

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 595 500**

51 Int. Cl.:

<b>A61K 38/48</b>	(2006.01)
<b>C12N 9/52</b>	(2006.01)
<b>C12N 1/20</b>	(2006.01)
<b>A61P 1/00</b>	(2006.01)
<b>A61P 17/00</b>	(2006.01)
<b>A23L 2/52</b>	(2006.01)
<b>A23J 3/34</b>	(2006.01)

12

## TRADUCCIÓN DE PATENTE EUROPEA

T3

- 86 Fecha de presentación y número de la solicitud internacional: **05.11.2012 PCT/EP2012/071816**
- 87 Fecha y número de publicación internacional: **13.06.2013 WO13083338**
- 96 Fecha de presentación y número de la solicitud europea: **05.11.2012 E 12781098 (4)**
- 97 Fecha y número de publicación de la concesión europea: **29.06.2016 EP 2788018**

54 Título: **Nuevas proteasas que pueden hidrolizar péptidos y proteínas de gluten a un pH ácido, procedentes del actinomiceto Actinoallomurus**

30 Prioridad:

**06.12.2011 EP 11425291**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**30.12.2016**

73 Titular/es:

**FONDAZIONE ISTITUTO INSUBRICO DI RICERCA  
PER LA VITA (100.0%)  
Via Roberto Lepetit, 34  
21040 Gerenzano (VA), IT**

72 Inventor/es:

**CAVALETTI, LINDA;  
CARRANO, LUCIA;  
ABBONDI, MONICA;  
BRUNATI, MARA y  
TARAVELLA, ANNA**

74 Agente/Representante:

**CURELL AGUILÁ, Mireia**

ES 2 595 500 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 proteasas que pueden hidrolizar péptidos y proteínas de gluten a un pH ácido, procedentes del actinomiceto *Actinoallomurus*.

5

**Campo de la invención**

La presente invención se refiere a una nueva familia de endopeptidasas que presentan propiedades catalíticas únicas que les permiten degradar polipéptidos de gran tamaño, incluyendo los ricos en prolina. La presente invención se refiere además a métodos para producir la composición de enzimas, así como una composición farmacéutica y un complemento alimenticio que contiene la composición de enzimas y la utilización de la misma en la degradación de polipéptidos.

10

**Antecedentes de la invención**

15

La enfermedad celíaca (EC) es un trastorno crónico del tracto gastrointestinal en el que la ingestión de gluten, presente en productos alimentarios preparados a partir del trigo, el centeno, la cebada y variedades cruzadas de los mismos, conduce a daños en la mucosa del intestino delgado mediante un mecanismo autoinmunitario en individuos genéticamente susceptibles (Green P.H.R., Cellier C., "Celiac Disease", N. Engl. J. Med. 357:1731-1743, 2007; Kagnoff M.F., "Celiac disease: pathogenesis of a model immunogenetic disease", J. Clin. Invest. 117:41-9, 2007). Los mecanismos mediante los que el gluten induce sus efectos patogénicos se han explicado en los últimos años. Participan mecanismos inmunitarios tanto innatos como adaptativos y son responsables del daño final en las mucosas.

20

25

El gluten consiste de gliadinas y gluteninas, las fracciones insolubles en agua/sal de proteínas de reserva presentes en los granos de cereal. Se crea una red de gluten mediante la interacción entre las dos proteínas en el caso de que se mezclen harina y agua en la preparación de la masa.

30

35

Tras la ingestión, el gluten se dirige a una digestión parcial por enzimas proteolíticos gástrico-pancreáticos o de borde en cepillo que resulta en muchos péptidos de longitud diferente (entre unos cuantos y más de 30 aminoácidos) que son resistentes a la digestión adicional debido al alto contenido en residuos de prolina ya que muchas proteasas son incapaces de cortar los enlaces peptídicos situados en los extremos N-terminal o C-terminal de la prolina (Hausch F., Shan L., Santiago N.A., Gray G.M., Khosla C., "Intestinal digestive resistance of immunodominant gliadin peptides", Am. J. Physiol. Gastrointest. Liver Physiol. 283:996-1003, 2002; Shan L., Molberg O., Parrot I., Hausch F., Filiz F., Gray G.M., Sollid L.M., Khosla C., "Structural basis for gluten intolerance in celiac sprue", Science 297:2275-2279, 2002). La falta de enzimas de corte específicos de la prolina no es una deficiencia enzimática específica en los sujetos celíacos, tal como se sugería anteriormente, sino que es propia del aparato digestivo de los mamíferos, que no ha evolucionado para consumir proteínas con un contenido tan elevado de proteína en la dieta.

40

45

50

Los péptidos del gluten no digeridos pueden pasar la barrera epitelial mediante mecanismos todavía no totalmente conocidos, aunque recientemente se ha propuesto un paso paracelular mediado por zonulina y una ruta transcelular mediante transcitosis y retrotranscitosis (Drago S., El Asmar R., Di Pierro M., Grazia Clemente M., Tripathi A., Sapone A., Thakar M., Iacono G., Carroccio A., D'Agate C., Not T., Zampini L., Catassi C., Fasano A., "Gliadin, zonulin and gut permeability: Effects on celiac and non-celiac intestinal mucosa and intestinal cell lines", Scand. J. Gastroenterol. 41:408-19, 2006; Schumann M., Richter J.F., Wedell I., Moos V., Zimmermann-Kordmann M., Schneider T., Daum S., Zeitz M., Fromm M., Schulzke J.D., "Mechanisms of epithelial translocation of the alpha(2)-gliadin-33mer in celiac sprue", Gut 57:747-754, 2008; Matysiak-Budnik T., Moura I.C., Arcos-Fajardo M., Lebreton C., Ménard S., Candalh C., Ben-Khalifa K., Dugave C., Tamouza H., van Niel G., Bouhnik Y., Lamarque D., Chaussade S., Malamut G., Cellier C., Cerf-Bensussan N., Monteiro R.C., Heyman M., "Secretory IgA mediates retrotranscytosis of intact gliadin peptides via the transferrin receptor in celiac disease", J. Exp. Med. 205:143-154, 2008).

55

60

Una vez en la lámina propia (LM), la desamidación de la glutamina en residuos glutamato por la transglutaminasa tisular (TGt, el autoantígeno en la EC) refuerza su presentación a las células T CD4+ DQ2 o DQ8 (Molberg O., McAdam S., Lundin K.E., Kristiansen C., Arentz-Hansen H., Kett K., Sollid L.M., "T cells from celiac disease lesions recognize gliadin epitopes deamidated in situ by endogenous tissue transglutaminase", Eur. J. Immunol. 31:1317-23, 2001), produciendo una respuesta proinflamatoria con interferón-gamma (IFN- $\gamma$ ) como citoquina efectora principal. También se ha demostrado que varios péptidos de gluten provocan daños mucosales con independencia de un reconocimiento específico por parte de los linfocitos T CD4+, induciendo una respuesta inmunológica innata mediante la regulación positiva de la expresión de IL-15, la ciclooxigenasa-2 y los marcadores de activación CD25 y CD83 en las células mononucleares LM: entre ellas, las mejor caracterizadas son los péptidos p31-43/49 de las  $\alpha$ 1-gliadinas (PGQQQPFPPQPY/PQPQPF) (Ciccocioppo R., Di Sabatino A., Corazza G.R., "The immune recognition of gluten in celiac disease", Clin. Exp. Immunol. 140:408-16, 2005).

65

Se estima que la EC afecta a aproximadamente 1% de las poblaciones europea y norteamericana. En un estudio de Finlandia se observaron tasas incrementadas (1:47) en personas de edad avanzada (Vilppula A., Kaukinen K.,

Luostarinen L., Krekelä I., Patrikainen H., Valve R., Mäki M., Collin P., "Increasing prevalence and high incidence of celiac disease in elderly people: a population-base study", *BMC Gastroenterol.* 29:9-49, 2009). Sin embargo, muchos estudios indican que la EC está difundida por todo el mundo con valores de prevalencia similares (Barada K., Bitar A., Mokadem M.A., Hashash J.G., Green P., "Celiac disease in Middle Eastern and North African countries: a new burden?", *World J. Gastroenterol.* 16:1449-57, 2010; Dalgic B., Sari S., Basturk B., Ensari A., Egritas O., Bukulmez A., Baris Z., Turkish Celiac Study Group, "Prevalence of celiac disease in healthy Turkish school children", *Am. J. Gastroenterol.* 106:1512-7, 2011; Makharia G.K., Verma A.K., Amarchand R., Bhatnagar S., Das P., Goswami A., Bhatia V., Ahuja V., Datta Gupta S., Anand K., "Prevalence of celiac disease in the northern part of India: a community based study", *J. Gastroenterol. Hepatol.* 26:894-900, 2011; Wang X.Q., Liu W., Xu C.D., Mei H., Gao Y., Peng H.M., Yuan L., Xu J.J., "Celiac disease in children with diarrhea in 4 cities in China", *J. Pediatr. Gastroenterol. Nutr.* 53:368-70, 2011).

Actualmente no se dispone de terapias y el único remedio para la enfermedad es una dieta estrictamente libre de gluten durante toda la vida, necesaria no sólo para prevenir el daño mucosal específico de la EC y los trastornos consecuentes relacionados con la mala absorción (como la anemia deficiente en hierro o la osteoporosis) sino también otras enfermedades autoinmunitarias que se han asociado a la EC, como la diabetes de tipo 1 y la tiroiditis autoinmunitaria, o complicaciones más graves, como los linfomas de células T asociados a enteropatía.

La evitación total del gluten (el umbral seguro de ingesta de gluten generalmente se indica en 50 mg/día, aunque se consideran más seguros 10 mg/día) mantiene la EC en remisión en todos los pacientes excepto en un pequeño porcentaje (2-5%), los cuales sufren una forma no sensible. Sin embargo, dicha dieta resulta muy exigente para el paciente, que se encuentra restringido en sus actividades habituales y sufre de aislamiento social. La utilización de gluten como aditivo en procesos alimentarios es generalizada y es la causa principal de la ingestión involuntaria de gluten, provocando que esta dieta resulte muy difícil de mantener.

Por dichos motivos sería fuertemente bienvenida por los pacientes de EC cualquier alternativa que les permite asumir en su dieta diaria por lo menos cantidades mínimas de gluten.

La utilización de enzimas proteolíticas exógenas para la destoxicación del gluten es una de las estrategias más prometedoras para el tratamiento de la EC. Diferentes maneras de aplicación pueden aprovechar el potencial de estos enzimas: el tratamiento de harinas que contienen gluten, antes o durante la fermentación de la masa, avanzando de esta manera hacia la producción de "nuevos alimentos", así como el consumo concomitante de gluten y enzimas proteolíticos adecuados, que actuaría de esta manera como "complemento alimentario", de manera similar a la utilización de lactasa para la intolerancia a la lactosa. Resulta evidente que se requieren diferentes propiedades enzimáticas para cumplir los diferentes objetivos.

El enfoque enzimático al tratamiento de la EC se basa en la demostración de Shan et al. (*Science*, 2002) de la capacidad de algunos enzimas microbianos de cortar los péptidos de gluten en residuos específicos y eliminar las secuencias del péptido específicamente tóxicas/inmunotóxicas. En particular, demostraron que una endoproteasa específica de protilo exógena obtenida de *Flavobacterium meningosepticum* (FM-PEP) resultaba útil en la digestión de los péptidos gliadina. La adición de una PEP in vitro en presencia de extractos de membrana de borde en cepillo (MBC) o durante la perfusión in vivo de intestino delgado de rata causaba una rápida degradación del péptido 33-mero inmunodominante ("33-mero") y una pérdida de su capacidad de estimular las células T específicas de gliadina (Hausch et al., 2002).

Otros enzimas de la misma familia (EC 3.4.21.26) de otras cepas bacterianas (es decir, *Sphingomonas capsulata* y *Myxococcus xanthus*) han sido evaluadas con dicho objetivo por los mismos autores (Shan L., Marti T., Sollid L.M., Gray G.M., Khosla C., "Comparative biochemical analysis of three bacterial prolyl endopeptidases: implications for coeliac sprue", *Biochem. J.* 383:311-318, 2004). Globalmente, dichos estudios han demostrado diferencias sustanciales entre los tres enzimas con respecto a la longitud de cadena y la especificidad de subsitio y han confirmado el potencial de terapia enzimática oral, aunque han suscitado preocupación con respecto a su posible eficacia in vivo debido a restricciones a la especificidad de sustrato, el pH de la actividad (actividad óptima a pH prácticamente neutro en lugar de pH ácido, según se requiere para actuar en el estómago), el largo tiempo necesario para completar la digestión de los péptidos tóxicos y la resistencia a la degradación por la pepsina. Se ha sugerido en consecuencia una combinación de dos enzimas con actividad gástrica y especificidad de sustrato complementaria (Gass J., Bethune M.T., Siegel M., Spencer A., Khosla C., "Combination enzyme therapy for gastric digestion of dietary gluten in patients with celiac sprue" *Gastroenterology* 133:472-480, 2007): PEP de *S. capsulata* asociado a EP-B2, la isoforma 2 de la endoproteasa B específica de glutamina de *Hordeum vulgare*, una cisteína proteasa derivada de semillas de cebada en germinación que se activa a pH ácido y por la pepsina (Bethune M.T., Strop P., Tang Y., Sollid L.M., Khosla C. "Heterologous expression, purification, refolding, and structural-functional characterization of EP-B2, a self-activating barley cysteine endoprotease", *Chem. Biol.* 13:637-47, 2006), ha demostrado ser una herramienta terapéutica potencialmente más potente. Otro estudio informó de que una PEP derivada de *Aspergillus niger*, que despliega su actividad principal bajo condiciones ácidas en el estómago, puede iniciar la degradación de la gliadina antes de alcanzar la luz intestinal (Stepniak D., Spaenij-Dekking L., Mitea C., Moester M., de Ru A., Baak-Pablo R. van Veelen P., Edens L., Koning F. "Highly efficient gluten degradation with a newly identified prolyl endoprotease: implications for celiac disease", *Am. J. Physiol. Gastrointest. Liver Physiol.*,

291:G621-9, 2006).

Estos resultados han sido el objeto de varios documentos de patente. En el documento nº WO2003/068170 (EP nº 572127), los inventores reivindican que la administración de una dosis eficaz de glutanasa en un paciente celíaco o con dermatitis herpetiforme reduce los niveles de oligopéptidos del gluten tóxicos, atenuando o eliminando de esta manera los efectos dañinos del gluten. Se proporciona un respaldo adicional a este enfoque en el documento nº WO2005/107786 (EP nº 1740197), en el que se dan a conocer formulaciones farmacéuticas de enzimas glutenasa para la utilización en el tratamiento de los pacientes celíacos o con dermatitis herpetiforme. El documento nº WO2005/027953 (EP nº 16663298) describe un tratamiento con una nueva endoproteasa específica de prolilo de *Aspergillus niger* (AN-PEP) que resulta de utilidad en la digestión de los péptidos del gluten tóxicos. El documento nº WO2005/019251 proporciona leucina aminopeptidasa (LAP) de dos especies fúngicas diferentes, *Trichophyton rubrum* y *Aspergillus fumigatus*, en combinación con dipeptidil-peptidasa IV (DppIV). Estos enzimas han sido evaluados para el corte del 33-mero bajo condiciones de pH neutro ya que la actividad óptima de los LAP se ha estimado en aproximadamente 7,0, con un intervalo de actividad entre pH 6 y 8, impidiendo o limitando de esta manera una posible degradación de la gliadina en el líquido gástrico.

También se ha demostrado que las tridipeptidil peptidasas de *A. fumigatus* pueden degradar proteínas a pH ácido (Reichard U., Lechenne B., Asif A.R., Streit F., Grouzmann E., Jousson O., Monod M. "Sedolisins, a new class of secreted proteases from *Aspergillus fumigatus* with endoprotease or tripeptidyl-peptidase activity at acidic pHs", Appl. Environ. Microb. 72:1739-48, 2006). En el documento nº WO2011/077359 se proporciona un kit compuesto de una prolil-proteasa (AfuS28) y por lo menos una tripeptidil proteasa perteneciente a la familia de la sedolisina destinada a la utilización como complemento alimenticio útil también para el tratamiento de la EC.

Resulta evidente que el valor terapéutico de un enzima o composición de enzimas en el tratamiento de la EC se relaciona con que los enzimas: a) sean resistentes a la degradación por otros enzimas gastrointestinales, b) resulten eficientes en el medio en el que el 33-mero y los demás péptidos tóxicos son producidos, y c) muestren una actividad proteolítica rápida y elevada hacia los péptidos del gluten. Estos enzimas deberían mantenerse activos a pH ácido y deberían poder acceder a una composición compleja de gluten dificultada por otros componentes de los alimentos normales, horneados o cocinados.

### Sumario de la invención

Los solicitantes en la presente invención han identificado una nueva familia de enzimas y proporcionan un enzima o una composición de dicha familia de enzimas que presenta propiedades catalíticas únicas.

Los solicitantes han identificado un *Actinoallomurus* sp. como productor de dicha familia de enzimas. Las endopeptidasas de la invención se producen mediante el cultivo de *Actinoallomurus* en un medio de cultivo apropiado, por ejemplo que contenga soja. La cepa, originalmente identificada con el código de solicitante A8, ha sido aislada a partir de un suelo italiano y depositada el 24 de junio de 2011 en el DSMZ (Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH; ahora Leibniz-Institut DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH), Inhoffenstrasse 7b, D-38124 Braunschweig, Alemania, bajo las disposiciones del Tratado de Budapest. La cepa recibió el número de acceso DSM 24988.

Las endopeptidasas de la presente invención son capaces de hidrolizar determinados oligopéptidos del gluten que son resistentes al corte por enzimas gástricos y pancreáticos y cuya presencia en la luz intestinal resulta en efectos tóxicos. El tratamiento enzimático puede eliminar dichos péptidos y los efectos tóxicos de los mismos; puede degradar la gliadina y de esta manera destoxificar el gluten sin ninguna adición de PEP. Las peptidasas de la invención presentan actividades en un amplio abanico de pH y son capaces de ejercer actividad proteolítica entre pH 3 y pH 8. Dicha actividad se define como actividad de glutenasa.

Dichos enzimas glutenasa proporcionan métodos para evitar los síntomas de la celiaquía y/o la dermatitis herpetiforme mediante la reducción de los niveles de oligopéptidos del gluten tóxicos en los alimentos, antes o después de la ingestión por el paciente. Estos enzimas pueden utilizarse conjuntamente con otros enzimas proteolíticos, tales como proteinasas y aminopeptidasas para producir eficazmente hidrolizados de proteínas utilizados para alimentos y bebidas, y medicinas.

Dichos enzimas son secretados por diferentes cepas de *Actinollomurus* mediante el cultivo de los mismos en medios adecuados y su actividad puede evaluarse mediante la utilización de tanto sustratos cromogénicos como el análisis zimográfico.

Además, los enzimas de la invención pueden producirse en las células hospedadoras mediante la introducción de ácidos nucleicos codificantes de dichos enzimas, el cultivo de las células en un medio de cultivo bajo condiciones adecuadas para producir y recuperar los enzimas.

**Descripción de los dibujos**

Figura 1: árbol filogenético inferido a partir de las secuencias de las glutenasas de la invención y las proteasas sedolisina o cumamolisisina.

5 Las secuencias proteicas se alinearon con el programa CLUSTALW y se calculó el árbol de máxima probabilidad con valores bootstrap utilizando el software MEGA5. Las endopeptidasas de *Actinoallomurus* sp. se muestran en negrita. Se utilizó SedE como grupo externo. A continuación, se visualizó el árbol con el software MEGA5. Los valores de bootstrap iguales o superiores al 50% se indican en los nodos. La barra de la escala indica el número de sustituciones de aminoácido por cada sitio.

10 SedE=sedolisina E de *A. fumigatus* Af293 (EAL86850); SedB=sedolisina B de *Aspergillus fumigatus* Af293 (CAE17674); TPPa=tripeptidil peptidasa A de *A. oryzae* RIB40 (BAC56232); SedD=sedolisina D de *A. fumigatus* Af293 (CAE17675); SedC=sedolisina C de *A. fumigatus* Af293 (CAE46473); SedA=sedolisina A de *A. fumigatus* Af293 (CAE51075); KumaA=cumamolisisina A de *Allicyclobacillus sendaiensis* (Q8GB88); KumaB=cumamolisisina de *Bacillus* sp. MN-32 (Q8RR56); sedolisina A=sedolisina de *Pseudomonas* sp. 101 (P42790); sedolisina B=sedolisina B de *Xanthomonas* sp. T22 (Q60106); Endopep- son las glutenasas de la invención. Los números de acceso de secuencia de proteína se indican entre paréntesis.

20 Figura 2: perfil de actividad de endopep-140 (>100 kDa) y endopep-40 (<100 kDa) medido en el sustrato fluorescente Succinil-Ala-Ala-Pro-Phe-AMC bajo diversas condiciones de pH. Se midió la fluorescencia tras 2 h a 37°C. Se proporcionan los valores de pH específicos en el eje X; el porcentaje de actividad relativa se proporciona en el eje Y.

25 Figura 3: actividad de endopeptidasas de endopep-140 (>100 kDa) y endopep-40 (<100 kDa) medida sobre el sustrato fluorescente Succinil-Ala-Ala-Pro-Phe-AMC a pH 3 y a pH 5. La fluorescencia se midió en diversos intervalos de tiempo a 37°C. Los tiempos de incubación se proporcionan en el eje X; las unidades arbitrarias de fluorescencia (ufr), en el eje Y.

30 Figura 4: SDS-PAGE al 8% teñido con azul de Coomassie de las glutenasas recombinantes.

Se indujo la expresión de proteínas con IPTG 0,2  $\mu$ M a 22°C durante la noche. S1 indica la cepa de *E. coli* transformada con pET28b-endopep-140; S3 indica la cepa de *E. coli* transformada con pET28b-endopep-40. Los pesos moleculares del marcador utilizado se indican a la izquierda.

35 A la izquierda: carril 1=marcador de peso molecular; carril 2=S1 no inducido; carril 3=S1 inducido, recolección a las 3 h; carril 4=S1 inducido, recolección de durante la noche (d.n.); carril 5=blanco; carril 6=S3 no inducido; carril 7=S3 inducido, recolección de las 3 h; carril 8=S3 inducido, recolección de d.n.

40 Figura 5: degradación del péptido gliadina 33-mero por la endopep-140.

Los productos de reacción del curso temporal se separaron mediante HPLC de fase inversa.

45 Figura 5A: registros de EM, recuentos iónico  $m/z=300-2000$ ; figura 5B: perfil de UV a 230 nm;  $\mu$ UA son microunidades de absorbancia. Se comparan los perfiles obtenidos en los tiempos 0, 30 min. y 2 h. Únicamente se encontraba presente en el tiempo 0 una señal, correspondiente al péptido 33-mero intacto ( $[M-2H]=1957$ ). Esta señal se redujo efectivamente con el tiempo. Aparecieron varios picos tras la incubación de 30 min. y su cantidad se incrementó en el tiempo 2 h. Todos los iones moleculares observados se informan en la Tabla 2. El tamaño del pico más característico se subraya en la figura.

50 Figura 6: degradación del péptido gliadina 33-mero por la endopep-140 en presencia de pepsina (1 mg/ml).

Los productos de reacción del curso temporal se separaron mediante HPLC de fase inversa.

55 Figura 6A: registros de EM, recuentos iónicos  $m/z=300-2000$ ; figura 6B: perfil de UV a 230 nm,  $\mu$ UA son las microunidades de absorbancia. Se compararon los perfiles obtenidos en los tiempos 0, 30 min. y 2 h. Aunque sólo se encontraba presente en el tiempo 0 una señal, correspondiente al péptido 33-mero intacto ( $[M-2H]=1957$ ), aparecieron varios picos tras una incubación de 30 min. y su cantidad se incrementó en el tiempo 2 h. Todos los iones moleculares observados se indican en la Tabla 3. El tamaño de los picos más característicos se subraya en la figura. La señal en ( $[M-H]=1068$ ) corresponde a un péptido de nueve aminoácidos presente en la secuencia del 33-mero. La señal 689,7 se debe a la pepsina.

60 Figura 7: análisis de HPLC-EM. El péptido gliadina 33-mero se incubó con pepsina. figura 7A: registros de EM; figura 7B: perfil de UV a 230 nm,  $\mu$ UA son microunidades de absorbancia. Se compararon los perfiles obtenidos en los tiempos 0 y 2 h. No se observó prácticamente ninguna degradación del 33-mero.

65

Figura 8: proteólisis de gliadina, SDS-PAGE teñido con azul de Coomassie.

Digestión de la gliadina por dos glutenasas diferentes en presencia o en ausencia de pepsina. Se incubó la gliadina durante 2 h a 37°C.

A la izquierda: carril 1=marcador de peso molecular; carril 2=gliadina; carril 3=gliadina+endopep-40 (<100 kDa); carril 4=gliadina+endopep-40 (<100 kDa) + pepsina; carril 5=reacción 3 detenida tras una incubación de 10 min.; carril 6=gliadina+endopep-140 (>100 kDa); carril 7=gliadina+endopep-140 (>100 kDa)+pepsina.

### Breve descripción de las tablas

Tabla 1. Proteínas putativas purificadas a partir del secretoma de *Actinoallomurus*.

Tabla 2. Péptidos liberados por la digestión del 33-mero con endopep-140.

Tabla 3. Péptidos liberados por la digestión del 33-mero con endopep-40.

### Descripción detallada de la invención

Los actinomicetos son bacterias Gram-positivas filamentosas conocidas principalmente por su capacidad para producir metabolitos bioactivos secundarios. Los actinomicetos han sido utilizados como fuente de hidrolasas, aunque sólo unas cuantas, particularmente las alcalifílicas, han sido exploradas hasta el momento para su potencial enzimático (Endo A., Murakawa S., Shimizu H., Shiraishi Y. "Purification and properties of collagenase from a *Streptomyces* species", J. Biochem. 102:163-70, 1987; Sakurai Y., Inoue H., Nishii W., Takahashi T., Iino Y., Yamamoto M., Takahashi K., "Purification and Characterization of a Major Collagenase from *Streptomyces parvulus*", Biosci. Biotechnol. Biochem. 73:21-28, 2009; Mehta V.J., Thumar J.T., Singh S.P., "Production of alkaline protease from an alkaliphilic actinomycete" Bioresource Technology 97:1650-1654, 2006).

