgvysa isdkapnahv vvi gyprfyk
181 lgttciglse tkrtainkas dhlntvl aqr aaahgftfgd vrtttfghel csgepwlhsv
241 nwl nigesyh ptaagqsggy l pvlngaa

```

FIGURA 37 (SEQ ID N° 44)

```

1 ccggcgggcc cgtgcaggag cagcagccgg ccggcgatgt cctcggggct cgtcttcate
61 aggcggtcca tgcggtggc gaccggcgcc gtgtagttag ccgggacctc gtcccaggtg
121 ccggcgccga tctggcgggt ggtgcggtgc gggccgcgcc gaggggagac gtaccagaag
181 cccatcgtca cgttctccgg ctgctggtcg ggctcgtccg ccgctccgtc cgtcgcctcg
241 ccgagcacct tctcggcgag gtcggcgctg gtcgcccgtca ccgtgacgtc ggcgccccgg
301 ctccagcgcg agatcagcag cgtccagccg tcccccctcc ccagcgtcgc gctgcggtcg
361 tgcgctcggg cgatccgag cccgcgccgg ccggcgccga ccgctggtgc gccggaccgt
421 acgctgctca tgttcgcccg gtgctgagta ggctgctcac ccgtggcga acgcccagg
481 aacagcgcgt cgcgcgcgtc ggcgggggag tgcctgctgt ccacgttgag ccggtatcggc
541 agggcttctg gcgggttcac gacatgtctg ccatgatcgg gcaccggcc gccgctgca
601 cccgctttcc cgggcacgca cgcaggggc tttctcggc tcttccgtcc gaacttgaac
661 gagtgcagc catctcttgg catggacact tccagtcaac gcgctagct gctaccaegg
721 ttgtggcagc aatctctgta agggaggttc catgagactt tccgacttg tggcttctc
781 gatttcgctc gtcctcggc ccggcgccgc cctcaccggg gcagcagccg cccagggcgg
841 ccaaccggcc gccgcgagc gctatgtggc cctcggcgac tctactctc cgggggtcgg
901 agcgggcagc tacatcagct cgcggggcga ctgcaagcgc agcaggaagg cccatcccta
961 cctgtggggg gccgcccact cgcctccac gttcgaactt accgcctgtt ccggcgcccg
1021 taagggtgat gttctctccg gacagctcgg ccgctcagc tccggcaccg gcctcgtctc
1081 gatcagcatc ggcggcaacg acgcccgttt cgcgcacacc atgacgacct gtgtgctcca
1141 gtcgagagc tctgctgtg cgggatcgc caccgcccag gcgtactcgc actcagcct
1201 gccggcaag ctgcagggcg tctactggc aatcagcgac aaggcgcga acgcccagct
1261 cgtcgtcatc ggtaccgcc gcttctacaa gctcggcacc acctgcctcg gcctgtccga
1321 gaccaagcgg acggcgatca acaaggcctc cgcaccctc aacaccgtcc tgcgccagcg
1381 cgcggccgcc cacggcttca cctcggcga cgtacgacc accttcaccg gccacgagct
1441 gtctccggc agcccctggc tgcacagcgt caactggctg aacatggcg agtctgacca
1501 cccaccggcg gccggccagt ccggtgcta cctgcccgtc ctcaaccggc ccgctgacc
1561 tcaggcggaa ggagaagaag aaggagcggg yyyagacgag gagtgggagg ccccggcga
1621 cgggttcccc gtccccgtct ccgtctcgt cccgttccc caagtaccg agaacggcc
1681 cgcgtcggac gtggcccga ccggactccg caccctccag cgcagggcac tctcgaaccg
1741 gccggtctg tgcgtcgtc taccaccac gccgtcctgg cgcgagcgt cgcgcccga
1801 cgggaaggac agcgtcccgc acccccgatc ggagaccgac ccttcccgg tcaaccaccg
1861 gtagccgacc tccgcccga gcccccgac cgtgaaccgt gccgtgaacc cgggtgcccg
1921 gtcgtcggc ggcggacagg ccccagata gttgggtcgc gagctcacc cgggtcacctc
1981 caccgactgc gtcgcccggc

```

FIGURA 38 (SEQ ID N° 45)

```

1 mrrsritayv ts1llavgca ltgaatagas paaaatgyva lgdsyssgv agsylvssgd
61 ckrsskaypy lwqaaahspss fsfmacsgar tgdvlanqlg tlnsstglvs ltiggnndagf
121 sdvmttcvliq sdsaclsrin takayvdstl pggldsvyta istkapsahv avlgyprfyk
181 lggscclagls etkrseinda adylnsaiak raadhgftfg dvkstftghe icssstwlhs
241 ldl1lniqsy hptaacqsgg ylpvmnsva
    
```

FIGURA 39 (SEQ ID N° 46)

```

1 ccaccgccgg gtcggcggcg agtctcctgg cctcggtcgc ggagaggttg gccgtgtagc
61 cgttcagcgc ggcgccgaac giccttctca ccgtgccgcc gtactcgttg atcaggccct
121 tgccttctct cgacgcggcc ttgaagccgg tgcccttctt gagcgtgacg atgtagctgc
181 ccttgatcgc ggtgggggag ccggcggcga gcaccgtgcc ctgggcccgg gttggcctggg
241 cgggcagtgc ggtgaatccg cccacgaggg cgcgggtcgc cacggcgggt atcgcggcga
301 tccggatctt ctgtcatcgc agctgtgcca tacgagggag tcctcctctg gccagcggcg
361 gccttgggtg gggcgcacgg ctgtgggggg tgcgcgcgtc atcacgcaca cggccctgga
421 gcgtcgtgtt ccgccctggg ttgagtaaa cctcggccat ctacgggggt ggctcaaggg
481 agttgagacc ctgtcatgag tctgacatga gcacgcgaatc aacggggccg tagcaccccc
541 ggggcgaccc cggaaagtgc cgagaagtct tggcatggac acttcctgtc aacaogcgtg
601 gctggtacga cggttacggc agagatcctg ctaaagggag gttccatgag acggttccga
661 attacggcat acgtgacctc actcctcctc gccgtcggct gcgccctcac cggggcagcg
721 acggcgcagg cgtccccagc cgcgcgggcc acgggctatg tggccctcgg cactcgtac
781 tgcctcggtg tcggcgcggg cagctacctc agctccagcg gcgactgcaa gcgcagttcg
841 aaggcctatc cgtacctctg gcaggccggc cattcacctc cgtcgttcag tttcatggct
901 tgctcggggc ctcgtacggg tgatgtcctg gccaatcagc tcggcacctc gaactcgtcc
961 accggcctgg tctcctcac catcggaggg aacgacggcg gcttctccga cgtcatgacg
1021 acctgtgtgc tccagtcgca cagcgcctgc ctctcccga tcaaacggc gaaggcgtac
1081 gtcgactcca cctgcccgg ccaactcgac agcgtgtaca cggcgatcag cacgaaggcc
1141 ccgtcggccc atgtggcgtg gctgggttac ccccgcttct acaaaactgg cggctcctgc
1201 ctccggggcc tctcggagac caagccgtcc gccatcaag acggcgccga ctatctgaac
1261 agcggccatc ocaagcgcgc cgcggaccac ggcttcacct tcggcgactc caagagcacc
1321 ttaccgggcc atgagatctg ctccagcagc acctggctgc acagtctcga cctgctgaac
1381 atcggccagt cctaccacc gaccgcccgc ggccagtccg gcggctatct gccggtcatg
1441 aacagcgtgg cctgagctcc cacggcctga atttttaagg cctgaatctt taaggcgagg
1501 gtgaaccgga agcggaggcc ccgtcccgtg gggctctcgt cgcacaggtc accgagaacg
1561 gcacggagtt ggacgtcgtg cgcaccgggt cgcgcacctc gacggcgatc tcggtcggga
1621 tcgttccgct cgtgtcgtac gtggtgacga acacctgctt ctgctgggtc ttccgcccgc
1681 tcgcccggaa ggacagcgtc tccagcccg gatccgggac ctgccccttc ttggtaaccc
1741 agcggtactc cacctcgacc ggcaccggc ccaccgtgaa ggtcgcggtg acgtggggcg
1801 cctgggcccgt gggcggcggg caggcaccgg agtagtcggt gtgcacggcg gtgaccgtca
1861 ccttacggga ctggcggcg ggggtcgtcg tacgcgcgcc gccaccggcg cctcccggag
1921 tggagcccga gctgtggtcg ccccgcctg cggcgttgtc gtcctcgggg gttttcgaac
    
```

FIGURA 40 (SEQ ID N° 47)

```

1 mgsppraatr rrlflgipal vltaltlvtl avptgretlw mwceatqdw clgvpvdsrg
61 qpaedgefil lspvqaatwg nyyalgdsys sgdgadyyp gtavkggcwr sanaypelva
121 eaydfaghls flacsqgrgy amldaidevg sqldwmspht slvtigiggn dlqfstvlkt
181 cmrvrpllds kactdgedai rkrmakfett feelisevrt rapdarilvv gyprifpeep
241 tgayytiltas nqrwlneti qefnqqlaeav avhdeeiias ggvgsvfevd vyhaldghei
301 gsdepwvngv qlrdlatgvt vdrstfhpna aghravgerv ieqietgpgr plyatfavva
361 gatvdtlage vg

```

FIGURA 41 (SEQ ID N° 48)

```

1 ctgcagacac ccgccccgcc ttctcccgga tcgtcatggt cggcgactcc ctcagcgaca
61 ccggcaagat gtactccaag atgcgcggct acctgccgtc ctccccgccg tactacgagg
121 gccgttctc gaacggcccc gtctggctgg agcagctgac gaagcagttc cccggcctga
181 cgatcgccaa cgaggccgag gggggcgcga ccgcagtcgc ctacaacaag atctcctgga
241 accccaagta ccaggtcatt aacaacctcg actacgaggt caccacagttc ttgcagaggg
301 actcgttcaa gcccgacgac ctgggtcatcc tctgggtggg cgccaacgac taactggcct
361 acggttggaa cacggagcag gacgccaagc ggggtcgcga cgccatctcg gacgcggcaa
421 accgcatggt cctgaacggc gcgaagcaga tctgtctggt caacctgcc gacctgggcc
481 agaaccgctc cggccgctcc cagaaggctc togaggccgt ctgcacgtg tccgcctacc
541 acaacaagct gtcctcaac ctgcccggc agctcgcccc gacgggcatg gtcaagctgt
601 tcgagatcga caagcagttc gcgggatgc tgcgcgacct ccagaacttc ggctgagcg
661 acgtggagaa cccgtgctac gacggcggct acgtgtggaa gccgttcgcc acccggctccg
721 tctcgaccga ccggcagctg tcggccttct cggcccagga gcgcctggcg atcgtggca
781 acccgtcctt gccacaggcg gtagcttcgc cgatggccc cgcctcggcc tcgccctca
841 actgcgaggg caagatggtc tgggaccag tccacccac caccgtggtc caccgcgcc
901 tctcggagcg cggccacc ttcatcgaya cccagtacga gttcctcgcc cactagtcta
961 gagatccc

```

FIGURA 42

1. L131
2. *S. avermitilis*
3. *T. fusca*
4. Consenso

