



(19)



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 277 615**

(51) Int. Cl.:

C12N 15/55 (2006.01)

C12N 9/20 (2006.01)

C12N 15/63 (2006.01)

C12R 1/90 (2006.01)

(12)

TRADUCCIÓN DE PATENTE EUROPEA

T3

(86) Número de solicitud europea: **02719712 .8**

(86) Fecha de presentación : **22.01.2002**

(87) Número de publicación de la solicitud: **1360306**

(87) Fecha de publicación de la solicitud: **12.11.2003**

(54) Título: **Una secuencia de ADN de la enzima fosfolipasa A₁ del ciliado *Tetrahymena* y uso de la misma.**

(30) Prioridad: **22.01.2001 DE 101 05 152**

(45) Fecha de publicación de la mención BOPI:
16.07.2007

(45) Fecha de la publicación del folleto de la patente:
16.07.2007

(73) Titular/es: **Cilian AG.**
Johann-Krane-Weg 42
48149 Münster, DE

(72) Inventor/es: **Hartmann, Marcus;**
Grenningloh, Marco y
Tiedtke, Arno

(74) Agente: **Carpintero López, Francisco**

ES 2 277 615 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

Una secuencia de ADN de la enzima fosfolipasa A₁ del ciliado *Tetrahymena* y uso de la misma.

5 La invención se refiere a un ácido nucleico que codifica la fosfolipasa A₁ y al uso del mismo según el preámbulo de las reivindicaciones 1 a 5.

10 Las levaduras, bacterias y células de mamíferos tienen una gran importancia para la preparación y producción biotecnológicas de principios activos recombinantes mediante la expresión heteróloga de proteínas extrañas. A este respecto se usan sistemas de expresión bacterianos basados en *E. coli* o *B. subtilis* para la producción de péptidos o proteínas recombinantes, como por ejemplo, la insulina, la interleucina-2, el activador del plasminógeno tisular, proteasas y lipasas. En las bacterias gram negativas, los sistemas de expresión se basan fundamentalmente en el uso de elementos genéticos como el operón *lac* o el operón del triptófano. A este respecto, las proteínas extrañas al huésped se producen en el interior de la célula en cuerpos de inclusión o, en caso del uso de sistemas de expresión basados en genes de β -lactamasa, en el espacio periplasmático. La producción de proteínas recombinantes en el medio de fermentación circundante no está establecida. En las bacterias gram positivas hasta ahora sólo se incorporan y se expresan en sistemas de expresión casi exclusivamente proteínas de la propia célula.

20 Las levaduras como *S. cerevisiae*, *Hansenula polymorpha*, *Kluyveromyces lactis* o *Pichia Pastoris* se emplean igualmente para la expresión heteróloga de proteínas recombinantes, como por ejemplo el factor XIIIa humano, la proquimosina bovina, la fitasa o antígenos superficiales. Los sistemas de expresión se basan aquí en vectores lanzadera (vectores con una parte de levaduras y una parte bacteriana) que (según el tipo de levadura) están contruidos sobre los elementos genéticos de la galactoquinasa-epimerasa, la metanoloxidasa, la fosfatasa ácida o la alcoholdehidrogenasa. Por lo general, la proteína recombinante se produce en el citoplasma de la célula. Al usar secuencias señal propias de levaduras, como la del factor alfa, las proteínas expresadas también pueden excretarse al medio de fermentación. La glicosilación de proteínas excretadas tiene lugar según el tipo "rico en manosa", frecuentemente se producen hiperglicosilaciones en las proteínas, lo que puede conducir a la formación de anticuerpos en el paciente.

30 Las células de mamíferos, como diversos tipos de células de roedores (células CHO, células C127) o monos (células Vero, CV-1 o COS) se emplean igualmente para la expresión heteróloga de proteínas recombinantes. Los sistemas de expresión se basan aquí en virus recombinantes (vector BPV) o en vectores lanzadera. Para regular la expresión se emplean los sistemas víricos de potenciador/promotor de SV40 o elementos amplificadores celulares. Las proteínas recombinantes, como la eritropoyetina, se excretan al medio de fermentación, ya que, por lo general, los genes extraños ya llevan en sí secuencias señal y éstas son comprendidas por el sistema de expresión y usadas para el control del destino. Para la producción biotecnológica de enzimas glicosiladas extracelulares se emplean además protozoos del género *Tetrahymena*. *Tetrahymena* crece en medios de fermentación de bajo coste mediante el empleo de procedimientos de fermentación convencionales. Para la transformación de tales células de *Tetrahymena* se dispone de vectores que se basan en elementos del ADNr de *Tetrahymena*. Para la expresión heteróloga de proteínas bacterianas en *Tetrahymena* se emplean construcciones de ADN a partir de genes de *Tetrahymena*. Si se dispone de los elementos genéticos adecuados para la regulación de la transcripción, del control del destino y de la glicosilación de proteínas extrañas, *Tetrahymena* es un sistema de expresión ideal para la producción de proteínas recombinantes terapéuticas a bajo coste.

45 Los sistemas de expresión de bacterias gram negativas usados hasta ahora conducen por lo general a la formación de cuerpos de inclusión en la célula, acompañada de una desnaturalización de la proteína. Para la obtención de la proteína recombinante, las células deben romperse y la proteína desnaturalizada e inactiva debe replegarse para ser funcional. Esto ocasiona costosas etapas adicionales en el procedimiento y reduce el rendimiento de la proteína deseada. La glicosilación, de importancia para las proteínas eucarióticas, no tiene lugar en absoluto. En el caso del uso de sistemas de expresión de bacterias gram positivas es problemática además la degradación de la proteína diana por elevadas actividades proteolíticas en el caldo de fermentación.

55 En el caso del uso de levaduras para la expresión heteróloga, la proteína diana deseada se produce igualmente sólo en la célula, de la que ha de retirarse mediante disgregación de la misma. Esto ocasiona etapas adicionales de elevados coste y duración en el procedimiento, como en los sistemas de expresión bacterianos. Si se usan péptidos señal propios de levaduras, las proteínas extrañas no se procesan ni glicosilan correctamente al excretarse.

60 Si, por el contrario, se emplean sistemas de células de mamíferos para la producción de proteínas recombinantes, las proteínas deseadas se hallan extracelularmente en el medio de fermentación, correctamente procesadas y glicosiladas. Sin embargo, la desventaja aquí es, por un lado, la baja tasa de expresión debido al procesamiento erróneo y a una traducción ineficiente de los genes, que deben introducirse en el genoma de la línea celular productora mediante vectores víricos. Por otro lado, los medios de fermentación con suero para células de mamíferos son extremadamente costosos. Además, la tecnología de fermentación para las líneas celulares, que son sensibles a las fuerzas de cizalladura, es también muy compleja e igualmente costosa, debido a las construcciones para la aireación sin formación de burbujas. Otros problemas resultan del elevado riesgo de infección de las líneas celulares por micoplasmas y virus.