Además, la secuencia genómica completa de *Streptomyces coelicolor* A3(2) predice la presencia de 60 proteasas y peptidasas putativas secretadas de entre 819 proteínas secretadas potenciales (Bentley S.D., Chater K.F., Cerdeño-Tárraga A.M., Challis G.L., Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H., Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M., Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S., Huang C.H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S., Rabbinowitsch E., Rajandream M.A., Rutherford K., Rutter S., Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K., Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J., Hopwood D.A., "Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2)", Nature 417:141-7, 2002) proporcionando a los actinomicetos valor adicional como fuente potencial de nuevas enzimas hidrolíticas. Entre los actinomicetos, las cepas acidofílicas ofrecen una oportunidad más elevada de producir enzimas con propiedades mejor adaptadas a los requisitos para resultar eficaces en la EC, es decir actividad a pH ácido. Recientemente se han descrito por primera vez varios nuevos géneros de actinomicetos acidofílicos (es decir, *Catenulispora*, *Actinospica*, *Rugosimonospora* y *Streptacidiphilus*), así como bacterias filamentosas con propiedades acidofílicas pertenecientes a un linaje bacteriano previamente desconocido (es decir, *Ktedonobacter*) (Busti E., Cavaletti L., Monciardini P., Schumann P., Rohde M., Sosio M., Donadio S., "Catenulispora acidiphila gen. nov., sp. nov., a novel, mycelium-forming actinomycete, and proposal of *Catenulisporaceae* fam. nov.", Int. J. Syst. Evol. Bacteriol. 56:1741-1746, 2006; Cavaletti L., Monciardini P., Bamonte R., Schumann P., Rohde M., Sosio M., Donadio S., "New lineage of filamentous, spore-forming, gram-positive bacteria from soil" Appl. Env. Microbiol. 72:4360-4369, 2006; Cavaletti L., Monciardini P., Schumann P., Rohde M., Bamonte R., Busti E., Sosio M., Donadio S., "Actinospica acidiphila gen. nov., sp. nov. and *Actinospica robiniae* gen. nov., sp. nov.; proposal for *Actinospicaceae* fam. nov. and *Catenulisporinae* subordo. nov. in the order *Actinomycetales*", Int. J. Syst. Evol. Bacteriol. 56:1747-1753, 2006; Monciardini P., Cavaletti L., Ranghetti A., Schumann P., Rohde M., Bamonte R., Sosio M., Mezzelani A., Donadio S., "Novel members of the family *Micromonosporaceae*, *Rugosimonospora acidiphila* gen. nov., sp. nov. and *Rugosimonospora africana* sp. nov" Int. J. Syst. Evol. Bacteriol. 59:2752-2758, 2009; Kim S.B., Lonsdale J., Seong C.M., Goodfellow M., "Streptacidiphilus gen. nov., acidophilic actinomycetes with wall chemotype I and emendation of the family *Streptomycetaceae* (Waksman and Henrici (1943)<sup>AL</sup>) emend. Rainey et al. 1997)", Antonie van Leeuwenhoek 83:107-116, 2003).

*Actinoallomurus*, al igual que muchos otros actinomicetos, puede crecer en un medio que contiene proteínas como única fuente de nitrógeno y de carbono. Esta capacidad de crecer en un medio proteico depende de la acción sinérgica de las endo- y exo-proteasas secretadas, ya que sólo los aminoácidos y péptidos cortos pueden ser asimilados mediante transportadores membranales. Las cepas de *Actinoallomurus* presentan un crecimiento óptimo a pH ligeramente ácido, sugiriendo de esta manera que los enzimas proteolíticos pueden expresarse a este pH y pueden ser capaces de digerir proteínas complejas bajo condiciones ácidas. Según el conocimiento de los solicitantes, ningún otro informe describe la producción de proteasas que actúan a pH ácido procedentes de actinomicetos.

La cepa *Actinoallomurus* sp. DSM 24988 fue aislada por los solicitantes a partir de un suelo italiano, se almacenó en



sedolisina) producida por *Actinoallomurus* que digiere los polipéptidos de longitud completa y degrada un fragmento de gliadina que es conocido que es resistente a la acción de las proteasas. Por lo tanto puede utilizarse por lo menos una endoproteasa para el tratamiento de la enfermedad celíaca o cualquier enfermedad del proceso digestivo, tal como la mala absorción. Además, debido a que el enzima es resistente a la pepsina, una combinación de dichas dos actividad proteolíticas podría resultar en una degradación más extensa de los polipéptidos y proteínas, tales como la gliadina. Los solicitantes han descubierto que el 33-mero resulta degradado en péptidos de seis aminoácidos (Tablas 2 y 3). Aunque los enzimas de la invención son activos por sí solos, la adición a los mismos de una o más peptidasas que actúan sobre los péptidos ricos en prolina, tales como las proliil-endoproteasas y/o las x-proliil-dipeptidil aminopeptidasas y/o proliil-aminopeptidasas puede resultar en una acción más rápida.

La familia de proteínas S8/S53 de la presente invención sola u opcionalmente en combinación con otras proteasas puede resultar útil en la industria alimentaria, tal como, aunque sin limitación, para degradar sustratos para amargor, el tratamiento de carnes, la industria de los jabones, o para degradar priones, virus y proteínas tóxicas o contaminantes en péptidos cortos y/o aminoácidos libres.

De esta manera, la presente invención proporciona una composición enzimática, que comprende por lo menos una endopeptidasa S8/S53 activa a un pH de entre 3 y 8 (ambos inclusive), seleccionada de entre el grupo que consiste de:

- a) endopep-140 que comprende la SEC ID nº 1, un fragmento biológicamente activo de la misma, una variante alélica natural de la misma, o una secuencia que presenta una identidad de secuencias de por lo menos 70%, 80%, 90% o 95%,
- b) endopep-40 que comprende la SEC ID nº 2, un fragmento biológicamente activo de la misma, una variante alélica natural de la misma, o una secuencia que presenta una identidad de por lo menos 70%, 80%, 90% o 95%,
- c) endopep-120 que comprende la SEC ID nº 3, un fragmento biológicamente activo de la misma, una variante alélica natural de la misma, o una secuencia que presenta una identidad de por lo menos 70%, 80%, 90% o 95%,
- d) endopep-60 que comprende la SEC ID nº 4, un fragmento biológicamente activo de la misma, una variante alélica natural de la misma, o una secuencia que presenta una identidad de por lo menos 70%, 80%, 90% o 95%,
- e) ndopep-41 que comprende la SEC ID nº 5, un fragmento biológicamente activo de la misma, una variante alélica natural de la misma, o una secuencia que presenta una identidad de por lo menos 70%, 80%, 90% o 95%.

La expresión "peptidasa o peptidasas de la invención", "proteasa o proteasas de la invención" o "endopeptidasa o endopeptidasas de la invención", tal como se utilizan en la presente memoria identifica una o más de las endopeptidasas definidas anteriormente.

La presente invención incluye además una endopeptidasa derivada de cualquiera de las secuencias indicadas anteriormente, en la que cualquiera de los aminoácidos puede modificarse respecto de los residuos correspondientes mostrados en SEC ID nº 1, 2, 3, 4 o 5, manteniendo todavía su actividad biológica y funciones fisiológicas, o un fragmento biológicamente activo de la misma. El objetivo de la presente invención incluye además cualquier variante de peptidasa que contiene sustituciones, deleciones, modificaciones de cadena lateral y/o inserciones en determinadas posiciones dentro de la secuencia de aminoácidos de la secuencia de aminoácidos nativa que conserva la actividad biológica y la función fisiológica de dicha secuencia nativa de aminoácidos. Los ejemplos de sustituciones son bien conocidos para el experto en la materia y se indican en, por ejemplo, el documento nº WO2011/077359.

En otro aspecto, la presente invención se refiere a proteasas aisladas de la invención y a fragmentos biológicamente activos de las mismas (o derivados, partes, análogos o homólogos de las mismas). La expresión fragmento biológicamente activo se refiere a regiones de las proteasas de la invención, los cuales resultan necesarios para actividades de proteasa específicas.

En una forma de realización adicional, la proteasa de la invención es una proteasa que comprende una secuencia de aminoácidos que presenta una identidad de por lo menos 70%, todavía más preferentemente 80%, todavía más preferentemente 90% y todavía más preferentemente 95% de la identidad respecto a la secuencia de aminoácidos que comprende la SEC ID nº 1, 2, 3, 4 o 5 y conserva la actividad de las proteasas que comprenden las SEC ID nº 1, 2, 3, 4 o 5.

El término "identidad" y "homología" en referencia a una secuencia de nucleótidos o de aminoácidos se utilizan intercambiamente en la presente memoria y se refieren al grado en que dos secuencias polinucleotídicas o



polipeptídicas son idénticas u homólogas residuo a residuo en una región particular de comparación. La alineación y el porcentaje de identidad u homología pueden determinarse utilizando cualquier programa informático adecuado conocido de la técnica, por ejemplo los descritos en Current Protocols in Molecular Biology (Ausubel F.M. et al., "Commercially Available Software", Current Protocols in Molecular Biology, complemento 30, sección 7.7.18, Tabla 7.7.1, 1987). Entre los programas preferentes se incluyen el programa GCG Pileup, FASTA (Pearson R. y Lipman D.J., "Improved Tools for Biological Sequence Analysis", Proc. Natl. Acad. Sci. USA 85:2444-2448, 1988) y BLAST (Altschul S.F., Gish W., Miller W., Myers E.W., Lipman D.J., "Basic local alignment search tool", J. Mol. Biol. 215:403-410, 1990).

La invención proporciona además proteasas de la invención fusionadas operativamente con otro polipéptido que el experto en la materia denomina proteínas quiméricas o etiquetadas. El polipéptido puede fusionarse con el extremo N-terminal y/o C-terminal de la proteasa de la invención. El experto en la materia puede preparar y producir diferentes proteínas de fusión etiquetadas mediante técnicas de ADN recombinante estándares o técnicas convencionales, incluyendo sintetizadores automáticos de ADN. Dichas proteínas de fusión pueden facilitar la purificación de la proteasa recombinante de la invención o su expresión y secreción en las células huésped. Estas técnicas son bien conocidas por el experto en la materia.

Los solicitantes han demostrado que los péptidos grandes, tales como el péptido gliadina 33-mero, pueden ser degradados a pH ácido por las endopeptidasas. La endopep-140 secretada actúa cortando la cadena polipeptídica en varios puntos, generando diferentes péptidos pequeños, de seis aminoácidos (Tabla 2, figura 5). Esta endopeptidasa muestra una preferencia para sitios que presentan tirosina o leucina o fenilalanina en la posición P1 y prolina en la posición P2 y P1'. El incremento de la cantidad de endopeptidasa y/o el tiempo de incubación conduce a una degradación completa del polipéptido 33-mero. Según los pesos moleculares detectados, aparentemente los residuos glutamina se transforman en residuos glutámicos, sugiriendo de esta manera la implicación de una actividad amidotransferasa. Las endopeptidasas de la composición enzimática de la invención pueden funcionar en presencia de pepsina de mamífero (Tabla 2, figura 6).

Se llevaron a cabo los mismos análisis con endopep-40 y los resultados obtenidos, mostrados en la Tabla 3, indican que esta glutenasa se comportaba de manera similar.

La endopep-140 contiene un dominio estructural con elevada homología con otras endopeptidasas de la invención. Por ejemplo, la alineación de BLAST proporciona identidades=223/370 (60%), positivos=261/370 (71%) entre endopep-140 y endopep-40 o identidades=227/431 (53%), positivos=273/431 (63%) entre endopep-140 y endopep-41. Se obtuvieron resultados similares mediante la alineación de BLAST entre endopep-40 y endopep-41, con identidades=230/419 (55%), positivos=278/419 (66%). Los positivos son los residuos aminoácidos que son idénticos o muy similares entre sí en sus características fisicoquímicas según se determina con la matriz BLOSUM62.

Las endopeptidasas de la composición enzimática de la invención pueden producirse mediante cultivo de una cepa de *Actinoallomurus* natural o un mutante de la misma que es capaz de producir por lo menos una endopeptidasa S8/S53 definida anteriormente o una célula hospedadora obtenida mediante técnicas de ADN recombinante.

El término "recombinante", utilizado en referencia a una célula, indica que la célula replica un ácido nucleico heterólogo, o expresa un péptido o proteína codificado por un ácido nucleico heterólogo. Los genes no presentes dentro de la forma nativa (no recombinante) de la célula o presentes en la forma nativa de la célula, en la que los genes son modificados y reintroducidos en la célula por medios artificiales, pueden encontrarse contenidos en células recombinantes.

El experto en la materia reconocerá que dichas células pueden ser organismos transgénicos unicelulares o multicelulares.

La invención incluye además métodos basados en el cultivo de células que contienen un ácido nucleico endógeno a la célula que ha sido modificada sin eliminar el ácido nucleico de la célula; entre dichas modificaciones se incluyen las obtenidas mediante sustitución génica, mutación específica de sitio y técnicas relacionadas, por ejemplo mediante tratamiento con agentes mutagénicos o con radiaciones ionizantes.

La expresión "variante alélica" se refiere a cualquiera de entre dos o más formas alternativas de un gen que ocupa el mismo locus cromosómico. La variación alélica aparece naturalmente mediante mutación y puede resultar en polimorfismo fenotípico dentro de las poblaciones. Las mutaciones génicas pueden ser silenciosas (ningún cambio en el polipéptido codificado) o pueden codificar polipéptidos que presentan una secuencia de aminoácidos alterada. La expresión variante alélica se refiere además a una proteína codificada por una variante alélica de un gen.

De esta manera, la presente invención proporciona un método para producir la composición enzimática de la invención que comprende:

A) cultivar una cepa de *Actinoallomurus* natural capaz de producir por lo menos una endopeptidasa S8/S53 anteriormente definida y recuperar la endopeptidasa o endopeptidasas a partir del lote de cultivo, o

5 B) cultivar una cepa de *Actinoallomurus* derivada de una cepa de *Actinoallomurus* natural capaz de producir por lo menos una endopeptidasa S8/S53 anteriormente definida mediante técnicas convencionales de mutación y/o selección, que mantienen la capacidad de producir por lo menos una de la endopeptidasa o endopeptidasas anteriormente definidas y recuperar la endopeptidasa o endopeptidasas a partir del lote de cultivo, o

C) una técnica de ADN recombinante que comprende las etapas de:

10 1) introducir en una célula hospedadora un ácido nucleico codificante de por lo menos una endopeptidasa de la clase S8/S53 subtilisina quexina sedolisina seleccionada de entre el grupo que consiste de:

15 a) endopept-140 que comprende SEC ID nº 1, un fragmento biológicamente activo de la misma, una variante alélica natural de la misma, o una secuencia que presenta una identidad de por lo menos 70%, 80%, 90% o 95%,

20 b) endopept-40 que comprende SEC ID nº 2, un fragmento biológicamente activo de la misma, una variante alélica natural de la misma, o una secuencia que presenta una identidad de por lo menos 70%, 80%, 90% o 95%,

c) endopept-120 que comprende SEC ID nº 3, un fragmento biológicamente activo de la misma, una variante alélica natural de la misma, o una secuencia que presenta una identidad de por lo menos 70%, 80%, 90% o 95%,

25 d) endopept-60 que comprende SEC ID nº 4, un fragmento biológicamente activo de la misma, una variante alélica natural de la misma, o una secuencia que presenta una identidad de por lo menos 70%, 80%, 90% o 95%, y

30 e) endopept-41 que comprende SEC ID nº 5, un fragmento biológicamente activo de la misma, una variante alélica natural de la misma, o una secuencia que presenta una identidad de por lo menos 70%, 80%, 90% o 95%,

35 2) cultivar la célula de la etapa 1) en un medio de cultivo bajo condiciones adecuadas para producir la endopeptidasa o endopeptidasas, y

3) recuperar la endopeptidasa o endopeptidasas a partir del lote de cultivo.

40 Pueden producirse una o más endopeptidasas de la invención al llevar a cabo un método tal como se ha definido anteriormente. En el caso de que se produzcan separadamente endopeptidasas individuales mediante la realización del método anteriormente definido, la presente invención incluye opcionalmente la etapa de combinar dos o más de las endopeptidasas obtenidas para proporcionar una composición enzimática que contiene una mezcla de dichas endopeptidasas. Dichas endopeptidasas pueden obtenerse en forma de endopeptidasas aisladas. Con la expresión "endopeptidasa aislada", tal como se utiliza en la presente memoria, se hace referencia a una forma purificada de la endopeptidasa que se encuentra sustancialmente libre de otras proteínas o material celular de la célula a partir de la

45 que se derivó la endopeptidasa.

50 Todas las expresiones de ácidos nucleicos de ADN/ARN también se refieren a las técnicas de manipulación relacionadas utilizadas en la presente memoria bien conocidas por el experto en biología molecular y hasta ahora la totalidad de dichos términos se utiliza en su amplio y común significado tal como se informa en Maniatis T., Fritsch E.F., Sambrook J., "Molecular cloning: a laboratory manual", Cold Spring Harbor, NY, 1982, o Ausubel F.M., "Current Protocols in Molecular Biology", John Wiley & Sons, New York, NY, 1993, o Sambrook et al., "Molecular Cloning: a Laboratory Manual" 2a ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

55 Los ácidos nucleicos codificantes de las endopeptidasas de la composición enzimática de la invención son un objetivo adicional de la invención que incluye los ácidos nucleicos las secuencias de los cuales se proporcionan en la presente memoria y se definen como SEC ID nº 7, 8, 9, 10 o 11 o fragmentos de las mismas. La invención incluye además ácidos nucleicos mutantes o variantes o parte de dichos ácidos nucleicos.

60 De acuerdo con lo anteriormente expuesto, entre los ácidos nucleicos de la presente invención se incluye una secuencia polinucleotídica mostrada en SEC ID nº 7, 8, 9, 10 o 11, o una secuencia que presenta una identidad de por lo menos 50%, preferentemente de 60%, más preferentemente de por lo menos 70%, todavía más preferentemente de 80%, todavía más preferentemente de 90% y todavía más preferentemente de por lo menos 95% respecto a la secuencia de ácidos nucleicos que comprende SEC ID nº 7, 8, 9, 10 o 11 y codifica una secuencia de aminoácidos que sigue manteniendo la actividad biológica y función fisiológica de dicha secuencia de

65 aminoácidos.

Las secuencias caracterizadas por identidades al nivel de nucleótido se indican en la presente memoria también como "secuencias de ácidos nucleicos homólogas" o variaciones de la misma. Las secuencias de nucleótidos homólogas codifican aquellas secuencias codificantes de isoformas de proteasas de la invención. Las isoformas pueden expresarse en el mismo organismo como resultado de, por ejemplo, el procesamiento alternativo del ARN. Alternativamente, las isoformas pueden estar codificadas por genes diferentes.

Entre las secuencias de nucleótidos homólogas se incluyen además, aunque sin limitación, variaciones alélicas naturales y mutaciones de las secuencias de nucleótidos indicadas en la presente memoria. Entre las secuencias de ácidos nucleicos homólogas se incluyen aquellas secuencias de ácidos nucleicos que codifican las sustituciones de aminoácidos conservadoras en SEC ID nº 1, 2, 3, 4 o 5.

En la invención, entre las secuencias de nucleótidos homólogas pueden incluirse secuencias de nucleótidos codificantes de una proteasa de la invención de otras especies pertenecientes a *Actinoallomurus*, así como de otros géneros diferentes de *Actinoallomurus*, tales como, por ejemplo, *Catenulispora*, *Actinospica*, *Ktedonobacter*, *Streptomyces*, *Streptacidiphilus*, *Micromonospora* y *Rugosimonospora*.

Los solicitantes determinaron mediante análisis de BLAST que los genomas de *Catenulispora*, *Ktedonobacter* y *Streptomyces* contienen genes que codifican proteínas que presentan una homología superior a 50% respecto a la proteasa correspondiente de la invención; por ejemplo, las proteínas putativas con números de acceso ACU72534 y ACU72320 de *Catenulispora acidiphila* presentan identidades=261/400 (65%), positivos=295/400 (74%) con endopep-40 SEC ID nº 2 e identidades=705/1.342 (53%), positivos=878/1342 (65%) con endopep-140 SEC ID nº 1, respectivamente, o la proteína putativa con número de acceso EFH81837 de *Ktedonobacter racemifer* presenta identidades=244/402 (61%), positivos=299/402 (74%) con endopep-40 SEC ID nº 2 o la secuencia de la proteína e14 de *Streptomyces* sp. con número de acceso EFF91270 también presenta identidades=230/371 (62%), positivos=274/371 (74%) con endopep-40 SEC ID nº 2. Los genes codificantes de endopeptidasa homólogos pueden encontrarse presentes en cepas pertenecientes a géneros filogenéticamente relacionados con los anteriormente indicados, las secuencias genómicas de los cuales, no se encuentran disponibles en bases de datos públicas.

Un "fragmento biológicamente activo" de la endoproteasa de la invención puede prepararse mediante aislamiento de un fragmento de ácidos nucleicos de SEC ID nº 7, 8, 9, 10 o 11 que codifica una endoproteasa con la misma actividad biológica que las endoproteasas de la invención, que expresa la parte codificada de endoproteasa (por ejemplo, mediante expresión recombinante in vitro) y evaluación de la actividad del fragmento codificado de endoproteasa. La invención comprende además moléculas de ácidos nucleicos que difieren de las secuencias de ácidos nucleicos mostradas en SEC ID nº 7, 8, 9, 10 o 11 debido a la degeneración del código genético y, de esta manera, codifican las mismas proteasas que se encuentran codificadas por las secuencias de ácidos nucleicos mostradas en SEC ID nº 7, 8, 9, 10 o 11.

Las técnicas de manipulación génica y de expresión de proteínas son conocidas por el experto en la materia y también pueden encontrarse en Kriegler M., "Gene Transfer and Expression: A Laboratory Manual", Stockton Press, NY, 1990.

Los solicitantes indican con la expresión "actividad biológica" o "actividad funcional", la función natural o normal de las proteasas de la invención, por ejemplo la capacidad de degradar otras proteínas. Los residuos aminoácidos que se encuentran conservados entre las proteasas de la invención se predice que resultarán particularmente difíciles de alterar. Los aminoácidos para los que pueden realizarse sustituciones conservadoras son bien conocidas de la técnica. Tal como reconoce el experto en la materia, cada codón en un ácido nucleico (aparte de AUG, que habitualmente es el único codón para la metionina) puede modificarse y rendir una molécula funcionalmente idéntica mediante técnicas estándares. Además, las sustituciones, deleciones o adiciones individuales que alteran, añadir o delecionan un único aminoácido o un porcentaje reducido de aminoácidos (típicamente menos de 5%, más típicamente menos de 1%) en una secuencia codificada son "mutaciones conservadoras", en el que las alteraciones resultan en la sustitución de un aminoácido por un aminoácido químicamente similar.

Otro aspecto de la invención se refiere a moléculas de ácidos nucleicos codificantes de las proteasas de la invención que contienen cambios en residuos aminoácidos que no resultan esenciales para la actividad. Dichas proteasas de la invención difieren en secuencia de aminoácidos de SEC ID nº 1, 2, 3, 4 o 5, aunque conservan actividad biológica. En una forma de realización, la molécula de ácidos nucleicos aislada comprende una secuencia de nucleótidos codificante de una proteasa, en la que la proteasa comprende una secuencia de aminoácidos con una identidad de por lo menos aproximadamente 50% respecto a las secuencias de aminoácidos de SEC ID nº 1, 2, 3, 4 o 5.

La expresión "ácido nucleico aislado" o "secuencia polinucleotídica aislada", tal como se utiliza en la presente memoria, identifica una molécula de ácidos nucleicos que se separa de otras moléculas de ácidos nucleicos que se encuentran presentes en la célula a partir de la cual se obtiene el ácido nucleico y se encuentra sustancialmente libre de otro material celular o material del medio de cultivo en el caso de que dicha molécula de ácidos nucleicos se obtenga de un microorganismo natural, un microorganismo derivado del mismo o de microorganismos obtenidas por medios recombinantes.

Una molécula de ácidos nucleicos aislada codificante de una endoproteasa de la invención homóloga respecto a la proteína de SEC ID nº 1, 2, 3, 4 o 5 puede crearse mediante la introducción de una o más sustituciones, adiciones o deleciones de nucleótidos en la secuencia de ácidos nucleicos de SEC ID nº 7, 8, 9, 10 o 11, de manera que se introducen una o más sustituciones, adiciones o deleciones de aminoácidos en la proteasa codificada. Pueden introducirse mutaciones en SEC ID nº 7, 8, 9, 10 o 11, mediante técnicas estándares, tales como la mutagénesis dirigida a sitio, la mutagénesis mediada por PCR y la transposición del ADN. Alternativamente, pueden introducirse mutaciones aleatoriamente a lo largo de la totalidad o una parte de una secuencia codificante de la proteasa de la invención, tal como mediante mutagénesis por saturación, y los mutantes resultantes pueden cribarse para actividad biológica de la proteasa de la invención con el fin de identificar los mutantes que conservan actividad.

La composición de enzima glutenasa según la presente invención también puede utilizarse en forma inmovilizada y utilizarse, por ejemplo, para el tratamiento de productos alimentarios líquidos. La composición enzimática de la invención también puede utilizarse en la industria de las frutas y cervecera para la limpieza y mantenimiento de equipos. Por ejemplo, un producto alimentario líquido que contiene gluten se deja que fluya a lo largo de una matriz permeable para el gluten en el que se incluye la composición enzimática de la invención. Se extrae el gluten del producto alimentario y se digiere mediante la acción de los enzimas. La composición enzimática de la invención también puede contribuir a la energía disponible del alimento. Por ejemplo, una proteína que comprende prolina parcial o totalmente digerible se degrada total o parcialmente con la composición enzimática de la invención, resultando en la disponibilidad de más alimento digerible para el ser humano o animal. Por lo tanto, se mejora la tasa de crecimiento y/o la tasa de conversión de alimento (es decir, el peso de alimento ingerido respecto a la ganancia de peso) del ser humano o animal.