```

1 (1) -----MRLTRSLSAASVIVFALLLALLGISPAQAAG----- 50
2 (1) -----MRRSRTAYVTSLLLAVGCALTGAATAQASPA-----
3 (1) VGSGPRAATRRLFLGIPALVLTALTLVLAVPTGRETLWRMWCEATQDW
4 (1) MRRSRFLA ALILLTLA AL GAA ARAAP

1 (32) -----P-AYVALGDSYSSGNGAGSYID 100
2 (33) -----AAATGYVALGDSYSSGNGAGSYLS
3 (51) CLGVPVDSRQPAEDGEFLLSPVQAATWGNYYALGDSYSSGDGARDIYP
4 (51) A A YVALGDSYSSG GAGSY

1 (53) SSGD---CHRSNNAYPARWAAANAP---SSFTFAACSGAVTTDVIN--- 150
2 (57) SSGD---CKRSSKAYPYLWQAAHSP---SSFSFMACSGARTGDVLA---
3 (101) GTAVKGGCWRSANAYPELVAEAYDFA---GHLSFLACSGQRYAMLDAIDE
4 (101) SSGD C RSTKAYPALWAAAAA SSFSF ACSGARTYDVLA

1 (93) --NQLGALNAST--GLVSIITIGGNDAGFADAMTTCVTS-----SDSTCL 200
2 (97) --NQLGTLNSST--GLVSLITIGGNDAGFSDVMTTCVLO-----SDSACL
3 (149) VGSQLDWNSPRT--SLVTITIGGNDLGFSTVLKTCMVR-----VPLLDS
4 (151) QL LNS T LVSITIGGNDAGFAD MTTCVL SDSACL

1 (133) NRLATATNYINTLLA-----RLDAVYSQIKARAPNARVVVLGYPRMY 250
2 (137) SRINTAKAYVDSLPG-----QLDSVYTAISTKAPSAHVAVLGYPRFY
3 (191) KACTDQEDAIRKMAKF-----ETTFEELISEVRTRAPDARILVVGYPRIF
4 (201) RIA AK YI TLPA RLDSVYSI TRAP ARVVVLGYPRIY

1 (176) LASNPWYCLGLSNTKRAINTTADTLNSVISSRATAH-----GF 300
2 (180) KLG-SCLAGLSETKRSAINDAADYLNSAIKRAADH-----GF
3 (237) PEEPTGAYTTLTASNQRWLNETIQEFNQQLAEAVAVHDEEIAASGGVGSV
4 (251) SG LGLS TKRAINDAAD LNSVIKRAADH GF

1 (215) RFGDVRPTFNNHELFFGNDWLHSLTLP-----VWESYH 350
2 (218) TFGDVKSTFTGHEICSSSTWLHSLDLLN-----IGQSYH
3 (287) EFVDVYHALDGHEIGSDEPWVNGVQLRDLATG-----VTVDRSTFH
4 (301) TFGDV TF GHELCSA PWLHSLTLP V SYH

1 (248) PTSTGHQSGYLPVLNANSST----- 395
2 (252) PTRAGQSGGYLPVMNSVA-----
3 (328) PNAAGHRAVGERVIEQIETGPRPLYATFAVVAGATVDIAGEVG
4 (351) PTA GHAAGYLPVLSI T

```



**FIGURA 43**

SEQ ID N° 17 que es la secuencia de aminoácidos de una lípido aciltransferasa de *Candida parapsilosis*;

```

MRYFAAFLL INTISAFVLA PKKPSQDDFY TFPQGYEAQP LGSILKTRNV PNPLTNVFTF VKVQNAWQLL
VRSEDTFGNP NAIVTTLIQP FNAKKDKLVS YQTFEDSGKL DCAPSYAIQY GSDISTLTQ GEMYISALL
DQGYVVTPD YEGPKSTFTV GLQSGRATLN SLRATLKSGN LTGVSSDAET LLWGYSGGSL ASGWAAAIQK
EYAPELSKNL LGAALGGFVT NITATAEAVD SGPFAGIISN ALAGIGNEYF DFKNYLLKKV SPLLSITYRL
GNTHCLLDGG IAYFGKSFFS
RIIRYFPDGW DLVNQEPIKT ILQDNGLVYQ PKDLTPQIPL FIYHGTLDAI VPIVNSRKTFF QQWCDWGLKS
GEYNEDLTNG HITESIVGAP AALTWIINRF NGQPPVDGCQ HNVRSANLEY EGTPOSIKNY FEAAALHAILG
FDLGPDKRD KVTLGGLLKL ERFAP
    
```

**FIGURA 44**

SEQ ID N° 18 que es la secuencia de aminoácidos de una lípido aciltransferasa de *Candida parapsilosis*;

```

MRYFAIAFLL INTISAFVLA PKKPSQDDFY TFPQGYEAQP LGSILKTRNV PNPLTNVFTF VKVQNAWQLL
VRSEDTFGNP NAIVTTLIQP FNAKKDKLVS YQTFEDSGKL DCAPSYAIQY GSDISTLTQ GEMYISALL
DQGYVVTPD YEGPKSTFTV GLQSGRATLN SLRATLKSGN LTGVSSDAET LLWGYSGGSL ASGWAAAIQK
EYAPELSKNL LGAALGGFVT NITATAEAVD SGPFAGIISN ALAGIGNEYF DFKNYLLKKV SPLLSITYRL
GNTHCLLDGG IAYFGKSFFS RIIRYFPDGW DLVNQEPIKT ILQDNGLVYQ PKDLTPQIPL FIYHGTLDAI
VPIVNSRKTFF QQWCDWGLKS GEYNEDLTNG HITESIVGAP AALTWIINRF NGQPPVDGCQ HNVRSANLEY
PGTPOSIKNY FEAAALHAILG FDLGPDKRD KVTLGGLLKL ERFAPHHHHH H
    
```

FIGURA 45

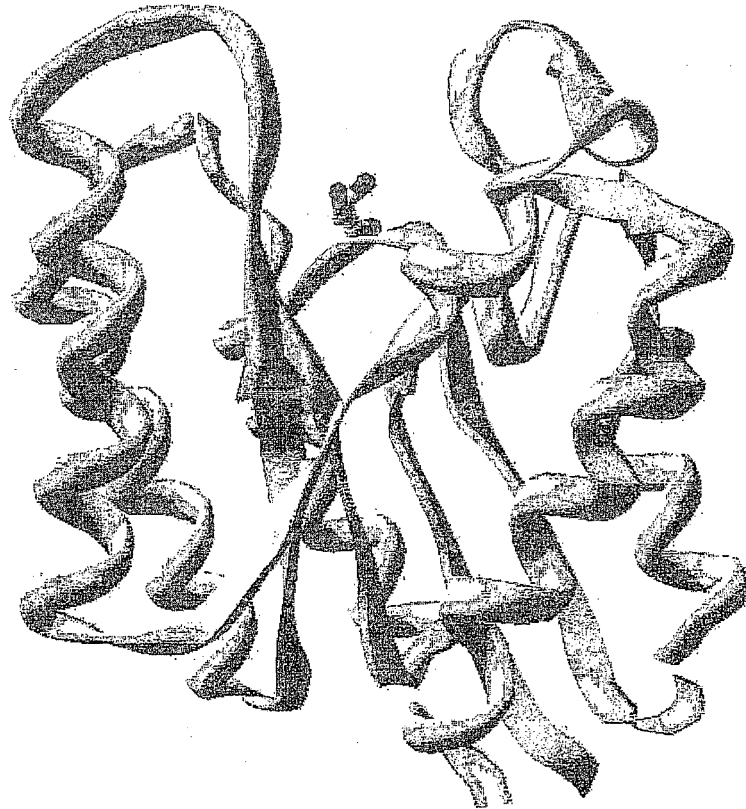


FIGURA 46

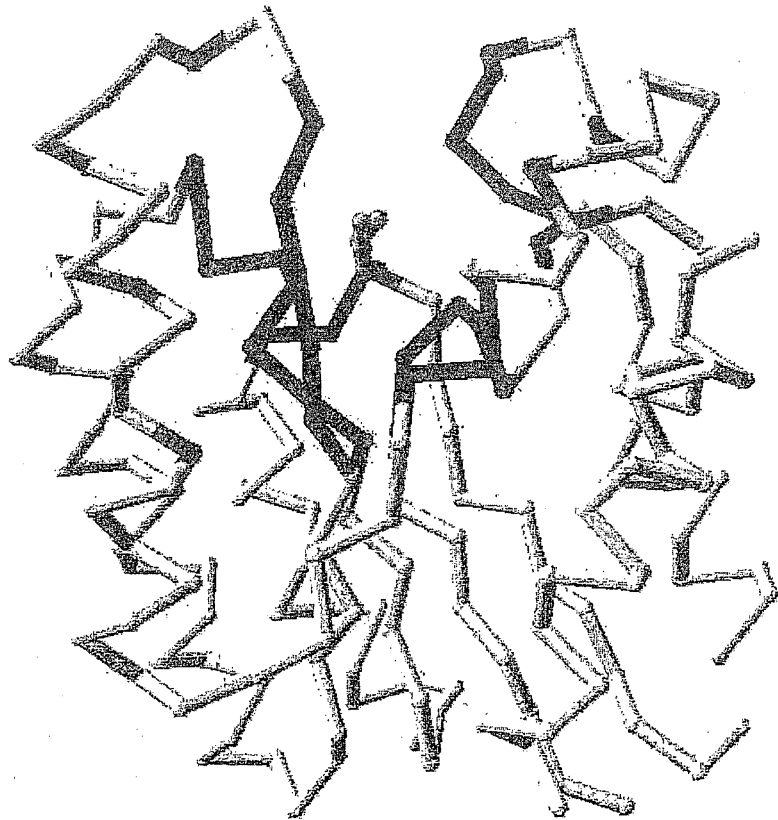
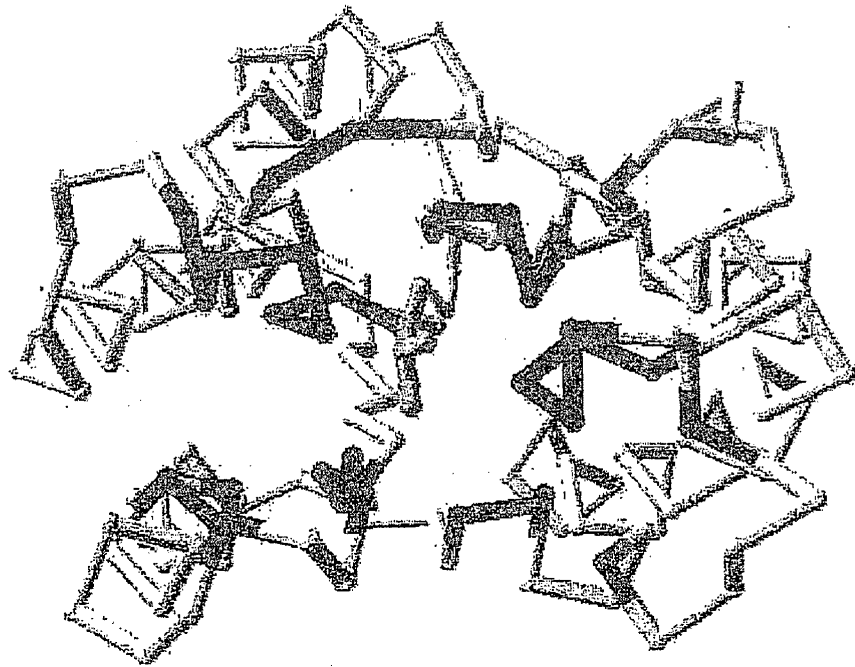


FIGURA 47









1DE0m L T F P A Y L E N A A K L F T A K G A K V I L S S Q T P N N P W E T G T F V N S P T R  
h4h4h4h4 h4 h4h4h4h4h4h4h4h4h4h4 s5s5s5 s5 s5 s5  
1IYNm Q Q T E Q T L R Q I L Q D V K a A N A E P I I m q i R L P A N Y G R - - - - - R Y N E A  
h4h4h4h4h4 h4 h4h4h4h4h4h4h4h4h4 s5s5s5 s5 s5 s5 s5s5s5s5s5? s5s5s5  
P10480m K R V R D A I S D A A N R N V L N G A K E I L L F N L P d l g q n P S A R S Q K V V E A A S H V S A  
  
1DE0 F V E Y A E L A A E V A - - - - - G V E Y V D H W S Y V D S I Y E T L G N A t v n  
h5h5h5h5h5 h5 h5h5h5h5h5h5 h5 h5h5h5h5h5h5h5 s6 s6 s6s6s6s6s6s6h6h6h6h6h6h6 h h h h  
1IYN N E A F S A I Y P K L A k e - - - - - f D V P L L P F F M E E V Y L K P Q W - - - -  
h5h5h5h5h5 h5 h5h5h5h5h5h5h5 h5 s6 s6 s6s6? h6h6h6h6h6 s  
P10480 S A Y H N Q L L L N L A r q l a p t g m v k l f e i D K Q F A E M L R D P Q N F G L S D Q R N a c y  
  
1DE0m F V E Y A E L A A E V A - - - - - G V E Y V D H W S Y V D S I Y E T L G N A t v n - -  
h5h5h5h5 h5 h5h5h5h5h5h5h5 h5 s6 s6 s6 s6s6h6h6h6h6h6h6h6h6 h h h h - - - -  
1IYNm F S A I Y P K L A k e - - - - - f D V P L L P F F M E E V Y L K P Q W - - - -  
h5h5h5h5h5 h5 h5h5h5 h5 s6 s6 s6 s6 s6? h6 h6 h6 h6 h6 h6  
P10480m Y H N Q L L L N L A r q l a p t g m v k l f e i D K Q F A E M L R D P Q N F G L S D Q R N a c y g g  
  
1DE0 - - - - -  
1IYN - - - - -  
P10480 g g s y v w k p f a s i s a s t d s q l s a f n p q e r l a i a g n p l l a q a v a s p m a a i r s a  
1DE0m - - - - -  
1IYNm - - - - -  
P10480m s y v w k p f a s i s a s t d s q l s a f n p q e r l a i a g n p l l a q a v a s p m a a i r s a s t  
  
1DE0 - - - - - s Y F P I D H T H T S P A G A E V V A E A F L K A V V C T G T S L K S V L T T T S F E G  
h h s? s? s? h?  
1IYN - - - - - M Q D D G I H P N R D A Q P F I A D W M A K Q L Q P L V N H D S L E  
s s s h?  
P10480 s t l n c e g k m f w d Q V H P T T V V H A A L S E P A A T F I E S Q Y E F L A H -  
  
1DE0m - - - - - s Y F P I D H T H T S P A G A E V V A E A F L K A V V C T G T S L K S V L T T T S F E G T C  
h h s? s? s? h?  
1IYNm - - - - - M Q D D G I H P N R D A Q P F I A D W M A K Q L Q P L V N H D S L E  
s s s h?  
P10480m l n c e g k m f w d Q V H P T T V V H A A L S E P A A T F I E S Q Y E F L A H -



# ES 2 449 474 T3

**FIGURA 51**

