



19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 366 830**

51 Int. Cl.:  
**C12N 9/48** (2006.01)  
**A61K 38/48** (2006.01)  
**A61K 38/54** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **04786408 .7**  
96 Fecha de presentación : **25.08.2004**  
97 Número de publicación de la solicitud: **1673457**  
97 Fecha de publicación de la solicitud: **28.06.2006**

54 Título: **Nuevas proteínas fúngicas y ácidos nucleicos que codifican las mismas.**

30 Prioridad: **25.08.2003 US 498318 P**

45 Fecha de publicación de la mención BOPI:  
**25.10.2011**

45 Fecha de la publicación del folleto de la patente:  
**25.10.2011**

73 Titular/es: **FUNZYME BIOTECHNOLOGIES SA**  
**1 Chemin d'Alcire**  
**1228 Plan-les-Ouates, CH**

72 Inventor/es: **Stocklin, Reto;**  
**Grouzmann, Eric y**  
**Monod, Michael**

74 Agente: **Elzaburu Márquez, Alberto**

ES 2 366 830 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 proteínas fúngicas y ácidos nucleicos que codifican las mismas.

## CAMPO DE LA INVENCIÓN

5 La presente descripción se refiere a nuevos polipéptidos y a los ácidos nucleicos que los codifican, que tienen propiedades catalíticas únicas. Más particularmente, la descripción se refiere a ácidos nucleicos que codifican nueva leucina aminopeptidasa (LAP) y otros polipéptidos amino y carboxi-peptidasas, que se referirán en la presente memoria colectivamente como EXOX, así como a vectores, células anfitrionas, anticuerpos y métodos recombinantes para producir estos ácidos nucleicos y polipéptidos. Estos genes se han identificado en dos especies fúngicas diferentes, *Trichophyton rubrum* y *Aspergillus fumigatus*.

## 10 ANTECEDENTES DE LA INVENCIÓN

Las bacterias, levaduras y hongos filamentosos, así como las células especializadas de las plantas, invertebrados y vertebrados, expresan proteínas de membrana útiles para el transporte de aminoácidos, dipéptidos y tripéptidos. Lubkowitz *et al.*, *Microbiology* 143:387-396 (1997); Hauser *et al.*, *Mol. Membr. Biol.* 18(1):105-112 (2001); Stacey *et al.*, *Trends Plant Sci.* 7(6):257-263 (2002); Rubio-Aliaga y Daniel, *Trends Pharmacol. Sci.* 23(9):434-440 (2002). En levaduras, hongos filamentosos y plantas también se conocen transportadores que también aceptan oligopéptidos grandes (4-5 restos de aminoácidos). La digestión de proteínas en aminoácidos se ha investigado en microorganismos usados en la industria de fermentación de alimentos. Las bacterias del género *Lactobacillus* (O'Cuinn *et al.*, *Biochem. Soc. Trans.* 27(4):730-734 (1999)) y los hongos del género *Aspergillus* (Doumas *et al.*, *Appl. Environ. Microbiol.* 64:4809-4815 (1998)) secretan endoproteasas y exoproteasas, que cooperan muy eficazmente en la digestión de proteínas.

La actividad aminopeptidasa, que también puede jugar un papel en el desarrollo de los hongos durante la infección, se ha detectado en el micelio y el sobrenadante del cultivo de una especie de hongos (De Bersaques y Dockx, *Arch. Belg. Dermatol. Syphiligr.* 29:135-140 (1973); Danew y Friedrich, *Mykosen* 23:502-511 (1980)), sin embargo, hasta el momento no se ha aislado ni caracterizado ninguna aminopeptidasa o carboxipeptidasa de dermatofitos.

25 La descripción se basa en parte en el descubrimiento de polipéptidos aislados que contienen la forma madura de una secuencia de aminoácidos seleccionada de SEQ ID NOs: 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33 y 35. La descripción también proporciona polipéptidos aislados que contienen una secuencia de aminoácidos seleccionada de SEQ ID NOs. 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33 y 35, así como polipéptidos aislados que son al menos 90% idénticos a los polipéptidos que tienen estas secuencias, en el que el polipéptido opcionalmente tiene actividad aminopeptidasa o carboxipeptidasa. Por ejemplo, el polipéptido puede ser una leucina aminopeptidasa tal como ruLAP2.

También se proporcionan polipéptidos aislados que tienen una o más sustituciones de aminoácidos conservativas. Dichos polipéptidos pueden poseer actividad aminopeptidasa.

35 La descripción también engloba polipéptidos que son variantes alélicas naturales de la secuencia seleccionada del grupo que consiste en SEQ ID NOs: 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33 y 35. Estas variantes alélicas incluyen secuencias de aminoácidos que son las traducciones de secuencias de ácido nucleico que se diferencian en uno o más nucleótidos de las secuencias de ácido nucleico seleccionadas del grupo que consiste en SEQ ID NOs: 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33 y 35. El polipéptido variante en el que cualquier aminoácido cambiado en la secuencia elegida se cambia para proporcionar una sustitución conservativa.

40 También se describe un método para eliminar aminoácidos particulares de péptidos, por ejemplo etiquetas de proteínas recombinantes, en el que el polipéptido activo que elimina el aminoácido es un polipéptido que tiene una secuencia de aminoácidos al menos 90% idéntica a un polipéptido que tiene la secuencia de aminoácidos seleccionada del grupo que consiste en SEQ ID NOs: 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33 y 35, o un fragmento biológicamente activo de éste.

45 Cualquiera de los polipéptidos puede ser natural. Además, cualquiera de estos polipéptidos puede estar en una composición que incluye un vehículo, y la composición puede estar en un kit que incluye uno o más contenedores.

También se proporcionan dermatofitos que contienen los polipéptidos. Por ejemplo, los dermatofitos adecuados incluyen *Epidermophyton floccosum*, *Microsporium audouinii*, *Microsporium ferrugineum*, *Trichophyton concentricum*, *Trichophyton kanei*, *Trichophyton megninii*, *Trichophyton mentagrophytes*, *Trichophyton raubitschekii*, *Trichophyton rubrum*, *Trichophyton schoenleinii*, *Trichophyton soudanense*, *Trichophyton tonsurans*, *Trichophyton violaceum*, *Trichophyton yaoundei*, *Microsporium canis*, *Microsporium equinum*, *Microsporium nanum*, *Microsporium persicolor*, *Trichophyton equinum*, *Trichophyton simii*, *Trichophyton verrucosum*, *Microsporium gypseum*, *Trichophyton ajelloi* y *Trichophyton terrestre*.

También se describen sobrenadantes de cultivo microbiano que contienen los polipéptidos de la invención.

La descripción también se refiere al uso de terapéuticos en la fabricación de un medicamento para tratar un síndrome asociado con una enfermedad humana, en el que el terapéutico incluye los polipéptidos descritos y la enfermedad se selecciona de una patología asociada con estos polipéptidos.

5 La descripción también se refiere a métodos para degradar un sustrato polipeptídico. Dichos métodos incluyen poner en contacto el sustrato polipeptídico con uno o más de los polipéptidos, que se han aislado. Por ejemplo, el sustrato polipeptídico puede ser una proteína de longitud completa. Además, el uno o más polipéptidos aislados pueden usarse para digerir secuencialmente el sustrato polipeptídico. El sustrato polipeptídico puede seleccionarse de caseína desnaturalizada, gliadina, gluten, albúmina de suero bovino o fragmentos de éstos. Por ejemplo, el polipéptido aislado puede ser una aminopeptidasa, que puede ser una leucina aminopeptidasa tal como ruLAP2.

10 La descripción también se refiere a métodos para identificar un agente terapéutico potencial para usarse en el tratamiento de infecciones fúngicas, en el que la infección fúngica está relacionada con la expresión aberrante o interacciones fisiológicas aberrantes de los polipéptidos de la invención. Dichos métodos incluyen proporcionar una célula que expresa el polipéptido y que tiene una propiedad o función atribuible al polipéptido, poner en contacto la célula con una composición que comprende una sustancia candidata, y determinar si la sustancia altera la propiedad o función atribuible al polipéptido. Si no se observa ninguna alteración en presencia de la sustancia cuando la célula se pone en contacto con una composición en ausencia de la sustancia, la sustancia se identifica como agente terapéutico potencial. Por ejemplo, la propiedad o función atribuible al polipéptido puede ser actividad aminopeptidasa o carboxipeptidasa.

20 La descripción también se refiere a métodos para tratar un estado patológico en un mamífero por la administración de un polipéptido al mamífero en una cantidad que es suficiente para aliviar el estado patológico. Típicamente, el polipéptido tiene una secuencia de aminoácidos al menos 90% idéntica a un polipéptido que contiene la secuencia de aminoácidos seleccionada de SEQ ID NOs: 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33 y 35, o un fragmento biológicamente activo de éste. El estado patológico que se va a tratar incluye una infección fúngica, enfermedad celíaca, mala absorción en el tracto digestivo, esprúe, una reacción alérgica y una deficiencia enzimática. Por ejemplo, la reacción alérgica puede ser una reacción al gluten.

25 Además se describen métodos para tratar un estado patológico en un mamífero por la administración de un inhibidor de proteasa al mamífero en una cantidad que es suficiente para aliviar el estado patológico. El inhibidor de proteasa incluye una secuencia de aminoácidos al menos 90% idéntica a un polipéptido que contiene la secuencia de aminoácidos seleccionada de SEQ ID NOs: 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33 y 35, o un fragmento biológicamente activo de éste. Por ejemplo, el estado patológico puede ser una infección fúngica.

30 La descripción se refiere además a polipéptidos aislados que tienen una secuencia de aminoácidos seleccionada de SEQ ID NOs: 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33 y 35. Estos polipéptidos pueden producirse cultivando una célula en condiciones que dan lugar a la expresión del polipéptido. En algunas realizaciones, la célula incluye un vector que contiene una molécula aislada de ácido nucleico que tiene una secuencia de ácido nucleico seleccionada del grupo que consiste en SEQ ID NOs: 2, 5, 8, 11, 14, 17, 20, 23, 26, 29, 32 y 34. Opcionalmente, la célula puede ser una célula fúngica, una célula bacteriana, una célula de insecto (con o sin un baculovirus), una célula de planta y una célula de mamífero.

35 La descripción también proporciona moléculas aisladas de ácido nucleico que contienen una secuencia de ácido nucleico seleccionada de SEQ ID NOs: 2, 5, 8, 11, 14, 17, 20, 23, 26, 29, 32 y 34. Por ejemplo, dichas moléculas de ácido nucleico pueden ser naturales.

40 Además se describen moléculas de ácido nucleico que se diferencian en un único nucleótido de una secuencia de ácido nucleico seleccionada de SEQ ID NOs: 2, 5, 8, 11, 14, 17, 20, 23, 26, 29, 32 y 34 así como a las moléculas aisladas de ácido nucleico que codifican la forma madura de un polipéptido que tiene una secuencia de aminoácidos seleccionada del grupo que consiste en SEQ ID NOs: 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33 y 35. Además, las moléculas de ácido nucleico pueden ser unas que hibridan en condiciones astringentes con la secuencia de nucleótidos seleccionada del grupo que consiste en SEQ ID NOs: 2, 5, 8, 11, 14, 17, 20, 23, 26, 29, 32 y 34 o un complemento de esa secuencia de nucleótidos. Las moléculas de ácido nucleico pueden estar incluidas en un vector, que incluye además un promotor unido de manera operativa con dicha molécula de ácido nucleico. También se proporcionan células que incluyen el vector.

45 La descripción también proporciona métodos para producir los polipéptidos de la invención. Los métodos incluyen cultivar una célula en condiciones que den lugar a la expresión del polipéptido y la célula incluye un vector que tiene una molécula aislada de ácido nucleico que contiene una secuencia de ácido nucleico seleccionada de SEQ ID NOs: 2, 5, 8, 11, 14, 17, 20, 23, 26, 29, 32 y 34. En algunos casos, la célula se selecciona de una célula fúngica, una célula bacteriana, una célula de insecto, una célula de planta o célula de mamífero.

50 La descripción también se refiere a métodos para producir una proteína cultivando un dermatofito que contiene el polipéptido en condiciones suficientes para la producción de la proteína y aislando la proteína del cultivo de dermatofito. Por ejemplo, la proteína puede ser una proteína secretada. Asimismo, la proteína también puede ser una aminopeptidasa o una carboxipeptidasa. Específicamente, la aminopeptidasa puede ser una leucina

aminopeptidasa, tal como ruLAP2. Adicionalmente, el dermatofito puede seleccionarse de *Epidermophyton floccosum*, *Microsporum audouinii*, *Microsporum ferrugineum*, *Trichophyton concentricum*, *Trichophyton kanei*, *Trichophyton megninii*, *Trichophyton mentagrophytes*, *Trichophyton raubitschekii*, *Trichophyton rubrum*, *Trichophyton schoenleinii*, *Trichophyton soudanense*, *Trichophyton tonsurans*, *Trichophyton violaceum*, *Trichophyton yaoundei*, *Microsporum canis*, *Microsporum equinum*, *Microsporum nanum*, *Microsporum persicolor*, *Trichophyton equinum*, *Trichophyton mentagrophytes*, *Trichophyton simii*, *Trichophyton verrucosum*, *Microsporum gypseum*, *Trichophyton ajelloi* y *Trichophyton terrestre*.

Las proteínas producidas pueden aplicarse a sustratos polipeptídicos. En algunos casos, la proteína producida puede degradar el polipéptido o puede digerir secuencialmente una sustancia polipeptídica de longitud completa. Opcionalmente, la longitud del el sustrato polipeptídico puede ser de 2 a 200 aminoácidos.

En algunos casos, la proteína producida añade uno o más aminoácidos al sustrato polipeptídico. En otros casos, la proteína producida elimina uno o más aminoácidos del sustrato polipeptídico para formar un sustrato polipeptídico modificado y la proteína producida añade posteriormente uno o más aminoácidos al sustrato polipeptídico modificado, formando de esta manera un producto polipeptídico que comprende una secuencia de aminoácidos diferente de la del sustrato polipeptídico.

La descripción también proporciona métodos para tratar micosis en un paciente que las sufre. Dichos métodos incluyen administrar una cantidad eficaz de un inhibidor con la actividad de una proteína EXOX seleccionada de SEQ ID NOs: 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33 y 35. Por ejemplo, la proteína EXOX puede incluir la SEQ ID NO: 2.

La descripción proporciona además métodos para degradar un sustrato polipeptídico. Estos métodos incluyen poner en contacto el sustrato polipeptídico con uno o más de los polipéptidos aislados. Opcionalmente, el sustrato polipeptídico es una proteína de longitud completa, y el uno o más polipéptidos aislados pueden ser polipéptidos que digieran secuencialmente el sustrato polipeptídico. El sustrato polipeptídico puede seleccionarse de caseína desnaturalizada, gliadina, gluten, albúmina de suero bovino o fragmentos de éstos. Por ejemplo, en algunos casos, el polipéptido aislado es una aminopeptidasa. La aminopeptidasa puede ser una leucina aminopeptidasa, tal como ruLAP2.

Adicionalmente, el método opcionalmente poner en contacto el sustrato polipeptídico con una o más proteasas. En algunos casos, las proteasas se seleccionan de tripsina, pronasa, quimiotripsina y proteinasaK.

Además se describen métodos para eliminar aminoácidos del extremo amino de una proteína. Los métodos incluyen poner en contacto la proteína con uno o más de los polipéptidos aislados. En algunos casos, el extremo amino de una proteína incluye una etiqueta His. En otros casos, el extremo amino de una proteína incluye una etiqueta Xaa-Pro. Opcionalmente, Xaa es un aminoácido que incluye al menos dos grupos nucleofílicos vecinos, con ejemplos que incluyen serina, treonina o cisteína.

Además se describen polipéptidos aislados que pueden tener actividad proteolítica inversa.

Además se describen métodos para añadir uno o más aminoácidos a un sustrato polipeptídico. El método incluye poner en contacto el sustrato polipeptídico con uno o más de los polipéptidos aislados.

#### DESCRIPCIÓN BREVE DE LOS DIBUJOS

La FIG. 1 es una fotografía de una transferencia Western de una preparación de sobrenadante de *T. rubrum* ensayada con anti-Alp de *A. oryzae* (Panel A, izquierda) y antisuero Mep (Panel C, derecha). El Panel B muestra un gel de SDS-PAGE al 10% teñido con azul de Coomassie. En el carril 1, las proteínas de 0,25 ml de sobrenadante de cultivo de *T. rubrum* se precipitaron con TCA antes se cargarlas en el gel de SDS-PAGE. Se cargaron 0,2 g de ALP de *A. oryzae* y MEP recombinantes purificadas en el carril 2 y carril 3, respectivamente. La masa molecular de los estándares de proteína se muestra en el margen izquierdo.

La FIG. 2 es una fotografía de un gel de SDS-PAGE que ilustra un perfil proteico de ruLAP2 recombinante (1, 2), fuLAP2 (3, 4), ruLAP1 (5, 6) y fuLAP1 (7, 8) producidas en *P. pastoris*. Se cargó 1 g de cada LAP recombinante purificada en un gel de SDS-PAGE al 10%. Los carriles 2, 4, 6 y 8 muestran las proteínas desglicosiladas por tratamiento con N-glicosidasa F. El gel se tiñó con azul brillante de Coomassie R-250.

La FIG. 3 es una fotografía de una transferencia Western de sobrenadante de cultivo de *T. rubrum* y LAP recombinantes usadas como controles ensayadas con antisuero anti-ruLAP2 (carriles 1-4) y anti-ruLAP1 (carriles 5-8). En los carriles 1, 2, 5 y 6 las proteínas de 0,25 ml de sobrenadante de cultivo de *T. rubrum* se precipitaron con TCA antes de cargarlas en el gel de SDS-PAGE. Se cargó 0,1 g de ru-LAP2 (carriles 3, 4) y ruLAP1 (carriles 7, 8) recombinantes purificadas como un control. Se usó N-glicosidasa F para la desglicosilación de las proteínas. La masa molecular de los estándares de proteína se muestra en el margen izquierdo.

La FIG. 4 es un gráfico de la actividad enzimática AMPP (aminopeptidasa P) de *T. rubrum* a varios valores de pH. Parece que AMPP tiene actividad en un rango amplio de valores de pH, de pH 6 a 11.

La FIG. 5 es un gráfico de la actividad enzimática AMPP de *T. rubrum* a varias temperaturas- La enzima presenta actividad a temperaturas que varían de 25 a 60 C con una temperatura óptima de 50C.

La FIG. 6 es un gráfico que muestra la digestión de gliadina 14mer con ruLAP1 durante 4h a 37<sup>o</sup>C con una proporción E/S (p:p) de 1/50.

- 5 La FIG. 7 es un gráfico que muestra la digestión de gliadina 14mer con DPPIV solo y con una mezcla DPPIV/ruLAP2.

La FIG. 8 es un gráfico que muestra la digestión de gliadina 33mer con ruDPPIV durante 4h a 37<sup>o</sup>C con una proporción E/S (p:p) de 1/50.

La FIG. 9 es un gráfico que muestra la digestión de gliadina 33mer con una mezcla DPPIV/ruLAP2.

- 10 Las FIGS. 10A y 10B son espectros de masas de Gly-Ser-proNPY (A) antes y (B) después de la digestión con ruLAP2.

Las FIGS. 11A y 11B son espectros de masas de Ala -proNPY (A) antes y (B) después de la digestión con ruLAP2.

Las FIGS. 12A y 12B son espectros de masas de TG47 (A) antes y (B) después de la digestión con ruLAP2.

Las FIGS. 13A y 13B son espectros de masas de desMet-G-CSF (A) antes y (B) después de la digestión con DPPIV.

- 15 La FIG. 14 es un alineamiento de las secuencias de aminoácidos deducidas de aminopeptidasas de la subfamilia M28E.

La FIG. 15 es un alineamiento de las secuencias de aminoácidos deducidas de aminopeptidasas de la subfamilia M28A.

#### DESCRIPCIÓN DETALLADA DE LA INVENCIÓN

- 20 Tal y como se usa en la presente memoria, el término proteasa es sinónimo de peptidasa, enzima proteolítica e hidrolasa peptídica. Las proteasas incluyen todas las enzimas que catalizan la escisión de los enlaces peptídicos (CO-NH) de las proteínas, digiriendo estas proteínas en péptidos o aminoácidos libres. Las exopeptidasas actúan cerca de los extremos de las cadenas polipeptídicas en el extremo amino (N) o carboxi (C). Las que actúan en un extremo N libre liberan un único resto de aminoácidos y se denominan aminopeptidasas. Una gran variedad de proteasas altamente específicas están implicadas en varios procesos biológicos y fisiológicos diferentes. Así, éstas representan dianas de primera línea para nuevas aplicaciones de fármacos así como para las degradaciones peptídicas y/o proteicas controladas.

- Los dermatofitos son hongos patógenos para los seres humanos y animales, que causan infecciones cutáneas. Vanbreuseghem *et al.*, GUIDE PRACTIQUE DE MYCOLOGIE MEDICALE ET VETERINAIRE. (1978); Kwong-Chong y Bennet, MEDICAL MYCOLOGY (1992); Weitzman y Summerbell, Clin. Microbiol. Rev. 8:240-259 (1995). Los ejemplos de dermatofitos incluyen, pero no están limitados a, *T. ajelloi*, *A. uncinatum*, *K. ajelloi*, *T. asteroides*, *T. mentagrophytes*, *T. concentricum*, *T. cruris*, *E. floccosum*, *T. dankalienese*, *G. dankaliensis*, *T. equinum*, *T. equinum var. autotrophicum*, *T. equinum var. equinum*, *T. erinacei*, *T. fischeri*, *T. flavescens*, *T. floccosum*, *E. floccosum*, *T. glorieae*, *T. gourvilii*, *T. granulare*, *T. granulorum*, *T. gypseum*, *T. inguinale*, *T. interdigitale*, *T. intertriginis*, *T. kanei*, *T. krajdenii*, *T. long fusum*, *T. megninii*, *A. quinckanum*, *A. benhamiae*, *A. vanbreuseghemii*, *T. pedís*, *T. proliferans*, *T. quickaneum*, *T. radiolatum*, *T. mentagrophytes var. erinacei*, *T. mentagrophytes var. interdigitale*, *T. mentagrophytes var. mentagrophytes*, *T. mentagrophytes var. nodulare*, *T. mentagrophytes var. quinckeanum*, *T. niveum*, *T. nodulare*, *T. persicolor*, *M. persicolor*, *T. phaseoliforme*, *T. proliferans*, *T. purpureum*, *T. quinckeanum*, *T. radiolatum*, *T. raubitschekii*, *T. rubrum*, *S. ruber*, *T. schoenleinii*, *T. simii*, *A. simii*, *T. soudanense*, *T. sulphureum*, *T. tonsurans*, *A. insingulare*, *A. lenticularum*, *A. quadrifidum*, *T. tonsurans*, *T. sulphureum*, *T. terrestre*, *T. tonsurans var. sulphureum*, *T. tonsurans var. tonsurans subvar. perforans*, *T. vanbreuseghemii*, *T. verrucosum*, *T. violaceum*, *T. yaoundei*, *E. floccosum*, *M. audouinii*, *M. ferrugineum*, *T. kanei*, *T. megninii*, *T. mentagrophytes*, *T. raubitschekii*, *T. schoenleinii*, *T. soudanese*, *T. violaceum*, *M. canis*, *M. equinum*, *M. nanum*, *M. persicolor*, *T. verrucosum* y *M. gypseum*. Entre las especies patógenas aisladas en hospitales y prácticas privadas en Europa, las observadas más comúnmente son *Trichophyton rubrum*, *T. mentagrophytes* y *Microsporum canis*. Monod *et al.*, Dermatology, 205:201-203 (2002). De hecho, los dermatofitos pueden crecer exclusivamente en el estrato córneo, uñas o pelo, y digieren los componentes de la cubierta celular cornificada. Hasta la fecha, todos los dermatofitos investigados producen actividad proteolítica *in vitro* y muchos investigadores indican el aislamiento y caracterización de una o dos endoproteasas secretadas de una especie individual. Para una revisión, véase Monod *et al.*, Int. J. Med. Microbiol. 292:405-419 (2002). En particular, se ha mostrado que *M. canis* posee dos familias de genes que codifican endoproteasas de la familia S8 (subtilisinas) y M36 (fungalisininas) como se clasifica en la base de datos de enzimas proteolíticas MEROPS (en <http://merops.sanger.ac.uk/>). Brouta *et al.*, Infect. Immun. 70:5676-5683 (2002); Descamps *et al.*, J Invest Dermatol. 70:830-835 (2002). Un miembro de cada familia de genes de *M. canis* aislado codificaba una de las dos endoproteasas previamente caracterizadas de sobrenadantes de cultivos. Mignon *et al.*, Med. Mycol. 36:395-404 (1998); Brouta *et al.*, Med. Mycol. 39:269-275 (2001). Se mostró que ambas enzimas eran queratinolíticas y que se

producían durante la infección en gatos. Mignon *et al.*, Med. Mycol. 36:395-404 (1998); Brouta *et al.*, Med. Mycol. 39:269-275 (2001). Esta actividad proteolítica permite a los dermatofitos crecer exclusivamente en el estrato córneo, uñas o pelo, y usar los componentes digeridos de la cubierta de la célula cornificada, es decir, aminoácidos únicos o péptidos cortos, como nutrientes para el crecimiento *in vivo*.

- 5 En la presente memoria se describen dos nuevas leucina aminopeptidasas (LAP) del dermatofito *T. rubrum*, ruLAP1 y ruLAP2. *T. rubrum* es una especie del género *Trichophyton* que incluye, por ejemplo, *T. ajelloi*, *T. asteroides*, *T. mentagrophytes*, *T. concentricum*, *T. cruris*, *T. dankalienese*, *T. equinum*, *T. equinum var. autotrophicum*, *T. equinum var. equinum*, *T. erinacei*, *T. fischeri*, *T. flavescens*, *T. floccosum*, *T. gloriae*, *T. gourvilii*, *T. granulare*, *T. granulatum*, *T. gypseum*, *T. inguinale*, *T. interdigitale*, *T. intertriginis*, *T. kanei*, *T. krajdienii*, *T. longifusum*, *T. megninii*, *T. pedis*, *T. proliferans*, *T. quickaneum*, *T. radiolatum*, *T. mentagrophytes var. erinacei*, *T. mentagrophytes var. interdigitale*, *T. mentagrophytes var. mentagrophytes*, *T. mentagrophytes var. nodulare*, *T. mentagrophytes var. quinckeanum*, *T. niveum*, *T. nodulare*, *T. persicolor*, *T. phaseoforme*, *T. proliferans*, *T. purpureum*, *T. purpureum*, *T. quinckeanum*, *T. radiolatum*, *T. raubitschekii*, *T. schoenleinii*, *T. simii*, *T. soudanense*, *T. sulphureum*, *T. sulphureum*, *T. tonsurans*, *T. sulphureum*, *T. terrestre*, *T. tonsurans var. sulphureum*, *T. tonsurans var. tonsurans subvar. perforans*, *T. vanbreuseghemii*, *T. verrucosum*, *T. violaceum*, *T. yaoundei*, *T. kanei*, *T. raubitschekii*, *T. soudanense*. Las propiedades de ambas LAP se compararon con las de las enzimas secretadas codificadas por los genes ortólogos del hongo oportunista *Aspergillus fumigatus*, fuLAP1 y fuLAP2 y la LAP microsomal disponible comercialmente de riñón porcino (pkLAP) (MEROPS>familia M1). Todas estas enzimas presentan una actividad leucina aminopeptidasa. También, las aminopeptidasas de *A. fumigatus* fuLAP1 y fuLAP2 presentan aproximadamente 70% de identidad de aminoácidos con los ortólogos de *A. oryzae* indicados en las Patentes de los Estados Unidos Nos. 6.127.161 y 5.994.113, que se incorporan en la presente memoria por referencia. Además, ruLAP2 parece ser única porque (i) ruLAP1 y ruLAP2 presentan aproximadamente 50% de identidad de aminoácidos con los ortólogos de *A. fumigatus* fuLAP1 y fuLAP2 y con los ortólogos de *A. oryzae* indicados en las patentes de los Estados Unidos #6.127.161 y 5.994.113; (ii) una mezcla de ruLAP2 y una endoproteasa semejante a tripsina que se origina del sistema de expresión de *P. pastoris* digiere secuencialmente una cadena polipeptídica de longitud completa tal como caseína desnaturalizada; (iii) una mezcla de ruLAP2 y ruDPPIV (otra exoproteasa de *T. rubrum*) degrada un fragmento de gliadina que se sabe que es resistente a la acción de las proteasas, proporcionando de esta manera evidencia de que ruLAP2 sola o en combinación con ruDPPIV podría usarse para el tratamiento de la enfermedad celíaca o de cualquier enfermedad del tracto digestivo tal como mala absorción; (iv) ruLAP2 en combinación con otras proteasas (mezclas) es útil en la industria alimentaria, tal como degradación de sustratos para amargor, degradación de theves, tratamiento de carne, industria de jabones, degradación de priones, degradación de virus, y degradación de proteínas tóxicas o contaminantes; (v) y, como ruLAP2 y/o otras proteasas secretadas por el hongo son necesarias para que los dermatofitos crezcan en el sustrato cornificado de la uña, los inhibidores de ruLAP2 y/o de otras proteasas secretadas por el hongo serían un nuevo método de tratamiento para las micosis.
- 35 Esta descripción proporciona nuevos ácidos nucleicos y proteínas fúngicos, que tienen actividad leucina aminopeptidasa. LAP juegan un papel en diversas funciones incluyendo, pero sin limitarse a, coagulación de la sangre, muerte celular controlada, diferenciación de los tejidos, invasión tumoral, y en el ciclo de infección de varios microorganismos y virus patógenos lo que hace de estas enzimas una diana valiosa y una herramienta potente para nuevos productos farmacéuticos. Además de tener una función en la fisiología, las aminopeptidasas también tienen aplicaciones comerciales, principalmente en las industrias de detergente y alimentarias. Los microorganismos, tales como los hongos, son una fuente excelente de estas enzimas debido a su amplia diversidad bioquímica y a su susceptibilidad a manipulación genética. Los microorganismos degradan las proteínas y utilizan los productos de degradación como nutrientes para su crecimiento. Así, las nuevas LAP identificadas en la presente memoria son útiles en una multitud de aplicaciones industriales incluyendo pero sin limitarse a hidrólisis de proteínas en la industria alimentaria, degradación de subproductos (por ejemplo, plumas); degradación de priones; degradación de proteínas para proteómica; hidrólisis de polipéptidos para análisis de aminoácidos; limpieza de heridas (por ejemplo, atacando el tejido muerto); limpieza y/o preparación de prótesis; suavizantes textiles; jabones; limpieza o desinfección de tanques sépticos o cualquier contenedor (tal como tanques de retención, botellas, etc.) que contienen proteínas que deben eliminarse o esterilizarse; y limpieza de instrumentos quirúrgicos.
- 50 Esta descripción proporciona nuevas enzimas y mezclas de enzimas, es decir, una mezcla de más de una enzima que digiere estructuras proteicas insolubles, tales como la cubierta de células cornificadas en péptidos cortos y aminoácidos libres. De hecho, además de las endoproteasas de la familia S8 y M36, *T. rubrum* secreta dos LAP cada una con una actividad de sustrato diferente. Cada una de ruLAP1 y ruLAP2 pertenece a la misma familia de LAP (MEROPS>M28). Las propiedades de ambas LAP se compararon con las de las enzimas secretadas codificadas por los genes ortólogos del hongo oportunista *A. fumigatus*, fuLAP1 y fuLAP2 y la LAP microsomal disponible comercialmente de riñón porcino (pkLAP) (MEROPS>familia M1). Todas estas enzimas presentan actividad leucina aminopeptidasa. Además, ruLAP2 tiene una estructura primaria original y es única en que es capaz, en presencia de ruDPPIV, de digerir secuencialmente una cadena polipeptídica, tal como un fragmento de gliadina que se sabe que es resistente a otras proteasas. ruLAP2 parcialmente purificada también es capaz, en presencia de una endoproteasa semejante a tripsina que se origina del sistema de expresión de *P. pastoris*, de digerir secuencialmente una cadena polipeptídica de longitud completa, tal como caseína desnaturalizada.

La descripción se basa, en parte, en el aislamiento de nuevas secuencias de ácido nucleico que codifican nuevos polipéptidos. Los nuevos ácidos nucleicos y sus polipéptidos codificados se refieren individualmente como ruLAP1, ruLAP2, fuLAP1 y fuLAP2. Los ácidos nucleicos, y sus polipéptidos codificados, se designan colectivamente en la presente memoria como "EXOX".

5 Los nuevos ácidos nucleicos EXOX incluyen los ácidos nucleicos cuyas secuencias se proporcionan en las Tablas 1A, 1B, 2A, 2B, 3A, 3B, 4A, 4B, 5A, 5B, 6A, 6B, 7A, 7B, 8A, 8B, 9A, 9B, 10A, 10B, 11A, 11B y 12A, o un fragmento, derivado, análogo u homólogo de éstos. Las nuevas proteínas EXOX incluyen los fragmentos de proteína cuyas secuencias se proporcionan en las Tablas 1C, 2C, 3C, 4C, 5C, 6C, 7C, 8C, 9C, 10C, 11C y 12B. Los ácidos nucleicos y proteínas EXOX individuales se describen más adelante.

10 Usando un método genético inverso, dos aminopeptidasas secretadas por *T. rubrum* se han caracterizado en comparación con ortólogos de *A. fumigatus* y la aminopeptidasa microsomal pLAP de riñón porcino. Las cuatro enzimas fúngicas identificadas en la presente memoria (ruLAP1, fuLAP1, ruLAP2, y fuLAP2) así como pLAP comparten una preferencia común para Leu-AMC como sustrato y funcionan como leucina aminopeptidasas. Además, la aminopeptidasa pLAP, que actúa también con una eficacia extremadamente alta frente a Ala-AMC, también se denomina alanina aminopeptidasa (MEROPS>M1.001).

15 Los ácidos nucleicos EXOX, que codifican las proteínas EXOX, incluyen los ácidos nucleicos cuyas secuencias se proporcionan en la presente memoria o fragmentos de éstas. La descripción también incluye ácidos nucleicos mutantes o variantes cualquiera de cuyas bases puede cambiarse de la base correspondiente mostrada en la presente memoria, mientras que todavía codifica una proteína que mantiene sus actividades y funciones fisiológicas semejantes a EXOX o un fragmento de dicho ácido nucleico. También se describen ácidos nucleicos cuyas secuencias son complementarias a las descritas en la presente memoria, incluyendo fragmentos de ácidos nucleicos que son complementarios a cualquiera de los ácidos nucleicos recién descritos. También se describen ácidos nucleicos o fragmentos de ácidos nucleicos, o complementos de éstos, cuyas estructuras incluyen modificaciones químicas. Dichas modificaciones incluyen, como ejemplo no limitativo, bases modificadas y ácidos nucleicos cuyas estructuras centrales de azúcar fosfato están modificados o derivatizados. Estas modificaciones se realizan al menos en parte para incrementar la estabilidad química del ácido nucleico modificado, de manera que puedan usarse, por ejemplo, como ácidos nucleicos de unión antisentido en aplicaciones terapéuticas en un sujeto.

20 Las proteínas EXOX incluyen las proteínas EXO cuyas secuencias se proporcionan en la presente memoria. La descripción también incluye proteínas mutantes o variantes cualquiera de cuyos restos puede cambiarse del resto correspondiente mostrado en la presente memoria, mientras que todavía codifica una proteína que mantiene sus actividades y funciones fisiológicas semejantes a EXO o un fragmento funcional de ésta. La descripción engloba además anticuerpos y fragmentos de anticuerpo, tal como  $F_{ab}$  o  $(F_{ab})_2$ , que se unen inmunoespecíficamente a cualquiera de las proteínas.

25 Los ácidos nucleicos y las proteínas EXOX son útiles en aplicaciones terapéuticas potenciales tales como el tratamiento de infecciones fúngicas. Los ácidos nucleicos, proteínas EXOX e inhibidores también tienen otras funciones que incluyen pero no están limitadas a: (i) reactivo de biotecnología para mejorar la producción de proteínas, por ejemplo, eliminación de etiquetas, producción de aminoácidos raros; (ii) desarrollo de fármacos para determinadas indicaciones patológicas, por ejemplo, enfermedad celíaca (intolerancia al gluten); (iii) desarrollo de fármacos para condiciones dermatológicas, por ejemplo, agentes anti-micosis, tratamiento de verrugas, cicatrización de heridas; (iv) cosmetología, por ejemplo, con herramientas de peeling, depilación, dermoabrasión y dermaplaning; (v) industria alimentaria, por ejemplo, producción de suplementos nutricionales, edulcorantes, generación de alimentos hipoalergénicos por predigestión; (vi) agente desinfectante, por ejemplo, descontaminación de contaminantes basados en proteínas tales como priones o virus (por digestión de la proteína de la cubierta), limpieza de instrumentos quirúrgicos o preparación de ítems para cirugía tales como prótesis o dispositivos médicos; (vii) desinfección o reciclado de determinados desechos, por ejemplo, plumas, huesos, pelo y pelaje; (viii) agente de limpieza, por ejemplo, champú o detergente líquido.

30 Los inhibidores de los EXO, específicamente de ruLAP2, también pueden usarse como agentes anti-micóticos fúngicos para tratar las micosis. Las LAP en sí mismas también pueden usarse para tratar enfermedades del tracto digestivo, tal como mala absorción o enfermedad celíaca, que está causada por el gluten del trigo. Gluten es el término característico para la mezcla proteica de glutelinas y gliadinas (prolaminas) encontrada en cereales. Debido a sus propiedades físico-químicas inherentes tales como actuar como un agente de unión y de extensión, el gluten se usa comúnmente como un aditivo en los alimentos. La detección del gluten es importante en el control de calidad y la selección de alimentos para individuos con enfermedades relacionadas con o causadas por la intolerancia al gluten, incluyendo, enteropatía por intolerancia al gluten, enfermedad celíaca, esprúe y reacciones alérgicas relacionadas, en las que es necesaria una dieta sin el gluten contenido en trigo, centeno, cebada y en algunos casos avena.

#### Exoproteasa Ácidos Nucleicos y Polipéptidos

La actividad aminopeptidasa de *T. rubrum* demostrada aquí y estudios previos en subtilisinas y metaloproteasas secretadas por *M. canis* muestran que los dermatofitos secretan una batería de proteasas similar a las de las

especies de *Aspergillus* en un medio que contiene proteína como única fuente de carbono y de nitrógeno. Además, dos genes, ruDPPIV y ruDPPV:EMBL AF082514 para ruDPPV, que codifica dipeptidil-aminopeptidasas altamente similares a DPPIV y DPPV de *A. fumigatus* y *A. oryzae* (Beauvais *et al.*, J. Biol. Chem. 272:6238-6244 (1997); Beauvais *et al.*, Infec. Immun. 65:3042-3047 (1997); Doumas *et al.*, Appl. Environ. Microbiol. 64:4809-4815 (1998); Doumas *et al.*, J. Food Mycol. 2:271-279 (1999)) se aislaron de bibliotecas genómicas y de ADNc de *T. rubrum*. Las estructuras intrón-exón de los genes de *T. rubrum* que codifican estas proteasas son similares a los genes homólogos aislados de *A. fumigatus* y *A. oryzae*. Estos resultados no son sorprendentes ya que los teleomorfos de las especies de *Aspergillus* y los teleomorfos de las especies de dermatofito están muy relacionados, ya que pertenecen al mismo grupo taxonómico de Ascomycetes que producen *asci* de prototunicate en *cleistothecia* (clase Eurotiomycetes). A diferencia de los genes que codifican las subtilisinas y las fungalisinas, *ruLAP1* y *ruLAP2* no son miembros de grandes familias de genes en el genoma de *T. rubrum*.

RuLAP1 presenta aproximadamente 50% de identidad de aminoácidos con fuLAP1 y/o LAP1 (Véanse las Tablas 19A y 20. Estas tres enzimas pertenecen estructuralmente a la misma subfamilia M28E como las leucil aminopeptidasas de *Aeromonas* y *Vibrio* (MEROPS>M28.002). Además, ruLAP2 presenta aproximadamente 50% de identidad de aminoácidos con fuLAP2 y/o LAP2 (Véanse las Tablas 19B y 21). Estas tres enzimas pertenecen estructuralmente a la misma subfamilia M28A como la proteasa vacuolar Y de *S. cerevisiae* (MEROPS>M28.001) y la aminopeptidasa secretada de *Streptomyces griseus* (MEROPS>M28.00X). Además, los miembros de las subfamilias M28A y M28E comparten similitudes bajas. Sin embargo, los aminoácidos de los dos sitios de unión a Zn<sup>++</sup> en estas aminopeptidasas están conservados y se identificaron en las LAP fúngicas caracterizadas en la presente memoria (Véanse las Tablas 20 y 21). En las aminopeptidasas secretadas de *S. griseus* y *Aeromonas proteolytica*, los dos restos de aminoácidos His y Asp se unen a un primer ión Zn<sup>++</sup> y dos restos adicionales His y Glu se unen a un segundo ión Zn<sup>++</sup>, mientras que un segundo resto Asp hace de puente entre los dos iones Zn<sup>++</sup>. Greenblatt *et al.*, J. Mol. Biol. 265:620-636 (1997); Hasselgren *et al.*, J. Biol. Inorg. Chem. 6:120-127 (2001). La sustitución de Zn<sup>++</sup> por diferentes iones divalentes en la aminopeptidasa secretada de *S. griseus* se ve influida por Ca<sup>++</sup> y tiene efectos variables. Ben-Meir *et al.*, Eur. J. Biochem 212:107-112 (1993); Lin *et al.*, J. Biol. Inorg. Chem. 2:744-749 (1997); Hasselgren *et al.*, J. Biol. Inorg. Chem. 6:120-127 (2001). Se encontró que las aminopeptidasas eran sensibles a diferentes iones. Como la aminopeptidasa de *S. griseus*, ruLAP2 y fuLAP2 se activan fuertemente por Co<sup>++</sup>.

RuLAP2 y fuLAP2 poseen actividades proteolíticas sustancialmente diferentes a pesar de un alto porcentaje de identidad de secuencia. En particular, ruLAP2 es capaz de hidrolizar eficazmente Asp- y Glu-7-amino-4-metilcumarina (AMC) y ruLAP2 es la única LAP identificada hasta ahora que es capaz, primero en presencia de ruDPPIV, de digerir un péptido de gliadina que se sabe que es resistente a la digestión por proteasas gástricas y pancreáticas, o segundo, en la forma de un extracto parcialmente purificado que contiene una endoproteasa semejante a tripsina que se origina del sistema de expresión de *P. pastoris*, de digerir una cadena polipeptídica de longitud completa tal como caseína desnaturalizada. La capacidad de una LAP de degradar un polipéptido largo no es predecible únicamente tomando como base su capacidad de escindir restos aminoacil-AMC. Las propiedades particulares de las enzimas de los dermatofitos se han observado con endoproteasas secretadas por *M. canis*. La subtilisina de 31,5 kDa de *M. canis* y las metaloproteasa de 43,5 kDa de *M. canis* son capaces de digerir queratina azul a diferencia de las proteasas secretadas homólogas de *A. fumigatus* y *A. oryzae*. Como los dermatofitos evolucionaron de su hábitat natural en el suelo, han desarrollado una estrategia de infección usando proteasas particulares para degradar los tejidos queratinizados. Las propiedades únicas de ruLAP2 podrían reflejar organismos altamente especializados que parasitan el estrato córneo y las uñas.

Además de las LAP descritas en la presente memoria, también se ha aislado una serie de nuevas proteasas del hongo patógeno *T. rubrum* y se describen más adelante. Como las LAP, estas proteasas están todas caracterizadas como exoproteasas. Incluyen: dos carboxipeptidasas, una prolilaminopeptidasa, una aminopeptidasa P, una prolidasa, y una dipeptidilpeptidasa IV. También se han caracterizado dos nuevas proteasas adicionales: una leucina aminopeptidasa (caLAP1) de *Microsporium canis* y meLAP1, una leucina aminopeptidasa de *Trichophyton mentagrophytes*.

### ruLAP2

ruLAP2 es una leucina aminopeptidasa de *T. rubrum*. Un ácido nucleico de ruLAP2 de 1.757 nucleótidos (SEQ ID NO: 1) se muestra en la Tabla 1A.



