

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 583 259**

51 Int. Cl.:

C07K 14/195 (2006.01)

C12N 9/88 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **01.12.2010 E 10788057 (7)**

97 Fecha y número de publicación de la concesión europea: **20.04.2016 EP 2507258**

54 Título: **Nuevas liasas alfa-amidantes de peptidil alfa-hidroxiglicina**

30 Prioridad:

01.12.2009 EP 09177593

04.12.2009 US 266711 P

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

20.09.2016

73 Titular/es:

NOVO NORDISK A/S (100.0%)

Novo Allé

2880 Bagsværd, DK

72 Inventor/es:

SHAW, ALLAN CHRISTIAN

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 583 259 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

Nuevas liasas alfa-amidantes de peptidil alfa-hidroxiglicina

Campo de la invención

5 La presente invención se refiere a polipéptidos aislados que tienen actividad liasa alfa-amidante de peptidil- α -hidroxiglicina, a métodos para preparar tales polipéptidos y al uso de tales polipéptidos en procedimientos para la producción de péptidos α -amidados C-terminales.

Antecedentes de la invención

10 En los organismos multicelulares, ciertos péptidos ("precursores"), tales como los neuropéptidos, se modifican de forma post-traducciona a través de una serie de etapas enzimáticas que escinden y modifican adicionalmente sus-
tratos peptídicos para producir péptidos biorreactivos totalmente funcionales. El proceso comienza en el aparato de trans-Golgi y continúa como gránulos secretores inmaduros. Una modificación post-traducciona de etapa tardía muy importante para muchos de estos péptidos es la amidación alfa carboxiterminal.

15 La amidación en alfa de residuos C-terminales es esencial para la actividad de varias hormonas peptídicas implicadas en el metabolismo humano o animal. Diversas hormonas peptídicas se utilizan en la actualidad como fármacos en el tratamiento de seres humanos, por ejemplo, para el control de la obesidad y/o la diabetes, o están en desarrollo como fármacos potenciales. Un ejemplo de una hormona peptídica de este tipo es la amilina (por ejemplo, Sym-
lin[®], acetato de pramlintida, que es un análogo de la amilina humana). La amilina humana es un péptido con 37 residuos de aminoácidos que se puede utilizar para tratar o prevenir la obesidad y/o la diabetes. Por consiguiente, el extremo C-terminal de la amilina se tiene que amidar con el fin de obtener una actividad biológica completa. Asimismo, el péptido YY (PYY) se tiene que amidar en alfa para obtener una actividad biológica completa.

20 *E. coli* y levadura se emplean extensamente para la expresión recombinante de péptidos de origen eucariota. Sin embargo, debido a la naturaleza de las hormonas peptídicas con un grupo α -amida C-terminal, no se pueden expresar en una forma activa, empleando los sistemas de expresión microbianos del estado de la técnica basados en *E. coli* y levadura, ya que las enzimas alfa amidantes no se expresan de forma natural en estos organismos. Por lo tanto, se tiene que introducir una amida alfa C-terminal en los péptidos expresados de forma recombinante utilizando, por ejemplo, una modificación *ex vivo* con enzimas alfa amidantes.

25 La modificación enzimática de una Gly C-terminal a una α -amida, de precursores de péptidos por medio de una monooxigenasa α -amidante de peptidilo (PAM) bifuncional, se encuentra en varios organismos eucariotas. Se han descrito múltiples variantes de transcritos por corte y empalme de forma alternativa que codifican diferentes isoformas, para esta enzima. La enzima se ha descrito exclusivamente para organismos multicelulares (metazoos). La conversión en un péptido de un residuo Gly C-terminal a una α -amida es un proceso en dos etapas, en el que el dominio N-terminal de PAM (denominado PHM) cataliza la conversión de Glicina a α -hidroxiglicina y el dominio C-terminal de PAM (denominado PAL) cataliza la conversión de la α -hidroxiglicina a α -amida. En organismos eucariotas, los dos dominios catalíticos actúan secuencialmente para catalizar péptidos neuroendocrinos en productos activos alfa-amidados. Dos puentes disulfuro están muy conservados en los dominios PAL de organismos eucariotas.

30 Si bien puede ser posible sintetizar por medios químicos pequeños péptidos que contienen un grupo amida en el extremo C-terminal (amida alfa), péptidos mayores alfa-amidados son difíciles de producir y caros. Enzimas que realizan la amidación en alfa son por tanto útiles en la conversión de péptidos precursores recombinantes en péptidos maduros.

40 El documento US4708934 describe una monooxigenasa α -amidante de peptidil-glicina extraída a partir de líneas celulares de carcinoma medular de tiroides y muestras de tejidos.

El documento US5789234 describe la producción de una enzima alfa-amidante por técnicas de ADN recombinante.

45 El documento WO90/08194 se refiere a un procedimiento para la producción de péptido alfa amidado en el extremo C-terminal a partir de un péptido precursor mediante el uso de una enzima alfa amidante del extremo C-terminal eucariota. También se describe un método para la expresión eucariota de estas enzimas alfa amidantes del extremo C-terminal.

El documento WO89/02460 describe una enzima PAM de origen bovino, su clonación, su ADNc y expresión mediante tecnología de ADN recombinante.

50 El documento EP0448513 describe un procedimiento para la expresión recombinante de una monooxigenasa alfa-hidroxilante de peptidilglicina obtenida a partir de *Xenopus laevis*, que comprende cultivar células de insecto transfectadas con un baculovirus recombinante al que se ha incorporado un ADN que codifica la monooxigenasa alfa-hidroxilante de peptidilglicina para producir la enzima.

El documento EP0465404 describe una enzima (PHL; PAL) obtenida a partir de *Xenopus laevis* que cataliza la escisión del enlace N-C en el resto α -hidroxiglicina de un péptido α -hidroxilado en el extremo C-terminal, la clonación de

la enzima y su expresión recombinante en células de insecto.

El documento 20060292672 describe una línea celular para expresar PAM o uno de sus dos dominios catalíticos.

5 El documento EP2172550 describe un derivado de enzima recombinante alfa-amidante en el extremo C-terminal que carece de la formación de uno de los cinco enlaces disulfuro que se producen normalmente en una enzima alfa-amidante del extremo C-terminal obtenida a partir de *Xenopus laevis* y un método de producción de dicho derivado de forma recombinante en *E. coli*, en donde el cuerpo de inclusión obtenido se solubiliza y se somete a un procedimiento de replegamiento.

Compendio de la invención

10 La presente invención proporciona un método para producir un péptido α -amidado, que comprende las etapas de: (i) permitir que un péptido diana con un residuo de glicina C-terminal reaccione tanto con una monooxigenasa alfa-hidroxilante de peptidilglicina (PHM) como con un polipéptido procedente de un organismo procariota con actividad liasa alfa-amidante de peptidil- α -hidroxiglicina (PAL), en condiciones adecuadas para la actividad enzimática, en donde la reacción con dicha PHM y dicho polipéptido en dicho péptido diana se lleva a cabo ya sea en dos etapas distintas o simultáneamente; (ii) recuperar el péptido α -amidado en el extremo C-terminal.

15 La invención también se refiere a enzimas, que son capaces de catalizar la conversión de una α -hidroxiglicina a una α -amida en un péptido (actividad liasa alfa-amidante de peptidil- α -hidroxiglicina).

Las enzimas se obtienen a partir de organismos procariotas y tienen diferentes propiedades fisicoquímicas y estructurales que las que se describen para las enzimas PAL eucariotas.

20 La invención también proporciona enzimas que tienen actividad liasa alfa-amidante de peptidil- α -hidroxiglicina caracterizadas porque se pueden expresar en *E. coli* como proteínas solubles, enzimáticamente activas.

La liasa alfa-amidante de peptidil- α -hidroxiglicina tiene una secuencia de aminoácidos que no comprende residuos de cisteína, o como máximo 1 residuo de cisteína o como máximo, 2 residuos de cisteína.

En una realización, la liasa alfa-amidante de peptidil- α -hidroxiglicina es la enzima de tipo silvestre o comprende de 1-30 modificaciones de aminoácidos en comparación con el polipéptido de tipo silvestre.

25 En otra realización, el polipéptido procedente de un organismo procariota con actividad liasa alfa-amidante de peptidil- α -hidroxiglicina (PAL) se ha expresado en *E. coli* recombinante.

En aún otra realización del método de la invención, la reacción se realiza en condiciones en las que están presentes iones Cu^{2+} y/o ácido ascórbico cuando se utiliza dicho polipéptido.

Descripción de los dibujos

30 **Figura 1:** Mapa vectorial representativo de pET11a del plásmido C que codifica el dominio de tipo PAL bacteriano (sin péptidos señal previstos) procedente de *Erythrobacter* (SEQ ID NO: 1) con una pareja de fusión N-terminal (SEQ ID NO: 7) y un sitio de escisión de la proteasa 3C de HRV14 en la región enlazadora intercalada (SEQ ID NO: 8) (La proteína de fusión completa está marcada con la característica Proteína 3). Se indican los sitios de las enzimas de restricción NdeI, XhoI y BamHI. La región promotora T7, el gen de resistencia a la ampicilina, la región re-
35 presora lacI y el origen del sitio de replicación también se muestran en el mapa del vector.

Figura 2: SDS-PAGE que muestra el perfil de expresión del dominio de tipo PAL de *Erythrobacter* marcado con RL9_THEMA TAP (A) y PAL rata (B). Uind: control de cultivo celular no inducido (control neg.), Ind: después de la inducción con IPTG 0,5 mM durante 3 horas a 30°C, Sup: fracción soluble, Pel: fracción insoluble. La flecha indica la posición del peso molecular de las bandas de PAL sobre los geles.

40 Aunque los niveles de expresión son comparables, el dominio de tipo PAL de *Erythrobacter* se expresa como una proteína altamente soluble, en contraste con la PAL de rata, que es insoluble y requerirá un replegamiento para obtener una enzima funcional.

45 **Figura 3:** A: Cromatograma de FPLC, que muestra el perfil de purificación en SP Sefarosa FF del dominio de tipo PAL de *Erythrobacter* marcado con RL9_THEMA. La proteína de fusión eluye con una conc. de NaCl de ~0,25 M indicada por la flecha. La línea de puntos indica la curva de conductividad, la línea continua indica la señal de absorción de UV a 280 nm. Los números en el eje X indican fracciones y ml.

B: gel SDS-PAGE que muestra las fracciones del pico principal en el cromatograma de la Figura 3A. Apl: Aplicación cargada en la columna, Ft: Fracciones del flujo continuo. El marcador TAP proporciona una captura eficaz y de pureza elevada, después una etapa de cromatografía de intercambio de cationes.

50 **Figura 4A:** Cromatograma de FPLC que muestra la separación en Q Sefarosa HP del dominio maduro de tipo PAL de *Exiguobacterium* liberado del marcador RL9_THEMA, después del procesamiento de la proteína de fusión con la

proteasa 3C de HRV14. Un pico principal eluye a NaCl -0,4 M como indica la flecha. La línea de puntos indica la curva de conductividad, la línea continua indica la señal de UV a 280 nm. Los números en el eje X indican fracciones y ml.

5 **Figura 4B:** análisis SDS-PAGE de la fracción eluida de la separación FPLC mostrada en la Figura 4A. Apl: Aplicación cargada en la columna, Ft: Fracciones del flujo continuo. En el carril de la aplicación, dos bandas representan el dominio maduro de tipo PAL de *Exiguobacterium* (~30 kDa) y el marcador de purificación liberado (~18 kDa). El pico principal de la purificación contiene casi exclusivamente PAL madura (flecha negra), mientras que el marcador TAP liberado no se une a la columna de intercambio aniónico y está presente en la fracción de flujo continuo (flecha punteada).

10 **Figura 5:** análisis UPLC de ácido α -hidroxi hipúrico sintético y benzamida que representa α -hidroxiglicina (Gly(OH)) y α -amida (-NH₂), respectivamente (marcadas por debajo de los picos relevantes). El ácido α -hidroxi hipúrico se incubó con enzimas de tipo PAL bacterianas con o sin marcador TAP en el extremo N-terminal durante 3 horas a 37°C, con adición de cofactores relevantes. Las flechas en negrita señalan las conversiones enzimáticas más significativas a benzamida.

15 **Figura 5.A:** Análisis del dominio de tipo PAL de *Erythrobacter* con (+TAP) (SEQ ID NO: 7) o sin (-TAP) marcador de purificación, sometido a ensayo empleando condiciones de tampón ácidas (MES pH 5,5) o básicas (Tris pH 7,5). Los controles (ctrl) negativos (neg.) son los cromatogramas para muestras sin adición de enzima. La conversión de ácido alfa-hidroxihipúrico a benzamida puede ocurrir espontáneamente a pH elevado (Fig. 5A6/7) como se ha descrito anteriormente en la bibliografía. Sin embargo, basándose en las áreas relativas de ácido α -hidroxi hipúrico y de las áreas del pico de sustrato de benzamida, se concluye que la adición de enzima puede catalizar la conversión, de manera más eficaz a pH 7,5, tanto con marcador TAP (Fig. 5A2/7) como sin marcador TAP (Fig. 5A4/7).

20 **Figura 5.B:** Análisis del dominio de tipo PAL de *Chthoniobacter* con marcador de purificación (+TAP) (SEQ ID NO: 7) sometido a ensayo utilizando condiciones de tampón ácidas (MES pH 5,5) o básicas (Tris pH 7,5). La enzima de *Chthoniobacter* es más activa en el intervalo de pH más bajo en comparación con la enzima de *Erythrobacter*, ya que el área del pico de benzamida se incrementa significativamente después de la incubación con la enzima a pH 5,5 (Fig. 5B1/7) en comparación con el ctrl. neg. (Fig. 5B3/7).

25 **Figura 6:** Espectros MS extraídos y desconvolucionados de péptido modelo análogo de amilina que confirma la actividad de PAL de dominios de tipo PAL bacterianos, mediante incubaciones conjuntas con PHM de rata. Eje X: relación masa carga en Dalton. Eje Y: intensidad relativa de los picos de MS. Las masas de los picos prominentes de MS se indican en el espectro; Control del péptido precursor (que contiene un residuo Gly C-terminal): 13934.66 Da, péptido intermedio con α -hidroxiglicina (que comprende una α -hidroxiglicina C-terminal): 13951 Da, péptido α -amidado: 13876.62 Da

30 A: Control no tratado, B: PHM de rata sola (2 horas), C: dominio de tipo PAL de *Planctomyces*, proteína 10 (2 horas), D: dominio de tipo PAL de *Erythrobacter*, proteína 3 (2 horas), E: dominio de tipo PAL de *Planctomyces*, proteína 10 (5 horas), F: dominio de tipo PAL de *Erythrobacter*, proteína 3 (5 horas).

35 **Figura 7:** Espectros MS extraídos y desconvolucionados de péptido purificado de Pramlintida extendido con alfa hidroxi-glicina C-terminal, incubado con cuatro dominios diferentes bacterianos de tipo PAL (SEQ ID NO: 1-4). Eje X: relación masa carga en Dalton. Eje Y: intensidad relativa de los picos de MS. Las masas de los dos picos prominentes de MS se indican en el espectro; los picos del péptido intermedio de α -hidroxiglicina (que comprende una α -hidroxiglicina C-terminal: ~4023 Da) y los picos de péptido α -amidado (~3949 Da) se determinaron todos muy cerca de las masas isotópicas promedio calculadas. Se indican las relaciones enzima sustrato (p/p) para cada reacción.

A: *Erythrobacter* (SEQ ID NO: 1) i) control tratado con PHM (ctrl), ii) 1:1250, iii) 1:500, iv) 1:100

B: *Planctomyces* (SEQ ID NO:4) i) control tratado con PHM (ctrl), ii) 1:1250, iii) 1:500, iv) 1:100.

C: *Chthoniobacter* (SEQ ID NO:3) i) control tratado con PHM (ctrl), ii) 1:100, iii) 1:50, iv), 1:25.

40 D: *Exiguobacterium* (SEQ ID NO:2) i) control tratado con PHM (ctrl), ii) 1:100, iii) 01:50, iv) 01:25

Definiciones

45 **"PAM"** o **"monooxigenasa alfa-amidante de peptidil-glicina"** se refiere a la enzima bifuncional que cataliza la conversión de un residuo de glicina C-terminal a una α -hidroxiglicina y la conversión de una α -hidroxiglicina a una α -amida. La enzima también se conoce como peptidilglicina 2-hidroxilasa, enzima alfa-amidante de peptidilo, monooxigenasa alfa-hidroxilante de peptidilglicina y liasa alfa-amidante de peptidil-alfa-hidroxiglicina.

"PHM" o **"monooxigenasa alfa-hidroxilante de peptidilglicina"** es una enzima, capaz de catalizar la conversión de un residuo de glicina C-terminal a una α -hidroxiglicina. Otros términos para PHM son: peptidilglicina 2-hidroxilasa, monooxigenasa alfa-amidante de peptidilglicina, peptidilglicina alfa-hidroxilasa, monooxigenasa alfa-hidroxilante de peptidilglicina, monooxigenasa alfa-hidroxilante de peptidilglicina, peptidilglicina alfa-monooxigenasa, EC 1.14.17.3 y

peptidilglicina monooxigenasa.

Tal como se usa en el presente documento, "**enzima PAL**", "**PAL**" o "**liasa alfa-amidante de peptidil- α -hidroxiglicina**" es una enzima capaz de catalizar la conversión de una α -hidroxiglicina a una α -amida. Sinónimos de PAL son: peptidilamidoglicolato liasa, desalquilasa amidante de α -hidroxiglicina, hGAD, PGL, peptidilamidoglicolato peptidilamida-liasa, EC 4.3.2.5, y peptidilhidroxiglicina N-C-liasa (PHL). La actividad de la enzima PAL se puede mostrar tal y como se describe en el Ensayo (I).

Por "**una enzima capaz de catalizar la conversión de una α -hidroxiglicina C-terminal a una α -amida**" se entiende una enzima que es capaz de catalizar la reacción $R\text{-Gly(OH)} \rightarrow R\text{-NH}_2$, en donde R es un péptido, una proteína o un compuesto químico.

En la presente memoria, la expresión "una enzima de la invención" significa un polipéptido de la invención que tiene actividad liasa alfa-amidante de peptidil- α -hidroxiglicina.

Sustratos no peptídicos, tales como el ácido alfa-hidroxihipúrico también sirven como sustratos para PAL.

La expresión "**de tipo PAL**" se entiende que indica una enzima que tiene la misma actividad que la enzima PAL eucariota conocida.

Por "**péptido diana**" se entiende el péptido que se modifica en un proceso de α -amidación para obtener un grupo α -amida C-terminal. El péptido diana debe comprender un residuo Gly en el extremo C-terminal. El péptido diana se puede describir por tener la fórmula $R^1\text{-X-Gly}$, en la que X representa cualquier aminoácido y en donde X es el aminoácido que se va a convertir en una amida de aminoácido, es decir, para el que $-\text{COOH}$ se va a convertir en CO-NH_2 en la reacción del proceso enzimático de α -amidación, R^1 representa una porción restante del péptido y Gly representa un residuo de glicina C-terminal.

Un ejemplo de un péptido diana es un precursor de amilina, que además de la secuencia de amilina comprende un residuo Gly en el extremo C-terminal. Otros ejemplos no limitativos de precursores de péptidos con una extensión de Gly, pertinentes para la invención incluyen el neuropéptido Y (NPY), el péptido YY (PYY), PYY-3-36, el polipéptido pancreático (PP), el péptido similar a glucagón (GLP-1), la gastrina, la calcitonina, el péptido relacionado con calcitonina (CGRP), el péptido liberador de gastrina, la vasopresina, la oxitocina, la neuroquinina A, la secretina, la pancreastatina, la pro-opiomelanocortina (POMC), la hormona estimulante de melanocitos alfa (MSH alfa), la hormona estimulante de melanocitos gamma (MSH gamma 1) y el péptido bisagra amidado (HP-N) o análogos funcionales de los mismos.

La expresión "**polipéptido aislado**" o "**polinucleótido aislado**", tal y como se emplea en esta memoria, se refiere a un polipéptido o a un polinucleótido que está aislado de su fuente natural.

Tal y como se emplea en este documento, los términos "**purificado**" o "**recuperado**" se refieren a la eliminación de contaminantes de una muestra. Por ejemplo, la enzima PAL de la invención se purifica mediante la eliminación de proteínas contaminantes y otros compuestos dentro de una solución o preparación. En algunos aspectos la enzima PAL de la invención se expresa utilizando bacterias y estas enzimas PAL recombinantes se purifican mediante la eliminación de otros componentes de la célula hospedadora y el porcentaje de PAL recombinante aumenta de este modo en la muestra.

La expresión "**polipéptido sustancialmente puro**" indica en esta memoria una preparación de polipéptido que contiene como máximo 20%, como máximo 10% o como máximo a 5% en peso de otro material polipeptídico con el que está asociado de forma natural o recombinante. Por lo tanto, se prefiere que el polipéptido sustancialmente puro sea al menos 80% puro, al menos 90% puro o preferiblemente al menos 95% puro en peso del material del polipéptido total presente en la preparación. Los polipéptidos de la presente invención están preferiblemente en forma sustancialmente pura, es decir, que la preparación de polipéptido está esencialmente exenta de otro material polipeptídico con el que está asociado de forma natural o recombinante. Esto se puede lograr, por ejemplo, mediante la preparación del polipéptido por métodos recombinantes bien conocidos o por métodos de purificación clásicos.

En un aspecto "**% de pureza**" se define como la cantidad de proteína de interés, dividida por la cantidad de proteína de interés + la cantidad de contaminantes de la célula hospedadora X 100. Se puede determinar por análisis de SDS-PAGE o separaciones por HPLC hasta cantidades determinadas.

En un aspecto, el polipéptido de la invención tiene una pureza de al menos 1%, por ejemplo al menos 5%, al menos 10%, al menos 20%, al menos 40%, al menos 60%, al menos 80%, al menos 90%.

El término "**recuperado**" o "**recuperación**" tal como se utiliza en el presente documento para el polipéptido de la invención, significa en una realización que el polipéptido/enzima no está asociado con niveles significativos (por ejemplo, como máximo 1%, como máximo 2%, como máximo 3%, como máximo 5%, como máximo 10% o como máximo 25%) de moléculas biológicas extrañas y no deseadas, contenidas en el sistema en el que se produce el polipéptido, por ejemplo, un cultivo celular. Un polipéptido recuperado se refiere a un polipéptido de la invención que ha pasado por una etapa de pureza debido a la intervención humana (ya sea automática, manual o ambas). Se

entiende que en la invención también se recuperan polipéptidos de la invención y polipéptidos aislados de la invención que están presentes en una composición. En otras palabras, el término "recuperado" o "aislado" no pretende excluir mezclas artificiales o sintéticas con otros compuestos o materiales.

5 Los términos "**proteína**", "**péptido**" y "**polipéptido**" se usan indistintamente en este documento. Cuando un péptido es una porción de una proteína, los expertos en la técnica entienden el uso del término en su contexto. Las expresiones "una enzima de la invención" y "un polipéptido de la invención" se usan indistintamente en el presente documento.

10 Por "**modificación**" de una secuencia de aminoácidos se entiende una sustitución, una delección y/o una adición (incluyendo inserciones) de uno o varios aminoácidos en la secuencia. En aspectos adicionales, también incluye sustituciones de una o varias cadenas laterales de aminoácidos.

15 La expresión "**vector de expresión**" se define en esta memoria como una molécula de ADN lineal o circular que comprende un polinucleótido que codifica un polipéptido de la presente invención y está ligado funcionalmente a nucleótidos adicionales que facilitan su expresión. El término "**plásmido**", "**vector de expresión**" y "**vector**" se usan indistintamente ya que plásmido es la forma de vector utilizada más comúnmente en la actualidad. Sin embargo, la invención se entiende que incluye otras formas de vectores de expresión de este tipo que sirven para funciones equivalentes. Tal como se usa en el presente documento, "**vector de expresión**" o "**vector**" se refiere a una estructura artificial de ADN que contiene una secuencia de ADN que está ligada funcionalmente a una secuencia de control adecuada, capaz de efectuar la expresión del ADN en un hospedador adecuado. Tales secuencias de control, por ejemplo, pueden incluir un promotor para efectuar la transcripción, una secuencia operadora opcional para controlar dicha transcripción, una secuencia que codifica sitios de unión a ribosoma adecuados del ARNm y secuencias que controlan la terminación de la transcripción y de la traducción. El vector, por ejemplo, puede ser un plásmido, un fago o simplemente un inserto genómico potencial. Una vez transformado en un hospedador adecuado, el vector, por ejemplo, se puede replicar y funcionar independientemente del genoma anfitrión, o se puede integrar, en algunos casos, en el propio genoma.

25 El grado de similitud entre dos secuencias de aminoácidos se describe por el parámetro "**identidad**" ("% de identidad"). La identidad en el contexto de secuencias de aminoácidos de la invención se puede determinar por cualquier técnica/programa adecuado, normalmente mediante un análisis de alineación de Needleman-Wunsch (véase Needleman y Wunsch, J. Mol. Biol. (1970) 48:443-453) utilizando la matriz de puntuación BLOSUM50 con una penalización por hueco inicial de -12 y una penalización por extensión de -2. El resultado de Needleman-Wunsch marcado como la "identidad más larga" (obtenido usando la opción -nobrief) se utiliza como el porcentaje de identidad, se puede calcular de la siguiente manera: $(\text{Residuos idénticos} \times 100) / (\text{Longitud de la alineación} - \text{Número total de huecos en la alineación})$. Debido a que la alineación de Needleman-Wunsch proporciona una medición general o global de la identidad entre dos secuencias, se debe reconocer que las secuencias diana que pueden ser porciones o subsecuencias de secuencias peptídicas más grandes, se pueden utilizar de una manera análoga para completar secuencias o, alternativamente, se pueden utilizar valores de alineación locales para evaluar las relaciones entre las subsecuencias, tal como se determina, por ejemplo, mediante una alineación de Smith-Waterman (J. Mol. Biol. (1981) 147:195-197), que se puede obtener a través de programas disponibles. Otros métodos de alineación local que pueden ser adecuados para el análisis de la identidad, incluyen programas que aplican algoritmos heurísticos locales de alineación, tales como los programas FastA y BLAST.

40 Cuando en esta memoria se emplea la expresión "**secuencia codificante**" significa una secuencia de nucleótidos, que especifica directamente la secuencia de aminoácidos de su producto proteico. Los límites de la secuencia codificante están determinados generalmente por un marco de lectura abierto que normalmente empieza con el codón de iniciación ATG o codones de iniciación alternativos tales como GTG y TTG, y termina con un codón de terminación tal como TAA, TAG y TGA. La secuencia codificante puede ser ADN, ADNc o una secuencia de nucleótidos recombinante o sintética. El término "**ADNc**" se define en esta memoria como una molécula de ADN que se puede preparar mediante transcripción inversa de una molécula de ARNm maduro, cortado y empalmado, obtenido a partir de una célula eucariota. El ADNc carece de secuencias de intrones que están normalmente presentes en el ADN genómico correspondiente. El primer transcrito de ARN inicial es un precursor del ARNm que se procesa a través de una serie de etapas antes de aparecer como ARNm cortado y empalmado maduro. Estas etapas incluyen la eliminación de secuencias de intrones mediante un proceso denominado corte y empalme. El ADNc obtenido a partir de ARNm carece, por lo tanto, de cualquier secuencia de intrón.

55 Las expresiones "**molécula de ácido nucleico codificante**", "**secuencia de ácido nucleico codificante**", "**secuencia de ADN codificante**" y "**ADN codificante**" se refieren al orden o a la secuencia de desoxirribonucleótidos a lo largo de una hebra de ácidos desoxirribonucleicos. El orden de estos desoxirribonucleótidos determina el orden de los aminoácidos a lo largo de la cadena proteica. Por tanto, la secuencia de ADN codifica la secuencia de aminoácidos de una proteína, por ejemplo, una enzima. La expresión "**estructura artificial de ácido nucleico**" tal como se utiliza en esta memoria, se refiere a una molécula de ácido nucleico, ya sea monocatenario o bicatenario, que se aísla a partir de un gen de origen natural o que se modifica para contener segmentos de ácidos nucleicos de una manera que de otro modo no existiría en la naturaleza, o que es sintética. La expresión estructura artificial de ácido nucleico es sinónima de la expresión "casete de expresión", cuando la estructura artificial de ácido nucleico contiene las secuencias de control requeridas para la expresión de una secuencia codificante de la presente invención.

