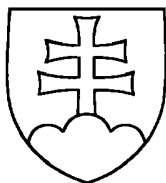


SLOVENSKÁ REPUBLIKA

(19)

SK



ÚRAD  
PRIEMYSELNÉHO  
VLASTNÍCTVA  
SLOVENSKEJ REPUBLIKY

ZVEREJNENÁ PRIHLÁŠKA  
VYNÁLEZU

(21) Číslo dokumentu:

246-99

(22) Dátum podania:

(31) Číslo prioritnej prihlášky:

(32) Dátum priority:

(33) Krajina priority:

(40) Dátum zverejnenia: 10.04.2000

(86) Číslo PCT:

(13) Druh dokumentu: A3

(51) Int. Cl.7 :

C 12N 1/20,  
C 12N 15/09,  
15/10, 15/29,  
C 12N 15/31,  
C 12N 15/82,  
A 01G 13/00,  
A 01H 1/00,  
3/00, 4/00

(71) Prihlasovateľ: DOW AgroSciences LLC, Indianapolis, IN, US;  
WISCONSIN ALUMNI RESEARCH FOUNDATION, Madison, WI, US;

(72) Pôvodca vynálezu: Ensign Jerald C., Madison, WI, US;  
Bowen David J., Oregon, WI, US;  
Petell James, Zionsville, IN, US;  
Fatig Raymond, Zionsville, IN, US;  
Schoonover Sue, Brownsburg, IN, US;  
Ffrench-Constant Richard H., Madison, WI, US;  
Rocheleau Thomas A., Madison, WI, US;  
Blackburn Michael B., Madison, WI, US;  
Hey Timothy D., Zionsville, IN, US;  
Merlo Donald J., Carmel, IN, US;  
Orr Gregory L., Indianapolis, IN, US;  
Roberts Jean L., Arcadia, IN, US;  
Strickland James A., Lebanon, IN, US;  
Guo Lining, Brownsburg, IN, US;  
Ciche Todd A., Madison, WI, US;  
Sukhapinda Kitisri, Zionsville, IN, US;

(74) Zástupca: Patentservis Bratislava, a. s., Bratislava, SK;

(54) Názov prihlášky vynálezu: **Insekticídne proteínové toxíny z *Photorhabdus***

(57) Anotácia:  
Sú opísané proteíny izolované z baktérií rodu *Photorhabdus*, ktoré sú toxické pre hmyz. *Photorhabdus luminescens* (pôvodne *Xenorhabdus luminescens*) sa objavil v cicavčích klinických vzorkách ako bakteriálny symbiont enteropatogénových nematód rodu *Heterohaditis*. Tieto proteínové toxíny sa môžu geneticky spracovať a aplikovať do potravy hmyzích lariev alebo do rastlín na ničenie škodlivého hmyzu.

## Insekticídne proteínové toxíny z *Photorhabdus*

### Oblasť techniky

Vynález sa týka toxínov, ktoré sa izolovali z baktérií, a použitia uvedených toxínov ako insekticídov.

### Doterajší stav techniky

Mnoho druhov hmyzu sa považuje za škodcov v domácnostiach, pre výletníkov, záhradníkov, hospodárov a ďalších, ktorí investujú do poľnohospodárskych produktov, pričom sú tieto často zničené, alebo sú aspoň znížené výnosy ako výsledok škodcovstva hmyzu na žatevných poliach. Zvlášť v oblastiach, kde je vegetačná sezóna krátka, môže zjavné škodcovstvo hmyzu znamenať dramatické zníženie výnosov na poli a stratu celého zisku pestovateľov. Vzácne zásoby jednotlivých poľnohospodárskych produktov vedú k stále vyšším kalkuláciám výrobcov potravín a tiež konečných spotrebiteľov rastlinných potravín a produktov, ktoré z týchto rastlín pochádzajú.

V prevencii pred ohrozením úrody a obilnín hmyzom a eliminácií obťažovania hmyzími škodcami, sú odkázaní na silné organické pesticídy a insekticídy, ktoré sú známe zreteľnými toxickými účinkami. Tieto synteticky pripravené produkty postihujú celú populáciu, pretože sú príliš drsné na životné prostredie a pôsobia na všetky činitele. Podobne v nepoľnohospodárskych oblastiach by sa mohli osadníci uspokojiť tým, že sa zbavia hmyzu v domovoch alebo vonku, a mohli by jesť bez toho, aby museli zabíjať hmyz.

Extenzívne používanie chemických insekticídov vyzdvihlo enviromentálne a zdravotné záujmy hospodárov, spoločností produkujúcich insekticídy, vládnych agentúr, verejných záujmových

skupín a verejnost' všeobecne. Vývoj menej rušivých stratégií riadenia škodcov je poháňaný jednak spoločenskými podnikmi pre životné prostredie a jednak vývojom biologických prostriedkov, ktoré využívajú mechanizmy riadenia hmyzu. Biologická kontrola účinkov predstavuje sľubnú alternatívu chemickým insekticídmi.

Organizmy na každej úrovni evolučného vývoja majú spôsoby zvýšenia vlastného úspechu a prežitia. Použitie biologických molekúl ako nástrojov ochrany a agresie je známe v ríši živočíchov aj rastlín. Okrem toho relatívne nové nástroje genetického inžinierstva dovoľujú modifikácie biologických insekticídov uskutočňovaných špeciálnymi roztokmi na špeciálne problémy.

Jeden taký činiteľ, *Bacillus thuringiensis* (Bt), patrí k účinným insekticídnym činidlám a ako taký sa komerčne veľmi využíva. V skutočnosti je insekticíd Bt-baktérie proteín, ktorý má tak obmedzenú toxicitu, že sa môže aplikovať na úrodu pre ľudskú výživu v deň žatvy. Pre organizmy, ktoré nie sú cieľové, je Bt toxín stráviteľným netoxickým proteínom.