**Tabla 1A. Secuencia de nucleótidos genómica de ruLAP2 (SEQ ID NO:1).**

ATGAAGTCGCAACTGTTGAGCCTGGCTGTGGCCGTCACAACCATCTCCCAGGGCGTTGTTGGTCAAGAG  
 CCCTTCGGATGGCCTTCAAGCCTATGGTCACTCAGGTGAGTTGCTCTCAACAGATCGATCGATCGATC  
 TACCTTTGTCCCTGTCACATCAAACCTCCAGCAGAGCCAAAGAAACAGACACAATGTTCTGGGGAATTC  
 TTATGGGCTAATGTAAATGTATAGGATGACCTGCAAAACAAGATAAAGCTCAAGGATATCATGGCAGGC  
 GTCGAGAAGCTGCAAAGCTTTTCTGATGCTCATCCTGAAAAGAACCGAGTGTGTTGGTGGTAATGGCCAC  
 AAGGACACTGTAGAGTGGATCTACAATGAGATCAAGGCCACTGGCTACTACGATGTGAAGAAGCAGGAG  
 CAAGTACACCTGTGGTCTCATGCCGAGGCTGCTCTCAATGCCAATGGCAAGGACCTCAAGGCCAGCGCC  
 ATGTCCTACAGCCCTCCTGCCAGCAAGATCATGGCTGAGCTTGTGTTGCCAAGAACAATGGCTGCAAT  
 GCTGTATGTGCCATACACTTTCTATACGTCACATTCTCTCTAGAATGAAGAGCACGGGAGAACCTA  
 TATGTATACAGACTGATTACCCAGCGAACACTCAGGGCAAGATCGTCCCTCGTTGAGCGTGGTGTCTGCA  
 GCTTCGGCGAGAAGTCTGCTCAGGCTGGTGATGCAAAGGCTGCTGGTGCCATTGTCTACAACAACGTCC  
 CCGGATCCCTTGCTGGCACTCTTGGTGGCCTTGACAAGCGCCATGTCCCAACCGCTGGTCTTTCCAGG  
 AGGATGGAAAGAACCTTGCTACCCTCGTTGCTTCTGGTAAGATTGATGTCACCATGAACGTTATCAGTC  
 TGTTTGAGAACC GAACCACGTAAGTAGCTCAACGGCTGATCCAGCATCAATTGTCTCGAGTATATACTA  
 AATCGATACCTCATAGCTGGAACGTATTGCTGAGACCAAGGGAGGAGACCACAACAACGTTATCATGC  
 TCGGTGCTCACTCCGACTCCGTCGATGCCGGCCCTGGTATTAACGACAACGGCTCGGGCTCCATTGGTA  
 TCATGACCGTTGCCAAAGCCCTCACCAACTTCAAGCTCAACAACCGCGTCCGCTTTGCTGGTGGACCG  
 CTGAGGAATTCGGTCTCCTTGGAAGCACCTTCTACGTCAACAGCCTCGATGACCGTGAGCTGCACAAGG  
 TCAAGTTGTACCTCAACTTCGACATGATCGGGCTCTCCCAACTTCGCCAACCAGATCTACGACGGTGACG  
 GTTCGGCCTACAACATGACCGGCCCCGCTGGCTCTGCTGAAATCGAGTACCTGTTGAGAAAGTTCTTTG  
 ACGACCAGGGTATCCCACACCAGCCCACTGCCTTCACTGGCCGATCCGACTACTCTGCTTTCATCAAGC  
 GCAACGTGCCCCGCTGGCGGCCTTCTCACTGGAGCCGAGGTTGTCAAGACCCCGAGCAAGTCAAGTTGT  
 TCGGTGGTGAAGGCTGGCGTTGCCTATGACAAGAACTACCATCGCAAGGGCGACACCGTTGCCAACATCA  
 ACAAGGGAGCTATCTTCTTAACACTCGAGCCATCGCCTACGCTATCGCCGAGTATGCCCGATCCCTCA  
 AGGGATTCCCAACCCGCCAAAGACCGGCAAGCGTGACGTCAACCCCAAGTATTCTAAGATGCCTGGTG  
 GTGGCTGCGGACACCACACTGTCTTCATGTAA

Un marco de lectura abierto (“ORF”) de ruLAP2 descrito de 1.488 nucleótidos empieza con un codón de inicio ATG en la posición 1 (subrayado en la Tabla 1B).

Tabla 1B. Secuencia de nucleótidos de ruLAP2 (SEQ ID NO: 2).

```

ATGAAGTCGCAACTGTTGAGCCTGGCTGTGGCCGTCACAACCATCTCCCAGGGCGTTGTTGGTCAAGAG
CCCTTCGGATGGCCTTTCAAGCCTATGGTCACTCAGGATGACCTGCAAAAACAAGATAAAGCTCAAGGAT
ATCATGGCAGGCGTCGAGAAGCTGCAAAGCTTTTCTGATGCTCATCCTGAAAAGAACCGAGTGTTTGGT
GGTAATGGCCACAAGGACACTGTAGAGTGGATCTACAATGAGATCAAGGCCACTGGCTACTACGATGTG
AAGAAGCAGGAGCAAGTACACCTGTGGTCTCATGCCGAGGCTGCTCTCAATGCCAATGGCAAGGACCTC
AAGGCCAGCGCCATGTCTACAGCCCTCCTGCCAGCAAGATCATGGCTGAGCTTGTGTTGTTGCCAAGAAC
AATGGCTGCAATGCTACTGATTACCCAGCGAACACTCAGGGCAAGATCGTCTCGTTGAGCGTGGTGTCTC
TGCAGCTTCGGCGAGAAGTCTGCTCAGGCTGGTGTGCAAAGGCTGCTGGTGCCATTGTCTACAACAAC
GTCCCGGATCCCTTGCTGGCACTCTTGGTGGCCTTGACAAGCGCCATGTCCCAACCGCTGGTCTTTCC
CAGGAGGATGGAAAGAACCCTTGCTACCCTCGTTGCTTCTGGTAAGATTGATGTCACCATGAACGTTATC
AGTCTGTTTGAGAACCGAACCACCTGGAACGTCAATGCTGAGACCAAGGGAGGAGACCACAACAACGTT
ATCATGCTCGGTGCTCACTCCGACTCCGTCGATGCCGGCCCTGGTATTAACGACAACGGCTCGGGCTCC
ATTGGTATCATGACCGTTGCCAAAGCCCTCACCAACTTCAAGCTCAACAACGCCGTCGCTTTGCCTGG
TGGACCGCTGAGGAATTCGGTCTCCTTGGAAGCACCTTCTACGTCAACAGCCTCGATGACCGTGAGCTG
CACAAGGTCAAGTTGTACCTCAACTTCGACATGATCGGCTCTCCCAACTTCGCCAACCAGATCTACGAC
GGTGACGGTTCGGCCTACAACATGACCGGCCCGCTGGCTCTGCTGAAATCGAGTACCTGTTTCGAGAAG
TTCTTTGACGACCAGGGTATCCCACACCAGCCACTGCCTTCACTGGCCGATCCGACTACTCTGCTTTC
ATCAAGCGCAACGTGCCCGCTGGCGGCCTCTTCACTGGAGCCGAGGTTGTCAAGACCCCCGAGCAAGTC
AAGTTGTTTCGGTGGTGGGCTGGCGTTGCCTATGACAAGAACTACCATCGCAAGGGCGACACCGTTGCC
AACATCAACAAGGGAGCTATCTTCTTAACACTCGAGCCATCGCCTACGCTATCGCCGAGTATGCCCGA
TCCCTCAAGGGATTCCCAACCCGCCAAAGACCGGCAAGCGTGACGTCAACCCCAAGTATTCTAAGATG
CCTGGTGGTGGCTGCGGACACCACACTGTCTTCATGTAA
    
```

Un ácido nucleico de ruLAP2 descrito (SEQ ID NO: 2) codifica una proteína que tiene 495 restos de aminoácidos (SEQ ID NO: 3), que se presenta en la Tabla 1C usando el código de aminoácidos de una letra.

Tabla 1C. Secuencia de la proteína codificada ruLAP2 (SEQ ID NO:3).

```

MKSQLLSLAVAVTTISQGVVQEPFGWPFKPMVTQDDLQNKIKLKDIMAGVEKLQSFSDAHPEKNRVFG
GNGHKDTVEWIYNEIKATGYDVKKQEQVHLWSHAEEALNANGKDLKASAMSYSPASKIMAE LVVAKN
NGCNATDYPANTQGKIVLVERGVCSFGEKSAQAGDAKAAGAI VYNNVPGSLAGTLGGLDKRHVPTAGLS
QEDGKNLATLVASGKIDVTMNVISLFENRTTNVIAETKGGDHNNVIMLGAHSDSVDAGPGINDNGSGS
IGIMTVAKALTNFKLNNAVRFAWWTAEFGLLGSTFYVNSLDDRELHKVKLYLNFDMIGSPNFANQIYD
GDGSAYNMTGPAGSAEIEYLFKFFDDQGI PHQPTAFTGRSDYSAFIKRNVPAGGLFTGAEVVKTPQV
KLFGEAGVAYDKNYHRKGD TVANINKGAI FLNTRAIAYAIAEYARSLKGFPTRPKTGKRVDVNPQYSKM
PGGGCGHHTVFM
    
```

5 La ruLAP2 descrita tiene homología con las secuencias de aminoácidos mostradas en los datos BLAST listados en las Tablas 1D, 1E y 1F.

Se usaron las opciones de programa siguientes:

**tblastn** – compara la “Secuencia 1” de la proteína frente a la “Secuencia 2” de nucleótidos que se ha traducido en los seis marcos de lectura

10 **blastx** – compara la “Secuencia 1” de nucleótidos frente a la “Secuencia 2” de proteína

**blastp** – para comparaciones proteína – proteína

15 En todos los alineamientos BLAST en la presente memoria, el “valor E” o valor “Esperado” es una indicación numérica de la probabilidad de que las secuencias alineadas podrían haber conseguido su similitud frente a la secuencia de búsqueda BLAST sólo por casualidad, en la base de datos en la que se buscó. El valor Esperado (E) es un parámetro que describe el número de logros que uno puede “esperar” ver sólo por casualidad cuando se busca en una base de datos de un tamaño particular. Disminuye exponencialmente con la Puntuación (S) que se asigna a una equivalencia entre dos secuencias. Esencialmente, el valor E describe el ruido de fondo aleatorio que existe para equivalencias entre secuencias.

Tabla 1D. Resultados TBLASTN para ruLAP2

Índice/Identificador Génico	Proteína/Organismo	Longitud (aa)	Identidad (%)	Positivos (%)	Esperado
gi469363	Gen de aminopeptidasa Y de <i>Saccharomyces cerevisiae</i>	32421	170/477 (35%)	239/437 (55%)	8e-65
gi15839805	CDC15551 de <i>Mycobacterium tuberculosis</i> , sección 33 de 280 del genoma completo	18857	152/424 (35%)	225/424 (53%)	5e-57
gi9949032	PAO1 de <i>Pseudomonas aeruginosa</i> , sección de 281 de 529 del genoma completo	12547	129/317 (40%)	180/317 (56%)	1e-56

Tabla 1E. Resultados BLASTX para ruLAP2

Índice/Identificador Génico	Proteína/Organismo	Longitud (aa)	Identidad (%)	Positivos (%)	Esperado
gi28918599	Proteína hipotética/ <i>Neurospora crassa</i>	508	219/467 (46%)	287/467 (61%)	e-112
gi584764	APE3 LEVADURA; Precursor de aminopeptidasa/ <i>Saccharomyces cerevisiae</i>	537	170/477 (35%)	239/437 (55%)	1e-65
gi23017467	Proteína hipotética/ <i>Thermobifida fusca</i>	514	151/460 (32%)	237/460 (51%)	5e-61
gi15839805	Hidrolasa/ <i>Mycobacterium tuberculosis</i> CDC15551	493	152/424 (35%)	225/424 (53%)	6e-58

Tabla 1F. Resultados BLASTP para ruLAP2

Índice/Identificador Génico	Proteína/Organismo	Longitud (aa)	Identidad (%)	Positivos (%)	Esperado
Gi28918599	Proteína hipotética/ <i>Neurospora crassa</i>	508	219/467 (46%)	287/467 (61%)	e-105
Gi584764	APE3 LEVADURA; Precursor de aminopeptidasa/ <i>Saccharomyces cerevisiae</i>	537	169/477 (35%)	237/477 (49%)	2e-64
Gi15839805	Hidrolasa/ <i>Mycobacterium tuberculosis</i> CDC15551	493	152/424 (35%)	225/424 (53%)	5e-57
Gi23017467	Proteína hipotética/ <i>Thermobifida fusca</i>	514	151/460 (32%)	237/460 (51%)	1e-56

**ruLAP1**

- 5 ruLAP1 es una leucina aminopeptidasa de *T. rubrum*. Un ácido nucleico de ruLAP1 de 1.256 nucleótidos se muestra en la Tabla 2A (SEQ ID NO: 4).

**Tabla 2A. Secuencia de nucleótidos genómica de ruLAP1 (SEQ ID NO:4).**

ATGAAGCTCCTCTCTGTTCTTGGCGCTGAGCGCTACCGCTACCTCCGTCCTCGGAGCTAGCATTCTGTT  
 GATGCCCGGGCCGAGAAGTTCCTCATCGAAGTTGCCCTGGTGAGACTCGCTGGGTTACCGAGGAGGAG  
 AAGTGGGAGCTTAAGCGGGTATGTACCACTATCCTACGCAAAGTTGTATTTTCACTAGATAATATTGG  
 TTATTAACACCCATTCTAGAAGGGTCAAGACTTCTTTGACATCACTGACGAGGAGGTTGGATTCACTGC  
 TGCTGTTGCACAGCCAGCCATTGCCTACCCAACCTCCATCCGCCATGCTAATGCTGTTAACGCCATGAT  
 TGCTACCCTCTCCAAGGAGAACATGCAGCGCGATCTGACCAAGCTCAGCTCGTTCCAAACCGCTTACTA  
 TAAGGTTGACTTTGGCAAGCAGTCTGCCACCTGGCTCCAGGAGCAAGTCCAGGCTGCCATCAATACCGC  
 TGGTGCCAATCGCTACGGAGCCAAGGTCGCCAGCTTCCGACACAACCTTCGCTCAGCACAGCATCATTGC  
 CACTATTCGGCCGCTCCCTGAAGTCGTTGTCGTCGGTGCTCACCAAGACAGCATCAACCAACGCAG  
 CCCATGACCGGCCGCGCTCCAGGTGCCGATGACAACGGCAGTGGCTCCGTCACCATCCTTGAGGCCCT  
 CCGTGGTGTCTCCGGGACCAGACCATCCTCCAGGGCAAGGCTGCCAACACCATTGAGTTCCTACTGGTA  
 CGCCGGTGAGGAAGCTGGTCTTCTGGGCTCCAGGCCATCTTCGCCAACTACAAACAGACCGGCAAGAA  
 GGTCAAGGGCATGCTCAACCAGGACATGACCGGTTACATCAAGGGAATGGTCGACAAGGGTCTCAAGGT  
 GTCCTTCGGTATCATCACCGACAACGTCAACGCTAACTTGACCAAGTTCGTCCGCATGGTCATCACCAA  
 GGTAAAGCTTCAACTCTTGATAAATATATTTTTTCATCGATGAAATGATGTCCTAATAATGCTTAAGTACT  
 GCTCAATCCCAACCATCGACACCCGCTGCGGCTATGCTTGCTCTGACCACGCCTCTGCCAACCGCAATG  
 GCTACCCATCTGCCATGGTTGCCGAGTCTCCATCGATCTCCTCGACCCTCACCTCCACACTGACTCTG  
 ACAACATTAGCTACCTCGACTTCGACCACATGATCGAGCAGCTAAGCTCATTGTCCGGCTTCGTCCTG  
 AGCTCGCTAAGTAA

Un marco de lectura abierto ("ORF") de ruLAP1 descrito de 1.122 nucleótidos empieza con un codón ATG (subrayado en la Tabla 2B) en la posición 1.

**Tabla 2B. Secuencia de nucleótidos de ruLAP1 (SEQ ID NO:5).**

ATGAAGCTCCTCTCTGTTCTTGGCGCTGAGCGCTACCGCTACCTCCGTCCTCGGAGCTAGCATTCTGTT  
 GATGCCCGGGCCGAGAAGTTCCTCATCGAAGTTGCCCTGGTGAGACTCGCTGGGTTACCGAGGAGGAG  
 AAGTGGGAGCTTAAGCGGAAGGGTCAAGACTTCTTTGACATCACTGACGAGGAGGTTGGATTCACTGCT  
 GCTGTTGCACAGCCAGCCATTGCCTACCCAACCTCCATCCGCCATGCTAATGCTGTTAACGCCATGATT  
 GCTACCCTCTCCAAGGAGAACATGCAGCGCGATCTGACCAAGCTCAGCTCGTTCCAAACCGCTTACTAT  
 AAGGTTGACTTTGGCAAGCAGTCTGCCACCTGGCTCCAGGAGCAAGTCCAGGCTGCCATCAATACCGCT  
 GGTGCCAATCGCTACGGAGCCAAGGTCGCCAGCTTCCGACACAACCTTCGCTCAGCACAGCATCATTGCC  
 ACTATTCGGCCGCTCCCTGAAGTCGTTGTCGTCGGTGCTCACCAAGACAGCATCAACCAACGCAGC  
 CCCATGACCGGCCGCGCTCCAGGTGCCGATGACAACGGCAGTGGCTCCGTCACCATCCTTGAGGCCCTC  
 CGTGGTGTCTCCGGGACCAGACCATCCTCCAGGGCAAGGCTGCCAACACCATTGAGTTCCTACTGGTAC  
 GCCGGTGAGGAAGCTGGTCTTCTGGGCTCCAGGCCATCTTCGCCAACTACAAACAGACCGGCAAGAAG  
 GTCAAGGGCATGCTCAACCAGGACATGACCGGTTACATCAAGGGAATGGTCGACAAGGGTCTCAAGGTG  
 TCCTTCGGTATCATCACCGACAACGTCAACGCTAACTTGACCAAGTTCGTCCGCATGGTCATCACCAAG  
 TACTGCTCAATCCCAACCATCGACACCCGCTGCGGCTATGCTTGCTCTGACCACGCCTCTGCCAACCGC  
 AATGGCTACCCATCTGCCATGGTTGCCGAGTCTCCATCGATCTCCTCGACCCTCACCTCCACACTGAC  
 TCTGACAACATTAGCTACCTCGACTTCGACCACATGATCGAGCAGCTAAGCTCATTGTCCGGCTTCGTC  
 ACTGAGCTCGCTAAGTAA

5 Un ácido nucleico de ruLAP1 descrito (SEQ ID NO: 5) codifica una proteína que tiene 377 restos de aminoácidos (SEQ ID NO: 6), que se presenta en la Tabla 2C usando el código de aminoácidos de una letra.

**Tabla 2C. Secuencia de la proteína codificada ruLAP1 (SEQ ID NO:6).**

```

MKLLSVLALSATATSVLGASIPVDARA EKFLIELAPGETRWVTEEEKWELKRKGQDFDI TDEEVGFTA
AVAQPAIAYPTSIRHANAVNAMIATLSKENMQRDLTKLSSFQTAYYKVDVFGKQSATWLQEQQVQAAINTA
GANRYGAKVASFRHNFAQHSIIATIPGRSPEVVVVGAHQDSINQSRPMTGRAPGADDNGSGSVTILEAL
RGVLRDQTILOGKAANTIEFHUYAGEEAGLLGSQAI FANYKQTGKKVKGMLNQDMTG YIKGMVDKGLKV
SFGIITDNVNANLTKFVRMVITKYCS IPTIDTRCGYACSDHASANRNGYPSAMVAESP IDLLDPHLHTD
SDNISYLD FDMIEHAKLIVGFVTE LAK

```

La ruLAP1 descrita tiene homología con las secuencias de aminoácidos mostradas en los datos BLAST listados en las Tablas 2D, 2E y 2F. Estos datos se analizaron por el programa por parejas blast.

**Tabla 2D: Resultados TBLASTN para ruLAP1**

Índice/Identificador Génico	Proteína/Organismo	Longitud (aa)	Identidad (%)	Positivos (%)	Esperado
>gi1762234	Poliquétido sintasa PKSL2/ <i>Aspergillus parasiticus</i>	9894	131/247 (53%) 40/76 (52%) 20/24 (83%)	171/247 (69%) 57/76 (75%) 22/24 (91%)	1e-95
>gi23393798	Leucina aminopeptidasa (Lap1)/ <i>Aspergillus sojae</i>	2547	77/159 (48%) 63/148 (42%) 14/30 (46%)	97/159 (61%) 89/148 (60%) 23/30 (76%)	4e-64
>gi927685	Cromosoma IV lambda3641 y cósmido 9831 y 9410 de <i>Saccharomyces cerevisiae</i>	78500	137/350 (39%)	201/350 (57%)	3e-62
>gi7413486	ARNm parcial para aminopeptidasa de <i>Agaricus</i>	1089	130/346 (37%)	189/346 (54%)	2e-55

5

**Tabla 2E: Resultados BLASTX para ruLAP1**

Índice/Identificador Génico	Proteína/Organismo	Longitud (aa)	Identidad (%)	Positivos (%)	Esperado
>gi23393799	Leucina aminopeptidasa/ <i>Aspergillus sojae</i>	377	126/248 (50%) 37/78 (47%) 13/24 (54%)	162/248 (65%) 55/78 (70%) 20/24 (83%)	5e-87
>gi28918132	Proteína hipotética/ <i>Neurospora crassa</i>	402	115/247 (46%) 43/77 (55%) 18/24 (75%)	153/247 (61%) 58/77 (75%) 23/24 (95%)	8e-86
>gi6320623	ORF hipotético; ydr415cp/ <i>Saccharomyces cerevisiae</i>	374	96/254 (37%) 36/77 (46%)	143/254 (56%) 49/77 (63%)	7e-55
>gi28916832	Proteína hipotética/ <i>Neurospora crassa</i>	409	96/226 (42%) 31/66 (46%)	135/226 (59%) 41/66 (62%)	4e-54

Índice/Identificador Génico	Proteína/Organismo	Longitud (aa)	Identidad (%)	Positivos (%)	Esperado
>gi23393799	Leucina aminopeptidasa/ <i>Aspergillus sojae</i>	377	175/348 (50%)	234/348 (67%)	4e-99
>gi28918132	Proteína hipotética/ <i>Neurospora crassa</i>	402	175/345 (50%)	230/345 (66%)	2e-97
>gi6320623	ORF hipotético; ydr415cp/ <i>Saccharomyces cerevisiae</i>	374	140/351 (39%)	201/351 (57%)	7e-65
>gi28916832	Proteína hipotética/ <i>Neurospora crassa</i>	409	129/296 (43%)	178/296 (60%)	3e-58

**fuLAP2**

fuLAP2 es una leucina aminopeptidasa de *A. fumigatus*. Un ácido nucleico de fuLAP2 de 1.557 nucleótidos se muestra en la Tabla 3A (SEQ ID NO: 7).

<p>ATGAAGCTGCTCTACCTCACATCGTTTGCCTCTCTGGCCGTGGCCAATGGCCCAGGATGGGACTGGAAG                  CCCCAGTTTCATCCGGTTAGTGTTCCTCTCGCCGGGTTTGTCTGCTGTATGCTAACAGCATCCTGTCTA                  TTACAGAAAGTCTCGCCCAAATGATCCATTTGTGGGATCTTCTGCAGGGCGCTCAACAGCTGGAAGAC                  TTCGCCTATGCCTACCCCGAGCGCAACCGCGTCTTTGGTGGACGGGCCACGAGGACACCGTCAACTAC                  CTCTACCGTGAGTTGAAGAAAACCGGCTACTACGACGTTTACAAGCAGCCCCAGGTTACCAGTGGACC                  CGAGCCGACCAGGCTCTCACCGTCGACGGCCAGTCTATGACGCCACAACCATGACTTACAGCCCCAGC                  GTAACGCCACGGCGCCGCTGGCAGTGGTGAACAACCTGGGCTGCGTCGAGGCTGACTATCCCGCCGAT                  CTGACGGGCAAGATTGCTCTGATCTCGCGGGGCGAGTGCACCTTTGCGACCAAATCCGTCTTGAGCGCC                  AAGGCCGGGGCGGCGGGCACTCGTGTAACAATATCGAGGGTTCGATGGCGGGAACTCTGGGCGGC                  GCGACCAGCGAGCTGGGTGCCTACGCTCCCATCGCCGGCATCAGCCTCGCGGACGGACAGGCGCTGATC                  CAGATGATCCAGGCGGGCACGGTGCAGCCAACCTGTGGATCGACAGCCAGGTCGAGAACCCTACCACC                  TACAACGTGATCGCGCAGACCAAGGGCGGGCACCCCAACAACGTCTGCGCGTGGGTGGCCACACGGAC                  TCGGTGAGGGCCGGGCCCGGCATCAACGACGACGGCTCCGGCATCATCAGCAACCTCGTCGTGCGCAAG                  GCGCTGACCCGCTTCTCGGTCAAGAACCGGGTGGCTTCTGCTTCTGGACGGCGGAGGAGTTGCGCCTG                  CTGGGCAGCAACTACTACGTCAACAGCCTCAATGCCACCGAGCAGGCCAAGATCCGCCTGTATCTCAAC                  TTCGACATGATCGCCTCCCCAACTACGCCCTGATGATCTATGACGGCGACGGCTCGGCCTTCAACCTG                  ACGGGGCCGGCCGGCTCGGCGCAGATCGAGCGGCTCTTCGAGGACTACTACAGTCGATCCGCAAGCCG                  TTCGTGCCGACCGAGTTCAACGGCCGCTCCGACTACCAGGCCTTTATTCTCAACGGCATCCCCGCGGA                  GGCCTCTTACCAGGCGCGGAGGCGATCAAGACCGAGGAACAGGCCAAATTGTTTGGCGGCCAGGCCGGC                  GTGGCTCTGGACGCCAACTACCACGCCAAGGGTGACAACATGACTAATCTCAACCGCGAGGCTTTCCTG                  ATCAATTCCAGGGCGACGGCCTTTGCCGTGGCGACGTACGCCAACAGCCTTGACTCGATCCCCCACGC                  AACATGACCACCGTGGTCAAGCGGTGCGAGCTGGAGCAAGCCATGAAGAGGACCCCGCACACGCACACC                  GGCGGAACAGGATGCTACAAGGACCGGGTTGAGCAGTAG</p>
--

- 5 Un marco de lectura abierto (“ORF”) de fuLAP2 descrito de 1.497 nucleótidos empieza con un codón ATG (subrayado en la Tabla 3B) en la posición 1.

**Tabla 3B. Secuencia de nucleótidos de fuLAP2 (SEQ ID NO:8).**

ATGAAGCTGCTCTACCTCACATCGTTTTGCCTCTCTGGCCGTGGCCAATGGCCCAGGATGGGACTGGAAG  
 CCCCAGATTTCATCCGAAAGTCCGCCCCAAATGATCCATTTGTGGGATCTTCTGCAGGGCGCTCAACAG  
 CTGGAAGACTTCGCCTATGCCTACCCCGAGCGCAACCCGCTCTTTGGTGGACGGGCCACGAGGACACC  
 GTCAACTACCTCTACCGTGAGTTGAAGAAAACCGGCTACTACGACGTTTACAGCAGCCCCAGGTTAC  
 CAGTGGACCCGAGCCGACCAGGCTCTACCGTCGACGGCCAGTCTTATGACGCCACAACCATGACTTAC  
 AGCCCCAGCGTAAACGCCACGGCGCCGCTGGCAGTGGTGAACAACCTGGGCTGCGTCGAGGCTGACTAT  
 CCGCCGATCTGACGGGCAAGATTGCTCTGATCTCGCGGGGCGAGTGCACCTTTGCGACCAAATCCGTC  
 TTGAGCGCCAAGGCCGGGGCGGCGGCGGCACTCGTGTACAACAATATCGAGGGTTCGATGGCGGGA  
 CTGGGCGGCGGACAGCGAGCTGGTGCCTACGCTCCCATCGCCGCATCAGCCTCGCGGACGGACAG  
 GCGCTGATCCAGATGATCCAGGCGGCGAGGTCAGAGCCAACTGTGGATCGACAGCCAGTTCGAGAAC  
 CGTACCCTACAACGTGATCGCGCAGACCAAGGGCGGCGGCAACCAACAACGTCGTCGCGTGGGTGGC  
 CACACGGACTCGGTCGAGGCCGGGCCGCGCATCAACGACGACGGCTCCGGCATCATCAGCAACCTCGTC  
 GTCGCCAAGGCGCTGACCCGCTTCTCGGTCAAGAACGCGGTGCGCTTCTGCTTCTGGACGGCGGAGGAG  
 TTCGGCCTGCTGGGCAGCAACTACTACGTCAACAGCCTCAATGCCACCGAGCAGGCCAAGATCCGCCTG  
 TATCTCAACTTCGACATGATCGCCTCCCCAACTACGCCCTGATGATCTATGACGGCGACGGCTCGGCC  
 TTCAACCTGACGGGGCCGGCCGGCTCGGCGCAGATCGAGCGGCTCTTCGAGGACTACTACAGTCGATC  
 CGCAAGCCGTTCTGTCGCGACCGAGTTCAACGGCCGCTCCGACTACCAGGCCTTTATTCTCAACGGC  
 CCGCGGGAGGCTCTTACCAGGCGGCGGAGGCGATCAAGACCGAGGAAACAGGCCCAATTGTTTGGCGGC  
 CAGGCCGGCTGGCTCTGGACGCCAACTACCACGCCAAGGGTGACAACATGACTAATCTCAACCGCGAG  
 GCTTCTGATCAATTCCAGGCGGCGGCTTTGCCGTGGCGACGTCGCAACAGCCTTGACTCGATC  
 CCCCCAGCAACATGACCACCGTGGTCAAGCGGTGCGAGCTGGAGCAAGCCATGAAGAGGACCCCGCAC  
 ACGCACACCGGCGGAACAGGATGCTACAAGGACCGGGTTGAGCAGTAG

Un ácido nucleico de fuLAP2 descrito (SEQ ID NO: 8) codifica una proteína que tiene 498 restos de aminoácidos (SEQ ID NO: 9), que se presenta en la Tabla 3C usando el código de aminoácidos de una letra.

**Tabla 3C. Secuencia de la proteína codificada fuLAP2 (SEQ ID NO:9).**

MKLLYLTSFASLAVANGPGWDWKPRVHPKVLPQMIHLWDLQGAQQLDFAYAYPERNRVFGGRAHEDT  
 VNYLYRELKKTGYDVKQPQVHQWTRADQALTVDGQSYDATMTYSPSVNATAPLAVVNNLGCVEADY  
 PADLTGKIALISRGECTFATKSVLSAKAGAAAALVYNNIEGSMAGTLLGGATSELGAYAPIAGISLADGQ  
 ALIQMIQAGTVTANLWIDSQVENRTTYNVIAQTKGGDPNNVVALGGHTDSVEAGPGINDDGGIISNLV  
 VAKALTRFSVKNAVRFCFWTAEFGLLGSNYVNSLNATEQAKIRLYLNFDMIASPNYALMIYDGDGSA  
 FNLTGPAQSAQIERLFEDYYSIRKPFVPTFNGRSDYQAFILNGIPAGGLFTGAEAIKTEEQALFGG  
 QAGVALDANYHAKGDNMTNLNREAFILNSRATAFAVATYANSLDISIPRNMVVVKRSQLEQAMKRTPH  
 THTGGTGCKDRVEQ

5 La fuLAP2 descrita tiene homología con las secuencias de aminoácidos mostradas en los datos BLAST listados en las Tablas 3D, 3E y 3F. Estos datos se analizaron por el programa POR PAREJAS BLAST.

**Tabla 3D: Resultados TBLASTN para fuLAP2**

Índice/Identificador Génico	Proteína/Organismo	Longitud (aa)	Identidad (%)	Positivos (%)	Esperado
>gi469463	<i>Saccharomyces cerevisiae</i> /gen de aminopeptidasa Y	2272	184/464 (39%)	243/464 (52%)	7e-69
>gi9949032	PAO1 de <i>Pseudomonas aeruginosa</i> , sección de 281 de 529 del genoma completo	12547	165/445 (37%)	231/445 (51%)	9e-67
>gi23017467	CDC1551 de <i>Mycobacterium tuberculosis</i> , sección 33 de 280 del genoma completo	18857	166/426 (38%)	218/426 (51%)	2e-62

Tabla 3E: Resultados BLASTX para fuLAP2

Índice/Identificador Génico	Proteína/Organismo	Longitud (aa)	Identidad (%)	Positivos (%)	Esperado
>gi28918599	Proteína hipotética/ <i>Neurospora crassa</i>	508	250/479 (52%)	314/479 (65%)	e-131
>gi23017467	Proteína hipotética/ <i>Thermobifida fusca</i>	514	173/465 (37%)	251/465 (53%)	4e-74
>gi584764	APE3 LEVADURA; Precursor de aminopeptidasa/ <i>Saccharomyces cerevisiae</i>	537	184/464 (39%)	243/464 (52%)	8e-70
>gi15598135	Aminopeptidasa probable/PAO1 de <i>Pseudomonas aeruginosa</i>	536	165/445 (37%)	231/445 (51%)	1e-67
>gi15839805	Hidrolasa/ <i>Mycobacterium tuberculosis</i> CDC15551	493	166/426 (38%)	218/426 (51%)	3e-63

Tabla 3F. Resultados BLASTP para fuLAP2

Índice/Identificador Génico	Proteína/Organismo	Longitud (aa)	Identidad (%)	Positivos (%)	Esperado
>gi28918599	Proteína hipotética/ <i>Neurospora crassa</i>	508	250/469 (52%)	314/479 (65%)	e-128
>gi23017467	Proteína hipotética/ <i>Thermobifida fusca</i>	514	173/465 (37%)	251/465 (53%)	3e-71
>gi584764	APE3 LEVADURA; Precursor de aminopeptidasa/ <i>Saccharomyces cerevisiae</i>	537	183/464 (39%)	243/464 (52%)	6e-70
>gi15598135	Aminopeptidasa probable/PAO1 de <i>Pseudomonas aeruginosa</i>	536	164/445 (36%)	230/445 (51%)	3e-65

**fuLAP1**

- 5 fuLAP1 es una leucina aminopeptidasa de *A. fumigatus*. Un ácido nucleico de fuLAP1 de 1.298 nucleótidos se muestra en la Tabla 4A (SEQ ID NO: 10).



**Tabla 4A. Secuencia de nucleótidos genómica de fuLAP1 (SEQ ID NO:10).**

```

ATGAAAGTTCTTACAGCTATTGCGCTGAGCGCAATAGCTTTCACAGGGGCTGTAGCTGCAGTGATTACT
CAGGAAGCATTCTTAAACAACCCCCGCATCCATCATGACCAGGAGAAGTACTTGATCGAACTGGCCCCCT
TATCGAACACGATGGGTGACTGAAGAGGAGAAATGGGCATTGAAATTGGTACCATACTTCCCCAAAATT
TGGGTCTCCAAGTCCACGGGCGACTAACTGCACGATTGCTTGAAGGACGGCGTGAATTTTATCGATATC
ACAGAAGAGCACAAACACCGGATTTTACCCGACTCTCCACAGCGCCAGCTATGTGAAATATCCACCGAAG
ATGCAGTATGCAGAAGAAGTGGCTGCTCTTAACAAGAATTTATCGAAAGAAAACATGAAGGCCAACCTG
GAACGATTACATCATTTCATACTCGCTATTACAAATCTCAGACGGGAATCCGATCGGCAACGTGGCTG
TTCGACCAAGTTCAGAGAGTTGTCTCTGAGTCTGGAGCCGCTGAGTATGGTGCAACTGTTGAGCGATTCT
TCTCATCCATGGGGTCAGTTCAGCATTATTGCCCGAATACCCGGCCGAACGAACAAGACTGTGGTGCTG
GGCGCCCATCAGGACAGCATCAATTTGTTTCTCCCGTCAATCTTGGCTGCTCCCGGTGCTGATGACGAT
GGAAGTGGAACTGTCACCATTCTTGAAGCGTTGCGCGGTCTGCTGCAGTCAGACGCCATTGCCAAGGGT
AATGCATCCAATACTGTTCGAGTTCCTGAGTCTGAGTCTGCAGAAGAAGGCGGAATGCTGGGCTCCCAGGCA
ATATTTTCCAATTACAAGCGGAATAGGCGGGAAATCAAAGCCATGCTCCAGCAAGACATGACTGGCTAC
GTCCAGGGAGCTTTGAACGCGGTGTTGAGGAAGCCATAGGAATTATGGTTCGATTATGTCGACCAGGGC
CTCACACAGTTTCTCAAGGACGTTGTTACAGCGGTAAGCCTCAGTTGTCCCCACGAAAAGCTGTTTAG
TCGACAAATGAAATTGACGGCTGCATTAGTACTGCTCTGTGGGTTACCTGGAGACGAAGTGGGATATG
CCTGCTCCGACCACACCTCGGCCAGTAAATATGGTTATCCCGCGGCTATGGCGACAGAAGCAGAGATGG
AAAATACCAATAAGAAGATACATACTACCGACGACAAGATCAAGTATTTGAGCTTCGATCATATGTTGG
AGCATGCCAAGTTGAGTCTTGGCTTCGCTTTCGAATTGGCATTGCGCCGTTTTAA

```

Un marco de lectura abierto ("ORF") de fuLAP1 descrito de 1.167 nucleótidos empieza con un codón ATG en la posición 1 (subrayado en la Tabla 4B).

**Tabla 4B. Secuencia de nucleótidos de fuLAP1 (SEQ ID NO:11).**

```

ATGAAAGTTCTTACAGCTATTGCGCTGAGCGCAATAGCTTTCACAGGGGCTGTAGCTGCAGTGATTACT
CAGGAAGCATTCTTAAACAACCCCCGCATCCATCATGACCAGGAGAAGTACTTGATCGAACTGGCCCCCT
TATCGAACACGATGGGTGACTGAAGAGGAGAAATGGGCATTGAAATTGGACGGCGTGAATTTTATCGAT
ATCACAGAAGAGCACAAACACCGGATTTTACCCGACTCTCCACAGCGCCAGCTATGTGAAATATCCACCG
AAGATGCAGTATGCAGAAGAAGTGGCTGCTCTTAACAAGAATTTATCGAAAGAAAACATGAAGGCCAAC
CTGGAACGATTACATCATTTCATACTCGCTATTACAAATCTCAGACGGGAATCCGATCGGCAACGTGG
CTGTTTCGACCAAGTTCAGAGAGTTGTCTCTGAGTCTGGAGCCGCTGAGTATGGTGCAACTGTTGAGCGA
TTCTCTCATCCATGGGGTCAGTTCAGCATTATTGCCCGAATACCCGGCCGAACGAACAAGACTGTGGTG
CTGGGCGCCCATCAGGACAGCATCAATTTGTTTCTCCCGTCAATCTTGGCTGCTCCCGGTGCTGATGAC
GATGGAAGTGGAACTGTCACCATTCTTGAAGCGTTGCGCGGTCTGCTGCAGTCAGACGCCATTGCCAAG
GGTAATGCATCCAATACTGTTCGAGTTCCTGAGTCTGAGTCTGCAGAAGAAGGCGGAATGCTGGGCTCCCAG
GCAATATTTTCCAATTACAAGCGGAATAGGCGGGAAATCAAAGCCATGCTCCAGCAAGACATGACTGGC
TACGTCCAGGGAGCTTTGAACGCGGTGTTGAGGAAGCCATAGGAATTATGGTTCGATTATGTCGACCAG
GGCCTCACACAGTTTCTCAAGGACGTTGTTACAGCGTACTGCTCTGTGGGTTACCTGGAGACGAAGTGC
GGATATGCCTGCTCCGACCACACCTCGGCCAGTAAATATGGTTATCCCGCGGCTATGGCGACAGAAGCA
GAGATGGAAAATACCAATAAGAAGATACATACTACCGACGACAAGATCAAGTATTTGAGCTTCGATCAT
ATGTTGGAGCATGCCAAGTTGAGTCTTGGCTTCGCTTTCGAATTGGCATTGCGCCGTTTTAA

```

- 5 Un ácido nucleico de fuLAP1 descrito (SEQ ID NO: 11) codifica una proteína que tiene 388 restos de aminoácidos (SEQ ID NO: 12), que se presenta en la Tabla 4C usando el código de aminoácidos de una letra.

**Tabla 4C. Secuencia de la proteína codificada fuLAP1 (SEQ ID NO:12).**

```

MKVLTAI ALSAIAFTGAVAAVITQEAFLNNPRIHHDQEKYLI ELAPYRTRRWVTEEEKWALKLDGVNFI D
ITEEHNTGFYPTLHSASYVKYPPKMQYAEVAALNKNLSKENMKANLERFTSFHTRYYSQTGIRSATW
LFDQVQRVSES GAAEYGATVERF SHPWGQFSI IARI PGR TNKT VVLGAHQDS INLFLPSILAAPGADD
DGS GTVTI LEALRGLLQSDAI AKGNASNTVE FHWYSAEEGMLGSQAI FSNYKRNREI KAMLQDDMTG
YVQ GALNAGVEEAIGIMVDYVDQGLTQFLKDVVTAYCSVGYLET KCGYACSDHTSASKYGYPAAMATEA
EMENTNKKIHTTDDKI KYLSFDHMLEHAKLSLGF AFELAFAPF

```

La fuLAP1 descrita tiene homología con las secuencias de aminoácidos mostradas en los datos BLAST listados en las Tablas 4D, 4E y 4F. Estos datos se analizaron por el programa POR PAREJAS BLAST.

**Tabla 4D: Resultados TBLASTN para fuLAP1**

Índice/Identificador Génico	Proteína/Organismo	Longitud (aa)	Identidad (%)	Positivos (%)	Esperado
>gi1762234	Poliquétido sintasa PKSL2/ <i>Aspergillus parasiticus</i>	9894	208/249 (80%) 61/84 (72%) 46/62 (74%)	226/249 (90%) 67/84 (79%) 55/62 (88%)	e-169
>gi23393798	Leucina aminopeptidasa (Lap1)/ <i>Aspergillus sojae</i>	2547	66/110 (60%) 68/152 (44%) 37/75 (49%) 15/30 (50%)	82/110 (74%) 92/152 (60%) 52/75 (69%) 21/30 (70%)	7e-82
>gi927685	Cromosoma IV lambda3641 y cósmido 9831 y 9410 de <i>Saccharomyces cerevisiae</i>	78500	152/341 (44%)	207/341 (60%)	1e-71
>gi5832144	Biblioteca de ADNc de la cepa T4 de <i>Botrytis cinerea</i> en condición de carencia de nitrógeno	780	89/134 (66%) 27/53 (50%)	106/134 (79%) 33/53 (62%)	7e-58

5

**Tabla 4E: Resultados BLASTX para fuLAP1**

Índice/Identificador Génico	Proteína/Organismo	Longitud (aa)	Identidad (%)	Positivos (%)	Esperado
>gi28918132	Proteína hipotética/ <i>Neurospora crassa</i>	402	208/352 (59%)	255/352 (72%)	e-116
>gi23393799	Leucina aminopeptidasa/ <i>Aspergillus sojae</i>	377	183/355 (51%)	241/355 (67%)	3e-97
>gi6320623	ORF hipotético; ydr415cp/ <i>Saccharomyces cerevisiae</i>	374	152/341 (44%)	207/341 (60%)	2e-72
>gi18250467	Aminopeptidasa/ <i>Agaricus bisporus</i>	384	139/352 (39%)	186/352 (52%)	1e-58

**Tabla 4F. Resultados BLASTP para fuLAP1**

Índice/Identificador Génico	Proteína/Organismo	Longitud (aa)	Identidad (%)	Positivos (%)	Esperado
>gi28918132	Proteína hipotética/ <i>Neurospora crassa</i>	402	208/352 (59%)	255/352 (72%)	e-116
>gi23393799	Leucina aminopeptidasa (LAP1)/ <i>Aspergillus sojae</i>	377	183/355 (51%)	241/355 (67%)	6e-98
>gi6320623	ORF hipotético; ydr415cp/ <i>Saccharomyces cerevisiae</i>	374	152/341 (44%)	207/341 (60%)	3e-73
>gi18250467	Aminopeptidasa/ <i>Agaricus bisporus</i>	384	140/352 (39%)	190/352 (53%)	7e-59

**ruCBPS1**

ruCBPS1 es una carboxipeptidasa de *T. rubrum*. La secuencia de ADN genómico de un ácido nucleico de ruCBPS1 de 2.106 nucleótidos (SEQ ID NO: 13) se muestra en la Tabla 5A.

**Tabla 5A. Secuencia de nucleótidos genómica de ruCBPS1 (SEQ ID NO:13).**

ATGGTGTCAATCTGCGGAGTGGCAGCCTGCCTGCTGACAGTTGCTGGCCATCTTGCGCAGGCTCAGTTC CCACAAAACCGGAGGGAGTCACTGTCCTGGAGTCGAAATTCGGCAGCGGTGCTCGCATCACTTATAAG GAGGTCCGTTAGCTGCATAGAAAAGTCCACGTGAAGACGCTGTAGCTAACAATCCACTAGCCTGGCCTCT GTGAGACGACAGAGGGCGTCAAGTCGTACGCCGATATGTCCATCTGCCTCCAGGCACGCTCAGGGACT TCGGTGTCGAGCAGGACTACCCTATCAACACCTTTTTTTGGTCTTTGAGGCAAGAAAGGACCCTGAAA ATGCCCTCTCGGCATCTGGATGAACGGTGGCCCGGGTAGCTCGTCGATGTTTGGAAATGATGACTGAGA ACGGGCCTTGCTTCGTCAATGCAGACTCCAATTCTACTCGCCTGAACCCTCATTCTTGGAAACAATGAAG GTATGCCATCAGCTTCTGATGGAAAACATAATATTGCTAACATTGTACTTTCTGTGACTAGTCAATATG CTGTATATAGACCAGCCAGTGCAGGTCGGTCTGTCTACGACACTTTGGCCAACCTCACCAGGAATCTA GTCACGGATGAGATCACGAACTGAAACCCGGAGAACCTATCCGGAACAGAATGCCACTTTCCTGGTA GGTACATATGCAAGCCGCAATATGAACACCACTGCACACGGAAC TAGGCATGCTGCCATGGCTCTCTGG CACTTCGCCCAAGTCTGGTTCCAAGAGTCCCAGGATATCACCTAGGAACAACAAGATCAGCATTGCT ACCGAATCCTACGGCGGCCGTTATGGCCCGCCTTTACTGCCTTCTTTGAAGAGCAGAACCAGAAGATC AAGAACGGCACATGGAAGGGACACGAGGGAACATGCACGTGCTGCATCTCGACACCCTCATGATCGTC AACGGATGCATCGACCGTCTTGTTCCAATGGCCGGCATATCCGCAAATGGCGTACAACAACACATATAGC ATCGAGGCAGTCAACGCCAGCATTCATGCAGGAATGCTGGATGCCCTTACC CGGACGGTGGCTGTGCGA GACAAGATTAACCACTGCCGCTCCCTCTCTTCTGTGTTGATCCTGAGAATCTCGGCATCAACTCAACC GTCAATGATGCTGCAAGGATGCCGAGACATTCTGCTCCAATGATGTTGCGGATCCCTACCTCAAGTTC TCTGGCCGCAACTACTATGACATCGGACAGCTTGACCCAGCCCATTCCAGCACCATTTTACATGGCC TGGCTAAATCAGCCGCATGTGCAGGCAGCACTGGGTGTGCCACTTAACTGGACACAGTCAAACGATGTT GTGTCTACCGCATTCCGTGCAATTGGTGAATACCTCGGCCAGGGTGGCTGGAGAACCTGGCTTATTTG CTGGAGAATGGCATCAAGGTTTCGCTTGTTTACGGTGATCGGGACTACGCATGCAACTGGTTCGGTGGT GAGCTCTCAAGTCTGGGAATCAACTACACTGACACCCACGAATCCATAATGCCGGCTATGCAGGTATC CAGATCAATAGCAGCTACATCGGTGGACAGGTGAGGCAGTACGGCAACCTCTCCTTCGCCC GCGTCTAC GAGGCCGGCCATGAGGTGCCATCGTACCAACCCGAGACTGCATGCAGATATCCACCGTCCCTGTTC AACAAAGGATATCGCTACTGGAACCAAGGACAGTCCATCGCGCATGGACGGAGGCAAGTTTACGGCACC TCGGCCCTGCGGACTCGTTTGGTTTCAAGAACAACAACTCCACCGCAGCACGTCCACTTCTGTCAATC TTAGACACCAGCACTGCACCAAGGAGCAGATCCAGTCAAGTGGAGAACGGCACTGCCCGGTACGCAGC TGGATCATTGTCGACTCCAACCTCGACCTCTCTGTTCCCGAGGTAGTTGGCTCAGGGGAACCCACGCCA ACCCCTATGCCTGGAGGGGCTACTACACTATCTGCTCACGGGTCTTGTATGGCGTGACATATGGGCT GTTATTGTTGTAGCTGTTATAGACTGGCAATGTAA
--

Un ácido nucleico de ruCBPS1 de 1.989 (SEQ ID NO: 14) se muestra en la Tabla 5B. Un marco de lectura abierto ("ORF") de ruCBPS1 descrito empieza con un codón de inicio ATG en la posición 1 (subrayado en la Tabla 5B).