```

10      20      30      40      50
60
.....*.....|.....*.....|.....*.....|.....*.....|.....*.....|.....*.....|
1IVN_A      4 LLILGDSL SAG-----YRMSASAAWPALLNDKWqsk---
----- 34
P10480      28
IVMFGDSLSDTgkmyskmrgylpssppyyeGRFSNGPVWLEQLTNEFPGLTianeaeagpp 87

120
.....*.....|.....*.....|.....*.....|.....*.....|.....*.....|.....*.....|
1IVN_A      35 -tsvVNASISGDT-----
SQQGLARLPALLKQHQP RW 65
P10480      88 tavaY NKISWNPkyq-----
VINNLDYEVTQFLQKDSFKPDDL 125

180
.....*.....|.....*.....|.....*.....|.....*.....|.....*.....|.....*.....|
1IVN_A      66 VLVELGGNDG-----
LRGFQQPQQTEQT 87
P10480      126 VILWVGANDY-----LA--
YGNWTEQDAKVRDA 152

240
.....*.....|.....*.....|.....*.....|.....*.....|.....*.....|.....*.....|
1IVN_A      88 LRQILQDVKaANAEPllmqiRLPANYGR-----
----- 115
P10480      153 ISDAANRMV-LNGAK-----EILLFNLPdlg-----
----qnp 180

300
.....*.....|.....*.....|.....*.....|.....*.....|.....*.....|.....*.....|
1IVN_A      116 -----RYNEAFSAIYPKL Ake-----
fDVPllPFME 142
P10480      181 SARSQKVVEAASHVSAYHNQLLLNLArqlaptg-----
mvklfeidKQFAEMLRD 230

360
.....*.....|.....*.....|.....*.....|.....*.....|.....*.....|.....*.....|
1IVN_A      143 EVYLKQW-----
----- 150
P10480      231
PQNFGLSQQRNacyggsyvwkpfasrsastdsqleafnpqerlaiagnpllaqavaspm 290

