



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 294 033**

51 Int. Cl.:
C07K 14/435 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Número de solicitud europea: **01974472 .1**

86 Fecha de presentación : **10.10.2001**

87 Número de publicación de la solicitud: **1326890**

87 Fecha de publicación de la solicitud: **16.07.2003**

54 Título: **Fosfatasa alcalina de camarón.**

30 Prioridad: **10.10.2000 GB 0024827**

45 Fecha de publicación de la mención BOPI:
01.04.2008

45 Fecha de la publicación del folleto de la patente:
01.04.2008

73 Titular/es: **Biotec Pharmacon ASA**
Strandgaten 3
9008 Tromsø, NO

72 Inventor/es: **Lanes, Olav;**
Nilsen, Inge y
Överbö, Kersti

74 Agente: **Elzaburu Márquez, Alberto**

ES 2 294 033 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

Fosfatasa alcalina de camarón.

La presente invención se refiere a moléculas de ácidos que codifican la enzima fosfatasa alcalina.

La fosfatasa alcalina (E.C. 3.1. 3.1), con los nombres alternativos fosfomonoesterasa alcalina, fosfomonoesterasa o glicero fosfatasa, son fosfohidrolasas de monoéster ortofosfórico con una actividad enzimática óptima en condiciones alcalinas. Ejemplos de fosfatasa alcalina (ALP - siglas en inglés) son ADN, ARN y ribo-, así como desoxirribonucleósido-trifosfatos. La hidrólisis produce un alcohol y un ortofosfato. En otras palabras, las ALFs desfosforilan ADN, ARN, rNTPs y dNTPs. También se ha reseñado la desfosforilación de proteínas por parte de diversas ALPs. Las ALPs están repartidas por la naturaleza, encontrándose en organismos que oscilan desde bacterias a seres humanos. Organismos complejos contienen, habitualmente, ALFs tanto específicas como no específicas para el tejido. Los polipéptidos difieren en tamaño de 15 a 170 kDa. Algunas de estas proteínas están unidas o "ancladas" a membranas celulares. Lo más comúnmente, las enzimas tienen una necesidad por cationes de metales divalentes tales como Mg^{2+} o Zn^{2+} .

El actual uso de ALPs en biología molecular se enfoca, pero no se limita a tres aplicaciones principales del análisis o la preparación de ADN: 1) desfosforilación del ADN vector tras digestiones de enzimas de restricción para minimizar el auto-ligamientos del vector de clonación, favoreciendo así el ligamiento del inserto al vector y creando una construcción recombinante, 2) desfosforilación de dNTPs después de amplificaciones de PCR (reacción en cadena de la polimerasa), en uso combinado con una exonucleasa de cadena sencilla que hidroliza cebadores para dar dNTPs, para omitir la necesidad de una limpieza adicional antes de la secuenciación directa de ADN de productos de la PCR, ó 3) desfosforilación de extremos de ADN para el subsiguiente marcaje con ^{32}P utilizando $[\gamma\text{-}^{32}P]NTP$ y T4 polinucleótido cinasa. Las reacciones de la enzima ALP descritas son etapas intermedias en procedimientos de análisis del ADN. Otras importantes aplicaciones conocidas en la técnica incluyen aquellas en las que las actividades de la enzima se utilizan como informadores, tal como en el análisis con enzima unida a inmunosorbente (ELISA - siglas en inglés), en sistemas de fusión de genes o de suministro de genes o en conjugación con oligonucleótidos utilizados como sondas de hibridación.

En productos comercialmente disponibles se utilizan tres ALPs principales:

- i) fosfatasa alcalina del intestino de terneros (CIAP - siglas en inglés)
- ii) fosfatasa alcalina de camarones (SAP - siglas en inglés) del camarón ártico *Pandalus borealis*
- iii) fosfatasa alcalina bacteriana (BAP - siglas en inglés)

Las enzimas CIAP y SAP animales tienen actividades específicas similares de 2000-4000 unidades/mg de proteína, en comparación con 50 unidades/mg de proteína BAP, que hace a las dos enzimas anteriores más atractivas en calidad de enzimas eficaces. La enzima SAP es inactivada por un calor moderado tal como se comenta más abajo y, así, se prefiere para varias aplicaciones en las que la actividad de la enzima necesita ser eliminada antes de etapas adicionales en los procedimientos.

Se informó de un mutante de la BAP tratado por ingeniería genética y sensible a la temperatura [Shandilya, H. y Chatterjee, D.K., 1995. An engineered thermosensitive alkaline phosphatase for dephosphorylating DNA. Focus, 17 (3): 93-95], no dándose detalles para describir la ingeniería. Esta enzima mutante (TsAP), vendida por Life Technologies, Inc., es inactivada (95% o más) por el calor (65°C durante 15 min), en presencia solamente de EDTA. La temperatura de reacción recomendada para el mutante y la enzima de tipo salvaje es también 65°C. TsAP tiene una actividad al menos 40 veces mayor que la BAP de tipo salvaje. Con referencia a la alta actividad específica, TsAP es casi equiparable a otras ALPs tales como CIAP y SAP.

Se han purificado y caracterizado dos ALPs térmicamente lábiles a partir de un microorganismo psicrófilico (de Prada, P. y Brenchley, J.E., 1997, Purification and characterization of two extracellular alkaline phosphatases from a psychrophilic *Arthrobacter* isolate. Appl. Env. Microbiol., 63(7): 2928-2931]. Las enzimas que variaban con respecto a las especificidades y propiedades cinéticas del sustrato, exhibían diferentes inestabilidades térmicas de las cuales la más lábil se asemeja a la inestabilidad de la enzima SAP. No se informó de ninguna actividad específica en unidades/mg de proteína ni de estructuras primarias.

Se aisló y caracterizó una ALP adaptada al frío de bacalao del Atlántico [Ásgeirsson, B., Hartemink, R. y Chlebowski, J. F., 1995, Alkaline phosphatase from atlantic cod (*Gadus morhua*). Kinetic and structural properties which indicate adaption to low temperatures. Comp. Biochem. Physiol., 110B(2): 315-329]. La enzima mostró una inestabilidad térmica similar a SAP. No se proporcionó ninguna estructura primaria de la proteína/gen.

Un estudio de isozimas ALP de trucha [Whitmore, D.H. y Goldberg, E., 1972, Trout intestinal alkaline phosphatases II. The effect of temperature upon enzymatic activity *in vitro* and *in vivo*. J. Exp. Zool., 182: 59-68] demostró que las temperaturas del entorno afectan al modelo de isozima y que algunas isoformas son térmicamente inestables.

Camarones de la región de aguas calientes fuera de Taiwan contienen varias ALPs [Lee, A.-C. y Chuang, N.-N., 1991, Characterization of different molecular forms of alkaline phosphatase in the hepatopancreas from shrimp *Penaeus monodon* (Crustáceos: Decapoda). Comp. Biochem. Physiol., 99B(4): 845-850], y se sacó la conclusión que las enzimas eran térmicamente estables. No se proporcionaron estructuras primarias.

Se encontró que una actividad de fosfatasa alcalina del hepatopáncreas del camarón del Ártico *Pandalus borealis* estaba contenida en las aguas residuales de procesamiento de la industria de los camarones [Olsen, R.L., Johansen, A. y Myrnes, B., 1990, Recovery of enzymes from shrimp waste. Process Biochem. 25:67-68], y la enzima fue purificada más tarde del hepatopáncreas [Olsen, R.L., Øverbø, K. y Myrnes, B., 1991, Alkaline phosphatase from hepatopancreas of shrimp (*Pandalus borealis*): a dimeric enzyme with catalytically active subunits. Comp. Biochem. Physiol. 99B(4): 755-761]. Esta proteína purificada tenía un peso molecular aparente de 65 kDa (cada subunidad) y demostró ser una enzima dimerica con subunidades catalíticamente activas en contraposición a la mayoría de otras ALPs de animales que requieren una dimerización para la actividad. De acuerdo con el informe, la SAP tiene un punto isoeléctrico de 3,7. La enzima de camarón separa eficazmente el fosfato 5' terminal de cualesquiera extremos de la cadena de ADN (extremos 5' y 3' colgantes o romos) producidos por endonucleasas de restricción, a pesar de que los extremos 5' sobresalientes son más reactivos que los extremos romos o los extremos 5' recesivos.

Con relación a la CIAP, la enzima SAP tiene un ligero desplazamiento a temperaturas más bajas para la actividad, pero no se considera que sea ciertamente activa al frío. Con una actividad de enzima máxima a aproximadamente 40°C (45°C para CIAP), casi el 40% de la actividad es retenida a 10°C o a 50°C, en comparación con una actividad de 10% y 90%, respectivamente, para la CIAP. Aunque la temperatura para la actividad máxima es próxima a 40°C, la enzima SAP comienza a perder actividad cuando se pre-incuba durante un periodo de 15 min más allá de 37°C. Después de la pre-incubación a 65°C, la actividad de la SAP se reduce en un 95% o más, y la actividad es indetectable después de la pre-incubación a 70°C. En comparación, después de tratamientos térmicos similares, la CIAP retiene un 40% y un 20% de su actividad.

Así, con relación a su competidor comercial, la enzima SAP es térmicamente inestable y activa al frío, haciéndola particularmente adecuada para uso en protocolos de laboratorio de múltiples etapas, en que una simple etapa de calentamiento puede desactivar la enzima, de modo que no forma parte en etapas ulteriores del método.

BIOTEC ASA, Tromsø, Noruega, produce la enzima SAP comercial a partir de las aguas residuales de la industria de los camarones. A bordo de los pesqueros de arrastre, los camarones recién recogidos son congelados en grandes bloques. Al desembarcar, los camarones son descongelados cuidadosamente mediante agua fría recirculada. Se utilizan aproximadamente 1000 l de agua para 4000 kg de camarones. Durante el proceso de congelación y descongelación, el hepatopáncreas de los camarones se parte y el contenido se libera al agua. Luego se concentran las aguas residuales y se utilizan varias etapas cromatográficas para purificar la enzima SAP.

El productor suministra la SAP para el mercado mundial a través de compañías bien conocidas tales como USB, Boehringer-Mannheim o Amersham Pharmacia Biotech.

Debido a su alta actividad específica, su "versatilidad" en relación con sus extremos de ADN y su inestabilidad a la temperatura relativa, la SAP se utiliza frecuentemente en la desfosforilación de vectores de clonación antes de reacciones de ligamiento, y en el tratamiento de mezclas de productos de amplificación por PCR, antes de las reacciones de secuenciación del ADN, tal como se describe en las patentes de EE.UU. 5.741.676 y 5.756.285.

La actual producción de la SAP padece de una calidad variable de las aguas residuales, que de nuevo afecta a la eficacia de producción. Dos factores provocan esta variación; i) variación estacional natural de la producción de enzimas en los camarones; y ii) la manipulación de la fuente de camarones antes de o durante la congelación y la manipulación de los camarones o del agua durante o después del proceso de descongelación.

Existe también una preocupación sobre la disponibilidad futura de las aguas residuales; i) como fuente natural, no se garantiza que los camarones estén disponibles en todo momento; y ii) la industria de los camarones busca ahora posibles nuevas vías para congelar los camarones, es decir congelación sencilla, en la que se ha sometido a ensayo que las aguas residuales contienen pequeñas cantidades de enzimas tales como SAP.

Además de estos problemas prácticos, el mercado demanda un producto de SAP recombinante, ya que los productos recombinantes se prefieren con frecuencia en las técnicas de biología molecular, particularmente en que es una cuestión la pureza del producto, por ejemplo en la producción de productos terapéuticos basados en ADN o en la ciencia forense. La SAP tiene propiedades enzimáticas y fisicoquímicas muy ventajosas y es una enzima preferida para los protocolos de laboratorio en los que es útil. Por lo tanto, existe una apreciable necesidad de una fuente sintética o recombinante de SAP, la cual se produce de una manera uniforme y pura. Sin embargo, no se han dilucidado previamente secuencias ni de ADN ni de aminoácidos para SAP.

Con el fin de aislar el gen y de clonar y producir subsiguientemente una SAP recombinante se han realizado varios intentos para obtener una secuencia N-terminal de la proteína purificada. No ha tenido éxito ninguno de estos intentos. Los análisis de la secuencia han revelado múltiples (cuatro o más) aminoácidos alternativos para cada posición N-terminal específica. Así, no ha estado disponible información alguna de la secuencia de la proteína para uso, incluso

con sondas de oligonucleótidos altamente degeneradas, para buscar el gen de SAP. Solamente se puede especular sobre la o las razones de los extremos N no homólogos en una preparación de enzimas aparentemente homogénea. Es posible que las isoformas de SAP sean producidas por diferentes genes, mediante corte y empalme alternativo o variando modificaciones post-traducción de la proteína, o que la SAP sea atacada por proteasas en el extremo N, sin ninguna reducción detectable en el peso molecular.

Se ha encontrado por parte de los autores de esta invención que la homología limitada entre la fosfatasa alcalina del hepatopáncreas de camarones árticos (*Pandalus borealis*) y fosfatasa alcalina de especies que han sido secuenciadas, evita el aislamiento rutinario de secuencias de ácidos nucleicos que codifican fosfatasa alcalina de *Pandalus borealis*.

A pesar de estas dificultades reales para obtener un consenso coherente de la secuencia de aminoácidos para la SAP, los autores de esta invención han sido capaces, sorprendentemente, de aislar un ADNc que codifica SAP a partir de una genoteca de ADNc de *Pandalus borealis* y, así, dilucidar la secuencia codificante de SAP. La invención se refiere a la provisión de una fuente alternativa de la enzima.

La secuencia de SEQ ID nº 1 corresponde a la secuencia completa de ADNc de la enzima, menos una pequeña parte en el extremo 5'. No obstante, se piensa que el producto de expresión de una molécula de ácido nucleico que comprende esta secuencia tiene toda la proporción, o al menos una proporción significativa (es decir, al menos 60%, 70%, 80% o, típicamente, al menos 90%) de la actividad de la enzima nativa. Los pocos aminoácidos N-terminales que se apartan de esta secuencia no contienen sitios de unión para el sustrato o cofactor.

Por lo tanto, una referencia en esta memoria a SAP o a una fosfatasa alcalina térmicamente inestable se considera que incluye también moléculas que son más pequeñas que ALPs que se producen en la naturaleza, pero que tienen la misma o esencialmente la misma actividad. Ventajosamente, productos de expresión de este tipo y otras enzimas ALP recombinantes de acuerdo con la invención pueden tener una actividad mayor que la enzima nativa según se aísla. Preferiblemente, la actividad específica es mayor que el valor publicado para SAP de 1900 U mg⁻¹. En esta memoria se describe un ensayo adecuado para la actividad enzimática [Olsen *et al.*].

Así, en un aspecto, la presente invención proporciona una molécula de ácido nucleico que comprende una secuencia de nucleótidos según se recoge en SEQ ID nº 1 o la secuencia complementaria de dicha secuencia, en donde dicha secuencia de nucleótidos codifica o es complementaria a una secuencia que codifica una fosfatasa alcalina térmicamente inestable.

Una molécula de ácido nucleico de acuerdo con la invención puede, así, ser ADN, ADNc o ARN de cadena sencilla o doble.

Una molécula de ácido nucleico de la invención puede ser una molécula de ácido nucleico aislada (en otras palabras, aislada o separada de los componentes con los que normalmente se encuentra en la naturaleza) o puede ser una molécula de ácido nucleico recombinante o sintética.

Se hace referencia en esta memoria a secuencias que consisten esencialmente en la secuencia de SEQ ID nº 1 y sus variantes. Así, pueden estar también presentes regiones flanqueantes en cualquier extremo. En particular, puede estar presente una región flanqueante 5' de 1 a 81 ó 1-60 nucleótidos, más particularmente 1-30 nucleótidos, que pueden incorporar una secuencia de señalización para dirigir el tratamiento de la forma natural activa de la enzima durante la transcripción/traducción. Regiones flanqueantes 3' contendrán, típicamente, 1-60, por ejemplo 1-30 nucleótidos adicionales.

El complemento de una secuencia dada será una secuencia sencilla precisa que se puede determinar de acuerdo con las reglas normales del apareamiento de bases.

Una secuencia de nucleótidos que codifica una fosfatasa alcalina térmicamente inestable puede tener al menos un 65%, más preferiblemente al menos un 75%, por ejemplo al menos un 85% o al menos un 95% de identidad de la secuencia con la SEQ ID nº 1 o su complemento. De igual manera, una secuencia de nucleótidos de este tipo puede hibridarse preferiblemente bajo condiciones estrictas a SEQ ID nº 1 o su complemento. La presente invención se basa en la identificación de la secuencia de aminoácidos y la secuencia de ADNc para SAP. Se entiende bien en la técnica que existirán o se pueden generar variantes de estas secuencias, por ejemplo variantes alélicas y secuencias relacionadas modificadas por sustitución, adición o delección de bases, sencilla o múltiple, que sean funcionalmente equivalentes a las secuencias recientemente identificadas.

Una fosfatasa alcalina eucariótica térmicamente inestable es codificada por la molécula de ácido nucleico de la presente invención. Derivados de esta enzima pueden incorporar sustituciones conservativas de aminoácidos, debido a que aminoácidos procedentes de la enzima que se produce de forma natural pueden ser reemplazados por aminoácidos con características funcionales similares en términos, por ejemplo, de tamaño, lipofilicidad y polaridad. Grupos de aminoácidos funcionalmente relacionados de este tipo son comprendidos por el experto en la técnica. También pueden estar incluidas adiciones, delecciones o sustituciones adicionales que no afectan significativamente a la actividad catalítica de la enzima. Típicamente, los derivados de la enzima tendrán 1-50, preferiblemente 1-30, por ejemplo 1-15 alteraciones no conservativas en comparación con la enzima nativa.