Iné známe triedy činiteľov biologického ovládania hmyzu sú isté rody hlístovcov, známe ako prenášače parazitov, prostredníctvom dedičnosti bakteriálnych symbiontov zabíjajúcich hmyz. Hlístovce obsahujúce insekticídne baktérie napádajú larvy hmyzu. Baktérie potom zabíjajú larvy. Hlístovce sa množia v uhynutých larvách. Potomstvo takto vyliahnúch hlístovcov, obsahujúcich baktérie, môže potom napádať ďalšie larvy.

V minulosti boli insekticídne hlístovce rodov *Steinernema* a *Heterorhabditis* používané ako činitele na ovládanie hmyzu. Zrejme každý rod hlístovcov je hostiteľom špeciálnych druhov baktérií. *Photorhabdus luminescens* je baktéria, ktorá symbioticky žije v hlístovcoch rodu *Heterorhabditis*.

Aj keď tieto hlístovce sú účinnými insekticídnymi činiteľmi,

je súčasne ťažké, drahé a neefektívne produkovať, udržiavať a distribuovať hlístovce na reguláciu hmyzu.

Je známa technika, ktorou je možné izolovať insekticídny toxín z *Photorhabdus luminescens*. Jeho účinnosť sa prejaví len vtedy, keď sú vstreknuté do lariev hmyzu rádiv *Lepidoptera* a *Coleoptera*. Toto spôsobovalo nemožnosť efektívneho využitia insekticídnych vlastností hlístovcov alebo ich bakteriálnych symbiontov. Mohlo by to však byť dôležité a praktickejšie s menším laboratórnym úsilím pre široké spektrum metód aplikácie insekticídneho toxínu, ktorý by si mohol udržať svoje biologické vlastnosti aj po dodaní. Bolo by žiadúce vyvíjať toxíny s orálnou aktivitou produkovanou rodom *Photorhabdus*. Izolácia a použitie týchto toxínov je žiadúce z hľadiska svojich účinkov. Predtým ako sa objavili výskumy účinkov, tieto toxíny neboli izolované alebo charakterizované.

#### Podstata vynálezu

Prírodné toxíny sú komplexy proteínov, ktoré sú produkované a vylučované rastúcimi bakteriálnymi bunkami rodu *Photorhabdus*. Predmetom záujmu sú proteíny produkované druhmi *Photorhabdus luminescens*. Proteínové komplexy, s veľkosťou molekúl približne 1 000 kDa, môžu byť separované SDS-PAGE gélovou analýzou na množstvo proteínových zložiek. Toxíny neobsahujú hemolyzín, lipázu, typ C fosfolipázu, alebo nukleázu. Toxíny vykazujú významnú toxicitu pri expozícii na množstvo druhov hmyzu.

Predkladaným vynálezom sa pripravuje ľahko využiteľný insekticídny proteín ako aj expresia toxínu do heterológneho systému.

Predkladaný vynález tiež poskytuje metódy podávania insekticídnych toxínov, ktoré sú aktívne funkčné a účinné proti mnohým druhom hmyzu.



Predmet, výhody a črty vynálezu budú zrejmé z nasledujúcej špecifikácie.

Stručný opis obrázkov na výkresoch

Obr.1 znázorňuje zostavu izolátov klonovanej DNA, používanej ako časti génových sekvencií toxínu predkladaného vynálezu.

Obr.2 je mapa troch plazmidov používaných v procese sekvencovania.

Obr.3 je mapa znázorňujúca vzájomné vzťahy niektorých častí fragmentov DNA.

Obr.4 znázorňuje analýzy homológov medzi proteínovými sekvenciami proteínov TcbA<sub>ii</sub> a TcaB<sub>ii</sub>.

Obr.5 je fenogram kmeňov *Photorhabdus*. Vzťahy kmeňov *Photorhabdus* boli definované rep-PCR.

Vrchná os obr. 5 vyjadruje percentuálnu podobnosť kmeňov založenú na skóre produktov rep-PCR (tj.0,0 [žiadna podobnosť] až 1,0 [100 % podobnosť]). Na pravej osi indikujú počty a písmená rôzne testované kmene: 14=W-14, Hm=Hm, H9=H9, 7=WX-7, 1=WX-1, 2=WX-2, 88=HP88, NC-1=NC-1, 4=WX-4, 9=WX-9, 8=WX-8, 10=WX-10, WIR=WIR, 3=WX-3, 11=WX-11, 5=WX-5, 6=WX-6, 12=WX-12, x14=WX-14, 15=WX-15, Hb=Hb, B2=B2, 48 až 52= ATCC 43948 až ATCC 43952. Vertikálne čiary oddelujúce horizontálne čiary indikujú stupeň príbuznosti (ako je vidieť z extrapolovaného priesečníka vertikálnej čiary s hornou osou) medzi kmeňmi alebo skupinami kmeňov na základe horizontálnych čiar (napríklad kmeň W-14 je s kmeňmi H9 a Hm podobný približne na 60 %).

Obr.6 je zobrazenie genómových máp kmeňa W-14.

Obr. 6A je zobrazenie tca a tcb loci a primárnych génových produktov.

Obr. 7 je fenogram kmeňov *Photarhabdus* definovaný pomocou rep-PCR. Vrchná os obr. 7 vyjadruje percentuálnu podobnosť kmeňov založenú na skóre produktov rep-PCR (tj. 0,0 [žiadna podobnosť] až 1,0 [100 % podobnosť]). Na pravej osi indikujú počty a písmená rôzne testované kmene. Vertikálne čiary oddelujúce horizontálne čiary indikujú stupeň príbuznosti (ako je vidieť z extrapolovaného priesečníka vertikálnej čiary s hornou osou) medzi kmeňmi alebo skupinami kmeňov na základe horizontálnych čiar (napríklad kmeň *Indicus* je približne na 30 % podobný kmeňom MP1 a HB Oswego). Na fenograme sú tieto kmene *Photarhabdus*: 14=W-14, Hm=Hm, H9=H9, 7=WX-7, 1=WX-1, 2=WX-2, 88=HP88, NC-1=NC-1, 4=WX-4, 9=WX-9, 8=WX-8, 10=WX-10, WIR=WIR, 3=WX-3, 11=WX-11, 5=WX-5, 6=WX-6, 12=WX-12, x14=WX-14, 15=WX-15, Hb=Hb, B2=B2, 48 =ATCC 43948, 49 = ATCC 43949, 50 = ATCC 43950, 51 = ATCC 43951, 52 = ATCC 43952.