65 En total, el uso de células de mamíferos para la preparación biotecnológica de proteínas recombinantes tiene como consecuencia unos costes muy elevados, condiciones de seguridad y bajos rendimientos.

Las desventajas anteriormente mencionadas en la producción de proteínas recombinantes no son válidas en el empleo de ciliados como *Tetrahymena*. Así por ejemplo, algunas hidrolasas ácidas implicadas en la digestión de partículas alimenticias se excretan fuera de la célula en grandes cantidades y complejamente glicosiladas.

Alam y col. describen en J. Euk. Microbiol. 43 (4), 1996, páginas 295 a 303, la clonación de un gen que codifica la alfa-glucosidasa ácida de *Tetrahymena pyriformis*. Sin embargo, la proteína sólo se excreta fuera de la célula en una pequeña proporción. Además, en la solicitud de patente internacional PCT/EP00/01853 se describe el gen de una β -hexosaminidasa de *Tetrahymena thermophila*, de la que sin embargo se sabe que sólo se excreta fuera de la célula aproximadamente un 80%

Sin embargo, hasta ahora no es posible conseguir la expresión de proteínas eucarióticas glicosiladas en *Tetrahymena* que igualmente se excretan exclusivamente al medio de fermentación. La causa es que hasta ahora no se conocían las secuencias de ADN de las proteínas extracelulares propias de *Tetrahymena* que son necesarias para la construcción de vectores de expresión y que excretan exclusivamente la proteína extraña al huésped al medio de fermentación circundante. Se conocen las secuencias de ADN de una proteína que codifica la β -hexosaminidasa de *Tetrahymena thermophila*. Se ha solicitado la patente de una secuencia tal con los distintivos oficiales DE 19958979.8, DE 19909189.7 así como PCT/EP00/01853. Sin embargo, la desventaja de estas secuencias es que los pre/propéptidos que contienen sólo dirigen aproximadamente el 80 por ciento de una proteína extraña al huésped al medio de fermentación circundante. La causa es que, en condiciones naturales, aproximadamente el 20 por ciento de la enzima β -hexosaminidasa se halla en la membrana y sólo aproximadamente el 80 por ciento de la enzima producida naturalmente se excreta fuera de la célula. Por esta razón, los pre/propéptidos de la β -hexosaminidasa, cuando se posicionan delante de una proteína extraña al huésped mediante procedimientos de ingeniería génica, dirigen aproximadamente el 20 por ciento de dicha proteína extraña a la membrana citoplasmática, en la superficie de *Tetrahymena thermophila*. Esto supone una desventaja considerable en la técnica del procedimiento para la preparación del principio activo recombinante. Por una parte, así se reduce el rendimiento, ya que una parte de la proteína expresada permanece en las células unida a la membrana y, por tanto, no puede purificarse el total de la proteína expresada a partir del caldo de fermentación. Por otro lado, la proteína extraña al huésped en la membrana puede tener efectos tóxicos sobre las células huésped y retardar así el crecimiento celular.

Además, hasta la fecha no se conoce ningún promotor constitutivo de *Tetrahymena* que consiga una transcripción uniforme o continua de proteínas heterólogas. Hasta ahora solamente se conocen promotores de genes de histonas y tubulinas (Bannon y col., 1984, Gaertig y col., 1993). Sin embargo, una desventaja decisiva de estos promotores es que se activan dependiendo del ciclo celular. Los genes de proteínas heterólogas que están unidos a tales promotores dependientes del ciclo celular sólo se expresan en células en crecimiento o división. Esto supone considerables desventajas en cuanto a la técnica de procedimiento, ya que la proteína deseada sólo se produce en la fase logarítmica de crecimiento. En la fase estacionaria de crecimiento, en la que se alcanza la máxima densidad celular en el proceso de producción y, con ello, la máxima capacidad de rendimiento del organismo de expresión (*Tetrahymena*), apenas tiene lugar crecimiento y, con ello, sólo una baja expresión de la proteína heteróloga.

Es objetivo de la invención proporcionar un sistema que permita la producción de proteínas heterólogas en un sistema de expresión después de la transformación de *Tetrahymena*, fuera de las células en el medio de fermentación.

El objetivo se consigue mediante un sistema en el que se usa un ácido nucleico con una secuencia designada como ID SEC N° 1 que codifica una fosfolipasa A₁ (ID SEC N° 7). Ventajosamente el producto de expresión de ese ADN se excreta fuera la célula en grandes cantidades en las condiciones de cultivo. La proteína expresada se excreta en gran medida al medio de cultivo circundante y no se halla fija a la membrana. La secuencia de ácido nucleico según la invención contiene un promotor que consigue una transcripción constitutiva, es decir independiente del ciclo celular, de los genes de proteínas heterólogas situados a continuación del mismo. A este respecto, la transcripción constitutiva tiene la ventaja de que las proteínas se expresan continuamente en el organismo huésped, sin influencia del ciclo celular. Por tanto, durante la fase estacionaria de crecimiento con un crecimiento menor, también puede tener lugar la transcripción del gen extraño y expresarse las proteínas heterólogas.

La secuencia de ADN según la invención de la fosfolipasa A₁ contiene preferentemente una región secuencia arriba de la PLA₁ (ID SEC N° 2) que porta elementos promotores para la iniciación de la transcripción, un péptido señal y un propéptido, otros elementos genéticos para el control del destino de las proteínas, así como en particular, una región secuencia abajo de la PLA₁ (ID SEC N° 3) que contiene elementos genéticos para la terminación de la transcripción. El uso de estos ácidos nucleicos en un vector hace posible la expresión de las proteínas heterólogas expresadas independientemente del ciclo celular y su transporte dirigido fuera de la célula al medio de cultivo circundante sin que, a este respecto, las proteínas expresadas se incorporen en las membranas citoplasmáticas, con lo que, por tanto, dichas proteínas pueden aislarse sin disgregar las células a partir del caldo de fermentación.

La figura 1 muestra un ácido nucleico que codifica la región secuencia arriba (ID SEC N° 2), la región codificante (ID SEC N° 1) y la región secuencia abajo (ID SEC N° 3) de la fosfolipasa A₁ de ciliados.

La figura 2 muestra un producto de expresión correspondiente al ácido nucleico según ID SEC N° 1. La proteína según ID SEC N° 7 es igualmente objeto de la presente invención.

Es también objeto de la invención, en particular, la secuencia señal (ID SEC N° 6) de la proteína según la invención. Ésta consiste preferentemente en los aminoácidos 1 a 110 de la proteína según la invención (ID SEC N° 5). Un ácido nucleico que codifica el fragmento N-terminal también es objeto de la invención (ID SEC N° 3). Éste consiste preferentemente en un fragmento de los ácidos nucleicos según la invención (ID SEC N° 4), en particular con la secuencia de ácido nucleico 1 a 155 según la figura 1.