ES 2 294 033 T3

Por “funcionalmente equivalente” se quiere dar a entender secuencias de ácidos nucleicos que codifican polipéptidos (o los polipéptidos propiamente dichos) con actividad de fosfatasa alcalina, preferiblemente con una actividad específica similar o mayor en comparación con la enzima nativa aislada de las aguas residuales del procesamiento de *Pandalus borealis*. En otras palabras, secuencias de ácidos nucleicos funcionalmente equivalentes codifican polipéptidos que pueden catalizar la separación de grupos 5' fosfato terminales a partir de extremos de las cadenas de ADN (extremos colgantes o romos 5' y 3') producidos, por ejemplo, por endonucleasas de restricción.

Variaciones en la secuencia de nucleótidos que codifican fosfatasa alcalina pueden producirse entre diferentes poblaciones de *Pandalus borealis* dentro de las especies, entre poblaciones similares de diferente origen geográfico y también dentro del mismo organismo.

Para los fines de definir el nivel de severidad, una severidad media comprende una hibridación y/o un lavado llevado a cabo en tampón 0,2 x SSC - 2 x SSC, SDS al 0,1% (p/v) a 42°C hasta 65°C, mientras que una severidad elevada comprende una hibridación y/o un lavado llevado a cabo en tampón 0,1 x SSC - 0,2 x SSC, SDS al 0,1% (p/v) a una temperatura de al menos 55°C. Las condiciones para hibridaciones y lavados son bien comprendidos por un experto normal en la técnica.

Se pueden obtener secuencias de nucleótidos que se hibridan a SEQ ID nº 1 o su complemento, o a partes de los mismos (es decir, a sondas de hibridación derivadas de SEQ ID nº 1), bajo condiciones de alta severidad y que, preferiblemente, codifican o son complementarios a una secuencia que codifica una enzima fosfatasa alcalina o parte de la misma. Condiciones de alta severidad se pueden determinar fácilmente de acuerdo con técnicas bien conocidas en la técnica según se describe, por ejemplo, en Sambrook *et al.*, 1989, Molecular Cloning, A Laboratory Manual, 2ª edición. Secuencias hibridantes incluidas dentro del alcance de la invención son las que se unen bajo condiciones no estrictas (6 x SSC/formamida al 50% a la temperatura ambiente) y se lavan en condiciones de alta severidad (por ejemplo 0,1 x SSC, 68°C), en que SSC = NaCl 0,15 M, citrato de sodio 0,015 M, pH 7,2.

Una sonda de hibridación puede ser una parte de la secuencia SEQ ID nº 1 (o secuencia complementaria) que tiene una longitud de bases y una composición suficientes para que funcione para que se hibride a la muestra o a secuencias de ácido nucleico de ensayo, para determinar si se produce o no una hibridación bajo condiciones de alta severidad. Así, la sonda puede tener una longitud de al menos 15 bases, preferiblemente una longitud de al menos 30, 40, 50, 75, 100 ó 200 bases. Longitudes representativas de la sonda incluyen, así, 30-500 bases, por ejemplo 30-300, 50-200, 50-150, 75-100.

La identidad de la secuencia de nucleótidos se puede determinar utilizando el programa BestFit del paquete de software de Genetics Computer Group (GCG) versión 10 de la Universidad de Wisconsin. El programa utiliza el algoritmo de homología local de Smith y Waterman con los valores por defecto: penalidad por creación del intervalo = 50, penalidad por extensión del intervalo = 3, emparejamiento medio = 10.000, mal emparejamiento medio = -9.000.

Métodos para producir variantes de las secuencias específicas recogidas en esta memoria, por ejemplo mediante mutagénesis dirigida al lugar, mutagénesis aleatoria o escisión enzimática y/o ligamiento de ácidos nucleicos son bien conocidos en la técnica, como lo son métodos para determinar si el ácido nucleico, así modificado, tiene una homología significativa con la secuencia sujeto, por ejemplo por hibridación.

Se describe en esta memoria una molécula de ácido nucleico aislada que codifica o es complementaria a una secuencia que codifica una fosfatasa alcalina térmicamente inestable, comprendiendo dicha fosfatasa alcalina la secuencia de aminoácidos de SEQ ID nº 2 o variantes de dicha secuencia de aminoácidos que son al menos en un 50%, de preferencia al menos en un 60%, por ejemplo al menos en un 70% u 80% idénticas a ella.

Se describe en esta memoria un ADNc, preferiblemente derivado de *Pandalus borealis*, que codifica una fosfatasa alcalina térmicamente inestable que comprende la secuencia de aminoácidos de la Figura 2 (SEQ ID nº 2) o una variante o fragmento de la misma, que es térmicamente inestable y que tiene actividad de fosfatasa alcalina.

Se puede obtener un ADNc, de preferencia derivado de *Pandalus borealis* que codifica una fosfatasa alcalina térmicamente inestable que consiste en una secuencia de aminoácidos de la que uno o más residuos aminoácidos se suprimen, se sustituyen con o se añaden a la secuencia de aminoácidos de SEQ ID nº 2.

Un polipéptido codificado por una molécula de ácido nucleico de la invención según se define en esta memoria. La provisión de una molécula de ácido nucleico de acuerdo con la invención permite, así, obtener enzimas fosfatasa alcalina recombinantes y variantes de las mismas en cantidades y con una pureza previamente no disponible, permitiendo con ello un suministro sostenible de la enzima en una forma que puede fabricarse totalmente bajo condiciones controladas o de acuerdo con condiciones certificables según se especifica en cada momento.

Se pueden obtener polipéptidos recombinantes que comprenden (o consisten esencialmente en) una secuencia de aminoácidos que constituye una enzima fosfatasa alcalina térmicamente inestable codificada por la secuencia de nucleótidos tal como se muestra en la figura 1 (SEQ ID nº 1) o una variante funcionalmente equivalente de la misma.

El término “polipéptido”, tal como se utiliza en esta memoria, incluye tanto proteína de longitud completa como secuencias de péptidos más cortas.

“Funcionalmente equivalente”, tal como se utiliza antes en relación con las secuencias de aminoácidos de polipéptidos, define polipéptidos relacionados con o derivados de las secuencias de polipéptidos antes mencionadas, en que la secuencia de aminoácidos ha sido modificada por una sustitución, adición o delección de aminoácidos sencilla o múltiple, y también secuencias en que los aminoácidos han sido químicamente modificados, incluido por glicosilación o desglicosilación, pero que, no obstante, conservan actividad catalítica. Variantes funcionalmente equivalentes de este tipo pueden producirse en forma de variaciones biológicas naturales, o se pueden preparar utilizando técnicas conocidas, por ejemplo polipéptidos recombinantes funcionalmente equivalentes se pueden preparar utilizando las técnicas conocidas de mutagénesis dirigida al lugar, mutagénesis aleatoria o escisión enzimática y/o ligamiento de aminoácidos.

Los polipéptidos sintéticos se pueden preparar mediante expresión en una célula hospedante que contiene una molécula de ADN recombinante que comprende una secuencia de nucleótidos según se describe antes ampliamente, operativamente enlazada a una secuencia del control de la expresión, o un vehículo de clonación o vector de ADN recombinante que contiene una molécula de ADN recombinante de este tipo, por ejemplo un plásmido, cósmido, molécula bacteriófaga o cromosoma artificial de levaduras. Alternativamente, los polipéptidos pueden ser expresados por inyección directa de una molécula de ADN desnudo de acuerdo con la invención en una célula hospedante.

El polipéptido sintético, así expresado, puede ser un polipéptido de fusión que comprende una porción catalíticamente funcional de una enzima fosfatasa alcalina y un polipéptido adicional codificado por el ADN de la molécula recombinante fusionada a ella. Por ejemplo, puede ser deseable producir una proteína de fusión que comprende una fosfatasa alcalina sintética u otro polipéptido de acuerdo con la invención acoplado a una proteína tal como β -galactosidasa, fosfatasa, glutatión-S-transferasa, ureasa y similares. La mayoría de las proteínas de fusión se forman mediante la expresión de un gen recombinante, en el que dos secuencias codificantes han sido unidas entre sí con marcos de lectura en fase. También puede ser deseable fusionar péptidos sintéticos de este tipo a proteínas o péptidos con una función de unión de ligandos, por ejemplo anticuerpos, para producir, por ejemplo, anticuerpos informadores enlazados a enzimas, como es bien conocido en la técnica. Todos los derivados de fusión o híbridos de este tipo de moléculas de ácidos nucleicos que codifican fosfatasa alcalina y sus secuencias de aminoácidos respectivas quedan abarcados por la presente invención. Técnicas de expresión de ADN y de polipéptidos recombinantes adecuadas de este tipo se describen, por ejemplo, en Sambrook *et al.*, 1989. Alternativamente, los polipéptidos se pueden producir por medios químicos, tal como el proceso de síntesis en fase sólida de Merrifield bien conocida.

Más particularmente, se puede obtener un polipéptido recombinante que comprende (o consiste esencialmente en):

(a) la totalidad o parte de una secuencia de aminoácidos según se muestra en SEQ ID n° 2; o

(b) la totalidad o parte de una secuencia de aminoácidos que tiene al menos un 60% de identidad de la secuencia con la totalidad o parte de SEQ ID n° 2.

El término “recombinante” distingue estas moléculas de polipéptidos que se producen en la naturaleza y significa que la molécula ha sido generada utilizando una “tecnología de ADN recombinante”, una frase bien entendida en la técnica. Se piensa que la enzima con la secuencia de aminoácidos de SEQ ID n° 2 es una versión truncada de la enzima que se produce en la naturaleza en *Pandalus borealis* y, por lo tanto, si el correspondiente ácido nucleico se utiliza en la producción recombinante, entonces la enzima recombinante producida no será, en cualquier caso, la misma que la molécula no recombinante, producida de forma natural.

En particular, la secuencia de aminoácidos puede exhibir al menos un 65%, 70%, 75%, 80%, 85%, 90%, 95% ó 98% de identidad con el polipéptido de SEQ ID n° 2. Alternativamente, la secuencia de aminoácidos puede exhibir al menos un 70%, 75%, 80%, 85%, 90%, 95% ó 98% de similitud con el polipéptido de SEQ ID n° 2.

En particular, se pueden producir variaciones en el extremo N. Los polipéptidos también pueden incluir 4-20, por ejemplo 5-10 aminoácidos adicionales en el extremo N. Debe también señalarse que la secuencia incompleta de ADNc de SEQ ID n° 1 corresponde a la secuencia de aminoácidos de SEQ ID n° 2, pero comienza solamente en el residuo 7, ácido aspártico. En el polipéptido recombinante pueden producirse variaciones de aminoácidos, en particular sustituciones en los primeros 6 aminoácidos en comparación con la secuencia dada en la fig. 2. Además, el residuo lisina N-terminal de la fig. 2 puede ser reemplazado por arginina.

Se puede generar una secuencia de ADNc correspondiente a la secuencia de aminoácidos de SEQ ID n° 2, y ésta se constituye por la SEQ ID n° 1 junto con la secuencia de nucleótidos derivada en el extremo 5'. Esta región es proporcionada también, en su mayor parte, por el cebador delantero de SEQ ID n° 7, derivada ella misma de un fragmento secuenciado de SAP nativa. Debido a la degeneración del código genético, no es posible especificar, a partir de los aminoácidos KAYWNK, la secuencia con mayor precisión que AArGCnTAyTGGAyAAAr [SEQ ID n° 19], en donde r = A o G, n = T, C, A o G e y = C o T. A la secuencia completa del ADNc correspondiente a SEQ ID n° 2, que comprende SEQ ID n° 19 y SEQ ID n° 1, se alude en esta memoria como SEQ ID n° 20.

Sin embargo, el trabajo de secuenciación llevado a cabo como parte del Ejemplo 4, ha establecido la secuencia de ADNc completa para la fosfatasa alcalina de *Pandalus borealis*. Así, se han resuelto las incertidumbres discutidas anteriormente con respecto a SEQ ID n° 19, y el ADNc que codifica KAYWNK es como sigue: AAAGCATATTG GAACAAA [SEQ ID n° 21]. Esta secuencia puede reemplazar a SEQ ID n° 19 dentro de SEQ ID n° 15 para producir,

ES 2 294 033 T3

junto con la parte de SEQ ID n° 17 que codifica la secuencia de aminoácidos NPITEED, codificando la secuencia de ADNc precisa la enzima recombinante. A esta secuencia completa se alude como SEQ ID n° 22.

La identidad o similitud de la secuencia de aminoácidos se puede determinar utilizando el programa BestFit del paquete de software de Genetics Computer Group (GCG) versión 10 de la Universidad de Wisconsin. El programa utiliza el algoritmo de homología local de Smith y Waterman con los valores por defecto: penalidad por creación del intervalo = 8, penalidad por extensión del intervalo = 2, emparejamiento medio = 2.912, mal emparejamiento medio = -2.003.

Una "parte" de la secuencia de aminoácidos de SEQ ID n° 2 o sus variantes, según se define antes, puede comprender al menos 20 aminoácidos contiguos, de preferencia al menos 30, 40, 50, 70, 100, 150, 200, 300, 400 ó 450 aminoácidos contiguos. Partes con más de 200, en particular más de 300 aminoácidos pueden considerarse "partes significativas", y partes de este tipo tendrán típicamente el sitio de actividad enzimática tal como se identifica en esta memoria y, de preferencia, el sitio de unión del co-factor.

El polipéptido recombinante es, de preferencia, funcionalmente activo de acuerdo con las definiciones antes dadas, por ejemplo es enzimáticamente activo. Alternativamente, puede no ser por sí mismo funcionalmente activo, pero puede proporcionar regiones con propiedades funcionales del conjunto, por ejemplo puede representar el sitio activo o el sitio de unión del co-factor requerido para la actividad enzimática.

Como se ha discutido antes, la SEQ ID n° 1 y la SEQ ID n° 2 no corresponden a la secuencia completa de ADNc o de aminoácidos de la SAP nativa. Los polipéptidos truncados que exhiben la misma o esencialmente la misma actividad enzimática que la SAP nativa y conservan esa inestabilidad térmica de la enzima constituyen un aspecto adicional de la presente invención. Se pueden obtener las moléculas de ADNc completas o los polipéptidos recombinantes que incorporan las secuencias truncadas definidas en esta memoria o variantes funcionalmente equivalentes de las mismas.

Se puede obtener un polipéptido con la misma o esencialmente la misma actividad catalítica y propiedades térmicas que SAP, pero que carece de su región N-terminal. Preferiblemente, estos polipéptidos carecen del aminoácido N-terminal y 2-50, de preferencia 5-30, por ejemplo 5-10 aminoácidos que están adyacentes al mismo en la enzima nativa. Por "esencialmente la misma actividad catalítica" se quiere dar a entender que el polipéptido tiene al menos un 75%, preferiblemente al menos un 85%, más preferiblemente al menos un 90% de la actividad de SAP nativa. Aparte de estas partes N-terminales ausentes, los polipéptidos tienen preferiblemente al menos un 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% ó 98% de identidad con el polipéptido de SEQ ID n° 2.

Una fosfatasa alcalina se puede considerar térmicamente inestable si su actividad catalítica es indetectable después de pre-incubación a 65°C durante 15, preferiblemente 10 minutos.

Un ensayo adecuado para la actividad catalítica de enzimas de la invención sigue al de Olsen *et al.* [R.L. Olsen, K. Øverbø y B. Myrnes, (1991) Alkaline phosphatase from hepatopancreas of shrimp (*Pandalus borealis*): a dimeric enzyme with catalytically active subunits. Comp. Biochem. Physiol. 99B: 755-761]. La actividad de fosfatasa alcalina se determina incubando la enzima con 6 mM de fosfato de p-nitrofenol a 37°C en tampón glicina/NaOH 0,1 M pH 10,4 con 1 mM de MgCl₂ y 1 mM de ZnCl₂ en un volumen total de 0,5 ml. Al cabo de 15 min, la reacción se detiene mediante 0,5 ml de NaOH 2M y la cantidad de p-nitrofenol liberada se determina midiendo la absorbancia a 405 nm ($\epsilon_{405} = 18,5 \text{ mM}^{-1} \text{ cm}^{-1}$). Una unidad de actividad de la enzima producirá 1 μmol de p-nitrofenol por min. La SAP no se ve negativamente afectada por el zinc, y no existe una demanda significativa de magnesio por encima de una concentración mínima.

Variantes de SEQ ID n° 1 o SEQ ID n° 2 que codifican o comprenden proteínas que conservan y exhiben las propiedades catalíticas de SAP, pero que son térmicamente inestables - es decir, no padecen una pérdida completa de actividad después de calentamiento a 70°C durante 15 min.

También se proporcionan métodos para preparar moléculas de ácidos nucleicos recombinantes de acuerdo con la invención, que comprenden insertar las moléculas de ácidos nucleicos que contienen las secuencias de nucleótidos de la invención en otra molécula de ácido nucleico, por ejemplo en un ácido nucleico vector, por ejemplo ADN vector.

Así, la invención proporciona, además, un vector de expresión recombinante, que es un ADN vector que posee un fragmento de ADN que comprende una cualquiera de las moléculas de nucleótidos de la invención.

En el vector recombinante, el ADN vector es capaz de propagación en una célula hospedante tal como *Escherichia coli*, *Saccharomyces cerevisiae* o *Pichia pastoris*.

Vectores de expresión de la invención pueden incluir secuencias control apropiadas tales como, por ejemplo, elementos de control de la traducción (por ejemplo codones de inicio y detención, sitios de unión ribosomales) y de la transcripción (por ejemplo regiones de promotor-operador, secuencias de detención de la terminación), enlazados en un marco de lectura de emparejamiento con las moléculas de ácidos nucleicos de la invención.