#### Podrobný opis vynálezu

Predkladaný vynález je zameraný na objav špecifickej triedy insekticídnych proteínových toxínov rodu *Photarhabdus*, ktoré majú orálnu toxicitu proti hmyzu. Jedinečnou črtou *Photarhabdus* je jeho bioluminiscencia. *Photarhabdus* sa môže izolovať z rôznych zdrojov. Jedným takým zdrojom sú hlístovce, predovšetkým hlístovce rodu *Heterorhabditis*. Ďalším takým zdrojom sú ľudské klinické vzorky otvorených rán, pozri Farmer a kol., 1989 J.Clin. Microbiol. 27, str. 1594-1600. Tieto saprofytické kmene sú uložené v americkej zbierke kultúr (American Type Culture Collection, Rockville, MD) pod značkou ATCC č. 43948, 43949, 43950, 43951 a 43952 a texte sú uvedené ako odkaz. Je možné, že aj iné zdroje môžu v sebe skrývať baktérie *Photarhabdus*, produkujúce insekticídne toxíny. Tieto zdroje v životnom prostredí môžu byť terestriálne ako aj akvatické.

Rod *Photorhabdus* je taxonomicky definovaný ako člen čelade *Enterobacteriaceae* napriek tomu, že má isté, pre túto čelaď atypické znaky. Kmene tohoto rodu sú napríklad neschopné odbúravať dusičnany, produkujú žltý a červený pigment a sú schopné bioluminiscencie. Tento posledne menovaný znak je inak v rámci čelade *Enterobacteriaceae* neznámy. Ako samostatný rod z rodu *Xenorhabdus* bol *Photorhabdus* opísaný len nedávno (Boemare a kol., 1993, Int. J. Syst. Bacteriol., 43, 249-255). Toto rozdelenie je založené na štúdiách DNA-DNA hybridizácie, fenotypických rozdielov (napríklad prítomnosť (*Photorhabdus*) alebo neprítomnosť (*Xenorhabdus*) katalázy a bioluminiscencie) a hostiteľskou čeladou hlístovca (*Xenorhabdus*; *Steinernematidae*; *Photorhabdus*; *Heterorhabditidae*). Porovnávanie analýzy bunkových mastných kyselín podporuje toto oddelenie rodu *Photorhabdus* od *Xenorhabdus* (Janse a kol., 1990, Lett. Appl. Microbiol 10, 131-135, Suzuki a kol., 1990, J. Gen. Appl. Microbiol., 36, 393-401).

Preto, aby sa určilo, že zbierka kmeňov zahrnutých v tejto práci obsahovala kmene *Photorhabdus*, sa vlastnosti kmeňov charakterizovali na základe znakov, ktoré definujú *Photorhabdus* a ktoré ho oddeľujú ho od iných druhov čelade *Enterobacteriaceae* a od rodu *Xenorhabdus* (Farmer, 1984,, Bergey's Manual of Systemic Bacteriology, Vol. 1, str. 510 - 511; Akhurst a Boemare, 1988, J. Gen. Microbiol., 134, str. 1835 - 1845; Boemare a kol., 1993, Int. J. Sust. Bacteriol., 43, str. 249 - 255, ktoré sú sem včlenené formou odkazov). Študované znaky boli nasledovné: Grammovo farbenie negatívnych tyčiniek, veľkosť organizmu, pigmentácia kolónie, telové inklúzie, prítomnosť kataláz, schopnosť redukovať dusičnany, bioluminiscencia, absorbcia farbiva, hydrolýza želatiny, rast na vybraných médiách, teplota rastu, prežívanie v anaerobných podmienkach a pohyblivosť. Na potvrdenie toho, že všetky tieto kmene patria k samostatnému rodu *Photorhabdus* sa použili analýzy mastných kyselín.

V súčasnosti bakteriálny rod *Photorhabdus* zahrňa samostatne

definované druhy, *Photorhabdus luminescens* (ATCC typový kmeň č. 29 999, Poinar a kol, 1977, *Nematologica* 23, 97 - 102). V literatúre bolo opísaných veľa príbuzných kmeňov (napríklad Akhust a kol, 1988, *J. Gen. Microbiol.*, 134, 1835-1845; Boemare a kol., 1993, *Int. J. Syst. Bacteriol.*, 43, str. 249 - 255; Putz a kol, 1990, *Appl. Environ. Microbiol.*, 56, 181 - 186). Bol tu opísaný veľký počet kmeňov *Photorhabdus*. Tieto kmene sú spísané v tabuľke 18 v príkladoch. Pretože v súčasnosti je v rámci rodu *Photorhabdus* definovaný len 1 druh (*luminescens*), luminiscenčné druhové znaky sa použili na charakteristiku jednotlivých kmeňov. Ako je možné vidieť na obr. 5., tieto kmene sú celkom odlišné. Preto v budúcnosti nebude nečakané, ak budú definované ďalšie druhy *Photorhabdus*, ktoré budú mať niektoré vlastnosti druhu *luminescens*, ale aj niektoré iné vlastnosti, ktoré nie sú v súčasnosti definované ako znak *Photorhabdus luminescens*. Ale do rozsahu tohto vynálezu je zaradený ktorýkoľvek z druhov alebo kmeňov *Photorhabdus*, produkujúci proteíny, ktoré sa vyznačujú funkčnou aktivitou, ako prostriedky regulujúce hmyz. Ostatné znaky a vlastnosti sú bezvýznamné.