La secuencia de ácido nucleico de la región que no se transcribe (región secuencia arriba) (ID SEC N° 2), por delante de la región codificante de la PLA₁ de *Tetrahymena*, se halla entre la posición -275 y la posición -1 (representada en letras minúsculas). La región que no se transcribe determinada comprende 275 bases. Como elementos de un promotor se hallan una caja TATA en la posición -49 a -55 (en negrita), así como una caja CAAT putativa entre la base -133 y la base -136 (en negrita). La región de la secuencia codificante del ADNc se representa en letras mayúsculas. La numeración de la secuencia comienza con el codón de iniciación ATG. Las regiones conocidas a partir de la secuenciación de la proteína están enmarcadas y el codón de parada está subrayado. La proteína madura se codifica a partir de la base 331. El protocolo de secuencia desde la base 1 hasta la base 330 representa la pre/prosecuencia (ID SEC N° 8) de la PLA₁. El protocolo de secuencia desde la base 331 hasta la base 963 representa la secuencia de la PLA₁ madura (ID SEC N° 9). En la posición 961 se encuentra la parada de traducción TGA, en la posición 1039 la señal de poliadenilación AATAAA. La secuencia de ácido nucleico desde la posición 964 hasta la posición 1134, que se halla a continuación de la secuencia codificante de la PLA₁ de *Tetrahymena*, representa la región secuencia abajo de la PLA₁ (ID SEC N° 3) que no se transcribe (igualmente representada en letras minúsculas). De la posición 964 a la posición 1101 se halla la región conocida a partir de la secuenciación del ADNc, que se confirmó igualmente mediante PCR inversa. El último codón del ARNm (ttt, posición 1098-1101) se añade después de la transcripción de la cola de poli-A.

Otro aspecto de la invención es el uso de una secuencia de ácido nucleico de hidrolasas ácidas según la invención o partes de la misma para la expresión homóloga o heteróloga de proteínas y péptidos recombinantes, así como para la recombinación homóloga o heteróloga (noqueado, sustitución génica).

Es objeto de la invención un procedimiento para la expresión homóloga o heteróloga de proteínas y péptidos, así como para la recombinación homóloga o heteróloga (noqueado, sustitución génica) en el que se transfectan ciliados con un ácido nucleico según la invención.

Los ácidos nucleicos o partes de los mismos pueden combinarse, en particular, con los potenciadores, promotores, operadores, orígenes, terminadores, resistencias a antibióticos, usuales para la expresión homóloga o heteróloga de proteínas, o con otros ácidos nucleicos o fragmentos o secuencias de ADN de cualquier tipo, de viroides, virus, bacterias, arqueos, protozoos, hongos, plantas, animales o humanos.

El ácido nucleico según la invención se encuentra contenido, en particular, en un vector, un plásmido, un cósmido, un cromosoma o minicromosoma, un transposón, un elemento IS, un ADNr o cualquier otro tipo de ADN o ARN circular o lineal.

Es también objeto de la invención un procedimiento en el que el ácido nucleico según la invención codificante de la fosfolipasa A₁ o partes del mismo se combinan con potenciadores como, por ejemplo, la región NF-1 (un potenciador de citomegalovirus), promotores como, por ejemplo, los promotores *lac*, *trc*, *tac* o *tac*, los promotores de las clases II y III del sistema RNAP de T7, los promotores de los bacteriófagos T7 y SP6, los promotores *aprE*, amilasa o *spac* para los sistemas de expresión de *Bacillus*, los promotores AOX1, AUG1 ó 2 ó GAPp (*Pichia*) para los sistemas de expresión de levaduras, el promotor RSV (virus SV40), el promotor CMV (citomegalovirus), el promotor AFP (adenovirus) o los promotores de metalotioneína para los sistemas de expresión de mamíferos, los promotores del virus Sindbis o los promotores del virus del bosque Semlike para células de insectos, promotores para sistemas de expresión de células de insectos como hsp70, DS47, actina 5C o copia, promotores específicos de plantas como el promotor 35S (virus del mosaico de la coliflor), el promotor de la amilasa o el promotor de la patatina de clase I, operadores como por ejemplo, el operador *tet*, péptidos señal como las pre/prosecuencias señal a-MF (*Saccharomyces*), orígenes, terminadores, resistencias a antibióticos y principios activos como ampicilina, kanamicina, estreptomycin, cloranfenicol, penicilina, anfotericina, cicloheximida, 6-metilpurina, paromomicina, higromicina, α -amanitina, marcadores de auxotrofia como el gen de la dihidrofolatoreductasa, usuales para la expresión homóloga o heteróloga de proteínas, o con otros ácidos nucleicos o fragmentos o secuencias de ADN de cualquier tipo, de viroides, virus, bacterias, arqueos, protozoos, hongos, plantas, animales o humanos.

En particular el ácido nucleico según la invención o partes del mismo se incorporan en un vector, un plásmido, un cósmido, un cromosoma o minicromosoma, un transposón, un elemento IS, un ADNr o cualquier otro tipo de ADN o ARN circular o lineal.

Es evidente para el experto que según la invención pueden emplearse también ácidos nucleicos que sean homólogos al menos en el 40% al ácido nucleico según ID SEC N° 1. También la proteína según ID SEC N° 2 puede modificarse sin perder funcionalidad. Así por ejemplo, pueden llevarse a cabo las sustituciones de aminoácidos denominadas conservadoras. Además pueden intercambiarse, por ejemplo, aminoácidos hidrófobos entre sí.

Para la purificación y el aislamiento de la fosfolipasa A₁ de *Tetrahymena* y para la determinación de la secuencia pueden usarse los siguientes procedimientos.

Obtención de la PLA₁

La PLA₁ se obtuvo a partir de sobrenadantes sin células del cultivo de *Tetrahymena thermophila*. Para ello, las células se fermentaron en un fermentador de 2 litros (Biostat MD, Braun Diessel Biotech, Melsungen, Alemania), controlado por medio de una unidad de control digital DCU. A este respecto, el fermentador se hizo funcionar en primer lugar de forma discontinua durante 24 horas y a continuación de forma continua. Una recogida del sobrenadante sin células se garantizó mediante un módulo de perfusión con un tamaño de poros de aproximadamente 0,3 µm (S6/2, Enka, Wuppertal).

La fermentación se llevó a cabo con los siguientes parámetros:

- el volumen de trabajo fue de 2 litros
- la tasa de perfusión fue de 2 litros/día
- el número de revoluciones del agitador se limitó a 800 rpm
- la temperatura se mantuvo constante a 30°C
- el valor de pH se mantuvo constante a pH 7
- el título de inoculación fue de 50.000 células/ml

Para la fermentación se usó la cepa SB 1868 VII. Esta es una cepa natural de *Tetrahymena thermophila*.

La fermentación se realizó durante un periodo de 264 horas y el material recogido se ensayó en cuanto a su actividad de PLA₁.