Vectores de acuerdo con la invención pueden incluir plásmidos y virus (incluidos virus tanto bacteriófagos como eucarióticos), de acuerdo con técnicas bien conocidas y documentadas en la técnica, y se pueden expresar en una diversidad de sistemas de expresión diferentes, también bien conocidos y documentados en la técnica.

Se conoce una diversidad de técnicas y se pueden utilizar para introducir vectores de este tipo en células procarióticas o eucarióticas para la expresión, o en células de la línea germinal o somáticas para formar animales transgénicos. Técnicas de transformación o transfección adecuadas se describen bien en la bibliografía.

Una construcción genética que comprende una molécula de ácido nucleico de la invención según se define en esta memoria, operativamente enlazada a una secuencia de promotor, constituye un aspecto adicional de la presente invención.

La invención incluye también células hospedantes procarióticas o eucarióticas, transformadas o transfectadas, u organismos transgénicos no humanos que contienen una molécula de ácido nucleico recombinante de acuerdo con la invención según se define antes. Por conveniencia, en esta memoria no se hace distinción alguna entre transformación y transfección. Células hospedantes de este tipo pueden incluir, por ejemplo, células procarióticas tales como *E. coli*, *Streptomyces* y otras bacterias, células eucarióticas tales como levaduras (por ejemplo, *Pichia pastoris*) o el sistema de baculovirus-células de insectos, células de mamíferos transformadas y animales transgénicos no humanos y plantas. Se prefieren células hospedantes procarióticas.

Así, la invención proporciona también un transformante, que es una célula transformada con el vector recombinante de la invención.

En un aspecto adicional, la presente invención proporciona una célula u organismo recombinante que tiene una molécula de ácido nucleico que comprende una secuencia de nucleótidos según se recoge en SEQ ID nº 1, y que expresa o es capaz de expresar una fosfatasa alcalina térmicamente inestable. La célula u organismo es "recombinante" debido a que contiene material genético (en este caso que codifica una fosfatasa alcalina térmicamente inestable) que no es nativo para ese organismo, así se excluye el camarón *Pandalus borealis*, ya que expresa una fosfatasa alcalina térmicamente inestable, el gen, por lo tanto, se encuentra de forma natural en esa especie.

Visto de manera alternativa, la célula u organismo que expresa una fosfatasa alcalina térmicamente inestable puede considerarse "modificado", en el sentido de que dicha célula u organismo de ese genotipo o fenotipo no existe en la naturaleza. Así, la célula u organismo puede también denominarse una célula u organismo que se produce de forma "no natural", debido a la presencia en él de material genético no nativo y su capacidad de producir una proteína (una fosfatasa alcalina térmicamente inestable) no producida por miembros de esa especie o tipo de célula que se produce de forma natural.

Las células y organismos antes descritos se pueden denominar transformantes, debido a que ellos o sus antepasados han sido transformados por introducción de material genético que codifica una fosfatasa alcalina térmicamente inestable tal como se describe en esta memoria.

La invención también proporciona otro transformante, que es *Escherichia coli* o *Saccharomyces cerevisiae* transformado con el vector recombinante. Cepas de bacterias, levaduras y otras células cultivadas de expresión, adecuadas, son conocidas en la técnica.

La invención también proporciona un procedimiento para producir una fosfatasa alcalina térmicamente inestable, que comprende cultivar los transformantes de la invención en un medio de cultivo y aislar la fosfatasa alcalina a partir del transformante cultivado. Más particularmente, la invención proporciona un método para la fabricación de una fosfatasa alcalina térmicamente inestable, que comprende transformar una célula hospedante con un vector de expresión que comprende una secuencia de nucleótidos según se recoge en SEQ ID nº 1 o la secuencia complementaria de dicha secuencia, mantener la célula en un medio de cultivo y aislar la fosfatasa alcalina térmicamente inestable expresada por dicha célula. Métodos para cultivar células son bien conocidos en la técnica y, típicamente, permitirá una multiplicación de las células y una expresión, simultánea o subsiguiente, del gen ALP. El producto génico se puede acumular intracelular o extracelularmente. Métodos para recolectar y purificar proteínas generadas de esta manera son bien conocidos en la técnica.

Se ha realizado un trabajo adicional y se ha refinado la información precisa de la secuencia. Los cambios en las secuencias de ADNc y proteína revisadas son mínimos. La secuencia de ADNc revisada se muestra en la fig. 4 y se llama SEQ ID nº 15. Al igual que incorpora una secuencia 5' adicional (cuya secuencia de aminoácidos equivalente estaba siempre presente en SEQ ID nº 2), carece de 9 nucleótidos tal como se muestra en la comparación de las secuencias en la figura 7. Según se compara, las secuencias tienen una identidad de 99,4%.

A nivel de proteína, existe un cambio ligeramente mayor, ya que las correcciones creaban un cambio en el marco de lectura en una pequeña parte de la proteína. La secuencia de aminoácidos revisada se muestra en la figura 5 y se denomina SEQ ID nº 16. En la figura 8 se muestra una comparación con la secuencia de aminoácidos original. En comparación, las secuencias son aproximadamente un 96% idénticas.

Así, todas las referencias que figuran antes en esta memoria a SEQ ID nº 2 pueden ser reemplazadas por SEQ ID nº 16.

Además, se ha dilucidado recientemente una secuencia para el ADNc completo de SAP y, así, la molécula completa corresponde a SEQ ID nº 17, seguida de la SEQ ID nº 15, y la secuencia completa de aminoácidos de SAP (más la secuencia señal) corresponde a SEQ ID nº 18, seguida de la SEQ ID nº 2 ó 16. Se piensa que de los 47 aminoácidos de SEQ ID nº 18, los últimos 7 aminoácidos (NPITEED) o algunos de ellos, por ejemplo PITEED o TEED, se encuentran en el extremo N de la proteína SAP madura. Se cree que los aminoácidos restantes sean un polipéptido señal. La fosfatasa alcalina recombinante tiene preferiblemente la secuencia de aminoácidos NPITEED en su extremo N.

La invención se describirá ahora adicionalmente en los siguientes ejemplos no limitantes, en que:

figura 1 muestra la secuencia de ADNc que codifica SAP procedente de *Pandalus borealis* (menos una pequeña parte en el extremo 5'). Están subrayadas la secuencia del codón de detención y la secuencia del ADNc situada más abajo, que incluyen la cola poliA;

figura 2 muestra la secuencia de aminoácidos codificada por el ADNc de SAP mostrado en la figura 1 (e incorpora una región N-terminal adicional). Están subrayadas y especificadas homologías con los fragmentos de proteínas secuenciados. La secuencia del sitio activo de la enzima (de acuerdo con fosfatasas alcalinas conocidas) se muestra en negrillas; y

figura 3 muestra la secuencia de SAP alineada a sus homólogos de puntuación superior de ALPs no específicas para el tejido encontradas en la base de datos Swiss-Prot; ratón (número de acceso P09242), ser humano (n.a. P05186), pollo (n.a. Q92058) y una ALP del gusano de seda *Bombyx mori* (n.a. P29523).

figura 4 muestra la SAP codificadora del ADNc revisado de *Pandalus borealis*. Está indicada la cola poliA (A)₂₃. Cinco posiciones de nucleótidos en la secuencia 5' (minúsculas) están desviadas por el cebador de PCR degenerado utilizado para la amplificación y secuenciación y, así, las posiciones no están especificadas.

figura 5 muestra la secuencia de aminoácidos codificada por el ADNc de SAP mostrado en la figura 4. Homologías con los fragmentos de proteína secuenciados están subrayadas y especificadas. La secuencia del sitio activo de la enzima propuesta se muestra en negrillas. El residuo lisina N-terminal (minúsculas) podría ser, alternativamente, arginina, debido a la desviación del cebador de PCR degenerado, como se comenta con respecto a la figura 4 anterior.

figura 6 corresponde a la figura 3 anterior, pero incorpora la secuencia de aminoácidos revisada de la figura 5.

figura 7 muestra una alineación de la secuencia de ADNc revisada de la figura 4 (parte superior) con la secuencia original de la figura 1 (que comprende adicionalmente el extremo 5' que originariamente sólo se presenta en la secuencia de aminoácidos equivalente de la figura 2). La comparación se efectuó utilizando el programa de Needle, que utiliza el algoritmo de alineación global de Needleman-Wunsch, para encontrar la alineación óptima (incluidos los huecos) de dos secuencias. [Needleman *et al.* J. Mol. Biol. (1983) 48, págs. 443-453], matriz Blossum 62 con penalizaciones de los huecos de Hueco abierto = 10,0 y Hueco extendido = 0,5. El % de identidad global = 99,4%.

figura 8 muestra una secuencia de alineación de la secuencia de aminoácidos revisada de la figura 5 (parte superior) con la secuencia original de la figura 2. La comparación se efectuó utilizando el programa de Needle antes descrito, matriz Blossum 62 con penalizaciones de los huecos de Hueco abierto = 10,0 y Hueco extendido = 0,5. El % de identidad global = 96,03%.

figura 9 da la secuencia de ADNc [SEQ ID nº 17] y de aminoácidos [SEQ ID nº 18] de la secuencia señal propuesta y la región N-terminal de fosfatasa alcalina de *Pandalus borealis*. La secuencia de aminoácidos de SAP continúa entonces con KAYWNK....., según se describe en las figuras 2 y 4.

figura 10 es una foto de una SDS-PAGE de extractos de proteínas de células totales procedentes de los cultivos de expresión del Ejemplo 4, en que: pistas 1 y 10: patrón de proteínas de intervalo amplio, pistas 2 y 9: SAP 1,5 µg, pista 3: cultivo control negativo, no inducido, pista 4: cultivo control negativo, inducido, pista 5: cultivo 1 de SAP, no inducido, pista 6: cultivo 1 de SAP, inducido, pista 7: cultivo 2 de SAP, no inducido, pista 8: cultivo 2 de SAP, inducido. Las proteínas en el patrón de proteínas de intervalo amplio son: miosina (200 kD), β-galactosidasa (116 kD), fosforilasa b (97,4 kD), albúmina de suero (66 kD), ovoalbúmina (45 kD), anhidrasa carbónica (31 kD), inhibidor de tripsina (21,5 kD), lisozima (14,4 kD) y aprotinina (6,5 kD).

figura 11 es una foto de un inmunoborrón de extractos de proteínas de células totales procedentes de los cultivos de expresión del Ejemplo 4, en que: pistas 1 y 8: SAP 50 ng, pista 2: cultivo control negativo, no inducido, pista 3: cultivo control negativo, inducido, pista 4: cultivo 1 de SAP, no inducido, pista 5: cultivo 1 de SAP, inducido, pista 6: cultivo 2 de SAP, no inducido, pista 7: cultivo 2 de SAP, inducido.

Ejemplos

Ejemplo 1

5 *Secuenciación de la proteína SAP y aislamiento del gen SAP*

Aislamiento, fragmentación y análisis de la secuencia de la proteína

10 Fosfatasa alcalina de camarón (SAP) se purificó hasta una homogeneidad aparente según se describe [R. Olsen *et al.*, 1991]. Una banda de proteína sencilla de aproximadamente 55 kDa se detectó mediante tinción con coomassie, así como tinción de plata tras electroforesis en un sistema de gel Bis-Tros NuPAGE al 10% (Novex), utilizando un tampón de desplazamiento de ácido 2-(N-morfolino)etanosulfónico 50 mM que contenía SDS al 0,1%.

15 Se liofilizó 1 mg de proteína y se sometió a un servicio de análisis comercial de la secuencia (Innovagen, Suecia), en que la proteína se fragmentó mediante tripsina y los fragmentos se separaron mediante HPLC de fase inversa. Se produjeron más de 30 fragmentos, y 4 fragmentos se seleccionaron para el análisis adicional. El análisis de la secuencia, mediado por espectrometría de masas, de los fragmentos seleccionados (H23:18, H23:30, 5ReRP6:26, 5ReRP6:17) produjo 12, 10, 8 y 5 aminoácidos en secuencia, respectivamente.

20 *Síntesis de oligonucleótidos derivados de secuencias de fragmentos de proteínas*

Basándose en las secuencias de aminoácidos se predijeron codones estándares y subsiguientemente se hicieron a medida desoxirribonucleótidos degenerados de secuencias complementarias delanteras e inversas (Eurogentec, Bélgica).

25 *Síntesis de ADNc de camarones*

30 Se diseccionaron camarones de *P. borealis* recién recogidos y los hepatopáncreas individuales se almacenaron en nitrógeno líquido hasta su uso. Se aisló ARNm de un hepatopáncreas sencillo mediante el uso del sistema PolyAtract® 1000 (Promega). Se utilizó ARNm aislado para la síntesis de la primera cadena de ADNc, antes de efectuar la síntesis de la segunda cadena y la amplificación de ADNc, siguiendo las instrucciones dadas en el kit de síntesis de ADNc por PCR Smart™ (Clontech) y el kit de PCR para ADNc Advantage (Clontech).

35 *Amplificación por PCR y análisis de la secuencia de un fragmento interno del gen SAP*

Se utilizó una pequeña parte alícuota del ADNc sintetizado como molde en PCRs cebadas mediante combinaciones por pares (direcciones delantera e inversa) de los oligonucleótidos derivados de proteínas. Las reacciones de amplificación, con composiciones estándares de mezcla, se efectuaron en un termociclador de gradiente de Eppendorf. Los productos de la PCR se detectaron después de electroforesis en gel de agarosa. La combinación de cebadores de los oligonucleótidos 17F y 26R dieron un producto de amplificación distinto de aproximadamente 600 pb, el cual se secuenció utilizando el kit de secuenciación del ciclo terminador radio marcado Thermo Sequenase (Amersham), explotando los mismos cebadores de la PCR para la secuenciación.

45 *Amplificación rápida de extremos de ADNc (RACE - siglas en inglés)*

Basándose en la secuencia de ADN encontrada en el producto de PCR, se sintetizaron nuevos cebadores delantero e inverso para las reacciones de RACE 5' y 3', según se describe para el uso con el kit de amplificación de ADNc Maratón™ (Clontech) en la amplificación a partir del molde de ADNc.

50 Mediante el uso del cebador delantero MalpF, específico para el gen SAP, en combinación con el cebador API contenido en el kit, que tenía una identidad de la secuencia con respecto al adaptador ligado, se produjo un fragmento 3'-RACE de 1,4 kb. De manera similar, el cebador MalpR, específico para el gen inverso, dio origen a un producto 5'-RACE de 0,5 kb. El análisis de la secuencia de ADN del producto 3'-RACE confirmó que el producto RACE solapaba al producto de la PCR de 600 pb. El producto 3'-RACE se subclonó en el vector pCR-Script (Stratagene) para producir el plásmido pMalpF para la secuenciación ulterior del gen ADNc situado más abajo de la región de la secuencia contenida en el producto de PCR de 600 pb.

El producto de PCR de 600 pb y el inserto pMalpF se secuenciaron varias veces en ambas direcciones utilizando nuevos cebadores específicos para genes.

60 *Amplificación por PCR de alta fidelidad del gen completo de ADNc*

Se realizó una reacción PCR final utilizando los cebadores SapF y SapR y la Pfu polimerasa (Stratagene) y ADNc como molde para un producto de amplificación del gen de ADNc de SAP de 1,5 kb para la confirmación de la secuencia obtenida. Se encontró que el gen de ADNc contenía un marco de lectura abierto que codifica un polipéptido de 478 aminoácidos y una secuencia 3' que incluye una señal putativa para la poliadenilación del transcrito - véanse las figuras 1 y 2.

ES 2 294 033 T3

Perfiles de ciclación térmica utilizados en la PCR

Se utilizaron los siguientes perfiles de ciclación para reacciones de amplificación:

5			
	-	amplificación total de ADNc	1 x: 95°C, 2 min
10			16x: 95°C, 5 s
			65°C, 5 s
			68°C, 6 min
15	-	PCR de fragmento interno de 600 pb	1x: 94°C, 2 min
			36x: 94°C, 10 s
20			51°C, 10 s
			72°C, 1 min
25			1x: 72°C, 5 min
	-	5'-RACE	1x: 94°C, 30 s
30			5x: 94°C, 5 s
			72°C, 4 min
35			1x: 94°C, 30 s
			5x: 94°C, 5 s
			70°C, 4 min
40			1x: 94°C, 30 s
			23x: 94°C, 5 s
			68°C, 4 min
45	-	3'-RACE	1x: 94°C, 30 s
			36x: 94°C, 5 s
			68°C, 4 min
50	-	PCR de alta fidelidad	1x: 94°C, 3 min
			40x: 94°C, 10 s
55			55°C, 15 s
			72°C, 3 min
60			1x: 72°C, 5 min

Resultados

Se estimó originalmente que la subunidad de proteína SAP tenía una masa molecular de 65 kDa. La hoja de producto USB describe una proteína SAP de 59 kDa, y en el sistema SDS-PAGE de los inventores, la SAP migra de una manera similar a una proteína de 55 kDa de peso molecular estándar. Las discrepancias se explican probablemente por variaciones en sistemas tampón y en las cualidades del gel utilizadas.

ES 2 294 033 T3

El ADNc aislado codifica un polipéptido con identidades de secuencia a los fragmentos analizados de la proteína SAP - véase la figura 2.

