

PŘIHLÁŠKA VYNÁLEZU

zveřejněná podle § 31 zákona č. 527/1990 Sb.

(19)
ČESKÁ
REPUBLIKA



ÚŘAD
PRŮMYSLOVÉHO
VLASTNICTVÍ

(22) Přihlášeno: **22.01.2002**
(32) Datum podání prioritní přihlášky: **22.01.2001**
(31) Číslo prioritní přihlášky: **2001/10105152**
(33) Země priority: **DE**
(40) Datum zveřejnění přihlášky vynálezu:
(Věstník č: 3/2004)
(86) PCT číslo: **PCT/EP2002/000578**
(87) PCT číslo zveřejnění: **WO 2002/057459**

(21) Číslo dokumentu:

2003-1991

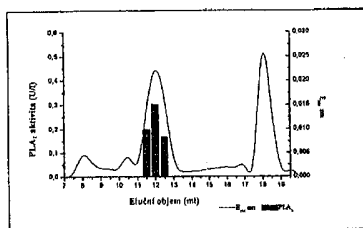
(13) Druh dokumentu: **A3**

(51) Int. Cl. :
C12N 15/55
C12N 9/20
C12N 15/63
W(C12N 9/20, C12R 1:90)

- (71) Přihlašovatel:
CILIAN AG, Münster, DE
- (72) Původce:
Hartmann Marcus, Münster, DE
Grenningloh Marco, Münster, DE
Tiedtke Arno, Münster, DE
- (74) Zástupce:
Čermák Karel Dr., Národní 32, Praha 1, 11000

(54) Název přihlášky vynálezu:
**Sekvence DNA kódující enzym fosfolipáza A1 z
nálevníka Tetrahymena a její použití**

- (57) Anotace:
Nukleová kyselina kódující fosfolipázu A₁, z nálevníků
(ciliat). Konkrétně se jedná o fosfolipázu A₁m která má
sekvenci aminokyselin SEQ ID NO:7.



CZ 2003 - 1991 A3

Sekvence DNA kódující enzym fosfolipáza A₁ z nálevníka
Tetrahymena a její použití

65799

Oblast techniky

Vynález se týká nukleové kyseliny kódující fosfolipázu A₁ a jejího použití podle nároků 1 až 5.

Dosavadní stav techniky

Kvasinky, bakterie a savčí buňky mají velký význam pro biotechnologickou přípravu a produkci rekombinantních účinných látek pomocí heterologní exprese cizích genů. Bakteriální expresní systémy založené na *E.coli* nebo *B.subtilis* jsou používány pro produkci rekombinantních peptidů nebo proteinů, jako je insulin, interleukin-2, tkáňový aktivátor plasminogenu, proteázy a lipázy. U Gram-negativních bakterií jsou expresní systémy většinou založeny na použití genetických prvků, jako je lac operon nebo tryptofanový operon. Proteiny cizorodé pro hostitele jsou produkovány buďto do „inkluzních tělísek“ uvnitř buňky, nebo, jestliže je expresní systém založen na použití β -laktamázových genů, do periplasmatického prostoru. Zatím nebyla provedena produkce rekombinantních proteinů do obklopujícího fermentačního média. Až dosud byly u Gram-pozitivních bakterií téměř výlučně zaváděny do expresních systémů a exprimovány buňkám vlastní proteiny.

Kvasinky, jako je *S.cerevisiae*, *Hansenula polymorpha*, *Kluyveromyces lactis* nebo *Pichia pastoris*, jsou rovněž používány pro heterologní expresi rekombinantních proteinů, jako je lidský faktor XIIIa, bovinní pro-chymosin, fytáza nebo povrchové antigeny. V tomto případě jsou expresní systémy založeny na kyvadlových vektorech (vektory mající

jak kvasinkový, tak bakteriální podíl), které jsou založeny (v závislosti na druhu kvasinky) na genetických prvcích galaktokinázy, epimerázy, methanol oxidázy, kyselé fosfatázy nebo alkohol dehydrogenázy. Zpravidla je rekombinantní protein produkován do cytoplasmu buňky. Jestliže jsou použity vlastní kvasinkové signální sekvence, jako je alfa faktor, rovněž mohou být exprimované proteiny sekretovány do fermentačního média. Glykosylace sekretovaných proteinů je provedena jako typ „high manose“ a často probíhají hyperglykosylace na proteinu, které mohou vést k tvorbě protilátek u pacienta.

Savčí buňky, jako jsou různé buněčné typy z hlodavců (CHO buňky, C127 buňky) nebo z opic (vero, CV-1 nebo COS buňky), jsou rovněž používány pro heterologní expresi rekombinantních proteinů. V tomto případě jsou expresní systémy založeny na rekombinantních virech (BPV vektor) nebo na kyvadlových vektorech. Pro regulaci exprese jsou používány virové SV40 zesilovací/promotorové systémy nebo buněčné zesilující elementy. Rekombinantní proteiny, jako je erythropoietin, jsou sekretovány do fermentačního média, neboť cizí geny obvykle nesou své vlastní signální sekvence, které jsou rozpoznávány expresním systémem a jsou používány k nasměrování.

Kromě toho jsou pro biotechnologickou produkci glykosylovaných extracelulárních enzymů používány protozoa (prvoci) rodu *Tetrahymena*. *Tetrahymena* roste na nenákladném fermentačním médiu za použití standardních fermentačních metod. Pro transformaci takových buněk *Tetrahymena* jsou dostupné vektory, které jsou založeny na elementech rDNA. z *Tetrahymena* Pro heterologní expresi bakteriálních proteinů v *Tetrahymena* jsou používány konstrukty DNA připravené z genů z *Tetrahymena*. Jestliže jsou dostupné vhodné genetické prvky pro regulaci transkripce, nasměrování a

glykosylaci cizích proteinů, je *Tetrahymena* ideálním expresním systémem pro nenákladný způsob výroby terapeutických rekombinantních proteinů.