Okrem toho, ako sa tu prezentuje, baktérie rodu *Photorhabdus* produkujú proteíny, ktoré majú funkčnú aktivitu takú, ako sa tu definuje. Zvlášť zaujímavé sú proteíny produkované druhom *Photorhabdus luminescens*. V rámci tohto vynálezu sa neobmedzuje len na použitie kmeňov, ktoré sú tu opísané. Tieto kmene po prvýkrát ukazujú, že proteíny produkované rôznymi izolátmi *Photorhabdus* sú toxické pri expozícii na hmyz. V rámci vynálezu sú špecifikované kmene, z toho niektoré mutanty, a tiež niektoré kmene alebo druhy rodu *Photorhabdus*, u ktorých je opísaná funkčná aktivita.

Je niekoľko termínov, ktoré sa tu použili a majú zvláštny význam a sú nasledovné:

Pod "funkčnou aktivitou" sa rozumie, že funkcia proteínových toxínov ako prostriedkov na reguláciu hmyzu je v tom, že proteíny

sú orálne aktívne, majú toxické účinky, alebo sú schopné narušiť alebo zabrániť kŕmeniu, čo má alebo nemá za následok úhyn hmyzu. Keď hmyz príde do kontaktu s účinným množstvom toxínu cestou expresnej transgénnej rastliny, formulovaným zložením proteínu (-inov), rozstrekovateľnou (-ými) zložkou (-ami) proteínu, návnadovým alebo iným spôsobom podania, typickými výsledkami sú úhyn hmyzu alebo hmyz nie je schopný požírať zdroj, ktorý robí toxín dostupným pre hmyz.

Termínom "genetický materiál" sa rozumie zahrnutie všetkých génov, nukleovej kyseliny DNA a RNA.

"Homológom" sa rozumie sekvencia aminokyseliny, ktorá je identifikovaná ako vlastná homológia k referenčnej W-14 aminokyselinovej toxínovej polypeptidovej sekvencii.

"Homológiou" sa rozumie aminokyselinová sekvencia, ktorá má index podobnosti aspoň 33 % a/alebo index identity aspoň 26 % k referenčnej W-14 aminokyselinovej toxínovej polypeptidovej sekvencii ako je to vyhodnotené GAP algoritmom s použitím B10sum 62 proteínovej vyhodnocovacej matrice (Wisconsin Package Version 9.0, Genetics Computer Group (GCC), Madison, WI).

"Identitou" sa rozumie aminokyselinová sekvencia, ktorá obsahuje identické zvyšky na danej pozícii, po porovnaní s referenčnou W-14 aminokyselinovou toxínovou polypeptidovou sekvenciou GAP algoritmom.

Pod termínom "oligonukleotidy" sa rozumie makromolekulové zloženie krátkych reťazcov nukleotidov buď RNA alebo DNA. Takáto dĺžka by mohla byť najmenej jeden nukleotid, ale typicky je v rozpätí od okolo 10 do asi 12 nukleotidov. Určenie dĺžky oligonukleotidov je známe v rámci v laboratórnej techniky a nebolo by tu žiadne obmedzenie. Oligonukleotidy však môžu byť menšie ako 10 alebo väčšie ako 12.

Pod termínom "toxín *Photorhabdus*" sa rozumie každý proteín produkovaný kmeňom mikroorganizmu *Photorhabdus*, ktorý je funkčne aktívny proti hmyzu, pričom sa toxín *Photorhabdus* môže vyrobiť vo forme spraya, expresiou na transgénne rastliny, návnady, s cestou prenosu prostredníctvom baculovírusu, alebo prenosom prostredníctvom iných aplikovateľných hostiteľov alebo prenosového systému.

Termíny "toxický" alebo "toxicita" sa tu používajú v zmysle toxínov produkovaných *Photorhabdus*, ktoré majú "funkčnú aktivitu" takú, aká sa tu definuje.

"Skráteným peptidom" sa tu rozumie každý peptid, ktorý je fragmentom(ami) sledovaných peptidov, ktoré majú funkčnú aktivitu.

"Hlavnou sekvenciálnou homológiou" sa rozumie buď: DNA fragment, ktorý má nukleotidovú sekvenciu dostatočne podobnú inému DNA fragmentu na produkciu proteínu, ktorý má podobné biochemické vlastnosti; alebo polypeptid, ktorý má aminokyselinové sekvencie dostatočne podobné inému polypeptidu na vykazanie podobných biochemických vlastností.

Fermentačné bujóny vybraných kmeňov, ktorých zoznam je v tabuľke 20 sa použili na stanovenie nasledovného: rozsah produkcie insekticídneho toxínu rodu *Photorhabdus*, insekticídneho spektra týchto toxínov a stanovenia zdroja suroviny na purifikáciu toxínových komplexov. Kmene, ktoré sa tu opisujú ukazujú na to, že sú orálne toxické proti množstvu čeladi hmyzu. Tieto čelade hmyzu zahrňujú, ale neobmedzujú sa len na tieto čelade: *Coleoptera*, *Homoptera*, *Lepidoptera*, *Diptera*, *Acarina*, *Hymenoptera* a *Dictyoptera*.

Tak ako aj u iných bakteriálnych toxínov, počet mutácií baktérií v populácii zapríčiňuje, že mnohé príbuzné toxíny sú trochu rozdielne v existujúcich sekvenciách. Toxíny, ktoré sú

predmetom záujmu, sú tie, ktoré produkujú komplexy proteínov toxické na množstvo druhov hmyzu expozíciou, ako je tu opísané. Prednostne sú toxíny aktívne proti: *Lepidoptera*, *Coleoptera*, *Homoptera*, *Diptera*, *Hymenoptera*, *Dictyoptera* a *Acarina*. Vynálezy majú v úmysle zachytiť proteínové toxíny, ktoré sú homologické proteínovými toxínmi, produkovanými opisovanými kmeňmi a všetkými z nich pochádzajúcimi kmeňmi, ako aj všetky proteínové toxíny produkované rodom *Photorhabdus*. Tieto homologické proteíny môžu mať rozdielne sekvencie, ale nie rozdielnu funkciu od toxínov, ktoré sa tu opisujú. Homologickými toxínmi sa myslia proteínové komplexy zahŕňajúce 300 kDa až 2000 kDa a obsahujúce najmenej dve (2) podjednotky, kde podjednotkou je peptid, ktorý môže alebo nemusí byť rovnaký ako ďalšia podjednotka. Identifikujú sa rôzne proteínové podjednotky a sú tu uvádzané ako príklady. Typickými sú podjednotky proteínov medzi okolo 18 kDa až okolo 230 kDa, medzi okolo 160 kDa až okolo 230 kDa, 100 kDa až 160 kDa, okolo 80 kDa až asi 100 kDa, a okolo 50 kDa až asi 80 kDa.