Purificación de la PLA₁

Para la purificación de la PLA₁ se usó 1 litro de sobrenadante de cultivo sin células a partir de la fermentación. A éste se le añadieron 140 g de sulfato amónico y se concentró a un volumen de 50 ml mediante una unidad de ultrafiltración (Pellikon XL, tamaño de exclusión 3 kDa, Millipore). A continuación, la muestra se purificó por cromatografía de interacción hidrófoba (Fractogel EMD Phenyl 1650 de 20 x 1,6, Merck, Darmstadt). A este respecto, el caudal fue de 5 ml/min y el producto de elución se recogió en fracciones de 5 ml. La actividad enzimática se midió por la desacilación de un fosfolípido marcado radiactivamente (L-3-fosfatidilcolina, 1-palmitoilo-2-[1-¹⁴C]linoelilo). La figura 3 muestra el perfil de elución obtenido, el gradiente de acetato sódico y las actividades enzimáticas en las fracciones individuales.

Las tres fracciones con las máximas actividades enzimáticas se reunieron y se retamponaron mediante una unidad de ultrafiltración en el tampón inicial (bis-tris 20 mM, pH 6,5) para la cromatografía de intercambio aniónico (AEC). A continuación, la muestra se aplicó a la columna (Q-Sepharose-Hiload-16/10, Pharmacia, Suecia) y la PLA₁ se eluyó de la columna con un gradiente lineal de NaCl (velocidad de flujo = 3 ml/min) y se recogió en fracciones de 5 ml. La figura 4 muestra el perfil de elución obtenido, el gradiente de NaCl y las actividades enzimáticas de las fracciones individuales.

De la fracción con la máxima actividad de PLA₁ se tomaron 200 µl y se separaron mediante cromatografía de exclusión por tamaños (SEC). Para ello se usó una columna Superdex HR 75 30/10 (Pharmacia, Suecia). El caudal en esta cromatografía fue de 0,6 ml/min, el producto eluido se recogió en fracciones de 200 µl. La figura 5 muestra el perfil de elución y las actividades enzimáticas de las fracciones individuales.

La pureza de las fracciones obtenidas se analizó mediante electroforesis unidimensional en gel. A este respecto se determinaron dos bandas distintas a aproximadamente 26 y aproximadamente 28 kDa. Una separación de estas dos bandas mediante electroforesis bidimensional en gel resultó en un desdoblamiento de las dos bandas en 2 o 3 manchas con puntos isoelectricos diferentes. Éstos se hallaron a pH 6,3 y 5,7 para las proteínas de 26 kDa, y a pH 6,3, 5,7 y 5,3 para las proteínas de 28 kDa. Un examen posterior de estas manchas mediante un análisis de huella peptídica dió como resultado que dichas manchas son isoformas de la misma proteína.

Análisis molecular biológico de la PLA₁

Después de comprobar la pureza de las proteínas, se transfirieron muestras de las mismas a una membrana de PVDF y se secuenciaron a partir del extremo N-terminal. Adicionalmente, otra muestra se digirió con tripsina e igualmente se secuenció. Con ayuda de las secuencias de proteínas así obtenidas se prepararon cebadores de oligonucleótidos que se emplearon en una reacción de PCR con transcriptasa inversa (3' RACE, amplificación rápida de extremos de ADNc). Mediante esta técnica de PCR se consiguió amplificar el ADNc de la fosfolipasa A₁ y secuenciarlo a continuación.

ES 2 277 615 T3

La secuencia obtenida tiene una longitud de 633 bases o 729 bases, el peso molecular de la proteína madura derivado de dicha secuencia es de aproximadamente 22,4 kDa. En la secuencia deducida volvieron a encontrarse también al 100% los oligopéptidos de 22 aminoácidos (N-terminal) o 18 aminoácidos (dentro de la proteína) determinados a partir de la secuenciación de proteínas. Adicionalmente a la secuencia de la proteína madura, pudo editarse también la secuencia del pre/propéptido mediante 5' RACE (amplificación rápida de extremos de ADNc) (fig. 2). Éste consiste en un péptido de 110 aminoácidos de longitud que porta, por un lado, la secuencia señal y, por otro, el propéptido que inactiva la enzima y que no se escinde hasta alcanzar el sitio definitivo de acción de la enzima. Las comparaciones de secuencias dieron como resultado que, con la excepción de una secuencia consenso de cinco aminoácidos (GxSxG), presente en todas las lipasas y fosfolipasas y discutida como sitio de unión para lípidos o fosfolípidos, no hay ninguna homología con fosfolipasas A₁ ya conocidas. Se determinaron además las secuencias secuencia arriba y secuencia abajo de la PLA₁ mediante PCR inversa (fig.1). Para ello se cortó el ADN genómico con endonucleasas de restricción, se ligó con ligasa T4 y finalmente se amplificó con cebadores inversos. Para la amplificación de la región secuencia arriba de la PLA₁ mediante PCR inversa se usó ADN genómico cortado con la endonucleasa de restricción *SspI*. A este respecto, pudo editarse corriente arriba una región de 275 bases y pudieron identificarse elementos promotores. En la posición -136 se halla una caja CAAT, que presenta una distancia al inicio de la traducción comparable a las de las cajas CCAAT del gen de la histona (-141 o -151) de *Tetrahymena* encontradas por Brunk y Sadler (1990). En la posición -55 se identificó una caja TATA, que fija el sitio exacto del inicio de la transcripción en los genes eucarióticos. Su secuencia concuerda con las secuencias consenso encontradas en eucariotas. Para la amplificación mediante PCR inversa de la región secuencia abajo, que contiene el terminador para la transcripción de la PLA₁ de *Tetrahymena*, se usó ADN genómico cortado con *BamHI*. Así pudieron determinarse otras 222 bases adicionalmente a la región secuencia abajo conocida a partir del experimento 3' RACE (fig. 3).

25

30

35

40

45

50

55

60

65

REIVINDICACIONES

1. Ácido nucleico compuesto de regiones codificantes y no codificantes con la secuencia ID SEC N° 4, codificando las regiones codificantes la fosfolipasa A₁ (PLA₁) de *Tetrahymena thermophila* (ID SEC N° 5) y representando las regiones no codificantes las regiones secuencia arriba y secuencia abajo de este gen.

2. Ácido nucleico según la reivindicación 1, presentando las regiones no codificantes secuencia arriba y secuencia abajo las secuencias ID SEC N°s 2 y 3 y la región codificante las secuencias ID SEC N°s 1, 8 y/o 9.

3. Ácido nucleico con una de las secuencias ID SEC N°s 1, 2, 3, 8 ó 9.

4. Secuencia señal N-terminal de la PLA₁ (ID SEC N° 5) según la reivindicación 1, con la secuencia ID SEC N° 6.

5. Proteína fosfolipasa A₁ (PLA₁) con la secuencia de aminoácidos según ID SEC N° 7.

6. Uso de un ácido nucleico según las reivindicaciones 1 a 3 para la expresión homóloga o heteróloga de proteínas y péptidos recombinantes, así como para la recombinación homóloga o heteróloga (noqueado, sustitución génica).