#### Métodos de producción de las endopeptidasas y caracterización de las mismas

La presente invención se refiere además a métodos para producir los enzimas de la presente invención que comprenden cultivar (a) una cepa, que se encuentra en su forma de tipo salvaje, o (b) una forma derivada de la misma mediante técnicas comunes de mutación, por ejemplo mediante tratamiento con agentes mutagénicos o con radiaciones ionizantes, o (c) una célula hospedadora que es capaz de producir el polipéptido de la invención y recuperar dicho polipéptido a partir del lote de cultivo. En los métodos de producción de la presente invención, los medios de cultivo celular se seleccionan diferentemente entre los conocidos de la técnica de producción de proteínas según el cultivo de cepas de tipo salvaje, cepas derivadas de las mismas, o células huésped recombinantes. El cultivo tiene lugar en un medio nutritivo adecuado que comprende fuentes de carbono y nitrógeno y sales inorgánicas, utilizando procedimientos conocidos de la técnica. Los medios adecuados se encuentran disponibles de proveedores comerciales o pueden prepararse según composiciones publicadas (por ejemplo en catálogos de la American Type Culture Collection). Por ejemplo, un medio que contiene soja es más adecuado para permitir la producción de enzimas de la invención por microorganismos de tipo salvaje.

Las células pueden cultivarse mediante cultivo en matraz de agitación, fermentación a pequeña escala o a gran escala (incluyendo fermentaciones continuas, por lotes, por lotes alimentados o en estado sólido) en fermentadores de laboratorio o industriales realizada en un medio adecuado y bajo condiciones que permitan que el polipéptido se exprese y/o se aisle. En el caso de que el polipéptido se secrete al medio nutritivo, el polipéptido puede recuperarse directamente del medio. En el caso de que no se secrete el polipéptido, puede recuperarse a partir de lisados celulares. Se producen varias actividades proteolíticas y el cultivo se suspende al alcanzar un máximo la producción de endopeptidasa. Tras completar el cultivo, se separa el medio de cultivo mediante filtración para separar los cuerpos microbianos, y el filtrado se procesa de la manera habitual para la recolección de endopeptidasa mediante varios procedimientos en combinación, tales como la ultrafiltración, la concentración bajo presión reducida, la precipitación en solución hipersalina ("salting out"), la precipitación con solvente orgánico, la diálisis, la filtración en gel, la cromatografía de adsorción, la cromatografía de intercambio iónico, el electroenfoque y la liofilización. Los procedimientos adecuados deberían seleccionarse considerando las propiedades físicas y químicas deseadas de las endopeptidasas.

Un ejemplo representativo de producción de endopeptidasas de la presente invención mediante el cultivo de una cepa de *Actinoallomurus* de tipo salvaje se proporciona en el Ejemplo 1, posteriormente en la presente memoria. La endopeptidasa o endopeptidasas de la invención pueden purificarse en el grado deseado de pureza que puede depender del uso pretendido y de la actividad específica de la endopeptidasa o endopeptidasas.

#### Expresión heteróloga en células hospedadoras

Pueden utilizarse células recombinantes y microorganismos para producir las endopeptidasas de la presente invención. La célula hospedadora puede ser cualquiera de las células huésped que resultarán familiares para el experto en la materia, incluyendo células procarióticas, células eucarióticas, células de mamífero, células de insecto, células fúngicas, células de levadura y/o células vegetales. La selección de un huésped apropiado se encuentra comprendido dentro de los conocimientos del experto en la materia.

Los microorganismos útiles son células bacterianas, tales como las bacterias Gram-positivas, incluyendo, aunque sin limitación, una célula de *Bacillus* (por ejemplo *Bacillus subtilis* y *Bacillus cereus*) o una célula de *Streptomyces*, o

5 células de bacterias del ácido láctico, o bacterias Gram-negativas, tales como *E. coli* y *Pseudomonas*. Entre las células productoras preferentes se incluyen las de organismos que es conocido que se consideran generalmente como seguras, tales como los procariotas *Lactobacillus*, *Pyrococcus*, *Bacillus*, *Streptomyces* y el eucariota *Aspergillus*. Entre las células más preferentes se incluyen las que ya se utilizan en la preparación de alimentos, tales como *Lactobacillus* spp. y/o *Aspergillus oryzae*.

La producción extracelular de los enzimas puede conseguirse de microorganismos tales como *Aspergillus oryzae*, *Lactobacillus casei*, *Kluyveromyces lactis* o *Streptomyces lividans*.

10 La introducción de un vector en una célula hospedadora bacteriana puede llevarse a cabo, por ejemplo, mediante transformación de protoplasto (por ejemplo Chang S., Cohen S.N., "High frequency transformation of *Bacillus subtilis* protoplasts by plasmid DNA", Mol. Gen. Genet. 168:111-115, 1979), utilizando células competentes (por ejemplo Dubnau D., Davidoff-Abelson R., "Fate of transforming DNA following uptake by competent *Bacillus subtilis*. I. Formation and properties of the donor-recipient complex", J. Mol. Biol. 56:209-221, 1971), electroporación (por ejemplo Shigekawa K., Dower W.J., "Electroporation of eukaryotes and prokaryotes: a general approach to the introduction of macromolecules into cells", Biotechniques 6:742-751, 1988) o la conjugación (por ejemplo Koehler T.M., Thorne C.B., "*Bacillus subtilis* (natto) plasmid pLS20 mediates interspecies plasmid transfer", J. Bacteriol. 169:5771-5278, 1987).

20 Son organismos multicelulares útiles para la utilización como células huésped, por ejemplo, plantas, partes de plantas, semillas o células vegetales. Los organismos multicelulares productores preferentes son, por ejemplo, cultivos alimentarios transgénicos, tales como cereales o vegetales, tales como el tomate, que contienen los ácidos nucleicos codificantes de las proteasas de la invención.

25 Un ejemplo representativo de producción de endopeptidasas de la presente invención en las células huésped recombinantes se proporciona en el Ejemplo 2, posteriormente en la presente memoria.

#### Formulaciones farmacéuticas

30 Los agentes endopeptidasas de la presente invención pueden administrarse solos o incorporados en una diversidad de formulaciones. Una composición farmacéutica de la invención se formula para que resulte compatible con su vía deseada de administración, que preferentemente es la administración oral. Por ejemplo, debido a que algunas de las endopeptidasas de la composición enzimática de la invención se secretan, puede administrarse por vía oral una preparación en bruto obtenida a partir de medio de cultivo celular de *Actinoallomurus* u otros microorganismos recombinantes unicelulares útiles, tales como bacterias Gram-positivas, incluyendo, aunque sin limitación, una célula de *Bacillus*, una célula de *Streptomyces*, células de bacterias del ácido láctico y bacterias Gram-negativas, tales como *E. coli*. Entre las bacterias del ácido láctico se incluyen, aunque sin limitación, especies de los géneros *Lactococcus*, *Lactobacillus*, *Leuconostoc*, *Streptococcus*, *Pediococcus* y *Enterococcus*. Alternativamente, la composición enzimática de la invención puede administrarse por vía oral después de la purificación.

40 Dichas formulaciones pueden prepararse con los ingredientes apropiados para generar una preparación en forma líquida, por ejemplo en forma de una solución, emulsión o en forma sólida, tal como tabletas o cápsulas, o semisólida. La formulación de la composición enzimática puede administrarse de una diversidad de maneras, incluyendo aquellas particularmente útiles para mezclar con un alimento. Los componentes enzimáticos puede estar activos antes o durante la ingestión y pueden tratarse, por ejemplo, mediante un encapsulado adecuado, para controlar la temporización de la actividad.

50 Con el fin de preparar una composición farmacéutica apropiada de las endopeptidasas de la presente invención, puede utilizarse cualquier método de estabilización de material químico o biológico conocido de la técnica, que comprende aquellos basados en la irradiación o la modulación de la temperatura o combinaciones de los mismos.

55 Para tratar la enfermedad celíaca, las composiciones farmacéuticas utilizadas preferentemente se formulan de manera que liberen su actividad en el líquido gástrico. Este tipo de formulaciones proporcionará una actividad óptima en el lugar correcto, por ejemplo la liberación de las endoproteasas de la invención en el estómago.

Alternativamente puede administrarse en el paciente, un microorganismo, tal como un cultivo bacteriano o de levadura, capaz de producir los agentes activos. Dicho cultivo puede mezclarse con preparaciones alimentarias o formularse, por ejemplo, como cápsula entérica.

60 La presente invención proporciona además un complemento alimenticio que comprende la composición enzimática de la presente invención. La expresión "complemento alimentario" en el contexto de la presente invención es intercambiable con las expresiones aditivo alimentario, un complemento dietético y un complemento nutricional.

65 El complemento alimenticio de la invención puede formularse, prepararse, suministrarse y dispensarse tal como se ha indicado e otros documentos anteriores referentes al campo de la presente invención (documentos WO2011/077359, WO2003/068170, WO2005/107786) que proporcionan métodos de tratamiento y/o prevención de

un síndrome asociado a una enfermedad humana, seleccionando dicha enfermedad de entre el grupo que comprende la enfermedad celíaca, la dermatitis herpetiforme, la mala absorción en el tracto digestivo, reacciones alérgicas, deficiencias enzimáticas, infecciones fúngicas, la enfermedad de Crohn, micosis y el esprúe.

5 A título de ejemplo, el complemento alimenticio de la invención puede ser un producto recubierto o no con un enzima granulado que puede mezclarse fácilmente con componentes alimentarios, alternativamente, algunos complementos alimentarios de la invención pueden formar un componente de una premezcla. Alternativamente, los complementos alimentarios de la invención puede ser un líquido estabilizado, o una suspensión acuosa o de base aceitosa. La composición enzimática de la invención puede suministrarse mediante la expresión de los enzimas directamente en cultivos de alimentos transgénicos (tal como, por ejemplo, plantas transgénicas, semillas y similares), tales como granos, cereales, maíz, soja, semillas de colza y altramuces.

15 La composición farmacéutica o el complemento alimenticio de la invención puede proporcionarse antes de las comidas, inmediatamente antes de las comidas, con las comidas o inmediatamente después de las comidas, de manera que las endoproteasas de la composición enzimática de la presente invención sean liberadas o activadas en la luz gastrointestinal superior, en donde las endoproteasas pueden complementar los enzimas gástricos y pancreáticos para detoxificar el gluten ingerido y evitar que los péptidos dañinos pasen a la capa de enterocitos.

20 La composición enzimática de la presente invención presenta numerosas aplicaciones en la industria del procesamiento alimentario, en particular pueden utilizarse en la preparación de complementos alimentarios tal como se ha indicado en los documentos de la técnica anterior indicados anteriormente.

25 A partir de la presente descripción resulta evidente que un objetivo adicional de la presente invención consiste en proporcionar un método para degradar oligopéptidos del gluten que son resistentes al corte por enzimas gástricos y pancreáticos y cuya presencia en la luz interna resulta en efectos tóxicos, que comprende poner en contacto dichos oligopéptidos del gluten con una composición enzimática o por lo menos una endopeptidasa aislada de la presente invención.

30 En particular, un aspecto de dicho método consiste en el tratamiento o la prevención de la enfermedad celíaca, la dermatitis herpetiforme y/o cualquier otro trastorno asociado a la intolerancia del gluten, que comprende administrar en el paciente que lo necesita, una cantidad eficaz de una composición enzimática o por lo menos una endopeptidasa aislada de la presente invención, preferentemente incorporada en una formulación farmacéutica, complemento alimenticio o bebida.

## 35 Ejemplos

### Ejemplo 1: identificación y caracterización de endopeptidasas de *Actinoallomurus* sp. DSM 24988

#### 40 1.1 Cultivo de la cepa y producción de proteínas

La cepa de *Actinoallomurus* utilizada según la presente invención derivada de la colección de cepas de los solicitantes.

45 Se mantuvo *Actinoallomurus* sp. DSM 24988 en medio agar ISP2 (Shirling y Gottlieb, 1966) acidificado a pH 5,5 con HCl. El contenido microbiano de una placa se raspó y se inoculó en un matraz Erlenmeyer de 50 ml que contenía 15 ml de medio AF5 que está compuesto de: (g/l) dextrosa 20, extracto de levadura 2, harina de soja 8, NaCl 1 y MES 10. El medio se preparó en agua destilada y se ajustó el pH a 5,5 antes de esterilizar a 121°C durante 20 min. El matraz inoculado se cultivó a 28°C, en un matraz rotatorio funcionando a 200 rpm. Tras 5 a 6 días de incubación, se inoculó 5% del cultivo en una segunda serie de matraces Erlenmeyer de 500 ml que contenían 100 ml del mismo medio de fermentación. Se llevó a cabo la producción de proteínas en matraces incubados durante 15 días a 28°C en un agitador rotatorio funcionando a 200 rpm. Se realizó un seguimiento de la producción de la proteína mediante bioensayo tal como se indica posteriormente.

55 Se produjeron varias actividades proteolíticas y se suspendió el cultivo al alcanzar un máximo la producción de endopeptidasas, seguido de los ensayos indicados posteriormente. La fermentación se recolectó tras 15 días y se centrifugó el caldo a 4.000 rpm durante 15 minutos. Se separó el medio de cultivo mediante filtración para separar los cuerpos microbianos y se procesó el filtrado. Se seleccionaron procedimientos adecuados considerando las propiedades físicas y químicas deseadas de las endopeptidasas.

#### 60 1.2 Determinación de las actividades proteolíticas

Se midieron las actividades enzimáticas a diferentes pH en tampón de acetato (acetato amónico-AMC, concentración final de 50 mM, pH 4,0 a 8,0; ácido acético 20 mM, pH 3,0) a 37°C en un volumen total de 0,2 ml.

65 Se determinó la actividad enzimática de endopeptidasa mediante la medición de la hidrólisis de Succinil-Ala-Ala-Pro-Phe-AMC (Amino-Metil-Coumarina) como sustrato (Bachem AG, Hauptstrasse 144, 4416 Bubendorf, Suiza).

Se prepararon soluciones madre de sustrato a una concentración de 10 mM y se almacenaron a -20°C. Se preparó el sustrato a partir de 10 a 20 µl de solución de metanol al 60% que contenía Succinil-Ala-Ala-Pro-Phe-AMC 2 mM y 50 a 150 µl de solución tampón de acetato. La mezcla de reacción contenía una concentración de 0,2 mM de sustrato y la preparación enzimática (entre 10 y 40 µl o 1,0 µg en cada ensayo) en 200 µl de tampón de acetato a diferentes valores de pH.

Al sustrato, que había sido previamente calentado a 37°C durante 10 minutos, se añadieron 10 a 40 µl de la solución enzimática y la reacción se llevó a cabo a 37°C durante hasta 2 horas. Se midió la AMC liberada con un lector de microplacas Fusion (Perkin Elmer Italia SpA, Monza, Italia) a una longitud de onda de excitación de 360 nm y una longitud de onda de emisión de 460 nm. A título de control, el sustrato se incubó sin enzima.

La actividad enzimática para liberar 1 mmol de AMC en 1 minuto se define como 1 Unidad.

### 1.3 Purificación de endopeptidasa

Se separó el micelio del medio de cultivo mediante filtración a través de papel (Miracloth, de Calbiochem, USA). A continuación, se centrifugaron 200 ml de sobrenadante durante 10 minutos a 5.000 rpm para eliminar los residuos; después se centrifugó adicionalmente el sobrenadante a 10.000 rpm durante 20 min. a 4°C para separar la fracción insoluble. Se precipitaron las proteínas del sobrenadante con etanol (1:4 v/v) o saturación de 20%-75% de sulfato amónico. Se resuspendió el pellet y se dializó, al aparecer se concentró utilizando un Centricon Plus-70 con un valor de corte de 5 kDa (Millipore, Bedford, Mass., USA). La presencia de proteasas capaz de hidrolizar la Succinil-Ala-Ala-Pro-Phe-AMC se subrayó al someterla a ensayo a pH 5. Las alícuotas de la suspensión de proteínas tras la diálisis se sometieron a ebullición para someterlas a análisis de EM con inyección directa ("shotgun"). El procedimiento de ebullición resulta necesario para evitar la autodigestión. Se detectaron por lo menos 5 proteínas pertenecientes a la familia de proteasas S8/S53.

Las proteínas resuspendidas se cromatografiaron en resina de intercambio iónico Amberlite IRA900 (Alfa Cesar GmbH, Karlsruhe, 76185 Alemania) o celulosa DE-52 (Whatman Inter. Ltd., Maidstone, Inglaterra) u otra matriz; la fracción obtenida se sometió a ensayo para la hidrólisis de Succinil-Ala-Ala-Pro-Phe-AMC y las fracciones activas agrupadas se sometieron a análisis de SDS-PAGE 1D y zimográfico.

Las proteínas se fraccionaron adicionalmente mediante filtración de exclusión por tamaño en una escala de cortes de peso molecular (300 kDa, 100 kDa, 50 kDa, 30 kDa, 10 kDa (Nanosep Pall, Michigan, USA; Vivaspin Sartorius GmbH, 37070 Goettingen, Alemania). Se sometió a ensayo su actividad tal como se indica. Se obtuvieron fracciones activas con diferente peso molecular. Se sometieron a ebullición alícuotas de dichas fracciones y se analizaron mediante EM con inyección directa.

Se separaron las proteínas mediante electroforesis en un gel homogéneo al 10% o disco de gel al 4-15% o de poliacrilamida para cualquier kd (Bio-Rad, Hercules, 9640, CA, USA) seguido de tinción con azul brillante de Coomassie R-250 (Bio-Rad) o plata (Sigma, Aldrich, USA) o mediante tinción de actividad enzimática (zimografía) realizada mediante incubación del gel a 37°C con sustrato fluorescente Succinil-Ala-Ala-Pro-Phe-AMC 50-100 µM al pH deseado.

Se purificaron dos endopeptidasas pertenecientes a la familia S8/S53 hasta que el análisis electroforético mostrase la presencia de una única banda detectada tanto mediante tinción de plata como zimografía.

La primera endopeptidasa se obtuvo tras la filtración de peso molecular: se retuvo con un valor de corte de >100 kDa. El análisis de EM de inyección directa de la alícuota sometida a ebullición correspondiente reveló la presencia de endopep-140, SEC ID nº 1 (Tabla 1A), con un peso molecular de 142.449,4 y un pl teórico de 5,18; se detectaron trazas de las dos otras proteínas. Esta fracción de proteína resultó activa en el intervalo de pH de 4 a 8, con un pH óptimo de 5 (figura 2). Al someterla a ensayo a pH 3, el enzima mostró 95% de su actividad a pH 5 tras 30 minutos de incubación a 37°C (figura 3). Esta fracción de proteína resultó más eficiente para la digestión de 33-mero que para la hidrólisis de Succinil-Ala-Ala-Pro-Phe-AMC, tal como se muestra en el Ejemplo 3.

La segunda endopeptidasa, indicada como endopep-40, SEC ID nº 2, se obtuvo en la fracción de proteína de valor de corte de <100 kDa (Tabla 1B), con un peso molecular de 39.908,4 y un pl teórico de 5,94. Esta proteína resultó ser activa en el intervalo de pH de 4 a 6 con un óptimo de pH de 5. Al someterla a ensayo a pH 3, el enzima mostró 85% de su actividad a pH 5 tras 30 minutos de incubación a 37°C (figura 3). La banda fluorescente destacada mediante zimografía mostraba un peso molecular de aproximadamente 40 kDa. Se caracterizó para su actividad enzimática tanto la proteína activa eluida de la banda de gel como del retenido correspondiente de corte de peso molecular. El análisis de EM de inyección directa reveló la presencia de otras proteínas; las señales que indicaban la presencia de secuencias de endopep-140 podrían deberse a polipéptidos de menor tamaño (<100 kDa) formadas por su degradación.

Las otras tres endopeptidasas (indicadas como endopep-120, SEC ID nº 3; endopep-60, SEC ID nº 4; endopep-41,

SEC ID nº 5) no se purificaron hasta la homogeneidad. Su presencia se detectó mediante EM de inyección directa en el secretoma y en las fracciones activas de las primeras etapas de purificación. El análisis de EM de inyección directa permitió identificar su secuencia de aminoácidos, informada en la presente memoria. Los solicitantes denominaron las proteínas según su peso molecular inferido. De esta manera, endopep-120 presenta un peso molecular de 112.012,8 y un pI teórico de 6,75; endopep-60 presenta un peso molecular de 59.646,0 y un pI teórico de 6,02, y endopep-41 presenta un peso molecular de 41.646,1 y un pI teórico de 5,22.

#### 1.4 Análisis de EM de proteínas: experimentos de EM de inyección directa

Actualmente las metodologías proteómicas resultan de importancia primordial para los estudios de caracterización de biomarcadores o proteínas impulsados por resultados.

El enfoque proteómico clásico se basa en la electroforesis en gel bidimensional (2DG), en la que se aíslan puntos de proteína de interés y se identifican mediante espectrometría de masas (EM).

El enfoque 2DG presenta una resolución relativamente alta, que se encuentra limitada, sin embargo, por la dificultad para detectar determinadas clases de proteínas. Entre ellas se incluyen proteínas membranales debido a su baja solubilidad en el tampón de la electroforesis en gel, proteínas con un peso molecular bajo (<10 kDa) o alto (>200 kDa), así como aquellas con un punto isoelectrico extremo (pI<4 o >9). Una limitación adicional de este enfoque reside en la dificultad del análisis de proteínas menos representadas y en el hecho de que resulta tedioso y laborioso.

Estos problemas se resuelven utilizando una nueva metodología proteómica basada en cromatografía capilar bidimensional acoplada con espectrometría de masas en tándem (2DC-EM/EM), también denominada MudPIT (Multidimensional Protein Identification Technology) (Washburn, M.P., Wolters D., Yates J.R. 3rd, "Large-scale analysis of the yeast proteome by multidimensional protein identification technology", Nat. Biotechnol. 19:242-7, 2001). Implica la generación de péptidos a partir de la digestión enzimática de una mezcla compleja de proteínas, su separación mediante dos columnas de micro-HPLC y el análisis directo de los picos eluidos mediante EM/EM. La 2DC-EM/EM combina la separación mediante intercambio iónico y de fase inversa de mezclas de péptidos obtenidas de la digestión total de proteínas totales (o prefraccionadas). En particular, en primer lugar se separan mezclas de péptidos mediante cromatografía de intercambio iónico (columna SCX, 5 µm, 0,3 DI x 150 mm) utilizando siete etapas de concentración creciente de cloruro amónico (0 a 1.000 mM). Cada etapa salina se carga directamente en la columna de fase inversa (C<sub>18</sub>, 0,180 DI x 100 mm) y se separa con un gradiente de acetonitrilo: eluyente A, ácido fórmico al 0,1% en agua; eluyente B, ácido fórmico al 0,1% en acetonitrilo. Los péptidos eluidos de la columna C<sub>18</sub> se analizan directamente con un espectrómetro de masas con trampa de iones; el límite de detección es de aproximadamente 10 fmoles o menos. Los espectros se adquieren en modo positivo (típicamente en el intervalo de 400 a 1.600 m/z) utilizando exclusión dinámica para el análisis de EM/EM. Se utilizan diferentes proteasas para digerir el extracto de proteínas de las muestras (tripsina, pepsina o proteinasa K).

La identificación de las proteínas correspondientes se obtiene seguidamente mediante una búsqueda automatizada en la base de datos con software adecuado, tal como el algoritmo SEQUEST (Eng. J.K., McKormack A.L., Yates J.R., "An Approach to Correlate Tandem Mass Spectra Data of Peptides with Amino Aminoacid Sequences in a Protein Database", J. Am. Soc. Mass Spectrom. 5:976-984, 1994) para la manipulación de los datos de espectros de masas. Los espectros de masas experimentales se correlacionan con las secuencias de los péptidos obtenidos mediante comparación con los espectros de masas teóricos en la base de datos de proteínas descargados del NCBI ([www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)) u otros sitios de internet.

#### 1.5 Criterios para la identificación de proteínas

Las identificaciones de péptidos se aceptaron en el caso de que pudiesen establecerse con una probabilidad superior a 90,0% tal como especifica el algoritmo Peptide Prophet (Keller A., Nesvizhskii A.I., Kolker E., Aebersold R., "Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search", Anal. Chem. 15:5383-92, 2002). Se aceptaron las identificaciones de péptidos en el caso de que pudiesen establecerse con una probabilidad superior a 95,0% y contuviesen por lo menos un péptido identificado. Debido a que no se encontró una identidad superior a 50% en las bases de datos públicas de proteínas, se secuenció el genoma de *Actinoallomurus* y se tradujo en los 6 marcos de lectura posibles. La alineación posterior con BLAST de los péptidos detectados frente a las proteínas putativas de *Actinoallomurus* identificó las proteínas homólogas conocidas.

Se identificó putativamente un total de 94 proteínas en el secretoma de *Actinoallomurus*. Las enzimas constituían una fracción significativa de ellas, se detectaron 17 enzimas proteolíticas, y también se encontraron glucosidasas, lipasas, fosfatasa ácida y diesterasas. No pudo asignarse ninguna función mediante análisis BLAST a 25 secuencias.

Entre las proteasas, pudieron detectarse cinco proteínas pertenecientes a la familia de la peptidasa S8/S53, dos de las cuales se encontraban entre las proteínas más representadas.



Se llevaron a cabo los mismos análisis en todas las etapas de purificación posteriores hasta que sólo se detectase unas cuantas secuencias.