Až dosud používané Gram-negativní bakteriální expresní systémy obvykle vedou k tvorbě „inkluzních tělísek“ v buňce, spojené s denaturací proteinů. Pro získání rekombinantního proteinu musí buňky být lyzovány a denaturovaný inaktivovaný protein musí být zpětně poskládán, aby byl funkční. Toto představuje dodatečné nákladné stupně způsobu a snižuje výtěžek požadovaného proteinu. Glykosylace, která je u eukaryontních proteinů významná, je kompletně vynechána. Jestliže je použit bakteriální expresní systém, je degradace dalším problémem cílového proteinu díky vysoké proteolytické aktivitě ve fermentační půdě.

Jestliže jsou použity kvasinky pro heterologní expresi, požadovaný cílový protein je rovněž produkován pouze do buňky, ze které musí být odstraněn pomocí lyze buněk. Podobně jako v bakteriálním expresním systému, způsobuje tato skutečnost dodatečné časově a finančně nákladné stupně způsobu. Jestliže jsou použity vlastní kvasinkové signální sekvence, nejsou cizí proteiny správně sestřihány a glykosylovány pro sekreci.

Naopak, jestliže je pro produkci rekombinantních proteinů použit savčí buněčný systém, jsou požadované proteiny nalézány ve fermentačním médiu v extracelulárním stavu, správně sestřihány a glykosylovány. Nicméně, nevýhodou je, že na jedné straně dochází k nízké hladině exprese díky nedokonalé úpravě a neúčinné translaci genů, které byly zavedeny do genomu produkční buněčné linie prostřednictvím virálních vektorů. Na druhé straně jsou fermentační média pro savčí buňky obsahující sérum extrémně finančně nákladná. Kromě toho je fermentační technologie

pro buněčné linie citlivé na stříh komplikovaná a podobně nákladná díky konstrukci bezbublinové aerace. Další problémy vznikají z vysokého rizika mykoplasmatické a virové infekce buněčných linií. Shrnutí, použití savčích buněk pro biotechnologický způsob přípravy rekombinantních proteinů má přináší velmi vysoké náklady, bezpečnostní požadavky a nízké výtěžky.

S použitím nálevníků, jako je *Tetrahymena*, odpadají výše uvedené nevýhody při produkci rekombinantních proteinů. Tak například, některé kyselé hydrolázy, které působí při trávení částic potravy, jsou vylučovány z buněk ve velkých množstvích a s komplexní glykosylací.

Alam et al., J.Euk.Microbiol., 43 (4), 1996, 295-303 popisuje klonování genu, který kóduje kyselou α -glukosidázu z *Tetrahymena pyriformis*. Nicméně, pouze malý podíl proteinu je vylučován z buněk. Kromě toho, Mezinárodní patentová přihláška PCT/EP 00/01853 popisuje gen pro β -hexosaminidázu z *Tetrahymena thermophila*, o které je nicméně známo, že je vylučována z buňky pouze z 80 %.

Až dosud nebylo možné exprimovat glykosylované eukaryontní proteiny v *Tetrahymena* s výlučnou sekrecí do fermentačního prostředí. Je to proto, že sekvence DNA pro extracelulární proteiny vlastní *Tetrahymena*, které jsou nezbytné pro konstrukci expresních vektorů a které jedinečně exportují cizí protein do obklopujícího fermentačního média, nebyly až dosud známy. Jsou známy sekvence DNA, které kódují protein β -hexosaminidázu z *Tetrahymena thermophila*. Takové sekvence byly popsány v přihlášce vynálezu číslo DE 199 58 979.8, DE 199 09 189.7 a PCT/EP 00/01853. Nicméně, nevýhodou těchto sekvencí je, že v nich obsažené pre/propeptidy způsobují uvolňování požadovaného pro hostitele cizího proteinu do obklopujícího fermentačního média pouze z 80 %. Je to dáno skutečností,

že za normálních podmínek je enzym β -hexosaminidáza přítomna uvnitř membrány přibližně až z 20 % a pouze 80 % přirozeně produkovaného enzymu je vylučováno z buněk. Z tohoto důvodu pre/propeptidy β -hexosaminidázy, jestliže jsou metodami genetického inženýrství umístěny před protein cizí pro hostitele, nasměrují pouze přibližně 20 % cizorodého proteinu do cytoplasmatické membrány na povrchu *Tetrahymena thermophila*. Toto je spojeno s významnou nevýhodou technologie postupu pro způsob přípravy rekombinantních účinných látek. Na jedné straně, je výtěžek snižován, protože část exprimovaného proteinu zůstává v buňkách vázána k membráně, a proto není možné purifikovat veškerý exprimovaný protein z fermentační pudy. Na druhé straně pro hostitele cizí protein může v buněčné membráně projevovat toxické účinky pro hostitele, a tak snižovat buněčný růst.

Kromě toho, žádné konstitutivní promotory z *Tetrahymena*, které způsobují odpovídající nebo kontinuální transkripci heterologních proteinů, nejsou až dosud známy. Dosud jsou známy pouze promotory histonových a tubulinových genů (Bannon et al., 1984, Gaertig et al., 1993). Nicméně, zásadní nevýhodou těchto promotorů je, že jejich aktivace je závislá na buněčném cyklu. Geny pro heterologní protein, které jsou spojeny s takovými na buněčném cyklu závislejšími promotory, jsou exprimovány pouze v rostoucích a dělicích se buňkách. Toto má závažné nevýhody pro technologii postupu, neboť požadovaný protein je tak produkován pouze v logaritmické fázi růstu. Ve stacionární fázi růstu, ve kterém je dosaženo nejvyšší hustoty buněk a tím nejvyššího výkonu exprimovaného organismu (*Tetrahymena*) ve způsobu produkce, je stěží zachován buněčný růst, a tak se uskutečňuje pouze nízká exprese heterologního proteinu.