- 5 La expresión "**ligado funcionalmente**" indica en esta memoria una configuración en la que una secuencia de control se coloca en una posición apropiada en relación con la secuencia codificante de la secuencia de polinucleótidos, de tal manera que la secuencia de control dirige la expresión de la secuencia codificante de un polipéptido. El término "**expresión**" incluye cualquier etapa implicada en la producción del polipéptido, incluyendo, por ejemplo, la transcripción, la modificación postranscripcional, la traducción, la modificación postraduccional y la secreción.
- 10 La expresión "**obtenido a partir de**" tal como se utiliza en esta memoria en conexión con obtener un polipéptido o un polinucleótido a partir de una fuente dada (es decir, un organismo biológico), significa que el polinucleótido (el polinucleótido que codifica el polipéptido) es idéntico o es una variante de una secuencia de polinucleótidos presente de forma natural en ese organismo fuente, independientemente de si la secuencia de polinucleótidos se ha insertado o si el polipéptido es producido por un organismo distinto del organismo fuente. "Obtenido a partir de" también significa "identificado a partir de" en el contexto de la invención. "Obtenido a partir de" en el contexto de la presente invención también significa la identificación de las enzimas de la invención a partir de bases de datos con secuencias de nucleótidos/proteínas bacterianas, es decir, mediante la realización de una búsqueda asistida por ordenador en una base de datos de proteínas, por ejemplo, Uniprot, trEMBL o Ref-SeqP.
- 15 Con la expresión de **tipo silvestre**, tal como se usa en el presente documento en el contexto de la presente invención, se entiende la forma (por ejemplo, un gen o una secuencia proteica) tal y como aparece en la naturaleza. También se incluyen proteínas codificadas por una secuencia de nucleótidos deducida mediante búsquedas en bases de datos que contienen datos de secuencias de nucleótidos/proteínas bacterianas. En una realización, la expresión de tipo silvestre incluye la secuencia peptídica sin péptido señal ni péptido líder.
- 20 Con el término "**maduro**", tal como se usa en el presente documento (enzima, polipéptido o secuencia de aminoácidos de la invención), se entiende la secuencia funcional mínima putativa de un polipéptido a la que no se han añadido extensiones de aminoácidos naturales o artificiales (por ejemplo, péptidos señal o parejas de fusión).
- 25 La expresión "**célula hospedadora**", tal como se usa en el presente documento, incluye cualquier tipo de célula que es susceptible de transformación, transfección, transducción y similares con una estructura artificial de ácido nucleico o un vector de expresión que comprende un polinucleótido de la presente invención.
- La expresión "**marcador de la purificación**" significa una secuencia peptídica fusionada con una enzima, ya sea en el extremo N-terminal o C-terminal de la enzima y se utiliza para la purificación de la enzima.
- 30 La expresión marcador "**TAP**" se refiere a un marcador de proteína termoestable alcalina, obtenido a partir de bacterias termófilas que, cuando se fusiona con una secuencia peptídica de una enzima, ya sea en el extremo N-terminal o C-terminal, se puede utilizar para la purificación de la enzima, como se da a conocer en los documentos de solicitudes de patentes internacionales publicadas con los números WO 2006/108826 y WO 2008/043847.
- 35 La expresión "**enzima de fusión**", "**proteína de fusión**" o "**enzima marcada**" significa una enzima que tiene una "**pareja de fusión**" fijada a cualquiera de los extremos C-terminal o N-terminal de la enzima. Un ejemplo de una pareja de fusión es un marcador proteico que puede aumentar el nivel de expresión, la solubilidad o la purificación de la proteína de fusión.
- Con la expresión "**enlazador**" se entiende una secuencia de aminoácidos que une la pareja de fusión, por ejemplo, un marcador de purificación y la enzima entre sí. La secuencia del enlazador, por ejemplo, puede comprender una secuencia que favorece un mejor plegamiento de la proteína diana y/o un sitio de escisión para la escisión del marcador de purificación.
- 40 Una "**estructura de hélice**" se caracteriza por tener una secuencia de aminoácidos que se traduce en una estructura en espiral, estabilizada por enlaces de hidrógeno intercatenarios.
- 45 "**% de solubilidad**" se define como la cantidad de proteína soluble procedente de un lisado de células hospedadoras dividida por la cantidad de proteína soluble + insoluble procedente de un lisado de células hospedadoras X 100. Se puede determinar mediante una comparación basada en análisis SDS-PAGE de las fracciones insolubles y solubles de lisados celulares.
- 50 En el presente contexto, la expresión "**enzima funcional**" se entiende que indica una proteína con una función similar a la enzima natural. La proteína puede ser estructuralmente similar a la enzima natural y se puede obtener a partir de la enzima natural mediante la adición de uno o varios aminoácidos a uno o a ambos extremos C-terminal y N-terminal de la enzima natural, la sustitución de uno o varios aminoácidos en uno o en una cantidad de sitios diferentes en la secuencia de aminoácidos natural, la delección de uno o varios aminoácidos en uno o en ambos extremos de la enzima natural o en uno o varios sitios en la secuencia de aminoácidos, o la inserción de uno o varios aminoácidos en uno o varios sitios de la secuencia de aminoácidos natural.
- 55 "**Aplicación**" significa una muestra que contiene la enzima que se carga en una columna de purificación. "**Flujo a través**" significa la parte de la aplicación que contiene proteínas de la célula hospedadora y contaminantes que no se unen a la columna de purificación. "**Pico principal**" se refiere al pico en un cromatograma de purificación que tiene la mayor intensidad UV y que contiene la proteína. "**mUA**" son las unidades de miliabsorbancia. "**UV 280**" es

la absorbancia a una longitud de onda de 280 nm a la que se van a absorber las proteínas, medida en unidades de miliabsorbancia. "IPTG" es isopropil- β -D-tiogalactopiranosido. **SDS-PAGE** es electroforesis en gel de dodecilsulfato sódico-poliacrilamida. **FPLC** (cromatografía líquida rápida de proteínas), es una forma de cromatografía líquida similar a la cromatografía líquida de alta resolución (HPLC) que se utiliza para separar o purificar proteínas a partir de mezclas complejas. **LC-MS** (cromatografía líquida - espectrometría de masas) es una técnica analítica que combina las capacidades de separación física de la cromatografía líquida (o HPLC) con las capacidades de análisis de masas de la espectrometría de masas.

Aminoácidos: En el presente contexto, las indicaciones de tres letras o de una letra de los aminoácidos se han empleado en su significado convencional, tal y como se indica en la tabla 1. Salvo que se indique explícitamente, los aminoácidos mencionados en esta memoria son L-aminoácidos. Además, los extremos izquierdo y derecho de una secuencia de aminoácidos de un péptido son, respectivamente, los extremos N-terminal y C-terminal, a menos que se especifique lo contrario.

Tabla 1: Abreviaturas de aminoácidos:

Aminoácido	Código de tres letras	Código de una letra
Glicina	Gly	G
Prolina	Pro	P
Alanina	Ala	A
Valina	Val	V
Leucina	Leu	L
Isoleucina	Ile	I
Metionina	Met	M
Cisteína	Cys	C
Fenilalanina	Phe	F
Tirosina	Tyr	Y
Triptófano	Trp	W
Histidina	His	H
Lisina	Lys	K
Arginina	Arg	R
Glutamina	Gln	Q
Asparagina	Asn	N
Ácido glutámico	Glu	E
Ácido aspártico	Asp	D
Serina	Ser	S
Treonina	Thr	T

Descripción detallada de la invención

La modificación enzimática de precursores de péptidos con una Gly en el extremo C-terminal a un péptido alfa-amidado por medio de una monooxigenasa alfa-amidante de peptidilo bifuncional (PAM), se encuentra en varios organismos eucariotas. Los sustratos naturales de PAM perteneciente a neurohormonas o neurotoxinas y las enzimas (PAM, PHM y PHL) que participan en la alfa amidación de péptidos se han descrito exclusivamente para organismos multicelulares (metazoos). En los organismos que pertenecen a las bacterias o Archea, hasta ahora no se ha demostrado que posean polipéptidos que puedan llevar a cabo las etapas enzimáticas necesarias en un péptido para convertir una Gly C-terminal en un resto alfa amida.

Las enzimas eucariotas PAL, PHM y PAM empleadas para la conversión industrial *in vitro* de precursores de péptidos extendidos con glicina C-terminal en péptidos alfa amidados, se proporcionan mediante expresión recombinante en sistemas de expresión de células de mamífero, debido a un rendimiento insuficiente de enzima activa cuando se expresan en otros sistemas de expresión, tales como, por ejemplo, *E. coli* en el que se expresan como proteínas insolubles que requieren un repliegamiento.

La expresión en sistemas de expresión de células de mamífero proporciona un rendimiento limitado en comparación con la expresión en sistemas de origen bacteriano, fúngico o de levadura. Además, la sostenibilidad de los procesos

de expresión en células de mamífero es limitada, el tiempo de cultivo es largo frecuentemente con un bajo rendimiento de la expresión por unidad de tiempo. La expresión de proteínas en *E. coli* generalmente da como resultado ciclos de fermentación más cortos y rendimientos por célula que son varios factores más grandes, en comparación con los sistemas de expresión de mamífero.

- 5 Para las enzimas eucariotas PAL, PHM o PAM es difícil diseñar un procedimiento económicamente sostenible para proporcionar grandes cantidades de enzima α -amidante para uso en la producción industrial de péptidos α -amidados recombinantes, debido a la necesidad de un sistema de expresión en células de mamífero para la producción de la enzima correctamente plegada y activa.

10 Las enzimas eucariotas PAL, PHM o PAM tienen muchos residuos Cys como base de su integridad estructural y cuando estas enzimas se expresan en el entorno citosólico reductor de células *E. coli*, no pueden formar enlaces disulfuro correctos como las enzimas naturales. Kolhekar A.S. et al. Biochemistry 2002, 41, 12384-12394 describen la caracterización del dominio catalítico de PAL de rata en términos de puentes disulfuro. Los dos puentes disulfuro en PAL están totalmente conservados entre dominios PAL caracterizados procedentes de eucariotas (p12388), lo que indica que los puentes disulfuro tienen importancia en la integridad estructural de la enzima. Según Kolhekar A.S. et al. Biochemistry 2002, 41, 12384-12394 parece que la reducción de los puentes disulfuro con B-mercaptoetanol da como resultado una reducción de la actividad enzimática, lo que apoya la importancia estructural de los puentes disulfuro en los dominios PAL eucariotas.

20 En el documento WO90/08194 se describe un intento de aumentar la actividad enzimática mediante el tratamiento de una enzima α -amidante expresada en *E. coli*, con un agente reductor, tal como ditiotreitól o 2-mercaptoetanol, en combinación con un agente desnaturalizante, y luego oxidando la proteína reducida, pero este método falla en la mejora de la actividad enzimática.

25 El documento EP2172550 describe un derivado de enzima α -amidante del extremo C-terminal recombinante que carece de la formación de al menos uno de los cinco enlaces disulfuro que se producen normalmente en una enzima α -amidante del extremo C-terminal (actividad PHM) obtenida a partir de *Xenopus laevis*, y un método de producción de dicho derivado de forma recombinante en *E. coli*.

La presente invención tal y como se describe en esta memoria, proporciona enzimas que son capaces de catalizar la conversión de una α -hidroxiglicina a una α -amida. Las enzimas se obtienen a partir de organismos procariontes y tienen diferentes propiedades fisicoquímicas y características estructurales que las descritas para las enzimas PAL eucariotas.

- 30 En particular, se proporciona una forma sostenible para producir enzimas α -amidantes, fundamentales para la economía de proceso de un procedimiento recombinante. La expresión de las nuevas enzimas de la invención en *E. coli* da lugar a rendimientos finales más altos, en comparación con las enzimas PAM/PAL de expresión en mamífero.

35 Una característica de las enzimas es que se pueden expresar con alto rendimiento como proteínas solubles en *E. coli* y son muy fáciles de manejar en el procesamiento aguas abajo, en contraste con sus homólogas eucariotas, ya que no requieren un replegamiento para obtener la actividad.

40 Otra característica importante de las enzimas es que se conserva su actividad a pesar de la fusión a parejas de fusión N-terminales. Un ejemplo de estas parejas de fusión son los marcadores TAP usados en esta invención, que permiten una purificación más fácil de las enzimas y una extracción más fácil de la enzima después de la etapa enzimática. Otro ejemplo es el marcador His6. Una cantidad de enzimas marcadas con TAP se clonaron y se expresaron en *E. coli*. Por lo tanto, las enzimas se pueden expresar con un alto rendimiento como proteínas de fusión marcadas con TAP en *E. coli* y son muy fáciles de manejar en el procesamiento aguas abajo, en contraste con sus homólogas eucariotas.

45 Por consiguiente, las enzimas no requieren potencialmente un replegamiento que lleva mucho tiempo y es caro. Se proporciona de este modo un método de producción con alto rendimiento de las enzimas en una célula hospedadora de no mamífero, ejemplificada por *E. coli*.

A continuación se describen diferentes aspectos de la invención.

50 En un aspecto la enzima PAL comprende el siguiente motivo 1: Xaa₁ Val Xaa₂ Asp Arg Xaa₃ Xaa₄ Xaa₅ Arg Xaa₆ Gin Xaa₇ Xaa₈ Xaa₉ Xaa₁₀ Xaa₁₁ Gly Xaa₁₂ Xaa₁₃ Xaa₁₄ Xaa₁₅ Xaa₁₆ Trp, en donde Xaa₁ y Xaa₇ pueden ser cualquier aminoácido natural excepto Cys y Xaa₂, Xaa₃, Xaa₄, Xaa₅, Xaa₆, Xaa₈, Xaa₉, Xaa₁₀, Xaa₁₁, Xaa₁₂, Xaa₁₃, Xaa₁₄, Xaa₁₅, Xaa₁₆ pueden ser cualquier aminoácido de origen natural.

En un aspecto Xaa₄ es Asn. En un aspecto Xaa₁₄ es Leu. En un aspecto Xaa₇ es Val o Ile. En un aspecto Xaa₈ es Phe o Leu. En un aspecto Xaa₉ es Asp o Ser.

55 En un aspecto de la invención, la enzima PAL comprende el siguiente motivo 2: Asp Gly Tyr Xaa₁₇ Asn Xaa₁₈ Arg Xaa₁₉ Xaa₂₀ Xaa₂₁ Phe Xaa₂₂ Xaa₂₃ Xaa₂₄ Gly Xaa₂₅ Xaa₂₆ Xaa₂₇ Xaa₂₈ Xaa₂₉ Xaa₃₀ Xaa₃₁ Xaa₃₂ Xaa₃₃ Gly Xaa₃₄ Xaa₃₅ Xaa₃₆ Gly Xaa₃₇ Phe, en donde Xaa₁₇ a Xaa₃₇ se seleccionan independientemente a partir de un aminoácido

de origen natural, siempre que Xaa₁₇ no sea Cys.

El motivo 1 se localiza en una región de sitio activo y contiene un residuo Arg en una posición similar a Arg706 en una región de sitio activo en PAL de rata (*Rattus norvegicus*). Los aminoácidos en la región que rodea el residuo Arg706 que está incluido en el motivo 1, muestran una identidad de secuencia limitada entre PAL de rata y las enzimas de la invención, pero la propia Arg706 está conservada en las enzimas caracterizadas de la invención (SEQ ID NO 1-4). Chufan et al. 2009 determinaron que la V_{máx} relativa del mutante PALcc R706A de rata era solo el 3% de la de PALcc de rata de tipo silvestre y muestran que este residuo está situado en el bolsillo del sitio activo, lo que confirma el papel esencial de Arg706 (pág. 969) para la catálisis enzimática de las enzimas PAL (Chufán EE, De M, Eipper BA, Mains RE, Amzel LM. Amidation of Bioactive Peptides: The Structure of the Lyase Domain of the Amidating Enzyme. Structure. 2009 15 de Jul; 17(7):965-73).

El motivo 2 está localizado en otra región del sitio activo, que contiene un residuo Tyr en una posición similar a Tyr654 en PAL de rata, que está altamente conservado y un residuo esencial para la catálisis enzimática como se muestra en estudios de mutagénesis y estructurales (De, M., Bell, J., Blackburn, N.J., Mains, R.E., y Eipper, B.A. (2006). Role for an essential tyrosine in peptide amidation. J. Biol. Chem. 281, 20873-20882).

En un aspecto, la invención aprovecha un polipéptido aislado capaz de catalizar la conversión de una α -hidroxiglicina a una α -amida, en donde el polipéptido se selecciona a partir del grupo que consiste en: un polipéptido que comprende o que consiste en una secuencia de aminoácidos que tiene al menos 70% o al menos 75%, tal como, por ejemplo, al menos 80%, al menos 85%, al menos 90%, al menos 95, al menos 96% o al menos 98%, o 100% de identidad con la secuencia de aminoácidos seleccionada a partir del grupo que consiste en (a) los aminoácidos 2-306 de SEQ ID NO: 1; (b) los aminoácidos 3-336 de SEQ ID NO: 2; (c) los aminoácidos 3-305 de SEQ ID NO: 3; (d) los aminoácidos 3-279 de SEQ ID NO: 4; (e) SEQ ID NO: 13; (f) SEQ ID NO: 15; (g) SEQ ID NO: 19; (h) SEQ ID NO: 20; (i) SEQ ID NO: 21; (j) SEQ ID NO: 22; y (j) SEQ ID NO: 23.

En un aspecto, las enzimas se obtienen a partir de un organismo procariota, por ejemplo, una bacteria.

En un aspecto de la invención, la enzima se obtiene a partir de una especie de *Burkholderia*, *alpha proteobacterium* *Methanosarcina*, *Sorangium*, *Mesorhizobium*, *Bradyrhizobium* o *Solibacter*.

En un aspecto de la invención, la enzima se obtiene a partir de una especie de *Erythrobacter*. En un aspecto, la enzima se obtiene a partir de una especie de *Exiguobacterium*. En un aspecto, la enzima se obtiene a partir de una especie de *Chthoniobacter*. En un aspecto, la enzima se obtiene a partir de una especie de *Planctomyces*.

En un aspecto de la invención, la enzima se obtiene a partir de una especie de *Erythrobacter*, *Exiguobacterium*, *Chthoniobacter* o *Planctomyces*. En un aspecto, la enzima de la invención se obtiene a partir de un organismo procariota seleccionado a partir del grupo que consiste en una especie de *Erythrobacter*, en una especie de *Exiguobacterium*, en una especie de *Chthoniobacter*, en una especie de *Planctomyces*, en una especie de *Burkholderia*, en una especie de *alfa Proteobacterium*, en una especie de *Methanosarcina*, en una especie de *Sorangium*, en una especie de *Salinispora*, en una especie de *Mesorhizobium*, en una especie de *Bradyrhizobium* y en una especie de *Solibacter*.

En un aspecto, la enzima PAL es una enzima de tipo silvestre. En un aspecto la enzima PAL es la secuencia madura sin un péptido señal.

En un aspecto, la enzima PAL de la invención comprende hasta 30 modificaciones en comparación con la enzima de tipo silvestre. En un aspecto, la enzima comprende 1-30 modificaciones en comparación con la enzima de tipo silvestre. En un aspecto, la enzima comprende 5-25 modificaciones en comparación con la enzima de tipo silvestre. En un aspecto, la enzima comprende 10-20 modificaciones en comparación con la enzima de tipo silvestre. En un aspecto, la enzima comprende 12-18 modificaciones en comparación con la enzima de tipo silvestre. En un aspecto el número de modificaciones como se describe en este documento, se compara con la secuencia madura de tipo silvestre.

En un aspecto, la enzima de la invención comprende hasta 30 modificaciones. En un aspecto, la enzima comprende 1-30 modificaciones. En un aspecto, la enzima comprende 5-25 modificaciones. En un aspecto, la enzima comprende 10-20 modificaciones. En un aspecto, la enzima comprende 12-18 modificaciones. Se entiende que las secuencias ejemplificadas en el presente documento, por ser procedentes de una fuente procariota, pueden ser la base para tales modificaciones (por ejemplo, SEQ ID NO: 1).

En una realización de la invención, la enzima se modifica de modo que al menos un residuo Cys está sustituido o se deletiona, por ejemplo, por otro aminoácido adecuado tal como Ala, Ser o Val.

En una realización, el polipéptido de la invención comprende una pareja de fusión.

Un ejemplo de estas parejas de fusión es un marcador TAP, por ejemplo, los marcadores TAP utilizados en esta invención, que permiten una purificación más fácil de las enzimas y una extracción más fácil de la enzima después de la etapa enzimática. Otro ejemplo es un marcador His, p. ej., el marcador His6.

- 5 En una realización, la enzima de acuerdo con la invención puede comprender un marcador de purificación. En una realización, la enzima de acuerdo con la invención comprende un marcador de purificación que comprende una proteína ribosómica altamente básica, obtenida a partir de bacterias termófilas, tal y como se describe en los documentos de solicitudes de patentes internacionales publicadas con los números WO 2006/108826 y WO 2008/043847.
- En una realización, el marcador permanece en la enzima después de la purificación.
- En una realización de la invención, el marcador de purificación son marcadores de afinidad inmovilizados sobre metal tales como His6 o His8, marcadores de transferasa de glutatión, marcadores recuperados con anticuerpos, tales como el marcador FLAG, el marcador HA, el marcador myc, biotina o estreptavidina.
- 10 En una realización, el marcador comprenderá una secuencia enlazadora que comprende un sitio de escisión para la escisión *in vitro* del marcador de purificación para proporcionar la enzima.
- En una realización de la invención, el polipéptido de la invención comprende la pareja de fusión SEQ ID NO: 7 o SEQ ID NO: 11.
- 15 El enlazador puede tener, por ejemplo, 1-30, 1-25, 1-20 o 1-15 de residuos de aminoácidos y, en un aspecto, el enlazador puede comprender residuos de aminoácidos tales como Leu, Pro y Ala, lo que incrementa la formación de hélice alfa u otras características que dan como resultado una rigidez estructural. En otro aspecto el enlazador comprenderá residuos Gly y Ser, que proporcionan flexibilidad al enlazador.
- En una realización, el enlazador se puede seleccionar entre el grupo a continuación:
- Ser Ser Ser Gly Gly Ser Gly Ser Glu Val Leu Phe Gln (SEQ ID NO: 8)
- 20 Ser Ser Ser Gly Ser Gly Glu Val Leu Phe Gln (SEQ ID NO: 9)
- SER SER SER GLY GLY SER GLY GLY SER GLY (SEQ ID NO: 10)
- Ser Ser Gly Gly Ser Gly Ser Glu Val Leu Phe Gln (SEQ ID NO: 12)
- Ser Ser Ser Gly Gly Ser Gly Ser Glu Thr Leu Phe Gln (SEQ ID NO: 14)
- SER SER SER GLY GLY SER GLY GLY SER (SEQ ID NO: 16)
- 25 El enlazador puede estar fijado a cualquier extremo C-terminal o N-terminal de la enzima. El sitio de escisión puede ser cualquier sitio de escisión que permite la escisión *in vitro* del marcador de purificación desde la enzima.
- En una realización de la invención, la enzima comprende SEQ ID No. 1 y es capaz de catalizar la conversión de una α -hidroxiglicina a una α -amida. En un aspecto de la invención, la enzima comprende SEQ ID No. 2 y es capaz de catalizar la conversión de una α -hidroxiglicina a una α -amida. En un aspecto de la invención, la enzima comprende
- 30 SEQ ID No. 3 y es capaz de catalizar la conversión de una α -hidroxiglicina a una α -amida. En un aspecto de la invención, la enzima comprende SEQ ID No. 4 y es capaz de catalizar la conversión de una α -hidroxiglicina a una α -amida. En un aspecto de la invención, la enzima comprende SEQ ID No. 13 y es capaz de catalizar la conversión de una α -hidroxiglicina a una α -amida. En un aspecto de la invención, la enzima comprende SEQ ID No. 15 y es capaz de catalizar la conversión de una α -hidroxiglicina a una α -amida.
- 35 En un aspecto de la invención, la enzima PAL comprende como máximo 2 residuos de cisteína. En un aspecto de la invención, la enzima PAL no comprende residuos de cisteína. En un aspecto de la invención, la enzima PAL comprende 1 residuo de cisteína.
- En un aspecto de la invención, la primera etapa del proceso de α -amidación se lleva a cabo sometiendo la proteína a una enzima con actividad PHM y permitiendo de este modo que el residuo Gly C-terminal se convierta en α -hidroxiglicina y luego permitiendo que la enzima PAL catalice la conversión de la α -hidroxiglicina a una α -amida.
- 40 La actividad PHM eucariota depende de ácido ascórbico y Cu^{2+} y la actividad PAL eucariota depende de Zn^{2+} y otros iones divalentes como cofactores. Una dependencia similar de Zn^{2+} para la actividad se observó en las enzimas de tipo PAL de la invención.
- En un aspecto de la invención, la α -amidación de un péptido diana se lleva a cabo por un procedimiento que comprende las etapas de: i) someter un péptido diana que tiene un residuo Gly C-terminal, a una enzima con actividad PHM con la adición de Cu^{2+} y/o ácido ascórbico, permitiendo de ese modo que el residuo Gly C-terminal del péptido se convierta a α -hidroxiglicina; y ii) permitir que una enzima PAL catalice la conversión de la α -hidroxiglicina a la α -amida de dicho péptido, en donde la reacción con dicha PHM y dicha enzima de la invención sobre dicho péptido se lleva a cabo en dos etapas distintas o simultáneamente.
- 45 En un aspecto de la invención, la α -amidación de un péptido diana se lleva a cabo mediante un procedimiento que
- 50 En un aspecto de la invención, la α -amidación de un péptido diana se lleva a cabo mediante un procedimiento que

- comprende las etapas de i) someter un péptido diana con un residuo Gly C-terminal a una enzima con actividad PHM en condiciones tales que la enzima PHM convierte el residuo Gly C-terminal en α -hidroxiglicina; y ii) utilizar una enzima PAL procedente de un organismo procarionta para catalizar la conversión de la α -hidroxiglicina a la α -amida de dicho péptido. En un aspecto, el péptido diana se somete a la enzima con actividad PHM en presencia de Cu^{2+} y/o ácido ascórbico. En un aspecto, el péptido diana se somete a una enzima con actividad PHM y a una enzima PAL al mismo tiempo. En un aspecto, el péptido diana se somete a una enzima con actividad PHM y la enzima PAL al mismo tiempo y en presencia de Cu^{2+} y/o ácido ascórbico.
- En un aspecto, la enzima PAL es capaz de convertir ácido α -hidroxihipúrico a benzamida. La actividad de la enzima se puede mostrar midiendo la conversión de ácido α -hidroxihipúrico a benzamida, como se describe en Katopodis AG et al., *Biochemistry*. 1990, 29(26):6115-6120 o como se describe en el Ensayo (I).
- La actividad biológica de ciertos péptidos se incrementa significativamente cuando los péptidos están α -amidados en el extremo C-terminal. Ejemplos de péptidos diana que se benefician de estar α -amidados en el extremo C-terminal son amilina, neuropéptido Y (NPY), péptido YY (PYY), PYY-3-36, polipéptido pancreático (PP), gastrina, calcitonina, péptido relacionado con calcitonina (CGRP), péptido liberador de gastrina, vasopresina, oxitocina, neuroquinina A, secretina, pancreastatina, pro-opiomelanocortina (POMC), hormona estimulante de melanocitos alfa (MSH alfa), hormona estimulante de melanocitos gamma (MSH gamma 1) y péptido bisagra amidado (HP-N).
- En una realización, el péptido diana se selecciona a partir del grupo que consiste en amilina, neuropéptido Y (NPY), péptido YY (PYY), PYY-3-36, polipéptido pancreático (PP), péptido similar a glucagón (GLP-1), gastrina, calcitonina, péptido relacionado con calcitonina (CGRP), péptido liberador de gastrina, vasopresina, oxitocina, neuroquinina A, secretina, pancreastatina, pro-opiomelanocortina (POMC), hormona estimulante de melanocitos alfa (MSH alfa), hormona estimulante de melanocitos gamma (MSH gamma 1) y péptido bisagra amidado (HP-N) o análogos funcionales de los mismos.
- En un aspecto, la diana es un precursor de amilina extendido con Gly C-terminal o análogos funcionales del mismo. En un aspecto, el péptido diana es un precursor de GLP-1 extendido con Gly C-terminal o análogos funcionales del mismo. En un aspecto, el péptido diana es un precursor de PYY extendido con Gly C-terminal o análogos funcionales del mismo.
- En un aspecto de la invención, la enzima se utiliza junto con una enzima con actividad PHM.
- En un aspecto, la enzima PAL se utiliza junto con una enzima con actividad PHM y en presencia de Cu^{2+} y/o ácido ascórbico.
- Cuando se prepara el péptido α -amidado con las enzimas de acuerdo con el método de la invención, se permite que las enzimas reaccionen con un precursor del péptido, cuyo precursor tiene un residuo Gly C-terminal. El residuo Gly se puede convertir en α -hidroxiglicina mediante un proceso enzimático en presencia de ácido ascórbico y/o Cu^{2+} . Un ejemplo de una enzima capaz de convertir Gly en α -hidroxiglicina es una enzima con actividad PHM. Las enzimas utilizadas en el método de acuerdo con la invención catalizan la conversión de la α -hidroxiglicina a la α -amida.
- En un aspecto, el péptido α -amidado se utiliza para la preparación de un medicamento, tal como, por ejemplo, para el tratamiento o la prevención de la obesidad, la hiperglucemia incluyendo la hiperglucemia inducida por estrés, la diabetes de tipo 2, la tolerancia alterada a la glucosa, la diabetes de tipo 1, quemaduras, heridas de cirugía, otras enfermedades o lesiones en donde es necesario un efecto anabólico en el tratamiento, infarto de miocardio, accidente cerebrovascular, cardiopatía coronaria, otros trastornos cardiovasculares, tratamiento de pacientes diabéticos y no diabéticos en estado crítico y polineuropatía.
- En otro aspecto, el péptido α -amidado se utiliza como medicamento para retrasar o prevenir la progresión de una enfermedad en la diabetes de tipo 2.
- En un aspecto de la invención, se proporciona el péptido α -amidado que es para emplear como medicamento para el tratamiento o la prevención de la obesidad, la hiperglucemia incluyendo la hiperglucemia inducida por estrés, la diabetes de tipo 2, la tolerancia alterada a la glucosa, la diabetes de tipo 1 y quemaduras, heridas de cirugía y otras enfermedades o lesiones en donde es necesario un efecto anabólico en el tratamiento, infarto de miocardio, accidente cerebrovascular, cardiopatía coronaria y otros trastornos cardiovasculares.
- En un aspecto, la invención comprende un método para producir un péptido α -amidado que comprende permitir que un péptido diana reaccione con una enzima PHM y una enzima PAL, y purificar el péptido α -amidado. Las enzimas utilizadas en el método de la invención se preparan empleando un ácido nucleico aislado que codifica la enzima de la invención. En un aspecto, se proporciona un ácido nucleico recombinante que comprende un promotor y dicho ácido nucleico aislado.
- En un aspecto, se proporciona dicho ácido nucleico recombinante, en donde el ácido nucleico codifica además un marcador de purificación.
- En un aspecto, se proporciona un vector que comprende dicho ácido nucleico recombinante. En un aspecto, se

proporciona una célula hospedadora que comprende dicho ácido nucleico recombinante. En un aspecto de la invención, el ácido nucleico recombinante está presente en el genoma de la célula hospedadora o en un vector que se replica de manera autónoma en la célula hospedadora.