<b>Tabla 5B. Secuencia de nucleótidos de ruCBPS1 (SEQ ID NO:14).</b>
<p> <u>ATGGTGT</u>CATTCTGCGGAGTGGCAGCCTGCCTGCTGACAGTTGCTGGCCATCTTGCGCAGGCTCAGTTC                      CCACCAAACCGGAGGGAGTCACTGTCCTGGAGTCGAAATTCGGCAGCGGTGCTCGCATCACTTATAAG                      GAGCCTGGCCTCTGTGAGACGACAGAGGGCGTCAAGTCGTACGCCGATATGTCCATCTGCCTCCAGGC                      ACGCTCAGGGACTTCGGTGTGAGCAGGACTACCCTATCAACACCTTTTTTTGGTTCTTTGAGGCAAGA                      AAGGACCCTGAAAATGCCCCTCTCGGCATCTGGATGAACGGTGGCCCGGGTAGCTCGTCGATGTTTGA                      ATGATGACTGAGAACGGGCCTTGCTTCGTCAATGCAGACTCCAATTCTACTCGCCTGAACCCCTATTCT                      TGGAACAATGAAGTCAATATGCTGTATATAGACCAGCCAGTGCAGGTCGGTCTGTCCACGACACTTTG                      GCCAACTTCACCAGGAATCTAGTCACGGATGAGATCACGAACTGAAACCCGGAGAACCCTATTCCGGAA                      CAGAATGCCACTTTCTGGTAGGTACATATGCAAGCCGCAATATGAACACCACTGCACACGGAAGTAGG                      CATGCTGCCATGGCTCTCTGGCACTTCGCCCAAGTCTGGTTCGAAGAGTTCACAGGATATCACCCTAGG                      AACAACAAGATCAGCATTGCTACCGAATCTACGGCGGCCGTTATGGCCCGGCCTTTACTGCCTTCTTT                      GAAGAGCAGAACCAGAAGATCAAGAACGGCACATGGAAGGGACACGAGGGAACATGCACGTGCTGCAT                      CTCGACACCCTCATGATCGTCAACGGATGCATCGACCGTCTTGTCCAATGGCCGGCATATCCGCAAATG                      GCGTACAACAACACATATAGCATCGAGGCAGTCAACGCCAGCATTATGCAGGAATGCTGGATGCCCTC                      TACCGCGACGGTGGCTGTGCGAGACAAGATTAACCACTGCCGCTCCCTCTCTTCTGTGTTGATCCTGAG                      AATCTCGGCATCAACTCAACCGTCAATGATGTCTGCAAGGATGCCGAGACATTCTGCTCCAATGATGTT                      CGCGATCCCTACCTCAAGTTCTCTGGCCGCAACTACTATGACATCGGACAGCTTGACCCAGCCATTTC                      CCAGCACCATTTTACATGGCCTGGCTAAATCAGCCGCATGTGCAGGCAGCACTGGGTGTGCCACTTAAC                      TGGACACAGTCAAACGATGTTGTGTCTACCGCATTCCGTGCAATTGGTGACTACCCTCGGCCAGGGTGG                      CTGGAGAACCTGGCTTATTTGCTGGAGAATGGCATCAAGGTTTCGCTTGTTTACGGTGATCGGGACTAC                      GCATGCAACTGGTTCGGTGGTGAGCTCTCAAGTCTGGGAATCAACTACACTGACACCCACGAATCCAT                      AATGCCGGCTATGCAGGTATCCAGATCAATAGCAGCTACATCGGTGGACAGGTGAGGCAGTACGGCAAC                      CTCTCCTTCGCCCGCGTCTACGAGGCCGGCCATGAGGTGCCATCGTACCAACCCGAGACTGCACTGCAG                      ATATCCACCGTTCCTGTTCACAAGGATATCGCTACTGGAACCAAGGACACGTATCGCGCATGGAC                      GGAGGCAAGTTTTACGGCACCTCGGGCCCTGCGGACTCGTTTGGTTTTCAAGAACAACCTCCACCGCAG                      CACGTCCACTTCTGTATATCTTAGACACCAGCAGCTGCACTCAACCAAGGAGCAGATCCAGTCAAGTGA                      GAGTGCAGCTGCGCGTACGAGCTGGATCATTGTGCTGCACTCAACTCGACCTCTCTGTCCCGAGGTAGTT                      GGCTCAGGGGAACCCACGCCAACCCTATGCCTGGAGGGGCTACTACACTATCTGCTACGGGTTCTTG                      TATGGCGTGACATTATGGGCTGTTATTGTTGTAGCTGTTATAGAGCTGGCAATGTAA                 </p>

5 Un ácido nucleico de ruCBPS1 descrito (SEQ ID NO: 14) codifica una proteína que tiene 662 restos de aminoácidos (SEQ ID NO: 15), que se presenta en la Tabla 5C usando el código de aminoácidos de una letra.

<b>Tabla 5C. Secuencia de la proteína codificada ruCBPS1 (SEQ ID NO:15).</b>
<p>                     MVSFCGVAACLLTVAGHLAQAQFPKPEGVTVLESKFGSGARITYKEPGLCETTEGVKSYAGYVHLPPG                      TLRDFGVEQDYPINTFFWFPEARKDPENAPLGIWMNGGPGSSSMFGMMTENGPCFVNADSNSTRLNPHS                      WNNEVNMLYIDQPVQVGLSYDTLANFTRNLVTEITKLKPGEP I PEQNATFLVGTYASRNMNTTAHGTR                      HAAMALWHFAQVWFQEFPGYHPRNKNKIS IATESYGGRYGPAFTAF FEEQNQKIKNGTWKGHEGTMHVLH                      LDTLMIVNGCIDRLVQWPAYPQMAYNNTYSIEAVNAS IHAGMLDALYRDGGCRDKINHCRSLSSVFDPE                      NLGINSTVNDVCKDAETFCSDNVRDPYLKFSGRNYD I GQLDPSFP PAFPYMAWLNQPHVQAALGVPLN                      WTQSNVSTAFRAIGDYPRPGWLENLAYLLENGIKVSLVYGDYACNWFGGELSSLGINYTDTHEFH                      NAGYAGIQINSSYIGGQVRQYGNLS FARVYEAGHEVPSYQPETALQIFHRSLFNKDIATGKDTSSRMD                      GPKFYGTSGPADSFGPKNPPPQHVHFCCHILDSTCTKEQIQSVENGTAAVRSWIIVDSNSTSLFPEVV                      GSGEPTPTMPGGATLSAHGFLYGVTLWAVIVVAVIELAM                 </p>

La ruCBPS1 descrita tiene homología con las secuencias de aminoácidos mostradas en los datos BLAST listados en las Tablas 5D, 5E y 5F. Estos datos se analizaron por el programa POR PAREJAS BLAST.

Tabla 5D: Resultados TBLASTN para ruCBPS1

Índice/Identificador Génico	Proteína/Organismo	Longitud (aa)	Identidad (%)	Positivos (%)	Esperado
gi/32410708	Cepa OR74A de <i>Neurospora crassa</i>	1947	222/632 (35%)	321/632 (50%)	1e-90
gi/3046860	gen cpy1 de <i>Schizosaccharomyces pombe</i> para carboxipeptidasa Y	4308	137/481 (28%)	204/481 (42%)	6e-41
gi/18152938	Gen de <i>Pichia angusta</i> para carboxipeptidasa Y (CPY)	2214	141/520 (27%)	228/520 (43%)	4e-40
gi/4028157	Gen de <i>Pichia angusta</i> para el precursor de la carboxipeptidasa Y (CPY)	2509	140/520 (26%)	226/520 (43%)	7e-40
gi/170828	Gen de <i>Candida albicans</i> para el precursor de la carboxipeptidasa Y (CPY)	1985	131/482 (27%)	205/482 (42%)	3e-36

Tabla 5E: Resultados BLASTX para ruCBPS1

Índice/Identificador Génico	Proteína/Organismo	Longitud (aa)	Identidad (%)	Positivos (%)	Esperado
gi/15004616	carboxipeptidasa S1/ <i>Aspergillus oryzae</i>	555	209/535 (39%)	294/535 (54%)	1e-98
gi/435818	carboxipeptidasa S1, CPD-S1/ <i>Penicillium janthinellum</i>	423	159/498 (31%)	234/498 (46%)	6e-64
gi/995456	prepro-carboxipeptidasa Z/ <i>Absidia zychnae</i>	460	147/506 (29%)	219/506 (43%)	8e-48
gi/3046861	carboxipeptidasa Y/ <i>Schizosaccharomyces pombe</i>	1002	137/481 (28%)	204/481 (42%)	7e-42
gi/18152939	carboxipeptidasa Y/ <i>Pichia angusta</i>	537	141/520 (27%)	228/520 (43%)	4e-41
gi/4028158	precursor de la carboxipeptidasa Y; carboxipeptidasa vacuolar/ <i>Pichia angusta</i>	541	140/520 (26%)	226/520 (43%)	7e-41
gi/7597001	precursor de la carboxipeptidasa Y/ <i>Candida albicans</i>	542	131/482 (27%)	206/482 (42%)	2e-37

Tabla 5F: Resultados BLASTP para ruCBPS1

Índice/Identificador Génico	Proteína/Organismo	Longitud (aa)	Identidad (%)	Positivos (%)	Esperado
gi/15004616	carboxipeptidasa S1/ <i>Aspergillus oryzae</i>	555	210/537 (39%)	296/537 (55%)	2e-95
gi/435818	carboxipeptidasa S1, CPD-S1/ <i>Penicillium janthinellum</i>	423	159/498 (31%)	234/498 (46%)	2e-60
gi/995456	prepro-carboxipeptidasa Z/ <i>Absidia zychnae</i>	460	146/500 (29%)	217/500 (43%)	6e-47
gi/19115337	carboxipeptidasa Y/ <i>Schizosaccharomyces pombe</i>	1002	136/481 (28%)	204/481 (42%)	7e-41

**ruCBPS1'**

ruCBPS1' es una carboxipeptidasa de *T. rubrum*. La secuencia de ADN genómico de un ácido nucleico de ruCBPS1' de 2.030 nucleótidos (SEQ ID NO: 16) se muestra en la Tabla 6A.

**Tabla 6A. Secuencia de nucleótidos genómica de ruCBPS1' (SEQ ID NO:16).**

```

ATGCGCTTTGCTGCTAGCATTGCCGTGGCCCTGCCAGTCATTACGCGGCGAGTGTCTCAAGGCTTCCCT
CCACCCGTTAAGGGCGTCACCGTGGTCAAATCCAAGTTCGACGAAAACGTAAAGATCACATACAAGGAG
GTATGTGTTTACATCATTTCACATCCAGATCTTATATCCTTACAATAAATCTGGCTAACTCACTGGAT
AGAATGACATATGTGAAACCACTCAAGGAGTTAGATCATTACCGGTCATGTCCACCTTCTCCAGACA
ACGATGACTTTGGTGTCTACCGGAACTACTCCATCAACACATTCTTCTGGTTCTTTGAAGCTCGTGAAG
ACCCTAAGAATGCTCCTCTCTCCATCTGGCTGAACGGTGGTCCGGGATCGTCATCCATGATTGGACTCT
TCCAGGAAAACGGTCCATGCTGGGTCAATGAAGACTCTAAATCTACCACCAACAATTCATTTTCATGGA
ACAATAAGTAAATATGCTCTACATTGATCAGCCAAACCAAGTCGGTTTCAGTTATGACGTACCTACCA
ACATCACTTACTCTACCATCAATGATACAATATCTGTTGCGGACTTCTCTAACGGTGTCCCTGCGCAA
ATCTTTCTACGTTGGTTGGAACCGGCAGCAGCCAGAACCCTTGGGCAACTGCCAATAACACTGTGAACG
CTGCTCGTTCTATCTGGCACTTTGCACAAGTGTGGTCCAGGAATCCCTGAACACAAGCCTAACATA
ACAAGATCAGTATTTGGACAGAGTCTATGGAGGAAGATATGGTCCCTCATTGCGCTTACTTCCAGG
AACAGAACGAAAAGATCAAAAACCATAACCATTACTGAAGAAGGAGAGATGCATATTTCTGAACCTCGACA
CCCTCGGTATCATCAACGGCTGCATCGATCTTATGTTCCAAGCAGAAAGTTATGCTGAATCCCATACA
ACAACACCTATGGCATCAAAGCTTATACCAAGGAGAAGCGTGACGCTATATTACACGACATCCACCGTC
CTGACGGCTGCTTCGACAAGGTTACCAAGTGCCGTGAGGCCGCGAAAGAAGGAGACCTCACCTTCTACA
GCAACAATGCAACCGTCAACACAATCTGTGCGGATGCTAACTCTGCCTGCGACAAATATCTAATGGATC
CTTCCAAAGAGACCAATCTTGGTTACTATGATATTGCTCATCCTCTTCAGGATCCCTTCCCCCACCAT
TCTATAAGGGCTTCCCTCAGCCAATCCAGCGTTCTATCTGACATGGGATCGCCAGTCAACTTCTCCCAAT
ACGCCAAGCTGTGGGAAAATCATTCCATGGAGTTGGCGACTACGCTCGCCCTGATGTGCGCGGCTTCA
CCGGTGACATTGCTTATCTTCTCGAGAGCGGAGTCAAGGTTGCTCTCGTCTATGGTGACAGAGACTACA
TCTGCAATTGGTTCCGGTGGTGAGCAGTCAAGTCTTGGCTTGAACACTACTGGCACCCAAAGACTTCCACA
GGGCAAAATATGCCGATGTCAAGGTCAACTTTCATACGTCGGAGGCGTAGTGCCTCAACATGGAAACT
TCTCTTTACCAGAGTTTCGAGGCCGGTCAAGTCCCTGGTTACCAACCCGAGACTGCCCTCAAGA
TCTTTGAGCGCATCATGTTCAACAAGGATATTTCTACCGGTGAGATCGACATTGCTCAGAAAACCAGACT
ACGGTACCACTGGAACCTGAGTCTACGTTCCATATCAAAAACGATATCCCTCCTTCGCTGAGCCGACCT
GCTACCTCCTCAGTGTGACGGAACCTGTACCCCGGAGCAGCTTAATGCTATTAAGGATGGAACCTGCAG
TTGTTGAGAATACTATTATTAAGAGCCCTGCTGCGTCGAAGGGGAACCCTCCACCAACCAGACCTCAT
CTCCACAGCAGCCCTACCGTGGAAAGTGCCATGCTAAAGGCTCCTGTGGCAATGCTAGCAATATCAG
CTCTCACTGTCTTGTCTTCTTCTTGTAG
    
```

5

Un ácido nucleico de ruCBPS1' de 1.959 (SEQ ID NO: 17) se muestra en la Tabla 6B. Un marco de lectura abierto ("ORF") de ruCBPS1' descrito empieza con un codón de inicio ATG en la posición 1 (subrayado en la Tabla 6B).

**Tabla 6B. Secuencia de nucleótidos de ruCBPS1' (SEQ ID NO:17).**

```

ATGCGCTTTGCTGCTAGCATTGCCGTGGCCCTGCCAGTCATTACGCGGCGAGTGCTCAAGGCTTCCCT
CCACCCGTTAAGGGCGTCACCGTGGTCAAATCCAAGTTCGACGAAAACGTAAAGATCACATACAAGGAG
AATGACATATGTGAAACCACTCAAGGAGTTAGATCATTACCGGTCATGTCCACCTTCTCCAGACAAC
GATGACTTTGGTGTCTACCGGAACTACTCCATCAACACATTCTTCTGGTTCTTTGAAGCTCGTGAAGAC
CCTAAGAATGCTCCTCTCTCCATCTGGCTGAACGGTGGTCCGGGATCGTCATCCATGATTGGACTCTTC
CAGGAAAACGGTCCATGCTGGGTCAATGAAGACTCTAAATCTACCACCAACAATTCAATTTTCATGGAAC
AATAAAGTAAATATGCTCTACATTGATCAGCCAAACCAAGTCGGTTTCAGTTATGACGTACCTACCAAC
ATCACTTACTCTACCATCAATGATACAATATCTGTTGCGGACTTCTCTAACGGTGTCCCTGCGCAAAT
CTTTCTACGTTGGTTGGAACCGGCAGCAGCCAGAACCCTTGGGCAACTGCCAATAACACTGTGAACGCT
GCTCGTTCTATCTGGCACTTTGCACAAGTGTGGTTCCAGGAATTCCTGAACACAAGCTTAAACAATAAC
AAGATCAGTATTTGGACAGAGTCCATATGGAGGAAGATATGGTCCCTCATTTCGCCTCTTACTCCAGGAA
CAGAACGAAAAGATCAAAAACCATAACCATTACTGAAGAAGGAGAGATGCATATTCTGAACCTCGACACC
CTCGGTATCATCAACGGCTGCATCGATCTTATGTTCCAAGCAGAAAAGTTATGCTGAATTCCCATACAAC
AACACCTATGGCATCAAAGCTTATACCAAGGAGAAGCGTGACGCTATATTACACGACATCCACCGTCTT
GACGGCTGCTTCGACAAGGTTACCAAGTGCCGTGAGGCCGCGAAAAGAAGGAGACCCTCACTTCTACAGC
AACAAATGCAACCGTCAACACAATCTGTGCGGATGCTAACTCTGCCTGCGACAAATATCTAATGGATCCT
TTCCAAGAGACCAATCTTGTTACTATGATATTGCTCATCCTCTTCAGGATCCCTTCCCCCACCATTTC
TATAAGGGCTTCTCAGCCAATCCAGCGTCTATCTGACATGGGATCGCCAGTCAACTTCTCCCAATAC
GCCAAGCTGTGGGAAAATCATTCCATGGAGTTGGCGACTACGCTCGCCCTGATGTGCGCGGCTTACC
GGTGACATTGCTTATCTTCTCGAGAGCGGAGTCAAGGTTGCTCTCGTCTATGGTGACAGAGACTACATC
TGCAATTGGTTCGGTGGTGAGCAGGTGAGTCTTGGCTTGAACACTGACGACCCCAAGACTTCCACAGG
GCAAAATATGCCGATGTCAAGGTCAACTCTTCATACGTCGGAGGCGTAGTGCGTCAACATGGAACTTC
TCTTTCACCAGAGTTTTCGAGGCCGGTTCATGAAGTCCCTGGTTACCAACCCGAGACTGCCCTCAAGATC
TTTGAGCGCATCATGTTCAACAAGGATATTTCTACCGGTGAGATCGACATTGCTCAGAAAACCAGACTAC
GGTACCCTGGAAGTGAAGTCTACGTTCCATATCAAAAACGATATCCCTCCTTCGCCTGAGCCGACCTGC
TACCTCCTCAGTGTGACGGAACCTGTACCCCGGAGCAGCTTAATGCTATTAAGGATGGAAGTGCAGTT
GTTGAGAATACTATTATTAAGAGCCCTGCTGCGTGAAGGGGAACCCTCCACCAACCACGACCTCATCT
CCCACAGCAGCCCTACCGCTGGAAGTGCCATGCTAAAGGCTCCTGTGGCAATGCTAGCAATATCAGCT
CTCACTGTCTTGCTTTCTTCTTGTAG
    
```

Un ácido nucleico de ruCBPS1' descrito (SEQ ID NO: 17) codifica una proteína que tiene 652 restos de aminoácidos (SEQ ID NO: 18), que se presenta en la Tabla 6C usando el código de aminoácidos de una letra.

**Tabla 6C. Secuencia de la proteína codificada ruCBPS1' (SEQ ID NO:18).**

```

MRFAASIAVALPVIHAASAQGFPPPVKGVTVVKS KFDENVKI TYKENDI CETTQGVRSFTGHVHLPDND
DDFGVYRNY SINTFFWFFEAREDPKNAPLSIWLNGGPGSSSMIGLFQENGPCWVNEDSKSTTNNSFSWN
NKVNMLYIDQPNQVGFSDVPTNITYSTINDTISVADFSNGVPAQNLSTLVGTGSSQNPWATANNTVNA
ARSIWHFAQVWFQEFPEHKPNNNKISIWTESYGGRYGPSFASYFQEQNEKIKNHTITEEGEMHILNLDT
LGIINGCIDLMFQAESYAEFPYNNTYGIKAYTKEKRDAILHDIHRPDGCFDKVTKCREAAKEGDPHFYS
NNATVNTICADANSACDKYLMDFQETNLGYDIAHPLQDPFPPPFYKGFSLQSSVLSDMGSPVNFSQY
AQAVGKSFHGVGDYARPDVRGFTGDIAYLLESGVKVALVYGDRDYICNWFGEQVSLGLNYTGTQDFHR
AKYADVKNSSYVGGVVRQHGNSFTRVFEAGHEVPGYQPETALKIFERIMFNKDISTGEIDIAQKPDY
GTTGTESTFHIKNDIPPSPEPTCYLLSADGTCTPEQLNAIKDGTAVVENYIIKSPAASKGNPPPTTSS
PTAAPTAGSAMLKAPVAMLAISALTVLAFFL
    
```

5 La ruCBPS1' descrita tiene homología con las secuencias de aminoácidos mostradas en los datos BLAST listados en las Tablas 6D, 6E y 6F. Estos datos se analizaron por el programa POR PAREJAS BLAST.

Tabla 6D: Resultados TBLASTN para ruCBPS1'

Índice/Identificador Génico	Proteína/Organismo	Longitud (aa)	Identidad (%)	Positivos (%)	Esperado
gi/32410708	Cepa OR74A de <i>Neurospora crassa</i>	1947	246/632 (38%)	337/632 (53%)	e-104
gi/3046860	gen cpy1 de <i>Schizosaccharomyces pombe</i> para carboxipeptidasa Y	4308	137/480 (28%)	215/480 (44%)	1e-45
gi/18152938	Gen de <i>Pichia angusta</i> para carboxipeptidasa Y (CPY)	2214	139/508 (27%)	227/508 (44%)	2e-42

Tabla 6E: Resultados BLASTX para ruCBPS1'

Índice/Identificador Génico	Proteína/Organismo	Longitud (aa)	Identidad (%)	Positivos (%)	Esperado
gi/15004616	carboxipeptidasa S1/ <i>Aspergillus oryzae</i>	555	221/567 (38%)	310/567 (54%)	e-102
gi/435818	carboxipeptidasa S1, CPD-S1/ <i>Penicillium janthinellum</i>	423	174/499 (34%)	258/499 (51%)	4e-77
gi/995456	prepro-carboxipeptidasa Z/ <i>Absidia zychae</i>	460	155/491 (31%)	243/491 (49%)	2e-58
gi/19115337	carboxipeptidasa Y/ <i>Schizosaccharomyces pombe</i>	1002	137/480 (28%)	215/480 (44%)	1e-46
gi/4028158	precursor de la carboxipeptidasa Y; carboxipeptidasa vacuolar/ <i>Pichia angusta</i>	541	139/508 (27%)	226/508 (44%)	2e-43

Tabla 6F: Resultados BLASTP para ruCBPS1'

Índice/Identificador Génico	Proteína/Organismo	Longitud (aa)	Identidad (%)	Positivos (%)	Esperado
gi/15004616	carboxipeptidasa S1/ <i>Aspergillus oryzae</i>	555	222/567 (39%)	310/567 (54%)	7e-98
gi/435818	carboxipeptidasa S1, CPD-S1/ <i>Penicillium janthinellum</i>	423	174/499 (34%)	259/499 (51%)	1e-71
gi/995456	prepro-carboxipeptidasa Z/ <i>Absidia zychae</i>	460	156/491 (31%)	244/491 (49%)	2e-57
gi/19115337	carboxipeptidasa Y/ <i>Schizosaccharomyces pombe</i>	1002	137/480 (28%)	215/480 (44%)	4e-44

**ruPAP**

- 5 ruPAP es una prolilaminopeptidasa de *T. rubrum*. La secuencia de ADN genómico de un ácido nucleico de ruPAP de 1.795 nucleótidos (SEQ ID NO: 19) se muestra en la Tabla 7A.



**Tabla 7A. Secuencia de nucleótidos genómica de ruPAP (SEQ ID NO:19).**

ATGCAAGCAGCAAAATTGTTGAGCCGGTACTGGCAAAATGTACCTGGTTAGTGCAGCTAATCTTGAGTC  
 ACATCATGCATAGTTAACCGAGTATCACAAACAATCTACTATTGCGTTTTTGCTAATGGCTACCATAG  
 GAAGACTGAGGGTATCTGAGCTCCTTTTCGATGTCCCTTTAGACTACTCAAACCCGTCTTCCACTTCGC  
 TCCGGTTGTTTCGCCAGGAGTGTGCAGCGGCGAATTCAGGGTCTCTCTCGATGATAAAGACAGACAGC  
 TACCCTNGGATTGTTTTCCTGCAGGGTGGACCAGGAGGCTTGCCACAACCTCAGGAGGTAGGCTGG  
 GTTGGGCCATTGCTGGATCGAGGATTCCAGGTGAGTCTCCAGAATCGGGATGAGTAAGTGTAGAACC  
 TTGTTGAATTTCTTGATTAGATCCTTCTCCTTGACCAGCGAGGAACAGGGCTTCAACCCCTATAACCG  
 CTGCGACGCTTGCTCTTCAGGGAAACGCAGTAAAGCAAGCCGAATATCTTAGGCTATTCCGTGCCGATA  
 ATATCGTGCAGACTGTGAAGCAGTGCCTAAACTATTGACTGCTTATTACCTCCAGATAAGCAGAAAT  
 GGAGCGTCTTGCCAGAGTTTTGGAGGATTCTGTGCCGTACCGTATGTTTCTAAGTAGTGAGTAACTA  
 CTCCTTCAAATCCACCTGCTATAGATTGTCGTGCAAATCTAACCTTCATCATCTAGTCCAGGGACTT  
 AAAGAAGTCTTCACAACTGGTGGATTACCCCTCTTGTGTCAAAGCCTGATCCTGTGTACGAGAGGACC  
 TACGGTAAGTTGGGATAGATTGGGCTATTTTTAGTTTAAATACAGCTGACATCTACAGACAAGGTCCA  
 GTCCCGGAATAAAGTGTACTATTCCACTTTCCCGAAGACGAAGATCGAGTGCAGGATTATACTCAAGCA  
 TCTCCAACCCACGATGTTAAGCTCCCGATGGCTCACCGTAACTCCGGAACGCTTTCTCCAGCTAGG  
 AATTCATTTTGAATGAAAGGTACGCCATACTTCGCAGGTGACTTCTCGTAACCAATGACTAACATATG  
 CATATAGGGGGCATCGGCTTAGTTCATAGTATGATACCATCAATAACTTACATTATACTTATTCACTGA  
 CTAACAATGTGCAAATATCAGGCATAATTTGAAGTGCATTAATGAACTGGAATACTTTGGCTTCTCA  
 CACGACCTACTTTATCTCTGATTGAGAACGACACGAGTGCAGACAACGGCATTCTATATGCCATAATGC  
 ATGAATCTATCTACTGCCAAGGGTAAAACGTCTCTCCTGATCGAGTCAATATCAGAATCTAACGTGATA  
 CCGTAGGGAGGCCTCAAACCTGGGCTGCCGAAAGACTACTACCAAAGTTCTCTGGCTTCCGAGGCGCTCA  
 TAATCCTGATGGCATCTACTTCACTGGGGAGATGGTATACAAACTGGTTTGGAGTCCACAGAAGT  
 CGGCCAGCTCAAAGAGGTAGCCGATATTCTTGCTTCTACAATGACTGGCCGAGTTGTATGATAAGGA  
 ACAGCTCGCGCGCAACGAGGTGCCAGTGTATTCCGCTACATATGTGAGGATATGTACGTGCACCTCAG  
 CTACGCCAACGAAACAGCTGCCACTATTACAATTGCAAACAGTTCATACCAACACGATGTACCACAA  
 CGGACTGCGTTTCCAGATTCCGCTGAACTTATTGCGCAGCTGTTTGCTCTTTCGTGATGATACGATTGACTA  
 G

Un ácido nucleico de ruPAP de 1.326 (SEQ ID NO: 20) se muestra en la Tabla 7B. Un marco de lectura abierto (“ORF”) de ruPAP descrito empieza con un codón de inicio ATG en la posición 1 (subrayado en la Tabla 7B).

**Tabla 7B. Secuencia de nucleótidos de ruPAP (SEQ ID NO:20).**

ATGCAAGCAGCAAAATTGTTGAGCCGGTACTGGCAAAATGTACCTGGAAGACTGAGGGTATCTGAGCTC  
 CTTTTCGATGTCCCTTTAGACTACTCAAACCCGTCTTCCACTTCGCTCCGGTTGTTTCGCCAGGAGTGTG  
 CAGCGGCGAATTCAGGGTCTCTCTCGATGATAAAGACAGACAGCTACCCTGGATTGTTTTCCTGCAG  
 GGTGGACCAGGAGGAGCTTGCCACAACCTCAGGAGGTAGGCTGGGTTGGGCCATTGCTGGATCGAGGA  
 TTCCAGATCCTTCTCCTTGACCAGCGAGGAACAGGGCTTTCAACCCCTATAACCGCTGCGACGCTTGCT  
 CTTCAGGGAAACGCAGTAAAGCAAGCCGAATATCTTAGGCTATTCCGTGCCGATAATATCGTGCGAGAC  
 TGTGAAGCAGTGCCTAAACTATTGACTGCTTATTACCTCCAGATAAGCAGAAATGGAGCGTCTTGGC  
 CAGAGTTTTGGAGGATTCTGTGCCGTACCGTATGTTTCTAATCCTGAGGGACTTAAAGAAGTCTTCA  
 ACTGGTGGATTACCCCTCTTGTGTCAAAGCCTGATCCTGTGTACGAGAGGACCTACGACAAGGTCCAG  
 TCCCGGAATAAAGTGTACTATTCCACTTTCCCGAAGACGAAGATCGAGTGCAGGATTATACTCAAGCAT  
 CTCCAAACCCACGATGTTAAGCTCCCGATGGCTCACCGTAACTCCGGAACGCTTTCTCCAGCTAGGA  
 ATTCATTTTGAATGAAAGGCATAATTTGAAGTGCATTAATGAACTGGAATACTTTGGCTTCTCACA  
 CGACCTACTTTATCTCTGATTGAGAACGACACGAGTGCAGACAACGGCATTCTATATGCCATAATGCAT  
 GAATCTATCTACTGCCAAGGGGAGGCCTCAAACCTGGGCTGCCGAAAGACTACTACCAAAGTCTCTGGC  
 TTCCGAGGCGCTCATAATCCTGATGGCATCTACTTCACTGGGGAGATGGTATACAAACTGGTTTGGAG  
 TCGTCCACAGAAGTCCGGCCAGCTCAAAGAGGTAGCCGATATTCTTGCTTCTACAATGACTGGCCGAG  
 TTGTATGATAAGGAACAGCTCGCGCGCAACGAGGTGCCAGTGTATTCCGCTACATATGTGAGGATATG  
 TACGTGCACCTCAGCTACGCCAACGAAACAGCTGCCACTATTACAATTGCAAACAGTTCATACCAAC  
 ACGATGTACCACAACGGACTGCGTTCAGATTCCGCTGAACTTATTGCGCAGCTGTTTGCTCTTTCGTGAT  
 GATACGATTGACTAG

Un ácido nucleico de ruPAP descrito (SEQ ID NO: 20) codifica una proteína que tiene 441 restos de aminoácidos (SEQ ID NO: 21), que se presenta en la Tabla 7C usando el código de aminoácidos de una letra.

<b>Tabla 7C. Secuencia de la proteína codificada ruPAP (SEQ ID NO:21).</b>
<p><b>MQAAKLLSRYWQNPGRRLRVSELLFDVPLDYSNPSSTSLRRLFARSVQRRIPGSSLDDKDRQLPWIVFLQ GGPGGACQPQEVGWVGPLLDGRGFQILLDDQRGTGLSTPITAATLALQGNVQAEYLRLFRADNIVRD CEAVRKLTLTAYYPPDKQKWSVLGQSFGGFCAVTVSNPEGLKEVFTTGGPLPPLVSKPDPVYERTYDKVQ SRNKVYYSTFPEDEDVRVRIILKHLQTHDVKLPDGSPLTPERFLQLGIHFGMKGIILKGINELEYFGFLT RPTLSLIENDTSADNGILYAIMHESIYCQGEASNWAAERLLPKFSGFRGAHNPDIYFTGEMVYKHWFE SSTELGQLKEVADILASYNWPLDYDKEQLARNEVPVYSATYVEDMYVHFSYANETAATIHNCQPFITN TMYHNGLRSDSAELIAQLFALRDDTID</b></p>

5 La ruPAP descrita tiene homología con las secuencias de aminoácidos mostradas en los datos BLAST listados en las Tablas 7D, 7E y 7F. Estos datos se analizaron por el programa POR PAREJAS BLAST.

**Tabla 7D: Resultados TBLASTN para ruPAP**

Índice/Identificador Génico	Proteína/Organismo	Longitud (aa)	Identidad (%)	Positivos (%)	Esperado
gi/14329656	Gen papA de <i>Aspergillus niger</i> para proil aminopeptidasa A	3752	151/307 (49%)	190/307 (61%)	e-118
gi/32414442	Cepa OR74A de <i>Neurospora crassa</i>	1449	212/477 (44%)	285/477 (59%)	e-100
gi/604877	Gen de <i>Aeromonas sobria</i> para proil aminopeptidasa	1740	175/420 (41%)	239/420 (56%)	4e-77

**Tabla 7E: Resultados BLASTX para ruPAP**

Índice/Identificador Génico	Proteína/Organismo	Longitud (aa)	Identidad (%)	Positivos (%)	Esperado
gi/18307408	proil aminopeptidasa A/ <i>Aspergillus niger</i>	442	266/442 (60%)	334/442 (75%)	e-152
gi/14456054	proil aminopeptidasa putativa/ <i>Aspergillus nidulans</i>	365	211/366 (57%)	263/366 (71%)	e-114
gi/22507295	proil aminopeptidasa/ <i>Talaromyces emersonii</i>	300	181/301 (60%)	226/301 (75%)	4e-99
gi/1236731	proil aminopeptidasa/ <i>Aeromonas sobria</i>	425	175/420 (41%)	239/420 (56%)	4e-78

**Tabla 7F: Resultados BLASTP para ruPAP**

Índice/Identificador Génico	Proteína/Organismo	Longitud (aa)	Identidad (%)	Positivos (%)	Esperado
gi/18307408	proil aminopeptidasa A/ <i>Aspergillus niger</i>	442	267/443 (60%)	336/443 (75%)	e-157
gi/14456054	proil aminopeptidasa putativa/ <i>Aspergillus nidulans</i>	365	211/366 (57%)	263/366 (71%)	e-116

Índice/Identificador Génico	Proteína/Organismo	Longitud (aa)	Identidad (%)	Positivos (%)	Esperado
gi/22507295	prolil aminopeptidasa/ <i>Talaromyces emersonii</i>	300	181/301 (60%)	226/301 (75%)	e-102
gi/1236731	prolil aminopeptidasa/ <i>Aeromonas sobria</i>	425	175/420 (41%)	239/420 (56%)	2e-78

**ruAMPP**

ruAMPP es una aminopeptidasa P de *T. rubrum*. La secuencia de ADN genómico de un ácido nucleico de ruAMPP de 2.418 nucleótidos (SEQ ID NO: 22) se muestra en la Tabla 8A.

<b>Tabla 8A. Secuencia de nucleótidos genómica de ruAMPP (SEQ ID NO:22).</b>
<p>ATGCCGCCACCACCGGTTGACACGACCCAGCGTCTCGCAAAGCTGCGAGAGCTGATGGCTCAGAACAAG                      GTCGATGTATATAGTATGCAATTCAGATACACCATTAAAGCTCCCTTGATAATAACAGTCGTATACTCA                      TTCTTCTTTCTTCTACTCCTCGCCTTAAAGTTGTGCTTCGGAAGACAGCCATCAGTCGGAGTACATTG                      CTCCATGTGATGGGCGTCGAGGTTAGACCTGTCCCTCCATAAAAGAATACCTACCCGTAATACCAGCCG                      GCAGACGCTCATACTGATCACTGCAGCTTTCATATCCAGCTTCACCTGGCTCGGCAGGATGTGCCATCGT                      CTCTATGAGTAAAGCTGCTCTGTCTACAGACGGCAGATACTTCAGCCAAGCTGCAAAAACAGCTCGATGC                      CAACTGGATCCTGTTGAAGCGAGGTGTGAGGGTGTCCCAACCCTGGGAAGAATGGTATATCTGCCCTG                      GTATCGACTTTTCCGGTATAATGGTTGACAGGCTGGATATAGGACCGCTGAGCAGGCCGAGACACGGCA                      AGGTTGTGGGTGTTGACCCGTCACCTATTACGGCAGGTGAGAATCTACAGTATGCGTCTCTTACAAGTG                      TCATCGTACTAACTGTATGTTATAGCGGATGCACGAAAGCTTCTCAGACGTTGAAGACCACCGGAGG                      CTCCTTGGTTGGAATTGATCAGAACCCTGATTGATGCCGTCTGGGGAGATGAACGTCCTGCACGGCCTGC                      CAACCAAATTACGGTACAGCCTGTTGAGCGCGCGGGAAAGTCATTCGAGGAGAAAGTGAAGACCTGCG                      AAAGGAATTGACTGCGAAGAAGAGGTCTGCTATGTTTATTTGTATGACGCTAGATCTATTTTGTATCAA                      ACATATACTAACAAACGCAATATAGCCACCTTGATGAGATTGCATGGCTCTTCAACCTCCGTGGAAGC                      GAGTAAGTTTCTATATAAATGGTATCTTTCACTTATACAAAAGCCATGCTGACTGGTGTAGTATTCC                      ATATAACCCCGTCTTTTCTCGTACGCAATTGTGACGCCCTCAGTTGCGGAACCTATGTGATGAGAG                      CAAGCTGTCTCCAGAAGCCAGAAAACATCTCGAAGGCAAGGTCGTTCTCAAGCCATACGAGTCCATCTT                      CCAAGCTTCAAAGTCCTCGCCGAATCAAAGGCATCGGCTAGCAGCGGTTCTCTGCGGAAGTCTTGTGTT                      GTCTAACAAGGCTTCGTGGTCTTTGAGCCTCGCCCTCGGTGGGGAACAGAAGCTCGTTGAGGTTCGAAG                      TCCCATCACTGACGCCAAAGCCATCAAGAACGAAGTTGAACTGGAAGGATTAGAAAATGCCATATCCG                      AGACGGTGCAGCTCTGATCGAGTACTTCGCCTGGCTTGAAAATGCATTGATCAAAGAAGGTGCCAAGCT                      AGACGAAGTAGATGGAGCCGACAACTCTTCGAGATCCGCAAGAAATATGACCTCTTCGTCCGCCAATC                      CTTTCGACACCATCTCTTACCGGTGCTAACGGTGTACCATTCATTACAAACCCGAGAAGTCAACTTG                      CGCTATCATTGACCCGAAGGCTATGTACCTGTGTGACTCTGGTGGCCAATACCTTGATGGTACTACTGA                      TACTACCCGAACCTCCACTTTGGAGAGCCCACGGAGTTCAGAAGAAGGCTTATGCACTTGTCTAAA                      GGGACATATCAGCATTGACAATGCCATTTTCCCAAAGGAACCACCGGATACGCCATTGACTCGTTTGC                      TCGACAGCATTGTGGAAGGAGGCTGATTACCTCCACGGCACCAGTTCATGGTGTGGCTCATTTTT                      GGTACGGGTTTTCTTTTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTATTATTACTTCTCTTAGGCTAACAC                      ATTCTCTCTAAGAACGTCCATGAGGGACCTATGGGCATAGGAAGCCGTGCTCAGTACGCTGAAGTCCCT                      CTCTCTGCCAGCAATGTTCTTTCCAACGGTAGGATTTCTGCATCTCATCTTTCTTGAATCCTACTAATT                      GCAAAATAGAGCCTGGATATTATGAAGACGGCAACTTCGGCATTTCGTCTCGAGAGTAAGTTCAATGACT                      GCGTATTCTAGTTTTTTTTCATACTGACGGCCTCTTTAGACCTCGTAATCTGCAAGGAGGTCCAGACTGCA                      CACAAATTCGGCGACAAGCCCTTCTCGGATTTGAGTCCATCACCTGGTACCTTTCTGCCAAAAACTC                      CTTGATGCTTCTCTCTTGACCGAAGCTGAGAGAAAGTGGGTGAATGATTACCATGCGAAAGTCTGGGAG                      AAGACCAGTCCCTTCTTTGAGAAGGACGAGTTAACAACCGCCTGGCTAAAGCGCGAGACACAACCTATT                      TAA</p>

5 Un ácido nucleico de ruAMPP de 1.878 (SEQ ID NO: 23) se muestra en la Tabla 8B. Un marco de lectura abierto ("ORF") de ruAMPP descrito empieza con un codón de inicio ATG en la posición 1 (subrayado en la Tabla 8B).

**Tabla 8B. Secuencia de nucleótidos de ruAMPP (SEQ ID NO:23).**

ATGCCGCCACCACCGGTTGACACGACCCAGCGTCTCGCAAAGCTGCGAGAGCTGATGGCTCAGAACAAG  
 GTCGATGTATATATTGTGCCTTCGGAAGACAGCCATCAGTCGGAGTACATTGCTCCATGTGATGGGCGT  
 CGAGCTTTCATATCCAGCTTCACTGGCTCGGCAGGATGTGCCATCGTCTCTATGAGTAAAGCTGCTCTG  
 TCTACAGACGGCAGATACTTCAAGCAAGCTGCAAAAACAGCTCGATGCCAACTGGATCCTGTTGAAGCGA  
 GGTGTCGAGGGTGTCCCAACCTGGGAAGAATGGACCGCTGAGCAGGCCGAGACACGGCAAGGTTGTGGG  
 TCGGATGCACGAAAGCTTCTCAGACGTTGAAGACCACCGGAGGCTCCTTGGTTGGAATTGATCAGAAC  
 CTGATTGATGCCGTCTGGGGAGATGAACGTCCTGCACGGCCTGCCAACCAAATTACGGTACAGCCTGTT  
 GAGCGCGCGGAAAGTCAATTCGAGGAGAAAGTGAAGACCTGCGAAAGGAATTGACTGCGAAGAAGAGG  
 TCTGCTATGGTTATTTTCGAGTAAGTTTCTATATAAATGGTATCTTTCACTTTATACAAAAAGCCATGCT  
 GACTGGTGTAGTATTCCATATAACCCCGTCTTTTTCTCGTACGCAATTGTGACGCCCTCAGTTGCGGAA  
 CTCTATGTCTGAGAGCAAGCTGTCTCCAGAAGCCAGAAAACATCTCGAAGGCAAGGTCGTTCTCAAG  
 CCATAGCTCCATCTTCCAAGCTTCCAAAGTCTCGCCGAATCAAAGGCATCGGCTAGCAGCGGTTCC  
 TCTGGGAAGTCTTGTGTCTAACAAGGCTTCGTGGTCTTTGAGCCTCGCCCTCGGTGGGGAACAGAAC  
 GTCGTTGAGGTTGCAAGTCCCATCACTGACGCCAAAGCCATCAAGAACGAAGTTGAACTGGAAGGATTC  
 AGAAAATGCCATATCCGAGACGGTGCAGCTCTGATCGAGTACTTCGCCTGGCTTGAAAATGCATTGATC  
 AAAGAAGGTGCCAAGCTAGACGAAGTAGATGGAGCCGACAACTCTTCGAGATCCGCAAGAAATATGAC  
 CTCTTCGTCGGCAACTCCTTCGACACCATCTCTTCTACCGGTGCTAACGGTGTACTACCATTATTACAAA  
 CCCGAGAAGTCAACTTGCCTATCATTGACCCGAAGGCTATGTACCTGTGTGACTCTGGTGGCCAATAC  
 CTTGATGGTACTACTGATACTACCCGAACTCTCCACTTTGGAGAGCCACGGAGTTCAGAAAGGCT  
 TATGCACTTGTCTAAAGGGACATATCAGCATTGACAATGCCATTTTCCCAAAGGAACCACCGGATAC  
 GCCATTGACTCGTTTTGCTCGACAGCATTGTGGAAGGAGGGTCTGGATTACCTCCACGGCACCAGGTCAT  
 GGTGTTGGCTCATTGAAAGTCCATGAGGGACCTATGGGCATAGGAAGCCGTGCTCAGTACGCTGAA  
 GTTCTCTCTCGCCAGCAATAGCCTGGATATTATGAAGACGGCAACTTCGGCATTCGCTCGAGAGTA  
 AGTTCATGACTGCGTATTCTAGTTTTTTCATACTGACGGCCTCTTTAGACCTCGTAATCTGCAAGGAG  
 GTCCAGACTGCACACAAATTCGCGCACAAGCCCTTCTCGGATTTGAGTCCATCACCTGGTACCTTTC  
 TGCCAAAACCTCCTTGATGCTTCTCTCTTGACCGAAGCTGAGAGAAAGTGGGTGAATGATTACCATGCG  
 AAAGTCTGGGAGAAGACCAGTCCCTTCTTTGAGAAGGACGAGTTAACAACCGCCTGGCTAAAGCGCGAG  
 ACACAACCTATTAA

Un ácido nucleico de ruAMPP descrito (SEQ ID NO: 23) codifica una proteína que tiene 625 restos de aminoácidos (SEQ ID NO: 24), que se presenta en la Tabla 8C usando el código de aminoácidos de una letra.