Podstata vynálezu

Cílem vynálezu je poskytnout systém, který umožňuje produkci heterologních proteinů z buněk do fermentačního média v expresním systému, kterým je transformovaný nálevník *Tetrahymena*

Tohoto cíle je dosaženo systémem, v němž je použito nukleové kyseliny, která má sekvenci SEQ ID NO:1, kódující fosfolipázu A₁ (SEQ ID NO:7). Během kultivačních podmínek produkt exprese této DNA výhodně exportován z buňky ve velkém množství. Exprimovaný protein je uvolňován do obklopujícího kultivačního média ve velkém množství a není obsažen v membráně. Sekvence nukleové kyseliny podle vynálezu obsahuje promotor, který způsobuje konstitutivní, tj. na buněčném cyklu nezávislou, transkripci po směru genů pro heterologní proteiny. Taková konstitutivní transkripce má tu výhodu, že proteiny jsou kontinuálně exprimovány pomocí heterologní exprese v hostitelském organismu, aniž by byla ovlivněna buněčným cyklem. Z tohoto důvodu je transkripce cizího genu účinná a heterologní protein je exprimován rovněž během stacionární růstové fáze bez buněčného růstu.

Sekvence DNA pro fosfolipázu A₁ podle vynálezu obsahuje výhodně oblast „proti směru“ PLA₁ (SEQ ID NO:2), která nese prvky promotoru pro iniciaci transkripce, signální peptid a propeptid, další genetické elementy pro nasměrování proteinu, a zejména PLA₁ oblast „po směru“ (SEQ ID NO:3), která obsahuje genetické prvky pro ukončení transkripce. Použití těchto nukleových kyselin ve vektoru umožňuje expresi heterologně exprimovaných proteinů nezávisle na buněčném cyklu a jejich transport ven z buňky do obklopujícího kultivačního média, aniž by exprimované proteiny byly inkorporovány do cytoplasmatické membrány, a

takové proteiny mohou být izolovány z fermentační pŕdy bez buněčné lyze.

Obrázek 1 uvádí nukleovou kyselinu kódující oblast „proti směru“ (SEQ ID NO:2), kódující oblast (SEQ ID NO:1) a oblast „po směru“ (SEQ ID NO:3) pro fosfolipázu A₁ z nálevníků.

Obrázek 2 uvádí produkt exprese odpovídající nukleové kyselině podle SEQ ID NO:1. Vynález se také týká proteinu podle SEQ ID NO:7.

Konkrétně se vynález týká signální sekvence (SEQ ID NO:6) proteinu podle vynálezu. Výhodně se jedná o aminokyseliny 1 až 110 proteinu podle vynálezu (SEQ ID NO:5). Vynález se rovněž týká nukleové kyseliny kódující N-terminální fragment (SEQ ID NO:3). Jedná se přednostně o fragment nukleových kyselin podle vynálezu (SEQ ID NO:4), zejména mající nukleové kyseliny 1 až 155 podle Obrázku 1.

Sekvence nukleové kyseliny netranslatové oblasti („oblast proti směru“) (SEQ ID NO:2) umístěná opačně do protisměru k oblasti kódující PLA₁ z *Tetrahymena*, je umístěna mezi pozicí -275 a pozicí -1 (znázorněno malými písmeny). Stanovená netranslatovaná oblast obsahuje 275 bazí. Jako prvky promotoru jsou přítomny TATA box v pozicích -49 až -55 (tištěno tučně), předpokládaný CAAT box mezi bazí -133 a bazí -136 (tištěno tučně). Rozsah kódující sekvence cDNA je znázorněn velkými písmeny. Číslování sekvence začíná start kodónem ATG. Oblasti známé ze sekvenování proteinu jsou v rámečku a stop kodón je podtržen. Zralý protein je kódován od baze 331. Sekvenční pořadí od baze 1 do baze 330 představuje pre/pro sekvenci (SEQ ID NO:8) PLA₁. Sekvenční pořadí od baze 331 do baze 963 je sekvencí zralé PLA₁. (SEQ ID NO:9). V pozici 961 je translační stop kodón TGA a v pozici 1039 je polyadenylační signál AAT AAA. Sekvence nukleové kyseliny od pozice 964 do

pozice 1134, která je uvedena pod kódující sekvencí pro PLA₁ z *Tetrahymena*, znázorňuje sekvenci „oblasti po směru“ PLA₁. (SEQ ID NO:3), která není translatována (rovněž je uvedena malými písmeny). Od pozice 964 až do pozice 1101 je oblast známá ze sekvenování cDNA, která byla rovněž potvrzena inverzní PCR. Po transkripci je připojen poly(A)-konec k poslednímu kodónu mRNA (ttt, pozice 1098-1101).

Dalším aspektem vynálezu je použití sekvence nukleové kyseliny pro kyselé hydrolázy podle vynálezu nebo jejich částí pro homologní nebo heterologní expresi rekombinantních proteinů a peptidů a pro homologní nebo heterologní rekombinaci („knock-out“ - vyřazení, „gene replacement“- genová substituce), ve které jsou nálevníci tranfektováni nukleovou kyselinou podle vynálezu.

Nukleové kyseliny nebo jejich části mohou být kombinovány, konkrétně se zesilovači, promotory, operátory, počátky, terminátory, geny rezistence na antibiotika, které jsou běžné pro homologní nebo heterologní expresi proteinů, nebo s jinými nukleovými kyselinami nebo fragmenty DNA, nebo se všemi typy sekvencí z viroidů, virů, bakterií, archezoí, protozoí, hub, rostlin, zvířat nebo lidí.

