

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **3 018 269**

51 Int. Cl.:

**C12N 9/18**

(2006.01)

12

## TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **13.12.2018 PCT/IB2018/060042**

87 Fecha y número de publicación internacional: **20.06.2019 WO19116311**

96 Fecha de presentación y número de la solicitud europea: **13.12.2018 E 18836866 (6)**

97 Fecha y número de publicación de la concesión europea: **29.01.2025 EP 3724324**

54 Título: **Biocatalizadores de carboxiesterasa**

30 Prioridad:

**13.12.2017 US 201762598181 P**

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

**14.05.2025**

73 Titular/es:

**GLAXOSMITHKLINE INTELLECTUAL PROPERTY  
DEVELOPMENT LIMITED (100.00%)  
GSK Medicines Research Centre, Gunnels Wood  
Road  
Stevenage SG1 2NY, GB**

72 Inventor/es:

**BROWN, KRISTIN K.;  
DORR, BRENT M.;  
FUERST, DOUGLAS E.;  
HONICKER, KATHERINE JOYCE;  
JORDAN, LYDIA SANCHEZ;  
MORRISON, JAMES PATRICK;  
PLOTNIKOV, NIKOLAY V.;  
SCHOBER, MARKUS y  
VOLADRI, RAMA**

74 Agente/Representante:

**PEREZ CID, Ana**

### Observaciones:

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

ES 3 018 269 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

Biocatalizadores de carboxiesterasa

5 Campo de la invención

La presente divulgación se refiere a biocatalizadores mejorados de carboxiesterasa y métodos de uso de los biocatalizadores para producir amidas

10 Antecedentes de la invención

La formación de enlaces de amida es una de las reacciones más frecuentes en la síntesis orgánica y comúnmente se encuentran amidas en ingredientes farmacéuticos activos, moléculas biológicamente activas, polímeros, péptidos y proteínas sintéticos. Un estudio realizado por el Novartis Institute for BioMedical Research encontró que la formación de enlaces de amida y productos químicos relacionados de acilación representa el 21,3% de todas las reacciones químicas realizadas en la síntesis de productos farmacéuticos en los últimos 40 años (Schneider, *et al.*, *J. Med Chem.* 59, 4385-4402, 2016). Los métodos tradicionales de síntesis de amidas usan sustratos de ácido carboxílico y amina y requieren reactivos de acoplamiento estequiométrico. A medida que la reacción avanza a través de un compuesto intermedio activado altamente reactivo, pueden presentarse reacciones secundarias indeseables que llevan a la formación de subproductos no deseados, tal como ureas. La economía atómica deficiente y la cantidad significativa de residuos (que contienen metales y con frecuencia tóxicos) generados dan como resultado que sea costoso el proceso de formación de amidas. Otros inconvenientes de los métodos tradicionales de síntesis de amidas incluyen falta de enantioselectividad y quimioselectividad, uso de reactivos de acoplamiento explosivos o tóxicos y el requisito de proteger otros grupos funcionales presentes en los reactivos.

Se han desarrollado planteamientos catalíticos químicos que remueven el requisito de reactivos de acoplamiento estequiométrico y, por lo tanto, mejoran la economía atómica y reducen la cantidad de desechos generados (revisado en Pattabiraman and Bode, *Nature*, 480, 471-479, 2011 y de Figueiredo, *et al.*, *Chemical Reviews*, 12029-12122, 2016). La catálisis con ácido borónico representa el planteamiento más antiguo de la amidación química, en el que la activación transitoria del ácido carboxílico por un boronato de arilo permite la formación catalítica de la amida. Sin embargo, estos métodos padecen de una muy mala tolerancia a los disolventes, además de un alcance limitado del sustrato y la necesidad frecuente de altas temperaturas, lo que limita su uso más extendido. Los estudios más recientes incluyen el uso de amidación catalizada con metales, en la que las sales metálicas sirven como ácidos de Lewis para la activación transitoria de ácidos carboxílicos para apoyar la amidación. Hasta la fecha, estos estudios padecen de muchas de las mismas deficiencias que la catálisis con boronato, lo que requiere altas temperaturas, cargas de catalizador y un alcance limitado de disolventes y tolerancia al sustrato. También se han explorado métodos basados en redox que emplean carbenos N-heterocíclicos (NHC) o catalizadores metálicos, lo que permite la conversión oxidativa de alcoholes, aldehídos, cetonas o nitrilos en sus amidas correspondientes. Desafortunadamente, tanto los catalizadores metálicos como los de NHC son caros, bastante tóxicos, con frecuencia requieren codisolventes peligrosos y, por lo general, sufren una mala tolerancia a los grupos funcionales.

Se han usado lipasas como biocatalizadores para generar enlaces de amida en disolventes orgánicos, al activar directamente y luego al acoplar un material de partida de éster a una amina. Ventajosamente, estas enzimas son habitualmente altamente enantioselectivas y tolerantes térmicamente (para una revisión, ver Gotor, *Bioorg Med Chem*, 7, 2189-2197, 1999). Sin embargo, la mayoría de las lipasas actualmente estudiadas parecen tener una estrecha especificidad de sustrato y, además, se deben usar en disolventes orgánicos secos para evitar la hidrólisis no deseada. Este problema de especificidad es especialmente pronunciado en la síntesis de amidas terciarias, donde se ha demostrado que muy pocas enzimas tienen actividades incluso marginales (estudiado en van Pelt, *Green, Chem.* 13, 1791-1798, 2011).

Para superar estas limitaciones, habitualmente, se emplea un proceso de evolución dirigida en el que las variantes enzimáticas se expresan y estudian de una manera de alto rendimiento. Sin embargo, estas enzimas se derivan con frecuencia de cepas de *Pseudomonas* o *Bacillus*, y no se pueden expresar fácilmente en cepas de laboratorio para las que existen técnicas de manipulación genética robustas, tal como *E. coli* o *S. cerevisiae*. Además, el requisito de disolventes secos y tamices moleculares hace que la evolución dirigida sea extremadamente desafiante, debido tanto al alto contenido de agua de los lisados celulares como al desafío técnico de secar cientos de reacciones en paralelo.

60 Breve descripción de los dibujos

La Figura 1 muestra una alineación de secuencias de la secuencia de polinucleótido que codifica para el codón de *E. coli* optimizado para la enzima de carboxiesterasa tipo silvestre, Esterasa 2 de *A. acidocaldarius* (SEQ ID NO: 1) contra cada una de las secuencias de polinucleótidos que codifica para las carboxiesterasas modificadas

genéticamente a las que se hace referencia a continuación en la Tabla 3 y que se muestran en el Listado de Secuencias posterior (SEQ ID NOs: 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, y 125). Todas estas secuencias de polinucleótidos divulgadas son idénticas entre sí en un 95-99%.

La Figura 2 muestra un alineamiento de secuencias de la secuencia de polipéptido derivada de la enzima de carboxiesterasa tipo silvestre, Esterasa 2 de *A. acidocaldarius* (SEQ ID NO: 2) contra cada una de las secuencias de carboxiesterasa de polipéptido modificadas genéticamente a las que se hace referencia a continuación en la Tabla 3 y que se muestran en el Listado de Secuencias a continuación (SEQ ID NOs: 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, y 126). Todas estas secuencias de polipéptidos divulgadas son idénticas entre sí en un 95-99%.

#### Breve descripción de la invención

La presente invención proporciona un polipéptido de carboxiesterasa que comprende una secuencia de aminoácidos que es al menos 93% o más idéntica a la secuencia de aminoácidos expuesta en SEQ ID NO: 4, o un fragmento funcional del mismo, en donde el residuo correspondiente a X198 en SEQ ID NO: 4 es L, y el polipéptido de carboxiesterasa tiene actividad enzimática mejorada con respecto al polipéptido de carboxiesterasa de SEQ ID NO: 2.

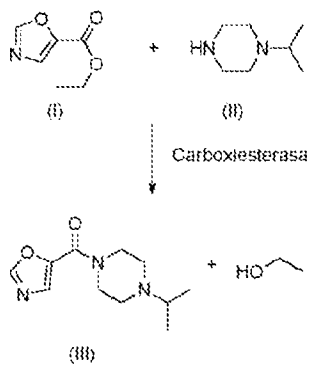
En vista de las limitaciones de la técnica anterior, se han desarrollado una serie de mutantes a partir de la enzima de carboxiesterasa tipo silvestre altamente termotolerante, Esterasa 2 de *A. acidocaldarius* (SEQ ID NO: 2), en donde estas enzimas mutadas poseen una actividad de amidación mejorada en más de 785 000 veces en relación con la enzima tipo silvestre. Debido a esta actividad dramáticamente alterada, estas enzimas mutantes poseen una tolerancia sustancial al agua y a los alcoholes, lo que permite la síntesis directa, a escala, de amidas a partir de precursores simples de ésteres y aminas. Esta estrategia de síntesis directa acorta la síntesis de amidas en 1-2 pasos de química, reduce el uso de disolventes orgánicos y remueve el uso de agentes activadores estequiométricos, lo que representa colectivamente una mejora dramática en el estado de la técnica química.

La presente divulgación proporciona polipéptidos, polinucleótidos que codifican para los polipéptidos y métodos de uso de los polipéptidos, en particular, para la conversión biocatalítica de oxazol-5-carboxilato de etilo (el "sustrato de éster") en (4-isopropilpiperazin-1-il)(oxazol-5-il)metanona (un "producto") en presencia de 1-isopropil-piperazina (un "sustrato de amina"). La presente divulgación también proporciona polipéptidos, polinucleótidos que codifican para los polipéptidos y métodos de uso de los polipéptidos, en particular, para la conversión biocatalítica de sustrato de éster de oxazol-5-carboxilato de etilo (el "sustrato de éster") en ((2S,6R)-2,6-dimetilmorfolino)(oxazol-5-il)metanona (otro producto) en presencia de *cis*-2,6-dimetilmorfolina, un sustrato de amina. Los productos son materiales de partida para productos farmacéuticos que son de interés en el desarrollo para el tratamiento de la enfermedad pulmonar obstructiva crónica (COPD). Específicamente, los productos se pueden usar para sintetizar inhibidores de fosfoinosítido 3-cinasa  $\delta$  (inhibidores de PI3K $\delta$ ), que son una clase de fármacos que se usan para tratar la inflamación, las enfermedades autoinmunes y el cáncer. Las composiciones de la invención también se pueden usar como materiales de partida para otros tipos de productos farmacéuticos.

En tanto que el polipéptido tipo silvestre, enzima de carboxiesterasa, Esterasa 2 de *A. acidocaldarius* (SEQ ID NO: 2) solo actúa sobre el sustrato de éster con una eficiencia muy baja (<1% del sustrato convertido en producto), como se evidencia en la Tabla 3, las carboxiesterasas modificadas genéticamente (E.C.3.1.1) de la presente divulgación son capaces de llevar a cabo una fácil conversión del sustrato de éster en los productos en presencia de un sustrato de amina. Por lo tanto, en un aspecto, la presente divulgación se refiere a carboxiesterasas mejoradas capaces de convertir oxazol-5-carboxilato de etilo, el sustrato de éster, en (4-isopropilpiperazin-1-il)(oxazol-5-il)metanona, un producto, en presencia de 1-isopropil-piperazina, el "sustrato de amina", a niveles medibles por aproximadamente 0,1% de conversión mediante una técnica de análisis, tal como absorbancia HPLC-UV.

#### Descripción detallada de la invención

La presente divulgación proporciona biocatalizadores altamente eficientes capaces de mediar transformaciones que implican la amidación de determinados aceptores del grupo amida, por ejemplo, la síntesis del compuesto de la fórmula III. Los biocatalizadores son polipéptidos amidados modificados genéticamente que pueden convertir el sustrato de la fórmula I, oxazol-5-carboxilato de etilo (el "sustrato de éster"), en el producto de la fórmula III, (4-isopropilpiperazin-1-il)(oxazol-5-il)metanona, (un "producto") en presencia de un sustrato de amina de la fórmula II (1-isopropil-piperazina), de la siguiente manera:



En determinadas realizaciones, las carboxiesterasas modificadas genéticamente se derivan de la carboxiesterasa de origen natural de la Esterasa 2 de *A. acidocaldarius*, que es una enzima de carboxiesterasa que cataliza la hidrólisis de un éster a través de la formación y resolución de una amina intermedia de acil-enzima. La carboxiesterasa de SEQ ID NO: 4 difiere de la enzima de origen natural derivada de la carboxiesterasa tipo silvestre, Esterasa 2 de *A. acidocaldarius* (SEQ ID NO: 2) al tener una sustitución de glutamato (E) en la posición del residuo X198 con leucina (L) y tiene actividad medible para el sustrato de éster, oxazol-5-carboxilato de etilo (fórmula I). La carboxiesterasa de SEQ ID NO: 4 se ha modificado genéticamente para mediar la conversión eficiente del sustrato de éster de la fórmula I. al producto de la fórmula III. en presencia de un sustrato de amina, tal como 1-isopropil-piperazina (fórmula II). La conversión se puede llevar a cabo bajo condiciones suaves (30°C con un alto % de conversión), lo que hace que el proceso sea aplicable a la producción en gran volumen de amidas de la fórmula III y fórmula V.

#### Abreviaturas y definiciones

A los efectos de las descripciones en la presente, las abreviaturas usadas para los aminoácidos codificados genéticamente son convencionales y son las siguientes en la Tabla 1:

**Tabla 1**

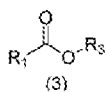
| Aminoácido   | Tres letras | Una letra |
|--------------|-------------|-----------|
| alanina      | ALA         | A         |
| arginina     | ARG         | R         |
| asparagina   | ASN         | N         |
| aspartato    | ASP         | D         |
| cisteína     | CYS         | C         |
| glutamato    | GLU         | E         |
| glutamina    | GLN         | Q         |
| glicina      | GLY         | G         |
| histidina    | HIS         | H         |
| isoleucina   | ILE         | I         |
| leucina      | LEU         | L         |
| lisina       | LYS         | K         |
| metionina    | MET         | M         |
| fenilalanina | PHE         | F         |
| prolina      | PRO         | P         |
| serina       | SER         | S         |
| treonina     | THR         | T         |
| triptófano   | TRP         | W         |
| tirosina     | TYR         | Y         |
| valina       | VAL         | V         |

Cuando se usan las abreviaturas de tres letras, a menos que estén precedidas específicamente por una "L" o una "D" o sean claras a partir del contexto en el que se usa la abreviatura, el aminoácido puede estar en la configuración L o D sobre el carbono α (Cα). Por ejemplo, mientras que "Ala" designa alanina sin especificar la configuración sobre el carbono α, "D-Ala" y "L-Ala" designan D-alanina y L-alanina, respectivamente. Cuando se usan las abreviaturas de una letra, las letras mayúsculas designan aminoácidos en la configuración L sobre el carbono α y las letras minúsculas designan aminoácidos en la configuración D sobre el carbono α. Por ejemplo, "A" designa L-alanina y "a" designa D-alanina. Cuando las secuencias de péptidos se presentan como una cadena de

abreviaturas de una o tres letras (o mezclas de las mismas), las secuencias se presentan en la dirección N→C de acuerdo con la convención.

Los términos técnicos y científicos usados en las descripciones en la presente tendrán los significados comúnmente entendidos por un experto en la técnica, a menos que se defina específicamente lo contrario. Por consiguiente, se pretende que los siguientes términos tengan los siguientes significados. Todas las patentes de los Estados Unidos y las solicitudes de patente de los Estados Unidos publicadas, incluidas todas las secuencias divulgadas dentro de estas patentes y solicitudes de patente, a las que se hace referencia en la presente se incorporan expresamente mediante esta referencia.

"Producto secundario ácido" o "producto secundario de hidrólisis" se refiere al ácido carboxílico resultante de la reacción de un sustrato de éster con agua como resultado de una enzima de carboxiesterasa. Los subproductos ácidos son moléculas de la fórmula general (3) en la que R<sub>3</sub> es -H. R<sub>1</sub> se describe anteriormente.



"Aminoácido o residuo ácido" se refiere a un aminoácido o residuo hidrófilo que tiene una cadena lateral que exhibe un valor pK de menos de aproximadamente 6 cuando el aminoácido se incluye en un péptido o polipéptido. Los aminoácidos ácidos habitualmente tienen cadenas laterales cargadas negativamente a pH fisiológico debido a la pérdida de un ion hidrógeno. Genéticamente, los aminoácidos ácidos codificados incluyen L-Glu (E) y L-Asp (D).

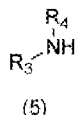
"Alquilo" pretende incluir grupos alquilo de la longitud designada en una configuración recta o ramificada. Los grupos alquilo de ejemplo son metilo, etilo, propilo, isopropilo, butilo, sec-butilo, terc-butilo, pentilo, isopentilo, hexilo, isohexilo y similares. Los grupos alquilo están insustituídos o sustituidos con uno a tres grupos elegidos independientemente de: halógeno, hidroxilo, carboxi, aminocarbonilo, amino, alcoxi de C<sub>1-4</sub> y alquiltio de C<sub>1-4</sub>.

"Amida" pretende significar un grupo funcional que contiene un grupo carbonilo unido a un átomo de nitrógeno. Una amida también se refiere a cualquier compuesto que contiene el grupo funcional amida. Las amidas se derivan de un ácido carboxílico y una amina.

"Amidar o "Amidación" pretende significar la formación de un grupo funcional amida a partir de un compuesto que contiene carbonilo, habitualmente resultante de la reacción de una funcionalidad de ácido carboxílico y amina, pero también formado aquí a partir de una funcionalidad éster y amina.

"Amina" pretende significar un grupo funcional que contiene un átomo de nitrógeno hibridado sp<sup>3</sup>. Una amina también se refiere a cualquier compuesto que contiene el grupo funcional amina.

"Sustrato de amina" se refiere a un compuesto amino que es capaz de desplazar la cadena lateral de alcohol de un sustrato de éster bajo la acción de una carboxiesterasa, generando de esta forma un producto de amida. Los sustratos de amina son moléculas de la fórmula general (5), en las que cada uno de R<sub>3</sub> y R<sub>4</sub>, cuando se toman independientemente, es un grupo alquilo o arilo que está insustituído o sustituido con al menos un grupo enzimáticamente no inhibidor. Los grupos R<sub>3</sub> y R<sub>4</sub>, cuando se toman conjuntamente, pueden formar un anillo que no está sustituido, sustituido o fusionado a otros anillos. Los sustratos de amina habituales que se pueden usar con la invención incluyen, pero no se limitan a, porciones de piperazinilo o morfolino cíclicos, así como aminas primarias o aromáticas. En el contexto de la presente divulgación, un sustrato de amina incluye, entre otros, los compuestos de la fórmula II, 1-isopropil-piperazina, y de la fórmula IV, *cis*-2,6-dimetilmorfolina.



"Aminoácido" o "residuo" como se usa en el contexto de los polipéptidos divulgados en la presente se refiere al monómero específico en una posición de secuencia (por ejemplo, P5 indica que el "aminoácido" o "residuo" en la posición 5 de SEQ ID NO: 2 es una prolina).

"Diferencia de aminoácidos" o "diferencia de residuos" se refiere a un cambio en el residuo en una posición específica de una secuencia de polipéptidos en comparación con una secuencia de referencia. La posición de la secuencia de polipéptido en la que está presente un aminoácido o cambio de aminoácido particular ("diferencia de residuo") a veces se describe en la presente como "X<sub>n</sub>" o "posición n", donde n se refiere a la posición del residuo

con respecto a la secuencia de referencia.

Por ejemplo, una diferencia de residuo en la posición X8, donde la secuencia de referencia tiene una serina, se refiere a un cambio del residuo en la posición X8 a cualquier residuo que no sea serina. Como se divulga en la presente, una enzima puede incluir una o más diferencias de residuos con respecto a una secuencia de referencia, donde las múltiples diferencias de residuos habitualmente se indican mediante una lista de las posiciones especificadas donde se realizan cambios con respecto a la secuencia de referencia (por ejemplo, "una o más diferencias de residuos en comparación con SEQ ID NO: 4 en las siguientes posiciones de residuos: X27, X30, X35, X37, X57, X75, X103, X185, X207, X208, X271, X286, o X296").

Una mutación de sustitución específica, que es un reemplazo del residuo específico en una secuencia de referencia con un residuo especificado diferente puede indicarse mediante la notación convencional "X(número)Y", donde X es el identificador de una letra individual del residuo en la secuencia de referencia, "número" es la posición del residuo en la secuencia de referencia, e Y es el identificador de una letra individual de la sustitución del residuo en la secuencia modificada genéticamente.

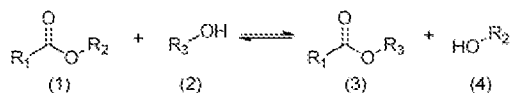
"Aminoácido o residuo alifático" se refiere a un aminoácido o residuo hidrófobo que tiene una cadena lateral de hidrocarburo alifático. Los aminoácidos alifáticos codificados genéticamente incluyen L-Ala (A), L-Val (V), L-Leu (L) y L-Ile (I).

"Aminoácido o residuo aromático" se refiere a un aminoácido o residuo hidrófilo o hidrófobo que tiene una cadena lateral que incluye al menos un anillo aromático o heteroaromático. Los aminoácidos aromáticos codificados genéticamente incluyen L-Phe (F), L-Tyr (Y) y L-Trp (W). Aunque debido al pKa de su átomo de nitrógeno heteroaromático L-His (H) a veces se clasifica como un residuo básico, o como un residuo aromático ya que su cadena lateral incluye un anillo heteroaromático, en la presente la histidina se clasifica como un residuo hidrófilo o como un "residuo restringido" (ver posteriormente).

"Arolo" pretende significar un grupo aromático, que incluye fenilo y naftilo. "Arolo" está insustituido o sustituido con uno a cinco sustituyentes seleccionados independientemente de fluoro, hidroxilo, trifluorometilo, amino, alquilo de C<sub>1-4</sub> y alcoxi de C<sub>1-4</sub>.

"Aminoácido o residuo básico" se refiere a un aminoácido o residuo hidrófilo que tiene una cadena lateral que exhibe un valor pKa mayor que aproximadamente 6 cuando el aminoácido se incluye en un péptido o polipéptido. Los aminoácidos básicos habitualmente tienen cadenas laterales cargadas positivamente a pH fisiológico debido a la asociación con el ion hidronio. Los aminoácidos básicos codificados genéticamente incluyen L-Arg (R) y L-Lys (K).

"Carboxiesterasa" se usa para referirse a un polipéptido que tiene una capacidad enzimática de interconvertir la cadena lateral de un sustrato de éster (1) con la de un compuesto donante (2), convirtiendo el sustrato de éster (1) en su correspondiente producto de éster (3) y la forma de alcohol libre de la cadena lateral de éster (4).



"Optimizado por codones" se refiere a cambios en los codones del polinucleótido que codifica para una proteína con respecto a los usados preferentemente en un organismo particular de modo que la proteína codificada se exprese de manera eficiente en el organismo de interés. Aunque el código genético es degenerado en el sentido de que la mayoría de los aminoácidos están representados por varios codones, llamados "sinónimos" o codones "sinónimos", es bien sabido que el uso de codones por organismos particulares no es aleatorio y está sesgado hacia tripletes de codones particulares. Este sesgo de uso de codones puede ser mayor en referencia a un gen dado, genes de función común u origen ancestral, proteínas altamente expresadas frente a proteínas de bajo número de copias y las regiones codificantes de proteínas agregadas del genoma de un organismo. En algunas realizaciones, los polinucleótidos que codifican para las enzimas carboxiesterasa pueden tener codones optimizados para una producción óptima a partir del organismo hospedador seleccionado para la expresión.

"Ventana de comparación" se refiere a un segmento conceptual de al menos aproximadamente 20 posiciones de nucleótidos o residuos de aminoácidos contiguos en donde una secuencia se puede comparar con una secuencia de referencia de al menos 20 nucleótidos o aminoácidos contiguos y en donde la porción de la secuencia en la ventana de comparación puede comprender adiciones o deleciones (es decir, espacios) de 20 por ciento o menos en comparación con la secuencia de referencia (que no comprende adiciones o deleciones) para la alineación óptima de las dos secuencias. La ventana de comparación puede ser más larga que 20 residuos contiguos e incluye, opcionalmente, 30, 40, 50, 100 o ventanas más largas.

Las sustituciones o mutaciones de aminoácidos "conservadoras" se refieren a la intercambiabilidad de residuos que tienen cadenas laterales similares y, por lo tanto, en general implican la sustitución del aminoácido en el polipéptido con aminoácidos dentro de la misma clase de aminoácidos definida o similar. Sin embargo, tal como se usa en la presente, en algunas realizaciones, las mutaciones conservadoras no incluyen sustituciones de un residuo hidrófilo a hidrófilo, hidrófobo a hidrófobo, que contiene hidroxilo a que contiene hidroxilo, o pequeño a pequeño, si la mutación conservadora puede ser en cambio una sustitución de un residuo alifático a uno alifático, no polar a no polar, polar a polar, ácido a ácido, básico a básico, aromático a aromático o restringido a restringido. Además, tal como se usa en la presente, A, V, L o I pueden mutarse de forma conservadora a otro residuo alifático o a otro residuo no polar. La Tabla 2 a continuación muestra ejemplos de sustituciones conservadoras.

Tabla 2

| Residuo    | Posibles mutaciones conservadoras                                    |
|------------|----------------------------------------------------------------------|
| A, L, V, I | Otros alifáticos (A, L, V, I)<br>Otros no polares (A, L, V, I, G, M) |
| G, M       | Otros no polares (A, L, V, I, G, M)                                  |
| D, E       | Otros ácidos (D, E)                                                  |
| K, R       | Otros básicos (K, R)                                                 |
| P          | ninguna                                                              |
| N, Q, S, T | Otro polar                                                           |
| H, Y, W, F | Otros aromáticos (H, Y, W, F)                                        |
| C          | Ninguna                                                              |

"Aminoácido o residuo restringido" se refiere a un aminoácido o residuo que tiene una geometría restringida. En la presente, los residuos restringidos incluyen L-Pro (P) y L-His (H). La histidina tiene una geometría restringida porque tiene un anillo de imidazol relativamente pequeño. La prolina tiene una geometría restringida, porque también tiene un anillo de cinco miembros.

"Secuencia de control" se define en la presente para incluir todos los componentes, que son necesarios o ventajosos para la expresión de un polinucleótido y/o polipéptido de la presente divulgación. Cada secuencia de control puede ser nativa o extraña al polinucleótido de interés. Estas secuencias de control incluyen, pero no se limitan a, un líder, secuencia de poliadenilación, secuencia de propéptido, promotor, secuencia de péptido señal y terminador de la transcripción.

"Conversión" se refiere a la transformación enzimática de un sustrato en el producto correspondiente.

"Correspondiente a", "referencia a" o "relativo a" cuando se usa en el contexto de la numeración de una secuencia de aminoácidos o polinucleótidos dada se refiere a la numeración de los residuos de una secuencia de referencia especificada cuando la secuencia de aminoácidos o polinucleótidos dada se compara con la secuencia de referencia. En otras palabras, el número de residuo o la posición de residuo de un polímero dado se designa con respecto a la secuencia de referencia en lugar de por la posición numérica real del residuo dentro de la secuencia de aminoácidos o polinucleótidos dada de la secuencia no de referencia. Por ejemplo, una secuencia de aminoácidos dada, tal como la de una carboxiesterasa modificada, se puede alinear con una secuencia de referencia mediante la introducción de espacios para optimizar las coincidencias de residuos entre las dos secuencias. En estos casos, aunque los espacios están presentes, la numeración del residuo en la secuencia de aminoácidos o polinucleótidos dada se realiza con respecto a la secuencia de referencia con la que se ha alineado.

"Cisteína" o L-Cys (C) es inusual ya que puede formar puentes disulfuro con otros aminoácidos L-Cys (C) u otros aminoácidos que contienen sulfanilo o sulfhidrilo. Los "residuos similares a cisteína" incluyen cisteína y otros aminoácidos que contienen porciones sulfhidrilo que están disponibles para la formación de puentes disulfuro. La capacidad de L-Cys (C) (y otros aminoácidos con cadenas laterales que contienen -SH) para existir en un péptido en la forma de -SH libre reducido o puenteado con disulfuro oxidado afecta si L-Cys (C) contribuye con un carácter hidrófobo o hidrófilo neto a un péptido. Si bien L-Cys (C) exhibe una hidrofobicidad de 0,29 de acuerdo con la escala de consenso normalizada de Eisenberg (Eisenberg *et al.*, 1984, *supra*), debe entenderse que para los fines de la presente divulgación, L-Cys (C) se clasifica en su propio grupo único.

"Deleción" se refiere a la modificación del polipéptido mediante la remoción de uno o más aminoácidos del polipéptido de referencia. Las deleciones pueden comprender la remoción de 1 o más aminoácidos, 2 o más aminoácidos, 5 o más aminoácidos, 10 o más aminoácidos, 15 o más aminoácidos, o 20 o más aminoácidos, hasta el 10% de la cantidad total de aminoácidos, hasta el 20% de la cantidad total de aminoácidos, o hasta el 30% de la cantidad total de aminoácidos que componen el polipéptido mientras se retiene la actividad enzimática y/o se retienen las propiedades mejoradas de una enzima de carboxiesterasa modificada. Las deleciones se pueden dirigir a las porciones internas y/o porciones terminales del polipéptido. En diversas realizaciones, la eliminación

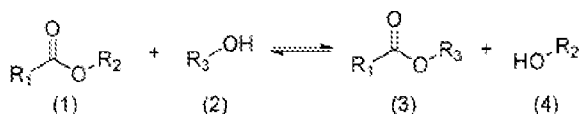
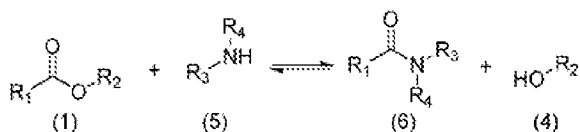
puede comprender un segmento continuo o puede ser discontinua.

"Derivado de", tal como se usa en la presente en el contexto de las enzimas modificadas genéticamente, identifica la enzima de origen y/o el gen que codifica para esta enzima, en el que se basó la modificación. Por ejemplo, la enzima de carboxiesterasa modificada de SEQ ID NO: 4 se obtuvo mutando la carboxiesterasa de SEQ ID NO: 2. Por lo tanto, esta enzima de carboxiesterasa modificada de SEQ ID NO: 4 es "derivado de" el polipéptido de SEQ ID NO: 2.

Una "carboxiesterasa modificada", como se usa en la presente, se refiere a una proteína de tipo carboxiesterasa que se ha modificado sistemáticamente, a través de la inserción de nuevos aminoácidos en su secuencia de referencia, la delección de aminoácidos presentes en su secuencia de referencia o la mutación de aminoácidos en su secuencia de referencia en aminoácidos alternativos, ya sea a través de un proceso de mutagénesis aleatoria seguido de la selección de mutantes que tienen una propiedad particular o a través de la introducción intencional de cambios de aminoácidos particulares en la secuencia de la proteína.

"Éster" pretende querer decir un grupo funcional que contiene un grupo carbonilo unido a un átomo de oxígeno que a su vez está unido a un átomo de carbono. Un éster también se refiere a cualquier compuesto que contenga el grupo funcional éster. Los ésteres se derivan de un ácido carboxílico y un alcohol.

Un "sustrato de éster" se refiere específicamente a los compuestos de la fórmula (1) que contienen un éster, que reacciona con una carboxiesterasa modificada. En el contacto de la presente divulgación, un sustrato de éster incluye, entre otros, el compuesto de la fórmula I, oxazol-5-carboxilato de etilo.



"Fragmento", tal como se usa en la presente, se refiere a un polipéptido que tiene una delección amino-terminal y/o carboxi-terminal, pero donde la secuencia de aminoácidos restante es idéntica a las posiciones correspondientes en la secuencia. Los fragmentos pueden tener al menos 14 aminoácidos de longitud, al menos 20 aminoácidos de longitud, al menos 50 aminoácidos de longitud o más, y hasta 70%, 80%, 90%, 95%, 98% y 99%, o más, del polipéptido de carboxiesterasa de longitud completa, por ejemplo, el polipéptido de SEQ ID NO: 4.

Un "fragmento funcional" o un "fragmento biológicamente activo", usado indistintamente, en la presente se refiere a un polipéptido que tiene una o más delecciones amino-terminales y/o carboxi-terminales y/o delecciones internas, pero donde la secuencia de aminoácidos restante es idéntica a las posiciones correspondientes en la secuencia con la que se está comparando (por ejemplo, una enzima de *T. fusca* modificada genéticamente de longitud completa de la presente invención) y que retiene sustancialmente toda la actividad del polipéptido de longitud completa.

"Halógeno" pretende incluir los átomos de halógeno, flúor, cloro, bromo y yodo.

El polinucleótido "heterólogo" se refiere a cualquier polinucleótido que se introduce en una célula hospedadora mediante técnicas de laboratorio, e incluye polinucleótidos que se remueven de una célula hospedadora, se someten a manipulación de laboratorio y luego se reintroducen en una célula hospedadora.

"Rigurosidad de hibridación" se refiere a condiciones de hibridación, tal como condiciones de lavado, en la hibridación de ácidos nucleicos. En general, las reacciones de hibridación se realizan bajo condiciones de menor rigurosidad, seguidas de lavados de rigurosidad variable pero mayor. El término "hibridación moderadamente rigurosa" se refiere a condiciones que permiten que el ADN diana se una a un ácido nucleico complementario que tiene aproximadamente 60% de identidad, preferentemente, aproximadamente 75% de identidad, aproximadamente 85% de identidad con el ADN diana, o con más de aproximadamente 90% de identidad con el polinucleótido diana. Las condiciones moderadamente rigurosas de ejemplo son condiciones equivalentes a la hibridación en 50% de formamida, solución 5x Denhart, solución salina 5x-fosfato de sodio-EDTA (SSPE), 0,2% de dodecilsulfato de sodio (SDS) a 42°C, seguido de lavado en 0,2x SSPE, 0,2% de SDS, a 42°C. "Hibridación de

alta rigurosidad" se refiere en general a condiciones que son de aproximadamente 10°C o menos de la temperatura de fusión térmica  $T_m$  como se determina bajo la condición de solución para una secuencia de polinucleótidos definida. En algunas realizaciones, una condición de alta rigurosidad se refiere a condiciones que permiten la hibridación de solo aquellas secuencias de ácido nucleico que forman híbridos estables en NaCl 0,018 M a 65°C (es decir, si un híbrido no es estable en NaCl 0,018 M a 65°C, no será estable bajo condiciones de alta rigurosidad, como se contempla en la presente). Se pueden proporcionar condiciones de alta rigurosidad, por ejemplo, mediante hibridación en condiciones equivalentes a 50% de formamida, solución 5x Denhart, 5xSSPE, 0,2% de SDS a 42°C, seguido de lavado en 0,1 x SSPE y 0,1 % de SDS a 65°C. Otra condición de alta rigurosidad es la hibridación en condiciones equivalentes a la hibridación en 5X SSC que contiene 0,1% (p:v) SDS a 65°C y el lavado en 0,1x SSC que contiene 0,1% de SDS a 65°C. Otras condiciones de hibridación de alta rigurosidad, así como condiciones moderadamente rigurosas, se describen en las referencias citadas anteriormente.

"Aminoácido o residuo hidrófilo" se refiere a un aminoácido o residuo que tiene una cadena lateral que exhibe una hidrofobicidad inferior a cero de acuerdo con la escala de hidrofobicidad de consenso normalizada de Eisenberg *et al.*, 1984, *J. Mol. Biol.* 179:125-142. Los aminoácidos hidrófilos codificados genéticamente incluyen L-Thr (T), L-Ser (S), L-His (H), L-Glu (E), L-Asn (N), L-Gln (Q), L-Asp (D), L-Lys (K) y L-Arg (R). "Aminoácido o residuo hidrófobo" se refiere a un aminoácido o residuo que tiene una cadena lateral que exhibe una hidrofobicidad mayor que cero de acuerdo con la escala de hidrofobicidad de consenso normalizada de Eisenberg *et al.*, 1984, *J. Mol. Biol.* 179:125-142. Genéticamente, los aminoácidos hidrófobos codificados incluyen L-Pro (P), L-Ile (I), L-Phe (F), L-Val (V), L-Leu (L), L-Trp (W), L-Met (M), L-Ala (A) y L-Tyr (Y).

"Aminoácido o residuo que contiene hidroxilo" se refiere a un aminoácido que contiene una porción hidroxilo (OH). Los aminoácidos que contienen hidroxilo codificados genéticamente incluyen L-Ser (S) L-Thr (T) y L-Tyr (Y).

"Propiedad enzimática mejorada" se refiere a cualquier propiedad enzimática hecha mejor o más deseable para un propósito particular en comparación con la propiedad encontrada en una enzima de referencia. Para los polipéptidos de carboxiesterasa modificados genéticamente descritos en la presente, la comparación se hace en general con la enzima de carboxiesterasa tipo silvestre, aunque en algunas realizaciones, la carboxiesterasa de referencia puede ser otra carboxiesterasa modificada mejorada. Las propiedades enzimáticas para las que se puede realizar una mejora incluyen, pero no se limitan a, actividad enzimática (que se puede expresar en términos de porcentaje de conversión del sustrato en un período de tiempo), estabilidad térmica, estabilidad al disolvente, perfil de actividad de pH, requisitos de coenzima, refractariedad a los inhibidores (por ejemplo, inhibición del producto), estereoespecificidad y supresión de la producción de productos secundarios ácidos.

"Actividad enzimática aumentada" o "actividad aumentada" se refiere a una propiedad mejorada de una enzima modificada genéticamente, que puede representarse por un aumento en la actividad específica (por ejemplo, producto producido/tiempo/peso de proteína) o un aumento en el porcentaje de conversión del sustrato en el producto (por ejemplo, porcentaje de conversión de la cantidad de partida de sustrato en producto en un período de tiempo específico usando una cantidad específica de carboxiesterasa), en comparación con una enzima de referencia. Los métodos de ejemplo para determinar la actividad enzimática se proporcionan en los Ejemplos. Cualquier propiedad relacionada con la actividad enzimática puede verse afectada, incluidas las propiedades enzimáticas clásicas de  $K_m$ ,  $V_{max}$  o  $k_{cat}$ , cambios que pueden conducir a una mayor actividad enzimática. Las mejoras en la actividad enzimática pueden ser de aproximadamente 1,5 veces la actividad enzimática de la enzima tipo silvestre o modificada genéticamente correspondiente, hasta 2 veces, 5 veces, 10 veces, 20 veces, 25 veces, 50 veces, 75 veces, 100 veces, 1000 veces, 10 000 veces, 100 000 veces, 500 000 veces, 785 000 veces o más actividad enzimática que la enzima de origen natural (por ejemplo, una carboxiesterasa) u otra enzima modificada genéticamente de la cual se derivaron las enzimas que exhiben una mayor actividad. En realizaciones específicas, las enzimas carboxiesterasa modificadas genéticamente de la presente divulgación exhiben una actividad enzimática mejorada en el intervalo de 1,5 a 50 veces, 1,5 a 100 veces o más que la de la enzima de carboxiesterasa original (es decir, la carboxiesterasa tipo silvestre o modificada genéticamente de la que se derivaron). El experto en la técnica entiende que la actividad de cualquier enzima está limitada por la difusión, de modo que la velocidad de recambio catalítico no puede exceder la velocidad de difusión del sustrato, incluidas las coenzimas requeridas. El máximo teórico del límite de difusión es en general de aproximadamente  $10^8$  a  $10^9$  ( $M^{-1} S^{-1}$ ). Por lo tanto, cualquier mejora en la actividad enzimática de la carboxiesterasa tendrá un límite superior relacionado con la velocidad de difusión de los sustratos sobre los que actúa la enzima de carboxiesterasa. La actividad de carboxiesterasa se puede medir mediante cualquiera de los ensayos estándar usados para medir carboxiesterasas, tal como el cambio en la concentración de sustrato o producto, o el cambio en la concentración del sustrato de amina. Las comparaciones de las actividades enzimáticas se realizan usando una preparación definida de enzima, un ensayo definido bajo una condición establecida y uno o más sustratos definidos, como se describe con más detalle en la presente. En general, cuando se comparan las enzimas en los lisados celulares, se determina el número de células y la cantidad de proteína analizada, así como el uso de sistemas de expresión idénticos y células huésped idénticas para minimizar las variaciones en la cantidad de enzima producida por las células huésped y presente en los lisados.

"Inserción" se refiere a la modificación del polipéptido mediante la adición de uno o más aminoácidos al polipéptido de referencia. En algunas realizaciones, las enzimas carboxiesterasa modificadas genéticamente mejoradas comprenden inserciones de uno o más aminoácidos en el polipéptido de carboxiesterasa de origen natural, así como inserciones de uno o más aminoácidos en otros polipéptidos de carboxiesterasa mejorados. Las inserciones pueden ser en las porciones internas del polipéptido, o en el extremo carboxi o amino. Las inserciones como se usan en la presente incluyen proteínas de fusión como se conoce en la técnica. La inserción puede ser un segmento contiguo de aminoácidos o estar separada por uno o más de los aminoácidos en el polipéptido de origen natural.

"Polipéptido aislado" se refiere a un polipéptido que está sustancialmente separado de otros contaminantes que lo acompañan naturalmente, por ejemplo, proteínas, lípidos y polinucleótidos. El término abarca polipéptidos que se han removido o purificado de su entorno o sistema de expresión de origen natural (por ejemplo, célula hospedadora o síntesis *in vitro*). Las enzimas carboxiesterasa mejoradas pueden estar presentes dentro de una célula, presentes en el medio celular o preparadas en diversas formas, tal como lisados o preparaciones aisladas. Como tal, en algunas realizaciones, la enzima de carboxiesterasa mejorada puede ser un polipéptido aislado.

"Sustitución no conservativa" se refiere a la sustitución o mutación de un aminoácido en el polipéptido con un aminoácido con propiedades de cadena lateral significativamente diferentes. Las sustituciones no conservativas pueden usar aminoácidos entre, en lugar de dentro de, los grupos definidos listados anteriormente. En una realización, una mutación no conservadora afecta: (a) la estructura de la estructura principal del péptido en el área de la sustitución (por ejemplo, prolina por glicina); (b) la carga o hidrofobicidad; o (c) el volumen de la cadena lateral.

"Aminoácido no polar" o "Residuo no polar" se refiere a un aminoácido o residuo hidrófobo que tiene una cadena lateral que no está cargada a pH fisiológico y que tiene enlaces en los que el par de electrones compartidos en común por dos átomos en general se mantiene por igual por cada uno de los dos átomos (es decir, la cadena lateral no es polar). Los aminoácidos no polares codificados genéticamente incluyen L-Gly (G), L-Leu (L), L-Val (V), L-Ile (I), L-Met (M) y L-Ala (A).

"Porcentaje de conversión" se refiere al porcentaje del sustrato que se convierte en el producto dentro de un período de tiempo bajo condiciones especificadas. Por lo tanto, por ejemplo, la "actividad enzimática" o "actividad" de un polipéptido de carboxiesterasa se puede expresar como "porcentaje de conversión" del sustrato al producto.

"Porcentaje de identidad de secuencia", "por ciento de identidad" y "por ciento idéntico" se usan en la presente para referirse a comparaciones entre secuencias de polinucleótidos o secuencias de polipéptidos, y se determinan comparando dos secuencias alineadas óptimamente en una ventana de comparación, en donde la porción del polinucleótido o secuencia de polipéptidos en la ventana de comparación puede comprender adiciones o deleciones (es decir, espacios) en comparación con la secuencia de referencia para la alineación óptima de las dos secuencias. El porcentaje se calcula determinando la cantidad de posiciones en las que se produce la base de ácido nucleico o el residuo de aminoácido idénticos en ambas secuencias o una base de ácido nucleico o un residuo de aminoácido se alinea con un espacio para proporcionar la cantidad de posiciones coincidentes, al dividir la cantidad de posiciones coincidentes por la cantidad total de posiciones en la ventana de comparación y multiplicando el resultado por 100 para proporcionar el porcentaje de identidad de secuencia. La determinación de la alineación óptima y el porcentaje de identidad de secuencia se realiza usando los algoritmos BLAST y BLAST 2.0 (ver, por ejemplo, Altschul, *et al.*, 1990, *J. Mol. Biol.* 215:403-410 y Altschul, *et al.*, 1977, *Nucleic Acids Res.* 3389-3402). El software para realizar análisis BLAST está disponible públicamente a través del sitio web del Centro Nacional de Información Biotecnológica.

De forma breve, los análisis BLAST implican primero identificar pares de secuencias de alta puntuación (HSP) mediante la identificación de palabras cortas de longitud W en la secuencia de consulta, que coinciden o satisfacen alguna puntuación umbral T de valor positivo cuando se alinean con una palabra de la misma longitud en una secuencia de base de datos. T se conoce como el umbral de puntuación de la palabra del vecindario (Altschul, *et al.*, supra). Estos resultados iniciales de palabras del vecindario actúan como semillas para iniciar búsquedas para encontrar HSP más largos que los contengan. Los aciertos de palabras se extienden en ambas direcciones a lo largo de cada secuencia hasta donde se pueda aumentar la puntuación de alineación acumulativa. Las puntuaciones acumulativas se calculan usando, para las secuencias de nucleótidos, los parámetros M (puntuación de recompensa para un par de residuos coincidentes; siempre >0) y N (puntuación de penalización para residuos no coincidentes; siempre <0). Para las secuencias de aminoácidos, se usa una matriz de puntuación para calcular la puntuación acumulativa. La extensión de los golpes de palabra en cada dirección se detiene cuando: la puntuación de alineación acumulativa cae en la cantidad X desde su valor máximo alcanzado; la puntuación acumulativa va a cero o menos, debido a la acumulación de una o más alineaciones de residuos de puntuación negativa; o se alcanza el final de cualquiera de las secuencias. Los parámetros del algoritmo BLAST W, T y X determinan la sensibilidad y la velocidad de la alineación. El programa BLASTN (para secuencias de nucleótidos) usa por defecto una longitud de palabra (W) de 11, una expectativa (E) de 10, M=5, N=-4 y una comparación de ambas cadenas. Para las secuencias de aminoácidos, el programa BLASTP usa por defecto una longitud de

palabra (W) de 3, una expectativa (E) de 10 y la matriz de puntuación BLOSUM62 (ver Henikoff y Henikoff, 1989, *Proc Natl Acad Sci USA* 89:10915).