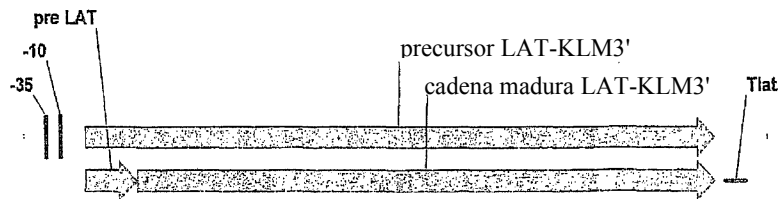
370      380      390      400
.....*.....|.....*.....|.....*.....|.....*.....|
1IVN_A      151 -----MQDDGI-----HPNRDAQPFADWM 170
P10480      291 arsa stlncegrMFWDQV-----HPTTVVHAALSEPA 322

```

FIGURA 52

		1	50
P10480	(1)	MKKWFVCLLGLVALTVQAADSRPAFSRIVMFGDSLSDTGKMYSKMRGYLP	
A. sal	(1)	-----ADTRPAFSRIVMFGDSLSDTGKMYSKMRGYLP	
A. hyd	(1)	-----ADSRPAFSRIVMFGDSLSDTGKMYSKMRGYLP	
Consenso	(1)	AD* RPAFSRIVMFGDSLSDTGKMYSKMRGYLP	
		51	100
P10480	(51)	SSPPYYEGRFNSNGPVWLEQLTNEFPGLTIANEAEAGGPTAVAYNKISWNP	
A. sal	(33)	SSPPYYEGRFNSNGPVWLEQLTKQFPGLTIANEAEAGGATAVAYNKISWNP	
A. hyd	(33)	SSPPYYEGRFNSNGPVWLEQLTKQFPGLTIANEAEAGGATAVAYNKISWNP	
Consenso	(51)	SSPPYYEGRFNSNGPVWLEQLT* * FPGLTIANEAEAGG* TAVAYNKISWNP	
		101	150
P10480	(101)	YQVINNLDYEVTFLOKDSFKPDDLVLWVGANDYLAYGWNTQDAKRV	
A. sal	(83)	YQVINNLDYEVTFLOKDSFKPDDLVLWVGANDYLAYGWNTQDAKRV	
A. hyd	(83)	YQVINNLDYEVTFLOKDSFKPDDLVLWVGANDYLAYGWNTQDAKRV	
Consenso	(101)	YQVINNLDYEVTFLOKDSFKPDDLVLWVGANDYLAYGWNTQDAKRV	
		151	200
P10480	(151)	DAISDAANRMVINGAKIILLFNLPDLGQNPSARSQKVVEAASHVSAYHNQ	
A. sal	(133)	DAISDAANRMVINGAKQILLFNLPDLGQNPSARSQKVVEAVSHVSAYHNK	
A. hyd	(133)	DAISDAANRMVINGAKQILLFNLPDLGQNPSARSQKVVEAVSHVSAYHNQ	
Consenso	(151)	DAISDAANRMVINGAK* ILLFNLPDLGQNPSARSQKVVEA* SHVSAYHN*	
		201	250
P10480	(201)	LLLNLARQLAFTGMVKLFEIDKQFAEMLRDPQNFGLSDQRNACYGGSYVW	
A. sal	(183)	LLLNLARQLAFTGMVKLFEIDKQFAEMLRDPQNFGLSDVENPCYDGGYVW	
A. hyd	(183)	LLLNLARQLAFTGMVKLFEIDKQFAEMLRDPQNFGLSDVENPCYDGGYVW	
Consenso	(201)	LLLNLARQLAFTGMVKLFEIDKQFAEMLRDPQNFGLSD* * N* CV* G* YVW	
		251	300
P10480	(251)	KPFASRSASTDSQLSAFNPQERLAIAGNPLLAQAVASPMARRSASTLNCE	
A. sal	(233)	KPFATRSVSTDRQLSAFSPQERLAIAGNPLLAQAVASPMARRSASPLNCE	
A. hyd	(233)	KPFATRSVSTDRQLSAFSPQERLAIAGNPLLAQAVASPMARRSASPLNCE	
Consenso	(251)	KPFA* RS* STD* QLSAF* PQERLAIAGNPLLAQAVASPMARRSAS* LNCE	
		301	336
P10480	(301)	GKMFWDQVHPTTVVHAALSEAATFIESQYEFIAH-	
A. sal	(283)	GKMFWDQVHPTTVVHAALSERAATFIEQYEFIAHG	
A. hyd	(283)	GKMFWDQVHPTTVVHAALSERAATFIANQYEFIAH-	
Consenso	(301)	GKMFWDQVHPTTVVHAALSE* AATFI* * QYEFIAH*	

FIGURA 53



Construcción génica para la expresión de KLM3'  
1225 pb

FIGURA 54

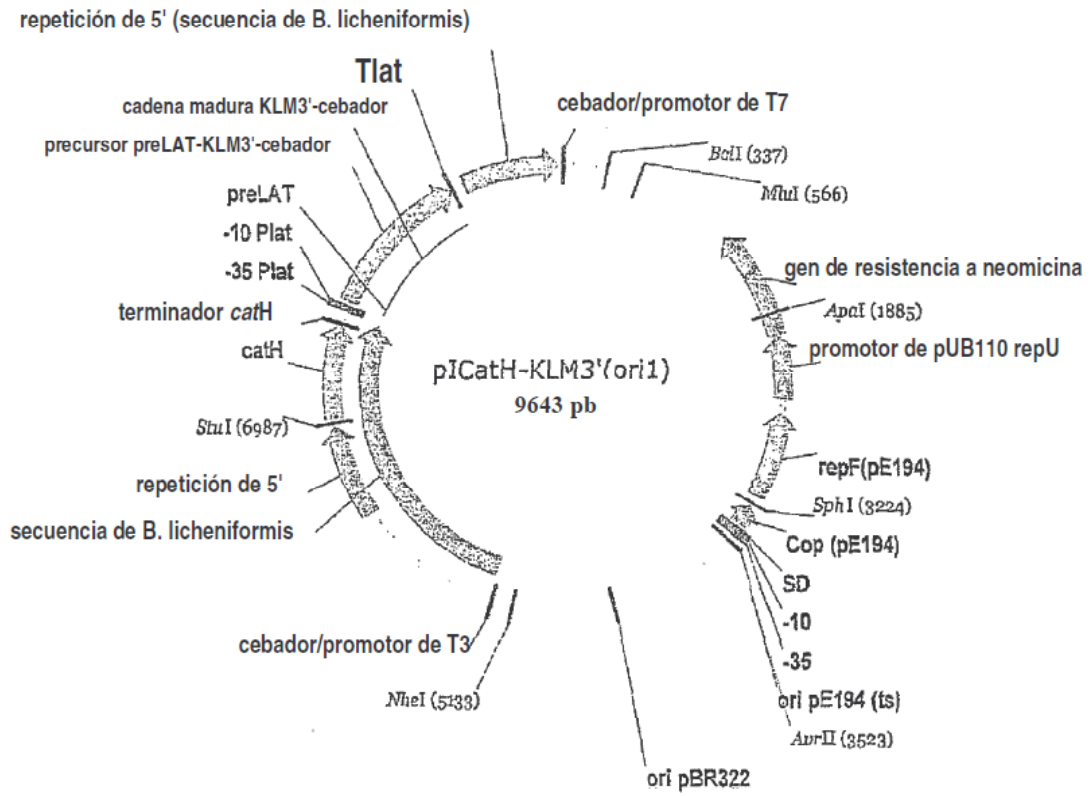


FIGURA 55

-35-

1 GCTTTTCTTT TCGAAGAAA TATAGGGAAA ATGGTACTTG TTTAAAATTC GGAATATTTA  
CGAAAAGAAA ACCCTCTTTT ATATCCCTTT TACCATGABC AATTTTTAAG CCTTATAAAT  
-10 M K Q Q K R L  
61 TACAAATATCA TATGTTTCAC ATTCGAAAGGG GAGGAGAAATC AAGRAACAAC AAAAACGGCT  
ATGTTATAGT ATACAAAGTG TAACTTTCCC CTCCTCTTAG TACTTTGTGTG TTTTGGCCGA  
\* Y A R L L T L L F A L I F L L P H S A A  
121 TTACGCCCGA TTGCTGACGC TGTATTTGC GCTCATCTTC TTGCTGCCTC ATTTGCGAGC  
AATGGGGGCT AACCACTGCG ACAATAAOC CGAGTRGAG AACGACGGG TAAGACCTCG  
\* S A A D T R P A F S R I V M F G D S L S  
181 TTCAGCAGCA GATACAAGC CGGCCTTTAG CCGGATCGTC ATGTTGGAG ATAGCCCTGAG  
AAGTCGTCTG CTATGTTCTG GCCGCAATC GGCTAGCAG TACAACCTC TATCGGACTC  
\* D T G K M Y S K M R G Y L P S S P F Y Y  
241 CGATACGGGC AAAATGTATA GCAAATGAG AGGCTATCTT CCGTCAAGCC CGCCGIATTA  
GCTATGCCCG TTTTACATAT CTTTTACTC TCCGATAGAA GGCAGTTCGG CGCCGCAATAT  
\* E G R F S N G P V W L E Q L T R Q F P G  
301 TGAAGCCCGC TTTAGCAATG CACCGGCTCG GCTGGAACAA CTGACCAAC AATTTCCGGG  
ACTTCCGGCG AAATCGTTAC CTGGCCAGAC CGACCTTGT TACTGCTTTG TTAAGGCCCC  
\* L T I A N E A E G G A T A V A Y N K T S  
361 ACTGACGATC GCTAATGAG CAGAAGGAGG AGCAACAGCC GTCCGCTATA ACATAATCAG  
TGACTGCTAG CGATTACTTC GTCTTCTCC TCGTTGTCCG CAGCCGATAT TGTPTTAGTC  
\* W D P R Y Q V I N N L D Y E V T Q F L Q  
421 CTGGGACCCG AAATATCAGG TCATCAACAA CCTGGACTAT GAAGTCACAC AGTTTCTTCA  
GACCCCTGGCC TTTATAGTCC AGTAGTTGTT GGACCTGATA CTTCACTGTG TCAAGAAGAT  
\* K D S F K P D D L V I L W V G A N D Y L  
481 GAAGACAGC TTTAAACCGG ATGATCTGCT CATCCTTTGG CTCGGCCCA ATGATTTATCT  
CTTCTGTGTC AAATTTGGCC TACTAGACCA GTAGGAACC CAGCCGCGGT TACTAATAGA  
\* A Y G W N T E Q D A K R V R D A I S D A  
541 GCGGTATGGC TGGAACACAG AACCAAGATGC CAAAAGAGTC AGAGATGCCA TCAGCGATGC  
CCGATACCG ACCTTEGTC TTGTTCTACG GTTTTCTCAG TCTCTACGGT AGTCGCTACG  
\* A N R M V L N G A K Q I L L F N L P D L  
601 CGCTAATAGA ATGCTCCCTGA ACGGCGCCAA ACAANTCCTG CTGTTTAACT TGCCGGATCT  
GCGATTATCT TACCAGGACT TGCCGCGGTT TGTTTAGGAC GACAAATTGG ACCGCCATAGA  
\* G Q N P S A R S Q K V V E A V S H V S A  
661 GGGACAAAAT CCGAGCCGCA GAAGCCAAA AGTCGTGAA GCAGTACGCC ATGTCAGCGC  
CCCTGTTTTA GGCTCGCGST CTTGCTTTT TCAGCAGCTT CGTCACTCGG TACAGTCCGG  
\* Y H N K L L L N L A R Q L A P T G M V K  
721 CTATCATAC AAATGCTGC TGAACCTGGC AAGACAATG GCACCGACGG GAATGGTTAA  
GATAGTATTG TTTGACGAGC ACTTGGACCG TCTGTTAAC CGTGGCTGCC CTTACCAATT  
\* L F E I D R Q F A E M L R D P Q N F G L  
781 ATTGTTTGAA ATTGACAAA AGTTTGCCGA AATGCTGAGA GATCCGCAAA ATTTTGGCCT  
TAACAAACTT TAACGTGTTG TCAAACGGCT TTACGACTCT CTAGGCGTTT TAAAACCGEA  
\* S D V E N P C Y D G G Y V W K P F A T R  
841 GAGCGATGTC CAAAACCGCT CCTATGATGG CCGATATGTC TGGAAACCGT TTGCCCAAG  
CTCGCTACAG CTTTGGGCA CGATACTACC GCCATACAG ACCTTTGGCA AACGGTGTTC  
\* S V S T D R Q L S A F S P Q E R L A I A  
901 AAGCTCAGC ACGGATAGAC AACTGTCAGC GTTTAGCCCG CAAGAAAGAC TGGCAATCGC  
TTCGCACTCG TGCCATCTG TTGACAGTCG CAAATCGGGC GTTCTTCTG ACCGTTAGCG  
\* G N P L L A Q A V A S P M A R R S A S P  
961 CGGAAATCCG CTTTGGCAC AAGCAGTTGC TTCACCGATG GCAAGAAGAT CAGCAAGCCC  
GCCTTAGGC GAAACCGTG TTCGTCAACG AAGTGGCTAC CTTTCTCTA GTCGTTCCGG  
\* L N C E G K M F W D Q V H P T T V V H A  
1021 GCTGAATTGC GAAGCCARA TGTTTGGGA TCAGGTCCAT CCGACAACAG TTGTCATGTC  
CGACTTAAC CTTCCGTTTT ACAAAACCTT AGTCCAGGTA GGCTGTGTC AACAGETACG  
\* A L S E R A A T F I E T Q Y E F L A H G  
1081 TGCCCTTCA GAAAGAGCGG CGACGTTTAT CGAARACAG TATGAATTTG TGCCCATGCG  
ACGGGAAAGT CTTTCTGCC GCTGCAATA GCTTGTGTC ATACTTAAAG ACCGGTACCC  
\* stop  
1141 CTGAGTTAAC AGAGGACGGA TTTCTGAGG GAAATCCGTT TTTTATTTT AAGCTTGGAG  
GACTCAATTG TCTCCTGCC AAAGGACTTC CTTTAGGCBA AAAAAAATA TTCGAACCTC  
1201 ACAAGGTAAA GGATAAACC TCGAG  
TGTCCATTT CCTATTTTGG AGCTC

**FIGURA 56**

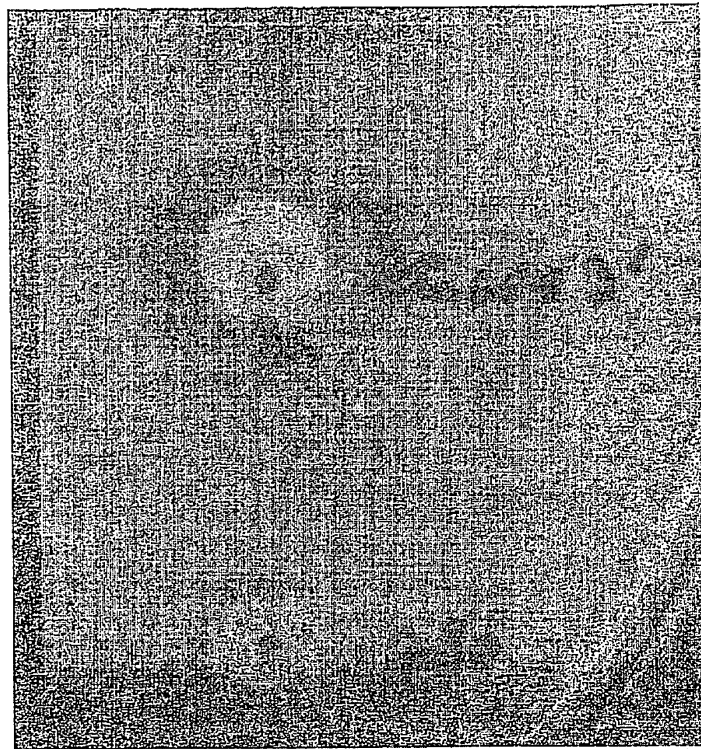


FIGURA 57 (SEQ ID N° 49)