La secuencia de la proteína SAP se utilizó como secuencia de consulta en búsquedas de homología en bases de datos disponibles al público. Los programas de software BLAST y ClustalW se utilizaron para la búsqueda de homología y alineaciones de múltiples secuencias, respectivamente. La secuencia de SAP se alineó a sus homólogos de puntuación superior de ALPs no específicas para el tejido encontradas en ratón (número de acceso P09242), ser humano (n.a. P05186), pollo (n.a. Q92058) y una ALP del gusano de seda *Bombyx mori* (n.a. P29523). La alineación se muestra en la figura 3. No se encuentra ningún equivalente de secuencia del ADNc mediante búsquedas de homología en bases de datos disponibles al público. Sin embargo, la puntuación de la secuencia de proteína derivada es relativamente elevada con un cierto número de ALPs conocidas con una identidad del 44% y una similitud del 60% en las secuencias de aminoácidos. Las homologías de las proteínas se investigaron en la base de datos SWALL (base de datos de secuencias de proteínas no redundante; Swissprot+Trembl+TremblNew) que operan con el programa Fasta3 con parámetros por defecto (matriz blosum 62 penalización de hueco abierto = -12 y penalización de la extensión del hueco = -2, y KTUP = 2), utilizando el servidor de internet del European Bioinformatics Institute. [W.R. Pearson y D.J. Lipman (1988), "Improved tools for biological sequence analysis", PNAS 85: 2444-2448; W.R. Pearson, (1990), "Rapid and sensitive sequence comparison with FASTP and FASTA", Methods in Enzymology 183:63-98].

También, la proteína codificada por el ADNc tiene el motivo VTDSAASAT, que corresponde al patrón de sitio activo de ALP definido por Prosite PS00 123, contenido en su secuencia. Así, el ADNc aislado de un transcrito de ARNm procedente del hepatopáncreas de camarón representa el gen para SAP de *P. borealis*.

Existe una discrepancia de un aminoácido entre el fragmento de proteína secuenciado 5ReRP6:26 y la correspondiente secuencia derivada de ADNc. Esto podría explicarse por variaciones alélicas en el gen para SAP o por un error en el análisis de la secuencia de proteínas.

Secuencias utilizadas en el Ejemplo 1

Fragmentos de proteínas

H23:18	DINFRYASAAP(V)K [SEQ ID nº 3]
H23:30	HLITDWLDDK [SEQ ID nº 4]
5ReRP6:26	VIMGGERR [SEQ ID nº 5]
5ReRP6:17	AYWNK [SEQ ID nº 6]

Oligonucleótidos

17F:	GCITAYTGGAAYAAR [SEQ ID nº 7]
26R:	CKYTCICCCICCATDATIAC [SEQ ID nº 8],

en que D = A + G + T

I = inosina

K = G + T

R = A + G

Y = C + T

MalpF:	ATTCGTGGGAAGAATTCGACTTTGC [SEQ ID nº 9]
MalpR:	GATCTGCCAGCCTCCTGGAACCA [SEQ ID nº 10]
SapF:	AARGCNTAYTGGAAYAARGAT [SEQ ID nº 11]
SapR:	GAAGGTAATCATCTACATCTCA [SEQ ID nº 12]

ES 2 294 033 T3

Se ha repetido el Ejemplo 1 y se ha refinado la información de la secuencia. Tal como se refleja en las figuras 4 y 5, se encontró que el gen para ADNc contenía un marco de lectura abierto que codifica una proteína de 475 aminoácidos. El ADNc de la fig. 4 tiene una masa molecular teórica de 53 kDa en comparación con una estimación de la masa de la SAP nativa purificada basada en SDS-PAGE de 54-55 kDa. A partir de las alineaciones de múltiples secuencias, se piensa que la SAP contiene 5-10 aminoácidos adicionales en su extremo N, llevando la masa molecular próxima a 54 kDa. Como se ha comentado antes, se piensa que esta región N-terminal comprende algo o la totalidad del heptapéptido NPITEED, véase también el Ejemplo 4.

Ejemplo 2

Composición de aminoácidos de la secuencia de proteínas derivada de ADNc en comparación con SAP nativa

Proteína SAP liofilizada se disolvió en HCl 6 M y se hidrolizó a 110°C durante 24 h. Después de la evaporación de HCl, la muestra se resuspendió en tampón citrato de sodio 0,2 M y se sometió a cromatografía HPLC para la identificación y la determinación del porcentaje molar de los aminoácidos en el producto hidrolizado.

Nota: este sistema no diferencia entre aspartato y asparagina o entre glutamato y glutamina. Así, se combinan las cifras para estos aminoácidos se combinan en la proteína SAP nativa. También, los números de residuos cisteína están supuestamente subestimados, ya que la proteína no se oxidó.

		% en	
		SAP nativa derivada de ADNc	
25	Alanina	8,3	9,6
	Aspartato + Asparagina	(13,7)	13,6
	Aspartato	10,0	-
30	Asparagina	3,7	-
	Glutamato + Glutamina	(9,0)	11,7
35	Glutamato	6,7	-
	Glutamina	2,3	-
40	Cisteína	1,0	0,2
	Fenilalanina	4,1	4,1
	Glicina	8,4	8,4
45	Histidina	3,7	2,8
	Isoleucina	5,2	5,4
50	Lisina	4,6	4,2
	Leucina	7,3	8,0
	Metionina	1,6	1,9
55	Prolina	2,7	3,1
	Arginina	5,2	5,2
	Serina	4,6	4,6
60	Treonina	8,7	8,5
	Valina	5,4	4,9
65	Triptófano	1,6	-
	Tirosina	4,1	3,6

ES 2 294 033 T3

Conclusión: para la mayoría de los aminoácidos las relaciones molares de cada aminoácido en la secuencia de proteínas derivada de ADNc son próximas a las relaciones encontradas en la proteína nativa purificada. Los aparentemente pequeños desacuerdos se comentan arriba.

5 Así, en comparación con la proteína SAP, el ADNc aislado codifica una proteína con una composición de aminoácidos muy similar o idéntica.

Este Ejemplo y el análisis se han repetido para la secuencia de ADNc revisada y los resultados se presentan más abajo:

10 *Composición de aminoácidos de la secuencia de proteínas derivada de ADNc en comparación con SAP nativa*

% en

SAP nativa derivada de ADNc

Alanina	8,8	9,6
Aspartato + Asparagina	-	13,6
Aspartato	10,1	-
Asparagina	3,4	-
Glutamato + Glutamina	-	11,7
Glutamato	7,1	-
Glutamina	2,3	-
Cisteína	0,8	0,2
Fenilalanina	4,0	4,1
Glicina	8,0	8,4
Histidina	3,3	2,8
Isoleucina	5,4	5,4
Lisina	4,6	4,2
Leucina	7,6	8,0
Metionina	2,1	1,9
Prolina	2,7	3,1
Arginina	4,8	5,2
Serina	4,6	4,6
Treonina	9,0	8,5
Valina	5,4	4,9
Triptófano	1,5	-
Tirosina	4,0	3,6

Ejemplo 3

Expresión recombinante de SAP en Pichia pastoris

5 La SAP que codifica ADNc se utiliza para expresar el carácter heterólogo de la enzima en un microorganismo adecuado para la producción de enzima recombinante mediante fermentación. Están disponibles varios sistemas de expresión y se prefieren sistemas de expresión eucarióticos, ya que éstos pueden expresar la enzima a niveles razonables en su estado nativo.

10 Un sistema de este tipo es la levadura metilotrófica *Pichia pastoris* (Invitrogen), de la que se ha informado que expresa proteínas recombinantes a niveles de hasta 12 g/L mediante expresión intracelular [Clare, J.J., Raiymont, F.B., Ballantine, S.P., Sreekrishna, K. y Romanos, M.A. (1991): High-level expression of tetanus toxin fragment c in *Pichia pastoris* strains containing multiple tandem integrations of the gene. *Bio/Technology* 9, 455-460] y hasta 2,6 g/L son secretados al medio. [Paifer, E., Margolles, E., Cremata, J., Montesino, R., Herrera, L. y Delgado, J.M. (1994): Efficient expression and secretion of recombinant alpha amylase in *Pichia pastoris* using two different signal sequences. *Yeast* 10, 1416-1419].

El gen insertado es regulado por el promotor AOX1 (alcohol oxidasa), de manera que la expresión puede ser inducida por metanol.

20 El ADNc de SAP es insertado en el vector pPIC9K comercialmente disponible. Después, la SAP recombinante puede ser expresada como un producto secretado mediante fusión a la secuencia de péptido señal derivada del factor α de *S. cerevisiae* contenido en el vector. Cuando se integra en una cepa *his* (KM71 o GS115), las cepas se seleccionan en cuanto a recombinantes mediante selección de clones *his* + en un medio exento de histidina, ya que el genotipo *his* es complementado por el inserto del vector. Además, se rastrean recombinantes en cuanto múltiples insertos de genes mediante la resistencia a concentraciones incrementadas de geneticina (G418, Invitrogen), adquirida por el gen de resistencia a la kanamicina portado por el plásmido.

30 En detalle, el ADNc de SAP generado por PCR se utiliza como un molde en una nueva reacción de PCR, con el fin de generar un nuevo producto de PCR que contiene sitios de restricción apropiados para la integración en el vector pPIC9K en marco con la secuencia señal de factor α . En calidad de cebador delantero para la reacción PCR se utiliza un cebador "17F" modificado [SEQ ID nº 13], que reemplaza al codón ATG (Met) por un codón lisina (AAG), con el fin de reconstruir un residuo lisina putativo escindido por tripsina. Se construye el cebador inverso [SEQ ID nº 14] con el fin de fijar un sitio de restricción *NotI* en la posición 3' con respecto al codón de detención de la secuencia de ADNc de SAP.

40 El producto de PCR resultante se corta con *NotI* y se liga en un vector pPIC9K previamente cortado con *SnaBI* y *NotI* (dando como resultado el plásmido pPIC9K-SAP). Dado que el vector termina en extremos romos por *SnaBI* en el sitio de unión 5' y los dos extremos están en el marco de lectura, no es necesaria ninguna modificación adicional en el extremo 5' del ADNc modificado.

45 El vector se propaga luego mediante transformación de una cepa de *E. coli* competente, tal como TOP10F' (Invitrogen) o equivalente, y los recombinantes se seleccionan por su resistencia a ampicilina. El plásmido pP109K-SAP aislado puede ser luego linearizado mediante corte de restricción con *SacI* con el fin de crear una casete de vector a insertar en el gen AOX1 de *Pichia pastoris*. Después de la transformación de células de *Pichia pastoris* (cepas GS115 o KM71) con una construcción pPIC9K-SAP linearizada, se seleccionan recombinantes a partir de mutantes *his* + de acuerdo con el protocolo del fabricante.

50 Colonias aisladas se testan en cuanto a la expresión de SAP recombinante en el medio mediante células en desarrollo en presencia de metanol de acuerdo con el método recomendado por el fabricante. Si se desea, se consiguen aumentos adicionales en los niveles de expresión mediante el rastreo de cepas recombinantes en cuanto a clones que contienen múltiples copias del gen recombinante. Esto se hace cultivando mutantes *his* + en presencia de niveles incrementados del antibiótico G418 (Invitrogen). Es probable que clones que se desarrollan a altos niveles de G418 contengan múltiples copias del gen.

55 Todos los métodos de cultivo, transformación y selección se describen en detalle en protocolos disponibles del suministrador del sistema de *Pichia pastoris* (Invitrogen, Holanda).

60 La construcción de SAP modificada proporciona un producto recombinante en el que se añaden 5 aminoácidos (EAEAY) a partir de la secuencia de reconocimiento de la señal restante de la peptidasa KEX2 señal de *Pichia pastoris*. Esta secuencia N-terminal resulta en un producto final que está ligeramente truncado, pero en un consenso razonable con otras fosfatasa alcalinas, que todavía contienen una secuencia de reconocimiento para la peptidasa señal de *Pichia pastoris*.

65

*Cebadores utilizados en el Ejemplo 3**Delantero*

5 AAGGCITAAYTGGAAYAAR [SEQ ID n° 13]

en que

I = Inosina

10 R = A + G

Y = C + T

15 *Inverso*

TACAGCGGCCCGCCATCTOAI1TTTCG [SEQ ID n° 14]

20 Ejemplo 4

*Expresión de SAP en Escherichia coli TOP10*25 *Determinación de la secuencia señal de SAP*

La secuencia señal N-terminal de SAP se determinó mediante 5'-RACE, utilizando un ADNc como molde. ARNm se aisló a partir de 100 mg de hepatopáncreas de camarones (*Pandalus borealis*), utilizando el Mini Kit de ARNm Oligotex Direct (Qiagen), según se describe por el fabricante. Se preparó ADNc de hepatopáncreas de camarones utilizando el kit de síntesis de ADNc por PCR SMART (Clontech), siguiendo el protocolo descrito por el fabricante. El ADNc era de extremos romos y se purificó añadiendo 40 µg de proteinasa K a 50 µl del ADNc y se incubó a 45°C durante 45 min y a 90°C durante 8 min, y luego se añadieron 15 U de T4-ADN polimerasa (Promega) y se incubaron a 16°C durante 30 min y a 72°C durante 10 min. Finalmente, el ADNc se extrajo dos veces en fenol/cloroformo y precipitó etanol. Los adaptadores de RACE se ligaron al ADNc según se describe en el kit de amplificación de ADNc Marathon (Clontech). La reacción 5'-RACE se efectuó en un volumen final de 50 µl, utilizando la mezcla de ADNc polimerasa Advantage (Clontech) y un cebador AP1 (suministrado con el kit de amplificación de ADNc Marathon) y un cebador específico para SAP (5'-GCG TGG TGC ATA TGG TCA ATC CGT CC-3') y 5 µl de ADNc ligado a adaptador 50 veces diluido como molde. La reacción PCR de RACE se realizó en un termociclador Biometra TGradient con una etapa de desnaturalización inicial de 94°C durante 30 segundos, seguida de 5 ciclos de 94°C durante 5 s y 72°C durante 3 min, 5 ciclos de 94°C durante 5 s y 70°C durante 3 min y 30 ciclos de 94°C durante 5 s y 68°C durante 3 min. El fragmento de SAP de 1200 pb generado se purificó a partir del gel de agarosa, se re-amplificó utilizando una PCR y se secuenció.

45 *Clonación de SAP en el vector de expresión pBAD/gIII A*

Para clonar el gen para SAP en el vector pBAD/gIII A, el gen para SAP se amplificó por PCR con dos cebadores que contenían un sitio SacI y HindIII, respectivamente (subrayados). Basándose en la secuencia de SAP de longitud completa (secuencia señal más de la secuencia principal), los cebadores se diseñaron para amplificar el gen SAP con los aminoácidos NPITEEDKAYWNK..... como extremo N. Además, se cambiaron tres bases en el cebador 1 (de C a A, de A a C y de A a C), con el fin de optimizar a tres de los codones N-terminales (representados con letras minúsculas en la secuencia del cebador) (cebador 1: 5'-ATG GAG CTC AAC CCa ATc ACc GAA GAA GAC AA-3' cebador 2: 5'-ATG AAG CTT TCA TTT TTC GTC ACA GAA AGT G-3'). La PCR se llevó a cabo en un volumen final de 50 µl que contenía 10 µM de cada uno de los dos cebadores, 1,5 U de Pfu-polimerasa (Promega) y tampón suministrado, dNTP 0,2 mM y 10 ng de ADNc de hepatopáncreas de camarones como molde. La PCR se hizo con una etapa de desnaturalización inicial de 94°C durante 30 s, seguida de 30 ciclos de 94°C durante 15 s, 55°C durante 1 min y 72°C durante 3 min. El producto de PCR se purificó con el kit de purificación por PCR Qiaquick (Qiagen).

Tanto el vector pBAD/gIII A como el producto de PCR de SAP se trataron con 5 U de enzimas de restricción SacI y HindIII, respectivamente. El producto de PCR de SAP se ligó en el vector en una reacción de 10 µl utilizando 1 U de T4-ADN ligasa (USB) y aproximadamente 1 mg tanto del plásmido como del producto de PCR. La mezcla de reacción se incubó a 16°C durante 16 h. La mezcla de ligamiento se utilizó luego para transformar células TOP10 de *E. coli*, añadiendo la mezcla de ligamiento a 40 µl de células TOP10 de *E. coli* electrocompetentes y se electroporaron a 1800 V. Inmediatamente después de la electroporación, se añadió 1 ml de medio SOC y se incubó a 37°C durante 1 h. Las células se extendieron luego en placas LB que contenían 100 mg/ml de ampicilina para la selección de células transformadas.

Expresión de SAP en TOP10 de E. coli

Para la expresión se seleccionaron dos colonias TOP10 de *E. coli* que contenían el vector pBAD/gIII A recombinante que contenía el gen para SAP. Como control negativo se utilizó una cepa TOP10 de *E. coli* que contenía un vector pBAD/gIII A vacío.

Tres matraces Erlenmeyer que contengan 50 ml de medio LB se inocularon con 2,5 ml de cultivos de una noche en medio LB suministrado con 100 µg/ml de ampicilina. Las células se dejaron desarrollar a 30°C con sacudimiento hasta que las células alcanzaron una densidad de $DO_{600} = 0,5$. Cada uno de los tres cultivos se escindieron luego en dos (2 x 25 ml), en que se indujo uno de los dos cultivos añadiendo L-arabinosa hasta una concentración final de 0,05% (p/v). Todos los cultivos de células se hicieron desarrollar luego a 30°C con sacudimiento durante 3 h antes de recolectar las células.

Análisis de los cultivos de expresión

De cada uno de los cultivos de expresión se centrifugó una muestra de 200 µl y los sobrenadantes se separaron. A los sedimentos de las células se añadieron 100 µl de tampón de muestra 1X SDS-PAGE y se incubaron a 100°C durante 3 min y se centrifugaron a 18.000 g durante 3 min. Los lisados de células se aplicaron luego directamente y se analizaron en un gel de SDS-PAGE. Se electromanchó ulteriormente un gel de SDS-PAGE adicional para transferir las proteínas a una membrana de manchado con nitrocelulosa (0,45 µm, Bio-Rad) y se inmuno-tiñó utilizando un anticuerpo dirigido contra SAP purificada a partir del agua de tratamiento de camarones y conjugado de IgG anti-conejo de cabra (H + L) (IgG humana adsorbida) y peroxidasa de rábano picante (Bio-Rad). Se siguieron protocolos estándares para SDS-PAGE, electromanchado e inmuno-tinción.