- 5 Hay muchos otros algoritmos disponibles que funcionan de manera similar a BLAST para proporcionar un porcentaje de identidad para dos secuencias. La alineación óptima de las secuencias para la comparación se puede realizar, por ejemplo, mediante el algoritmo de homología local de Smith y Waterman, 1981, *Adv. Appl. Math.* 2:482, mediante el algoritmo de alineación de homología de Needleman y Wunsch, 1970, *J. Mol. Biol.* 48:443, mediante el método de búsqueda de similitud de Pearson y Lipman, 1988, *Proc. Natl. Acad. Sci. USA* 85:2444, mediante implementaciones computarizadas de estos algoritmos (GAP, BESTFIT, FASTA y TFASTA en el paquete de software GCG Wisconsin), o mediante inspección visual (ver en general, *Current Protocols in Molecular Biology*, F. M. Ausubel, *et al.*, eds., Current Protocols, una alianza conjunta entre Greene Publishing Associates, Inc. y John Wiley & Sons, Inc., (1995 Supplement) (Ausubel)). Además, la determinación del alineamiento de secuencias y el porcentaje de identidad de secuencia puede emplear los programas BESTFIT o GAP en el paquete de software GCG Wisconsin (Accelrys, Madison WI), usando los parámetros predeterminados proporcionados.
- 10
- 15 "pH estable" se refiere a un polipéptido que mantiene una actividad similar (más de, por ejemplo, 60% a 80%) después de la exposición a pH bajo o alto (por ejemplo, 4,5-6 u 8-12) durante un período de tiempo (por ejemplo, 0,5-24 h) en comparación con la enzima no tratada.
- 20 "Aminoácido o residuo polar" se refiere a un aminoácido o residuo hidrófilo que tiene una cadena lateral que no está cargada a pH fisiológico, pero que tiene al menos un enlace en el que el par de electrones compartidos en común por dos átomos se mantiene más estrechamente por uno de los átomos. Los aminoácidos polares codificados genéticamente incluyen L-Asn (N), L-Gln (Q), L-Ser (S) y L-Thr (T).
- 25 "Codones de sesgo de uso de codones altos, óptimos y preferidos" se refiere, indistintamente, a codones que se usan con mayor frecuencia en las regiones codificantes de proteínas que otros codones que codifican el mismo aminoácido. Los codones preferidos se pueden determinar en relación con el uso de codones en un gen individual, un conjunto de genes de función u origen común, genes altamente expresados, la frecuencia de codones en las regiones codificantes de proteínas agregadas de todo el organismo, la frecuencia de codones en las regiones codificantes de proteínas agregadas de organismos relacionados, o combinaciones de los mismos. Los codones cuya frecuencia aumenta con el nivel de expresión génica son habitualmente codones óptimos para la expresión. Se conoce una variedad de métodos para determinar la frecuencia de codones (por ejemplo, uso de codones, uso de codones sinónimos relativos) y la preferencia de codones en organismos específicos, incluido el análisis multivariante, por ejemplo, usando análisis de grupos o análisis de correspondencia, y la cantidad efectiva de codones usados en un gen (ver GCG CodonPreference, Genetics Computer Group Wisconsin Package; Codon W, John Peden, University of Nottingham; McInerney, J. O, 1998, *Bioinformatics* 14:372-73; Stenico, *et al.*, 1994, *Nucleic Acids Res.* 22:2437-46; Wright, F., 1990, *Gene* 87:23-29). Las tablas de uso de codones están disponibles para una lista creciente de organismos (ver, por ejemplo, Wada, *et al.*, 1992, *Nucleic Acids Res.* 20:2111-2118; Nakamura, *et al.*, 2000, *Nucl. Acids Res.* 28:292; Duret, *et al.*, *supra*; Henaut y Danchin, "Escherichia coli and Salmonella", 1996, Neidhardt, *et al.*, Eds., ASM Press, Washington DC, p. 2047-2066. La fuente de datos para obtener el uso de codones puede basarse en cualquier secuencia de nucleótidos disponible capaz de codificar una proteína. Estos conjuntos de datos incluyen secuencias de ácido nucleico que se sabe que codifican proteínas expresadas (por ejemplo, secuencias codificantes de proteínas completas-CDS), etiquetas de secuencia expresada (EST) o regiones codificantes predichas de secuencias genómicas (véase, por ejemplo, Mount, D., *Bioinformatics: Sequence and Genome Analysis*, Capítulo 8, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 2001; Uberbacher, E.C., 1996, *Methods Enzymol.* 266:259-281; Tiwari, *et al.*, 1997, *Comput. Appl. Biosci.* 13:263-270).
- 30
- 35
- 40
- 45
- 50 "Proteína", "polipéptido" y "péptido" se usan indistintamente en la presente para indicar un polímero de al menos dos aminoácidos unidos covalentemente por un enlace de amida, independientemente de la longitud o modificación postraduccional (por ejemplo, glicosilación, fosforilación, lipidación, miristilación, ubiquitinación, etc.). Dentro de esta definición se incluyen D- y L-aminoácidos, y mezclas de D- y L-aminoácidos.
- 55 "Secuencia de referencia" se refiere a una secuencia definida con la que se compara otra secuencia (por ejemplo, alterada). Una secuencia de referencia puede ser un subconjunto de una secuencia más grande, por ejemplo, un segmento de un gen de longitud completa o secuencia de polipéptido. En general, una secuencia de referencia tiene al menos 20 residuos de nucleótidos o aminoácidos de longitud, al menos 25 residuos de longitud, al menos 50 residuos de longitud, o la longitud completa del ácido nucleico o polipéptido. Debido a que dos polinucleótidos o polipéptidos pueden cada uno (1) comprender una secuencia (es decir, una porción de la secuencia completa) que es similar entre las dos secuencias, y (2) puede comprender además una secuencia que es divergente entre las dos secuencias, las comparaciones de secuencia entre dos (o más) polinucleótidos o polipéptido se realizan habitualmente comparando secuencias de los dos polinucleótidos sobre una ventana de comparación para identificar y comparar regiones locales de similitud de secuencia.
- 60

El término "secuencia de referencia" no pretende limitarse a secuencias tipo silvestre, y puede incluir secuencias alteradas o modificadas genéticamente. Por ejemplo, en algunas realizaciones, una "secuencia de referencia" puede ser una secuencia de aminoácidos alterada o modificada genéticamente de forma previa. Por ejemplo, una "secuencia de referencia basada en SEQ ID NO: 2 que tiene un residuo de glicina en la posición X12" se refiere a una secuencia de referencia correspondiente a SEQ ID NO: 2 con un residuo de glicina en X12 (la versión inalterada de SEQ ID NO: 2 tiene un aspartato en X12).

"Aminoácido pequeño" o "residuo pequeño" se refiere a un aminoácido o residuo que tiene una cadena lateral que se compone de un total de tres o menos carbonos y/o heteroátomos (excluyendo el carbono  $\alpha$  y los hidrógenos). Los aminoácidos o residuos pequeños se pueden clasificar adicionalmente como aminoácidos o residuos pequeños alifáticos, no polares, polares o ácidos, de acuerdo con las definiciones anteriores. Los aminoácidos pequeños codificados genéticamente incluyen L-Ala (A), L-Val (V), L-Cys (C), L-Asn (N), L-Ser (S), L-Thr (T) y L-Asp (D).

"Estable al disolvente" o "estabilidad al disolvente" se refiere a un polipéptido que mantiene una actividad similar (más de, por ejemplo, 60% a 80%) después de la exposición a concentraciones variables (por ejemplo, 5-99%) de un disolvente, (por ejemplo, alcohol isopropílico, dimetilsulfóxido, tetrahidrofurano, 2-metiltetrahidrofurano, acetona, tolueno, acetato de butilo, éter ter-butílico de metilo, acetonitrilo, etc.), durante un período de tiempo (por ejemplo, 0,5-24 h) en comparación con la enzima no tratada.

"Identidad sustancial" se refiere a una secuencia de polinucleótidos o polipéptidos que tiene al menos 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 91, 93, 94, 95, 96, 97, 98, 99 o más por ciento de identidad de secuencia, en comparación con una secuencia de referencia en una ventana de comparación de al menos 20 posiciones de residuos, frecuentemente en una ventana de al menos 30-50 residuos, en donde el porcentaje de identidad de secuencia se calcula comparando la secuencia de referencia con una secuencia que incluye delecciones o adiciones que totalizan 20 por ciento o menos de la secuencia de referencia en la ventana de comparación. En realizaciones específicas aplicadas a polipéptidos, el término "identidad sustancial" significa que dos secuencias de polipéptidos, cuando se alinean de manera óptima, tal como por los programas GAP o BESTFIT usando ponderaciones de espacio predeterminados, comparten al menos 80 por ciento de identidad de secuencia, preferentemente al menos 89 por ciento de identidad de secuencia, al menos 95 por ciento de identidad de secuencia o más (por ejemplo, 99 por ciento de identidad de secuencia). Preferentemente, las posiciones de los residuos que no son idénticas difieren por sustituciones conservadoras de aminoácidos.

"Polipéptido sustancialmente puro" se refiere a una composición en la que la especie de polipéptido es la especie predominante presente (es decir, en una base molar o en peso es más abundante que cualquier otra especie macromolecular individual en la composición), y es en general una composición sustancialmente purificada cuando la especie objeto comprende al menos aproximadamente 50 por ciento de la especie macromolecular presente en mol o % en peso. En general, una composición de carboxiesterasa sustancialmente pura comprenderá aproximadamente 60 % o más, aproximadamente 70% o más, aproximadamente 80% o más, aproximadamente 90% o más, aproximadamente 95% o más, aproximadamente 96% o más, aproximadamente 97% o más, aproximadamente 98% o más o aproximadamente 99% o más de todas las especies macromoleculares por mol o % en peso presentes en la composición. En algunas realizaciones, la especie objeto se purifica hasta una homogeneidad esencial (es decir, las especies contaminantes no se pueden detectar en la composición mediante métodos de detección convencionales) en donde la composición consiste esencialmente en una especie macromolecular individual. Las especies de disolventes y las especies de iones elementales no se consideran especies macromoleculares. En algunas realizaciones, el polipéptido de carboxiesterasa mejorado aislado es una composición de polipéptido sustancialmente pura.

"Sustrato", tal como se usa en la presente, se refiere a un compuesto reactivo con carboxiesterasa, que consiste en un compuesto que contiene un éster (1), una amina (5) o un alcohol (2). En el contexto de la presente divulgación, un sustrato para la carboxiesterasa incluye, entre otros, el compuesto de la fórmula I y el compuesto de la fórmula II, como se describe adicionalmente en la presente.

"Termoestable" o "estabilidad térmica" se usan indistintamente para hacer referencia a un polipéptido que es resistente a la inactivación cuando se expone a un conjunto de condiciones de temperatura (por ejemplo, 40-80°C) durante un período de tiempo (por ejemplo, 0,5-24 h) en comparación con la enzima no tratada, reteniendo así un determinado nivel de actividad residual (más de 60% a 80%, por ejemplo) después de la exposición a temperaturas elevadas.

Como se usa en la presente, "estable al disolvente" se refiere a la capacidad de un polipéptido para mantener una actividad similar (por ejemplo, más de, por ejemplo, 60% a 80%) después de la exposición a concentraciones variables (por ejemplo, 5-99%) de disolvente (por ejemplo, alcohol isopropílico, tetrahidrofurano, 2-metiltetrahidrofurano, acetona, tolueno, acetato de butilo, éter ter-butílico de metilo, etc.), durante un periodo de tiempo (por ejemplo, 0,5-24 h) en comparación con la enzima no tratada.

## Descripción detallada de las realizaciones de la invención

En las realizaciones de la presente, las carboxiesterasas modificadas genéticamente se mejoran en su capacidad de convertir el sustrato de éster, oxazol-5-carboxilato de etilo en el producto, (4-isopropilpiperazin-1-il)(oxazol-5-il)metanona, en presencia de un sustrato de amina, 1-isopropil-piperazina, en comparación con la enzima de carboxiesterasa tipo silvestre, Esterasa 2 de *A. acidocaldarius* (SEQ ID NO: 2). Las carboxiesterasas, incluidas las descritas en la presente, son enzimas independientes que carecen de cofactores químicos, y que son enzimas solubles en agua que se pueden formular como enzima disuelta, como enzima inmovilizada en una resina, o como polvo liofilizado en presencia de una o más sales.

En algunas realizaciones, la mejora en la actividad enzimática es con respecto a otra carboxiesterasa modificada, tal como el polipéptido de SEQ ID NO: 4. La actividad mejorada en el sustrato de éster se puede manifestar por un aumento en la cantidad de sustrato convertido en producto (por ejemplo, porcentaje de conversión) por la enzima modificada genéticamente con respecto a una enzima de referencia (por ejemplo, SEQ ID NO: 2) bajo condiciones definidas. La actividad mejorada puede incluir una mayor velocidad de formación de producto que da como resultado un aumento en la conversión del sustrato de éster en el producto en presencia de un sustrato de amina en un tiempo definido bajo una condición definida. El aumento de la actividad (por ejemplo, el aumento del porcentaje de conversión y/o la velocidad de conversión) también se puede caracterizar por la conversión del sustrato en la misma cantidad de producto con una menor cantidad de enzima. La cantidad de producto se puede evaluar mediante una variedad de técnicas, por ejemplo, separación de la mezcla de reacción (por ejemplo, mediante cromatografía) y detección del producto separado mediante absorbancia UV o espectroscopía de masas en tándem (MS/MS) (ver, por ejemplo, Ejemplo 4). Una condición de reacción definida de ejemplo para comparación con la actividad de SEQ ID NO: 2 es aproximadamente 40 g/L de oxazol-5-carboxilato de etilo (el sustrato de éster), 44 g/L de 1-isopropil-piperazina (el sustrato de amina) y 20 g/L de un polipéptido de carboxiesterasa correspondiente a una secuencia de aminoácidos elegida de: SEQ ID NOs: 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 92, o 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, o 126, en donde la enzima se prepara en presencia de sulfato de sodio y se procesa en presencia de aproximadamente 10 g/L a aproximadamente 20 g/L de agua en metil-isobutil-cetona (MIBK), como se indica a continuación en la descripción de las condiciones de reacción para las carboxiesterasas listadas en la Tabla 3. Las condiciones de reacción definidas para la comparación con determinadas carboxiesterasas modificadas genéticamente también se proporcionan en la descripción para las carboxiesterasas listadas en la Tabla 3 y en las descripciones correspondientes en el Ejemplo 7. En algunas realizaciones, las carboxiesterasas modificadas genéticamente tienen al menos 1,5 veces, 2 veces, 3 veces, 4 veces, 5 veces, 6 veces, 7 veces, 8 veces, 9 veces, 10 veces, 15 veces, 20 veces, 30 veces, 40 veces, 50 veces, 75 veces, 100 veces, 150 veces, 200 veces, 300 veces, 400 veces, 500 veces, 1000 veces, 1.500 veces, 2.000 veces, 10.000 veces, 100.000 veces, 500.000 veces, 785.000 veces o más que la actividad del polipéptido de SEQ ID NO: 2 bajo las condiciones de reacción definidas.

En algunas realizaciones, la actividad enzimática mejorada también se asocia con otras mejoras en la propiedad enzimática. En algunas realizaciones, la mejora en la propiedad enzimática es con respecto a la estabilidad térmica, tal como a 60°C o más.

En algunas realizaciones, la actividad enzimática mejorada está asociada con mejoras en la estabilidad al disolvente, tal como cuando se ejecuta en 98% (volumen/volumen) en metil-isobutil-cetona (MIBK) o éter metílico de ter-butilo (TBME).

En algunas realizaciones, los polipéptidos de carboxiesterasa modificados genéticamente de la presente divulgación son capaces de convertir el sustrato de éster, oxazol-5-carboxilato de etilo, en el producto, (4-isopropilpiperazin-1-il)(oxazol-5-il)metanona con una actividad que es mayor que la actividad del polipéptido de SEQ ID NO: 2 en presencia de un sustrato de amina, por ejemplo, 1-isopropil-piperazina, y comprende una secuencia de aminoácidos que es al menos 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o más, idéntica a la secuencia de referencia de SEQ ID NO: 2, o un fragmento funcional del mismo.

En algunas realizaciones, los polipéptidos de carboxiesterasa modificados genéticamente de la presente divulgación son capaces de convertir el sustrato de éster, oxazol-5-carboxilato de etilo en el producto, (4-isopropilpiperazin-1-il)(oxazol-5-il)metanona en presencia de un sustrato de amina, tal como 1-isopropil-piperazina, con una actividad que es mayor que el polipéptido de SEQ ID NO: 2, y comprende una secuencia de aminoácidos que es al menos 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o más idéntica a una secuencia de referencia listada en la Tabla 3, por ejemplo, SEQ ID NO: 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, o 126, o un fragmento funcional del mismo, como se describe más adelante.

- En algunas realizaciones, los polipéptidos de carboxiesterasa modificados genéticamente comprenden una secuencia de aminoácidos que tiene una o más diferencias de residuos en comparación con una secuencia de referencia de carboxiesterasa. Las diferencias de residuos pueden ser sustituciones no conservadoras, sustituciones conservadoras o una combinación de sustituciones no conservadoras y conservadoras. Con respecto a las diferencias de residuos y las descripciones de posiciones de residuos, las carboxiesterasas proporcionadas en la presente se pueden describir en referencia a la secuencia de aminoácidos de las carboxiesterasas de origen natural de Esterasa 2 de *A. acidocaldarius* (SEQ ID NO: 2), la carboxiesterasa de SEQ ID NO: 2, o una carboxiesterasa modificada, tal como el polipéptido de SEQ ID NO: 4. Para las descripciones en la presente, la posición del residuo de aminoácido en la secuencia de referencia se determina en la carboxiesterasa a partir del residuo de metionina (M) de inicio (es decir, M representa la posición del residuo 1), aunque el experto en la técnica entenderá que este residuo de metionina de inicio puede removerse mediante maquinaria de procesamiento biológico, tal como en una célula hospedadora o sistema de traducción *in vitro*, para generar una proteína madura que carece del residuo de metionina de inicio.
- En algunas realizaciones, las diferencias de residuos pueden ocurrir en una o más de las siguientes posiciones de residuos: X2, X7, X9, X10, X19, X20, X22, X27, X28, X29, X30, X33, X34, X35, X36, X37, X38, X46, X48, X54, X57, X66, X75, X85, X86, X87, X96, X103, X139, X160, X176, X181, X183, X185, X188, X190, X197, X198, X205, X207, X208, X212, X216, X248, X249, X255, X263, X266, X270, X271, X278, X280, X286, X290 y X296.
- En algunas realizaciones, las diferencias de residuos o combinaciones de los mismos, se asocian con propiedades enzimáticas mejoradas. En algunas realizaciones, los polipéptidos de carboxiesterasa pueden tener, adicionalmente, 1-2, 1-3, 1-4, 1-5, 1-6, 1-7, 1-8, 1-9, 1-10, 1-11, 1-12, 1-13, 1-14, 1-15, 1-16, 1-17, 1-18, 1-19, 1-20, 1-22, 1-24, 1-26, 1-28, 1-30, 1-35, 1-40, 1-45, 1-50, 1-55, 1-60, o 1-62 diferencias de residuos en posiciones de residuos distintas de las posiciones específicas indicadas por "Xn" listadas anteriormente. En algunas realizaciones, el número de diferencias puede ser 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 22, 24, 26, 28, 30, 35, 40, 45, 50, 55, 60, o 62 diferencias de residuos en las otras posiciones de residuos de aminoácidos. En algunas realizaciones, las diferencias de residuos en otras posiciones de residuos comprenden sustituciones con residuos de aminoácidos conservadores.
- En las realizaciones de la presente, las diferencias de residuos en comparación con SEQ ID NO: 2 en las posiciones de los residuos que afectan la unión del sustrato en la carboxiesterasa permite la acomodación del sustrato de éster de la fórmula estructural (I), que se describe más adelante, en particular, el sustrato de éster, oxazol-5-carboxilato de etilo. Sin limitarse a la teoría, al menos dos regiones, una primera región de unión al sustrato y una segunda región de unión al sustrato, interactúan con diferentes elementos estructurales del sustrato de éster. La primera región de unión comprende el residuo X85, X185, X214, X215 y X254, la segunda región de unión comprende las posiciones de residuo X30, X33, X34, X37, X82, X205, X210, X283, X286 y X287, mientras que las posiciones X83, X84, X155, X156, X206, X214 y X282 se superponen a los dos sitios. Estos residuos se determinaron mediante cristalografía de rayos X. Por consiguiente, los polipéptidos de carboxiesterasa en la presente tienen una o más diferencias de residuos en las posiciones de residuos que comprenden X30, X33, X34, X37, X85, X185, X205 y X286. En algunas realizaciones, los polipéptidos de carboxiesterasa en la presente tienen al menos dos o más, o tres o más diferencias de residuos en las posiciones de residuos especificadas asociadas con la unión al sustrato.
- En otras realizaciones, el polipéptido de carboxiesterasa mejorado comprende una secuencia de aminoácidos que es al menos 80% idéntica a la secuencia de aminoácidos expuesta en SEQ ID NO: 4, en donde la secuencia de aminoácidos de carboxiesterasa mejorada incluye la característica de que: el residuo correspondiente a X198 se elige de: un residuo no polar, un residuo aromático y un residuo alifático. En aun otras realizaciones, los polipéptidos de carboxiesterasa mejorados incluyen la siguiente característica: X198 se elige de: F, L, I, Y y M. En algunas realizaciones, el polipéptido de carboxiesterasa mejorado puede comprender una secuencia de aminoácidos que comprende una o más diferencias de residuos en comparación con la secuencia de SEQ ID NO: 4 en las siguientes posiciones de residuos correspondientes a: X27, X30, X35, X37, X57, X66, X75, X103, X207, X208, X271 X286, y X296. La guía para la elección de varios residuos de aminoácidos que pueden estar presentes en las posiciones de residuos especificadas se proporciona en la descripción detallada que sigue.
- En otras realizaciones, el polipéptido de carboxiesterasa mejorado comprende una secuencia de aminoácidos que es al menos 80% idéntica a la secuencia de aminoácidos expuesta en SEQ ID NO: 4, en donde la secuencia de aminoácidos incluye al menos una característica elegida de: el residuo correspondiente a X27 es un residuo restringido; el residuo correspondiente a X30 es un residuo alifático; el residuo correspondiente a X35 se elige de un residuo básico y un residuo polar; el residuo correspondiente a X37 se elige de un residuo alifático y un residuo polar; el residuo correspondiente a X57 es un residuo no polar; el residuo correspondiente a X75 se elige de un residuo básico y un residuo polar; el residuo correspondiente a X103 se elige de un residuo no polar y un residuo aromático; el residuo correspondiente a X185 se elige de un residuo alifático, un residuo no polar y un residuo aromático; el residuo correspondiente a X207 se elige de un residuo ácido y un residuo polar; el residuo correspondiente a X208 se elige de un residuo alifático, un residuo básico y un residuo polar; el residuo

correspondiente a X271 se elige de un residuo ácido y un residuo polar; el residuo correspondiente a X286 se elige de un residuo alifático, un residuo polar y un residuo pequeño; y el residuo correspondiente a X296 se elige de un residuo alifático y un residuo básico.

- 5 En aun otras realizaciones, el polipéptido de carboxiesterasa mejorado comprende una secuencia de aminoácidos que es al menos 80% idéntica a la secuencia de aminoácidos expuesta en SEQ ID NO: 4, en donde la secuencia de aminoácidos incluye al menos una característica elegida de: X27 es P; X30 se elige de I, L y V; X35 es H; X37 se elige de I, L, T y V; X57 es M; X75 es R; X103 se elige de F, M y W; X185 se elige de F, I y M; X207 es E; X208 se elige de R, L y H; X271 es D; X286 se elige de M, V y G; y X296 se elige de V, L y R. En algunas realizaciones, los polipéptidos de carboxiesterasa mejorados comprenden una secuencia de aminoácidos que incluye las siguientes características: X35 se elige de un residuo básico y un residuo polar; y X185 se elige de un residuo polar y un residuo alifático. En realizaciones alternativas, el polipéptido de carboxiesterasa mejorado comprende una secuencia de aminoácidos que incluye las siguientes características: X35 es H; y X185 se elige de F, I y M.

- 15 En algunas realizaciones, el polipéptido de carboxiesterasa mejorado comprende una diferencia de residuos en comparación con la secuencia de aminoácidos expuesta en SEQ ID NO: 24, al menos una posición de residuo elegida de: X9, X19, X34, X35, X37, X46, X48, X66, X87, X103, X139, X190, X207, X216, X263, X271, X278 y X296. En otras realizaciones, el polipéptido de carboxiesterasa mejorado comprende una secuencia de aminoácidos que incluye al menos una de las siguientes características: el residuo correspondiente a X9 es un residuo aromático; el residuo correspondiente a X19 se elige de un residuo básico y un residuo polar; el residuo correspondiente a X34 se elige de un residuo restringido, un residuo ácido y un residuo polar; el residuo correspondiente a X35 se elige de un residuo polar; el residuo correspondiente a X46 es un residuo alifático; el residuo correspondiente a X48 es un residuo alifático; el residuo correspondiente a X66 es un residuo alifático; el residuo correspondiente a X87 se elige de un residuo alifático y un residuo pequeño; el residuo correspondiente a X103 se elige de un residuo aromático; el residuo correspondiente a X139 es un residuo básico; el residuo correspondiente a X190 es un residuo aromático; ; el residuo correspondiente a X207 es un residuo básico; el residuo correspondiente a X216 se elige de un residuo aromático, un residuo básico y un residuo polar; el residuo correspondiente a X263 se elige de un residuo alifático y un residuo polar; el residuo correspondiente a X271 se elige de un residuo ácido y un residuo polar; el residuo correspondiente a X278 se elige de un residuo alifático y un residuo aromático, y el residuo correspondiente a X296 se elige de un residuo alifático y un residuo básico.

- En aun otras realizaciones, el polipéptido de carboxiesterasa mejorado comprende una secuencia de aminoácidos que es al menos 80% idéntica a la secuencia de aminoácidos expuesta en SEQ ID NO: 24, en donde la secuencia de aminoácidos incluye al menos una característica elegida de: X9 es Y; X19 es R; X34 se elige de E, N o P; X35 es S, X37 es T; X46 se elige de I, L o V; X48 es L; X66 es V; X87 es A; X103 se elige de W o F; X139 es R; X190 es Y; X207 es E; X216 se elige de N y W; X263 se elige de T y A; X271 es D; X278 se elige de W y L; y X296 se elige de V, L y R. En otras realizaciones más, el polipéptido de carboxiesterasa mejorado comprende una secuencia de aminoácidos que incluye las siguientes características: X9 es un residuo aromático y X87 es un residuo alifático.
- 40 En realizaciones alternativas, el polipéptido de carboxiesterasa mejorado comprende una secuencia de aminoácidos que incluye las siguientes características: X9 es Y; y X87 es A

- En otras realizaciones, el polipéptido de carboxiesterasa mejorado comprende una secuencia de aminoácidos que comprende una diferencia de residuos en comparación con la secuencia de aminoácidos expuesta en SEQ ID NO: 54, al menos una posición de residuo elegida de: X20, X28, X29, X30, X33, X34, X188, X216 y X286. En realizaciones alternativas, el polipéptido de carboxiesterasa mejorado comprende una secuencia de aminoácidos que incluye al menos una característica elegida de: X20 se elige de un residuo alifático y un residuo básico; el residuo correspondiente a X28 se elige de un residuo ácido, un residuo polar y un residuo restringido; el residuo correspondiente a X29 se elige de un residuo ácido y un residuo polar; el residuo correspondiente a X30 es un residuo alifático; el residuo correspondiente a X33 es un residuo aromático; el residuo correspondiente a X34 es un residuo pequeño; el residuo correspondiente a X188 se elige de un residuo pequeño y un residuo aromático; el residuo correspondiente a X216 es un residuo polar, y el residuo correspondiente a X286 se elige de un residuo alifático, un residuo pequeño, un residuo no polar y un residuo polar.

- 55 En algunas realizaciones, el polipéptido de carboxiesterasa mejorado comprende la secuencia de aminoácidos expuesta en SEQ ID NO: 54, en donde la secuencia de aminoácidos incluye al menos una mutación específica elegida de: X20 se elige de I y R; X28 se elige de D, P y S; X29 es D; X30 es V; X33 es W; X34 es G; X188 se elige de G y F; X216 es N y X286 se elige de S, M, V, G y A. En realizaciones alternativas, el polipéptido de carboxiesterasa mejorado comprende una secuencia de aminoácidos que incluye la siguiente característica: X216 es un residuo polar. En aun otras realizaciones, el polipéptido de carboxiesterasa mejorado comprende una secuencia de aminoácidos que incluye la siguiente característica: X216 es N.

En otras realizaciones, el polipéptido de carboxiesterasa mejorado comprende una secuencia de aminoácidos que comprende una diferencia de residuos en comparación con la secuencia de aminoácidos expuesta en SEQ ID NO:

68 en una posición de residuo elegida de: X10, X20, X22, X28, X30, X33, X36, X37, X46, X66, X75, X103, X197, X263, X266, X280, y X290. En otra realización, el polipéptido de carboxiesterasa mejorado comprende una secuencia de aminoácidos que incluye al menos una característica elegida de: el residuo correspondiente a X10 es un residuo alifático; el residuo correspondiente a X20 se elige de un residuo alifático y un residuo básico; el residuo correspondiente a X22 es un residuo aromático; el residuo correspondiente a X28 se elige de un residuo ácido, un residuo polar y un residuo restringido; el residuo correspondiente a X30 es un residuo alifático; el residuo correspondiente a X33 es un residuo aromático; el residuo correspondiente a X36 es un residuo alifático o aromático; el residuo correspondiente a X37 es un residuo aromático o pequeño; el residuo correspondiente a X46 es un residuo básico; el residuo correspondiente a X66 es un residuo polar; el residuo correspondiente a X75 es un residuo básico; el residuo correspondiente a X103 es un residuo aromático; el residuo correspondiente a X197 es un residuo alifático; el residuo correspondiente a X263 es un residuo básico; el residuo correspondiente a X266 es un residuo polar; el residuo correspondiente a X280 se elige de un residuo alifático y un residuo polar; y el residuo correspondiente a X290 se elige de un residuo alifático y un residuo aromático.

En aun otras realizaciones, el polipéptido de carboxiesterasa mejorado comprende una secuencia de aminoácidos que es al menos 80% idéntica a la secuencia de aminoácidos expuesta en SEQ ID NO: 68, en donde la secuencia de aminoácidos incluye al menos una característica elegida de: X10 se elige de L y M; X20 se elige de I y R; X22 es W; X28 se elige de D, P y S; X30 es V; X33 es W; X36 se elige de F, I y M; X37 se elige de G e Y; X46 es R; X66 es T; X75 es R; X103 es W; X197 es L; X263 es R; X266 es T; X280 se elige de M y T; y X290 se elige de W e I. En una realización alternativa, el polipéptido de carboxiesterasa mejorado comprende una secuencia de aminoácidos que incluye al menos una característica elegida de: X30 es un residuo alifático, X33 es un residuo aromático, X75 es un residuo básico y X103 es un residuo aromático. En aun otra realización, el polipéptido de carboxiesterasa mejorado comprende una secuencia de aminoácidos que incluye las siguientes características: X30 es V; X33 es W; X75 es R; y X103 es W.

En otras realizaciones, el polipéptido de carboxiesterasa mejorado comprende una secuencia de aminoácidos que comprende una diferencia de residuos en comparación con la secuencia de aminoácidos expuesta en SEQ ID NO: 68 en una posición de residuo elegida de: X28, X38, X46, X54, X66, X75, X85, X86, X96, X160, X176, X183, X188, X205, X212, X248, X249, X255, X270, y X286. En otra realización, el polipéptido de carboxiesterasa mejorado comprende una secuencia de aminoácidos que incluye al menos una característica elegida de: el residuo correspondiente a X28 se elige de un residuo ácido, un residuo polar, un residuo pequeño y un residuo restringido; el residuo correspondiente a X38 se elige de un residuo alifático y un residuo básico; el residuo correspondiente a X46 se elige de un residuo ácido y un residuo básico; el residuo correspondiente a X54 se elige de un residuo ácido y un residuo polar; el residuo correspondiente a X66 es un residuo polar; el residuo correspondiente a X75 es un residuo básico; el residuo correspondiente a X85 se elige de un residuo aromático o básico y un residuo pequeño; el residuo correspondiente a X86 es un residuo polar; el residuo correspondiente a X96 se elige de un residuo no polar y un residuo alifático; el residuo correspondiente a X160 se elige de un residuo polar y un residuo restringido; el residuo correspondiente a X176 se elige de un residuo alifático, un residuo aromático o básico y un residuo no polar; el residuo correspondiente a X183 es un residuo no polar; el residuo correspondiente a X188 se elige de un residuo aromático y uno pequeño; el residuo correspondiente a X205 es un residuo aromático; el residuo correspondiente a X212 es un residuo ácido; el residuo correspondiente a X248 es un residuo alifático; el residuo correspondiente a X249 es un residuo aromático; el residuo correspondiente a X255 es un residuo polar; el residuo correspondiente a X270 se elige de un residuo alifático y un residuo polar; y el residuo correspondiente a X286 se elige de un residuo alifático, un residuo no polar, un residuo pequeño y un residuo polar.

En aun otras realizaciones, el polipéptido de carboxiesterasa mejorado comprende una secuencia de aminoácidos que es al menos 80% idéntica a la secuencia de aminoácidos expuesta en SEQ ID NO: 68, en donde la secuencia de aminoácidos incluye al menos una característica elegida de: X28 se elige de C, D, S, H, P, G y R; X38 se elige de E y L; X46 se elige de K, R y Q; X54 se elige de R, Q y S; X66 se elige de L, T y V; X75 es R; X85 se elige de G y H; X86 es T; X96 se elige de M y L; X160 se elige de T y P; X176 se elige de M, L y H; X183 es Q; X188 se elige de G y F; X205 es F; X212 es D; X248 es V; X249 es W; X255 es N; X270 se elige de N y L; y X286 se elige de M, V, G, N y S. En una realización alternativa, el polipéptido de carboxiesterasa mejorado comprende una secuencia de aminoácidos que incluye al menos una característica elegida de: X28 es un residuo polar, X38 es un residuo básico y X85 es un residuo pequeño. En aun otra realización, el polipéptido de carboxiesterasa mejorado comprende una secuencia de aminoácidos que incluye las siguientes características: X28 es C; X38 es E; y X85 es G.

En otras realizaciones, el polipéptido de carboxiesterasa mejorado comprende una secuencia de aminoácidos que comprende una diferencia de residuos en comparación con la secuencia de aminoácidos expuesta en SEQ ID NO: 100 en una posición de residuo elegida de: X7, X22, X36, X38, X46, X54, X66, y X75. En otra realización, el polipéptido de carboxiesterasa mejorado comprende una secuencia de aminoácidos que incluye al menos una característica elegida de: el residuo correspondiente a X7 es un residuo alifático; el residuo correspondiente a X22 se elige de un residuo alifático y un residuo aromático; el residuo correspondiente a X36 se elige de un residuo polar y un residuo no polar; el residuo correspondiente a X38 es un residuo aromático; el residuo correspondiente

a X46 se elige de un residuo polar y un residuo básico; el residuo correspondiente a X54 se elige de un residuo polar y un residuo básico; el residuo correspondiente a X66 es un residuo polar; y el residuo correspondiente a X75 se elige de un residuo básico y un residuo no polar.

5 En aun otras realizaciones, el polipéptido de carboxiesterasa mejorado comprende una secuencia de aminoácidos que es al menos 80% idéntica a la secuencia de aminoácidos expuesta en SEQ ID NO: 100, en donde la secuencia de aminoácidos incluye al menos una característica elegida de: X7 es L; X22 se elige de W y L; X36 se elige de T y M; X38 es W; X46 se elige de K y Q; X54 se elige de S, Q y K; X66 se elige de G y T; y X75 se elige de M y R. En una realización alternativa, el polipéptido de carboxiesterasa mejorado comprende una secuencia de aminoácidos que incluye al menos una característica elegida de: X36 es un residuo polar, X38 es un residuo aromático y X75 es un residuo básico. En aun otra realización, el polipéptido de carboxiesterasa mejorado comprende una secuencia de aminoácidos que incluye las siguientes características: X36 es T; X38 es W; y X75 es R.

15 En otras realizaciones, el polipéptido de carboxiesterasa mejorado comprende una secuencia de aminoácidos que comprende una diferencia de residuos en comparación con la secuencia de aminoácidos expuesta en SEQ ID NO: 114 en una posición de residuo elegida de: X2, X181 y X286. En otra realización, el polipéptido de carboxiesterasa mejorado comprende una secuencia de aminoácidos que incluye al menos una característica elegida de: el residuo correspondiente a X2 se elige de un residuo alifático, un residuo básico, un residuo polar y un residuo aromático; el residuo correspondiente a X181 es un residuo básico; y el residuo correspondiente a X286 se elige de un residuo polar y un residuo no polar.

25 En aun otras realizaciones, el polipéptido de carboxiesterasa mejorado comprende una secuencia de aminoácidos que es al menos 80% idéntica a la secuencia de aminoácidos expuesta en SEQ ID NO: 114, en donde la secuencia de aminoácidos incluye al menos una característica elegida de: X2 se elige de L, Q, R y H; X181 es Q; y X286 se elige de C y S. En una realización alternativa, el polipéptido de carboxiesterasa mejorado comprende una secuencia de aminoácidos que incluye al menos una característica elegida de: X286 es un residuo no polar. En aun otra realización, el polipéptido de carboxiesterasa mejorado comprende una secuencia de aminoácidos que incluye la siguiente característica: X286 es C.

30 En algunas realizaciones, el polipéptido de carboxiesterasa mejorado comprende una secuencia de aminoácidos que corresponde a las secuencias de aminoácidos expuestas en SEQ ID NO: 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, o 126.

35 En un aspecto adicional, la presente divulgación proporciona polinucleótidos que codifican para cada uno de los polipéptidos de carboxiesterasa modificados genéticamente mejorados descritos anteriormente. En algunas realizaciones, los polinucleótidos pueden ser parte de un vector de expresión que tiene una o más secuencias de control para la expresión del polipéptido de carboxiesterasa. En una realización alternativa, el polinucleótido corresponde a cualquiera de las secuencias de nucleótidos expuestas en SEQ ID NO: 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, o 125.

45 En otro aspecto, la presente divulgación proporciona células hospedadoras que comprenden los polinucleótidos que codifican para las carboxiesterasas modificadas genéticamente o vectores de expresión capaces de expresar las carboxiesterasas modificadas genéticamente. En algunas realizaciones, la célula hospedadora puede ser una célula hospedadora bacteriana, tal como *E. coli*. Las células hospedadoras se pueden usar para la expresión y el aislamiento de las enzimas carboxiesterasa modificadas genéticamente descritas en la presente, o, alternativamente, se pueden usar directamente para la conversión del sustrato de éster en producto. En algunas realizaciones, las amidas modificadas genéticamente, en la forma de células enteras, extractos crudos, polipéptidos aislados o polipéptidos purificados, se pueden usar individualmente o como una combinación de diferentes amidas modificadas genéticamente.

55 El experto en la técnica apreciará que, tras la producción de una enzima, en particular, dependiendo de la línea celular usada y la secuencia de aminoácidos particular de la enzima, pueden ocurrir modificaciones postraduccionales. Por ejemplo, estas modificaciones postraduccionales pueden incluir la escisión de determinadas secuencias líder, la adición de diversas porciones de azúcar en diversos patrones de glicosilación y fosforilación, desamidación, oxidación, aleatorización de enlaces disulfuro, isomerización, recorte de lisina de extremo C y ciclación de glutamina de extremo N. La presente invención abarca el uso de enzimas carboxiesterasa modificadas genéticamente que se han sometido a, o pueden haber sufrido, una o más modificaciones postraduccionales. Por lo tanto, las carboxiesterasas modificadas genéticamente de la invención pueden incluir una que haya sufrido una modificación postraducciona, tal como se describe en la presente.

La desamidación es una reacción enzimática que convierte principalmente la asparagina (N) en ácido iso-aspartico

(iso-aspartato) y ácido aspártico (aspartato) (D) en una proporción de aproximadamente 3:1. Esta reacción de desamidación está, por lo tanto, relacionada con la isomerización del aspartato (D) a iso-aspartato. La desamidación de la asparagina y la isomerización del aspartato, ambas implican la succinimida intermedia. En un grado mucho menor, la desamidación se puede presentar con residuos de glutamina de una manera similar.

La oxidación se puede presentar durante la producción y el almacenamiento (es decir, en presencia de condiciones oxidantes) y da como resultado una modificación covalente de una proteína, inducida directamente por especies reactivas de oxígeno o indirectamente por reacción con subproductos secundarios del estrés oxidativo. La oxidación ocurre principalmente con residuos de metionina, pero se puede presentar en residuos de triptófano y cisteína libre.

La aleatorización de enlaces de disulfuro se puede presentar durante la producción y las condiciones básicas de almacenamiento. Bajo determinadas circunstancias, los enlaces disulfuro pueden romperse o formarse incorrectamente, lo que resulta en residuos de cisteína no apareados (-SH). Estos sulfhidrilos libres (no emparejados) (-SH) pueden promover el reordenamiento.

Es probable que la glutamina (Q) y el glutamato (ácido glutámico) (E) de extremo N en las carboxiesterasas modificadas genéticamente formen piroglutamato (pGlu) a través de la ciclación. La mayor parte de la formación de pGlu ocurre en la fabricación, pero se puede formar de forma no enzimática, dependiendo del pH y la temperatura de las condiciones de procesamiento y almacenamiento.

El recorte de lisina de extremo C es una reacción enzimática catalizada por carboxipeptidasas, y se observa comúnmente en enzimas. Las variantes de este proceso incluyen la remoción de lisina de las enzimas de la célula hospedadora recombinante.

En la presente invención, no se sabe que las modificaciones y cambios postraduccionales en la secuencia de aminoácidos primaria descritos anteriormente den como resultado cambios significativos en la actividad de las enzimas carboxiesterasa modificadas genéticamente.

La Tabla 3 a continuación proporciona polipéptidos de carboxiesterasa modificados genéticamente de ejemplo, con cada fila listando dos SEQ ID NOs, el número impar se refiere a la secuencia de nucleótidos que codifica para la secuencia de aminoácidos proporcionada por el número par. Las diferencias de residuos se basan en la comparación con la secuencia de referencia de SEQ ID NO: 2, una carboxiesterasa correspondiente a la Esterasa 2 de *A. acidocaldarius* tipo silvestre, a la que se hace referencia en el Ejemplo 6. En la columna Actividad, los niveles de actividad creciente (es decir, "+" "++" "+++", etc.), se definen de la siguiente manera: "-" indica menos de 1% de conversión de sustrato a producto pero no más de 0,9% de conversión (175 µL de lisado, 100mM de éster, 100mM de isopropil-piperazina, 2% de agua en TBME); "+" indica al menos 1,1 a 80 veces la actividad de SEQ ID NO: 2, pero no mayor que la actividad de SEQ ID NO: 4 (175 µL de lisado, 100mM de éster, 100mM de isopropil-piperazina, 2% de agua en TBME); "++" indica al menos 1,1 a 11 veces la actividad de SEQ ID NO: 4, pero no mayor que la actividad de SEQ ID NO: 18 (150µL de lisado, 200mM de éster, 200mM de isopropil-piperazina, 2% de agua en MIBK); "+++" indica al menos 1,1 a 5 veces la actividad de SEQ ID NO: 18 pero no mayor que la actividad de SEQ ID NO: 54 (120µL de lisado, 300mM de éster, 300mM de isopropil-piperazina, 2% de agua en MIBK); "++++" indica al menos 1,5 a 2 veces la actividad de SEQ ID NO: 54, pero no mayor que la actividad de SEQ ID NO: 68 (90µL de lisado, 300mM de éster, 300mM de isopropil-piperazina, 2% de agua en MIBK); "+++++" indica al menos 1,1 a 2,0 veces la actividad de SEQ ID NO: 68 (50µL de lisado, 300mM de éster, 300mM de isopropil-piperazina, 2% de agua en MIBK); "\$" indica al menos 1,1 a 2 veces la actividad de SEQ ID NO: 68, pero no mayor que la actividad de SEQ ID NO: 100 (50µL de lisado, 354mM de éster, 425mM de isopropil-piperazina, 2% de agua en MIBK); "\$\$" indica al menos 1,1 a 5 veces la actividad de SEQ ID NO: 100, pero no mayor que la actividad de SEQ ID NO: 114 (50µL de lisado, 354mM de éster, 425mM de isopropil-piperazina, 2% de agua en MIBK); "\$\$\$" indica al menos 1,1 a 2 veces la actividad de SEQ ID NO: 114 (50µL de lisado, 354mM de éster, 354mM de isopropil-piperazina, 2% de agua en MIBK). En cada caso, la actividad se determinó usando una cantidad variable de lisado cargado en una selección de liofilización y actividad de múltiples pocillos, luego se hizo reaccionar con la concentración observada de sustrato y en el sistema disolvente observado en un volumen de 200 µL durante 16 horas, como se describe en el Ejemplo 5.