Ako je povedané vyššie, niektoré kmene *Photorhabdus* sa môžu izolovať z hlístovcov. Niektoré hlístovce, predĺžené cylindrické červy kmeňa *Nematoda*, majú vyvinutú schopnosť využiť larvy hmyzu ako vhodné životné prostredie pre svoj rast. Larvy hmyzu poskytujú zdroj potravy pre rastúce hlístovce a aj prostredie, v ktorom sa množia. Dramatický účinok, nasledujúci po invázii istých hlístovcov do lariev je smrť lariev. Smrť lariev je výsledkom výskytu baktérií v niektorých hlístovcoch, ktoré produkujú insekticídny toxín, zastavujúci rast lariev a inhibujúci aktivitu požierania.

Je zaujímavé, že každý rod hlístovca, ktorý je parazitom hmyzu, obsahuje zvláštne druhy baktérií, ktoré sú jedinečne adaptované na symbiotický rast s týmito hlístovcami. Od obdobia, kedy bol tento výskum iniciovaný sa názov bakteriálneho rodu *Xenorhabdus* reklasifikoval do *Xenorhabdus* a *Photorhabdus*. Baktérie rodu *Photorhabdus* sú charakterizované ako symbionty hlístovcov *Heterorhabditus*, a druhy *Xenorhabdus* sú symbiotanmi

druhov *Steinernema*. Táto zmena nomenklatúry sa v tomto opise zohľadňuje, ale žiadnym spôsobom táto zmena v nomenklatúre nemeší rozsah vynálezu, ktorý sa tu opisuje.

Peptidy a gény, ktoré sú tu určené, sú pomenované v súhlase so súčasnými smernicami publikovanými v "Inštrukciách pre autorov" v *Journal of Bacteriology* str. i-xii (Jan.1996), ktoré sú tu začlenené prostredníctvom odkazov. Nasledujúce peptidy a gény sa izolovali z kmeňa *Photorhabdus* W-14.



Tabuľka 1

Peptidová/génová nomenklatúra toxínového komplexu

1 Názov peptidu	2 Peptid Sekvencia ID. Čís.*	3 Názov génu	4 Gén Sekvencia ID. Čís.*
<u>tca genómová oblasť</u>			
TcaA	34 <sup>C</sup>	<i>tcaA</i>	33
TcaA <sub>i</sub>	pro-peptid	<i>tcaA</i>	-
TcaA <sub>ii</sub>	[15] <sup>a</sup> , 34 <sup>C</sup>	<i>tcaA</i>	-
TcaA <sub>iii</sub>	[4] <sup>a</sup> , 35 <sup>C</sup>	<i>tcaA</i>	-
TcaA <sub>iv</sub>	[62] <sup>a</sup>	<i>tcaA</i>	-
TcaB	[3] <sup>a</sup> , (19, 20) <sup>b</sup> , 26 <sup>C</sup>	<i>tcaB</i>	25
TcaB <sub>i</sub>	[3] <sup>a</sup> , (19, 20) <sup>b</sup> , 28 <sup>C</sup>	<i>tcaB</i>	27
TcaB <sub>ii</sub>	[5] <sup>a</sup> , 30 <sup>C</sup>	<i>tcaB</i>	29
TcaC	[2] <sup>a</sup> , 32 <sup>C</sup>	<i>tcaC</i>	31
<u>tcb genómová oblasť</u>			
TcbA	12 <sup>C</sup> , [16] <sup>a</sup> , (21, 22, 23, 24) <sup>b</sup>	<i>tcbA</i>	11
TcbA <sub>i</sub>	pro-peptid	<i>tcbA</i>	-
TcbA <sub>ii</sub>	[1] <sup>a</sup> , (21, 22, 23, 24) <sup>b</sup> , 53 <sup>C</sup>	<i>tcbA</i>	52
TcbA <sub>iii</sub>	[40] <sup>a</sup> , 55 <sup>C</sup>	<i>tcbA</i>	54
<u>tcc genómová oblasť</u>			
Tcca	[8] <sup>a</sup> , 57 <sup>C</sup>	<i>tcca</i>	56
Tccb	[7] <sup>a</sup> , 59 <sup>C</sup>	<i>tccb</i>	58
Tccc	61 <sup>C</sup>	<i>tccc</i>	60
<u>tcd genómová oblasť</u>			
TcdA	(17, 18, 37, 38, 39, 42, 43) <sup>b</sup> , 47 <sup>C</sup>	<i>tcdA</i>	(36) <sup>d</sup> , 46
TcdA <sub>i</sub>	pro-peptid	<i>tcdA</i>	-
TcdA <sub>ii</sub>	[13] <sup>a</sup> , (17, 18, 37, 38, 39) <sup>b</sup> , 49 <sup>C</sup>	<i>tcdA</i>	48
TcdA <sub>iii</sub>	[41] <sup>a</sup> , (42, 43) <sup>b</sup> , 51 <sup>C</sup>	<i>tcdA</i>	50
TcdB	[14] <sup>a</sup>	<i>tcdB</i>	-

<sup>a</sup> Sekvencia ID ČÍS. v hranatých zátvorkách sú N-zakončené peptidy

<sup>b</sup> Čísla v guľatých zátvorkách sú N-zakončenia vnútorných peptidových triptických fragmentov

<sup>c</sup> odvodený od génovej sekvencie

<sup>d</sup> vnútorný génový fragment

Vyššie uvedené sekvencie sú zoskupené podľa genómových oblastí. Presnejšie, baktéria *Photobacterium luminescens* (W-14) má najmenej štyri rozdielne genómové oblasti - *tca*, *tcb*, *tcc* a *tcd*. Ako je vidieť v tabuľke 1, peptidové produkty sú produkované z týchto rozdielnych genómových oblastí. Ďalej, ako je uvedené v príkladoch, najmä v príkladoch 15 a 21 sú jednotlivé génové produkty produkované z troch oblastí spojené s aktivitou hmyzu. Je tiež značná homológia medzi týmito štyrmi genómovými oblasťami.