5 Los datos obtenidos mediante el análisis de EM de las fracciones activas sometidas a ebullición mostraron pocas secuencias de proteínas, aunque se encontraba presente una única banda en el gel de SDS-PAGE teñido con plata. Tal como se muestra en la Tabla 1, se detectó endopep-140 (SEC ID nº 1) y endopep-40 (SEC ID nº 2) como las proteínas responsables de las dos actividades indicadas. La presencia de endopep-140 en la muestra identificada como <100 kDa posiblemente se debe a la degradación parcial de la proteína en polipéptidos de menor tamaño.

10 El análisis *in silico* indicó que tanto endopep-140 como endopep-40 eran proteínas glucosiladas que contenían un péptido de señal que conducía a la secreción, sugiriendo que pueden producirse en forma de preproenzimas.

15 El análisis *in silico* agrupó la totalidad de las cinco endopeps en la familia de S8/S53 e indicó que todas eran proteínas glucosiladas. Se encontró una homología elevada entre endopep-140, endopep-40 y endopep-41 (SEC ID nº 5), permitiendo su agrupamiento en un subgrupo tal como se muestra con el árbol filogenético de la figura 1. Se detectó una secuencia de péptido de señal también para endopep-60 (SEC ID nº 4), mientras que se encontraba ausente en endopep-120 (SEC ID nº 3) y endopep-41.

20 El árbol filogenético mostrado en la figura 1 subraya la novedad de estas glutenasas S8/S53, que no se agrupan con el enzima cumamolisinina ni con sedolisina ni con tripeptidil-peptidasa (sedA a sedD) dados a conocer en el documento nº WO2011/077359, que pertenecen a la familia de S53.

25 La novedad de dichos enzimas recibe respaldo adicional de la especificidad de sustrato, tal como se muestra en el Ejemplo 3. El patrón de degradación del 33-mero obtenido mediante digestión con endopep-140 o endopep-40 mostraba una actividad endoproteolítica.

## Ejemplo 2: producción de endopeptidasa en células huésped recombinantes

### 30 2.1 Cepas y plásmidos

En el presente estudio se estudió *Actinoallomus* sp. DSM 24988. Todos los experimentos de subclonación de plásmidos se llevaron a cabo en *E. coli* DH10B (Invitrogen, Carlsbad, CA) utilizando el plásmido pTZ57R/T (Fermentas, UAB, Lituania).

35 Se utilizó *Escherichia coli* BL21 (DE3) Star (Novagen Italia, Podenzano, PC) para producir peptidasas heterólogas (recombinantes).

### 40 2.2 Producción de proteasas recombinantes

Se produjeron proteasas de *Actinoallomurus* recombinantes y se purificaron a partir de *Escherichia coli* BL21 (DE3) Star como el sistema de expresión.

45 Para construir cepas de *E. coli* que producían secuencias de nucleótidos de los genes de Endopep-140 y Endopep-40 se amplificaron utilizando los juegos de cebadores siguientes:

Endopep-140	5'-AAAAAGCTTCAGCTACAGGTGTGGTCCGG-3'	SEC ID nº 12
Rendopep-140	5'-AAAAAAACATATGCCCGATCTTCCCACCC-3'	SEC ID nº 13
Endopep-40	5'-AAAAAGCTTCAGAAGGCTCCGGTGCC-3'	SEC ID nº 14
50 Rendopep-40	5'-AAAAAAACATATGTCACGACGCGTGACCG-3'	SEC ID nº 15

Se clonaron los productos de PCR en el vector pTZ57R/T y se secuenciaron antes de la clonación en el vector de expresión con el fin de verificar que no se había introducido ninguna mutación durante las ampliaciones por PCR.

55 A continuación, se clonaron los genes de endopep en los sitios *NdeI-HindIII* del plásmido pET28b (Novagen) y se transformaron en el huésped de expresión BL21(DE3) Star.

60 Las células transformadas se cultivaron en cultivos de 50 ml de medio LB que contenía 30 µg/ml de canamicina a 37°C hasta alcanzar una DO600 de entre 0,6 y 1. Se indujo la expresión de las glutenasas con la adición de isopropil-β-D-tiogalactósido 0,2 µM (Sigma) y los cultivos se incubaron adicionalmente a 22°C durante la noche. Tal como se muestra en la figura 4, se obtuvieron bandas de proteínas correspondientes al peso molecular deseado de endopep-140 y endopep-40.

### 65 2.3 Purificación de una o más endopeptidasas producidas heterológamente

Las células que expresaban las endopeptidasas recombinantes se centrifugaron a 5.000 g durante 20 minutos. El

pellet se resuspendió en 8 ml de solución tampón (fosfato sódico 20 mM, NaCl 500 mM, pH 7,8) y después se añadió lisozima 1 mg/ml para lisar las células completamente.

Se purificaron proteínas etiquetadas con His a partir de los extractos celulares en bruto mediante cromatografía de afinidad para Ni<sup>2+</sup> inmovilizada utilizando el sistema de purificación de Ni-NTA (Invitrogen) siguiendo las instrucciones del fabricante. Se hicieron migrar alícuotas de cinco microlitros de las fracciones eluidas a través de SDS-PAGE y se tiñeron con azul de Coomassie o tinción de plata con el fin de verificar la presencia y el grado de pureza de las endopeptidasas expresadas. La figura 4 muestra los niveles de expresión obtenidos para endopep-140 y endopep-40 tras la inducción en comparación con la cepa de control no inducida.

La actividad enzimática de los extractos totales obtenidos de las cepas de *E. coli* transformadas con pET28b vacío y pET28b-endopep-140 o pET28b-endopep-40, respectivamente, se sometió a ensayo con Succinil-Ala-Ala-Pro-Phe-AMC como sustrato. SE detectó la actividad enzimática en los extractos a partir de cepas pET28b-endopep, mientras que no se observó actividad con el control negativo.

### **Ejemplo 3: actividad biológica de endopeptidasa.**

#### 3.1 Degradación de péptido tóxico 33-mero de gliadina a pH ácido

Una solución de péptido inmunotóxico 33-mero de gliadina (50 µM) se incubó a 37°C durante hasta 2 horas en presencia de 4 µU de endopep-140 o 2 µU de endopep-40 en presencia o en ausencia de pepsina 1 mg/ml. La reacción se llevó a cabo en ácido acético 20 mM, pH 3, volumen total: 300 µl. Se realizó un seguimiento de la reacción en diferentes tiempos entre 0 y 120 min. (t0, t3, t15, t30, t60 y t120 min.) a 37°C. Se realizó un seguimiento de la desaparición del péptido 33-mero y de la aparición de los productos de degradación mediante análisis de HPLC-EM. Se obtuvieron alícuotas de 50 µl y se detuvo la actividad enzimática con 50 µl de H<sub>2</sub>O:CH<sub>3</sub>CN 50:50 (+ácido fórmico al 0,1%). Las muestras se sometieron a HPLC-EM, se analizaron en un espectrómetro de masas LTQ-XL (Thermo Fisher Scientific, San Jose, California, USA). Los perfiles de HPLC-EM obtenidos con endopep-140, endopep-40 y pepsina, y pepsina sola se muestran en las figuras 5, 6 y 7, respectivamente. La desaparición del péptido 33-mero resultó evidente tras 2 h de incubación.

Todos los péptidos detectados tras la incubación del péptido 33-mero con 4 µU de endopep-140 o 2 µU de endopep-40 se informan en las Tablas 2 y 3, respectivamente.

Tras 2 horas de incubación, el pico más intenso observado presentaba un estado de carga de 1 y correspondía a [MH]<sup>+</sup> de 748,4, indicando la presencia de péptidos pequeños. Esta señal también fue la primera en aparecer. Basándose en la secuencia del 33-mero, cuatro péptidos (1-6, 8-13, 15-20 y 22-27) correspondían a dicha masa molecular, estando todos compuestos de 6 residuos. Según el análisis de EM del 33-mero digerido, se encuentran presentes residuos glutámicos en lugar de residuos de glutamina en la secuencia del péptido hidrolizado, sugiriendo de esta manera que se ha producido una desamidación del sustrato.

El análisis mostró que tras 2 horas de incubación del 33-mero con pepsina prácticamente no se observaba hidrólisis (figura 7 y Tablas 2 y 3), concordando con los datos de la literatura.

El incremento de la cantidad de endopep-140 hasta 8 µU o de endopep-40 hasta 9 µU resultó en la desaparición completa del 33-mero tras 2 h a pH 3 (no mostrado).

Incluso en presencia de pepsina se observó el mismo patrón de degradación, aunque la cantidad de péptido intacto tras 2 h era más elevado en su ausencia, sugiriendo que las endopeptidasas no resultan destruidas por la pepsina.

#### 3.2 Análisis de HPLC/EM

Se llevó a cabo un análisis de HPLC utilizando un espectrómetro de masas con trampa de iones. Se utilizó un LTQ-XL o Advantage acoplado con bomba Accela o Surveyor (ThermoFisher Scientific, San Jose, CA, USA) para caracterizar los hidrolizados de la proteína enzimática producidos por la mezcla de enzimas. Los péptidos se separaron utilizando un ThermoFisher Scientific Hypersil gold o una columna simétrica C18 (Waters, Milford, Mass., USA) en combinación con un gradiente de ácido fórmico al 0,1% en agua Milli Q (Millipore, Bedford, Mass., USA) (solución A) y ácido fórmico al 0,1% en acetónitrilo (solución B) para la elución. El gradiente se inició con 95% de solución A y se incrementó hasta 60% de solución B. Se obtuvo información detallada de los péptidos individuales mediante la utilización del algoritmo de EM/EM "dependiente de escaneo", que es un algoritmo característico de un espectrómetro de masas con trampa de iones. Se realizó un seguimiento del análisis de la banda completa de escaneo mediante análisis "zoom scan". Se utilizó el 33-mero (M=3910) para ajustar para la sensibilidad óptima en el modo EM y para la fragmentación óptima en el modo EM/EM, realizando una infusión constante de 60 µg/ml, resultando en especies principalmente de doble carga y de triple carga en modo EM, y una energía de colisión óptima de aproximadamente 35% en el modo EM/EM.

**Ejemplo 4: actividad biológica de la endopeptidasa**4.1 Degradación de la gliadina y mezcla hidrolizada de gluten

- 5 Debido a sus características estructurales específicas, las prolil oligopeptidasas no pueden digerir los péptidos grandes, habitualmente de más de 30 aminoácidos. Además, la tripeptidil-proteasa sedolisina no puede hidrolizar la gliadina sin adición de prolil proteasa. Dicha limitación es una desventaja evidente para un enzima, que se pretende que hidrolice tan rápida y eficientemente como resulte posible todos los potenciales péptidos ricos en prolina tóxicos.
- 10 La endopep-140 y la endopep-40 de *Actinoallomurus* se incubaron con gliadinas para comprobar si las endoproteasas de *Actinoallomurus* eran capaces de digerir las proteínas completas. Los productos de hidrólisis formados se analizaron mediante SDS-PAGE.

15 Se preparó una solución de gliadina mediante disolución de gliadina en polvo (Sigma G3375) en etanol al 70% a 70°C a una concentración de 10 mg/ml (20 µM). Las moléculas de gliadina intactas consisten de una mezcla de proteínas de 40 a 55 kDa. Se incubó gliadina a 37°C con 10 µl (5 µU) de las endoproteasas de *Actinoallomurus* en ácido acético 20 mM, pH 3, en un volumen final de 0,3 ml. Las reacciones se llevaron a cabo en un Thermomixer (Eppendorf AG, Hamburg, Alemania) tanto en ausencia como en presencia de pepsina a una concentración de 1 mg/ml (figura 8). Se realizó un seguimiento de las reacciones en diferentes intervalos de tiempo; se extrajeron muestras de 30 µl de la mezcla de incubación en los tiempos 0 min., 30 min., 1 y 2 h y se mantuvieron a 4°C hasta la SDS-PAGE. Todos los materiales utilizados para la SDS-PAGE y la tinción se obtuvieron de Bio-Rad (Bio-Rad Laboratories, Inc., Hercules, CA, USA). Se prepararon muestras utilizando tampón para muestras siguiendo las instrucciones del fabricante y se separaron en geles Bis-Tris para cualquier kd utilizando el sistema tampón SDS siguiendo las instrucciones del fabricante. La tinción se llevó a cabo utilizando Coomassie R250 o plata.

25 Tal como puede observarse en la figura 8, la gliadina es cortada por las endopeptidasas derivadas de *Actinoallomurus* en péptidos pequeños hasta una digestión prácticamente completa en 2 horas. La degradación del producto puede observarse como una reducción de la intensidad de la banda en el gel de SDS. El presente experimento también muestra que la pepsina no afecta a la eficacia de la digestión.

30 Tabla 1. Proteínas putativas purificadas a partir del secretoma de *Actinoallomurus*.

35 Se filtraron fracciones purificadas de *Actinoallomurus* en un filtro de 100 kDa de corte de peso molecular y se sometieron a análisis de EM con inyección directa. Los péptidos identificados se alinearon frente a proteínas inferidas mediante traducción del genoma de *Actinoallomurus* (Id\_proteína: código interno para la secuencia putativa de proteína) y el posterior análisis *in silico* mediante BLAST puso de manifiesto las proteínas conocidas homólogas (número de acceso de BLAST: código de secuencia; nota en BLAST: nombre de proteína; n.r.: ningún resultado). Los números de los espectros concordantes proporcionan una medida semicuantitativa de las cantidades de proteína, indicadas como frecuencia (F\_media).

40 A: datos obtenidos de la fracción >100 kDa; B: datos obtenidos de la fracción <100 kDa.

A

Id_proteína	Número de acceso de BLAST	Nota de BLAST	F_media	Notas
Seq_48773	YP_003835151.1	Peptidasa S8/S53	101,9	Endopep-140
Seq_100531	dbj BAJ32040.1	Lipasa putativa	20,2	
Seq_293141	n.r.	n.r.	10,14	

B

Id_proteína	Número de acceso de BLAST	Nota de BLAST	F_media	Notas
Seq_44766	ZP_06921971.1	Proteína secretada	40,14	
Seq_72108	YP_003114375.1	Peptidasa S8/S53	30,25	Endopep-40
Seq_48773	YP_003835151.1	Peptidasa S8/S53	20,15	Endopep-140
Seq_293141	n.r.	n.r.	10,15	

45 Tabla 2. Péptidos liberados por la digestión de 33-mero con endopep-140.

50 Todos los pesos moleculares detectados mediante HPLC-EM tras la digestión del péptido gliadina 33-mero (50 µM) con endopep-140 4 µU en presencia y en ausencia de pepsina (1 mg/ml). El control se obtuvo mediante incubación del 33-mero con pepsina sola. Incubación a 37°C, ácido acético 20 mM, pH 3.



ES 2 595 500 T3

	1956,64			
T <sub>30</sub> + pepsin	1088,73	1 [MH] <sup>+</sup>	25-33	
	1197,45	1 [MH] <sup>+</sup>	1-10	
	748,45	1 [MH] <sup>+</sup>	1-6; 8-13; 15-20; 22-27	
	1306,27; 1956,64	3 [M-3H] <sup>+</sup> ; 2 [M-2H] <sup>+</sup>	1-33	70% residual
T <sub>2h</sub>	1088,73	1 [MH] <sup>+</sup>	25-33	
	1197,45	1 [MH] <sup>+</sup>	1-10	
	748,45	1 [MH] <sup>+</sup>	1-6; 8-13; 15-20; 22-27	
	1306,27; 1956,64	3[M-3H] <sup>+</sup> ; 2 [M-2H] <sup>+</sup>	1-33	20% residual
T <sub>2h</sub> + pepsin	1088,8	1 [MH] <sup>+</sup>	25-33	
	1197,45	1 [MH] <sup>+</sup>	1-10	
	748,45	1 [MH] <sup>+</sup>	1-6; 8-13; 15-20; 22-27	
	1306,27; 1956,64	3 [M-3H] <sup>+</sup> ; 2 [M-2H] <sup>+</sup>	1-33	30% residual
Control T <sub>2h</sub>	1306,27; 1956,64	3 [M-3H] <sup>+</sup> ; 2 [M-2H] <sup>+</sup>	1-33	95% residual

**Listado de secuencias**

- 5 <110> Fondazione Istituto Insubrico di Ricerca per la Vita (FIIRV)
- <120> Nuevas proteasas que pueden hidrolizar péptidos y proteínas del gluten a pH ácido del actinomiceto *Actinoallomurus*
- 10 <130> G72535
- <160> 15
- <170> PatentIn versión 3.5
- 15 <210> 1
- <211> 1390
- <212> PRT
- <213> *Actinoallomurus*
- 20 <220>
- <221> PROPEP
- <222> (1)..(1390)
- <400> 1

ES 2 595 500 T3

Met Pro Asp Leu Pro Thr Gln Arg Arg Ser Pro Arg Arg Ser Arg Arg  
 1 5 10 15

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

Leu Ala Leu Ala Val Thr Thr Leu Thr Leu Ile Ser Met Pro Ser Ala  
 35 40 45

Asn Ala Ala Pro Ala Lys Pro Val Val Pro His Ala Lys Ala Ser Ser  
 50 55 60

Pro Thr Ala Gln Gly Gln His Pro Thr Ser Arg Val Cys Arg Thr Pro  
 65 70 75 80

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

Ile Gly Pro Lys Gly Val Gln Ala Asn Asp Ala Pro Ala Gly Phe Gly  
 100 105 110

Pro Ala Asp Leu Leu Asn Ala Tyr Asn Leu Pro Ser Ala Gly Ser Thr  
 115 120 125

Glu Thr Val Ala Ile Val Asp Ala Phe Asp Asn Pro Asn Ala Glu Ala  
 130 135 140

Asp Leu Gly Val Tyr Arg Gln Gln Tyr Gly Leu Pro Ala Cys Thr Thr  
 145 150 155 160

ES 2 595 500 T3

Ala Asn Gly Cys Phe Lys Lys Ile Asp Gln Arg Gly Gly Thr Ser Tyr  
 165 170 175

Pro Pro Pro Asp Ala Gly Trp Ala Gly Glu Ile Ala Leu Asp Ile Asp  
 180 185 190

Met Val Ser Ala Ile Cys Pro Thr Cys Lys Ile Leu Leu Val Glu Ala  
 195 200 205

Asp Asp Asn Tyr Met Asp Asn Leu Gly Thr Ala Val Asn Gln Ala Val  
 210 215 220

Ser Gln Gly Ala Lys Phe Val Ser Asn Ser Tyr Gly Gly Ser Glu Gly  
 225 230 235 240

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

Ala Ile Thr Ala Ser Ser Gly Asp Ser Gly Tyr Gly Ala Ser Tyr Pro  
 260 265 270

Ala Ala Ser Pro Tyr Val Thr Ser Val Gly Gly Thr Ser Leu Ala Lys  
 275 280 285

Asp Thr Ser Ser Ser Arg Gly Trp Thr Glu Ser Thr Trp Asn Gly Ala  
 290 295 300

Gly Ser Gly Cys Ser Ala Tyr Glu Thr Lys Pro Ser Phe Gln Lys Asp  
 305 310 315 320

Thr Gly Cys Ala Arg Arg Thr Ile Ala Asp Val Ser Ala Val Ala Asp  
 325 330 335

Pro Asn Thr Gly Val Ala Val Tyr Asn Gly Gly Trp His Val Tyr Gly  
 340 345 350

Gly Thr Ser Val Ser Ala Pro Ile Ile Ala Ser Val Tyr Ala Leu Gly  
 355 360 365

Gly Ala Pro Ala Ala Gly Ser Ala Pro Asn Ser Phe Pro Tyr Asp His  
 370 375 380

Pro Ala Asp Leu Asn Asp Val Thr Ser Gly Ser Asn Gly Ser Cys Ser  
 385 390 395 400

Pro Ala Tyr Leu Cys Thr Ala Gly Thr Gly Tyr Asp Gly Pro Thr Gly

ES 2 595 500 T3

					405					410						415
Leu	Gly	Thr	Pro	Asn	Gly	Ile	Ala	Ala	Phe	Arg	Ser	Gly	Pro	His	Ala	
			420					425					430			
Val	Val	Thr	Gly	Thr	Val	Thr	Asp	Gly	Thr	Ala	Pro	Leu	Ala	Ser	Ala	
		435					440					445				
Gln	Val	Lys	Ala	Gly	Asp	Val	Thr	Ala	Thr	Thr	Asp	Gly	Gln	Gly	His	
	450					455					460					
Tyr	Thr	Leu	Ser	Val	Pro	Pro	Gly	Thr	Tyr	Asp	Val	Ser	Ala	Ser	Lys	
465					470					475					480	
Phe	Gly	Tyr	Thr	Gly	Lys	Thr	Ile	Ser	Gly	Val	Gln	Val	Ala	Asp	Gly	
				485					490					495		
Gln	Thr	Val	Thr	Glu	Asp	Phe	Ala	Leu	Thr	Ala	Lys	Ala	Arg	Val	Asp	
			500					505					510			
Val	Ser	Gly	Leu	Val	Arg	Asp	Gly	Ser	Gly	His	Gly	Trp	Pro	Val	Tyr	
		515					520					525				
Ala	Thr	Val	Arg	Val	Lys	Asp	Gln	Pro	Thr	Ala	Val	Ala	Tyr	Thr	Asp	
	530					535					540					
Pro	Lys	Thr	Gly	Arg	Tyr	Thr	Leu	Ser	Leu	Pro	Thr	Asp	Asp	Thr	Tyr	
545					550					555					560	
Thr	Leu	Gln	Val	Asp	Pro	Leu	Tyr	Pro	Gly	Tyr	Thr	Gln	Asp	Ser	His	
				565					570					575		
Asp	Val	Thr	Val	Thr	Ser	Ala	Asp	Val	Val	His	Asp	Val	Asn	Val	Ala	
			580					585					590			
Val	Asp	Gln	Thr	Thr	Cys	Ser	Ala	Pro	Gly	Tyr	Ser	Tyr	His	Tyr	Ala	
		595					600					605				
Gly	Ala	Thr	Gln	Pro	Phe	Asp	Gly	Thr	Thr	Ala	Pro	Ala	Gly	Trp	Thr	
	610					615					620					
Val	Asp	Asp	Lys	Val	Gly	Asn	Gly	Gln	Thr	Trp	Val	Phe	Asn	Asp	Pro	
625					630					635					640	
Gly	Ser	Arg	Gly	Asn	Lys	Thr	Gly	Gly	Thr	Gly	Gly	Phe	Ala	Val	Ile	
				645					650					655		



ES 2 595 500 T3

Asp Ser Asp His Tyr Gly Ser Gly Asn Thr Gln Asp Ser Thr Leu Phe  
 660 665 670  
 Ser Pro Val Thr Asp Phe Ser Ala Arg Thr His Pro Ile Leu Ser Phe  
 675 680 685  
 Arg Ser Asp Trp Tyr Gly Tyr Thr Gly Gln Ala Gly Asp Leu Asp Leu  
 690 695 700  
 Ser Val Asp Gly Gly Thr Thr Trp Thr Asn Leu Lys His Tyr Thr Ala  
 705 710 715 720  
 Ser Glu Arg Gly Pro Arg Thr Glu Thr Met Asp Leu Ser Ala Ala Ala  
 725 730 735  
 Gly Lys Ser Ala Val Gln Val Arg Phe His Phe Thr Gly Lys Phe Gly  
 740 745 750  
 Tyr Tyr Trp Gln Val Asp Asp Val Phe Leu Gly Asp Arg Thr Cys Asp  
 755 760 765  
 Pro Thr Ser Gly Gly Leu Val Val Gly Gln Val Thr Asp Ala Asn Thr  
 770 775 780  
 Gly Ala Gly Val Ala Gly Ala Thr Val Thr Ser Val Asp Arg Pro Ala  
 785 790 795 800  
 Glu His Ala Thr Ser Ala Ala Thr Pro Asp Asp Pro Asn Leu Gly Asp  
 805 810 815  
 Gly Leu Phe Trp Met Phe Ser Ser Val Thr Gly Ser His Lys Phe Thr  
 820 825 830  
 Ala Thr Ala Gly Asn Tyr Asp Ser His Asp Val Thr Val Asn Val Ala  
 835 840 845  
 Ala Asn Trp Ala Thr Ala Ala Asp Ile Ala Leu Ser Ala Pro Arg Leu  
 850 855 860  
 Thr Val Thr Pro Gly Ala Ile Ser Lys Thr Val Gly Trp Gln Lys Ala  
 865 870 875 880  
 Ala Thr Ala Ala Leu Thr Leu Lys Asn Thr Gly Thr Ser Pro Val Thr  
 885 890 895  
 Ala Lys Ile Gly Glu Gln Pro Gly Gly Ser Thr Thr Ala Thr Ala Gly  
 900 905 910

ES 2 595 500 T3

Ala Pro Leu Gln Arg Val Lys Gly Asp Phe Ala Ser Gly Arg Leu Gln  
 915 920 925

Gln Gly Lys Ala Ser Gly Ala Lys Ala Gly Val Thr Pro Tyr Ala Pro  
 930 935 940

Pro Trp Thr Ala Ile Ala Asp Tyr Ala Ser Pro Ile Met Asp Asn Gly  
 945 950 955 960

Ala Val Ala Leu Asn Gly Lys Ile Tyr Ser Val Ala Gly Val Asp Gly  
 965 970 975

Ala Asn Val Leu Asn Lys Ala Tyr Ala Tyr Asp Pro Gly Thr Gln Ala  
 980 985 990

Trp Thr Ala Ile Ala Pro Leu Ala Thr Gly Arg Glu Ala Pro Gln Ala  
 995 1000 1005

Thr Thr Thr Gly Gly Lys Leu Tyr Val Thr Gly Gly Trp Gly Ser  
 1010 1015 1020

Thr Gly Ala Ala Val Ala Lys Thr Glu Ile Phe Asp Pro Ser Ser  
 1025 1030 1035

Gly Ala Trp Ser Ala Gly Ala Asp Asn Pro Lys Pro Tyr Ala Gly  
 1040 1045 1050

Ser Gly Ala Ala Val Leu Asp Gly Lys Ile Tyr Val Val Gly Gly  
 1055 1060 1065

Cys Leu Ala Thr Cys Gly Thr Lys Asp Val Gln Val Tyr Asp Pro  
 1070 1075 1080

Ala Ala Asn Ser Trp Ser Ser Gly Pro Ala Tyr Pro Glu Asn Thr  
 1085 1090 1095

Ala Trp Leu Gly Cys Ala Gly Ile Asp Gly Lys Leu Tyr Cys Ala  
 1100 1105 1110

Gly Gly Ser Ala Ala Ala Ser Thr Lys His Gly Tyr Val Leu Asp  
 1115 1120 1125

Pro Ala Ser Gly Thr Trp Ser Pro Ile Ala Asp Leu Pro Ile Asp  
 1130 1135 1140

Leu Trp Ala Met Gly Tyr Ser Ala Ala Asn Gly Lys Leu Ile Val  
 1145 1150 1155

ES 2 595 500 T3

Ser Gly Gly Val Thr Asn Gly Ala Ser Thr Leu Thr Asn Gln Gly  
 1160 1165 1170

Phe Ala Tyr Asp Pro Ser Ala Asn Ser Trp Thr Ala Leu Pro Asn  
 1175 1180 1185

Ser Asn Asn Ala Leu Tyr Arg Gly Ala Ser Ala Cys Gly Phe Tyr  
 1190 1195 1200

Lys Ile Gly Gly Ser Leu Gly Gln Phe Asn Ala Val Lys Ser Gly  
 1205 1210 1215

Glu Val Leu Pro Gly Tyr Asp Gln Cys Ala Ser Thr Ala Asp Val  
 1220 1225 1230

Pro Trp Leu Ser Glu Asp Lys Thr Glu Val Thr Ile Gln Pro Gly  
 1235 1240 1245

Gln Ser Val Lys Val Asn Val Thr Leu Asp Ala Asn Val Pro Ala  
 1250 1255 1260

Ile Thr Gln Pro Gly Thr Tyr Thr Ala Gln Leu Thr Val Gly Ala  
 1265 1270 1275

Lys Thr Pro Tyr Ala Ile Pro Pro Val Ala Val Thr Met Thr Val  
 1280 1285 1290

Asn Pro Pro Gly Thr Trp Gly Lys Ile Thr Gly Thr Leu Thr Gly  
 1295 1300 1305

Ala Gly Cys Thr Gly Ser Pro Ala Pro Leu Thr Gly Ala Thr Leu  
 1310 1315 1320

Gln Ile Asp Ser Trp Ala Ala Ser Tyr Thr Leu Lys Thr Asp Lys  
 1325 1330 1335

Asn Gly Gln Tyr Ala Leu Trp Leu Asp Val Arg Asn Asn Pro Leu  
 1340 1345 1350

Thr Leu Ile Ala Ala Lys Asp Gly Trp Ala Pro Gln Thr Arg Asn  
 1355 1360 1365

Val Lys Ile Thr Lys Leu Thr Ser Thr Thr Ala Asp Phe Thr Leu  
 1370 1375 1380

Lys Pro Asp His Thr Cys Ser  
 1385 1390

ES 2 595 500 T3

<211> 398  
 <212> PRT  
 <213> *Actinoallomurus* sp.