5 En una realización, la célula hospedadora es de mamífero, tal como, por ejemplo, células CHO. En una realización, la célula hospedadora es una célula hospedadora de no mamífero. En un aspecto, una bacteria, un hongo, por ejemplo la levadura, se utiliza como célula hospedadora para la producción del polipéptido de la invención. Las células hospedadoras adecuadas son, por ejemplo, especies de *Escherichia*, por ejemplo *E. coli*, de *Bacillus*, de *Saccharomyces*, por ejemplo *Saccharomyces cerevisiae*, o de *Pichia*. En un aspecto, la célula hospedadora es *E. coli*.

10 La enzima con actividad liasa alfa-amidante de peptidil- α -hidroxiglicina (PAL) se puede producir por un método que comprende las etapas de: (i) cultivar una célula hospedadora de expresión recombinante de origen no mamífero, que comprende una estructura artificial de ácido nucleico que comprende una secuencia de nucleótidos que codifica una enzima de la invención, tal y como se describe en el presente documento, en condiciones adecuadas para la expresión de la enzima; y (ii) recuperar la enzima a partir (a) del material sobrenadante después de la rotura celular y centrifugación y/o (b) el medio de crecimiento; en donde la célula hospedadora no tiene origen de mamífero, por
15 ejemplo, una cepa de *Escherichia coli*, y en donde la enzima es soluble cuando se recupera en la etapa (ii). En una realización, la enzima está en una forma catalíticamente activa cuando se recupera en la etapa (ii) de modo que la enzima no requiere una etapa de replegamiento para obtener la actividad catalítica (actividad del tipo PAL, tal y como se describe en el presente documento). La enzima PAL para su uso en el método de la invención se puede producir, por ejemplo, en *E. coli* en donde no comprende una etapa de replegamiento del polipéptido. En un aspecto,
20 la enzima PAL que se produce en una célula hospedadora recombinante tal como se describe en el presente documento, por ejemplo, *E. coli*, es soluble al menos 50% o al menos 80%, es decir, sin la necesidad de replegamiento para obtener la actividad catalítica.

25 En un aspecto, la célula hospedadora es *E. coli*, que comprende un vector que comprende un ácido nucleico recombinante que codifica la enzima PAL, un marcador de purificación y un promotor, y la célula hospedadora se mantiene en condiciones adecuadas para la producción de dicha enzima.

En un aspecto, la enzima se purifica a partir del caldo de fermentación mediante el uso de una etapa de cromatografía de intercambio catiónico que da como resultado una pureza de aproximadamente el 80%, por ejemplo, de al menos el 80%.

30 En *E. coli* la mayoría de las proteínas abundantes se encuentra en una agrupación con un pI que oscila desde 4-7 y 8-10 en el que la mayoría de las proteínas, así como las proteínas más abundantes se encuentran en el intervalo de pI 4-7. Los marcadores TAP que se pueden usar para purificar las enzimas son altamente alcalinos y cuando se fusionan con una enzima aumentarán significativamente la carga global positiva y el pI de la enzima, de manera que se distingue claramente del grueso de los contaminantes de la célula hospedadora. Esto permitirá que la enzima se eluya a unas concentraciones salinas o con un pI tal, que los contaminantes de la célula hospedadora no serán capaces de unirse a una matriz de intercambio catiónico dada. En una realización, el marcador de purificación tiene un pI superior a aproximadamente 9, por ejemplo, superior a 9. En una realización, el marcador de purificación tiene un pI superior a aproximadamente 10, es decir, superior a 10. En una realización, el pI del marcador de purificación estará entre aproximadamente 9 y aproximadamente 12,5, es decir, en el intervalo de 9-12,5 y en un aspecto adicional, el pI es de aproximadamente 10, es decir, el pI es 10.

40 Cualquier matriz de intercambio catiónico adecuada se puede utilizar en el método de acuerdo con la invención y una lista no limitante de material adecuado para la columna de intercambio catiónico es: SP-Sefarosa XL Amersham n° de cat. 17-5073-01; Streamline SP XL Amersham n° de cat. 17-5076-01; Streamline Direct CST Amersham n° de cat. 17-5266-03; Obelix SP Amersham n° de cat. 11-0010-86; S-Support Unosphere, BioRad n° de cat. 156-0113; SP-Sefarosa High Performance Amersham n° de cat. 17-1087-03; Source30S Amersham n° de cat. 17-1273-02 y
45 Toyopearl SP650S Toso-Haas n° de cat. 08437.

Los marcadores TAP contribuirán de manera diferente a la carga global de una enzima específica en función del pI y la carga del marcador de purificación elegido. Por lo tanto, la purificación de una proteína diana específica se puede optimizar mediante la elección de un marcador de purificación que permita la elución de la proteína de fusión a una concentración salina o a un pH en el que coeluyan solo cantidades mínimas de los contaminantes de la célula hospedadora.
50

Los residuos de aminoácidos en el enlazador se pueden seleccionar a partir de residuos de aminoácidos de este tipo que proporcionarán una estructura menos flexible a la enzima marcada. De este modo, la interferencia entre la enzima y el marcador de purificación se puede minimizar. En una realización, el enlazador puede comprender elementos estructurales tales como una estructura de hélice alfa.

55 La enzima marcada expresada, producida por las células se puede recuperar desde el medio de cultivo por procedimientos convencionales que incluyen la separación de las células hospedadoras del medio, mediante centrifugación o filtración, liberando la proteína de fusión mediante rotura mecánica de la célula, tal como ultrasonidos o presión, precipitando los componentes proteínicos del material sobrenadante o filtrado por medio de una sal, p. ej.,

sulfato de amonio. Después de una etapa de captura mediante cromatografía de intercambio catiónico, la enzima marcada se puede eluir en un gradiente salino y se recogen las fracciones eluidas que contienen la proteína de fusión.

5 Un polipéptido expresado como una enzima sin marcar, se puede recuperar por medio de una purificación, tal como p. ej., cromatografía de intercambio aniónico, cromatografía de interacción hidrófoba y cromatografía de filtración en gel (véase, por ejemplo, Scopes, R., Protein Purification, Springer-Verlag, NY, 1982).

La pureza del polipéptido de la invención se puede evaluar, por ejemplo, mediante un análisis de geles PAGE teñidos con Coomassie, utilizando el programa informático de análisis de imágenes de gel o análisis de cromatogramas de HPLC con UV 214 nm.

10 Después de la primera etapa de purificación, el marcador de purificación se puede eliminar por escisión directamente con una enzima de procesamiento adecuada (por ejemplo, EK). Si la concentración de sal es demasiado alta, la enzima se puede desalar antes de la escisión. El sitio de escisión puede ser cualquier sitio de escisión que permita una escisión eficaz *in vitro* después de aislar la enzima purificada. El sitio de escisión de enterocinasa utilizado más comúnmente tiene la secuencia DDDDK, en donde la escisión se produce después de K. Otros sitios de escisión no
15 limitativos de enzimas procesadoras incluyen el sitio de escisión del Factor Xa, que es más comúnmente IEGR, en donde la escisión se produce después de R; el sitio de escisión de la trombina, que es más comúnmente LVPRG o LVPRGS en donde la escisión se produce después de R; el sitio de escisión de la proteasa del virus etc del tabaco (TEV), que es más comúnmente ENLYFQG/S, en donde la escisión se produce después de Q y el sitio de escisión de la proteasa 3C de HRV14, que es más comúnmente LEVLFQ/GP, en donde la escisión se produce después de
20 Q. Se hace referencia al documento de solicitud de patente WO 2006/108826.

Las etapas después de la escisión pueden incluir una purificación adicional en columna de intercambio catiónico como en la primera etapa. En tal situación, el marcador de purificación liberado por la enzima de procesamiento tendrá un *pI* extremadamente elevado, lo que conduce a una unión muy eficaz a la matriz de intercambio catiónico. La enzima escindida se puede recoger entonces en el flujo a través de la columna, mientras que el marcador de
25 purificación liberado por escisión y los contaminantes restantes muy cargados, procedentes de la línea celular de producción, quedarán retenidos en la columna de intercambio catiónico.

Las etapas de purificación después de la escisión también pueden comprender otros medios de purificación tales como cromatografía de intercambio aniónico, cromatografía de interacción hidrófoba y cromatografía de filtración en gel (véase, por ejemplo, Scopes, R., Protein Purification, Springer-Verlag, NY, 1982).

30 En un aspecto de la invención, la enzima se purifica hasta tener al menos aproximadamente 90 a 95% de homogeneidad (es decir, hasta 90 a 95% de homogeneidad), hasta al menos aproximadamente 98% de homogeneidad (es decir, hasta 98% de homogeneidad). La pureza se puede determinar por electroforesis en gel, por ejemplo, mediante un análisis de los aminoácidos u otros métodos basados en HPLC.

El ácido nucleico recombinante que codifica la enzima con o sin el marcador, también se puede preparar sintéticamente por métodos convencionales establecidos, por ejemplo, el método de fosoramidita descrito por Beaucage y
35 Caruthers, Tetrahedron Letters 22 (1981), 1859-1869, o el método descrito por Matthes et al., EMBO Journal 3 (1984), 801-805. De acuerdo con el método de la fosoramidita, los oligonucleótidos se sintetizan, por ejemplo, en un sintetizador automático de ADN, se purifican, se reasocian, se ligan y se clonan en vectores adecuados. Las secuencias de ADN que codifican la enzima, con o sin el marcador también se pueden preparar por reacción en cadena de la polimerasa, tal como corte y empalme mediante PCR de extensión por solapamiento, usando cebadores
40 específicos, por ejemplo, tal y como se describe en el documento US 4.683.202, Saiki et al., Science 239 (1988), 487 - 491, o Sambrook et al., supra.

Por otra parte, el ácido nucleico recombinante puede ser de origen mixto sintético y genómico, mixto sintético y ADNc o de origen mixto genómico y ADNc, preparado mediante ligación de fragmentos de origen sintético, genómico o de ADNc (según sea apropiado), los fragmentos correspondientes a varias partes de toda la estructura artificial
45 de ácido nucleico, de acuerdo con técnicas convencionales.

Las secuencias de ADN que codifican la enzima, normalmente se insertan en un vector recombinante que puede ser cualquier vector, que puede estar sometido convenientemente a procedimientos de ADN recombinante, y la elección del vector dependerá frecuentemente de la célula hospedadora en la que se va a introducir. Por tanto, el vector puede ser un vector de replicación autónoma, es decir, un vector que existe como una entidad extracromosómica cuya replicación es independiente de una replicación cromosómica, por ejemplo, un plásmido. Alternativamente, el vector puede ser uno que, cuando se introduce en una célula hospedadora, se integra en el genoma de la célula hospedadora y se replica junto con el o los cromosomas en los que se ha integrado.

El vector es preferiblemente un vector de expresión en el que la secuencia de ADN que codifica la enzima está ligada funcionalmente a segmentos adicionales requeridos para la transcripción del ADN. En general, el vector de expresión se obtiene a partir de ADN plasmídico o vírico, o puede contener elementos de ambos. La expresión "ligado funcionalmente" indica que los segmentos están dispuestos de modo que actúan conjuntamente para sus fines previstos, por ejemplo, la transcripción se inicia en un promotor y procede a través de la secuencia de ADN que codifica

el polipéptido de la invención, como, por ejemplo, una proteína de fusión.

Los vectores de expresión para uso en la expresión de la enzima comprenderán un promotor capaz de dirigir la transcripción de un gen clonado o ADNc. El promotor puede ser cualquier secuencia de ADN, que muestre actividad transcripcional en la célula hospedadora de elección y que se puede obtener a partir de genes que codifican proteínas homólogas o heterólogas para la célula hospedadora.

5

Ejemplos de promotores adecuados para dirigir la transcripción del ADN en células de mamífero son el promotor SV40 (Subramani et al., Mol. Cell. Biol. 1 (1981), 854 -864), el promotor MT-1 (gen de la metalotioneína) (Palmiter et al., Science 222 (1983), 809-814), el promotor CMV (Boshart et al., Cell 41:521-530, 1985) o el promotor tardío principal de adenovirus 2 (Kaufman y Sharp, Mol. Cell. Biol., 2:1304-1319, 1982).

10 Ejemplos de promotores adecuados para uso en células hospedadoras de levadura incluyen promotores procedentes de genes glicolíticos de la levadura (Hitzeman et al., J. Biol. Chem. 255 (1980), 12073 - 12080; Alber y Kawasaki, J. Mol. Appl. Gen. 1 (1982), 419 - 434) o genes de la alcohol deshidrogenasa (Young et al., in Genetic Engineering of Microorganisms for Chemicals (Hollaender et al., compiladores), Plenum Press, New York, 1982), o el promotor TPI1 (documento US 4.599.311) o el promotor ADH2-4c (Russell et al., Nature 304 (1983), 652 - 654).

15 Ejemplos de promotores adecuados para uso en células hospedadoras de hongos filamentosos son, por ejemplo, el promotor ADH3 (McKnight et al., The EMBO J. 4 (1985), 2093 - 2099) o el promotor tpiA. Ejemplos de otros promotores útiles son los obtenidos a partir del gen que codifica la amilasa TAKA de *A. oryzae*, la proteinasa aspártica de *Rhizomucor miehei*, la alfa-amilasa neutra de *A. niger*, la alfa-amilasa estable al ácido de *A. niger*, la glucoamilasa (gluA) de *A. niger* o *A. awamori*, la lipasa de *Rhizomucor miehei*, la proteasa alcalina de *A. oryzae*, la triosa fosfato isomerasa de *A. oryzae* o la acetamidasa de *A. nidulans*. Los preferidos son los promotores amilasa TAKA y gluA. Los promotores adecuados están mencionados, por ejemplo, en los documentos EP 238 023 y EP 383 779.

20

Ejemplos de promotores adecuados para uso en células hospedadoras bacterianas incluyen el promotor del gen de la amilasa maltogénica de *Bacillus stearothermophilus*, el gen de la alfa-amilasa de *Bacillus licheniformis*, el gen de la amilasa BAN de *Bacillus amyloliquefaciens*, el gen de la proteasa alcalina de *Bacillus subtilis* o el gen de la xilosidasa de *Bacillus pumilus* o los promotores del fago lambda P_R o P_L o los promotores utilizados para la expresión en *E. coli*, por ejemplo, lac, trp, phoA, araBAD, tac, bacteriófago T7 y cspA.

25

El vector también puede comprender un marcador seleccionable, por ejemplo, un producto génico que complementa un defecto en la célula hospedadora, tal como el gen que codifica la dihidrofolato reductasa (DHFR) o el gen TPI de *Schizosaccharomyces pombe* (descrito por P.R. Russell, Gene 40, 1985, págs. 125-130), o un gen marcador que confiere resistencia a un fármaco, por ejemplo, ampicilina, kanamicina, tetraciclina, cloranfenicol, neomicina, higromicina o metotrexato. Para los hongos filamentosos, los marcadores seleccionables incluyen amdS, pyrG, argB, niaD y sC.

30

Las secuencias de ADN que codifican la enzima, si es necesario, también pueden estar conectadas funcionalmente a un terminador adecuado, tal como el terminador de la hormona de crecimiento humana (Palmiter et al., Science 222, 1983, págs. 809-814) o los terminadores TPI1 (Alber y Kawasaki, J. Mol. Appl. Gen. 1, 1982, págs. 419-434) o ADH3 (McKnight et al., The EMBO J. 4, 1985, págs. 2093-2099). Los vectores de expresión también pueden contener un conjunto de sitios de corte y empalme de ARN localizados aguas abajo del promotor y aguas arriba del sitio de inserción de la secuencia de polipéptido en sí, como, por ejemplo, una proteína de fusión. Sitios de corte y empalme de ARN preferidos se pueden obtener a partir de genes de adenovirus y/o de inmunoglobulina. También está contenida en los vectores de expresión una señal de poliadenilación localizada aguas abajo del sitio de inserción. Las señales de poliadenilación particularmente preferidas incluyen la señal de poliadenilación temprana o tardía de SV40 (Kaufman y Sharp, *ibid.*), la señal de poliadenilación de la región 5 Elb de adenovirus, el terminador del gen de la hormona de crecimiento humana (DeNoto et al Nucl. Acids Res. 9:3719-3730, 1981). Los vectores de expresión también pueden incluir una secuencia líder vírica no codificante, tal como el líder 2 tripartito de adenovirus, situado entre el promotor y los sitios de corte y empalme de ARN; y secuencias potenciadoras, tales como el potenciador de SV40.

35

40

45

Para dirigir la enzima en la vía secretora de las células hospedadoras, una secuencia señal secretora (también conocida como una secuencia líder, secuencia prepro o secuencia pre) se puede proporcionar en el vector recombinante. La secuencia señal secretora se une a las secuencias de ADN que codifican la enzima en el marco de lectura correcto. Las secuencias señal secretoras están situadas comúnmente en 5' de la secuencia de ADN que codifica el péptido. La secuencia señal secretora puede ser aquella que está asociada normalmente con la enzima o que puede ser de un gen que codifica otra proteína secretada.

50

Para la secreción a partir de células de levadura, la secuencia señal secretora puede codificar cualquier péptido señal, que asegure una dirección eficaz de los polipéptidos expresados en la vía secretora de la célula. El péptido señal puede ser un péptido señal de origen natural, o una parte funcional del mismo, o puede ser un péptido sintético. Se ha encontrado que los péptidos señal adecuados son el péptido señal del factor alfa (véase el documento US 4.870.008), el péptido señal de la amilasa salivar de ratón (véase, O. Hagenbuchle et al., Nature 289, 1981, págs. 643-646), un péptido señal de carboxipeptidasa modificado (véase, L.A. Valls et al., Cell 48, 1987, págs. 887-897), el

55

péptido señal BAR1 de la levadura (véase, el documento WO 87/02670) o el péptido señal de la proteasa aspártica 3 de levadura (YAP3) (véase, M. Egel-Mitani et al., *Yeast* 6, 1990, págs. 127-137).

Para una secreción eficaz en la levadura, una secuencia que codifica un péptido líder también se puede insertar aguas abajo de la secuencia señal y aguas arriba de la secuencia de ADN que codifica los polipéptidos. La función del péptido líder es permitir que el péptido expresado sea dirigido desde el retículo endoplásmico al aparato de Golgi y además a una vesícula secretora para la secreción en el medio de cultivo (es decir, exportación de los polipéptidos a través de la pared celular o al menos a través de la membrana celular en el espacio periplásmico de la célula de levadura). El péptido líder puede ser el líder del factor alfa de levadura (cuyo uso se describe, por ejemplo, en los documentos US 4.546.082, US 4.870.008, EP 16 201, EP 123 294, EP 123 544 y EP 163 529). Alternativamente, el péptido líder puede ser un péptido líder sintético, es decir, un péptido líder que no se encuentra en la naturaleza. Los péptidos líder sintéticos se pueden construir, por ejemplo, tal y como se describe en los documentos WO 89/02463 o WO 92/11378.

Para el uso en hongos filamentosos, el péptido señal se puede obtener convenientemente a partir de un gen que codifica una amilasa o glucoamilasa de *Aspergillus* sp., un gen que codifica una lipasa o proteasa de *Rhizomucor miehei* o una lipasa de *Humicola lanuginosa*. El péptido señal se obtiene preferiblemente a partir de un gen que codifica la amilasa TAKA de *A. oryzae*, la alfa-amilasa neutra de *A. niger*, la amilasa estable en medio ácido de *A. niger* o la glucoamilasa de *A. niger*. Péptidos señal adecuados se describen, por ejemplo, en los documentos EP 238 023 y EP 215 594.

La célula hospedadora en la que se introduce el ADN que codifica la enzima puede ser cualquier célula que es capaz de producir la presente enzima e incluye bacterias, levaduras, hongos y células eucariotas superiores.

Ejemplos de células hospedadoras bacterianas que durante el cultivo son capaces de producir el polipéptido de la invención son bacterias gram positivas tales como cepas de *Bacillus*, tales como cepas de *B. Subtilis*, *B. licheniformis*, *B. lentus*, *B. brevis*, *B. stearothermophilus*, *B. alkalophilus*, *B. amyloliquefaciens*, *B. coagulans*, *B. circulans*, *B. lautus*, *B. megaterium* o *B. thuringiensis*, o cepas de *Streptomyces*, tales como *S. lividans* o *S. murinus*, o bacterias gram negativas tales como cepas de *Escherichia coli*. La transformación de las bacterias puede efectuarse mediante transformación de protoplastos o usando células competentes de una manera conocida per se (véase, Sambrook et al., supra).

Ejemplos de células de levadura adecuadas, incluyen células de *Saccharomyces* spp. o *Schizosac-charomyces* spp., En particular, cepas de *Saccharomyces cerevisiae* o *Saccharomyces kluyveri*. Los métodos para transformar células de levadura con ADN heterólogo y producir polipéptidos heterólogos a partir de las mismas, se describen, por ejemplo, en los documentos US 4.599.311, 4.931.373, 4.870.008, 5.037.743 y 4.845.075, todos ellos se incorporan en esta memoria como referencia. Las células transformadas se seleccionan a través de un fenotipo determinado por un marcador seleccionable, comúnmente la resistencia a fármacos o la capacidad de crecer en ausencia de un nutriente particular, por ejemplo, leucina. Un vector preferido para su uso en levaduras es el vector POT1 descrito en el documento US 4.931.373. Otros ejemplos de células de levadura adecuadas son cepas de *Kluyveromyces*, tales como *K. Lactis*, *Hansenula*, por ejemplo *H. polymorpha* o *Pichia*, por ejemplo *P. pastoris* (véase, Gleeson et al., *J. Gen. Microbiol.* 132, 1986, págs. 3459-3465; documento US 4.882.279).

Ejemplos de otras células fúngicas son células de hongos filamentosos, por ejemplo, *Aspergillus* spp., *Neurospora* spp., *Fusarium* spp. o *Trichoderma* spp., en particular cepas de *A. oryzae*, *A. nidulans* o *A. niger*. El uso de *Aspergillus* spp. para la expresión de proteínas se describe, por ejemplo, en los documentos EP 272 277, EP 238 023, EP 184 438. La transformación de *F. oxysporum* se puede realizar, por ejemplo, tal y como se describe en Malardier et al., 1989, *Gene* 78: 147-156. La transformación de *Trichoderma* spp. se puede realizar, por ejemplo, como se describe en el documento EP 244 234.

Cuando un hongo filamentoso se usa como célula hospedadora, se puede transformar con el ácido nucleico recombinante de la invención, de forma convenientemente la integración del ácido nucleico en el cromosoma hospedador para obtener una célula hospedadora recombinante. Esta integración se considera generalmente que es una ventaja ya que el ácido nucleico recombinante es más probable que se mantenga estable en la célula. La integración del ácido nucleico recombinante en el cromosoma hospedador se puede realizar de acuerdo con métodos convencionales, por ejemplo, mediante recombinación homóloga o heteróloga.

La célula hospedadora transformada o transfectada se cultiva a continuación en un medio nutritivo adecuado, en condiciones que permiten la expresión de la enzima, después de lo cual, la totalidad o parte de la enzima resultante se puede recuperar a partir del cultivo, por ejemplo, la enzima sin marcar. El medio usado para cultivar las células puede ser cualquier medio convencional que sea adecuado para el cultivo de las células hospedadoras, tal como medio mínimo o complejo que contiene suplementos apropiados. Los medios adecuados están disponibles en proveedores comerciales o se pueden preparar de acuerdo con recetas publicadas (por ejemplo, en catálogos de la "American Type Culture Collection").

Ejemplos

Todos los ejemplos de enzimas de la invención, identificadas mediante el uso de bases de datos y las caracteriza-

ciones siguientes, contenían un péptido señal previsto con alto grado de seguridad, usando la señal P (Henrik Nielsen et al., "Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites" Protein Engineering 10, 1-6 (1997)). Ya que el péptido señal no forma parte normalmente de la enzima funcional madura, éste se retiró de la secuencia original de acuerdo con la predicción. Las enzimas en los ejemplos 1-6 descritos a continuación, se diseñaron con una extensión de GlyPro N-terminal para facilitar la eliminación de la pareja de fusión TAP con una proteasa 3C de HRV14 que reconocía la secuencia GluValLeuPheGln / GlyPro.

Ejemplo 1

Clonación de proteínas de fusión con la proteína ribosómica L9 como marcador de purificación y dominio de tipo PAL de *Erythrobacter* sp. SD-21

Cuatro variantes de la secuencia del dominio de *Erythrobacter* de tipo PAL se clonaron, se expresaron y se evaluaron. Las variantes se prepararon para evaluar si las diferentes parejas de fusión, enlazadores o la extensión hacia el extremo N-terminal del dominio de tipo PAL, afectaban a la expresión y la solubilidad.

Proteína 1:

SEQ ID NO: 13 codifica un dominio de tipo PAL obtenido a partir de la bacteria *Erythrobacter* sp. SD-21. El peso molecular de la proteína se calculó en 33467,1 Da. Cuando se fusiona con un marcador de purificación de proteína alcalina termoestable (TAP) (SEQ ID NO: 7) a través de un enlazador flexible rico en Gly/Ser (SEQ ID NO: 10), las enzimas de fusión tendrán un peso molecular de 51231,7 Da y representan la proteína 1 tal y como se describe en la tabla 1.

Para obtener el plásmido A que codificaba la proteína 1, en un fragmento génico XhoI/BamHI que codificaba SEQ ID NO: 13 se optimizaron codones para la expresión en *E. coli* y se obtuvo como un fragmento génico sintetizado con sitios de clonación XhoI en el extremo 5' terminal y BamHI en el extremo 3' terminal (GeneScript). En el extremo 5' se introdujo un fragmento corto que codificaba la pareja enlazadora (SEQ ID NO: 10), directamente aguas arriba del fragmento de nucleótidos que codificaba el dominio de tipo PAL (SEQ ID NO: 13).

El fragmento XhoI/BamHI se ligó en pET11 (Novagen) un vector de expresión que ya codificaba un marcador de purificación comprendido en un fragmento NdeI/XhoI (Proteína ribosómica L9 de *T. maritima*, descrita en los documentos de solicitudes de patente internacional, publicados con los números WO 2006/108826 y WO 2008/043847 (SEQ ID NO 7) utilizando el Sistema de ligación rápida LigaFast® (Promega) por métodos descritos por el fabricante. Los productos de la ligación se utilizaron para transformar células de *E. coli* competentes para TOP10 (Invitrogen) y se incubaron durante una noche en medio LB (Luria-Bertani) en placas de agar con 150 microgramos por mililitro de ampicilina. El plásmido A que codificaba la proteína de fusión de interés se obtuvo a partir de clones positivos del cultivo, después de la propagación del plásmido en medio líquido y minipreparaciones convencionales de plásmidos. La secuencia de nucleótidos correcta del plásmido A se verificó mediante secuenciación del ADN con cebadores específicos de la secuencia del promotor/terminador de T7.

Proteína 2:

Otra variante del dominio de tipo PAL de *Erythrobacter* sp. SD-21 (plásmido B) se obtuvo mediante la ligación de un fragmento sintético XhoI/NcoI de ~110 pb (Geneart) que comprendía cambios menores en la región enlazadora y la parte N-terminal del dominio de tipo PAL de *Erythrobacter* en un sitio XhoI/NcoI en el plásmido A, dando lugar de este modo a la proteína 2. La proteína 2 comprende el mismo marcador de purificación SEQ ID NO 7 que el que se describe para la proteína 1, pero tiene un enlazador con un sitio de escisión para la proteasa 3C de HRV14 (ELTFQ) (SEQ ID NO: 14) y cambios en el extremo N-terminal del dominio de tipo PAL (SEQ ID NO: 15). El peso molecular teórico de la proteína de fusión es 52092,6 Da y 33823,5 Da para el dominio de tipo PAL liberado con la proteasa 3C de HRV14.

Proteína 3:

Otra variante del dominio de tipo PAL de *Erythrobacter* (plásmido C, ejemplificado en el mapa del vector en la Fig. 1) se obtuvo mediante la ligación de un fragmento sintético XhoI/NcoI (Geneart) que comprendía cambios menores en la región enlazadora y parte del extremo N-terminal del dominio de tipo PAL de *Erythrobacter* en un sitio XhoI/NcoI en el plásmido B, dando lugar de este modo a la proteína 3. La proteína 3 comprende el marcador de purificación de SEQ ID NO: 7, pero tiene un enlazador ligeramente modificado (SEQ ID NO: 8) que permite la escisión con la proteasa 3C de HRV14, en comparación con la proteína 2 y un dominio de tipo PAL de *Erythrobacter* truncado en 4 aminoácidos (SEQ ID NO: 1) en comparación con la proteína 2 y un aminoácido en comparación con la proteína 1 y que comprende un residuo Gly N-terminal (SEQ ID NO: 1). El peso molecular teórico de la proteína de fusión es 51591,1 Da y 33324 Da para el dominio de tipo PAL liberado por la proteasa 3C de HRV14.