```

1   ATGAAACAC AAAACGGCT TTACGCCGA TTGCTGACG TGTATTTGC
   TACTTTGTTG TTTTGGCCGA AATGCGGGCT AACGACTGCG ACPATAAACG

51  GCTCATCTT TTGCTGCCTC ATTCTGCAGC TTCAGCAGCA GATACAAGAC
   CGAGTAGAAG AACGACGGAG TAAGACGTGG AAGTCGTCTG CTATGTCTCTG

101 CGGCGTTTAG CCGGATCGTC ATGTTTGGAG ATAGCCTGAG CGATACGGGC
   GCCGCARATC GGCCTAGCAG TACAACCTC TATCGGACTC GCTATGCCCCG

151 AAAATCTATA GCAAAATGAG AGGCTATCTT CCGTCAAGCC CGCCGTATTA
   TTTTACATAT CGTTTACTTC TCCSATAGAA GGCAGTTCGG GCGGCATAAT

201 TGAAGGCCGC TTTAGCAATG GACCGGTCTG GCTGGACAA CTGACGAAAC
   ACTTCCGGCC AAATCGTTAC CTGGCCAGAC CGACCTTGTG GACTGCTTTG

251 AATTTCCGGG ACTGACGATC GCTAATGAAG CAGAAGGAGG AGCAACAGCG
   TTAAGGGCCC TGACTGCTAG CGATTACTTC GTCTTCTCTC TCEPTGTCGG

301 GTCGCCTATA ACAAATCAG CTGGGACCCC AAATATCAGG TCATCAACAA
   CAGCGGATAT TCTTTTACTC GACCCGTGGC TTTATAGTCC AGTAGTTGTT

351 CCTGGACTAT GAGTFCACAC AGTTTCTTCA GAAAGACAGC TTTAAACGGG
   GGACCTGATA CTTCACTGTG TCAAGAGAGT CTTTCTGTCT AAATTTGGCC

401 ATGATCTGGT CATCCTTTGG GTCGGCGCCA ATGATTATCT GCGGTATGGC
   TACTAGACCA GTAGAAACC CAGCCGCGGT TACTAATAGA CCGCATACCG

451 TGGAACACAG AACAAAGATC CAAAAGAGTC AGAGATGCCA TCAGCGATGC
   ACCTTGTGTC TTGTTCTACG GTTTTCTCAG TCTCTACGGT AGTGCCTACG

501 CGCTAATAGA ATGCTCCTGA ACGGCGCCAA ACAAACTCTG CTGTTTAAAC
   GCGATTATCT TACCAGGACT TCGCGCGGTT TGTTTAGGAC GACAAATGCG

551 TGCCGGATCT GGGACAAAT CCGAGCGCCA GAAGCCAAA AGTCGTGAA
   ACEGCTAGA CCTCTTTTA GGCPCCGGTT CTTGCGTTTT TCACAGCTT

601 GCAGTCAGCC ATGTCAGCCG CTATCATAAC AACTGCTGTC TGAACCTGGC
   CGTCAGTCGG TACAGTCGGG GATAGTATTG TTTGACGAGC ACTTGGACCG

651 AAGACAATTG GCACCGACGG GAATGCTTA ATGTTTGAA ATTGACAAC
   TTCTGTTAAC CGTGGCTGCC CTTACCAAT TAACAACCT TAAGTGTGTTG

701 AGTTTGGCGA AATGCTGAGA GATCCGCAA ATTTTGGCCT GAGCGATGTC
   TCARACGGCT TTACGACTCT CTAGCCGTTT TAAACCCGA CTCGCTACAG

751 GAAAACCCGT GCTATGATGG CCGATATGTC TGGAAACCGT TTGCCACAAG
   CTTTGGGCA CGATACTACC GCCTATACAG ACCTTTGGCA AACGGTGTTC

901 AAGCGTCAGC ACGGATAGAC AACTGTCAGC GTTTAGCCCG CAAGAAAGAC
   TTCGCAGTCG TGCCATCTG TTGACAGTCG CAAATCGGGC GTTCTTTCTG

851 TGGCAATCGC CGGAAATCCG CTTTGGGAC AAGCAGTTGC TTCACCGATG
   ACCGTTAGCG GCCFTTAGGC GAAAACCGTG TTCGTCAACG AAGTGGCTAC

901 GCAAGAGAT CAGCAAGCCC GCTGAATTGC GAAGCABAA TGTPTTGGGA
   CGTTCTTCTA GTCGTTGGG GACTTAAAG CTTCCGTTTT ACAAAACCTT

951 TCAGGTCCAT CCGACACAG TTGTCCATGC TGCCCTTCA GAAAGGCGG
   AGTCCAGGTA GGCTGTTGTC AACAGGTACG ACGGGAAAGT CTTTCTCGCC

1001 CGACGTTTAT CGAAACACAG TATGAATTC TGGCCCATGG CTGA
   GCTGCAATA GCTTTGCTGC ATACTTAAAG ACCGGGTACC GACT

```

FIGURA 58 (SEQ ID N° 50)

```

1 ATGAAAAAAT GGTTFCTGTG TTTATTGGGA TTGGTCGGCG TGACAGTTCA GGCAGCCGAC
51 AGCCCTCCCG CCTTCTCCCG GATCGTGTATG TTGGCGGACA GCCTCTCCGA TACCGGCAAAG
121 ATGTACAGCA AGATGCGCGG TTACCTCCCC TCCAGC0000 CCTACTATGA GGGCCGCTTC
181 TCCACCGGGC CCGTCTGGCT GGAGCAGCTG ACCAACGAST TCCCGGGCCT GACCATAGCC
241 AACGAGGCGG AAGGCGGACC GACCGCCGTG GCTTACAACA AGATCTCCTG GAATCCCAAG
301 TATCAGGTCA TCAACAACCT GGACTACGAG GTCACCCAGT TCCTGCAGAA AGACAGCTTC
361 AAGCCGGACG ATCTGGTGAT CCTCTGGGTC GCGGCCAAGC ACTATCTGGC CTATGGCTGG
421 AACACAGAGC AGGATGCCAA GCGGGTGGCG GACGCCATCA GCGATCGGGC CAACCCGATG
481 GTGCTGAACG GCGCCAAAGG AATACTGCTG TTCAACCTGC CGGATCTGGG CCAGAACC000
541 TCGGCCCGCA GCCAGAAAGT GGTGAGGGCG GCCAGCCATG TCPCCGCCTA CCACAAGCC00
601 CTGCTGCTGA ACCTGGCAGC CCAGCTGGCT CCCACCGGCA TGGTGAAGCT GTTCGAGATC
661 GACAAGCACT TTGCCAGAT GCTGCGTGAT CCGCAGAACT TCGCCCTGAG CGACCAGAGG
721 AACGCTGCT ACGGTGGCAG CTATGTATGG AAGCCGTTTG CCTCCCGCAG CGCCAGCACC
781 GACAGCCAGC TCPCCGCCTT CAACCCGCAE GAGCGCCTCG CCATCGCCGG CAACCCGCTG
841 CTGGCCAGG CCGTGGCCAG CCCATGGCT GCGCGCAGCG CCAGC0000 CAACTGTGAG
901 GGCAAGATGT TCTGGGATCA GGTCCACCCC ACCACTGTG TGACAGCCCG CCTGAGCGAG
961 CCCGCCGCA CCTTCATCGA GAGCCAGTAC GAGTTCCTCG CCCAC

```

FIGURA 59 (SEQ ID N° 51)

```

1 ATGAAAAAAT GGTTFGTTG TTTATTGGSG TTGATCGCGC TGACAGTTCA GGCAGCCGAC
51 ACTCGCCCGC CCTTCTCCCG GATCGTGTATG TTGGCGGACA GCCTCTCCGA TACCGGCAA
121 ATGTACAGCA AGATGCGCGG TTACCTCCCC TCCAGC0000 CCTACTATGA GGGCCGCTTC
181 TCCACCGGAC CCGTCTGGCT GGAGCAGCTG ACCAAGCAGT TCCCGGGCTT GACCATCGCC
241 AACGAGCGG AAGGCGGTGC CACTGCGGTG GCTTACAACA AGATCTCCTG GAATCCCAAG
301 TATCAGGTCT ACAACAACCT GGACTACGAG GTCACCCAGT TCCTGCAGAA AGACAGCTTC
361 AAGCCGGACG ATCTGGTGAT CCTCTGGGTC GGTGCCAATG ACTATCTGGC ATATGGCTGG
421 AATACGGAGC AGGATGCCAA GCGAGTTCGC GATGCCATCA GCGATCGGGC CAACCCGATG
481 GTACTGAACG GTGCCAAGCA GATACTGCTG TTCAACCTGC CGGATCTGGG CCAGAACC000
541 TCAGCCCGCA GTCAGAAAGT GGTGAGGGCG GTCAGCCATG TCPCCGCCTA TCACAACAAG
501 CTGCTGCTGA ACCTGGCAGC CCAGCTGGCC CCCACCGGCA TGGTGAAGCT GTTCGAGATC
561 GACAAGCAAT TTGCCAGAT GCTGCGTGAT CCGLAGAACT TCGCCCTGAG CGACGTGAG
721 AACCCCTGCT ACEACGGCGG CTATGTGTGG AAGCCGTTTG CCACCCGCAE GGTGAGCACC
781 GACCGCCAGC TCPCCGCCTT CAGTCCGCAE GAACGCTCG CCATCGCCGG CAACCCGCTG
841 CTGGCAGAG CCGTGGCCAG TCCTATGGCC GCGCGCAGCG CCAGC0000 CAACTGTGAG
901 GGCAAGATGT TCTGGGATCA GGTACACCCG ACCACTGTG TGACAGCAGC CCTGAGCGAG
961 CGCGCCGCA CCTTCATCGA GACCCAGTAC GAGTTCCTCG CCCACGGAAG A

```

FIGURA 60 (SEQ ID N° 52)

```

1 ATGCCGAAGC CTGCCCTTG CCGTGTATG ACCGCGACAG TCGCCGCGGT CCGCACGCTC
51 GCCCTCGGCC TCACCGACGC CACCGCCAC GCCGCGCCCG CCCAGGCCAC TCCGACCCTG
121 GACTACGTG CCTCGGCGA CAGCTACAGC GCGGGCTCCG GCGTCTTGCC GGTGACCC00
181 GCCAACCTGC TCTGTCTGCG CTCGACGGCC AACTAC0000 ACGTCATCGC GGACACGACC
241 GCGGCGCGCC TCACGGACGT CACCTGCGGC GCGCGGAGAG CCGCCGACTT CACGCGGGCC
301 CAGTACCGGG GCGTCCGACC CCAGTTGGAC GCGCTCGGCA CCGGCACGGA CCTGTTGACG
361 CTCACCATCG GCGGCAACGA CAACAGCACC TTCATCAAGC CCATCAGGCG CTGCGGCAGC
421 CCGGCTGTC TCAGCGGCGG CAAGGGCAGC CCCTGCAAGG ACAGGCAAGG CACCTCCTTC
481 GACGACGAGA TCGAGGCCAA CACGTACCCC GCGCTCAAGG AGGCGCTGCT CCGGTCGCGC
541 GCCAGGGCTC CCCACGCCAG GGTGCGGGCT CTCGGCTACC CGTGGATCAC CCGGGCCACC
601 CCGGACCCGT CTCGCTTCTT GAAGCTCC00 CTCGCGCGCG GTTAGCTGCC CTACCTGCGG
661 GCCATCCAGG CACACCTCA CAACGCGGTC CCGCGGGCGG CCGAGGAGAC CCGAGCCACC
721 TACGTGGACT TCTCCGGGGT GTCCGACGCG CAGCAGCGCT GCGAGGCC00 CCGCACCCCG
781 TGGATCGAAC CGCTGCTCTT CCGGCACAGC CTCGTTCCCG TCCACCCCAA GCGCCCTGGG
841 GAGCGGCGCA TCGCCGAGCA CACGATGAGC GTCCTCGGCC TGGACTGA

```

FIGURA 61 (SEQ ID N° 53)