Tabla 3

| SEQ ID NO | Mutaciones con respecto a SEQ ID NO: 2 | Número de diferencias de residuos | Actividad |
|-----------|----------------------------------------|-----------------------------------|-----------|
| 1/2       |                                        | -                                 | -         |
| 3/4       | E198L;                                 | 1                                 | +         |

# ES 3 018 269 T3

|       |                                                                       |    |       |
|-------|-----------------------------------------------------------------------|----|-------|
| 5/6   | E198I;                                                                | 1  | +     |
| 7/8   | E198Y;                                                                | 1  | +     |
| 9/10  | E198F;                                                                | 1  | +     |
| 11/12 | E198M;                                                                | 1  | +     |
| 13/14 | S185M; E198L; T207E; K296V;                                           | 4  | ++    |
| 15/16 | F37L; S185F; E198L; T207E; G208H;                                     | 5  | ++    |
| 17/18 | S185F; E198L;                                                         | 2  | ++    |
| 19/20 | F37L; P75R; S185F; E198L; G208H; K296L;                               | 6  | ++    |
| 21/22 | S185F; E198L; T207E; G208H;                                           | 4  | ++    |
| 23/24 | S35H; S185F; E198L;                                                   | 3  | ++    |
| 25/26 | A27P; F30L; D103W; E198I; Q286G;                                      | 5  | ++    |
| 27/28 | G57M; D103W; E198L; G208H; Q286G; K296R;                              | 6  | ++    |
| 29/30 | E198L; Q286G;                                                         | 2  | ++    |
| 31/32 | D103W; E198I; Q286G;                                                  | 3  | ++    |
| 33/34 | D103W; E198L; K271D; Q286G;                                           | 4  | ++    |
| 35/36 | Q9Y; S35H; F37T; D103W; S185F; E198L;                                 | 6  | +++   |
| 37/38 | V48L; S185F; E198L; D216N;                                            | 4  | +++   |
| 39/40 | Q9Y; A19R; Q34N; S35H; V87A; S185F; E198L; K271D; K296R;              | 9  | +++   |
| 41/42 | Q9Y; Q34E; S35H; V87A; S185F; E198L; K271D;                           | 7  | +++   |
| 43/44 | Q9Y; A19R; Q34E; S35H; A46V; V87A; S185F; E198L; K271D; E278Q; K296R; | 11 | +++   |
| 45/46 | Q9Y; Q34N; S35H; A46V; V87A; S185F; E198L; E278Q; K296R;              | 9  | +++   |
| 47/48 | S35H; R66V; S185F; P190Y; E198L; T207E; E263T;                        | 7  | +++   |
| 49/50 | Q9Y; Q34E; S35H; A46V; V87A; A139R; S185F; E198L;                     | 8  | +++   |
| 51/52 | Q9Y; Q34N; S35H; V87A; S185F; E198L; K296R;                           | 7  | +++   |
| 53/54 | Q9Y; S35H; V87A; S185F; E198L;                                        | 5  | +++   |
| 55/56 | Q9Y; S35H; V87A; S185F; E198L; Q286S;                                 | 6  | ++++  |
| 57/58 | Q9Y; F30V; S35H; V87A; S185F; E198L;                                  | 6  | ++++  |
| 59/60 | Q9Y; Q34E; S35H; V87A; S185F; E198L;                                  | 6  | ++++  |
| 61/62 | Q9Y; S35H; V87A; S185F; Y188F; E198L;                                 | 6  | ++++  |
| 63/64 | Q9Y; Q34G; S35H; V87A; S185F; E198L;                                  | 6  | ++++  |
| 65/66 | Q9Y; P20I; S35H; V87A; S185F; E198L;                                  | 6  | ++++  |
| 67/68 | Q9Y; S35H; V87A; S185F; E198L; D216N;                                 | 6  | ++++  |
| 69/70 | Q9Y; Q29D; S35H; V87A; S185F; E198L;                                  | 6  | ++++  |
| 71/72 | Q9Y; Q33G; S35H; V87A; S185F; E198L;                                  | 6  | ++++  |
| 73/74 | Q9Y; Q28S; S35H; V87A; S185F; E198L;                                  | 6  | ++++  |
| 75/76 | Q9Y; F30V; S35H; V87A; S185F; I197L; E198L; D216N;                    | 8  | +++++ |
| 77/78 | Q9Y; S35H; L36I; F37G; V87A; S185F; E198L; D216N; L280T; L290I;       | 10 | +++++ |
| 79/80 | Q9Y; V10M; S35H; F37G; V87A; S185F; E198L; D216N;                     | 8  | +++++ |
| 81/82 | Q9Y; Q28S; S35H; V87A; S185F; E198L; D216N;                           | 7  | +++++ |
| 83/84 | Q9Y; Q28S; S35H; V87A; S185F; I197L; E198L; D216N; N266T;             | 9  | +++++ |
| 85/86 | Q9Y; Q28S; S35H; V87A; S185F; I197L; E198L; D216N;                    | 8  | +++++ |
| 87/88 | Q9Y; S35H; L36M; F37G; V87A; S185F; E198L; D216N; L280M; L290I        | 10 | +++++ |
| 89/90 | Q9Y; F30V; Q33W; S35H; P75R; V87A; D103W; S185F; E198L; D216N;        | 10 | +++++ |

# ES 3 018 269 T3

|         |                                                                                                           |    |        |
|---------|-----------------------------------------------------------------------------------------------------------|----|--------|
| 91/92   | Q9Y; P20I; F30V; Q33W; S35H; A46R; R66T; V87A; D103W; S185F; E198L; D216N; E263R;                         | 13 | +++++  |
| 93/94   | Q9Y; Y22W; S35H; P75R; V87A; S185F; E198L; D216N; L290I;                                                  | 9  | +++++  |
| 95/96   | Q9Y; S35H; A46K; V87A; L160T; S185F; E198L; D216N; Q286N;                                                 | 9  | \$     |
| 97/98   | Q9Y; S35H; A46K; V87A; L160T; S185F; E198L; D216N; Q286S;                                                 | 9  | \$     |
| 99/100  | Q9Y; Q28C; S35H; P38E; W85G; V87A; S185F; E198L; D216N;                                                   | 9  | \$     |
| 101/102 | Q9Y; R66L; V86T; V87A; V96M; L160P; Y183Q; S185F; Y188F; E198L; L205F; L212D; D216N; R255N; V270L; Q286S; | 16 | \$     |
| 103/104 | Q9Y; R66L; V86T; V87A; V96M; S185F; E198L; L205F; D216N; A249W; Q286S;                                    | 11 | \$     |
| 105/106 | Q9Y; Q28C; S35H; D54R; P75R; V87A; I176M; S185F; E198L; L205F; D216N; T248V;                              | 12 | \$     |
| 107/108 | I7L; Q9Y; Q28C; S35H; P38E; W85G; V87A; S185F; E198L; D216N;                                              | 10 | \$\$   |
| 109/110 | Q9Y; Q28C; S35H; P38E; A46K; W85G; V87A; S185F; E198L; D216N;                                             | 10 | \$\$   |
| 111/112 | Q9Y; Q28C; S35H; L36T; P38W; D54Q; R66T; W85G; V87A; S185F; E198L; D216N;                                 | 12 | \$\$   |
| 113/114 | Q9Y; Q28C; S35H; L36T; P38W; P75R; W85G; V87A; S185F; E198L; D216N;                                       | 11 | \$\$   |
| 115/116 | Q9Y; Y22W; Q28C; S35H; L36M; P38W; P75R; W85G; V87A; S185F; E198L; D216N;                                 | 12 | \$\$   |
| 17/118  | Q9Y; Q28C; S35H; L36T; P38W; P75M; W85G; V87A; S185F; E198L; D216N;                                       | 11 | \$\$   |
| 119/120 | Q9Y; Q28C; S35H; P38E; W85G; V87A; S185F; E198L; D216N;                                                   | 9  | \$\$   |
| 121/122 | Q9Y; Q28C; S35H; L36T; P38W; P75R; W85G; V87A; S185F; E198L; D216N; Q286C;                                | 12 | \$\$\$ |
| 123/124 | P2L; Q9Y; Q28C; S35H; L36T; P38W; P75R; W85G; V87A; S185F; E198L; D216N;                                  | 12 | \$\$\$ |
| 125/126 | Q9Y; Q28C; S35H; L36T; P38W; P75R; W85G; V87A; L181Q; S185F; E198L; D216N;                                | 12 | \$\$\$ |

5 Como se señala anteriormente, en algunas realizaciones, el polipéptido de carboxiesterasa mejorado comprende una secuencia de aminoácidos que es al menos aproximadamente 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o más idéntica a una secuencia de referencia de SEQ ID NO: 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, o 126. En algunas realizaciones, los polipéptidos de carboxiesterasa mejorados pueden tener 1-2, 1-3, 1-4, 1-5, 1-6, 1-7, 1-8, 1-9, 1-10, 1-11, 1-12, 1-13, 1-14, 1-15, 1-16, 1-17, 1-18, 1-19, 1-20, 1-22, 1-24, 1-26, 1-28, 1-30, 1-35, 1-40, 1-45, 1-50, 1-55, 1-60, o 1-62 diferencias de residuos en comparación con la carboxiesterasa representada por SEQ ID NO: 2. En algunas realizaciones, el número de diferencias de residuos puede ser 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 22, 24, 26, 28, 30, 35, 40, 45, 50, 55, 60 o 62 diferencias en comparación con SEQ ID NO: 2.

15 En algunas realizaciones, el polipéptido de carboxiesterasa mejorado comprende una secuencia de aminoácidos que es al menos aproximadamente 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 8% o 99% idéntica a una secuencia de referencia basada en SEQ ID NO: 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, o 126, con la condición de que la secuencia mejorada de aminoácidos de carboxiesterasa comprenda cualquiera del conjunto de diferencias de residuos contenidas en cualquiera de las secuencias de polipéptidos listadas en la Tabla 3, en comparación con SEQ ID NO: 2. En algunas realizaciones, los polipéptidos de carboxiesterasa mejorados pueden tener adicionalmente 1-2, 1-3, 1-4, 1-5, 1-6, 1-7, 1-8, 1-9, 1-10, 1-11, 1-12, 1-14, 1-15, 1-16, 1-18, 1-20, 1-22, 1-24,

1-26, 1-30, 1-35, 1-40, 1-45, 1-50, 1-55, 1-60, o 1-62 diferencias de residuos en otras posiciones de residuos de aminoácidos en comparación con la secuencia de referencia. En algunas realizaciones, el número de diferencias puede ser 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 18, 20, 22, 24, 26, 30, 35, 40, 45, 50, 55 o 62 diferencias de residuos en otras posiciones de residuos.

5

En algunas realizaciones, las diferencias de residuos en otras posiciones de residuos comprenden sustituciones con residuos de aminoácidos conservadores. En algunas realizaciones, los polipéptidos de carboxiesterasa mejorados capaces de convertir el sustrato de éster, oxazol-5-carboxilato de etilo en presencia de un sustrato de amina a niveles de producto detectables por HPLC-UV a 230 nm en MIBK saturado con agua comprende una secuencia de aminoácidos elegida de la secuencia de SEQ ID NO: 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, o 126.

10

En algunas realizaciones, el polipéptido de carboxiesterasa modificado genéticamente es capaz de convertir el sustrato de éster en producto con una actividad 2 veces, 5 veces, 10 veces, 20 veces, 25 veces, 50 veces, 75 veces, 100 veces, 1000 veces, 10 000 veces, 100 000 veces, 500 000 veces, 785 000 veces o mayor que el polipéptido de SEQ ID NO: 2. En algunas realizaciones, el polipéptido de carboxiesterasa modificado genéticamente es capaz de convertir el sustrato de éster en un producto con una actividad de 50 a 100 veces o mayor que el polipéptido de SEQ ID NO: 2 y comprende una secuencia de aminoácidos correspondiente a SEQ ID NO: 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, o 126.

15

20

En algunas realizaciones, el polipéptido de carboxiesterasa modificado genéticamente es capaz de convertir el sustrato de éster en producto con aproximadamente 1,1 a 5 veces o más actividad que el polipéptido de SEQ ID NO: 24. En algunas realizaciones, el polipéptido de carboxiesterasa modificado genéticamente es capaz de convertir el sustrato de éster en producto con aproximadamente 1,1 a 5 veces o más actividad que el polipéptido de SEQ ID NO: 24 y comprende una secuencia de aminoácidos correspondiente a la secuencia de SEQ ID NO: 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, o 126.

25

30

En algunas realizaciones, el polipéptido de carboxiesterasa modificado genéticamente es capaz de convertir el sustrato de éster en un producto con aproximadamente 1,1 a 5 veces o más actividad que el polipéptido de SEQ ID NO: 54. En algunas realizaciones, el polipéptido de carboxiesterasa modificado genéticamente es capaz de convertir el sustrato de éster en un producto con aproximadamente 1,1 a 5 veces o más actividad que el polipéptido de SEQ ID NO: 54 y comprende una secuencia correspondiente a la secuencia de SEQ ID NO: 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, o 126.

35

En algunas realizaciones, el polipéptido de carboxiesterasa modificado genéticamente es capaz de convertir el sustrato de éster en producto con aproximadamente 1,1 a 6 veces o más actividad que el polipéptido de SEQ ID NO: 68. En algunas realizaciones, el polipéptido de carboxiesterasa modificado genéticamente es capaz de convertir el sustrato de éster en producto con aproximadamente 1,1 a 5 veces o más actividad que el polipéptido de SEQ ID NO: 68 y comprende una secuencia de aminoácidos correspondiente a la secuencia de SEQ ID NO: 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, o 126.

40

45

En algunas realizaciones, el polipéptido de carboxiesterasa modificado genéticamente es capaz de convertir el sustrato de éster en producto con aproximadamente 1,1 a 5 veces o más actividad que el polipéptido de SEQ ID NO: 100. En algunas realizaciones, el polipéptido de carboxiesterasa modificado genéticamente es capaz de convertir el sustrato de éster en un producto con aproximadamente 1,7 veces o más actividad que el polipéptido de SEQ ID NO: 100 y comprende una secuencia de aminoácidos correspondiente a la secuencia de SEQ ID NO: 108, 110, 112, 114, 116, 118, 120, 122, 124, o 126.

50

En algunas realizaciones, el polipéptido de carboxiesterasa modificado genéticamente es capaz de convertir el sustrato de éster en un producto con aproximadamente 1,1 a 2 veces mayor actividad que el polipéptido de SEQ ID NO: 114. En algunas realizaciones, el polipéptido de carboxiesterasa modificado genéticamente es capaz de convertir el sustrato de éster en producto con aproximadamente 1,1 a 5 veces o más actividad que el polipéptido de SEQ ID NO: 114 y comprende una secuencia de aminoácidos correspondiente a la secuencia de SEQ ID NO: 122, 124, o 126.

55

60

En algunas realizaciones, los polipéptidos de carboxiesterasa modificados genéticamente mejorados pueden comprender delecciones en residuos de aminoácidos específicos de los polipéptidos de carboxiesterasa modificados genéticamente descritos en la presente. Por lo tanto, para todas y cada una de las realizaciones de los polipéptidos de carboxiesterasa de la divulgación, las eliminaciones pueden comprender uno o más

aminoácidos, 2 o más aminoácidos, 3 o más aminoácidos, 4 o más aminoácidos, 5 o más aminoácidos, 6 o más aminoácidos, 8 o más aminoácidos, 10 o más aminoácidos, 15 o más aminoácidos, o 20 o más aminoácidos, hasta el 5% de la cantidad total de aminoácidos, hasta el 10% de la cantidad total de aminoácidos, hasta el 20% de la cantidad total de aminoácidos, o hasta el 30% de la cantidad total de aminoácidos de los polipéptidos de carboxiesterasa, siempre que se mantenga la actividad funcional de la actividad de carboxiesterasa. En algunas realizaciones, las delecciones pueden comprender hasta 1-2, 1-3, 1-4, 1-5, 1-6, 1-7, 1-8, 1-9, 1-10, 1-11, 1-12, 1-14, 1-15, 1-16, 1-18, 1-20, 1-22, 1-24, 1-26, 1-30, 1-35, 1-40, 1-45, 1-50, 1-55, 1-60, o 1-62 residuos de aminoácidos. En algunas realizaciones, el número de delecciones puede ser de hasta 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 18, 20, 22, 24, 26, 30, 35, 40, 45, 50, 55, 60 o 62 aminoácidos. Como se describe en la presente, los polipéptidos de carboxiesterasa de la divulgación pueden estar en la forma de polipéptidos de fusión en los que los polipéptidos de carboxiesterasa se fusionan con otros polipéptidos, tal como, a modo de ejemplo y no de limitación, etiquetas de anticuerpo (por ejemplo, epítipo myc), secuencias de purificación (por ejemplo, etiquetas His para la unión a metales) y señales de localización celular (por ejemplo, señales de secreción). Por lo tanto, los polipéptidos de carboxiesterasa se pueden usar con o sin fusiones a otros polipéptidos.

Los polipéptidos descritos en la presente no se restringen a los aminoácidos codificados genéticamente. Además de los aminoácidos codificados genéticamente, los polipéptidos descritos en la presente pueden comprender, ya sea en su totalidad o en parte, aminoácidos no codificados de origen natural y/o sintéticos. Determinados aminoácidos no codificados comúnmente encontrados de los cuales los polipéptidos descritos en la presente pueden estar comprendidos incluyen, pero no se limitan a: los D-estereoisómeros de los aminoácidos codificados genéticamente; ácido 2,3-diaminopropiónico (Dpr); ácido  $\alpha$ -aminoisobutírico (Aib); ácido  $\epsilon$ -aminohexanoico (Aha); ácido  $\delta$ -aminovalérico (Ava); N-metilglicina o sarcosina (MeGly o Sar); ornitina (Orn); citrulina (Cit); t-butilalanina (Bua); t-butilglicina (Bug); N-metili soleucina (Melle); fenilglicina (Phg); ciclohexilalanina (Cha); norleucina (Nle); naftilalanina (Nal); 2-clorofenilalanina (Ocf); 3-clorofenilalanina (Mcf); 4-clorofenilalanina (Pcf); 2-fluorofenilalanina (Of); 3-fluorofenilalanina (Mff); 4-fluorofenilalanina (Pff); 2-bromofenilalanina (Obf); 3-bromofenilalanina (Mbf); 4-bromofenilalanina (Pbf); 2-metilfenilalanina (Omf); 3-metilfenilalanina (Mmf); 4-metilfenilalanina (Pmf); 2-nitrofenilalanina (Onf); 3-nitrofenilalanina (Mnf); 4-nitrofenilalanina (Pnf); 2-cianofenilalanina (Ocf); 3-cianofenilalanina (Mcf); 4-cianofenilalanina (Pcf); 2-trifluorometilfenilalanina (Otf); 3-trifluorometilfenilalanina (Mtf); 4-trifluorometilfenilalanina (Ptf); 4-aminofenilalanina (Paf); 4-yodofenilalanina (Pif); 4-aminometilfenilalanina (Pamf); 2,4-diclorofenilalanina (Opef); 3,4-diclorofenilalanina (Mpcf); 2,4-difluorofenilalanina (Opff); 3,4-difluorofenilalanina (Mpff); pirid-2-ilalanina (2pAla); pirid-3-ilalanina (3pAla); pirid-4-ilalanina (4pAla); naft-1-ilalanina (1nAla); naft-2-ilalanina (2nAla); tiazolilalanina (taAla); benzotienilalanina (bAla); tienilalanina (tAla); furilalanina (fAla); homofenilalanina (hPhe); homotirosina (hTyr); homotriptófano (hTrp); pentafluorofenilalanina (5ff); estililalanina (sAla); autrilalanina (aAla); 3,3-difenilalanina (Dfa); ácido 3-amino-5-fenipentanoico (Afp); penicilamina (Pen); ácido 1,2,3,4-tetrahidroisoquinolin-3-carboxílico (Tic);  $\beta$ -2-tienilalanina (Thi); sulfóxido de metionina (Mso); N(w)-nitroarginina (nArg); homolisina (hLys); fosfonometilfenilalanina (pmPhe); fosfoserina (pSer); fosfotreonina (pThr); ácido homoaspártico (hAsp); ácido homoglutámico (hGlu); ácido 1-aminociclopent-(2 o 3)-en-4-carboxílico; ácido pipercolico (PA), ácido azetidín-3-carboxílico (ACA); ácido 1-aminociclopentan-3-carboxílico; alilglicina (aOly); propargilglicina (pgGly); homoalanina (hAla); norvalina (nVal); homoleucina (hLeu), homovalina (hVal); homoisoleucina (hIle); homoarginina (hArg); N-acetil lisina (AcLys); ácido 2,4-diaminobutírico (Dbu); ácido 2,3-diaminobutírico (Dab); N-metilvalina (MeVal); homocisteína (hCys); homoserina (hSer); hidroxiprolina (Hyp) y homoprolina (hPro). Los aminoácidos no codificados adicionales de los que pueden comprender los polipéptidos descritos en la presente serán evidentes para los expertos en la técnica (ver, por ejemplo, los diversos aminoácidos proporcionados en Fasman, 1989, *CRC Practical Handbook of Biochemistry and Molecular Biology*, CRC Press, Boca Raton, FL, en pp. 3-70 y las referencias citadas en el mismo). Estos aminoácidos pueden estar en la configuración L o D.

Los expertos en la técnica reconocerán que los aminoácidos o residuos que portan grupos protectores de cadena lateral también pueden estar comprendidos en los polipéptidos descritos en la presente. Los ejemplos no limitantes de estos aminoácidos protegidos, que en este caso pertenecen a la categoría aromática, incluyen (grupos protectores listados entre paréntesis), pero no se limitan a: Arg(tos), Cys(metilbencilo), Cys(nitropiridinsulfenilo), Glu( $\delta$ -benciléster), Gln(xantilo), Asn(N-xantilo $\delta$ ), His(bom), His(bencilo), His(tos), Lys(fmoc), Lys(tos), Ser(O-bencilo), Thr(O-bencilo) y Tyr(O-bencilo).

Los aminoácidos no codificantes que están restringidos conformacionalmente de los cuales pueden estar compuestos los polipéptidos descritos en la presente incluyen, pero no se limitan a, N-metil aminoácidos (configuración L); ácido 1-aminociclopent-(2 o 3)-en-4-carboxílico; ácido pipercolico; ácido azetidín-3-carboxílico; homoprolina (hPro); y ácido 1-aminociclopentan-3-carboxílico.

Como se describió anteriormente, las diversas modificaciones introducidas en el polipéptido de origen natural para generar una enzima de carboxiesterasa modificada pueden dirigirse para afectar una propiedad específica de la enzima, tal como actividad, especificidad para su sustrato y termoestabilidad, etc.

En otro aspecto, la presente divulgación proporciona polinucleótidos que codifican para los polipéptidos de

carboxiesterasa mejorados. Los polinucleótidos pueden estar unidos operativamente a una o más secuencias reguladoras heterólogas que controlan la expresión génica para crear un polinucleótido recombinante capaz de expresar el polipéptido de carboxiesterasa. Las construcciones de expresión que contienen un polinucleótido heterólogo que codifica para la carboxiesterasa modificada se pueden introducir en células hospedadoras apropiadas para expresar el polipéptido de carboxiesterasa correspondiente.

Debido al conocimiento de los codones correspondientes a los diversos aminoácidos, la disponibilidad de una secuencia proteica proporciona una descripción de todos los polinucleótidos capaces de codificar para el sujeto. La degeneración del código genético, donde los mismos aminoácidos están codificados por codones alternativos o sinónimos, permite que se produzca una cantidad extremadamente grande de ácidos nucleicos, todos los cuales codifican para los polipéptidos de carboxiesterasa mejorados divulgados en la presente. Por lo tanto, después de haber identificado una secuencia de aminoácidos particular, los expertos en la técnica podrían producir cualquier cantidad de ácidos nucleicos diferentes simplemente modificando la secuencia de uno o más codones de una manera que no cambie la secuencia de aminoácidos de la proteína. En este sentido, la presente divulgación contempla específicamente todas y cada una de las posibles variaciones de polinucleótidos que podrían realizarse seleccionando combinaciones basadas en las posibles elecciones de codones, y todas estas variaciones deben considerarse específicamente divulgadas para cualquier polipéptido divulgado en la presente, incluidas las secuencias de aminoácidos presentadas en la Tabla 3.

En algunas realizaciones, los polinucleótidos se pueden seleccionar y/o modificar genéticamente para que comprendan codones que se seleccionan para que se ajusten a la célula hospedadora en la que se produce la proteína. Por ejemplo, los codones preferidos usados en bacterias se usan para expresar el gen en bacterias; los codones preferidos usados en levaduras se usan para la expresión en levaduras; y los codones preferidos usados en mamíferos se usan para la expresión en células de mamíferos. Dado que no es necesario reemplazar todos los codones para optimizar el uso de codones de la carboxiesterasa (por ejemplo, debido a que la secuencia natural puede tener codones preferidos y debido a que el uso de codones preferidos puede no ser necesario para todos los residuos de aminoácidos), los polinucleótidos con codones optimizados que codifican para los polipéptidos de carboxiesterasa pueden contener codones preferidos en aproximadamente 40%, 50%, 60%, 70%, 80% o más del 90% de las posiciones de codones de la región codificante de longitud completa.

En algunas realizaciones, el polinucleótido codifica para un polipéptido de carboxiesterasa que comprende una secuencia de aminoácidos que es al menos 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o más idéntica a la secuencia de referencia de SEQ ID NO:4, o un fragmento funcional de la misma, en donde el polipéptido es capaz de convertir el sustrato de éster, en presencia de un sustrato de amina con una actividad que mejora en comparación con la actividad de la carboxiesterasa de SEQ ID NO: 2 derivado de *Esterasa 2* de *A. acidocaldarius*.

En algunas realizaciones, el polinucleótido codifica para un polipéptido de carboxiesterasa que comprende una secuencia de aminoácidos que tiene al menos aproximadamente 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% o 99% o más de identidad de secuencia con el polipéptido que comprende una secuencia de aminoácidos correspondiente a SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, o 126, o un fragmento funcional del mismo, en donde el polipéptido tiene al menos una propiedad mejorada en la conversión del sustrato de éster, oxazol-5-carboxilato de etilo al producto, (4-isopropilpiperazin-1-il)(oxazol-5-il)metanona en presencia de un sustrato de amina, 1-isopropil-piperazina. En algunas realizaciones, el polipéptido de carboxiesterasa codificado tiene una actividad que es igual o mayor que la actividad del polipéptido de SEQ ID NO: 2.

En algunas realizaciones, el polinucleótido codifica para un polipéptido de carboxiesterasa que comprende una secuencia de aminoácidos que es al menos aproximadamente 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% o 99% idéntica a una secuencia de referencia basada en SEQ ID NO: 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, o 126, o un fragmento funcional del mismo, con la condición de que la secuencia de aminoácidos de carboxiesterasa mejorada comprenda cualquiera del conjunto de diferencias de residuos contenidas en cualquiera de las secuencias de polipéptidos listadas en la Tabla 3, en comparación con SEQ ID NO: 2.

En algunas realizaciones, los polinucleótidos que codifican para los polipéptidos de carboxiesterasa mejorados se seleccionan de SEQ ID NO: 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, o 125.

En algunas realizaciones, los polinucleótidos son capaces de hibridarse bajo condiciones altamente rigurosas con un polinucleótido que comprende SEQ ID NO: 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39,

41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, o 125, o un complemento del mismo, donde los polinucleótidos de hibridación altamente rigurosa codifican para un polipéptido de carboxiesterasa capaz de convertirse en producto en presencia de un sustrato de amina con una actividad que es igual o mayor que el polipéptido de SEQ ID NO:2.

En algunas realizaciones, los polinucleótidos codifican para los polipéptidos descritos en la presente, pero tienen aproximadamente 80% o más de identidad de secuencia, aproximadamente 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91 %, 92%, 93%, 94%, 95%, 96%, 97%, 98% o 99% o más de identidad de secuencia a nivel de nucleótidos con un polinucleótido de referencia que codifica para la carboxiesterasa modificada descrita en la presente. En algunas realizaciones, el polinucleótido de referencia se selecciona de SEQ ID NO: 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, o 125.

En algunas realizaciones, el polipéptido de carboxiesterasa comprende la secuencia de aminoácidos expuesta en SEQ ID NO: 122. En otra realización, la presente divulgación proporciona una secuencia de polinucleótidos que codifica para la secuencia de polipéptidos de carboxiesterasa expuesta en SEQ ID NO: 122. En aun otra realización, la presente divulgación proporciona un polinucleótido que codifica para un polipéptido de carboxiesterasa, en donde el polinucleótido comprende la secuencia de polinucleótidos expuesta en SEQ ID NO: 121. En aun otra realización, el polipéptido de carboxiesterasa consiste en la secuencia de polipéptidos expuesta en SEQ ID NO: 122. En otra realización, el polipéptido de carboxiesterasa consiste en los residuos 2-310 de SEQ ID NO: 122.

Las carboxiesterasas y polinucleótidos mejorados que codifican para estos polipéptidos se pueden preparar usando métodos comúnmente usados por los expertos en la técnica. Como se señaló anteriormente, la secuencia de aminoácidos de origen natural y el polinucleótido correspondiente que codifica para la enzima de carboxiesterasa tipo silvestre, Esterasa 2 de *A. acidocaldarius*, de la cual la secuencia original, SEQ ID NO: 2, está disponible en WO02/057411 (ver SEQ ID NO: 10). En algunas realizaciones, la secuencia de polinucleótidos original tiene codones optimizados para mejorar la expresión de la carboxiesterasa en una célula hospedadora específica. Las carboxiesterasas modificadas genéticamente se pueden obtener al someter el polinucleótido que codifica para la carboxiesterasa de origen natural a métodos de mutagénesis y/o evolución dirigida. Una técnica de evolución dirigida de ejemplo es la mutagénesis y/o el reordenamiento de ADN, como se describe en Stemmer, 1994, *Proc Natl Acad Sci USA* 91:10747-10751; WO 95/22625; WO 97/0078; WO 97/35966; WO 98/27230; WO 00/42651; WO 01/75767; y la patente de los Estados Unidos 6,537,746.

Otros procedimientos de evolución dirigida que se pueden usar incluyen, entre otros, el proceso de extensión escalonada (StEP), la recombinación *in vitro* (Zhao, *et al.*, 1998, *Nat. Biotechnol.* 16:258-261), PCR mutagénica (Caldwell, *et al.*, 1994, *PCR Methods Appl.* 3:S136-S140), y mutagénesis de casete (Black, *et al.*, 1996, *Proc Natl Acad Sci USA* 93:3525-3529). Las técnicas de mutagénesis y evolución dirigida útiles para los fines de la presente también se describen en las siguientes referencias: Ling, *et al.*, 1997, "Approaches to DNA mutagenesis: an overview", *Anal. Biochem.* 254(2):157-78; Dale, *et al.*, 1996, "Oligonucleotide-directed random mutagenesis using the phosphorothioate method," *Methods Mol. Biol.* 57:369-74; Smith, 1985, "In vitro mutagenesis," *Ann. Rev. Genet.* 19:423-462; Botstein, *et al.*, 1985, "Strategies and applications of *in vitro* mutagenesis", *Science* 229:1193-1201; Carter, 1986, "Site-directed mutagenesis", *Biochem. J.* 237:1-7; Kramer, *et al.*, 1984, "Point Mismatch Repair," *Cell* 38:879-887; Wells, *et al.*, 1985, "Cassette mutagenesis: an efficient method for generation of multiple mutations at defined sites," *Gene* 34:315-323; Minshull, *et al.*, 1999, "Protein evolution by molecular breeding," *Curr Opin Chem Biol* 3:284-290; Christians, *et al.*, 1999, "Directed evolution of thymidine kinase for AZT phosphorylation using DNA family shuffling," *Nature Biotech* 17:259-264; Cramer, *et al.*, 1998, "DNA shuffling of a family of genes from diverse species accelerates directed evolution," *Nature* 391:288-291; Cramer, *et al.*, 1997, "Molecular evolution of an arsenate detoxification pathway by DNA shuffling," *Nature Biotech* 15:436-438; Zhang, *et al.*, 1997, "Directed evolution of an effective fructosidase from a galactosidase by DNA shuffling and screening," *Proc Natl Acad Sci USA* 94:45-4-4509; Cramer, *et al.*, 1996, "Improved green fluorescent protein by molecular evolution using DNA shuffling," *Nature Biotech* 14:315-319; y Stemmer, 1994, "Rapid evolution of a protein *in vitro* by DNA shuffling," *Nature* 370:389-391.

En algunas realizaciones, los clones obtenidos después del tratamiento de mutagénesis se seleccionan para carboxiesterasas que tienen una propiedad enzimática mejorada deseada. La medición de la actividad de la enzima de carboxiesterasa de las bibliotecas de expresión se puede realizar mediante el uso de técnicas estándar, tal como la separación del producto (por ejemplo, mediante HPLC) y la detección del producto mediante la medición de la absorbancia UV del sustrato y los productos separados y/o mediante la detección mediante el uso de espectroscopía de masas en tándem (por ejemplo, MS/MS). Los ensayos de ejemplo se describen en el Ejemplo 4 a continuación. La velocidad de aumento en el producto deseado por unidad de tiempo indica la actividad relativa (enzimática) del polipéptido de carboxiesterasa en una cantidad fija del lisado (o un polvo liofilizado hecho de este). Cuando la propiedad enzimática mejorada deseada es la estabilidad térmica, la actividad enzimática se puede

medir después de someter las preparaciones enzimáticas a una temperatura definida y medir la cantidad de actividad enzimática restante después de los tratamientos térmicos. Luego se aíslan los clones que contienen un polinucleótido que codifica para las carboxiesterasas deseadas, se secuencian para identificar los cambios en la secuencia de nucleótidos (si los hay) y se usan para expresar la enzima en una célula hospedadora.

5

En otras realizaciones, que son bien conocidas en la técnica, las enzimas se pueden diversificar genéticamente mientras se mantienen sus actividades diana, tal como mediante la técnica de deriva neutra.

10

Cuando se conoce la secuencia del polipéptido modificado genéticamente, los polinucleótidos que codifican para la enzima se pueden preparar mediante métodos estándar en fase sólida, de acuerdo con métodos sintéticos conocidos. En algunas realizaciones, los fragmentos de hasta aproximadamente 100 bases se pueden sintetizar individualmente, luego unirse (por ejemplo, mediante métodos de litigio enzimático o químico, o métodos mediados por polimerasa) para formar cualquier secuencia continua, deseada. Por ejemplo, los polinucleótidos y oligonucleótidos de la invención se pueden preparar mediante síntesis química usando, por ejemplo, el método clásico de fosforamidita descrito por Beaucage, *et al.*, 1981, *Tet Lett* 22:1859-69, o el método descrito por Matthes, *et al.*, 1984, *EMBO J.* 3:801-05, por ejemplo, como se practica habitualmente en métodos sintéticos automatizados. De acuerdo con el método de fosforamidita, los oligonucleótidos se sintetizan, por ejemplo, en un sintetizador automático de ADN, se purifican, se hibridan, se ligan y se clonan en vectores apropiados. Además, esencialmente cualquier ácido nucleico puede obtenerse de cualquiera de una variedad de fuentes comerciales, The Great American Gene Company, Ramona, CA, ExpressGen Inc, Chicago, IL, Operon Technologies Inc, Alameda, CA y muchos otros.

15

20

Las enzimas carboxiesterasa modificadas genéticamente expresadas en una célula hospedadora se pueden recuperar de las células y /o el medio de cultivo usando una o más de las técnicas conocidas para la purificación de proteínas, que incluyen, entre otros, tratamiento con lisozima, tratamiento con sonido, filtración, precipitación salina, ultracentrifugación y cromatografía. Las soluciones adecuadas para la lisis y la extracción de alta eficiencia de proteínas de bacterias, tal como *E. coli.*, están disponibles comercialmente bajo el nombre comercial CellLytic BTM de Sigma-Aldrich de St. Louis, MO.

25

30

Las técnicas cromatográficas para el aislamiento del polipéptido de carboxiesterasa incluyen, entre otras, cromatografía de fase inversa, cromatografía líquida de alto rendimiento, cromatografía de intercambio iónico, electroforesis en gel y cromatografía de afinidad. Las condiciones para purificar una enzima particular dependerán, en parte, de factores tal como la carga neta, la hidrofobicidad, la hidrofiliidad, el peso molecular, la forma molecular, etc., y serán evidentes para los expertos en la técnica. En algunas realizaciones, las carboxiesterasas modificadas genéticamente se pueden expresar como proteínas de fusión con etiquetas de purificación, tal como etiquetas His que tienen afinidad por los metales, o etiquetas de anticuerpo para la unión a anticuerpos, por ejemplo, etiqueta de epitopo myc.

35

40

En algunas realizaciones, se pueden usar técnicas de afinidad para aislar las enzimas carboxiesterasa mejoradas. Para la purificación por cromatografía de afinidad, se puede usar cualquier anticuerpo que se una específicamente al polipéptido de carboxiesterasa. Para la producción de anticuerpos, se pueden inmunizar varios animales hospedadores, que incluyen, pero no se limitan a, conejos, ratones, ratas, etc., mediante inyección con un polipéptido modificado genéticamente. El polipéptido puede estar unido a un portador adecuado, tal como BSA, por medio de un grupo funcional de cadena lateral o enlazadores unidos a un grupo funcional de cadena lateral. Se pueden usar varios adyuvantes para aumentar la respuesta inmunológica, dependiendo de la especie hospedadora, que incluye, pero no se limita a, de Freund (completo e incompleto), geles minerales, tal como hidróxido de aluminio, sustancias tensioactivas, tal como lisolecitina, polioles plurónicos, polianiones, péptidos, emulsiones de aceite, hemocianina de lapa californiana, dinitrofenol y adyuvantes humanos potencialmente útiles, tal como BCG (*bacilo de Calmette Guérin*) y *Corynebacterium parvum*.

50

En un aspecto adicional, los polipéptidos de carboxiesterasa mejorados descritos en la presente se pueden usar en un proceso para la amidación de determinados aceptores de grupos amida (por ejemplo, un sustrato de éster) en presencia de un sustrato de amina.

55

En algunas realizaciones, las carboxiesterasas mejoradas se pueden usar en un proceso para preparar una amida, en donde los componentes se combinan y contienen:

60

(a) un éster de la forma  $R_1-COOR_2$ , en donde  $R_1$  se elige de: un carbono  $sp^3$  con 0 a 3 sustituyentes alquilo; y un anillo aromático, y  $R_2$  se elige de: un grupo metilo; un grupo etilo; y cadenas alquilo de 1-6 carbonos;

(b) un sustrato de amina;

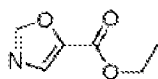
(c) un polipéptido de carboxiesterasa mejorado, y;

(d) un disolvente.

En una realización, el disolvente es un disolvente orgánico que contiene hasta 3 equivalentes molares de agua, con respecto al sustrato de éster, en una cantidad de aproximadamente 0,5% (vol/vol) a aproximadamente 3% (vol/vol).

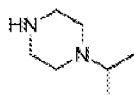
En algunas realizaciones del proceso, las carboxiesterasas mejoradas se eligen de: SEQ ID NO: 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, y 126.

En otras realizaciones, los polipéptidos de carboxiesterasa mejorados se pueden usar en un proceso para preparar una molécula de amida, en donde los componentes se combinan y contienen: (a) un sustrato de éster de la forma  $R_1-COOR_2$ , en donde  $R_1$  se elige de: un carbono  $sp^3$  con 0 a 3 sustituyentes alquilo; y un anillo aromático, y  $R_2$  se elige de: un grupo metilo; un grupo etilo y una cadena de alquilo de 1-6 carbonos; (b) un sustrato de amina; (c) un polipéptido de carboxiesterasa mejorado descrito anteriormente; y (d) un disolvente que en otra realización de este proceso, se usa un disolvente orgánico. En una realización, el disolvente orgánico se elige de: tolueno; 2-metiltetrahidrofurano; tetrahidrofurano; dimetilacetamida; metil-isobutil-cetona (MIBK); diclorometano; éter metílico de ter-butilo; éter metílico de ciclopentilo; metil ciclohexano; diclorometano; acetonitrilo; metil etil cetona; acetato de isopropilo; etanol; isopropanol; acetato de etilo; heptano; xantano; y 2-metiltetrahidrofurano (2-Me-THF); y agua. En aun otra realización, el disolvente orgánico contiene hasta 3 equivalentes molares de agua con respecto al sustrato de éster en una cantidad de aproximadamente 0,5% (vol/vol) a aproximadamente 3% (vol/vol). En otra realización de este proceso, el polipéptido de carboxiesterasa en el paso (c) se prepara en presencia de una sal para estabilizar su forma física durante la reacción. En aun otra realización, la sal se adiciona como un componente de reacción adicional. En una realización de este proceso, el éster es oxazol-5-carboxilato de etilo que tiene la fórmula:



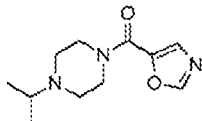
I.

el sustrato de amina es 1-isopropil-piperazina que tiene la fórmula



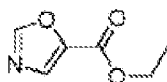
II.

y la amida es (4-isopropilpiperazin-1-il)(oxazol-5-il)metanona que tiene la fórmula:



III.

En aun otra realización de este proceso, el éster es oxazol-5-carboxilato de etilo que tiene la fórmula:



I.

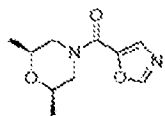
y el sustrato de amina es *cis*-2,6-dimetilmorfolina que tiene la fórmula:



IV;

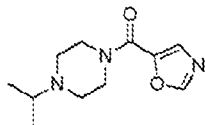
5

y la amida es ((2S,6R)-2,6-dimetilmorfolino)(oxazol-5-il)metanona que tiene la fórmula:



V.

- 10 En otra realización, esta reacción en este proceso comprende: aproximadamente 50 g/L de oxazol-5-carboxilato de etilo, 44 g/L de 1-isopropil-piperazina y aproximadamente 25 g/L de un polipéptido de carboxiesterasa correspondiente a una secuencia de aminoácidos de: SEQ ID NO: 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, o 126, en donde la
- 15 carboxiesterasa se prepara en presencia de sulfato de sodio y se procesa en presencia de aproximadamente 10 g/L a aproximadamente 20 g/L de agua en MIBK. En algunas realizaciones, la invención es una amida que se elabora mediante estos procesos usando las carboxiesterasas mejoradas. En otra realización, la invención es una amida, 4-isopropilpiperazin-1-il)(oxazol-5-il)metanona, que tiene la fórmula:

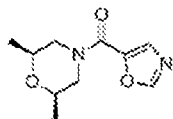


III.,

20

en donde la amida de la fórmula III. se prepara mediante los procesos descritos anteriormente.

- 25 En una realización alternativa, la invención es una amida, ((2S,6R)-2,6-dimetilmorfolino)(oxazol-5-il)metanona, que tiene la fórmula:



V.,

- 30 en donde la amida de la fórmula V, se prepara mediante los procesos descritos anteriormente.

Los compuestos de la fórmula I (Astatech, AS-23210), fórmula II (Oakwood Products Inc, OAK-008910) y fórmula IV (Oakwood Products Inc, OAK-091224), se adquirieron de proveedores comerciales.

- 35 En algunas realizaciones, el proceso comprende poner en contacto o incubar el sustrato de éster, oxazol-5-carboxilato de etilo, con una carboxiesterasa mejorada en presencia de un sustrato de amina bajo condiciones de reacción adecuadas para convertir el sustrato de éster en el producto, (4-isopropilpiperazin-1-il)(oxazol-5-

il)metanona, con una velocidad de conversión y/o actividad de aproximadamente 50 a aproximadamente 785 000 veces o más que la de SEQ ID NO: 2. Los ejemplos de polipéptidos comprenden una secuencia de aminoácidos correspondiente a SEQ ID NO: 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, o 126.

En algunas realizaciones de los procesos anteriores, el disolvente de reacción para llevar a cabo el proceso se elige de metil-isobutil-cetona (MIBK), tolueno, éter metílico de ter-butilo (TBME) o 2-metil tetrahidrofurano (2-Me-THF).

En algunas realizaciones de los procesos anteriores, la preparación enzimática para la reacción incluye una sal elegida de uno de fosfato de potasio (KPi), sulfato de potasio o sulfato de sodio.