Ako je uvedené ďalej v príkladoch, *tcbA* gén bol expresovaný v *E. coli* ako dva možné biologicky aktívne proteínové fragmenty (*TcbA<sub>i</sub>* a *TcbA<sub>ii/iii</sub>*). Gén *tcdA* bol tiež expresovaný v *E. coli*. Ako je uvedené v príklade 16, kde bol natívny nespracovaný *TcbA* toxín spracovaný s endogénnou metaloproteázou alebo s obsahom hmyzieho čreva, ktorý obsahoval proteázy, *TcbA* proteínový toxín bol spracovaný do malých podjednotiek, ktoré boli menšie ako je veľkosť natívnych peptidov, pričom vzrástla aktivita liskavky. Najmenšie toxínové peptidy zostali asociované ako časť toxínového komplexu. Je vhodné v niektorých situáciách zvýšiť aktiváciu toxínu(ov) proteolytickým spracovaním alebo použitím skrátených peptidov. Takto je vhodnejšie použiť v niektorých aplikáciách skrátený(é) peptid(y), napríklad v komerčných transgénových rastlinných aplikáciách.

Navyše ku kmeňu W-14, sú iné druhy v rode *Photobacterium*, ktoré majú rôznu funkčnú aktivitu (podrobnejšie pozri tabuľky 20 a 36). Aj keď je tam rôzna aktivita, v niektorých prípadoch majú aminokyselinové sekvencie hlavnú sekvenciálnu homológiu. Okrem toho molekulové sondy ukazujú, že niektoré gény obsiahnuté v kmeňoch sú homologické s génmi obsiahnutými v kmeni W-14. Skutočne všetky kmene tu uvedené majú jednu alebo viac homológií toxínových génov W-14. Údaje o protilátkach v príklade 26 a údaje o N-koncových sekvenciách v príklade 25 ďalej podporujú záver, že existuje homológia a identita (založená na aminokyselinovej sekvencii) medzi proteínovým(i) toxínom(mi), ktoré sú produkované

týmito kmeňmi. Na molekulovej úrovni sondy génu W-14 ukázali, že homológy alebo samotné W-14 gény (tabuľky 37, 38 a 39) sú rozptýlené v rode *Photorhabdus*. Ďalej je možné, že existujú nové toxínové gény v iných kmeňoch, ktoré nie sú homologické s W-14, ale udržiavajú celkové proteínové atribúty (pozri špeciálne príklady 14 a 25).

Aj keď je homológia alebo identita medzi toxínovými génmi produkovanými kmeňmi *Photorhabdus*, samotné kmene sú celkom odlišné. Použitím technológie polymerázovej reťazovej reakcie ďalej diskutovanej v príklade 22, väčšina tu uvedených kmeňov je celkom odlišiteľná. Napríklad ako je vidieť na obr. 5, bolo percento relatívnej podobnosti niektorých reťazcov, napríklad HP88 a NC-1, okolo 0,1, čo ukazuje podstatnú rozdielnosť. Preto aj keď hmyzie toxínové gény alebo génové produkty, ktoré sú produkované kmeňmi, sú tie isté alebo podobné, samotné kmene sú rozdielne.

Vzhľadom na údaje tu ďalej diskutované v príkladoch a v diskusii je zrejmé, že bola objavená nová a jednotná čelaď hmyzieho(ích) proteínového(ých) toxínu(ov). Ďalej bolo ukázané, že tento toxín(y) široko existuje v bakteriálnom rode *Photorhabdus*. Taktiež sa môže vyskytovať prípad, že tieto toxínové gény široko existujú v čeladi *Enterobacteraceae*. Protilátky pripravené, ako je opísané v príklade 21 alebo génové sondy pripravené ako je opísané v príklade 25, sa môžu použiť na ďalšie triedenie pre bakteriálne kmene v čeladi *Enterobacteraceae* na produkciu homologického(ých) toxínu(ov), ktorý má funkčnú aktivitu. Môže tiež nastať prípad, že existujú špecifické sady primérov, ktoré by uľahčili identifikáciu nových génov v rode *Photorhabdus* alebo v čeladi *Enterobacteraceae*.

Ako je uvedené vyššie, protilátky sa môžu použiť na rýchle triedenie baktérií druhu *Photorhabdus* alebo čelade *Enterobacteraceae* na homologické toxínové produkty ako je uvedené v príklade 26. Toto sa v stave techniky celkom podobá na použitie

analytického alebo triediaceho nástroja (pozri americký patent č. 5,430,137, ktorý je tu uvedený ako odkaz). Navyše sa v literatúre všeobecne akceptuje, že protilátky sú použité proti 6 až 20 segmentom aminokyselinového zvyšku, ktorý má tendenciu na zaberanie vystaveného povrchu polypeptidov (Current Protocols in Immunology, Coligan a kol., National Institute of Health, John Wiley & Sons, Inc.). Zvyčajne aminokyselina pozostáva z blízkych aminokyselinových zvyškov, avšak v určitých prípadoch môžu byť tvorené nepríbuznými aminokyselinami, ktoré sú viazané špecifickou konformáciou. Aminokyselinové segmenty prijaté protilátkami sú vysoko špecifické a zvyčajne poukazujú na epitopy. Aminokyselinové fragmenty môžu byť generované chemickým a/alebo enzýmovým štiepením natívneho proteínu, automatizovanou syntézou peptidu v tuhej fáze alebo produkciou z organizmov produkovaných genetickým inžinierstvom. Polypeptidové fragmenty sa môžu izolovať rôznymi metódami a/alebo kombináciou HPLC a FPLC chromatografických metód, ktoré sú známe v stave techniky. Výber polypeptidového fragmentu sa môže uľahčiť použitím algoritmov, napríklad Kyte a Doolittle, 1982, Journal of Molecular Biology 157: 105-132 a Chou a Fasman, 1974, Biochemistry 13: 222-245, ktoré predpovedajú tieto sekvencie najpravdepodobnejšie k vystaveným na povrchu proteínu. Na prípravu imunogénu s obsahom polypeptidového fragmentu, o ktorý je záujem, vo všeobecnosti sú polypeptidy kovalentne spojené použitím chemických reakcií na nosné proteíny tak, ako je kľúčovou dierkou viazaný hemocyanín prostredníctvom voľných amínových (lyzín), sulfydylových (cysteín), fenolových (tyrozín) alebo karboxylových (aspartát alebo glutamát) skupín. Imunogén sa s adjuvans injektuje do zvierat akými sú myši alebo králiky, alebo kurence na vyvolanie imunitnej odpovede proti imunogénu. Analýza protilátkového titra v antisére injektovaných živočíchov proti polypeptidovému fragmentu sa môže stanoviť množstvom imunologických metód ako je ELISA a Western blot. Alternatívne sa môžu monoklonálne protilátky pripraviť použitím buniek sleziny injektovaného zvierata na fúziu s bunkami tumoru na produkciu umrtných hybridomových buniek produkujúcich jednotlivé druhy protilátok.