```

1 TCASTCCAGG CCGAGGACGT CCATCGTGTG CTCGGCCATG CGCCGCTDGC CCAGGGCGFTT
61 GGGSTGGACG GGAACGAGGC TGTGCCGGAA GAGCAGGGST TCGATCCAGC GGGTGCCTGG
121 GGCCTCGCAG GCGTCTGTGGC CGTCCGGACAC CCGGGAGGAG TCCACCTAGG TGGCTCCGGT
181 CTCCTCGGCG GCCCGCCGGA CCGCGTCTGT GAGGTGTGCC TGGATGGCCC GCAGGTAGGG
241 CACGTACCCG GCGGCGAGGG GGAGCTTCAG GAAGCAGGAC GGGTCCGGCG TGGCCGGST
301 GATCCACGGG TAGCCGAGAG CCGCCACCCT GCGTGGGGA GCCCTGGCGC GGACGCCGAG
361 CAGCGCTTCC TTGAGCGCGG GGTACGTGTG GGCCTCGATC TCGTCTCGA AGGAGGTGCC
421 GTCCCTGTCC TTGACGGGGC TGCCCTTGCC GCCGCTGAGG ACACCCGCGG TGCCCGAGGC
481 CGTGATGGCG TTGATGAAGG TGCTGTGTGC GTTGCCCGCG ATGGTTAGCG TGACCAGGTC
541 CGTCCCGGTG CCGAGCGCGT CCAACTGGGG TGCGACGCCC GGGTACTGGG CCCCGGTGAA
601 GTCGCGGTC TCGCGCGCGC CGCAGGTGAC GTCCTGAGG CCGGCGCCCG TCGTGTCCSC
661 GATGACGTGG GGGTAGTGG CCGTCCAGCG CAGACAGGAC AGGTTGGCGG GGTGACGGG
721 CAGGACGCCG GAGCCGCGCC TGTAGCTTFC GCCGAGGGCG ACCTAGTCCA GGGTCCGAGT
781 GGCCTGGCGC GCGCGCGCGT GGGCGGTGGC GTCGGTAGAG CCGAGGGGGA CGGTGCCGAC
841 GCGCGCGACT GTCGCGGTCA TGACACGGCC AAGGGCAGGC TTCGGCAT
    
```

FIGURA 62 (SEQ ID N° 54)

```

1 ATGGATTACG AGAAGTTTCT GTTATTTGGG GATTCCATTA CTGAATTTGC TTTTAACTACT
61 AGGCCCATTE AAGATGGCAA AGATCAGTAT GCTCTTGGAG CCGCATTAGT CAACGAATAT
121 ACGAGAAAAA TGGATATTCT TCAAAGAGGG TTCAAAGGGT ACACTTCTAG ATGGGCGTTG
181 AAAATACTTC CTGAGATTTT AAAGCATGAA TCCAAATATG TCATGGCCAC AATATTTTTG
241 GGTGCCAACG ATGCATGCTC AGCAGGTCC CAAAGTGTCC CCTTCCCGA ATTTATCGAT
301 AATATTCGTC AATGGGTATC TTTGATGAG TCTTACCATA TCCGTCTTAT TATAATAGGA
361 CCGGGGCTAG TAGATAGAGA GAAGTGGGAA AAAGAAAAAT CTGAAGAAAT AGCTCTCGGA
421 TACTTCCGTA CCAACGAGAA CTTTGGCATT TATTCAGATG CCTTAGCAA ACTAGCCAAAT
481 GAGGAAAAAG TTCCCTTCGT GGCTTTGAAT PAGGCGTTTC AACAGGAGG TGGTGATGCT
541 TEGCAACAAC TGCTAACAGA TGGACTGCAC TTTTCCGGAA AAGGGTACAA AATTTTTCAT
601 GACGAATTAT TGAAGGTCAT TGAGACATTC TACCCCAAT ATCATCCCAA AAACATGCAG
661 TACAACTGA AAGATTGGAG AGATGTGCTA GATGATGGAT CTAACATAAT GTCTTGA
    
```

FIGURA 63 (SEQ ID N° 55)

```

atgaacctgc gtcaatggat gggcgccgcc acggctgccc ttgccttggg cttggccggg 60
tgcgggggcg gtgggaccya ccagagcggc aatcccaatg tcgccaaggt gcagcgcatg 120
gtggtgttcg gcgacagcct gagcgtatc ggcaectaca cccccgtcgc gcagggcgtg 180
ggcgcgcgca agttcaccaac caaccggggc ccgatctggg ccgagaccgt ggcggcgcaa 240
ctggcggtga cgtcaccgccc ggcgggtgatg ggctacgcca cctccgtgca gaattgcccc 300
aaggccggct gcttcgacta tgcgcagggc ggctcgcgcy tgaccgatcc gaacggcatc 360
ggccacaacg gcggcgcggg ggcgctgacc taccgggttc agcagcagct cgccaacttc 420
tacgcygcca gcaacaacac attcaacggc aataacgatg togtctctgt gctggccggc 480
agcaacgaca ttttctctcg gaccactgcy gcggccacca gcggctccgg cgtgacgccc 540
gccattgcca cggcccaggc gcagcagggc gcgacggacc tggctgggta tgtcaaggac 600
atgatcgcca aggggtgcgac gcaggtctac gtgttcaacc tgcccagacg cagcctgacg 660
ccggaacggc tggcaagcggc cagcaccggc caggcgtgcy tgcaacgctt ggtgggcaac 720
ttcaacacga cgctgcaaag cgggctggcc ggcaacctcg cgcgcatcat cgacttcaac 780
gcacaactga ccgcgcgcat ccagaatggc gectcgttcg gcttcgpcaa caccagcgcc 840
cgggctgcy acgccacaa gatcaatgcy ctggtgcga gcgcggcgcy cagctcgtg 900
ttctgctcgg ccaacacgct ggtggttcc ggtgcggacc agagctacct gttcggcagc 960
ggcgtgcacc cgaccacggc cggccatcgc ctgatcgcca gcaacgtgct ggcgcgctg 1020
ctgcggata acgtcgcga ctga 1044
    
```



FIGURA 64 (SEQ ID N° 56)

```

1  gtgatcgggt  cgtacgtggc  ggtgggggac  agcttcaccg  agggcgctcg  cgaccccggc
61  cccgacgggg  cgttcgtcgg  ctgggcccga  cggctcgcgc  tactgctcgc  ggaccggcgc
121  cccgagggcg  acttcacgta  caccgaacct  gccgtgcgcg  gcaggctcct  cgaccagatc
181  gtggcggaac  aggtcccgcg  ggtcgtcggc  ctccgcccgc  acctcgtctc  gtccgcccgc
241  ggccggcaacg  acatcatccg  gcccccgcac  gatcccgcag  aggtcgccga  gccggttcgag
301  ctggcgggtg  ccgcgctgac  cgcgcgggcc  ggaaccgtcc  tggtagccaa  cggggttcgac
361  acccgggggg  tgcccgtcct  caagcacctg  cggggcaaga  tcgcccagta  caacggggac
421  gtcgcgcgca  tcgcccagcc  ctacggctgc  ccggtgctcg  acctgtggtc  gctcgggggc
481  gtcaccggacc  gcaggggcgt  ggaccggcac  cggctgcacc  tctcggccga  ggggacaccc
541  cgggtggcgc  tgcgcgggg  gcaggccctg  ggctgcgcg  tcccggccga  ccctgaccag
601  ccctggccgc  ccctgcccgc  gccggccagc  ctccgacctc  gggcgacgca  cgtgcaactg
661  gcgcgcgagt  acctggtgac  gtggatcggg  cgcggctgac  ggggagagtc  gtcggggcag
721  cactgacgg  ccaaggggac  gctgtcgcgc  gaccgatca  agaccgggat  cggcgggtg
781  gcctga
    
```

FIGURA 65 (SEQ ID N° 57)

```

1  atgcagacga  accccgcgta  caccagtctc  gtccgcctcg  gcgactcctt  caccgagggc
61  atgtcggacc  tgcctgccga  cggctcctac  cgtggctggg  ccgacctcct  cgccaccggc
121  atggcggccc  gctcccgcgg  ctcccggtac  gccaacctgg  cggtagccgg  gaagctgac
181  ggacagatcg  tcgacgagca  ggtggacgtg  gcccccacca  tgggagccga  cgtgatcacg
241  ctggtcggcg  ggtcaacga  cacgctgcgg  cccaagtggc  acatggcccc  ggtcggggac
301  ctgctgaccc  agcccgtaga  acggctcggc  ccgcactggc  agcagctggt  gctgatgccc
361  agtcccggtc  gccagggtcc  ggtgctggag  cgttccggcc  cccgcatgga  ggccctgttc
421  gccgtgatcg  acgacctggc  cgggcccgcac  ggcgcccgtg  tcgtcgacct  gtaccggggc
481  cagtgcctcg  ccgaccctcg  gatgtgggac  gtggaccggc  tgcaactgac  ccgcccagggc
541  caccgcgggg  tcgcccgggg  ggtgtggcag  tcgctcggcc  acgagcccga  ggacccccag
601  tggcacgcgc  cgatcccggc  gaccgcccgc  ccggggtggg  tgaccgcccag  gaccggggac
661  gtcgggttcg  cccggcagca  cctgctgccc  tggataggcc  gcaggctgac  cgggcgctcg
721  tccgggggac  gccctcccgc  caagcggccc  gacctgctgc  cctaccgagg  ccccgcacgg
781  tga
    
```

FIGURA 66 (SEQ ID N° 58)

```

1  atgacccggg  gtcgtgacgg  gggtaggggg  ggcggcccca  ccaagcaocg  tgcctgtctc
61  gcggcgatcg  tcaccctgat  agtggcgatc  tccgcccgca  tatacgccgg  agcgtccggc
121  gacgacggca  gcagggacca  cgcgctgcag  gccggaggcc  gtctcccacg  aggagacgoc
181  gcccccgcgt  ccaccgggtg  ctgggtgggg  gccctgggcca  ccgcaaccgg  cgcggtccgag
241  ccgggcaacc  agaagcaocg  cctggcgggg  cgtccctgtc  gcaacgtcgt  gcacacctcg
301  gtcggcgcca  ccggcggcgc  gatcacccct  tcgaaacctg  acgggcagtc  gccgctgacc
361  gtcacacacg  cctgatcgc  cctggccgcc  gggcccga  ccgcccgcgc  gatcggcgac
421  accatgcgcc  ggctcacctt  cggcggcagc  gcccggtgga  tcaatcccgc  gggcggccag
481  gtgatgagcg  acaccgccc  cctcgcctac  cctacgggg  cgaacgtcct  ggtcaccaag
541  tactccccca  tcccgtccgg  gccggtgacc  taccatccgc  agggcccggc  gaccagctac
601  ctggccgacg  gcgaccgcac  ggccggactc  accgcccgtc  cgtacaccac  ccccaccgccc
661  tactggcgct  acctgaccgc  cctcgcctg  ctgagccaag  aggcggacgg  caccgctcgtg
721  gcgttcggcg  actccatcac  cgaagcggcc  cgtcgcgaga  ggaagcccaa  ccaaccgctgg
781  accgacgtcc  togccgcagc  cctgcaccag  gcggcggggc  acggcccggg  caagcccggc
841  taacagctcg  tcaacgaggg  catcagcggc  aaccggctcc  tgaccagcag  gccggggcgg
901  ccggccgaca  acccgagcgg  actgagccgg  ttccagcggg  acgtgctgga  acgacccaac
961  gtcaggcccg  tcgtcgtcgt  cctcggcgtc  aacgacgtcc  tgaaacgccc  ggaactcgcc
1021  gaccggcagc  ccacacctgac  cggccctgccc  accctcgtcg  accggggcga  ccgcccgggga
1081  ctgcggttcg  tcggcccccac  gatcacggcc  ttccggcgct  accggggcta  caccgaggcc
1141  ccgagacgca  tcgycagga  ggtcaaccag  gagatccgct  ccggccgggt  ctccgacacg
1201  gtcgtcgact  tcgacaaggc  cctgcccgcac  ccgtaccgac  cgcgcccgat  gccctccgac
1261  tacgacagcg  gcgaccacct  gcccccggcc  gacaaggggt  accgcccgat  gggcggggtc
1321  atcgacctgg  ccgctgtaa  gggcggggcg  ccggtcaagg  cgtag
    