5 <220>  
 <221> PROPEP  
 <222> (1)..(398)

<400> 2

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

Met Val Pro Phe Leu Ser Thr Ala Ala Asn Ala Ala Pro Gln Ala Ala  
 20 25 30

Pro Ala Ser Val Ser His Pro Phe His His Ser Cys Ala Thr Val Lys  
 35 40 45

Pro Gly Arg Ala Ser Cys Asn Ala Leu Val Arg Ser Asp Ile Ala Gln  
 50 55 60

Ser Ala Ala Thr Leu Ala His Gln Ala Ala Ala Pro Ser Gly Leu Ser  
 65 70 75 80

Pro Ala Asn Leu Gln Ser Ala Tyr Lys Leu Pro Ser Ser Thr Ala Gly  
 85 90 95

Ser Gly Gln Thr Val Ala Ile Val Asp Ala Tyr Asp Ala Pro Thr Ala  
 100 105 110

Glu Ala Asp Leu Asn Val Tyr Arg Ser Gln Phe Gly Leu Gly Ala Cys  
 115 120 125

Thr Thr Ala Asn Gly Cys Phe Lys Lys Val Asp Gln Asn Gly Gly Thr  
 130 135 140

Ser Tyr Pro Arg Lys Asp Gly Gly Trp Ala Gln Glu Ile Ser Leu Asp  
 145 150 155 160

Leu Asp Met Val Ser Ala Val Cys Pro Asn Cys Lys Ile Val Leu Val  
 165 170 175

Glu Ala Lys Thr Asn Ser Phe Ala Asn Leu Gly Thr Ala Glu Asn Thr  
 180 185 190

10

ES 2 595 500 T3

Ala Ala Ser Leu Ala Asn Val Ile Ser Asn Ser Tyr Gly Gly Ser Asp  
 195 200 205

Ala Ser Asp Ala Ser Tyr Gly Ser Tyr Tyr Asn His Pro Gly Lys Ala  
 210 215 220

Ile Thr Val Ser Ser Gly Asp Ala Gly Tyr Gly Val Glu Tyr Pro Ala  
 225 230 235 240

Ser Ser His Tyr Val Thr Ala Val Gly Gly Thr Ser Leu Arg Thr Ala  
 245 250 255

Ser Thr Ser Arg Gly Trp Ser Glu Thr Ala Trp Ser Gly Ala Gly Ser  
 260 265 270

Gly Cys Ser Ala Tyr Asn Thr Ala Leu Ser Gly Gln Ser Gly Leu Thr  
 275 280 285

Gly Cys Ser Arg Arg Ala Val Ala Asp Val Ser Ala Val Ala Asp Pro  
 290 295 300

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

Gly Trp Met Val Phe Gly Gly Thr Ser Val Ala Ala Pro Ile Ile Gly  
 325 330 335

Gly Val Tyr Gly Leu Ala Ala Asn Ala Ala Ser Ile Asp Asn Asn Tyr  
 340 345 350

Pro Tyr Ala His Thr Ser Ser Leu Phe Asp Val Thr Ser Gly Ser Asn  
 355 360 365

Gly Thr Cys Thr Thr Thr Lys Trp Cys Thr Ala Gly Thr Gly Trp Asp  
 370 375 380

Gly Pro Thr Gly Leu Gly Thr Pro Asn Gly Thr Gly Ala Phe  
 385 390 395

<210> 3  
 <211> 1094  
 <212> PRT  
 <213> *Actinoallomurus*

5

<220>  
<221> PROPEP  
<222> (1)..(1094)

5

<400> 3

ES 2 595 500 T3

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

Pro Thr Pro Glu Gly Ser Gly Arg His Trp Thr Ala Thr Pro Leu Ser  
 20 25 30

Gly Thr Val Val Gln Gly Phe Lys Ser Thr Thr Gly Arg Leu Ala Lys  
 35 40 45

Ser Asp Ala Gly Leu Leu Arg Leu Lys Ser Ser Lys Pro Val His Val  
 50 55 60

Met Val Lys Leu Asp Tyr Asp Ser Leu Ala Ala Tyr Arg Gly Gly Ile  
 65 70 75 80

Pro Gly Tyr Ala Ala Thr Ser Pro Ala Val Thr Gly His Thr Leu Asp  
 85 90 95

Pro Ala Ser Ala Gly Ala Arg Arg Tyr Gln Gly Tyr Ile Gly Gly Val  
 100 105 110

Glu Asn Arg Phe Arg Ala Ala Leu Gly Lys Arg Leu Pro Gly Ala Lys  
 115 120 125

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

Pro Gly Asp Lys Val Ala Asp Leu Leu Lys Leu Pro Gly Val Ala Ala  
 145 150 155 160

Val Gln Glu Asp Ala Val Ala Lys Pro Leu Thr Asp Ser Ser Pro Gly  
 165 170 175

Phe Ile Gly Ala Pro Thr Ile Tyr Lys Gln Leu Gly Gly Ser Asp Ser  
 180 185 190

Ser Gly Lys Gly Val Ile Val Gly Asp Leu Asp Ser Gly Ala Trp Pro  
 195 200 205

Glu His Pro Ser Tyr Lys Asp Ser Gly Lys Leu Pro Ala Pro Pro Pro  
 210 215 220

Thr Thr Asp Gly Ala Pro Arg Pro Cys Asp Phe Gly Asp Asn Pro Leu  
 225 230 235 240

Thr Pro Ala Asn Asp Pro Tyr Val Cys Asp His Lys Leu Ile Ser Gly  
 245 250 255

ES 2 595 500 T3

Gln Pro Phe Leu Asp Thr Tyr Asn Ala Val Val Gly Gly Glu Arg Phe  
 260 265 270

Pro Asp Ser Ala Arg Asp Ser Asp Gly His Gly Thr His Thr Ser Thr  
 275 280 285

Thr Ala Ala Gly Ser Ala Val Ser His Ala Asn Pro Leu Gly Ile Asp  
 290 295 300

Arg Gly Ala Ile His Gly Ile Ala Pro Ala Ala His Ile Ala Val Tyr  
 305 310 315 320

Lys Val Cys Gly Ala Gln Gly Cys Phe Gln Ser Asp Ser Val Ala Ala  
 325 330 335

Val Gln Arg Ala Ile Leu Asp Lys Val Arg Val Ile Asn Tyr Ser Ile  
 340 345 350

Ser Gly Gly Val Asp Pro Tyr Ser Asp Pro Val Glu Leu Ala Phe Leu  
 355 360 365

Asp Ala Tyr Ala Ala Gly Val Leu Val Ser Ala Ser Ala Gly Asn Asp  
 370 375 380

Gly Pro Thr Ala Gly Thr Val Asn His Asn Gly Pro Trp Val Thr Thr  
 385 390 395 400

Val Ala Ala Ser Thr Gln Gln Arg Thr Phe Gln Ser Thr Val Thr Leu  
 405 410 415

Gln Ala Gly Gly Ala Ser Leu Lys Leu Ala Gly Ser Ser Ile Thr Ser  
 420 425 430

Gly Ile Thr Ser Pro Leu Pro Val Val Leu Ala Ser Ala Ala Pro Tyr  
 435 440 445

Gly Asp Pro Asp Cys Asp Thr Gln Ala Ala Pro Gly Thr Phe Thr Gly  
 450 455 460

Lys Ile Val Ala Cys Arg Leu Leu Asn Arg Gly Arg Ile Met Lys Gly  
 465 470 475 480

Tyr Asn Val Phe Lys Gly Gly Ala Ala Gly Met Leu Leu Tyr Asn Asp  
 485 490 495

Thr Leu Ser Gln Thr Met Thr Asp Asn His Trp Leu Pro Thr Val His  
 500 505 510



ES 2 595 500 T3

Leu Glu Lys Pro Gln Ala Asp Ala Leu Leu Ala Phe Leu Ser Gly His  
 515 520 525

Thr Gly Thr Thr Ala Thr Phe Pro Gln Gly Ala Lys Ala Asn Gly Gln  
 530 535 540

Gly Asp Val Met Thr Ala Phe Ser Ser Arg Gly Pro Gly Gly Asp Phe  
 545 550 555 560

Leu Lys Pro Asp Val Thr Ala Pro Gly Leu Gln Ile Met Gly Gly Gln  
 565 570 575

Thr Pro Val Pro Asp Asp Pro Ser Leu Gly Pro Pro Gly Thr Leu Tyr  
 580 585 590

Gln Val Ile Ala Gly Thr Ser Met Ser Ala Pro His Val Thr Gly Ala  
 595 600 605

Ala Ala Leu Leu Phe Ala Leu His Pro Ser Trp Thr Pro Gly Gln Val  
 610 615 620

Lys Ser Ala Leu Glu Thr Thr Gly Lys Thr Ser Val Val Lys His Asp  
 625 630 635 640

Arg Lys Thr Pro Ala Asp Pro Phe Asp Leu Gly Gly Gly Arg Ile Asp  
 645 650 655

Leu Thr Lys Ala Gly Asp Pro Gly Leu Thr Ile Asp Glu Thr Ala Ala  
 660 665 670

Asn Tyr Ala Ala Ser Ala Thr Asp Pro Leu His Arg Ile Asp Leu Asn  
 675 680 685

Val Pro Ser Val Asn Ala Pro Val Met Pro Gly Ala Ile Met Thr Thr  
 690 695 700

Arg Thr Val Lys Asn Val Ser Gly Lys Thr Met Thr Tyr Gly Thr Ser  
 705 710 715 720

Gly Thr Thr Val Lys Gly Ala Ser Ile Ser Val Ser Pro Gly Thr Phe  
 725 730 735

Thr Val Lys Pro Gly Lys Thr Ala Arg Leu Arg Ile Thr Ile Ala Ala  
 740 745 750

ES 2 595 500 T3

Pro Ala Leu Ala Asn Gly Gln Tyr Phe Gly Arg Val Asp Leu Arg Gln  
755 760 765

Arg Asn Gly Gly His Asp Leu His Leu Pro Val Ala Phe Val Arg Lys  
770 775 780

Gln Gly Val Val Ser Leu Asn Gln Thr Cys Thr Pro Ser Val Ile Ala  
785 790 795 800

Leu Asn Ser Gly Arg Ser Thr Cys Asp Val Asp Val Glu Asn Thr Ser  
805 810 815

Leu Ala Asp Thr Lys Val Ile Ala Ala Ser His Leu Asp Thr Arg Leu  
820 825 830

Arg Leu Thr Ala Val Thr Gly Ala Thr Lys Val Gly Ala His Asp Val  
835 840 845

Leu Ala Gln Ala Asp Leu Ala Arg Arg Gln Pro Asp Lys Pro Gln Ile  
850 855 860

Ala Pro Gly Ala Thr Pro Ala Gly Tyr Leu Pro Leu Asp Ala Phe Gly  
865 870 875 880

Ile Pro Pro Thr Pro Ile Gly Asp Glu Gln Ser Leu Asn Leu Thr Thr  
885 890 895

Pro Ala Phe Thr Phe Ala Gly Arg Thr Tyr Thr Ser Leu Gly Val Val  
900 905 910

Ser Asp Gly Tyr Thr Val Ala Gly Gly Ala Thr Pro Asp Asp Val Ala  
915 920 925

Ala Thr Pro Gln Thr Leu Pro Asn Pro Ala Arg Pro Asn Gly Glu Leu  
930 935 940

Ala Pro Phe Trp Thr Asp Leu Asp Gly Ala Gly Ala Pro Gly Val Tyr  
945 950 955 960

Ala Ala Arg Leu Thr Asp Gly Thr Ser Thr Trp Ile Val Val Glu Trp  
965 970 975

Arg Leu Asn Val Phe Gly Thr Asn Ser Leu Arg Val Phe Gln Gln Trp  
980 985 990

Ile Gly Val Asn Gly Ala Glu Asp Ile Ser Tyr Thr Tyr Asp Pro Asn  
995 1000 1005

ES 2 595 500 T3

Asn Leu Pro Ala Ala Pro Pro Ala Gly Tyr Gly Leu Thr Val Gly  
 1010 1015 1020

Ala Glu Asn Asp Glu Gly Thr Ala Gly Ser Gln Ile Ser Gly Thr  
 1025 1030 1035

Pro Thr Glu Asp Leu Arg Val Thr Ser Thr Pro Gly Ala Ala Gly  
 1040 1045 1050

Gly Ser Leu Lys Tyr Ser Phe Thr Leu Lys Gly Thr Gly Arg Gly  
 1055 1060 1065

Asn Ala Pro Val Thr Thr Leu Val Ser Thr Pro Leu Val Arg Gly  
 1070 1075 1080

Val Thr Ala Glu Val Asp Asn Ile Thr Val Asn  
 1085 1090

<210> 4  
 <211> 570  
 5 <212> PRT  
 <213> *Actinoallomurus*

<220>  
 <221> PROPEP  
 10 <222> (1)..(570)

<400> 4  
 Met Arg Arg Ala Leu Val Leu Ala Ala Cys Ala Val Val Thr Ala Ala  
 1 5 10 15

Leu Ala Ala Pro Ala Gln Ala Gly Ser Ser Ser Lys Thr Gln Gly Lys  
 20 25 30

Thr Thr Glu Tyr Val Val Leu Tyr Lys Asp Gly Val Ser Leu Glu Lys  
 35 40 45

Ala His Ser Ala Val Lys Ala Ala Gly Gly Thr Ile Val Lys Glu Asn  
 50 55 60

Lys Ala Ile Gly Val Ala Thr Val Arg Ser Ala Ser Thr Ala Phe Leu  
 65 70 75 80

Thr Asp Ala Arg Lys Gln Ser Ala Val Asp Gly Val Ala Thr Asn Arg  
 85 90 95

Ala Val Gly Glu Ala Pro Lys Val Ala Arg Ala Ala Val Asn Lys Ser  
 100 105 110

ES 2 595 500 T3

Gln Ala Val Glu Lys Glu Gly Arg Val Gly Gly His Ala Gly Ser Ser  
 115 120 125

Ser His Lys Pro Ser Ala Glu Pro Leu Ala Asp Arg Gln Trp Asp Met  
 130 135 140

Lys Gln Ile His Ala Thr Thr Asp Gly Ser Tyr Lys Lys Glu Pro Gly  
 145 150 155 160

Asp Arg Arg Val Leu Val Gly Val Ile Asp Thr Gly Ile Asp Gly Thr  
 165 170 175

His Pro Asp Ile Ala Pro Asn Phe Asp Lys Ser Leu Ser Arg Asn Phe  
 180 185 190

Thr Thr Asp Ile Pro Val Asp Ala Asn Gly Thr Glu Val Asp Gly Pro  
 195 200 205

Cys Glu His Pro Ser Cys Val Asp Pro Val Asp Glu Asp Asp Asn Glu  
 210 215 220

His Gly Thr His Val Ala Ser Thr Ile Ala Ser Pro Leu Asn Gly Leu  
 225 230 235 240

Gly Ile Ala Gly Val Ala Pro Asn Val Thr Leu Val Asn Val Arg Ala  
 245 250 255

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

Thr Tyr Ser Ala Asp His Gly Ile Asp Val Val Asn Met Ser Phe Tyr  
 275 280 285

Thr Asp Pro Trp Leu Phe Asn Cys Thr Asn Asn Pro Ala Asp Ser Pro  
 290 295 300

Glu Asn Gln Ala Glu Gln Arg Thr Val Ile Gln Ala Ser Glu Arg Ala  
 305 310 315 320

Leu Ala Tyr Ala His Arg His Gly Val Thr Leu Val Ala Ala Ala Gly  
 325 330 335

Asn Gly Ala Thr Asp Tyr Thr Lys Thr Ile Thr Asp Ala Ser Ser Pro  
 340 345 350

Asp Tyr Pro Ser Val Pro Gly Glu Ala Pro Tyr Ser Arg Thr Ile Pro

ES 2 595 500 T3

355		360		365											
Pro	Ser	Cys	Ile	Ser	Glu	Pro	Ser	Glu	Gly	Gln	Asn	Val	Leu	Ala	Val
	370					375					380				
Ser	Ala	Leu	Gly	Ile	Ser	Thr	Arg	Lys	Ser	Tyr	Tyr	Ser	Asp	Tyr	Gly
385					390					395					400
Asn	Gly	Tyr	Val	Ala	Val	Ser	Ala	Pro	Gly	Gly	Asp	Ser	Tyr	Asp	Thr
				405					410					415	
Ala	Asp	Gln	Lys	Ala	Asp	Val	Thr	His	Ala	Ile	Leu	Ala	Ala	Tyr	Pro
			420					425					430		
Lys	Ser	Leu	Ala	Val	Ala	Arg	Gly	Glu	Leu	Asn	Ala	Asp	Gly	Thr	Pro
		435					440					445			
Asn	Val	Pro	Tyr	Val	Val	Arg	Ser	Cys	Lys	Gly	Ser	Thr	Cys	Ala	Tyr
	450					455					460				
Tyr	Gln	Tyr	Leu	Gln	Gly	Thr	Ser	Met	Ala	Ser	Pro	His	Ala	Thr	Gly
465					470					475					480
Val	Val	Ala	Leu	Ile	Val	Ser	Arg	Tyr	Gly	Lys	Pro	Asp	Arg	Val	His
				485					490					495	
Gly	Gly	Leu	Thr	Leu	Ser	Pro	Asp	Arg	Val	Lys	Ser	Ile	Leu	Glu	Gly
			500					505					510		
Thr	Ala	Thr	Glu	His	Ala	Cys	Pro	Asp	Pro	Arg	Ala	Phe	Thr	Tyr	Thr
		515					520					525			
Arg	Gln	Val	Lys	Gln	Ser	Asp	Gly	Thr	Tyr	Arg	Thr	Val	Thr	Ala	Thr
	530					535					540				
His	Thr	Cys	Glu	Gly	Ser	Lys	Ser	His	Asn	Gly	Phe	Tyr	Gly	His	Gly
545					550					555					560
Ile	Ile	Asp	Ala	Leu	Gly	Ala	Val	Thr	His						
				565					570						

<210> 5  
 <211> 416  
 <212> PRT  
 <213> *Actinoallomurus*

5

<220>  
<221> PROPEP  
<222> (1)..(416)

5 <400> 5

ES 2 595 500 T3

Met Asn Pro Arg Gln Arg Phe Ala Ile Val Leu Ala Ala Leu Ile Thr  
1 5 10 15

Met Ile Phe Ser Leu Val Ser Val Pro Ser Ala Gly Ala Ala Val Ser  
20 25 30

Ala Ser Lys Leu Ser Val Gly Phe Gly Cys Ala Ser Ala Ala His Ala  
35 40 45

Gly Gln Leu His Cys Phe Gly Arg Ile Arg Ala His Arg Ala Ser Asn  
50 55 60

Gly Lys Ile Ala Pro Leu Thr Val Thr Ser Pro Thr Gly Leu Leu Pro  
65 70 75 80

Ala Asp Leu Gln Ser Ala Tyr Lys Val Ala Gly Leu Asn Gly Gly Gly  
85 90 95

Arg Thr Val Ala Ile Val Asp Ala Gln Asp Asn Pro Lys Ala Glu Ala  
100 105 110

Asp Leu Ala His Tyr Arg Ser Gln Leu Gly Leu Pro Ala Cys Thr Thr  
115 120 125

Ala Asn Gly Cys Phe Lys Lys Val Asn Gln Asn Gly Gln Ala Ser Pro  
130 135 140

Leu Pro Ala Ala Asp Tyr Gly Trp Ala Glu Glu Ile Ser Leu Asp Leu  
145 150 155 160

Asp Met Val Ser Ala Ile Cys Pro Ser Cys His Ile Leu Leu Val Glu  
165 170 175

Ala Asn Ala Pro Asp Asp Thr Ser Leu Gly Thr Ala Val Asp Thr Ala  
180 185 190

Ala Ala Thr Ser Gly Val Val Ala Ile Ser Asn Ser Tyr Gly Gly Ala  
195 200 205

Glu Asp Ser Thr Ile Leu Ala Ala Asp Ala His Phe Asn His Pro Gly  
210 215 220

Ile Ala Val Thr Ala Ser Ser Gly Asp Ser Gly Tyr Gly Val Ser Trp  
225 230 235 240

ES 2 595 500 T3

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

Lys Ala Ser Asn Ala Arg Gly Trp Thr Glu Thr Ala Trp Ser Gly Ala  
 260 265 270

Gly Ser Gly Cys Ser Ala Tyr Glu Pro Lys Pro Ser Trp Gln His Asp  
 275 280 285

Thr Ala Cys Ala Lys Arg Thr Val Ala Asp Val Ser Ala Val Ala Asp  
 290 295 300

Pro Ala Thr Gly Val Gly Val Tyr Asp Thr Tyr Asn Ser Cys Gly Thr  
 305 310 315 320

Ser Ser Phe Cys Asp Phe Leu Ile Ser Leu Gly Leu Val Gln Gly Leu  
 325 330 335

Asp Gly Trp Ala Ala Val Gly Gly Thr Ser Ala Ser Ser Pro Ile Ile  
 340 345 350

Ala Ser Val Tyr Ala Leu Ala Gly Asn Thr Gly Ser Thr Thr Tyr Gly  
 355 360 365

Ser Tyr Pro Tyr Ala His Thr Ser Ala Leu Phe Asp Val Thr Ser Gly  
 370 375 380

Ser Asn Gly Ser Cys Gly Gly Thr Tyr Leu Cys Thr Ala Gly Thr Gly  
 385 390 395 400

Tyr Asp Gly Pro Thr Gly Leu Gly Thr Pro Asn Gly Thr Gly Gly Phe  
 405 410 415

<210> 6  
 <211> 33  
 5 <212> PRT  
 <213> Secuencia artificial

<220>  
 <223> Péptido 33-mero gliadina

10 <400> 6  
 Leu Gln Leu Gln Pro Phe Pro Gln Pro Gln Leu Pro Tyr Pro Gln Pro  
 1 5 10 15

Gln Leu Pro Tyr Pro Gln Pro Gln Leu Pro Tyr Pro Gln Pro Gln Pro  
 20 25 30

Phe

15 <210> 7  
 <211> 4173



ES 2 595 500 T3

<212> ADN  
 <213> *Actinoallomurus*

5 <220>  
 <221> gen  
 <222> (1)..(4173)