**Tabla 8C. Secuencia de la proteína codificada ruAMPP (SEQ ID NO:24).**

MPPPPVDTTQRLAKLRELMAQNKVDVYIVPSEDSHQSEYIAPCDGRRAFISSFTGSAGCAIVSMSKAAL  
 STDGRYFSQAAKQLDANWILLKRGVEGVPTWEEWTAEQAE TRQCGSDARKLSQTLKTTGGSLVGLDQN  
 LIDAVWGDERPARPANQITVQPVERAGKSFEEKVEDLRKELTAKKRSAMVISSKFLYKWYLSLYTKSHA  
 DWCSIPYNPVFFSYAIVTPSVAELYVDESKLSPEARKHLEGKVVLPYESIFQASKVLAESKASASSGS  
 SGKFLLSNKASWSLSLALGGEQNVVEVRSPITDAKAIKNEVELEGFRKCHIRDGAALIEYFAWLENALI  
 KEGAKLDEVDGADKLF EIRKKYDLFVGNSFDTISSTGANGATIHYKPEKSTCAIIDPKAMYLCDSSGQY  
 LDGTTDTTRTLHFGEPTFQKKAYALVLKGHISIDNAIFPKGTTGYAIDSFARQHLWKEGLDYLHGTGH  
 GVGSFLNVHEGPMGIGSRAQYAEVPLSASNLDIMKTATSAFVSRVSSMTAYSSFFILTASLDLVICKE  
 VQTAHKFGDKPFLGFESITLVPFCQKLLDASLLTEAERKWVNDYHAKVWEKTSPPFEKDELTTAWLKRE  
 TQPI

5 La ruAMPP descrita tiene homología con las secuencias de aminoácidos mostradas en los datos BLAST listados en las Tablas 8D, 8E y 8F. Estos datos se analizaron por el programa POR PAREJAS BLAST.

Tabla 8D: Resultados TBLASTN para ruAMPP

Índice/Identificador Génico	Proteína/Organismo	Longitud (aa)	Identidad (%)	Positivos (%)	Esperado
gi/32403169	Cepa OR74A de <i>Neurospora crassa</i>	1845	339/630 (53%)	433/630 (68%)	0.0
gi/20453016	Gen de aminopeptidasa P de <i>Drosophila melanogaster</i>	12647	268/638 (42%)	369/638 (57%)	e-127
gi/17571207	<i>Drosophila melanogaster</i> (ApepP) en el cromosoma 2	12001	268/638 (42%)	369/638 (57%)	e-127
gi/4583560	Gen Daminopep-p de <i>Drosophila melanogaster</i>	2358	268/638 (42%)	369/638 (57%)	e-127

Tabla 8E: Resultados BLASTX para ruAMPP

Índice/Identificador Génico	Proteína/Organismo	Longitud (aa)	Identidad (%)	Positivos (%)	Esperado
gi/25529603	X-Pro aminopeptidasa, forma citosólica/ <i>Drosophila melanogaster</i>	613	268/638 (42%)	369/638 (57%)	e-127
gi/4107172	aminopeptidasa P/ <i>Drosophila melanogaster</i>	613	258/638 (40%)	369/638 (57%)	e-124
gi/15384991	Xaa-Pro aminopeptidasa 2/ <i>Lycopersicon esculentum</i>	654	268/674 (39%)	365/674 (54%)	e-120
gi/8489879	aminopeptidasa P citosólica/ <i>Homo sapiens</i>	623	254/646 (39%)	358/646 (55%)	e-119
gi/2584787	semejante a aminopeptidasa P/ <i>Homo sapiens</i>	623	254/646 (39%)	357/646 (55%)	e-119

Tabla 8F: Resultados BLASTP para ruAMPP

Índice/Identificador Génico	Proteína/Organismo	Longitud (aa)	Identidad (%)	Positivos (%)	Esperado
gi/30923284	Peptidasa probable C22G7.01c	598	291/629 (46%)	384/629 (61%)	e-156
gi/25529603	X-Pro aminopeptidasa, forma citosólica/ <i>Drosophila melanogaster</i>	613	268/638 (42%)	369/638 (57%)	e-124
gi/15384991	Xaa-Pro aminopeptidasa 2/ <i>Lycopersicon esculentum</i>	654	268/674 (39%)	365/674 (54%)	e-123
gi/8489879	aminopeptidasa P citosólica/ <i>Homo sapiens</i>	623	254/646 (39%)	358/646 (55%)	e-122
gi/2584787	semejante a aminopeptidasa P/ <i>Homo sapiens</i>	623	254/646 (39%)	357/646 (55%)	e-122
gi/4107172	aminopeptidasa P/ <i>Drosophila melanogaster</i>	613	258/638 (40%)	369/638 (57%)	e-121

Índice/Identificador Génico	Proteína/Organismo	Longitud (aa)	Identidad (%)	Positivos (%)	Esperado
gi/18777778	aminopeptidasa citoplásmica/ <i>Rattus norvegicus</i> P	623	253/645 (39%)	353/645 (54%)	e-120
gi/18875372	aminopeptidasa citosólica/ <i>Mus musculus</i> P	623	250/645 (38%)	354/645 (54%)	e-118
gi/15384989	Xaa-Pro aminopeptidasa 1/ <i>Lycopersicon esculentum</i>	655	264/674 (39%)	361/674 (53%)	e-117

**ruPLD**

ruPLD es una prolidasa de *T. rubrum*. La secuencia de ADN genómico de un ácido nucleico de ruPLD de 2.344 nucleótidos (SEQ ID NO: 25) se muestra en la Tabla 9A.

<b>Tabla 9A. Secuencia de nucleótidos genómica de ruPLD (SEQ D NO:25).</b>
ATCAACCTCACCTCTTCACCGTCTCACGCCCTTCGTCCCGTCCAACCTCTTCATTTGCCCCTCTCTATGA TAACCAACAAACATCCGCTGTTATGTAATCGAACCCGCGGTTAGCCATCCCTAGCCCCGCGTTTTCTCC CAGCATCAATACGACCGAAATGAAGACAGACGGGGAAGACGAGGCAAAACAATAACACATCAACAATTT AACCCGTTGCCGTCTTCTACCCATCTTGTCTACGCATCGTCCAACCTTTCTTGCCCTATATCAGCCGA ACTCGGCCATCATGGATATCCACGTCGACAAATACCCGGCTAAGAGTACGCCAGGCGCGTCGCCGAGA AGCTCAAGGCCGCGGGGCACGGCTCTACCGGCATCATCTTCGTTCGAAGGCCAAAAGGAGCATATTATCG ATGATAGCGACGAGCCGTTTCACCTCCGGTGAGCCGTGGGAATACACTCGACTGGGCGGAATAAGCTAA CAAAAGGGTGTGATAGTCAACGCCGAAACTTCCTCTATCTGTCCGGCTGTCTTGAGGCCGAGTGCTCCG TTGCATACAACATCGAGAAAGATGAGCTTACATTGTTTCATTCCACCAGTCGACCCAGCCTCGGTTATGT GGTCCGGCCTCCCTCTTGAGCCCGCCGAAGCCTTGAAGCAGTTCGATGTTGATGCCGTGCTCCTCACAA CTGAGATAAACAACATCTCGCGAAGTGTGGGGGCGAGAAGGTCTTACCATTGCAGACAGAGTTTGCC CGGAGGTCTCCTTCTCATCCTTCAAGCACAACGACACCCGATGCCCTGAAGCTTGCCATCGAGTCTGCG GTATAGTGAAAGACGAGTATGAAATTGGTCTTCTCCGACGTGCTAATGAGGTCTCCAGCCAAGCTCATA TTGAAGTGATGAAAGCCGCAACCAAGTCAAAGAACGAGAGAGAGCTCTATGCTACTCTCAACTATGTCT GCATGTCTAATGGCTGCTCCGACCAGTCTTACCATCCAATTCTTGATGTGGCCCCAATGCTGCCACTC TCCACTACACCAAGAACAACGGTGACCTAACTAACCCGGCTACCGGGATTAAGGACCAGCTCGTACTTA TCGACGCTGGATGCCAGTACAAGGCGTACTGTGCAGATATCACTCGTGCATTTCCCCTTGTCGGGCAAA TCACCACGGAGGGCCGCGAGATCTATGATATTGCCTTGGAGATGCAGAAAGTCGCGTTTGGCATGATCA AACCTAATGTTTTGTTTCGACGACATGCATGCTGCGGTCCACCGGTTGCGATCAAGGGGCTGCTCAAGA TTGGCATTCTACTGGCTCTGAGGATGAGATTTTCGATAAGGGAATCAGCACTGCCTTTTTCCCACATG GTCTAGGCCACCATCTCGGCATGGACACTCAGATGTTGGAGGAAACCCTAACCCGGCTGACCCGAATC GCATGTTTTAAATACTTGCCTCTGCGAGGCACTGTTCCAGAGGGATCCGTCATTACAATTGAGCCCGGTG TAAGTGTGAATCGAGTAGTGTCTCCGCCGAATGTTTCACATAACATTTACTAACCTTGTCTTAGGTCT ACTTCTGCCGTTACATCATTTGAGCCATTCTTACTAACCCCGAGACCAGCAAGTACATCAACTCCGAAG TTCTAGACAAGTACTGGGCTGTTGGAGGTGTACGTATCGAGGACAACGTCGTCGTCGCCGCCAATGGCT TTGAGAACCCTGACCACGGTGCCAAAGGAGCCCAGGAGGTGCAACGCATTGTCCAGGAGGGTGCTAAAT AATTATGTTTTTATTTCAGTACACCGAGTGGTTCGGACACACGCAGGAGCATGTACATATTTATGATCTAC CCAGTTGATTTGCTACCAAAAAAGAACCACCACAGCCCTATTTATTGATATTACATAGTAGGAATAAA GGCCACTTTGCCACCGCGAATAATAACAATAAGAAAAGCAACTACTCGTACAACCAGCCTAGAAAGCT CTAGACCTCTTTCTCGCTGGGCCCTTGAATGCCGGGCTACTGGTGTATCACGCTCCCTGGCCCTCTTC TCCTTCATGTCCAACACCCGATTAAGCAATCGAAACTGAACTGGGGATGCTCAAGACACAATGCCTTG AACTGCTCTTACGATCATGACGCAGCACATCACTCATCTTAGCCAGAAGCGAGCAACCCGGTCTCTG ATAGCAGTGTCTCCGGCGTGGTATGGCTGTACACGTATCTCGCATACTCGATCTCACCCGTAGCACTA CTCTCGATGCTACCAATCTTGTCTGAGCAAGCAGTTTGAGTTTTTTCGTTTTCCGAGCTTTTCCGCCA

5 Un ácido nucleico de ruPLD de 1.401 (SEQ ID NO: 26) se muestra en la Tabla 9B. Un marco de lectura abierto ("ORF") parcial de ruPLD descrito se obtuvo según se consideró por la ausencia de un codón de inicio ATG en la posición 1.

**Tabla 9B. Secuencia de nucleótidos de ruPLD (SEQ ID NO:26).**

```

CCGAACTCGGCCATCATGGATATCCACGTCGACAAATACCCGGCTAAGAGTCACGCCAGGCGCGTCGCC
GAGAAGCTCAAGGCCGCGGGGCACGGCTCTACCGGCATCATCTTCGTCGAAGGCCAAAAGGAGCATATT
ATCGATGATAGCGACGAGCCGTTTCACTTCCGTCAACGCCGAAACTTCCTCTATCTGTCCGGCTGTCTT
GAGGCCGAGTGTCCGTTGCATACAACATCGAGAAAGATGAGCTTACATTGTTTATTCCACCAGTCGAC
CCAGCCTCGGTTATGTGGTCCGGCTCCCTCTTGAGCCCGCCGAAGCCTTGAAGCAGTTCGATGTTGAT
GCCGTGCTCCTCACAACCTGAGATAAACTATCTCGCGAAGTGTGGGGGCGAGAAGGTCTTCACCATT
GCAGACAGAGTTTGCCCGGAGGTCTCCTTCTCATCCTTCAAGCACAACGACACCGATGCCCTGAAGCTT
GCCATCGAGTCTCGCGTATAGTGAAAGACGAGTATGAAATTGGTCTTCTCCGACGTGCTAATGAGGTC
TCCAGCCAAGCTCATATTGAAGTGATGAAAGCCGCAACCAAGTCAAAGAACGAGAGAGAGCTCTATGCT
ACTCTCAACTATGTCTGCATGTCTAATGGCTGCTCCGACCAGTCTTACCATCCAATTCTTGCATGTGGC
CCCAATGCTGCCACTCTCCACTACACCAAGAACAACGGTGACCTAACTAACCCGGCTACCGGGATTAAG
GACCAGTCGTAATTATCGACGCTGGATGCCAGTACAAGGCGTACTGTGCAGATATCACTCGTGCATTC
CCCTTGCCGGCAAATTCACCACGGAGGGCCGCGCAGATCTATGATATTGCCCTTGGAGATGCAGAAAGTC
GCGTTTGCCATGATCAAACCTAATGTTTGTTCGACGACATGCATGCTGCCGTCACCGGGTTGCGATC
AAGGGGCTGCTCAAGATTGGCATTCTCACTGGCTCTGAGGATGAGATTTTCGATAAGGGAATCAGCACT
GCCTTTTTCCACATGGTCTAGGCCACCATCTCGGCATGGACACTCACGATGTTGGAGGAAACCTAAC
CCGGCTGACCCGAATCGCATGTTTAAATACTGCGTCTGCGAGGCACTGTTCCAGAGGGATCCGTCATT
ACAATTGAGCCCGGTGCTACTTCTGCCGTTACATCATTGAGCCATTCTTACTAACCCCGAGACCAGC
AAGTACATCAACTCCGAAGTCTAGACAAGTACTGGGCTGTTGGAGGTGTACGTATCGAGGACAACGTC
GTCGTCGCGCCAATGGCTTTGAGAACCTGACCACGGTGCCAAAGGAGCCCGAGGAGTTCGAACGCATT
GTCCAGGAGGGTGCTAAATAA
    
```

Un ácido nucleico de ruPLD parcial descrito (SEQ ID NO: 26) codifica una proteína con una secuencia parcial que tiene 466 restos de aminoácidos (SEQ ID NO: 27), que se presenta en la Tabla 9C usando el código de aminoácidos de una letra.

**Tabla 9C. Secuencia de la proteína codificada ruPLD (SEQ ID NO:27).**

```

PNSAIMDIHVDKYPKSHARRVAEKLKAAGHGSTGIIFVEGQKEHIIDDSDEPFHFRQRRNFLYLSGCL
EAECVAVNIEKDELTLFIPPVDPASVMWSGLPLEPAEALKQFDVDAVLLTTEINNYLAKCGGEKVFTI
ADRVCPEVSFSSFKHNDTDALKLAIESCRIVKDEYEIGLLRRANEVSSQAHIEVMKAATKSKNERELYA
TLNYVCMNSGSDQSYHPILACGPNAATLHYTKNNGDLTNPATGIKDQLVLIDAGCQYKAYCADITRAF
PLSGKFTTEGRQIYDIALEMQKVAFGMIKPNVLFDDMHAHVHVAIKGLLKIGILTSEDEIFDKGIST
AFFPHGLGHHLGMDTHDVGNNPNPADPNRMFKYLRLRGTVPESVITIEPGVYFCRYIIEPFLTNPETS
KYINSEVLDKYWAVGGVRIEDNVVVRANGFENLTTVPKEPEEVERIVQEGAK
    
```

5

La ruPLD parcial descrita tiene homología con las secuencias de aminoácidos mostradas en los datos BLAST listados en las Tablas 9D, 9E y 9F. Estos datos se analizaron por el programa POR PAREJAS BLAST.

**Tabla 9D: Resultados TBLASTN para ruPLD**

Índice/Identificador Génico	Proteína/Organismo	Longitud (aa)	Identidad (%)	Positivos (%)	Esperado
gi/14272360	Gen pepP para prolidasa de <i>Aspergillus nidulans</i> , exones 1-3	2632	199/348 (57%)	249/348 (71%)	e-143
gi/32420910	Cepa OR74A de <i>Neurospora crassa</i>	2562	235/457 (51%)	324/457 (70%)	e-136
gi/3114965	ARNm para prolidasa, forma 1, de <i>Suberites domuncula</i>	1688	157/464 (33%)	235/464 (50%)	4e-66

Índice/Identificador Génico	Proteína/Organismo	Longitud (aa)	Identidad (%)	Positivos (%)	Esperado
gi/22531161	Proteína semejante a X-Pro dipeptidasa de <i>Arabidopsis thaliana</i>	1672	160/477 (33%)	242/477 (50%)	2e-64

Tabla 9E: Resultados BLASTX para ruPLD

Índice/Identificador Génico	Proteína/Organismo	Longitud (aa)	Identidad (%)	Positivos (%)	Esperado
gi/14272361	prolidasa/ <i>Emericella nidulans</i>	496	267/463 (57%)	336/463 (72%)	e-153
gi/3114966	prolidasa/ <i>Suberites domuncula</i>	501	157/464 (33%)	235/464 (50%)	1e-66
gi/22531162	proteína semejante a X-Pro dipeptidasa/ <i>Arabidopsis thaliana</i>	486	160/477 (33%)	242/477 (50%)	6e-65
gi/30582223	peptidasa D/ <i>Homo sapiens</i>	493	152/452 (33%)	231/452 (51%)	2e-63
gi/20271451	peptidasa D/ <i>Homo sapiens</i>	493	152/452 (33%)	230/452 (50%)	3e-63

Tabla 9F: Resultados BLASTP para ruPLD

Índice/Identificador Génico	Proteína/Organismo	Longitud (aa)	Identidad (%)	Positivos (%)	Esperado
gi/14272361	prolidasa/ <i>Emericella nidulans</i>	496	267/463 (57%)	336/463 (72%)	e-158
gi/3114966	prolidasa/ <i>Suberites domuncula</i>	501	158/466 (33%)	235/466 (50%)	6e-67
gi/22531162	proteína semejante a X-Pro dipeptidasa/ <i>Arabidopsis thaliana</i>	486	159/477 (33%)	241/477 (50%)	6e-64
gi/30584879	peptidasa D de <i>Homo sapiens</i>	494	152/452 (33%)	231/452 (51%)	2e-63
gi/15929143	peptidasa D/ <i>Homo sapiens</i>	493	152/452 (33%)	231/452 (51%)	2e-63
gi/20271451	peptidasa D/ <i>Homo sapiens</i>	493	152/452 (33%)	230/452 (50%)	4e-63

**caLAP2**

caLAP2 es una leucina aminopeptidasa de *Microsporium canis*. Un ácido nucleico de caLAP2 de 1.730 nucleótidos (SEQ ID NO: 28) se muestra en la Tabla 10A.



Tabla 10A. Secuencia de nucleótidos genómica de caLAP2 (SEQ ID NO:28).
<p>                     ATGAAGACACAGTTGTTGAGTCTGGGAGTTGCCCTCACGGCCATCTCTCAGGGCGTTATTGCTGAGGAT                      GCCTTGAAGTGGCCATTCAAGCCGTTGGTTAATGCTGTGAGTATATACACAAGATCGATCGATCGTCCT                      CTTGTCCCTGTCACTTATCGCTCTACAGTAAGCAAAAATACTGGAGAATCATGTGCTGATGTAAATGTA                      TAGGATGACCTGCAAAACAAGATTAAGCTCAAGGATCTTATGGCTGGCGTACAGAACTCCAAGACTTC                      GCCTACGCTCACCCCTGAGAAGAATCGAGTATTCCGGTGGTGCTGGCCACAAGGATACCGTCGACTGGATC                      TACAATGAGCTCAAGGCTACCGGCTACTACGATGTGAAGATGCAGCCACAAGTCCACCTGTGGTCTCAT                      GCTGAGGCAGCTGTCAATGCCAATGGCAAGGATCTCACTGCCAGTGCCATGTCTACAGCCCTCCAGCC                      GACAAGATCACTGCCGAGCTTGTCTTGGCCAAGAACATGGGATGCAATGCTGTATGTGCGCCCTTTTC                      CATTCTATATACGACTGGTTCGTTGGAAATTCAGAAGAGCTGACAATTGCAAACAGACTGATTACCCA                      GAGGGTACCAAGGGCAAGATTGTCCTCATCGAGCGTGGTGTCTGCAGCTTTGGCGAGAAGTCCGCTCAG                      GCTGGCGATGCAAAGGCTATTGGTGCCATCGTCTACAACAACGTCCTGGAAGCTTGGCCGGCACCCCTG                      GGTGGCCCTTGACAACCGCCATGCTCCAACCTGCTGGAATCTCTCAGGCTGATGGAAAGAACCTCGCTAGC                      CTTGTGCGCTCTGGCAAGGTTACCGTCACCATGAACGTTATCAGCAAGTTTGAGAACAGGACTACGTGA                      GTATTGTTCCATACTTTGGTCAACAATGATATATACACGTAACAACACTGCTCTATAGCTGGAACGTCA                      TTGCCGAGACCAAGGGAGGAGACCACAACAACGTCATCATGCTCGGTTCTCACTCTGACTCTGTGCGACG                      CCGGCCCTGGTATCAACGACAACGGCTCCGGTACCATTGGTATCATGACCGTTGCCAAAGCCCTCACCA                      ACTTCAAGGTCAACAACGCCGTCGGCTTCGGCTGGTGGACCGCCGAGGAGTTCGGCCTTCTCGGCAGCA                      CTTTCTACGTCGACAGCCTTGACGACCGTGAAGTGCACAAGGTCAAGCTGTACCTCAACTTCGACATGA                      TTGGCTCCCCCAACTTCGCCAACCCAGATCTACGACGGAGACGGCTCCGCCTACAACATGACTGGCCCCG                      CCGGATCTGCTGAAATCGAGTACCTGTTGAGAAAGTTCCTTCGATGACCAGGGAATCCCACACCAGCCCA                      CCGCCTTACCGGCCGCTCCGACTACTCTGCCTTCATCAAGCGCAACGTCCTGCCGGAGGTCTGTTTA                      CTGGTGCTGAGGTCGTCAAGACCGCCGAGCAGGCTAAGCTATTTGGCGGCGAGGCTGGCGTTGCTTATG                      ACAAGAACTACCACGGCAAGGGCGACACTGTAGACAACATCAACAAGGGTGTATCTACCTCAACTC                      GAGGAATCGCGTATGCCACTGCTCAGTATGCTAGTTCGCTGCGCGGATTCCTCAACCCGCCAAAGACGG                      GTAAGCGTGACGTGAGCCCCCGTGGCCAGTCTATGCCTGGTGGTGGATGCGGACACCACAGCGTCTTCA                      TGTA                 </p>

Un marco de lectura abierto ("ORF") de caLAP2 descrito de 1.488 nucleótidos empieza con un codón de inicio ATG en la posición 1 (subrayado en la Tabla 10B).

**Tabla 10B. Secuencia de nucleótidos de caLAP2 (SEQ ID NO:29).**

ATGAAGACACAGTTGTTGAGTCTGGGAGTTGCCCTCACGGCCATCTCTCAGGGCGTTATTGCTGAGGAT  
 GCCTTGAAGTGGCCATTCAAGCCGTTGGTTAATGCTGATGACCTGCAAACAAGATTAAGCTCAAGGAT  
 CTTATGGCTGGCGTACAGAACTCCAAGACTTCGCCTACGCTCACCCGTGAGAAGAATCGAGTATTCGGT  
 GGTGCTGGCCACAAGGATACCGTCGACTGGATCTACAATGAGCTCAAGGCTACCGGCTACTACGATGTG  
 AAGATGCAGCCACAAGTCCACCTGTGGTCTCATGCTGAGGCAGCTGTCAATGCCAATGGCAAGGATCTC  
 ACTGCCAGTGCCATGTCTACAGCCCTCCAGCCGACAAGATCACTGCCGAGCTTGTCTCTGGCCAAGAAC  
 ATGGGATGCAATGCTACTGATTACCCAGAGGGTACCAAGGGCAAGATTGTCTCTCGAGCGTGGTGTCTC  
 TGCAGCTTTGGCGAGAAGTCCGCTCAGGCTGGCGATGCAAAGGCTATTGGTGCCATCGTCTACAACAAC  
 GTCCCTGGAAGCTTGGCCGGCACCCTGGGTGGCCTTGACAACCGCCATGCTCCAAGTGTGGAATCTCT  
 CAGGCTGATGGAAGAACCTCGCTAGCCTTGTGCGCTCTGGCAAGGTTACCGTCACCATGAACGTTATC  
 AGCAAGTTTGAAGACAGGACTACCTGGAACGTCATTGCCGAGACCAAGGGAGGAGACCACAACAACGTC  
 ATCATGCTCGGTTCTCACTCTGACTCTGTGACGCGCCGGCCTGGTATCAACGACAACGGCTCCGGTACC  
 ATTGGTATCATGACCGTTGCCAAAGCCCTACCAACTTCAAGGTCAACAACGCCGTCGCTTCGGCTGG  
 TGGACCGCCGAGGAGTTCGGCCTTCTCGGCAGCACTTCTACGTCGACAGCCTTGACGACCGTGAAGT  
 CACAAGGTCAAGCTGTACCTCAACTTCGACATGATTGGCTCCCCCAACTTCGCCAACCAGATCTACGAC  
 GGAGACGGCTCCGCCTACAACATGACTGGCCCCGCCGGATCTGCTGAAATCGAGTACCTGTTTCGAGAAG  
 TTCTTCGATGACCAGGGAATCCCACACCAGCCACCGCCTTACCAGGCCGCTCCGACTACTCTGCCTTC  
 ATCAAGCGCAACGTCCTGCGGAGGTCTGTTTACTGGTGTGAGGTCGTCAAGACCGCCGAGCAGGCT  
 AAGCTATTTGGCGGCGAGGCTGGCGTTGCTTATGACAAGAATAACCACGGCAAGGGCGACACTGTAGAC  
 AACATCAACAAGGGTGTATCTACCTCAACTCGAGGAATCGCGTATGCCACTGCTCAGTATGCTAGT  
 TCGCTGCGCGGATTCCCAACCCGCCCAAAGACGGGTAAGCGTGACGTGAGCCCCCGTGGCCAGTCTATG  
 CCTGGTGGTGGATGCGGACACCACAGCGTCTTCATGTAA

Un ácido nucleico de caLAP2 descrito (SEQ ID NO: 29) codifica una proteína que tiene 495 restos de aminoácidos (SEQ ID NO: 30), que se presenta en la Tabla 10C usando el código de aminoácidos de una letra.

**Tabla 10C. Secuencia de la proteína codificada caLAP2 (SEQ ID NO:30).**

MKTQLLSLGVALTAI SQGVIAEDALNWPFKPLVNADDLQNKIKLKDLMAGVQKLQDFAYAHPEKRVFVG  
 GAGHKD'TVDWIYNELKATGYDYDKMQPQVHLWSHAEAAVNANGKDLTASAMSYSPADKITAELVLAKN  
 MGCNATDYPEGTKGKIVLIERGVCSFGEKSAQAGDAKAI GAI VYNNVPGSLAGTLGGLDNRHAPTAGIS  
 QADGKNLASLVASGKVTVTMNVISKFENRTTWNVIAETKGGDHNNVIMLGSHSDSVDAGPGINDNGSGT  
 IGIMTVAKALTNFKVNNAVRFGWWTAEFGLLGSTFYVDSLDDRELHKVKLYLNFDMIGSPNFANQIYD  
 GDGSAYNMTGPAGSAEIEYLFKPFDDQGI PHQPTAFTGRSDYSAFIKRNVPAGGLFTGAEVVKTAEQA  
 KLFGEAGVAYDKNYHGKGD'TVDNINKGAIYLNTRGIAAYATAQYASSLRGFPTRPKTGKRDVSPRGQSM  
 PGGCGHHSVFM

5 La caLAP2 descrita tiene homología con las secuencias de aminoácidos mostradas en los datos BLAST listados en las Tablas 10D, 10E y 10F. Estos datos se analizaron por el programa POR PAREJAS BLAST.

**Tabla 10D: Resultados TBLASTN para caLAP2**

Índice/Identificador Génico	Proteína/Organismo	Longitud (aa)	Identidad (%)	Positivos (%)	Esperado
gi/600025	Genes RIF1, DPB3, Yml27 y SNF5 de <i>Saccharomyces cerevisiae</i> (s288c)	32421	182/477 (38%)	254/477 (53%)	8e-77
gi/469463	Gen de aminopeptidasa Y de <i>Saccharomyces cerevisiae</i>	2272	182/477 (38%)	254/477 (53%)	8e-77
gi/16033407	Gen del precursor de leucina aminopeptidasa de <i>Bacillus licheniformis</i>	2054	132/474 (27%)	215/474 (45%)	3e-27

Tabla 10E: Resultados BLASTX para caLAP2

Índice/Identificador Génico	Proteína/Organismo	Longitud (aa)	Identidad (%)	Positivos (%)	Esperado
gi/1077010	precursor de aminopeptidasa Y, vacuolar/ <i>Saccharomyces cerevisiae</i>	537	182/477 (38%)	254/477 (53%)	9e-78
gi/6319763	Aminopeptidasa yscIII; Ape3p/ <i>Saccharomyces cerevisiae</i>	563	182/477 (38%)	254/477 (53%)	9e-78
gi/31791596	lipoproteína aminopeptidasa probable LPQL/ <i>Mycobacterium bovis</i>	500	188/485 (38%)	269/485 (55%)	3e-77
gi/15839805	hidrolasa/ <i>Mycobacterium tuberculosis</i>	493	187/481 (38%)	268/481 (55%)	6e-77

Tabla 10F: Resultados BLASTP para caLAP2

Índice/Identificador Génico	Proteína/Organismo	Longitud (aa)	Identidad (%)	Positivos (%)	Esperado
gi/6319763	aminopeptidasa yscIII; Ape3p/ <i>Saccharomyces cerevisiae</i>	563	182/477 (38%)	254/477 (53%)	5e-78
gi/1077010	precursor de aminopeptidasa Y, vacuolar/ <i>Saccharomyces cerevisiae</i>	537	182/477 (38%)	254/477 (53%)	8e-78
gi/15839805	hidrolasa/ <i>Mycobacterium tuberculosis</i>	493	187/481 (38%)	268/481 (55%)	1e-71
gi/31617182	lipoproteína aminopeptidasa probable LPQL/ <i>Mycobacterium bovis</i>	500	188/485 (38%)	269/485 (55%)	2e-71
gi/15598135	aminopeptidasa probable/ <i>Pseudomonas aeruginosa</i>	536	166/445 (37%)	242/445 (54%)	2e-65

**meLAP2**

- 5 meLAP2 es una leucina aminopeptidasa de *Trichophyton mentagrophytes*. Un ácido nucleico de meLAP2 de 1.775 nucleótidos (SEQ ID NO: 31) se muestra en la Tabla 11A.

Tabla 11A. Secuencia de nucleótidos genómica de meLAP2 (SEQ ID NO:31).

ATGAAGTCGCAACTGTTGAGCCTAGCCGTGGCCGTCACCACCATTTCAGGGCGTTGTTGGTCAAGAG  
 CCCTTGGATGGCCCTTCAAGCCTATGGTCACTCAGGTGAGTTGCTGTCAACAGATCGATCGATCGATC  
 TACCTTCGTCCCTGTCACCTATAACTCCACAGCAGGACCAAGAAAACACAAGTTTTCCGGGGAATTCTT  
 ATGTGCTGATGTAATGTATAGGATGACCTGCAAAACAAGATTAAGCTCAAGGATATCATGGCAGGTGT  
 CGAGAAGCTGCAAAGCTTTCTGATGCTCATCCTGAAAAGAACCGAGTGTTGCGGTGTAATGGCCACAA  
 GGACACTGTCGAGTGGATCTACAATGAGCTCAAGGCCACCGGCTACTACAATGTGAAGAAGCAGGAGCA  
 GTTACACCTGTGGTCTCACGCTGAGGCCGCTCTCAGTGCCAATGGCAAGGACCTCAAGGCCAGCGCCAT  
 GTCGTACAGCCCTCCTGCCAACAGATCATGGCCGAGCTTGTCGTTGCCAAGAACAATGGCTGCAATGC  
 TGTAAAGTGCCATACACTTCCTATACATCACATTCACCTTTAGAATGAAGAGCGCGGGAGAACTGATTTTT  
 TTTTTTTTTTTTTTTTTTTTTTTGTAACAGACCATTACCCAGAGAACACTCAGGGAAAGATAGTCCTCATT  
 CAGCGTGGTGTCTGCAGCTTCGGCGAGAAGTCTTCTCAGGCTGGTGTGCGAAGGCTATTGGTGGCCGTT  
 GTCTACAACAACGTCCCCGGATCCCTTGCTGGCACTCTGGTGGCCTTGACAAGCGCCATGTCCCAACC  
 GCTGGTCTTTCCAGGAGGATGGAAAGAATCTTGCTAGCCTCGTTGCTTCTGGCAAGGTTGATGTCACC  
 ATGAACGTTGTGAGTCTGTTTGAGAACC GAACCGTAAGTAACCTCAACGTCATATCCAGCATTAACTCT  
 TCAGGAGTATATACTAATTCGGTATCTCACAGCTGGAACGTCATTGCTGAGACCAAGGGAGGAGACC  
 ACAACAATGTTGTCATGCTTGGTGCTCACTCCGACTCCGTCGATGCCGGCCCCGGTATCAACGACAACG  
 GCTCCGGCTCCATTGGTATCATGACCGTTGCCAAAGCCCTTACTAAGCTCAACAACGCCGTTT  
 GCTTTGCCTGGTGGACCGCTGAGGAATTCGGTCTCCTTGGAAAGCACCTTCTACGTCGACAGCCTTGATG  
 ACCGTGAGCTGCACAAGGTCAAGCTGTACCTCAACTTCGACATGATCGGCTCTCCCAACTTCGCCAACC  
 AGATCTACGACGGTGACGGTTCGGCCTACAACATGACTGGTCCCGCTGGCTCTGCTGAAATCGAGTACC  
 TGTTGAGAAAGTTCTTTGACGACCAGGGTCTCCACACCAGCCACTGCCTTACCGGCCGATCCGACT  
 ACTCTGCATTCAAGCGCAACGTCCCCGCTGGAGGTCTTTTCACTGGTGGCCGAGGTTGTCAAGACCC  
 CCGAGCAAGTTAAGCTGTTGGTGGTGGAGGCTGGCGTTGCCATGACAAGAAGTACCATGGCAAGGGTG  
 ACACCGTTGCCAACAATCAACAAGGGAGCTATCTTCTTAACACTCGAGCAATCGCCTACTCTGTGGCCG  
 AGTATGCTCGATCCCTCAAGGGCTTCCCAACCCGCCAAAGACCGGCAAGCGTGCCGTC AACCTCAGT  
 ATGCTAAGATGCCTGGTGGTGGTTGCGGACACCACACTGCTTCATGTAA

Un marco de lectura abierto ("ORF") de meLAP2 descrito de 1.488 nucleótidos empieza con un codón de inicio ATG en la posición 1 (subrayado en la Tabla 11B).

Tabla 11B. Secuencia de nucleótidos de meLAP2 (SEQ ID NO:32).
<p>ATGAAGTCGCAACTGTTGAGCCTAGCCGTGGCCGTCACCACCATTCCAGGGCGTTGTTGGTCAAGAG                      CCCTTTGGATGGCCCTTCAAGCCTATGGTCACTCAGGATGACCTGCAAAACAAGATTAAGCTCAAGGAT                      ATCATGGCAGGTGTCGAGAAGCTGCAAAGCTTTTCTGATGCTCATCCTGAAAAGAACCGAGTGTTCGGT                      GGTAATGGCCACAAGGACTGTCGAGTGGATCTACAATGAGCTCAAGGCCACCGGCTACTACAATGTG                      AAGAAGCAGGAGCAGGTACACCTGTGGTCTCACGCTGAGGCCGCTCTCAGTGCCAATGGCAAGGACCTC                      AAGGCCAGCGCCATGTGCTACAGCCCTCCTGCCAACAAGATCATGGCCGAGCTTGTGCTGGCAAGAAC                      AATGGCTGCAATGCTACCGATTACCCAGAGAACAACACTCAGGGAAAGATAGTCCTCATTAGCGTGGTGTG                      TGCAGCTTCGGCGAGAAGTCTTCTCAGGCTGGTGATGCGAAGGCTATTGGTGCCGTTGTCTACAACAAC                      GTCCCCGGATCCCTTGTGTCGACCTCTTGGTGGCCTTGACAAGCGCCATGTCCCAACCGCTGGTCTTTCC                      CAGGAGGATGGAAAGAATCTTGCTAGCCTCGTTGCTTCTGGCAAGTTGATGTACCATGAACGTTGTG                      AGTCTGTTTGAAGAACCAGACCTGGAACGTCATTGCTGAGACCAAGGGAGGAGACCACAACAATGTT                      GTCATGCTTGGTGCTCACTCCGACTCCGTCGATGCCGGCCCCGGTATCAACGACAACGGCTCCGGCTCC                      ATTGGTATCATGACCGTTGCCAAGCCCTTACTAACTTCAAGCTCAACAACGCCGTTGCTTTGCCTGG                      TGGACCGCTGAGGAATTCGGTCTCCTTGGAAAGCACCTTCTACGTCGACAGCCTTGATGACCGTGAGCTG                      CACAAGGTCAAGCTGTACCTCAACTTCGACATGATCGGCTCTCCCAACTTCGCCAACCAGATCTACGAC                      GGTGACGGTTCGGCCTACAACATGACTGGTCCCGCTGGCTCTGCTGAAATCGAGTACCTGTTGAGAGAAG                      TTCTTTGACGACCAGGGTCTCCACACCAGCCACTGCCTTACCAGGCCGATCCGACTACTCTGCATTCT                      ATCAAGCGCAACGTCCTCCGCTGGAGGTCTTTTCACTGGTGCCGAGGTTGTCAAGACCCCCGAGCAAGTT                      AAGCTGTTCCGGTGGTGAGGCTGGCGTTGCCTATGACAAGAACTACCATGGCAAGGGTGACACCGTTGCC                      AACATCAACAAGGGAGCTATCTTCTTAACACTCGAGCAATCGCCTACTCTGTGGCCGAGTATGCTCGA                      TCCCTCAAGGGCTTCCCAACCCGCCCAAGACCGGCAAGCGTGCCGTCAACCCTCAGTATGCTAAGATG                      CCTGGTGGTGGTTGCGGACACCACACTGTCTTCATGTAA</p>

Un ácido nucleico de meLAP2 descrito (SEQ ID NO: 32) codifica una proteína que tiene 495 restos de aminoácidos (SEQ ID NO: 33), que se presenta en la Tabla 11C usando el código de aminoácidos de una letra.

Tabla 11C. Secuencia de la proteína codificada meLAP2 (SEQ ID NO:33).
<p>MKSQLLSLAVAVTTISQGVVQEPFGWPFKPMVTQDDLQNKIKLKDIMAGVEKLQSFSDAHPEKNRVFG                      GNGHKDTVEWYINELKATGYYNVKKQEQVHLWVSHAEALSANGKDLKASAMSYSPANKIMAEVVAKN                      NGCNATDYPENTQKIVLIQRGVCSFGEKSSQAGDAKIGAVVYNNVPGSLAGTLGGLDKRHVPTAGLS                      QEDGKNLASLVASGKVDVTMNVVSLFENRTTWNVIAETKGGDHNNVVMGHAHSDSVDAGPGINDNGSGS                      IGIMTVAKALTNFKLNNAVRFWWTAEEFLLGSTFYVDSLDDRELHKVKLYLNFDMIGSPNFANQIYD                      GDGSAYNMTGPAGSAEIEYLFKFFDDQGLPHQPTAFTGRSDYSAFIKRNVPAGGLFTGAEVVKTPEQV                      KLFGGEAGVAYDKNYHGKGDVTANINKGAI FLNTRAIAYSVAEYARSLKGFPTRPKTGKRAVNPQYAKM                      PGGGCGHHTVFM</p>

5 La meLAP2 descrita tiene homología con las secuencias de aminoácidos mostradas en los datos BLAST listados en las Tablas 11D, 11E y 11F. Estos datos se analizaron por el programa POR PAREJAS BLAST.

Tabla 11D: Resultados TBLASTN para meLAP2

Índice/Identificador Génico	Proteína/Organismo	Longitud (aa)	Identidad (%)	Positivos (%)	Esperado
gi/600025	Genes RIF1, DPB3, YmL27 y SNF5 de <i>Saccharomyces cerevisiae</i> (s288c)	32421	182/479 (37%)	251/479 (52%)	2e-70
gi/469463	Gen de aminopeptidasa Y de <i>Saccharomyces cerevisiae</i>	2272	182/477 (37%)	251/479 (52%)	2e-70

Tabla 11E: Resultados BLASTX para meLAP2

Índice/Identificador Génico	Proteína/Organismo	Longitud (aa)	Identidad (%)	Positivos (%)	Esperado
gi/1077010	precursor de aminopeptidasa Y, vacuolar/ <i>Saccharomyces cerevisiae</i>	537	180/479 (37%)	251/479 (52%)	8e-71
gi/6319763	aminopeptidasa yscIII; Ape3p/ <i>Saccharomyces cerevisiae</i>	563	180/479 (37%)	251/479 (52%)	8e-71
gi/15839805	hidrolasa/ <i>Mycobacterium tuberculosis</i>	493	159/440 (36%)	236/440 (53%)	1e-63
gi/31791596	lipoproteína aminopeptidasa probable LPQL/ <i>Mycobacterium bovis</i>	500	159/440 (36%)	236/440 (53%)	1e-63
gi/15598135	aminopeptidasa probable/ <i>Pseudomonas aeruginosa</i>	536	158/445 (35%)	237/445 (53%)	1e-62
gi/1045225	N-acetilpuromicina N-acetilhidrolasa/ <i>Streptomyces anulatus</i>	485	154/477 (32%)	218/477 (45%)	4e-48
gi/29831415	aminopeptidasa putativa/ <i>Streptomyces avermitilis</i>	315	95/244 (38%)	131/244 (53%)	2e-37

Tabla 11F: Resultados BLASTP para meLAP2

Índice/Identificador Génico	Proteína/Organismo	Longitud (aa)	Identidad (%)	Positivos (%)	Esperado
gi/6319763	aminopeptidasa yscIII; Ape3p/ <i>Saccharomyces cerevisiae</i>	563	179/479 (37%)	248/479 (51%)	9e-71
gi/1077010	precursor de aminopeptidasa Y, vacuolar/ <i>Saccharomyces cerevisiae</i>	537	179/479 (37%)	248/479 (51%)	9e-71
gi/31617182	lipoproteína aminopeptidasa probable LPQL/ <i>Mycobacterium bovis</i>	500	159/440 (36%)	236/440 (53%)	2e-62
gi/15839805	hidrolasa/ <i>Mycobacterium tuberculosis</i>	493	159/440 (36%)	236/440 (53%)	2e-62

**ruDPPIV**

- 5 ruDPPIV es una dipeptidilpeptidasa IV de *T. rubrum*. Un ácido nucleico de ruDPPIV de 2.326 nucleótidos (SEQ ID NO: 34) se muestra en la Tabla 12A. Un marco de lectura abierto ("ORF") de ruDPPIV descrito empieza con un codón de inicio ATG en la posición 1 (subrayado en la Tabla 12A).