```

FIGURA 67 (SEQ ID N° 59)

```

1 atgacgagca tctcagggcc gaggttggcg cggcggatcg cggccggcgc ggcgtacggc
61 ggcggcggca tggccctggc gggagcggcg ggggtcggte tgggtggggc ctaggtgcag
121 ctggccagac gcagggtggg ggtggccacg ccgaccgggg tggcgaacgc gcagggatcg
181 tacggcggca ccctgccacc ggcggggcac ccgcccgtgc ggctgatgat gctggggcag
241 tccacggccc cggggcaggc cgtgcaccgg gccgggcaga cggcggcgc gctgctggcg
301 tccgggctcg cggcggggcc ggaacggccg gtgcggctgg ggtcggtcgc ccagccgggg
361 gcgtgctcgg acgacctgga ccggcagggt gcctgggtgc tggccgagcc ggaaccgggtg
421 cccgacatct gcgtgatcat ggtcggggcc aacgacgtca cccaccggat gccggcgacc
481 cgctcgggtg ggcacctgtc ctggcgggta cggcggctgc gcaccggccg tggcggaggtg
541 gtggtcggca cctgtccggc cctggggcac atcgagcggg tggcggacc gctgcccgtgg
601 ctggcccggc gggcctcacg gcagctcggc gccggcacaga ccacggcgc cgtcggagcag
661 ggcggcggca cgggtcggct gggcggaccg ctgggtccgg agttcggcga gaaccggcgg
721 gagctcttcg gccccgacaa ctaccacccc tccggcgggg ggtaccgcc cggcggcag
781 gcggtactgc cctcgggtgg cgcggcgtcc ggcctgtggc cggcggcaga ggaaccaccg
841 gacgcgctgc gcggcggggg ctccctcggc gtggcggcgg cggcggcggg gcggcggctc
901 gagcgggta cggaggtcgc cggcccatg cctaccgggg ctggggggcc cttggcggctt
961 ctgaaaggcc ggaagcggcg tgggtgtcg gagcgggaac cgtccagccc gtcggcgggtt
1021 tga

```

FIGURA 68 (SEQ ID N° 60)

```

1 atgggtcag ggacggacca cggacggcg tacggccctc gccggggcgc tctcggctc
61 gccgccctga ccggcccgct cctggcggcg ggcgtggcgg cctgcgactc cgtggggcgc
121 gactcaaccg ctccctccgg cagcccgtcg aagcggacga ggaacggccc cgcctgggac
181 accagcccgg cgtccgtcgc cggcgtgggc gactccatca cggcggcctt cgaacccgtg
241 gcggtgctgt cggactgccc gagggtgtcg tggcggaccg gcagcagcgc gaaggtcag
301 tgcctggcgc tacggctgct ggggaaggcg gacggcggcc agcacagctg gaactaccgg
361 gtcaccccgg cccggatggc ggaacctgacc gctcaggtga cggggggcgg gcaagcggag
421 ccggagctgg tggcgggtat ggcggggcgg aacgacgcgt gccgggtccc gacctcggcg
481 atgacggcgg tggcggactt ccggggcggc ttcgaggagg ngatggcccac cctggcgaag
541 aagctcccca aggcgcaggt gtactgtctg agcatcccgg acctcaagcg gctctggtcc
601 cagggcccga ccaaccctct gggcaagcag gtgtggaagc tggcctgtg cccgtcagtg
661 ctggggcagc cggatccctt ggactcggcg gcgaccctgc ggcgcaacac ggtgcggcag
721 cgggtggcgg actacaacga ggtgctcggg gaggtctgog cgaaggaccg gcgggtcccg
781 agcagcagc gcggcgtgca cgaattccgg ttcggcacgg accagttgag ccactgggac
841 tgggtccacc cgaattggtga cggccaggcc cggctggcgg agatcgcta cggcggcgtc
901 accgcgaaga atccctga

```

FIGURA 69 (SEQ ID N° 61)

```

1 ttcatacaaa cgaatgcaca acaccggcca tccgggtcat cctgatcgt gggaatgggt
61 gacaagcctt ccggtgacga aagggtcctg ctacatcaga aatgacagaa atcctgctca
121 ggggaggtcc atgagactgt ccgacgcgc gggcacggcg tccgctctcc tcccaacccc
181 ggcctcggca ctctcggcgg cagcggccgc cgtgtccgcg ccggcaatcc agggcaccga
241 ctacgtggcc ctccggcact ctaactcctc ggggtctggc gggggcagct acgacagcag
301 cagtggtccc tgtaaagcga gcccaagtc ctaccggcc cttggggcgg cctcgcaac
361 cggtagcggg ttcaacttca ccgctgttc gggcggccc caggagagag tgtggccaa
421 gacgtgacc ccggtcaact ccggcaccga cctgggtcagc attaccatcg gcggcaacga
481 cgggggcttc gccgacacca tgaccactcg caacctccag ggcgagagcg cgtgctggc
541 gcggatcggc aaggcggcgg ctaacatcca gcagacgctg cccggcccagc tggaccaggt
601 ctacagagcc atcgacagcc gggcccggcg agcccaggtc gtcgtcctgg gctaccggcg
661 cttctacaag ctggggcggca gctgcggcgt cggctctctg gagaagtccc gcggggccat
721 caecggcccg gccgagcaca tcaacggcgt caccggcaag cggcggcggg aocagggctg
781 cgccctcggg caggtcaaca gaccttcgc cgggcaagag cttgctcgg gcggcccctg
841 gctgcacagc gtcacccctc ccgtggagaa ctctaccac cccacggcca acggacagtc
901 caagggctac ctgcccgtcc tgaactcggc caactgatct cgggctact ccggccctga
961 cgaagtcccg ccccggggcg gggctcggc gtagggtcgc gtaccggcgt gcggcgtcgc
1021 gccgggtggc ccggcgtacg tggcggcggc cccggacggc gtcgggtc

```

FIGURA 70 (SEQ ID N° 62)

```

1  ATGAAAAAT  GGTTCGTGTG  TTTATTGGGA  TTGGTCGCGC  TGACAGTTCA
   TACTTTTTTA  CCAAACACAC  AATAAACCCT  AACCAGCGCG  ACTGTCAAGT

51  GGCAGCCGAC  AGTCGCCCCG  CCTTTCCCGC  GATCGTGATG  TTCGGCGACA
   CCGTCGSCCT  TCAGCGGGGC  GGRAAAGGGC  CTAGCACTAC  AAGCCCTGTG

101  GCCTCTCCGA  TACCGGCAA  ATGTACAGCA  AGATGCGCGG  TTACCTCCCC
   CGGAGAGGCT  ATGGCCGTTT  TACATGTCTG  TCTACGCGCC  AATGAGGGGG

151  TCCAGCCCGC  CCTACTATGA  GGCCTGTTTC  TCCAACGGAC  CCGTCTGGCT
   AGGTCCGGCG  GGATGATACT  CCCGGCAAAG  AGGTTGCCTG  GGCAGACCGA

201  GGAGCAGCTG  ACCAAACAGT  TCCCGGGTCT  GACCATCGCC  AACGAAGCGG
   CCTCGTCGAC  TGGTTGTCA  AAGGCCCAGA  CTGGTAGCGG  TTGCTTCGGC

251  AAGGCGGTGC  CACTGCCGTG  GCTTACAACA  AGATCTCCTG  GAATCCCAAG
   TTCGCCACG  GTGACGGCAC  CGAATGTGT  TCTAGAGGAC  CTTAGGGTTC

301  TATCAGGTCA  TCACACACCT  GGACTACGAG  GTCACCCAGT  TCTTGAGAA
   ATAGTCCAGT  AGTTGTTGGA  CCTGATGCTC  CAGTGGGTCA  AGAAGCTCTT

351  AGACAGCTTC  AAGCCGGACG  ATCTGGTGT  CCTCTGGTTC  GGTGCCAATG
   TCTGTGAGAG  TTCGGCCTGC  TAGACCACTA  GGAGACCCAG  CCACGGTTAC

401  ACTATCTGSC  CTRTGGCTGG  AACACGGAGC  AGGATGCCAA  GCGGGTTCGC
   TGATAGACCG  GATACCGACC  TTGTGCTCG  TCCTACGGTT  GCGCCAGCGG

451  GATGCCATCA  GCGATGCGGC  CAACCGCATG  GFACTGAACG  GTGCCAAGCA
   CTACGCTAGT  CGCTACGCCG  GTTGGCGTAC  CATGACTTGC  CAGGTTCTGT

501  GATACTGCTG  TTCAACCTGC  CGGATCTGGG  CCAGAACCCG  TCAGCTCGCA
   CTAAGACGAC  AAGTTGGACG  GCCTAGACCC  GGTCTTGGGG  AGTCGAGCGT

551  GTCAGRAGGT  GGTGAGGGCG  GTCAGCCATG  TCTCCGCTTA  TCACAACCGG
   CAGTCTTCCA  CCAGCTCCGC  CAGTCGGTAC  AGAGCGGGAT  AGTGTGCTC

601  CTGCTGCTGA  ACCFGGCAGC  CCAGCTGGCC  CCCACCGGCA  TGSTAAAGCT
   GACGACGACT  TGCACCGTGC  GSTCGACCGG  GGGTGGCCGT  ACCATTTCGA

651  GTTCGAGATC  GACRAGCAAT  TTGCCGAGAT  GCTGCGTGAT  CCGCRAACT
   CAAGCTCTAG  CTGTTCTGTA  AACGGCTCTA  CGACGCACTA  GCGCTTTGA

701  TCGGCTGAG  CGACGTCGAG  AACCCCTGCT  ACGACGGCGG  CTATGTGTGG
   AGCCGGACTC  GCTGCAGCTC  TTEGGGACGA  TGCTGCCGCC  GATACACACC

751  AAGCCGTTTG  CCACCCGAG  CGTCAGCACC  GACCGCCAGC  TCTCCGCTT
   TTCGGCAAAC  GGTGGCGCTC  GCAGTCTGTG  CTGGCGGTTC  AGAGCGGGAA

801  CAGTCCGACG  GAACGCCCTG  CCATCGCCGG  CAACCCGCTG  CTGGCACAGG
   GTCAGGCCGC  CTTGCGGAGC  GGTAGCGGCC  GTTGGGCGAC  GACCGTGTCC