7. Procedimiento para la expresión homóloga o heteróloga de proteínas y péptidos, así como para la recombinación homóloga o heteróloga (noqueado, sustitución génica), en el que se transfectan ciliados con un ácido nucleico según una de las reivindicaciones 1 a 3.

8. Procedimiento según la reivindicación 7, en el que los ácidos nucleicos o parte de los mismos se combinan con los potenciadores, promotores, operadores, orígenes, terminadores, resistencias a antibióticos, usuales para la expresión homóloga o heteróloga de proteínas, o con otros ácidos nucleicos o fragmentos o secuencias de ADN de cualquier tipo, de viroides, virus, bacterias, arqueozoos, protozoos, hongos, plantas, animales o humanos.

9. Procedimiento según las reivindicaciones 7 y/u 8, en el que el ácido nucleico está contenido en un vector, un plásmido o un cósmido.

10. Vectores, plásmidos o cósmidos que contienen al menos uno de los ácidos nucleicos según al menos una de las reivindicaciones 1 a 3.

Fig. 1

tca	gat	atg	ata	tgc	tct	tta	taa	aaa	gca	gag	att	ttt	aat	aga	aga	atc	aat	gac	tca	tga	att	-241
aag	aat	ttt	ttt	tat	taa	aaa	gca	gag	att	ttt	aat	aga	aga	atc	aat	gac	tca	tga	att	-181		
taa	taa	aga	ttt	tca	agt	gtt	ttc	taa	tac	cga	cta	gct	tta	taa	att	cac	tta	tta	atc	-121		
aac	gat	ata	aaa	att	ata	tta	aca	aat	caa	taa	ata	aaa	aaa	taa	ata	aaa	aca	aaa	caa	-61		
ATG	AAC	AAG	ACT	CTC	ATC	TTA	GCT	TTA	GTT	GTT	GTT	TTG	GCT	TTA	ACT	GCC	ACC	ACC	TTG	-1		
GTT	GCT	TTC	CAC	AAC	CAC	TCT	CAC	AAC	ATC	AGA	GTT	GAC	TAA	GAC	CCC	GCC	ACT	CTC	TTC	60		
AAG	CAA	TTC	AAG	CAA	ACT	TAC	AAT	AAG	AAG	TAT	GCT	GAT	GCT	ACT	TTC	GAA	ACC	TAC	AGA	120		
TTC	GGT	GTC	TTC	ACC	CAA	AAC	TTA	GAA	ATC	GTC	AAG	ACT	GAC	TCT	ACT	TTC	GGT	GTC	ACC	180		
TAA	TTC	ATG	GAC	TTA	ACT	CCT	GCT	GAA	TTC	GCT	CAA	CAA	TTC	CTC	ACT	TTA	CAC	GAA	AAG	240		
GTT	AAC	AGC	ACC	GAA	GTT	TAC	AGA	GCT	TAA	GGT	GAA	GCT	ACC	GAA	GTT	GAC	TGG	ACT	GCC	300		
AAG	GGT	AAG	GTC	ACC	CCT	GTT	AAG	AAC	TAA	GGT	TCT	TGT	GGT	TCC	TGC	TGG	GCT	TTC	TCC	360		
ACC	ATT	GGT	GCC	GTT	GAA	TCT	GCT	CTT	TGG	ATT	GCT	GGT	CAA	GGT	GAA	TAA	AAC	ACT	CTT	420		
AAC	CTT	GCT	GAA	TAA	GAA	TAA	GTT	GAC	TGT	GCT	AAG	TCC	CCC	AAG	TAC	GAC	TCT	GAA	GGT	480		
TGC	AAC	GGT	GGT	TGG	ATG	GTT	GAA	GGT	TTC	AAG	TAC	ATC	ATC	GAC	AAC	AAG	ATC	TCC	TAA	540		
ACT	GCT	AAC	TAT	CCC	TAC	ACT	GCT	AAG	GAT	GGT	AAG	TGC	AAG	GAC	ACC	TCT	TCC	TTC	AAG	600		
AAG	TTC	TCT	ATT	TCT	AAG	TAC	GCT	GAA	ATC	CCC	TAA	GGT	GAC	TGC	AAC	TCC	CTC	AAC	TCT	660		
GCC	TTA	GAA	CAA	GGT	CCT	ATC	TCC	GTT	GCT	GTT	GAT	GCC	ACC	AAC	TTC	TAA	TTC	TAC	ACT	720		
TCT	GGT	GTC	TTT	AAA	AAC	TGC	AAG	GCC	AAC	CTC	AAC	CAC	GGT	GTC	CTC	TTA	GTT	GCC	AAC	780		
GTT	GAC	TCT	TCT	CTC	AAG	ATC	AAG	AAC	TCC	TGG	GGT	CCT	TCT	TGG	GGT	GAA	AAG	GGT	TTC	840		
ATC	AGA	TTA	GCT	GCC	GGT	AAC	ACT	TGC	GGT	GTC	TGC	AAT	GCT	GCC	TCT	TAC	CCT	ATT	GTT	900		
TGA	aaa	aac	ata	atc	caa	att	aaa	aaa	aat	tac	tca	aaa	ctg	ata	ata	taa	aaa	att	aat	960		
ttt	cat	aat	ttt	aat	gta	aat	aaa	tac	ctt	tat	att	tga	cgt	ttt	gta	ctc	aaa	ata	aat	1020		
taa	agt	taa	caa	acc	ata	ttt	att	taa	ttc	tac	ttt	tca	att	ttt	aaa	aat	ata			1080		
																				1140		

Fig. 2:

La secuencia de aminoácidos de la fosfolipasa A1 de Tetrahymena: pre/propéptido (aminoácidos 1 a 110):

M	N	K	T	L	I	L	A	L	V	G	V	L	A	L	15
T	A	T	T	L	V	A	F	L	N	H	S	H	N	i	30
R	V	D	Q	D	P	A	T	L	F	K	Q	F	K	Q	45
T	Y	N	K	K	Y	A	D	P	T	F	E	T	Y	R	60
F	G	V	F	T	Q	N	L	E	I	V	K	T	D	S	75
T	F	G	V	T	Q	F	M	D	L	T	P	A	E	F	90
A	Q	Q	F	L	T	L	H	E	K	V	N	S	E	E	105
V	Y	R	A	Q											110

Proteína madura (del extremo N hasta el extremo C, aminoácidos 111 a 342):