**Tabla 12A. Secuencia de nucleótidos de ruDPPIV (SEQ ID NO:34).**

ATGAAGCTCCTCTCGCTACTTATGCTGGCGGGCATCGCCCAAGCCATCGTTCTCCTCGTGAGCCCCGT  
 TCACCAACTGGTGGCGGCAACAAGCTGTTGACCTACAAGGAGTGTGTCCCTAGAGCTACTATCTCTCCA  
 AGGTCGACGTCCCTTGCCCTGGATTAACAGTGAAGAAGATGGCCGGTACATCTCCAGTCCGACGATGGA  
 GCATTGATCCTCCAGAACATCGTCACGAACACCAACAAGACTCTCGTGGCCGACAGACAAGGTACCCAAG  
 GGTACTATGACTACTGGTCAAGCCAGACCTTTCTGCTGTCTTATGGGCAACCAATTACACCAAGCAG  
 TACCGTCACTCTTACTTTGCCAACTACTTCATTCTAGACATCAAAAAGGGATCGTTGACCCCTCTAGCC  
 CAGGACCAGGCTGGTGACATCCAGTATGCTCAATGGAGCCCCATGAACAACCTCTATCGCCTATGTCCGT  
 GRAAACGACCTGTATATCTGGAACAATGGCAAGACCAAGCGTATTACCGAAAATGGCGGCCCGGATATC  
 TTCAATGGTGTCCCTGACTGGGTATACGAGGAAGAAATCTTCGGGGACCGGTTGCTCTTTGGTTCTCA  
 CCTGACGGTGAATACCTTGCGTACCTCCGCTTTAACGAGACTGGAGTCCCAGCTACACTATTCCGTAC  
 TACAAGAACAAGCAAAAAGATTGCCCCCTGCCTACCAAGGGAGCTGGAGATCCGTTACCCTAAAGTCTCT  
 GCGAAGAACCCAACCGTGCAGTTCACCTGTTAAACATTGCTTCATCCAGGAGACAACCTATCCCAGTT  
 ACTGCGTTCCCGAAAACGATCTTGTGATCGGTGAGGTTGCTTGGCTCAGCAGTGGCCATGATAGTGTA  
 GCATATCGTGCTTTCAACCGTGTCCAGGATAGAGAAAAGATTGTCAGCGTCAAGGTTGAGTCCAAGGAA  
 TCCAAGGTTATTTCGCGAAAAGAGATGGCACCAGCGCTGGATCGACAACCTTCTCTCATGTATATATCG  
 GAAACGTTAACGGCAAGGAGTACTACGTCGATATATCTGATGCTTCTGGCTGGGCACATATCTACCTCT  
 ACCCGGTTGATGGAGGAAAGGAGATTGACTAACAAGGGAGAATGGGAAGTCGTTGCCATTCTCAAGG  
 TTGACACGAAGAAGAAGCTGATCTACTTACCTCTACCAAATATCACAGCACCCTCGACACGTCTACT  
 CTGTCTCGTATGACACAAAGGTCATGACCCCTCTCGTCAACGATAAGGAGGCTGCGTACTACACTGCAT  
 CCTTCTCGGCCAAGGGTGGTTACTATATCTTGTCTTACCAAGGTCCAAATGTTCCATACCAAGAACTTT  
 ACTCCACCAAGGACAGTAAGAAGCCTCTCAAGACAATCACTAGCAATGATGCATGCTCGAGAAGCTGA  
 AGGAGTACAAGCTCCCCAAGGTTAGCTTCTTTGAGATCAAGCTTCCATCTGGTGAAACCCCTTAATGTTA  
 AGCAACGCCTACCACCTAACTTCAACCCACACAAGAAGTACCCCGTCTCTTCACTCCGTATGGTGGCC  
 CTGGTGCCCAAGAGGTAAGCCAGGCATGGAATTCATTGGACTTCAAGTCTTACATTACATCTGACCCTG  
 AGCTTGAATACGTTACCTGGACTGTTGACAACCGTGGAAACCGGCTACAAGGGCCGCAAGTTCGCGACGG  
 CCGTAGCTAAGCGTCTCGGTTTCTCGAAGCCCAGGACCAGGCTTTTGCTGCTAAGGAGGTGCTGAAAA  
 ACCGTTGGGCTGATAAGGACCATATTGGAATCTGGGGCTGNAGCTATGGCGGCTTCTGACCCTAAGA  
 CCCTCGAGACCGACAGTGGTGTATTCACTTTTGGTATCAGTACTGCTCCTGTCTCTGATTTTCAGACTCT  
 ACGACAGCATGTACTGAGCGTTACATGAAGACCGTTGAACCTAAACGCTGACGGCTACAGTGAGACCG  
 CCGTGCACAAGGTTGATGGCTTTAAGAACCTCAAAGGTCATTACTCATCCAGCATGGAACCGGTGACGA  
 CAACGTCCTACTTCCAAAACGCCGCTGTCCTTTCCAACACCCTGATGAACGGCGGTGTAACCTGCAGACAA  
 GTTGACTACTCAGTGGTTACTGACTCGGACCACGGCATCAGATACGATATGGACTCCACTTACCAGTA  
 CAAGCAGCTTTCTAAGATGGTCTACGACCAGAAGCAACGAAGGCCAGAAAGCCACCAATGCACCAATG  
 GAGCAAGAGAGTTTTGGCTGCCCTGTTTGGTGAGAGGGCAGAGGAATGA

Un ácido nucleico de ruDPPIV descrito (SEQ ID NO: 34) codifica una proteína que tiene 775 restos de aminoácidos (SEQ ID NO: 35), que se presenta en la Tabla 12B usando el código de aminoácidos de una letra.

**Tabla 12B. Secuencia de la proteína codificada ruDPPIV (SEQ ID NO:35).**

MKLLSLLMLAGIAQAI VPPREPRSPTGGGNKLLTYKECVPRATISPRSTSLAWINSEEDGRYISQSDDG  
 ALILQNI VTNKTLVAADKVPKGYDYWFKPDL S AVLWATNYTKQYRHSYFANYFILD I KKGSLTPLA  
 QDQAGDIQYAQWSPMNNSIAYVRXNDLYIWNNGKTKRITENGGPDI FNGVPDWVYEEEI FGDRFALWFS  
 PDGEYLAYLRFNETGVPTYTIPYYKNKQKIAPAYPRELEIRYPKVS AKNPTVQFHL LN I ASSQETTI PV  
 TAFPENDLVIGEVAWLSSGHD SVAYRAFNRVQDREKIVSVKVESKESKVI RERDGTGDWIDNLLSMSYI  
 GNVNGKEYYVDISDASGWAHIYLYPVDGGKEIALTKGEWEVVA I LKVDTKKKLIYFTSTKYHSTTRHVV  
 SVSYDTKVM TPLVNDKEAAYTASFSAKGGYI LSYQGNVPYQELYSTKDSKKPLKITSN DALLEKL  
 KEYKLPKVSFFEIKLPSGETLNVKQRLPPNFNPHKKYPVLF T PYGGPGAQEV SQAWNSLDFKSYITSDP  
 ELEYVTWTV DNRGTGYKGRKFRSAVAKRLGFLEAQDQVFAAKEVLKNRWADKDHIGIWGXSYGGFLTAK  
 TLETDSGVFTFGISTAPVSDFRLYDSMYTERYMKTVELNADGYSETAVHKVDGFKNLKGHYLIQHGTGD  
 DNVHFQNAAVLSNTLMNGGVTADKLTQWFTSDHGI RYDMDSTYQYKQLSKM VYDQKQRRPESPPMHQ  
 WSKRVLAALFGERAE

La ruDPPIV descrita tiene homología con las secuencias de aminoácidos mostradas en los datos BLAST listados en las Tablas 10C, 10D y 10E. Estos datos se analizaron por el programa POR PAREJAS BLAST.

**Tabla 12C: Resultados TBLASTN para ruDPPIV**

Índice/Identificador Génico	Proteína/Organismo	Longitud (aa)	Identidad (%)	Positivos (%)	Esperado
gi/2351699	Gen de dipeptidil-peptidasa IV (Dpp4) de <i>Aspergillus fumigatus</i>	2352	469/761 (61%)	585/761 (76%)	0.0
gi/2924304	Gen de DppIV de <i>Aspergillus oryzae</i>	4771	448/769 (58%)	568/769 (73%)	0.0
gi/32422540	Cepa OR74A de <i>Neurospora crassa</i>	2688	256/720 (35%)	374/720 (51%)	e-114
gi/14330262	Gen de dapB de <i>Aspergillus niger</i> para dipeptidil aminopeptidasa tipo IV, exones 1-3	3989	224/637 (35%)	333/637 (52%)	e-111
gi/1621278	ARNm para dipeptidil-peptidasa IV de <i>Xenopus laevis</i>	3337	244/752 (32%)	375/752 (49%)	e-100
gi/6978772	Dipeptidil peptidasa 4 (Dpp4) de <i>Rattus norvegicus</i>	4835	246/742 (33%)	373/742 (50%)	8e-98

**Tabla 12D: Resultados BLASTX para ruDPPIV**

Índice/Identificador Génico	Proteína/Organismo	Longitud (aa)	Identidad (%)	Positivos (%)	Esperado
gi/2351700	dipeptidil-peptidasa IV/ <i>Aspergillus fumigatus</i>	765	218/341 (63%)	270/341 (79%)	0.0
gi/2924305	proil dipeptidil peptidasa/ <i>Aspergillus oryzae</i>	771	213/344 (61%)	270/344 (78%)	0.0
gi/1621279	dipeptidil-peptidasa IV/ <i>Xenopus laevis</i>	748	118/349 (33%)	186/349 (53%)	8e-93
gi/535388	dipeptidil-peptidasa IV/ <i>Homo sapiens</i>	766	125/375 (33%)	191/375 (50%)	3e-90

5

**Tabla 12E: Resultados BLASTP para ruDPPIV**

Índice/Identificador Génico	Proteína/Organismo	Longitud (aa)	Identidad (%)	Positivos (%)	Esperado
gi/2351700	dipeptidil-peptidasa IV/ <i>Aspergillus fumigatus</i>	765	468/761 (61%)	585/761 (76%)	0.0
gi/2924305	proil dipeptidil peptidasa/ <i>Aspergillus oryzae</i>	771	448/769 (58%)	568/769 (73%)	0.0
gi/14330263	dipeptidil aminopeptidasa tipo IV/ <i>Aspergillus niger</i>	901	261/733 (35%)	387/733 (52%)	e-114
gi/19114882	dipeptidil aminopeptidasa/ <i>Schizosaccharomyces</i>	793	258/742	396/742	e-106



Índice/Identificador Génico	Proteína/Organismo	Longitud (aa)	Identidad (%)	Positivos (%)	Esperado
	<i>pombe</i>		(34%)	(53%)	
gi/3660	dipeptidil aminopeptidasa B/ <i>Saccharomyces cerevisiae</i>	841	254/750 (33%)	370/750 (49%)	2e-95

Un aspecto de la descripción está relacionado con moléculas de ácido nucleico aisladas que codifican polipéptidos EXOX o partes biológicamente activas de éstos. También se incluyen fragmentos de ácido nucleico suficientes para usarse como sondas de hibridación para identificar ácidos nucleicos que codifican EXOX (por ejemplo, ARNm EXOX) y fragmentos para usarse como cebadores de PCR para la amplificación y/o mutación de moléculas de ácido nucleico EXOX. Tal y como se usa en la presente memoria, el término "molécula de ácido nucleico" se pretende que incluya moléculas de ADN (por ejemplo, ADNc o ADN genómico), moléculas de ARN (por ejemplo, ARNm), análogos del ADN o ARN generado usando análogos de nucleótidos, y derivados, fragmentos y homólogos de éstos. La molécula de ácido nucleico puede ser monocatenaria o bicatenaria.

Un ácido nucleico EXOX puede codificar un polipéptido EXOX maduro. Tal y como se usa en la presente memoria, una forma "madura" de un polipéptido o proteína descrito en la presente memoria es el producto de un polipéptido o forma precursora o proproteína natural. El polipéptido, precursor o proproteína natural incluye, como ejemplo no limitativo, el producto génico de longitud completa, codificado por el gen correspondiente. Alternativamente, puede estar definido como el polipéptido, precursor o proproteína codificado por un ORF descrito en la presente memoria. La forma "madura" del producto aparece, de nuevo como ejemplo no limitativo, como resultado de una o más etapas de procesamiento naturales ya que pueden tener lugar en el interior de la célula, o célula anfitriona, en la que aparece el producto génico. Los ejemplos de dichas etapas de procesamiento que dan lugar a una forma "madura" de un polipéptido o proteína incluyen la escisión del resto de metionina N-terminal codificado por el codón de inicio de un ORF, o la escisión proteolítica de un péptido señal o secuencia líder. Así, una forma madura que aparece a partir de un polipéptido o proteína precursor que tiene restos 1 a N, en el que el resto 1 es la metionina N-terminal, tendrá los restos 2 a N que permanecen después de la eliminación de la metionina N-terminal. Alternativamente, una forma madura que aparece a partir de un polipéptido o proteína precursor que tiene restos 1 a N, en el que se escinde una secuencia señal N-terminal del resto 1 al resto M, tendrá los restos desde el resto M+1 al resto N que permanecen. Además, tal y como se usa en la presente memoria, una forma "madura" de un polipéptido o proteína puede aparecer a partir de una etapa de modificación posterior a la traducción distinta de un evento de escisión proteolítico. Dichos procesos adicionales incluyen, como ejemplo no limitativo, glicosilación (tipos N, O y W), miristoilación, fosforilación, sulfatación, ciclación N-terminal, o amidación C-terminal. En general, un polipéptido o proteína maduro puede resultar de la operación de sólo uno de estos procesos, o una combinación de cualquiera de ellos.

El término "sondas", tal y como se utiliza en la presente memoria, se refiere a secuencias de ácido nucleico de longitud variable, preferiblemente entre al menos aproximadamente 10 nucleótidos (nt), 100 nt, o tantos como aproximadamente, por ejemplo, 6.000 nt, dependiendo del uso específico. Las sondas se usan en la detección de secuencias de ácido nucleico idénticas, similares o complementarias. Las sondas con una longitud mayor se obtienen generalmente a partir de una fuente natural o recombinante, son altamente específicas, y más lentas para hibridar que las sondas de oligómero con una longitud menor. Las sondas pueden ser mono o bicatenarias y estar diseñadas para tener especificidad en PCR, tecnologías de hibridación basadas en membranas o tecnologías semejantes a ELISA.

El término molécula de ácido nucleico "aislada", tal y como se utiliza en la presente memoria, es una, que se separa de otras moléculas de ácido nucleico, que están presentes en la fuente natural del ácido nucleico. Preferiblemente, un ácido nucleico "aislado" no contiene secuencias que de forma natural flanquean el ácido nucleico (por ejemplo, secuencias localizadas en el extremo 5' y 3' del ácido nucleico) en el ADN genómico del organismo del que se obtiene el ácido nucleico. Por ejemplo, las moléculas de ácido nucleico EXOX aisladas pueden contener menos de aproximadamente 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0,5 kb ó 0,1 kb de secuencias de nucleótido que de forma natural flanquean la molécula de ácido nucleico en el ADN genómico de la célula/tejido/especie del que se obtiene el ácido nucleico. Además, una molécula de ácido nucleico "aislada", tal como una molécula de ADNc, puede no contener sustancialmente otro material celular o medio de cultivo cuando se produce por técnicas recombinantes, o precursores químicos u otros productos químicos cuando se sintetiza químicamente. Particularmente, esto significa que el ácido nucleico o proteína es al menos aproximadamente 50% pura, más preferiblemente al menos aproximadamente 85% pura y lo más preferiblemente al menos aproximadamente 99% pura.

Tal y como se usa en la presente memoria, el término "recombinante" cuando se usa respecto a una célula indica que la célula replica un ácido nucleico heterólogo, o expresa un péptido o proteína codificado por un ácido nucleico heterólogo. Las células recombinantes pueden contener genes que no se encuentran en la forma nativa (no recombinante) de la célula. Las células recombinantes también pueden contener genes encontrados en la forma nativa de la célula en la que los genes se modifican y se reintroducen en la célula por medios artificiales. El término también engloba a las células que contienen un ácido nucleico endógeno para la célula que se ha modificado sin eliminar el ácido nucleico de la célula; dichas modificaciones incluyen las obtenidas por reemplazo de genes,

mutación específica de sitio, y técnicas relacionadas. Un experto en la técnica reconocerá que estas células pueden usarse para organismos transgénicos unicelulares o multicelulares, por ejemplo hongos transgénicos que producen EXOX.

5 Una molécula de ácido nucleico, por ejemplo, una molécula de ácido nucleico que tiene la secuencia de nucleótidos de SEQ ID NOs: 2, 5, 8, 11, 14, 17, 20, 23, 26, 29, 32 ó 34 o un complemento de esta secuencia de nucleótidos mencionada anteriormente, puede aislarse usando técnicas de biología molecular estándar y la información de la secuencia proporcionada en la presente memoria. Usando toda o una parte de la secuencia de ácido nucleico de SEQ ID NOs: 2, 5, 8, 11, 14, 17, 20, 23, 26, 29, 32 ó 34 como sonda de hibridación, las moléculas EXOX pueden aislarse usando técnicas de hibridación y clonación estándar (por ejemplo, como se describe en Sambrook *et al.*, (eds.), MOLECULAR CLONING: A LABORATORY MANUAL 2ª Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989; y Ausubel *et al.*, (eds.), CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, Nueva York, NY, 1993).

15 Un ácido nucleico puede amplificarse usando ADNc, ARNm o alternativamente ADN genómico, como molde y cebadores oligonucleotídicos apropiados según las técnicas de amplificación por PCR estándar. El ácido nucleico así amplificado puede clonarse en un vector apropiado y caracterizarse por análisis de secuencia de ADN. Además, los oligonucleótidos correspondientes a las secuencias de nucleótidos EXOX pueden prepararse por técnicas sintéticas estándar, por ejemplo, usando un sintetizador de ADN automático.

20 Tal y como se usa en la presente memoria, el término "oligonucleótido" se refiere a una serie de restos de nucleótido unidos, teniendo dicho oligonucleótido un número suficiente de bases nucleotídicas para usarse en una reacción de PCR. Una secuencia oligonucleotídica corta puede basarse en, o diseñarse a partir de, una secuencia genómica o de ADNc y se usa para amplificar, confirmar o revelar la presencia de un ADN o ARN idéntico, similar o complementario en una célula o tejido particular. Los oligonucleótidos comprenden partes de una secuencia de ácido nucleico que tiene aproximadamente 10 nt, 50 nt ó 100 nt de longitud, preferiblemente aproximadamente 15 nt a 30 nt de longitud. Opcionalmente, un oligonucleótido que comprende una molécula de ácido nucleico con una longitud menor de 100 nt comprenderá además al menos 6 nucleótidos contiguos de SEQ ID NOs: 2, 5, 8, 11, 14, 17, 20, 23, 26, 29, 32 ó 34 o un complemento de ésta. Los oligonucleótidos pueden sintetizarse químicamente y también pueden usarse como sondas.

30 Opcionalmente, una molécula de ácido nucleico aislada de la invención comprende una molécula de ácido nucleico que es un complemento de la secuencia de nucleótidos mostrada en SEQ ID NOs: 2, 5, 8, 11, 14, 17, 20, 23, 26, 29, 32 ó 34 o una parte de esta secuencia de nucleótidos (por ejemplo, un fragmento que puede usarse como una sonda o cebador o un fragmento que codifica una parte biológicamente activa de un polipéptido EXOX). Una molécula de ácido nucleico que es complementaria a la secuencia de nucleótidos mostrada en SEQ ID NOs: 2, 5, 8, 11, 14, 17, 20, 23, 26, 29, 32 ó 34 es una que es lo suficientemente complementaria a la secuencia de nucleótidos mostrada en SEQ ID NOs: 2, 5, 8, 11, 14, 17, 20, 23, 26, 29, 32 ó 34 que puede unirse por enlace de hidrógeno con poco o ningún error de emparejamiento a la secuencia de nucleótidos mostrada en SEQ ID NOs: 2, 5, 8, 11, 14, 17, 20, 23, 26, 29, 32 ó 34, formando de esta manera un dúplex estable.

40 Tal y como se usa en la presente memoria, el término "complementario" se refiere al emparejamiento de bases de Watson-Crick o Hoogsteen entre unidades de nucleótidos de una molécula de ácido nucleico. El término "unión" significa la interacción física o química entre dos polipéptidos o compuestos o polipéptidos o compuestos asociados o combinaciones de éstos. La unión incluye interacciones iónicas, no iónicas, de van der Waals, hidrofóbicas, y semejantes. Una interacción física puede ser bien directa o indirecta. Las interacciones indirectas pueden ser a través o debidas a los efectos de otro polipéptido o compuesto. La unión directa se refiere a las interacciones que no tienen lugar a través, o debidas a, el efecto de otro polipéptido o compuesto, pero que en su lugar carecen de otros intermedios químicos sustanciales.

45 Los fragmentos proporcionados en la presente memoria se definen como secuencias de al menos 6 (contiguos) ácidos nucleicos o al menos 4 (contiguos) aminoácidos, una longitud suficiente como para permitir la hibridación específica en el caso de los ácidos nucleicos o el reconocimiento específico de un epítipo en el caso de aminoácidos, respectivamente, y son como mucho alguna parte menor de una secuencia de longitud completa. Los fragmentos pueden obtenerse de cualquier parte contigua de una secuencia de ácido nucleico o aminoácidos elegida. Los derivados son secuencias de ácido nucleico o secuencias de aminoácidos formadas a partir de los compuestos nativos bien directamente o por modificación o sustitución parcial. Los análogos son secuencias de ácido nucleico o secuencias de aminoácidos que tienen una estructura similar a, pero no idéntica a, el compuesto nativo pero se diferencian de éste respecto a determinados componentes o cadenas laterales. Los análogos pueden ser sintéticos o de un origen evolutivo diferente y pueden tener una actividad metabólica similar u opuesta comparado con el tipo salvaje. Los homólogos u ortólogos son secuencias de ácido nucleico o secuencias de aminoácidos de un gen particular que se obtienen de diferentes especies.

60 Los derivados y análogos pueden tener longitud completa u otra distinta de longitud completa, si el derivado o análogo contiene un ácido nucleico o aminoácido modificado, como se describe más adelante. Los derivados o análogos de los ácidos nucleicos o proteínas incluyen, pero no están limitados a, moléculas que comprenden regiones que son sustancialmente homólogas a los ácidos nucleicos o proteínas de la invención, en varias

realizaciones, con al menos aproximadamente 70%, 80% ó 95% de identidad (con una identidad preferida de 80-95%) sobre una secuencia de ácido nucleico o aminoácidos de tamaño idéntico o cuando se compara con una secuencia alineada en la que el alineamiento se hace por un programa informático de homología conocido en la técnica, o cuyo ácido nucleico codificador es capaz de hibridar con el complemento de una secuencia que codifica las proteínas mencionadas anteriormente en condiciones astringentes, moderadamente astringentes o de baja astringencia. Véase, por ejemplo, Ausubel *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, Nueva York, NY, 1993, y más adelante.

Una “secuencia de ácido nucleico homóloga” o “secuencia de aminoácido homóloga”, o variaciones de éstas, se refieren a secuencias caracterizadas por una homología a nivel de nucleótidos o nivel de aminoácidos como se ha discutido anteriormente. Las secuencias de nucleótidos homólogas codifican las secuencias que codifican las isoformas de polipéptidos EXOX. Las isoformas pueden expresarse en el mismo organismo como resultado, por ejemplo, de corte y empalme alternativo del ARN. Alternativamente, las isoformas pueden estar codificadas por genes diferentes. En la descripción, las secuencias de nucleótidos homólogas pueden incluir secuencias de nucleótidos que codifican un polipéptido EXOX de especies distintas de los hongos. Las secuencias de nucleótidos homólogas también incluyen, pero no están limitadas a, variaciones y mutaciones alélicas naturales de las secuencias de nucleótidos mostradas en la presente memoria. Las secuencias de ácido nucleico homólogas incluyen las secuencias de ácido nucleico que codifican sustituciones de aminoácidos conservativas (véase más adelante) en SEQ ID NOs: 2, 5, 8, 11, 14, 17, 20, 23, 26, 29, 32 ó 34, así como un polipéptido que posee actividad biológica EXOX. Más adelante se describen varias actividades biológicas de las proteínas EXOX.

Un polipéptido EXOX está codificado por el maco de lectura abierto (“ORF”) de un ácido nucleico EXOX. Una cadena de ácidos nucleicos que comprende un ORF no está interrumpida por un codón de parada. Un ORF que representa la secuencia codificadora para una proteína completa empieza con un codón de “inicio” ATG y termina con uno de los tres codones de “parada”, concretamente, TAA, TAG o TGA. Para los propósitos de esta invención, un ORF puede ser cualquier parte de una secuencia codificadora, con o sin un codón de inicio, un codón de parada, o ambos. Para que un ORF se considere como un buen candidato para codificar una proteína celular de manera fiable, se establece habitualmente un requerimiento de tamaño mínimo, por ejemplo, una cadena de ADN que codificaría una proteína de 50 aminoácidos o más.

Las secuencias de nucleótidos determinadas a partir de la clonación de los genes EXOX fúngicos permite la generación de sondas y cebadores diseñados para usarse en la identificación y/o clonación de homólogos EXOX en otras especies, así como homólogos EXOX de otros hongos. La sonda/cebador típicamente comprende un oligonucleótido sustancialmente purificado. El oligonucleótido típicamente comprende una región de secuencia de nucleótido que hibrida en condiciones astringentes con al menos aproximadamente 12, 25, 50, 100, 150, 200, 250, 300, 350 ó 400 secuencia de nucleótido de cadena con sentido consecutivas de SEQ ID NOs: 2, 5, 8, 11, 14, 17, 20, 23, 26, 29, 32 ó 34; o una secuencia de nucleótido de cadena antisentido de SEQ ID NOs: 2, 5, 8, 11, 14, 17, 20, 23, 26, 29, 32 ó 34; o un mutante natural de SEQ ID NOs: 2, 5, 8, 11, 14, 17, 20, 23, 26, 29, 32 ó 34.

“Un polipéptido que tiene una parte biológicamente activa de un polipéptido EXOX” se refiere a polipéptidos que presentan una actividad similar, pero no necesariamente idéntica a, una actividad de un polipéptido que incluye formas maduras, como se mide en un ensayo biológico particular, con o sin dependencia de la dosis. Un fragmento de ácido nucleico que codifica una “parte biológicamente activa de EXOX” puede prepararse aislando una parte de SEQ ID NOs: 2, 5, 8, 11, 14, 17, 20, 23, 26, 29, 32 ó 34 que codifica un polipéptido que tiene una actividad biológica EXOX (las actividades biológicas de las proteínas EXOX se describen más adelante), expresando la parte codificada de la proteína EXOX (por ejemplo, por expresión recombinante *in vitro*) y evaluando la actividad de la parte codificada de EXOX.

#### **Variantes de Ácido Nucleico y Polipéptido EXOX**

Además, se describen moléculas de ácido nucleico que se diferencian de las secuencias de nucleótidos mostradas en SEQ ID NOs: 2, 5, 8, 11, 14, 17, 20, 23, 26, 29, 32 ó 34 debido a la degeneración del código genético y así codifican las mismas proteínas EXOX que están codificadas por las secuencias de nucleótido mostradas en SEQ ID NOs: 2, 5, 8, 11, 14, 17, 20, 23, 26, 29, 32 ó 34. Una molécula de ácido nucleico aislada tiene una secuencia de nucleótidos que codifica una proteína que tiene una secuencia de aminoácidos mostrada en SEQ ID NOs: 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33 ó 35. Además de las secuencias de nucleótidos EXOX fúngicas mostradas en SEQ ID NOs: 2, 5, 8, 11, 14, 17, 20, 23, 26, 29, 32 y 34, los expertos en la técnica apreciarán que pueden existir polimorfismos de la secuencia de ADN que dan lugar a cambios en las secuencias de aminoácidos de los polipéptidos EXOX en una población de varias especies. Dichos polimorfismos genéticos en los genes EXOX pueden existir entre las especies fúngicas individuales en una población debido a la variación alélica natural. Tal y como se usa en la presente memoria, los términos “gen” y “gen recombinante” se refieren a moléculas de ácido nucleico que comprenden un marco de lectura abierto (ORF) que codifica una proteína EXOX, preferiblemente una proteína EXOX fúngica. Dichas variaciones alélicas naturales pueden resultar típicamente en 1-5% de variación en la secuencia de nucleótidos de los genes EXOX.

Las moléculas de ácido nucleico correspondientes a variantes y homólogos alélicos naturales de los ADNc EXOX pueden aislarse tomando como base su homología con los ácidos nucleicos EXOX fúngicos descritos en la presente

memoria usando los ADNc fúngicos, o una parte de éstos, como sonda de hibridación según las técnicas de hibridación estándar en condiciones de hibridación astringentes.

De acuerdo con esto, una molécula de ácido nucleico aislada de la invención tiene una longitud de al menos 6 nucleótidos e hibrida en condiciones astringentes con la molécula de ácido nucleico que comprende la secuencia de nucleótidos de SEQ ID NOs: 2, 5, 8, 11, 14, 17, 20, 23, 26, 29, 32 ó 34.

Opcionalmente, el ácido nucleico tiene una longitud de al menos 10, 25, 50, 100, 250, 500, 750, 1.000, 1.500 ó 2.000 o más nucleótidos. Alternativamente, una molécula de ácido nucleico aislada hibrida con la región codificadora. Tal y como se usa en la presente memoria, el término "hibrida en condiciones astringentes" se pretende que describa las condiciones para la hibridación y lavado en las que las secuencias de nucleótidos al menos 60% homólogas entre sí típicamente permanecen hibridadas entre sí.

Los homólogos u otras secuencias relacionadas (por ejemplo, ortólogos, parálogos) pueden obtenerse por hibridación con baja, moderada o alta astringencia con toda o parte de la secuencia fúngica particular como una sonda usando los métodos muy conocidos en la técnica para la hibridación y clonación de ácidos nucleicos.

Tal y como se usa en la presente memoria, la frase "condiciones de hibridación astringentes" se refiere a condiciones en las que una sonda, cebador u oligonucleótido hibridará con su secuencia diana, pero no con otras secuencias. Las condiciones astringentes son dependientes de la secuencia y serán diferentes en diferentes circunstancias. Las secuencias más largas hibridan específicamente a temperaturas más altas que las secuencias más cortas. Generalmente, las condiciones astringentes se seleccionan para ser aproximadamente 5°C menores que el punto de fusión térmico ( $T_m$ ) para la secuencia específica a una fuerza iónica y pH definidos. La  $T_m$  es la temperatura (en fuerza iónica, pH y concentración de ácido nucleico definidos) a la que el 50% de las sondas complementarias a la secuencia diana hibridan con la secuencia diana en equilibrio. Como las secuencias diana están presentes generalmente en exceso, a  $T_m$ , el 50% de las sondas están ocupadas en equilibrio. Típicamente, las condiciones astringentes serán aquellas en las que la concentración de sal es menor de aproximadamente 1,0 M ión sodio, típicamente aproximadamente 0,01 a 1,0 M ión sodio (u otras sales) a pH 7,0 a 8,3 y la temperatura es al menos aproximadamente 30°C para sondas, cebadores u oligonucleótidos cortos (por ejemplo, 10 nt a 50 nt) y al menos aproximadamente 60°C para sondas, cebadores y oligonucleótidos más largos. Las condiciones astringentes también pueden conseguirse con la adición de agentes desestabilizantes, tal como formamida.

Las condiciones astringentes son conocidas para los expertos en la técnica y pueden encontrarse en Ausubel *et al.*, (eds.), CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, NY (1989), 6.3.1-6.3.6. Preferiblemente, las condiciones son tales que las secuencias al menos aproximadamente 65%, 70%, 75%, 85%, 90%, 95%, 98% ó 99% homólogas entre sí típicamente permanecen hibridadas entre sí. Un ejemplo no limitativo de condiciones de hibridación astringentes son hibridación en una tampón con alta sal que comprende 6X SSC, 50 mM Tris-HCl (pH 7,5), 1 mM EDTA, 0,02% PVP, 0,02% Ficoll, 0,02% BSA, y 500 mg/ml de ADN de esperma de salmón desnaturalizado a 65°C, seguido de uno o más lavados en 0,2X SSC, 0,01% BSA a 50°C. Una molécula de ácido nucleico aislada de la invención que hibrida en condiciones astringentes con las secuencias de SEQ ID NOs: 2, 5, 8, 11, 14, 17, 20, 23, 26, 29, 32 ó 34 corresponde a una molécula de ácido nucleico natural. Tal y como se usa en la presente memoria, una molécula de ácido nucleico "natural" se refiere a una molécula de ARN o ADN que tiene una secuencia de nucleótidos que ocurre en la naturaleza (por ejemplo, codifica una proteína natural).

Se proporciona una secuencia de ácido nucleico que puede hibridar con la molécula de ácido nucleico que comprende la secuencia de nucleótidos de SEQ ID NOs: 2, 5, 8, 11, 14, 17, 20, 23, 26, 29, 32 ó 34 o fragmentos, análogos o derivados de éstos, en condiciones de astringencia moderada. Un ejemplo no limitativo de condiciones de hibridación de astringencia moderada son hibridación en 6X SSC, 5X disolución de Denhardt, 0,5% SDS y 100 mg/ml de ADN de esperma de salmón desnaturalizado a 55°C, seguido de uno o más lavados en 1X SSC, 0,1% SDS a 37°C. Otras condiciones de astringencia moderada que pueden usarse son muy conocidas en la técnica. Véase, por ejemplo, Ausubel *et al.*, (eds.), 1993, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, NY y Kriegler, 1990; GENE TRANSFER AND EXPRESSION, A LABORATORY MANUAL, Stockton Press, NY.

Se proporciona un ácido nucleico que puede hibridar con la molécula de ácido nucleico que comprende las secuencias de nucleótidos de SEQ ID NOs: 2, 5, 8, 11, 14, 17, 20, 23, 26, 29, 32 ó 34 o fragmentos, análogos o derivados de éstos, en condiciones de astringencia baja. Un ejemplo no limitativo de condiciones de hibridación de astringencia baja son hibridación en 35% formamida, 5X SSC, 50 mM Tris-HCl (pH 7,5), 5 mM EDTA, 0,02% PVP, 0,02% Ficoll, 0,2% BSA, 100 mg/ml de ADN de esperma de salmón desnaturalizado, 10% (p/v) de sulfato de dextrano a 40°C, seguido de uno o más lavados en 2X SSC, 25 mM Tris-HCl (pH 7,4), 5 mM EDTA y 0,1% SDS a 50°C. Otras condiciones de astringencia baja que pueden usarse son muy conocidas en la técnica (por ejemplo, como se emplean para hibridaciones cruzadas entre especies). Véase, por ejemplo, Ausubel *et al.*, (eds.), 1993, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, NY y Kriegler, 1990, GENE TRANSFER AND EXPRESSION, A LABORATORY MANUAL, Stockton Press, NY; Shilo y Weinberg, *Proc Natl Acad Sci USA* 78:6789-6792 (1981).

### Mutaciones Conservativas

Además de las variantes alélicas naturales de secuencias EXOX que pueden existir en la población, el experto en la técnica apreciará además que pueden introducirse cambios por mutación en las secuencias de nucleótidos de SEQ ID NOs: 2, 5, 8, 11, 14, 17, 20, 23, 26, 29, 32 ó 34, dando lugar de esta manera a cambios en las secuencias de aminoácidos de las proteínas EXOX codificadas, sin alterar la capacidad funcional de dichas proteínas EXOX. Por ejemplo, las sustituciones de nucleótidos que dan lugar a sustituciones de aminoácidos en restos de aminoácidos “no esenciales” pueden hacerse en la secuencia de SEQ ID NOs: 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33 ó 35. Un resto de aminoácido “no esencial” es un resto que puede alterarse de las secuencias de tipo salvaje de las proteínas EXOX sin alterar su actividad biológica, mientras que un resto de aminoácido “esencial” se requiere para dicha actividad biológica.

Tal y como se usa en la presente memoria, el término “actividad biológica” o “actividad funcional” se refiere a la función natural o normal de las proteínas EXO, por ejemplo la capacidad de degradar otras proteínas. Se predice que los restos de aminoácidos que están conservados entre las proteínas EXOX son particularmente no adecuados para alteración. Los aminoácidos para los que pueden hacerse sustituciones conservativas son muy conocidos en la técnica. Un experto en la técnica reconocerá que cada codón en un ácido nucleico (excepto AUG, que es habitualmente el único codón para metionina) puede modificarse para rendir una molécula funcionalmente idéntica por técnicas estándar. Además, las sustituciones, deleciones o adiciones individuales que alteran, añaden o delecionan un único aminoácido o un pequeño porcentaje de aminoácidos (típicamente menos de 5%, más típicamente menos de 1%) en una secuencia codificada son “mutaciones conservativas” en las que las alteraciones resultan en la sustitución de un aminoácido con un aminoácido químicamente similar.

Otro aspecto de la descripción está relacionado con moléculas de ácido nucleico que codifican proteínas EXOX que contienen cambios en los restos de aminoácidos que no son esenciales para la actividad. Dichas proteínas EXOX se diferencian en la secuencia de aminoácidos de SEQ ID NOs: 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33 ó 35 aunque siguen reteniendo la actividad biológica. En un aspecto, la molécula de ácido nucleico aislada comprende una secuencia de nucleótidos que codifica una proteína, en el que la proteína comprende una secuencia de aminoácidos al menos aproximadamente 45% homóloga a las secuencias de aminoácidos de SEQ ID NOs: 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33 ó 35. Preferiblemente, la proteína codificada por la molécula de ácido nucleico es al menos aproximadamente 60% homóloga a SEQ ID NOs: 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33 ó 35; más preferiblemente al menos aproximadamente 70% homóloga a SEQ ID NOs: 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33 ó 35; aún más preferiblemente al menos aproximadamente 80% homóloga a SEQ ID NOs: 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33 ó 35; incluso más preferiblemente al menos aproximadamente 90% homóloga a SEQ ID NOs: 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33 ó 35; y lo más preferiblemente al menos aproximadamente 95% homóloga a SEQ ID NOs: 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33 ó 35.

Una molécula de ácido nucleico aislada que codifica una proteína EXOX homóloga a la proteína de SEQ ID NOs: 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33 ó 35 puede crearse introduciendo una o más sustituciones, adiciones o deleciones de nucleótidos en la secuencia de nucleótidos de SEQ ID NOs: 2, 5, 8, 11, 14, 17, 20, 23, 26, 29, 32 ó 34 de manera que se introducen una o más sustituciones, adiciones o deleciones de aminoácidos en la proteína codificada.

Las mutaciones pueden introducirse en SEQ ID NOs: 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33 ó 35 por técnicas estándar, tal como mutagénesis dirigida a sitio, mutagénesis mediada por PCR y transposición de ADN. Preferiblemente, las sustituciones conservativas de aminoácidos se hacen en uno o más restos de aminoácidos predichos, no esenciales. Las sustituciones de una sola base están entre los cambios más comunes en el ADN humano. Estos cambios de bases pueden ocurrir en las regiones codificadoras o no codificadoras del ADN. Si ocurren en la región codificadora, pueden ser sustituciones conservativas o no conservativas. Una “sustitución conservativa de aminoácido” es un nuevo aminoácido que tiene propiedades similares y es una en la que el resto de aminoácidos se reemplaza con un resto de aminoácido que tiene una cadena lateral similar. Las sustituciones no conservativas se refieren a un nuevo aminoácido, que tiene propiedades diferentes. Las familias de restos de aminoácidos que tienen cadenas laterales similares se han definido en la técnica. Estas familias incluyen los aminoácidos con cadenas laterales básicas (por ejemplo, lisina, arginina, histidina), cadenas laterales ácidas (por ejemplo, ácido aspártico, ácido glutámico), cadenas laterales polares no cargadas (por ejemplo, glicina, asparagina, glutamina, serina, treonina, tirosina, cisteína), cadenas laterales no polares (por ejemplo, alanina, valina, leucina, isoleucina, prolina, hidroxiprolina, fenilalanina, metionina, triptófano), cadenas laterales ramificadas beta (por ejemplo, treonina, valina, isoleucina) y cadenas laterales aromáticas (por ejemplo, tirosina, fenilalanina, triptófano, histidina). Así, para una sustitución conservativa, un resto predicho de aminoácido no esencial en la proteína EXOX se reemplaza con otro resto de aminoácido de la misma familia de cadenas laterales. Alternativamente, en otra realización, las mutaciones pueden introducirse aleatoriamente a lo largo de toda o parte de una secuencia codificadora EXOX, tal como por mutagénesis de saturación, y los mutantes resultantes pueden cribarse para actividad biológica EXOX para identificar los mutantes que retienen la actividad. Después de la mutagénesis de SEQ ID NOs: 2, 5, 8, 11, 14, 17, 20, 23, 26, 29, 32 ó 34, la proteína codificada puede expresarse por cualquier tecnología recombinante conocida en la técnica y puede determinarse la actividad de la proteína.

El grado de relación de las familias de aminoácidos también puede determinarse tomando como base las interacciones de las cadenas laterales. Los aminoácidos sustituidos pueden ser restos “fuertes” totalmente

conservados o restos “débiles” totalmente conservados, El grupo “fuerte” de los restos de aminoácidos conservados puede ser uno cualquiera de los grupos siguientes: STA, NEQK, NHQK, NDEQ, QHRK, MILV, MILF, HY, FYW, en el que los códigos de aminoácidos de una letra están agrupados para los aminoácidos que pueden ser sustituidos entre sí. Asimismo, el grupo “débil” de los restos conservados puede ser uno cualquiera de los siguientes: CSA, ATV, SAG, STNK, STPA, SGND, SNDEQK, NDEQHK, NEQHRK, HFY, en el que las letras en cada grupo representan el código de aminoácidos de una letra.

La proteína EXOX mutante puede ensayarse para (i) la capacidad de formar interacciones proteína:proteína con otras proteínas EXOX, otras proteínas de la superficie celular, o partes biológicamente activas de éstas, (ii) formación de complejos entre una proteína EXOX mutante y un ligando EXOX; o (iii) la capacidad de una proteína EXOX mutante de unirse a una proteína diana intracelular o parte biológicamente activa de ésta; (por ejemplo, proteínas avidina).

En otra realización más, una proteína EXOX mutante puede ensayarse para la capacidad de regular una función biológica específica (por ejemplo, actividad proteolítica).

### Polipéptidos EXOX

Un polipéptido como se describe incluye un polipéptido que incluye la secuencia de aminoácidos de polipéptidos EXOX cuyas secuencias se proporcionan en SEQ ID NOs: 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33 y 35. La descripción también enseña una proteína mutante o variante de la que cualquiera de sus restos puede cambiarse de los restos correspondientes mostrados en SEQ ID NOs: 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33 ó 35 codificando todavía una proteína que mantiene sus actividades EXOX y funciones fisiológicas, o un fragmento funcional de ésta.

En general, una variante EXOX que conserva la función semejante a EXOX incluye cualquier variante en la que los restos en una posición particular en la secuencia se han sustituido por otros aminoácidos e incluye además la posibilidad de insertar un resto o restos adicionales entre dos restos de la proteína parental así como la posibilidad de delecionar uno o más restos de la secuencia parental.

Un aspecto de la descripción está relacionado con proteínas EXOX aisladas y partes biológicamente activas de éstas, o derivados, fragmentos, análogos u homólogos de éstas. Las partes biológicamente activas se refieren a regiones de las proteínas EXOX que son necesarias para la función normal, por ejemplo, actividad aminopeptidasa. También se proporcionan fragmentos de polipéptido adecuados para usarse como inmunógenos para producir anticuerpos anti-EXOX. En un aspecto, las proteínas EXOX nativas pueden aislarse de células, fuentes de tejidos o sobrenadantes de cultivos por un esquema de purificación apropiado usando técnicas de purificación de proteínas apropiadas. En otro aspecto, las proteínas EXOX se producen por técnicas de ADN recombinantes. Alternativamente a la expresión recombinante, una proteína o polipéptido EXOX puede sintetizarse químicamente usando técnicas de síntesis de péptidos estándar.

Un polipéptido o proteína o parte biológicamente activa de éste “aislado” o “purificado” no contiene sustancialmente material celular u otras proteínas contaminantes de la fuente de células o tejido del que se obtiene la proteína EXOX, o no contiene sustancialmente precursores químicos u otros productos químicos cuando se sintetiza químicamente. El lenguaje “no contiene sustancialmente material celular” incluye preparaciones de proteínas EXOX en las que la proteína se separa de los componentes celulares de las células de la que se aísla o se produce por recombinación. En una realización, el lenguaje “no contiene sustancialmente material celular” incluye preparaciones de proteínas EXOX que tienen menos de aproximadamente 30% (en peso seco) de proteínas no EXOX (también referidas en la presente memoria como “proteína contaminante”), más preferiblemente menos de aproximadamente 20% de proteínas no EXOX, aún más preferiblemente menos de aproximadamente 10% de proteínas no EXOX y lo más preferiblemente menos de aproximadamente 5% de proteínas no EXOX. Cuando la proteína EXOX o la parte biológicamente activa de ésta se produce por recombinación, preferiblemente tampoco contiene sustancialmente ningún constituyente del medio de cultivo, por ejemplo, los componentes del medio de cultivo pueden representar menos de aproximadamente 20%, más preferiblemente menos de aproximadamente 10%, y lo más preferiblemente menos de aproximadamente 5%, de la preparación de proteína EXOX.

El lenguaje “no contiene sustancialmente precursores químicos u otros productos químicos” incluye preparaciones de proteínas EXOX en las que la proteína se separa de los precursores químicos u otros productos químicos que están implicados en la síntesis de la proteína. En un aspecto, el lenguaje “no contiene sustancialmente precursores químicos u otros productos químicos” incluye preparaciones de proteínas EXOX que tienen menos de aproximadamente 30% (en peso seco) de precursores químicos o productos químicos no EXOX, más preferiblemente menos de aproximadamente 20% de precursores químicos o productos químicos no EXOX, aún más preferiblemente menos de aproximadamente 10% de precursores químicos o productos químicos no EXOX y lo más preferiblemente menos de aproximadamente 5% de precursores químicos o productos químicos no EXOX. Además, “no contiene sustancialmente precursores químicos u otros productos químicos” incluiría subproductos de oxidación. Un experto en la técnica sabrá cómo evitar la oxidación, por ejemplo, manteniendo los productos químicos en un entorno sin oxígeno.