En algunas realizaciones, la condición de reacción para llevar a cabo el proceso puede comprender una temperatura de aproximadamente 15°C a una temperatura de aproximadamente 30°C. En una realización, el sustrato de amina usado en el proceso puede ser una amina quiral o una amina aquiral. Un sustrato de amina aquiral tiene la ventaja de no estar limitado en su reacción a un estereoisómero específico, requiriendo así menos del sustrato de amina. Se pueden usar varios sustratos de amina adecuados, que incluyen, a modo de ejemplo y no limitante, 1-isopropil-piperazina y *cis*-2,6-dimetilmorfolina. En algunas realizaciones, se pueden usar otros sustratos de amina, incluyendo, entre otros,  $\alpha$ -feniletanamina (también denominada 1-feniletanamina), y sus enantiómeros (S)-1-feniletanamina y (R)-1-feniletanamina, 2-amino-4-fenilbutano, glicina, ácido L-glutámico, L-glutamato, glutamato monosódico, ácido L-aspartico, L-lisina, L-ornitina,  $\beta$ -alanina, taurina, n-octilamina, ciclohexilamina, 1,4-butanodiamina, 1,6-hexanodiamina, ácido 6-aminoheptanoico, ácido 4-aminobutírico, tiramina, y bencilamina, 2-aminobutano, 2-amino-1-butanol, 1-amino-1-feniletano, 1-amino-1-(2-metoxi-5-fluorofenil)etano, 1-amino-1-fenilpropano, 1-amino-1-(4-hidroxifenil)propano, 1-amino-1-(4-bromofenil)propano, 1-amino-1-(4-nitrofenil)propano, 1-fenil-2-aminopropano, 1-(3-trifluorometilfenil)-2-aminopropano, 2-aminopropanol, 1-amino-1-fenilbutano, 1-fenil-2-aminobutano, 1-(2,5-dimetoxi-4-metilfenil)-2-aminobutano, 1-fenil-3-aminobutano, 1-(4-hidroxifenil)-3-aminobutano, 1-amino-2-metilciclopentano, 1-amino-3-metilciclopentano, 1-amino-2-metilciclohexano, 1-amino-1-(2-naftil)etano, 3-metilciclopentilamina, 2-metilciclopentilamina, 2-etilciclopentilamina, 2-metilciclohexilamina, 3-metilciclohexilamina, 1-aminotetralina, 2-aminotetralina, 2-amino-5-metoxitetralina, y 1-aminoindano, incluyendo ambos isómeros individuales (R) y (S) cuando sea posible.

En algunas realizaciones, el proceso para convertir el sustrato de éster, oxazol-5-carboxilato de etilo, comprende poner en contacto el sustrato de éster a aproximadamente 36 mL/L con aproximadamente 20 g/L de una carboxiesterasa descrita en la presente en MIBK y una temperatura de aproximadamente 30°C en presencia de 43 mL/L de 1-isopropil-piperazina, en donde al menos 80%, 85%, 90%, 92%, 94%, 96% o 98% o más del sustrato de éster se convierte en producto en 24 horas. En algunas realizaciones, el polipéptido de carboxiesterasa capaz de llevar a cabo la reacción anterior comprende una secuencia de aminoácidos correspondiente a SEQ ID NO: 122.

En algunas realizaciones, los procesos anteriores pueden comprender además el paso de aislar el compuesto de la fórmula estructural III., o el compuesto de la fórmula estructural V., del disolvente de reacción.

También se proporcionan en la presente composiciones de las carboxiesterasas y sustratos/productos. En algunas realizaciones, las composiciones pueden comprender el compuesto de la fórmula III., o el compuesto V., y una carboxiesterasa mejorada de la divulgación.

Cualquiera o más de las carboxiesterasas modificadas genéticamente mejoradas pueden ser parte de la composición.

## Ejemplos

Se ilustran diversas características y realizaciones de la divulgación en los siguientes ejemplos representativos, que pretenden ser ilustrativos y no limitantes.

**Ejemplo 1:** Adquisición del gen de carboxiesterasa tipo silvestre Esterasa 2 de *A. acidocaldarius* y construcción de vectores de expresión

Los genes que codifican para carboxiesterasa (CE) se diseñaron para la expresión en *E. coli* en función de la secuencia de aminoácidos informada de la carboxiesterasa, *Esterasa 2* de *A. acidocaldarius* (SEQ ID NO: 2), y un algoritmo de optimización de codones descrito en el Ejemplo 1 de la solicitud US 2008/0248539. Los oligonucleótidos se sintetizaron por separado y luego se unieron usando oligonucleótidos, en general compuestos de 42 nucleótidos. El gen se clonó luego en el vector de expresión pCK 110900 (representado como la Figura 3 en la solicitud de los Estados Unidos 2006/195947, ambas incorporadas a la presente mediante esta referencia en su

totalidad y para todos los fines) bajo el control de un promotor lac. Este vector de expresión también contiene el origen de replicación P15a y el gen de resistencia al cloranfenicol. Los plásmidos resultantes se transformaron en *E. coli* W3110 usando métodos estándar. El gen de codón optimizado y el polipéptido codificado se listan, respectivamente, como SEQ ID NOs: 1 y 2 en la Tabla 3 y el Listado de Secuencias a continuación.

5

Del mismo modo, los genes que codifican para las carboxiesterasas modificadas genéticamente de la presente divulgación listados en la Tabla 3 (SEQ ID NOs: 3-94) se clonaron en el vector pCK 110900 para la expresión en *E. coli* W3110.

10 **Ejemplo 2:** Producción de polvos de carboxiesterasa - procedimiento en matraz de agitación

Se inoculó una colonia microbiana individual de *E. coli* que contiene un plásmido que codificaba para una carboxiesterasa de interés en 50 mL de caldo Luria Bertoni que contiene 30 µg/mL de cloranfenicol y 1% de glucosa. Las células se cultivaron durante la noche (al menos 16 horas) en una incubadora a 30°C con agitación a 250 rpm. El cultivo se diluyó en 1000 mL de caldo Terrific que contiene 30 µg/mL de cloranfenicol para dar una OD600 aproximada de 0,2 y se dejó crecer a 30°C con agitación a 250 rpm. La expresión del gen de carboxiesterasa se indujo mediante la adición de isopropil β D-tiogalactósido (IPTG) a una concentración final de 1 mM cuando la OD600 del cultivo es de 0,6 a 0,8 y luego se continuó la incubación durante la noche (al menos 16 h). Las células se cosecharon por centrifugación (3738 RCF, 20 min, 4°C) y el sobrenadante se descartó. Los gránulos se congelaron durante 2 horas a -80°C. Luego, los sedimentos se descongelaron y se resuspendieron en 3 mL de amortiguador de sulfato de sodio (que consiste en 15 g/L de sulfato de sodio anhidro en agua) por gramo de masa de sedimento final (por ejemplo, 10 g de sedimento congelado suspendido en 30 mL de amortiguador de sulfato de sodio). Los residuos celulares se removieron mediante centrifugación (15 777 RCF, 40 min, 4°C). El sobrenadante del lisado transparente se recolectó, se agrupó y se liofilizó para proporcionar un polvo seco de enzima de carboxiesterasa cruda.

25

**Ejemplo 3:** Producción de polvos de carboxiesterasa - procedimiento de fermentación

Una alícuota de stock de trabajo congelado (*E. coli* que contiene plásmido con el gen de carboxiesterasa de interés) se removió del congelador y se dejó descongelar a temperatura ambiente. Se inoculó 300µL de stock de trabajo en una etapa de siembra primaria de 250 ml de caldo M9YE (1,0 g/L de cloruro de amonio, 0,5 g/L de cloruro de sodio, 6,0 g/L de fosfato monoácido de disodio, 3,0 g/L de fosfato diácido de potasio, 2,0 g/L de extracto de levadura Tastone-154, 1 L/L de agua desionizada) que contiene 30 µg/ml de cloranfenicol y 1% de glucosa en matraces de 1 L y se dejó crecer a 26°C con agitación a 220 rpm. Cuando la OD600 del cultivo fue de 0,5 a 1,0, los matraces se removieron de la incubadora y se usaron inmediatamente para inocular una etapa de siembra secundaria.

35

Se llevó a cabo una etapa de siembra secundaria en fermentadores de 5L a escala de laboratorio usando 4L de medio de crecimiento (0,88 g/L de sulfato de amonio, 0,98 g/L de citrato de sodio; 12,5 g/L de fosfato ácido de dipotasio trihidratado, 6,25g/L de fosfato ácido de potasio, 3,3g/L de extracto de levadura Springer 0251, 0,083 g/L de citrato de amonio férrico, 0,5 mL/L de antiespumante y 8,3 mL/L de una solución de oligoelementos que contiene 2 g/L de cloruro de calcio dihidratado, 2,2 g/L de sulfato de zinc heptahidratado, 0,5 g/L de sulfato de manganeso monohidratado, 1 g/L de sulfato cuproso heptahidratado, 0,1 g/L de molibdato de amonio tetrahidratado y 0,02 g/L de tetraborato de sodio) esterilizado a 120°C durante 40 minutos. Los fermentadores se inocularon con 2 ml de semilla primaria de OD 0,5-1,0 y se incubaron a 30°C, 300 rpm y 0,5vvm de aireación. Cuando la OD600 del cultivo fue de 0,5-1,0 OD600, la semilla secundaria se transfirió inmediatamente a una fermentación de etapa final.

45

La fermentación de la etapa final se llevó a cabo a escala de laboratorio en fermentadores de 10L usando 6L de medio de crecimiento (0,88 g/L de sulfato de amonio, 0,98 g/L de citrato de sodio; 12,5 g/L de fosfato ácido de dipotasio trihidratado, 6,25g/L de fosfato ácido de potasio, 3,3g/L de extracto de levadura Springer 0251, 0,083 g/L de citrato de amonio férrico, 0,5 mL/L de antiespumante y 8,3 mL/L de una solución de oligoelementos que contiene 2 g/L de cloruro de calcio dihidratado, 2,2 g/L de sulfato de zinc heptahidratado, 0,5 g/L de sulfato de manganeso monohidratado, 1 g/L de sulfato cuproso heptahidratado, 0,1 g/L de molibdato de amonio tetrahidratado y 0,02 g/L de tetraborato de sodio) esterilizado a 121 C durante 40 minutos y complementado después de la esterilización con 20 g/L de glucosa monohidratada, 0,48g/L de cloruro de amonio y 0,204g/L de sulfato de magnesio heptahidratado. Los fermentadores se inocularon con 500 ml de semilla secundaria OD600 0,5-1,0 y se incubaron a 30C y 1,6vvm de aireación. El oxígeno disuelto se controló a 30% mediante agitación a velocidad variable de 300-950 rpm. El pH se mantuvo a 7,0 mediante la adición de hidróxido de amonio al 20% v/v. El crecimiento del cultivo se mantuvo mediante la adición de una solución de alimentación que contiene 500 g/L de monohidrato de glucosa, 12 g/L de cloruro de amonio y 5,1 g/L de heptahidrato de sulfato de magnesio.

60

Después de que el cultivo alcanzó una OD600 de 80 +/- 10, se indujo la expresión de carboxiesterasa mediante la adición de isopropil-β-D-tiogalactósido (IPTG) a una concentración final de 1 mM y se continuó la fermentación durante otras 24 horas. Luego, el cultivo se enfrió a 8°C y se mantuvo a esa temperatura hasta la cosecha. Las

células se recolectaron por centrifugación a 5000 G durante 40 minutos en una centrífuga Sorvall RC12BP a 4°C. Los sedimentos celulares cosechados se congelaron a -80°C y se almacenaron hasta el procesamiento y la recuperación aguas abajo, como se describe a continuación.

- 5 Los gránulos se congelaron durante 2 horas a -80°C. Luego, los sedimentos se descongelaron y se resuspendieron en 3 mL de amortiguador de sulfato de sodio (que consiste en 15 g/L de sulfato de sodio anhidro en agua) por gramo de masa de sedimento final (por ejemplo, 10 g de sedimento congelado suspendido en 30 mL de amortiguador de sulfato de sodio). Después de la resuspensión, las células se filtraron a través de una malla de 200 µm antes de pasar dos veces a través del microfluidizador a 12000 psig (82737,087 kPa man.). Los residuos
- 10 celulares se removieron mediante centrifugación (15 777 RCF, 40 min, 4°C). El sobrenadante del lisado transparente se recolectó, se agrupó y se liofilizó para proporcionar un polvo seco de enzima de carboxiesterasa cruda. El polvo de carboxiesterasa se almacena a -80°C.

- 15 **Ejemplo 4:** Métodos analíticos de alto rendimiento para la identificación de variantes de la Esterasa 2 de *A. acidocaldarius* capaces de convertir el sustrato de éster en amida bajo condiciones acuosas

- Método UPLC para determinar la conversión del sustrato éster I en amida III: La conversión enzimática del sustrato de éster de la fórmula I. (disponible comercialmente, número CAS 118994-89-1) a la amida de la fórmula III se determinó usando una UPLC Agilent 1290 equipada con una columna Agilent Zorbax RRHD Eclipse Plus Phenyl-Hexyl (3,0 x 50 mm, 1,8 µm) usando un gradiente de NH<sub>4</sub>Ac 5 mM en agua (fase móvil A) y acetonitrilo (fase móvil B) a una velocidad de flujo de 2 mL/min a una temperatura de columna de 60°C. A partir de una relación 99,9:0,1 de A:B, el método sigue una retención de 0,25 minutos, seguida de un gradiente de 0,05 minutos a 80:20 A:B, seguido de un gradiente de 0,5 minutos a 60:40 A:B, luego un gradiente de purga de 0,1 minutos a 0:100 A:B, una retención de 0,2 minutos a 0:100 A:B, y un gradiente de 0,1 minutos a 99,9:0,1 A:B, y finalmente una retención de 0,3 minutos a 99,9:0,1 A:B. La elución del compuesto se monitoreó a 210 y 230 nm, con éster eluyendo a 0,56 min, amida eluyendo a 0,52 min, y el subproducto ácido de la reacción eluyendo como un pico estrecho cerca del frente del disolvente, a 0,14 min.
- 20
- 25

- Método UPLC para determinar la conversión del sustrato de éster de la fórmula I. en la amida de la fórmula III.: La conversión enzimática del sustrato de éster de la fórmula I. a la amida de la fórmula III. se determinó usando una UPLC Agilent 1290 equipada con una columna Agilent Zorbax SB-C18 RRHD (3,0 x 50 mm, 1,8 µm) usando un gradiente de 0,05% de TFA en agua (fase móvil A) y 0,05% de TFA en acetonitrilo (fase móvil B) a una velocidad de flujo de 2 mL/min a una temperatura de columna de 60°C. A partir de una relación 99,9:0,10 de A:B, el método sigue una retención de 0,25 minutos, seguida de un gradiente de 0,0,25 minutos a 80:20 A:B, seguido de un gradiente de 0,1 minutos a 100:0 A:B, seguido de una retención de 0,1 minutos, seguido de un gradiente de 0,1 minutos a 99,9:0,1 A:B, seguido de una retención de 0,2 minutos. La elución del compuesto se monitoreó a 210 y 230 nm, con éster eluyendo a 0,53 min, amida a 0,23 min, y el subproducto ácido de la reacción eluyendo como un pico estrecho cerca del frente del disolvente, a 0,2 min con un volumen de inyección de 1 µL.
- 30
- 35

- Método UPLC para determinar la conversión del sustrato de éster de la fórmula I. en la amida de la fórmula V.: La conversión enzimática del sustrato de éster de la fórmula I. a la amida de la fórmula V. se determinó usando una UPLC Agilent 1290 equipada con una columna Agilent Zorbax SB-C 18 (3,0 x 50 mm, 1,8 µm) usando un gradiente de 0,05% de TFA en agua (fase móvil A) y 0,05% de TFA en acetonitrilo (fase móvil B) a una velocidad de flujo de 1,5 mL/min a una temperatura de columna de 60°C. A partir de una relación 80:20 de A:B, el método sigue una retención de 0,9 minutos, seguida de un gradiente de 0,1 minutos a 0:100 A:B, seguido de un gradiente de 0,1 minutos a 80:20 A:B, seguido de una retención de 0,4 minutos. La elución del compuesto se monitoreó a 210 y 230 nm, con éster eluyendo a 0,5 min, amida a 0,39 min, y el subproducto ácido de la reacción eluyendo como un pico estrecho cerca del frente del disolvente, a 0,14 min.
- 40
- 45

- 50 **Ejemplo 5:** Selección de alto rendimiento para la identificación de variantes de la Esterasa 2 de *A. acidocaldarius* capaces de convertir el sustrato de éster en amida

- El gen que codifica para la Esterasa 2 de *A. acidocaldarius* (SEQ ID NO: 2), construido como se describe en el Ejemplo 1, se mutagenizó usando los métodos descritos a continuación y la población de moléculas de ADN alteradas se usó para transformar una cepa hospedadora de *E. coli* adecuada. Los transformantes resistentes a los antibióticos se seleccionaron y procesaron para identificar aquellos que expresan una carboxiesterasa con una capacidad mejorada para amidar el sustrato de éster de la fórmula I. al compuesto de las fórmulas (III) y (V) en presencia de sustratos de amina de las fórmulas (II) o (IV), respectivamente. La selección celular, el crecimiento, la expresión inducida de enzimas variantes de carboxiesterasa y la recolección de sedimentos celulares fueron como se describe a continuación.
- 55
- 60

Las colonias de *E. coli* recombinantes que portaban un gen que codifica para la carboxiesterasa se seleccionaron usando un selector de colonias robótico de dispositivos moleculares Q-PIX (Genetix USA, Inc., Boston, MA) en

placas de microtitulación de 96 pocillos de poca profundidad que contienen en cada pocillo 180  $\mu$ L de caldo LB, 1% de glucosa y 30  $\mu$ g/mL de cloranfenicol (CAM). Las células se cultivaron durante la noche a 30°C con agitación a 200 rpm. Luego se transfirió una alícuota de 20  $\mu$ L de este cultivo a placas de 96 pocillos profundos que contienen 380  $\mu$ L de caldo TB y 30  $\mu$ g/mL de CAM. Después de la incubación de las placas de pocillos profundos a 30°C con agitación a 250 rpm durante 2-3 h, se indujo la expresión génica recombinante dentro de las células cultivadas mediante la adición de IPTG a una concentración final de 1 mM. Después, las placas se incubaron a 30°C con agitación a 250 rpm durante 18 h.

Las células se sedimentaron mediante centrifugación (3738 RCF, 10 min, 4°C), se resuspendieron en 200  $\mu$ L de amortiguador de lisis y se lisaron mediante agitación a temperatura ambiente durante 2 horas. En el caso de las condiciones de selección liofilizadas, el amortiguador de lisis contiene sulfato de sodio 100 mM (14,2 g/L), 1 mg/mL de lisozima, 500  $\mu$ g/mL de sulfato de polimixina B (PMBS) y 12,5U/mL Benzonase. En el caso de las condiciones de selección acuosas, el amortiguador de lisis consistió en fosfato de potasio 10 mM, pH 7,0, 1 mg/mL de lisozima y 500  $\mu$ g/mL de PMBS. Después de sellar las placas con sellos de nylon permeables al aire, se agitaron vigorosamente durante 2 horas a temperatura ambiente. Los residuos celulares se sedimentaron mediante centrifugación (3738 RCF, 10 min., 4°C) y el sobrenadante transparente se analizó directamente o se almacenó a 4°C hasta su uso.

Para la selección en condiciones semi-acuosas, con carboxiesterasas modificadas genéticamente en etapa temprana, se adicionó una alícuota de 120  $\mu$ L de solución de sustrato (720 mL/L de DMSO, 90 mL/L de etanol de prueba 200, 19,67 mL/L de isopropil-piperazina, 23,5 mL/L de sustrato de éster y 8,5 mL/L de HCl 6N) a cada pocillo de una placa de pocillos profundos Costar, seguido de la adición de 80  $\mu$ L del sobrenadante de lisado recuperado usando un instrumento robótico Biomek FX (Beckman Coulter, Fullerton, CA). Resultó una solución que comprendía 100 mM de sustrato de éster, 100 mM de isopropil-piperazina, 45% de DMSO, 5% de EtOH a un pH final de 9,0. Las placas se termosellaron con cinta de termosellado laminada de aluminio/polipropileno a 165°C durante 4 segundos y luego se agitaron durante la noche (al menos 16 horas) a 50°C. Las reacciones se extinguieron mediante la adición de 200  $\mu$ L de Acetonitrilo mediante un Biomex FX. Las placas se resellaron, se agitaron durante 5 min y luego se centrifugaron a 3738 RCF durante 10 min. Una muestra de 20  $\mu$ L de sustrato se transfirió a una placa de polipropileno de pozos poco profundos (Costar #3365) que contiene 180  $\mu$ L de 75% de acetonitrilo en agua, se selló, se agitó durante 10 min y luego se analizó como se describe en el Ejemplo 4.

Para la selección en condiciones orgánicas, con carboxiesterasas modificadas genéticamente de etapa temprana, se adicionó una alícuota de 150  $\mu$ L de sobrenadante de lisado recuperado a un estante de 96 pocillos de aluminio (F158359, Unchained Labs, Pleasanton, CA) cargado con insertos de vial de vidrio de 1 mL (S11168, Unchained Labs, Pleasanton, CA). Este aparato luego se sometió a liofilización, se calentó suavemente a temperatura ambiente y se cargó con 4  $\mu$ L de agua destilada usando un Dispensador de Reactivos Multidrop Combi (Thermo Scientific, Waltham, MA) seguido de 200  $\mu$ L de solución de sustrato orgánico (10,93 mL/L de isopropil-piperazina, 23,5 mL/L de sustrato de éster) en éter metílico de ter-butilo (tBME). Las placas se sellaron luego mediante el uso de una lámina de teflón (S11690-2, Unchained Labs, Pleasanton, CA) en capas debajo de dos juntas de goma (S13086, Unchained Labs, Pleasanton, CA) y una tapa de estante de metal (F158424, Unchained Labs, Pleasanton, CA), unida por siete tornillos (C151943-050, Unchained Labs, Pleasanton, CA). Luego, estos constructos se incubaron, con agitación, a 50°C durante la noche (al menos 16 horas). Las reacciones se extinguieron mediante la adición de 200  $\mu$ L de alcohol isopropílico mediante un Biomek FX, luego se sellaron, se agitaron durante 10 min y se centrifugaron a 235 RCF durante 2 minutos para sedimentar los sólidos residuales. Luego se transfirió una muestra de 200  $\mu$ L a una placa de pocillos profundos Costar que contiene 200  $\mu$ L de isopropanol en cada pocillo, se termoselló con cinta de termosellado laminada de aluminio/polipropileno a 165°C durante 4 segundos, se agitó durante 10 min y luego se centrifugó a 3738 RCF durante 10 min. Una muestra de 40  $\mu$ L de sustrato se transfirió a una placa de polipropileno de pocillos poco profundos (Costar #3365) que contiene 160  $\mu$ L de isopropanol, se selló, se agitó durante 10 min y luego se analizó como se describe en el Ejemplo 4.

Para la selección en condiciones orgánicas, con carboxiesterasas modificadas genéticamente en etapa tardía, se adicionó una alícuota de 120  $\mu$ L de sobrenadante de lisado recuperado a una rejilla de 96 pocillos de aluminio (F158359, Unchained Labs, Pleasanton, CA) cargada con insertos de vial de vidrio de 1 mL (S11168, Unchained Labs, Pleasanton, CA). Este aparato luego se sometió a liofilización, se calentó suavemente a temperatura ambiente y se cargó con 200  $\mu$ L de solución de sustrato orgánico (53,2mL/L isopropil-piperazina, 43 mL/L de sustrato de éster) en metil isobutil éster (MIBK) seguido de 4  $\mu$ L de agua destilada usando un Dispensador de Reactivos Multidrop Combi (Thermo Scientific, Waltham, MA). Las placas se sellaron luego mediante el uso de una lámina de teflón (S11690-2, Unchained Labs, Pleasanton, CA) en capas debajo de una juntas de goma (S13086, Unchained Labs, Pleasanton, CA) y una tapa de estante de metal (F158424, Unchained Labs, Pleasanton, CA), unida por cinco tornillos (C151943-050, Unchained Labs, Pleasanton, CA). Luego, estos constructos se incubaron, con agitación, a 15°C durante la noche (al menos 16 horas). Las reacciones se removieron de la incubadora y se centrifugaron a 235 RCF. Se transfirió una muestra de 20  $\mu$ L de sobrenadante a una placa de polipropileno de

pocillos poco profundos (Costar #3365) que contiene 180 µL de isopropanol (con 5 g/L de naftaleno) en cada pocillo, se termoselló con cinta de termosellado laminada de aluminio/polipropileno a 165°C durante 4 segundos, se agitó durante 10 min y luego se centrifugó a 3738 RCF durante 10 min. Una muestra de 10 µL de sustrato se transfirió a una placa de polipropileno de pocillos poco profundos (Costar #3365) que contiene 190 µL de isopropanol, se selló, se agitó durante 10 min y luego se analizó como se describe en el Ejemplo 4.

**Ejemplo 6:** Amidación en metil isobutil éster (MIBK) del sustrato de éster, fórmula I. y sustrato de amina, fórmula II., por carboxiesterasas modificadas genéticamente derivadas de Esterasa 2 de *A. acidocaldarius*

Las carboxiesterasas de etapa temprana mejoradas descritas en la Tabla 3 se evaluaron a escala preparativa en MIBK de la siguiente manera. Se adicionaron 10 mg de polvo de enzima liofilizado a un vial de HPLC de 1,5 mL, junto con 15 µL de agua destilada. Posteriormente, se adicionaron 485 µL de una solución de sustrato (42,93 mL isopropil-piperazina/L MIBK, 36,4 mL éster/L MIBK) y los viales se sellaron. La reacción se agitó en un agitador de viales calentado Eppendorf Thermomixer C a 50°C y 850 rpm durante 16 horas. Las reacciones se extinguieron mediante la adición de 500 µL de isopropanol. Luego se transfirió una muestra de 50 µL a una placa de polipropileno de pocillos poco profundos (Costar #3365) que contiene 150 µL de isopropanol, se agitó durante 10 minutos y luego se centrifugó a 3738 RCF durante 10 min. Una muestra de 10 µL de sustrato se transfirió a una placa de polipropileno de pocillos poco profundos (Costar #3365) que contiene 190 µL de isopropanol, se selló, se agitó durante 10 min y luego se analizó, como se describe en el Ejemplo 4.

Las carboxiesterasas de etapa tardía mejoradas descritas en la Tabla 3 se evaluaron a escala preparativa en MIBK de la siguiente manera. Se adicionó 11,25mg de polvo de enzima liofilizado a un vial de HPLC de 1,5 mL. Posteriormente, se adicionó 750µL de una solución de sustrato (53,2 mL de isopropil-piperazina/L MIBK, 43 mL de éster/L MIBK), junto con 15 µL de agua destilada, y los viales se sellaron. La reacción se agitó en un agitador de viales calentado Eppendorf Thermomixer C a 15°C y 850 rpm durante 16 horas. Las reacciones se extinguieron mediante la remoción de 20ul como se describe en la selección de carboxiesterasa modificada en etapa tardía en el Ejemplo 5. La Tabla 3 proporciona SEQ ID NO: correspondiente a las variantes de carboxiesterasa que se analizaron de esta manera, así como la cantidad de diferencias de residuos de aminoácidos de la carboxiesterasa tipo silvestre Esterasa 2 de *A. acidocaldarius* (SEQ ID NO: 2).

**Ejemplo 7:** Amidación en metil isobutil éster (MIBK) del sustrato de éster I y el sustrato de amina II mediante carboxiesterasas modificadas genéticamente derivadas de Esterasa de *A. acidocaldarius*

El siguiente ejemplo ilustra un proceso a escala de gramos usado para aumentar la conversión del sustrato de éster, oxazol-2-carboxilato de etilo, el compuesto de la fórmula I y el sustrato de amina, 1-isopropil-piperazina, el compuesto de la fórmula II. Este proceso aprovecha la mezcla de líquidos mejorada a gran escala para aumentar la conversión del sustrato en producto. El monitoreo de rutina permite la captura y el aislamiento de más del 70% del rendimiento total del producto.

El proceso de reacción a gran escala contiene los siguientes componentes de reacción:

|                                  |                         |
|----------------------------------|-------------------------|
| oxazol-2-carboxilato de etilo:   | 20 G (98%) (141,7 mmol) |
| 1-isopropil-piperazina:          | 23,62 G (184,2 mmol)    |
| carboxiesterasa (SEQ ID NO: 90): | 14 G                    |
| MIBK:                            | 400 ml                  |
| agua:                            | 11 ml                   |

Proceso. A un reactor CLR de 2L equipado con agitador superior se adicionaron 340 mL de MIBK y se ajustó para agitar a temperatura ambiente. Luego se adicionaron 11 mL de agua, seguido de un calentamiento suave a 30°C. Se cargó 14 g de carboxiesterasa, seguido de un lavado con MIBK de 20 mL. Luego se cargó 23,62 g de 1-isopropil-piperazina, seguido de un lavado con MIBK de 20 mL. Finalmente, se adicionaron 20 g de oxazol-2-carboxilato de etilo, seguido de un lavado final de 20 mL de MIBK. Después de 16 horas, la conversión superó el 90%.

Los sólidos residuales se filtraron y se lavaron con 80 mL de MIBK para aislar cualquier material adsorbido. Luego, los lavados se agruparon y se lavaron con 20 mL de 10%p/p de cloruro de sodio, se agitaron y se separaron. La fase orgánica se recolectó y se concentró a 60 mL bajo vacío, luego se enfrió a 15°C, momento en el cual comenzó a producirse la cristalización. Se cargó 70 mL de n-heptano en el transcurso de 10 minutos, se dejó que la cristalización avanzara hasta completarse en el transcurso de 2 horas, y luego se recolectó el material cristalino por filtración. El producto cristalino se lavó dos veces con 80 mL de MIBK: n-heptano 1: 4 y se secó a 50°C bajo vacío. Este proceso dio como resultado un rendimiento general (p/p) del 66% evaluado por la masa del producto.

**Ejemplo 8:** Amidación en metil isobutil éster (MIBK) del sustrato de éster I y el sustrato de amina IV mediante

carboxiesterasas modificadas genéticamente derivadas de Esterasa de *A. acidocaldarius*

Se realizaron reacciones a escala de vial para producir la amida de la fórmula V. usando carboxiesterasas modificadas genéticamente. Se adicionaron 25mg de polvo de enzima liofilizado a un vial de HPLC de 1,5 mL, junto con 10 µL de agua destilada. Posteriormente, se adicionó 490 µL de una solución de sustrato (26,72 mL de (2S,6R)-2,6-dimetilmorfolina/L de MIBK, 24,27mL de éster/L de MIBK) y los viales se sellaron. La reacción se agitó en un agitador de viales calentado Eppendorf Thermomixer C a 50°C y 800 rpm durante 16 horas. Se adicionó una alícuota de 100 µL de cada solución de reacción a una placa de polipropileno de pocillos poco profundos (Costar #3365) que contiene 100 µL de isopropanol, se selló, se agitó durante 10 minutos y luego se centrifugó a 3738 RCF durante 10 minutos. Luego, una alícuota de 10 µL del sobrenadante se diluyó en 190 µL de acetonitrilo que contiene 25% de agua en una placa de polipropileno de pocillos poco profundos (Costar #3365), se selló y luego se agitó a 850 rpm durante 5 minutos. Las reacciones se analizaron como se describe en el Ejemplo 4. En todos los casos observados, la actividad enzimática coincidió estrechamente con la de las reacciones descritas en el Ejemplo 6 y en la Tabla 3.

Listado de secuencias

>S00032511

SEQ ID NO:1

20 TIPO: ADN, ORGANISMO: ALICYCLOBACILLUS

ATGCCATTAGATCCTGTGATTCAACAAGTCCTCGATCAACTGAACCGTATGCCAGCCCCTGACTACAAGCACTT  
AAGCGCTCAACAGTTTCGTTACAGCAATCACTGTTTCCACCAGTTAAGAAAGAACCGGTGCGCAGAAAGTTTCG  
CGAATTCGACATGGATCTGCCGGGCGGTACCTGAAAGTCCGTATGTACCGTCCGGAAGGTGTTGAACCACC  
25 ATATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGTTGGTGACCTGGAAACGCATGATCCGGTGTGT  
CGTGTGTTGGCGAAAGATGGACGCGCAGTGTTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTTT  
CCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGATC  
CAGCACGCATTGCAGTTGGCGGCGATTGAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAAA  
GAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTCCACTGGTTATGATCCAGCACATCCAC  
30 CAGCAAGTATCGAAGAGAATGCGGAGGGTTACCTGTTAACCAGGAGGCATGATGCTGTGGTTTCGTGACCAGT  
ACCTGAACAGCCTTGAGGAAGTCACTACCCATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACC  
ACCGGCGTACATTGCAACCGCACAGTACGATCCCTGCGCGATGTCGGCAAACCTTTACGCTGAAGCCCTGAA  
CAAAGCGGGCGTTAAGGTGGAGATCGAGAACTTCGAGGATCTGATTCATGGCTTCGCGCAGTTCTATTCTCTG  
AGCCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

>S00039799

SEQ ID NO:3

TIPO: ADN, ORGANISMO: ARTIFICIAL

ATGCCATTAGATCCTGTGATTCAACAAGTCCTCGATCAACTGAACCGTATGCCAGCCCCTGACTACAAGCACTT  
AAGCGCTCAACAGTTTCGTTACAGCAATCACTGTTTCCACCAGTTAAGAAAGAACCGGTGCGCAGAAAGTTTCG  
CGAATTCGACATGGATCTGCCGGGCGGTACCTGAAAGTCCGTATGTACCGTCCGGAAGGTGTTGAACCACC  
ATATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGTTGGTGACCTGGAAACGCATGATCCGGTGTGT  
CGTGTGTTGGCGAAAGATGGACGCGCAGTGTTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTTT  
45 CCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGATC  
CAGCACGCATTGCAGTTGGCGGCGATTGAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAAA  
GAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTCCACTGGTTATGATCCAGCACATCCAC  
CAGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAACCAGGAGGCATGATGCTGTGGTTTCGTGACCAGTA  
CCTGAACAGCCTTGAGGAAGTCACTACCCATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCA  
50 CCGGCGTACATTGCAACCGCACAGTACGATCCCTGCGCGATGTCGGCAAACCTTTACGCTGAAGCCCTGAAC  
AAAGCGGGCGTTAAGGTGGAGATCGAGAACTTCGAGGATCTGATTCATGGCTTCGCGCAGTTCTATTCTCTGA  
GCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

>S00039972

SEQ ID NO:5

TIPO: ADN, ORGANISMO: ARTIFICIAL

ATGCCATTAGATCCTGTGATTCAACAAGTCCTCGATCAACTGAACCGTATGCCAGCCCCTGACTACAAGCACTT  
AAGCGCTCAACAGTTTCGTTACAGCAATCACTGTTTCCACCAGTTAAGAAAGAACCGGTGCGCAGAAAGTTTCG  
60 CGAATTCGACATGGATCTGCCGGGCGGTACCTGAAAGTCCGTATGTACCGTCCGGAAGGTGTTGAACCACC  
ATATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGTTGGTGACCTGGAAACGCATGATCCGGTGTGT  
CGTGTGTTGGCGAAAGATGGACGCGCAGTGTTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTTT  
CCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGATC

CAGCACGCATTGCAGTTGGCGGCGATTGAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAAA  
GAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTCCACTGGTTATGATCCAGCACATCCAC  
CAGCAAGTATCATTGAGAAATGCGGAGGGTTACCTGTTAACCAGGAGGCATGATGCTGTGGTTTCGTGACCAAGTA  
CCTGAACAGCCTTGAGGAACTGACTACCCATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCA  
5 CCGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTACGCTGAAGCCCTGAAC  
AAAGCGGGCGTTAAGGTGGAGATCGAGAACTTCGAGGATCTGATTCATGGCTTCGCGCAGTTCTATTCTCTGA  
GCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

>S00039989

10 SEQ ID NO:7

TIPO: ADN, ORGANISMO: ARTIFICIAL

ATGCCATTAGATCCTGTGATTCAACAAGTCCTCGATCAACTGAACCGTATGCCAGCCCCTGACTACAAGCACTT  
AAGCGCTCAACAGTTTCGTTTACAGCAATCACTGTTTCCACCAGTTAAGAAAGAACCGGTGCGCAGAAAGTTG  
15 CGAATTCGACATGGATCTGCCGGGCCGTACCCTGAAAGTCCGTATGTACCGTCCGGAAGGTGTTGAACCACC  
ATATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGTTGGTGACCTGGAAACGCATGATCCGGTGTGT  
CGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTTT  
CCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCATCTTGATC  
CAGCACGCATTGCAGTTGGCGGCGATTGAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAAA  
20 GAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTCCACTGGTTATGATCCAGCACATCCAC  
CAGCAAGTATCTATGAGAATGCGGAGGGTTACCTGTTAACCAGGAGGCATGATGCTGTGGTTTCGTGACCAGTA  
CCTGAACAGCCTTGAGGAACTGACTACCCATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCA  
CCGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTACGCTGAAGCCCTGAAC  
AAAGCGGGCGTTAAGGTGGAGATCGAGAACTTCGAGGATCTGATTCATGGCTTCGCGCAGTTCTATTCTCTGA  
25 GCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

>S00040129

SEQ ID NO:9

TIPO: ADN, ORGANISMO: ARTIFICIAL

ATGCCATTAGATCCTGTGATTCAACAAGTCCTCGATCAACTGAACCGTATGCCAGCCCCTGACTACAAGCACTT  
AAGCGCTCAACAGTTTCGTTTACAGCAATCACTGTTTCCACCAGTTAAGAAAGAACCGGTGCGCAGAAAGTTG  
CGAATTCGACATGGATCTGCCGGGCCGTACCCTGAAAGTCCGTATGTACCGTCCGGAAGGTGTTGAACCACC  
ATATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGTTGGTGACCTGGAAACGCATGATCCGGTGTGT  
35 CGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTTT  
CCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCATCTTGATC  
CAGCACGCATTGCAGTTGGCGGCGATTGAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAAA  
GAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTCCACTGGTTATGATCCAGCACATCCAC  
CAGCAAGTATCTTTGAGAATGCGGAGGGTTACCTGTTAACCAGGAGGCATGATGCTGTGGTTTCGTGACCAGTA  
40 CCTGAACAGCCTTGAGGAACTGACTACCCATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCA  
CCGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTACGCTGAAGCCCTGAAC  
AAAGCGGGCGTTAAGGTGGAGATCGAGAACTTCGAGGATCTGATTCATGGCTTCGCGCAGTTCTATTCTCTGA  
GCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

45 >S00040188

SEQ ID NO:11

TIPO: ADN, ORGANISMO: ARTIFICIAL

ATGCCATTAGATCCTGTGATTCAACAAGTCCTCGATCAACTGAACCGTATGCCAGCCCCTGACTACAAGCACTT  
AAGCGCTCAACAGTTTCGTTTACAGCAATCACTGTTTCCACCAGTTAAGAAAGAACCGGTGCGCAGAAAGTTG  
50 CGAATTCGACATGGATCTGCCGGGCCGTACCCTGAAAGTCCGTATGTACCGTCCGGAAGGTGTTGAACCACC  
ATATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGTTGGTGACCTGGAAACGCATGATCCGGTGTGT  
CGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTTT  
CCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCATCTTGATC  
55 CAGCACGCATTGCAGTTGGCGGCGATTGAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAAA  
GAACGTGGTGGTCCGGCAATTGCGTTTCAACTGCTGATCTATCCCTCCACTGGTTATGATCCAGCACATCCAC  
CAGCAAGTATCATGGAGAATGCGGAGGGTTACCTGTTAACCAGGAGGCATGATGCTGTGGTTTCGTGACCAGT  
ACCTGAACAGCCTTGAGGAACTGACTACCCATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACC  
ACCGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTACGCTGAAGCCCTGAA  
60 CAAAGCGGGCGTTAAGGTGGAGATCGAGAACTTCGAGGATCTGATTCATGGCTTCGCGCAGTTCTATTCTCTG  
AGCCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

>S00044917

SEQ ID NO:13

TIPO: ADN, ORGANISMO: ARTIFICIAL

ATGCCATTAGATCCTGTGATTCAACAAGTCCTCGATCAACTGAACCGTATGCCAGCCCCTGACTACAAGCACTT  
AAGCGCTCAGCAGTTTCGTTACAGCAATCACTGTTGCCACCAGTTAAGAAAGAACCGGTCGAGAAGTTCCG  
5 CGAATTCGACATGGATCTGCCGGGCCGTACCCTGAAAGTCCGTATGTACCGTCCGGAAGGTGTTGAACCACC  
ATATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGTTGGTGACCTGGAAACGCATGATCCGGTGTGT  
CGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTTT  
CCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGATC  
CAGCACGCATTGCAGTTGGCGGCGATTGAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAAA  
10 GAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTTACTGGTTATGATCCAGCACATCCACC  
AGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAGAGCATGGCATGATGCTGTGGTTTCGTGACCAGTAC  
CTGAACAGCCTTGAGGAAGTCACTACCCATGGTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCAC  
CGCGGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTACGCTGAAGCCCTGAACA  
AAGCGGGCGTTAAGGTGGAGATCGAGAAGTTCGAGGATCTGATTGCTTGGCTTCGCGCAATTTATTCTCTGAG  
15 CCCAGGCGCAACCAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

>S00044931

SEQ ID NO:15

TIPO: ADN, ORGANISMO: ARTIFICIAL

20 ATGCCATTAGATCCTGTGATTCAACAAGTCCTCGATCAACTGAACCGTATGCCAGCCCCTGACTACAAGCACTT  
AAGCGCTCAGCAGTTTCGTTACAGCAATCACTGTTTCCACCAGTTAAGAAAGAACCGGTCGAGAAGTTCCG  
CGAATTCGACATGGATCTGCCGGGCCGTACCCTGAAAGTCCGTATGTACCGTCCGGAAGGTGTTGAACCACC  
ATATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGTTGGTGACCTGGAAACGCATGATCCGGTGTGT  
25 CGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTTT  
CCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGATC  
CAGCACGCATTGCAGTTGGCGGCGATTGAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAAA  
GAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTTACTGGTTATGATCCAGCACATCCACC  
AGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAACCAGGAGGCATGATGCTGTGGTTTCGTGACCAGTAC  
30 CTGAACAGCCTTGAGGAAGTCACTACCCATGGTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCAC  
CGCGGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTACGCTGAAGCCCTGAACA  
AAGCGGGCGTTAAGGTGGAGATCGAGAAGTTCGAGGATCTGATTGCTTGGCTTCGCGCAGTTCTATTCTCTGAG  
CCCAGGCGCAACCAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

35 >S00044943

SEQ ID NO:17

TIPO: ADN, ORGANISMO: ARTIFICIAL

40 ATGCCATTAGATCCTGTGATTCAACAAGTCCTCGATCAACTGAACCGTATGCCAGCCCCTGACTACAAGCACTT  
AAGCGCTCAACAGTTTCGTTACAGCAATCACTGTTGCCACCAGTTAAGAAAGAACCGGTCGAGAAGTTCCG  
CGAATTCGACATGGATCTGCCGGGCCGTACCCTGAAAGTCCGTATGTACCGTCCGGAAGGTGTTGAACCACC  
ATATCGGGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGTTGGTGACCTGGAAACGCATGATCCGGTGTG  
TCGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTT  
TCCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGAT  
45 CCAGCAGCATGATTGGCGGCGATTGAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAA  
AGAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTTACTGGTTATGATCCAGCACATCCAC  
CAGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAACGCATGGCATGATGCTGTGGTTTCGTGACCAGTA  
CCTGAACAGCCTTGAGGAAGTCACTACCCATGGTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCA  
CCGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTACGCTGAAGCCCTGAAC  
50 AAAGCGGGCGTTAAGGTGGAGATCGAGAAGTTCGAGGATCTGATTGCTTGGCTTCGCGCAGTTCTATTCTCTGA  
GCCAGGCGCAACCCTCGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

>S00045046

SEQ ID NO:19

55 TIPO: ADN, ORGANISMO: ARTIFICIAL

ATGCCATTAGATCCTGTGATTCAACAAGTCCTCGATCAACTGAACCGTATGCCAGCCCCTGACTACAAGCACTT  
AAGCGCTCAACAGTTTCGTTACAGCAATCACTGTTCCCACCAGTTAAGAAAGAACCGGTCGAGAAGTTCCG  
60 CGAATTCGACATGGATCTGCCGGGCCGTACCCTGAAAGTCCGTATGTACCGACCGGAAGGTGTTGAACCACC  
ATATCCGGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGTTGGTGACCTGGAAACGCATGATCCGGTGTG  
TCGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTT  
TCCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGAT  
CCAGCACGCATTGCAGTTGGCGGCGATTGAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAA  
AGAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTTACTGGTTATGATCCAGCACATCCAC

# ES 3 018 269 T3

CAGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAGAGCATGGCATGATGCTGTGGTTTCGTGACCAGTA  
CCTGAACAGCCTTGAGGAACTGACTCACCCTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCA  
CCGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTTACGCTGAAGCCCTGAAC  
AAAGCGGGCGTTAAGGTGGAGATCGAGAACTTCGAGGATCTGATTCATGGCTTCGCGCAGTTCTATTCTCTGA  
5 GCCCAGGCGCAACCAAGGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