G	E	A	T	E	V	D	W	T	A	K	G	K	V	T	125
P	V	K	N	Q	G	S	C	G	S	C	W	A	F	S	140
T	I	G	A	V	E	S	A	L	L	I	A	G	Q	G	155
E	Q	N	T	L	N	L	A	E	Q	E	L	V	D	C	170
A	K	S	P	K	Y	D	S	E	G	C	N	G	G	W	185
M	V	E	G	F	K	Y	I	I	D	N	K	I	S	Q	200
T	A	N	Y	P	Y	T	A	K	D	G	K	C	K	D	215
T	S	S	F	K	K	F	S	I	S	K	Y	A	E	I	230
P	Q	G	D	C	N	S	L	N	S	A	L	E	Q	G	245
P	I	S	V	A	V	D	A	T	N	F	Q	F	Y	T	260
S	G	V	F	K	N	C	K	A	N	L	N	H	G	V	275
L	L	V	A	N	V	D	S	S	L	K	I	K	N	S	290
W	G	P	S	W	G	E	K	G	F	I	R	L	A	A	305
G	N	T	C	G	V	C	N	A	A	S	Y	P	I	V	320
Parada															335

Figura 3 : HIC I, perfil de elución, gradiente de acetato sódico y actividad enzimática de la PLA1 eluida

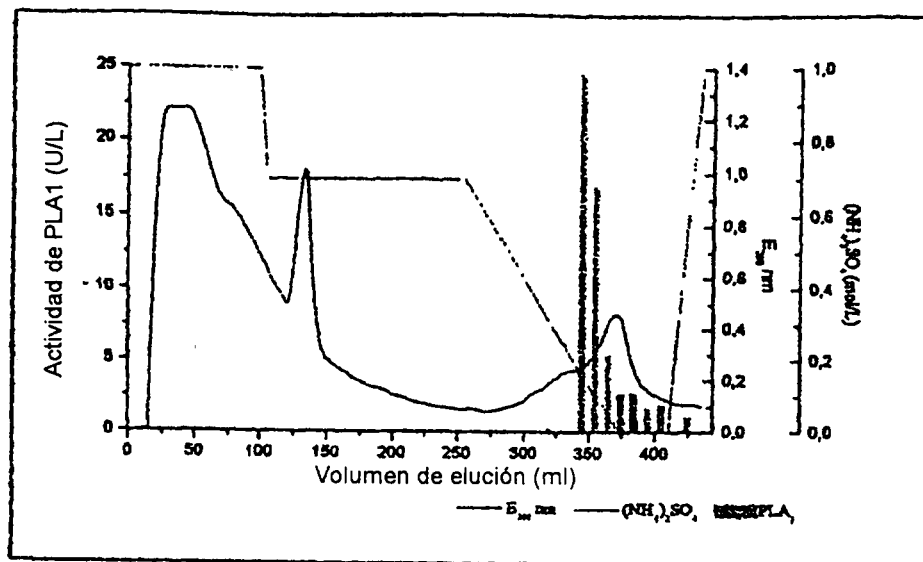


Figura 4: AEC, perfil de elución, gradiente de NaCl y actividad enzimática de la PLA1 eluida

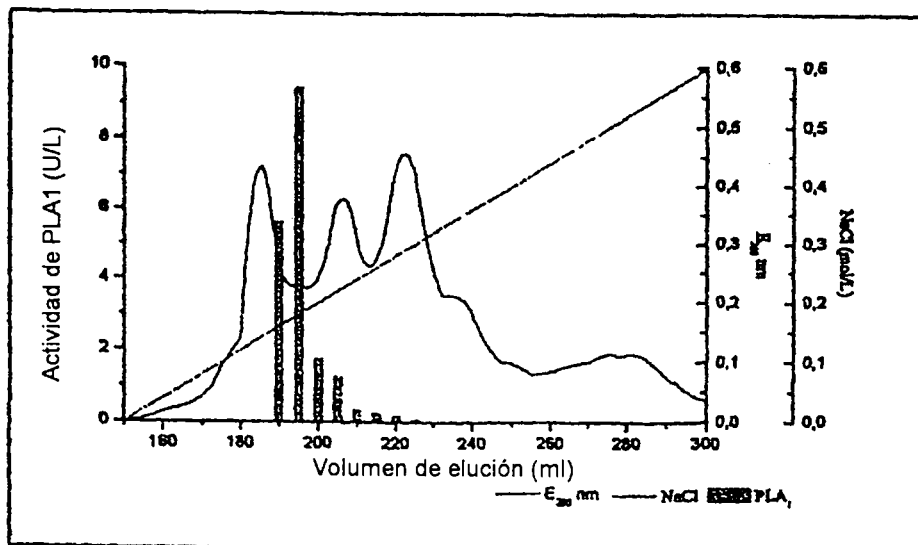
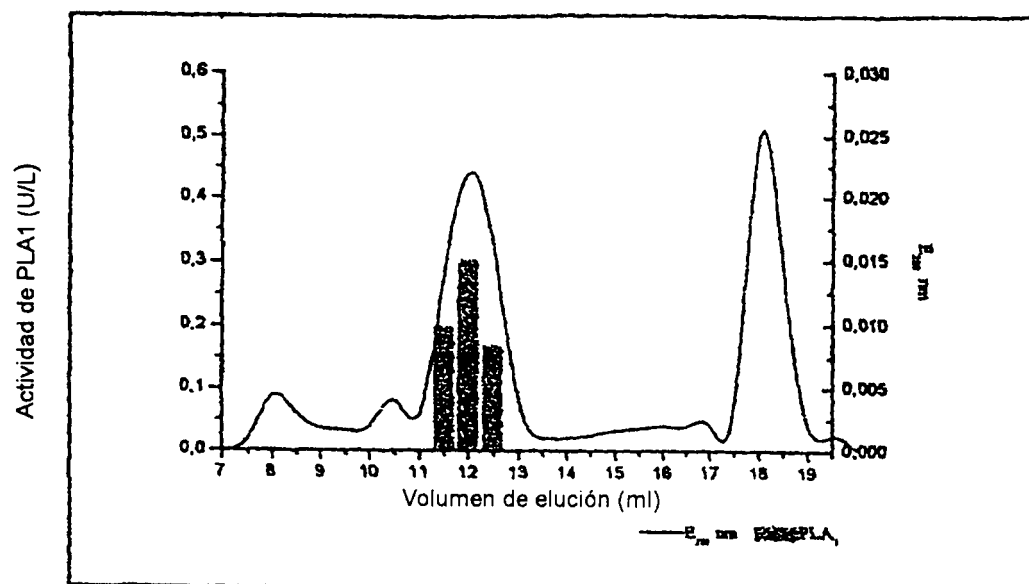