Hybridómové bunky sa triedia s použitím imunologických metód na oddelenie línií, ktoré produkujú špecifickú protilátku na vybraný polypeptidový fragment. Purifikácia protilátok z rôznych zdrojov sa môže uskutočniť pomocou kolón antigénovej afinity alebo protilátkovej afinity alebo chromatografickou metódou HPLC alebo FPLC.

Tu opísané toxíny sú celkom jedinečné v tom, že toxíny majú funkčnú aktivitu, ktorá je kľúčom k objaveniu stratégie hmyzieho správania sa. V objavovaní stratégie riadenia hmyzieho správania sa, je možné pozdržať alebo zmariť proces degradácie proteínu vnesením proteínu priamo do organizmu s vylúčením tráviaceho traktu. V takých prípadoch sa uplatňovanie proteínu v organizme udrží až do jeho denaturácie, nešpecifickej degradácie alebo eliminácie imunitným systémom vo vyšších organizmoch. Vnesenie insekticídneho toxínu do hmyzu je možné aplikovať len v laboratóriu a len na veľký hmyz, ktorý sa ľahko injektuje. Pozorovanie, že insekticídne proteínové toxíny, ktoré sa tu opisujú, preukazujú toxickú aktivitu po orálnom užití alebo po kontakte s toxínmi, pripúšťa vývoj plánovania hmyzieho správania sa, ktorý je založený výhradne na schopnosti začlenenia proteínových toxínov do diéty hmyzu. Výsledkom takého plánu môže byť produkcia hmyzích návnad.

Toxíny *Photobacterium* môžu pôsobiť na hmyz v purifikovanej forme. Toxíny sa tiež môžu prenášať v množstve od asi 100 mg/l bujónu. Toto sa mení podľa znalosti pomerov zloženia, podmienok zdroja inokula, techník izolácie toxínu a podobne. Toxíny môžu byť podané ako výpotky sekrécie alebo bunkové proteíny skutočne sa nachádzajúce v heterológnych prokaryotických alebo eukaryotických hostiteľoch. Baktérie sú typické pre hostiteľov, v ktorých proteíny vznikajú. Eukaryotickí hostitelia by mohli zahŕňať, ale neobmedzujú sa len na rastliny, hmyz a kvasinky. Alteratívne sa toxíny môžu produkovať v baktériách alebo transgénových rastlinách na poli alebo v hmyze baculovírusovým vektorom. Typické je, že toxíny sú vnesené do hmyzu zavedením

jedného alebo viacerých toxínov do krmiva hmyzu.

Úplná letalita požierania hmyzu je potrebná, ale nie je vyžadovaná preto, aby sa dosiahla potrebná toxicita. Ak sa hmyz vyhne toxínu alebo prestane požierať, je toto vyhýbanie sa v rámci niektorých aplikácií potrebné, aj keď sú účinky subletálne. Napríklad, ak je žiadaná odolnosť hmyzu voči transgénovým žatevným rastlinám, je neochota na hmyz požierať rastliny potrebnějšía, ako letátna toxicita hmyzu, pretože konečným cieľom je ochrana rastlín nie zabíjanie hmyzu.

Je veľa ďalších ciest, ako môžu byť toxíny včlenené do diéty hmyzu. Ako príklad je možné falšovať zdroje krmiva lariev s toxickým proteínom, rozprašovaním krmiva s roztokom proteínu tak, ako je tu opísané. Alternatívne sa môže purifikovaný proteín pomocou genetického inžinierstva fúzovať do inej neškodnej baktérie, ktorá by potom mohla rásť v kultúre a buď sa aplikovať do potravinového zdroja alebo by mohla pretrvávajúť v pôde, v oblasti, v ktorej bolo žiadúce hmyz vyhubiť. Proteín by mohol byť tiež pomocou genetického inžinierstva včlenený do potravinového zdroja hmyzu. Napríklad, najdôležitejším zdrojom mnohých lariev hmyzu je rastlinný materiál.