>S00045575  
SEQ ID NO:21  
TIPO: ADN, ORGANISMO: ARTIFICIAL

10 ATGCCATTAGATCCTGTGATTCAACAAGTCCTCGATCAACTGAACCGTATGCCAGCCCCTGACTACAAGCACTT  
AAGCGCTCAACAGTTTCGTTCCACAGCAACACCTGTTCCACCAGTTAAGAAAGAACCGGTGCGAGAAGTTCCG  
CGAATTCGACATGGATCTGCCGGGCCGTACCCTGAAAGTCCGTATGTACCGCCCGGAAGGTGTTGAACCACC  
ATATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGTTGGTGACCTGGAAACGCATGATCCGGTGTGT  
15 CGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTTT  
CCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGATC  
CAGCACGCATTGCAGTTGGCGGCGATTGAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAAA  
GAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTTACTGGTTATGATCCAGCACATCCACC  
AGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAACCGGAGGCATGATGCTGTGGTTTCGTGACCAGTAC  
20 CTGAACAGCCTTGAGGAACTGACTCACCCTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCAC  
CGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTACGCTGAAGCCCTGAACA  
AAGCGGGCGTTAAGGTGGAGATCGAGAACTTCGAGGATCTGATTCATGGCTTCGCGCAGTTCTATTCTCTGAG  
CCCAGGCGCAACCAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

25 >S00047040  
SEQ ID NO:23  
TIPO: ADN, ORGANISMO: ARTIFICIAL

30 ATGCCATTAGATCCTGTGATTCAACAAGTCCTCGATCAACTGAACCGTATGCCAGCACCTGACTACAAGCACTT  
AAGCGCTCAACAGTTTCGTTCCACAGCAATCACTGTTTCCACCAGTTAAGAAAGAACCGGTGCGAGAAGTTCCG  
CGAATTCGACATGGATCTGCCGGGCCGTACCCTGAAAGTCCGTATGTACCGTCCGGAAGGTGTTGAACCACC  
ATATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGTTGGTGACCTGGAAACGCATGATCCGGTGTGT  
CGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTTT  
CCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGATC  
35 CAGCACGCATTGCAGTTGGCGGCGATTGAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAAA  
GAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCATGACTGGTTATGATCCAGCACATCCAC  
CAGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAGAGGGAGGCATGATGCTGTGGTTTCGTGACCAGT  
ACCTGAACAGCCTTGAGGAACTGACTCACCCTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACC  
ACCGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTACGCTGAAGCCCTGAA  
40 CAAAGCGGGCGTTAAGGTGGAGATCGAGAACTTCGAGGATCTGATTCATGGCTTCGCGCAGTTCTATTCTCTG  
AGCCCAGGCGCAACCGTGCAGTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

>S00048311  
SEQ ID NO:25  
45 TIPO: ADN, ORGANISMO: ARTIFICIAL

ATGCCATTAGATCCTGTGATTCAACAAGTCCTCGATCAACTGAACCGTATGCCAGCCCCTGACTACAAGCACTT  
AAGCCCGCAACAGTTGCGTTCCACAGCAATCACTGTTTCCACCAGTTAAGAAAGAACCGGTGCGAGAAGTTCCG  
CGAATTCGACATGGATCTGCCGGGTCGTACCCTGAAAGTCCGTATGTACCGTCCGGAAGGTGTTGAACCACC  
50 ATATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGTTGGTGACCTGGAAACGCATGATCCGGTGTGT  
CGTGTGTTGGCGAAATGGGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTTT  
CCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGATC  
CAGCACGCATTGCAGTTGGCGGCGATTGAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAAA  
GAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTCCACTGGTTATGATCCAGCACATCCAC  
55 CAGCAAGTATCATTGAGAATGCGGAGGGTTACCTGTTAACCGGAGGCATGATGCTGTGGTTTCGTGACCAGTA  
CCTGAACAGCCTTGAGGAACTGACTCACCCTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCA  
CCGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTACGCTGAAGCCCTGAAC  
AAAGCGGGCGTTAAGGTGGAGATCGAGAACTTCGAGGATCTGATTCATGGCTTCGCGGGGTTCTATTCTCTG  
60 AGCCCAGGCGCAACCAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

>S00048430  
SEQ ID NO:27  
TIPO: ADN, ORGANISMO: ARTIFICIAL

ATGCCATTAGATCCTGTGATTCAACAAGTCCTCGATCAACTGAACCGTATGCCAGCCCCTGACTACAAGCACTT  
 AAGCGCTCAACAGTTTCGTTACAGCAGTCACTGTTCCACCAGTTAAGAAAGAACCGGTCGCAGAAGTTCCG  
 CGAATTCGACATGGATCTGCCGATGCGTACCCTGAAAGTCCGTATGTACCGTCCGGAAGGTGTTGAACCACC  
 ATATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGTTGGTGACCTGGAAACGCATGATCCGGTGTGT  
 5 CGTGTGTTGGCGAAATGGGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTTT  
 CCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGATC  
 CAGCACGCATTGCAGTTGGCGGCGATTGAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAAA  
 GAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTCCACTGGTTATGATCCAGCACATCCAC  
 CAGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAACCCATGGCATGATGCTGTGGTTTCGTGACCAGTA  
 10 CCTGAACAGCCTTGAGGAACTGACTCACCCTATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCA  
 CCGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTACGCTGAAGCCCTGAAC  
 AAAGCGGGCGTTAAAGTGGAGATCGAGAACTTCGAGGATCTGATTATGGCTTCGCGGGGTTCTATTCTCTGA  
 GCCCAGGCGCAACCAGAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

15 >S00048441  
 SEQ ID NO:29  
 TIPO: ADN, ORGANISMO: ARTIFICIAL

ATGCCATTAGATCCTGTGATTCAACAAGTCCTCGATCAACTGAACCGTATGCCAGCCCCTGACTACAAGCACTT  
 20 AAGCGCTCAACAGTTTCGTTACAGCAATCACTGTTCCGCCAGTTAAGAAAGAACCGGTCGCAGAAGTTCCG  
 CGAATTCGACATGGATCTGCCGGGCCGTACCCTGAAAGTCCGTATGTACCGTCCGGAAGGTGTTGAACCACC  
 ATATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGTTGGTGACCTGGAAACGCATGATCCGGTGTGT  
 CGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTTT  
 CCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCCGCCGATTTCCATCTTGAT  
 25 CCAGCACGCATTGCAGTTGGCGGCGATTGAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAA  
 AGAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTCCACTGGTTATGATCCAGCACATCCA  
 CCAGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAACCAGGAGGCATGATGCTGTGGTTTCGTGACCAG  
 TACCTGAACAGCCTTGAGGAACTGACTCACCCTATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTAC  
 CACCGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTACGCTGAAGCCCTGA  
 30 ACAAGCGGGCGTTAAAGTGGAGATCGAGAACTTCGAGGATCTGATTATGGCTTCGCGGGGTTCTATTCTCT  
 GAGCCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

>S00048480  
 SEQ ID NO:31  
 35 TIPO: ADN, ORGANISMO: ARTIFICIAL

ATGCCATTAGATCCTGTGATTCAACAAGTCCTCGATCAACTGAACCGTATGCCAGCCCCTGACTACAAGCACTT  
 AAGCGCTCAACAGTTTCGTTACAGCAATCACTGTTCCGCCAGTTAAGAAAGAACCGGTCGCAGAAGTTCCG  
 CGAATTCGACATGGATCTGCCGGGCCGTACCCTGAAAGTCCGTATGTACCGTCCGGAAGGTGTTGAACCACC  
 40 ATATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGTTGGTGACCTGGAAACGCATGATCCGGTGTGT  
 CGTGTGTTGGCGAAATGGGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTTT  
 CCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCCGCCGATTTCCATCTTGAT  
 CCAGCACGCATTGCAGTTGGCGGCGATTGAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAA  
 AGAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTCCACTGGTTATGATCCAGCACATCCA  
 45 CCAGCAAGTATCATTGAGAATGCGGAGGGTTACCTGTTAACCAGGTGGCATGATGCTGTGGTTTCGTGACCAGT  
 ACCTGAACAGCCTTGAGGAACTGACTCACCCTATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACC  
 ACCGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTACGCTGAAGCCCTGAA  
 CAAAGCGGGCGTTAAAGTGGAGATCGAGAACTTCGAGGATCTGATTATGGCTTCGCGGGGTTCTATTCTCTG  
 AGCCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

50 >S00048593  
 SEQ ID NO:33  
 TIPO: ADN, ORGANISMO: ARTIFICIAL

ATGCCATTAGATCCTGTGATTCAACAAGTCCTCGATCAACTGAACCGTATGCCAGCCCCTGACTACAAGCACTT  
 AAGCGCTCAACAGTTTCGTTACAGCAATCACTGTTCCACCAGTTAAGAAAGAACCGGTCGCAGAAGTTCCG  
 CGAATTCGACATGGATCTGCCGGGCCGTACCCTGAAAGTCCGTATGTACCGTCCGGAAGGTGTTGAACCACC  
 ATATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGTTGGTGACCTGGAAACGCATGATCCGGTGTGT  
 CGTGTGTTGGCGAAATGGGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTTT  
 60 CCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCCGCCGATTTCCATCTTGAT  
 CCAGCACGCATTGCAGTTGGCGGCGATTGAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAA  
 AGAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTCCACTGGTTATGATCCAGCACATCCA  
 CCAGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAACCAGGTGGCATGATGCTGTGGTTTCGTGACCAG  
 TACCTGAACAGCCTTGAGGAACTGACTCACCCTATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTAC

# ES 3 018 269 T3

CACCGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTTACGCTGAAGCCCTGA  
ACAAAGCGGGCGTTGATGTGGAGATCGAGAACTTCGAGGATCTGATTCATGGCTTCGCGGGGTTCTATTCTCT  
GAGCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

5 >S00052854  
SEQ ID NO:35  
TIPO: ADN, ORGANISMO: ARTIFICIAL

ATGCCATTAGATCCTGTGATTCAATATGTCCTCGATCAACTGAACCGTATGCCAGCCCCCTGACTACAAGCACTT  
10 AAGCGCTCAACAGTTTCGTTACAGCAACACCTGACCCACCAGTTAAGAAAGAACCGGTGCGAGAAGTTGCG  
CGAATTCGACATGGATCTGCCGGGCGGTACCCTGAAAGTCCGTATGTACCGCCCCGGAAGGTGTTGAACCACC  
ATATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGTTGGTGACCTGGAAACGCATGATCCGGTGTGT  
CGTGTGTTGGCGAAATGGGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTTT  
CCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGATC  
15 CAGCACGCATTGCAGTTGGCGGCGATTGAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAAA  
GAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTTTACTGGTTATGATCCAGCACATCCACC  
AGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAACCGGAGGCATGATGCTGTGGTTTCGTGACCAGTAC  
CTGAACAGCCTTGAGGAACTGACTCACCCTATGGTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCAC  
CGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTTACGCTGAAGCCCTGAACA  
20 AAGCGGGCGTTAAGGTGGAGATCGAGAACTTCGAGGATCTGATTCATGGCTTCGCGCAGTTCTATTCTCTGAG  
CCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

>S00053124  
SEQ ID NO:37  
25 TIPO: ADN, ORGANISMO: ARTIFICIAL

ATGCCATTAGATCCTGTGATTCAACAAGTCCTCGATCAACTGAACCGTATGCCAGCCCCCTGACTACAAGCACTT  
AAGCGCTCAACAGTTTCGTTACAGCAAAGCCTGTTTCCACCAGTTAAGAAAGAACCGGTGCGAGAAGTTGCG  
CGAATTCGACATGGATCTGCCGGGCGGTACCCTGAAAGTCCGTATGTACCGCCCCGGAAGGTGTTGAACCACC  
30 ATATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGTTGGTGACCTGGAAACGCATGATCCGGTGTGT  
CGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTTT  
CCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGATC  
CAGCACGCATTGCAGTTGGCGGCGATTGAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAAA  
GAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTTTACTGGTTATGATCCAGCACATCCACC  
35 AGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAACCGGAGGCATGATGCTGTGGTTTCGTAACCAGTAC  
CTGAACAGCCTTGAGGAACTGACTCACCCTATGGTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCAC  
CGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTTACGCTGAAGCCCTGAACA  
AAGCGGGCGTTAAGGTGGAGATCGAGAACTTCGAGGATCTGATTCATGGCTTCGCGCAGTTCTATTCTCTGAG  
CCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

40 >S00053286  
SEQ ID NO:39  
TIPO: ADN, ORGANISMO: ARTIFICIAL

ATGCCATTAGATCCTGTGATTCAATATGTCCTCGATCAACTGAACCGTATGCCACGTCCTGACTACAAGCACTTA  
45 AGCGCTCAACAGTTTCGTTACAGAAACCACCTGTTTCCACCAGTTAAGAAAGAACCGGTGCGAGAAGTTGCG  
GAATTCGACATGGATCTGCCGGGCGGTACCCTGAAAGTCCGTATGTACCGCCCCGGAAGGTGTTGAACCACCA  
TATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGCGGGTGACCTGGAAACGCATGATCCGGTGTGT  
CGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTTT  
50 CCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGATC  
CAGCACGCATTGCAGTTGGCGGCGATTGAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAAA  
GAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTTTACTGGTTATGATCCAGCACATCCACC  
AGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAACCGGAGGCATGATGCTGTGGTTTCGTGACCAGTAC  
CTGAACAGCCTTGAGGAACTGACTCACCCTATGGTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCAC  
55 CGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTTACGCTGAAGCCCTGAACA  
AAGCGGGCGTTGATGTGGAGATCGAGAACTTCGAGGATCTGATTCATGGCTTCGCGCAGTTCTATTCTCTGAG  
CCCAGGCGCAACCCGTGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

>S00053297  
60 SEQ ID NO:41  
TIPO: ADN, ORGANISMO: ARTIFICIAL

ATGCCATTAGATCCTGTGATTCAATATGTCCTCGATCAACTGAACCGTATGCCAGCCCCCTGACTACAAGCACTT  
AAGCGCTCAACAGTTTCGTTACAGGAACACCTGTTTCCACCAGTTAAGAAAGAACCGGTGCGAGAAGTTGCG

CGAATTCGACATGGATCTGCCGGGCCGTACCCTGAAAGTCCGTATGTACCGCCCGGAAGGTGTTGAACCACC  
 ATATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGCGGGTGACCTGGAAACGCATGATCCGGTGTG  
 TCGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTT  
 TCCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGAT  
 5 CCAGCACGCATTGCAGTTGGCGGCGATTGAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAA  
 AGAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTTTACTGGTTATGATCCAGCACATCCAC  
 CAGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAACC GGAGGCATGATGCTGTGGTTTCGTGACCAGTA  
 CCTGAACAGCCTTGAGGAAGTACTCACCATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCA  
 CCGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTACGCTGAAGCCCTGAAC  
 10 AAAGCGGGCGTTGATGTGGAGATCGAGAACTTCAGGATCTGATTCATGGCTTCGCGCAGTTCTATTCTCTGA  
 GCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

>S00053319

SEQ ID NO:43

15 TIPO: ADN, ORGANISMO: ARTIFICIAL

ATGCCATTAGATCCTGTGATTCAATATGTCCTCGATCAACTGAACCGTATGCCACGTCCTGACTACAAGCACTTA  
 AGCGCTCAACAGTTTCGTTACAGGAACACCTGTTCCACCAGTTAAGAAAGAACCGGTCGTGGAAGTTTCG  
 GAATTCGACATGGATCTGCCGGGCCGTACCCTGAAAGTCCGTATGTACCGCCCGGAAGGTGTTGAACCACCA  
 20 TATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGCGGGTGACCTGGAAACGCATGATCCGGTGTGT  
 CGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTTT  
 CCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGATC  
 CAGCACGCATTGCAGTTGGCGGCGATTGAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAAA  
 GAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTTTACTGGTTATGATCCAGCACATCCACC  
 25 AGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAACC GGAGGCATGATGCTGTGGTTTCGTGACCAGTAC  
 CTGAACAGCCTTGAGGAAGTACTCACCATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCAC  
 CGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTACGCTGAAGCCCTGAACA  
 AAGCGGGCGTTGATGTGGAGATCGAGAACTTCAGGATCTGATTCATGGCTTCGCGCAGTTCTATTCTCTGAG  
 CCCAGGCGCAACCCGTGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG  
 30

>S00053506

SEQ ID NO:45

TIPO: ADN, ORGANISMO: ARTIFICIAL

ATGCCATTAGATCCTGTGATTCAATATGTCCTCGATCAACTGAACCGTATGCCAGCCCCTGACTACAAGCACTT  
 AAGCGCTCAACAGTTTCGTTACAGAACACCTGTTCCACCAGTTAAGAAAGAACCGGTCGTGGAAGTTTCG  
 CGAATTCGACATGGATCTGCCGGGCCGTACCCTGAAAGTCCGTATGTACCGCCCGGAAGGTGTTGAACCACC  
 ATATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGCGGGTGACCTGGAAACGCATGATCCGGTGTG  
 TCGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTT  
 40 TCCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGAT  
 CCAGCACGCATTGCAGTTGGCGGCGATTGAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAA  
 AGAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTTTACTGGTTATGATCCAGCACATCCAC  
 CAGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAACC GGAGGCATGATGCTGTGGTTTCGTGACCAGTA  
 CCTGAACAGCCTTGAGGAAGTACTCACCATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCA  
 45 CCGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTACGCTGAAGCCCTGAAC  
 AAAGCGGGCGTTAAGGTGGAGATCGAGAACTTCAGGATCTGATTCATGGCTTCGCGCAGTTCTATTCTCTGA  
 GCCAGGCGCAACCCGTGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

>S00054201

50 SEQ ID NO:47

TIPO: ADN, ORGANISMO: ARTIFICIAL

ATGCCATTAGATCCTGTGATTCAACAAGTCCTCGATCAACTGAACCGTATGCCAGCCCCTGACTACAAGCACTT  
 AAGCGCTCAACAGTTTCGTTACAGCAACACCTGTTCCACCAGTTAAGAAAGAACCGGTCGAGAAAGTTTCG  
 55 CGAATTCGACATGGATCTGCCGGGCCGTACCCTGAAAGTCCGTATGTACGTGCCGGAAGGTGTTGAACCACC  
 ATATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGTTGGTGACCTGGAAACGCATGATCCGGTGTGT  
 CGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTTT  
 CCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGATC  
 CAGCACGCATTGCAGTTGGCGGCGATTGAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAAA  
 60 GAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTTTACTGGTTATGATTATGCACATCCACC  
 AGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAGAAGGAGGCATGATGCTGTGGTTTCGTGACCAGTAC  
 CTGAACAGCCTTGAGGAAGTACTCACCATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCAC  
 CGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTACGCTACCGCCCTGAACA  
 AAGCGGGCGTTAAGGTGGAGATCGAGAACTTCAGGATCTGATTCATGGCTTCGCGCAGTTCTATTCTCTGAG

CCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

>S00054685

SEQ ID NO:49

5 TIPO: ADN, ORGANISMO: ARTIFICIAL

ATGCCATTAGATCCTGTGATTCAATATGTCCTCGATCAACTGAACCGTATGCCAGCCCCTGACTACAAGCACTT  
AAGCGCTCAACAGTTTCGTTACAGGAACACCTGTTCCACCAGTTAAGAAAGAACCGGTGCGGAAGTTCCG  
CGAATTCGACATGGATCTGCCGGGCCGTACCCTGAAAGTCCGTATGTACCGCCCGGAAGGTGTTGAACCACC  
10 ATATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGCGGGTGACCTGGAAACGCATGATCCGGTGTG  
TCGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTT  
TCCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCGCGATTTCCATCTTGAT  
CCAGCACGCATTGCAGTTGGCGGCGATTACAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAA  
AGAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTTTACTGGTTATGATCCAGCACATCCAC  
15 CAGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAACCGGAGGCATGATGCTGTGGTTTCGTGACCAGTA  
CCTGAACAGCCTTGAGGAAGTACTCACCATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCA  
CCGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCCGGCAAACCTTACGCTGAAGCCCTGAAC  
AAAGCGGGCGTTAAGGTGGAGATCGAGAACTTCGAGGATCTGATTCATGGCTTCGCGCAGTTCTATTCTCTGA  
GCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG  
20

>S00054833

SEQ ID NO:51

25 TIPO: ADN, ORGANISMO: ARTIFICIAL

ATGCCATTAGATCCTGTGATTCAATATGTCCTCGATCAACTGAACCGTATGCCAGCCCCTGACTACAAGCACTT  
AAGCGCTCAACAGTTTCGTTACAGAACACCTGTTCCACCAGTTAAGAAAGAACCGGTGCGAGAAGTTCCG  
CGAATTCGACATGGATCTGCCGGGCCGTACCCTGAAAGTCCGTATGTACCGCCCGGAAGGTGTTGAACCACC  
30 ATATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGCGGGTGACCTGGAAACGCATGATCCGGTGTG  
TCGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTT  
TCCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGAT  
CCAGCACGCATTGCAGTTGGCGGCGATTACAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAA  
AGAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTTTACTGGTTATGATCCAGCACATCCAC  
CAGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAACCGGAGGCATGATGCTGTGGTTTCGTGACCAGTA  
35 CCTGAACAGCCTTGAGGAAGTACTCACCATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCA  
CCGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCCGGCAAACCTTACGCTGAAGCCCTGAAC  
AAAGCGGGCGTTAAGGTGGAGATCGAGAACTTCGAGGATCTGATTCATGGCTTCGCGCAGTTCTATTCTCTGA  
GCCAGGCGCAACCCGTGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

40 >S00054853

SEQ ID NO:53

TIPO: ADN, ORGANISMO: ARTIFICIAL

ATGCCATTAGATCCTGTGATTCAATATGTCCTCGATCAACTGAACCGTATGCCAGCCCCTGACTACAAGCACTT  
45 AAGCGCTCAACAGTTTCGTTACAGCAACACCTGTTCCACCAGTTAAGAAAGAACCGGTGCGAGAAGTTCCG  
CGAATTCGACATGGATCTGCCGGGCCGTACCCTGAAAGTCCGTATGTACCGCCCGGAAGGTGTTGAACCACC  
ATATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGCGGGTGACCTGGAAACGCATGATCCGGTGTG  
TCGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTT  
TCCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGAT  
50 CCAGCACGCATTGCAGTTGGCGGCGATTACAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAA  
AGAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTTTACTGGTTATGATCCAGCACATCCAC  
CAGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAACCGGAGGCATGATGCTGTGGTTTCGTGACCAGTA  
CCTGAACAGCCTTGAGGAAGTACTCACCATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCA  
CCGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCCGGCAAACCTTACGCTGAAGCCCTGAAC  
55 AAAGCGGGCGTTAAGGTGGAGATCGAGAACTTCGAGGATCTGATTCATGGCTTCGCGCAGTTCTATTCTCTGA  
GCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

>S00058614

SEQ ID NO:55

60 TIPO: ADN, ORGANISMO: ARTIFICIAL

ATGCCATTAGATCCTGTGATTCAATATGTCCTCGATCAACTGAACCGTATGCCAGCCCCTGACTACAAGCACTT  
AAGCGCTCAACAGTTTCGTTACAGCAACACCTGTTCCACCAGTTAAGAAAGAACCGGTGCGAGAAGTTCCG  
CGAATTCGACATGGATCTGCCGGGCCGTACCCTGAAAGTCCGTATGTACCGCCCGGAAGGTGTTGAACCACC

ATATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGCGGGTGACCTGGAAACGCATGATCCGGTGTG  
TCGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACAACAAGTT  
TCCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGAT  
CCAGCACGCATTGCAGTTGGCGGCGATTACAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAA  
5 AGAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTTTACTGGTTATGATCCAGCACATCCAC  
CAGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAACCGGAGGCATGATGCTGTGGTTTCGTGACCAGTA  
CCTGAACAGCCTTGAGGAACTGACTACCCATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCA  
CCGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTACGCTGAAGCCCTGAAC  
AAAGCGGGCGTTAAGGTGGAGATCGAGAACTTCGAGGATCTGATTATGGCTTCGCGAGTTTCTATTCTCTGA  
10 GCCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

>S00059197  
SEQ ID NO:57  
TIPO: ADN, ORGANISMO: ARTIFICIAL

15 ATGCCATTAGATCCTGTGATTCAATATGTCCTCGATCAACTGAACCGTATGCCAGCCCCCTGACTACAAGCACTT  
AAGCGCTCAACAGGTTTCGTTTACAGCAACACCTGTTTCCACCAGTTAAGAAAGAACCGGTGCGAGAAGTTTCG  
CGAATTCGACATGGATCTGCCGGGCCGTACCTGAAAGTCCGTATGTACCGCCCCGAAGGTGTTGAACCACC  
ATATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGCGGGTGACCTGGAAACGCATGATCCGGTGTG  
20 TCGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACAACAAGTT  
TCCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGAT  
CCAGCACGCATTGCAGTTGGCGGCGATTACAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAA  
AGAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTTTACTGGTTATGATCCAGCACATCCAC  
CAGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAACCGGAGGCATGATGCTGTGGTTTCGTGACCAGTA  
25 CCTGAACAGCCTTGAGGAACTGACTACCCATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCA  
CCGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTACGCTGAAGCCCTGAAC  
AAAGCGGGCGTTAAGGTGGAGATCGAGAACTTCGAGGATCTGATTATGGCTTCGCGCAGTTCTATTCTCTGA  
GCCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

30 >S00059305  
SEQ ID NO:59  
TIPO: ADN, ORGANISMO: ARTIFICIAL

ATGCCATTAGATCCTGTGATTCAATATGTCCTCGATCAACTGAACCGTATGCCAGCCCCCTGACTACAAGCACTT  
35 AAGCGCTCAACAGTTTTCGTTTACAGGAACACCTGTTTCCACCAGTTAAGAAAGAACCGGTGCGAGAAGTTTCG  
CGAATTCGACATGGATCTGCCGGGCCGTACCTGAAAGTCCGTATGTACCGCCCCGAAGGTGTTGAACCACC  
ATATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGCGGGTGACCTGGAAACGCATGATCCGGTGTG  
TCGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACAACAAGTT  
TCCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGAT  
40 CCAGCACGCATTGCAGTTGGCGGCGATTACAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAA  
AGAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTTTACTGGTTATGATCCAGCACATCCAC  
CAGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAACCGGAGGCATGATGCTGTGGTTTCGTGACCAGTA  
CCTGAACAGCCTTGAGGAACTGACTACCCATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCA  
CCGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTACGCTGAAGCCCTGAAC  
45 AAAGCGGGCGTTAAGGTGGAGATCGAGAACTTCGAGGATCTGATTATGGCTTCGCGCAGTTCTATTCTCTGA  
GCCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

>S00059320  
SEQ ID NO:61  
50 TIPO: ADN, ORGANISMO: ARTIFICIAL

ATGCCATTAGATCCTGTGATTCAATATGTCCTCGATCAACTGAACCGTATGCCAGCCCCCTGACTACAAGCACTT  
AAGCGCTCAACAGTTTTCGTTTACAGCAACACCTGTTTCCACCAGTTAAGAAAGAACCGGTGCGAGAAGTTTCG  
CGAATTCGACATGGATCTGCCGGGCCGTACCTGAAAGTCCGTATGTACCGCCCCGAAGGTGTTGAACCACC  
55 ATATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGCGGGTGACCTGGAAACGCATGATCCGGTGTG  
TCGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACAACAAGTT  
TCCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGAT  
CCAGCACGCATTGCAGTTGGCGGCGATTACAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAA  
AGAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTTTACTGGTTTTGATCCAGCACATCCAC  
60 CAGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAACCGGAGGCATGATGCTGTGGTTTCGTGACCAGTA  
CCTGAACAGCCTTGAGGAACTGACTACCCATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCA  
CCGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTACGCTGAAGCCCTGAAC  
AAAGCGGGCGTTAAGGTGGAGATCGAGAACTTCGAGGATCTGATTATGGCTTCGCGCAGTTCTATTCTCTGA  
GCCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

>S00059754  
SEQ ID NO:63  
TIPO: ADN, ORGANISMO: ARTIFICIAL

5 ATGCCATTAGATCCTGTGATTCAATATGTCCTCGATCAACTGAACCGTATGCCAGCCCCTGACTACAAGCACTT  
AAGCGCTCAACAGTTTCGTTACAGCAACACCTGTTCCACCAGTTAAGAAAGAACCGGTGCGAGAAGTTCCG  
CGAATTCGACATGGATCTGCCGGGCCGTACCCTGAAAGTCCGTATGTACCGCCCGGAAGGTGTTGAACCACC  
ATATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGCGGGTGACCTGGAAACGCATGATCCGGTGTG  
10 TCGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTT  
TCCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGAT  
CCAGCACGCATTGCAGTTGGCGGCGATTACAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAA  
AGAACGTGGTGGTCCAGCAATTGCGTTTCACTGCTGATCTATCCCTTTACTGGTTATGATCCAGCACATCCAC  
CAGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAACC GGAGGCATGATGCTGTGGTTTCGTATT CAGTA  
15 CCTGAACAGCCTTGAGGAACTGACTCACCCTATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCA  
CCGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTACGCTGAAGCCCTGAAC  
AAAGCGGGCGTTAAGGTGGAGATCGAGAACTTCGAGGATCTGATT CATGGCTTCGCGCAGTTCTATTCTCTGA  
GCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

20 >S00059759  
SEQ ID NO:65  
TIPO: ADN, ORGANISMO: ARTIFICIAL

25 ATGCCATTAGATCCTGTGATTCAATATGTCCTCGATCAACTGAACCGTATGCCAGCCATTGACTACAAGCACTTA  
AGCGCTCAACAGTTTCGTTACAGCAACACCTGTTCCACCAGTTAAGAAAGAACCGGTGCGAGAAGTTCCG  
GAATTCGACATGGATCTGCCGGGCCGTACCCTGAAAGTCCGTATGTACCGCCCGGAAGGTGTTGAACCACCA  
TATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGCGGGTGACCTGGAAACGCATGATCCGGTGTGT  
CGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTTT  
CCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGATC  
30 CAGCACGCATTGCAGTTGGCGGCGATTACAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAAA  
GAACGTGGTGGTCCAGCAATTGCGTTTCACTGCTGATCTATCCCTTTACTGGTTATGATCCAGCACATCCACC  
AGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAACC GGAGGCATGATGCTGTGGTTTCGTGACCAGTAC  
CTGAACAGCCTTGAGGAACTGACTCACCCTATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCAC  
CGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTACGCTGAAGCCCTGAACA  
35 AAGCGGGCGTTAAGGTGGAGATCGAGAACTTCGAGGATCTGATT CATGGCTTCGCGCAGTTCTATTCTCTGAG  
CCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

>S00059902  
SEQ ID NO:67  
40 TIPO: ADN, ORGANISMO: ARTIFICIAL

ATGCCATTAGATCCTGTGATTCAATATGTCCTCGATCAACTGAACCGTATGCCAGCCCCTGACTACAAGCACTT  
AAGCGCTCAACAGTTTCGTTACAGCAACACCTGTTCCACCAGTTAAGAAAGAACCGGTGCGAGAAGTTCCG  
CGAATTCGACATGGATCTGCCGGGCCGTACCCTGAAAGTCCGTATGTACCGCCCGGAAGGTGTTGAACCACC  
45 ATATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGCGGGTGACCTGGAAACGCATGATCCGGTGTG  
TCGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTT  
TCCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGAT  
CCAGCACGCATTGCAGTTGGCGGCGATTACAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAA  
AGAACGTGGTGGTCCAGCAATTGCGTTTCACTGCTGATCTATCCCTTTACTGGTTATGATCCAGCACATCCAC  
50 CAGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAACC GGAGGCATGATGCTGTGGTTTCGTAATCAGTA  
CCTGAACAGCCTTGAGGAACTGACTCACCCTATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCA  
CCGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTACGCTGAAGCCCTGAAC  
AAAGCGGGCGTTAAGGTGGAGATCGAGAACTTCGAGGATCTGATT CATGGCTTCGCGCAGTTCTATTCTCTGA  
GCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

55 >S00060141  
SEQ ID NO:69  
TIPO: ADN, ORGANISMO: ARTIFICIAL

60 ATGCCATTAGATCCTGTGATTCAATATGTCCTCGATCAACTGAACCGTATGCCAGCCCCTGACTACAAGCACTT  
AAGCGCTCAAGATTTTCGTTACAGCAACACCTGTTTCCACCAGTTAAGAAAGAACCGGTGCGAGAAGTTCCG  
CGAATTCGACATGGATCTGCCGGGCCGTACCCTGAAAGTCCGTATGTACCGCCCGGAAGGTGTTGAACCACC  
ATATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGCGGGTGACCTGGAAACGCATGATCCGGTGTG  
TCGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTT

TCCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGAT  
 CCAGCACGCATTGCAGTTGGCGGCGATTACAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAA  
 AGAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTTTACTGGTTATGATCCAGCACATCCAC  
 CAGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAACCGGAGGCATGATGCTGTGGTTTCGTGACCAGTA  
 5 CCTGAACAGCCTTGAGGAACTGACTCACCCTATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCA  
 CCGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTACGCTGAAGCCCTGAAC  
 AAAGCGGGCGTTAAGGTGGAGATCGAGAACTTCGAGGATCTGATTCATGGCTTCGCGCAGTTCTATTCTCTGA  
 GCCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

10 >S00060676  
 SEQ ID NO:71  
 TIPO: ADN, ORGANISMO: ARTIFICIAL

ATGCCATTAGATCCTGTGATTCAATATGTCCTCGATCAACTGAACCGTATGCCAGCCCCCTGACTACAAGCACTT  
 15 AAGCGCTCAACAGTTTCGTTCAAGTCAACACCTGTTTCCACCAGTTAAGAAAGAACCGGTGCGAGAAGTTGCG  
 CGAATTCGACATGGATCTGCCGGGCCGTACCCTGAAAGTCCGTATGTACCGCCCCGGAAGGTGTTGAACCACC  
 ATATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGCGGGTGACCTGGAAACGCATGATCCGGTGTG  
 TCGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTT  
 TCCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGAT  
 20 CCAGCACGCATTGCAGTTGGCGGCGATTACAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAA  
 AGAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTTTACTGGTTATGATCCAGCACATCCAC  
 CAGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAACCGGAGGCATGATGCTGTGGTTTCGTGACCAGTA  
 CCTGAACAGCCTTGAGGAACTGACTCACCCTATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCA  
 CCGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTACGCTGAAGCCCTGAAC  
 25 AAAGCGGGCGTTAAGGTGGAGATCGAGAACTTCGAGGATCTGATTCATGGCTTCGCGCAGTTCTATTCTCTGA  
 GCCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

>S00060772  
 SEQ ID NO:73  
 30 TIPO: ADN, ORGANISMO: ARTIFICIAL

ATGCCATTAGATCCTGTGATTCAATATGTCCTCGATCAACTGAACCGTATGCCAGCCCCCTGACTACAAGCACTT  
 AAGCGCTAGTCAGTTTCGTTACAGCAACACCTGTTTCCACCAGTTAAGAAAGAACCGGTGCGAGAAGTTGCG  
 CGAATTCGACATGGATCTGCCGGGCCGTACCCTGAAAGTCCGTATGTACCGCCCCGGAAGGTGTTGAACCACC  
 35 ATATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGCGGGTGACCTGGAAACGCATGATCCGGTGTG  
 TCGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTT  
 TCCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGAT  
 CCAGCACGCATTGCAGTTGGCGGCGATTACAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAA  
 AGAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTTTACTGGTTATGATCCAGCACATCCAC  
 40 CAGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAACCGGAGGCATGATGCTGTGGTTTCGTGACCAGTA  
 CCTGAACAGCCTTGAGGAACTGACTCACCCTATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCA  
 CCGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTACGCTGAAGCCCTGAAC  
 AAAGCGGGCGTTAAGGTGGAGATCGAGAACTTCGAGGATCTGATTCATGGCTTCGCGCAGTTCTATTCTCTGA  
 GCCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

45 >S00066575  
 SEQ ID NO:75  
 TIPO: ADN, ORGANISMO: ARTIFICIAL

ATGCCATTAGATCCTGTGATTCAATATGTGCTCGATCAACTGAACCGTATGCCAGCCCCCTGACTACAAGCACTT  
 AAGCGCTCAACAGGTTTCGTTACACAACACCTGTTTCCACCAGTTAAGAAAGAACCGGTGCGAGAAGTTGCG  
 CGAATTCGACATGGATCTGCCGGGCCGTACCCTGAAAGTCCGTATGTACCGCCCCGGAAGGTGTTGAACCACC  
 ATATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGCGGGTGACCTGGAAACGCATGATCCGGTGTG  
 TCGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTT  
 55 TCCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGAT  
 CCAGCACGCATTGCAGTTGGCGGCGATTACAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAA  
 AGAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTTTACTGGTTATGATCCAGCACATCCAC  
 CAGCAAGTCTTCTTGAGAATGCGGAGGGTTACCTGTTAACCGGAGGCATGATGCTGTGGTTTCGTGAATCAGTA  
 CCTGAACAGCCTTGAGGAACTGACTCACCCTATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCA  
 CCGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTACGCTGAAGCCCTGAAC  
 60 AAAGCGGGCGTTAAGGTGGAGATCGAGAACTTCGAGGATCTGATTCATGGCTTCGCGCAGTTCTATTCTCTGA  
 GCCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

>S00067402

# ES 3 018 269 T3

SEQ ID NO:77

TIPO: ADN, ORGANISMO: ARTIFICIAL

5 ATGCCATTAGATCCTGTGATTCAATATGTGCTCGATCAACTGAACCGTATGCCAGCCCCTGACTACAAGCACTT  
AAGCGCTCAACAGTTTCGTTACAGCAACACATTGGTCCACCAGTTAAGAAAGAACCGGTGCGAGAAGTTTCG  
CGAATTCGACATGGATCTGCCGGGGCCGTACCCTGAAAGTCCGTATGTACCGCCCCGGAAGGTGTTGAACCACC  
ATATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGCGGGTGACCTGGAAACGCATGATCCGGTGTG  
TCGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTT  
10 TCCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGAT  
CCAGCACGCATTGCAGTTGGCGGCGATTACAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAA  
AGAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTTTACTGGTTATGATCCAGCACATCCAC  
CAGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAACCGGAGGCATGATGCTGTGGTTTCGTAATCAGTA  
CCTGAACAGCCTTGAGGAACTGACTACCCATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCA  
CCGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTTACGCTGAAGCCCTGAAC  
15 AAAGCGGGCGTTAAGGTGGAGATCGAGAACCTCGAGGATACGATTCATGGCTTCGCGCAGTTTTATTCTATTA  
GCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

>S00067403

SEQ ID NO:79

20 TIPO: ADN, ORGANISMO: ARTIFICIAL

ATGCCATTAGATCCTGTGATTCAATATGTCCTCGATCAACTGAACCGTATGCCAGCCCCTGACTACAAGCACTT  
AAGCGCTCAACAGTTTCGTTACAGCAACACATGGGTCCACCAGTTAAGAAAGAACCGGTGCGAGAAGTTTCG  
CGAATTCGACATGGATCTGCCGGGGCCGTACCCTGAAAGTCCGTATGTACCGCCCCGGAAGGTGTTGAACCACC  
25 ATATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGCGGGTGACCTGGAAACGCATGATCCGGTGTG  
TCGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTT  
TCCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGAT  
CCAGCACGCATTGCAGTTGGCGGCGATTACAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAA  
AGAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTTTACTGGTTATGATCCAGCACATCCAC  
30 CAGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAACCGGAGGCATGATGCTGTGGTTTCGTAATCAGTA  
CCTGAACAGCCTTGAGGAACTGACTACCCATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCA  
CCGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTTACGCTGAAGCCCTGAAC  
AAAGCGGGCGTTAAGGTGGAGATCGAGAACCTCGAGGATATGATTCATGGCTTCGCGCAGTTCTATTCTATTAG  
CCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

>S00067700

SEQ ID NO:81

TIPO: ADN, ORGANISMO: ARTIFICIAL

40 ATGCCATTAGATCCTGTGATTCAATATATGCTCGATCAACTGAACCGTATGCCAGCCCCTGACTACAAGCACTTA  
AGCGCTCAACAGTTTCGTTACAGCAACACATTGGGTCCACCAGTTAAGAAAGAACCGGTGCGAGAAGTTTCG  
GAATTCGACATGGATCTGCCGGGGCCGTACCCTGAAAGTCCGTATGTACCGCCCCGGAAGGTGTTGAACCACCA  
TATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGCGGGTGACCTGGAAACGCATGATCCGGTGTGT  
CGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTTT  
45 CCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGATC  
CAGCACGCATTGCAGTTGGCGGCGATTACAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAAA  
GAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTTTACTGGTTATGATCCAGCACATCCACC  
AGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAACCGGAGGCATGATGCTGTGGTTTCGTAATCAGTAC  
CTGAACAGCCTTGAGGAACTGACTACCCATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCAC  
50 CGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTTACGCTGAAGCCCTGAACA  
AAGCGGGCGTTAAGGTGGAGATCGAGAACCTCGAGGATCTAATTTCATGGCTTCGCGCAGTTCTATTCTCTGAG  
CCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

>S00068003

55 SEQ ID NO:83

TIPO: ADN, ORGANISMO: ARTIFICIAL

ATGCCATTAGATCCTGTGATTCAATATGTCCTCGATCAACTGAACCGTATGCCAGCCCCTGACTACAAGCACTT  
AAGCGCTAGTCAGTTTCGTTACAGCAACACCTGTTTCCACCAGTTAAGAAAGAACCGGTGCGAGAAGTTTCG  
60 CGAATTCGACATGGATCTGCCGGGGCCGTACCCTGAAAGTCCGTATGTACCGCCCCGGAAGGTGTTGAACCACC  
ATATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGCGGGTGACCTGGAAACGCATGATCCGGTGTG  
TCGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTT  
TCCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGAT  
CCAGCACGCATTGCAGTTGGCGGCGATTACAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAA

AGAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTTTACTGGTTACGATCCAGCACATCCA  
CCAGCAAGTATTCTTGAGAATGCGGAGGGTTACCTGTTAACCAGGAGGCATGATGCTGTGGTTTCGTAATCAGT  
ACCTGAACAGCCTTGAGGAACTGACTCACCCATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACC  
ACCGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGTAAACTTTACGCTGAAGCCCTGAA  
5 CAAAGCGGGCGTTAAGGTGGAGATCGAGAACTTCGAGGATCTGATTCATGGCTTCGCGCAGTTCTATTCTCTG  
AGCCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

>S00068079

SEQ ID NO:85

10 TIPO: ADN, ORGANISMO: ARTIFICIAL

ATGCCATTAGATCCTGTGATTCAATATGTCCTCGATCAACTGAACCGTATGCCAGCCCCCTGACTACAAGCACTT  
AAGCGCTAGTCAGTTTCGTTTACAGCAACATCTGTTTCCACCAGTTAAGAAAGAACCGGTCGCAGAAAGTTCGC  
GAATTCGACATGGATCTGCCGGGCGGTACCCTGAAAGTCCGTATGTACCGCCCGGAAGGTGTTGAACCACCA  
15 TATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGCGGGTGACCTGGAACGCATGATCCGGTGTGT  
CGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTTT  
CCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGATC  
CAGCACGCATTGCAGTTGGCGGCGATTGAGCGGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAAA  
GAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTTTACTGGTTACGATCCAGCACATCCAC  
20 CAGCAAGTCTTCTTGAGAATGCGGAGGGTTACCTGTTAACCAGGAGGCATGATGCTGTGGTTTCGTAATCAGTA  
CCTGAACAGCCTTGAGGAACTGACTCACCCATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCA  
CCGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGTAAACTTTACGCTGAAGCCCTGACC  
AAAGCGGGCGTTAAGGTGGAGATCGAGAACTTCGAGGATCTGATTCATGGCTTCGCGCAGTTCTATTCTCTGA  
GCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

25

>S00068124

SEQ ID NO:87

TIPO: ADN, ORGANISMO: ARTIFICIAL

ATGCCATTAGATCCTGTGATTCAATATGTCCTCGATCAACTGAACCGTATGCCAGCCCCCTGACTACAAGCACTT  
AAGCGCTAGTCAGTTTCGTTTACAGCAACACCTGTTTCCACCAGTTAAGAAAGAACCGGTCGCAGAAAGTTCG  
CGAATTCGACATGGATCTGCCGGGCGGTACCCTGAAAGTCCGTATGTACCGCCCGGAAGGTGTTGAACCACC  
ATATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGCGGGTGACCTGGAACGCATGATCCGGTGTG  
TCGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTT  
35 TCCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGAT  
CCAGCACGCATTGCAGTTGGCGGCGATTGAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAA  
AGAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTTTACTGGTTATGATCCAGCACATCCAC  
CAGCAAGTCTTCTTGAGAATGCGGAGGGTTACCTGTTAACCAGGAGGCATGATGCTGTGGTTTCGTAATCAGTA  
CCTGAACAGCCTTGAGGAACTGACTCACCCATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCA  
40 CCGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGTAAACTTTACGCTGAAGCCCTGAAC  
AAAGCGGGCGTTAAGGTGGAGATCGAGAACTTCGAGGATCTGATTCATGGCTTCGCGCAGTTCTATTCTCTGA  
GCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