Figura 5: SEC, perfil de elución y actividades enzimáticas de la PLA1



ES 2 277 615 T3

LISTA DE SECUENCIAS

<110> Cilian AG
 <120> Una secuencia de ADN de la enzima fosfolipasa A₁ del ciliado *Tetrahymena* y el uso de la misma:
 <130> 020414wo ME/BM
 <140> PCT/EP02/00578
 <141> 22-01-2002
 <160> 9
 <170> PatentIn versión 2.1
 <210> 1
 <211> 963
 <212> ADN
 <213> *Tetrahymena thermophila*
 <400> 1

 atgaacaaga ctctcatctt agcttttagtt gttgtttttg ctttaactgc caccaccttg 60
 gttgcttttc acaaccactc tcacaacatc agagttgact aagaccccg cactctcttc 120
 aagcaattca agcaaaactta caataagaag tatgctgatg ctactttcga aacctacaga 180
 ttcggtgtct tcaccccaaaa cttagaatat gteaagactg actctacttt cgggtgtcacc 240
 taattcatgg acttaactcc tgctgaattc gctcaacaat tcctcacttt acacgaaaag 300
 gttaacagca ccgaagttta cagagcttaa ggtgaagcta ccgaagttga ctggactgcc 360
 aagggttaagg tcacccctgt taagaactaa ggttcttggg gttcctgctg ggctttctcc 420
 accattgggtg ccgttgaatc tgctctttgg attgctgggc aaggtgaata aaacactctt 480
 aaccttgctg aataagaata agttgactgt gctaagtcac ccaagtacga ctctgaaggt 540
 tgcaacgggtg gttggatggg tgaagggttc aagtacatca tcgacaacaa gatctcctaa 600
 actgctaact atccctacac tgctaaggat ggtaagtgca aggacacctc ttccttcaag 660
 aagttctcta tttctaagta cgctgaaatc ccctaagggt actgcaactc cctcaactct 720
 gccttagaac aaggtcctat ctccgttgct gttgatgcca ccaacttcta attctacact 780
 tctgggtgtct ttaaaaaactg caaggccaac ctcaaccacg gtgtccctct agttgccaac 840
 gttgactctt ctctcaagat caagaactcc tggggtcctt cttgggggtga aaagggtttc 900
 atcagattag ctgccggtaa cacttgccgt gtctgcaatg ctgcctctta ccctattgtt 960
 tga 963

 <210> 2
 <211> 275
 <212> ADN
 <213> *Tetrahymena thermophila*
 <400> 2

 aatattttatc aatgctactt ataattcttt tagtagaga tatgatatgc tctttctctg 60
 ctagacttaa cttatgacat ttgaactttt aataaaagaa ttttttttat taaaaagcag 120
 agatttttaa tagaagaatc aatgactcat gaatttaata aagattttca agtgttttct 180
 aataccgact agctttataa attcacttat taatcaacga tataaaaatt atattaacaa 240
 atcaataaat aaaaaaataa ataaaaacaa aacaa 275

 <210> 3
 <211> 171
 <212> ADN
 <213> *Tetrahymena thermophila*

ES 2 277 615 T3

<400> 3

```

5      aaaaacataa tccaaattaa aaaaaattac tcaaaactga taatataaaa aattaatttt 60
      cataatttta atgtaaataa atacctttat atttgacgtt ttgtactcaa aataaattaa 120
      agttaacaaa ccatatttat ttaattctac ttttcaattt ttaaaaatat a          171

```

<210> 4

10 <211> 1408

<212> ADN

<213> *Tetrahymena thermophila*

15 <400> 4

```

20      aatattttatc aatgctactt ataattcttt tagtatgaga tatgatatgc tctttctctg 60
      ctagacttaa cttatgacat ttgaactttt aataaaagaa ttttttttat taaaaagcag 120
      aqatttttaa taqaaqaatc aatqactcat qaatttaata aaqattttca aqtqttttct 180
      aataccgact agctttataa attcacttat taatcaacga tataaaaatt atattaacaa 240
      atcaataaat aaaaaaataa ataaaaacaa aacaaatgaa caagactctc atcttagctt 300
      tagttgttgt tttggcttta actgccacca ccttggttgc tttccacaac cactctcaca 360
      acatcagagt tgactaagac cccgccactc tcttcaagca attcaagcaa acttacaata 420
25      agaagtatgc tgatgctact ttcgaaacct acagattcgg tgtcttcacc caaaacttag 480
      aaatcgtcaa gactgactct actttcggtg tcacctaat catggactta actcctgctg 540
      aattcgctca acaattcctc actttcacga aaagggttaac agcaccgaag tttacagagc 600
      ttaaggtgaa gctaccgaag ttgactggac tgccaagggt aaggtcaccc ctgttaagaa 660
      ctaaggttct tgtggttcct gctgggcttt ctccaccatt ggtgccgttg aatctgctct 720
30      ttgattgct ggtcaagggt aataaaacac tcttaacctt gctgaataag aataagttga 780
      ctgtgctaag tcccccaagt acgactctga aggttgcaac ggtggttggg tgggtgaagg 840
      tttcaagtac atcatcgaca acaagatctc ctaaactgct aactatccct acactgctaa 900
      ggatggtaag tgcaaggaca cctcttcctt caagaagttc tctatttcta agtacgctga 960
      aatcccctaa ggtgactgca actccctcaa ctctgcctta gaacaagggt ctatctccgt 1020
35      tgctgttgat gccaccaact tctaattcta cacttctggt gtctttaaaa actgcaaggc 1080
      caacctcaac caccgtgtcc tcttagttgc caacgttgac tcttctctca agatcaagaa 1140
      ctctgggggt ccttcttggt gtgaaaaggg tttcatcaga ttagctgccg gtaacacttg 1200
      csgtgtctgc aatgctgcct cttaccctat tgtttgaaaa aacataatcc aaattaaaaa 1260
      aaattactca aaactgataa tataaaaaat taattttcat aattttaatg taataaataa 1320
40      cctttatatt tgacgttttg tactcaaaat aaattaaagt taacaaacca tattttattha 1380
      attctacttt tcaattttta aaaatata          1408