<400> 7  
 atgcccgatc ttcccaccca acggcgttcc ccacgcaggt cacgccggcc cgcggtcctt 60  
 gcgaggctcg cgggcctgcg cgtcggcgcg ctggtcctcg cgctcgccgt caccacgctc 120  
 accctcatct ccatgccctc ggccaacgcc gccccggcca agcccgtggt cccgcacgcc 180  
 aaggcgagca gtcccaccgc acagggccag cacccgacca gccgcgtgtg ccggaccccc 240  
 ggccgcgcca atgagatgac ctgcctcgcc atggtcctgt acgacgtcat cgggccgaag 300  
 ggcgtccagg cgaacgacgc gccggccggg ttcggcccgg ccgacctgct gaacgcctac 360  
 aacctgcct cggccggatc gacggagacg gtcgcgatcg tcgacgcgtt cgacaaccgg 420  
 aacgccgagg cggacctggg cgtctaccgc cagcagtacg ggctgccggc ctgcacgacg 480  
 gccaacggct gttcaagaa gatcgaccaa cgcggcggca ccagctatcc gccgccggac 540  
 gccggctggg ccggagagat cgcgctggac atcgacatgg tcagcgccat ctgcccgacc 600  
 tgcaagatcc tcctggtcga ggccgacgac aactacatgg acaacctggg cacggccgctc 660  
 aaccaggcgg tcagccaggg cgccaagtcc gtctccaaca gctacggcgg cagtgagggc 720  
 tccgacgagg ggcaggccga cgacgcgtac ttccaccacg acggcgtcgc catcaccgcg 780  
 agcagcggcg acagcggcta cggggcgagc taccggcccg cctccccgta cgtcacctcg 840  
 gtcggcggca cgtcgtggc caaggacacg tcctcctcgc gcggctggac cgagagcacc 900  
 tggaacggcg cgggcagcgg ctgctccgcg tacgagacca agccctcgtt ccagaaggac 960  
 accggctgcg cgcgccgtac gatcgccgac gtctccgcgg tggccgacct gaacaccggc 1020  
 gtcgcggtct acaacggcgg ctggcacgct tacggcggca cgagcgtgtc cgcctcgatc 1080  
 atcgcgtcgg tgtacgccct cggcggcgcg ccggcggcgg gctcggcgcc caactccttc 1140  
 ccgtacgacc acccggccga cctcaacgac gtgaccagcg gcagcaacgg gagctgtagt 1200  
 cccgcctacc tctgcacggc ggtaccggc tacgacggc cgaccgggct cggcaccctc 1260  
 aacggcatcg ccgcgttccg ctccgggccc cacgccgtcg tcaccggtac ggtcaccgac 1320  
 ggcaccgccc cgctggcgtc cgcccagggt aaggcggggc acgtcacggc gacgacggac 1380  
 ggccaggggc actacaccct gacggttccg ccggggacct acgacgtgtc cgcgagcaag 1440

ES 2 595 500 T3

ttcgggtaca	ccggtaaagac	gatctcgggg	gtgcaggctcg	ccgacggcca	gaccgtcacc	1500
gaggacttcg	cgctgaccgc	caaggcacgc	gtcgacgtct	ccggactcgt	acgggacggc	1560
tcgggtcacg	gctggccggg	ctacgcgacc	gtacgggtca	aggaccagcc	caccgcggtc	1620
gcctacaccg	acccgaagac	cgggcggtag	accctcagcc	tgcccacgga	cgacacgtac	1680
acgctgcagg	tggatccgct	gtatccgggt	tacacgcagg	actcgcacga	cgtgaccgtg	1740
acctcggcgg	acgtggtcca	cgacgtcaac	gtcgcgggtg	accagacgac	gtgcagcgca	1800
ccgggctaca	gctaccacta	cgcgcggagcc	accagccgt	tcgacggcac	gaccgcgccg	1860
gcgggctgga	ccgtcgatga	caaggctcggc	aacggccaga	cctgggtgtt	caacgacccg	1920
ggctcccgtg	gcaacaagac	gggcgggacc	ggcggcttcg	ccgtcatcga	cagtgaccac	1980
tacggctcgg	gcaacacgca	ggacagcacg	ctgttcagcc	cggtcaccga	cttcagcgcg	2040
cgtaccacc	cgatcctgag	cttcgcgagc	gactggtacg	gctacaccg	ccaggccggc	2100
gacctcgacc	tgagcgtcga	cggcggcacg	acgtggacca	acctgaagca	ctacacggcc	2160
agcgagcgcg	ggccgcgtac	ggagacgatg	gacctgtccg	cggcggcccg	gaagagcgcc	2220
gtgcaggatga	ggttccactt	caccggcaag	ttcggctact	actggcaggt	cgacgacgtc	2280
ttcctcggcg	accgtacctg	tgacccgacc	tcgggcggcc	tggtcgtcgg	ccagggtgacc	2340
gacgcgaaca	ccggcgcggg	cgtggccggg	gccacggcca	cctcggtgga	ccggcggcc	2400
gagcacgcca	cctcggcggc	gacgccggac	gacccaacc	tcggcgacgg	tctcttctgg	2460
atgttctcct	ccgtcaccgg	cagccacaag	ttcaccgcca	cggcggggaa	ctacgactcc	2520
catgacgtca	cggtgaacct	cgcgcgcaac	tgggcgaccg	cggccgacat	cgcgctgtcc	2580
gcgccacggc	tcacggtcac	tccgggcgcg	atcagcaaga	cggtcggctg	gcagaaggcg	2640
gccaccgcgg	cgctgaccct	gaagaacacc	ggcacgtcgc	cggtgaccgc	caagatcggg	2700
gaacagccgg	gtggcagcac	gaccgccacc	gcgggtgcac	cgctgcagcg	ggtgaagggc	2760
gacttcgcca	gtggacgact	gcagcagggc	aaggcgtcgg	gtgccaaaggc	cggcgtgaca	2820
ccgtacgcgc	cgcctgggac	ggccatcgcc	gactacgcct	cgccgatcat	ggacaacggc	2880
gcgggtggcac	tgaacggcaa	gatctactcg	gtggcggggc	tcgacggcgc	gaacgtgctg	2940
aacaaggcgt	acgcctacga	tcccggcacc	caggcctgga	ccgccatcgc	cccgctcgcc	3000
acgggacgtg	aggctcccca	ggcgacgacc	accggcggaa	agctgtacgt	caccggcggc	3060
tggggctcga	cgggcgcggc	ggtggccaag	acggagatct	tcgatccgtc	ctcggggcgc	3120
tggtcggccg	gcgcggacaa	cccgaagccg	tacgccggct	ccggcgcggc	cgtgctcgac	3180
ggcaagatct	acgtcgtcgg	cgggtgcctc	gccacctgcg	gtacgaagga	cgtgcaggtc	3240
tacgaccggg	cggccaactc	ctggagctcg	ggaccggcct	atccggagaa	caccgcgtgg	3300
ctcggctcgc	ccggcatcga	cggcaagctg	tactgcgcgg	gcggctccgc	ggcgcgagc	3360

ES 2 595 500 T3

accaagcacg gctacgtgct cgacccggcc tcgggcacct ggtcgccgat cgccgacctg 3420  
 ccgatcgacc tgtgggggat gggctactcg gcggcgaacg ggaagctgat cgtctccggt 3480  
 ggcgtcacca acggggccag cacgctcacc aatcagggct tcgcctacga cccgtcggcg 3540  
 aactcctgga cggccctgcc caactccaac aacgccctgt accgcggcgc gtcggcctgc 3600  
 gggttctaca agatcggagg atcgtctggg cagttcaacg cggccaagag cggtgagggtg 3660  
 ctgcccggct atgaccagtg cgcctcgacc gccgacgttc cgtggctgtc ggaggacaag 3720  
 accgaggtga cgatccagcc cggccagagc gtcaagggtca acgtgaccct cgacgcgaac 3780  
 gtcccggcga tcaactcagcc cggcacgtac accgcgcagc tcaccgtcgg ggccaagacg 3840  
 ccgtacgcga tcccgccggt cgccgtcacg atgacgggtga acccgccggg cacctgggga 3900  
 aagatcaccg gaacgctcac cggagccggc tgtacggggt ctcccgcacc gctgaccggg 3960  
 gcgaccctac agatcgactc ctgggctgcg tcgtacacgc tcaagaccga caagaacggt 4020  
 cagtacgcgc tctggctcga cgtccgcaac aaccgctga cgctgatcgc ggccaaggac 4080  
 ggatggggcg cgcagaccag aaacgtcaag atcaccaaac tgacctcgac cacggccgat 4140  
 ttcactctga aaccgacca cacctgtagc tga 4173

<210> 8  
 <211> 1197  
 <212> ADN  
 <213> *Actinoallomurus*

5

<220>  
 <221> gen  
 <222> (1)..(1197)

10

<400> 8  
 atgtcacgac gcgtgaccgg gaccatactg ggcggggtga tcctcgccat ggtccccttc 60  
 ctttccaccg cggccaacgc cgcacccag gccgcgccgg cttccgtctc ccaccgcttc 120  
 caccactcct gcgccacggt gaagccgggt cgggcgagct gcaatgccct cgtacgcagc 180  
 gacatcgccc agagcgcggc gaccctcgcg caccaagcgg ccgccccatc cgggctctcg 240  
 ccggccaacc tgagagcgc ctacaagctg ccgtcctcca cggccggatc cggccagacc 300  
 gtcgcgatcg tcgacgccta tgacgccccg accgcogaag cggacttgaa cgtgtaccga 360  
 agccagttcg gactcggcgc gtgcacgacc gccaacggct gtttcaagaa ggtcgaccag 420  
 aacggcggca cgtcctatcc gaggaaggac ggcggctggg cgcaggagat ctccctggac 480  
 ctcgacatgg tctccgcggt ctgccccaac tgcaagatcg ttctcgtcga ggccaagacc 540  
 aactcgttcg ccaacctggg taccgcccag aacaccgagg cgagtctcgc gaacgtcatc 600  
 agcaacagct acggcggctc ggacgcctct gacgcgagct atggctcgtc ctacaaccac 660

ES 2 595 500 T3

ccgggcaagg ccatcacggt cagctccggc gacgccggct acggcgtgga gtaccgggcc 720  
 tegtcccact acgtgaccgc cgtcggcggc acctcgtgac gcaccgcgag caccagccgc 780  
 ggctggagcg agaccgcgtg gagcggcggc ggcagtggtt gctcggccta caacaccgcg 840  
 ctgtccggcc agtccggcct caccggctgc tcccggcggc ccgtcggcga cgtctccgcc 900  
 gtggccgacc cggccaccgg cgtcggcgtc tacgacagca cggcctacca gggccagagc 960  
 ggctggatgg tcttcggcgg caccagcgtc gccgcaccga tcatcggtag cgtgtacggc 1020  
 ctgcgccca acgcccggag catcgacaac aactaccctt acgcccacac cagctcgttc 1080  
 ttcgacgtca cgtcgggcag caacggcacc tgcaccacca ccaagtggtag caccggcggc 1140  
 accggctggg acgccccac cggcctcggg acgcccgaac gcaccggagc cttctga 1197

<210> 9  
 <211> 3285  
 <212> ADN  
 <213> *Actinoallomurus*

5

<220>  
 <221> gen  
 <222> (1)..(3285)

10

<400> 9  
 atgagtgccg ccgttgtgat cttcggcggc gcaccggcga cggcggcggc gacgccggaa 60  
 ggctccggca gacactggac ggccacgccg ctcagcggca ccgtcgtcca gggcttcaag 120  
 tcgacgaccg gacggctcgc caagagcgat gccgggctgc ttcgcctgaa gtcgagcaag 180  
 cccgtgcacg tcatggtcaa gctcgactac gactcaactc cgcctaccg gggcgggatt 240  
 cccgggtacg ccgccacgag cccggcggtg accgggcaca cgtcggaccg ggcgagcggc 300  
 ggcgcgcggc gctaccaggg ctacatcggg ggcgtcgaga accgcttccg cgcgcacctg 360  
 ggcaagcggc tgccggcggc gaaggcggc ggcgggctgc gcacgggtga cggcgggtctg 420  
 gcggtgacgc tgcccggtag caaggtggcc gacctgctca agctgcccgg cgtggccggc 480  
 gtccaggagg acgcggtggc caagccactg accgactcca gccccggctt catcggcggc 540  
 ccgaccatct acaagcaact cggcggcagc gacagctccg gaaagggcgt catcgtcggc 600  
 gacctggaca gcggcgcctg gcccgagcac ccctcgtaca aggacagcgg caagctgccc 660  
 gcgccgccgc ccaccacgga cgggtgcgca cggcctgctg acttcggcga caatccgctg 720  
 acgcccggca acgaccgta cgtctgcgac cacaagctga tctcgggcca gccgttctctg 780  
 gacacctaca acgcggtcgt cggcggggag aggttccccg acagcggccc cgaactccgac 840  
 gggcacggca cgcacacctc gaccaccggc gccggttcgg cggtagacca cgcgaacccc 900  
 ctcggcatcg accgcggcgc catccacggc atcgggcccg ccgccacat cgcgctctac 960  
 aaggtgtgcg gcgcccaggg ctgcttccag tccgactcgg tggccggcgt gcagcggggc 1020

ES 2 595 500 T3

atcctcgaca aggtccgggt gatcaactac tcgatctccg gcggcgtcga cccatacagc 1080  
 gatccggctcg aactggcctt cctcgacgcg tacgcggccg gcgtgctcgt ctcggcctcg 1140  
 gccggcaacg acgggcccac cgcgggcacc gtcaaccaca acgggcccgtg ggtgaccacg 1200  
 gtggccgcgt ccacgcagca ggggacctc cagtcgaccg tcacgctgca ggcgggcggc 1260  
 gcgagcctga agctggccgg ctcgtcgatc accagcggga tcacctgcc gcttcccgtc 1320  
 gtgctcgcgt ccgcgggccc gtacggcgac ccggactgtg acacgcaggc cgtcccggc 1380  
 acgttcaccg gcaagatcgt cgcctgccgg ctgctcaacc gcggccggat catgaagggc 1440  
 tacaacgtct tcaagggcgg cgcggccggg atgctgctct acaacgacac gctgtcccag 1500  
 acgatgaccg acaaccactg gctgccgacc gtgcacctgg agaagccgca ggccgacgcg 1560  
 ctgctggcct tcctttccgg ccacaccggc accaccgcca cgttccccca gggcgcaag 1620  
 gcgaacggcc agggcgacgt catgaccgcg ttctcctcgc gcggcccggg cggcgacttc 1680  
 ctcaagcccg acgtcaccgc gcccggcctg cagatcatgg gcggccagac gccggttccg 1740  
 gacgaccct cgtgggccc gcccggcacc ctctaccagg tgatcgccgg tacctcgatg 1800  
 tcggcgccgc acgtcaccgg ggcggcggcg ctgctgttcg ccctgcatcc gagctggacg 1860  
 ccggggcagg tcaagtggc gctggagacg accggcaaga cgtctgtggt caagcacgac 1920  
 cgcaagacgc ccgccgacce gttcgacctc ggcggtgggc gcatcgacct caccaaggcc 1980  
 ggtgaccccg ggctgaccat cgacgagacc gcggcgaact acgcggcctc ggcgaccgac 2040  
 ccgctgcacc ggatcgacct gaacgtgccg tctgtgaacg caccggtgat gcccggcgcg 2100  
 atcatgacca cccgtacggt caagaacgtc tcgggaaaga cgatgacgta cggcacctcg 2160  
 ggtacgacgg tgaagggcgc ctccatctcc gtctccccg gcacgttac cgtcaagccg 2220  
 ggcaagaccg cccggctgcg gatcacgatc gccgcgccgg cgtcgccaa tgggcagtac 2280  
 ttcggccggg tcgacctgcg tcagcgaaac ggcggccacg acctgcacct gccggtcgcg 2340  
 ttcgtccgca agcagggcgt ggtgagcctg aaccagacct gcacgccctc ggtgatcgcg 2400  
 ctgaacagcg gccgatcgac gtgcgacgtc gacgtcgaga acacctctct cgccgacacc 2460  
 aaggtcatcg ccgcgagcca cctggacacg cggctgcgac tgaccgcggt gaccggcgcg 2520  
 acgaaggtgg gcgcgcacga cgtcctggcc caggccgacc tggcccggc ccagcccgac 2580  
 aagccgcaga tcgccccgg ggcacgccg gccggctacc tgccgctcga cgccttcggc 2640  
 atcccccgga cgccgatcgg cgacgagcag tcgctcaacc tcaccacgcc ggcggtcacc 2700  
 ttcgccggca ggacctacac cagcctgggc gtggtctccg acggctacac ggtcgccggc 2760  
 ggagcgaccc ccgatgacgt ggccgcgacg ccgcagacct tgccgaacct ggcacggccg 2820  
 aacggcgaac tggccccgtt ctggaccgac ctggacggcg ccggtgcccc gggcgtgtac 2880

ES 2 595 500 T3

gcggcccgcc tcaccgacgg cacgagcacc tggatcgtcg tggagtggcg gctcaacgtc 2940  
 ttcggcacga acagcctccg ggtgttccag cagtggatcg gggatgaacgg cgccgaggac 3000  
 atctcctaca cctacgaccc gaacaacctg ccggcgggcg cgcccgcggg ctacggcctg 3060  
 acggtcggcg cggagaacga cgagggtacc gccggttcgc agatctccgg caccgaccc 3120  
 gaggatctcc gcgtcacgag cacgccgggg gccgggggtg gttcgctgaa gtactccttc 3180  
 aactgaagg gaacggggccg gggcaacgcc ccggtgacga ccctggtctc cacgccgctc 3240  
 gtacgaggcg tgacggccga ggtcgacaac atcacggtga actga 3285

5 <210> 10  
 <211> 1713  
 <212> ADN  
 <213> *Actinoallomurus*

10 <220>  
 <221> gen  
 <222> (1)..(1713)

<400> 10  
 gtggggggcg ccctagtcct cgccgcctgc gccgtcgtca cggcgggcgt cgccggcacc 60  
 gcgcaggcgg gatcgagcag caagacgcag ggaaagacga ctgagtacgt cgtcctgtac 120  
 aaggacggag tctccctgga gaaggcgcac tcggccgtga aggcggccgg tggcaccatc 180  
 gtcaaggaga acaaggcgat cggcgtggcc accgtacgct cggccagcac ggcgttcctc 240  
 accgacgcgc gcaagcagtc ggcggtcgac ggcgtcgcga ccaaccgcgc ggtcggcgag 300  
 gcgcccgaag tcgcacgggc ggcgggtgaac aagagccagg ccgtggagaa ggagggccgc 360  
 gtcggggggc atgcggggtc ctctcgcac aagccgtcgg cggagccgct ggcggaccgt 420  
 cagtgggaca tgaagcagat ccacgcgacc acggacggct cgtacaagaa ggagccgggc 480  
 gaccggcgcg tgctcgtcgg cgtcatcgac accggcatcg acggcacgca cccggacatc 540  
 gcgccgaact tcgacaagtc gctgagccgg aacttcacca cggacatccc ggtcgacgcc 600  
 aacggcaccg aggtcgacgg cccgtgtgag caccgctcct gcgtggaccc ggtggacgag 660  
 gacgacaacg agcacggcac gcacgtcgcc tcgacgatcg cctcgcggct gaacggcctg 720  
 ggcatcgcgg gcgtggcgcc gaacgtgacg ctggtcaacg tacgggcccg gcaggactcg 780  
 ggctacttct tcctgcagcc cgtggtcgac gcgctgacgt actcggccga ccacggcatc 840  
 gacgtggtga acatgtcgtt ctacaccgac ccgtggctgt tcaactgcac caacaacccg 900  
 gccgactcgc cggagaacca ggccgagcag cgtacggta tccaggcgtc agaacgcgcc 960  
 ctggcgtacg cgcaccggca cgggtgcacc ctggtggcgg ccgccggcaa cggcgccacc 1020  
 gactacacca agacgatcac cgacgcgtcg agcccggact acccgtccgt gcccgcgag 1080  
 gcgccgtact cgcgtacgat cccgccgtcc tgcattctcc agccgagcga gggccagaac 1140

ES 2 595 500 T3

gtcctcgcgg tgcggccct gggcatcagc acgcgcaagt cgtactactc cgactacggc 1200  
aacggctacg tcgcggtgtc ggccccgggc ggcgactcct acgacacggc cgaccagaag 1260  
gcggacgtga cccacgcgat cctggcccg cctgcccgaagt ccctggccgt ggccccgggc 1320  
gaactgaacg cggacggcac cccgaacgtg ccgtacgtgg tccgctcgtg caagggtctg 1380  
acctgocgct actaccagta cctgcagggc acgtcgtatg cctctccgca cgcgaccggc 1440  
gtggtggcgc tgatcgtcag ccgctacggc aagcccgacc gcgtacacgg cggcctgacc 1500  
ctgtccccgg accgggtcaa gtccatcctg gagggaaacc ccaccgaaca cgcctgcccg 1560  
gaccacgcg ccttcacgta cacgcgccag gtcaagcagt ccgacggtag gtacaggacc 1620  
gtgaccgcga cccacacctg tgagggttcg aagagccaca acggcttcta cggccacggc 1680  
atcatcgtg ccctgggcgc cgtgacgcat taa 1713

<210> 11  
<211> 1251  
5 <212> ADN  
<213> *Actinoallomurus*

<220>  
10 <221> gen  
<222> (1)..(1251)

<400> 11  
ttgaaccccc gccaaagatt cgcgatcgtc ctcgcggccc tcatcacaat gatcttctcc 60  
ctggtttccg tcccctccgc cggcgcccg cgtcagcgcgt ccaagctcag cgtcggcttc 120  
ggctgocgct cggcggcgca cgcgggtcag ctgcaactgt tcggggggat ccgcccacc 180  
cgggcgagca acggcaagat cgcctccgctc acggtcacca gcccgaccgg actgctcccg 240  
gcggacctgc agtcggcgta caaggtggcc gggctgaacg gcggcggccg tacggtcgcg 300  
atcgtcgcag cgcaggacaa cccgaaggcc gagggcggacc tcgcccacta ccgctcccag 360  
ctcggcctgc ccgctgcac gaccgcgaac ggctgcttca agaaggtcaa ccagaacggc 420  
caggcgtcgc cgtgcccgc gcccgactac ggctgggccc aggagatcag cctcgcacctg 480  
gacatggtct cggcgtatct cccgagctgc cacatcctgc tcgtcagagg gaacgcccct 540  
gacgacacct cgtcggcac cgcggtcgc accgcccggc cgaccagcgg cgtcgtggcc 600  
atctccaaca gctacggagg cgcggaggac tcgaccatcc tcgcccggga cgcctccttc 660  
aaccacccgg gcatcgcggg cacggcgagc tccggcgact ccggctacgg cgtcagctgg 720  
ccggcctcgt cccagtacgt caccgcggtg ggcggcacga cgtgaacaa ggcgagcaac 780  
gcgcgoggct ggaccgagac cgcctggtcc ggcgcccggc cgggctgctc ggcgtacgag 840  
ccgaagccgt cctggcagca cgacaccgcc tgcgccaagc gcaccgtcgc ggacgtctcg 900

ES 2 595 500 T3

gcggtcgccc acccgccac cggcgtcggc gtctacgaca cctacaacag ctgcgggacc 960  
 agctcgttct gcgacttcct catctcgtc gggctcgtgc agggcctgga cggctgggcc 1020  
 gcggtcggcg ggacgagcgc gtccctcggc atcatcgcga gcgtgtacgc cctggccggc 1080  
 aacaccggca gcacgacgta cggctcgtac ccgtacgcgc acacgtccgc gctcttcgac 1140  
 gtcacgtccg gctcgaacgg aagctgtggc ggcacctacc tgtgcacggc cggaacggc 1200  
 tacgacggtc ccaccggtct gggcacgccc aacggaaccg gcggcttctg a 1251  
 <210> 12  
 <211> 28  
 <212> ADN  
 5 <213> Secuencia artificial  
 <220>  
 <223> cebador de PCR directo para Endopep140  
 10 <220>  
 <221> primer\_bind  
 <222> (1) . . (28)  
 <400> 12  
 15 aaaaagcttc agctacaggt gtggtcgg 28  
 <210> 13  
 <211> 29  
 <212> ADN  
 20 <213> Secuencia artificial  
 <220>  
 <223> cebador de PCR inverso para Endopep140  
 25 <220>  
 <221> primer\_bind  
 <222> (1) . . (29)  
 <400> 13  
 30 aaaaaaacat atgcccgatc ttcccacc 29  
 <210> 14  
 <211> 26  
 <212> ADN  
 35 <213> Secuencia artificial  
 <220>  
 <223> cebador de PCR directo para Endopep40  
 40 <220>  
 <221> primer\_bind  
 <222> (1) . . (26)  
 <400> 14  
 45 aaaaagcttc agaaggctcc ggtgcc 26  
 <210> 15  
 <211> 29  
 <212> ADN  
 50 <213> Secuencia artificial  
 <220>  
 <223> cebador de PCR inverso para Endopep40  
 55 <220>



ES 2 595 500 T3

<221> primer\_bind  
<222> (1) .. (29)

<400> 15  
5 aaaaaaacat atgtcacgac gcgtgaccg 29

**REIVINDICACIONES**

1. Composición enzimática que comprende por lo menos una endopeptidasa de la familia S8/S53 activa a un pH de entre 3 y 8 seleccionada de entre el grupo que consiste en:
- a) endopep-140 que comprende SEC ID nº 1, un fragmento biológicamente activo de la misma, una variante alélica natural de la misma, o una secuencia que presenta por lo menos 70%, 80%, 90% o 95% de identidad,
  - b) endopep-40 que comprende SEC ID nº 2, un fragmento biológicamente activo de la misma, una variante alélica natural de la misma, o una secuencia que presenta por lo menos 70%, 80%, 90% o 95% de identidad,
  - c) endopep-120 que comprende SEC ID nº 3, un fragmento biológicamente activo de la misma, una variante alélica natural de la misma, o una secuencia que presenta por lo menos 70%, 80%, 90% o 95% de identidad,
  - d) endopep-60 que comprende SEC ID nº 4, un fragmento biológicamente activo de la misma, una variante alélica natural de la misma, o una secuencia que presenta por lo menos 70%, 80%, 90% o 95% de identidad,
  - e) endopep-41 que comprende SEC ID nº 5, un fragmento biológicamente activo de la misma, una variante alélica natural de la misma, o una secuencia que presenta por lo menos 70%, 80%, 90% o 95% de identidad.
2. Composición enzimática según la reivindicación 1, en la que la(s) endopeptidasa(s) se selecciona(n) de entre el grupo que consiste en:
- a) endopep-140 que comprende SEC ID nº 1 o una secuencia que presenta por lo menos 95% de identidad,
  - b) endopep-40 que comprende SEC ID nº 2 o una secuencia que presenta por lo menos 95% de identidad,
  - c) endopep-120 que comprende SEC ID nº 3 o una secuencia que presenta por lo menos 95% de identidad,
  - d) endopep-60 que comprende SEC ID nº 4 o una secuencia que presenta por lo menos 95% de identidad, y
  - e) endopep-41 que comprende SEC ID nº 5 o una secuencia que presenta por lo menos 95% de identidad.
3. Composición enzimática según cualquiera de las reivindicaciones 1 y 2, en la que la endopeptidasa se selecciona de entre el grupo que consiste en:
- a) endopep-140 que comprende SEC ID nº 1 o una secuencia que presenta por lo menos 95% de identidad, y
  - b) endopep-40 que comprende SEC ID nº 2 o una secuencia que presenta por lo menos 95% de identidad.
4. Composición enzimática según cualquiera de las reivindicaciones 1 a 3, en la que la(s) endopeptidasa(s) puede(n) obtenerse a partir de una cepa de *Actinoallomurus*.
5. Composición enzimática según la reivindicación 4, en la que la(s) endopeptidasa(s) puede(n) obtenerse a partir de *Actinoallomurus sp.* DSM24988.
6. Composición enzimática según cualquiera de las reivindicaciones 1 a 5, que puede hidrolizar oligopéptidos del gluten que son resistentes al corte por enzimas gástricos y pancreáticos y cuya presencia en la luz intestinal resulta en efectos tóxicos.
7. Composición enzimática según cualquiera de las reivindicaciones 1 a 6, que comprende uno o más otras enzimas proteolíticas seleccionados de entre prolil-endoproteasas, x-prolil-dipeptidil aminopeptidasas y prolil-aminopeptidasas.
8. Composición enzimática según cualquiera de las reivindicaciones 1 a 7, en la que la(s) endopeptidasa(s) se fusiona(n) funcionalmente con otro polipéptido para formar proteína(s) quimérica(s) o derivada(s).
9. Endopeptidasa aislada de la familia S8/S53 activa a un pH de entre 3 y 8 seleccionada de entre:
- a) endopep-140 que comprende SEC ID nº 1 o una secuencia que presenta por lo menos 95% de identidad,
  - b) endopep-40 que comprende SEC ID nº 2 o una secuencia que presenta por lo menos 95% de identidad,
  - c) endopep-120 que comprende SEC ID nº 3 o una secuencia que presenta por lo menos 95% de identidad,
  - d) endopep-60 que comprende SEC ID nº 4 o una secuencia que presenta por lo menos 95% de identidad, y

e) endopep-41 que comprende SEC ID nº 5 o una secuencia que presenta por lo menos 95% de identidad.