Como se muestra en las figuras 10 y 11, SAP recombinante se expresa utilizando el vector pBAD/gIII A y células TOP10 de *E. coli* como hospedante. Las dos cepas de *E. coli* recombinantes que contenían vectores pBAD/gIII A recombinantes producen SAP recombinante, mientras que el control negativo no daba ninguna señal en la transferencia western (figura 11).

REIVINDICACIONES

- 5 1. Una molécula de ácido nucleico que comprende una secuencia de nucleótidos según se recoge en SEQ ID nº 1 o la secuencia complementaria de dicha secuencia, en donde dicha secuencia de nucleótidos codifica o es complementaria a una secuencia que codifica una fosfatasa alcalina térmicamente inestable.
- 10 2. Una construcción genética que comprende una molécula de ácido nucleico según la reivindicación 1, operativamente enlazada a una secuencia de promotor.
3. Un vector de expresión recombinante que comprende una molécula de ácido nucleico según la reivindicación 1, y que es capaz de propagación en una célula hospedante.
- 15 4. Un vector según la reivindicación 3, que es un plásmido o vector viral.
5. Una célula transformada con una construcción o vector según una cualquiera de las reivindicaciones 2 a 4.
- 20 6. Una célula según la reivindicación 5, que es una célula procariótica.
7. Una célula según la reivindicación 5, que es parte de un cultivo de células eucarióticas.
8. Una célula recombinante u organismo no humano con una molécula de ácido nucleico recombinante que comprende una secuencia de nucleótidos según se define en la reivindicación 1.
- 25 9. Una célula recombinante u organismo de acuerdo con la reivindicación 8 que es capaz de expresar una fosfatasa alcalina térmicamente inestable, que comprende la secuencia de aminoácidos de SEQ ID nº 2.
- 30 10. Un método para la fabricación de fosfatasa alcalina térmicamente inestable, que comprende transformar una célula hospedante con un vector de expresión según la reivindicación 3 ó 4, manteniendo la célula en un medio de cultivo y aislando la fosfatasa alcalina térmicamente inestable expresada por dicha célula.
11. Un método de acuerdo con la reivindicación 10, en el que la célula hospedante es *Pichia pastoris*.

Fig.1. [SEQ I.D. NO. 1]

GATGCCCAAGATGCACTGGATAAACAGCTGGGATAAACTGCGAGAGAGCAGGCCAAGAACGTGATCTTCTCCTAGG - 80
 AGACGGGATGTCCTCACTACAGTTACCGCAGCAAGAACTCTACAAAGGTGGATTGACCGGGAATTTGAACGGGAAAGA - 160
 TTTCTGGGAAGAAATTCGACTTTGCTGCTTTGAGTAAGACGTACAAATACGGACAAGCAAGTGACAGATAATGCTGCCAGT - 240
 GCCACAGCCTACCTCACTGGGTGAAGACCAACCAAGGAGTCAATCGGCCCTTGATGCCAATACCGTGAGAACAACTGCTC - 320
 TTACCAACTTGATGAATCCCTCTTCACTTACTCCATCGCCCAATGGTTCCAGGAGGCTGGCAGATCAACAGGTGTCGTGA - 400
 CATCGACCAAGGTAACCTCACTGCTACTCCTGAGGAACCTTATGCTCACTGCTGACAGAAATTTGGAAGAACCAAGTGAT - 480
 GTTGTCCACGACAGAGAAAGATCCAGAAATATGTGATGACATAGCTGAACAACCTGGTATTCAGAGAACCTGGAAAGAACTT - 560
 CAAGGTCAATCATGGGCGCGGTGCGGTGGAATCTTCTGAGGAAAGCGCTTGTATCGAAGATGGCAATTCACGGGGAAA - 640
 GGGAAAGATGGCAACACCTCATAAACCAATTTGGTTGGACGATAAGGCATCTCAGGAGCCACAGCATCGTATGTCTGGAAC - 720
 CGAGATGATCTATTGGCTGTTGACATTCGAAACACAGACTACCTTATGGACTGTTTTCATACACACATCTAGACACAGT - 800
 TTTAACAGAGACGCCGAAATGGATCCCACTTGGCCGAAATGACAAAGGTGCGCCATCGAAATGCTGACCAAGGATGAAA - 880
 ATGGCTTTTCTCTTAGTAGAAGGTGGACGGATTGACCATATGCACCACGCCAATCAGATCCGTCACTCGTTAGCGGAG - 960
 ACGCTGGACATGGAGGAAGCTGTATCCATGGCTCTCTCTATGACAGACCCAGAGGAAACAATCATTTCTGGGTACACAGC - 1040
 TGATCACGGTCACTCTCACTCCATCACTGGATACCGCAGACAGGAACACCGGATATTTTGGATTTTCGCCGGCATCAGCGACT - 1120
 TGGATGATAGGAGGTACACCATCCTGGACTATGGCAGTGGACCTGGATATCATATCAGGAAGATGGCAACGTTACGAA - 1200
 CCTACAGAGGAAGATTTGAAAGACATCAATTTCCGCTACCGCATCGGACCGGCCCAAGCATTCAGTTACCCATGATGGCAC - 1280
 AGACGTTGGTAAACGGACCTTTTCGCTCATCTCTTCACGGTCAACGGGTGCTCTATGAAAGAGAAATACATTCCTCC - 1360
 ACGCCTTGGCCTACGCAGCCTGCGTTGGTACAGGACGCACTTTCTGTGACGAAAATGAGATGTAGATGATTAACCTTCTT - 1440
AAATACATTAAGGATTGCAAAAAAAAAAAAAAAAAAAA - 1481

Fig.2. [SEQ I.D. NO. 2]

1 KAYWNKDAQDALDKQLGKREKQAKNVIFFLGDGMSLSTVTAARIYKGG
5ReP6:17
51 LTGKFEREKISWEEFDFAALSKTYNTDKQVTDSAASATAYLTGVKTNQGV
Sito activo
101 IGLDANTVRTNCSYQLDESLFTYSIAHWFQEAGRSTGVVTSTRVTHATPA
151 GTYAHVADRDWENDSDVVHDREDPEICDDIAEQLVFREPGKNFKVIMGGG
5ReP6:26
201 RRGFFPEEALDIEDGIPGEREDGKHLITDWLDDKASQGATASYVWNRDDL
H23:30
251 LAVDIRNTDYLMGLFSYTHLDTVLTRDAEMDPTLPENTKVAIEMLTKDEN
301 GFFLLVEGGRIDHMHANQIRQSLAETLDMEESCIHGSLYDRPRGNNHFW
351 VTADHGHTLTITGYADRNTDILDFAGISDLDDRRYTILDYGS GPGYHITE
401 DGKRYEPTEEDLKDINFRYSAAPKHSVTHDGTDVGIWVNGPFAHLFTVT
H23:18
451 GVYEENYIPHALAYAACVGTGRTFCDEK detención - 478

Fig.3.

```

      *      20      40      60      *
ppbt_ratón : -----KISPLVLAIGTCLTNSVPEKERDPSTVROQCEETLKHALKQKLN---TNVAVV : 54
ppbt_ser humano: -----KISPLVLAIGTCLTNSVPEKERDPKYRDOQCEETLKVALELQKLN---TNVAVV : 54
alp_pollo : -----EKAFLTLTLLAQCSASLVPEKERDPEYERQCEETLRDALRLCHLN---QNVAVV : 53
sap : -----KAYVNDKQDALDKQLGELP-----EKQAVV : 28
ppb_bonmo : NVVSVVALLAAGLVRAEDRYHPERLAAGEASATRSAAESEAFLVREAGEAERREREGAGIKQAAHNV : 74
              W AQ AKN

      80      100     120     140
ppbt_ratón : IFLLGDCGVSIVTAARILKQQLHNHTGSETRLENDKPFVLSKTYNTNAQVPSAGTATAYLGGVNAEGTV : 128
ppbt_ser humano: IFLLGDCGVSIVTAARILKQQLHNHTGSETRLENDKPFVLSKTYNTNAQVPSAGTATAYLGGVNAEGTV : 128
alp_pollo : ILFLGDCGVSIVTAARILKQQLHNRKGSLELNDKPFVLSKTYNTNAQVPSAGTATAYLGGVNAEGTV : 127
sap : IFLLGDCGSLSHVTAARIVYHGLTG-KPREKISVEEFDFAMSKTYNTDHFQUTDSALATAYLGGVNAEGTV : 101
ppb_bonmo : VHTLGDGHSVPLLAARTLLCORRGOTGSELSLHFLCEPTLGLSTYCVNAQVPSAGTATAYLGGVNAEGTV : 148
      FLGDGH T AAR G E F L KTY QV DS ATAYL GVK NG

      160     180     200     220
ppbt_ratón : GVSALTERTRQNTTQGN--EVTSLRLKARDACKSVGINVITIRVNHATPSAAVHSHADRDVSDNENPPEALS-Q : 199
ppbt_ser humano: GVSALTERSRQNTTQGN--EVTSLRLKARDACKSVGINVITIRVNHATPSAAVHSHADRDVSDNENPPEALS-Q : 199
alp_pollo : GVSAGVTRDRQNTTKGQ--EVTSLRLKARDACKSVGINVITIRVNHATPSAAVHSHADRDVSDNENPPEALS-Q : 198
sap : GLDANTVNTNGSYQDESLEFYSIAHFTQEKRSSTGVASTVTRATPLGTYAHVADRDVSDNENPPEALS-Q : 175
ppb_bonmo : GVTAVPKNCEASTDVTIRVQSAEALADCRVGIIVITIRVNHATPSAGTATAYLGGVNAEGTV : 222
      G A R C SI W G GVT TR HA P A A R W D

      240     260     280
ppbt_ratón : GCPDIAEQVHVN--IKDIDVIRGGGRKYHFNKNTDMAYESDEKLRGTPLDGLDLSISKSPF-PKHHSVTV : 270
ppbt_ser humano: GCPDIAEQVHVN--IKDIDVIRGGGRKYHFNKNTDMAYESDEKLRGTPLDGLDLSISKSPF-PKHHSVTV : 270
alp_pollo : GCPDIAEQVHVN--IPDIEPILGGGRKYHFNKNTDMAYESDEKLRGTPLDGLDLSISKSPF-PKHHSVTV : 269
sap : ICDDIAEQVTRFPGKHFVIRGGGRGFPEEALDIDGIP---GSPEDGKHITDGLDRAAGCATASVTV : 245
ppb_bonmo : KCPDIAEQVHVN--IKDIDVIRGGGRKYHFNKNTDMAYESDEKLRGTPLDGLDLSISKSPF-PKHHSVTV : 291
      C DIL QL VI GGGR P D E R DG L W K V

      300     320     340     360
ppbt_ratón : NRTETLALDPSRVYLLGAFEGDHOYLVNWNLTDPSSLSHVVVRLRIITNLKGFLLVEGGRIDECHHEGR : 344
ppbt_ser humano: NRTETLALDPSRVYLLGAFEGDHOYLVNWNLTDPSSLSHVVVRLRIITNLKGFLLVEGGRIDECHHEGR : 344
alp_pollo : HRETLALNVSRYVFLGAFEGDHOYLVNWNLTDPSSLSHVVVRLRIITNLKGFLLVEGGRIDECHHEGR : 343
sap : HRETLALNVSRYVFLGAFEGDHOYLVNWNLTDPSSLSHVVVRLRIITNLKGFLLVEGGRIDECHHEGR : 319
ppb_bonmo : HRETLALNVSRYVFLGAFEGDHOYLVNWNLTDPSSLSHVVVRLRIITNLKGFLLVEGGRIDECHHEGR : 365
      R LL D L GLF L P L E V L L GFFL VEGGRIDH HH

      380     400     420     440
ppbt_ratón : AKQAEHVAIVENDQAIQKAGNTSOK-DTLTWVTDSESVFTTGGITPCHNSIFGLAPHVSDDKKPFLLILGN : 417
ppbt_ser humano: AKQAEHVAIVENDQAIQKAGNTSSE-DTLTWVTDSESVFTTGGITPCHNSIFGLAPHVSDDKKPFLLILGN : 417
alp_pollo : AKQAEHVAIVENDQAIQKAGNTSPR-DTLTWVTDSESVFTTGGITPCHNSIFGLAPHVSDDKKPFLLILGN : 416
sap : IQRSQETLDMIESCTHGSLYDRPGMNFVTDSESVFTTGGITPCHNSIFGLAPHVSDDKKPFLLILGN : 392
ppb_bonmo : AMLALDETIENDRAVKVATDALXED-ESLVVTDSESVFTTGGITPCHNSIFGLAPHVSDDKKPFLLILGN : 437
      L E VTADH H GY R D Y

      460     480     500     5
ppbt_ratón : GPGYKVVD-CER-ENVSVDYAHNNYQASVPLRHEHGGEDVAVFVSPHALLHG--VNSQNTYIFVHAYV : 487
ppbt_ser humano: GPGYKVVD-CER-ENVSVDYAHNNYQASVPLRHEHGGEDVAVFVSPHALLHG--VNSQNTYIFVHAYV : 487
alp_pollo : GPGYKIVG-CER-ENVSVDYAHNNYQASVPLRHEHGGEDVAVFVSPHALLHG--VNSQNTYIFVHAYV : 486
sap : GPGYHITD-CER-YEPTEDLKDINFRTASLLPKHSVTFDGTGVVNGPFAHLFTVTGVYENYIFHALLAYV : 465
ppb_bonmo : GPGYKIVG-CER-ENVSVDYAHNNYQASVPLRHEHGGEDVAVFVSPHALLHG--VNSQNTYIFVHAYV : 509
      GPG G R P TH G DV G E PH A A

      20      540
ppbt_ratón : SCIGAMLDHAIAGSGSIPSPGALLPLAVLSLPTLF- : 524
ppbt_ser humano: SCIGAMLDHAIAGSGSIPSPGALLPLAVLSLPTLF- : 524
alp_pollo : SCIGPNRAHSSAARPAATATLPLVLLLLLLC----- : 519
sap : ASVGTGRTFDEK----- : 478
ppb_bonmo : ASHGPGRHVVSAAVPTALLSLLLAFTTLRHQCFL : 547
      C G C

```

Fig.4.

AArGCnTAyTGGAAYAArGATGCCCAAGATGCACTGGATAAACAGCTGGGGATAAAACTG 60
 CGAGAGAAGCAGGCCAAGAACGTGATCTTCTTCCTAGGAGACGGGATGTCCCTCAGTACA 120
 GTTACCGCAGCAAGAATCTACAAAGGTGGATTGACCGGGAAATTTGAACGGGAAAAGATT 180
 TCGTGGGAAGAATTCGACTTTGCTGCTTTGAGTAAGACGTACAATACGGACAAGCAAGTG 240
 ACAGATAGTGCTGCCAGTGCCACAGCCTACCTCACTGGGGTGAAGACCAACCAGGGAGTC 300
 ATCGGCCTTGATGCCAATACCGTGAGAACAACTGCTCTTACCAACTTGATGAATCCCTC 360
 TTCACTTACTCCATCGCCCATTTGGTTCCAGGAGGCTGGCAGATCAACAGGTGTCTGACA 420
 TCGACCAGGGTAACCTCATGCTACTCCTGCAGGAACTTATGCTCACGTGGCTGACAGAGAT 480
 TGGGAGAACGACAGTGATGTTGTCCACGACAGAGAAGATCCAGAAATATGTGATGACATA 540
 GCTGAACAACTGGTATTCAGAGAACC TGAAAGA ACTTCAAGGTCATCATGGGCGGCGGT 600
 CGGCGTGGAATTCCTTCTCTGAGGAAGCGCTTGATATCGAAGATGGCATTCAGGGGAAAGG 660
 GAAGATGGCAAACACCTCATAACCGATTGGTTGGACGATAAGGCATCTCAGGGAGCCACA 720
 GCATCGTATGTCTGGAACCGAGATGATCTATTGGCTGTTGACATTCGAAACACAGACTAC 780
 CTTATGGGACTGTTTTTCATACACACATCTAGACACAGTTTTTAACCAGAGACGCCGAAATG 840
 GATCCCACCTTGCCGGAAATGACAAAGGTGCGCCATCGAAATGCTGACGAAGGATGAAAAT 900
 GGCTTTTTCTCTCTTAGTAGAAGGTGGACGGATTGACCATATGCACCACGCCAATCAGATC 960
 CGTCAGTCGTTAGCGGAGACGCTGGACATGGAGGAAGCTGTATCCATGGCTCTCTCTATG 1020
 ACAGACCCAGAGGAAACAATCATTCTGGTCACAGCTGATCACGGTCACACTCTCACCATC 1080
 ACTGGATACGCAGACAGGAACACCGATATTTTGGATTTGCGCGGCATCAGCGACTTGGAT 1140
 GATAGGAGGTACACCATCCTGGACTATGGCAGTGGACCTGGATATCATATCACGGAAGAT 1200
 GGCAAACGTTACGAACCTACAGAGGAAGATTTGAAAGACATCAATTTCCGCTACGCATCG 1260
 GCAGCGCCCAAGCATTCAAGTTACCCATGATGGCACAGACGTTGGTATTTGGGTCAACGGA 1320
 CCTTTCGCTCATCTCTTCACGGGTGTCTATGAAGAGAATTACATTTCCCACGCCCTTGGCC 1380
 TACGCAGCCTGCGTTGGTACAGGACGCACTTTCTGTGACGAAAAATGAGATGTAGATGAT 1440
 TACCTTCTTAAGTACATTAAAGATTGC - (A)₂₃ 1490

Fig.5.