Konkrétně je nukleová kyselina podle vynálezu obsažena ve vektoru, plasmidu, kosmidu, chromozomu nebo minichromozomu, transposonu, elementu IS, rDNA, nebo ve všech typech cirkulární nebo lineární DNA nebo RNA.

Vynález se také týká způsobu, ve kterém nukleová kyselina nebo její části podle vynálezu, které kódují fosfolipázu A₁, jsou kombinovány v homologní nebo heterologní expresi s obvyklými zesilovači, jako je oblast NF-1 (cytomegalovirový zesilovač), promotory, jako je lac, trc, tic nebo tac promotor, promotory třídy II a III systému T7 RNAP (T7 RNA polymeráza), promotory bakteriofága T7 a SP6, aprE, amylázový nebo spac promotor pro systémy

exprese v druhu *Bacillus*, AOX1, AUG1 a 2 nebo GAPp promotory (*Pichia*) pro systémy exprese v kvasinkách, RSV promotor (virus SV40), CMV promotor (Cytomegalovirus), AFP promotor (adenoviry) nebo metallothioninové promotory pro expresi v savčích systémech, promotory viru Sindibis nebo promotory viru Semlike forest pro hmyzí buňky, promotory pro hmyzí systémy exprese, jako je hsp70, DS47, aktin 5C nebo copia, specifické rostlinné promotory, jako je promotor 35S, (virus kvěťákové mozaiky), amylázový promotor, nebo patatinový promotor třídy I, operátory, jako je tet operátor, signální peptidy, jako je a-MF prepro-signální sekvence (*Saccharomyces*), počátky, terminátory, geny pro rezistenci na antibiotika a jiné látky, jako je ampicillin, kanamycin, streptomycin, chloramfenikol, penicilin, amfotericin, cykloheximid, 6-methylpurin, paromycin, hygromycin, α -amanatin, auxotrofní markery, jako je gen pro dihydrofolát reduktázu, nebo jiné nukleové kyseliny nebo fragmenty DNA, nebo všechny typy sekvencí z viroidů, virů, bakterií, archeozoi, protozoi, hub, rostlin, zvířat nebo lidí.

Konkrétně, nukleová kyselina nebo její část podle vynálezu jsou včleněny do vektoru, plasmidu, kosmidu, chromozomu nebo minichromozomu, transposonu, IS elementu, rDNA, nebo všech typů cirkulární nebo lineární DNA nebo RNA.

Pro odborníka v oboru je zřejmé, že nukleové kyseliny mající alespoň 40% homologii s nukleovou kyselinou podle SEQ ID NO:1 mohou být použity podle vynálezu. Protein odpovídající SEQ ID NO:2 může být modifikován bez ztráty své funkce. Tak například mohou být provedeny tzv. konzervativní záměny aminokyselin. Například mohou být provedeny vzájemné výměny hydrofobních aminokyselin.

Pro čištění a izolaci fosfolipázy A_1 z *Tetrahymena* a pro stanovení její sekvence mohou být použity následující postupy.

Příklady provedení vynálezu

Získání PLA_1

PLA_1 byla získána z bezbuněčného supernatantu z *Tetrahymena thermophila*. Buňky byly fermentovány ve 2l fermentorech (Biostat MD, Braun Diessel Biotech, Melsungen, Germany), které byly řízeny digitální řídicí jednotkou (DCU, digital controlling unit). Fermentor nejprve pracoval po dobu 24 hodin vsádkovým systémem a pak kontinuálně. Sklizení bezbuněčného supernatantu bylo zajištěno promývacím modulem, který měl velikost pórů přibližně 0,3 μm (S6/2, Wuppertal).

Fermentace probíhala za následujících parametrů:

- . pracovní objem byl 2 litry;
- . rychlost promývání byla 2 litry/den;
- . otáčky míchadla byly stanoveny na 800 rpm (ot./min.);
- . teplota byla konstantní 30 °C;
- . hodnota pH byla udržována konstantní na pH 7;
- . inokulační titr byl 50 000 buněk/ml.

Pro fermentaci byl použit kmen SB1868 VII. Jedná se o kmen divokého typu (standardní) *Tetrahymena thermophila*. Fermentace probíhala po dobu 264 hodin a sklizený výtěžek byl testován na aktivitu PLA_1 .

Čištění PLA_1

Pro čištění PLA_1 byl použit 1 litr bezbuněčného supernatantu z fermentace, který byl smíchán se 140 g

síranu amonného a zakoncentrován pomocí ultrafiltrační jednotky (Pellikon XL, propustnost (size exclusion) 3 kDa, Millipore) na objem 50 ml. Následně byl vzorek purifikován hydrofóbní interakční chromatografií (20 x 1,6 Fractogel EMD Phenyl I 650, Merck, Darmstadt). Rychlost průtoku byla 5 ml/min a eluát byl sbírán do 5ml frakcí. Aktivita enzymu byla měřena deacylací radioaktivně značeného fosfolipidu (L-3-fosfatidylcholin, 1-palmitoyl-2-[1-¹⁴C]linoleoyl).

Obrázek 3 znázorňuje získaný eluční profil v natriumacetátovém gradientu a enzymové aktivity jednotlivých frakcí.

Tři frakce, které měly nejvyšší aktivitu enzymu, byly spojeny a opětovně pufrovány pomocí výchozího pufru (Bis-Tris 20 mM, pH 6,5) pro aniontově-výměnnou chromatografii (anion-exchange chromatography, AEC) s použitím ultrafiltrační jednotky. Následně byly vzorky nanесeny na sloupec (Q-Sepharose-Hiload-16/10, Pharmacia, Sweden) a PLA₁ byla ze sloupce eluována lineárním gradientem NaCl (průtoková rychlost = 3 ml/min) a sbírána do 5ml frakcí. Obrázek 4 znázorňuje získaný eluční profil v NaCl-gradientu a enzymové aktivity jednotlivých frakcí.