>S00068512

45 SEQ ID NO:89

TIPO: ADN, ORGANISMO: ARTIFICIAL

ATGCCATTAGATCCTGTGATTCAATATGTCCTCGATCAACTGAACCGTATGCCAGCCCCCGACTACAAGCACTT  
AAGCGCTCAACAAGTTTCGTTTATGGCAACACCTGTTTCCACCAGTTAAGAAAGAACCGGTCGCAGAAAGTTCG  
50 CGAATTCGACATGGATCTGCCGGGCGGTACCCTGAAAGTCCGTATGTACCGCCCGGAAGGTGTTGAACCACC  
ATATCGGGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGCGGGTGACCTGGAAACGCATGATCCGGTGTG  
TCGTGTGTTGGCGAAATGGGGACGCGCAGTGGTGTTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTT  
TCCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGAT  
CCAGCACGCATTGCAGTTGGCGGCGATTGAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAA  
55 AGAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTTTACTGGATATGATCCAGCACATCCAC  
CAGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAACCAGGAGGCATGATGCTGTGGTTTCGTAATCAGTA  
CCTGAACAGCCTTGAGGAACTGACTCACCCATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCA  
CCGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTTACGCTGAAGCCCTGAAC  
AAAGCGGGCGTTAAGGTGGAGATCGAGAACTTCGAGGATCTGATTCATGGCTTCGCGCAGTTCTATTCTCTGA  
60 GCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

>S00068639

SEQ ID NO:91

TIPO: ADN, ORGANISMO: ARTIFICIAL

ATGCCATTAGATCCTGTGATTCAATATGTCCTCGATCAACTGAACCGTATGCCAGCCATTGACTACAAGCACTTA  
 AGCGCTCAACAAGTTTCGTTTCATGGCAACACCTGTTTCCACCAGTTAAGAAAGAACCAGGTCCGCGAAGTTTCGC  
 GAATTCGACATGGATCTGCCGGGCGGTACCCTGAAAGTCCGTATGTACACGCCGGAAGGTGTTGAACCACCA  
 5 TATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGCGGGTGACCTGGAACGCATGATCCGGTGTGT  
 CGTGTGTTGGCGAAATGGGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTTT  
 CCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGATC  
 CAGCACGCATTGCAGTTGGCGGCGATTGACGAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAAA  
 GAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTTTACTGGATATGATCCAGCACATCCACC  
 10 AGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAACCGGAGGCATGATGCTGTGGTTTCGTAATCAGTAC  
 CTGAACAGCCTTGAGGAAGTACTCACCCTGTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCAC  
 CGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTTACGCTAGAGCCCTGAACA  
 AAGCGGGCGTTAAAGTGGAGATCGAGAACTTCGAGGATCTGATTGCTTTCGCGCAGTTCTATTCTCTGAG  
 CCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

15  
 >S00068752  
 SEQ ID NO:93  
 TIPO: ADN, ORGANISMO: ARTIFICIAL

20 ATGCCATTAGATCCTGTGATTCAATATGTCCTCGATCAACTGAACCGTATGCCAGCCCCTGACTGGAAGCACTT  
 AAGCGCTCAACAGTTTCGTTTCACAGCAACACCTGTTTCCACCAGTTAAGAAAGAACCAGGTCCGAGAAGTTTCG  
 CGAATTCGACATGGATCTGCCGGGCGGTACCCTGAAAGTCCGTATGTACCGCCCCGGAAGGTGTTGAACCACC  
 ATATCGGGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGCGGGTGACCTGGAACGCATGATCCGGTGTG  
 TCGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTT  
 25 TCCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGAT  
 CCAGCACGCATTGCAGTTGGCGGCGATTGACGCGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAA  
 AGAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTTTACTGGTTATGATCCAGCACATCCAC  
 CAGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAACCGGAGGCATGATGCTGTGGTTTCGTAATCAGTA  
 CCTGAACAGCCTTGAGGAAGTACTCACCCTGTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCA  
 30 CCGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTTACGCTGAAGCCCTGAAC  
 AAAGCGGGCGTTAAGGTGGAGATCGAGAACTTCGAGGATCTGATTGCTTTCGCGCAGTTCTATTCTATTA  
 GCCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

>S00176502  
 35 SEQ ID NO:95  
 TIPO: ADN, ORGANISMO: ARTIFICIAL

ATGCCATTAGATCCTGTGATTCAATATGTCCTCGATCAACTGAACCGTATGCCAGCCCCTGACTACAAGCACTT  
 AAGCGCTCAACAGTTTCGTTTCACAGCAACACCTGTTTCCACCAGTTAAGAAAGAACCAGGTCAAAGAAGTTTCG  
 40 CGAATTCGACATGGATCTGCCGGGCGGTACCCTGAAAGTCCGTATGTACCGCCCCGGAAGGTGTTGAACCACC  
 ATATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGCGGGTGACCTGGAACGCATGATCCGGTGTG  
 TCGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTT  
 TCCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGAT  
 CCAGCACGCATTGCAGTTGGCGGCGATTGACGAGGCGGCAACACCGCGGCCGTGACTAGTATTCTGGCGAA  
 45 AGAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTTTACTGGTTATGATCCAGCACATCCAC  
 CAGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAACCGGAGGCATGATGCTGTGGTTTCGTAATCAGTA  
 CCTGAACAGCCTTGAGGAAGTACTCACCCTGTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCA  
 CCGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTTACGCTGAAGCCCTGAAC  
 AAAGCGGGCGTTAAGGTGGAGATCGAGAACTTCGAGGATCTGATTGCTTTCGCGCAACTTCTATTCTCTGA  
 50 GCCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

>S00176595  
 SEQ ID NO:97  
 TIPO: ADN, ORGANISMO: ARTIFICIAL

55 ATGCCATTAGATCCTGTGATTCAATATGTCCTCGATCAACTGAACCGTATGCCAGCCCCTGACTACAAGCACTT  
 AAGCGCTCAACAGTTTCGTTTCACAGCAACACCTGTTTCCACCAGTTAAGAAAGAACCAGGTCAAAGAAGTTTCG  
 CGAATTCGACATGGATCTGCCGGGCGGTACCCTGAAAGTCCGTATGTACCGCCCCGGAAGGTGTTGAACCACC  
 ATATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGGTTGCGGGTGACCTGGAACGCATGATCCGGTGTG  
 60 TCGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTT  
 TCCAGCGGCGAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGAT  
 CCAGCACGCATTGCAGTTGGCGGCGATTGACGAGGCGGCAACACCGCGGCCGTGACTAGTATTCTGGCGAA  
 AGAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTTTACTGGTTATGATCCAGCACATCCAC  
 CAGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAACCGGAGGCATGATGCTGTGGTTTCGTAATCAGTA

# ES 3 018 269 T3

CCTGAACAGCCTTGAGGAACTGACTCACCCATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCA  
CCGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTTACGCTGAAGCCCTGAAC  
AAAGCGGGCGTTAAGGTGGAGATCGAGAACTTCGAGGATCTGATTATGGCTTCGCGAGCTTCTATTCTCTGA  
GCCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

5

>S00176980  
SEQ ID NO:99  
TIPO: ADN, ORGANISMO: ARTIFICIAL

10 ATGCCATTAGATCCTGTGATTCAATATGTCCTCGATCAACTGAACCGTATGCCAGCCCCTGACTACAAGCACTT  
AAGCGCTTGCCAGTTTCGTTACAGCAACACCTGTTTGAACCAAGTAAAGAAAGAACCGGTGCGAGAAGTTCCG  
CGAATTCGACATGGATCTGCCGGGGCCGTACCCCTGAAAGTCCGTATGTACCGCCCGGAAGGTGTTGAACCAAC  
ATATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGCGTTGCGGGTGACCTGGAAACGCATGATCCGGTGTG  
TCGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTT  
15 TCCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGAT  
CCAGCACGCATTGCAGTTGGCGGGCGATTACAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAA  
AGAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTTTACTGGTTATGATCCAGCACATCCAC  
CAGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAACCAGGAGGCATGATGCTGTGGTTTCGTAATCAGTA  
CCTGAACAGCCTTGAGGAACTGACTCACCCATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCA  
20 CCGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTTACGCTGAAGCCCTGAAC  
AAAGCGGGCGTTAAGGTGGAGATCGAGAACTTCGAGGATCTGATTATGGCTTCGCGCAGTTCTATTCTCTGA  
GCCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

>S00177019  
25 SEQ ID NO:101  
TIPO: ADN, ORGANISMO: ARTIFICIAL

ATGCCATTAGATCCTGTGATTCAATATGTCCTCGATCAACTGAACCGTATGCCAGCCCCTGACTACAAGCACTT  
AAGCGCTCAACAGTTTCGTTACAGCAAAGCCTGTTTCCACCAGTTAAGAAAGAACCGGTGCGAGAAGTTCCG  
30 CGAATTCGACATGGATCTGCCGGGGCCGTACCCCTGAAAGTCCGTATGTACCTGCCGGAAGGTGTTGAACCAAC  
ATATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGACCGCGGGTGACCTGGAAACGCATGATCCGATGTG  
TCGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTT  
TCCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGAT  
CCAGCACGCATTGCAGTTGGCGGGCGATTACAGCAGGCGGCAACCCGGCGGCCGTGACTAGTATTCTGGCGAA  
35 AGAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCCAGCCCTTTACTGGTTTGGATCCAGCACATCCA  
CCAGCAAGTATCCTTGAGAATGCGGAGGGTTACTTTTAAACCAGGAGGCATGATGGATTGGTTTCGTAATCAGTA  
CCTGAACAGCCTTGAGGAACTGACTCACCCATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCA  
CCGGCGTACATTGCAACCGCACAGTACGATCCCCTGAACGATGTCGGCAAACCTTTACGCTGAAGCCCTGAAC  
AAAGCGGGCCTGAAGGTGGAGATCGAGAACTTCGAGGATCTGATTATGGCTTCGCGAGCTTCTATTCTCTG  
40 AGCCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

>S00177053  
SEQ ID NO:103  
TIPO: ADN, ORGANISMO: ARTIFICIAL

45 ATGCCATTAGATCCTGTGATTCAATATGTCCTCGATCAACTGAACCGTATGCCAGCCCCTGACTACAAGCACTT  
AAGCGCTCAACAGTTTCGTTACAGCAAAGCCTGTTTCCACCAGTTAAGAAAGAACCGGTGCGAGAAGTTCCG  
CGAATTCGACATGGATCTGCCGGGGCCGTACCCCTGAAAGTCCGTATGTACCTGCCGGAAGGTGTTGAACCAAC  
ATATCCAGCACTGGTTTACTACCATGGTGGCGGTTGGACCGCGGGTGACCTGGAAACGCATGATCCGATGTG  
50 TCGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTT  
TCCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGAT  
CCAGCACGCATTGCAGTTGGCGGGCGATTACAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAA  
AGAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTTTACTGGTTATGATCCAGCACATCCAC  
CAGCAAGTATCCTTGAGAATGCGGAGGGTTACTTTTTAACCAGGAGGCATGATGCTGTGGTTTCGTAATCAGTA  
55 CCTGAACAGCCTTGAGGAACTGACTCACCCATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCA  
CCGGCGTACATTGCAACCTGGCAGTACGATCCCCTGCGCGATGTCGGCAAACCTTTACGCTGAAGCCCTGAAC  
AAAGCGGGCCTGAAGGTGGAGATCGAGAACTTCGAGGATCTGATTATGGCTTCGCGAGCTTCTATTCTCTGA  
GCCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

60 >S00177175  
SEQ ID NO:105  
TIPO: ADN, ORGANISMO: ARTIFICIAL

ATGCCATTAGATCCTGTGATTCAATATGTCCTCGATCAACTGAACCGTATGCCAGCCCCTGACTACAAGCACTT

AAGCGCTTGCCAGTTTCGTTCCACAGCAACACCTGTTTCCACCAGTTAAGAAAGAACCGGTGCGAGAAGTTGCG  
CGAATTCGACATGCGTCTGCCGGGCGGTACCCTGAAAGTCCGTATGTACCGCCCGGAAGGTGTTGAACCACC  
ATATCGTGCAGTGGTTTACTACCATGGTGGCGGTTGGGTTGCGGGTGACCTGGAAACGCATGATCCGGTGTG  
TCGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTT  
5 TCCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGAT  
CCAGCACGCATTGCAGTTGGCGGCGATTACAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAA  
AGAACGTGGTGGTCCAGCAATGGCGTTTCAACTGCTGATCTATCCCTTTACTGGTTATGATCCAGCACATCCAC  
CAGCAAGTATCCTTGAGAATGCGGAGGGTTACTTTTTAACCGGAGGCATGATGCTGTGGTTTCGTAATCAGTA  
CCTGAACAGCCTTGAGGAACTGACTCACCCTATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCA  
10 CCGGCGTACATTGCAGTGGCACAGTACGATCCCCTGCGCGATGTGCGGCAAACCTTACGCTGAAGCCCTGAAC  
AAAGCGGGCGTTAAGGTGGAGATCGAGAACTTCGAGGATCTGATTCATGGCTTCGCGCAGTTCTATTCTCTGA  
GCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

>S00187023

15 SEQ ID NO:107

TIPO: ADN, ORGANISMO: ARTIFICIAL

ATGCCATTAGATCCTGTGCTTCAATATGTGCTCGATCAACTGAACCGTATGCCAGCCCCTGACTACAAGCACTT  
AAGCGCTTGCCAGTTTCGTTCCACAGCAACACCTGTTTGAACCAGTTAAGAAAGAACCGGTGCGTGAAGTTGCG  
20 CGAATTCGACATGGACCTGCCGGGCGGTACCCTGAAAGTCCGTATGTACCGCCCGGAAGGTGTTGAACCACC  
ATATCCAGCACTGGTTTACTACCATGGTGGCGGTGGCGTTGCGGGTGACCTGGAAACGCATGATCCGGTGTG  
TCGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTT  
TCCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGAT  
CCAGCACGCATTGCAGTTGGCGGCGATTACAGCAGGCGGCAACCTCGCGGCCGTGACTAGTATTCTGGCGAA  
25 AGAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTTTACTGGTTATGATCCAGCACATCCAC  
CAGCAAGTATCCTGGAGAATGCGGAGGGTTACCTGTTAACCGGAGGCATGATGCTGTGGTTTCGTAATCAGTA  
CCTGAACAGCCTTGAGGAACTGACTCACCCTATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCA  
CCGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTGCGGCAAACCTTACGCTGAAGCCCTGAAC  
AAAGCGGGCGTTAAGGTGGAGATCGAGAACTTCGAGGATCTGATTCATGGCTTCGCGCAGTTCTATTCTCTGA  
30 GCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

>S00187052

SEQ ID NO:109

TIPO: ADN, ORGANISMO: ARTIFICIAL

35 ATGCCATTAGATCCTGTGATTCAATATGTGCTCGATCAACTGAACCGTATGCCAGCCCCTGACTACAAGCACTT  
AAGCGCTTGCCAGTTTCGTTCCACAGCAACACCTGTTTGAACCAGTTAAGAAAGAACCGGTCAAAGAAGTTGCG  
CGAATTCGACATGGACCTGCCGGGCGGTACCCTGAAAGTCCGTATGTACCGCCCGGAAGGTGTTGAACCACC  
ATATCCAGCACTGGTTTACTACCATGGTGGCGGTGGCGTTGCGGGTGACCTGGAAACGCATGATCCGGTGTG  
40 TCGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTT  
TCCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGAT  
CCAGCACGCATTGCAGTTGGCGGCGATTACAGCAGGCGGCAACCTCGCGGCCGTGACTAGTATTCTGGCGAA  
AGAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTTTACTGGTTATGATCCAGCACATCCAC  
CAGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAACCGGAGGCATGATGCTGTGGTTTCGTAATCAGTA  
45 CCTGAACAGCCTTGAGGAACTGACTCACCCTATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCA  
CCGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTGCGGCAAACCTTACGCTGAAGCCCTGAAC  
AAAGCGGGCGTTAAGGTGGAGATCGAGAACTTCGAGGATCTGATTCATGGCTTCGCGCAGTTCTATTCTCTGA  
GCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

50 >S0018766

SEQ ID NO:111

TIPO: ADN, ORGANISMO: ARTIFICIAL

ATGCCATTAGATCCTGTGATTCAATATGCTCTCGATCAACTGAACCGTATGCCAGCCCCTGACTACAAGCACTT  
AAGCGCTTGCCAGTTTCGTTCCACAGCAACACAGTTTTGGCCAGTTAAGAAAGAACCGGTGCGAGAAGTTGCG  
55 CGAATTCGACATGACGTGCCGGGCGGTACCCTGAAAGTCCGTATGTACACGCCGGAAGGTGTTGAACCACC  
ATATCCAGCACTGGTTTACTACCATGGTGGCGGTGGCGTTGCGGGTGACCTGGAAACGCATGATCCGGTGTG  
TCGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTT  
TCCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGAT  
60 CCAGCACGCATTGCAGTTGGCGGCGATTACAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAA  
AGAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTTTACTGGTTATGATCCAGCACATCCAC  
CAGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAACCGGAGGCATGATGCTGTGGTTTCGTAATCAGTA  
CCTGAACAGCCTTGAGGAACTGACTCACCCTATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCA  
CCGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTGCGGCAAACCTTACGCTGAAGCCCTGAAC

# ES 3 018 269 T3

AAAGCGGGCGTTAAGGTGGAGATCGAGAACTTCGAGGATCTGATTCATGGCTTCGCGCAGTTCTATTCTCTGA  
GCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

>S00187283

5 SEQ ID NO:113

TIPO: ADN, ORGANISMO: ARTIFICIAL

ATGCCATTAGATCCTGTGATTCAATATGTCCTCGATCAACTGAACCGTATGCCAGCCCCCTGACTACAAGCACTT  
AAGCGCTTGCCAGTTTCGTTACAGCAACACACGTTTTGGCCAGTTAAGAAAGAACCGGTGCGAGAAGTTGCG  
10 CGAATTCGACATGGATCTGCCGGGCGGTACCCTGAAAGTCCGTATGTACCGCCCCGGAAGGTGTTGAACCACC  
ATATCGCGCACTGGTTTACTACCATGGTGGCGGTGGCGTTGCGGGTGACCTGGAAACGCATGATCCGGTGTG  
TCGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACAACAAGTT  
TCCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGAT  
CCAGCACGCATTGCAGTTGGCGGGCGATTACAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAA  
15 AGAACGTGGTGGTCCAGCAATTGCGTTTCACTGCTGATCTATCCCTTTACTGGTTATGATCCAGCACATCCAC  
CAGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAACCGGAGGCATGATGCTGTGGTTTCGTAATCAGTA  
CCTGAACAGCCTTGAGGAAGTACTCACCCTATGGTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCA  
CCGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTACGCTGAAGCCCTGAAC  
AAAGCGGGCGTTAAGGTGGAGATCGAGAACTTCGAGGATCTGATTCATGGCTTCGCGCAGTTCTATTCTCTGA  
20 GCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

>S00187324

SEQ ID NO:115

TIPO: ADN, ORGANISMO: ARTIFICIAL

25 ATGCCATTAGATCCTGTGATTCAATATGTCCTCGATCAACTGAACCGTATGCCAGCCCCCTGACTGGAAGCACTT  
AAGCGCTTGCCAGTTTCGTTACAGCAACACATGTTTTGGCCAGTTAAGAAAGAACCGGTGCGCGAAGTTGCG  
CGAATTCGACATGGATCTGCCGGGCGGTACCCTGAAAGTCCGTATGTACCGCCCCGGAAGGTGTTGAACCACC  
ATATCGCGCACTGGTTTACTACCATGGTGGCGGTGGCGTTGCGGGTGACCTGGAACGCATGATCCGGTGTG  
30 TCGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACAACAAGTT  
TCCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGAT  
CCAGCACGCATTGCAGTTGGCGGGCGATTACAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAA  
AGAACGTGGTGGTCCAGCAATTGCGTTTCACTGCTGATCTATCCCTTTACTGGTTATGATCCAGCACATCCAC  
CAGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAACCGGAGGCATGATGCTGTGGTTTCGTAATCAGTA  
35 CCTGAACAGCCTTGAGGAAGTACTCACCCTATGGTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCA  
CCGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTACGCTGAAGCCCTGAAC  
AAAGCGGGCGTTAAGGTGGAGATCGAGAACTTCGAGGATCTGATTCATGGCTTCGCGCAGTTCTATTCTCTGA  
GCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

40 >S00187421

SEQ ID NO:117

TIPO: ADN, ORGANISMO: ARTIFICIAL

ATGCCATTAGATCCTGTGATTCAATATGTCCTCGATCAACTGAACCGTATGCCAGCCCCCTGACTACAAGCACTT  
45 AAGCGCTTGCCAGTTTCGTTACAGCAACACACGTTTTGGCCAGTTAAGAAAGAACCGGTGCGAGAAGTTGCG  
CGAATTCGACATGGATCTGCCGGGCGGTACCCTGAAAGTCCGTATGTACCGCCCCGGAAGGTGTTGAACCACC  
ATATATGGCACTGGTTTACTACCATGGTGGCGGTGGCGTTGCGGGTGACCTGGAACGCATGATCCGGTGTG  
TCGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACAACAAGTT  
TCCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGAT  
50 CCAGCACGCATTGCAGTTGGCGGGCGATTACAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAA  
AGAACGTGGTGGTCCAGCAATTGCGTTTCACTGCTGATCTATCCCTTTACTGGTTATGATCCAGCACATCCAC  
CAGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAACCGGAGGCATGATGCTGTGGTTTCGTAATCAGTA  
CCTGAACAGCCTTGAGGAAGTACTCACCCTATGGTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCA  
CCGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTACGCTGAAGCCCTGAAC  
55 AAAGCGGGCGTTAAGGTGGAGATCGAGAACTTCGAGGATCTGATTCATGGCTTCGCGCAGTTCTATTCTCTGA  
GCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

>S00188542

SEQ ID NO:119

60 TIPO: ADN, ORGANISMO: ARTIFICIAL

ATGCCATTAGATCCTGTGATTCAATATGTCCTCGATCAACTGAACCGTATGCCAGCCCCCTGACTACAAGCACTT  
AAGCGCTTGCCAGTTTCGTTACAGCAACACCTGTTGAACCAGTTAAGAAAGAACCGGTGCGAGAAGTTGCG  
CGAATTCGACATGGATCTGCCGGGCGGTACCCTGAAAGTCCGTATGTACCGCCCCGGAAGGTGTTGAACCACC

# ES 3 018 269 T3

ATATCCAGCACTGGTTTACTACCATGGTGGCGGTGGCGTTGCGGGTGACCTGGAAACGCATGATCCGGTGTG  
TCGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTT  
TCCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGAT  
CCAGCACGCATTGCAGTTGGCGGCGATTACAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAA  
5 AGAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTTTACTGGTTATGATCCAGCACATCCAC  
CAGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAACCGGAGGCATGATGCTGTGGTTTCGTAATCAGTA  
CCTGAACAGCCTTGAGGAACTGACTACCCATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCA  
CCGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTACGCTGAAGCCCTGAAC  
AAAGCGGGCGTTAAGGTGGAGATCGAGAACTTCGAGGATCTGATTATGGCTTCGCACAGTTCTATTCTCTGA  
10 GCCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

>S00207145  
SEQ ID NO:121  
TIPO: ADN, ORGANISMO: ARTIFICIAL

15 ATGCCATTAGATCCTGTGATTCAATATGTCCTCGATCAACTGAACCGTATGCCAGCCCCCTGACTACAAGCACTT  
AAGCGCTTGCCAGTTTCGTTACAGCAACACACGTTTTGGCCAGTTAAGAAAGAACCGGTGCGAGAAGTTGCG  
CGAATTCGACATGGATCTGCCGGGCGGTACCTGAAAGTCCGTATGTACCGCCCCGAAGGTGTTGAACCACC  
ATATCGCGCACTGGTTTACTACCATGGTGGCGGTGGCGTTGCGGGTGACCTGGAAACGCATGATCCGGTGTG  
20 TCGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTT  
TCCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGAT  
CCAGCACGCATTGCAGTTGGCGGCGATTACAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAA  
AGAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTTTACTGGTTATGATCCAGCACATCCAC  
CAGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAACCGGAGGCATGATGCTGTGGTTTCGTAATCAGTA  
25 CCTGAACAGCCTTGAGGAACTGACTACCCATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCA  
CCGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTACGCTGAAGCCCTGAAC  
AAAGCGGGCGTTAAGGTGGAGATCGAGAACTTCGAGGATCTGATTATGGCTTCGCGTGTCTATTCTCTGA  
GCCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

30 >S00208699  
SEQ ID NO:123  
TIPO: ADN, ORGANISMO: ARTIFICIAL

ATGTTGTTAGATCCTGTGATTCAATATGTCCTCGATCAACTGAACCGTATGCCAGCCCCCTGACTACAAGCACTT  
35 AAGCGCTTGCCAGTTTCGTTACAGCAACACACGTTTTGGCCAGTTAAGAAAGAACCGGTGCGAGAAGTTGCG  
CGAATTCGACATGGATCTGCCGGGCGGTACCTGAAAGTCCGTATGTACCGCCCCGAAGGTGTTGAACCACC  
ATATCGCGCACTGGTTTACTACCATGGTGGCGGTGGCGTTGCGGGTGACCTGGAAACGCATGATCCGGTGTG  
TCGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTT  
TCCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGAT  
40 CCAGCACGCATTGCAGTTGGCGGCGATTACAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAA  
AGAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCTGATCTATCCCTTTACTGGTTATGATCCAGCACATCCAC  
CAGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAACCGGAGGCATGATGCTGTGGTTTCGTAATCAGTA  
CCTGAACAGCCTTGAGGAACTGACTACCCATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCA  
CCGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTACGCTGAAGCCCTGAAC  
45 AAAGCGGGCGTTAAGGTGGAGATCGAGAACTTCGAGGATCTGATTATGGCTTCGCGCAGTTCTATTCTCTGA  
GCCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

>S00209368  
SEQ ID NO:125  
TIPO: ADN, ORGANISMO: ARTIFICIAL

50 ATGCCATTAGATCCTGTGATTCAATATGTCCTCGATCAACTGAACCGTATGCCAGCCCCCTGACTACAAGCACTT  
AAGCGCTTGCCAGTTTCGTTACAGCAACACACGTTTTGGCCAGTTAAGAAAGAACCGGTGCGAGAAGTTGCG  
CGAATTCGACATGGATCTGCCGGGCGGTACCTGAAAGTCCGTATGTACCGCCCCGAAGGTGTTGAACCACC  
55 ATATCGCGCACTGGTTTACTACCATGGTGGCGGTGGCGTTGCGGGTGACCTGGAAACGCATGATCCGGTGTG  
TCGTGTGTTGGCGAAAGATGGACGCGCAGTGGTGTGTTAGCGTTGACTACCGTCTGGCACCAGAACACAAGTT  
TCCAGCGGCAGTTGAAGACGCGTATGATGCACTGCAATGGATTGCAGAACGTGCAGCCGATTTCCATCTTGAT  
CCAGCACGCATTGCAGTTGGCGGCGATTACAGCAGGCGGCAACCTGGCGGCCGTGACTAGTATTCTGGCGAA  
AGAACGTGGTGGTCCAGCAATTGCGTTTCAACTGCAAACTATCCCTTTACTGGTTATGATCCAGCACATCCAC  
60 CAGCAAGTATCCTTGAGAATGCGGAGGGTTACCTGTTAACCGGAGGCATGATGCTGTGGTTTCGTAATCAGTA  
CCTGAACAGCCTTGAGGAACTGACTACCCATGGTTTAGTCCAGTGCTGTACCCGGATCTTAGCGGTTTACCA  
CCGGCGTACATTGCAACCGCACAGTACGATCCCCTGCGCGATGTCGGCAAACCTTACGCTGAAGCCCTGAAC  
AAAGCGGGCGTTAAGGTGGAGATCGAGAACTTCGAGGATCTGATTATGGCTTCGCGCAGTTCTATTCTCTGA  
GCCCAGGCGCAACCAAAGCACTGGTACGTATTGCCGAGAACTGCGCGATGCGTTGGCG

>S00032511  
SEQ ID NO:2  
TIPO: PROTEÍNA, ORGANISMO: ALICYCLOBACILLUS ACIDOCALDARIUS

5 MPLDPVIQQVLDQLNRMPAPDYKHLSAQQFRSQSLFPPVKKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPPY  
PALVYYHGGGWVVGDLETHDPVCRVLAKDGRAVVFSVDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPSTGYDPAHPPASIEENAEGYLLTGGMMLWFRDQYLSLEELT  
10 HPWFSPVLYPDLISGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVRIAEK  
LRDALA

>S00039799  
SEQ ID NO:4  
TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

15 MPLDPVIQQVLDQLNRMPAPDYKHLSAQQFRSQSLFPPVKKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPPY  
PALVYYHGGGWVVGDLETHDPVCRVLAKDGRAVVFSVDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPSTGYDPAHPPASILENAEGYLLTGGMMLWFRDQYLSLEELT  
20 HPWFSPVLYPDLISGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVRIAEK  
LRDALA

>S00039972  
SEQ ID NO:6  
TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

25 MPLDPVIQQVLDQLNRMPAPDYKHLSAQQFRSQSLFPPVKKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPPY  
PALVYYHGGGWVVGDLETHDPVCRVLAKDGRAVVFSVDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPSTGYDPAHPPASILENAEGYLLTGGMMLWFRDQYLSLEELT  
30 HPWFSPVLYPDLISGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVRIAEK  
LRDALA

>S00039989  
SEQ ID NO:8  
TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

35 MPLDPVIQQVLDQLNRMPAPDYKHLSAQQFRSQSLFPPVKKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPPY  
PALVYYHGGGWVVGDLETHDPVCRVLAKDGRAVVFSVDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPSTGYDPAHPPASILENAEGYLLTGGMMLWFRDQYLSLEELT  
40 HPWFSPVLYPDLISGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVRIAEK  
LRDALA

>S00040129  
SEQ ID NO:10  
TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

45 MPLDPVIQQVLDQLNRMPAPDYKHLSAQQFRSQSLFPPVKKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPPY  
PALVYYHGGGWVVGDLETHDPVCRVLAKDGRAVVFSVDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPSTGYDPAHPPASIFENAEGYLLTGGMMLWFRDQYLSLEELT  
50 HPWFSPVLYPDLISGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVRIAEK  
LRDALA

>S00040188  
SEQ ID NO:12  
TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

55 MPLDPVIQQVLDQLNRMPAPDYKHLSAQQFRSQSLFPPVKKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPPY  
PALVYYHGGGWVVGDLETHDPVCRVLAKDGRAVVFSVDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPSTGYDPAHPPASIMENAEGYLLTGGMMLWFRDQYLSLEELT  
60 HPWFSPVLYPDLISGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVRIAEK  
LRDALA

>S00044917  
SEQ ID NO:14  
TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

5 MPLDPVIQQVLDQLNRMPAPDYKHLAQQFRSQSLPPVKKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPPY  
PALVYYHGGGWVVGDLETHDPVCRVLAKDGRAVVFSDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYLLEHGMMLWFRDQYLSLEELT  
HPWFSPVLYPDLGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVRIAEK  
LRDALA

>S00044931  
SEQ ID NO:16

10 TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

15 MPLDPVIQQVLDQLNRMPAPDYKHLAQQFRSQSLFPPVKKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPPY  
PALVYYHGGGWVVGDLETHDPVCRVLAKDGRAVVFSDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYLLTGMMMLWFRDQYLSLEELT  
HPWFSPVLYPDLGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVRIAEK  
LRDALA

>S00044943  
SEQ ID NO:18

20 TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

25 MPLDPVIQQVLDQLNRMPAPDYKHLAQQFRSQSLFPPVKKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPPY  
RALVYYHGGGWVVGDLETHDPVCRVLAKDGRAVVFSDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYLLTGMMMLWFRDQYLSLEELT  
HPWFSPVLYPDLGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATLALVRIAEK  
LRDALA

>S00045046  
SEQ ID NO:20

30 TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

35 MPLDPVIQQVLDQLNRMPAPDYKHLAQQFRSQSLFPPVKKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPPY  
PALVYYHGGGWVVGDLETHDPVCRVLAKDGRAVVFSDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYLLEHGMMLWFRDQYLSLEELT  
HPWFSPVLYPDLGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVRIAEK  
LRDALA

>S00045575  
SEQ ID NO:22

40 TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

45 MPLDPVIQQVLDQLNRMPAPDYKHLAQQFRSQSLFPPVKKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPPY  
PALVYYHGGGWVVGDLETHDPVCRVLAKDGRAVVFSDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYLLTGMMMLWFRDQYLSLEELT  
HPWFSPVLYPDLGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVRIAEK  
LRDALA

>S00047040  
SEQ ID NO:24

50 TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

55 MPLDPVIQQVLDQLNRMPAPDYKHLAQQFRSQSLFPPVKKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPPY  
PALVYYHGGGWVVGDLETHDPVCRVLAKDGRAVVFSDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPMTGYDPAHPPASILENAEGYLLEGGMMLWFRDQYLSLEELT  
HPWFSPVLYPDLGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATVALVRIAEK  
LRDALA

>S00048311  
SEQ ID NO:26

60 TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

MPLDPVIQQVLDQLNRMPAPDYKHLSPQQLRSQSLFPPVKKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPPY  
PALVYYHGGGWVVGDLETHDPVCRVLAKWGRAVVFSDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPSTGYDPAHPPASIIENAEGYLLTGMMMLWFRDQYLSLEELT

HPWFSPVLYPDLSGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAFYSLSPGATKALVRIA  
EK  
LRDALA

>S00048430

5 SEQ ID NO:28

TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

MPLDPVIQQVLDQLNRMPAPDYKHLAQQFRSQSLFPPVKKEPVAEVREFDMDLPMRTLKVRMYRPEGVEPPY  
PALVYYHGGGWVVGDLETHDPVCRVLAKWGRAVVSVDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
10 AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPSTGYDPAHPPASILENAEGYLLTHGMMLWFRDQYLSLEELT  
HPWFSPVLYPDLSGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAFYSLSPGATRALVRIA  
EK  
LRDALA

>S00048441

15 SEQ ID NO:30

TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

MPLDPVIQQVLDQLNRMPAPDYKHLAQQFRSQSLFPPVKKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPPY  
PALVYYHGGGWVVGDLETHDPVCRVLAKDGRAVVSVDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
20 AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPSTGYDPAHPPASILENAEGYLLTGMMMLWFRDQYLSLEELT  
HPWFSPVLYPDLSGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAFYSLSPGATKALVRIA  
EK  
LRDALA

>S00048480

25 SEQ ID NO:32

TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

MPLDPVIQQVLDQLNRMPAPDYKHLAQQFRSQSLFPPVKKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPPY  
PALVYYHGGGWVVGDLETHDPVCRVLAKWGRAVVSVDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
30 AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPSTGYDPAHPPASILENAEGYLLTGMMMLWFRDQYLSLEELT  
HPWFSPVLYPDLSGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAFYSLSPGATKALVRIA  
EK  
LRDALA

>S00048593

35 SEQ ID NO:34

TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

MPLDPVIQQVLDQLNRMPAPDYKHLAQQFRSQSLFPPVKKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPPY  
PALVYYHGGGWVVGDLETHDPVCRVLAKWGRAVVSVDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
40 AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPSTGYDPAHPPASILENAEGYLLTGMMMLWFRDQYLSLEELT  
HPWFSPVLYPDLSGLPPAYIATAQYDPLRDVGKLYAEALNKAGVDVEIENFEDLIHGFAFYSLSPGATKALVRIA  
EK  
LRDALA

>S00052854

45 SEQ ID NO:36

TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

MPLDPVIQYVLDQLNRMPAPDYKHLAQQFRSQSLFPPVKKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPPY  
PALVYYHGGGWVVGDLETHDPVCRVLAKWGRAVVSVDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
50 AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPSTGYDPAHPPASILENAEGYLLTGMMMLWFRDQYLSLEELT  
HPWFSPVLYPDLSGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAFYSLSPGATKALVRIA  
EK  
LRDALA

>S00053124

55 SEQ ID NO:38

TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

MPLDPVIQQVLDQLNRMPAPDYKHLAQQFRSQSLFPPVKKEPVAELREFDMDLPGRTLKVRMYRPEGVEPPY  
PALVYYHGGGWVVGDLETHDPVCRVLAKDGRAVVSVDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
60 AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPSTGYDPAHPPASILENAEGYLLTGMMMLWFRNQYLSLEELT  
HPWFSPVLYPDLSGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAFYSLSPGATKALVRIA  
EK  
LRDALA

>S00053286

# ES 3 018 269 T3

SEQ ID NO:40

TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

5 MPLDPVIQYVLDQLNRMPRPDYKHLAQQFRSQNHLFPPVKKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPPY  
PALVYYHGGGWVAGDLETHDPVCRVLAKDGRAVVFSDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYLLTGGMMLWFRDQYLSLEELT  
HPWFSPVLYPDLGLPPAYIATAQYDPLRDVVGKLYAEALNKAGVDVEIENFEDLIHGFAQFYSLSPGATRALVRIAEL  
LRDALA

10 >S00053297

SEQ ID NO:42

TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

15 MPLDPVIQYVLDQLNRMPAPDYKHLAQQFRSQEHLFPPVKKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPPY  
PALVYYHGGGWVAGDLETHDPVCRVLAKDGRAVVFSDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYLLTGGMMLWFRDQYLSLEELT  
HPWFSPVLYPDLGLPPAYIATAQYDPLRDVVGKLYAEALNKAGVDVEIENFEDLIHGFAQFYSLSPGATKALVRIAEL  
LRDALA

20 >S00053319

SEQ ID NO:44

TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

25 MPLDPVIQYVLDQLNRMPRPDYKHLAQQFRSQEHLFPPVKKEPVVEVREFDMDLPGRTLKVRMYRPEGVEPPY  
PALVYYHGGGWVAGDLETHDPVCRVLAKDGRAVVFSDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYLLTGGMMLWFRDQYLSLEELT  
HPWFSPVLYPDLGLPPAYIATAQYDPLRDVVGKLYAEALNKAGVDVEIENFQDLIHGFAQFYSLSPGATRALVRIAEL  
LRDALA

30 >S00053506

SEQ ID NO:46

TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

35 MPLDPVIQYVLDQLNRMPAPDYKHLAQQFRSQNHLFPPVKKEPVVEVREFDMDLPGRTLKVRMYRPEGVEPPY  
PALVYYHGGGWVAGDLETHDPVCRVLAKDGRAVVFSDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYLLTGGMMLWFRDQYLSLEELT  
HPWFSPVLYPDLGLPPAYIATAQYDPLRDVVGKLYAEALNKAGVKVEIENFQDLIHGFAQFYSLSPGATRALVRIAEL  
LRDALA

40 >S00054201

SEQ ID NO:48

TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

45 MPLDPVIQQVLDQLNRMPAPDYKHLAQQFRSQHFLFPPVKKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPPY  
PALVYYHGGGWVAGDLETHDPVCRVLAKDGRAVVFSDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPFTGYDYAHPPASILENAEGYLLEGGMMLWFRDQYLSLEELT  
HPWFSPVLYPDLGLPPAYIATAQYDPLRDVVGKLYATALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVRIAEL  
LRDALA

50 >S00054685

SEQ ID NO:50

TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

55 MPLDPVIQYVLDQLNRMPAPDYKHLAQQFRSQEHLFPPVKKEPVVEVREFDMDLPGRTLKVRMYRPEGVEPPY  
PALVYYHGGGWVAGDLETHDPVCRVLAKDGRAVVFSDYRLAPEHKFPAAVEDAYDALQWIAERRADFHLDPARI  
AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYLLTGGMMLWFRDQYLSLEELT  
HPWFSPVLYPDLGLPPAYIATAQYDPLRDVVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVRIAEL  
LRDALA

60 >S00054833

SEQ ID NO:52

TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

MPLDPVIQYVLDQLNRMPAPDYKHLAQQFRSQNHLFPPVKKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPPY

PALVYYHGGGWVAGDLETHDPVCRVLAKDGRAVVSVDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYLLTGGMMLWFRDQYLNSEELT  
HPWFSPVLYPDLGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATRALVRIAEK  
LRDALA

5

>S00054853  
SEQ ID NO:54  
TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

10 MPLDPVIQYVLDQLNRMPAPDYKHLAQQFRSQHLPVPVKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPPY  
PALVYYHGGGWVAGDLETHDPVCRVLAKDGRAVVSVDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYLLTGGMMLWFRDQYLNSEELT  
HPWFSPVLYPDLGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVRIAEK  
LRDALA

15

>S00058614  
SEQ ID NO:56  
TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

20 MPLDPVIQYVLDQLNRMPAPDYKHLAQQFRSQHLPVPVKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPPY  
PALVYYHGGGWVAGDLETHDPVCRVLAKDGRAVVSVDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYLLTGGMMLWFRDQYLNSEELT  
HPWFSPVLYPDLGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVRIAEK  
LRDALA

25

>S00059197  
SEQ ID NO:58  
TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

30 MPLDPVIQYVLDQLNRMPAPDYKHLAQQFRSQHLPVPVKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPPY  
PALVYYHGGGWVAGDLETHDPVCRVLAKDGRAVVSVDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYLLTGGMMLWFRDQYLNSEELT  
HPWFSPVLYPDLGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVRIAEK  
LRDALA

35

>S00059305  
SEQ ID NO:60  
TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

40 MPLDPVIQYVLDQLNRMPAPDYKHLAQQFRSQHLPVPVKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPPY  
PALVYYHGGGWVAGDLETHDPVCRVLAKDGRAVVSVDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYLLTGGMMLWFRDQYLNSEELT  
HPWFSPVLYPDLGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVRIAEK  
LRDALA

45

>S00059320  
SEQ ID NO:62  
TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

50 MPLDPVIQYVLDQLNRMPAPDYKHLAQQFRSQHLPVPVKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPPY  
PALVYYHGGGWVAGDLETHDPVCRVLAKDGRAVVSVDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYLLTGGMMLWFRDQYLNSEELT  
HPWFSPVLYPDLGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVRIAEK  
LRDALA

55

>S00059754  
SEQ ID NO:64  
TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

60 MPLDPVIQYVLDQLNRMPAPDYKHLAQQFRSQHLPVPVKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPPY  
PALVYYHGGGWVAGDLETHDPVCRVLAKDGRAVVSVDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYLLTGGMMLWFRIQYLNSEELTH  
PWFSPLVLYPDLGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVRIAEK  
RDALA

>S00059759  
SEQ ID NO:66  
TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

5 MPLDPVIQYVLDQLNRMPAIDYKHLSAQQFRSQQHLFPPVKKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPPY  
PALVYYHGGGWVAGDLETHDPVCRVLAKDGRAVVFSDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYLLTGGMMLWFRDQYLNSEELT  
10 HPWFSPVLYPDLSGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVRIAEK  
LRDALA

>S00059902  
SEQ ID NO:68  
TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

15 MPLDPVIQYVLDQLNRMPAPDYKHLSAQQFRSQQHLFPPVKKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPPY  
PALVYYHGGGWVAGDLETHDPVCRVLAKDGRAVVFSDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYLLTGGMMLWFRNQYLNSEELT  
20 HPWFSPVLYPDLSGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVRIAEK  
LRDALA

>S00060141  
SEQ ID NO:70  
TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

25 MPLDPVIQYVLDQLNRMPAPDYKHLSAQQFRSQQHLFPPVKKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPPY  
PALVYYHGGGWVAGDLETHDPVCRVLAKDGRAVVFSDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYLLTGGMMLWFRDQYLNSEELT  
30 HPWFSPVLYPDLSGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVRIAEK  
LRDALA

>S00060676  
SEQ ID NO:72  
TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

35 MPLDPVIQYVLDQLNRMPAPDYKHLSAQQFRSQQHLFPPVKKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPPY  
PALVYYHGGGWVAGDLETHDPVCRVLAKDGRAVVFSDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYLLTGGMMLWFRDQYLNSEELT  
40 HPWFSPVLYPDLSGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVRIAEK  
LRDALA

>S00060772  
SEQ ID NO:74  
TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

45 MPLDPVIQYVLDQLNRMPAPDYKHLASQFRSQQHLFPPVKKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPPY  
PALVYYHGGGWVAGDLETHDPVCRVLAKDGRAVVFSDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYLLTGGMMLWFRDQYLNSEELT  
50 HPWFSPVLYPDLSGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVRIAEK  
LRDALA

>S00066575  
SEQ ID NO:76  
TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

55 MPLDPVIQYVLDQLNRMPAPDYKHLSAQQVRSQQHLFPPVKKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPPY  
PALVYYHGGGWVAGDLETHDPVCRVLAKDGRAVVFSDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYLLTGGMMLWFRNQYLNSEELT  
60 HPWFSPVLYPDLSGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVRIAEK  
LRDALA

>S00067402  
SEQ ID NO:78  
TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

5 MPLDPVIQYVLDQLNRMPAPDYKHLASQQFRSQQHIGPPVKKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPPY  
PALVYYHGGGWVAGDLETHDPVCRVLAKDGRAVVFSDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYLLTGGMMLWFRNQYLSLEELT  
HPWFSPVLYPDLSGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDTIHGFAQFYSSPGATKALVRIAEKL  
RDALA

>S00067403  
SEQ ID NO:80

10 TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

15 MPLDPVIQYVLDQLNRMPAPDYKHLASQQFRSQQHMGPPVKKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPP  
YPALVYYHGGGWVAGDLETHDPVCRVLAKDGRAVVFSDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPAR  
IAVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYLLTGGMMLWFRNQYLSLEELT  
HPWFSPVLYPDLSGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDMIHGFAQFYSSPGATKALVRIAEK  
LRDALA

>S00067700  
SEQ ID NO:82

20 TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

25 MPLDPVIQYMLDQLNRMPAPDYKHLASQQFRSQQHLGPPVKKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPP  
YPALVYYHGGGWVAGDLETHDPVCRVLAKDGRAVVFSDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPAR  
IAVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYLLTGGMMLWFRNQYLSLEELT  
HPWFSPVLYPDLSGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVRIAEK  
LRDALA