1 KAYWNKDAQDALDKQLGIKLRKQAKNVIFFLGDGMSLSTVTAARIYKGG
5ReRP6:17
51 LTGKFEREKISWEEFDFAALSKTYNTDKQVTDSAASATAYLTGVKTNQGV
Sitio activo
101 IGLDANTVRTNCSYQLDESLFTYSIAHWFQEAGRSTGVVTSTRVTHATPA
151 GTYAHVADRDWENDSDVVHDREDPEICDDIAEQLVFREPGKNFKVIMGGG
5ReRP6:26
201 RRGFFPEEALDIEDGIPGEREDGKHLITDWLDDKASQGATASYVWNRDDL
H23:30
251 LAVDIRNTDYLMGLFSYTHLDTVLTRDAEMDPTLPENTKVAIEMLTKDEN
301 GFFLLVEGGRIDHMHANQIRQSLAETLDMEEAVSMALSMTDPEETIILV
351 TADHGHTLTITGYADRNTDILDFAGISDLDDRRYTILDYGS GPGYHITED
401 GKRYEPTTEEDLKDINFRYASAAPKHSVTHDGTDVGIWVNGPFAHLFTGVY
H23:18
451 EENYIPHALAYAACVGTGRTFCDEK *

Fig.6.

```

      *           20           *           40           *           60           *           80
alp_ratón      : MISPFVLVAICT-----CLTNSFVPEKIDDPSTYRQQAQNTLRKNALELQK---LNTNVAKEVIMFLCD : 60
alp_ser humano : MISPFVLVAICT-----CLTNSLVPRKENDPKYTRDQAQNTLRKVALELQK---LNTNVAKEVIMFLCD : 60
alp_pollo      : -MRAPLTLTLLAQ-----LCSASLVPRERKDPKTYRQQAQNTLRDALRIQK---LNTNVAKEVIMFLCD : 59
SAP_camarón    : -----KATENKADQALDKQLCIX-----LREKQAKGVIMFLCD : 34
alp_bcamo      : HVVSVVAAAAAGLVRAEDRYHPFRLAACRASAAATSAASFSASFVRBAQRAITRRERICAGAKQAACRAKGVIMFLCD : 80
                  U      AQ                                ARK      FLCD

      *           100          *           120          *           140          *           160
alp_ratón      : CMCVSTVTAAITLRQQLHENTCETRLKNDKFPFVALSKGYNTNAQVPSACTATATCCVKANCTGVVSAAIERTRCN : 140
alp_ser humano : CMCVSTVTAAITLRQQLHENTCETRLKNDKFPFVALSKGYNTNAQVPSACTATATCCVKANCTGVVSAAIERTRCN : 140
alp_pollo      : CMCVSTVTAAITLRQQLHENTCETRLKNDKFPFVALSKGYNTNAQVPSACTATATCCVKANCTGVVSAAIERTRCN : 139
SAP_camarón    : CMLSTVTAAITTRCCLTC-KFIREKISVVEEDPAALSKGYNTNDKQVTSAAASAPAYDTGVCTMGCVILDAHTVTRNES : 113
alp_bcamo      : CMSVPLLAATYLLCGRRGCTGCEASLHFPQPTLCLARTTCVNAQVPSACTATATCCVKANCTGVVTAAPVPHDCI : 160
                  CH  T  AAR  C      R      F      L  KTT  QV  DS  ATATL  CVK  NC  G  A  R  C

      *           180          *           200          *           220          *           240
alp_ratón      : TTQGNVIT--SILRKAADACKSVGIIVTLFQVNLTPSAAYAHSAADQVTSQNMHPPIALS-QGCKDIATQIMHNTK--DI : 215
alp_ser humano : TTQGNVIT--SILRKAADACKSVGIIVTLFQVNLTPSAAYAHSAADQVTSQNMHPPIALS-QGCKDIATQIMHNTK--DI : 215
alp_pollo      : TTQGNVIT--SILRKAADACKSVGIIVTLFQVNLTPSAAYAHSAADQVTSQNMHPPIALS-QGCKDIATQIMHNTK--DI : 214
SAP_camarón    : YOLDKSLFTVSTAHNPQAGRSTGVVTSITVTELTACITAHVADQVENDSDVVEDKDPKITDDIARQIVTRKPGKNT : 193
alp_bcamo      : ASTDVTKRVSGLALALADKDVCTVTEFRTIASPACTFAKVANNKENDNDVRQICHQDVNRCPDIAHQITKMAPGNKT : 240
                  SI  W  C  CVT  TR  HA  P  A  A  R  W  D      C  DIA  QL

      *           260          *           280          *           300          *           320
alp_ratón      : DVIKCCGRKYHYPKHRTDVYELD KVARCTELDGLDLSIRSVTP-RKHSHYVNTITLHMDPSRVDTYLCLETPCD : 294
alp_ser humano : DVIKCCGRKYHYPKHRTDVYISD KPARCTELDGLDLDVTSRVP-RKHSHYVNTITLHMDPSRVDTYLCLETPCD : 294
alp_pollo      : IVILCCGRKYHYPKHRTSDVTPQERREKCTELDGLDGLVQAMHDTKP-ACKVAKTVVHRRRLALNVSRTDYLCLLETPCD : 293
SAP_camarón    : KVINCCGRKCFPPHIALDIDCIP-----GRMDCKHLLTDLLDQASQATASYVNRDDIADVDIRNTDYINCLFSTYH : 269
alp_bcamo      : KVIKCCGRREFLPTTQVPER-----GTRGLRTDGRLLINWQNDKSSQKVSXKYLUNRQILHRLCSSPPDYTLCLFICSH : 315
                  VI  CGGR  P  D  K      G  R  DG  L  W  K      W  R  LL      D  L  CLF

      *           340          *           360          *           380          *           400
alp_ratón      : HQYELRNHNLTPSLSENVEVALRIETKMLKCTFLVIGGRIDHGHGKAKQALHIAVENDQALCKACANTSQKDLTV : 374
alp_ser humano : HQYELRNHNLTPSLSENVEVALRIETKMLKCTFLVIGGRIDHGHGKAKQALHIAVENDQALCKACANTSQKDLTV : 374
alp_pollo      : HQYELRNHNLTPSLSENVEVALRIETKMLKCTFLVIGGRIDHGHGKAKQALHIAVENDQALCKACANTSQKDLTV : 373
SAP_camarón    : LDVLTDAEDHETLPETKVALHETLTDHETFLVIGGRIDHGHGKAKQALHIAVENDQALCKACANTSQKDLTV : 349
alp_bcamo      : LOYELKGDSTETLHALLTDVATRVLSRWERCTFLVIGGRIDHGHGKAKQALHIAVENDQALCKACANTSQKDLTV : 395
                  L      P  L  E  VA  L      GFPL  VIGGRIDH  HH      L  E      A  A

      *           420          *           440          *           460          *           480
alp_ratón      : VTADHSHVPTFCCTPRGNSIFGLAPHVSDTDEKPTTALLTGNCPYKVVVD-CERENVS-HVDYAHNHYQAQSAVPLRHX : 452
alp_ser humano : VTADHSHVPTFCCTPRGNSIFGLAPHVSDTDEKPTTALLTGNCPYKVVVD-CERENVS-HVDYAHNHYQAQSAVPLRHX : 452
alp_pollo      : VTADHSHVPTFCCTPRGNSIFGLAPHVSDTDEKPTTALLTGNCPYKVVVD-CERENVS-HVDYAHNHYQAQSAVPLRHX : 451
SAP_camarón    : VTADHSHVPTFCCTPRGNSIFGLAPHVSDTDEKPTTALLTGNCPYKVVVD-CERENVS-HVDYAHNHYQAQSAVPLRHX : 427
alp_bcamo      : VTADHSHVSPNCTSPRCTDULCTVR-SLSNRPFPVHLSITNCTGARIQONCVRPDVTIDANFLALRWATHTDVRLDSI : 474
                  VTADH  H      CY  R      D      Y  CPC      C  R      P

      *           500          *           520          *           540          *
alp_ratón      : THCCEDVAVFAKCPHAHLHGVHONTYIPVHLYTSCIGANLDHCAAGSGSAPSPGALLPLA7LSLPTLV- : 524
alp_ser humano : THCCEDVAVFAKCPHAHLHGVHONTYIPVHLYTSCIGANLDHCAAGSGSAPSPGALLPLA7LSLPTLV- : 524
alp_pollo      : THCCEDVAVFAKCPHAHLHGVHONTYIPVHLYTSCIGANLDHCAAGSGSAPSPGALLPLA7LSLPTLV- : 519
SAP_camarón    : THCCEDVAVFAKCPHAHLHGVHONTYIPVHLYTSCIGANLDHCAAGSGSAPSPGALLPLA7LSLPTLV- : 475
alp_bcamo      : THCCEDVAVFAKCPHAHLHGVHONTYIPVHLYTSCIGANLDHCAAGSGSAPSPGALLPLA7LSLPTLV- : 547
                  TH  C  DV      G      G  K  PH  A  A  C  G  C

```


Fig.7.

```

1      AArGChTAYTGGAAYAArGATGCCCAAGATGCACTGGATAAACAG 45
      |||
1      AArGChTAYTGGAAYAArGATGCCCAAGATGCACTGGATAAACAG 45

46      CTGGGGATAAACTGCGAGAGAAGCAGGCCAAGAACGTGATCTTC 90
      |||
46      CTGGGGATAAACTGCGAGAGAAGCAGGCCAAGAACGTGATCTTC 90

91      TTCCTAGGAGACGGGATGTCCCTCAGTACAGTTACCGCAGCAAGA 135
      |||
91      TTCCTAGGAGACGGGATGTCCCTCAGTACAGTTACCGCAGCAAGA 135

136     ATCTACAAAGGTGGATTGACCGGGAAATTGAACGGGAAAAGATT 180
      |||
136     ATCTACAAAGGTGGATTGACCGGGAAATTGAACGGGAAAAGATT 180

181     TCGTGGGAAGAATTCGACTTTGCTGCTTTGAGTAAGACGTACAAT 225
      |||
181     TCGTGGGAAGAATTCGACTTTGCTGCTTTGAGTAAGACGTACAAT 225

226     ACGGACAAGCAAGTGACAGATAGTGCTGCCAGTGCCACAGCCTAC 270
      |||
226     ACGGACAAGCAAGTGACAGATAGTGCTGCCAGTGCCACAGCCTAC 270

271     CTCAC TGGGGTGAAGACCAACCAGGGAGTCATCGGCCCTTGATGCC 315
      |||
271     CTCAC TGGGGTGAAGACCAACCAGGGAGTCATCGGCCCTTGATGCC 315

316     AATACCGTGAGAACAACCTGCTCTTACCAACTTGATGAATCCCTC 360
      |||
316     AATACCGTGAGAACAACCTGCTCTTACCAACTTGATGAATCCCTC 360

361     TTCACTTACTCCATCGCCCATTTGGTTCCAGGAGGCTGGCAGATCA 405
      |||
361     TTCACTTACTCCATCGCCCATTTGGTTCCAGGAGGCTGGCAGATCA 405

406     ACAGGTGTCGTGACATCGACCAGGGTAACTCATGCTACTCCTGCA 450
      |||
406     ACAGGTGTCGTGACATCGACCAGGGTAACTCATGCTACTCCTGCA 450

451     GGAAC TTATGCTCACGTGGCTGACAGAGATTGGGAGAACGACAGT 495
      |||
451     GGAAC TTATGCTCACGTGGCTGACAGAGATTGGGAGAACGACAGT 495

496     GATGTTGTCCACGACAGAGAAGATCCAGAAATATGTGATGACATA 540
      |||
496     GATGTTGTCCACGACAGAGAAGATCCAGAAATATGTGATGACATA 540

541     GCTGAACAACTGGTATTTCAGAGAACCTGGAAAGAACTTCAAGGTC 585
      |||
541     GCTGAACAACTGGTATTTCAGAGAACCTGGAAAGAACTTCAAGGTC 585

586     ATCATGGGCGGCGGTGCGCGTGGATTCTTTCTGAGGAAGCGCTT 630
      |||
586     ATCATGGGCGGCGGTGCGCGTGGATTCTTTCTGAGGAAGCGCTT 630

631     GATATCGAAGATGGCATTCCAGGGGAAAGGCAAGATGGCAAACAC 675
      |||
631     GATATCGAAGATGGCATTCCAGGGGAAAGGCAAGATGGCAAACAC 675

```

Fig.7 (Cont i).

```

676      CTCATAACCGATTGGTTGGACGATAAGGCATCTCAGGGAGCCACA 720
      |||||||
676      CTCATAACCGATTGGTTGGACGATAAGGCATCTCAGGGAGCCACA 720

721      GCATCGTATGTCTGGAACCGAGATGATCTATTGGCTGTTGACATT 765
      |||||||
721      GCATCGTATGTCTGGAACCGAGATGATCTATTGGCTGTTGACATT 765

766      CGAAACACAGACTACCTTATCGGACTGTTTTTCATACACACATCTA 810
      |||||||
766      CGAAACACAGACTACCTTATCGGACTGTTTTTCATACACACATCTA 810

811      GACACAGTTTTTAACAGAGACGCCGAAATGGATCCACCTTGCCG 855
      |||||||
811      GACACAGTTTTTAACAGAGACGCCGAAATGGATCCACCTTGCCG 855

856      GAAATGACAAAGGTCGCCATCGAAATGCTGACGAAGGATGAAAT 900
      |||||||
856      GAAATGACAAAGGTCGCCATCGAAATGCTGACGAAGGATGAAAT 900

901      GGCTTTTTCCTCTTAGTAGAAGGTGGACGGATTGACCATATGCAC 945
      |||||||
901      GGCTTTTTCCTCTTAGTAGAAGGTGGACGGATTGACCATATGCAC 945

946      CACGCCAATCAGATCCGTCAGTCGTTAGCGGAGACGCTGGACATG 990
      |||||||
946      CACGCCAATCAGATCCGTCAGTCGTTAGCGGAGACGCTGGACATG 990

991      GAGG . AAGCTGTATCCATGGCTCTCTCTATGACAGACCCAGAGGA 1034
      |||| |
991      GAGGAAAGCTGTATCCATGGCTCTCTCTATGACAGACCCAGAGGA 1035

1035      AACAAATCA . TTCT . GGTACAGCTGATCAGGTCACACTCTCACC 1077
      ||||| |||| |
1036      AACAAATCATTTCTGGGTACAGCTGATCAGGTCACACTCTCACC 1080

1078      ATCACTGGATACGCAGACAGGAACACCGATATTTTGGATTTCGCC 1122
      |||||||
1081      ATCACTGGATACGCAGACAGGAACACCGATATTTTGGATTTCGCC 1125

1123      GGCAATCAGCGACTTGGATGATAGGAGGTACACCATCCTGGACTAT 1167
      |||||||
1126      GGCAATCAGCGACTTGGATGATAGGAGGTACACCATCCTGGACTAT 1170

1168      GGCAGTGGACCTGGATATCATATCACGGAAGATGGCAAACGTTAC 1212
      |||||||
1171      GGCAGTGGACCTGGATATCATATCACGGAAGATGGCAAACGTTAC 1215

1213      GAACCTACAGAGGAAGATTGAAAGACATCAATTTCCGCTACGCA 1257
      |||||||
1216      GAACCTACAGAGGAAGATTGAAAGACATCAATTTCCGCTACGCA 1260

1258      TCGGCAGCGCCCAAGCATTCAATTACCCATGATGSCACAGACGTT 1302
      |||||||
1261      TCGGCAGCGCCCAAGCATTCAATTACCCATGATGSCACAGACGTT 1305

1303      GGTATTTGGGTCAACGGACCTTTCGCTCATCTCT . . . . . TCACG 1341
      |||||||
1306      GGTATTTGGGTCAACGGACCTTTCGCTCATCTCTTTCACGGTCACG 1350

1342      GGTGTCTATGAAGAGAATTACATTCCCCACGCCCTGGCCCTACGCA 1386
      |||||||
1351      GGTGTCTATGAAGAGAATTACATTCCCCACGCCCTGGCCCTACGCA 1395

1387      GCCTGCCGTTGGTACAGGACGCACTTTCTGTGACGAAAAATGAGAT 1431
      |||||||
1396      GCCTGCCGTTGGTACAGGACGCACTTTCTGTGACGAAAAATGAGAT 1440

```

Fig.7 (Cont ii).

```

1432 GTAGATGATTACCTTCTTAAGTACATTAAAGATTGCAAAAAAAAA 1476
      |||||||||||||||||||||||||||||||||||||||||||
1441 GTAGATGATTACCTTCTTAAGTACATTAAAGATTGCAAAAAAAAA 1485
      |||||||||||||||
1477 AAAAAAAAAAAAAA 1490
      |||||||||||
1486 AAAAAAAAAAAAAA 1499

```

Fig.8.

```

1 KAYWNKDAQDALDKQLGIKLRKQAKNVIFFLGDGMSLSTVTAAR 45
  |||||||||||||||||||||||||||||||||||||||||||
1 KAYWNKDAQDALDKQLGIKLRKQAKNVIFFLGDGMSLSTVTAAR 45

46 IYKGGLTGKFEREKISWEEFDFAALSKTYNTDKQVTD SAASATAY 90
  |||||||||||||||||||||||||||||||||||||||||||
46 IYKGGLTGKFEREKISWEEFDFAALSKTYNTDKQVTD SAASATAY 90

91 LTGVKTNQGVIGLDANTVRTNCSYQLDES LFTYSLAHWFQEAGRS 135
  |||||||||||||||||||||||||||||||||||||||||||
91 LTGVKTNQGVIGLDANTVRTNCSYQLDES LFTYSLAHWFQEAGRS 135

136 TGVVTSTRVTHATPAGTYAHVADRDWENDSDVVHDREDPEICDDI 180
  |||||||||||||||||||||||||||||||||||||||||||
136 TGVVTSTRVTHATPAGTYAHVADRDWENDSDVVHDREDPEICDDI 180

181 AEQLVFREP GKNFKVIMGGRRGFFPEEALDIEDGIPGEREDGKH 225
  |||||||||||||||||||||||||||||||||||||||||||
181 AEQLVFREP GKNFKVIMGGRRGFFPEEALDIEDGIPGEREDGKH 225

226 LITDWLDDKASQGATASYVWNRDDLAVDIRNTDYL MGLFSYTHL 270
  |||||||||||||||||||||||||||||||||||||||||||
226 LITDWLDDKASQGATASYVWNRDDLAVDIRNTDYL MGLFSYTHL 270

271 DTVLTRDAEMDPTLP EMTKVAIEMLT KDENGFFLLVEGGRIDKMH 315
  |||||||||||||||||||||||||||||||||||||||||||
271 DTVLTRDAEMDPTLP EMTKVAIEMLT KDENGFFLLVEGGRIDKMH 315

316 HANQIRQSLAETLDMEEAVSMALSMTDPE. ETIILVTADHGHTLT 359
  ||||||||||||||||: | |||||
316 HANQIRQSLAETLDMEE SCIHGSLYDRPRGNNHFWVTADHGHTLT 360

360 ITGYADRNTDILDFAGISDLDDRRYTILDYGS GPGYHITEDGKRY 404
  |||||||||||||||||||||||||||||||||||||||||||
361 ITGYADRNTDILDFAGISDLDDRRYTILDYGS GPGYHITEDGKRY 405

405 EPT EEDLKDINFRYASAAPKHSVTHDGT DVGIWVNGPFAHLF..T 447
  |||||||||||||||||||||||||||||||||||||||||||
406 EPT EEDLKDINFRYASAAPKHSVTHDGT DVGIWVNGPFAHLFTVT 450

448 GVYEENYIPHALAYAACVGTGR TFCDEK 475
  |||||||||||||||||||||||||||
451 GVYEENYIPHALAYAACVGTGR TFCDEK 478

```

Fig.9.

```

      1      11      21      31      41      51      61
1 ATGAGGGCTTTCTACTGTCTATTTGGGTTACTTTTAAACCCTAAGTCTTCAAGGAGAAGCAACACGTCAAC
1(+1) H R A F Y C L F G L L L T L S L Q G E A T R Q

      71 CAAGTTCCTATGATGGAGATGAGTACCACGAGAGATATCCGGGTCAATCCAACCCATAACAGAGAAGA
24(+1) P S S Y D G D E Y H E R Y P G Q S N P I T E E D

      141 C
48(+1)

```

Fig.10.