Clonación de la proteína de fusión con el marcador de purificación His6 y el dominio de tipo PAL de *Erythrobacter* sp. SD-21

Proteína 4

Otra variante del dominio de tipo PAL de *Erythrobacter* sp. SD-21 SEQ ID NO: 1, se obtuvo mediante la escisión de un fragmento XhoI/BamHI del vector C y ligando este fragmento en un vector que comprendía la pareja de fusión que codificaba una secuencia consecutiva SEQ ID NO: 9 y 10 para obtener el plásmido D. La proteína de fusión codificada comprende un marcador de 6 histidinas N-terminal (SEQ ID NO: 9) y un enlazador con un sitio de escisión para la proteasa 3C de HRV14 (SEQ ID NO: 12) en frente del dominio de tipo PAL de *Erythrobacter* SEQ ID NO: 1 y tiene un peso molecular teórico de 35645,4 Da.

Clonación de los dominios de PAL eucariota para un análisis comparativo

Con el fin de comparar los nuevos dominios de tipo PAL procedentes de las bacterias descritas en la presente patente, con los dominios de PAL bien conocidos, los dominios de PAL de *Rattus norvegicus* (plásmido E que codifica la proteína 5 y plásmido F que codifica la proteína 6 que comprende diferencias en el enlazador) y de *Xenopus laevis* (plásmido G que codifica la proteína 7) se clonaron esencialmente como se ha descrito para la proteína 1, utilizando la misma pareja de fusión (SEQ ID NO: 7) para evaluar los perfiles de expresión para PAL eucariota típica que contiene dos puentes disulfuro, descrita anteriormente (por ejemplo, Stoffers, D.A et al.: Proc. Natl. Acad. Sci. U.S.A. 86:735-739 (1989) y Mizuno Ket al.: Biochem. Biophys. Res. Commun. 148:546-552 (1987)).

Se prepararon las siguientes estructuras artificiales (Tabla 1):

Nombre de la Proteína	Nº de plásmido	Especie	Pareja de fusión	Enlazador	Secuencia de tipo PAL
1	A	<i>Erythrobacter</i>	SEQ ID NO: 7	SEQ ID NO: 10	SEQ ID NO: 13
2	B	<i>Erythrobacter</i>	SEQ ID NO: 7	SEQ ID NO: 14	SEQ ID NO: 15
3	C	<i>Erythrobacter</i>	SEQ ID NO: 7	SEQ ID NO: 8	SEQ ID NO: 1
4	D	<i>Erythrobacter</i>	SEQ ID NO: 11	SEQ ID NO: 12	SEQ ID NO: 1
5	E	<i>Rattus norvegicus</i>	SEQ ID NO: 7	SEQ ID NO: 9	SEQ ID NO: 5
6	F	<i>Rattus norvegicus</i>	SEQ ID NO: 7	SEQ ID NO: 16	SEQ ID NO: 5
7	G	<i>Xenopus laevis</i>	SEQ ID NO: 7	SEQ ID NO: 9	SEQ ID NO: 6

Las secuencias de nucleótidos se describen a continuación:

Plásmido nº: A (*Erythrobacter* 1): SEQ ID NO: 26

ATGAAAGTGATTCTGCTGCGTGATGTGCCGAAAAATTGGCAAAAAAGGCGAAATCAAAGAAGTGA
 GCGACGCGTACGCGCGTAACCTATCTGATTCGCGTGGCTTTGCGAAAGAATATACCGAAGGCCCT
 GGAACGTGCGATTAAACACGAAAAAGAAATCGAAAAACGCAAAAAAGAACGCGAACGTGAAGAA
 AGCGAAAAAATCCTGAAAGAACTGAAAAACGTACCCATGTGGTGAAAGTGAAAGCGGGCGAAG
 GCGGCAAAATTTTTGGCGCGGTGACCGCGGCGACCGTGGCGGAAGAAATTAGCAAACCACCG
 GCCTGAAACTGGATAAACGCTGGTCAAACCTGGACAAACCGATTAAAGAACTGGGCGAATATAGC
 CTGGAAGTGAGCCTGCCGGGTGGCGTGAAAGATACCATTAATAATTCGCGTGGAACGCGAAGAAA
 GCTCGAGCGGTGGCAGCGGCGGTAGCGGTGAAGCCCCGCGGTTACCATCGATGAAAGCTGGC
 CGGATATCCGGAAAGCGCCGTGTTTGGCGAACCGACCGCCATTGATGTGGATAGCCATGGCCA
 TATCTTTGTGCTGCATCGTCCCGTCCGGAATGGACCCAGCCGTTCCGAGCGATCCGATTAGCG
 AACCGACCGTGTTTCATGTTCCCGCGAACGGCAAACTGCTGAGCAAATGGGGCGCGGGCGAACT
 GGTTATGCCGCATGGTCTGAGCATTGATGGTGATAACAAAGTGTGGATTACCGATGTGGCGCGT
 GAACAGTTCTGCGTTTTACCCATGAAGGTGCGGAAGAAGTGGTCTGGGCACCCGTGGCGAAA
 CCGGCCAGGATGAAAGCCATTTCCGCCGTCCGGCGGATGTTACCTTTGTGGGTGATCGTGTGCT
 GGTGGCCGATGGCTATCTGAACCGTCGATTATGGTGTGGATCGCGCGGGCAACTTCCTGGAA
 CAGTGGGGTAAAGAAGGTGAAGATGCGGGCGAATTAATCTGCCGCATGCGATCGCGGCCGATA
 GCGAACGTATTTATGTGGCGGATCGTGAACCGCGCGTGTTCAGGTGCTGAGCCTGGATGGTGA
 ACCGCTGGCCCGTGGCGCCAGGATGGCACCGGCCATCCGTATGCCGTGAAACCGATTGGCGAG
 CGGCTATGTTCTGGCGATTGAAGGCCGCGATCGCGGGTTCGCAATACCGCCATTGGCCGCATT
 TATCGTGCGGATGGTGGCCTGGAACGTGTGTTTGTGCGGGCGTTGAACCGCATAACCGGCACCA
 GCCTGGGCCATGATGTGGCGATTGGTCCGGATGGTAGCGGTATATGGTCGACAACAAAGCGAA
 TCGTGTATTAAATTTGATCTGAGCCGCGCCGGCGTTGAAGAAGCGGATGCGGAT

Plásmido nº: B: (*Erythrobacter* 2): SEQ ID NO: 27

ATGAAAGTGATTCTGCTGCGTGATGTGCCGAAAAATTGGCAAAAAAGGCGAAATCAAAGAAGTGA
 GCGACGGCTACGCGCGTAACTATCTGATTCGCGGTGGCTTTGCGAAAGAATATACCGAAGGCT
 GGAACGTGCGATTAACACGAAAAAGAAATCGAAAAACGCAAAAAAGAACGCGAACGTGAAGAA
 AGCGAAAAATCCTGAAAGAACTGAAAAACGTACCCATGTGGTAAAAGTAAAAGCGGGCGAAG
 GCGGCAAAATTTTTGGCGCGGTGACCGCGGCGACCGTGGCGGAAGAAATTAGCAAACCACCG
 GCCTGAAACTGGATAAACGCTGGTTCAAACCTGGACAAACCGATTAAGAAGTGGGCGAATATAGC
 CTGGAAGTGAGCCTGCCGGGTGGCGTAAAAGATAACCATTAATAATTCGCGTGAACGCGAAGAAA
 GCTCGAGCGGCGGAGCGGTAGCGAAACCCTGTTTCAGGCGCGTGAAGAAGCGCCCGGTTA
 CCATTGATGAAAGCTGGCCGATATTCCGAAAGCGCGGTGTTTGGCGAACCGACCGCGATTGA
 TGTGGATAGCCATGGCCATATTTTTGTGCTGCATCGCGCGGGCCGCAATGGACCCAGCCGTTT
 CCGAGCGATCCGATTAGCGAACCGACCGTGTATGTTTGGCGGCAACGGCAAACCTGCTGAGCA
 AATGGGGCGCGGGCGAAGTGGTATGCCGCATGGCCTGAGCATTGATGGCGATAACAAAGTGT
 GGATTACCGATGTGGCGCGCAACAGGTGCTGCGCTTTACCCATGAAGGCGCGGAAGAACTGGT
 GCTGGGCACCCGCGGCGAAACCGGCCAGGATGAAAGCCATTTTGGCCGCCCGGGCGGATGTGAC
 CTTTGTGGGCGATCGCGTGTGGTGGCGGATGGCTATCTGAACCGCCGATTATGGTGTGTTGAT
 CGCGCGGCAACTTTCTGGAACAGTGGGGCAAAGAAGCGCAAGATGCGGGCGAATTTAACCTG
 CCGCATGCGATTGCGGCGGATAGCGAACGCATTTATGTGGCGGATCGCGAAAAACGCGCGCTG
 CAGGTGCTGAGCCTGGATGGCGAACCGCTGGCGCGCTGGCGCCAGGATGGCACCGGCCATCCG
 TATGCGGTGAAACCGATTGGCAGCGGCTATGTGCTGGCGATTGAAGGCCGCGATCGCGCGGGC
 CGAACACCGCGATTGGCCGATTTATCGCGCGGATGGCGGCCCTGGAACGCGTGTGATGCGG
 GCGTGGAAACCGCATACCGGCACAGCCTGGGCCATGATGTGGCGATTGGCCCGGATGGCAGCG
 CGTATATGGTCGACAACAAAGCGAACCGCGTGATTAATAATTTGATCTGAGCCGCGCGGGCGTGA
 AGAAGCGGATGCGGAT

Plásmido nº: C (Erythrobacter 3): SEQ ID NO: 28

ATGAAAGTGATTCTGCTGCGTGATGTGCCGAAAAATTGGCAAAAAAGGCGAAATCAAAGAAGTGA
 GCGACGGCTACGCGCGTAACTATCTGATTCGCGGTGGCTTTGCGAAAGAATATACCGAAGGCT
 GGAACGTGCGATTAACACGAAAAAGAAATCGAAAAACGCAAAAAAGAACGCGAACGTGAAGAA
 AGCGAAAAATCCTGAAAGAACTGAAAAACGTACCCATGTGGTAAAAGTAAAAGCGGGCGAAG
 GCGGCAAAATTTTTGGCGCGGTGACCGCGGCGACCGTGGCGGAAGAAATTAGCAAACCACCG
 GCCTGAAACTGGATAAACGCTGGTTCAAACCTGGACAAACCGATTAAGAAGTGGGCGAATATAGC
 CTGGAAGTGAGCCTGCCGGGTGGCGTAAAAGATAACCATTAATAATTCGCGTGAACGCGAAGAAA
 GCTCGAGCGGCGGAGCGGTAGCGAAGTGTGTTTCAGGGCCCGCCGTTACCATGATGAAA
 GCTGGCCGGATATTCCGAAAGCGCGGTGTTTGGCGAACCGACCGCGATTGATGTGGATAGCCA
 TGCCATATCTTTGTGCTGCATCGTGCCGGTGCAGAAATGGACCCAGCCGTTTCCGAGCGATCCG
 ATTAGCGAACCGACCGTTCATGTTCCGCCGCAACGGCAAACCTGCTGAGCAAATGGGGCGCGG
 GCGAATGTTATGCCGATGGTCTGAGCATTGATGGTGAACAAAGTGTGGATTACCGATTG
 GCGCGTGAACAGGTTCTGCGTTTTACCCATGAAGGTGCGGAAGAAGTGGTTCTGGGCAACCGTG
 GCGAAACCGGCCAGGATGAAAGCCATTTCCGCCGTCGCGGATGTTACCTTTGTGGGTGATCG
 TGTGCTGGTGGCCGATGGCTATCTGAACCGTCTGATTATGGTGTGTTGATCGCGCGGGCAACTTC
 TGGAACAGTGGGGTAAAGAAGGTGAAGATGCGGGCGAATTTAATCTGCCGCATGCGATCGCGG
 CGATAGCGAACGTATTTATGTGGCGGATCGTAAAACGCGCGTGTTCAGGTGCTGAGCCTGGAT
 GGTGAACCGCTGGCCCGTGGCGCCAGGATGGCACCGGCCATCCGATGCCGTGAAACCGATT
 GGCAGCGGCTATGTTCTGGCGATTGAAGGCCGCGATCGCGCGGGTTCGCAATACCGCCATTGGCC
 GCATTTATCGTGGGATGTTGGCTGGAACGTGTGTTTGTGATGCGGGCGTTGAACCGCATACCGG
 CACCGCTGGCCATGATGTGGCGATTGTCCGGATGGTAGCGCGTATATGGTCCGACAACAAA
 GCGAATCGTGTATTAAATTTGATCTGAGCCGCGCCGGCGTTGAAGAAGCGGATGCGGAT

Plásmido nº: D (Erythrobacter 4): SEQ ID NO: 29

ATGGGCAGCAGCCATCATCATCACTCACTGAGCGGCGGAGCGGTAGCGAAGTGTGTTTC
 AGGGCCCGCCGTTACCATTTGATGAAAGCTGGCCGATATTCCGAAAGCGCGGTGTTTGGCGA
 ACCGACCGCGATTGATGTGGATAGCCATGGCCATATCTTTGTGCTGCATCGTGCCGGTTCGCAAT
 GGACCCAGCCGTTTCCGAGCGATCCGATTAGCGAACCGACCGTGTTCATGTTCCGCCGCAACGG
 CAAACTGCTGAGCAAATGGGGCGCGGGCGAAGTGGTTATGCCGCATGGTCTGAGCATTGATGGT
 GATAACAAAGTGTGGATTACCGATGTGGCGCGTGAACAGTTCTGCGTTTTACCCATGAAGGTGC
 GGAAGAAGTGGTTCTGGGCAACCGTGGCGAAACCGGCCAGGATGAAAGCCATTTCCGGCCGTTCC
 GGCGGATGTTACCTTTGTGGGTGATCGTGTGCTGGTGGCCGATGGCTATCTGAACCGTCTGATTA
 TGGTGTGTTGATCGCGCGGGCAACTTCTGGAACAGTGGGGTAAAGAAGGTGAAGATCGGGGCG
 AATTTAATCTGCCGATCGATCGCGCCGATGAGCGAACGATTTTATGTGGCGGATCGTGAAC
 GCGCGTGTTCAGGTGCTGAGCCTGGATGGTGAACCGTGGCCCGTGGCGCCAGGATGGCACC
 GGCCATCCGATGCCGTGAAACCGATTGGCAGCGGCTATGTTCTGGCGATTGAAGGCCGCGATC
 GCGCGGGTTCGCAATACCGCCATTGGCCGATTTATCGTGGCGGATGGTGGCCTGGAACGTGTGTT
 TGATGCGGGCGTTGAACCGCATACCGGCACAGCCTGGGCCATGATGTGGCGATTGGTCCGGAT
 GGTAGCGCGTATATGGTCGACAACAAAGCGAATCGTGTATTAAATTTGATCTGAGCCGCGCCGG
 CGTTGAAGAAGCGGATGCGGAT

Plásmido nº: E (rata 1): SEQ ID NO: 30

ATGAAAGTGATTCTGCTGCGTGATGTGCCGAAAAATTGGCAAAAAAGGCGAAATCAAAGAAGTGA
 GCGACGGCTACGCGCGTAACATCTGATTCGCGTGGCTTTGCGAAAGAATATACCGAAGGCCT
 GGAACGTGCGATTAAACACGAAAAAGAAATCGAAAAACGCAAAAAAGAACGCGAACGTGAAGAA
 AGCGAAAAATCCTGAAAGAACTGAAAAACGTACCCATGTGGTAAAAGTAAAAGCGGGCGAAG
 GCGGCAAAATTTTTGGCGCGGTGACCGCGGCGACCGTGGCGGAAGAAATTAGCAAACCACCG
 GCCTGAAACTGGATAAACGCTGGTTCAAACCTGGACAAACCGATTAAAGAAGTGGGCGAATATAGC
 CTGGAAGTGAGCCTGCCGGGTGGCGTGAAAGATACCATTAAAAATTCGCGTGGAACGCGAAGAAA
 GCTCGAGCGGGAGCGGCGAAGTTCTGTTTTAGGGCGATTTTCATGTGGAAGAGGAACTGGATTG
 GCCGGGTGTTTTATCTGCTGCCGGGGCAGGTGAGCGGCGTTGCCCTGGATAGTAAAAACAACCTG
 GTGATTTTCCACCGTGGAGATCATGTGTGGGATGGTAACTCATTGATAGCAAATTTGTCTATCA
 GCAGCGCGGCCTGGGTCTATTGAAGAAGATACCATCCTGGTTATCGATCCGAATAATGCGGAA
 ATTCTGCAAAGTTCTGAAAAAACCTGTTTACCTGCCTCATGGGCTGAGCATCGATACCGATGG
 TAACTACTGGGTTACCGATGTAGCCCTGCATCAGGTCTTCAAACCTGGACCCACATTCAAAAGAAG
 TCCCGCTGCTGATTCTGGGACGTTCTATGCAGCCTGGCAGCGATCAGAACCATTTTTGCCAACCA
 ACAGATGTCGCCGTGGAGCCGTCGACTGGTGTGTTTTGTGTCGGACGGCTACTGTAACCTCC
 GTATCGTCCAGTTTAGCCCGTCAGGTAAATTTGTTACACAGTGGGGCGAAGAGTCGTCTGGCTCT
 AGTCCGCGCCCGGGCCAGTTTTCGGTGCCGCATAGCCTGGCCCTGGTGCCGCATCTGGACCAGC
 TGTGTGTGGCCGACCGTGAGAACGGCCGATTCAATGCTTTAAACTGATACAAAGGAATTTGTT
 CGCGAAATTAACACGCCAGCTTTGGACGTAATTTTTCGCGATTTTCATACATTCCAGGCTTTCTG
 TTTGCAGTCAACGGCAAACCATATTTGGCGATCAGGAACCGGTACAGGGTTTTGTTATGAATTTT
 TCATCGGGCGAAATTATTGATGTTTTTAAACCTGTGCGCAAACATTTTCGATATGCCACATGATATT
 GTGGCTAGCGAGGACGGTACGGTTTATATTGGGGATGCGCACACGAACACCGTGTGGAATTCAC
 CCTGACCGAAAAAATGGAACACCGTTCGGTC

Plásmido nº: F (rata 2): SEQ ID NO: 31

ATGAAAGTGATTCTGCTGCGTGATGTGCCGAAAAATTGGCAAAAAAGGCGAAATCAAAGAAGTGA
 GCGACGGCTACGCGCGTAACATCTGATTCGCGTGGCTTTGCGAAAGAATATACCGAAGGCCT
 GGAACGTGCGATTAAACACGAAAAAGAAATCGAAAAACGCAAAAAAGAACGCGAACGTGAAGAA
 AGCGAAAAATCCTGAAAGAACTGAAAAACGTACCCATGTGGTAAAAGTAAAAGCGGGCGAAG
 GCGGCAAAATTTTTGGCGCGGTGACCGCGGCGACCGTGGCGGAAGAAATTAGCAAACCACCG
 GCCTGAAACTGGATAAACGCTGGTTCAAACCTGGACAAACCGATTAAAGAAGTGGGCGAATATAGC
 CTGGAAGTGAGCCTGCCGGGTGGCGTGAAAGATACCATTAAAAATTCGCGTGGAACGCGAAGAAA
 GCTCGAGCGGGGTAGCGGGCGCTCTGGCGATTTTCATGTGGAAGAGGAACTGGATTGGCCGG
 GTGTTTATCTGCTGCCGGGGCAGGTGAGCGGCGTTGCCCTGGATAGTAAAAACAACCTGGTGT
 TTTCCACCGTGGAGATCATGTGTGGGATGGTAACTCATTGATAGCAAATTTGTCTATCAGCAGC
 GCGGCTGGGTCCTATTGAAGAAGATACCATCCTGGTTATCGATCCGAATAATGCGGAAATTTCTG
 CAAAGTTCTGAAAAAACCTGTTTACCTGCCTCATGGGCTGAGCATCGATACCGATGGTAACTA
 CTGGGTTACCGATGTAGCCCTGCATCAGGTCTTCAAACCTGGACCCACATTCAAAAGAAGGTCCGC
 TGCTGATTCTGGGACGTTCTATGCAGCCTGGCAGCGATCAGAACCATTTTTGCCAACCAACAGAT
 GTCGCCGTGGAGCCGTCGACTGGTGTGTTTTGTGTCGGACGGCTACTGTAACCTCCCGTATCG
 TCCAGTTTAGCCCGTCAGGTAAATTTGTTACACAGTGGGGCGAAGAGTCGTCTGGCTTAGTCCG
 CGCCCGGGCCAGTTTTCGGTGCCGCATAGCCTGGCCCTGGTGCCGCATCTGGACCAGCTGTGTG
 TGCCCGACCGTGAGAACGGCCGATTCAATGCTTTAAACTGATACAAAGGAATTTGTTGCGGAA
 ATTAACACGCCAGCTTTGGACGTAATGTTTTGCGGATTTTCATACATTCCAGGCTTTCTGTTTGA
 GTCAACGGCAAACCATATTTGGCGATCAGGAACCGGTACAGGGTTTTGTTATGAATTTTTCATCG
 GCGGAAATTTATTGATGTTTTTAAACCTGTGCGCAAACATTTTCGATATGCCACATGATATTGGCT
 AGCGAGGACGGTACGGTTTATATTGGGGATGCGCACACGAACACCGTGTGGAATTCACCCTGA
 CCGAAAAAATGGAACACCGTTCGGTC

5 Plásmido nº: G (Xenopus): SEQ ID NO: 32

ATGAAAGTGATTCTGCTGCGTGATGTGCCGAAAATTGGCAAAAAAGGCGAAATCAAAGAAGTGA
 GCGACGGCTACGCGCTAACTATCTGATCCGCGTGGCTTTGCGAAAGAATATACCGAAGGCCT
 GGAACGTGCGATTAACACGAAAAAGAAATCGAAAAACGCAAAAAAGAACGCGAACGTGAAGAA
 AGCGAAAAATCCTGAAAGAACTGAAAAACGTACCCATGTGGTAAAAGTAAAAGCGGGCGAAG
 GCGGCAAAATTTTTGGCGCGGTGACCGCGGGCAGCCGTGGCGGAAGAAATTAGCAAAACCACCG
 GCCTGAAACTGGATAAACGCTGGTTCAAACCTGGACAAACCGATTAAGAAGTGGGCGAATATAGC
 CTGGAAGTGAGCCTGCCGGGTGGCGTGAAAGATACCATTAAAATTCGCGTGGAACGCGAAGAAA
 GCTCGAGCGGGAGCGGGCAAGTTCTGTTTCAGGGCGATGTGCATCTGGAAGAAGATACCGACTG
 GCCGGGTGTGAACCTTAAAGTGGGCCAGGTGAGCGGGCTAGCCCTGGACCCGAAAAACAACCT
 GGCCATTTTTCATCGTGCGCATCATGTGTGGGATGAAAACAGCTTTGATCGCAACTTTGTGTATC
 AGCAGCGTGGCATTGGCCCCGATTAGGAAAGCACCATTCTGGTTGTTGATCCGAGCAGCAGCAA
 AGTGCTGAAAAGCACCGGCAAAAACCTGTTTTTCTGCCGCATGGCCTGACCATTGATCGTGATG
 GCAACTATTGGGTGACCGATGTGGCGCTGCATCAGGTGTTTAAACTGGGTGCGGGCAAAGAAA
 CCCGCTGCTGGTGCTGGGTCGTGCGTTTCAGCCGGCAGCGATCGTAAACATTTTTGCCAGCCG
 ACCGATGTTGCGGTTGATCCGATTACCGCAACTTTTTGTGGCGGATGGCTATTGCAACAGCCG
 TATTATGCAGTTTAGCCCCGAAACGGCATGTTTATTATGCAGTGGGGCGAAGAAACAGCAGCAAG
 TGCCGCTCCGGGCCAGTTTCTGATTCCGCACAGCCTGACCATGGTGCCGATCAGGGCCAGCT
 GTGCGTGCGGATCGTAAAACGGCCGTATTAGTGTCTTTCATGCGGAAACCGGCAATTTTTGTG
 AAACAAATCAAACATCAGGAATTTGGCCGCGAAGTTTTGCGGTGAGCTATGCGCCGGGTGGTG
 TGCTGTATGCGGTGAACGGCAAACCGTATTATGGCTATAGCGCACCGGTGCAGGGCTTTATGCT
 GAACTTTAGCAACGGCGATATTCTGGATACCTTTATTCCGGCGCGTAAAAACTTTGATATGCCGCA
 TGATATTGCGGCAGCGGATGATGGCACCGTGTATGTGGGCGATGCGCATGCGAACGCGGTGTG
 GAAATTTAGCCCCGAGCAAAGCGGAACATCGGTCCGTG

Expresión de estructuras artificiales:

5 Los plásmidos con la secuencia de ADN correcta se transformaron en la cepa de expresión de *E. coli* BL21 (DE3), que se sembró en placas de LB/ampicilina durante una noche. Células hospedadoras BL21(DE3) con plásmidos se propagaron en medio LB y la expresión de proteínas fue inducida usando el sistema de expresión T7/IPTG.

10 Células *E. coli* BL21 (DE3) transformadas con plásmidos que codificaban estructuras artificiales de fusión se cultivaron hasta una densidad óptica (DO_{600nm}) de ~0,4 - 0,6 a 37°C en medio LB/ampicilina, usando matraces de agitación. Si la temperatura de inducción sometida a ensayo era 18°C o 30°C, la temperatura se reducía en consecuencia hasta 18°C o 30°C durante aproximadamente 30 min, y se añadió IPTG 0,5 mM al cultivo durante 3 horas. Después de la inducción de proteínas, los cultivos se sedimentaron y las células se lisaron mediante ultrasonidos en tampón NaPO₄ 25 mM pH 7. El análisis por SDS-PAGE de muestras de lisados se llevó a cabo en tampón de muestra que contenía células inducidas y no inducidas, así como fracciones solubles e insolubles de células inducidas (obtenidas por tratamiento con ultrasonidos y centrifugación en un tampón que consistía en NaPO₄ 25 mM pH 7). El análisis se realizó para evaluar el nivel de expresión y la solubilidad de la variante.

15 *Perfil de expresión de variantes de Erythrobacter y comparación con dominios PAL eucariotas bien conocidos de Rattus norvegicus y Xenopus laevis*

20 Cuando se expresaban desde los plásmidos en BL21(DE3) durante 3 horas a 30°C en matraces de agitación aforados de 1 L, las variantes de proteínas de *Erythrobacter*, la proteína 3, la proteína 1 o la proteína 2, mostraron pesos moleculares esperados de aproximadamente 50 kDa y niveles de expresión muy similares. Se observó un alto grado de solubilidad (~80%) a pH neutro (en un tampón que contenía NaPO₄ 25 mM pH 7, ejemplificado para la proteína 2 en la Fig. 2A).

25 Se observaron los mismos niveles de expresión para PAL de rata (proteína 6) o PAL de *Xenopus* (proteína 7) usando el mismo marcador N-terminal y el mismo enlazador que la proteína 1. Sin embargo, en contraste con los dominios de tipo PAL de *Erythrobacter*, los dominios PAL de rata y *Xenopus* daban como resultado consistentemente una proteína insoluble, cuando se utilizaban las mismas condiciones de expresión (ejemplificado para la proteína 6 en la Fig. 2B). PAL de rata con un enlazador diferente (proteína 5) no cambiaba el perfil de expresión, pero aun así daba como resultado una proteína insoluble utilizando las mismas condiciones de expresión. Una comparación de los perfiles de expresión (proteína 3 frente a la proteína 6), con unas condiciones de expresión de 18°C, 30°C o 37°C, mostraba que la proteína de *Erythrobacter* era una proteína soluble a las tres temperaturas, mientras que PAL de rata era insoluble a las tres temperaturas. Esto muestra la ventaja de expresar PAL procarionta, que no contiene puentes disulfuro, en comparación con PAL eucariota, que es más probable que forme agregados durante la expresión. El dominio de tipo PAL de *Erythrobacter* también se puede preparar como una proteína de fusión soluble, a pesar de las variaciones en la región de 5 aa del extremo N-terminal de la proteína PAL y en la región enlazadora.

35 La variante marcada con His de la proteína de *Erythrobacter* (proteína 4) dio lugar a una proteína soluble cuando se expresaba tanto a 30°C como a 37°C, lo que demuestra que la alta solubilidad de la proteína de *Erythrobacter* no depende de la naturaleza del marcador de fusión N-terminal utilizado, sino que es una propiedad inherente de la proteína.