Las partes biológicamente activas de las proteínas EXOX incluyen péptidos que comprenden secuencias de aminoácidos lo suficientemente homólogas con u obtenidas a partir de las secuencias de aminoácidos de las proteínas EXOX (por ejemplo, la secuencia de aminoácidos mostrada en SEQ ID NOs: 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33 ó 35) que incluye menos aminoácidos que las proteínas EXOX de longitud completa y que presentan al menos una actividad de una proteína EXOX. Típicamente, las partes biológicamente activas comprenden un dominio o resto con al menos una actividad de la proteína EXOX. Una parte biológicamente activa de una proteína EXOX puede ser un polipéptido que tiene, por ejemplo, una longitud de 10, 25, 50, 100 o más restos de aminoácidos.

Además, otras partes biológicamente activas, en las que otras regiones de la proteína están delecionadas, pueden prepararse por técnicas recombinantes y evaluarse para una o más de las actividades funcionales de una proteína EXOX nativa.

La proteína EXOX tiene una secuencia de aminoácidos mostrada en SEQ ID NOs: 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33 ó 35. La proteína EXOX es sustancialmente homóloga a SEQ ID NOs: 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33 ó 35 y retiene la actividad funcional de la proteína de SEQ ID NOs: 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33 ó 35, aunque difiera en la secuencia de aminoácidos debido a la variación alélica natural o mutagénesis, como se describe con detalle más adelante. De acuerdo con esto, la proteína EXOX es una proteína que comprende una secuencia de aminoácidos al menos aproximadamente 90% homóloga a la secuencia de aminoácidos SEQ ID NOs: 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33 ó 35 y retiene la actividad funcional de las proteínas EXOX de SEQ ID NOs: 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33 ó 35. Tal y como se usa en la presente memoria, el término "actividad biológica" o "actividad funcional" se refiere a la función natural o normal de las proteínas EXOX, por ejemplo la capacidad de degradar otras proteínas.

### Determinación de la Homología Entre Dos o Más Secuencias

Para determinar el porcentaje de similitud u homología de dos secuencias de aminoácidos o de dos secuencias de ácido nucleico, las secuencias se alinean para propósitos de comparación óptima (por ejemplo, pueden introducirse huecos en la secuencia de una primera secuencia de aminoácidos o ácido nucleico para el alineamiento óptimo con una segunda secuencia de aminoácidos o ácido nucleico). Los restos de aminoácidos o nucleótidos en posiciones de aminoácidos o posiciones de nucleótidos correspondientes se comparan. Cuando una posición en la primera secuencia está ocupada por el mismo resto de aminoácido o nucleótido como la posición correspondiente en la segunda secuencia, las moléculas son homólogas en esa posición (es decir, tal y como se usa en la presente memoria la "homología" de aminoácidos o ácido nucleico es equivalente a "identidad" de aminoácidos o ácido nucleico).

La homología de la secuencia de ácido nucleico puede determinarse como el grado de identidad entre dos secuencias. La homología puede determinarse usando programas informáticos conocidos en la técnica, tal como el programa informático GAP proporcionado en el paquete de programas GCG. Véase Needleman y Wunsch, *J. Mol. Biol.* 48:443-452 1970. Usando el programa informático GCG GAP con los ajustes siguientes para comparación de secuencias de ácido nucleico: penalización por creación de HUECO de 5.0 y penalización por extensión de HUECO de 0.3, la región codificadora de las secuencias de ácido nucleico análogas referidas anteriormente presenta un grado de identidad preferiblemente de al menos 70%, 75%, 80%, 85%, 90%, 95%, 98% ó 99% con la parte CDS (codificadora) de la secuencia de ADN mostrada en SEQ ID NOs: 2, 5, 8, 11, 14, 17, 20, 23, 26, 29, 32 y 34.

El término "identidad de secuencia" se refiere al grado en el que dos secuencias de polinucleótido o polipéptido son idénticas en una base resto a resto sobre una región particular de comparación. El término "porcentaje de identidad de secuencia" se calcula comparando dos secuencias alineadas de manera óptima sobre esa región de comparación, determinando el número de posiciones en las que ocurre la base de ácido nucleico idéntica (por ejemplo, A, T, C, G, U, o I, en el caso de ácidos nucleicos) en ambas secuencias para rendir el número de posiciones emparejadas, dividiendo el número de posiciones emparejadas por el número total de posiciones en la región de comparación (por ejemplo, el tamaño de la ventana) y multiplicando el resultado por 100 para rendir el porcentaje de identidad de secuencia. El término "identidad sustancial" tal y como se usa en la presente memoria indica una característica de una secuencia de polinucleótidos, en la que el polinucleótido comprende una secuencia que tiene al menos 80% de identidad de secuencia, preferiblemente al menos 85% de identidad de secuencia y habitualmente 90 a 95% de identidad de secuencia, más habitualmente al menos 99% de identidad de secuencia comparado con una secuencia de referencia sobre una región de comparación.

### Proteínas Quiméricas y de Fusión

La descripción también proporciona proteínas quiméricas o de fusión EXOX. Tal y como se usa en la presente memoria, una "proteína quimérica" o "proteína de fusión" EXOX comprende un polipéptido EXOX unido de manera operativa a un polipéptido no EXOX. Un "polipéptido EXOX" se refiere a un polipéptido que tiene una secuencia de aminoácidos correspondiente a una proteína EXOX (SEQ ID NOs: 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33 ó 35), mientras que un "polipéptido no EXOX" se refiere a un polipéptido que tiene una secuencia de aminoácidos correspondiente a una proteína que no es sustancialmente homóloga a la proteína EXOX, por ejemplo, una proteína que es diferente de la proteína EXOX y que se obtiene del mismo organismo o de un organismo diferente. En una proteína de fusión EXOX, el polipéptido EXOX puede corresponder a toda o parte de una proteína EXOX. En un

aspecto, una proteína de fusión EXOX comprende al menos una parte biológicamente activa de una proteína EXOX. En otro aspecto, una proteína de fusión EXOX comprende al menos dos partes biológicamente activas de una proteína EXOX. En otro aspecto más, una proteína de fusión EXOX comprende al menos tres partes biológicamente activas de una proteína EXOX. En la proteína de fusión, el término "unido de manera operativa" se pretende que indique que el polipéptido EXOX y el polipéptido no EXOX están fusionados en marco entre sí. El polipéptido no EXOX puede fusionarse al extremo N-terminal y/o C-terminal del polipéptido EXOX.

En un aspecto, la proteína de fusión es una proteína de fusión GST-EXOX en la que las secuencias EXOX están fusionadas al extremo C-terminal de las secuencias GST (glutación S-transferasa). Dichas proteínas de fusión pueden facilitar la purificación de los polipéptidos EXOX recombinantes.

En otro aspecto, la proteína de fusión es una proteína EXOX que contiene una secuencia señal heteróloga en su extremo N-terminal. En determinadas células anfitrionas (por ejemplo, células anfitrionas de mamífero), la expresión y/o secreción de EXOX puede incrementarse mediante el uso de una secuencia señal heteróloga.

En otro aspecto más, la proteína de fusión es una proteína de fusión EXOX-inmunoglobulina en la que las secuencias EXOX están fusionadas con secuencias obtenidas de un miembro de la familia de proteínas inmunoglobulinas. Las proteínas de fusión EXOX-inmunoglobulina pueden incorporarse en composiciones farmacéuticas y administrarse a un sujeto para inhibir una interacción entre un ligando EXOX y una proteína EXOX en la superficie de una célula, para suprimir de esta manera la transducción de la señal mediada por EXOX *in vivo*. Las proteínas de fusión EXOX-inmunoglobulina pueden usarse para influir en la biodisponibilidad de un ligando EXOX conocido. La inhibición de la interacción ligando EXOX/EXOX puede ser útil terapéuticamente para el tratamiento de trastornos proliferativos y de diferenciación, así como para modular (por ejemplo, estimular o inhibir) la supervivencia celular. Además, las proteínas de fusión EXOX-inmunoglobulina pueden usarse como inmunógenos para producir anticuerpos anti-EXOX en un sujeto, para purificar ligandos EXOX y en ensayos de cribado para identificar moléculas que inhiben la interacción de EXOX con un ligando EXOX.

Una proteína quimérica o de fusión EXOX puede producirse por técnicas de ADN recombinante estándar. Por ejemplo, los fragmentos de ADN que codifican diferentes secuencias polipeptídicas se ligan entre sí en marco según técnicas convencionales, por ejemplo, empleando extremos romos o en bisel para la ligación, digestión con enzimas de restricción para proporcionar extremos apropiados, rellenando los extremos cohesivos según sea apropiado, tratamiento con fosfatasa alcalina para evitar uniones no deseadas y ligación enzimática. El gen de fusión puede sintetizarse por técnicas convencionales incluyendo sintetizadores de ADN automáticos. Alternativamente, la amplificación por PCR de fragmentos génicos puede realizarse usando cebadores de anclaje para proporcionar protuberancias complementarias entre dos fragmentos génicos consecutivos que pueden hibridarse posteriormente y reamplificarse para generar una secuencia génica quimérica (Véase, por ejemplo, Ausubel *et al.*, (eds.) CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, 1992). Además, muchos vectores de expresión están disponibles comercialmente que codifican ya un resto de fusión (por ejemplo, un polipéptido GST). Un ácido nucleico que codifica EXOX puede clonarse en dicho vector de expresión de manera que el resto de fusión está unido en marco con la proteína EXOX.

### Agonistas y Antagonistas EXOX

La descripción también se refiere a variantes de las proteínas EXOX que funcionan como agonistas EXOX (por ejemplo, miméticos) o como antagonistas EXOX. Las variantes de la proteína EXOX pueden generarse por mutagénesis (por ejemplo, mutación puntual o truncamiento discreto de la proteína EXOX). Un agonista de la proteína EXOX puede retener sustancialmente la misma, o un subconjunto de, las actividades biológicas de la forma natural de la proteína EXOX. Un antagonista de la proteína EXOX puede inhibir una o más de las actividades de la forma natural de la proteína EXOX, por ejemplo, uniéndose competitivamente a un miembro aguas abajo o arriba de una cascada de señalización celular, que incluye la proteína EXOX. Así, los efectos biológicos específicos pueden inducirse por tratamiento con una variante con función limitada. El tratamiento de un sujeto con una variante que tiene un subconjunto de las actividades biológicas de la forma natural de la proteína tiene menos efectos secundarios en un sujeto respecto al tratamiento con la forma natural de las proteínas EXOX.

Las variantes de las proteínas EXOX que funcionan como agonistas EXOX (por ejemplo, miméticos) o como antagonistas EXOX pueden identificarse por cribado de bibliotecas combinatorias de mutantes (por ejemplo, mutantes por truncamiento) de las proteínas EXOX para actividad agonista o antagonista de la proteína EXOX. En un aspecto, se genera una biblioteca variada de variantes EXOX por mutagénesis combinatoria a nivel de ácido nucleico y está codificada por una biblioteca de genes variados. Una biblioteca variada de variantes EXOX puede producirse, por ejemplo, ligando enzimáticamente una mezcla de oligonucleótidos sintéticos en secuencias génicas de manera que se puede expresar un conjunto degenerado de secuencias EXOX potenciales como polipéptidos individuales o, alternativamente, como un conjunto de proteínas de fusión mayores (por ejemplo, para exposición en fagos) que contiene el conjunto de secuencias EXOX en él. Existen varios métodos, que pueden usarse para producir bibliotecas de variantes EXOX potenciales a partir de una secuencia oligonucleotídica degenerada. La síntesis química de una secuencia génica degenerada puede realizarse en un sintetizador de ADN automático y el gen sintético ligarse en un vector de expresión apropiado. El uso de un conjunto de genes degenerado permite la provisión, en una mezcla, de todas las secuencias que codifican el conjunto deseado de secuencias EXOX



potenciales. Los métodos para sintetizar oligonucleótidos degenerados son muy conocidos en la técnica. Véase, por ejemplo, Narang, *Tetrahedron* 39:3 (1983); Itakura *et al.*, *Annu. Rev. Biochem.* 53:323 (1984); Itakura *et al.*, *Science* 198:1056 (1984); Ike *et al.*, *Nucl. Acids Res.* 11:477 (1983).

### Bibliotecas de Polipéptidos

5 Además, pueden usarse bibliotecas de fragmentos de las secuencias codificadoras de la proteína EXOX para generar una población variada de fragmentos EXOX para cribar y seleccionar posteriormente las variantes de una proteína EXOX. Puede generarse una biblioteca de fragmentos de secuencia codificadora tratando un fragmento de PCR bicatenario de una secuencia codificadora EXOX con una nucleasa en condiciones en las que ocurra el  
10 mellado sólo aproximadamente una vez por molécula, desnaturalizando el ADN bicatenario, renaturalizando el ADN para formar ADN bicatenario que puede incluir parejas sentido/antisentido de diferentes productos mellados, eliminando las partes monocatenarias de los dúplex reformados por tratamiento con nucleasa  $S_1$  y ligando la biblioteca de fragmentos resultante en un vector de expresión. Mediante este método, pueden obtenerse bibliotecas de expresión que codifican fragmentos N-terminales e internos de varios tamaños de las proteínas EXOX.

15 En la técnica se conocen varias técnicas para cribar productos génicos de bibliotecas combinatorias hechas por mutaciones o truncamiento puntual, y para cribar bibliotecas de ADNc para productos génicos que tienen una propiedad seleccionada. Dichas técnicas se pueden adaptar para el cribado rápido de las bibliotecas génicas generadas por la mutagénesis combinatoria de proteínas EXOX. Las técnicas usadas más ampliamente, que pueden adaptarse a análisis de alto rendimiento, para el cribado de bibliotecas génicas grandes incluyen típicamente clonar la biblioteca génica en vectores de expresión replicables, transformar células apropiadas con la biblioteca de  
20 vectores resultante, y expresar los genes combinatorios en condiciones en las que la detección de una actividad deseada facilita el aislamiento del vector que codifica el gen cuyo producto se ha detectado. La mutagénesis recursiva de conjunto (REM), una nueva técnica que incrementa la frecuencia de mutantes funcionales en las bibliotecas, puede usarse en combinación con los ensayos de cribado para identificar las variantes EXOX. Véase, por ejemplo, Arkin y Yourvan, *Proc. Natl. Acad. Sci. USA* 89:7811-7815 (1992); Delgrave *et al.*, *Protein Engineering* 6:327-331 (1993).

25 Las bibliotecas también pueden generarse por transposición de ADN. La transposición de ADN usa genes relacionados de diferentes especies o genes que están relacionados en su función, los fragmenta y reúne mediante recombinación. Entonces puede determinarse si los genes recombinados comprenden productos que se pueden usar o potencialmente interesantes. Cualquier gen recombinado que se encuentra que es útil se fragmenta de nuevo y reúne para formar nuevos genes recombinantes. Al hibridar y extenderse los diferentes fragmentos de diferentes especies y genes, se crea diversidad en la biblioteca. El proceso puede realizarse hasta que se encuentra una proteína de interés. Los factores importantes para crear genes recombinados con transposición de ADN incluyen la temperatura a la que ocurre la hibridación, la similitud de los genes y el tamaño de los fragmentos de ADN.

30 Stemmer *et al.*, *Nature* 370:389-391 (1994); Stemmer, *Proc. Natl. Acad. Sci. USA* 91:10747-10751 (1994); Pat. EEUU No. 5.603.793; Pat. EEUU No. 5.830.721; y Pat. EEUU No. 5.811.238, que se incorporan en la presente memoria por referencia, describen, por ejemplo, métodos de transposición *in vitro*, por ejemplo, por ciclos repetidos de mutagénesis, transposición y selección así como varios métodos para generar bibliotecas de péptidos expuestos y anticuerpos así como varias técnicas de unión de ADN después de la fragmentación de ADN, y su aplicación en la mutagénesis *in vitro* e *in vivo*. Además, en la técnica también se conocen varias aplicaciones de la tecnología de transposición de ADN. Además de las publicaciones indicadas anteriormente, véase Pat. EEUU No. 5.837.458, que proporciona la evolución de nuevas rutas metabólicas y el incremento del bioprocesamiento mediante técnicas de transposición recursiva, y Cramer *et al.*, *Nature Medicine* 2(1):100-103 (1996), que describe la transposición de anticuerpos para bibliotecas de fagos con anticuerpos. Véase también, WO95/22625, WO97/20078, WO96/33207, WO97/33957, WO98/27230, WO97/35966, WO98/31837, WO98/13487, WO98/13485 y WO98/42832.

### 45 Vectores de Expresión

Otro aspecto de la descripción se refiere a vectores, preferiblemente vectores de expresión, que contienen un ácido nucleico que codifica una proteína EXOX, o derivados, fragmentos, análogos u homólogos de ésta. Tal y como se usa en la presente memoria, el término "vector" se refiere a una molécula de ácido nucleico capaz de transportar otro ácido nucleico al que se ha unido. Un tipo de vector es un "plásmido", que se refiere a un bucle de ADN  
50 bicatenario circular en el que pueden ligarse segmentos de ADN adicionales. Otro tipo de vector es un vector viral, en el que segmentos de ADN adicionales pueden ligarse en el genoma viral. Determinados vectores son capaces de replicación autónoma en una célula anfitriona en la que se introducen (por ejemplo, los vectores bacterianos que tienen un origen de replicación bacteriano y vectores de mamíferos episomales). Otros vectores (por ejemplo, vectores de mamíferos no episomales) se integran en el genoma de una célula anfitriona después de la introducción en la célula anfitriona, y se replican de esta manera junto con el genoma anfitrión. Además, determinados vectores son capaces de dirigir la expresión de genes a los que están unidos de manera operativa. Dichos vectores se refieren en la presente memoria como "vectores de expresión". En general, los vectores de expresión de uso en técnicas de ADN recombinante están habitualmente en la forma de plásmidos. En la presente especificación, "plásmido" y "vector" pueden usarse indistintamente ya que el plásmido es la forma de vector usada más  
60 comúnmente. Sin embargo, la invención pretende incluir dichas otras formas de vectores de expresión, tales como

vectores virales (por ejemplo, retrovirus, adenovirus y virus adeno asociados con defectos en la replicación), que sirven para funciones equivalentes.

La producción de una proteína funcional está íntimamente relacionada con la maquinaria celular del organismo que produce la proteína. *E. coli* ha sido típicamente la “fábrica” de primera línea para la expresión de muchas proteínas porque su genoma se ha mapeado completamente y el organismo es fácil de manejar; crece rápidamente; requiere un medio de crecimiento barato y fácil de preparar; y secreta la proteína en el medio lo que facilita la recuperación de la proteína. Sin embargo, *E. coli* es un procarionte y carece de orgánulos intracelulares, tales como retículo endoplásmico y el aparato de golgi que están presentes en los eucariotas, que contienen enzimas que modifican las proteínas que se están produciendo. Muchas proteínas eucariotas pueden producirse en *E. coli* pero éstas pueden producirse de una forma no funcional, no terminada, ya que no ocurren la glicosilación ni las modificaciones posteriores a la traducción.

Por lo tanto, los investigadores han recurrido recientemente a sistemas de expresión de levaduras eucariotas, mamíferos y plantas para la producción de proteínas. Por ejemplo, la levadura metanoltrófica *P. pastoris* se ha convertido en un anfitrión potente para la expresión heteróloga de proteínas durante los últimos años y se ha establecido como un anfitrión eucariota alternativo para la expresión de proteínas humanas con tecnologías de alto rendimiento.

Como otro ejemplo, las plantas están siendo utilizadas como anfitriones de expresión para la expresión heteróloga a gran escala de proteínas y ofrecen ventajas potenciales de coste, posibilidad de escalado y seguridad sobre los sistemas de expresión tradicionales. Existen actualmente varios sistemas de expresión heterólogos de plantas incluyendo la expresión transitoria, cultivos de suspensión de células de plantas, virus de plantas recombinantes y sistemas transgénicos de cloroplastos. Aunque las proteínas expresadas en plantas tienen algunas variaciones respecto a las proteínas de mamíferos (por ejemplo, glicosilación), actualmente no hay evidencia de que estas diferencias resulten en reacciones adversas en pacientes humanos. Véase, por ejemplo, Julian *et al.*, Nat. Rev. Gen. 4:794-805 (2003).

Otro sistema de expresión heterólogo adecuado usa células de insecto, habitualmente en combinación con vectores de expresión de baculovirus. Los vectores de baculovirus disponibles para expresar proteínas en células de insecto cultivadas, por ejemplo, células Sf9, incluyen las serie pAc (Smith *et al.*, Mol. Cell. Biol. 3:2156-2165 (1983)) y la serie pVL (Lucklow y Summers, *Virology* 170:31-39 (1989)).

Las células anfitrionas también pueden usarse para producir animales transgénicos no humanos en los que se han introducido secuencias exógenas en su genoma. El animal transgénico es un animal no humano, preferiblemente un mamífero, más preferiblemente un roedor tal como una rata o un ratón, en el que una o más de las células del animal incluye un transgén. Otros ejemplos de animales transgénicos incluyen, por ejemplo, primates no humanos, ovejas, perros, vacas, cabras, pollos, anfibios. Los métodos para generar animales transgénicos mediante la manipulación del embrión y la microinyección, particularmente animales tales como ratones, se han convertido en convencionales en la técnica y se describen, por ejemplo, en las Patentes de EEUU Nos. 4.736.866; 4.870.009; y 4.873.191; y Hogan, 1986. En: MANIPULATING THE MOUSE EMBRYO, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. Métodos similares se usan para la producción de otros animales transgénicos.

#### **Sistema de Expresión de *Pichia Pastoris***

Una de estas levaduras eucariotas es la metanoltrófica *Pichia pastoris*. *P. pastoris* se ha desarrollado para convertirse en un anfitrión importante para la producción de proteínas extrañas ya que su promotor de alcohol oxidasa se aisló y clonó: La transformación de *P. pastoris* se indicó por primera vez en 1985. El sistema de expresión de proteínas heterólogas de *P. pastoris* lo desarrolló Phillips Petroleum, véase, por ejemplo, las Patentes de EEUU Nos. 4.855.231, 4.857.467, 4.879.231 y 4.929.555, cada una de las cuales se incorpora en la presente memoria por referencia. Este sistema está comercializado actualmente por Invitrogen. Comparado con otros sistemas de expresión eucariotas, *Pichia* ofrece muchas ventajas, porque no tiene el problema de endotoxinas asociado con las bacterias ni el problema de contaminación viral de proteínas producidas en cultivos de células animales. Además, *P. pastoris* puede utilizar metanol como una fuente de carbono en ausencia de glucosa. El sistema de expresión de *P. pastoris* usa el promotor de la alcohol oxidasa (AOX1) inducido por metanol, que controla el gen que codifica la expresión de la alcohol oxidasa, la enzima que cataliza la primera etapa del metabolismo del metanol. Este promotor se ha caracterizado e incorporado en una serie de vectores de expresión de *P. pastoris*. Como las proteínas producidas en *P. pastoris* se pliegan típicamente correctamente y se secretan en el medio, la fermentación de *P. pastoris* sometida a ingeniería genética proporciona una alternativa excelente a los sistemas de expresión de *E. coli*. Además, *P. pastoris* tiene la capacidad de glicosilar espontáneamente las proteínas expresadas, lo que también es una ventaja sobre *E. coli*. Se han producido varias proteínas usando este sistema, incluyendo el fragmento de la toxina del tétanos, pertactina de *Bordetella pertussis*, albúmina de suero humano y lisozima.

### Eliminación de Etiquetas con Proteínas EXOX

Se han desarrollado varios sistemas para permitir la purificación rápida y eficaz de proteínas recombinantes expresadas en bacterias. La mayoría se basa en la expresión de la proteína como una proteína de fusión con un dominio de glutatión-S-transferasa (GST), un péptido de unión a calmodulina (CBP) o una etiqueta His. Por ejemplo, la expresión de polipéptidos en marco con la glutatión S-transferasa (GST) permite la purificación de las proteínas de fusión de extractos bacterianos crudos en condiciones no desnaturizantes por cromatografía de afinidad en agarosa glutatión.

Además, este sistema de expresión con vectores incorpora generalmente un sitio de escisión específico de proteasa para facilitar la proteólisis de las proteínas de fusión bacterianas, que es, dependiendo del vector usado, un sitio de escisión de proteasa de la trombina, enteroquinasa o Factor Xa. La trombina escinde específicamente las proteínas diana que contienen la secuencia de reconocimiento Leu-Val-Pro-Arg↓Gly-Ser (SEQ ID NO: 44). El sitio de escisión de la enteroquinasa es Asp-Asp-Asp-Asp-Lys↓ (SEQ ID NO: 45). Como la enteroquinasa, el Factor Xa escinde el lado C-terminal de su secuencia de reconocimiento Ile-Glu-Gly-Arg↓ (SEQ ID NO: 46), y puede por lo tanto usarse para eliminar todas las secuencias codificadas por el vector de las construcciones diseñadas apropiadamente. Todas estas enzimas están ahora disponibles comercialmente con alta pureza para evitar la escisión secundaria que surge de las proteasas contaminantes. Estas enzimas se proporcionan en un kit que incluye habitualmente todas las herramientas para la captura de la enzima o biotiniladas para facilitar la eliminación de la enzima del medio de reacción de la escisión. Más recientemente, Qiagen también ha desarrollado el sistema TAGZyme para una eliminación eficaz de las etiquetas de His N-terminales de proteínas que implica exopeptidasas que escinden dipéptidos secuencialmente desde el extremo N-terminal hasta un resto de aminoácido que es un "punto de parada", que es ↓Lys-Xaa-, ↓Arg-Xaa-, ↓Xaa-Xaa-Pro-Xaa-, ↓Xaa-Pro-Xaa-Xaa- o ↓Gln-Xaa-.

Aunque no siempre es necesario eliminar la etiqueta corta de His de afinidad (independientemente del número de restos de His) de una proteína recombinante después de la purificación, existen algunas aplicaciones, tales como análisis estructural por cristalografía de rayos X o RMN, en las que la eliminación de la etiqueta es deseable. Lo mismo es también cierto para los restos residuales Gly-Ser del sitio de escisión de la trombina o cualquier aminoácido N-terminal residual suplementario que podría estar todavía presente y que podría estar relacionado con el sistema de expresión usado.

Un método más reciente para purificar por afinidad implica la utilización de una reacción de condensación entre un grupo carbonilo y una molécula con dos grupos nucleofílicos vecinos. Los ejemplos de aminoácidos con dos grupos nucleofílicos vecinos incluyen, por ejemplo, serina, treonina y cisteína. La purificación de una proteína o péptido implica la formación de un enlace covalente reversible formado entre, por ejemplo, un resto de cisteína, treonina o serina N-terminal y una resina apropiada. Véase Villain *et al.*, Chem. & Biol. 8:673-679 (2001). La adición de un par de restos, por ejemplo, Thr-Pro, Cys-Pro o Ser-Pro en el extremo N terminal de una proteína recombinante o de una proteína (péptido) obtenida por síntesis química, permite una purificación en dos etapas: (1) purificación por captura covalente; y (2) eliminación de la etiqueta di-péptido. Este método permite la recuperación eficaz de una proteína recombinante en su forma madura, sin la secuencia etiqueta dipéptido.

### Actividad Proteolítica Inversa de las Proteínas EXOX

Otro aspecto se refiere a métodos para añadir uno o más aminoácidos a aminoácidos, péptidos, oligopéptidos, polipéptidos o cualquier composición con una amina secundaria accesible, usando la actividad proteolítica inversa de una o más proteínas EXOX. Tal y como se usa en la presente memoria, el término "actividad proteolítica inversa" se refiere a la actividad enzimática que cataliza la adición de uno o más aminoácidos a un aminoácido, un péptido, un oligopéptido, un polipéptido o cualquier composición con una amina secundaria accesible. Un experto en la técnica reconocerá que, en condiciones termodinámicas adecuadas, las enzimas proteolíticas pueden tener actividad proteolítica inversa.

Un ejemplo de una enzima proteolítica con actividad proteolítica inversa es la tripsina, que es una serina proteasa pancreática con especificidad de sustrato basada en las cadenas laterales cargadas positivamente de lisina y arginina. La tripsina se usa ampliamente en la fabricación de la insulina humana a partir de la insulina porcina, que es similar a la forma humana excepto en que el último resto de aminoácido en la cadena B es alanina en lugar de treonina. La reacción de la insulina porcina con un éster de treonina en la presencia de tripsina rinde un éster de treonina de insulina humana eliminando la alanina terminal y añadiendo el éster de treonina. El tratamiento posterior del éster de treonina de la insulina humana con ácido trifluoroacético hidroliza el éster para rendir insulina humana.

En algunas realizaciones, las proteínas EXOX se usan para catalizar reacciones proteolíticas inversas. En algunos casos, las proteínas EXOX se incuban con un polipéptido y uno o más aminoácidos en condiciones que permiten la adición del uno o más aminoácidos al polipéptido.

Existen múltiples utilidades para usar las proteínas EXOX de la presente invención como enzimas proteolíticas inversas. Por ejemplo, la actividad proteolítica inversa de las proteínas EXOX puede usarse en la síntesis de una cadena polipeptídica. Las proteínas EXOX también pueden usarse como un agente de acoplamiento para añadir uno o más aminoácidos a otro aminoácido, un polipéptido o cualquier composición con una amina secundaria accesible.

### Composiciones Farmacéuticas

Las moléculas de ácido nucleico EXOX, las proteínas EXOX y los anticuerpos anti-EXOX (también referidos en la presente memoria como “compuestos activos”), y derivados, fragmentos, análogos y homólogos de éstos, pueden incorporarse en composiciones farmacéuticas adecuadas para administración. Dichas composiciones comprenden típicamente la molécula de ácido nucleico, proteína o anticuerpo y un vehículo farmacéuticamente aceptable. Tal y como se usa en la presente memoria “vehículo farmacéuticamente aceptable” se pretende que incluya cualquiera y todos los disolventes, medios de dispersión, recubrimientos, agentes antibacterianos y antifúngicos, agentes isotónicos y que retrasan la absorción, y semejantes, compatibles con la administración farmacéutica. Los vehículos adecuados se describen en la edición más reciente de Remington’s Pharmaceutical Sciences, un texto de referencia estándar en el campo, que se incorpora en la presente memoria por referencia. Los ejemplos preferidos de dichos vehículos o diluyentes incluyen, pero no están limitados a, agua, disolución salina, disoluciones de Ringer, disolución de dextrosa y albúmina de suero humana al 5%. Los liposomas y los vehículos no acuosos tal como aceites fijados también pueden usarse. El uso de dichos medios y agentes para sustancias farmacéuticamente activas es muy conocido en la técnica. Excepto en el momento, como cualquier medio o agente convencional, en que sea incompatible con el compuesto activo, se contempla el uso de éstos en las composiciones. En las composiciones también pueden incorporarse compuestos activos suplementarios.

Las tecnologías de encapsulación también se aplican ampliamente en muchas industrias. Los ejemplos incluyen productos farmacéuticos para la liberación controlada de los fármacos; pigmentos en alimento y bebidas; antioxidantes en alimentos; y la liberación controlada de feromonas de insecto en la agricultura. Las cápsulas, microcápsulas y microesferas son partículas esféricas pequeñas, que contienen un ingrediente activo en la matriz de la partícula o unido a la superficie de la partícula. Por ejemplo, se ha mostrado la encapsulación en micropartículas de alginato biodegradables. Las tecnologías de bioencapsulación están dirigidas a encapsular células, enzimas y materiales biológicamente activos.

Una composición farmacéutica de la invención se formula para ser compatible con su ruta pretendida de administración. Los ejemplos de rutas de administración incluyen administración parenteral, por ejemplo, intravenosa, intradérmica, subcutánea, oral (por ejemplo, inhalación), transdérmica (por ejemplo, tópica), transmucosal, y rectal. Las disoluciones o suspensiones usadas para aplicación parenteral, intradérmica o subcutánea pueden incluir los componentes siguientes: un diluyente estéril tal como agua para inyección, disolución salina, aceites fijados, polietilén glicoles, glicerina, propilén glicol u otros disolventes sintéticos; agentes antibacterianos tal como alcohol bencílico o metil parabenos; antioxidantes tales como ácido ascórbico o bisulfito sódico; agentes quelantes tal como ácido etilendiaminotetraacético (EDTA); tampones tales como acetatos, citratos o fosfatos, y agentes para el ajuste de la tonicidad tal como cloruro de sodio o dextrosa. El pH puede ajustarse con ácidos o bases, tal como ácido clorhídrico o hidróxido de sodio. La preparación parenteral puede incluirse en ampollas, jeringas desechables o viales de dosis múltiple hechos de vidrio o plástico.

Las composiciones farmacéuticas adecuadas para uso inyectable incluyen disoluciones acuosas estériles (solubles en agua) o dispersiones y polvos estériles para la preparación extemporánea de disoluciones inyectables estériles o dispersión. Para la administración intravenosa, los vehículos adecuados incluyen disolución salina fisiológica, agua bacteriostática, Cremophor EL™ (BASF, Parsippany, N.J.) o disolución salina tamponada con fosfato (PBS). En todos los casos, la composición debe ser estéril y debe ser fluida hasta el punto de que exista la posibilidad fácil de ponerla en jeringas. Debe ser estable en las condiciones de la fabricación y almacenamiento y debe protegerse frente a la contaminación por microorganismos, tales como bacterias, hongos o virus. El vehículo puede ser un disolvente o medio de dispersión que contiene, por ejemplo, agua, etanol, poliol (por ejemplo, glicerol, propilén glicol y polietilén glicol líquido y semejantes) y mezclas adecuadas de éstos. La fluidez apropiada puede mantenerse, por ejemplo, por el uso de un recubrimiento tal como lecitina, por el mantenimiento del tamaño de partícula requerido en el caso de dispersión y por el uso de tensioactivos. La prevención de la acción de microorganismos puede conseguirse con varios agentes antibacterianos y antifúngicos, por ejemplo, parabenos, clorobutanol, fenol, ácido ascórbico, timerosal, y semejantes. En muchos casos, será preferible incluir agentes isotónicos, por ejemplo, azúcares, polialcoholes tales como manitol, sorbitol o cloruro de sodio en la composición. La absorción prolongada de las composiciones inyectables puede conseguirse incluyendo en la composición un agente que retrase la absorción, por ejemplo, monoestearato de aluminio y gelatina.

Las disoluciones inyectables estériles pueden prepararse incorporando el compuesto activo (por ejemplo, una proteína EXOX o anticuerpo anti-EXOX) en la cantidad requerida en un disolvente apropiado con uno o una combinación de ingredientes enumerados anteriormente, según se requiera, seguido de esterilización por filtración. Generalmente, las dispersiones se preparan incorporando el compuesto activo en un vehículo estéril que contiene un medio de dispersión básico y los demás ingredientes requeridos de los enumerados anteriormente. En el caso de polvos estériles para la preparación de disoluciones inyectables estériles, los métodos de preparación son secado en vacío y liofilización que rinde un polvo del ingrediente activo más cualquier ingrediente deseado adicional de una disolución previamente esterilizada por filtración de éste.

Una preparación cruda de medio de cultivo celular de *T. rubrum* u hongos transgénicos que producen EXOX o EXOX purificada de *T. rubrum* u hongos transgénicos que producen EXOX puede administrarse oralmente porque las proteasas se secretan. Las composiciones orales incluyen generalmente un diluyente inerte o un vehículo

comestible. Pueden incluirse en cápsulas de gelatina o comprimirse en comprimidos. Para el propósito de la administración terapéutica oral, el compuesto activo puede incorporarse con excipientes y usarse en la forma de comprimidos, troches o cápsulas. Las composiciones orales también pueden prepararse usando un vehículo fluido para uso como un lavado bucal, en el que el compuesto en el vehículo fluido se aplica oralmente y se enjuaga y expectora o se traga. Los agentes de aglutinación farmacéuticamente compatibles y/o materiales adyuvantes pueden incluirse como parte de la composición. Los comprimidos, pastillas, cápsulas, troches y semejantes pueden contener cualquiera de los ingredientes siguientes, o compuestos de naturaleza similar: un aglutinante tal como celulosa microcristalina, goma de tragacanto o gelatina; un excipiente tal como almidón o lactosa, un agente disgregante tal como ácido algínico, Primogel, o almidón de maíz; un lubricante tal como estearato de magnesio o Sterotes; un deslizante tal como dióxido de silicio coloidal; un agente edulcorante tal como sacarosa o sacarina; o un agente saporífero tal como menta, salicilato de metilo o sabor a naranja.

Para la administración por inhalación, los compuestos se administran en la forma de un pulverizador en aerosol desde un contenedor o dispensador presurizado, que contiene un propelente adecuado, por ejemplo un gas tal como dióxido de carbono o un nebulizador.

La administración sistémica también puede ser por medios transmucosales o transdérmicos. Para la administración transmucosal o transdérmica, se usan en la formulación penetrantes apropiados para la barrera que se quiere permear. Dichos penetrantes se conocen generalmente en la técnica, e incluyen, por ejemplo, para administración transmucosal, detergentes, sales biliares y derivados del ácido fusídico. La administración transmucosal puede conseguirse mediante el uso de pulverizadores nasales o supositorios. Para la administración transdérmica, los compuestos activos se formulan en pomadas, ungüentos, geles o cremas como se conoce generalmente en la técnica.

Los compuestos también pueden prepararse en la forma de supositorios (por ejemplo, con bases de supositorio convencionales tal como manteca de cacao u otros glicéridos) o enemas de retención para administración rectal.

En un aspecto, los compuestos activos se preparan con vehículos que protegerán el compuesto frente a la eliminación rápida del cuerpo, tal como una formulación de liberación controlada, que incluye implantes y sistemas de administración microencapsulados. Pueden usarse polímeros biodegradables y biocompatibles, tal como acetato de etileno vinilo, polianhídridos, ácido poliglicólico, colágeno, poliortoésteres y ácido poliláctico. Los métodos para la preparación de dichas formulaciones serán evidentes para los expertos en la técnica. Los materiales también pueden obtenerse comercialmente, por ejemplo, de Alza Corporation y Nova Pharmaceuticals, Inc. También pueden usarse suspensiones de liposomas (incluyendo liposomas dirigidos a células infectadas con anticuerpos monoclonales frente a antígenos virales) como vehículos farmacéuticamente aceptables. Éstos pueden prepararse según los métodos conocidos por los expertos en la técnica, por ejemplo, como se describe en la Patente de EEUU No. 4.522.811.

Es especialmente ventajoso formular las composiciones orales o parenterales en formas de dosificación unitarias para facilitar la administración y la uniformidad de la dosificación. La forma de dosificación unitaria tal y como se usa en la presente memoria se refiere a unidades físicamente discretas adecuadas como dosificaciones unitarias para el sujeto que se va a tratar; cada unidad contiene una cantidad predeterminada del compuesto activo calculada para producir el efecto terapéutico deseado en asociación con el vehículo farmacéutico requerido. La especificación para las formas de dosificación unitarias de la invención está dictada por y es directamente dependiente de las características únicas del compuesto activo y el efecto terapéutico particular que se quiere conseguir, y las limitaciones inherentes en la técnica de formación de composiciones de dicho compuesto activo para el tratamiento de individuos.

Las moléculas de ácido nucleico pueden insertarse en vectores y usarse como vectores de terapia génica. Los vectores de terapia génica pueden administrarse a un sujeto, por ejemplo, por inyección intravenosa, administración local (véase, por ejemplo, la Patente de EEUU No. 5.328.470) o por inyección estereotáctica. Véase, por ejemplo, Chen *et al.*, *Proc. Natl. Acad. Sci. USA* 91:3054-3057 (1994). La preparación farmacéutica del vector de terapia génica puede incluir el vector de terapia génica en un diluyente aceptable, o puede comprender una matriz de liberación lenta en la que está incluido el vehículo de administración del gen. Alternativamente, cuando el vector de administración génica completo puede producirse intacto a partir de células recombinantes, por ejemplo, vectores retrovirales, la preparación farmacéutica puede incluir una o más células que producen el sistema de administración génica.

Las composiciones farmacéuticas pueden estar incluidas en un contenedor, paquete o dispensador junto con instrucciones para la administración.

## Ejemplos

### Ejemplo 1: Métodos y Materiales

#### Cepas y plásmidos

5 Un aislado clínico, *T. rubrum* CHUV 862-00, se usó en este estudio. *E. coli* LE392 se usó para la propagación del bacteriófago  $\lambda$ EMBL3 (Promega, Wallisellen, Suiza). Todos los experimentos de subclonación en plásmidos se realizaron en *E. coli* DH5 $\alpha$  usando el plásmido pMTL21. Chambers *et al.*, Gene 68:139-149 (1988). *P. pastoris* GSI 15 y el vector de expresión pKJ113 (Borg-von Zepelin *et al.*, Mol. Microbiol. 28:543-554 (1998)) se usaron para expresar peptidasas recombinantes. Se sabe en la técnica que *P. pastoris* puede utilizarse para expresar una multitud de proteínas recombinantes.

#### 10 Medio de crecimiento de *T. rubrum*

*T. rubrum* se creció en agar Sabouraud y medio líquido (Bio-Rad, Munich, Alemania) o, para estimular la producción de actividad proteolítica, en medio líquido que contiene proteína de soja al 0,2% (Supro 1711, Protein Technologies International, St. Louis, MO) como única fuente de nitrógeno y carbono. No se añadió sal a este medio. Los expertos en la técnica reconocerán que también es posible utilizar medio de crecimiento en el que se añade sal al medio. Un volumen de 100 ml de medio líquido se inocula con un tapón de micelio recién crecido en matraces de cultivo de tejidos de 800 ml. Los cultivos se incubaron 10 días a 30°C sin agitación.

#### Bibliotecas genómicas y de ADNc

20 Se preparó una biblioteca de ADN genómico de *T. rubrum* usando ADN aislado de micelio recién crecido. (Yelton *et al.*, Proc. Natl. Acad. Sci. USA. 81:1470-1474 (1984). El ADN se digirió parcialmente con Sau3A y los fragmentos de ADN que varían de 12 a 20 kb se aislaron de agarosa de bajo punto de fusión (Roche Diagnostics, Rotkreuz, Suiza) con agarosa (Roche Diagnostics). Estos fragmentos de ADN se insertaron en el bacteriófago XEMBL3 usando un sistema de clonación apropiado (Promega).

25 Se preparó una biblioteca de ADNc de *T. rubrum* en un plásmido pSPORT6 (Invitrogen Life Technologies; Rockville, Maryland, EEUU) usando el sistema de microcantidad de ARNm y 500  $\mu$ g de ARN total. El ARN se preparó a partir de cultivos de 10 días en medio líquido de proteína de soja (10x100ml). El micelio se molió en nitrógeno líquido hasta un polvo fino usando un mortero y maja y el ARN total se aisló usando un kit de purificación de ARN total RNeasy para plantas y hongos (Qiagen, Basilea, Suiza).

30 Se construyó previamente una biblioteca de ADNc de *A. fumigatus* con la cepa CHUVI 92-88 crecida 40 h a 30°C en medio líquido que contiene colágeno al 0,2% como una única fuente de nitrógeno y carbono (Monod *et al.*, 1991). El ARN total se extrajo como se ha descrito (Applegate y Monod) y el ARNm se purificó usando celulosa oligo(dT) (Sigma, Buchs, Suiza) según protocolos estándar (Sambrook *et al.*, 1989). Se preparó una biblioteca con este ARNm usando el fago lambda gTII (Promega) y los protocolos del fabricante.

La **Tabla 13** muestra los genes de *T. rubrum* y *A. fumigatus* que codifican aminopeptidasas.

Gen	ADN genómico (pb. desde el codón ATG hasta el de PARADA)	ADNc: longitud de ORF (pb.) desde el codón ATG	número de aa codificados desde el codón ATG	Intrones (pb del ADN genómico desde el codón ATG)
<i>ruLAP2</i>	1757	1488	495	3 intrones (pb 106-231; 556-632; 917-982) 4 exones que codifican 35, 108, 95, 257 aa
<i>fuLAP2</i>	1557	1497	498	1 intrón (pb 85-144) 2 exones que codifican 28, 470 aa
<i>ruLAP1</i>	1256	1122	373	2 intrones (pb 157-226; 968-1031) 3 exones que codifican 52, 247, 74 aa
<i>fuLAP1</i>	1298	1167	388	2 intrones (pb 187-252; 1000-1064) 3 exones que codifican 62, 249, 77 aa

### Clonación del gen LAP

Se inmovilizaron placas recombinantes ( $10^4$ ) de la biblioteca genómica en membranas de nilón GeneScreen (NEN Life science products, Boston, MA). Los filtros se hibridaron con sonda marcada con  $^{32}\text{P}$  usando condiciones de baja astringencia. Monod *et al.*, Mol. Microbiol. 13:357-368 (1994). Todas las placas positivas se purificaron y los ADN de bacteriófago asociados se aislaron como describe Grossberger. Gossberger, Nucleic Acid Res. 15:6737 (1987). Los fragmentos de hibridación de los bacteriófagos EMBL3 se subclonaron en pMTL2I según procedimientos estándar. La secuenciación de nucleótidos fue realizada por Microsynth (Balgach, Suiza).

### Aislamiento de ADNc por PCR estándar

Los ADNc de *T. rubrum* y *A. fumigatus* se obtuvieron por PCR usando ADN preparado a partir de 106 clones de las bibliotecas de ADNc. La PCR se realizó según condiciones estándar usando cebadores homólogos obtenidos de secuencias de ADN de los diferentes genes de peptidasa (Tabla 13). Doscientos ng de ADN, 10  $\mu\text{l}$  de cada oligonucleótido sentido y antisentido a una concentración de 42 mM y 8  $\mu\text{l}$  de mezcla de desoxinucleótidos (que contiene 10 mM de cada dNTP) se disolvieron en 100  $\mu\text{l}$  de tampón de PCR (10 mM Tris-HCl pH 8,3, 50 mM KCl y 1,5 mM  $\text{MgCl}_2$ ). A cada reacción se añadieron 2,5 unidades de ADN polimerasa AmpliTAQ (Perkin Elmer, Zurich, Suiza). Las mezclas de reacción se incubaron 5 min a  $94^\circ\text{C}$ , se sometieron a 25 ciclos de 0,5 min a  $94^\circ\text{C}$ , 0,5 min a  $55^\circ\text{C}$  y 0,5 min a  $72^\circ\text{C}$  y finalmente se incubaron 10 min a  $72^\circ\text{C}$ .