10. Endopeptidasa aislada de la familia S8/S53 según la reivindicación 9, que se selecciona de entre:

a) endopep-140 que comprende SEC ID nº 1 o una secuencia que presenta por lo menos 95% de identidad,

y

b) endopep-40 que comprende SEC ID nº 2 o una secuencia que presenta por lo menos 95% de identidad.

11. Composición enzimática según cualquiera de las reivindicaciones 1 a 8 o endopeptidasa aislada de la familia S8/S53 según cualquiera de las reivindicaciones 9 y 10 para la utilización como un medicamento para el tratamiento o la prevención de la celiaquía, la dermatitis herpetiforme y/o cualquier otro trastorno asociado a la intolerancia al gluten.

12. Utilización de una composición enzimática según cualquiera de las reivindicaciones 1 a 8 o por lo menos una endopeptidasa aislada de la familia S8/S53 según cualquiera de las reivindicaciones 9 y 10 para producir hidrolizados de proteínas utilizados para alimentos y bebidas.

13. Formulación farmacéutica que comprende como principio proteolítico activo, una composición enzimática según cualquiera de las reivindicaciones 1 a 8 o por lo menos una endopeptidasa aislada de la familia S8/S53 según cualquiera de las reivindicaciones 9 y 10.

14. Formulación farmacéutica según la reivindicación 13 que es una formulación farmacéutica oral.

15. Complemento alimenticio que comprende como principio proteolítico activo una composición enzimática según cualquiera de las reivindicaciones 1 a 8 o por lo menos una endopeptidasa aislada de la familia S8/S53 según cualquiera de las reivindicaciones 9 y 10.

16. Ácido nucleico aislado que codifica por lo menos una endopeptidasa según cualquiera de las reivindicaciones 1 a 10 que comprende por lo menos una secuencia polinucleotídica seleccionada de entre SEC ID nº 7, 8, 9, 10 y 11 o por lo menos una secuencia polinucleotídica que presenta por lo menos 70%, 80%, 90% o 95% de identidad respecto a cualquiera de SEC ID nº 7, 8, 9, 10 y 11.

17. Ácido nucleico aislado según la reivindicación 16 que comprende por lo menos una secuencia polinucleotídica que presenta por lo menos 95% de identidad respecto a cualquiera de SEC ID nº 7, 8, 9, 10 y 11.

18. Ácido nucleico aislado según cualquiera de las reivindicaciones 16 y 17 que comprende por lo menos una secuencia polinucleotídica seleccionada de entre SEC ID nº 7 y 8.

19. Procedimiento para producir una composición enzimática según la reivindicación 1, que comprende:

A) cultivar una cepa de *Actinoallomurus* natural que puede producir por lo menos una endopeptidasa de la familia S8/S53 según la reivindicación 1 en un medio de cultivo bajo unas condiciones adecuadas para producir la composición enzimática y recuperar la composición enzimática a partir del lote de cultivo, o

B) cultivar una cepa de *Actinoallomurus* derivada mediante procedimientos convencionales de mutación y/o selección a partir de una cepa natural según se define en A), que mantiene la capacidad de producir por lo menos una endopeptidasa de la familia S8/S53 como se define en la reivindicación 1, en un medio de cultivo bajo unas condiciones adecuadas para producir la composición enzimática y recuperar la composición enzimática a partir del lote de cultivo, o

C) introducir en una célula hospedadora un ácido nucleico que codifica por lo menos una endopeptidasa de la familia S8/S53 seleccionada de entre el grupo que consiste en:

a) endopep-140 que comprende SEC ID nº 1, un fragmento biológicamente activo de la misma, una variante alélica natural de la misma o una secuencia que presenta por lo menos 70%, 80%, 90% o 95% de identidad,

b) endopep-40 que comprende SEC ID nº 2, un fragmento biológicamente activo de la misma, una variante alélica natural de la misma, o una secuencia que presenta por lo menos 70%, 80%, 90% o 95% de identidad,

c) endopep-120 que comprende SEC ID nº 3, un fragmento biológicamente activo de la misma, una variante alélica natural de la misma, o una secuencia que presenta por lo menos 70%, 80%, 90% o 95% de

identidad,

5 d) endopep-60 que comprende SEC ID nº 4, un fragmento biológicamente activo de la misma, una variante alélica natural de la misma, o una secuencia que presenta por lo menos 70%, 80%, 90% o 95% de identidad, y

10 e) endopep-41 que comprende SEC ID nº 5, un fragmento biológicamente activo de la misma, una variante alélica natural de la misma, o una secuencia que presenta por lo menos 70%, 80%, 90% o 95% de identidad,

cultivar la célula en un medio de cultivo bajo unas condiciones adecuadas para producir la(s) endopeptidasa(s) y recuperar la(s) endopeptidasa(s) a partir del lote de cultivo.

15 20. Procedimiento según la reivindicación 19, párrafo A) o B), en el que la cepa de *Actinoallomurus* se sustituye por una cepa perteneciente al género *Catenulispora*, *Ktedonobacter* o *Streptomyces* que puede producir por lo menos una endopeptidasa según la reivindicación 1.

20 21. Procedimiento según la reivindicación 19, en el que la cepa de *Actinoallomurus* es *Actinoallomurus* sp. DSM 24988.

22. Procedimiento según la reivindicación 19, párrafo C), en el que el ácido nucleico comprende por lo menos una de las secuencias polinucleotídicas identificadas en cualquiera de las reivindicaciones 16 y 17.

25 23. Procedimiento según la reivindicación 22, en el que el ácido nucleico comprende por lo menos una de las secuencias polinucleotídicas SEC ID nº 7 y 8 o por lo menos una de las secuencias polinucleotídicas que presenta una identidad de por lo menos 95% respecto a cualquiera de las SEC ID nº 7 y 8.

30 24. Procedimiento según cualquiera de las reivindicaciones 19, párrafo C), 22 y 23, en el que la célula hospedadora es un microorganismo seleccionado de entre *Bacillus*, *Streptomyces*, *Lactobacillus*, *Pyrococcus*, *Pseudomonas* y *Escherichia coli*.

25. Cepa *Actinoallomurus* DSM 24988.

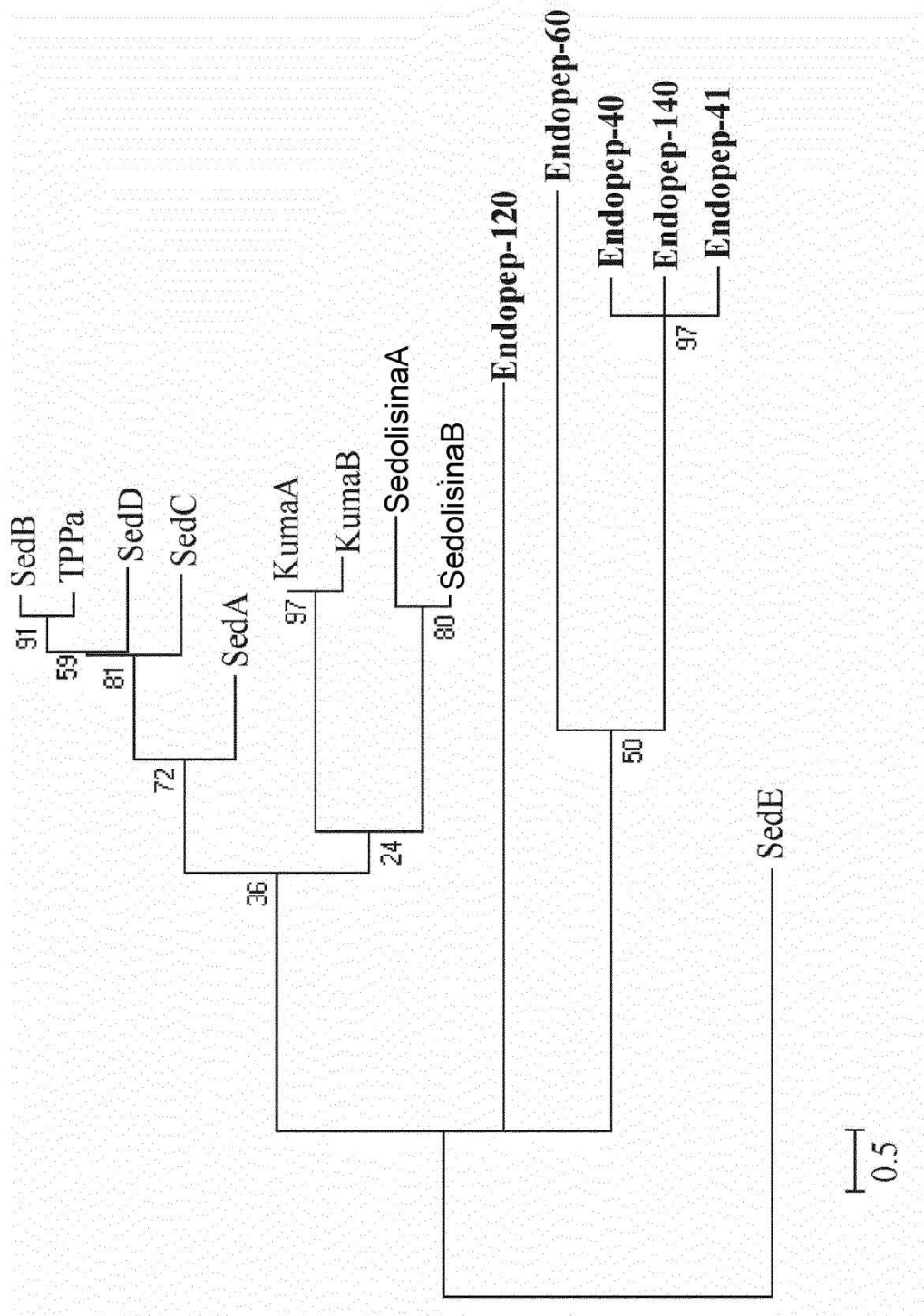
35 26. Composición enzimática según cualquiera de las reivindicaciones 1 a 8 o por lo menos una endopeptidasa aislada según cualquiera de las reivindicaciones 9 y 10 para la utilización en la degradación de oligopéptidos del gluten que son resistentes al corte por enzimas gástricos y pancreáticos y cuya presencia en la luz interna resulta en efectos tóxicos.

40 27. Composición enzimática según cualquiera de las reivindicaciones 1 a 8 o por lo menos una endopeptidasa aislada según cualquiera de las reivindicaciones 9 y 10 para la utilización según la reivindicación 26 en el tratamiento o la prevención de la celiaquía, la dermatitis herpetiforme y/o cualquier otro trastorno asociado a la intolerancia al gluten.

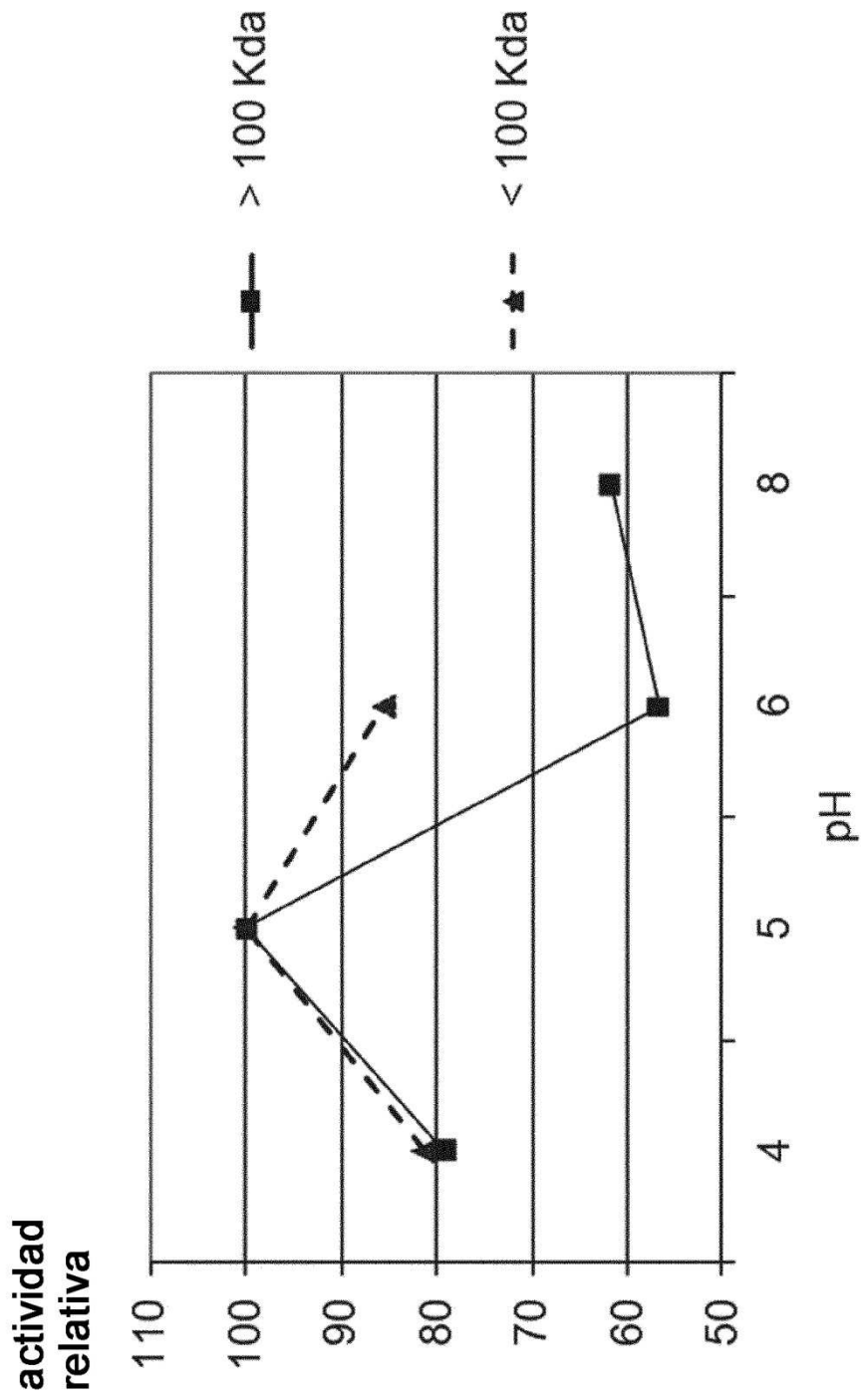
45 28. Composición enzimática según cualquiera de las reivindicaciones 1 a 8 o por lo menos una endopeptidasa aislada según cualquiera de las reivindicaciones 9 y 10 para la utilización según la reivindicación 27, en la que la composición enzimática o la endopeptidasa aislada se incorpora en una formulación farmacéutica, complemento alimenticio, bebida o producto para beber.

50 29. Utilización de la composición enzimática según cualquiera de las reivindicaciones 1 a 8 o por lo menos una endopeptidasa aislada según cualquiera de las reivindicaciones 9 y 10 en la preparación de complementos alimenticios adecuados para hidrolizar oligopéptidos del gluten que son resistentes al corte por enzimas gástricos y pancreáticos y cuya presencia en la luz interna resulta en efectos tóxicos.

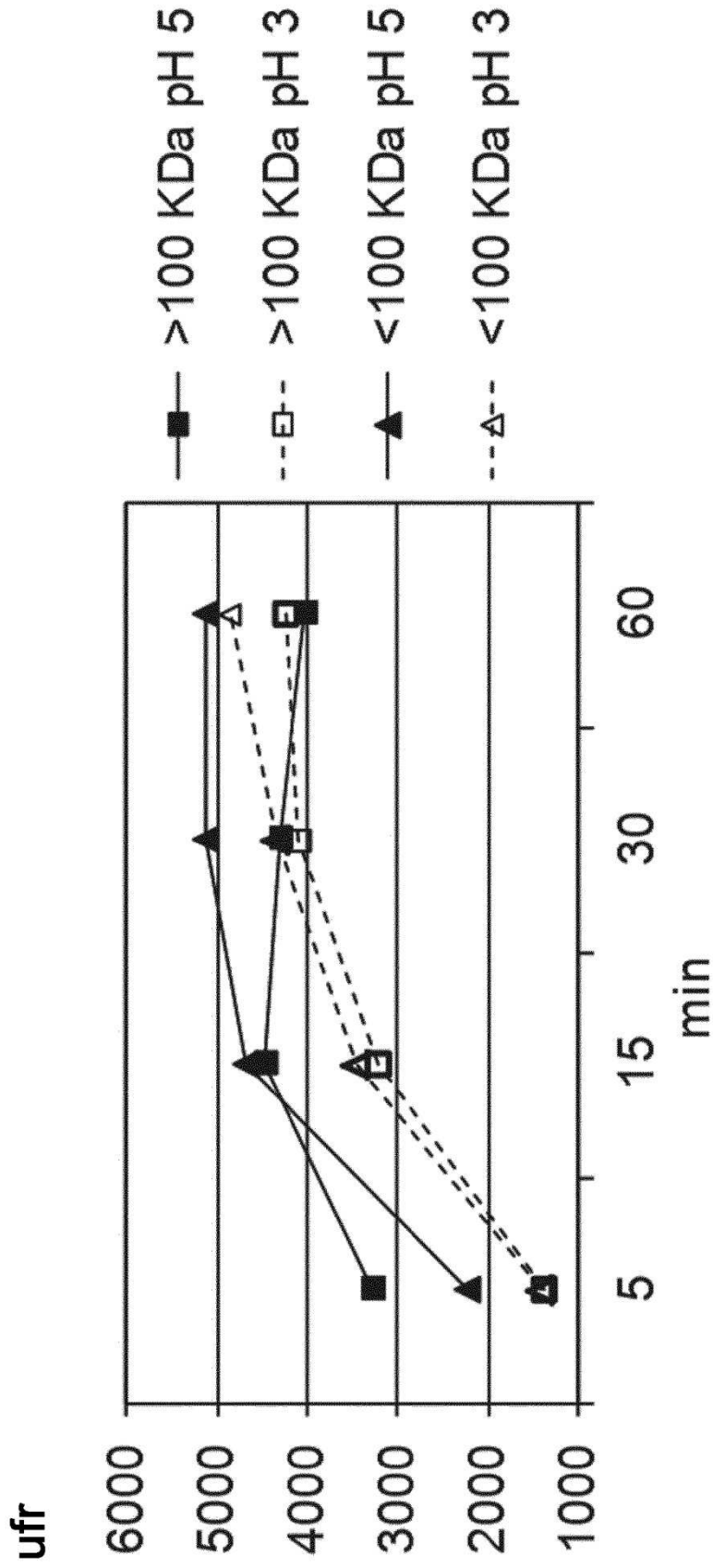
55 30. Utilización según la reivindicación 29 para el tratamiento de productos alimenticios líquidos, en la que la composición enzimática o la endopeptidasa aislada se encuentra en forma inmovilizada.