Purificación de proteínas de fusión de tipo PAL de Erythrobacter marcadas SEQ ID NO: 7 sobre SP Sefarosa FF:

Después de la expresión de las enzimas de tipo PAL, (proteína 1, proteína 2 o proteína 3) o de PAL de rata (proteína 6) en matraces de agitación aforados de 1-2 L durante 3 horas a 30°C, como se ha descrito anteriormente, se realizó una purificación mediante cromatografía de intercambio catiónico utilizando columnas de SP Sefarosa FF de 5 ml tal y como se describe a continuación:

Sedimentos del cultivo celular (procedentes de 80 ml de cultivo con DO600: ~1,6-1,8) de las variantes de la enzima, se rompieron mediante ultrasonidos en un total de 20 ml de tampón NaPO₄ 25 mM pH 7. Los restos celulares se centrifugaron (4000 rpm, 15 min). El material sobrenadante se filtró de forma estéril (filtros de 0,45 µm) y se diluyó hasta un volumen total de 40 ml con tampón NaPO₄ 25 mM pH 7 para obtener la aplicación de proteínas. La purificación se llevó a cabo utilizando un sistema de purificación de explorer 100 AKTA (GE Healthcare). Una columna de SP Sefarosa FF HiTrap preempaquetada con un volumen de columna de 5 ml (GE Healthcare, producto n°: 17-5157-01) se utilizó para la separación con un caudal de 3 ml/min usando los siguientes tampones:

Tampón A: fosfato de sodio 50 mM, pH 7

Tampón B: fosfato de sodio 50 mM, pH 7 + NaCl 1 M

La columna se equilibró inicialmente para 7 volúmenes de columna de tampón A. Después de cargar la aplicación, la proteína no unida se eliminó por lavado usando 7 volúmenes de columna de tampón A. Un gradiente lineal de 0-100% de tampón B para 20 volúmenes de columna, se utilizó para eluir las enzimas de la columna.

La aplicación, las fracciones de flujo a través y las fracciones que representaban la proteína eluida dentro del gradiente, se separaron mediante SDS-PAGE y los geles se analizaron después tñiendo con azul brillante de Coomassie o análisis LC-MS.

La proteína 1, la proteína 2 o la proteína 3 eluyeron de la columna con una concentración de sal de NaCl ~0,3 M en un solo pico, lo que, según el SDS-PAGE, tenía una pureza elevada de ~80% (Fig. 3A y B ejemplificadas para la proteína 3). No se observó ninguna cantidad significativa de proteína en la fracción de flujo a través, lo que indica que todas las variantes de Erythrobacter fueron capturadas de manera eficaz en la columna de SP Sefarosa FF.

En contraste, los intentos de purificar las proteínas de fusión de PAL de rata (proteína 6) o de *X. laevis* (proteína 7), utilizando exactamente las mismas condiciones para la preparación de muestras y la purificación, fallaron lo que indica que PAL eucariota adopta una confirmación plegada incorrectamente después de la expresión en *E. coli*, que no permite fácilmente la purificación mediante cromatografía de cationes.

Las fracciones que contenían enzimas de tipo PAL de Erythrobacter eluidas, se agruparon y se concentraron y se desalaron en un tampón que contenía Tris 50 mM pH 7,5 utilizando columnas Vivaspin MWCO 10.000 (Vivaspin) de acuerdo con las instrucciones del fabricante. El glicerol se añadió a una conc. total de 10% y las enzimas se almacenaron hasta su uso a 20°C.

Purificación y análisis LC-MS del dominio de tipo PAL madura de Erythrobacter (SEQ ID NO: 1) sin marcador de purificaciónEscisión con proteasa 3C de HRV14 marcada con TAP

Para retirar el marcador de purificación de la proteína 3, ~1,5 mg de la proteína de fusión presente en tampón Tris-HCl 50 mM pH 7,5, 10% de glicerol, TCEP 1 mM se escindieron durante una noche a 30°C, utilizando una proteasa 3C de HRV14 marcada con TAP (procedente del documento WO 2008/043847) con una relación de enzima a sustrato de 1:25 en un volumen de reacción de 3 ml (las relaciones se estimaron después de la medición de la concentración de proteína usando NanoDrop2000, Thermo Scientific, de acuerdo con las instrucciones del fabricante). Los geles SDS-PAGE teñidos con Coomassie y el análisis LC-MS se emplearon para confirmar que la escisión enzimática se producía utilizando el siguiente protocolo: El instrumento LC-MSD_TOF (Agilent Technologies), se utilizó con la configuración de MS recomendada por el fabricante, para analizar las muestras usando una columna analítica de Poroshell 300SB-C8, Micro Bore 1,0 x 75 mm, 5 micras (Agilent Technologies) en condiciones de HPLC convencional, con un flujo de 0,3 ml/min y una temperatura de columna de 40°C: se formó una elución en gradiente en una ejecución de 20 min, usando formiato de amonio 8,8 mM en 0,1% de ácido fórmico y agua (Tampón A) y acetonitrilo (Tampón B) del modo siguiente:

Tiempo (min)	% de tampón B
0	22
3	22
15	75

Tiempo (min)	% de tampón B
15,1	90
20	90

El análisis LC-MS de las digestiones con proteasa 3C de HRV14 de la proteína 3 durante una noche, dio como resultado dos fragmentos determinados en 33323 Da y 18284,78 Da, correspondientes al dominio de tipo PAL liberado (SEQ ID NO 1, masa calculada 33323,97 Da) y al marcador de purificación liberado y el enlazador con sitio para 3C de HRV14 (SEQ ID NO: 7 + SEQ ID NO: 8, masa calculada: 18285,18 Da), respectivamente. No se observaron productos de degradación visibles ni en el control no escindido ni en el producto de la digestión mediante análisis con SDS-PAGE, indicando una estabilidad elevada tanto del dominio de tipo PAL maduro, como de la proteína de fusión.

Purificación de los dominios maduros sobre Q-Sefarosa HP

10 Se diluyeron 1:3 las digestiones que contenían dominios PAL maduros liberados con NaPO₄ 25 mM pH 7 y la aplicación de la muestra se cargó en 1 ml de una columna de intercambio aniónico de HiTrap Q Sefarosa de alto rendimiento (HP) (GE Healthcare 17-1153 -01) y se separaron usando el sistema Explorer 100 de AKTA (GE Healthcare) con los siguientes tampones de purificación:

Tampón A: fosfato de sodio 50 mM, pH 7

15 Tampón B: fosfato de sodio 50 mM, pH 7 + NaCl 1 M

La columna se equilibró inicialmente para 7 volúmenes de columna de tampón A. Después de cargar la aplicación, la proteína no unida se eliminó mediante lavado usando 7 volúmenes de columna de tampón A. Un gradiente lineal de 0-100% de tampón B para 20 volúmenes de columna, se utilizó para eluir las enzimas de la columna.

Se observó un único pico en el gradiente con una conc. de NaCl de -0,3 M

20 El SDS-PAGE de la aplicación, el flujo a través de las fracciones y las fracciones que incluían la proteína eluida mostraban que el fragmento liberado SEQ ID NO: 7 + SEQ ID NO: 8 que representaba el marcador de purificación y el enlazador con 3C de HRV14, estaba presente en el flujo a través, y el dominio de tipo PAL liberado (SEQ ID NO: 1) estaba presente en el pico principal único, eluido desde el gradiente. La pureza de la proteína eluida se estimó en ~90%, mostrando que el dominio de tipo PAL maduro se puede purificar con pureza elevada en solo dos etapas cromatográficas.

Las fracciones que contenían enzima madura eluida (SEQ ID NO: 1) se agruparon entre si y se concentraron y se desalaron en un tampón que contenía Tris 50 mM pH 7,5, 10% de glicerol, utilizando columnas Vivaspín MWCO 10.000 Da (Vivaspín) y se almacenaron a 20°C hasta su uso.

30 El Ejemplo 1 muestra que las enzimas PAL de la invención se pueden preparar como una proteína soluble incluso con variaciones en la región de 5 aa del extremo N-terminal de la proteína PAL.

Ejemplo 2

Clonación, expresión de la proteína de fusión con la proteína ribosómica L9 como marcador de purificación y el dominio de tipo PAL de Exiguobacterium sp. (cepa ATCC BAA-1283/AT1b) y purificación de la proteína de fusión y del dominio de tipo PAL maduro:

35 La proteína 8 es un dominio de tipo PAL de Exiguobacterium (SEQ ID NO: 2) con un marcador de purificación N-terminal (SEQ ID NO 7) y un enlazador que contiene un sitio de la proteasa 3C de HRV14 (SEQ ID NO 8). La proteína de fusión tiene un peso molecular calculado de 56115 Da y el dominio de tipo PAL maduro tiene un peso molecular calculado de 37848,5 Da

Se prepararon las siguientes estructuras artificiales:

Nombre de la proteína	Nº de plásmido	Especie	Pareja de fusión	Enlazador	Secuencia de tipo PAL
8	B	Exiguobacterium	SEQ ID NO: 7	SEQ ID NO: 8	SEQ ID NO: 2

40

Plásmido nº: H (Exiguobacterium): SEQ ID NO: 33

ATGAAAGTGATTCTGCTGCGTGATGTGCCGAAAATTGGCAAAAAAGGCGAAATCAAAGAAGTGA
 GCGACGGCTACGCGCTAACTATCTGATTCGCGTGGCTTTGCGAAAGAATATACCGAAGGCCCT
 GGAACGTGCGATTAACACGAAAAAGAAATCGAAAAACGCAAAAAAGAACGCGAACGTGAAGAA
 AGCGAAAAATCCTGAAAGAACTGAAAAACGTACCCATGTGGTAAAAGTAAAAGCGGGCGAAG
 GCGGCAAAATTTTTGGCGCGGTGACCGCGGCGACCGTGGCGGAAGAAATTAGCAAACCACCG
 GCCTGAAACTGGATAAACGCTGGTTCAAACGGACAAAACCGATTAAGAAGTGGGCGAATATAGC
 CTGGAAGTGAGCCTGCCGGGTGGCGTGAAAGATACCATTAAAAATTCGCGTGGAACGCGAAGAAA
 GCTCGAGCGGCGGAGCGGTAGCGAAGTGCTGTTTCAGGGTCCGCGTACCGATCCGATTTTTAA
 AGATGAATATGATGAAAAAGCCAAAAGCAGCCGTTATACCAGCAGCTGGGTTTGGCCTGAAAAAG
 ATAGCGTTAGCCATCGTGGTGGTGAAGGTAGCGGTGTTAGCACCAGCCCGAGCGGTTATGTTTA
 TTATCTGCATCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT
 TTTTGGATCCGAATACCAATGAAATTTGGATGAATTTGGCGATAACCTGTTTCAGTCTCCGCATG
 GTATTGAAGTTGATAGCCAGAATAATTTGGGTGACCGATATTAATGCTGAATAAAGTGTAAAC
 TGGATGAACGTGGTAATGTTCTGGCCACCTTTGGTGATGATTATCGTCTGGGCACCGAAACCAGC
 CTGCGTATTCGTAATGAACTGCCGAATTTCCGGTCCGATGAATGAATACACCTTTGCACGTCCG
 ACCGATGTTACCGTTATGGAAGATGGCAGCTTTATTGTTGCAGATGGCTATCGCAATCATCGTATT
 GTGAAATTTAATCGTGATGGCAATATTCAGTGGGAAGTTGATGCATATGGTAGCTCTGATGGCGA
 ATTTAATCTGCCGCATGGCATTACCCATGATCAGAGCGGCAATTTTATGTTGCCGATCGCAATAA
 TGCACGTATTCAGGTGTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT
 GCCGTCCCTTATGGAATCGATGCAGGCAATGATGGTAATTTTATCTGGTGGATGGTGGCGATTAT
 CTGAATGGTGAACGTGAAACCCCGAAAAGCCAGATTGTTGTTCTGAGCCCGAAAGGTGAAGTTAT
 TGAACGTTTTGGTAGCTGGGGTAACAAAATGGGTGAGCTGCGTATTCGCGATGATCTGACCGTGC
 TGAAGATGGTACAATTTTTGTTGCCGAACTGCTGAATGAACGTCTGCAGAAAATTACCATTACCG
 AA

5 El plásmido H que codificaba la proteína 8 se obtuvo ligando un fragmento XhoI/BamHI sintético que codificaba el enlazador y la región de tipo PAL en un pET11a que ya codificaba un marcador TAP (SEQ ID NO: 7) y realizando las etapas de verificación tal y como se describen en el ejemplo 1. La expresión a 30°C durante 3 horas y el análisis con SDS-PAGE se realizó como se ha descrito en el Ejemplo 1.

Los niveles de expresión de la proteína 8 eran muy similares a los de las variantes de Erythrobacter. La proteína era de nuevo muy soluble, sin detectarse casi ninguna proteína en la fracción insoluble del lisado celular.

10 La captura mediante intercambio catiónico, la escisión con 3C de HRV14 y la purificación del dominio maduro de tipo PAL de Exiguobacterium se realizó esencialmente como se ha descrito en el Ejemplo 1. Después de la captura en SP Sefarosa FF, un pico principal eluía a una conc. de NaCl de -0,25 M y basándose en el análisis SDS-PAGE, la captura era altamente eficaz, así como la pureza inicial de la proteína de fusión purificada ~80%.

15 La purificación de la proteína madura de tipo PAL de Exiguobacterium (SEQ ID NO: 2) después de la escisión con la proteasa 3C de HRV14 de la proteína de fusión reunida y concentrada (usando una relación de enzima a sustrato de 1:25 durante una noche), se realizó con la misma eficacia que como se ha descrito en el Ejemplo 1. El dominio maduro de tipo PAL eluyó de la columna de Q Sefarosa HP a una conc. de NaCl de -0,5 M (Fig. 4A). El análisis con SDS-PAGE mostró que dos fragmentos correspondientes al dominio de tipo PAL maduro de Exiguobacterium (SEQ ID NO: 2) (banda superior) y el marcador TAP liberado (SEQ ID NO: 7-SEQ ID NO: 8) (banda inferior) estaban presentes en la aplicación (Fig. 4B). Después de la separación mediante intercambio aniónico, el dominio maduro de tipo PAL de Exiguobacterium se encontraba en el pico principal (fracciones 12-15) que eluía del gradiente y el marcador liberado estaba presente en las fracciones de flujo a través (Fig. 4B).

20 Las fracciones que contenían enzimas eluidas se agruparon entre sí y se concentraron de forma activa y se desalaron en un tampón que contenía Tris 50 mM pH 7,5, usando columnas Vivaspin MWCO 10.000 Da (Vivaspin) y se almacenaron a 20°C hasta su uso.

Ejemplo 3

25 Clonación, expresión y purificación de la proteína de fusión con la proteína ribosómica L9 como marcador de purificación y el dominio de tipo PAL de Chthoniobacter flavus Ellin428

30 La proteína 9 es una Chthoniobacter flavus Ellin428 (SEQ ID NO: 3) con un marcador de purificación N-terminal (SEQ ID NO: 7) y un enlazador que contiene un sitio de la proteasa 3C de HRV14 (SEQ ID NO: 8). La proteína de fusión tiene un peso molecular calculado de 51650,4 Da y el dominio de tipo PAL maduro con Gly-Pro N-terminal tiene un peso molecular calculado de 33383,3 Da

Se prepararon las siguientes estructuras artificiales:

Nombre de la proteína	Nº de plásmido	Especie	Pareja de fusión	Enlazador	Secuencia de tipo PAL
9	I	Chthoniobacter	SEQ ID NO: 7	SEQ ID NO: 8	SEQ ID NO: 3

Plásmido n°: I (Chthoniobacter): SEQ ID NO: 34

ATGAAAGTGATTCTGCTGCGTGATGTGCCGAAAAATTGGCAAAAAAGGCCGAAATCAAAGAAGTGA
GCGACGGCTACGCGCTAACTATCTGATTCGCGTGGCTTTGCGAAAGAATATACCGAAGGCCT
GGAACGTGCGATTAAACACGAAAAAGAAATCGAAAAACGCAAAAAAGAACGCGAACGTGAAGAA
AGCGAAAAATCCTGAAAGAACTGAAAAACGTACCCATGTGGTGAAAGTGAAAGCGGGCGAAG
GCGGCAAAATTTTTGGCGCGGTGACCGCGGCGACCGTGGCGGAAGAAATTAGCAAACCACCG
GCCTGAAACTGGATAAACGCTGGTCAAACCTGGACAAACCGATTAAAGAACTGGGCGAATATAGC
CTGGAAGTGAGCCTGCCGGGTGGCGTGAAAGATACCATTAATAATTCGCGTGGAACGCGAAGAAA
GCTCGAGCGGCGGCGAGCGGTAGCGAAGTGCTGTTTCAGGGTCCGGAAGCCGTTATGAAGTTGT
TCCGCATTGGCCTGTTCTGCCGGAAGGTCGTAGCCTGGGTGTTTGTGCAGGTGTTGGTGTTGATA
GCCATGGTAATGTGTTTGTGTTTCATCGTAATGAACGTAATTGGACCGCAGCATTTCCGGAAGAA
CCGATTGCAGAACCGACCATTTCTGTTTTGATGGTCAGAGCGGCAAACTGCTGACCGAATGGGG
TGCAGGTGAATTTATTATGCCGCATGGTCTGACCCTGGATCGTGAAGATAATGTTTGGCTGACCG
ATGTTGGTCGTCAGCAGGTTTTTAATATGCCCATGATGGTCATCTGCTGCTGACCCTGGGTGAA
CGTGGTGTGCAGGTAGCGATCAGACCCATTTAATCTGCCGACCGATGTTGCAGTGCTGCCTGA
TGAAGCTTTTATGTGTCTGATGGCTATCGTAATACCCGTGTGGTTAAATTTGATGCCGCAGGCC
ATTATCAGTTTGAATGGGGTGGTAAAGGCACCGAACCGGGTAAATTTCTGCTGCCTCATGGAGTT
GCAGTTGATTCTCATGGTCGTGTTTTGTTTGGCATCGTACCAATAGCCGCTGCAGGTTTTTGAT
CCGAAAGGTAATTTCTGGCCGAATGAAAGGTCGCGAGGTTGGTCGTCGATGGTGTAGCG
TTGCAGCAAATGATCATGTGTTTGTGATTGATGGTGGTGATCAGCTGCCGAATCAGCCGGAACAT
GCAAAAGCAGTTGAACTGGACCCTGAAGGTAATGTTGTTCCGCGTTTTGGTAGCTATGGTCGTGA
TCCGGGTCAGTTTCACTGCTGGGTGATGATTTGCAGTTGCACCGGATGGTTCTGTTTATGTTGGTG
ATGCCAAAGGTAACGCGTTCAGAAATTTGTTCCGGTGCATCCG

5 El plásmido I que codificaba la proteína 9 se obtuvo ligando un fragmento sintético XhoI/BamHI que codificaba el enlazador y la región de tipo PAL en un pET11a que ya codificaba un marcador de purificación (SEQ ID NO 7) y se realizaron las etapas de verificación descritas en el Ejemplo 1. La expresión a 30°C durante 3 horas y el análisis por SDS-PAGE se realizó como se han descrito en el Ejemplo 1.

Los niveles de expresión de la proteína 9 eran menores que los de las variantes de tipo PAL de Erythrobacter y Exiguobacterium. Sin embargo, la proteína era predominantemente soluble a pH neutro.

10 Mediante una captura con intercambio catiónico en SP Sefarosa FF, se observó un pico principal a una conc. de NaCl de ~0,25 M. La captura sin embargo era menos eficaz que la observada para los dos dominios anteriores, aunque la pureza era comparativamente alta, basándose en el análisis con SDS-PAGE.

15 Las fracciones que contenían enzimas eluidas se agruparon y se concentraron de forma activa y se desalaron en un tampón que contenía Tris 50 mM pH 7,5 usando columnas Vivaspin MWCO 10.000 Da (Vivaspin) y se almacenaron a 20°C hasta su uso.

Ejemplo 4

Clonación, expresión y purificación de la proteína de fusión con la proteína ribosómica L9 como marcador de purificación y el dominio de tipo PAL de Planctomyces mares DSM 8797

20 La proteína 10 es un dominio de tipo PAL de Planctomyces DSM 8797 (SEQ ID NO: 4) con un marcador de purificación N-terminal (SEQ ID NO: 7) y un enlazador que contiene un sitio para la proteasa 3C de HRV14 (SEQ ID NO 8). Para el dominio de tipo PAL de Planctomyces, los dos residuos Cys que aparecían en la secuencia de tipo silvestre, se sustituyeron por un residuo de alanina y valina.

La proteína de fusión tiene un peso molecular calculado de 48678, 56 Da y el dominio de tipo PAL maduro con Gly-Pro N-terminal tiene un peso molecular calculado de 30411,4 Da.

25 Se prepararon las siguientes estructuras artificiales:

Nombre de la proteína	Nº de plásmido	Especie	Pareja de fusión	Enlazador	Secuencia de tipo PAL
10	J	Planctomyces	SEQ ID NO: 7	SEQ ID NO: 8	SEQ ID NO: 4

Plásmido n°: J Planctomyces: SEQ ID NO: 35

ATGAAAGTGATTCTGCTGCGTGATGTGCCGAAAATTGGCAAAAAAGGCGAAATCAAAGAAGTGA
 GCGACGGCTACGCGCGTAACTATCTGATTCCGCGTGGCTTTGCGAAAAGAATATACCGAAGGCCT
 GGAACGTGCGATTAACACGAAAAAGAAATCGAAAAACGAAAAAGAACGCGAACGTGAAGAA
 AGCGAAAAATCCTGAAAAGAACTGAAAAAACGTACCCATGTGGTGAAAAGTAAAAGCGGGCGAAG
 GCGGCAAAATTTTTGGCGCGGTGACCGCGGCGACCGTGGCGGAAGAAATTAGCAAAACCACCG
 GCCTGAAACTGGATAAACGCTGGTTCAAACCTGGACAAAACCGATTAAGAAGTGGGCGAATATAGC
 CTGGAAGTGAGCCTGCCGGGTGGCGTGAAAGATACCATTAATAATTCGCGTGGAACGCGAAGAAA
 GCTCGAGCGGTGGTAGCGGTAGCGAAGTTCTGTTTCAGGGTCCGGCAGATAAAAATTGATTTTGA
 ACCGGCAGCCATTAATATTGAACTGCCGGAAGGTCTGGCACTGGGTCTGCAAGCGCAGTTGAT
 TTTGATAGCAAAGGTGCGATGTACCTGTTTCATCGTGGTCCGCGAGCCGATTCTGGTTTTTGATCAG
 AGCGGTAAATTTGTTTCGTAGCTGGGGTGATAAACTGATTAGCCAGGCACATGGCTTAAGAGTTGC
 ACCGGATGAAACCATTTGGGTTACCGATATTGGCAACCACATGGTGTTCAGTTTAACCCGGAAG
 GTAAACTGCTGCTGGCCCTGGGTGAGGCAGGTAAACCGGGTGATAGCCAGGATCAGTTTAATAA
 ACCGACCGATATTGCATTTGGTCCGCGAGGGCGAATTCTATATTTCTGACGGCTATGGTAATAGCC
 GTGTGATGAAATTTGCAGCCAATGGTAAAAATCTGGGTGAGTGGGGTACACCCGGTAAAGGTCC
 GGGTGAATTTAATCTGCCGCATAGCATTCTGGTTGATGCAAAAGGTGCGTGTCTGGTTGGTGATC
 GTGAAAATGATCGCGTGCGATCTTTGATCTGGAAGGCAATCTGCTGGAAATTTGGACCGTTTT
 GCACCGTATGGTATGGAATTTGATAGCCGTGGTAACCTGTTTGTTCAGATGGTCTGCAATAA
 AGTTCTGCAGCTGAATGCAAGCGGTAAAGTTGAAAATAGCTGGGGTAAAACCGGCAAAGAACCG
 GGTGAGTATAACCTGCCTCACATGCTGGCAGTCGACGCAGCAGGCAACCTGTTTGTGACCGAAA
 TTGGTGGTAAACGTCTGCAGAACTGCAGCGCAA

El plásmido J que codificaba la proteína 10 se obtuvo ligando un fragmento sintético XhoI/BamHI que codificaba el enlazador y la región de tipo PAL en un pET11a que ya codificaba un marcador de purificación (SEQ ID NO 7) y se realizaron las etapas de verificación descritas en el Ejemplo 1. La expresión en BL21(DE3) dio como resultado una proteína con una masa molecular esperada, como se determinó por análisis LC-MS tal y como se ha descrito en el Ejemplo 1, y los niveles de expresión y las características de solubilidad elevada de la proteína 10 eran similares a las de las variantes de *Erythrobacter* descritas en el Ejemplo 1.

La captura por intercambio catiónico se realizó como se ha descrito en el Ejemplo 1. Después de la captura en SP Sefarosa FF, se observó un pico principal a una conc. de NaCl de ~0,25 M y basada en un análisis por SDS-PAGE. La captura era eficaz con cantidades muy bajas de proteína detectada en el flujo a través y la pureza de la proteína del pico principal se estimó en ~80%.

Las fracciones que contenían enzimas eluidas se agruparon y se concentraron de forma activa y se desalaron en un tampón que contenía Tris 50 mM pH 7,5 usando columnas Vivaspin MWCO 10.000 Da (Vivaspin) y se almacenaron a 20°C hasta su uso.

Ejemplo 5

Ensayo de la actividad enzimática para estimar la conversión de ácido α -hidroxi hipúrico y benzamida

Reacción enzimática:

Un ensayo enzimático para medir la actividad de PAL se ha descrito previamente (Katopodis AG et al.: Biochemistry. (1990) 29(26):6115-20), el cual mide la conversión de ácido α -hidroxi hipúrico a benzamida. Con el fin de someter a ensayo la capacidad de los dominios PAL para convertir ácido α -hidroxi hipúrico a benzamida, las reacciones enzimáticas se configuraron del modo siguiente:

Enzima (descrita en los ejemplos anteriores): ~ 0,2 mg/ml

ácido α -hidroxi hipúrico (Sigma): 1 mg/ml

Zn₂SO₄: 1 mM

Tampón: MES pH 5,5 o Tris pH 7,5: 100 mM

La reacción se incubó durante 3 horas a 37°C y los picos correspondientes a ácido α -hidroxi hipúrico y a benzamida se evaluaron después de la separación utilizando UPLC (Ultra Performance Liquid Chromatography) en lugar de la HPLC descrita anteriormente.

Condiciones de separación de la UPLC

Configuración del instrumento Waters ACQ-TUV:

Columna: UPLC BEH300 C18, 1,7 mm, columna de 2,1 x 150 mm de Waters (parte nº 186003687)

Disolvente A: 0,1% de TFA,

Disolvente B: 90% de MeCN, 0,1% de TFA (v/v)

Tabla del gradiente

Etapa	Tiempo (min)	Caudal	% de A	% de B	Curva
1	inicial	0,3	95	5	
2	5	0,3	85	15	6
3	5,50	0,3	10	90	6
4	6,50	0,3	10	90	6
5	7	0,3	95	5	6
6	8	0,3	95	55	6
7	20	0,020	95	5	11

El análisis por UPLC de los niveles de ácido α -hidroxi hipúrico sintético y benzamida que representaban α -hidroxiglicina y α -amida, respectivamente, después de la incubación con enzimas, indica si las enzimas pueden convertir una α -hidroxiglicina a un grupo α -amida. Un aumento en el área del pico de benzamida en comparación con el área del pico de ácido α -hidroxi hipúrico, en comparación con un control sin enzima, verifica que la enzima tiene actividad PAL. Basándose en el análisis semicuantitativo de las áreas por debajo de los picos que representaban el ácido α -hidroxi hipúrico y la benzamida, se verificó que las cuatro enzimas (proteína 3, proteína 8, proteína 9 y proteína 10) pueden catalizar la transformación de α -hidroxiglicina a α -amida. Los dominios de tipo PAL de Erythrobacter y Exiguobacterium (proteína 3 y proteína 8) son predominantemente activos a pH más alto (pH 7,5), mientras que los dominios de tipo PAL de Chthoniobacter y Planctomyces eran significativamente más activos a pH bajo (pH 5,5), lo que indica las diferencias en el pH óptimo para las enzimas (como se ejemplifica para los dominios de tipo PAL de Erythrobacter y Chthoniobacter mostrados en la Fig. 5). Un análisis de variantes maduras y marcadas con TAP de la proteína de tipo PAL de Erythrobacter (proteína 3) y de la proteína de tipo PAL de Exiguobacterium (proteína 8) muestra que las proteínas de tipo PAL marcadas con TAP son enzimas funcionales con actividades comparables a las enzimas maduras sin una pareja de fusión N-terminal (SEQ ID NO: 7) (Fig. 5).

Ejemplo 6:

Prueba de la actividad de dominios de tipo PAL sobre un sustrato análogo de amilina extendido con Gly C-terminal marcado con TAP, co-tratado con PHM de rata.

Con el fin de someter a ensayo si los dominios bacterianos de tipo PAL son adecuados para la α -amidación de péptidos recombinantes extendidos con Gly de forma C-terminal, las reacciones enzimáticas se prepararon con dominios bacterianos de tipo PAL y un dominio PHM de rata (*Rattus norvegicus*) bien conocido. Se utilizó el dominio PHM de rata (que comprendía la secuencia de aminoácidos 36 a 497 de la secuencia de PAM de rata de 976 aa de longitud completa) para facilitar la formación de una α -hidroxiglicina C-terminal que se purificó después de una expresión recombinante transitoria en células HEK293 esencialmente como se ha descrito en Husten EJ et al. (1993) *J. Biol. Chem.*; 268(13): 9709-17.

Se diseñó un sustrato peptídico modelo, relevante para las enzimas que comprendía una secuencia como se representa en SEQ ID NO: 17. El péptido consiste en un marcador TAP N-terminal que comprende la proteína ribosómica L27 de *T. maritima*, un enlazador intercalado con un sitio para enterocinasa y un análogo de amilina humana con una extensión Gly C-terminal (que contiene las siguientes sustituciones de aminoácidos: Val17His, Ala25Pro, Ser28Pro y Ser29Pro). En un fragmento génico que comprendía esta secuencia se optimizaron los codones para una expresión en *E. coli* y se ligó en un vector pET11a y se expresó a 30°C durante 3 horas en BL21(DE3). La purificación se realizó usando una columna de SP Sefarosa FF con tampones y configuraciones fundamentalmente como se describen en el Ejemplo 1. Las fracciones se agruparon y se concentraron de forma activa en un tampón que contenía Tris 100 mM pH 7,5 a una conc. de 0,3 mg/ml.