Z frakcí, které měly nejvyšší PLA₁ aktivitu, bylo 200 µl odstraněno a odděleno SEC chromatografií (size exclusion chromatography). K tomuto účelu byl použit sloupec Superdex HR 75 30/10 (Pharmacia, Sweden). Rychlost promývání v této chromatografii byla 0,6 ml/min, eluát byl sbírán do 200 µl frakcí. Obrázek 5 uvádí získaný eluční profil a enzymové aktivity jednotlivých frakcí.

Získané frakce byla testovány na čistotu s použitím trojrozměrné gelové elektroforézy. Byly stanoveny dva odlišné pruhy při velikosti ~ 26 a ~ 28 kDa. Separace těchto dvou pruhů dvojrozměrnou gelovou elektroforézou

vedla k rozdělení dvou pruhů do 2 a 3 skvrn, v uvedeném pořadí, s odlišnými izoelektrickými body.

Proteiny o 26 kDa odpovídaly pH 6,3 a 5,7 a proteiny o 28 kDa odpovídaly pH 6,3, 5,7 a 5,3. Konečná analýza těchto skvrn pomocí mass fingerprint analýzy prokázala, že tyto skvrny jsou izoformami stejného proteinu.

Molekulárně-biologické ověření PLA₁

Jakmile byla prokázána čistota proteinu, byly vzorky proteinu nanесeny na PVDF membránu a podrobeny sekvenování od N konce. Kromě toho byly další vzorky štěpeny trypsinem a rovněž podrobeny sekvenování od počátku. S použitím takto získaných proteinových sekvencí byly připraveny oligonukleotidové primery, které byly podrobeny RT-PCR (reverse transcriptase PCR; 3'RACE, rapid amplification of cDNA ends). S použitím této techniky PCR byla fosfolipáza A₁ úspěšně amplifikována a následně sekvenována. Získaná sekvence měla délku 633 bazí a 729 bazí v uvedeném pořadí, a molekulová hmotnost zralého proteinu z ní odvozená je přibližně 22,4 kDa. V odvozené sekvenci oligopeptidy o 22 aminokyselinách (N-terminální) a o 18 aminokyselinách (uvnitř peptidu) stanovené sekvenováním proteinu byly opětovně nalézány ve 100 %. Kromě toho sekvence zralého proteinu, sekvence pre/propeptidu mohly rovněž být určeny pomocí 5'RACE (rapid amplification of cDNA ends) (Obrázek 2). Jedná se o peptid mající délku 110 aminokyselin, který nese jak signální sekvenci, tak sekvenci propeptidu, jenž enzym inaktivuje a je odštěpován pouze v místě konečné aktivity enzymu. Sekvenční porovnání poskytlo zjištění, že neexistuje žádná homologie s dříve známými fosfolipázami A₁, s výjimkou konsensus sekvence 5 aminokyselin (GxSxG), která je nalézána ve všech lipázách a fosfolipázách a je

považována za vazebné místo pro lipidy nebo fosfolipidy. Kromě toho, PLA₁ „sekvence po směru“ a „sekvence proti směru“ byly stanoveny inverzní PCR (Obrázek 1). Genomová DNA byly rozštěpena restričními endonukleázami, ligována T4 ligázou a nakonec amplifikována s inverzními primery. V inverzní PCR byla pro amplifikaci PLA₁ „oblasti proti směru“ ke štěpení genomové DNA použita restriční endonukleáza SspI. Proto mohla být stanovena „oblast proti směru“ o 275 bazích a byly identifikovány promotorové prvky. V pozici -136 se nalézá CAAT box, který má podobnou vzdálenost od translačního startu, jako CCAAT boxy histonových genů (-141 a -151) *Tetrahymena*, jak našli Brunk a Sadler (1990). TATA box, který fixuje exaktní počáteční bod transkripce v eukaryontních genech, byl identifikován v pozici -55. Jeho sekvence odpovídá konsensus sekvenci nalezené u eukaryont. Pro amplifikaci „oblasti po směru“, která obsahuje terminátor transkripce PLA₁ z *Tetrahymena*, metodou PCR, bylo použito štěpení DNA pomocí BamHI. Proto, kromě oblasti „po směru“ známé z 3'-RACE, mohlo být určeno dalších 222 bazí (Obrázek 3).

1. Nukleová kyselina obsahující kódující a nekódující oblasti mající sekvenci SEQ ID NO:4, ve které řečené kódující oblasti kódují fosfolipázu A₁ (PLA₁) z *Tetrahymena thermophila* (SEQ ID NO:5) a řečené nekódující oblasti znamenají oblasti „proti směru“ a „po směru“ od tohoto genu.
2. Nukleová kyselina podle nároku 1, ve které řečené nekódující oblasti „po směru“ a „proti směru“ mají sekvence SEQ ID NO:2 a 3 a kódující oblasti mají sekvence SEQ ID NO:1, 8 a/nebo 9.
3. N-terminální signální sekvence PLA₁ (SEQ ID NO:5) podle nároku 1 mající sekvenci SEQ ID NO:6.
4. Protein fosfolipázy A₁ (PLA₁) mající aminokyselinovou sekvenci podle SEQ ID NO:7.
6. Použití nukleové kyseliny podle nároků 1 až 3 nebo jejích částí pro homologní nebo heterologní expresi rekombinantních proteinů a peptidů a pro homologní nebo heterologní rekombinaci (vyřazení „knock-out“, substituce „gene replacement“).
7. Způsob homologní nebo heterologní exprese proteinů a peptidů a homologní nebo heterologní rekombinace („knock out“, „gene replacement“), ve kterém jsou nálevníci (ciliata) transfektováni nukleovou kyselinou podle jakéhokoliv z nároků 1 až 3.