### Producción de LAP recombinantes

Se construyeron plásmidos de expresión por clonación de los productos ADNc de PCR en el sitio de clonación múltiple del vector lanzadera *E. coli* – *P. pastoris* pKJII3. Los productos de PCR se purificaron usando un kit de purificación de PCR (Roche Diagnostics) y se digirieron por enzimas de restricción para las que se había diseñado previamente un sitio en el extremo 5' de los cebadores (Tabla 14). *P. pastoris* GSI 15 (Invitrogen) se transformó por electroporación con 10 pg de ADN plasmídico linealizado con *EcoRI* o *SmaI*. Los transformantes seleccionados en medio deficiente en histidina (1 M sorbitol, 1% (p/v) dextrosa, 1,3% (p/v) base de nitrógeno de levadura (YNB) sin aminoácidos,  $4 \times 10^{-5}\%$  (p/v) biotina,  $5 \times 10^{-3}\%$  aminoácidos (por ejemplo,  $5 \times 10^{-3}\%$  (p/v) de cada uno de ácido L-glutámico, L-metionina, L-lisina, L-leucina, L-isoleucina) 2% (p/v) agarosa) se cribaron para inserción de la construcción en el sitio *AOX1* en placas de metanol mínimas (1,34% (p/v) YNB sin aminoácidos,  $4 \times 10^{-5}\%$  (p/v) biotina, 0,5% (v/v) metanol, 2% (p/v) agarosa). Se asumió que los transformantes incapaces de crecer en el medio que sólo contiene metanol como una fuente de carbono contenían la construcción en la localización genómica correcta de la levadura por eventos de integración en el locus *AOX1* desplazando la región codificadora *AOX1*. Estos transformantes se crecieron hasta casi la saturación (DO 20 a 600 nm) a  $30^\circ\text{C}$  en 10 ml de medio de levaduras basado en glicerol (0,1 M tampón de fosfato de potasio a pH 6,0 que contiene 1% (p/v) de extracto de levadura, 2% (p/v) peptona, 1,34% (p/v) YNB sin aminoácidos, 1% (v/v) glicerol y  $4 \times 1\%$  (p/v) biotina). Las células se recogieron y se resuspendieron en 2 ml del mismo medio con 0,5% (v/v) metanol en lugar de glicerol y se incubaron durante 2 días. Después de 2 días de incubación, el sobrenadante se recogió y se ensayó para la producción de proteínas en geles SDS-PAGE. Las enzimas peptidasa recombinantes se produjeron en grandes cantidades en el sobrenadante de cultivo celular de 400 ml.

La Tabla 14 describe los materiales usados para la expresión de las diferentes LAP en *P. pastoris*.

Tabla 14

Gen	Cebadores oligonucleotídicos	Orientación	Secuencia de aminoácidos codificada	Producto de PCR (con sitios de clonación) <sup>‡</sup>	Vector
<i>ruLAP2</i>	GT TGT CGA CTT GTT GGT CAA GAG CCC TTC GGA TGG (SEQ ID NO: 47) CT TGC/GGC CGC TTA CAT GAA GAC AGT GTG GTG TCC (SEQ ID NO: 48)	sentido antisentido	(R)(L)VGQEPFGW (SEQ ID NO: 63) GHHTVFMSTOP (SEQ ID NO: 64)	<i>ruLAP2</i> (58-1485) <i>SaI</i> --- <i>NotI</i>	pKJ113 <i>XhoI</i> --- <i>NotI</i>
<i>fulLAP2</i> <sup>‡</sup>	GT TCT CGA GGC CCA GGA TGG GAC TGG AAG (SEQ ID NO: 49) CGC AAA GGT GCA CTC GCC CCG CGA (SEQ ID NO: 50) TCG CGG GGC GAG/ TGC ACC TTT GCG (SEQ ID NO: 51) CTT A/GA TCT CTA CTG CTC AAC CCG GTC CTT (SEQ ID NO: 52)	sentido antisentido sentido antisentido	(R)GPGWDWK (SEQ ID NO: 65) SRGECTFA (SEQ ID NO: 66) SRGECTFA (SEQ ID NO: 67) KDRVEQSTOP (SEQ ID NO: 68)	<i>fulLAP2a</i> (49-460) <i>XhoI</i> --- <i>ApaI</i> <i>fulLAP2b</i> (461-1494) <i>ApaI</i> --- <i>BglII</i>	pKJ113 <i>XhoI</i> --- <i>BamHI</i>
<i>ruLAP1</i>	GT TCT CGA GGC ATT CCT GTT GAT GCC CGG GCC G (SEQ ID NO: 53) CTT A/GA TCT TTA CTT AGC AAG CTC AGT GAC GAA GCC GAC (SEQ ID NO: 54)	sentido antisentido	(R)(G)IPVDARA (SEQ ID NO: 69) VGFVTELAKSTOP (SEQ ID NO: 70)	<i>ruLAP1</i> (61-1119) <i>XhoI</i> --- <i>BglII</i>	pKJ113 <i>XhoI</i> --- <i>BamHI</i>
<i>fulLAP1</i>	GT TCT CGA GGG GCT GTA GCT GCA GTG ATT (SEQ ID NO: 55) CTT A/GA TCT TTA AAA CGG CGC AAA TGC CAA (SEQ ID NO: 56)	sentido antisentido	(R)GAVAAVI (SEQ ID NO: 71) LAFAPFSTOP (SEQ ID NO: 72)	<i>fulLAP1</i> (46-1164) <i>XhoI</i> --- <i>BglII</i>	pKJ113 <i>XhoI</i> --- <i>BamHI</i>
<i>ruDPPV</i> <sup>‡</sup>	CT TCT CGA GTC GTT CCT CCT CGT GAG CCC CG (SEQ ID NO: 57) G TTC CAT GGT/ CAT GAC CTT TGT GTC ATA CGA GAC AG (SEQ ID NO: 58) GT TCC ATG GT/C ATG ACC CCT CTC GTC AAC GAT AAG G (SEQ ID NO: 59) CTT G/GA TCC TCA TTC CTC TGC CCT CTC ACC (SEQ ID NO: 60)	sentido antisentido sentido antisentido	(R)(V)VPPREPR (SEQ ID NO: 73) VSYDTKVM (SEQ ID NO: 74) VMTPLVNDK (SEQ ID NO: 75) GERAEESTOP (SEQ ID NO: 76)	<i>ruDPPV</i> <sub>a</sub> (49-1266) <i>XhoI</i> --- <i>RcaI</i> <i>ruDPPV</i> <sub>b</sub> (1267-2325) <i>RcaI</i> --- <i>BamHI</i>	pKJ111 <i>XhoI</i> --- <i>BamHI</i>
<i>ruDPPV</i>	CCG G/AA TTC TTT ACC CCA GAG GAC TTC (SEQ ID NO: 61) GAG T/CT AGA CTA GTA GTC GAA GTA AGA GTG (SEQ ID NO: 62)	sentido antisentido	(E)(F)FTPEDF (SEQ ID NO: 77) HSYFDYSTOP (SEQ ID NO: 78)	<i>ruDPPV</i> (58-2178) <i>EcoRI</i> --- <i>XbaI</i>	pPICZaA <i>EcoRI</i> --- <i>XbaI</i>

\* En paréntesis se muestran los aminoácidos codificados por las secuencias del sitio de restricción y añadidos al extremo N-terminal de las enzimas recombinantes.

<sup>‡</sup> Los números en paréntesis representan las posiciones de nucleótidos en los ADNc de *LAP* y *DPP*

<sup>‡</sup> Los fragmentos de PCR de *fulLAP2* y *ruDPPV* insertados extremo con extremo en los vectores lanzadera *E. coli*-*P. pastoris*



### Purificación de LAP recombinantes

Las proteínas secretadas del sobrenadante de cultivo de *P. pastoris* de 400 ml se concentraron por ultrafiltración usando una célula Amicon y una membrana YM30 Ultracel Amicon (punto de corte 30 kDa) (Millipore, Volketswil, Suiza). El concentrado se lavó con 50 mM Tris-HCl, pH 7,5 y se aplicó en una columna de Mono Q-Sefarosa (Amersham Pharmacia, Dübendorf, Suiza) equilibrada con el mismo tampón. Después de lavar la columna con 50 mM Tris-HCl, pH 7,5, la elución se realizó con un gradiente lineal de 0-0,5 M NaCl a una velocidad de caudal de 1 ml/min. Las diferentes fracciones eluidas de la columna Mono Q-Sefarosa se cribaron para actividad enzimática usando Leucina-7-amino-4-metilcumarina (Leu-AMC) como sustrato y las fracciones que contienen LAP se juntaron. Después de concentrar en una célula de ultrafiltración Amicon con una membrana YM30 Ultracel Amicon y de lavar con 20 mM Tris-HCl, pH 6,0, el extracto de LAP se cargó en una columna de FPLC de exclusión por tamaño Superosa 6 (Amersham Pharmacia) y la elución se realizó a una velocidad de caudal de 0,2 ml/min usando 20 mM Tris-HCl, pH 6,0 como eluyente. Las fracciones activas eluidas se juntaron. La enzima LAP se concentró hasta un volumen final de 0,4-1,0 ml en un concentrador Centricon con un punto de corte de 30 kDa (Millipore) a 4°C antes de caracterización funcional adicional.

En un esquema de purificación alternativo, cada etapa de purificación se realizó a 4°C. Las proteínas secretadas del sobrenadante de cultivo de *P. pastoris* de 400 ml se concentraron por ultrafiltración usando una célula Amicon y una membrana YM30 Ultracel Amicon (punto de corte de 30 kDa) (Millipore, Volketswil, Suiza). El concentrado se lavó con 100 ml de 20 mM de acetato de sodio, pH 6,0 y se aplicó en una columna Mono Q-Sefarosa (Amersham Pharmacia, Dübendorf, Suiza) equilibrada con el mismo tampón. Después de lavar la columna con tampón 20 mM Tris-HCl pH 6,0, la enzima se eluyó con un gradiente lineal de 0-0,2 M NaCl a una velocidad de caudal de 1 ml/min durante 142 min. Las diferentes fracciones eluidas de la columna Mono Q-Sefarosa se cribaron para actividad enzimática usando Leucina-7-amino-4-metilcumarina (Leu-AMC) como sustrato (véase más adelante) y las fracciones que contienen LAP se juntaron. Después de concentrar en una célula de ultrafiltración Amicon con una membrana YM30 Ultracel Amicon y de lavar con PBS, el extracto LAP se cargó en una columna de FPLC de exclusión por tamaño Superdex 200 (Amersham Pharmacia) usando tampón 20 mM acetato de sodio pH 6,0 y la elución se realizó a una velocidad de caudal de 0,2 ml/min. Las fracciones activas eluidas se juntaron. La enzima LAP se sometió a caracterización adicional después de concentrar hasta un volumen final de 0,4-1,0 ml en un concentrador Centricon con un punto de corte de 30 kDa (Millipore) a 4°C.

Una fracción que contiene actividad ruLAP2 eluye de MonoQ a 30-40 min (aprox. 50 mM NaCl) y a 65-70 min con superdex 200 = Pico 3. Sin embargo, una gran cantidad de actividad LAP2 no se retuvo y eluyó en el flujo directo a 1 M NaCl. Por lo tanto, después de desalar esta fracción con 20 mM acetato de sodio, la muestra se aplicó en la misma columna MonoQ con un gradiente más amplio entre 0 y 1 M NaCl durante 142 min a 0,5 ml/min. Un primer pico de actividad eluye a 7-15 min correspondiente a 70-140 mM NaCl y un segundo pico eluye a 150-250 mM NaCl (con más contenido en actividad). La fracción a 70-140 mM NaCl eluye a 78-80 min en Superdex y se juntó por lo tanto con el pico 3 obtenido anteriormente. La fracción a 150-250 mM NaCl proporciona dos fracciones activas que eluyen respectivamente a 44-49 min (Pico 1) y 50-63 min (Pico 2) en Superdex.

### Análisis del extracto proteico

Los extractos proteicos se analizaron por SDS-PAGE con un gel de separación de 12% poliacrilamida. Los geles se tiñeron con azul brillante de Coomassie R-250 (Bio-Rad). La digestión con N-glicosidasa F se realizó como se ha descrito previamente. Dumas *et al.*, Appl. Environ. Microbiol. 64:4809-4815 (1998).

### Transferencias Western

Las membranas se tiñeron en primer lugar con Rojo-Ponceau y las principales bandas de proteína se marcaron con una aguja. Las inmunotransferencias se realizaron usando antisuero de conejo y anti-IgG de conejo de cabra conjugado con fosfatasa alcalina (Bio-Rad) o anti-IgG de conejo de cabra conjugado con peroxidasa (Amersham Pharmacia) como anticuerpos de marcaje secundarios. Los antisueros de conejo frente a ruLAP1, ruLAP2, proteasa alcalina (ALP) secretada por *A. oryzae* y proteasa neutra (NPI) secretada por *A. oryzae* de la familia fungalisina (Dumas *et al.*, J. Food Mycol. 2:271-29 (1999)) se obtuvieron de Eurogentec (Lieja, Bélgica) usando enzima recombinante purificada.

### Ensayo de actividad aminopeptidasa

La actividad aminopeptidasa se determinó usando diferentes derivados fluorogénicos aminoacil-4-metilcumaril-7-amida de péptidos y el sustrato fluorogénico inactivado internamente Lys(Abz)-Pro-Pro-pNA para la determinación específica de actividad aminopeptidasa P. Stockel *et al.*, Adv. Exp. Med. Biol. 421:31-35 (1997). Todos los sustratos fueron de Bachem (Bubendorf, Suiza). Las disoluciones madre de sustrato se prepararon a 0,1 M según las recomendaciones del fabricante y se almacenaron a -20°C. La mezcla de reacción contenía una concentración de 5 mM de sustrato y preparación de enzima (entre 56 y 2,662 ng por ensayo dependiendo de la actividad de escisión de cada enzima para los sustratos) en 25 µl de tampón 50 mM Tris-HCl ajustado al pH óptimo para cada LAP (entre

7 y 8). Después de incubar a 37°C durante 60 min, la reacción se terminó añadiendo 5 µl de ácido acético glacial y la mezcla de reacción se diluyó con 3,5 ml de agua. La 7-amino-4-metilcumarina (AMC) liberada se midió usando un espectrofluorofotómetro (fluorómetro Perkin Elmer LS-5, Zurich, Suiza) a una longitud de onda de excitación de 370 nm y una longitud de onda de emisión de 460 nm. Se usó una curva estándar obtenida con AMC sintética para evaluar la AMC liberada. La dipropil-p-nitroanilida liberada se midió a una longitud de onda de excitación de 310 nm y una longitud de onda de emisión de 410 nm. Las actividades LA se expresaron en nmoles de AMC liberada o pNA/min/µg de proteína.

La Tabla 15 detalla la actividad hidrolítica de diferentes LAP frente a varias aminoacil-MCA comparación (%) con Leu-MCA usada como estándar.

Tabla 15

<u>Sustrato</u>	<u>ruLAP2</u>	<u>fuLAP2</u>	<u>ruLAP1</u>	<u>fuLAP1</u>	<u>pkLAP</u>
<b>Leu-AMC</b>	100.0	100.0	100.0	100.0	100.0
<b>Ile-AMC</b>	6.4	1.8	7.4	13.2	6.3
<b>Val-AMC</b>	4.8	0.8	4.9	27.6	4.0
<b>Ala-AMC</b>	33.3	11.7	5.2	4.7	584.7
<b>Gly-AMC</b>	3.3	2.2	5.1	0.8	74.8
<b>Ser-AMC</b>	26.1	10.3	5.9	10.3	24.6
<b>Thr-AMC</b>	0.9	0.1	1.7	5.1	4.4
<b>Cys-AMC</b>	14.9	2.1	18.5	5.0	35.5
<b>Met-AMC</b>	119.7	89.5	41.3	116.9	46.1
<b>Asn-AMC</b>	114.6	73.5	6.8	29.4	33.9
<b>Gln-AMC</b>	49.9	37.0	2.3	44.9	50.7
<b>Asp-AMC</b>	3.8	0.3	0.0	0.8	0.9
<b>Glu-AMC</b>	3.7	1.1	0.0	0.0	4.7
<b>Lys-AMC</b>	4.6	2.3	9.1	7.7	70.1
<b>Arg-AMC</b>	1.9	2.3	12.3	53.9	174.8
<b>His-AMC</b>	0.6	1.9	0.1	0.8	17.6
<b>Phe-AMC</b>	17.1	8.9	4.6	163.7	184.4
<b>Pro-AMC</b>	21.4	7.4	1.4	12.0	7.9
<b>Hyp-AMC</b>	14.2	13.3	0.3	3.9	1.7
<b>Gly-Pro-AMC</b>	7.2	74.1	0.0	5.4	16.7
<b>Pyr-AMC</b>	0.0	0.0	0.0	0.0	0.0
<b>Lys(Abz)Pro-PropNA</b>	0.0	0.0	0.0	0.0	0.0

#### Efecto de varios reactivos químicos en LAP

Los inhibidores y cationes metálicos se pre-incubaron con las enzimas durante 15 min a 37°C. Después, se añadió Leu-AMC a una concentración final de 5 mM. Después de una incubación adicional durante 60 min, la actividad enzimática se midió como se ha descrito anteriormente. Los inhibidores y sus concentraciones ensayadas en LAP purificadas fueron: 500 µM amastatina (Bachem), 40 µM benzamida (Sigma), 500 µM bestatina (Bachem), 5 mM/1 mM EDTA (Sigma), 100 µM E-64 (L-trans-epoxisuccinil-leu-4-guanidinobutilamida) (Bachem), 100 µM leupeptina (Sigma), 5 mM/1 mM orto-fenantrolina (Sigma), 500 µM ácido p-cloromercuribenzoico (Sigma), 100 µM pepstatina A (Sigma), 40 µM PMSF (Sigma), 20 µM TLCK (Roche Diagnostics) y 20 µM TPCK (Roche Diagnostics). CaCl<sub>2</sub>, MgCl<sub>2</sub>, MnCl<sub>2</sub>, CoCl<sub>2</sub>, ZnCl<sub>2</sub>, NiCl<sub>2</sub>, CuCl<sub>2</sub>, se ensayaron a concentraciones de 0,5 mM y 1 mM.

La Tabla 16 detalla la actividad hidrolítica de diferentes EXOX en presencia de varios inhibidores de proteasa usando Leu-MCA como sustrato para LAP. La actividad se proporciona como un porcentaje de la actividad de reacción enzimática control sin inhibidor.

Tabla 16

Inhibidor	ruLAP2	fuLAP2	ruLAP1	fuLAP1	pLAP
EDTA 5 mM	5	50	0	16	99
EDTA 1 mM	7	77	7	19	68
ortofenantrolina 5 mM	0	0	0	0	0
ortofenantrolina 1 mM	0	0	0	0	0
Bestatina 500 µM	55	88	0	11	24
Amastatina 500 µM	0	0	0	17	0
ácido p-cloromercuribenzoico 500 µM	21	96	32	90	59
E 64 100 µM	34	71	103	190	93
Leupeptina 100 µM	113	61	233	149	86
Pepstatina 100 µM	45	73	160	14	64
PMSF 40 µM	79	84	78	156	58
Benzamidina 40 µM	89	91	85	77	75
TLCK 20 µM	96	120	68	80	113
TPCK 20 µM	79	87	68	95	108

La Tabla 17 detalla la actividad hidrolítica de diferentes EXOX en presencia de varios cationes usando Leu-MCA como sustrato para LAP. La actividad se proporciona como un porcentaje de la actividad de reacción enzimática control sin ningún catión.

Tabla 17

	<b>ruLAP2</b>	<b>fuLAP2</b>	<b>ruLAP1</b>	<b>fuLAP1</b>	<b>pkLAP</b>
CaCl <sub>2</sub> 0.5 mM	126.6	110.0	151.7	54.9	177.4
CaCl <sub>2</sub> 1 mM	141.9	165.4	175.6	43.3	161.8
MgCl <sub>2</sub> 0.5 mM	121.2	97.6	129.9	68.5	130.1
MgCl <sub>2</sub> 1 mM	110.2	108.0	132.6	72.6	146.1
MnCl <sub>2</sub> 0.5 mM	77.5	84.3	120.7	25.9	157.6
MnCl <sub>2</sub> 1 mM	86.8	140.2	105.2	28.4	165.8
CoCl <sub>2</sub> 0.5 mM	591.2	378.0	210.2	104.3	876.1
CoCl <sub>2</sub> 1 mM	789.7	662.7	202.1	96.5	899.8
ZnCl <sub>2</sub> 0.5 mM	77.9	51.4	43.0	60.7	437.6
ZnCl <sub>2</sub> 1 mM	88.9	119.5	68.9	53.2	297.9
NiCl <sub>2</sub> 0.5 mM	130.5	98.4	74.8	51.7	1187.7
NiCl <sub>2</sub> 1 mM	147.9	149.3	58.1	37.2	1158.7
CuCl <sub>2</sub> 0.5 mM	50.9	68.9	40.1	25.8	1422.0
CuCl <sub>2</sub> 1 mM	34.7	73.6	13.7	17.0	1092.4

**pH óptimo de actividad de EXOX**

- 5 El pH óptimo para las actividades enzimáticas se determinó usando el sistema de tampón Ellis y Morrison. Ellis y Morrison, *Methods Enzymol.* 87:405-426 (1982). El tampón contenía tres componentes con diferentes valores de pKa mientras que la fuerza iónica del tampón permanecía constante a lo largo del intervalo completo de pH examinado. El pH del tampón se ajustó de 6 a 11 en incrementos de mitad de unidad de pH con 1 M HCl ó 1 M NaOH. Las condiciones del ensayo para la actividad en sustratos Leu-AMC fue la misma que anteriormente excepto que el tampón Tris/HCl se reemplazó con el tampón de Ellis y Morrison (composición) a los valores de pH indicados.
- 10 La Tabla 18 detalla las características de aminopeptidasas secretadas de *T. rubrum* y *A. fumigatus* nativas y recombinantes.

Tabla 18

Gen	Longitud del gen	Número de intrones	Prepro teína (aa)	Señal (aa)	Dominio maduro (aa)	Masa molecular de la cadena polipeptídica de la enzima madura (kDa)	Masa molecular de la enzima nativa/recombinante (kDa)	Masa molecular de la enzima recombinante después de glicosilación (kDa)	Número de sitios putativos de glicosilación	pl calculado (dominio maduro)*	Rendimiento de proteína recombinante (µg/ml)	Número de registro GenBank
ruLAP1	1256	2	373	19	354	38.804	31-33/38-40	38-40	3	6,39 (6,23)	40	AY496930
fuLAP1	1298	2	388	17	371	41.465	§NI/40	40	3	5,67 (5,67)	80	AY436356
ruLAP2	1757	3	495	18	477	51.487	58/58-65	52	4	7,32 (6,94)	40	AY496929
fuLAP2	1557	1	498	15	383	52.270	§NI/75-100	52	6	5,57 (5,46)	100	AY436357
ruDPPIV	2326	0	775	15	760	86.610	90/90	84	4	8,05	10	AY487021

§NI : significa no determinado

\* El valor en paréntesis corresponde al polipéptido de longitud completa sin prosequencia

**Temperatura óptima de actividad de EXOX**

Las condiciones de temperatura óptima se determinaron midiendo la actividad enzimática a su pH óptimo después de incubar cada una de las LAP con Leu-AMC (5 mM) a 20, 30, 40, 50, 60, 70 y 80°C durante 10, 30 y 60 min.

**Ensayos proteolíticos**

- 5 La actividad proteolítica se midió usando caseína marcada con resorufina en tampón fosfato (20 mM; pH 7,4). La mezcla de reacción contenía 0,02% de sustrato en un volumen total de 0,5 ml. Después de incubar a 37°C, el sustrato no digerido se precipitó con ácido tricloroacético (4% concentración final) y se separó del sobrenadante por centrifugación. La absorbancia a 574 nm del sobrenadante se midió después de alcalinización añadiendo 500 µl de tampón Tris (500 mM; pH 9,4). Para propósitos prácticos, una unidad (U) de actividad proteolítica se definió como la que produce una absorbancia de 0,001 por min.

**Ejemplo 2: actividad proteolítica secretada de *T. rubrum***

- 15 *T. rubrum* se creció a 30°C en un medio que contiene 0,2% proteína de soja como una única fuente de carbono y nitrógeno. Después de 14 días de crecimiento, se notó una clarificación concomitante del medio de cultivo y se detectó una actividad proteolítica sustancial (400 U ml<sup>-1</sup>) usando caseína marcada con resorufina como sustrato. Esta actividad proteolítica se inhibió un 15% y 85% por PMSF y orto-fenantrolina, respectivamente, avalando que se secretaron serina y metaloproteasas por *T. rubrum*. Los análisis por transferencia Western del sobrenadante del cultivo revelaron que *T. rubrum*, como *M. canis*, secretaba endoproteasas de la familia subtilisina (MEROPS>S8) y de la familia fungalisina (MEROPS>M36) similar a la proteasa alcalina ALP y la metaloproteasa neutra NPI secretada por *A. oryzae* (Véase la Figura 1). Además, se detectó una alta actividad en sustratos tales como Leu-AMC y Leu-pNA en el sobrenadante del cultivo de *T. rubrum*.

**Ejemplo 3: actividad aminopeptidasa secretada de *T. rubrum***

- 25 Las secuencias de nucleótidos de los genes de endoproteasa de *Microsporium canis* mostraron 50-70% de similitud con genes homólogos que codifican las subtilisinas y las fungalisinas secretadas por *A. oryzae* y *A. fumigatus*. Además, los genes de *M. canis* y *Aspergillus* mostraron estructuras colineales intrón-exón. Por lo tanto, las secuencias de ADN disponibles para los genes de *A. oryzae* y *Saccharomyces cerevisiae* que codifican aminopeptidasas se usaron para diseñar sondas para el cribado de una biblioteca de ADN genómico de *T. rubrum*. La caracterización de las aminopeptidasas secretadas por *T. rubrum* en comparación con las secretadas por el oportunista *A. fumigatus* se realizó usando proteínas recombinantes.

**Ejemplo 4: Clonación de los genes que codifican aminopeptidasas de *T. rubrum* y *A. fumigatus***

- 30 Las Tablas 19A y 19B detallan una comparación por parejas de varias LAP.

**Tabla 19A**

Enzima	%Similitud o Identidad <sup>a</sup>			
	ruLAP1	fuLAP1	orLAP1	LAP de <i>Vibrio</i>
ruLAP1		72	72	41
fuLAP1	50		70	39
orLAP1	48	49		42
LAP de <i>Vibrio</i>	22	21	23	

Tabla 19B

Enzima	%Similitud o Identidad <sup>a</sup>			
	ruLAP2	fuLAP2	orLAP2	aaY de <i>S. cer.</i>
M28A				
ruLAP2		69	71	53
fuLAP2	51		85	52
orLAP2	49	72		53
aaY de <i>S. cer.</i>	32	33	34	

<sup>a</sup> Los valores de porcentaje de similitud (esquina superior derecha) y porcentaje de identidad (esquina inferior izquierda) se obtuvieron con el programa Gap implementado en el paquete GCG del Genetics Computer Group, Universidad de Wisconsin, Madison.

5 La Figura 14 es un alineamiento de secuencias de aminoácidos deducidas de aminopeptidasas de la subfamilia M28E. Los sitios de procesamiento putativos de la secuencia señal están subrayados. Un sitio de procesamiento KR putativo en ruLAP1 se indica con un triángulo sólido. Los aminoácidos de los dos sitios de unión Zn<sup>++</sup> en la aminopeptidasa de *S. griseus* y conservados en otras LAP se indican con una flecha abierta. El alineamiento se realizó con el algoritmo Pileup implementado en el paquete GCG de la Universidad de Wisconsin y reformateado con Boxshade 3.2. AbispLAP1 es para LAP de *Agaricus bisporus*.

10 La Figura 15 es un alineamiento de secuencias de aminoácidos deducidas de aminopeptidasas de la subfamilia M28A. Los sitios de procesamiento putativos de la secuencia señal están subrayados. Dos restos de aminoácidos, His y Asp, conservados en las LAP fúngicas y que unen un primer ión Zn<sup>++</sup> en la aminopeptidasa de *S. griseus* se indican con triángulos abiertos. Dos restos adicionales His y Glu que unen un segundo ión Zn<sup>++</sup> se indican con rombos sólidos, mientras que el esto Asp que hace puente con los dos iones Zn<sup>++</sup> se indica con una flecha abierta. \* representa restos de metionina encontrados sólo en ruLAP2. El alineamiento se realizó con el algoritmo Pileup implementado en el paquete GCG de la Universidad de Wisconsin y reformateado con Boxshade 3.2.

15 Las secuencias de aminoácidos GPGINDDGSG (SEQ ID NO: 36) y DM(Q/M)ASPN (SEQ ID NO: 37) se encontraron en una aminopeptidasa secretada de *A. oryzae* de 52 kDa (Patente de EEUU 6.127.161) y la aminopeptidasa de *S. cerevisiae*. Nishizawa *et al.*, J. Biol. Chem. 269:13651-13655 (1994). A partir de estos datos, se usaron dos oligonucleótidos consenso (GGXATXAAYGAYGGXTCXGG (SEQ ID NO: 38) y TTXGGXGAXGXCXATCATRTC (SEQ ID NO: 39) como sentido y antisentido, respectivamente, para amplificar ADN de *T. rubrum*. Un producto de PCR de 220pb se obtuvo y secuenció. La secuencia de aminoácidos deducida mostró alta similitud con la secuencia de aminoácidos de las aminopeptidasas de *A. oryzae* y *S. cerevisiae*. Este fragmento de PCR de 220pb se usó como sonda para cribar una biblioteca de ADN genómico *T. rubrum* fago λ EMBL3 y se encontró una secuencia de nucleótidos que codifica una aminopeptidasa putativa (ruLAP2). Una secuencia de nucleótidos que codifica una aminopeptidasa secretada similar (fuLAP2) se encontró en la secuencia genómica de *A. fumigatus* (en la dirección de red www.TIGR.com).

20 Un fragmento de 1.200 pb que contiene la secuencia de nucleótidos del gen que codifica una aminopeptidasa de *A. oryzae* de 31 kDa (Patente de EEUU 5.994.113) se obtuvo por PCR de ADN genómico de *A. oryzae* usando los oligonucleótidos GCATTCCTGUGATGCCCGGGCCG (sentido) (SEQ ID NO: 40) y TTAGCTAGCAAGCTCAGTGACGAAGCCGAC (antisentido) (SEQ ID NO: 41). Este fragmento se usó como sonda para un segundo cribado de la biblioteca de ADN genómico de *T. rubrum*. Se encontró una secuencia de nucleótidos (EMBL) similar a las que codifican la aminopeptidasa de *A. oryzae* de 30 kDa y a otra aminopeptidasa secretada putativa de la secuencia genómica de *A. fumigatus* (en la dirección de red www.TIGR.com) en el ADN del fago λ EMBL3 de la biblioteca genómica de *T. rubrum*. Estas aminopeptidasas putativas de *T. rubrum* y *A. fumigatus* se denominaron ruLAP1 y fuLAP1, respectivamente.

25 Las secuencias de nucleótidos identificadas de ruLAP1, ruLAP2, fuLAP1 y fuLAP2 contenían cada una, una secuencia señal de 17-20 aminoácidos. La estructura intrón-exón de los genes de *T. rubrum* y *A. fumigatus* se determinaron mediante la secuenciación de un producto de PCR usando cebadores 5'-sentido y 3'-antisentido basados en ADN genómico aislado (Véase la Tabla 14) y ADN total de un conjunto de 10<sup>8</sup> clones de las bibliotecas de ADNc de *T. rubrum* y *A. fumigatus* como diana. El primero de los tres intrones en ruLAP2 estaba en una posición similar a la del intrón único de fuLAP2 (Véase la Tabla 13). Los genes ruLAP1 y fuLAP1 tienen estructuras colineales similares con dos intrones y tres exones.

45

**Ejemplo 5: Producción de aminopeptidasas recombinantes de *T. rubrum* y *A. fumigatus***

Los ADNc de *T. rubrum* y *A. fumigatus* obtenidos por RT-PCR se clonaron en pKJII3 (Borg-von Zepelin *et al.*, 1998) y se expresaron en *P. pastoris*. Dependiendo de la peptidasa producida, se obtuvieron aproximadamente 10-80 µg/ml de enzima activa en Leu-AMC (Véase la Tabla 18). En condiciones de cultivo idénticas, *P. pastoris* de tipo salvaje no secretó ninguna actividad leucina aminopeptidasa en el medio de cultivo. Los análisis por SDS-PAGE de ruLAP2, fuLAP1 y fuLAP2 recombinantes secretadas por los transformantes de *P. pastoris* mostraron una banda dispersa (Figura 2). Después de tratamiento con N-glicosidasa F, sólo apareció una banda principal con migración más rápida en los geles avalando que, a diferencia de ruLAP1, estas tres LAP eran glicoproteínas (Figura 2). La masa molecular aparente de cada LAP recombinante desglucosilada estaba cerca de la masa molecular calculada de la cadena polipeptídica deducida de la secuencia de nucleótidos de los genes que codifican la proteasa. Las estructuras primarias deducidas (secuencias de aminoácidos) de cada enzima recombinante se proporcionan en la Tabla 18.

**Ejemplo 6: Detección de ruLAP1 y ruLAP2 en sobrenadante de cultivo de *T. rubrum***

Usando antisuero anti-ruLAP1, se detectó una acumulación de un producto LAP1 con una movilidad electroforética mayor que la de ruLAP1 recombinante en el sobrenadante de cultivo de *T. rubrum* (Véase la Figura 3).

Usando antisuero anti-ruLAP2, los análisis de transferencia Western de un sobrenadante de cultivo de *T. rubrum* revelaron que *T. rubrum* secretaba LAP2 glicosilada con la misma movilidad electroforética que la de la enzima recombinante de *P. pastoris* (Véase la Figura 3).

**Ejemplo 7: Propiedades de LAP recombinantes**

Las aminopeptidasas ruLAP1, ruLAP2, fuLAP1, fuLAP2 así como la aminopeptidasa microsomal de riñón porcino (pkLAP) hidrolizaron cada una eficazmente Leu-AMC. Este sustrato se usó para determinar la temperatura y pH óptimos de actividad, y para caracterizar adicionalmente las enzimas midiendo el efecto de (i) varios inhibidores de peptidasa conocidos (Véase la Tabla 16) y (ii) diferentes iones divalentes (Véase la Tabla 17). Cada LAP fue capaz de escindir Leu-AMC a 20°C y tuvo una temperatura óptima que varió de 40 a 50°C. El pH óptimo fue entre 7,0 y 8,5 (Véase la Tabla 18). Un pre-tratamiento de 10 min a 80°C inactivó las enzimas totalmente e irreversiblemente.

Las aminopeptidasas ensayadas fueron inhibidas fuertemente o totalmente por amastatina (Véase la Tabla 16) a una concentración de 500 µM. RuLAP1, fuLAP1 y pkLAP también fueron inhibidas por bestatina, pero este inhibidor tuvo sólo un efecto inhibidor parcial en ruLAP2 y fuLAP2. De los agentes quelantes ensayados, orto-fenantrolina inhibió totalmente las cinco enzimas a concentraciones de 1 y 5 mM. FuLAP1, ruLAP2 y ruLAP1 fueron más sensibles a EDTA que las demás LAP. E64 y p-cloromercuribenzoato (inhibidores de cisteína proteasa) disminuyeron la actividad de ruLAP2 indicando la presencia de restos tiol críticos para actividad en la secuencia de aminoácidos de esta enzima. Leupeptina (inhibidor de serina/cisteína proteasa), PMSF (inhibidor de serina proteasa), benzamidina, TLCK y TPCK no tuvieron efectos inhibidores claros en todas las LAP ensayadas. Sorprendentemente, fuLAP1 y ruLAP1 presentaron alguna sensibilidad a 0,1 mM pepstatina (inhibidor de ácido aspártico proteasa).

Con la excepción de fuLAP1, que presenta una sensibilidad general a los iones divalentes, los iones Co<sup>++</sup> incrementaron la actividad de las LAP de 200% a 900% a una concentración de hasta 1 mM. Las cuatro LAP fúngicas mostraron sensibilidades variables a los iones divalentes. Por ejemplo, fuLAP2 se activó por Mn<sup>++</sup> y Ca<sup>++</sup>, mientras que fuLAP1 se inhibió por los mismos iones. La pkLAP microsomal, altamente activada por Zn, Ni y Cu<sup>++</sup>, se diferencia de las cuatro LAP fúngicas de la familia M28.

La actividad hidrolítica de las enzimas frente a diferentes aminoacil-AMC se comparó con Leu-AMC usada como referencia (Véase la Tabla 15). Según la aminopeptidasa ensayada, se detectaron varias preferencias para el diferente resto aminoacil. Por ejemplo, la aminopeptidasa pkLAP se diferencia de las cuatro LAP fúngicas por una eficacia extremadamente alta frente a Ala-AMC y Arg-AMC. ruLAP1 fue claramente la más selectiva para Leu-AMC. Sin embargo, se observaron otras actividades de escisión preferenciales más con ruLAP2, fuLAP1 y fuLAP2. Por ejemplo, Ser- y Pro-AMC fueron escindidas más eficazmente por ruLAP2, mientras que fuLAP1 prefirió Arg-, Val- y Phe-AMC. Sólo ruLAP2 escindió eficazmente Asp- y Glu-AMC. Ninguna de estas enzimas presentó una actividad aminopeptidasa P ya que no fueron capaces de escindir Lys(Abz)-Pro-Pro-pNA.

**Ejemplo 8: Aplicación de ruLAP2 junto con ruDPPIV en la digestión de péptidos gliadina**

La enfermedad celíaca (CD) es una enfermedad digestiva que daña el intestino delgado e interfiere con la absorción de nutrientes de los alimentos. Las personas que tienen enfermedad celíaca no pueden tolerar una proteína denominada gluten, que se encuentra en trigo, centeno y cebada. Cuando las personas con enfermedad celíaca toman alimentos que contienen gluten, su sistema inmune responde dañando el intestino delgado. La enfermedad tiene una frecuencia de ≈ 1:200 en la mayor parte de los grupos de la población mundial y el único tratamiento para la enfermedad celíaca es mantener una dieta de por vida estrictamente sin gluten. Para la mayor parte de la gente,



el seguimiento de esta dieta parará los síntomas, cicatrizará el daño intestinal existente y prevendrá el daño adicional.

- 5 Los componentes tóxicos principales del gluten del trigo son una familia de proteínas ricas en Pro y Gln denominadas gliadinas, que son resistentes a la degradación en el tracto gastrointestinal y que contienen varios epítomos que estimulan las células T. Existe alguna controversia acerca de los epítomos que inducen efectivamente una activación inmunológica de células T HLA-DQ2 positivas derivadas del intestino y periféricas (Vader *et al.*, Gastroenterology 122:1729-1737 (2002)) porque se han usado para estos estudios diferentes sistemas *in vitro*. La capacidad de los péptidos gliadina para inducir toxicidad en un modelo de cultivo de órgano de CD no se corresponde con la de estimular células T y viceversa. McAdam y Sollid, Gut 47:743-745 (2000). Además, la unión de muchos epítomos del gluten a HLA-DQ2 y HLA-DQ8 pero no todos se incrementa por la desamidación de determinados restos de glutamina en ácidos glutámicos mediante la acción de la enzima tisular del intestino delgado transglutaminasa, que potencia su capacidad de estimular a las células T. Molberg *et al.*, Nat. Med. 4:713-717 (1998). Sin embargo, la desamidación no es un requerimiento absoluto para la activación de las células T. Arentz-Hansen *et al.*, Gastroenterology 123:803-809 (2002).
- 10 Otras estrategias para tratar o prevenir CD, con la última esperanza de ser una alternativa para la dieta "sin gluten", se han sugerido en los últimos años, incluyendo la inhibición de la activación de las células T por compuestos que bloquean la unión del péptido a HLA-DQ2, inhibidores de la transglutaminasa tisular que evitan la desamidación del gluten (Sollid, Nat. Rev. Immunol. 2:647-655 (2002)) y suplemento con peptidasa peroral. Este último método se considera como una ayuda a la digestión completa de péptidos inmunoestimuladores por la implicación de las proli endopeptidasas bacterianas que tienen una amplia tolerancia para los péptidos que contienen prolina. Shan *et al.*, Science 297:2275-2279 (2002); Hausch *et al.*, Am. J. Physiol. Gastrointest Liver Physiol. 283:G996-G1003 (2002). Se identificó un fragmento relativamente grande de gliadina que es resistente a la degradación con enzimas digestivas. Además, se mostró que este péptido es un estimulador potente de diferentes clones de células T restringidos en HLA-DQ2 obtenidos de biopsias intestinales de pacientes CD estimulados con gluten, reconociendo cada uno de estos clones un epítomo diferente del 33mer. La proli endopeptidasa, que tiene una preferencia por el resto Pro-Xaa-Pro, es capaz de escindir el péptido gliadina 33mer y el efecto sinérgico de la aminopeptidasa del borde en cepillo disminuye rápidamente el potencial de estimulación de células T del péptido.
- 15 20 25

Aunque hay homólogos estables de este 33mer en cebada y centeno, estos restos peptídicos del gluten que están descritos como resistentes a la degradación gastrointestinal se usaron en nuestro caso como sustratos modelo para diferentes LAP, bien solos o en combinación con ruDPPIV: PQPQLPYPQPQLPY (SEQ ID NO: 42) (14mer) correspondiente al fragmento 82-95 de  $\alpha/\beta$  gliadina AIV (P04724) o LQLQFPQPQLPYPQPQLPYPQPQLPYPQPQPF (SEQ ID NO: 43) (33mer) correspondiente al fragmento 57-89 de gliadina MM1 (P 18573).

30

Una forma acetilada en el extremo N-terminal del 33mer (Ac-33mer) también se sintetizó como control para los experimentos de digestión con exopeptidasa para descartar cualquier escisión endoproteolítica por una enzima contaminante.

35

Las enzimas que se han evaluado incluyen: ruLAP1 (aminopeptidasa I de *Trichophyton rubrum*), ruLAP2 (aminopeptidasa II de *Trichophyton rubrum*), orLAP2 (aminopeptidasa II de *Aspergillus oryzae*), fuLAP2 (aminopeptidasa II de *Aspergillus fumigatus*), MicpKLAP (leucina aminopeptidasa microsomal de riñón porcino, Sigma), CytpKLAP (leucina aminopeptidasa citosólica de riñón porcino, Sigma) y ruDPPIV.

40

#### Síntesis de los péptidos:

Se realizó síntesis en fase sólida en un sintetizador peptídico 430A modificado a medida de Applied Biosystems, usando protocolos de neutralización *in situ*/activación con hexafluorofosfato de 2-(1H-benzotriazol-1-il)-1,1,3,3-tetrametiluronio (HBTU) para la elongación de cadena por química Boc en etapas en una resina estándar -O-CH<sub>2</sub>-fenilacetamidometilo. Schnölzer *et al.*, Int. J. Peptide Protein Res. 40:180-193 (1992).

45

Al final de la síntesis, los péptidos se desprotegeron y escindieron de la resina por tratamiento con HF anhidro durante 1hr a 0°C con 5% p-cresol como secuestrador. Después de la escisión, los péptidos se precipitaron con dietiléter enfriado en hielo, se disolvieron en acetonitrilo acuoso y se liofilizaron. Los péptidos se purificaron por RP-HPLC con una columna C<sub>18</sub> de Waters usando gradientes lineales de tampón B (90% acetonitrilo/10% H<sub>2</sub>O/0,1% ácido trifluoroacético) en tampón A (H<sub>2</sub>O/0,1% ácido trifluoroacético) y detección UV a 214 nm. Las muestras se analizaron por electropulverización espectrometría de masas con un instrumento Platform II (Micromass, Manchester, Inglaterra).

50

#### Condiciones de la reacción de degradación:

La incubación se realizó a 37°C en 50 mM Tris-HCl, pH 7,2 suplementado con 1 mM CoCl<sub>2</sub> con una concentración de sustrato de 1 mg/mL y una proporción E/S de 1:20. La reacción se paró por acidificación con CH<sub>3</sub>COOH y el

55

medio se analizó por RP-HPLC en una columna C<sub>8</sub> usando un gradiente 2%/min CH<sub>3</sub>CN en 0,1% TFA. Todos los picos se caracterizaron por ESI-MS.

#### Digestión del 14mer:

5 Como se muestra en la Figura 6, el 14mer no se digiere con ruLAP2 en 4h. No existe cambio en el perfil de HPLC cuando se compara con el control. De hecho, la digestión resulta sólo en la escisión de la Prolina N-terminal. Por otra parte, la adición de ruDPPIV resulta en una degradación completa en aminoácidos y dipéptidos, mientras que ruDPPIV sola no es capaz de hidrolizar el péptido (Figura 7).

#### Digestión del 33mer:

10 La digestión del 33mer con ruLAP2 sola resulta en degradación parcial (menos de 50%) del péptido en 4h (datos no mostrados). Este péptido no es un sustrato para ruDPPIV (Figura 8). Sin embargo, cuando se mezclan ambas enzimas, el 33mer se digiere totalmente (Figura 9) en aminoácidos y dipéptidos, algunos de los cuales pudieron identificarse por ESI-MS (Y, L, F, P, PY y PF).