El péptido modelo se incubó con los dominios de tipo PAL y otros componentes pertinentes, tal y como se describe a continuación.

5 μ L de PHM-His de rata (0,15 mg/ml)

40 5 μ L de dominio de tipo PAL (proteína 3 (0,4 mg/ml) o proteína 10 (0,6 mg/ml))

10 μ L de Tris 100 mM pH 7,5

Ácido ascórbico 10 mM

ZnSO₄ 10 µM

CuSO₄ 10 µM

30 µg/ml de catalasa

30 µL de análogo de amilina marcado con TAP ((V17H) (0,3 mg/ml)

- 5 Las muestras se incubaron a 37°C y la reacción enzimática se detuvo con 3 µL de ácido acético al 100% y se analizaron después de 1 h, 2 h y 4 h y 5 h por LC-MS en columnas Poroshell C8 SB300 de 1 mm x 7,5 mm de fase inversa mediante LC-MSD-TOF como se ha descrito en el Ejemplo 1.

10 Las masas isotópicas promedio previstas para el precursor de la proteína modelo SEQ ID NO 17 se calcularon (la metionina iniciadora se elimina con metionina aminopeptidasa de *E. coli*, debido a una alanina en la segunda posición de SEQ ID NO: 17):

Forma precursora (que contiene un residuo Gly C-terminal): 13934,66 Da

La forma intermedia (que comprende una α-hidroxiglicina C-terminal): 13951 Da

Forma alfa amidada en C-terminal: 13876,62 Da

15 Después del tratamiento del péptido modelo con PHM de rata durante 2 horas, solo se observó que la gran mayoría del péptido modelo estaba en la forma intermedia que comprendía una α-hidroxiglicina en el extremo C-terminal, tal y como se determinó por análisis LC-MS (Fig. 6B). Una pequeña cantidad de producto amidado en alfa también apareció debido a una conversión espontánea, tal y como se observó para sustratos sintéticos obtenidos a partir de benzamida en el Ejemplo 5. Sin embargo, después de la adición del dominio de tipo PAL, se observó una conversión significativa en la proteína amidada en alfa, pero con diferencias en la actividad específica de las dos enzimas sometidas a ensayo. Con la proteína 10 se observó una conversión total del precursor después de una incubación durante 1 h-2 h, utilizando las condiciones de incubación descritas (Fig. 6C). En comparación, una conversión casi completa a la forma amidada en alfa requería 5 horas de incubación con la enzima de la proteína 3 (Fig. 6D y 6F). Todos estos datos juntos muestran que los dominios bacterianos de tipo PAL pueden catalizar la conversión de un precursor de la proteína recombinante con extensión de Gly C-terminal, a una proteína totalmente amidada en alfa C-terminal, en presencia de una enzima con actividad PHM.

Ejemplo 7:

Prueba de la actividad de dominios de tipo PAL sobre un sustrato de un análogo de amilina, pramlintida, extendido con Gly C-terminal marcado con TAP

30 Dominios de tipo PAL bacterianos se sometieron a ensayo en el análogo peptídico de amilina humana extendido con Gly C-terminal, denominado pramlintida (que contiene las siguientes sustituciones de aminoácidos: Ala25Pro, Ser28Pro y Ser29Pro) (SEQ ID NO: 25), que se obtuvo por síntesis de péptidos en fase sólida y se liofilizó después de la purificación. Las masas isotópicas promedio teóricas de las tres formas de Pramlintida son:

Pramlintida con extensión de Gly: 4007,5 Da

35 Pramlintida con extensión de alfa hidroxiglicina: 4023,5 Da

Pramlintida amidada en alfa: 3949,4 Da

SEQ ID NO: 25: KCNTATCATQRLANFLVHSSNFGPILPPTNVGSNTYG

Preparación de Pramlintida extendida en el extremo C-terminal con alfa-hidroxi glicina usando PHM de rata

40 Pramlintida sintética con extensión de Gly C-terminal se disolvió en Tris 100 mM Tris pH 7,5, hasta una concentración final de 1 mg/ml. La solución se incubó con el dominio PHM de rata (utilizado en el Ejemplo 6) utilizando una relación de enzima frente a sustrato de 1:20 (p/p), tal y como se describe en la tabla siguiente:

PHM	Pramlintida	CuSO ₄ (µM)	Catalasa µg/ml	Ácido ascórbico (mM)
0,2 mg	4 mg	5	20	10

45 La reacción se llevó a cabo a 37°C durante 2 h y luego se detuvo mediante la adición de TFA hasta una concentración final de 0,1%. Después de la reacción, se analizaron 20 µl de muestra procedente de la mezcla de reacción mediante LC-MS (esencialmente como se ha descrito en el ejemplo 6, excepto que se utilizó una columna de fase inversa C18) y se verificó que se obtenía la forma alfa-hidroxiglicina intermedia de pramlintida.

La muestra tratada con PHM de rata se purificó mediante HPLC preparativa de fase inversa C18 usando una columna Agilent Zorbac C18 Extend en un instrumento de HPLC Agilent 1100 con las condiciones y los gradientes de acetonitrilo que figuran en la tabla siguiente.

Tampón A: 95% de H₂O 5% de ACN 0,1% de TFA

- 5 Tampón B: 70% de ACN 20% de isopropanol 0,1% de TFA.

Tiempo (min)	% de B	Caudal (ml/min)
0	0	0,3
4	0	0,3
4,5	10	0,3
15	100	0,3
16	100	0,3
17	10	0,3
20	10	0,3

- 10 El producto intermedio de Pramlintida extendido con alfa-hidroxiglicina C-terminal se purificó. Después de la recogida del péptido y la agrupación de fracciones relevantes, la muestra se secó empleando una centrifuga Therm DNA120 Speed Vac y el péptido se volvió a disolver en 1 ml de tampón Tris 100 mM pH 7,5. La cantidad final de Pramlintida extendida con alfa-hidroxiglicina se midió por absorción de UV280 y la concentración se ajustó a 0,5 mg/ml de péptido.

Estimación de la actividad del dominio bacteriano de tipo PAL en el péptido intermedio de Pramlintida con alfa-hidroxiglicina

- 15 Con el fin de someter a ensayo la actividad de los dominios de tipo PAL (SEQ ID NO: 1, 2, 3 y 4), se realizaron las siguientes reacciones en tampón Tris 100 mM 7,5, empleando SEQ ID NO: 25 alfa-hidroxilada preparada como sustrato peptídico:

Dominio de tipo PAL (ug)	Pramlintida_Gly (OH) (ug y mg/ml)	ZnSO ₄ (uM)	Relación entre enzima y sustrato (p/p)
2	50 ug 0,5 mg/ml	0,5	25
1	50 ug 0,5 mg/ml	0,5	50
0,5	50 ug 0,5 mg/ml	0,5	100
0,1	25 0,5 mg/ml	0,5	250
0,05	25 0,5 mg/ml	0,5	500
0,02	25 0,5 mg/ml	0,5	1250

Las mezclas de reacción se incubaron a 37°C durante 1 h y se detuvieron mediante la adición de TFA al 0,1% antes del análisis LC-MS y después se analizó la muestra mediante LC/MS.

- 20 Se observaron dos picos principales en los espectros desconvolucionados que representaban el producto intermedio de alfa-hidroxi-glicina y Pramlintida totalmente alfa-amidada.

- 25 Los cuatro dominios de tipo PAL bacteriana podían convertir el producto intermedio de Pramlintida con alfa-hidroxi glicina a Pramlintida alfa-amidada. Los dominios de tipo PAL de Erythrobacter (Fig 7A) y Planctomyces (Fig 7B) eran significativamente más eficaces en la catálisis de la reacción, ya que podían convertir totalmente el sustrato del péptido a péptido alfa-amidado usando una relación entre enzima y sustrato de 1:100, en las condiciones del ensayo. PAL de Exiguobacterium (Fig 7C) y Chthoniobacter (Fig 7D) no eran capaces de convertir todo el sustrato de péptido en péptido alfa amidado, pero tal y como se observaba para las intensidades de pico relativas de 4023 (Pramlintida extendida con alfa-hidroxi glicina) y 3948 Da (Pramlintida alfa-amidada) se puede concluir que catalizaban la reacción, pero con una tasa más lenta que requería mayores concentraciones de enzima. La diferencia entre los dos conjuntos de dominios de tipo PAL puede ser debida a la diferencia observada en la preferencia de pH descrita en el ejemplo 5.

Los resultados proporcionan una evidencia de que los cuatro dominios bacterianos de tipo PAL tienen actividad liasa alfa-amidante de peptidil hidroxiglicina y pueden catalizar la misma reacción tal y como se ha descrito anteriormente

para PAL eucariota sobre un sustrato peptídico típico, que requiere una alfa amidación para obtener una actividad biológica.

Ensayos

Ensayo (I)

5 Evaluación de la actividad

La actividad de la enzima PAL se puede medir tal y como se ha descrito en el Ejemplo 5. La actividad de la enzima PAL se puede mostrar mediante una medición de la conversión de ácido α -hidroxihipúrico en benzamida, tal y como se describe en Katopodis AG et al., *Biochemistry*. 1990, 29(26):6115-6120. En lugar de utilizar el método de HPLC como se ha descrito, es ventajoso utilizar el método UPLC adaptado. Las condiciones para el método UPLC de describen en el ejemplo 5.

10

LISTADO DE SECUENCIAS

<110> Novo Nordisk A/S

<120> Enzima

<130> 8086.204-WO

15

<140> -

< 141> 2010-11-26

<160> 18

<170> PatentIn versión 3.5

20

<210> 1

< 211> 306

< 212> PRT

< 213> Erythrobacter sp. SD-21

ES 2 583 259 T3

<400> 1

Gly Pro Pro Val Thr Ile Asp Glu Ser Trp Pro Asp Ile Pro Glu Ser
1 5 10 15

Ala Val Phe Gly Glu Pro Thr Ala Ile Asp Val Asp Ser His Gly His
20 25 30

Ile Phe Val Leu His Arg Ala Gly Arg Glu Trp Thr Gln Pro Phe Pro
35 40 45

Ser Asp Pro Ile Ser Glu Pro Thr Val Phe Met Phe Ala Ala Asn Gly
50 55 60

Lys Leu Leu Ser Lys Trp Gly Ala Gly Glu Leu Val Met Pro His Gly
65 70 75 80

Leu Ser Ile Asp Gly Asp Asn Lys Val Trp Ile Thr Asp Val Ala Arg
85 90 95

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

Gly Thr Arg Gly Glu Thr Gly Gln Asp Glu Ser His Phe Gly Arg Pro
115 120 125

Ala Asp Val Thr Phe Val Gly Asp Arg Val Leu Val Ala Asp Gly Tyr
130 135 140

Leu Asn Arg Arg Ile Met Val Phe Asp Arg Ala Gly Asn Phe Leu Glu
145 150 155 160

Gln Trp Gly Lys Glu Gly Glu Asp Ala Gly Glu Phe Asn Leu Pro His
165 170 175

ES 2 583 259 T3

Ala Ile Ala Ala Asp Ser Glu Arg Ile Tyr Val Ala Asp Arg Glu Asn
 180 185 190

Ala Arg Val Gln Val Leu Ser Leu Asp Gly Glu Pro Leu Ala Arg Trp
 195 200 205

Arg Gln Asp Gly Thr Gly His Pro Tyr Ala Val Lys Pro Ile Gly Ser
 210 215 220

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

Ala Ile Gly Arg Ile Tyr Arg Ala Asp Gly Gly Leu Glu Arg Val Phe
 245 250 255

Asp Ala Gly Val Glu Pro His Thr Gly Thr Ser Leu Gly His Asp Val
 260 265 270

Ala Ile Gly Pro Asp Gly Ser Ala Tyr Met Val Asp Asn Lys Ala Asn
 275 280 285

Arg Val Ile Lys Phe Asp Leu Ser Arg Ala Gly Val Glu Glu Ala Asp
 290 295 300

Ala Asp
 305

<210> 2
 <211> 336
 <212> PRT
 <213> Exiguobacterium sp. (cepa ATCC BAA-1283 / AT1b)

5

<400> 2
 Gly Pro Arg Thr Asp Pro Ile Phe Lys Asp Glu Tyr Asp Glu Lys Ala
 1 5 10 15

Lys Ser Ser Arg Tyr Thr Ser Ser Trp Val Trp Pro Glu Lys Asp Ser
 20 25 30

Val Ser His Arg Gly Gly Glu Gly Ser Gly Val Ser Thr Ser Pro Ser
 35 40 45

Gly Tyr Val Tyr Tyr Leu His Arg Gly Asp Gly Ser Tyr Ala Asn Glu
 50 55 60

Glu Leu Ile Thr Thr Pro Thr Ile Thr Val Phe Asp Pro Asn Thr Asn
 65 70 75 80

Glu Ile Val Asp Glu Phe Gly Asp Asn Leu Phe Gln Ser Pro His Gly

ES 2 583 259 T3

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

<210> 3
 <211> 305
 <212> PRT
 <213> Chtonibacter flavus Ellin428

ES 2 583 259 T3

<400> 3

Gly Pro Glu Ser Arg Tyr Glu Val Val Pro His Trp Pro Val Leu Pro
 1 5 10 15

Glu Gly Arg Ser Leu Gly Val Cys Ala Gly Val Gly Val Asp Ser His
 20 25 30

Gly Asn Val Phe Val Phe His Arg Asn Glu Arg Asn Trp Thr Ala Ala
 35 40 45

Phe Pro Glu Glu Pro Ile Ala Glu Pro Thr Ile Ser Val Phe Asp Gly
 50 55 60

Gln Ser Gly Lys Leu Leu Thr Glu Trp Gly Ala Gly Glu Phe Ile Met
 65 70 75 80

Pro His Gly Leu Thr Leu Asp Arg Glu Asp Asn Val Trp Leu Thr Asp
 85 90 95

Val Gly Arg Gln Gln Val Phe Lys Tyr Ala His Asp Gly His Leu Leu
 100 105 110

Leu Thr Leu Gly Glu Arg Gly Val Ala Gly Ser Asp Gln Thr His Phe
 115 120 125

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

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

His Tyr Gln Phe Glu Trp Gly Gly Lys Gly Thr Glu Pro Gly Lys Phe
 165 170 175

Arg Leu Pro His Gly Val Ala Val Asp Ser His Gly Arg Val Phe Val
 180 185 190

Cys Asp Arg Thr Asn Ser Arg Leu Gln Val Phe Asp Pro Lys Gly Lys
 195 200 205

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

Ser Val Ala Ala Asn Asp His Val Phe Val Ile Asp Gly Gly Asp Gln
 225 230 235 240

ES 2 583 259 T3

Leu Pro Asn Gln Pro Glu His Ala Lys Ala Val Glu Leu Asp Pro Glu
245 250 255

Gly Asn Val Val Pro Arg Phe Gly Ser Tyr Gly Arg Asp Pro Gly Gln
260 265 270

Phe Gln Leu Gly His Asp Ile Ala Val Ala Pro Asp Gly Ser Val Tyr
275 280 285

Val Gly Asp Ala Lys Gly Lys Arg Val Gln Lys Phe Val Pro Val His
290 295 300

Pro
305

<210> 4
<211> 279
<212> PRT
<213> Planctomyces maris DSM 8797

5

<400> 4
Gly Pro Ala Asp Lys Ile Asp Phe Glu Pro Ala Ala Ile Asn Ile Glu
1 5 10 15

Leu Pro Glu Gly Leu Ala Leu Gly Pro Ala Ser Ala Val Asp Phe Asp
20 25 30

Ser Lys Gly Arg Met Tyr Leu Phe His Arg Gly Pro Gln Pro Ile Leu
35 40 45

Val Phe Asp Gln Ser Gly Lys Phe Val Arg Ser Trp Gly Asp Lys Leu
50 55 60

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

Val Thr Asp Ile Gly Asn His Met Val Phe Gln Phe Asn Pro Glu Gly
85 90 95

Lys Leu Leu Leu Ala Leu Gly Gln Ala Gly Lys Pro Gly Asp Ser Gln
100 105 110

Asp Gln Phe Asn Lys Pro Thr Asp Ile Ala Phe Gly Pro Gln Gly Glu
115 120 125

Phe Tyr Ile Ser Asp Gly Tyr Gly Asn Ser Arg Val Met Lys Phe Ala
130 135 140

Ala Asn Gly Lys Asn Leu Gly Gln Trp Gly Thr Pro Gly Lys Gly Pro
145 150 155 160

ES 2 583 259 T3

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

Val Leu Val Gly Asp Arg Glu Asn Asp Arg Val Gln Ile Phe Asp Leu
 180 185 190

Glu Gly Asn Leu Leu Glu Ile Trp Thr Gly Phe Ala Pro Tyr Gly Met
 195 200 205

Glu Phe Asp Ser Arg Gly Asn Leu Phe Val Ala Asp Gly Arg Ala Asn
 210 215 220

Lys Val Leu Gln Leu Asn Ala Ser Gly Lys Val Glu Asn Ser Trp Gly
 225 230 235 240

Lys Thr Gly Lys Glu Pro Gly Glu Tyr Asn Leu Pro His Met Leu Ala
 245 250 255

Val Asp Ala Ala Gly Asn Leu Phe Val Thr Glu Ile Gly Gly Lys Arg
 260 265 270

Leu Gln Lys Leu Gln Arg Lys
 275

<210> 5
 <211> 324
 <212> PRT
 <213> Rattus norvegicus

5

<400> 5
 Gly Asp Phe His Val Glu Glu Glu Leu Asp Trp Pro Gly Val Tyr Leu
 1 5 10 15

Leu Pro Gly Gln Val Ser Gly Val Ala Leu Asp Ser Lys Asn Asn Leu
 20 25 30

Val Ile Phe His Arg Gly Asp His Val Trp Asp Gly Asn Ser Phe Asp
 35 40 45

Ser Lys Phe Val Tyr Gln Gln Arg Gly Leu Gly Pro Ile Glu Glu Asp
 50 55 60

Thr Ile Leu Val Ile Asp Pro Asn Asn Ala Glu Ile Leu Gln Ser Ser
 65 70 75 80

Gly Lys Asn Leu Phe Tyr Leu Pro His Gly Leu Ser Ile Asp Thr Asp
 85 90 95

ES 2 583 259 T3

Gly Asn Tyr Trp Val Thr Asp Val Ala Leu His Gln Val Phe Lys Leu
 100 105 110

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

Gln Pro Gly Ser Asp Gln Asn His Phe Cys Gln Pro Thr Asp Val Ala
 130 135 140

Val Glu Pro Ser Thr Gly Ala Val Phe Val Ser Asp Gly Tyr Cys Asn
 145 150 155 160

Ser Arg Ile Val Gln Phe Ser Pro Ser Gly Lys Phe Val Thr Gln Trp
 165 170 175

Gly Glu Glu Ser Ser Gly Ser Ser Pro Arg Pro Gly Gln Phe Ser Val
 180 185 190

Pro His Ser Leu Ala Leu Val Pro His Leu Asp Gln Leu Cys Val Ala
 195 200 205

Asp Arg Glu Asn Gly Arg Ile Gln Cys Phe Lys Thr Asp Thr Lys Glu
 210 215 220

Phe Val Arg Glu Ile Lys His Ala Ser Phe Gly Arg Asn Val Phe Ala
 225 230 235 240

Ile Ser Tyr Ile Pro Gly Phe Leu Phe Ala Val Asn Gly Lys Pro Tyr
 245 250 255

Phe Gly Asp Gln Glu Pro Val Gln Gly Phe Val Met Asn Phe Ser Ser
 260 265 270

Gly Glu Ile Ile Asp Val Phe Lys Pro Val Arg Lys His Phe Asp Met
 275 280 285

Pro His Asp Ile Val Ala Ser Glu Asp Gly Thr Val Tyr Ile Gly Asp
 290 295 300

Ala His Thr Asn Thr Val Trp Lys Phe Thr Leu Thr Glu Lys Met Glu
 305 310 315 320

His Arg Ser Val

<210> 6
 <211> 323
 <212> PRT
 <213> Xenopus laevis

ES 2 583 259 T3

<400> 6

Gly Asp Val His Leu Glu Glu Asp Thr Asp Trp Pro Gly Val Asn Leu
 1 5 10 15

Lys Val Gly Gln Val Ser Gly Leu Ala Leu Asp Pro Lys Asn Asn Leu
 20 25 30

Ala Ile Phe His Arg Gly Asp His Val Trp Asp Glu Asn Ser Phe Asp
 35 40 45

Arg Asn Phe Val Tyr Gln Gln Arg Gly Ile Gly Pro Ile Gln Glu Ser
 50 55 60

Thr Ile Leu Val Val Asp Pro Ser Ser Ser Lys Val Leu Lys Ser Thr
 65 70 75 80

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

Gly Asn Tyr Trp Val Thr Asp Val Ala Leu His Gln Val Phe Lys Leu
 100 105 110

Gly Ala Gly Lys Glu Thr Pro Leu Leu Val Leu Gly Arg Ala Phe Gln
 115 120 125

Pro Gly Ser Asp Arg Lys His Phe Cys Gln Pro Thr Asp Val Ala Val
 130 135 140

Asp Pro Ile Thr Gly Asn Phe Phe Val Ala Asp Gly Tyr Cys Asn Ser
 145 150 155 160

Arg Ile Met Gln Phe Ser Pro Asn Gly Met Phe Ile Met Gln Trp Gly
 165 170 175

Glu Glu Thr Ser Ser Asn Val Pro Arg Pro Gly Gln Phe Arg Ile Pro
 180 185 190

His Ser Leu Thr Met Val Pro Asp Gln Gly Gln Leu Cys Val Ala Asp
 195 200 205

Arg Glu Asn Gly Arg Ile Gln Cys Phe His Ala Glu Thr Gly Asn Phe
 210 215 220

Val Lys Gln Ile Lys His Gln Glu Phe Gly Arg Glu Val Phe Ala Val
 225 230 235 240

Ser Tyr Ala Pro Gly Gly Val Leu Tyr Ala Val Asn Gly Lys Pro Tyr
 245 250 255

ES 2 583 259 T3

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

Gly Asp Ile Leu Asp Thr Phe Ile Pro Ala Arg Lys Asn Phe Asp Met
275 280 285

Pro His Asp Ile Ala Ala Ala Asp Asp Gly Thr Val Tyr Val Gly Asp
290 295 300

Ala His Ala Asn Ala Val Trp Lys Phe Ser Pro Ser Lys Ala Glu His
305 310 315 320

Arg Ser Val

<210> 7

<211> 149

<212> PRT

<213> Thermotoga maritima

5

<400> 7

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

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

Arg Gly Phe Ala Lys Glu Tyr Thr Glu Gly Leu Glu Arg Ala Ile Lys
35 40 45

His Glu Lys Glu Ile Glu Lys Arg Lys Lys Glu Arg Glu Arg Glu Glu
50 55 60

Ser Glu Lys Ile Leu Lys Glu Leu Lys Lys Arg Thr His Val Val Lys
65 70 75 80

Val Lys Ala Gly Glu Gly Gly Lys Ile Phe Gly Ala Val Thr Ala Ala
85 90 95

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

Arg Trp Phe Lys Leu Asp Lys Pro Ile Lys Glu Leu Gly Glu Tyr Ser
115 120 125

Leu Glu Val Ser Leu Pro Gly Gly Val Lys Asp Thr Ile Lys Ile Arg
130 135 140

Val Glu Arg Glu Glu

145

10

ES 2 583 259 T3

<400> 13

Glu Ala Pro Pro Val Thr Ile Asp Glu Ser Trp Pro Asp Ile Pro Glu
 1 5 10 15

Ser Ala Val Phe Gly Glu Pro Thr Ala Ile Asp Val Asp Ser His Gly
 20 25 30

His Ile Phe Val Leu His Arg Ala Gly Arg Glu Trp Thr Gln Pro Phe
 35 40 45

Pro Ser Asp Pro Ile Ser Glu Pro Thr Val Phe Met Phe Ala Ala Asn
 50 55 60

Gly Lys Leu Leu Ser Lys Trp Gly Ala Gly Glu Leu Val Met Pro His
 65 70 75 80

Gly Leu Ser Ile Asp Gly Asp Asn Lys Val Trp Ile Thr Asp Val Ala
 85 90 95

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

Leu Gly Thr Arg Gly Glu Thr Gly Gln Asp Glu Ser His Phe Gly Arg
 115 120 125

Pro Ala Asp Val Thr Phe Val Gly Asp Arg Val Leu Val Ala Asp Gly
 130 135 140

Tyr Leu Asn Arg Arg Ile Met Val Phe Asp Arg Ala Gly Asn Phe Leu
 145 150 155 160

Glu Gln Trp Gly Lys Glu Gly Glu Asp Ala Gly Glu Phe Asn Leu Pro
 165 170 175

His Ala Ile Ala Ala Asp Ser Glu Arg Ile Tyr Val Ala Asp Arg Glu
 180 185 190

Asn Ala Arg Val Gln Val Leu Ser Leu Asp Gly Glu Pro Leu Ala Arg
 195 200 205

Trp Arg Gln Asp Gly Thr Gly His Pro Tyr Ala Val Lys Pro Ile Gly

ES 2 583 259 T3

210 215 220

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

Thr Ala Ile Gly Arg Ile Tyr Arg Ala Asp Gly Gly Leu Glu Arg Val
 245 250 255

Phe Asp Ala Gly Val Glu Pro His Thr Gly Thr Ser Leu Gly His Asp
 260 265 270

Val Ala Ile Gly Pro Asp Gly Ser Ala Tyr Met Val Asp Asn Lys Ala
 275 280 285

Asn Arg Val Ile Lys Phe Asp Leu Ser Arg Ala Gly Val Glu Glu Ala
 290 295 300

Asp Ala Asp
 305

<210> 14
 <211> 13
 <212> PRT
 <213> ARTIFICIAL

<220>
 <223> péptido

<400> 14
 Ser Ser Ser Gly Gly Ser Gly Ser Glu Thr Leu Phe Gln
 1 5 10

<210> 15
 <211> 310
 <212> PRT
 <213> Erythrobacter sp. SD-21

<400> 15
 Ala Arg Glu Glu Ala Pro Pro Val Thr Ile Asp Glu Ser Trp Pro Asp
 1 5 10 15

Ile Pro Glu Ser Ala Val Phe Gly Glu Pro Thr Ala Ile Asp Val Asp
 20 25 30

Ser His Gly His Ile Phe Val Leu His Arg Ala Gly Arg Glu Trp Thr
 35 40 45

Gln Pro Phe Pro Ser Asp Pro Ile Ser Glu Pro Thr Val Phe Met Phe
 50 55 60

Ala Ala Asn Gly Lys Leu Leu Ser Lys Trp Gly Ala Gly Glu Leu Val

ES 2 583 259 T3

<400> 16
 Ser Ser Ser Gly Gly Ser Gly Gly Ser
 1 5

5

<210> 17
 < 211> 129
 < 212> PRT
 < 213> ARTIFICIAL

<220>
 < 223> péptido

<400> 17
 Met Ala His Lys Lys Ser Gly Gly Val Ala Lys Asn Gly Arg Asp Ser
 1 5 10 15

Leu Pro Lys Tyr Leu Gly Val Lys Val Gly Asp Gly Gln Ile Val Lys
 20 25 30

Ala Gly Asn Ile Leu Val Arg Gln Arg Gly Thr Arg Phe Tyr Pro Gly
 35 40 45

Lys Asn Val Gly Val Gly Arg Asp Phe Thr Leu Phe Ala Leu Lys Asp
 50 55 60

Gly Arg Val Lys Phe Glu Thr Lys Asn Asn Lys Lys Tyr Val Ser Val
 65 70 75 80

Tyr Glu Glu Ser Ser Ser Asp Asp Asp Asp Lys Lys Cys Asn Thr Ala
 85 90 95

Thr Cys Ala Thr Gln Arg Leu Ala Asn Phe Leu His His Ser Ser Asn
 100 105 110

Asn Phe Gly Pro Ile Leu Pro Pro Thr Asn Val Gly Ser Asn Thr Tyr
 115 120 125

10

Gly

<210> 18
 < 211> 23
 < 212> PRT
 < 213> ARTIFICIAL

15

<220>
 < 223> péptido

20

<220>
 < 221> misc_feature
 < 222> (1)..(1)
 < 223> Xaa puede ser cualquier aminoácido natural excepto Cys

<220>
 < 221> misc_feature
 < 222> (3)..(3)
 < 223> Xaa puede ser cualquier aminoácido natural

ES 2 583 259 T3

<220>
< 221> misc_feature
< 222> (6)..(8)
< 223> Xaa puede ser cualquier aminoácido natural

5
<220>
< 221> misc_feature
< 222> (10)..(10)
< 223> Xaa puede ser cualquier aminoácido natural

10
<220>
< 221> misc_feature
< 222> (12)..(12)
< 223> Xaa puede ser cualquier aminoácido natural excepto Cys

15
<220>
< 221> misc_feature
< 222> (13)..(16)
< 223> Xaa puede ser cualquier aminoácido natural

20
<220>
< 221> misc_feature
< 222> (18)..(22)
< 223> Xaa puede ser cualquier aminoácido natural

<400> 18
Xaa Val Xaa Asp Arg Xaa Xaa Xaa Arg Xaa Gln Xaa Xaa Xaa Xaa Xaa
1 5 10 15

Gly Xaa Xaa Xaa Xaa Xaa Trp
20

REIVINDICACIONES

- 5 1. Un método para producir un péptido α -amidado que comprende las etapas de: (i) permitir que un péptido diana con un residuo de glicina C-terminal reaccione tanto con una monooxigenasa alfa-hidroxilante de peptidilglicina (PHM) como con un polipéptido procedente de un organismo procariota con actividad liasa alfa-amidante de peptidil- α -hidroxiglicina (PAL), en condiciones adecuadas para la actividad enzimática, en donde la reacción con dicha PHM y dicho polipéptido en dicho péptido diana se lleva a cabo ya sea en dos etapas distintas o simultáneamente; (ii) recuperar el péptido α -amidado en el extremo C-terminal.
- 10 2. El método según la reivindicación 1, en el que el polipéptido procedente de un organismo procariota tiene una secuencia de aminoácidos que no comprende residuos de cisteína, o como máximo 1 residuo de cisteína o como máximo 2 residuos de cisteína.
3. El método según cualquiera de las reivindicaciones 1-2, en el que dicho polipéptido es la enzima de tipo silvestre.
4. El método según cualquiera de las reivindicaciones 1-2, en el que dicho polipéptido comprende de 1-30 modificaciones de aminoácidos en comparación con el polipéptido de tipo silvestre.
- 15 5. El método según cualquiera de las reivindicaciones 1-4, en el que dicho polipéptido procedente de un organismo procariota con actividad liasa alfa-amidante de peptidil- α -hidroxiglicina (PAL) se ha expresado en una *E. coli* recombinante.
6. El método según cualquiera de las reivindicaciones 1-5, en el que el polipéptido procedente de un organismo procariota se ha preparado sin una etapa de replegamiento para obtener actividad catalítica.
- 20 7. El método según cualquiera de las reivindicaciones 1-6, en el que están presentes iones Cu^{2+} y/o ácido ascórbico cuando se utiliza dicho polipéptido.
8. El método según cualquiera de las reivindicaciones 1-7, en donde la enzima PHM es de un organismo mamífero.
- 25 9. El método según cualquiera de las reivindicaciones 1-8, en el que dicho péptido diana se selecciona a partir del grupo que consiste en amilina, neuropéptido Y (NPY), péptido YY (PYY), PYY-3-36, polipéptido pancreático (PP), péptido similar a glucagón (GLP-1), gastrina, calcitonina, péptido relacionado con calcitonina (CGRP), péptido liberador de gastrina, vasopresina, oxitocina, neuroquinina A, secretina, pancreastatina, pro-opiomelanocortina (POMC), hormona estimulante de melanocitos alfa (MSH alfa), hormona estimulante de melanocitos gamma (MSH gamma 1) y péptido bisagra amidado (HP-N) o análogos funcionales de los mismos.