8. Způsob podle nároku 9, ve kterém řečené nukleové kyseliny nebo jejich části jsou kombinovány se zesilovači, promotory, operátory, počátky, terminátory, geny pro antibiotickou rezistencí běžnými pro homologní nebo heterologní expresi proteinů, nebo s jinými nukleovými kyselinami nebo fragmenty DNA, nebo se všemi typy sekvencí z viroidů, virů, bakterií, archezoí, protozoí, hub, rostlin, zvířat a lidí.

11. Způsob podle nároku 10 a/nebo 11, ve kterém je nukleová kyselina je obsažena ve vektoru, plasmidu, kosmidu, chromozomu nebo minichromozomu, transposonu, IS elementu, rDNA, nebo v jakémkoliv jiném typu cirkulární nebo lineární DNA nebo RNA.

12. Vektory, plasmidy, kosmidy, chromozomy nebo minichromozomy, transposony, IS elementy, rDNA a jiné typy cirkulární nebo lineární DNA nebo RNA obsahující alespoň jednu z nukleových kyselin podle alespoň jednoho z nároků 1 až 3.

15.09.03

1/4

PV/403-1991

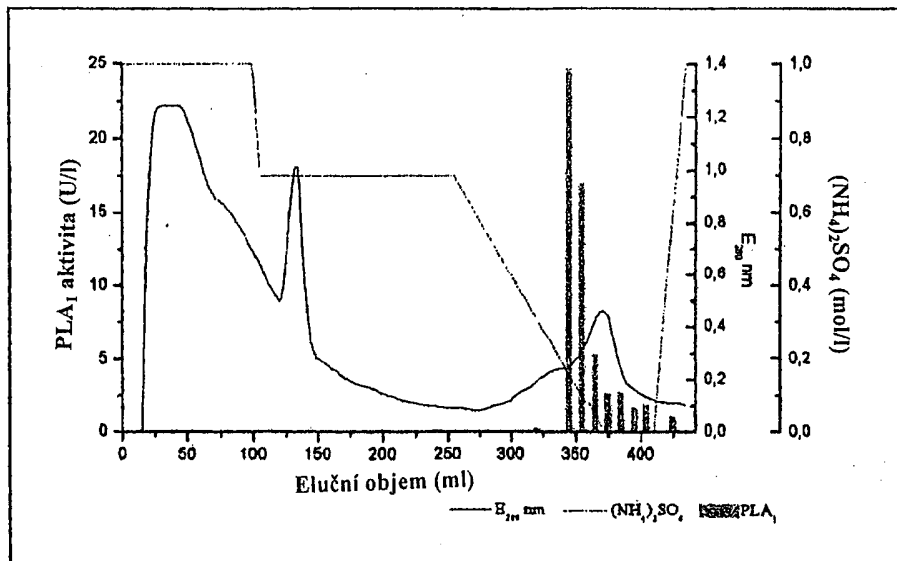
65791

Obr. 1

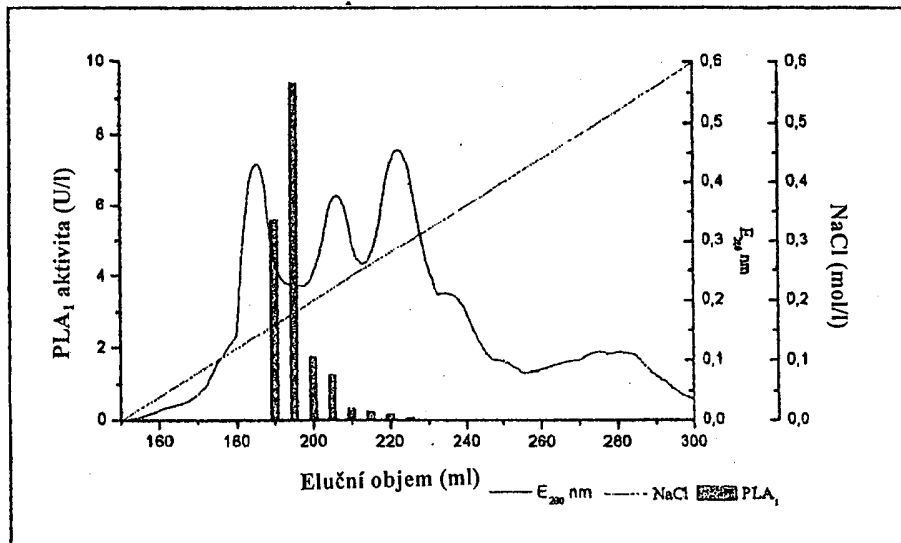
tga	gat	atg	ata	tgc	tct	ttc	tct	gct	aga	ctt	aac	tta	tga	cat	ttg	aac	ttt	taa	taa	-241
aag	aat	ttt	ttt	tat	taa	aaa	gca	gag	att	ttt	aat	aga	aga	atc	aat	gac	tca	tga	att	-181
taa	taa	aga	ttt	tca	agt	gtt	ttc	taa	tac	cga	cta	gct	tta	taa	alt	cac	tta	tta	atc	-121
aac	gat	ata	aaa	alt	ata	tta	aca	aat	caa	taa	ata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

[illegible]