>S00068003  
SEQ ID NO:84

30 TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

35 MPLDPVIQYVLDQLNRMPAPDYKHLASQFRSQQHIFPPVKKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPPY  
PALVYYHGGGWVAGDLETHDPVCRVLAKDGRAVVFSDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYLLTGGMMLWFRNQYLSLEELT  
HPWFSPVLYPDLSGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVRIAEK  
LRDALA

>S00068079  
SEQ ID NO:86

40 TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

45 MPLDPVIQYVLDQLNRMPAPDYKHLASQFRSQQHIFPPVKKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPPY  
PALVYYHGGGWVAGDLETHDPVCRVLAKDGRAVVFSDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYLLTGGMMLWFRNQYLSLEELT  
HPWFSPVLYPDLSGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVRIAEKL  
RDALA

>S00068124  
SEQ ID NO:88

50 TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

55 MPLDPVIQYVLDQLNRMPAPDYKHLASQFRSQQHIFPPVKKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPPY  
PALVYYHGGGWVAGDLETHDPVCRVLAKDGRAVVFSDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYLLTGGMMLWFRNQYLSLEELT  
HPWFSPVLYPDLSGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVRIAEK  
LRDALA

>S00068512  
SEQ ID NO:90

60 TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

MPLDPVIQYVLDQLNRMPAPDYKHLASQQVRSQHLFPPVKKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPPY  
RALVYYHGGGWVAGDLETHDPVCRVLAKWGRAVVFSDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYLLTGGMMLWFRNQYLSLEELT

HPWFSPVLYPDLSGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVRIAEK  
LRDALA

5 >S00068639  
SEQ ID NO:92  
TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

10 MPLDPVIQYVLDQLNRMPAIDYKHLAQQQVRSWQHLPVPVKEPVEVREFDMDLPGRTLKVRMYTPEGVEPPY  
PALVYYHGGGWVAGDLETHDPVCRVLAKDGRAVVFSDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGDSAGGNLA AVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYLLTGGMMLWFRNQYLNLSLEELT  
HPWFSPVLYPDLSGLPPAYIATAQYDPLRDVGKLYARALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVRIAEK  
LRDALA

15 >S00068752  
SEQ ID NO:94  
TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

20 MPLDPVIQYVLDQLNRMPAPDWKHLAQQQFRSQHLPVPVKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPPY  
YRALVYYHGGGWVAGDLETHDPVCRVLAKDGRAVVFSDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
RIAVGGDSAGGNLA AVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYLLTGGMMLWFRNQYLNLSLEE  
LTHPWFSPLVLYPDLSGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVRIAEK  
KLRLDALA

25 >S00176502  
SEQ ID NO:96  
TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

30 MPLDPVIQYVLDQLNRMPAPDYKHLAQQQFRSQHLPVPVKEPVEVREFDMDLPGRTLKVRMYRPEGVEPPY  
PALVYYHGGGWVAGDLETHDPVCRVLAKDGRAVVFSDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGDSAGGNLA AVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYLLTGGMMLWFRNQYLNLSLEELT  
HPWFSPVLYPDLSGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVRIAEK  
LRDALA

35 >S00176595  
SEQ ID NO:98  
TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

40 MPLDPVIQYVLDQLNRMPAPDYKHLAQQQFRSQHLPVPVKEPVEVREFDMDLPGRTLKVRMYRPEGVEPPY  
PALVYYHGGGWVAGDLETHDPVCRVLAKDGRAVVFSDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGDSAGGNLA AVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYLLTGGMMLWFRNQYLNLSLEELT  
HPWFSPVLYPDLSGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVRIAEK  
LRDALA

45 >S00176980  
SEQ ID NO:100  
TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

50 MPLDPVIQYVLDQLNRMPAPDYKHLAQQQFRSQHLPVPVKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPPY  
PALVYYHGGGWVAGDLETHDPVCRVLAKDGRAVVFSDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGDSAGGNLA AVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYLLTGGMMLWFRNQYLNLSLEELT  
HPWFSPVLYPDLSGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVRIAEK  
LRDALA

55 >S00177019  
SEQ ID NO:102  
TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

60 MPLDPVIQYVLDQLNRMPAPDYKHLAQQQFRSQSLFPPVKEPVAEVREFDMDLPGRTLKVRMYLPPEGVEPPY  
PALVYYHGGGWVAGDLETHDPMCRVLAKDGRAVVFSDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGDSAGGNLA AVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYLLTGGMMDWFRNQYLNLSLEELT  
THPWFSPLVLYPDLSGLPPAYIATAQYDPLNDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVRIAEK  
KLRLDALA

>S00177053

# ES 3 018 269 T3

SEQ ID NO:104

TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

5 MPLDPVIQYVLDQLNRMPAPDYKHLSAQQFRSQSLFPPVKKEPVAEVREFDMDLPGRTLKVRMYLPEGVEPPY  
PALVYYHGGGWVAGDLETHDPMCRVLAKDGRAVVFSDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYFLTGGMMLWFRNQYLSLEELT  
HPWFSPVLYPDLGLPPAYIATWQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAFYSLSPGATKALVRIA  
EKLRDALA

10 >S00177175

SEQ ID NO:106

TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

15 MPLDPVIQYVLDQLNRMPAPDYKHLSAQCQFRSQHLLFPPVKKEPVAEVREFDMRLPGRTLKVRMYRPEGVEPPY  
RALVYYHGGGWVAGDLETHDPVCRVLAKDGRAVVFSDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGDSAGGNLAAVTSILAKERGGPAMAFQLLIYPFTGYDPAHPPASILENAEGYFLTGGMMLWFRNQYLSLEEL  
THPWFSPLVLYPDLGLPPAYIAVAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVRIA  
EKLRLDALA

20 >S00187023

SEQ ID NO:108

TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

25 MPLDPVLQYVLDQLNRMPAPDYKHLSAQCQFRSQHLLFEPVKKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPPY  
PALVYYHGGGGVAGDLETHDPVCRVLAKDGRAVVFSDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYLLTGGMMLWFRNQYLSLEELT  
HPWFSPVLYPDLGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVRIA  
EKLRLDALA

30 >S00187052

SEQ ID NO:110

TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

35 MPLDPVIQYVLDQLNRMPAPDYKHLSAQCQFRSQHLLFEPVKKEPVKEVREFDMDLPGRTLKVRMYRPEGVEPPY  
PALVYYHGGGGVAGDLETHDPVCRVLAKDGRAVVFSDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYLLTGGMMLWFRNQYLSLEELT  
HPWFSPVLYPDLGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVRIA  
EKLRLDALA

40 >S00187266

SEQ ID NO:112

TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

45 MPLDPVIQYVLDQLNRMPAPDYKHLSAQCQFRSQHTFWPVKKEPVAEVREFDMQLPGRTLKVRMYTPEGVEPPY  
PALVYYHGGGGVAGDLETHDPVCRVLAKDGRAVVFSDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYLLTGGMMLWFRNQYLSLEELT  
HPWFSPVLYPDLGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVRIA  
EKLRLDALA

50 >S00187283

SEQ ID NO:114

TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

55 MPLDPVIQYVLDQLNRMPAPDYKHLSAQCQFRSQHTFWPVKKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPP  
YRALVYYHGGGGVAGDLETHDPVCRVLAKDGRAVVFSDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPAR  
IAVGGDSAGGNLAAVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYLLTGGMMLWFRNQYLSLEELT  
HPWFSPVLYPDLGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVRIA  
EKLRLDALA

60 >S00187324

SEQ ID NO:116

TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

MPLDPVIQYVLDQLNRMPAPDWKHLSAQCQFRSQHMFVWPVKKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPP

YRALVYYHGGGGVAGDLETHDPVCRVLAKDGRAVVSVDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPAR  
IAVGGSAGGNLAAVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYLLTGGMMLWFRNQYLNLSLEELT  
HPWFSPVLYPDLSGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVRIAEK  
LRDALA

5

>S00187421  
SEQ ID NO:118  
TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

10 MPLDPVIQYVLDQLNRMPAPDYKHLASACQFRSQHFTWPVKKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPP  
YMALVYYHGGGGVAGDLETHDPVCRVLAKDGRAVVSVDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPAR  
RIAVGGSAGGNLAAVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYLLTGGMMLWFRNQYLNLSLEE  
LTHPWFSPVLYPDLSGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVRIA  
EKLRDALA

15

>S00188542  
SEQ ID NO:120  
TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

20 MPLDPVIQYVLDQLNRMPAPDYKHLASACQFRSQHLEFVPVKKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPPY  
PALVYYHGGGGVAGDLETHDPVCRVLAKDGRAVVSVDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGSAGGNLAAVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYLLTGGMMLWFRNQYLNLSLEELT  
HPWFSPVLYPDLSGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVRIAEK  
LRDALA

25

>S00207145  
SEQ ID NO:122  
TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

30 MPLDPVIQYVLDQLNRMPAPDYKHLASACQFRSQHFTWPVKKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPPY  
YRALVYYHGGGGVAGDLETHDPVCRVLAKDGRAVVSVDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPAR  
IAVGGSAGGNLAAVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYLLTGGMMLWFRNQYLNLSLEELT  
HPWFSPVLYPDLSGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVRIAEK  
LRDALA

35

>S00208699  
SEQ ID NO:124  
TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

40 MLLDPVIQYVLDQLNRMPAPDYKHLASACQFRSQHFTWPVKKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPPY  
RALVYYHGGGGVAGDLETHDPVCRVLAKDGRAVVSVDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPARI  
AVGGSAGGNLAAVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYLLTGGMMLWFRNQYLNLSLEELT  
HPWFSPVLYPDLSGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVRIAEK  
LRDALA

45

>S00209368  
SEQ ID NO:126  
TIPO: PRT, ORGANISMO: SECUENCIA ARTIFICIAL

50 MPLDPVIQYVLDQLNRMPAPDYKHLASACQFRSQHFTWPVKKEPVAEVREFDMDLPGRTLKVRMYRPEGVEPPY  
YRALVYYHGGGGVAGDLETHDPVCRVLAKDGRAVVSVDYRLAPEHKFPAAVEDAYDALQWIAERAADFHLDPAR  
IAVGGSAGGNLAAVTSILAKERGGPAIAFQLLIYPFTGYDPAHPPASILENAEGYLLTGGMMLWFRNQYLNLSLEEL  
THPWFSPVLYPDLSGLPPAYIATAQYDPLRDVGKLYAEALNKAGVKVEIENFEDLIHGFAQFYSLSPGATKALVRIAE  
KLRLDALA

55

# REIVINDICACIONES

1. Un polipéptido de carboxiesterasa que comprende una secuencia de aminoácidos que es al menos 93% o más idéntica a la secuencia de aminoácidos expuesta en SEQ ID NO: 4, o un fragmento funcional del mismo, en donde el residuo correspondiente a X198 en SEQ ID NO: 4 es L, y el polipéptido de carboxiesterasa tiene actividad enzimática mejorada con respecto al polipéptido de carboxiesterasa de SEQ ID NO: 2.
2. El polipéptido como se reivindica en la reivindicación 1, donde la actividad enzimática mejorada es al menos 5 veces en relación con el polipéptido de SEQ ID NO: 2.
3. El polipéptido de carboxiesterasa como se reivindica en las reivindicaciones 1 o 2, donde la secuencia de aminoácidos del polipéptido de carboxiesterasa comprende una diferencia de residuos en comparación con la secuencia de aminoácidos expuesta en SEQ ID NO: 4 en al menos una posición de residuo seleccionada de: X27, X30, X35, X37, X57, X75, X103, X185, X207, X208, X271, X286, y X296.
4. El polipéptido de carboxiesterasa como se reivindica en la reivindicación 1, 2 o 3, donde la secuencia de aminoácidos del polipéptido de carboxiesterasa incluye la característica: el residuo correspondiente a X103 se elige de un residuo alifático y un residuo aromático.
5. El polipéptido de carboxiesterasa como se reivindica en la reivindicación 4, donde X103 se elige de un residuo aromático.
6. El polipéptido de carboxiesterasa como se reivindica en la reivindicación 5, donde X103 es W.
7. El polipéptido de carboxiesterasa como se reivindica en cualquiera de las reivindicaciones 1-7, donde la secuencia de aminoácidos del polipéptido de carboxiesterasa incluye la característica de que el residuo correspondiente a X185 se elige de un residuo no polar, un residuo alifático y un residuo aromático.
8. El polipéptido de carboxiesterasa como se reivindica en la reivindicación 7, donde X185 es un residuo aromático.
9. El polipéptido de carboxiesterasa como se reivindica en la reivindicación 8, donde X185 es F.
10. El polipéptido de carboxiesterasa como se reivindica en las reivindicaciones 1, 2 o 3, donde la secuencia de aminoácidos del polipéptido de carboxiesterasa incluye al menos una característica elegida de: X27 es P; X30 se elige de I, L y V; X35 es H; X37 se elige de I, L, T y V; X57 es M; X75 es R; X103 se elige de F, M y W; X185 se elige de F, I y M; X207 es E; X208 se elige de R, L y H; X271 es D; X286 se elige de M, V y G; y X296 se elige de V, L y R.
11. El polipéptido de carboxiesterasa como se reivindica en la reivindicación 1, donde la secuencia de aminoácidos corresponde a la secuencia de aminoácidos expuesta en SEQ ID NO: 4.
12. Un polinucleótido que codifica para un polipéptido como se reivindica en la reivindicación 11.
13. Un polinucleótido como se reivindica en la reivindicación 12, que tiene la secuencia de SEQ ID NO: 3.
14. Un vector que comprende un polinucleótido como se reivindica en la reivindicación 12 o 13.
15. Una célula hospedadora que comprende un vector como se reivindica en la reivindicación 14.
16. Una composición que comprende al menos una carboxiesterasa modificada como se reivindica en cualquiera de las reivindicaciones 1-11.

## Figura 1

Alineación por pares de todas las secuencias de ADN contra la secuencia de polinucleótidos que codifica el codón de *E. coli* optimizado para la enzima carboxiesterasa de tipo silvestre, Esterasa 2 de *A. acidocaldarius* (SEQ ID NO: 1)

|        |    | 80         | 90         | 100        | 110          | 120        | 130        | 140        |     |
|--------|----|------------|------------|------------|--------------|------------|------------|------------|-----|
| SEQ1   | 71 | ACTTAAGCGC | TCAACAGTTT | CGTTCACAGC | ATCACTCTTT   | TCCACCAAGT | AAGAAAGAAC | CGGTGCGAGA | 140 |
| SEQ3   | 71 | .....      | .....      | .....      | .....        | .....      | .....      | .....      | 140 |
| SEQ5   | 71 | .....      | .....      | .....      | .....        | .....      | .....      | .....      | 140 |
| SEQ7   | 71 | .....      | .....      | .....      | .....        | .....      | .....      | .....      | 140 |
| SEQ9   | 71 | .....      | .....      | .....      | .....        | .....      | .....      | .....      | 140 |
| SEQ11  | 71 | .....      | .....      | .....      | .....        | .....      | .....      | .....      | 140 |
| SEQ13  | 71 | .....      | .....      | .....      | .....        | .....      | .....      | .....      | 140 |
| SEQ15  | 71 | .....      | ..G.....   | .....      | .....        | G.....     | .....      | .....      | 140 |
| SEQ17  | 71 | .....      | ..G.....   | .....      | .....        | G.....     | .....      | .....      | 140 |
| SEQ19  | 71 | .....      | .....      | .....      | .....        | G.....     | .....      | .....      | 140 |
| SEQ21  | 71 | .....      | .....      | .....      | .....        | G.....     | .....      | .....      | 140 |
| SEQ23  | 71 | .....      | .....      | .....      | ..CAC.....   | .....      | .....      | .....      | 140 |
| SEQ25  | 71 | .....      | C..G.....G | .....      | .....        | .....      | .....      | .....      | 140 |
| SEQ27  | 71 | .....      | .....      | .....      | ..G.....     | .....      | .....      | .....      | 140 |
| SEQ29  | 71 | .....      | .....      | .....      | .....        | ..G.....   | .....      | .....      | 140 |
| SEQ31  | 71 | .....      | .....      | .....      | .....        | ..G.....   | .....      | .....      | 140 |
| SEQ33  | 71 | .....      | .....      | .....      | .....        | .....      | .....      | .....      | 140 |
| SEQ35  | 71 | .....      | .....      | .....      | ..CAC.....AC | C.....     | .....      | .....      | 140 |
| SEQ37  | 71 | .....      | .....      | .....      | ..TAC.....   | .....      | .....      | .....      | 140 |
| SEQ39  | 71 | .....      | .....      | .....      | A..CCAC..... | .....      | .....      | .....      | 140 |
| SEQ41  | 71 | .....      | .....      | .....      | G..CAC.....  | .....      | .....      | .....      | 140 |
| SEQ43  | 71 | .....      | .....      | .....      | G..CAC.....  | .....      | .....      | ..TG..     | 140 |
| SEQ45  | 71 | .....      | .....      | .....      | A..CCAC..... | .....      | .....      | ..TG..     | 140 |
| SEQ47  | 71 | .....      | .....      | .....      | ..CAC.....   | .....      | .....      | .....      | 140 |
| SEQ49  | 71 | .....      | .....      | .....      | G..CAC.....  | .....      | .....      | ..TG..     | 140 |
| SEQ51  | 71 | .....      | .....      | .....      | A..CCAC..... | .....      | .....      | .....      | 140 |
| SEQ53  | 71 | .....      | .....      | .....      | ..CAC.....   | .....      | .....      | .....      | 140 |
| SEQ55  | 71 | .....      | .....      | .....      | ..CAC.....   | .....      | .....      | .....      | 140 |
| SEQ57  | 71 | .....      | .....      | G.....     | ..CAC.....   | .....      | .....      | .....      | 140 |
| SEQ59  | 71 | .....      | .....      | .....      | G..CAC.....  | .....      | .....      | .....      | 140 |
| SEQ61  | 71 | .....      | .....      | .....      | ..CAC.....   | .....      | .....      | .....      | 140 |
| SEQ63  | 71 | .....      | .....      | .....      | ..CAC.....   | .....      | .....      | .....      | 140 |
| SEQ65  | 71 | .....      | .....      | .....      | ..CAC.....   | .....      | .....      | .....      | 140 |
| SEQ67  | 71 | .....      | .....      | .....      | ..CAC.....   | .....      | .....      | .....      | 140 |
| SEQ69  | 71 | .....      | .....      | G..T.....  | ..CAC.....   | .....      | .....      | .....      | 140 |
| SEQ71  | 71 | .....      | .....      | .....      | GGT.....     | ..CAC..... | .....      | .....      | 140 |
| SEQ73  | 71 | .....      | AGT.....   | .....      | ..CAC.....   | .....      | .....      | .....      | 140 |
| SEQ75  | 71 | .....      | .....      | G.....     | ..A.....     | ..CAC..... | .....      | .....      | 140 |
| SEQ77  | 71 | .....      | .....      | .....      | ..CACA..TGG  | .....      | .....      | .....      | 140 |
| SEQ79  | 71 | .....      | .....      | .....      | ..CACT..GG   | .....      | .....      | .....      | 140 |
| SEQ81  | 71 | .....      | AGT.....   | .....      | ..CAC.....   | .....      | .....      | .....      | 140 |
| SEQ83  | 71 | .....      | AGT.....   | .....      | ..CAT.....   | .....      | .....      | .....      | 140 |
| SEQ85  | 71 | .....      | AGT.....   | .....      | ..CAC.....   | .....      | .....      | .....      | 140 |
| SEQ87  | 71 | .....      | .....      | .....      | ..CACA..GG   | .....      | .....      | .....      | 140 |
| SEQ89  | 71 | .....      | AG.....    | .....      | ..TG.....    | ..CAC..... | .....      | .....      | 140 |
| SEQ91  | 71 | .....      | AG.....    | .....      | ..TG.....    | ..CAC..... | .....      | CGC.....   | 140 |
| SEQ93  | 71 | .....      | .....      | .....      | ..CAC.....   | .....      | .....      | .....      | 140 |
| SEQ95  | 71 | .....      | .....      | .....      | ..CAC.....   | .....      | .....      | AA.....    | 140 |
| SEQ97  | 71 | .....      | .....      | .....      | ..CAC.....   | .....      | .....      | AA.....    | 140 |
| SEQ99  | 71 | .....      | TGC.....   | .....      | ..CAC.....   | GA.....    | .....      | .....      | 140 |
| SEQ101 | 71 | .....      | .....      | .....      | ..AGC.....   | .....      | .....      | .....      | 140 |
| SEQ103 | 71 | .....      | .....      | .....      | ..AGC.....   | .....      | .....      | .....      | 140 |
| SEQ105 | 71 | .....      | TGC.....   | .....      | ..CAC.....   | .....      | .....      | E.....     | 140 |
| SEQ107 | 71 | .....      | TGC.....   | .....      | ..CAC.....   | GA.....    | .....      | E.....     | 140 |
| SEQ109 | 71 | .....      | TGC.....   | .....      | ..CAC.....   | GA.....    | AA.....    | .....      | 140 |
| SEQ111 | 71 | .....      | TGC.....   | .....      | ..CACAC..... | TGG.....   | .....      | .....      | 140 |
| SEQ113 | 71 | .....      | TGC.....   | .....      | ..CACAC..... | TGG.....   | .....      | .....      | 140 |
| SEQ115 | 71 | .....      | TGC.....   | .....      | ..CACA.....  | TGG.....   | .....      | C.....     | 140 |
| SEQ117 | 71 | .....      | TGC.....   | .....      | ..CACAC..... | TGG.....   | .....      | .....      | 140 |
| SEQ119 | 71 | .....      | TGC.....   | .....      | ..CAC.....   | GA.....    | .....      | .....      | 140 |
| SEQ121 | 71 | .....      | TGC.....   | .....      | ..CACAC..... | TGG.....   | .....      | .....      | 140 |
| SEQ123 | 71 | .....      | TGC.....   | .....      | ..CACAC..... | TGG.....   | .....      | .....      | 140 |
| SEQ125 | 71 | .....      | TGC.....   | .....      | ..CACAC..... | TGG.....   | .....      | .....      | 140 |

**Figura 1 (Continuación)**

|        |     | 150        | 160        | 170        | 180         | 190        | 200        | 210        |     |
|--------|-----|------------|------------|------------|-------------|------------|------------|------------|-----|
| SEQ1   | 141 | AGTTCGCGAA | TTCGACATGG | ATCTCCCGGG | CGGTACCCCTG | AAAGTCCGTA | TGTACCGTCC | GGAAGGTGTT | 210 |
| SEQ3   | 141 |            |            |            |             |            |            |            | 210 |
| SEQ5   | 141 |            |            |            |             |            |            |            | 210 |
| SEQ7   | 141 |            |            |            |             |            |            |            | 210 |
| SEQ9   | 141 |            |            |            |             |            |            |            | 210 |
| SEQ11  | 141 |            |            |            |             |            |            |            | 210 |
| SEQ13  | 141 |            |            |            |             |            |            |            | 210 |
| SEQ15  | 141 |            |            |            |             |            |            |            | 210 |
| SEQ17  | 141 |            |            |            |             |            |            |            | 210 |
| SEQ19  | 141 |            |            |            |             |            |            |            | 210 |
| SEQ21  | 141 |            |            |            |             |            | A.         |            | 210 |
| SEQ23  | 141 |            |            |            |             |            | C.         |            | 210 |
| SEQ25  | 141 |            |            |            | T.          |            |            |            | 210 |
| SEQ27  | 141 |            |            | AT         | G.          |            |            |            | 210 |
| SEQ29  | 141 |            |            |            |             |            |            |            | 210 |
| SEQ31  | 141 |            |            |            |             |            |            |            | 210 |
| SEQ33  | 141 |            |            |            |             |            |            |            | 210 |
| SEQ35  | 141 |            |            |            |             |            | C.         |            | 210 |
| SEQ37  | 141 | C.G.       |            |            |             |            | C.         |            | 210 |
| SEQ39  | 141 |            |            |            |             |            | C.         |            | 210 |
| SEQ41  | 141 |            |            |            |             |            | C.         |            | 210 |
| SEQ43  | 141 |            |            |            |             |            | C.         |            | 210 |
| SEQ45  | 141 |            |            |            |             |            | C.         |            | 210 |
| SEQ47  | 141 |            |            |            |             |            | GTG.       |            | 210 |
| SEQ49  | 141 |            |            |            |             |            | C.         |            | 210 |
| SEQ51  | 141 |            |            |            |             |            | C.         |            | 210 |
| SEQ53  | 141 |            |            |            |             |            | C.         |            | 210 |
| SEQ55  | 141 |            |            |            |             |            | C.         |            | 210 |
| SEQ57  | 141 |            |            |            |             |            | C.         |            | 210 |
| SEQ59  | 141 |            |            |            |             |            | C.         |            | 210 |
| SEQ61  | 141 |            |            |            |             |            | C.         |            | 210 |
| SEQ63  | 141 |            |            |            |             |            | C.         |            | 210 |
| SEQ65  | 141 |            |            |            |             |            | C.         |            | 210 |
| SEQ67  | 141 |            |            |            |             |            | C.         |            | 210 |
| SEQ69  | 141 |            |            |            |             |            | C.         |            | 210 |
| SEQ71  | 141 |            |            |            |             |            | C.         |            | 210 |
| SEQ73  | 141 |            |            |            |             |            | C.         |            | 210 |
| SEQ75  | 141 |            |            |            |             |            | C.         |            | 210 |
| SEQ77  | 141 |            |            |            |             |            | C.         |            | 210 |
| SEQ79  | 141 |            |            |            |             |            | C.         |            | 210 |
| SEQ81  | 141 |            |            |            |             |            | C.         |            | 210 |
| SEQ83  | 141 |            |            |            |             |            | C.         |            | 210 |
| SEQ85  | 141 |            |            |            |             |            | C.         |            | 210 |
| SEQ87  | 141 |            |            |            |             |            | C.         |            | 210 |
| SEQ89  | 141 |            |            |            |             |            | C.         |            | 210 |
| SEQ91  | 141 |            |            |            |             |            | ACC.       |            | 210 |
| SEQ93  | 141 |            |            |            |             |            | C.         |            | 210 |
| SEQ95  | 141 |            |            |            |             |            | C.         |            | 210 |
| SEQ97  | 141 |            |            |            |             |            | C.         |            | 210 |
| SEQ99  | 141 |            |            |            |             |            | C.         |            | 210 |
| SEQ101 | 141 |            |            |            |             |            | TG.        |            | 210 |
| SEQ103 | 141 |            |            |            |             |            | TG.        |            | 210 |
| SEQ105 | 141 | C.         | G.         |            |             |            | C.         |            | 210 |
| SEQ107 | 141 |            | C.         |            |             |            | C.         |            | 210 |
| SEQ109 | 141 |            | C.         |            |             |            | C.         |            | 210 |
| SEQ111 | 141 |            | C.         | G.         |             |            | ACS.       |            | 210 |
| SEQ113 | 141 |            |            |            |             |            | C.         |            | 210 |
| SEQ115 | 141 |            |            |            |             |            | C.         |            | 210 |
| SEQ117 | 141 |            |            |            |             |            | C.         |            | 210 |
| SEQ119 | 141 |            |            |            |             |            | C.         |            | 210 |
| SEQ121 | 141 |            |            |            |             |            | C.         |            | 210 |
| SEQ123 | 141 |            |            |            |             |            | C.         |            | 210 |
| SEQ125 | 141 |            |            |            |             |            | C.         |            | 210 |

Figura 1 (Continuación)

|        |     | 220         | 230        | 240        | 250        | 260       | 270        | 280         |
|--------|-----|-------------|------------|------------|------------|-----------|------------|-------------|
| SEQ1   | 211 | GAACCAACCAT | ATCCAGCACT | GGTTTACTAC | CATGGTGGCG | GTTGGGTGT | TGGTGACCTG | GAAACGCAATG |
| SEQ3   | 211 | .....       | .....      | .....      | .....      | .....     | .....      | .....       |
| SEQ5   | 211 | .....       | .....      | .....      | .....      | .....     | .....      | .....       |
| SEQ7   | 211 | .....       | .....      | .....      | .....      | .....     | .....      | .....       |
| SEQ9   | 211 | .....       | .....      | .....      | .....      | .....     | .....      | .....       |
| SEQ11  | 211 | .....       | .....      | .....      | .....      | .....     | .....      | .....       |
| SEQ13  | 211 | .....       | .....      | .....      | .....      | .....     | .....      | .....       |
| SEQ15  | 211 | .....       | .....      | .....      | .....      | .....     | .....      | .....       |
| SEQ17  | 211 | .....       | .....      | .....      | .....      | .....     | .....      | .....       |
| SEQ19  | 211 | .....       | ..GG.      | .....      | .....      | .....     | .....      | .....       |
| SEQ21  | 211 | .....       | ..G.       | .....      | .....      | .....     | .....      | .....       |
| SEQ23  | 211 | .....       | .....      | .....      | .....      | .....     | .....      | .....       |
| SEQ25  | 211 | .....       | .....      | .....      | .....      | .....     | .....      | .....       |
| SEQ27  | 211 | .....       | .....      | .....      | .....      | .....     | .....      | .....       |
| SEQ29  | 211 | .....       | .....      | .....      | .....      | .....     | .....      | .....       |
| SEQ31  | 211 | .....       | .....      | .....      | .....      | .....     | .....      | .....       |
| SEQ33  | 211 | .....       | .....      | .....      | .....      | .....     | .....      | .....       |
| SEQ35  | 211 | .....       | .....      | .....      | .....      | .....     | .....      | .....       |
| SEQ37  | 211 | .....       | .....      | .....      | .....      | .....     | .....      | .....       |
| SEQ39  | 211 | .....       | .....      | .....      | .....      | ..C G.    | .....      | .....       |
| SEQ41  | 211 | .....       | .....      | .....      | .....      | ..C G.    | .....      | .....       |
| SEQ43  | 211 | .....       | .....      | .....      | .....      | ..C G.    | .....      | .....       |
| SEQ45  | 211 | .....       | .....      | .....      | .....      | ..C G.    | .....      | .....       |
| SEQ47  | 211 | .....       | .....      | .....      | .....      | ..G. C.   | .....      | .....       |
| SEQ49  | 211 | .....       | .....      | .....      | .....      | ..C G.    | .....      | .....       |
| SEQ51  | 211 | .....       | .....      | .....      | .....      | ..C G.    | .....      | .....       |
| SEQ53  | 211 | .....       | .....      | .....      | .....      | ..C G.    | .....      | .....       |
| SEQ55  | 211 | .....       | .....      | .....      | .....      | ..C G.    | .....      | .....       |
| SEQ57  | 211 | .....       | .....      | .....      | .....      | ..C G.    | .....      | .....       |
| SEQ59  | 211 | .....       | .....      | .....      | .....      | ..C G.    | .....      | .....       |
| SEQ61  | 211 | .....       | .....      | .....      | .....      | ..C G.    | .....      | .....       |
| SEQ63  | 211 | .....       | .....      | .....      | .....      | ..C G.    | .....      | .....       |
| SEQ65  | 211 | .....       | .....      | .....      | .....      | ..C G.    | .....      | .....       |
| SEQ67  | 211 | .....       | .....      | .....      | .....      | ..C G.    | .....      | .....       |
| SEQ69  | 211 | .....       | .....      | .....      | .....      | ..C G.    | .....      | .....       |
| SEQ71  | 211 | .....       | ..GG.      | .....      | .....      | ..G. C.   | .....      | .....       |
| SEQ73  | 211 | .....       | .....      | .....      | .....      | ..C G.    | .....      | .....       |
| SEQ75  | 211 | .....       | .....      | .....      | .....      | ..C G.    | .....      | .....       |
| SEQ77  | 211 | .....       | .....      | .....      | .....      | ..C G.    | .....      | .....       |
| SEQ79  | 211 | .....       | .....      | .....      | .....      | ..C G.    | .....      | .....       |
| SEQ81  | 211 | .....       | .....      | .....      | .....      | ..C G.    | .....      | .....       |
| SEQ83  | 211 | .....       | .....      | .....      | .....      | ..C G.    | .....      | .....       |
| SEQ85  | 211 | .....       | .....      | .....      | .....      | ..C G.    | .....      | .....       |
| SEQ87  | 211 | .....       | .....      | .....      | .....      | ..C G.    | .....      | .....       |
| SEQ89  | 211 | .....       | ..GG.      | .....      | .....      | ..C G.    | .....      | .....       |
| SEQ91  | 211 | .....       | .....      | .....      | .....      | ..C G.    | .....      | .....       |
| SEQ93  | 211 | .....       | ..GG.      | .....      | .....      | ..C G.    | .....      | .....       |
| SEQ95  | 211 | .....       | .....      | .....      | .....      | ..C G.    | .....      | .....       |
| SEQ97  | 211 | .....       | .....      | .....      | .....      | ..C G.    | .....      | .....       |
| SEQ99  | 211 | .....       | .....      | .....      | ..G. C.    | ..C G.    | .....      | .....       |
| SEQ101 | 211 | .....       | .....      | .....      | ..ACC.     | ..C G.    | .....      | .....       |
| SEQ103 | 211 | .....       | .....      | .....      | ..ACC.     | ..C G.    | .....      | .....       |
| SEQ105 | 211 | .....       | ..GT.      | .....      | .....      | ..C G.    | .....      | .....       |
| SEQ107 | 211 | .....       | .....      | .....      | ..G. C.    | ..C G.    | .....      | .....       |
| SEQ109 | 211 | .....       | .....      | .....      | ..G. C.    | ..C G.    | .....      | .....       |
| SEQ111 | 211 | .....       | .....      | .....      | ..G. C.    | ..C G.    | .....      | .....       |
| SEQ113 | 211 | .....       | ..GC.      | .....      | ..G. C.    | ..C G.    | .....      | .....       |
| SEQ115 | 211 | .....       | ..GC.      | .....      | ..G. C.    | ..C G.    | .....      | .....       |
| SEQ117 | 211 | .....       | ..ATG.     | .....      | ..G. C.    | ..C G.    | .....      | .....       |
| SEQ119 | 211 | .....       | .....      | .....      | ..G. C.    | ..C G.    | .....      | .....       |
| SEQ121 | 211 | .....       | ..GC.      | .....      | ..G. C.    | ..C G.    | .....      | .....       |
| SEQ123 | 211 | .....       | ..GC.      | .....      | ..G. C.    | ..C G.    | .....      | .....       |
| SEQ125 | 211 | .....       | ..GC.      | .....      | ..G. C.    | ..C G.    | .....      | .....       |

Figura 1 (Continuación)

|        |   | 10         | 20          | 30            | 40         | 50            | 60           | 70         |    |
|--------|---|------------|-------------|---------------|------------|---------------|--------------|------------|----|
| SEQ1   | 1 | ATGCCATTAG | ATCCTGTGAT  | TCACAAAGTC    | CTCGATCAAC | TGAACCGTAT    | GCCAGCCCCCT  | GACTACAAGC | 70 |
| SEQ3   | 1 | .....      | .....       | .....         | .....      | .....         | .....        | .....      | 70 |
| SEQ5   | 1 | .....      | .....       | .....         | .....      | .....         | .....        | .....      | 70 |
| SEQ7   | 1 | .....      | .....       | .....         | .....      | .....         | .....        | .....      | 70 |
| SEQ9   | 1 | .....      | .....       | .....         | .....      | .....         | .....        | .....      | 70 |
| SEQ11  | 1 | .....      | .....       | .....         | .....      | .....         | .....        | .....      | 70 |
| SEQ13  | 1 | .....      | .....       | .....         | .....      | .....         | .....A.....  | .....      | 70 |
| SEQ15  | 1 | .....      | .....       | .....         | .....      | .....         | .....        | .....      | 70 |
| SEQ17  | 1 | .....      | .....       | .....         | .....      | .....         | .....        | .....      | 70 |
| SEQ19  | 1 | .....      | .....       | .....         | .....      | .....         | .....        | .....      | 70 |
| SEQ21  | 1 | .....      | .....       | .....         | .....      | .....         | .....        | .....      | 70 |
| SEQ23  | 1 | .....      | .....       | .....         | .....      | .....         | .....        | .....      | 70 |
| SEQ25  | 1 | .....      | .....       | .....         | .....      | .....         | .....        | .....      | 70 |
| SEQ27  | 1 | .....      | .....       | .....         | .....      | .....         | .....        | .....      | 70 |
| SEQ29  | 1 | .....      | .....       | .....         | .....      | .....         | .....        | .....      | 70 |
| SEQ31  | 1 | .....      | .....       | .....         | .....      | .....         | .....        | .....      | 70 |
| SEQ33  | 1 | .....      | .....       | .....         | .....      | .....         | .....        | .....      | 70 |
| SEQ35  | 1 | .....      | .....       | .....T.F..... | .....      | .....         | .....        | .....      | 70 |
| SEQ37  | 1 | .....      | .....       | .....         | .....      | .....         | .....        | .....      | 70 |
| SEQ39  | 1 | .....      | .....       | .....T.F..... | .....      | .....CGT..... | .....        | .....      | 70 |
| SEQ41  | 1 | .....      | .....       | .....T.F..... | .....      | .....         | .....        | .....      | 70 |
| SEQ43  | 1 | .....      | .....       | .....T.F..... | .....      | .....CGT..... | .....        | .....      | 70 |
| SEQ45  | 1 | .....      | .....       | .....T.F..... | .....      | .....         | .....        | .....      | 70 |
| SEQ47  | 1 | .....      | .....       | .....         | .....      | .....         | .....        | .....      | 70 |
| SEQ49  | 1 | .....      | .....       | .....T.F..... | .....      | .....         | .....        | .....      | 70 |
| SEQ51  | 1 | .....      | .....       | .....T.F..... | .....      | .....         | .....        | .....      | 70 |
| SEQ53  | 1 | .....      | .....       | .....T.F..... | .....      | .....         | .....        | .....      | 70 |
| SEQ55  | 1 | .....      | .....       | .....T.F..... | .....      | .....         | .....        | .....      | 70 |
| SEQ57  | 1 | .....      | .....       | .....T.F..... | .....      | .....         | .....        | .....      | 70 |
| SEQ59  | 1 | .....      | .....       | .....T.F..... | .....      | .....         | .....        | .....      | 70 |
| SEQ61  | 1 | .....      | .....       | .....T.F..... | .....      | .....         | .....        | .....      | 70 |
| SEQ63  | 1 | .....      | .....       | .....T.F..... | .....      | .....         | .....        | .....      | 70 |
| SEQ65  | 1 | .....      | .....       | .....T.F..... | .....      | .....AT.....  | .....        | .....      | 70 |
| SEQ67  | 1 | .....      | .....       | .....T.F..... | .....      | .....         | .....        | .....      | 70 |
| SEQ69  | 1 | .....      | .....       | .....T.F..... | .....      | .....         | .....        | .....      | 70 |
| SEQ71  | 1 | .....      | .....       | .....T.F..... | .....      | .....         | .....        | .....      | 70 |
| SEQ73  | 1 | .....      | .....       | .....T.F..... | .....      | .....         | .....        | .....      | 70 |
| SEQ75  | 1 | .....      | .....       | .....T.F..G   | .....      | .....         | .....        | .....      | 70 |
| SEQ77  | 1 | .....      | .....       | .....T.F..G   | .....      | .....         | .....        | .....      | 70 |
| SEQ79  | 1 | .....      | .....       | .....T.FA..G  | .....      | .....         | .....        | .....      | 70 |
| SEQ81  | 1 | .....      | .....       | .....T.F..... | .....      | .....         | .....        | .....      | 70 |
| SEQ83  | 1 | .....      | .....       | .....T.F..... | .....      | .....         | .....        | .....      | 70 |
| SEQ85  | 1 | .....      | .....       | .....T.F..... | .....      | .....         | .....        | .....      | 70 |
| SEQ87  | 1 | .....      | .....       | .....T.F..... | .....      | .....         | .....        | .....      | 70 |
| SEQ89  | 1 | .....      | .....       | .....T.F..... | .....      | .....C.....   | .....        | .....      | 70 |
| SEQ91  | 1 | .....      | .....       | .....T.F..... | .....      | .....AT.....  | .....        | .....      | 70 |
| SEQ93  | 1 | .....      | .....       | .....T.F..... | .....      | .....         | .....CG..... | .....      | 70 |
| SEQ95  | 1 | .....      | .....       | .....T.F..... | .....      | .....         | .....        | .....      | 70 |
| SEQ97  | 1 | .....      | .....       | .....T.F..... | .....      | .....         | .....        | .....      | 70 |
| SEQ99  | 1 | .....      | .....       | .....T.F..... | .....      | .....         | .....        | .....      | 70 |
| SEQ101 | 1 | .....      | .....       | .....T.F..... | .....      | .....         | .....        | .....      | 70 |
| SEQ103 | 1 | .....      | .....       | .....T.F..... | .....      | .....         | .....        | .....      | 70 |
| SEQ105 | 1 | .....      | .....       | .....T.F..... | .....      | .....         | .....        | .....      | 70 |
| SEQ107 | 1 | .....      | .....C..... | .....T.F..G   | .....      | .....         | .....        | .....      | 70 |
| SEQ109 | 1 | .....      | .....       | .....T.F..G   | .....      | .....         | .....        | .....      | 70 |
| SEQ111 | 1 | .....      | .....       | .....T.F..... | .....      | .....         | .....        | .....      | 70 |
| SEQ113 | 1 | .....      | .....       | .....T.F..... | .....      | .....         | .....        | .....      | 70 |
| SEQ115 | 1 | .....      | .....       | .....T.F..... | .....      | .....         | .....GG..... | .....      | 70 |
| SEQ117 | 1 | .....      | .....       | .....T.F..... | .....      | .....         | .....        | .....      | 70 |
| SEQ119 | 1 | .....      | .....       | .....T.F..... | .....      | .....         | .....        | .....      | 70 |
| SEQ121 | 1 | .....      | .....       | .....T.F..... | .....      | .....         | .....        | .....      | 70 |
| SEQ123 | 1 | ...TTG...  | .....       | .....T.F..... | .....      | .....         | .....        | .....      | 70 |
| SEQ125 | 1 | .....      | .....       | .....T.F..... | .....      | .....         | .....        | .....      | 70 |

Figura 1 (Continuación)

|        |     | 290        | 300        | 310        | 320        | 330        | 340        | 350        |
|--------|-----|------------|------------|------------|------------|------------|------------|------------|
| SEQ1   | 281 | ATCCGGTGTG | TCGTGTGTGT | GCSAAAGATG | GACGCGCAGT | GGTGTATTAC | GTTGACTACC | GTCTGGCACC |
| SEQ3   | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ5   | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ7   | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ9   | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ11  | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ13  | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ15  | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ17  | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ19  | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ21  | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ23  | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ25  | 281 | .....      | .....      | .....TGG   | .....      | .....      | .....      | .....      |
| SEQ27  | 281 | .....      | .....      | .....TGG   | .....      | .....      | .....      | .....      |
| SEQ29  | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ31  | 281 | .....      | .....      | .....TGG   | .....      | .....      | .....      | .....      |
| SEQ33  | 281 | .....      | .....      | .....TGG   | .....      | .....      | .....      | .....      |
| SEQ35  | 281 | .....      | .....      | .....TGG   | .....      | .....      | .....      | .....      |
| SEQ37  | 281 | .....      | .....      | .....TGG   | .....      | .....      | .....      | .....      |
| SEQ39  | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ41  | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ43  | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ45  | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ47  | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ49  | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ51  | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ53  | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ55  | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ57  | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ59  | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ61  | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ63  | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ65  | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ67  | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ69  | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ71  | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ73  | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ75  | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ77  | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ79  | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ81  | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ83  | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ85  | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ87  | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ89  | 281 | .....      | .....      | .....TGG   | .....      | .....      | .....      | .....      |
| SEQ91  | 281 | .....      | .....      | .....TGG   | .....      | .....      | .....      | .....      |
| SEQ93  | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ95  | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ97  | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ99  | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ101 | 281 | .....A     | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ103 | 281 | .....A     | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ105 | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ107 | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ109 | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ111 | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ113 | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ115 | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ117 | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ119 | 281 | .....      | .....      | .....TGG   | .....      | .....      | .....      | .....      |
| SEQ121 | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ123 | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |
| SEQ125 | 281 | .....      | .....      | .....      | .....      | .....      | .....      | .....      |

**Figura 1 (Continuación)**[illegible]

Figura 1 (Continuación)

|        |     | 430        | 440        | 450        | 460         | 475           | 480       | 490        |
|--------|-----|------------|------------|------------|-------------|---------------|-----------|------------|
| SEQ1   | 421 | GATTCACATC | TTGATCCAGC | ACGCATTGCA | GTGGGCGGCG  | ATTCAGCAGG    | CGGCACCTG | GCGGCGGTGA |
| SEQ3   | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ5   | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ7   | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ9   | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ11  | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ13  | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ15  | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ17  | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ19  | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ21  | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ23  | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ25  | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ27  | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ29  | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ31  | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ33  | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ35  | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ37  | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ39  | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ41  | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ43  | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ45  | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ47  | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ49  | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ51  | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ53  | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ55  | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ57  | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ59  | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ61  | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ63  | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ65  | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ67  | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ69  | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ71  | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ73  | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ75  | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ77  | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ79  | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ81  | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ83  | 421 | .....      | .....      | .....      | .....G..... | .....         | .....     | 490        |
| SEQ85  | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ87  | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ89  | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ91  | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ93  | 421 | .....      | .....      | .....      | .....G..... | .....         | .....     | 490        |
| SEQ95  | 421 | .....      | .....      | .....      | .....       | .....ACC..... | .....     | 490        |
| SEQ97  | 421 | .....      | .....      | .....      | .....       | .....ACC..... | .....     | 490        |
| SEQ99  | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ101 | 421 | .....      | .....      | .....      | .....       | .....C.....   | .....     | 490        |
| SEQ103 | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ105 | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ107 | 421 | .....      | .....      | .....      | .....       | .....C.....   | .....     | 490        |
| SEQ109 | 421 | .....      | .....      | .....      | .....       | .....C.....   | .....     | 490        |
| SEQ111 | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ113 | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ115 | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ117 | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ119 | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ121 | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ123 | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |
| SEQ125 | 421 | .....      | .....      | .....      | .....       | .....         | .....     | 490        |