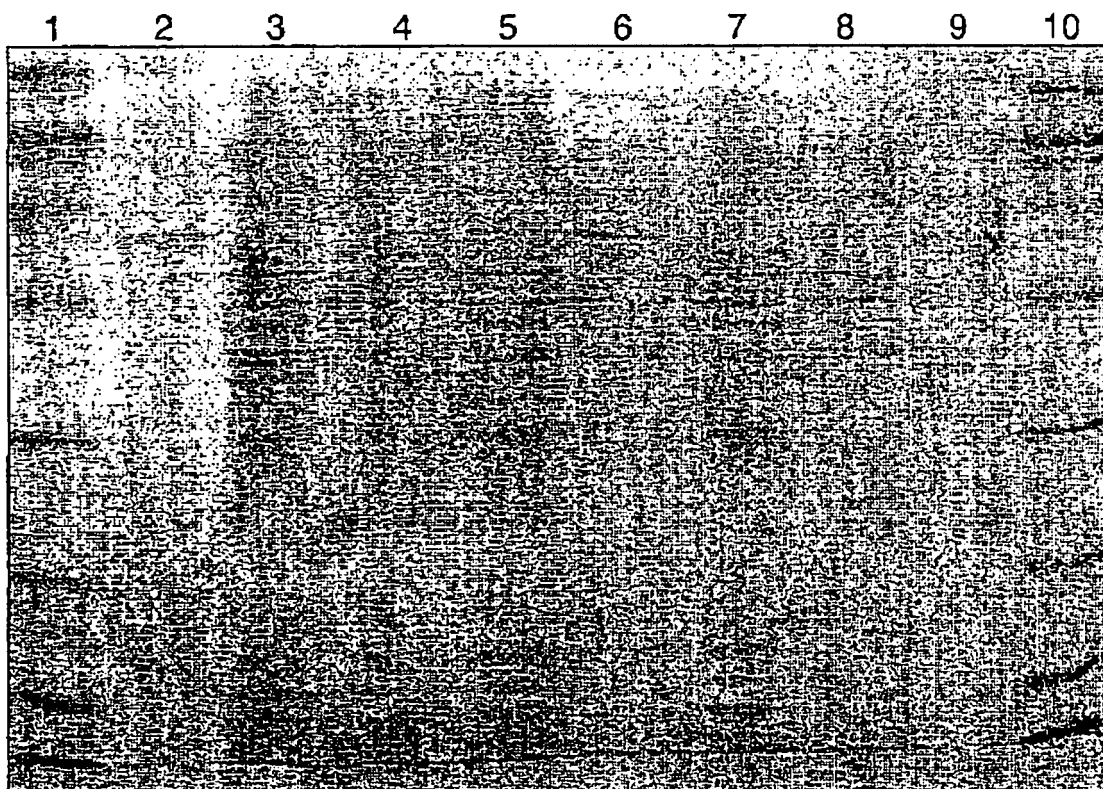
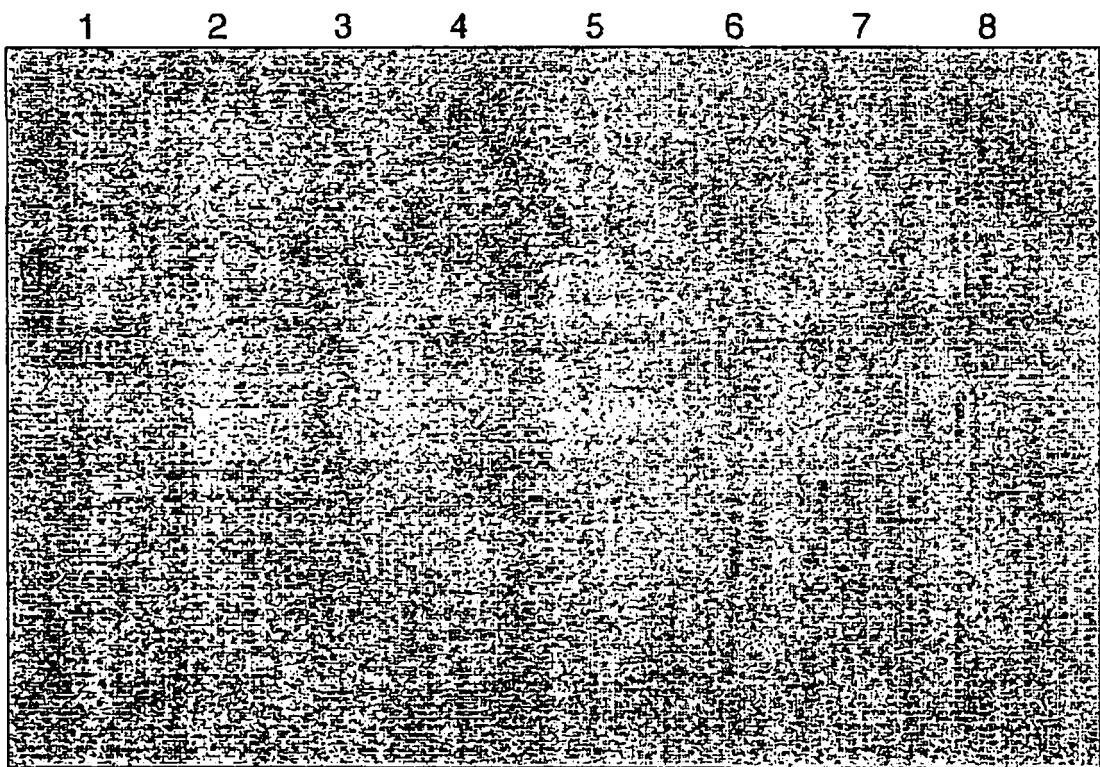


Fig.11.



ES 2 294 033 T3

LISTA DE SECUENCIAS

	<110> Norwegian Institute of Fisheries and Aquaculture Nilsen, Inge Overbo, Kersti	
	<120> Fosfatasa Alcalina de Camarón	
10	<130> 71985001	
	<140> PCT/GB01/04509	
	<141> 2001-10-10	
15	<160> 25	
	<170> PatentIn versión 3.0	
20	<210> 1	
	<211> 1481	
	<212> ADN	
25	<213> <i>Pandalus borealis</i>	
	<220>	
	<223> ADNc de SAP truncado en 5'	
30	<400> 1	
	gatgcccaag atgcactgga taaacagctg gggataaaac tgcgagagaa gcaggccaag 60 aacgtgatct tcttcctagg agacgggatg tccctcagta cagttaccgc agcaagaatc 120 35 taaaaaggtg gattgaccgg gaaatttgaa cgggaaaaga tttcgtggga agaattcgac 180 tttgctgctt tgagtaagac gtacaatac gacaagcaag tgacagatag tgctgccagt 240 gccacagcct acctcactgg ggtgaagacc aaccagggag tcatcggcct tgatgccaat 300 accgtgagaa caaactgctc ttaccaactt gatgaatccc tcttcactta ctccatcgcc 360 40 cattggttcc aggaggctgg cagatcaaca ggtgtcgtga catcgaccag ggtaactcat 420 gctactcctg caggaactta tgctcacgtg gctgacagag attgggagaa cgacagtgat 480 gttgtccacg acagagaaga tccagaaata tgtgatgaca tagctgaaca actggtattc 540 agagaacctg gaaagaactt caaggctcatc atggggcggc gtcggcgtgg attctttcct 600 45 gaggaagcgc ttgatatcga agatggcatt ccaggggaaa gggaagatgg caaacacctc 660 ataaccgatt gggtggacga taaggcatct caggggagcca cagcatcgta tgtctggaac 720 cgagatgatc tattggctgt tgacattcga aacacagact accttatggg actgttttca 780 50 tacacacatc tagacacagt tttaaccaga gacgccgaaa tggatccac cttgccggaa 840 atgacaaagg tcgccatcga aatgctgacg aaggatgaaa atggcctttt cctcttagta 900 gaaggtggac ggattgacca tatgcaccac gccaatcaga tccgtcagtc gttagcggag 960 acgctggaca tggaggaaag ctgtatccat ggctctctct atgacagacc cagaggaaac 1020 55 aatcatttct gggtcacagc tgatcacggc cacactctca ccatcactgg atacgcagac 1080 aggaacaccg atattttgga ttctgccggc atcagcgact tggatgatag gaggtacacc 1140 atcctggact atggcagtgg acctggatat catatcacgg aagatggcaa acgttacgaa 1200 cctacagagg aagatttgaa agacatcaat ttccgctacg catcggcagc gcccagcat 1260 60 tcagttaccc atgatggcac agacgttggc atttgggtca acggaccttt cgctcatctc 1320 ttcacggta cgggtgtcta tgaagagaat tacattcccc acgccttggc ctacgcagcc 1380 tgcgttggtg caggacgcac tttctgtgac gaaaaatgag atgtagatga ttaccttctt 1440 aagtacatta aagattgcaa aaaaaaaaaa aaaaaaaaaa a 1481	
65	<210> 2	
	<211> 478	

ES 2 294 033 T3

<212> PRT

<213> *Pandalus borealis*

5 <220>

<223> Producto de traducción de la SEQ. ID NO. 1.

<220>

10 <221> ACT_SITE

<222> (80)..(88)

<400> 2

15

Lys Ala Tyr Trp Asn Lys Asp Ala Gln Asp Ala Leu Asp Lys Gln Leu
1 5 10 15

20

Gly Ile Lys Leu Arg Glu Lys Gln Ala Lys Asn Val Ile Phe Phe Leu
20 25 30

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

25

Gly Gly Leu Thr Gly Lys Phe Glu Arg Glu Lys Ile Ser Trp Glu Glu
50 55 60

Phe Asp Phe Ala Ala Leu Ser Lys Thr Tyr Asn Thr Asp Lys Gln Val
65 70 75 80

30

Thr Asp Ser Ala Ala Ser Ala Thr Ala Tyr Leu Thr Gly Val Lys Thr
85 90 95

Asn Gln Gly Val Ile Gly Leu Asp Ala Asn Thr Val Arg Thr Asn Cys
100 105 110

35

Ser Tyr Gln Leu Asp Glu Ser Leu Phe Thr Tyr Ser Ile Ala His Trp
115 120 125

Phe Gln Glu Ala Gly Arg Ser Thr Gly Val Val Thr Ser Thr Arg Val
130 135 140

40

Thr His Ala Thr Pro Ala Gly Thr Tyr Ala His Val Ala Asp Arg Asp
145 150 155 160

Trp Glu Asn Asp Ser Asp Val Val His Asp Arg Glu Asp Pro Glu Ile
165 170 175

45

Cys Asp Asp Ile Ala Glu Gln Leu Val Phe Arg Glu Pro Gly Lys Asn
180 185 190

Phe Lys Val Ile Met Gly Gly Gly Arg Arg Gly Phe Phe Pro Glu Glu

50

55

60

65

ES 2 294 033 T3

	195	200	205
	Ala Leu Asp Ile Glu Asp Gly Ile Pro Gly Glu Arg Glu Asp Gly Lys		
5	210	215	220
	His Leu Ile Thr Asp Trp Leu Asp Asp Lys Ala Ser Gln Gly Ala Thr		
	225	230	235
	Ala Ser Tyr Val Trp Asn Arg Asp Asp Leu Leu Ala Val Asp Ile Arg		
	245	250	255
10	Asn Thr Asp Tyr Leu Met Gly Leu Phe Ser Tyr Thr His Leu Asp Thr		
	260	265	270
	Val Leu Thr Arg Asp Ala Glu Met Asp Pro Thr Leu Pro Glu Met Thr		
	275	280	285
15	Lys Val Ala Ile Glu Met Leu Thr Lys Asp Glu Asn Gly Phe Phe Leu		
	290	295	300
	Leu Val Glu Gly Gly Arg Ile Asp His Met His His Ala Asn Gln Ile		
	305	310	315
20	Arg Gln Ser Leu Ala Glu Thr Leu Asp Met Glu Glu Ser Cys Ile His		
	325	330	335
	Gly Ser Leu Tyr Asp Arg Pro Arg Gly Asn Asn His Phe Trp Val Thr		
	340	345	350
25	Ala Asp His Gly His Thr Leu Thr Ile Thr Gly Tyr Ala Asp Arg Asn		
	355	360	365
	Thr Asp Ile Leu Asp Phe Ala Gly Ile Ser Asp Leu Asp Asp Arg Arg		
	370	375	380
30	Tyr Thr Ile Leu Asp Tyr Gly Ser Gly Pro Gly Tyr His Ile Thr Glu		
	385	390	395
	Asp Gly Lys Arg Tyr Glu Pro Thr Glu Glu Asp Leu Lys Asp Ile Asn		
	405	410	415
35	Phe Arg Tyr Ala Ser Ala Ala Pro Lys His Ser Val Thr His Asp Gly		
	420	425	430
	Thr Asp Val Gly Ile Trp Val Asn Gly Pro Phe Ala His Leu Phe Thr		
	435	440	445
40	Val Thr Gly Val Tyr Glu Glu Asn Tyr Ile Pro His Ala Leu Ala Tyr		
	450	455	460
	Ala Ala Cys Val Gly Thr Gly Arg Thr Phe Cys Asp Glu Lys		
	465	470	475

45 <210> 3

<211> 13

<212> PRT

<213> Secuencia artificial

50

<200>

<223> Fragmento de proteína de SAP

55 <400> 3

Asp	Ile	Asn	Phe	Arg	Tyr	Ala	Ser	Ala	Ala	Pro	Val	Lys
1			5						10			

60 <210> 4

<211> 10

<212> PRT

65 <213> Secuencia artificial

<200>

ES 2 294 033 T3

<223> Fragmento de proteína de SAP

<400> 4

5 His Leu Ile Thr Asp Trp Leu Asp Asp Lys
1 5 10

<210> 5

10 <211> 8

<212> PRT

<213> Secuencia artificial

15 <200>

<223> Fragmento de proteína de SAP

20 <400> 5

Val Ile Met Gly Gly Glu Arg Arg
1 5

25 <210> 6

<211> 5

<212> PRT

30 <213> Secuencia artificial

<200>

<223> Fragmento de proteína de SAP

35 <400> 6

Ala Tyr Trp Asn Lys
1 5

40 <210> 7

<211> 15

<212> ADN

45 <213> Secuencia artificial

<220>

50 <223> Cebador

<220>

<221> modified_base

55 <222> (3)..(3)

<223> i

<400> 7

60 gcntaytgga ayaar

15

<210> 8

65 <211> 20

<212> ADN

<213> Secuencia artificial

ES 2 294 033 T3

<220>
 <223> Cebador

5 <220>
 <221> modified_base
 <222> (6)..(6)
 <223> i

10 <220>
 <221> modified_base
 <222> (9)..(9)
 15 <223> i

<220>
 <221> modified_base
 20 <222> (18)..(18)
 <223> i

25 <400> 8
 ckytcncnc ccatdatnac 20

30 <210> 9
 <211> 26
 <212> ADN
 <213> Secuencia artificial

35 <220>
 <223> Cebador

40 <400> 9
 atttcgtggg aagaattcga ctttgc 26

45 <210> 10
 <211> 23
 <212> ADN
 <213> Secuencia artificial

50 <220>
 <223> Cebador

55 <400> 10
 gatctgccag cctcctggaa cca 23

60 <210> 11
 <211> 21
 <212> ADN
 <213> Secuencia artificial

65 <220>
 <223> Cebador

ES 2 294 033 T3

	<220>	
	<221> misc_feature	
	<222> (6)..(6)	
5	<223> t, a, g o c	
	<400> 11	
10	aargcntayt ggaayaarga t	21
	<210> 12	
	<211> 22	
15	<212> ADN	
	<213> Secuencia artificial	
	<220>	
20	<223> Cebador	
	<400> 12	
25	gaaggtaatc atctacatct ca	22
	<210> 13	
	<211> 18	
30	<212> ADN	
	<213> Secuencia artificial	
	<220>	
35	<223> Cebador	
	<220>	
40	<221> modified_base	
	<222> (6)..(6)	
	<223> i	
45	<400> 13	
	aaggcntayt ggaayaar	18
50	<210> 14	
	<211> 26	
	<212> ADN	
	<213> Secuencia artificial	
55	<220>	
	<223> Cebador	
60	<400> 14	
	tacagcggcc gccatctcat ttttcg	26
65	<210> 15	
	<211> 1490	
	<212> ADN	