```

<210> 5

45 <211> 320

<212> PROT

<213> *Tetrahymena thermophila*

50

55

60

65

ES 2 277 615 T3

<400> 5

```

Met Asn Lys Thr Leu Ile Leu Ala Leu Val Gly Val Leu Ala Leu Thr
 1           5           10           15
5
Ala Thr Thr Leu Val Ala Phe His Asn His Ser His Asn Ile Arg Val
      20           25           30
10
Asp Gln Asp Pro Ala Thr Leu Phe Lys Gln Phe Lys Gln Thr Tyr Asn
      35           40           45
15
Lys Lys Tyr Ala Asp Pro Thr Phe Glu Thr Tyr Arg Phe Gly Val Phe
      50           55           60
20
Thr Gln Asn Leu Glu Ile Val Lys Thr Asp Ser Thr Phe Gly Val Thr
      65           70           75           80
25
Gln Phe Met Asp Leu Thr Pro Ala Glu Phe Ala Gln Gln Phe Leu Thr
      85           90           95
30
Leu His Glu Lys Val Asn Ser Thr Glu Val Tyr Arg Ala Gln Gly Glu
      100          105          110
35
Ala Thr Glu Val Asp Trp Thr Ala Lys Gly Lys Val Thr Pro Val Lys
      115          120          125
40
Asn Gln Gly Ser Cys Gly Ser Cys Trp Ala Phe Ser Thr Ile Gly Ala
      130          135          140
45
Val Glu Ser Ala Leu Leu Ile Ala Gly Gln Gly Glu Gln Asn Thr Leu
      145          150          155          160
50
Asn Leu Ala Glu Gln Glu Leu Val Asp Cys Ala Lys Ser Pro Lys Tyr
      165          170          175
55
Asp Ser Glu Gly Cys Asn Gly Gly Trp Met Val Glu Gly Phe Lys Tyr
      180          185          190
60
Ile Ile Asp Asn Lys Ile Ser Gln Thr Ala Asn Tyr Pro Tyr Thr Ala
      195          200          205
65
Lys Asp Gly Lys Cys Lys Asp Thr Ser Ser Phe Lys Lys Phe Ser Ile
      210          215          220
70
Ser Lys Tyr Ala Glu Ile Pro Gln Gly Asp Cys Asn Ser Leu Asn Ser
      225          230          235          240
75
Ala Leu Glu Gln Gly Pro Ile Ser Val Ala Val Asp Ala Thr Asn Phe
      245          250          255
80
Gln Phe Tyr Thr Ser Gly Val Phe Lys Asn Cys Lys Ala Asn Leu Asn
      260          265          270
85
His Gly Val Leu Leu Val Ala Asn Val Asp Ser Ser Leu Lys Ile Lys
      275          280          285
90
Asn Ser Trp Gly Pro Ser Trp Gly Glu Lys Gly Phe Ile Arg Leu Ala
      290          295          300
95
Ala Gly Asn Thr Cys Gly Val Cys Asn Ala Ala Ser Tyr Pro Ile Val
      305          310          315          320

```

<210> 6

65 <211> 110

<212> PROT

<213> *Tetrahymena thermophila*

ES 2 277 615 T3

<400> 6

```

Met Asn Lys Thr Leu Ile Leu Ala Leu Val Gly Val Leu Ala Leu Thr
 1           5           10           15

5 Ala Thr Thr Leu Val Ala Phe His Asn His Ser His Asn Ile Arg Val
           20           25           30

Asp Gln Asp Pro Ala Thr Leu Phe Lys Gln Phe Lys Gln Thr Tyr Asn
           35           40           45

10 Lys Lys Tyr Ala Asp Pro Thr Phe Glu Thr Tyr Arg Phe Gly Val Phe
           50           55           60

Thr Gln Asn Leu Glu Ile Val Lys Thr Asp Ser Thr Phe Gly Val Thr
65           70           75           80

15 Gln Phe Met Asp Leu Thr Pro Ala Glu Phe Ala Gln Gln Phe Leu Thr
           85           90           95

Leu His Glu Lys Val Asn Ser Thr Glu Val Tyr Arg Ala Gln
20           100           105           110

```

<210> 7

<211> 210

25 <212> PROT

<213> *Tetrahymena thermophila*

<400> 7

```

30 Gly Glu Ala Thr Glu Val Asp Trp Thr Ala Lys Gly Lys Val Thr Pro
   1           5           10           15

Val Lys Asn Gln Gly Ser Cys Gly Ser Cys Trp Ala Phe Ser Thr Ile
           20           25           30

35 Gly Ala Val Glu Ser Ala Leu Leu Ile Ala Gly Gln Gly Glu Gln Asn
           35           40           45

Thr Leu Asn Leu Ala Glu Gln Glu Leu Val Asp Cys Ala Lys Ser Pro
40           50           55           60

Lys Tyr Asp Ser Glu Gly Cys Asn Gly Gly Trp Met Val Glu Gly Phe
65           70           75           80

45 Lys Tyr Ile Ile Asp Asn Lys Ile Ser Gln Thr Ala Asn Tyr Pro Tyr
           85           90           95

Thr Ala Lys Asp Gly Lys Cys Lys Asp Thr Ser Ser Phe Lys Lys Phe
           100           105           110

50 Ser Ile Ser Lys Tyr Ala Glu Ile Pro Gln Gly Asp Cys Asn Ser Leu
           115           120           125

Asn Ser Ala Leu Glu Gln Gly Pro Ile Ser Val Ala Val Asp Ala Thr
55           130           135           140

Asn Phe Gln Phe Tyr Thr Ser Gly Val Phe Lys Asn Cys Lys Ala Asn
145           150           155           160

Leu Asn His Gly Val Leu Leu Val Ala Asn Val Asp Ser Ser Leu Lys
60           165           170           175

Ile Lys Asn Ser Trp Gly Pro Ser Trp Gly Glu Lys Gly Phe Ile Arg
           180           185           190

Leu Ala Ala Gly Asn Thr Cys Gly Val Cys Asn Ala Ala Ser Tyr Pro
65           195           200           205

Ile Val
210

```

ES 2 277 615 T3

<210> 8

<211> 330

<212> ADN

5 <213> *Tetrahymena thermophila*

<400> 8

```

10      atgaacaaga ctctcatctt agctttagtt gttgttttgg ctttaactgc caccaccttg 60
      gttgctttcc acaaccactc tcacaacatc agagttgact aagaccccg cactctcttc 120
      aagcaattca agcaaaactta caataagaag tatgctgatg ctactttcga aacctacaga 180
      ttcggtgtct tcacccaaaa cttagaaatc gtcaagactg actctacttt cgggtgcacc 240
      taattcatgg acttaactcc tgctgaattc gctcaacaat tcctcacttt acacgaaaag 300
      gttaacagca ccgaagttta cagagcttaa
      330

```

<210> 9

<211> 633

<212> ADN

20 <213> *Tetrahymena thermophila*

<400> 9

```

25      ggtgaagcta ccgaagtga ctggactgcc aagggttaagg tcaccctgt taagaactaa 60
      ggttcttggtg gttcctgctg ggctttctcc accattggtg ccggtgaatc tgctctttgg 120
      attgctgggc aagggtgaata aaacactctt aaccttgctg aataagaata agttgactgt 180
      gctaagtccc ccaagtacga ctctgaagggt tgcaacggtg gttggatggt tgaagggttc 240

30      aagtacatca tcgacaacaa gatctcctaa actgctaact atccctacac tgctaaggat 300
      ggtaagtgca aggacacctc ttccctcaag aagttctcta tttctaagta cgctgaaatc 360
      ccctaagggtg actgcaactc cctcaactct gccttagaac aaggtcctat ctccgttgct 420
35      gttgatgcc acaacttcta attctacact tctggtgtct ttaaaaactg caaggccaac 480
      ctcaaccacg gtgtcctctt agttgccaac gttgactctt ctctcaagat caagaactcc 540
      tggggtcctt cttggggtga aaagggtttc atcagattag ctgccggtaa cacttgcggt 600
      gtctgcaatg ctgcctctta cccattgttc tga
      633

```

40

45

50

55

60

65