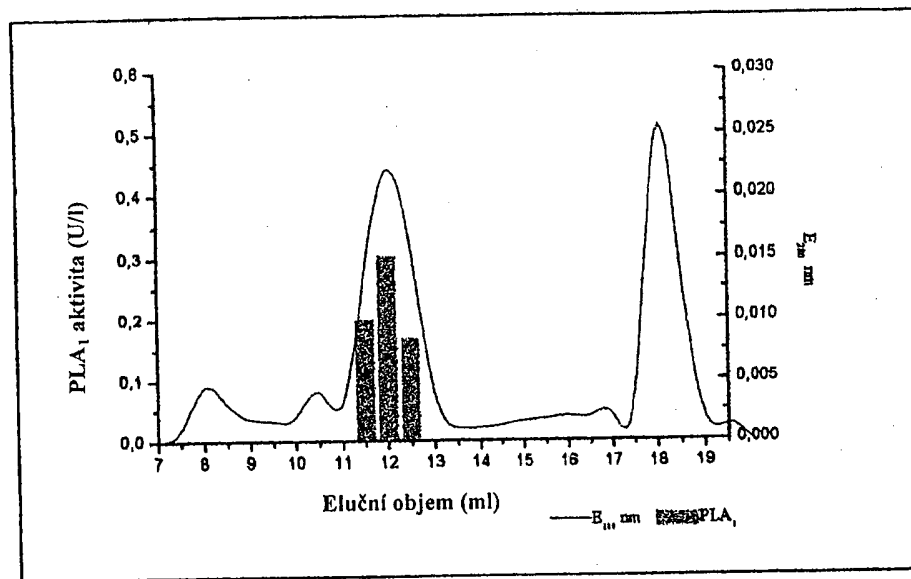
Obr. 3

HIC I, eluční profil, natriumacetátový gradient a enzymatická aktivita eluované PLA₁

Obr. 4

AEC, eluční profil, NaCl-gradient a enzymatická aktivita eluované PLA₁

Obr. 5
SEC, eluční profil a enzymatická aktivita PLA_1



16.09.03

P/2003-1991

Sekvenční protokolUpravený (20.12.2002

6791

<110> Cilian AG

<120> Sekvence DNA kódující enzym fosfolipáza A₁ z nálevníka *Tetrahymena* a její použití

<130> 020414wo ME/BM

<140> PCT/EP02/00578

<141> 2002-01-22

<160> 9

<170> PatentIn Ver. 2.1

<210> 1

<211> 963

<212> DNA

<213> *Tetrahymena thermophila*

<400> 1

```

atgaacaaga ctctcatctt agcttttagtt gttgttttgg ctttaactgc caccaccttg 60
gttgctttcc acaaccactc tcacaacatc agagttgact aagaccccg cactctcttc 120
aagcaattca agcaaactta caataagaag tatgtgatg ctactttcg aacctacaga 180
ttcgggtgtc tcacccaaaa cttagaaatc gtcaagactg actctacttt cgggtgcacc 240
taattcatgg acttaactcc tgetgaatc gctcaacaat tcctcacttt acacgaaaag 300
gttaacagca ccgaagttaa cagagcttaa ggtgaagcta ccgaagtga ctggactgcc 360
aagggttaagg tcacccctgt taagaactaa ggttcttggg gttcctgctg ggctttctcc 420
accattgggtg ccgttgaatc tgetcttggg attgctgggc aagggtgaata aaacactctt 480
aaccttgctg aataagaata agttgactgt gctaagtcac ccaagtacga ctctgaaggt 540
tgcaacgggtg gttggatggt tgaaggtttc aagtacatca tcgacaacaa gatctcctaa 600
actgctaact atccctacac tgctaaggat ggtaagtga aggacacctc ttccttcaag 660
aagttctcta tttctaagta cgtgaaatc ccctaagggt actgcaactc cctcaactct 720
gccttagaac aaggctctat ctccgttgct gttgatgcca ccaacttcta attctacact 780
tctgggtgtc ttaaaaactg caaggccaac ctcaaccacg gtgtcctctt agttgccaac 840
gttgactctt ctctcaagat caagaactcc tggggtcctt cttgggggtga aaagggtttc 900
atcagattag ctgccggtaa cacttgcggt gtctgcaatg ctgectctta ccctattggt 960
tga 963

```

<210> 2

<211> 275

<212> DNA

<213> *Tetrahymena thermophila*

<400> 2

```

aatatttata aatgctactt ataattcttt tagtatgaga 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 aaaaaataa ataaaaacaa aacaa 275

```

<210> 3

<211> 171

<212> DNA

<213> *Tetrahymena thermophila*

<400> 3

```

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

```

<210> 4
 <211> 1408
 <212> DNA
 <213> Tetrahymena thermophila

<400> 4
 aatatttatc aatgctactt ataattcttt tagtatgaga tatgatatgc tctttctctg 60
 ctagacttaa cttatgacat ttgaaacttt aataaaagaa ttttttttat taaaaagcag 120
 agatttttaa tagaagaatc aatgactcat gaatttaata aagattttca agtggtttct 180
 aataccgact agctttataa attcacttat taatcaacga tataaaaatt atattaacaa 240
 atcaataaat aaaaaataa ataaaaacaa aacaaatgaa caagactctc atcttagctt 300
 tagttgttgt tttggcttta actgccacca ccttggttgc tttccacaac cactctcaca 360
 acatcagagt tgactaagac cccgccactc tcttcaagca attcaagcaa acttacaata 420
 agaagtatgc tgatgctact ttcgaaacct acagattcgg tgtcttcacc caaaacttag 480
 aaatcgtaaa gactgactct actttcgggtg tcacctaatt catggactta actcctgctg 540
 aattcgctca acaattcctc actttcacga aaagggttaac agcaccgaag tttacagagc 600
 ttaagggtgaa gctaccgaag ttgactggac tgccaagggt aagggtcacc ctggttaagaa 660
 ctaagggttct tgtgggttct gctgggcttt ctccaccatt ggtgcccgtt aatctgctct 720
 ttggattgct ggtcaagggt aataaaacac tcttaacctt gctgaataag aataagttga 780
 ctgtgctaag tcccccaagt acgactctga aggttgcaac ggtggttga tggttgaagg 840
 tttcaagtac atcatcgaca acaagatctc ctaaaactgct aactatccct acactgctaa 900
 ggatggtaag tgcaaggaca cctcttctct caagaagttc tctatttcta agtacgctga 960
 aatccctcaa ggtgactgca actccctcaa ctctgcctta gaacaagggt cttatctcgt 1020
 tgctgttgat gccaccaact tctaattcta cacttctggt gtctttaaaa actgcaaggc 1080
 caacctcaac cacggtgtcc tcttagttgc caacgttgac tcttctctca agatcaagaa 1140
 ctctgggggt ccttcttggg gtgaaaagggt tttcatcaga ttagctgcog gtaacacttg 1200
 cgggtgtctgc aatgctgcct cttaccctat tgtttgaaaa aacataatcc aaattaaaaa 1260
 aaattactca aaactgataa tataaaaaat taattttcat aattttaatg taaataaata 1320
 cctttatatt tgacgttttg tactcaaaat aaattaaagt taacaaacca tattttatta 1380
 attctacttt tcaattttta aaaatata 1408