ES 2 294 033 T3

<213> *Pandalus borealis*

<220>

5 <223> Secuencia de ADNc de SAP revisada

<220>

<221> misc_feature

10 <222> (6)..(6)

<223> t, a, g o c

<400> 15

15	aargcntayt ggaayaarga tgcccaagat gcaactggata aacagctggg gataaaactg	60
	cgagagaagc aggccaaagaa cgtgatcttc ttcctaggag acgggatgtc cctcagtaca	120
	gttaccgcag caagaatcta caaaggtgga ttgaccggga aatttgaacg ggaaaagatt	180
20	tcgtgggaag aattcgactt tgctgctttg agtaagacgt acaatacggg caagcaagtg	240
	acagatagtg ctgccagtgc cacagcctac ctcaactggg tgaagaccaa ccaggaggatc	300
	atcggccttg atgccaatat cgtgagaaca aactgctctt accaacttga tgaatccctc	360
25	ttcacttact ccatcgccca ttgggtccag gaggtggca gatcaacagg tgcgtgaca	420
	tcgaccaggg taactcatgc tactcctgca qgaacttatg ctacgtggc tgacagagat	480
	tgggagaacg acagtgatgt tgtccacgac agagaagatc cagaaatatg tgatgacata	540
	gctgaacaac tgggtattcag agaacctgga aagaacttca aggtcatcat gggcggcggg	600
30	cggcgtggat tctttcctga ggaagcgctt gatatcgaag atggcattcc aggggaaagg	660
	gaagatggca aacacctcat aaccgattgg ttggacgata aggcattctca gggagccaca	720
	gcatcgtatg tctggaaccg agatgatcta ttggctgttg acattcgaaa cacagactac	780
35	cttatgggac tgttttcata cacacatcta gacacagttt taaccagaga cgccgaaatg	840
	gateccacct tgccggaaat gacaaaggct gccatcgaaa tgctgacgaa ggatgaaat	900
	ggctttttcc tcttagtaga aggtggacgg attgaccata tgcaccacgc caatcagatc	960
40	cgtcagtcgt tagcggagac gctggacatg gaggaagctg tatccatggc tctctctatg	1020
	acagacccag aggaaacaat cattctgggc acagctgac acggtcacac tctcaccatc	1080
	actggatacg cagacaggaa caccgatatt ttggatttcg ccggcatcag cgacttggat	1140
45	gataggagggt acaccatcct ggactatggc agtggacctg gatatcatat cacggaagat	1200
	ggcaaacggt acgaacctac agaggaagat ttgaaagaca tcaatttcog ctacgcatcg	1260
	gcagcgccca agcattcagt taccatgat ggcacagacg ttggtatttg ggtcaacgga	1320
	cctttcgtct atctcttcac ggggtgtctat gaagagaatt acattcccca cgccttggcc	1380
50	tacgcagcct gcgttggtac aggacgcact ttctgtgacg aaaaatgaga tgtagatgat	1440
	taccttctta agtacattaa agattgcaaa aaaaaaaaaa aaaaaaaaaa	1490

55 <210> 16

<211> 475

<212> PRT

<213> *Pandalus borealis*

60

<220>

<223> Producto de traducción de la SEQ ID NO. 15

65 <220>

<221> ACT_SITE

<222> (80)..(88)

ES 2 294 033 T3

<400> 16

	Lys	Ala	Tyr	Trp	Asn	Lys	Asp	Ala	Gln	Asp	Ala	Leu	Asp	Lys	Gln	Leu
	1				5					10					15	
5	Gly	Ile	Lys	Leu	Arg	Glu	Lys	Gln	Ala	Lys	Asn	Val	Ile	Phe	Phe	Leu
				20					25					30		
	Gly	Asp	Gly	Met	Ser	Leu	Ser	Thr	Val	Thr	Ala	Ala	Arg	Ile	Tyr	Lys
			35					40					45			
10	Gly	Gly	Leu	Thr	Gly	Lys	Phe	Glu	Arg	Glu	Lys	Ile	Ser	Trp	Glu	Glu
		50					55					60				
	Phe	Asp	Phe	Ala	Ala	Leu	Ser	Lys	Thr	Tyr	Asn	Thr	Asp	Lys	Gln	Val
	65					70					75					80
	Thr	Asp	Ser	Ala	Ala	Ser	Ala	Thr	Ala	Tyr	Leu	Thr	Gly	Val	Lys	Thr
15				85						90					95	
	Asn	Gln	Gly	Val	Ile	Gly	Leu	Asp	Ala	Asn	Thr	Val	Arg	Thr	Asn	Cys
				100					105					110		
	Ser	Tyr	Gln	Leu	Asp	Glu	Ser	Leu	Phe	Thr	Tyr	Ser	Ile	Ala	His	Trp
20			115					120					125			
	Phe	Gln	Glu	Ala	Gly	Arg	Ser	Thr	Gly	Val	Val	Thr	Ser	Thr	Arg	Val
		130					135					140				
	Thr	His	Ala	Thr	Pro	Ala	Gly	Thr	Tyr	Ala	His	Val	Ala	Asp	Arg	Asp
25		145				150					155					160
	Trp	Glu	Asn	Asp	Ser	Asp	Val	Val	His	Asp	Arg	Glu	Asp	Pro	Glu	Ile
				165						170					175	
	Cys	Asp	Asp	Ile	Ala	Glu	Gln	Leu	Val	Phe	Arg	Glu	Pro	Gly	Lys	Asn
				180					185					190		
30	Phe	Lys	Val	Ile	Met	Gly	Gly	Gly	Arg	Arg	Gly	Phe	Phe	Pro	Glu	Glu
		195						200					205			
	Ala	Leu	Asp	Ile	Glu	Asp	Gly	Ile	Pro	Gly	Glu	Arg	Glu	Asp	Gly	Lys
		210					215					220				
	His	Leu	Ile	Thr	Asp	Trp	Leu	Asp	Asp	Lys	Ala	Ser	Gln	Gly	Ala	Thr
	225					230					235					240
35	Ala	Ser	Tyr	Val	Trp	Asn	Arg	Asp	Asp	Leu	Leu	Ala	Val	Asp	Ile	Arg
				245						250					255	
	Asn	Thr	Asp	Tyr	Leu	Met	Gly	Leu	Phe	Ser	Tyr	Thr	His	Leu	Asp	Thr
			260						265					270		
	Val	Leu	Thr	Arg	Asp	Ala	Glu	Met	Asp	Pro	Thr	Leu	Pro	Glu	Met	Thr
		275						280					285			
40	Lys	Val	Ala	Ile	Glu	Met	Leu	Thr	Lys	Asp	Glu	Asn	Gly	Phe	Phe	Leu
		290					295					300				
	Leu	Val	Glu	Gly	Gly	Arg	Ile	Asp	His	Met	His	His	Ala	Asn	Gln	Ile
		305				310					315					320
	Arg	Gln	Ser	Leu	Ala	Glu	Thr	Leu	Asp	Met	Glu	Glu	Ala	Val	Ser	Met
				325						330					335	
45	Ala	Leu	Ser	Met	Thr	Asp	Pro	Glu	Glu	Thr	Ile	Ile	Leu	Val	Thr	Ala
			340						345					350		
	Asp	His	Gly	His	Thr	Leu	Thr	Ile	Thr	Gly	Tyr	Ala	Asp	Arg	Asn	Thr
			355					360					365			
	Asp	Ile	Leu	Asp	Phe	Ala	Gly	Ile	Ser	Asp	Leu	Asp	Asp	Arg	Arg	Tyr
50		370					375					380				
	Thr	Ile	Leu	Asp	Tyr	Gly	Ser	Gly	Pro	Gly	Tyr	His	Ile	Thr	Glu	Asp
		385				390					395					400
	Gly	Lys	Arg	Tyr	Glu	Pro	Thr	Glu	Glu	Asp	Leu	Lys	Asp	Ile	Asn	Phe
				405						410					415	
55	Arg	Tyr	Ala	Ser	Ala	Ala	Pro	Lys	His	Ser	Val	Thr	His	Asp	Gly	Thr
			420						425					430		
	Asp	Val	Gly	Ile	Trp	Val	Asn	Gly	Pro	Phe	Ala	His	Leu	Phe	Thr	Gly
			435					440					445			
	Val	Tyr	Glu	Glu	Asn	Tyr	Ile	Pro	His	Ala	Leu	Ala	Tyr	Ala	Ala	Cys
60		450					455					460				
	Val	Gly	Thr	Gly	Arg	Thr	Phe	Cys	Asp	Glu	Lys					
		465				470					475					

<210> 17

<211> 141

<212> ADN

<213> *Pandalus borealis*

ES 2 294 033 T3

<220>
 <223> ADNc que codifica el extremo N y péptido señal de SAP

5 <400> 17

```

      atgaggggctt tctactgtct atttgggtta cttttaaccc taagtcttca aggagaagca      60
      acacgtcaac caagttccta tgatggagat gagtaccacg agagatatcc gggccaatcc      120
10      aaccccataa cagaagaaga c                                     141
  
```

<210> 18
 <211> 47

15 <212> PRT
 <213> *Pandalus borealis*

<220>
 20 <223> Producto de traducción de la SEQ ID NO. 17

<220>
 <221> SEÑAL
 25 <222> (91)..(47)

<400> 18

```

      Met Arg Ala Phe Tyr Cys Leu Phe Gly Leu Leu Leu Thr Leu Ser Leu
30      1      5      10
      Gln Gly Glu Ala Thr Arg Gln Pro Ser Ser Tyr Asp Gly Asp Glu Tyr
      20      25      30
35      His Glu Arg Tyr Pro Gly Gln Ser Asn Pro Ile Thr Glu Glu Asp
      35      40      45
  
```

<210> 19
 <211> 18

40 <212> ADN
 <213> Secuencia artificial

<220>
 45 <221> misc_feature
 <222> (6)..(6)
 <223> t, a, g o c

50 <400> 19

```

      aargentayt ggaayaar
  
```

55 <210> 20
 <211> 1499
 <212> ADN
 <213> Secuencia artificial

60 <220>
 <223> Secuencia recombinante de la SEQ ID NO. 19 y la SEQ ID NO. 1

65 <220>
 <221> misc_feature

18

ES 2 294 033 T3

<222> (6)..(6)

<223> t, a g o c

5 <400> 20

	aargcntayt ggaayaarga tgcccaagat gcactggata aacagctggg gataaaactg	60
	cgagagaagc aggccaaagaa cgtgatcttc ttcctaggag acgggatgtc cctcagtaca	120
10	gttaccgcag caagaatcta caaagggtga ttgaccggga aatttgaacg ggaaaagatt	180
	tcgtgggaag aattcgactt tgctgctttg agtaagacgt acaatacggg caagcaagtg	240
	acagatagtg ctgccagtgc cacagcctac ctactgggg tgaagaccaa ccagggagtc	300
15	atcggccttg atgccaatac cgtgagaaca aactgctctt accaacttga tgaatccctc	360
	ttcacttact ccatcgccca ttggttccag gaggctggca gatcaacagg tgctgtgaca	420
	tcgaccaggg taactcatgc tactcctgca ggaacttatg ctacgtggc tgacagagat	480
	tgggagaacg acagtgatgt tgtccacgac agagaagatc cagaaatatg tgatgacata	540
20	gctgaacaac tggatttcag agaacctgga aagaacttca aggtcatcat gggcggcggt	600
	cggcgtggat tctttcctga ggaagcgctt gatatcgaag atggcattcc aggggaaaag	660
	gaagatggca aacacctcat aaccgattgg ttggacgata aggcattctca gggagccaca	720
25	gcatcgatg tctggaaccg agatgatcta ttggctgttg acattcgaaa cacagactac	780
	cttatgggac tgttttcata cacacatcta gacacagttt taaccagaga cgccgaaatg	840
	gatcccacct tgccggaaat gacaaaggc gccatcgaaa tgctgacgaa ggatgaaaat	900
30	ggctttttcc tcttagtaga aggtggacgg attgaccata tgcaccacgc caatcagatc	960
	cgtcagtcgt tagcggagac gctggacatg gaggaaagct gtatccatgg ctctctctat	1020
	gacagaccca gaggaacaa tcatttcttg gtcacagctg atcacggtca cactctcacc	1080
	atcactggat acgcagacag gaacaccgat attttggtt tgcccgcat cagcgacttg	1140
35	gatgatagga ggtacacat cctggactat ggcagtggac ctggatatca tatcacggaa	1200
	gatggcaaac gttacgaacc tacagaggaa gatttgaaag acalcaattt ccgctacgca	1260
	tcggcagcgc ccaagcattc agttacccat gatggcacag acgttggtat ttgggtcaac	1320
40	ggacctttcg ctcatctctt cacggtcacg ggtgtctatg aagagaatta cattccccac	1380
	gccttggcct acgcagcctg cgttggtaca ggacgcactt tctgtgacga aaaatgagat	1440
	gtagatgatt accttcttaa gtacattaaa gattgcaaaa aaaaaaaaaa aaaaaaaaaa	1499

45 <210> 21

<211> 18

<212> ADN

<213> *Pandalus borealis*

50

<220>

<223> Fragmento de N-terminal de SAP que codifica el péptido KAYWNK

55 <400> 21

aaagcatatt ggaacaaa

18

60 <210> 22

<211> 1511

<212> ADN

<213> Secuencia artificial

65

<220>

ES 2 294 033 T3

<223> Secuencia de ADNc recombinante del péptido señal y la SEQ ID NO. 21 que sustituye a la SEQ ID NO. 19 en la SEQ ID NO. 15

<400> 22

```

5      aacccccataa cagaagaaga caaagcatat tggaacaaag atgcccaaga tgcactggat      60
      aaacagctgg ggataaaact gcgagagaag caggccaaga acgtgatctt cttcctagga      120
10     gacgggatgt ccttcagtac agttaccgca gcaagaatct acaaaggagg attgaccggg      180
      aaatttgaac gggaaaagat ttcgtgggaa gaattcgact ttgctgcttt gagtaagacg      240
      tacaatacgg acaagcaagt gacagatagt gctgccagtg ccacagccta cctcactggg      300
      gtgaagacca accagggagt catcggcctt gatgccaata ccgtgagaac aaactgctct      360
15     taccaacttg atgaatccct cttcacttac tccatcgccc attggttcca ggaggctggc      420
      agatcaacag gtgtcgtgac atcgaccagg gtaactcatg ctactcctgc aggaacttat      480
      gctcacgtgg ctgacagaga ttgggagaac gacagtgatg ttgtccacga cagagaagat      540
20     ccagaaatat gtgatgacat agctgaacaa ctggtattca gagaacctgg aaagaacttc      600
      aaggatcatc tgggcggcgg tcggcggtga ttctttcctg aggaagcgtt tgatattcga      660
      gatggcattc caggggaaag ggaagatggc aaacacctca taaccgattg gttggacgat      720
      aaggcatctc agggagccac agcatcgtat gtctggaacc gagatgatct attggctggt      780
25     gacattcgaa acacagacta ccttatggga ctgttttcat acacacatct agacacagtt      840
      ttaaccagag acgccgaaat ggatcccacc ttgccggaaa tgacaaaggc cgccatcgaa      900
      atgctgacga aggatgaaaa tggctttttc ctcttagtag aagggtggacg gattgaccat      960
30     atgcaccacg ccaatcagat ccgtcagtcg ttagcggaga cgctggacat ggaggaagct      1020
      gtatccatgg ctctctctat gacagaccca gaggaacaa tcattctggt cacagctgat      1080
      cacggtcaca ctctcaccat cactggatac gcagacagga acaccgatat tttggatttc      1140
35     gccggcatca gcgacttggg tgaataggagg tacaccatcc tggactatgg cagtggacct      1200
      ggatatcata tcacggaaga tggcaaacgt tacgaacctc cagaggaaga tttgaaagac      1260
      atcaatttcc gctacgcacg ggcagcgccc aagcattcag ttacctatga tggcacagac      1320
      gttggtatct gggtaacagg acctttcgct catctcttca cgggtgtcta tgaagagaat      1380
40     tacattcccc acgccttggc ctacgcagcc tgcgttggta caggacgcac tttctgtgac      1440
      gaaaaatgag atgtagatga ttaccttctt aagtacatta aagattgcaa aaaaaaaaaa      1500
      aaaaaaaaaa a                                     1511

```

<210> 23

<211> 26

<212> ADN

50 <213> Secuencia artificial

<220>

<223> Cebador

55

<400> 23

gcgtggtgca tatgtcaat ccgtcc

26

60

<210> 24

<211> 32

<212> ADN

65 <213> Secuencia artificial

<220>

ES 2 294 033 T3

<223> Cebador

<400> 24

5 atggagctca acccaatcac cgaagaagac aa

32

<210> 25

10 <211> 30

<212> ADN

<213> Secuencia artificial

15 <220>

<223> Cebador

<400> 25

20 atgaagcttt catttttcgt cacagaaagt

30

25

30

35

40

45

50

55

60

65