15 El mismo patrón de HPLC se obtiene cuando ruDPPIV se mezcla con ruLAP2 o fuLAP2. Sin embargo, con ruLAP1 todavía están presentes algunos compuestos de alto peso molecular, pero representan menos del 10% del sustrato inicial.

Por otra parte, la incubación con aminopeptidasa microsomal de riñón porcino resulta sólo en una deleción parcial de Leu N-terminal y Phe C-terminal (debido a un contaminante carboxipeptidásico) y la adición de DPPIV no modifica el perfil. La aminopeptidasa citosólica de riñón porcino es totalmente inactiva frente al 33mer.

20 La estabilidad de la Ac-gliadina 33mer en los experimentos de digestión con LAP o DPPIV solas, o mezcladas entre sí, confirma que se requiere un grupo amino libre para la degradación completa de la gliadina 33mer por estas exopeptidasas.

#### Digestión con otras enzimas:

25 La digestión con Pronasa (E/S=1/25) durante 20h es sólo parcial (menos de 40%) y la adición de ruLAP2 (ambas enzimas a una proporción E/S (p:p) de 1:50) no mejora la hidrólisis. Por otra parte, la adición de DPPIV en las mismas condiciones resulta en una degradación completa del péptido debido a la acción complementaria de una aminopeptidasa y dipeptidilpeptidasa.

La quimiotripsina sola o suplementada con ruLAP o DPPIV no es capaz de degradar el péptido.

#### Ejemplo 9: Aplicación de ruLAP2 en el procesamiento de proteínas recombinantes expresadas fusionadas con otra proteína o con una Etiqueta N-terminal

30 LAP2 se evaluó en la escisión de Gly-Ser del extremo N-terminal de proNPY y de una Ala suplementaria del extremo N-terminal del mismo péptido. Con el fin de ampliar la aplicabilidad de LAP2 sola o conjuntamente con otra exopeptidasa en el procesamiento de proteínas recombinantes grandes, se incubó una proteína G-CSF recombinante (Cys<sup>17</sup>→Ser, Lys<sup>16,23,34,40</sup>→Arg) con una secuencia N-terminal Met-Thr-Pro-, sucesivamente con ruLAP2 y ruDPPIV para eliminar secuencialmente Met y el dipéptido Thr-Pro de la proteína de 175 restos.

#### 35 Digestión de Gly-Ser-proNPY con ruLAP2:

40 El péptido se incubó toda la noche a 37<sup>0</sup>C y 1 mg/ml en un tampón 50 mM Tris.HCl, 1 mM CoCl<sub>2</sub> con ruLAP2 a una proporción E/S de 1:20 y 1:100 (p:p). El material digerido se aisló por RP-HPLC y se caracterizó por ESI-MS. Como se muestra en la Figura 10, la incubación con ruLAP2 resulta en la escisión de dos restos N-terminales Gly-Ser con una pérdida teórica de 144.1 amu (encontrado 144.2). El mismo resultado se obtiene a una proporción E/S de 1:100. La digestión se detiene cuando la enzima llega a un resto Xaa-Pro-, que en el caso de proNPY es Tyr-Pro.

#### Digestión de Ala-proNPY con ruLAP2:

Las condiciones de incubación fueron las mismas que para Gly-Ser-proNPY. La Figura 11B muestra que la alanina N-terminal se eliminó casi totalmente (pérdida de masa molecular de 71 amu) de proNPY.

#### Escisión sucesiva de Met y Thr-Pro del extremo N-terminal de G-CSF:

45 El análogo mutante de G-CSF conocido como TG47 usado en estos experimentos es metionil- [C17S, K16,23,34,40R] G-CSF con una masa teórica de 18.894,90 para la proteína replegada.

**Digestión con ruLAP2:**

5 La disolución madre de G-CSF (1,9 mg/ml en PBS que contiene 0,1% Sarcosil) se diluyó 4 veces en 50 mM Tris-HCl a pH 7,2 suplementado con 1 mM CoCl<sub>2</sub> y se incubó con ruLAP2 (E/S=1/20 y 1:100, p:p) durante 15h a 37<sup>o</sup>C. La disolución se diluyó con 30% (v:v) acetonitrilo, se acidificó con ácido acético y la proteína se aisló por RP-HPLC para caracterización por MS. Como se muestra en la Figura 12A y B, la incubación de toda la noche resulta en la escisión completa de la metionina N-terminal con una pérdida de masa teórica de 131.2 amu. Con una proporción E/S (p:p) de 1:100, todavía están presentes trazas de material no escindido después de una incubación durante toda la noche.

10 Este experimento se repitió a una escala de 2 mg con el fin de aislar el material truncado en una columna de RP-HPLC semipreparativa, realizando la digestión con una proporción E/S de 1:25 (p:p) a 37<sup>o</sup>C durante 15h. El material aislado (0,8 mg) se caracterizó por ESI-MS (Figura 12B, desMet-G-CSF, masa molecular calculada a 18.763,7 amu; masa molecular medida a 18.762,5).

**Digestión de desMet-G-CSF con DPPIV:**

15 El material liofilizado se suspendió a una concentración de 1 mg/ml en 50 mM Tris-HCl, pH 7,5 que contiene 0,1% Sarcosil y se incubó toda la noche a 37<sup>o</sup>C con DPPIV a una proporción E/S de 1/20 (p:p). La proteína se aisló por RP-HPLC como anteriormente y se caracterizó por ESI-MS (Figuras 13A y B). La digestión con DPPIV (Figura 13B) resulta en la escisión del dipéptido N-terminal Thr-Pro (masa molecular calculada de 18.564,8 uma; masa molecular medida a 18.563). Todavía están presentes trazas de material no digerido en el medio de reacción.

20 Así, una aplicación secuencial de LAP2 y DPPIV resulta en la eliminación eficaz de una secuencia N-terminal de una proteína recombinante. La digestión con ruLAP2 se detiene cuando la enzima llega a un resto de aminoácido como "punto de parada", tal como Xaa-Pro-Xaa, o el resto Xaa-Pro, que puede introducirse específicamente como un "punto de parada" LAP2, se escinde posteriormente con DPPIV.

Sin embargo, la escisión inicial de los restos N-terminales es altamente dependiente de la secuencia ya que la etiqueta Met(His)<sub>6</sub> no se eliminó de Met(His)<sub>6</sub>-proNPY por incubación con LAP y DPPIV.

## REIVINDICACIONES

1. Una mezcla de enzimas aisladas y purificadas que consiste en
  - a) una dipeptidil peptidasa IV (DPP IV) aislada y purificada que comprende SEQ ID NO: 35, o una secuencia de aminoácidos que es al menos 95% idéntica a SEQ ID NO: 35 y
  - 5 b) una aminopeptidasa aislada y purificada procedente de un dermatofito seleccionado del grupo que consiste en *Epidermophyton floccosum*, *Microsporum audouinii*, *Microsporum ferrugineum*, *Trichophyton concentricum*, *Trichophyton kanei*, *Trichophyton megninii*, *Trichophyton mentagrophytes*, *Trichophyton raubitschekii*, *Trichophyton rubrum*, *Trichophyton schoenleinii*, *Trichophyton soudanense*, *Trichophyton tonsurans*, *Trichophyton violaceum*, *Trichophyton yaoundei*, *Microsporum canis*, *Microsporum equinum*, *Microsporum nanum*, *Microsporum persicolor*, *Trichophyton equinum*, *Trichophyton simii*, *Trichophyton verrucosum*, *Microsporum gypseum*, *Trichophyton ajelloi* y *Trichophyton terrestre*, que comprende la forma madura de una secuencia de aminoácidos que comprende bien
    - i) una leucina aminopeptidasa (LAP) aislada y purificada que comprende SEQ ID NO: 3, o una secuencia de aminoácidos que es al menos 95% idéntica a SEQ ID NO: 3; o
    - 15 ii) una leucina aminopeptidasa (LAP) aislada y purificada que comprende SEQ ID NO: 6, o una secuencia de aminoácidos que es al menos 95% idéntica a SEQ ID NO: 6.
2. La mezcla de enzimas aisladas y purificadas de la reivindicación 1, que comprende la aminopeptidasa de (i) y la aminopeptidasa de (ii).
3. La mezcla de enzimas aisladas y purificadas de las reivindicaciones 1 ó 2, en la que DPPIV comprende una secuencia de aminoácidos que comprende sustituciones de aminoácidos conservativas de menos de 5% de los restos de aminoácidos de SEQ ID NO: 35.
4. La mezcla de enzimas aisladas y purificadas de las reivindicaciones 1-3, en la que la leucina aminopeptidasa de (i) o (ii) comprende una secuencia de aminoácidos que comprende sustituciones de aminoácidos conservativas de menos de 5% de los restos de aminoácidos de SEQ ID NO: 3 o SEQ ID NO: 6.
- 25 5. La mezcla de enzimas aisladas y purificadas de cualquiera de las reivindicaciones 1 a 4, que comprende además una o más proteasas.
6. La mezcla de enzimas aisladas y purificadas de la reivindicación 5, en la que dicha una o más proteasas se seleccionan del grupo que consiste en tripsina, pronasa, quimiotripsina y proteinasaK.
7. La mezcla de enzimas aisladas y purificadas de cualquiera de las reivindicaciones 1-6, caracterizada porque la leucina aminopeptidasa (LAP) es ruLAP2 que consiste en SEQ ID NO: 3.
- 30 8. La mezcla de enzimas aisladas y purificadas de cualquiera de las reivindicaciones 1-7, para usarse en la degradación, hidrólisis o maduración de sustratos polipeptídicos o proteínas.
9. Un kit que comprende, en uno o más contenedores, la mezcla de enzimas aisladas y purificadas de cualquiera de las reivindicaciones 1-7.
- 35 10. Uso de la mezcla de enzimas aisladas y purificadas de cualquiera de las reivindicaciones 1-7, en la hidrólisis de proteínas, para la degradación de subproductos, proteínas tóxicas o contaminantes; para la degradación de priones o virus; para la degradación de proteínas para proteómica; para la degradación de sustrato cornificado; para la hidrólisis de polipéptidos para análisis de aminoácidos; para cosmetología tal como herramientas de peeling, depilación dermoabrasión y dermoplanning; para la limpieza y/o preparación de prótesis; para suavizantes textiles; para jabones; para la limpieza o desinfección de tanques sépticos o cualquier contenedor que contiene proteínas que deben eliminarse o esterilizarse; y para la limpieza de instrumentos quirúrgicos.
- 40 11. Uso de la mezcla de enzimas aisladas y purificadas de cualquiera de las reivindicaciones 1-7, en la fabricación de un medicamento para la limpieza de heridas, para la cicatrización de heridas o para tratar un síndrome asociado con una enfermedad humana, seleccionándose dicha enfermedad del grupo que consiste en enfermedad celíaca, mala absorción del tracto digestivo, una reacción alérgica, una deficiencia enzimática, una infección fúngica, micosis y esprúe.
- 45 12. El uso de la mezcla de enzimas aisladas y purificadas según la reivindicación 11, caracterizado porque la reacción alérgica es una reacción al gluten o a fragmentos de éste.
- 50 13. Un método *in vitro* para degradar un sustrato polipeptídico, comprendiendo el método poner el sustrato polipeptídico en contacto con la mezcla de enzimas aisladas y purificadas de cualquiera de las reivindicaciones 1-7.

14. El método de la reivindicación 13, caracterizado porque dicha mezcla de enzimas aisladas y purificadas digiere secuencialmente un sustrato polipeptídico de longitud completa o una proteína de longitud completa.
15. El método de cualquiera de las reivindicaciones 13 ó 14, caracterizado porque el sustrato polipeptídico se selecciona del grupo que consiste en caseína, gliadina, gluten, albúmina de suero bovino o fragmentos de éstas.
- 5 16. El método de cualquiera de las reivindicaciones 13 a 15, caracterizado porque la longitud del sustrato polipeptídico es de 2 a 200 aminoácidos.
17. Un método *in vitro* para eliminar aminoácidos del extremo amino de una proteína, que comprende poner dicha proteína en contacto con la mezcla de enzimas aisladas y purificadas de cualquiera de las reivindicaciones 1-7.
- 10 18. El método de la reivindicación 17, caracterizado porque el extremo amino de dicha proteína comprende una etiqueta His.

Figura 1

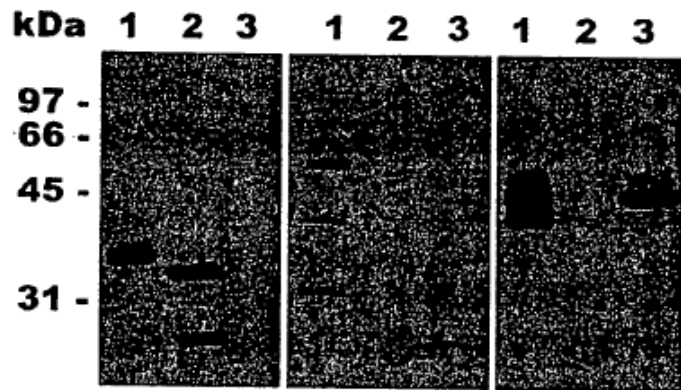


Figura 2

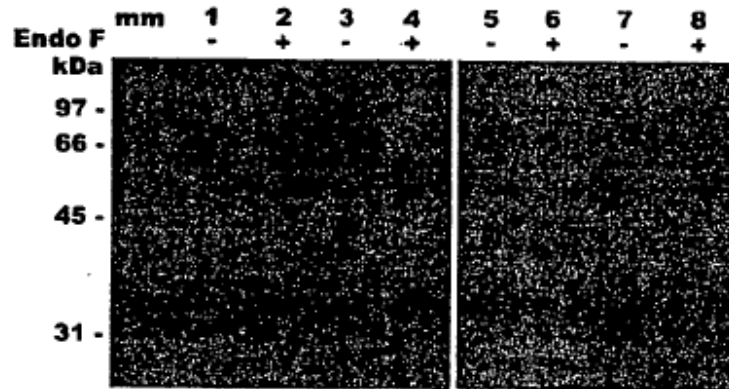


Figura 3

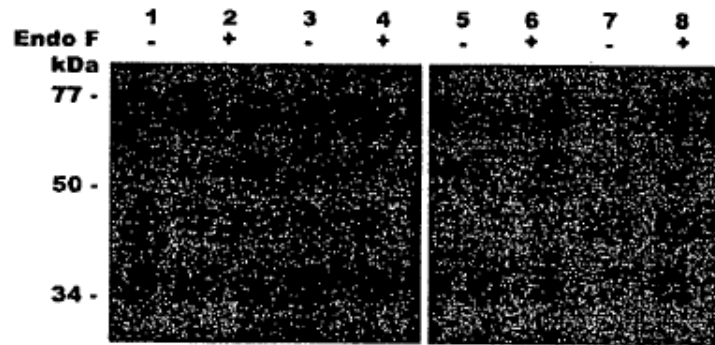




Figura 4; Determinación de la actividad AMPP de *T. rubrum* a diferentes pH usando Lys(Abz)-Pro-Pro-pNA como sustrato.

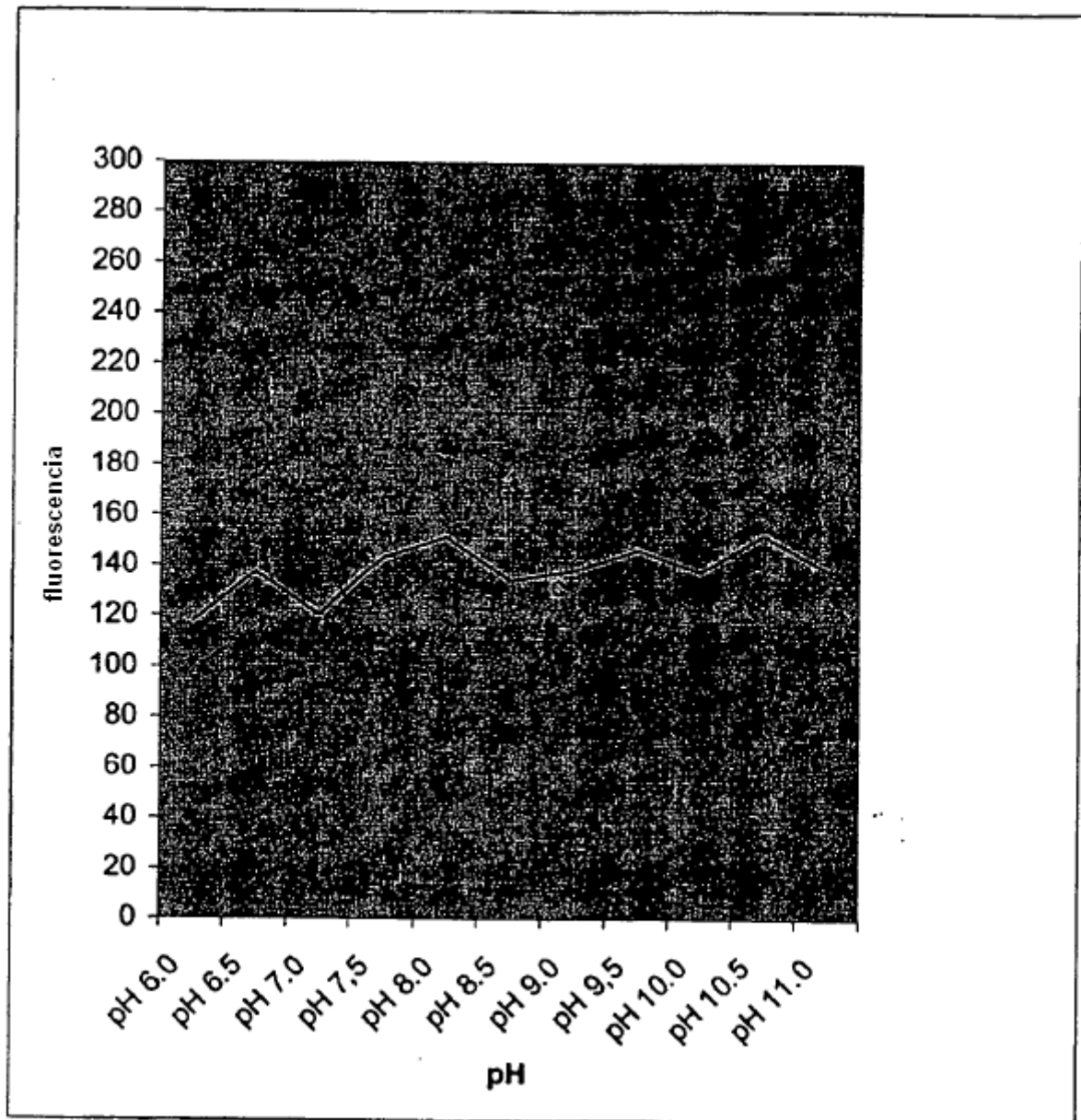
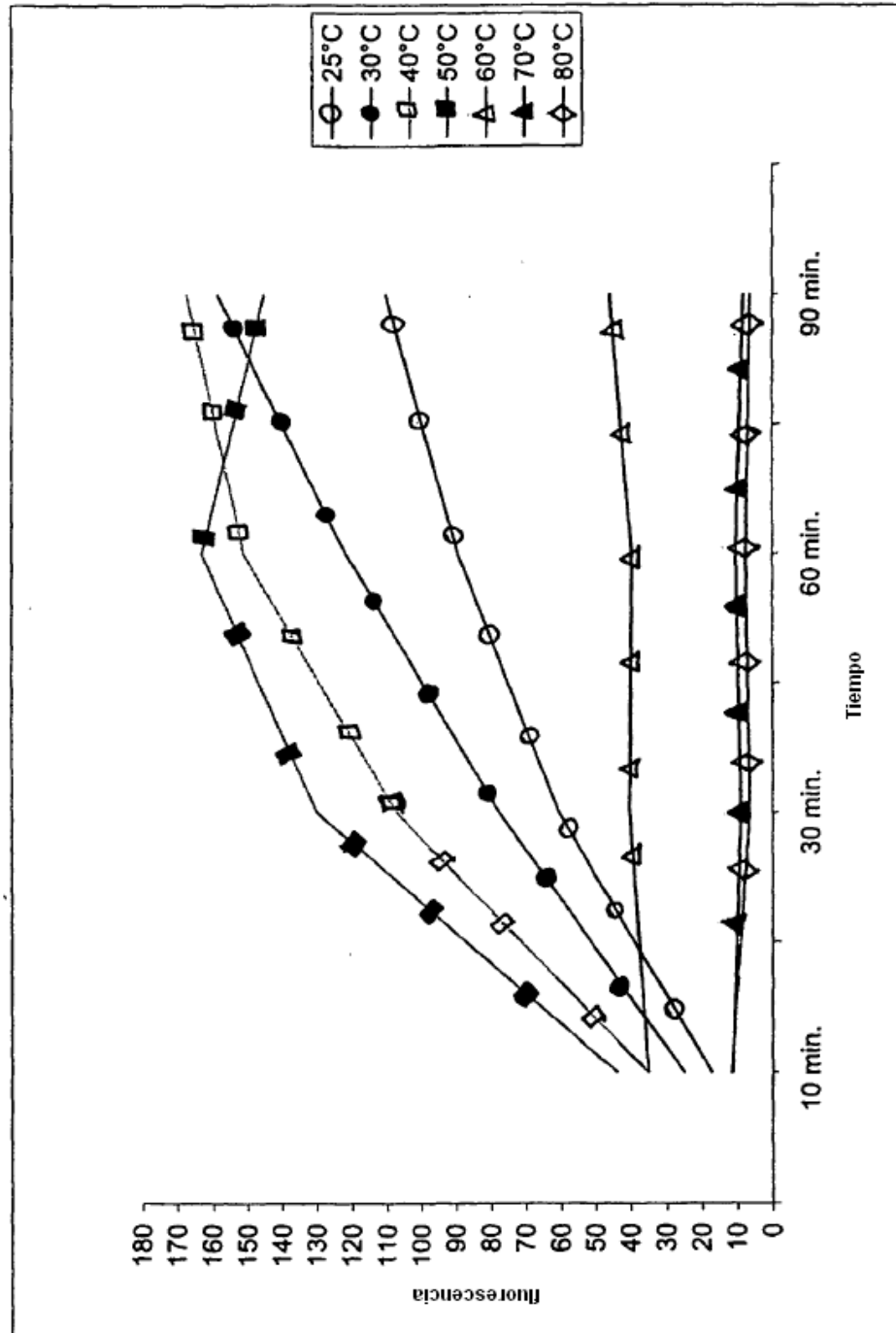
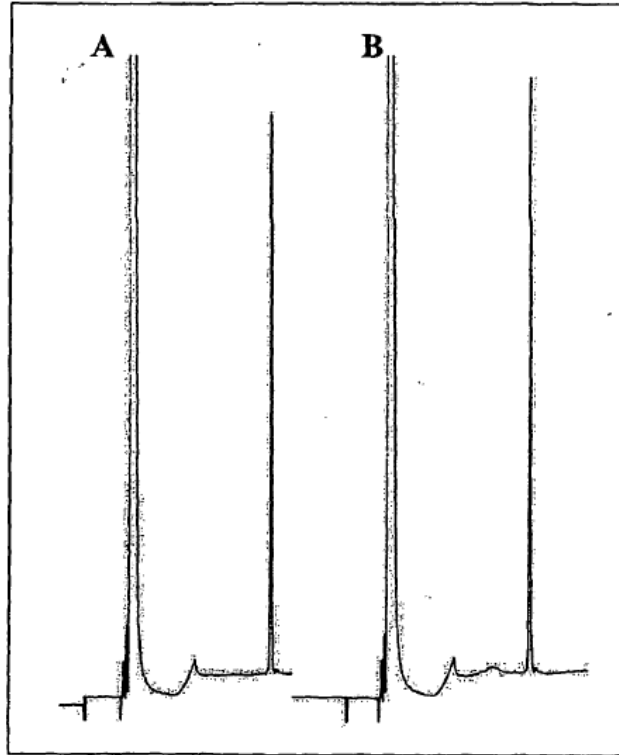


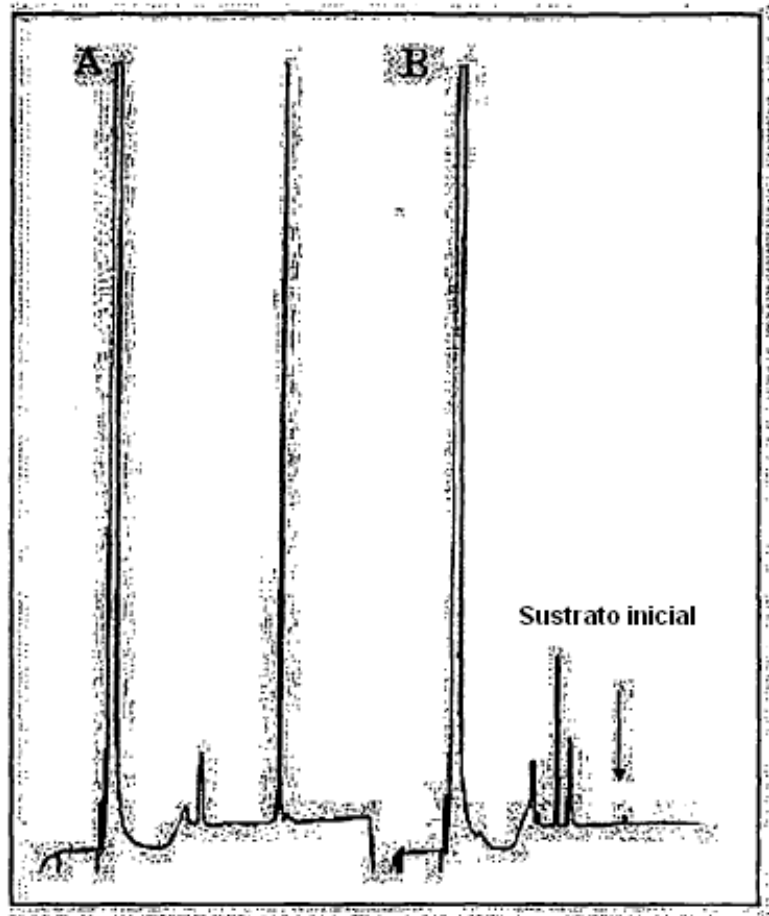
Figura 5: Actividad enzimática de AMPP de *T. rubrum* a diferentes temperaturas.



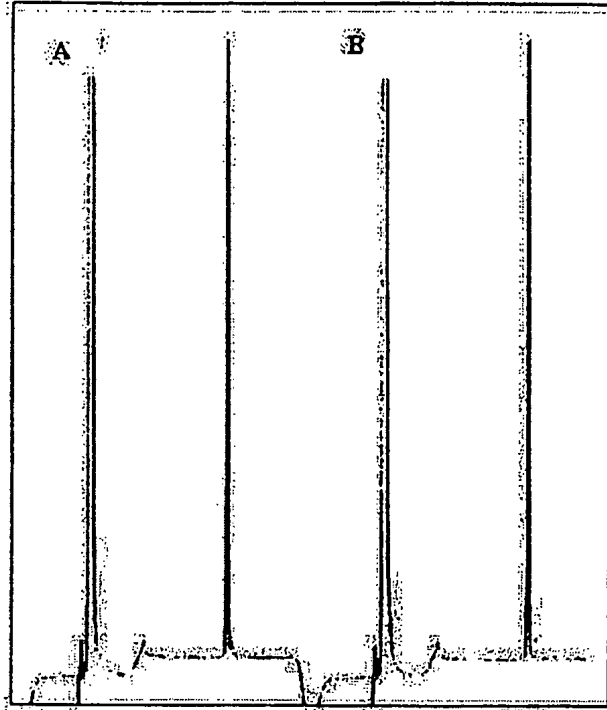
**Figura 6:** Digestión de gliadina 14mer sin (A) y con (B) ruLAP1 durante 4h a 37<sup>0</sup>C con una proporción E/S de 1/50 (p:p).



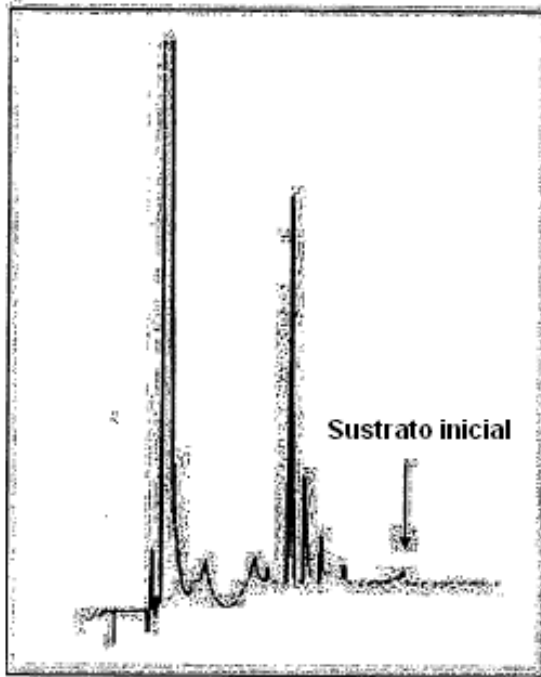
**Figura 7:** Digestión de gliadina 14mer durante 4h con ruDPPIV sola a una proporción E/S de 1/25 (p:p) (A) y con una mezcla de ruLAP1 y DPPIV a la misma proporción de 1/50 (p:p) (B).



**Figura 8:** Digestión de gliadina 33mer sin (A) y con (B) ruDPPIV durante 4h a 37<sup>0</sup>C con una proporción E/S de 1/50 (p:p).

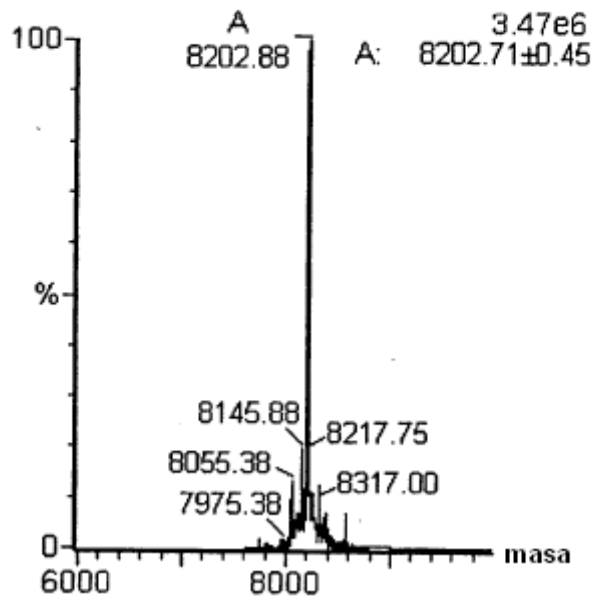


**Figura 9:** Digestión de gliadina 33mer a 37<sup>0</sup>C durante 4h con una mezcla de ruLAP1 y ruDPP1V a una proporción E/S de 1/50 (p:p) cada una.

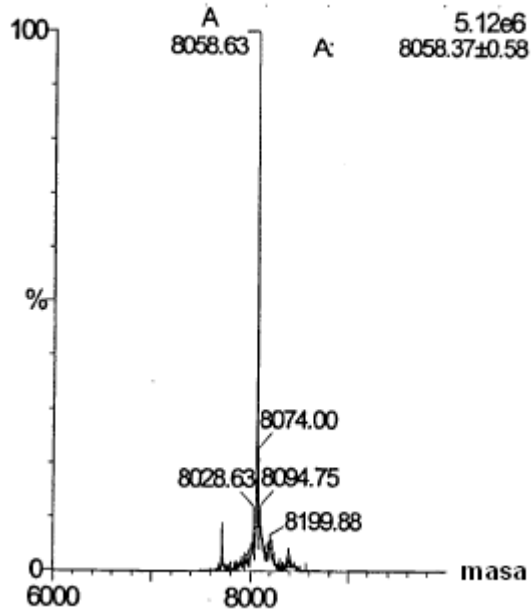


**Figura 10:** Espectro de masas de Gly-Ser-pro-NPY (masa molecular calculada a 8.203,12 Da) antes (A) y después de digestión (B) con ruLAP1, lo que resulta en la escisión de Gly-Ser (-144,1 Da)

10-A

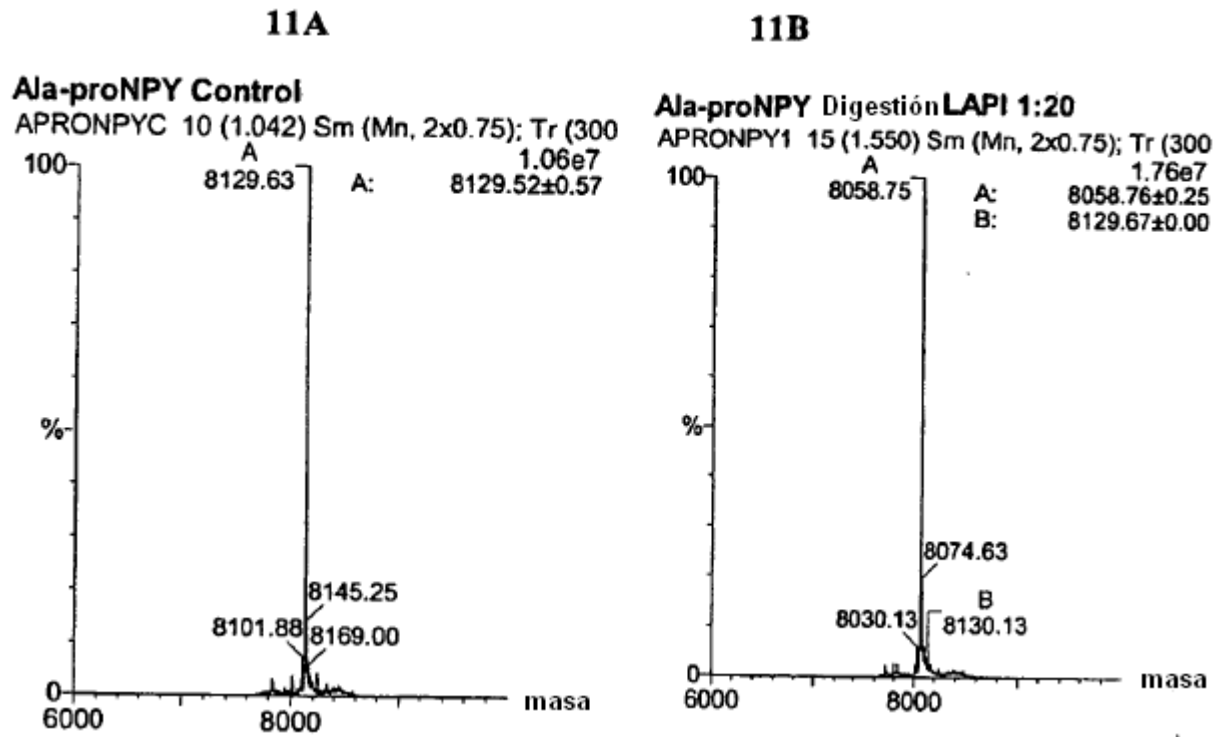


10-B

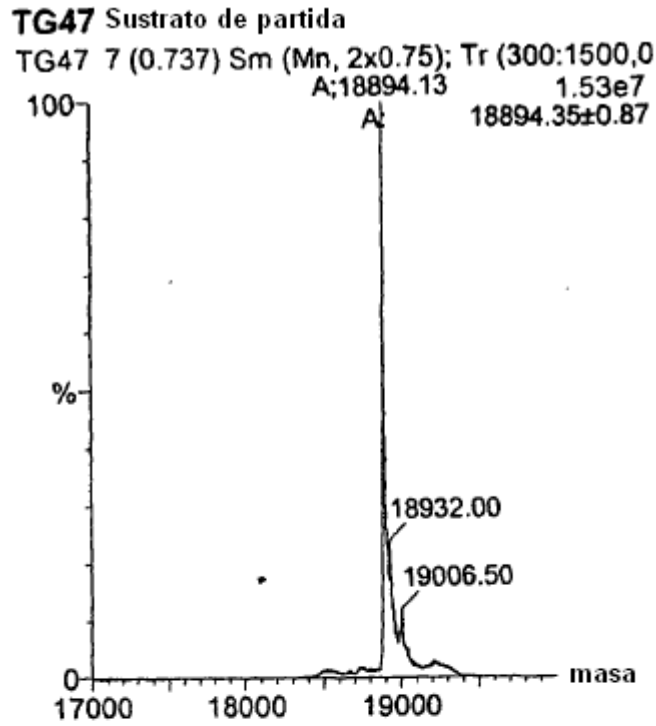




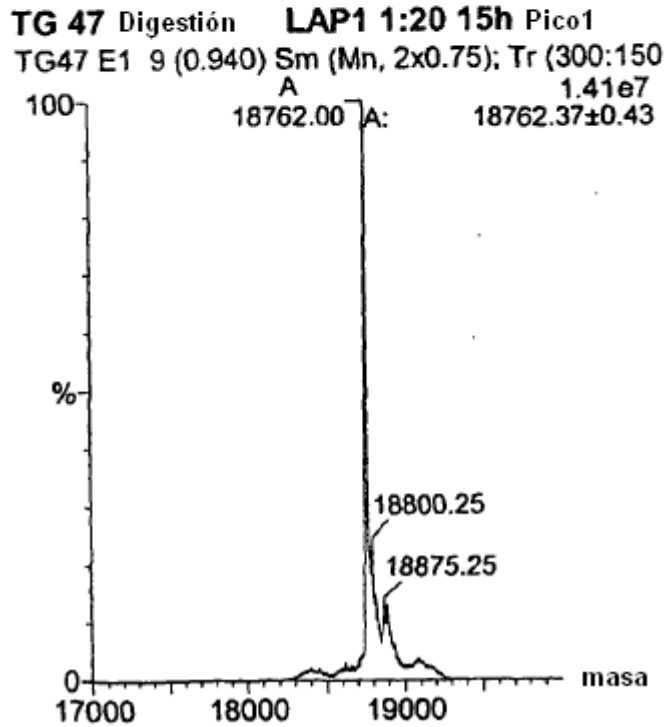
**Figura 11:** Espectro de masas de Ala-proNPY (A, masa molecular calculada a 8.130,0 uma) y del producto de digestión (B) con ruLAP1.



**Figura 12:** Espectro de masas de (A) TG47 (masa molecular calculada de 18.894,9 uma) y del producto de digestión con ruLAP1 (B), lo que resulta en la eliminación de la metionina N-terminal.

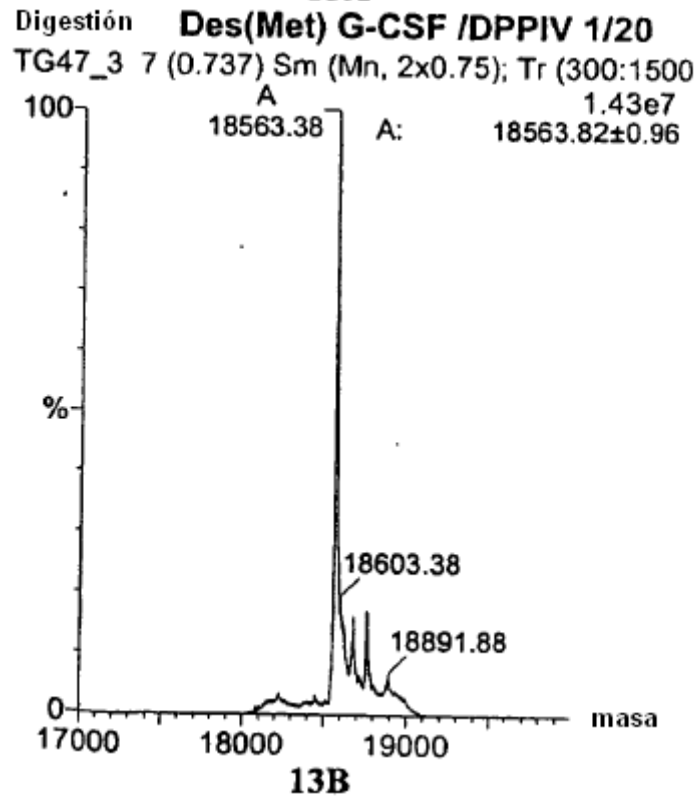
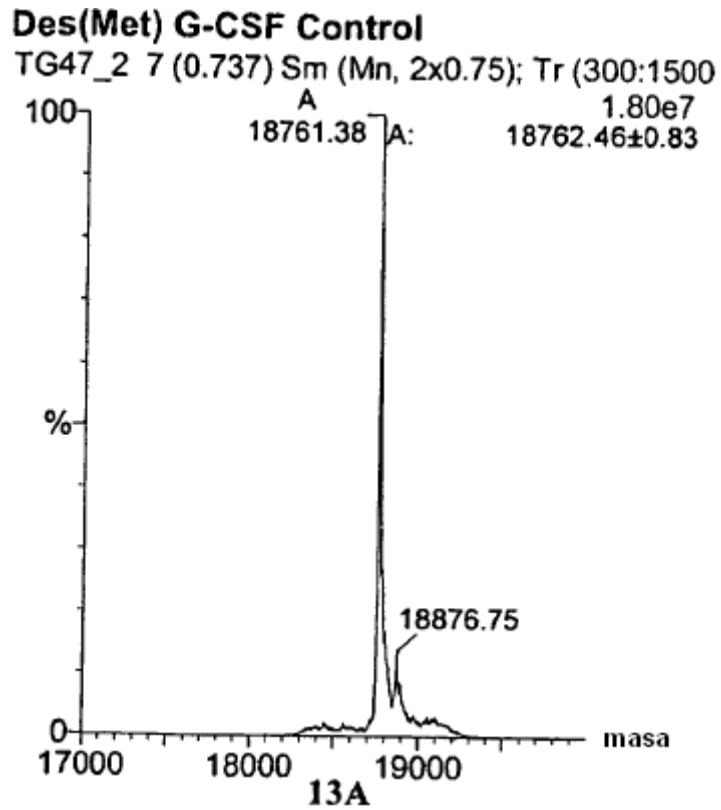


**12A**



**12B**

**Figura 13:** Caracterización por ESI-MS del producto de digestión de desMet-G-CSF sin (A) o con (B) ruDPPIV.



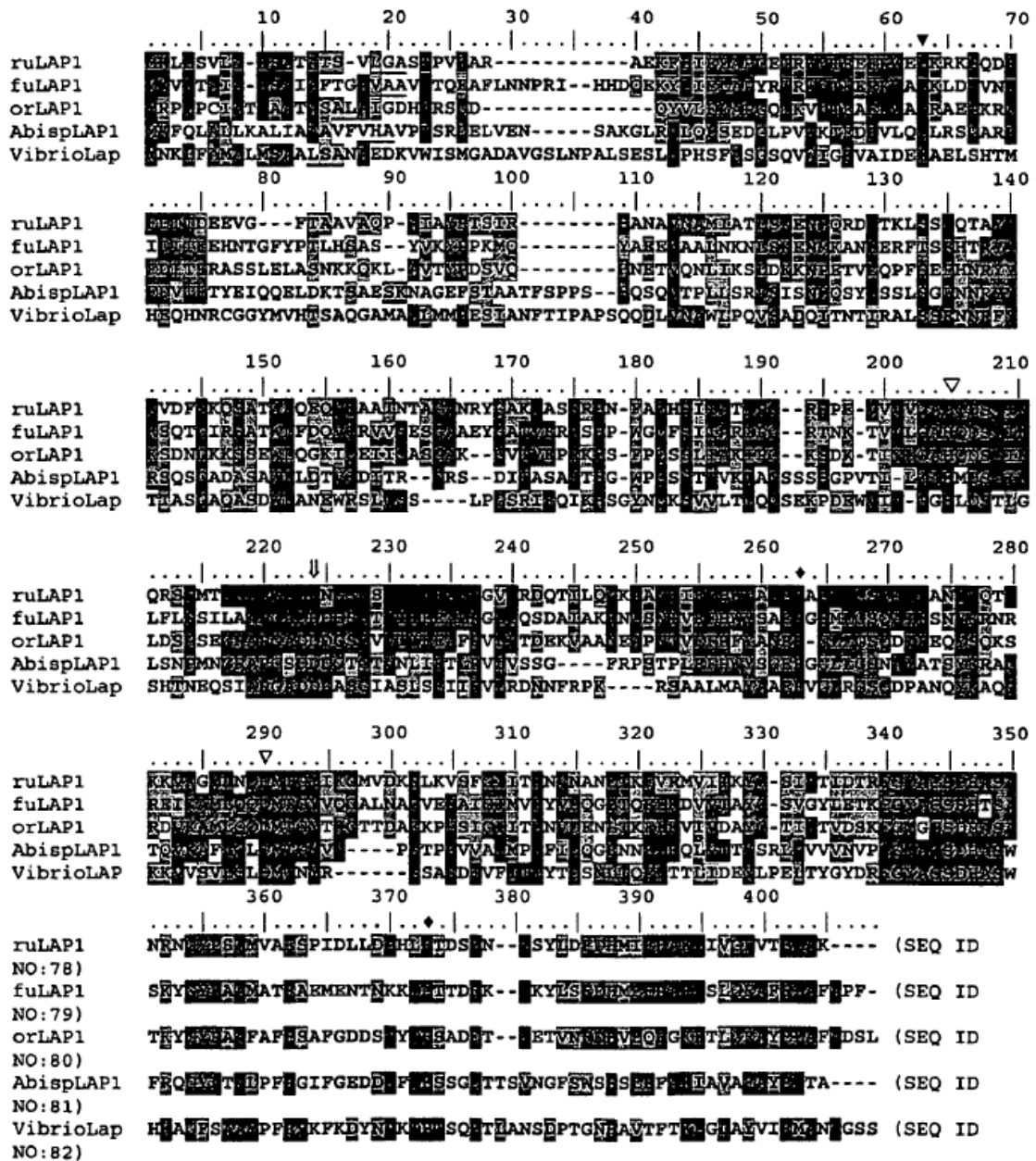


FIG. 14