Fig. 1/7

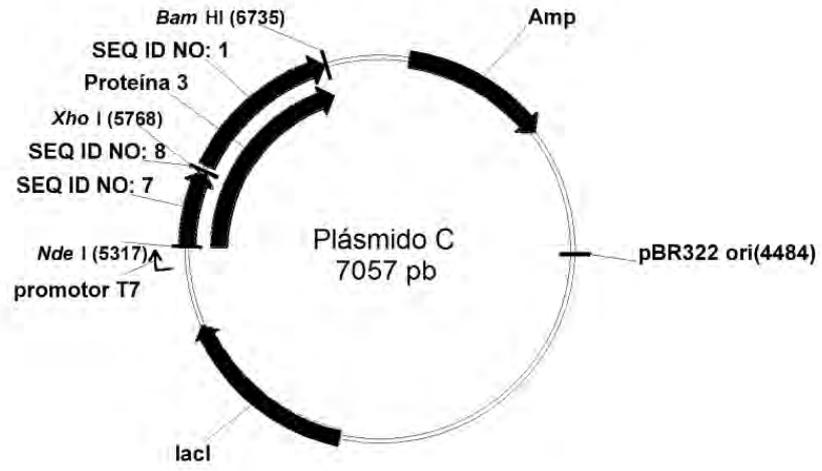


Fig. 2/7

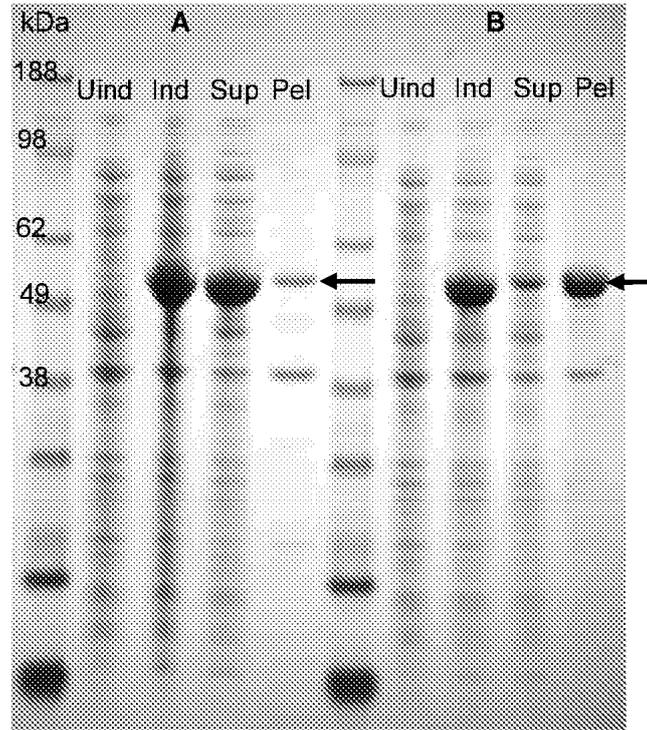


Fig. 3/7

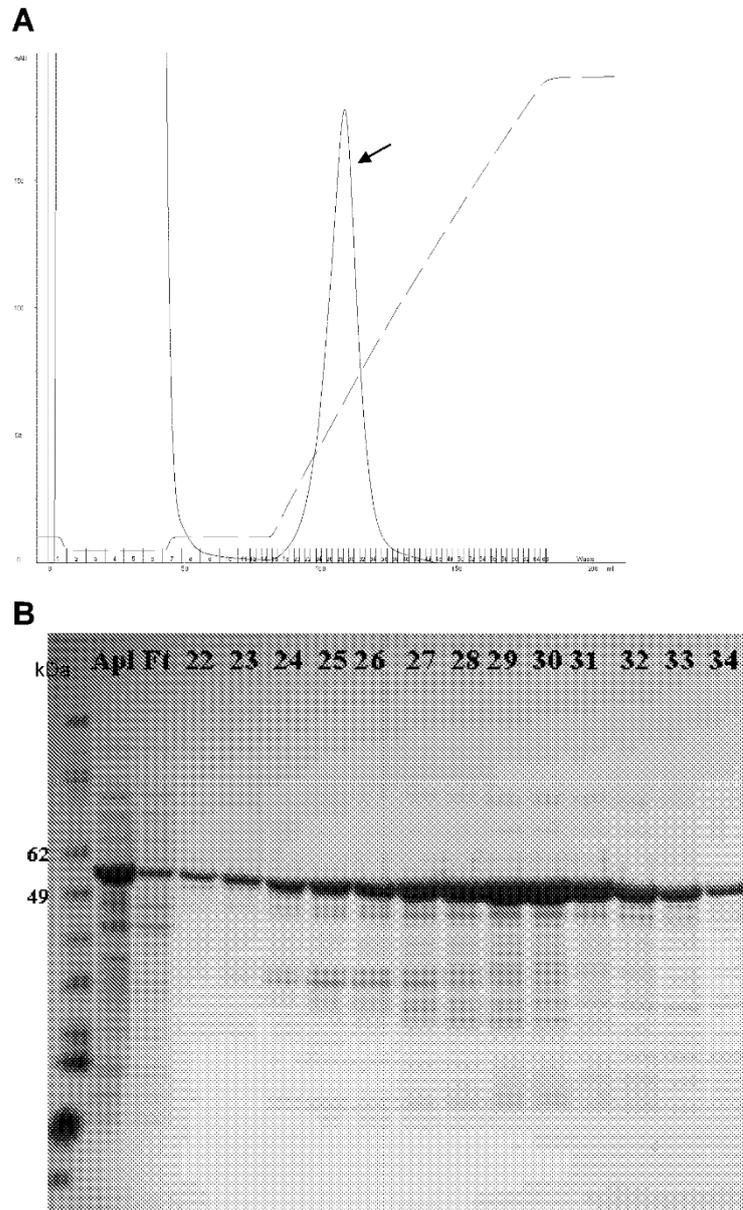
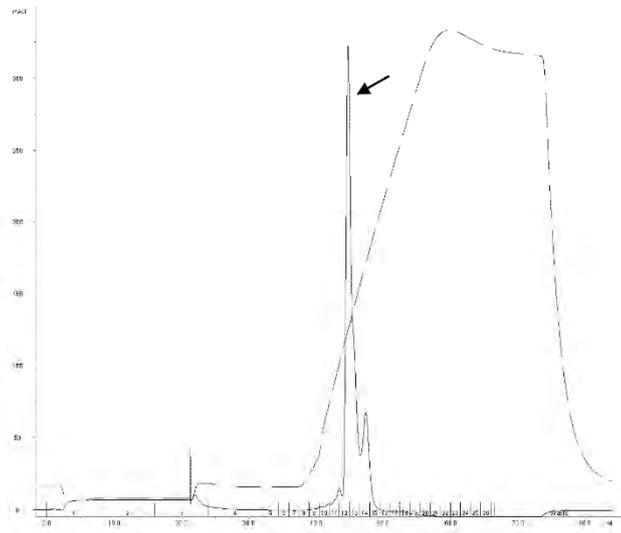


Fig. 4/7

A



B

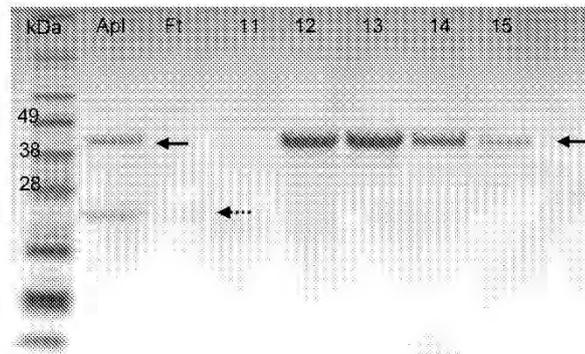


Fig. 5A1/7

+TAP, MES pH5.5

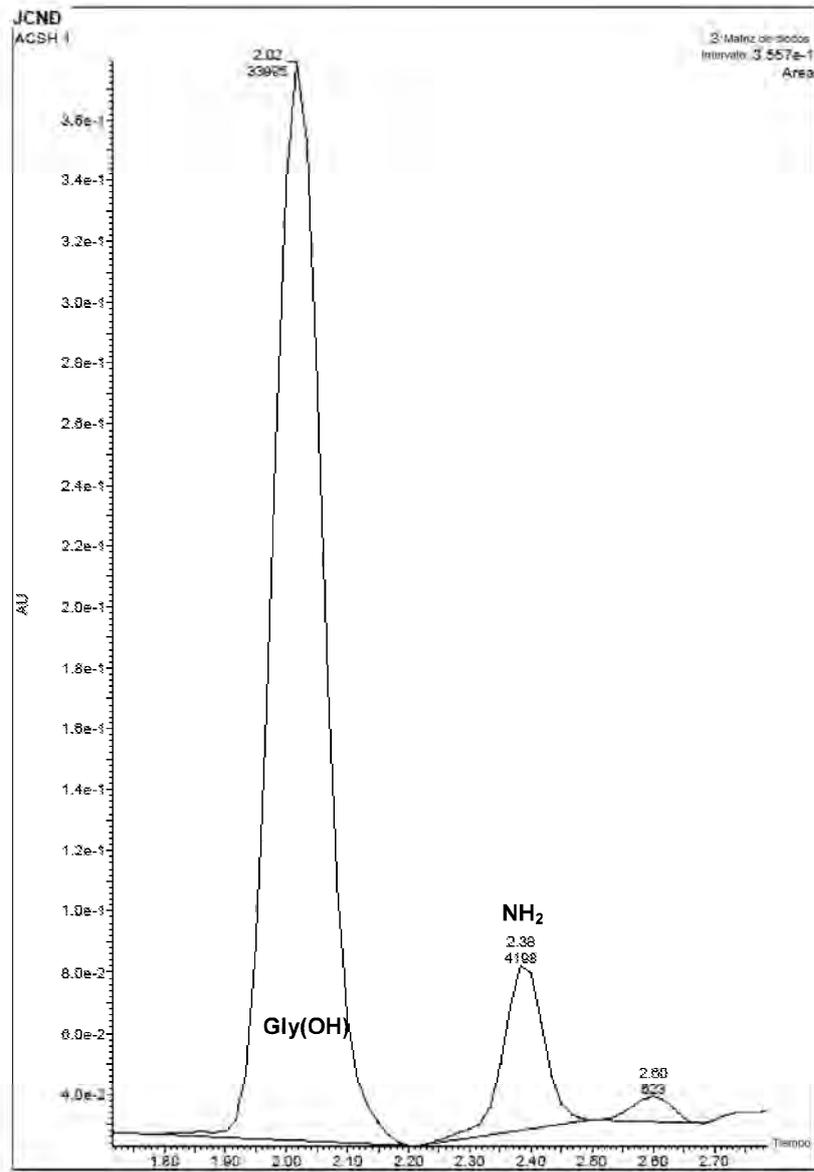


Fig. 5A2/7

+TAP, Tris pH7.5

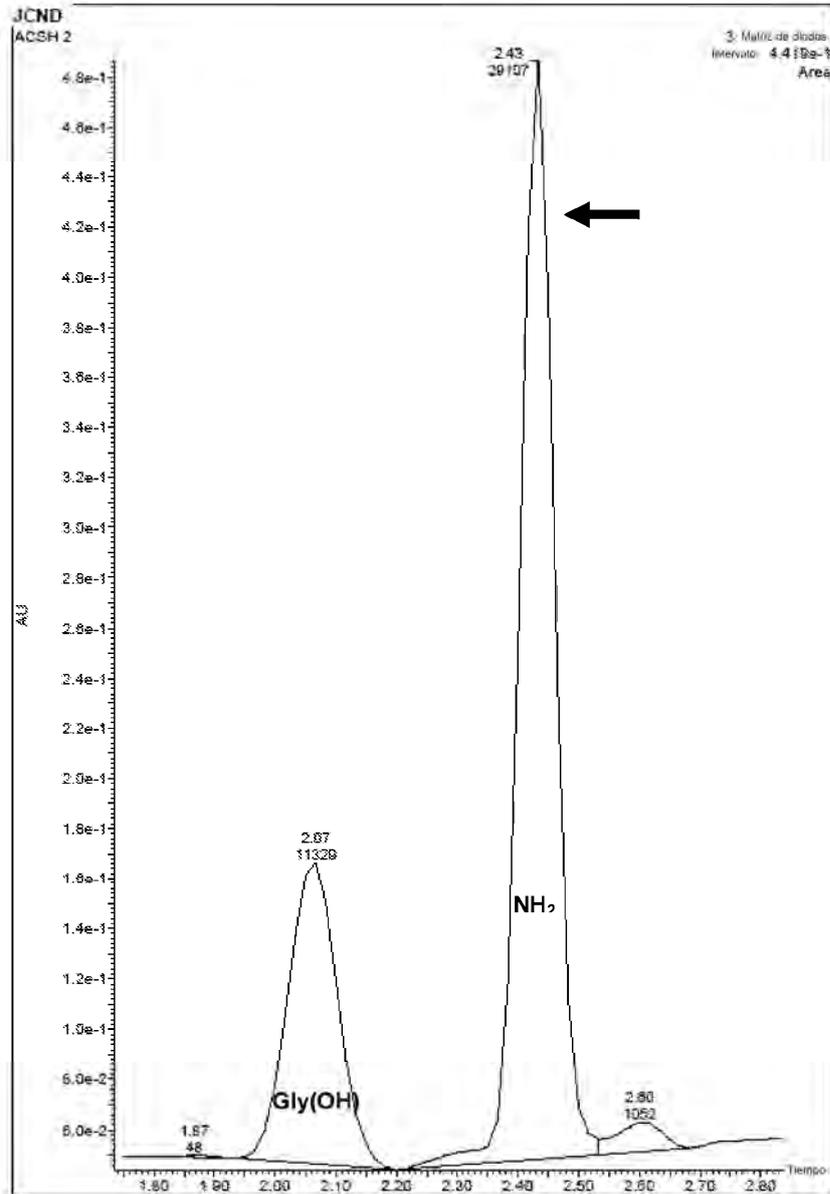


Fig. 5A3/7

-TAP, MES pH5.5

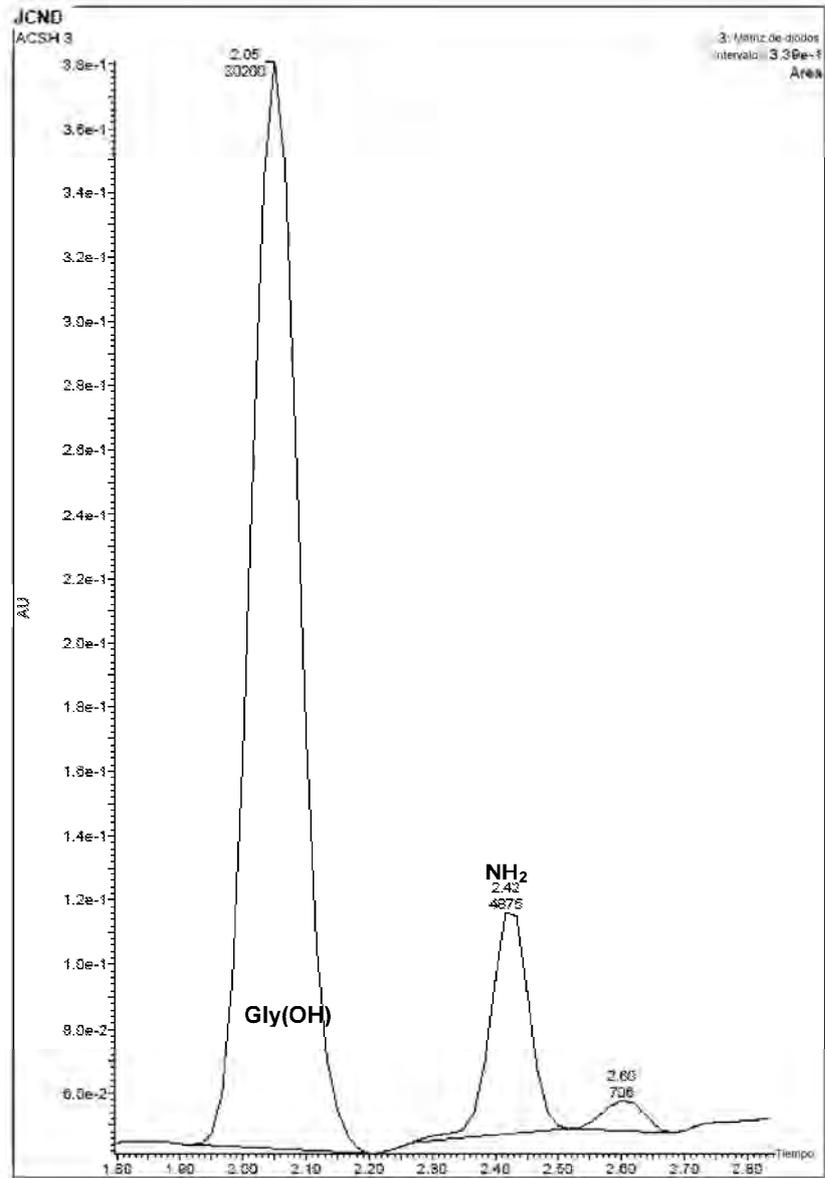


Fig. 5A4/7

-TAP, Tris pH7.5

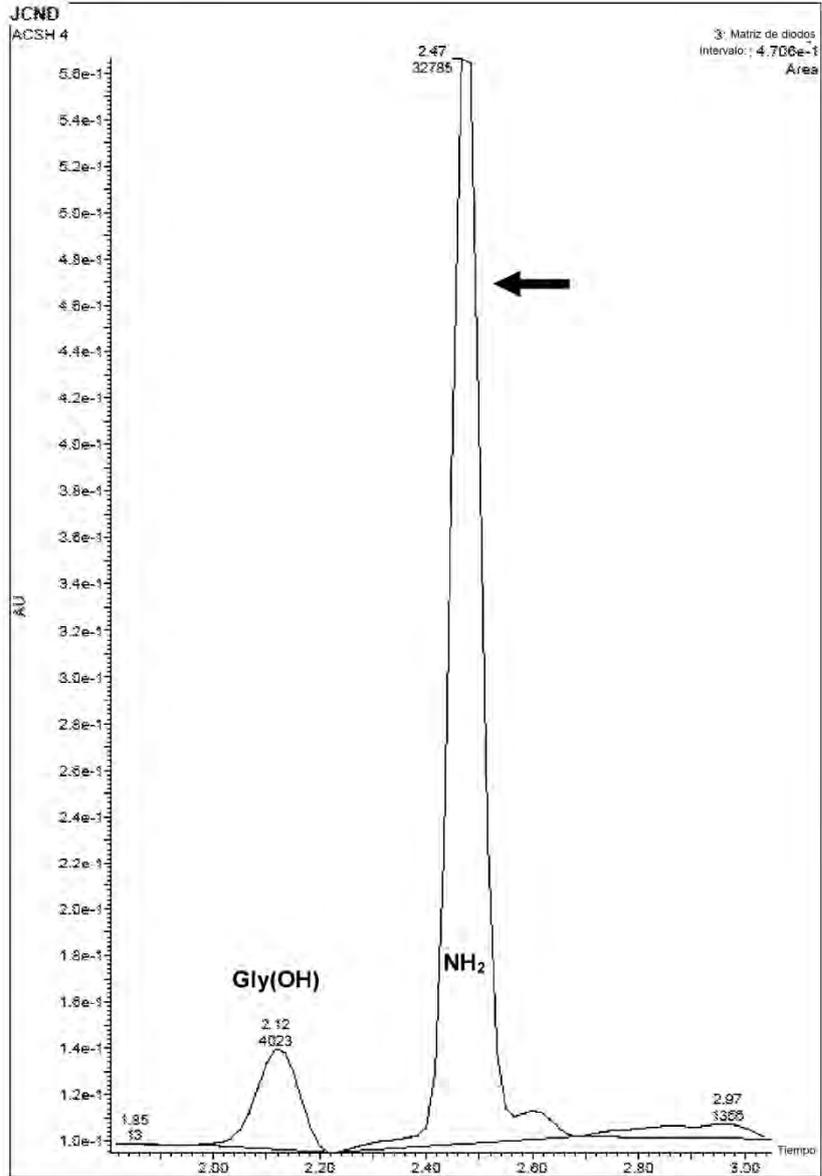


Fig. 5A5/7

MES pH 5.5, ctrl. neg.

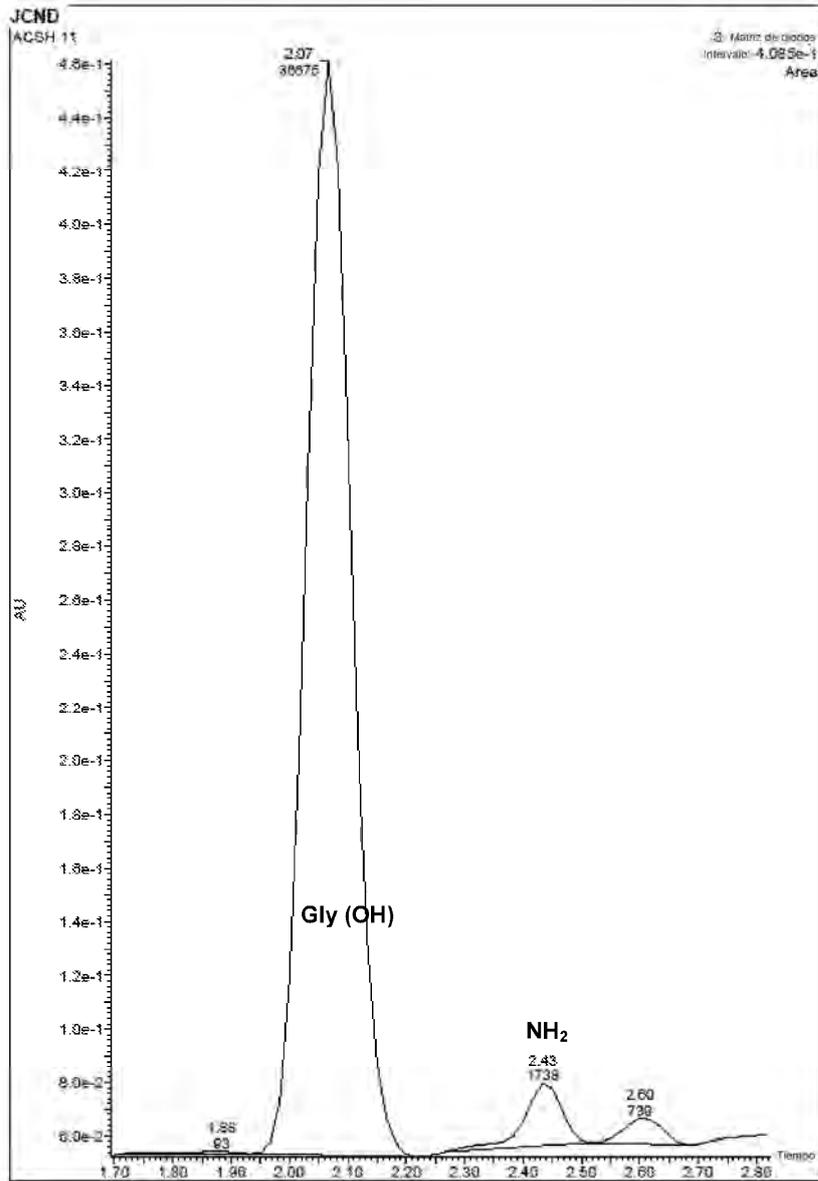


Fig. 5A6/7

Tris pH 7.5, ctrl. neg.