**Figura 1 (Continuación)**

|        |     | 520         | 520        | 520        | 530        | 540        | 550        | 560        |     |
|--------|-----|-------------|------------|------------|------------|------------|------------|------------|-----|
| SEQ1   | 491 | CTAGTATCTCT | GCCGAAAGAA | CGTGTGTGTC | CAGCAATATC | GTTTCAACTG | CTGATCTATC | CCTCACATGG | 560 |
| SEQ3   | 491 |             |            |            |            |            |            |            | 560 |
| SEQ5   | 491 |             |            |            |            |            |            |            | 560 |
| SEQ7   | 491 |             |            |            |            |            |            |            | 560 |
| SEQ9   | 491 |             |            |            |            |            |            |            | 560 |
| SEQ11  | 491 |             |            |            |            |            |            |            | 560 |
| SEQ13  | 491 |             |            |            |            |            |            | ..ATG..    | 560 |
| SEQ15  | 491 |             |            |            |            |            |            |            | 560 |
| SEQ17  | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ19  | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ21  | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ23  | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ25  | 491 |             |            |            |            |            |            |            | 560 |
| SEQ27  | 491 |             |            |            |            |            |            |            | 560 |
| SEQ29  | 491 |             |            |            |            |            |            |            | 560 |
| SEQ31  | 491 |             |            |            |            |            |            |            | 560 |
| SEQ33  | 491 |             |            |            |            |            |            |            | 560 |
| SEQ35  | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ37  | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ39  | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ41  | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ43  | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ45  | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ47  | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ49  | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ51  | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ53  | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ55  | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ57  | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ59  | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ61  | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ63  | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ65  | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ67  | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ69  | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ71  | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ73  | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ75  | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ77  | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ79  | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ81  | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ83  | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ85  | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ87  | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ89  | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ91  | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ93  | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ95  | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ97  | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ99  | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ101 | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ103 | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ105 | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ107 | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ109 | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ111 | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ113 | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ115 | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ117 | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ119 | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ121 | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ123 | 491 |             |            |            |            |            |            | ..TT..     | 560 |
| SEQ125 | 491 |             |            |            |            |            | ..AA..     | ..TT..     | 560 |

Figura 1 (Continuación)

|        |     | 570        | 580        | 590        | 600        | 610        | 620        | 630        |
|--------|-----|------------|------------|------------|------------|------------|------------|------------|
| SEQ1   | 561 | TTATGATCCA | GCACATCCAC | CAGCAGATAT | CGAGAGAGAT | GCGGAGGGGT | ACCTGTTRAC | CGGAGGCATG |
| SEQ3   | 561 |            |            |            | .CTT       |            |            |            |
| SEQ5   | 561 |            |            |            | .ATT       |            |            |            |
| SEQ7   | 561 |            |            |            | .T.T       |            |            |            |
| SEQ9   | 561 |            |            |            | .TTT       |            |            |            |
| SEQ11  | 561 |            |            |            | .ATC       |            |            |            |
| SEQ13  | 561 |            |            |            | .CTT       |            | .GA        | G          |
| SEQ15  | 561 |            |            |            | .CTT       |            | .GA        | GCAT       |
| SEQ17  | 561 |            |            |            | .CTT       |            |            |            |
| SEQ19  | 561 |            |            |            | .CTT       |            | .GCAT      |            |
| SEQ21  | 561 |            |            |            | .CTT       |            | .GA        | GCAT       |
| SEQ23  | 561 |            |            |            | .CTT       |            |            |            |
| SEQ25  | 561 |            |            |            | .ATT       |            |            |            |
| SEQ27  | 561 |            |            |            | .CTT       |            | .CAT       |            |
| SEQ29  | 561 |            |            |            | .CTT       |            |            |            |
| SEQ31  | 561 |            |            |            | .ATT       |            | .T         |            |
| SEQ33  | 561 |            |            |            | .CTT       |            | .T         |            |
| SEQ35  | 561 |            |            |            | .CTT       |            |            |            |
| SEQ37  | 561 |            |            |            | .CTT       |            |            |            |
| SEQ39  | 561 |            |            |            | .CTT       |            |            |            |
| SEQ41  | 561 |            |            |            | .CTT       |            |            |            |
| SEQ43  | 561 |            |            |            | .CTT       |            |            |            |
| SEQ45  | 561 |            |            |            | .CTT       |            |            |            |
| SEQ47  | 561 | .TAT       |            |            | .CTT       |            | .GA        | A          |
| SEQ49  | 561 |            |            |            | .CTT       |            |            |            |
| SEQ51  | 561 |            |            |            | .CTT       |            |            |            |
| SEQ53  | 561 |            |            |            | .CTT       |            |            |            |
| SEQ55  | 561 |            |            |            | .CTT       |            |            |            |
| SEQ57  | 561 |            |            |            | .CTT       |            |            |            |
| SEQ59  | 561 |            |            |            | .CTT       |            |            |            |
| SEQ61  | 561 | .T         |            |            | .CTT       |            |            |            |
| SEQ63  | 561 |            |            |            | .CTT       |            |            |            |
| SEQ65  | 561 |            |            |            | .CTT       |            |            |            |
| SEQ67  | 561 |            |            |            | .CTT       |            |            |            |
| SEQ69  | 561 |            |            |            | .CTT       |            |            |            |
| SEQ71  | 561 |            |            |            | .CTT       |            |            |            |
| SEQ73  | 561 |            |            |            | .CTT       |            |            |            |
| SEQ75  | 561 |            |            | .C         | .TCTT      |            |            |            |
| SEQ77  | 561 |            |            |            | .CTT       |            |            |            |
| SEQ79  | 561 |            |            |            | .CTT       |            |            |            |
| SEQ81  | 561 | .C         |            |            | .TCTT      |            |            |            |
| SEQ83  | 561 | .C         |            | .C         | .TCTT      |            |            |            |
| SEQ85  | 561 |            |            | .C         | .TCTT      |            |            |            |
| SEQ87  | 561 |            |            |            | .CTT       |            |            |            |
| SEQ89  | 561 | A          |            |            | .CTT       |            |            |            |
| SEQ91  | 561 | A          |            |            | .CTT       |            |            |            |
| SEQ93  | 561 |            |            |            | .CTT       |            |            |            |
| SEQ95  | 561 |            |            |            | .CTT       |            |            |            |
| SEQ97  | 561 |            |            |            | .CTT       |            |            |            |
| SEQ99  | 561 |            |            |            | .CTT       |            |            |            |
| SEQ101 | 561 | .T         |            |            | .CTT       |            | .T.T       |            |
| SEQ103 | 561 |            |            |            | .CTT       |            | .T.T       |            |
| SEQ105 | 561 |            |            |            | .CTT       |            | .T.T       |            |
| SEQ107 | 561 |            |            |            | .CTT       |            |            |            |
| SEQ109 | 561 |            |            |            | .CTT       |            |            |            |
| SEQ111 | 561 |            |            |            | .CTT       |            |            |            |
| SEQ113 | 561 |            |            |            | .CTT       |            |            |            |
| SEQ115 | 561 |            |            |            | .CTT       |            |            |            |
| SEQ117 | 561 |            |            |            | .CTT       |            |            |            |
| SEQ119 | 561 |            |            |            | .CTT       |            |            |            |
| SEQ121 | 561 |            |            |            | .CTT       |            |            |            |
| SEQ123 | 561 |            |            |            | .CTT       |            |            |            |
| SEQ125 | 561 |            |            |            | .CTT       |            |            |            |

Figura 1 (Continuación)

|        |     | 640        | 650        | 660        | 670        | 680        | 690        | 700        |     |
|--------|-----|------------|------------|------------|------------|------------|------------|------------|-----|
| SEQ1   | 631 | ATGCTGTGGT | TTCTGTACCA | GTACCTGAAC | AGCCTTGAGG | AACTGACTCA | CCCATGGTTF | AGTCCAGTGC | 700 |
| SEQ3   | 631 | .....      | .....      | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ5   | 631 | .....      | .....      | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ7   | 631 | .....      | .....      | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ9   | 631 | .....      | .....      | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ11  | 631 | .....      | .....      | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ13  | 631 | .....      | .....      | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ15  | 631 | .....      | .....      | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ17  | 631 | .....      | .....      | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ19  | 631 | .....      | .....      | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ21  | 631 | .....      | .....      | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ23  | 631 | .....      | .....      | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ25  | 631 | .....      | .....      | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ27  | 631 | .....      | .....      | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ29  | 631 | .....      | .....      | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ31  | 631 | .....      | .....      | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ33  | 631 | .....      | .....      | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ35  | 631 | .....      | .....      | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ37  | 631 | .....      | A..        | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ39  | 631 | .....      | .....      | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ41  | 631 | .....      | .....      | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ43  | 631 | .....      | .....      | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ45  | 631 | .....      | .....      | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ47  | 631 | .....      | .....      | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ49  | 631 | .....      | .....      | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ51  | 631 | .....      | .....      | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ53  | 631 | .....      | .....      | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ55  | 631 | .....      | .....      | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ57  | 631 | .....      | .....      | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ59  | 631 | .....      | .....      | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ61  | 631 | .....      | .....      | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ63  | 631 | .....      | ATT..      | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ65  | 631 | .....      | .....      | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ67  | 631 | .....      | A..T..     | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ69  | 631 | .....      | .....      | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ71  | 631 | .....      | .....      | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ73  | 631 | .....      | .....      | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ75  | 631 | .....      | A..T..     | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ77  | 631 | .....      | A..T..     | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ79  | 631 | .....      | A..T..     | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ81  | 631 | .....      | A..T..     | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ83  | 631 | .....      | A..T..     | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ85  | 631 | .....      | A..T..     | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ87  | 631 | .....      | A..T..     | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ89  | 631 | .....      | A..T..     | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ91  | 631 | .....      | A..T..     | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ93  | 631 | .....      | A..T..     | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ95  | 631 | .....      | A..T..     | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ97  | 631 | .....      | A..T..     | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ99  | 631 | .....      | A..T..     | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ101 | 631 | .....      | GAT..      | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ103 | 631 | .....      | A..T..     | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ105 | 631 | .....      | A..T..     | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ107 | 631 | .....      | A..T..     | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ109 | 631 | .....      | A..T..     | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ111 | 631 | .....      | A..T..     | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ113 | 631 | .....      | A..T..     | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ115 | 631 | .....      | A..T..     | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ117 | 631 | .....      | A..T..     | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ119 | 631 | .....      | A..T..     | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ121 | 631 | .....      | A..T..     | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ123 | 631 | .....      | A..T..     | .....      | .....      | .....      | .....      | .....      | 700 |
| SEQ125 | 631 | .....      | A..T..     | .....      | .....      | .....      | .....      | .....      | 700 |

Figura 1 (Continuación)

|        |     | 710         | 720          | 730          | 740           | 750          | 760          | 770        |
|--------|-----|-------------|--------------|--------------|---------------|--------------|--------------|------------|
| SEQ1   | 701 | ....{.....} | .....{.....} | .....{.....} | .....{.....}  | .....{.....} | .....{.....} | 770        |
| SEQ3   | 701 | TGTACCCGGA  | TCTTAGCGGT   | TTACCACCGG   | GCTACCTTGC    | ATCCGCACAG   | TACGATCCCC   | TGCGCGATGT |
| SEQ5   | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ7   | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ9   | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ11  | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ13  | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ15  | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ17  | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ19  | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ21  | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ23  | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ25  | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ27  | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ29  | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ31  | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ33  | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ35  | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ37  | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ39  | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ41  | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ43  | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ45  | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ47  | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ49  | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ51  | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ53  | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ55  | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ57  | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ59  | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ61  | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ63  | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ65  | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ67  | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ69  | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ71  | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ73  | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ75  | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ77  | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ79  | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ81  | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ83  | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ85  | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ87  | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ89  | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ91  | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ93  | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ95  | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ97  | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ99  | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ101 | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ103 | 701 | .....       | .....        | .....        | .....TGG..... | .....        | .....AA..... | 770        |
| SEQ105 | 701 | .....       | .....        | .....        | .....GTG..... | .....        | .....        | 770        |
| SEQ107 | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ109 | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ111 | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ113 | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ115 | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ117 | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ119 | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ121 | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ123 | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |
| SEQ125 | 701 | .....       | .....        | .....        | .....         | .....        | .....        | 770        |

Figura 1 (Continuación)

|        |     | 780        | 790        | 800        | 810         | 820        | 830        | 840        |
|--------|-----|------------|------------|------------|-------------|------------|------------|------------|
| SEQ1   | 771 | CGGCAAACTT | TACGCTGAAG | CCCTGAACAA | AGCGGGCGSTT | AGGCTGGAGA | TCGAGTACTT | CGAGGATCTG |
| SEQ3   | 771 | .....      | .....      | .....      | .....       | .....      | .....      | .....      |
| SEQ5   | 771 | .....      | .....      | .....      | .....       | .....      | .....      | .....      |
| SEQ7   | 771 | .....      | .....      | .....      | .....       | .....      | .....      | .....      |
| SEQ9   | 771 | .....      | .....      | .....      | .....       | .....      | .....      | .....      |
| SEQ11  | 771 | .....      | .....      | .....      | .....       | .....      | .....      | .....      |
| SEQ13  | 771 | .....      | .....      | .....      | .....       | .....      | .....      | .....      |
| SEQ15  | 771 | .....      | .....      | .....      | .....       | .....      | .....      | .....      |
| SEQ17  | 771 | .....      | .....      | .....      | .....       | .....      | .....      | .....      |
| SEQ19  | 771 | .....      | .....      | .....      | .....       | .....      | .....      | .....      |
| SEQ21  | 771 | .....      | .....      | .....      | .....       | .....      | .....      | .....      |
| SEQ23  | 771 | .....      | .....      | .....      | .....       | .....      | .....      | .....      |
| SEQ25  | 771 | .....      | .....      | .....      | .....       | .....      | .....      | .....      |
| SEQ27  | 771 | .....      | .....      | .....      | ..A..       | .....      | .....      | .....      |
| SEQ29  | 771 | .....      | .....      | .....      | ..A..       | .....      | .....      | .....      |
| SEQ31  | 771 | .....      | .....      | .....      | ..A..       | .....      | .....      | .....      |
| SEQ33  | 771 | .....      | .....      | .....      | G.T..       | .....      | .....      | .....      |
| SEQ35  | 771 | .....      | .....      | .....      | .....       | .....      | .....      | .....      |
| SEQ37  | 771 | .....      | .....      | .....      | .....       | .....      | .....      | .....      |
| SEQ39  | 771 | .....      | .....      | .....      | G.T..       | .....      | .....      | .....      |
| SEQ41  | 771 | .....      | .....      | .....      | G.T..       | .....      | .....      | .....      |
| SEQ43  | 771 | .....      | .....      | .....      | G.T..       | .....      | ..C..      | .....      |
| SEQ45  | 771 | .....      | .....      | .....      | .....       | .....      | ..C..      | .....      |
| SEQ47  | 771 | .....      | ..ACC..    | .....      | .....       | .....      | .....      | .....      |
| SEQ49  | 771 | .....      | .....      | .....      | .....       | .....      | .....      | .....      |
| SEQ51  | 771 | .....      | .....      | .....      | .....       | .....      | .....      | .....      |
| SEQ53  | 771 | .....      | .....      | .....      | .....       | .....      | .....      | .....      |
| SEQ55  | 771 | .....      | .....      | .....      | .....       | .....      | .....      | .....      |
| SEQ57  | 771 | .....      | .....      | .....      | .....       | .....      | .....      | .....      |
| SEQ59  | 771 | .....      | .....      | .....      | .....       | .....      | .....      | .....      |
| SEQ61  | 771 | .....      | .....      | .....      | .....       | .....      | .....      | .....      |
| SEQ63  | 771 | .....      | .....      | .....      | .....       | .....      | .....      | .....      |
| SEQ65  | 771 | .....      | .....      | .....      | .....       | .....      | .....      | .....      |
| SEQ67  | 771 | .....      | .....      | .....      | .....       | .....      | .....      | .....      |
| SEQ69  | 771 | .....      | .....      | .....      | .....       | .....      | .....      | .....      |
| SEQ71  | 771 | .....      | .....      | .....      | .....       | .....      | .....      | .....      |
| SEQ73  | 771 | .....      | .....      | .....      | .....       | .....      | .....      | .....      |
| SEQ75  | 771 | .....      | .....      | .....      | .....       | .....      | .....      | .....      |
| SEQ77  | 771 | .....      | .....      | .....      | .....       | .....      | ..AC..     | .....      |
| SEQ79  | 771 | .....      | .....      | .....      | .....       | .....      | .....      | ..A..      |
| SEQ81  | 771 | ..T..      | .....      | .....      | .....       | .....      | .....      | .....      |
| SEQ83  | 771 | ..T..      | .....      | ..C..      | .....       | .....      | .....      | .....      |
| SEQ85  | 771 | ..T..      | .....      | .....      | .....       | .....      | .....      | .....      |
| SEQ87  | 771 | .....      | .....      | .....      | .....       | .....      | .....      | ..A..      |
| SEQ89  | 771 | .....      | .....      | .....      | ..A..       | .....      | .....      | .....      |
| SEQ91  | 771 | .....      | ..AG..     | .....      | ..A..       | .....      | .....      | .....      |
| SEQ93  | 771 | .....      | .....      | .....      | .....       | .....      | .....      | .....      |
| SEQ95  | 771 | .....      | .....      | .....      | .....       | .....      | ..C..      | .....      |
| SEQ97  | 771 | .....      | .....      | .....      | .....       | .....      | .....      | .....      |
| SEQ99  | 771 | .....      | .....      | .....      | .....       | .....      | .....      | .....      |
| SEQ101 | 771 | .....      | .....      | .....      | ..C.G..     | .....      | .....      | .....      |
| SEQ103 | 771 | .....      | .....      | .....      | .....       | .....      | .....      | .....      |
| SEQ105 | 771 | .....      | .....      | .....      | .....       | .....      | .....      | .....      |
| SEQ107 | 771 | .....      | .....      | .....      | .....       | .....      | .....      | .....      |
| SEQ109 | 771 | .....      | .....      | .....      | .....       | .....      | .....      | .....      |
| SEQ111 | 771 | .....      | .....      | .....      | .....       | .....      | .....      | .....      |
| SEQ113 | 771 | .....      | .....      | .....      | .....       | .....      | .....      | .....      |
| SEQ115 | 771 | .....      | .....      | .....      | .....       | .....      | .....      | .....      |
| SEQ117 | 771 | .....      | .....      | .....      | .....       | .....      | .....      | .....      |
| SEQ119 | 771 | .....      | .....      | .....      | .....       | .....      | .....      | .....      |
| SEQ121 | 771 | .....      | .....      | .....      | .....       | .....      | .....      | .....      |
| SEQ123 | 771 | .....      | .....      | .....      | .....       | .....      | .....      | .....      |
| SEQ125 | 771 | .....      | .....      | .....      | .....       | .....      | .....      | .....      |

**Figura 1 (Continuación)**

[illegible]

Figura 1 (Continuación)

|        | 920                       | 930 |
|--------|---------------------------|-----|
| SEQ1   | 911 AACCTGGCGCA TGGCTGGCG | 930 |
| SEQ3   | 911 .....                 | 930 |
| SEQ5   | 911 .....                 | 930 |
| SEQ7   | 911 .....                 | 930 |
| SEQ9   | 911 .....                 | 930 |
| SEQ11  | 911 .....                 | 930 |
| SEQ13  | 911 .....                 | 930 |
| SEQ15  | 911 .....                 | 930 |
| SEQ17  | 911 .....                 | 930 |
| SEQ19  | 911 .....                 | 930 |
| SEQ21  | 911 .....                 | 930 |
| SEQ23  | 911 .....                 | 930 |
| SEQ25  | 911 .....                 | 930 |
| SEQ27  | 911 .....                 | 930 |
| SEQ29  | 911 .....                 | 930 |
| SEQ31  | 911 .....                 | 930 |
| SEQ33  | 911 .....                 | 930 |
| SEQ35  | 911 .....                 | 930 |
| SEQ37  | 911 .....                 | 930 |
| SEQ39  | 911 .....                 | 930 |
| SEQ41  | 911 .....                 | 930 |
| SEQ43  | 911 .....                 | 930 |
| SEQ45  | 911 .....                 | 930 |
| SEQ47  | 911 .....                 | 930 |
| SEQ49  | 911 .....                 | 930 |
| SEQ51  | 911 .....                 | 930 |
| SEQ53  | 911 .....                 | 930 |
| SEQ55  | 911 .....                 | 930 |
| SEQ57  | 911 .....                 | 930 |
| SEQ59  | 911 .....                 | 930 |
| SEQ61  | 911 .....                 | 930 |
| SEQ63  | 911 .....                 | 930 |
| SEQ65  | 911 .....                 | 930 |
| SEQ67  | 911 .....                 | 930 |
| SEQ69  | 911 .....                 | 930 |
| SEQ71  | 911 .....                 | 930 |
| SEQ73  | 911 .....                 | 930 |
| SEQ75  | 911 .....                 | 930 |
| SEQ77  | 911 .....                 | 930 |
| SEQ79  | 911 .....                 | 930 |
| SEQ81  | 911 .....                 | 930 |
| SEQ83  | 911 .....                 | 930 |
| SEQ85  | 911 .....                 | 930 |
| SEQ87  | 911 .....                 | 930 |
| SEQ89  | 911 .....                 | 930 |
| SEQ91  | 911 .....                 | 930 |
| SEQ93  | 911 .....                 | 930 |
| SEQ95  | 911 .....                 | 930 |
| SEQ97  | 911 .....                 | 930 |
| SEQ99  | 911 .....                 | 930 |
| SEQ101 | 911 .....                 | 930 |
| SEQ103 | 911 .....                 | 930 |
| SEQ105 | 911 .....                 | 930 |
| SEQ107 | 911 .....                 | 930 |
| SEQ109 | 911 .....                 | 930 |
| SEQ111 | 911 .....                 | 930 |
| SEQ113 | 911 .....                 | 930 |
| SEQ115 | 911 .....                 | 930 |
| SEQ117 | 911 .....                 | 930 |
| SEQ119 | 911 .....                 | 930 |
| SEQ121 | 911 .....                 | 930 |
| SEQ123 | 911 .....                 | 930 |
| SEQ125 | 911 .....                 | 930 |

Figura 2

Alineaciones por pares de todas las secuencias de aminoácidos contra la secuencia de aminoácidos de la enzima carboxiesterasa de tipo silvestre, Esterasa 2 de *A. acidocaldarius* (SEQ ID NO: 2)

|        |   | 10         | 20         | 30         | 40         | 50         | 60         | 70        |
|--------|---|------------|------------|------------|------------|------------|------------|-----------|
| SEQ2   | 1 | MPLDGVICQV | LDQLNAMDAD | DYRELSAQGF | RSQQSLPDPV | KKEIVAEVRE | PDMELPCRYL | NVMYRPECV |
| SEQ4   | 1 |            |            |            |            |            |            |           |
| SEQ6   | 1 |            |            |            |            |            |            |           |
| SEQ8   | 1 |            |            |            |            |            |            |           |
| SEQ10  | 1 |            |            |            |            |            |            |           |
| SEQ12  | 1 |            |            |            |            |            |            |           |
| SEQ14  | 1 |            |            |            |            |            |            |           |
| SEQ16  | 1 |            |            |            | L          |            |            |           |
| SEQ18  | 1 |            |            |            |            |            |            |           |
| SEQ20  | 1 |            |            |            | L          |            |            |           |
| SEQ22  | 1 |            |            |            |            |            |            |           |
| SEQ24  | 1 |            |            |            | H          |            |            |           |
| SEQ26  | 1 |            |            | P          | L          |            |            |           |
| SEQ28  | 1 |            |            |            |            |            | M          |           |
| SEQ30  | 1 |            |            |            |            |            |            |           |
| SEQ32  | 1 |            |            |            |            |            |            |           |
| SEQ34  | 1 |            |            |            |            |            |            |           |
| SEQ36  | 1 |            | Y          |            | H          | T          |            |           |
| SEQ38  | 1 |            |            |            |            | L          |            |           |
| SEQ40  | 1 |            | Y          | R          | NH         |            |            |           |
| SEQ42  | 1 |            | Y          |            | EH         |            |            |           |
| SEQ44  | 1 |            | Y          | R          | EH         | V          |            |           |
| SEQ46  | 1 |            | Y          |            | NH         | V          |            |           |
| SEQ48  | 1 |            |            |            | H          |            | V          |           |
| SEQ50  | 1 |            | Y          |            | EH         | V          |            |           |
| SEQ52  | 1 |            | Y          |            | RH         |            |            |           |
| SEQ54  | 1 |            | Y          |            | H          |            |            |           |
| SEQ56  | 1 |            | Y          |            | H          |            |            |           |
| SEQ58  | 1 |            | Y          | V          | H          |            |            |           |
| SEQ60  | 1 |            | Y          |            | EH         |            |            |           |
| SEQ62  | 1 |            | Y          |            | H          |            |            |           |
| SEQ64  | 1 |            | Y          |            | H          |            |            |           |
| SEQ66  | 1 |            | Y          | I          | H          |            |            |           |
| SEQ68  | 1 |            | Y          |            | H          |            |            |           |
| SEQ70  | 1 |            | Y          | D          | H          |            |            |           |
| SEQ72  | 1 |            | Y          |            | G          | H          |            |           |
| SEQ74  | 1 |            | Y          | S          | H          |            |            |           |
| SEQ76  | 1 |            | Y          | V          | H          |            |            |           |
| SEQ78  | 1 |            | Y          |            | HIG        |            |            |           |
| SEQ80  | 1 |            | YN         |            | H          | G          |            |           |
| SEQ82  | 1 |            | Y          | S          | H          |            |            |           |
| SEQ84  | 1 |            | Y          | S          | H          |            |            |           |
| SEQ86  | 1 |            | Y          | S          | H          |            |            |           |
| SEQ88  | 1 |            | Y          |            | HMG        |            |            |           |
| SEQ90  | 1 |            | Y          | V          | W          | H          |            |           |
| SEQ92  | 1 |            | Y          | I          | V          | W          | H          | R         |
| SEQ94  | 1 |            | Y          | W          |            | H          |            |           |
| SEQ96  | 1 |            | Y          |            | H          |            | K          |           |
| SEQ98  | 1 |            | Y          |            | H          |            | K          |           |
| SEQ100 | 1 |            | Y          | C          | H          | E          |            |           |
| SEQ102 | 1 |            | Y          |            |            |            |            | L         |
| SEQ104 | 1 |            | Y          |            |            |            |            | L         |
| SEQ106 | 1 |            | Y          | C          | H          |            | R          |           |
| SEQ108 | 1 |            | LY         | C          | H          | E          |            |           |
| SEQ110 | 1 |            | Y          | C          | H          | E          | K          |           |
| SEQ112 | 1 |            | Y          | C          | HT         | W          | Q          | T         |
| SEQ114 | 1 |            | Y          | C          | HT         | W          |            |           |
| SEQ116 | 1 |            | Y          | W          | HN         | W          |            |           |
| SEQ118 | 1 |            | Y          | C          | HT         | W          |            |           |
| SEQ120 | 1 |            | Y          | C          | H          | E          |            |           |
| SEQ122 | 1 |            | Y          | C          | HT         | W          |            |           |
| SEQ124 | 1 | L          | Y          | C          | HT         | W          |            |           |
| SEQ126 | 1 |            | Y          | C          | HT         | W          |            |           |

Figura 2 (Continuación)

|        |    | 30        | 50         | 100      | 110      | 120   | 130   | 140      |     |
|--------|----|-----------|------------|----------|----------|-------|-------|----------|-----|
| SEQ2   | 71 | .....     | .....      | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ4   | 71 | .....     | .....      | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ6   | 71 | .....     | .....      | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ8   | 71 | .....     | .....      | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ10  | 71 | .....     | .....      | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ12  | 71 | .....     | .....      | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ14  | 71 | .....     | .....      | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ16  | 71 | .....     | .....      | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ18  | 71 | .....     | .....      | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ20  | 71 | ...R..... | .....      | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ22  | 71 | .....     | .....      | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ24  | 71 | .....     | .....      | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ26  | 71 | .....     | .....      | .....    | ..W..... | ..... | ..... | .....    | 140 |
| SEQ28  | 71 | .....     | .....      | .....    | ..W..... | ..... | ..... | .....    | 140 |
| SEQ30  | 71 | .....     | .....      | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ32  | 71 | .....     | .....      | .....    | ..S..... | ..... | ..... | .....    | 140 |
| SEQ34  | 71 | .....     | .....      | .....    | ..S..... | ..... | ..... | .....    | 140 |
| SEQ36  | 71 | .....     | .....      | .....    | ..W..... | ..... | ..... | .....    | 140 |
| SEQ38  | 71 | .....     | .....      | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ40  | 71 | .....     | ..A.....   | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ42  | 71 | .....     | ..A.....   | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ44  | 71 | .....     | ..A.....   | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ46  | 71 | .....     | ..A.....   | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ48  | 71 | .....     | ..A.....   | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ50  | 71 | .....     | ..A.....   | .....    | .....    | ..... | ..... | ..R..... | 140 |
| SEQ52  | 71 | .....     | ..A.....   | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ54  | 71 | .....     | ..A.....   | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ56  | 71 | .....     | ..A.....   | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ58  | 71 | .....     | ..A.....   | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ60  | 71 | .....     | ..A.....   | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ62  | 71 | .....     | ..A.....   | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ64  | 71 | .....     | ..A.....   | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ66  | 71 | .....     | ..A.....   | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ68  | 71 | .....     | ..A.....   | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ70  | 71 | .....     | ..A.....   | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ72  | 71 | .....     | ..A.....   | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ74  | 71 | .....     | ..A.....   | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ76  | 71 | .....     | ..A.....   | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ78  | 71 | .....     | ..A.....   | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ80  | 71 | .....     | ..A.....   | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ82  | 71 | .....     | ..A.....   | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ84  | 71 | .....     | ..A.....   | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ86  | 71 | .....     | ..A.....   | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ88  | 71 | .....     | ..A.....   | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ90  | 71 | ...R..... | ..A.....   | .....    | ..W..... | ..... | ..... | .....    | 140 |
| SEQ92  | 71 | .....     | ..A.....   | .....    | ..W..... | ..... | ..... | .....    | 140 |
| SEQ94  | 71 | ...R..... | ..A.....   | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ96  | 71 | .....     | ..A.....   | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ98  | 71 | .....     | ..A.....   | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ100 | 71 | .....     | ..G.A..... | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ102 | 71 | .....     | ..TA.....  | ..M..... | .....    | ..... | ..... | .....    | 140 |
| SEQ104 | 71 | .....     | ..TA.....  | ..M..... | .....    | ..... | ..... | .....    | 140 |
| SEQ106 | 71 | ...R..... | ..A.....   | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ108 | 71 | .....     | ..G.A..... | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ110 | 71 | .....     | ..G.A..... | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ112 | 71 | .....     | ..G.A..... | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ114 | 71 | ...R..... | ..G.A..... | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ116 | 71 | ...R..... | ..G.A..... | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ118 | 71 | ...M..... | ..G.A..... | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ120 | 71 | .....     | ..G.A..... | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ122 | 71 | ...R..... | ..G.A..... | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ124 | 71 | ...R..... | ..G.A..... | .....    | .....    | ..... | ..... | .....    | 140 |
| SEQ126 | 71 | ...R..... | ..G.A..... | .....    | .....    | ..... | ..... | .....    | 140 |

Figura 2 (Continuación)

|        |     | 150         | 160        | 170        | 180        | 190       | 200        | 210       |
|--------|-----|-------------|------------|------------|------------|-----------|------------|-----------|
| SEQ2   | 141 | EPHLEPTARIA | VOCD8AGGNL | AAVTSILAKE | POCPALAPQL | LIYPTCYDT | ANDPASISEN | NECYLLTGM |
| SEQ4   | 141 | .....       | .....      | .....      | .....      | .....     | L.         | 210       |
| SEQ6   | 141 | .....       | .....      | .....      | .....      | .....     | I.         | 210       |
| SEQ8   | 141 | .....       | .....      | .....      | .....      | .....     | Y.         | 210       |
| SEQ10  | 141 | .....       | .....      | .....      | .....      | .....     | F.         | 210       |
| SEQ12  | 141 | .....       | .....      | .....      | .....      | .....     | M.         | 210       |
| SEQ14  | 141 | .....       | .....      | .....      | M.         | .....     | D.         | 210       |
| SEQ16  | 141 | .....       | .....      | .....      | F.         | .....     | EH.        | 210       |
| SEQ18  | 141 | .....       | .....      | .....      | F.         | .....     | L.         | 210       |
| SEQ20  | 141 | .....       | .....      | .....      | F.         | .....     | L.         | 210       |
| SEQ22  | 141 | .....       | .....      | .....      | F.         | .....     | L.         | 210       |
| SEQ24  | 141 | .....       | .....      | .....      | F.         | .....     | L.         | 210       |
| SEQ26  | 141 | .....       | .....      | .....      | F.         | .....     | I.         | 210       |
| SEQ28  | 141 | .....       | .....      | .....      | F.         | .....     | L.         | 210       |
| SEQ30  | 141 | .....       | .....      | .....      | F.         | .....     | L.         | 210       |
| SEQ32  | 141 | .....       | .....      | .....      | F.         | .....     | L.         | 210       |
| SEQ34  | 141 | .....       | .....      | .....      | F.         | .....     | L.         | 210       |
| SEQ36  | 141 | .....       | .....      | .....      | F.         | .....     | L.         | 210       |
| SEQ38  | 141 | .....       | .....      | .....      | F.         | .....     | L.         | 210       |
| SEQ40  | 141 | .....       | .....      | .....      | F.         | .....     | L.         | 210       |
| SEQ42  | 141 | .....       | .....      | .....      | F.         | .....     | L.         | 210       |
| SEQ44  | 141 | .....       | .....      | .....      | F.         | .....     | L.         | 210       |
| SEQ46  | 141 | .....       | .....      | .....      | F.         | .....     | L.         | 210       |
| SEQ48  | 141 | .....       | .....      | .....      | F.         | Y.        | L.         | 210       |
| SEQ50  | 141 | .....       | .....      | .....      | F.         | .....     | L.         | 210       |
| SEQ52  | 141 | .....       | .....      | .....      | F.         | .....     | L.         | 210       |
| SEQ54  | 141 | .....       | .....      | .....      | F.         | .....     | L.         | 210       |
| SEQ56  | 141 | .....       | .....      | .....      | F.         | .....     | L.         | 210       |
| SEQ58  | 141 | .....       | .....      | .....      | F.         | .....     | L.         | 210       |
| SEQ60  | 141 | .....       | .....      | .....      | F.         | .....     | L.         | 210       |
| SEQ62  | 141 | .....       | .....      | .....      | F.         | F.        | L.         | 210       |
| SEQ64  | 141 | .....       | .....      | .....      | F.         | .....     | L.         | 210       |
| SEQ66  | 141 | .....       | .....      | .....      | F.         | .....     | L.         | 210       |
| SEQ68  | 141 | .....       | .....      | .....      | F.         | .....     | L.         | 210       |
| SEQ70  | 141 | .....       | .....      | .....      | F.         | .....     | L.         | 210       |
| SEQ72  | 141 | .....       | .....      | .....      | F.         | .....     | L.         | 210       |
| SEQ74  | 141 | .....       | .....      | .....      | F.         | .....     | L.         | 210       |
| SEQ76  | 141 | .....       | .....      | .....      | F.         | .....     | L.         | 210       |
| SEQ78  | 141 | .....       | .....      | .....      | F.         | .....     | L.         | 210       |
| SEQ80  | 141 | .....       | .....      | .....      | F.         | .....     | L.         | 210       |
| SEQ82  | 141 | .....       | .....      | .....      | F.         | .....     | L.         | 210       |
| SEQ84  | 141 | .....       | .....      | .....      | F.         | .....     | L.         | 210       |
| SEQ86  | 141 | .....       | .....      | .....      | F.         | .....     | L.         | 210       |
| SEQ88  | 141 | .....       | .....      | .....      | F.         | .....     | L.         | 210       |
| SEQ90  | 141 | .....       | .....      | .....      | F.         | .....     | L.         | 210       |
| SEQ92  | 141 | .....       | .....      | .....      | F.         | .....     | L.         | 210       |
| SEQ94  | 141 | .....       | .....      | .....      | F.         | .....     | L.         | 210       |
| SEQ96  | 141 | .....       | F.         | .....      | F.         | .....     | L.         | 210       |
| SEQ98  | 141 | .....       | F.         | .....      | F.         | .....     | L.         | 210       |
| SEQ100 | 141 | .....       | F.         | .....      | F.         | .....     | L.         | 210       |
| SEQ102 | 141 | .....       | P.         | .....      | Q.         | F.        | F.         | 210       |
| SEQ104 | 141 | .....       | P.         | .....      | F.         | .....     | L.         | 210       |
| SEQ106 | 141 | .....       | P.         | .....      | F.         | .....     | L.         | 210       |
| SEQ108 | 141 | .....       | P.         | .....      | F.         | .....     | L.         | 210       |
| SEQ110 | 141 | .....       | P.         | .....      | F.         | .....     | L.         | 210       |
| SEQ112 | 141 | .....       | P.         | .....      | F.         | .....     | L.         | 210       |
| SEQ114 | 141 | .....       | P.         | .....      | F.         | .....     | L.         | 210       |
| SEQ116 | 141 | .....       | P.         | .....      | F.         | .....     | L.         | 210       |
| SEQ118 | 141 | .....       | P.         | .....      | F.         | .....     | L.         | 210       |
| SEQ120 | 141 | .....       | P.         | .....      | F.         | .....     | L.         | 210       |
| SEQ122 | 141 | .....       | P.         | .....      | F.         | .....     | L.         | 210       |
| SEQ124 | 141 | .....       | P.         | .....      | F.         | .....     | L.         | 210       |
| SEQ126 | 141 | .....       | P.         | .....      | Q.         | F.        | L.         | 210       |

Figura 2 (Continuación)

|        |     | 220               | 230        | 240        | 250         | 260         | 270         | 280        |
|--------|-----|-------------------|------------|------------|-------------|-------------|-------------|------------|
| SEQ2   | 211 | MLWFRDQILN        | SLEELTHFWF | SPVLYPDLSG | LPSAYIATAQ  | YDPLRDVGH   | YAEALNKAGV  | RVEIENFEDL |
| SEQ4   | 211 | .....             | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ6   | 211 | .....             | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ8   | 211 | .....             | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ10  | 211 | .....             | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ12  | 211 | .....             | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ14  | 211 | .....             | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ16  | 211 | .....             | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ18  | 211 | .....             | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ20  | 211 | .....             | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ22  | 211 | .....             | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ24  | 211 | .....             | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ26  | 211 | .....             | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ28  | 211 | .....             | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ30  | 211 | .....             | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ32  | 211 | .....             | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ34  | 211 | .....             | .....      | .....      | .....       | .....       | E.....      | .....      |
| SEQ36  | 211 | .....             | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ38  | 211 | .....N.....       | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ40  | 211 | .....             | .....      | .....      | .....       | .....       | D.....      | .....      |
| SEQ42  | 211 | .....             | .....      | .....      | .....       | .....       | D.....      | .....      |
| SEQ44  | 211 | .....             | .....      | .....      | .....       | .....       | D.....Q..   | .....      |
| SEQ46  | 211 | .....             | .....      | .....      | .....       | .....       | .....Q..    | .....      |
| SEQ48  | 211 | .....             | .....      | .....      | .....       | .....E..... | .....       | .....      |
| SEQ50  | 211 | .....             | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ52  | 211 | .....             | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ54  | 211 | .....             | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ56  | 211 | .....             | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ58  | 211 | .....             | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ60  | 211 | .....             | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ62  | 211 | .....             | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ64  | 211 | .....I.....       | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ66  | 211 | .....             | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ68  | 211 | .....N.....       | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ70  | 211 | .....             | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ72  | 211 | .....             | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ74  | 211 | .....             | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ76  | 211 | .....N.....       | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ78  | 211 | .....N.....       | .....      | .....      | .....       | .....       | .....T..... | .....      |
| SEQ80  | 211 | .....N.....       | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ82  | 211 | .....N.....       | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ84  | 211 | .....N.....       | .....      | .....      | .....       | .....T..... | .....       | .....      |
| SEQ86  | 211 | .....N.....       | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ88  | 211 | .....N.....       | .....      | .....      | .....       | .....       | .....M..... | .....      |
| SEQ90  | 211 | .....N.....       | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ92  | 211 | .....N.....       | .....      | .....      | .....       | .....R..... | .....       | .....      |
| SEQ94  | 211 | .....N.....       | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ96  | 211 | .....N.....       | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ98  | 211 | .....N.....       | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ100 | 211 | .....N.....       | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ102 | 211 | .....B.....N..... | .....      | .....      | .....N..... | .....       | .....L..... | .....      |
| SEQ104 | 211 | .....N.....       | .....      | .....      | .....W..... | .....       | .....       | .....      |
| SEQ106 | 211 | .....N.....       | .....      | .....      | .....V..... | .....       | .....       | .....      |
| SEQ108 | 211 | .....N.....       | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ110 | 211 | .....N.....       | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ112 | 211 | .....N.....       | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ114 | 211 | .....N.....       | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ116 | 211 | .....N.....       | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ118 | 211 | .....N.....       | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ120 | 211 | .....N.....       | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ122 | 211 | .....N.....       | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ124 | 211 | .....N.....       | .....      | .....      | .....       | .....       | .....       | .....      |
| SEQ126 | 211 | .....N.....       | .....      | .....      | .....       | .....       | .....       | .....      |

Figura 2 (Continuación)

|        |     | 280    | 290    | 310    |
|--------|-----|--------|--------|--------|
| SEQ2   | 281 | .....I | .....I | .....I |
| SEQ4   | 281 | .....I | .....I | .....I |
| SEQ6   | 281 | .....I | .....I | .....I |
| SEQ8   | 281 | .....I | .....I | .....I |
| SEQ10  | 281 | .....I | .....I | .....I |
| SEQ12  | 281 | .....I | .....I | .....I |
| SEQ14  | 281 | .....I | .....I | .....I |
| SEQ16  | 281 | .....I | .....I | .....I |
| SEQ18  | 281 | .....I | .....I | .....I |
| SEQ20  | 281 | .....I | .....I | .....I |
| SEQ22  | 281 | .....I | .....I | .....I |
| SEQ24  | 281 | .....I | .....I | .....I |
| SEQ26  | 281 | .....I | .....I | .....I |
| SEQ28  | 281 | .....I | .....I | .....I |
| SEQ30  | 281 | .....I | .....I | .....I |
| SEQ32  | 281 | .....I | .....I | .....I |
| SEQ34  | 281 | .....I | .....I | .....I |
| SEQ36  | 281 | .....I | .....I | .....I |
| SEQ38  | 281 | .....I | .....I | .....I |
| SEQ40  | 281 | .....I | .....I | .....I |
| SEQ42  | 281 | .....I | .....I | .....I |
| SEQ44  | 281 | .....I | .....I | .....I |
| SEQ46  | 281 | .....I | .....I | .....I |
| SEQ48  | 281 | .....I | .....I | .....I |
| SEQ50  | 281 | .....I | .....I | .....I |
| SEQ52  | 281 | .....I | .....I | .....I |
| SEQ54  | 281 | .....I | .....I | .....I |
| SEQ56  | 281 | .....I | .....I | .....I |
| SEQ58  | 281 | .....I | .....I | .....I |
| SEQ60  | 281 | .....I | .....I | .....I |
| SEQ62  | 281 | .....I | .....I | .....I |
| SEQ64  | 281 | .....I | .....I | .....I |
| SEQ66  | 281 | .....I | .....I | .....I |
| SEQ68  | 281 | .....I | .....I | .....I |
| SEQ70  | 281 | .....I | .....I | .....I |
| SEQ72  | 281 | .....I | .....I | .....I |
| SEQ74  | 281 | .....I | .....I | .....I |
| SEQ76  | 281 | .....I | .....I | .....I |
| SEQ78  | 281 | .....I | .....I | .....I |
| SEQ80  | 281 | .....I | .....I | .....I |
| SEQ82  | 281 | .....I | .....I | .....I |
| SEQ84  | 281 | .....I | .....I | .....I |
| SEQ86  | 281 | .....I | .....I | .....I |
| SEQ88  | 281 | .....I | .....I | .....I |
| SEQ90  | 281 | .....I | .....I | .....I |
| SEQ92  | 281 | .....I | .....I | .....I |
| SEQ94  | 281 | .....I | .....I | .....I |
| SEQ96  | 281 | .....I | .....I | .....I |
| SEQ98  | 281 | .....I | .....I | .....I |
| SEQ100 | 281 | .....I | .....I | .....I |
| SEQ102 | 281 | .....I | .....I | .....I |
| SEQ104 | 281 | .....I | .....I | .....I |
| SEQ106 | 281 | .....I | .....I | .....I |
| SEQ108 | 281 | .....I | .....I | .....I |
| SEQ110 | 281 | .....I | .....I | .....I |
| SEQ112 | 281 | .....I | .....I | .....I |
| SEQ114 | 281 | .....I | .....I | .....I |
| SEQ116 | 281 | .....I | .....I | .....I |
| SEQ118 | 281 | .....I | .....I | .....I |
| SEQ120 | 281 | .....I | .....I | .....I |
| SEQ122 | 281 | .....I | .....I | .....I |
| SEQ124 | 281 | .....I | .....I | .....I |
| SEQ126 | 281 | .....I | .....I | .....I |