851  CCGTTCGACG  TCCTATGGCC  CGCCGACGG  CCAGCCCCCT  CAACTGTGAG
   GGCACGGCTC  AGGATACCGG  GCGGCGTCCG  GGTGCGGGGA  GTTGCACACTC

901  GGCAAGATGT  TCTGGGATCA  GGTACACCCG  ACCACTGTGC  TGCACGACGC
   CCGTTCACA  AGACCCTAGT  CCATGTGGGC  TGGTGACAGC  ACSTGCCTCG

951  CCTGAGCGAG  CGCGCCGCCA  CCTTCATCGC  GAACCACTAC  GAGTTCCTCG
   GGACTCGCTC  GCGCGGCGGT  GSAAGTAGCG  CTTGGTCATG  CTCAAGGAGC

1001  CCCAC  TGA
      GGGTG  ACT
    
```

FIGURA 71 (SEQ ID N° 63)

```

1  ATGAAAAAT  GGTTFGTTF  TTTATTTGGG  TTGATCGGCG  TGACAGTTCA
   TACTTTTTTA  CCAACAAAC  AAATAACCCC  AACTAGCGCG  ACTGTCAAGT

51  GGCAGCCGAC  ACTCCCCCG  CCTTCTCCCG  GATCGGTGAT  TTCGGCGACA
   CCGTCCGCTG  TGAGCCGGGC  GGAAGAGGGC  CTAGCACTAC  AAGCCCTGT

101  GCCTCTCCGA  TACCGGCAA  ATGTACAGCA  AGATGCGGCG  TTACCTCCCC
   CCGAGAGGCT  ATGGCCGTTT  TACATGTCTG  TCTACGCGCC  AATGGAGGGG

151  TCCAGCCCGC  CCTACTATGA  GGGCCGTTF  TCCACGGAC  CCTCTGGCT
   AGGTCCGGCG  GGTGTACT  CCCGGCAAAG  AGGTTCCTG  GGCAGACCGA

201  GGAGCAGCTG  ACCAAGCAGT  TCCCGGTCT  GACCATCGCC  AACGAAGCGG
   CCTCGTCGAC  TGGTTCGTCA  AGGGCCAGA  CTGGTAGCGG  TTGTTCCGC

251  AAGGCGGTGC  CACTGCCGTG  GTTACAACA  AGATCTCCTG  GAATCCCAAG
   TTCGGCCAGC  GTGACGGCAC  CGAATGTGT  TCTAGAGGAC  CTAGGGTTG

301  TATCAGGTCA  TCAACACCT  GSACTACGAG  GTCACCCAGT  TCTTGCAGAA
   ATAGTCCAGT  AGTTFETGG  CCTGATGCTC  CAGTGGGTCA  AGAACCTTT

351  AGACAGTTTC  AAGCCGGACG  ATCTGGTGAT  CCTCTGGGTC  GGTGCCAATG
   TCTGTCAAG  TTCGCCCTGC  TAGCCACTA  GGAGACCCAG  CCRCGGTTAC

401  ACTATCTGGC  ATATGGGTGG  AATACGGAGC  AGGATGCCAA  GCGAGTTCCG
   TGATAGACCG  TATACCGACC  TTATGCCCTG  TCCTACGGTT  CGCTCAAGCG

451  GATGCCATCA  GCGATGGCG  CAACCCGATG  GTACTGAACG  GTGCCAAGCA
   CTACGGTAGT  CGCTACGCCG  GTTGGCGTAC  CATGACTTGC  CACGGTCTGT

501  GATACTGCTG  TTCAACTGCG  CGGATCTGGG  CCAGAACCCG  TCAGCCCGCA
   CTATGACGAC  AAGTTGGACG  GCCTAGACCC  GGTCTTGGGC  AGTCGGGCGT

551  GTCAGAAGGT  GGTGAGGGCG  GTCAGCCATG  TCTCCGCTTA  TCAACAACAG
   CAGTCTTCCA  CCAGCTCCGC  CAGTCCGTAC  AAGGGCGGAT  AGTGTGTGTC

601  CTGCTGCTGA  ACCTGGCAGC  CCAGCTGGCC  CCCACCGGCA  TGGTAAAGCT
   GACGACGACT  TGGACCTGCG  GGTCCACCCG  GGGTGGCCCT  ACCATTTGCA

651  GTTCGAGATC  GACAAGCAAT  TTGCCGAGAT  GCTGCGTGT  CCGCAGAACT
   CAAGCTCTAG  CTGTTCGTTA  AAGGGCTCTA  CGACGCACTA  GCGCTCTTGA

701  TCGGCCTGAG  CGACGTCGAG  AACCCCTGCT  ACGACGGCGG  CTATGTGTGG
   AGCCGGACTC  GGTGCAGCTC  TTGGGGACGA  TGCTGCCGCC  GATACACACC

751  AAGCCGTTTG  CCACCCGAG  CGTCAGCACC  GACCGCCAGC  TCTCCGCTT
   TTCGGCAAAC  GGTGGGCTC  GCAGTCTGTC  CTGGCGGTCG  AGAGGCGGAA

801  CAGTCCGAG  GAACGCCTCG  CCATCGCCCG  CAACCCGCTG  CTGGCACAGG
   GTCAGGCGTC  CTGCGGAGC  GGTAGCGGCC  GTTGGGCGAC  GACCCTGTCC

851  CCGTTGCCAG  TCCTATGGCC  CGCCGACGCG  CCAGCCCCCT  CAAGTGTGAG
   GGCRAACGGTC  AGGATACCGG  GCGGCGTCCG  GGTCCGGGGA  GTTGACACTC

901  GGCAAGTGT  TCTGGGATCA  GGTACACCCG  ACCACTGTGCG  TGCACCGCAG
   CCGTFTTACA  AGACCCTAGT  CCATGTGGGC  TGGTGACAGC  ACGTGCCTCG

951  CCTGAGCGAG  CCGCCGSCA  CCTTCATCGA  GACCCAGTAC  CAGTTCCTCG
   GSACTCCCTC  CCGCGGCGGT  GGLRGTAGCT  CTGGGTCATG  CTCAGGAGC

1001  CCCACGGATG  A
      GGGTGCCTAC  T
    
```

FIGURA 72 (SEQ ID N° 24)

```

1  ATGTTTAACT TTAAAAAGAA TTCTTAGT  GGATTATCGG CAGCTTTAAT
   TACAAATTC AATTTTCTT AAGAATCAA CCTAATAGCC GTCGAAATTA

51  GAGTATTAGC TTGTTTTCGG CAACCGCCTC TGCAGCTAGC GCCGACAGCC
   CFCATAATCG AACAAAAGCC GTTGGCGGAG ACGTCGATCG CGGCTGTCCG

101  GTCOCGCCTT TTCCCGGATC GTGATGTTCC GCGACAGCCT CTCGGATACC
   CAGGGCGGAA AAGGGCCTAG CACTACAAGC CGCTGTCCGA GAGGCTATGG

151  GGCAAAATGT ACAGCAAGAT GCGCGGTTAC CTCCTCCCA GCCCGCCCTA
   CCCTTTTACA TGTCGTTCTA CGCGCCAATG GAGGGGAGGT CGGCGGGAT

201  CTATGAGGCC CGTTTCTCCA ACGGACCCGT CTGGCTGGAG CAGCTGACCA
   GATACTCCCG GCAAAAGAGGT TGCTTGGGCA GACCGACCTC GTCGACTCGT

251  AACAGTTCCC GGGTCTGACC ATCGCCAAAG AAGCGGAAGG CGGTGCCACT
   TTGTCAAGGG CCCAGACTGG TAGCGGTTCC TTCGCCTTCC GCCACGGTGA

301  GCCGTGGGTT ACAACAAGAT CTCCTGGAAT CCCAAGTATC AGGTATCAA
   CGGCACCGAA TGTCTTCTA GAGGACCTTA GGGTTCATAG TCCAGTAGTT

351  CAACCTGGAC TACGAGGTCA CCCAGTCTT GCAGAAAGAC AGCTTCAAGC
   GTTGGACCTG ATGCTCCAGT GGTCAAGAA CGTCTTCTG TCGAAGTTCG

401  CGGACGATCT GGTGATCCTC TGGGTGCGTG CCPATGACTA TCTGGCCCTA
   GCCTGCTAGA CCACTAGGAG ACCCAGCCAC GGTFACTGAT AGACCGGATA

451  GGCTGGAACA CGSAGCAGGA TSCCAAGCGG GTTCGCGATG CCATCAGCGA
   CCGACCTTGT GCCTCGTCTT ACGGTTCCGC CAAGCGCTAC GGTACTCGCT

501  TCGCGCCAAC CGCATGGTAC TGAACGGTGC CAAGCAGATA CTGCTGTCCA
   ACGCCGGTTG GCGTACCATG ACTTGCCACG GTTCTCTAT GACGACAGT

551  ACCTGCCGGA TCTGGGCCAG AACCCGTCAG CTCGCAGTCA GAGGTTGCTC
   TGGACGGCCT AGACCCGGTC TTGGGCAATC GAGGCTCAGT CTTCACCCAG

601  GAGGCGGTCA GCCATGTCTC CECCTATCAC AACCACTGC TGCTGAACCT
   CTCGCCAGT CGGTACAGAG GCGGATAGTG TTGCTCGACG ACGACTTGG

651  GGCAGGCCAG CTGSCCCCA CCGCATGCTT AAGCTGTTT GRATGCGA
   CGGTGCGGTC GACCGGGGT GCGCGTACCA TTTGACAAG CTCTAGCTGT

701  AGCAATTTGC CGAGATGCTG CGTGATCCGC AGAACTTCGG CCFGAGCGAC
   TCGTTAAACG GCTCTACGAC GCACTAGGCG TCTTGAGCC GACTCGCTG

751  GTCGAGAACC CCTGCTACGA CGCGGCTAT GTGTGGAAGC CGTTTCCAC
   CAGCTCTTGG GGACGATGCT GCGGCCGATA CACACCTTCG GCAAACGGTE

801  CCGCAGCGTC AGCACCGACC GCCAGCTCTC CGCCTTCAGT CCGCAGGAA
   GCGCTGCGAG TCGTGGCTGG CCGTCGAGAG GCGGAAATCA GCGCTCCTTG

851  GCCTCGCCAT CGCCGGCAAC CCGCTGCTGG CACAGGCCGT TGCCAGTCTT
   CGGAGCGGTA GCGGCCGTTG GCGCAGCACC GTGTCCGGCA ACGGTCAGGA

901  ATGGCCCGCC GCAGCGCCAG CCCCCTCAAC TGTGAGGGCA AGATGTTCTG
   TACCGGGCGG CGTGCGGCTC GGGGAGTTG ACACCTCCGT TCTACAGAC

951  GGATCAGGTA CACCCGACCA CTGCTGTGCA CGCAGCCCTG AGCGAGCGCG
   CCTAGTCCAT GTGGGCTGGT GACAGCACGT GCCTCGGAC TCGCTCGCGC

1001 CCGCCACCTT CATCGCGAAC CAGTACGAGT TCCTCGCCCA CTGATGA
   GCGGTGGAA GTAGCGCTTG GTCATGCTCA AGGAGCGGCT GACTACT
    
```