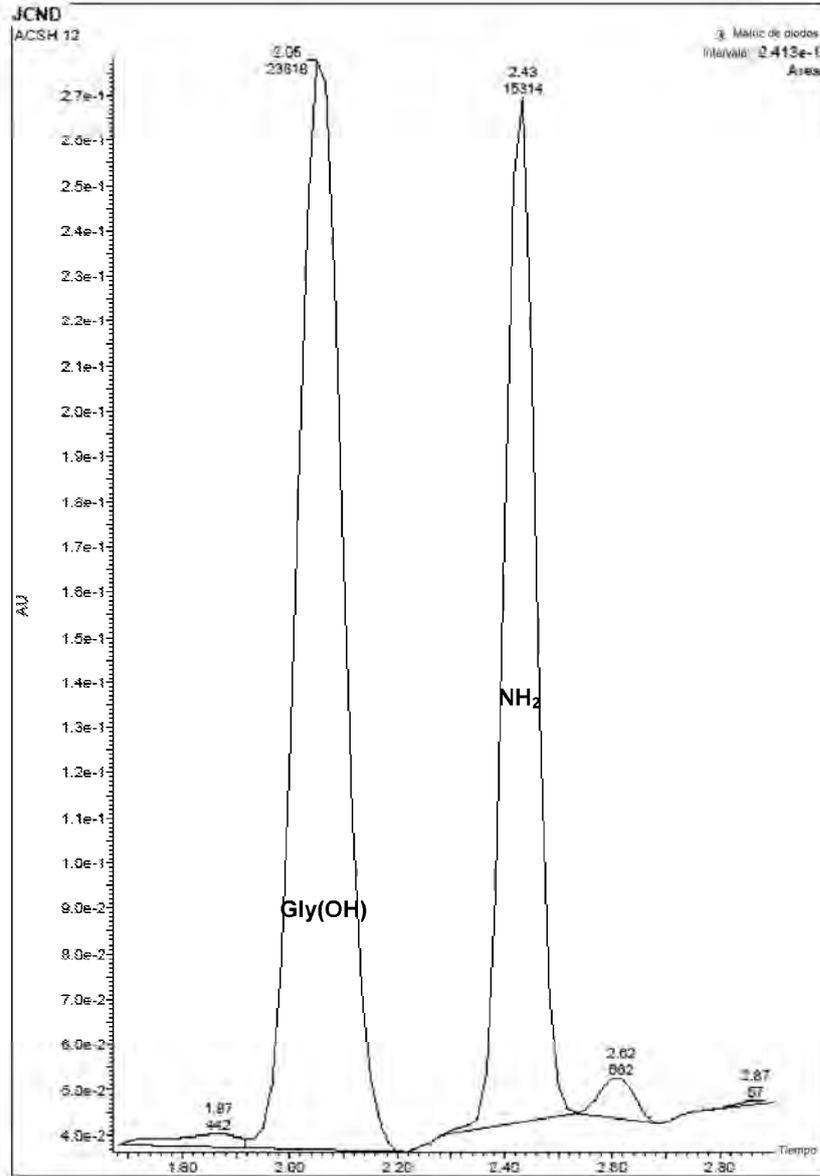


Fig. 5B1/7

+ marcador TAP MES pH 5.5

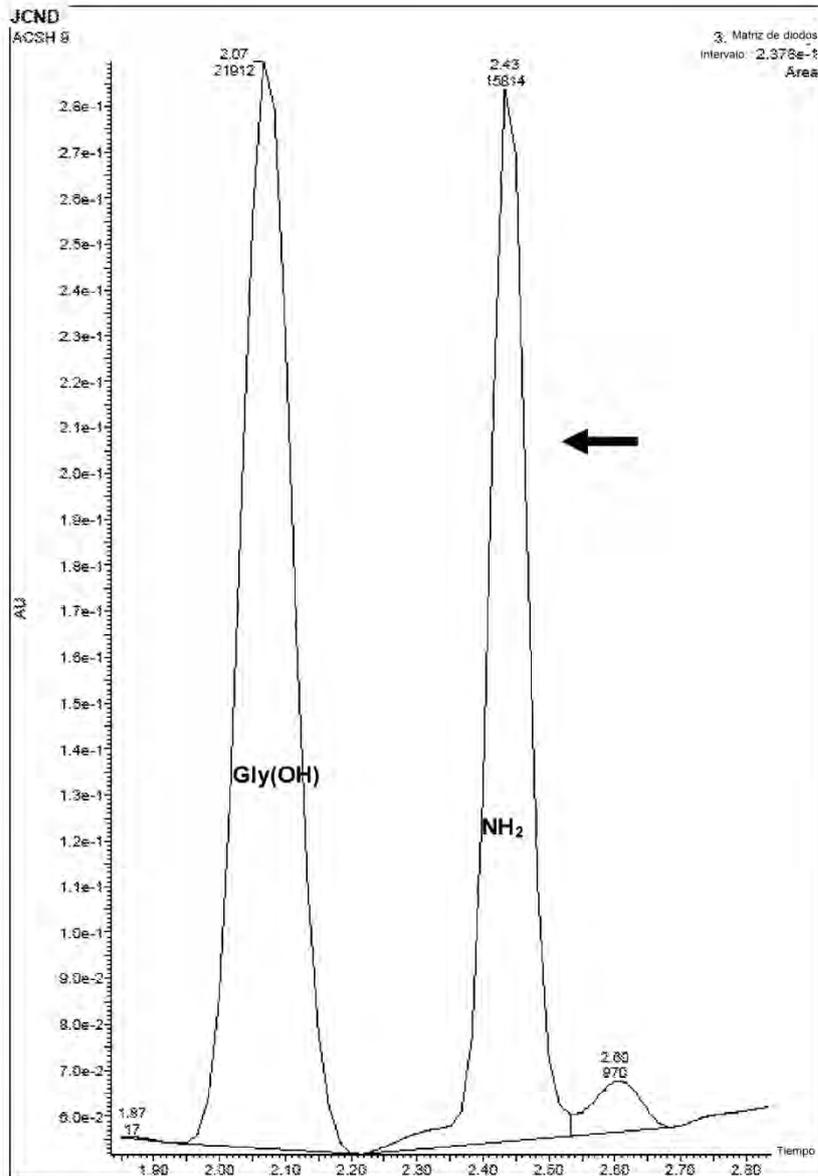


Fig. 5B2/7

+ marcador TAP, Tris pH 7.5

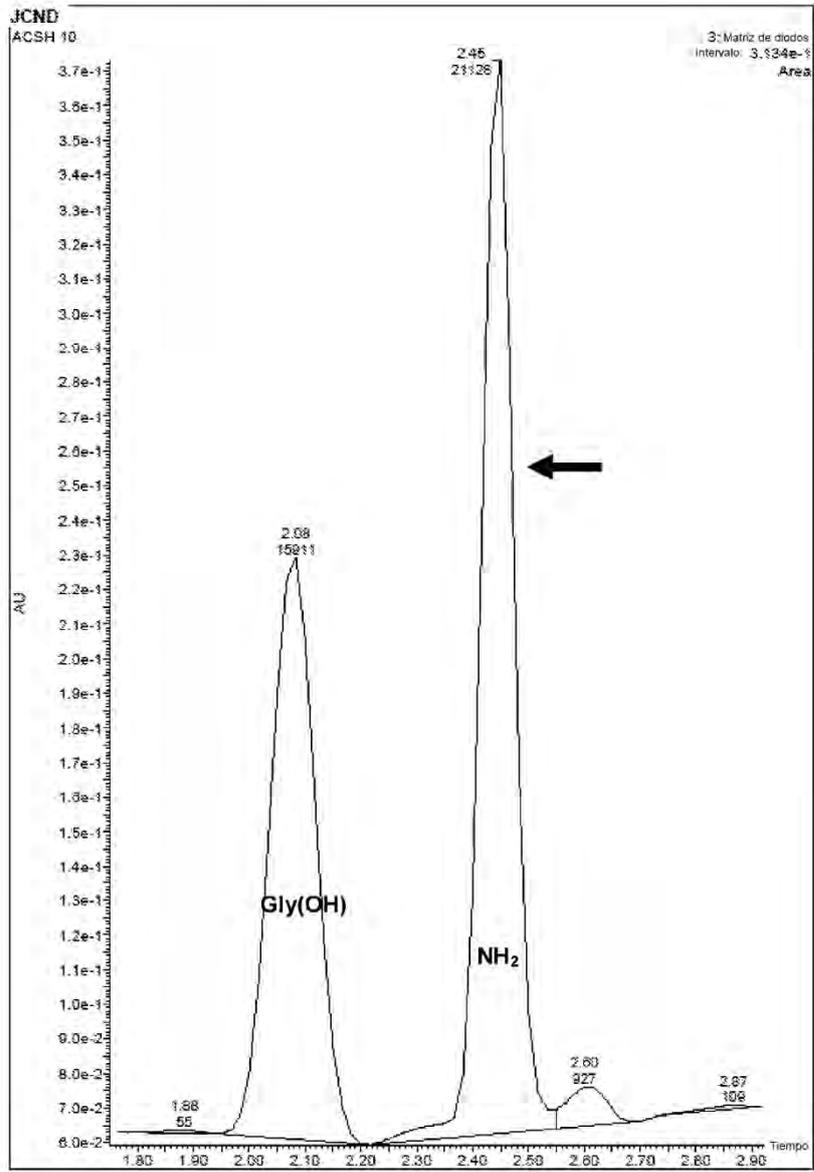


Fig. 5B3/7

Ctrl. neg., MES pH 5.5

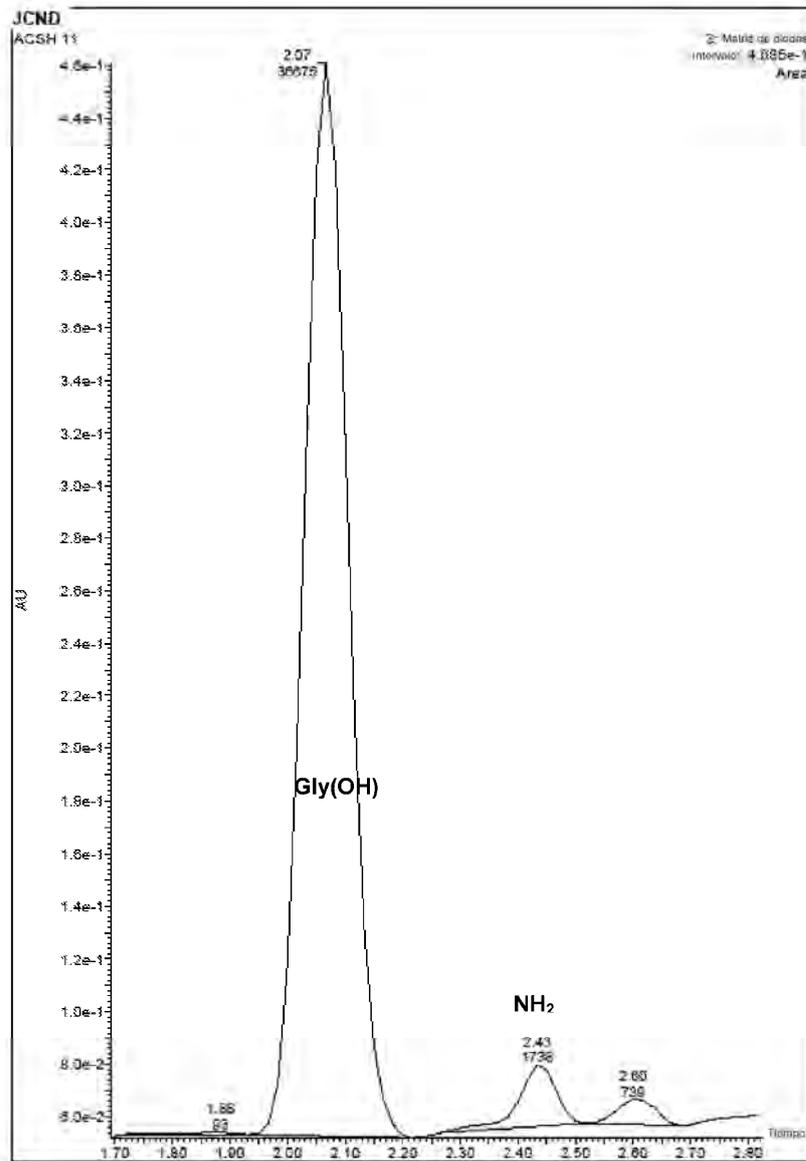


Fig. 5B4/7

Ctrl. neg., Tris pH 7.5

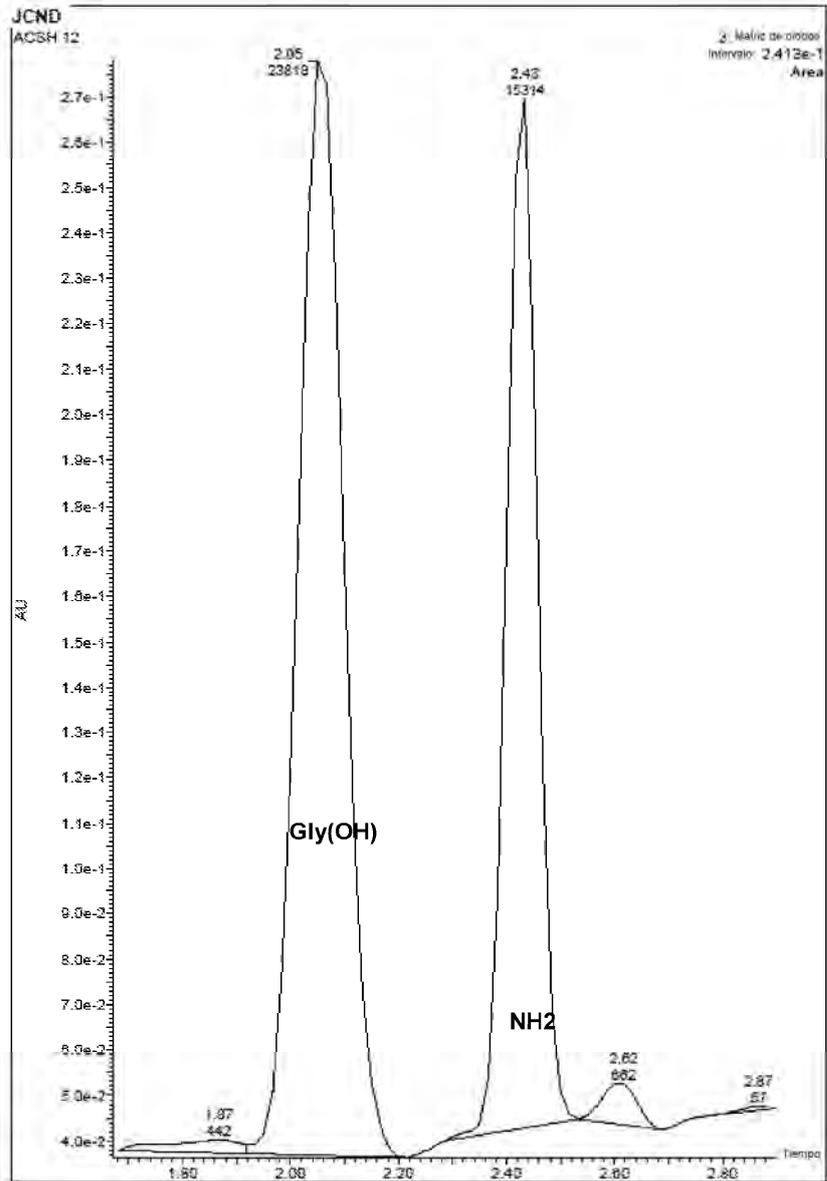


Fig. 6/7

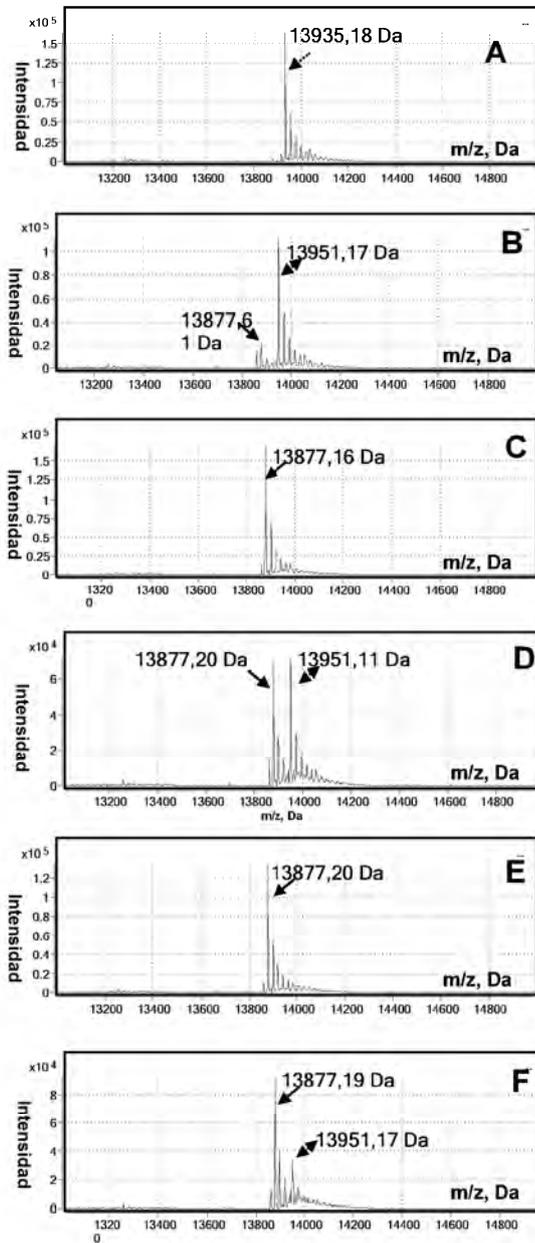
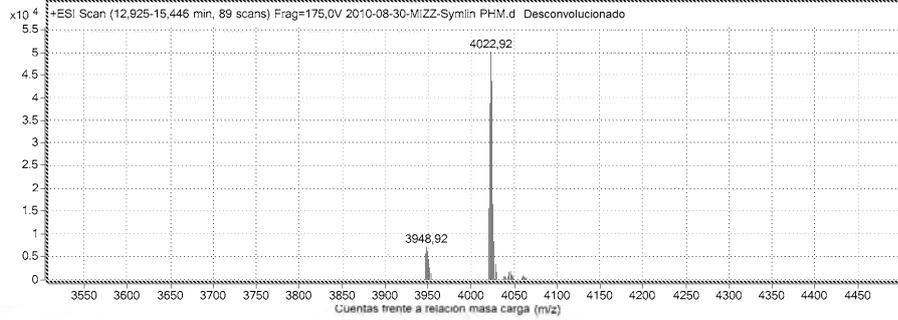


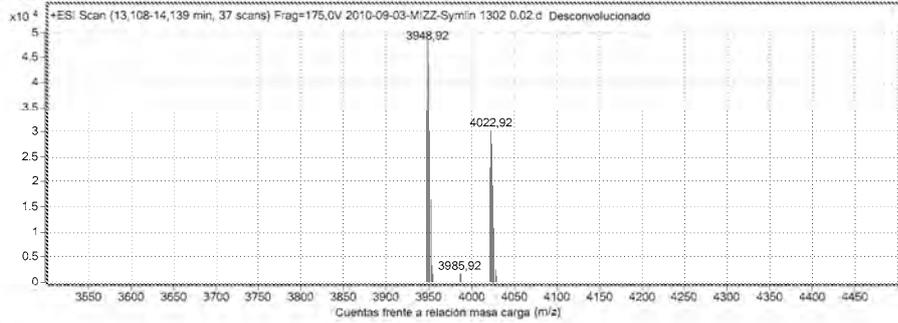
Fig. 7A7

Dominio de tipo PAL de Erythrobacter (SEQ ID NO: 1)

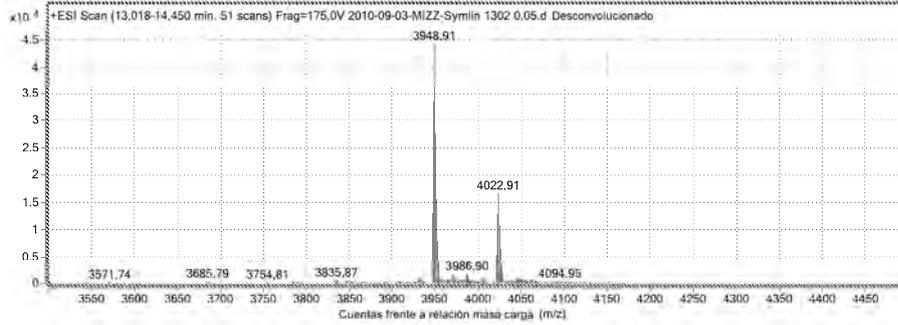
ctrl



1:1250



1:500



1:100

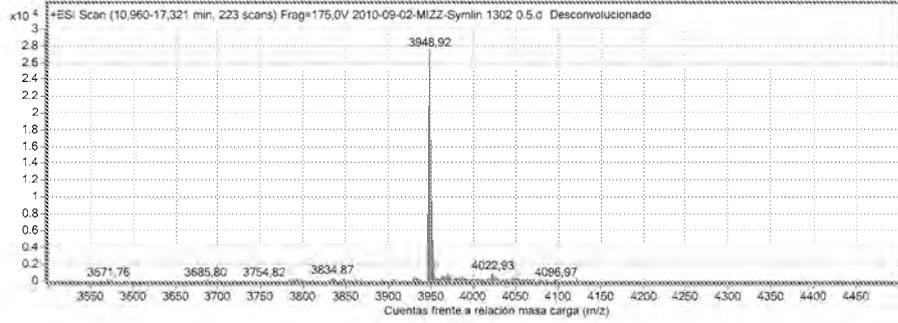
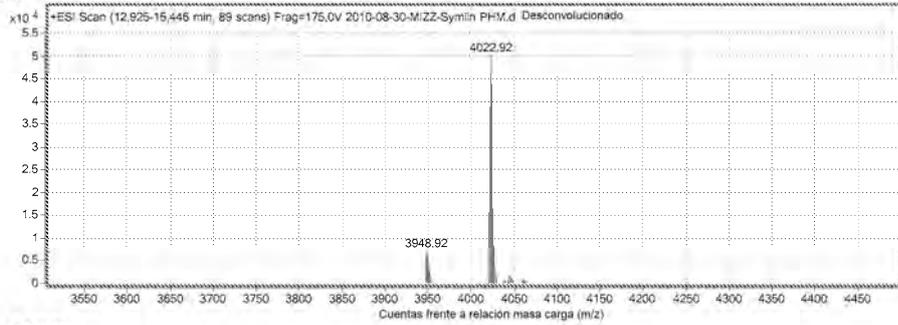


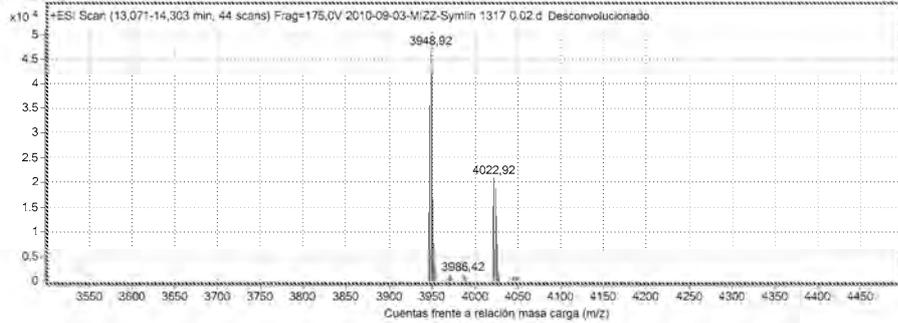
Fig. 7B/7

Dominio de tipo PAL de Planctomyces (SEQ ID NO: 4)

ctrl



1:1250



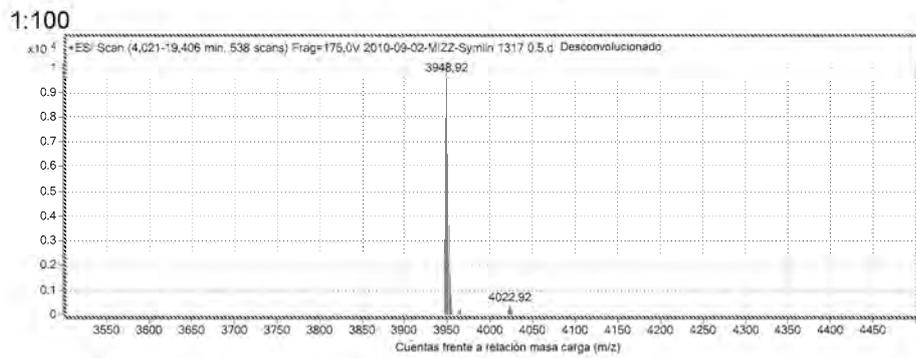
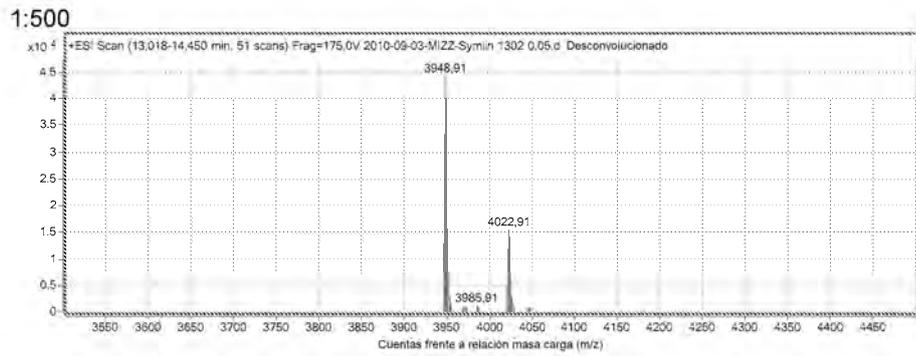
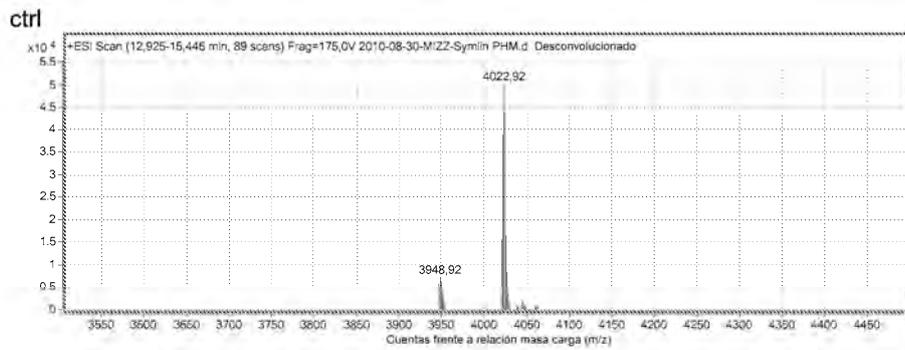


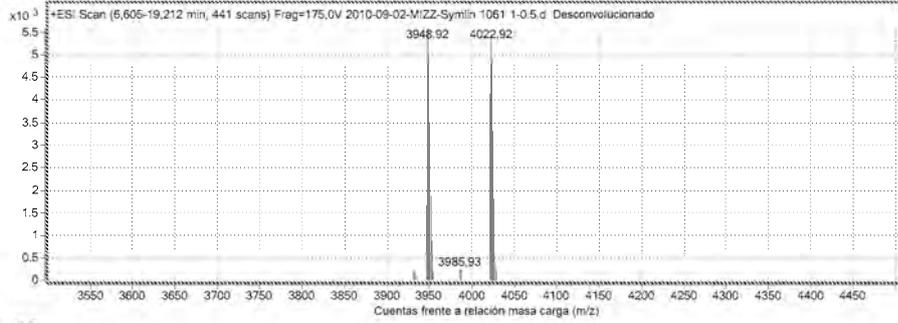
Fig. 7C/7

Dominio de tipo PAL de Exiguobacterium (SEQ ID NO: 2)

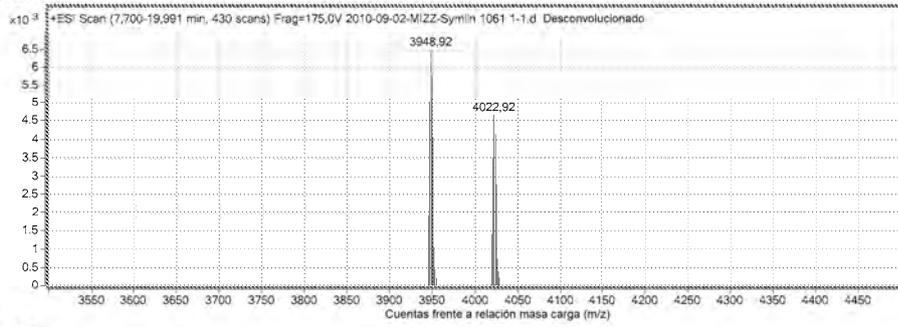


ES 2 583 259 T3

1:100



1:50



1:25

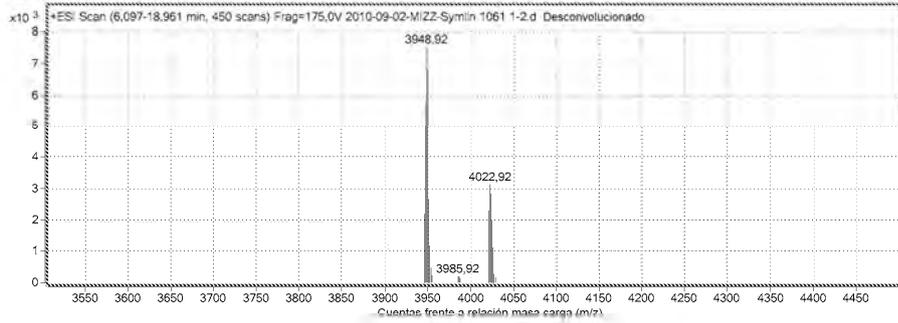
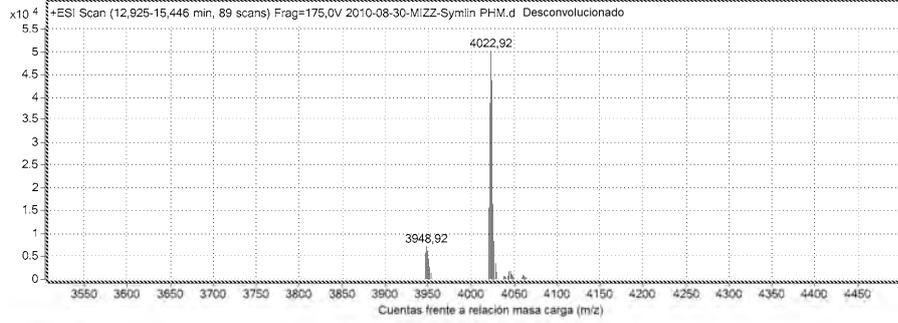


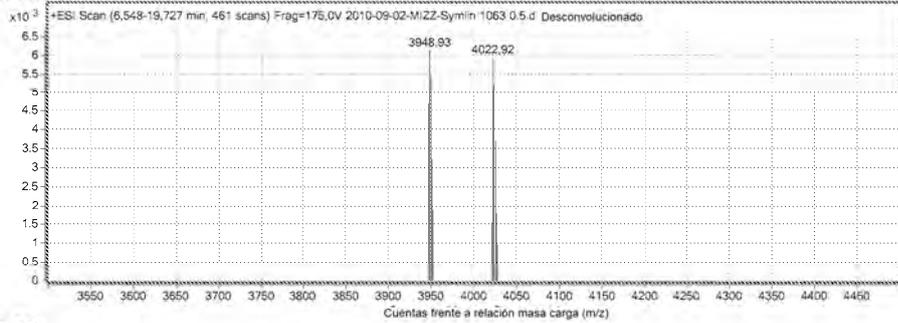
Fig. 7D/7

Dominio de tipo PAL de Chthoniobacter (SEQ ID NO: 3)

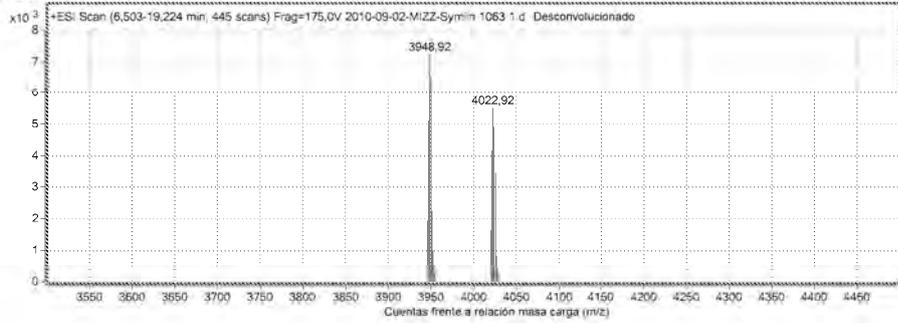
ctrl



1:100



1:50



1:25