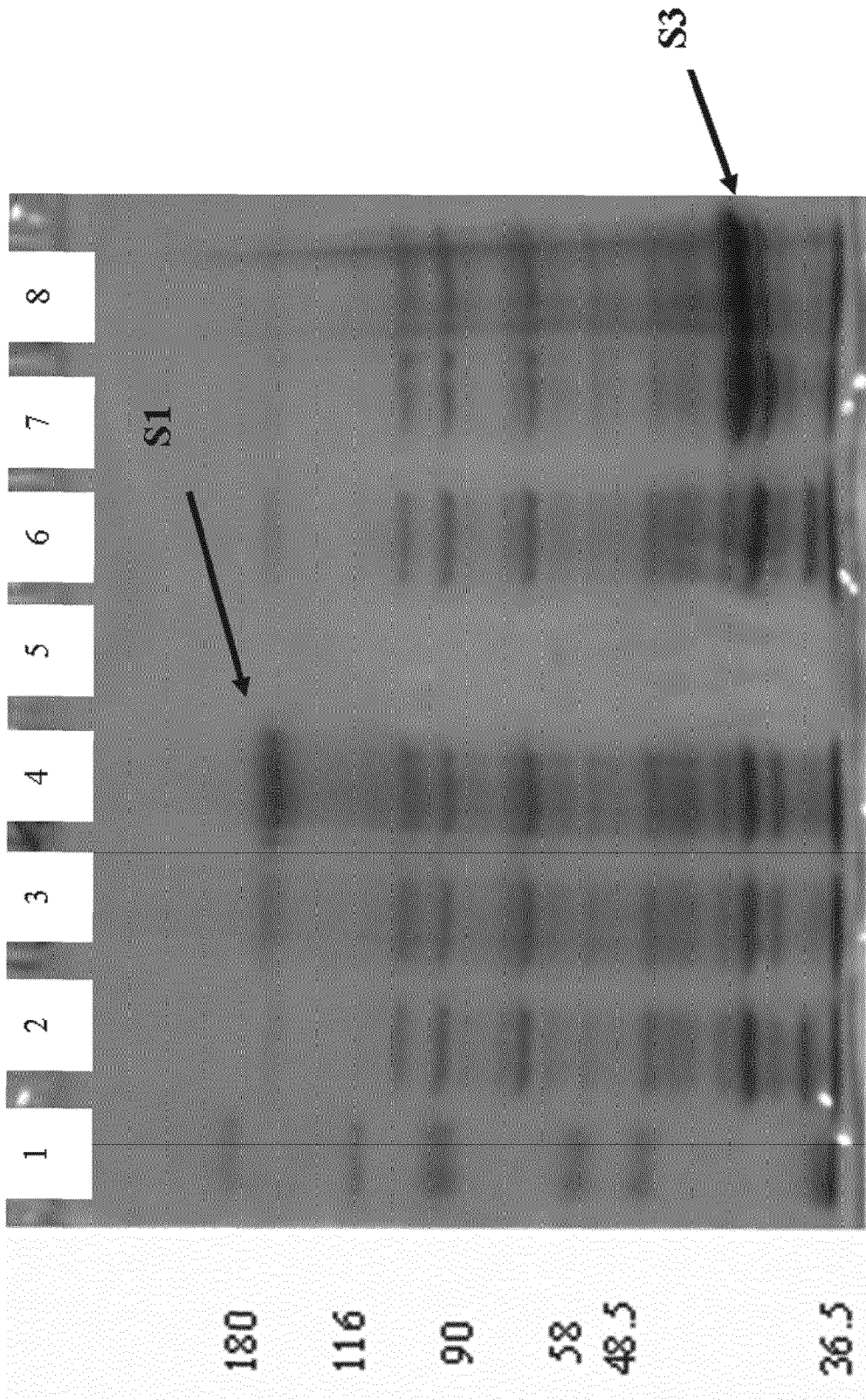
**Fig. 1**



*Fig. 2*

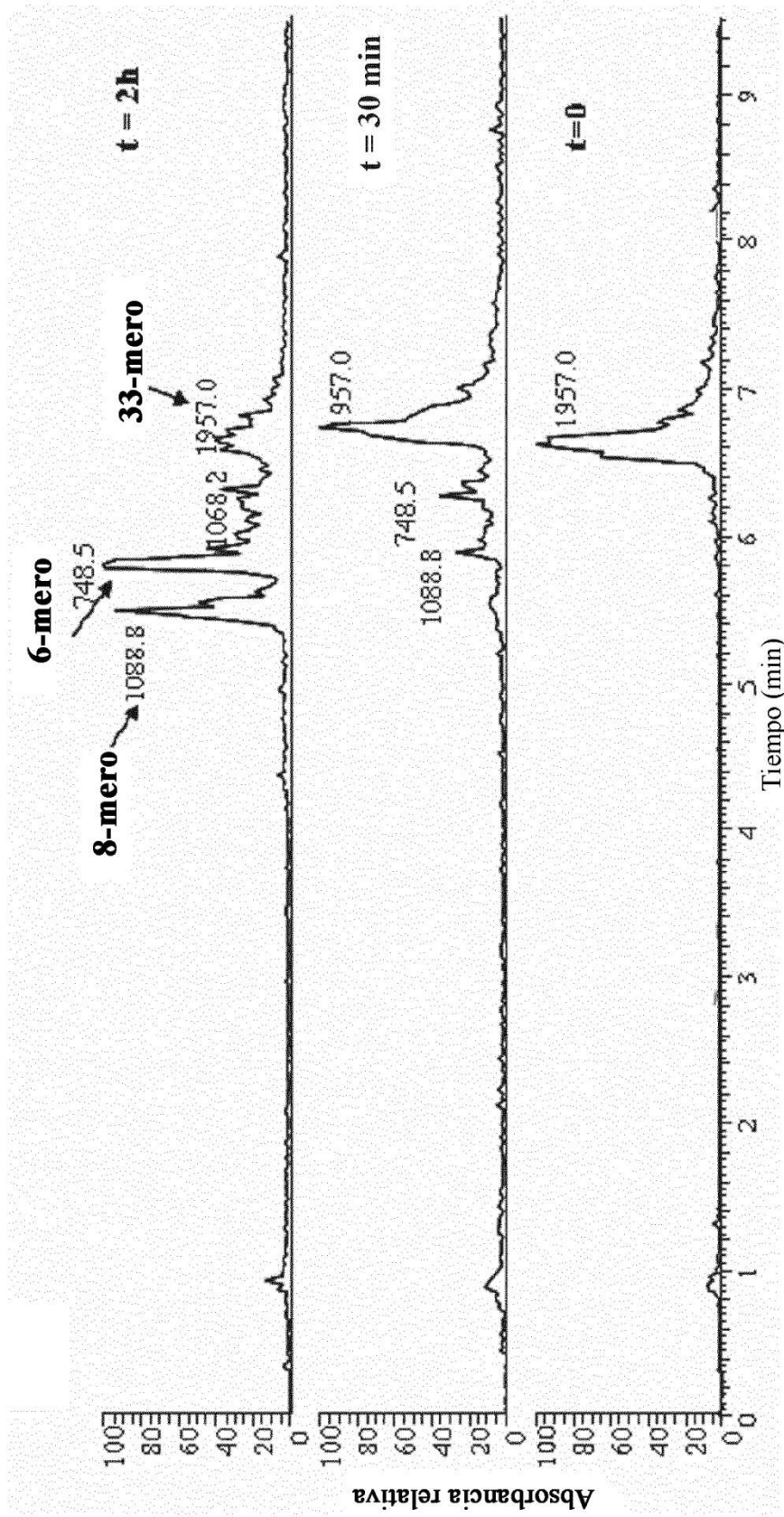


*Fig. 3*

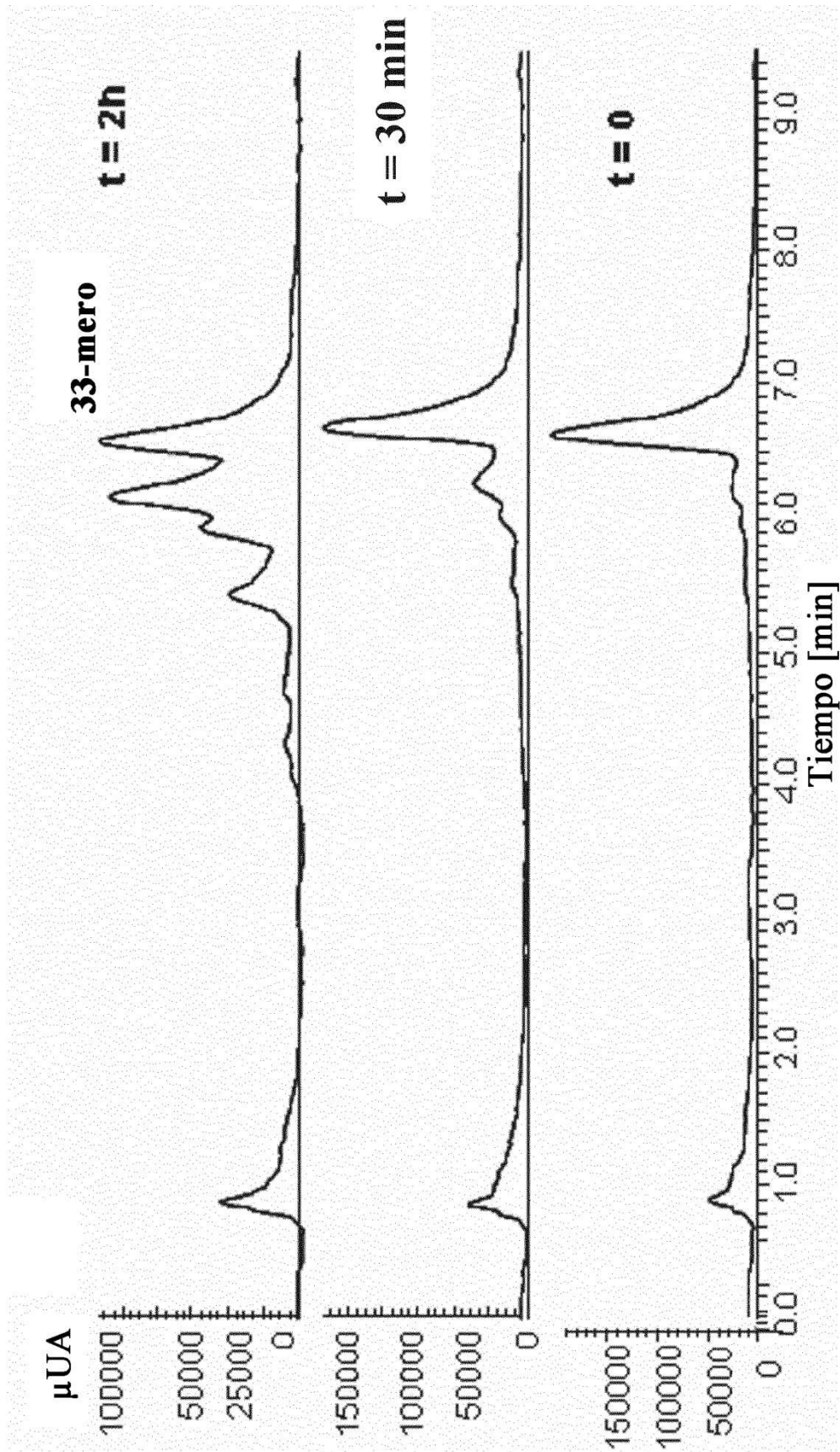


**Fig. 4**

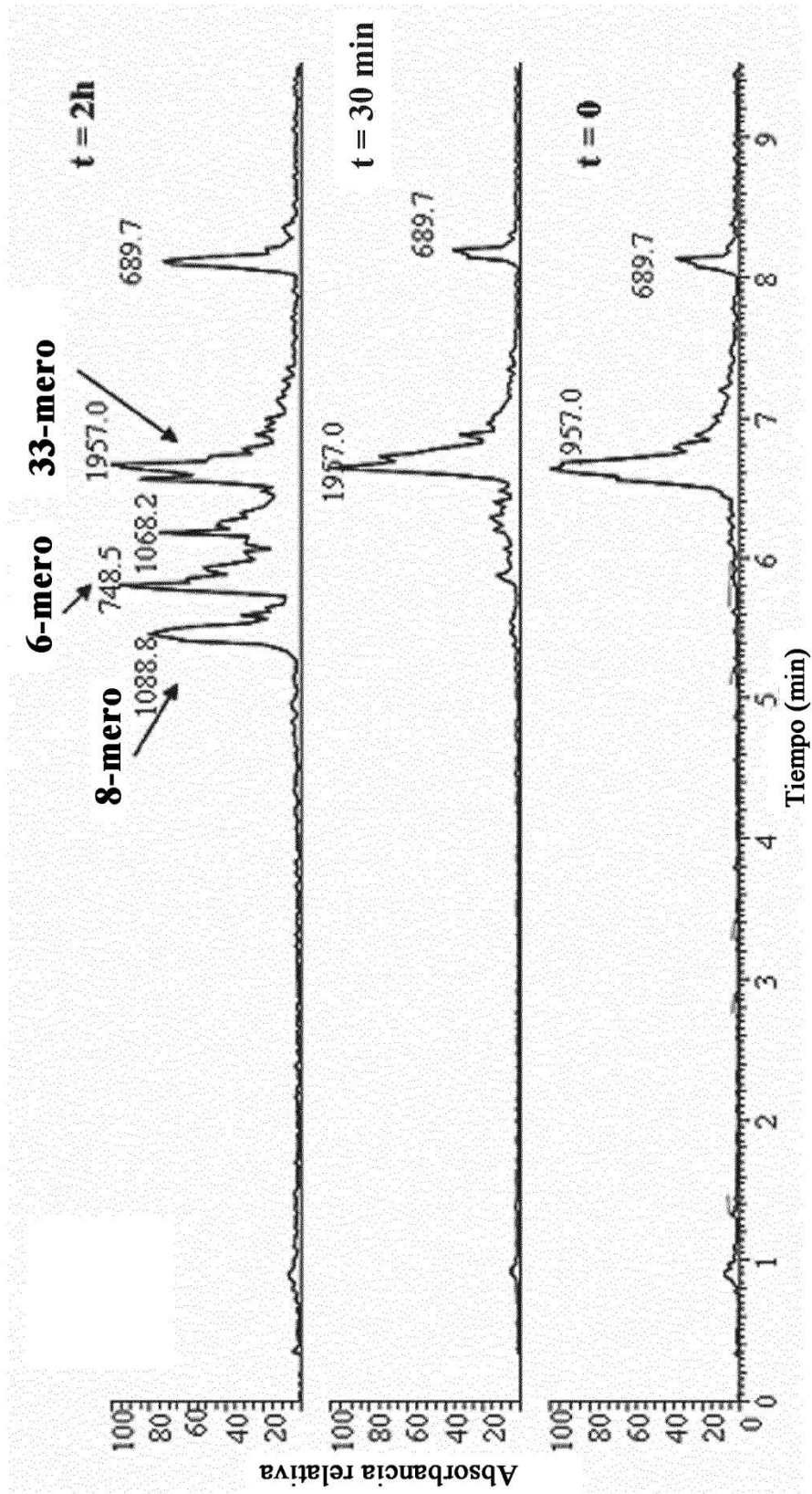




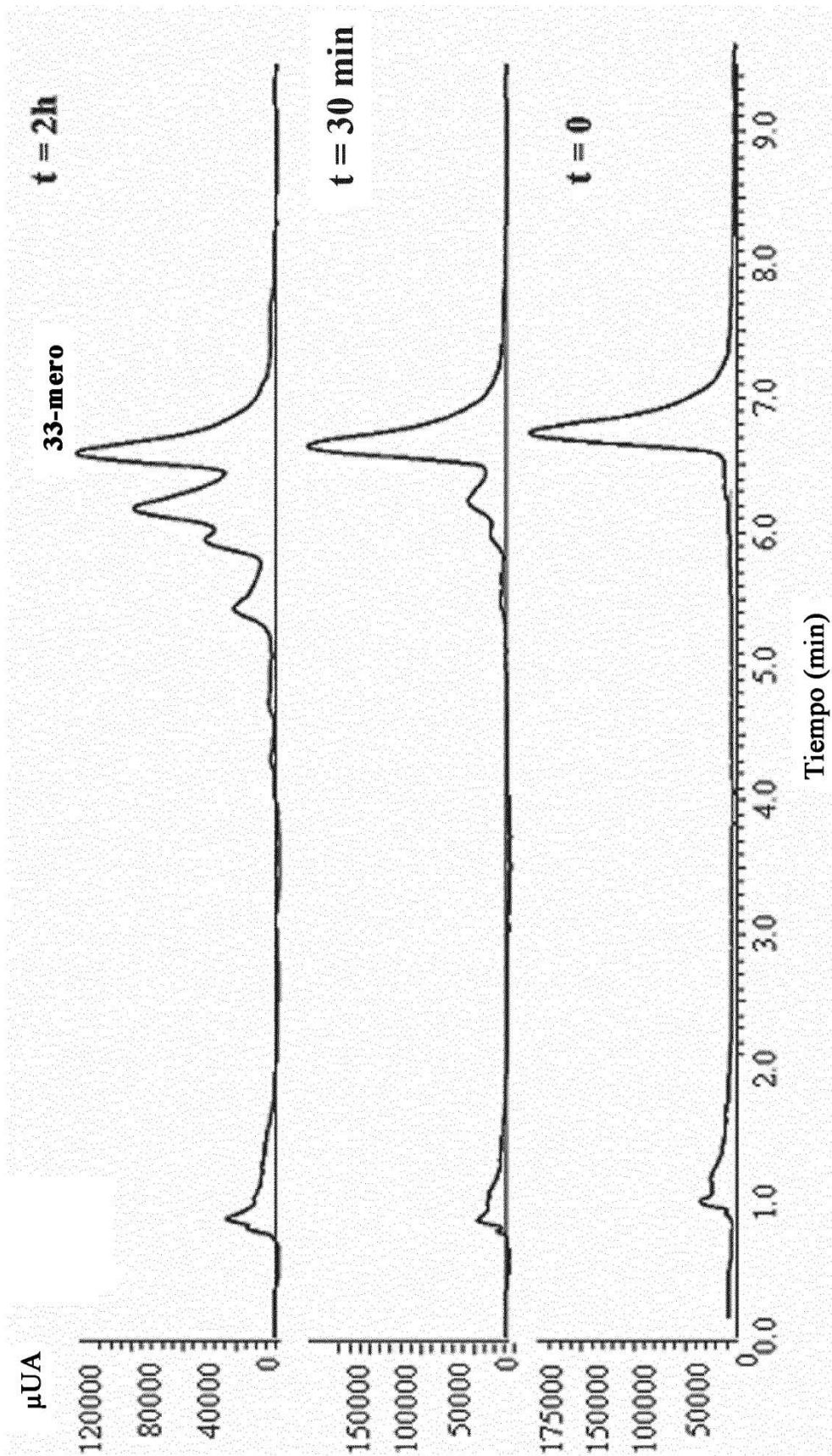
*Fig. 5A*



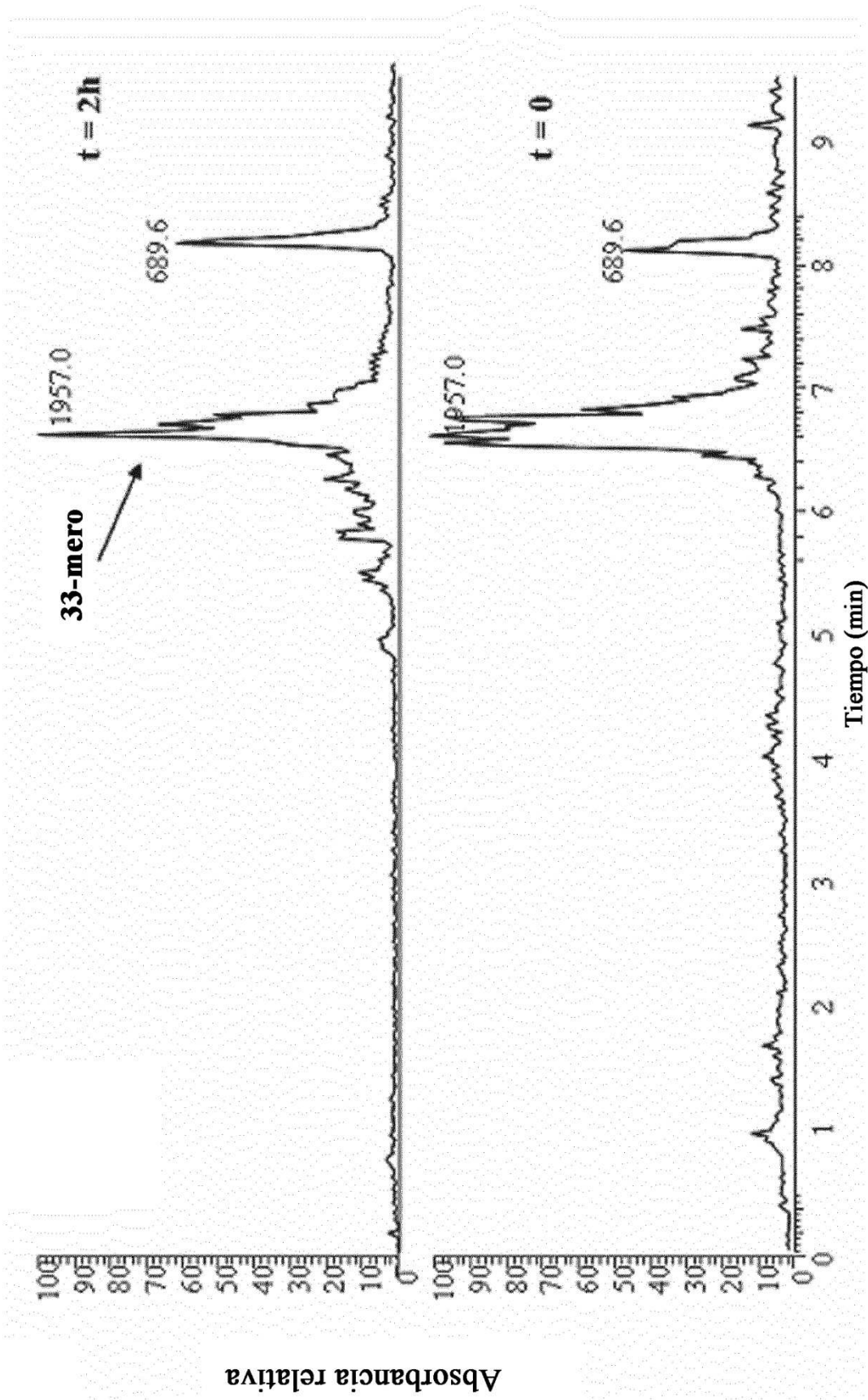
*Fig. 5B*



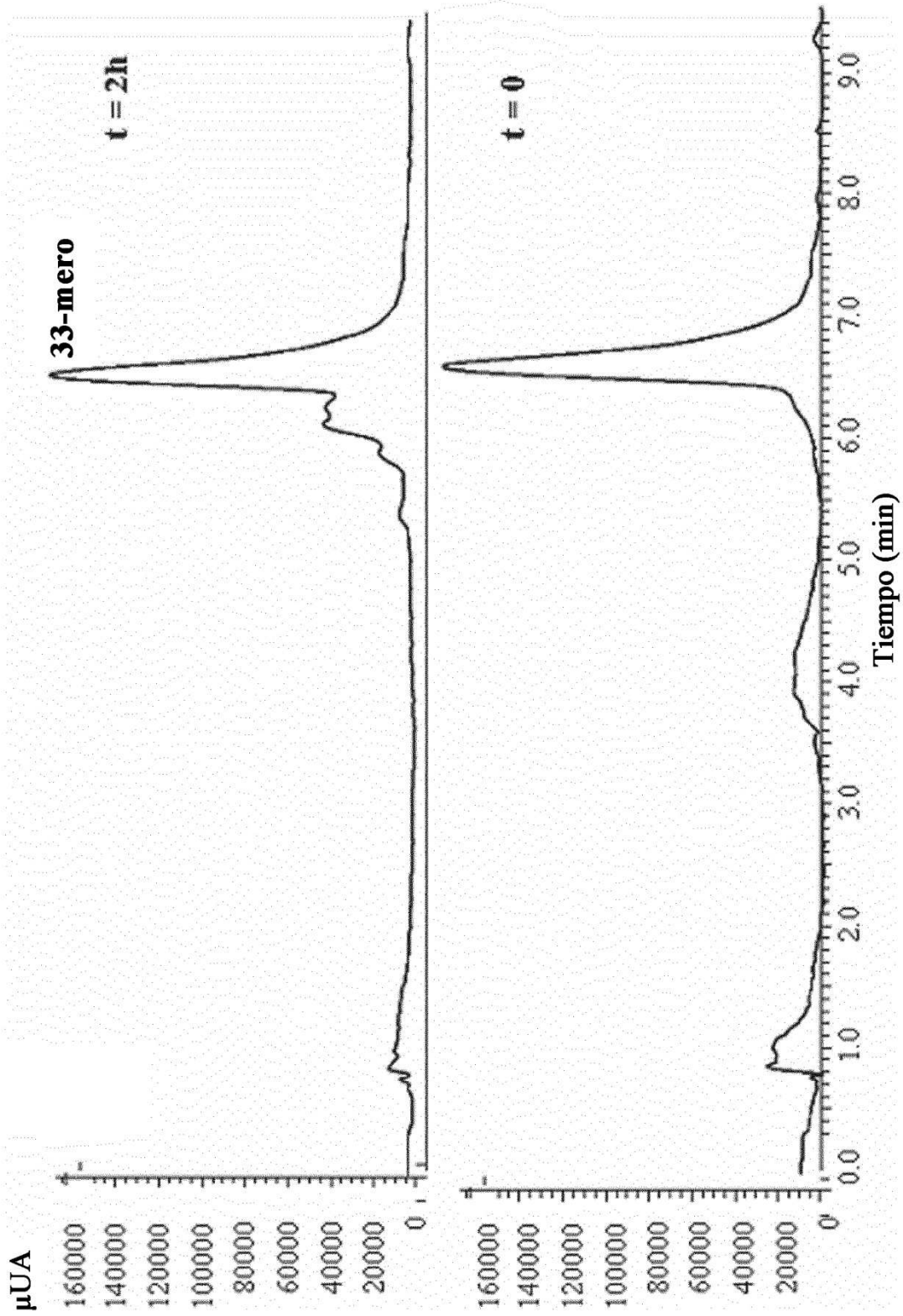
**Fig. 6A**



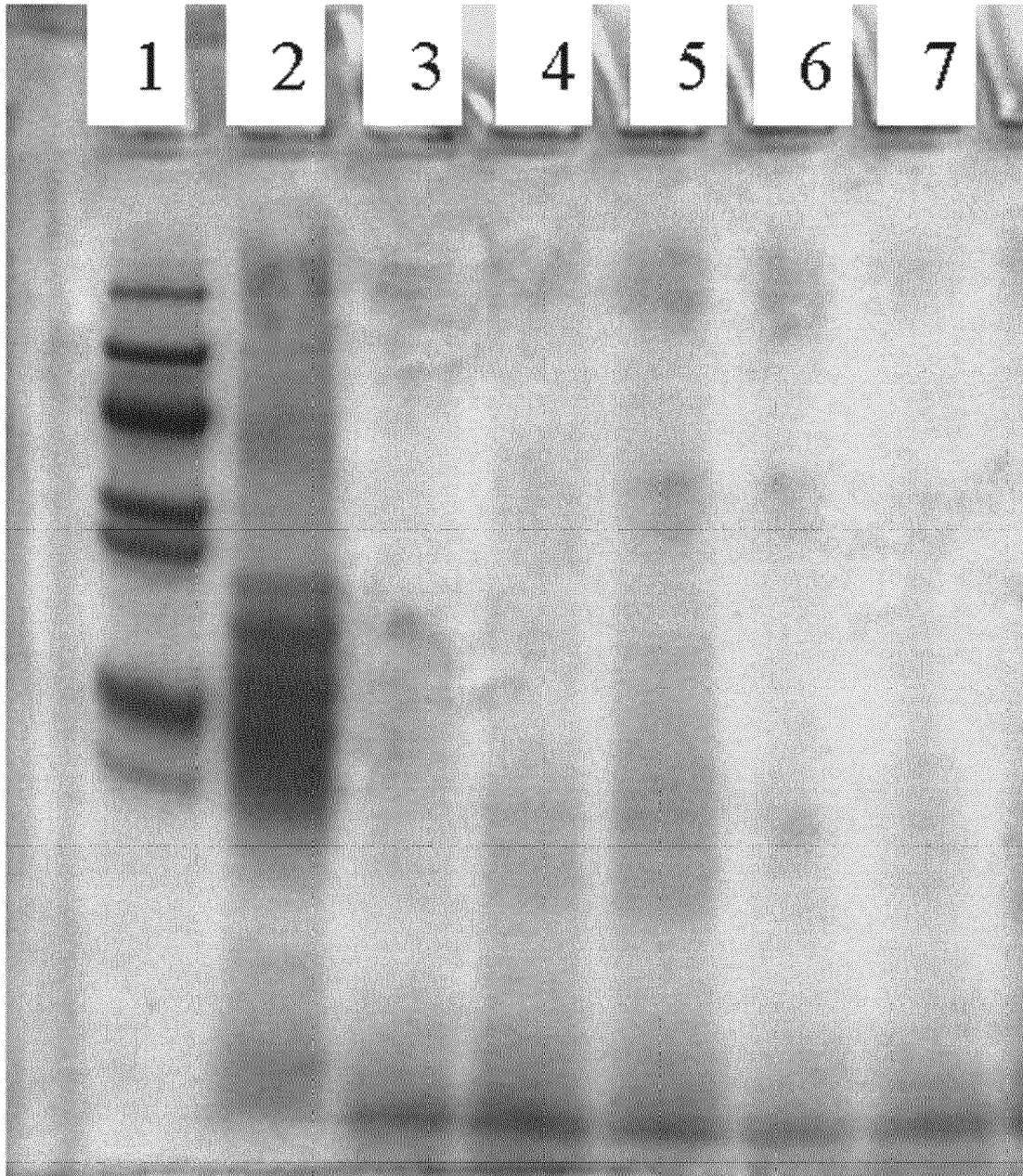
**Fig. 6B**



**Fig. 7A**



**Fig. 7B**



*Fig. 8*