<210> 5
 <211> 320
 <212> PRT
 <213> Tetrahymena thermophila

<400> 5
 Met Asn Lys Thr Leu Ile Leu Ala Leu Val Gly Val Leu Ala Leu Thr
 1 5 10 15
 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
 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
 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 Gly Glu
 100 105 110
 Ala Thr Glu Val Asp Trp Thr Ala Lys Gly Lys Val Thr Pro Val Lys
 115 120 125
 Asn Gln Gly Ser Cys Gly Ser Cys Trp Ala Phe Ser Thr Ile Gly Ala
 130 135 140

15.09.03

Val Glu Ser Ala Leu Leu Ile Ala Gly Gln Gly Glu Gln Asn Thr Leu
145 150 155 160

Asn Leu Ala Glu Gln Glu Leu Val Asp Cys Ala Lys Ser Pro Lys Tyr
165 170 175

Asp Ser Glu Gly Cys Asn Gly Gly Trp Met Val Glu Gly Phe Lys Tyr
180 185 190

Ile Ile Asp Asn Lys Ile Ser Gln Thr Ala Asn Tyr Pro Tyr Thr Ala
195 200 205

Lys Asp Gly Lys Cys Lys Asp Thr Ser Ser Phe Lys Lys Phe Ser Ile
210 215 220

Ser Lys Tyr Ala Glu Ile Pro Gln Gly Asp Cys Asn Ser Leu Asn Ser
225 230 235 240

Ala Leu Glu Gln Gly Pro Ile Ser Val Ala Val Asp Ala Thr Asn Phe
245 250 255

Gln Phe Tyr Thr Ser Gly Val Phe Lys Asn Cys Lys Ala Asn Leu Asn
260 265 270

His Gly Val Leu Leu Val Ala Asn Val Asp Ser Ser Leu Lys Ile Lys
275 280 285

Asn Ser Trp Gly Pro Ser Trp Gly Glu Lys Gly Phe Ile Arg Leu Ala
290 295 300

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

<210> 6
<211> 110
<212> PRT
<213> Tetrahymena thermophila

<400> 6
Met Asn Lys Thr Leu Ile Leu Ala Leu Val Gly Val Leu Ala Leu Thr
1 5 10 15

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

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

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
100 105 110

15.09.03

-4-

<210> 7
 <211> 210
 <212> PRT
 <213> Tetrahymena thermophila

<400> 7
 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
 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
 50 55 60
 Lys Tyr Asp Ser Glu Gly Cys Asn Gly Gly Trp Met Val Glu Gly Phe
 65 70 75 80
 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
 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
 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
 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
 195 200 205
 Ile Val
 210

<210> 8
 <211> 330
 <212> DNA
 <213> Tetrahymena thermophila

<400> 8
 atgaacaaga ctctcatctt agcttttagtt gttgttttgg ctttaactgc caccaccttg 60
 gttgctttcc acaaccactc tcacaacatc agagtttgact aagaccccg cactctcttc 120
 aagcaattca agcaaaacta caataagaag tatgctgatg ctactttcga aacctacaga 180
 ttcggtgtct tcacccaaaa cttagaaatc gtcaagactg actctacttt cgggtgtcacc 240
 taattcatgg acttaactcc tgctgaattc gctcaacaat tctctacttt acacgaaaag 300
 gttaacagca ccgaagttaa cagagcttaa 330

<210> 9
 <211> 633

15.09.03

<212> DNA

<213> *Tetrahymena thermophila*

<400> 9

```

ggggaagcta ccgaagttga ctggactgcc aagggttaagg tcacccctgt taagaactaa 60
gggtcttgtg gttcctgctg ggctttctcc accattggtg ccgttgaatc tgctctttgg 120
attgctgggc aagggtgaata aaacactctt aaccttgctg aataagaata agttgactgt 180
gctaagtccc ccaagtacga ctctgaaggc tgcaacggcg gttggatggt tgaaggtttc 240
aagtacatca tcgacaacaa gatctcctaa actgctaact atccctacac tgctaaggat 300
ggtaagtgcg aggacacctc ttccttcaag aagttctcta tttctaagta cgctgaaatc 360
ccctaagggtg actgcaactc cctcaactct gccttagaac aaggctctat ctccgttgct 420
gttgatgcca ccaacttcta attctacact tctgggtgtc ttaaaaactg caaggccaac 480
ctcaaccacg gtgtcctctt agttgccaac gttgactctt ctctcaagat caagaactcc 540
tggggtcctt cttgggggtga aaagggtttc atcagattag ctgccggtaa cacttgcggt 600
gtctgcaatg ctgcctctta ccctattggt tga 633

```