Zavedením genetického materiálu, ktorý kóduje insekticídne vlastností toxínov *Photorhabdus* do genómu rastlín, ktoré slúžia ako potrava pre špeciálnych hmyzích škodcov, dospelé jedince alebo larvy by uhynuli po konzumácii krmnej rastliny. Veľký počet zástupcov rodov *Monocotyledonus* a *Dictyledonus* sa transformovalo. Transgénoví poľnohospodárski škodcovia, tak ovocia ako aj zeleniny, sú predmetom komerčného záujmu. Títo škodcovia v sebe zahŕňajú škodcov kukurice, ryže, sóje, canola, slnečnice lucerny, pšenice, bavlny, podzemnice olejnej, paradajok, zemiakov a pod. Existuje niekoľko techník na zavedenie cudzieho genetického materiálu do rastlinných buniek preto, aby sa získali rastliny, ktoré sú stabilné a expresujú zavedené gény. Takéto techniky zahŕňajú akceleráciu genetického materiálu obaleného

v mikročasticiach priamo do buniek (U.S. Patenty 4 945 050 Cornell a 5 141 131 DowElanco). Rastliny sa môžu transformovať s použitím technológie na *Agrobaktérium*, pozri U.S. Patent 5 177 010 Univerzity Toledo, US Patent 5 104 310 od Texas A&M, európska patentová prihláška 013162B1, európske patentové prihlášky 120516, 159418B1 a 176 112 Schilperootu, americké patenty č. 5 149 645, 5 469 976, 5 464 763 a 4 940 838 a 4 693 976 od Schilperoot, európske patentové prihlášky 116718, 290799, 320500, všetky od MaxPlanck, európske patentové prihlášky 604662 a 627752 od Japan Tobacco, európske patentové prihlášky 0267159 a 0292435 a americký patent 5 231 019, všetky od Ciba Geigy, americké patenty 5 463 174 a 4 762 785 od Calgen a americké patenty 5 004 863 a 5 159 135 od Agracetus. Ďalšia transformačná technológia zahŕňa technológie, pozri americké patenty 5 302 523 a 5 464 765 obidve od Zeneca. Tiež sa používa technológia elektroporácie transformácií rastlín, pozri WO 87/06614 od Boyce Thompsonovom inštitútu, 5 472 869 a 5 384 253 obidve od Dekalba, WO9209696 a WO9321335 obidve od PGS. Všetky tieto transformačné patenty a publikácie sú zahrnuté v odkazoch. Okrem mnohých technologických postupov transformácie rastlín, sa môže meniť druh tkaniva, ktorý sa fúzuje s cudzími génmi. Také tkanivá, ako embryonálne tkanivo, kalusové tkanivo typu I a II, hypokotyl, meristém a pod., sú vhodné, ale nie sú obmedzené len na tieto druhy. Skoro všetky rastlinné tkanivá sa môžu transformovať pomocou dediferenciácie vhodnou technikou kvalifikovaným zručným odborníkom.

Inou premenou je voľba selektovateľného markera. Preferencia zvláštneho markera je na základe konania odborníka, ale hocijaké ďalšie selektovateľné merkery sa môžu použiť s ktorýmikoľvek génmi, ktoré sme tu neopisovali a ktoré by mohli mať funkciu ako selektovateľné markery. Takéto selektovateľné markery, obsahujú aminoglykozid fosfotransferázový gén transpónu Tn 5 (Aph II), ktorý kóduje odolnosť na antibiotiká, kanamycín, neomycín a G 418, ako aj také gény, ktoré kódujú odolnosť alebo toleranciu glyfozátu, hygromycínu, metotrexátu, fosfinotricínu (bialofoz),

imidazolínov, sulfonylureáz a triazolpyrimidínových herbicídov, ako je chlorosulfurón, bromoxynil, dalapon a pod., nie sú však na ne obmedzené.

Okrem selektovateľných markerov sa môže vyžadovať použitie reportérnych génov. V niektorých prípadoch reportérne gény sa môžu použiť bez selektovateľných markerov. Reportérne gény sú gény, ktoré nie sú vždy prítomné alebo expresované v prijímacom organizme alebo tkanive. Reportérne gény typicky kódujú proteíny, ktoré spôsobujú niektorú fenotypickú zmenu alebo enzymatickú vlastnosť. Príklady takýchto génov sú uvedené v K.Weising a kol. Ann. Rev. Genetics 22, 421, (1988), ktoré sú tu uvedené formou odkazu. Uprednostňovaný reporterový gén je glukuronidázový gén (GUS).

Bez ohľadu na transformačné techniky, je gén prednostne zavedený do génu transformačným vektorom, adaptovaným na expresiu toxínov *Photorhabdus* do rastlinných buniek zavedením vektora rastlinného promótoru. Okrem rastlinných promótorov, sa môžu účinne použiť promótorové zdroje do rastlinných buniek, na prenos cudzích génov. Napríklad sa môžu použiť promótorové vlastné baktériám, také ako oktopínový syntázový promótor, nopalínový syntázový promótor, promótor manopínovej syntázy, promótorové vlastné vírusom, také ako karfiolový mozaikový vírus (35 S a 19 S), upravený 35S, známy ako 35T (pozri PCT/US96/16582, WO 97/13402 zverejnený 17. apríla 1997, ktorý je tu uvedený ako odkaz) a podobne. Rastlinné promótorové v sebe zahŕňajú, ale neobmedzujú sa na ribulózu-1,6-bisfosfát (RUBP), karboxylázovú malú podjednotku (SSV), beta-konglycinínový promótor, fazeolínový promótor, ADH promótor, promótor teplotného šoku a špecifické tkanivové promótorové. Promótorové môžu tiež obsahovať isté zosilňovacie sekvenčné jednotky, ktoré môžu účinnosť transkripcie zlepšiť. Typické zosilňovače obsahujú v sebe, ale neobmedzujú sa na Adh-intron 1 a Adh-intron 6. Môžu sa použiť konštitutívne (základné) promótorové. Konštitutívne promótorové priamo pokračujú v génovej expresii vo všetkých bunkových typoch a v každom čase



(napríklad aktin, ubiquitin, CaMV 35S). Promótoory špecifické pre tkanivá sú zodpovedné za genetickú expresiu v špecifických bunkách alebo typoch pletív, takými ako sú listy alebo semená (napríklad zeín, oleozín, napin, ACP) a tieto promótoory sa môžu tiež použiť. Promótoory môžu byť tiež aktívne v určitej etape vývinu rastlín, a tiež v rastlinných tkanivách a orgánoch. Príklady takýchto promótorov obsahujú, ale nie sú obmedzené na peľovú špecifitu, embryonálnu špecifitu, špecifitu na kvety kukurice, špecifitu na bavlnené vlákno, koreňovú špecifitu, špecifické promótoory pre endosperm semien a iné.

Za istých okolností sa môže vyžadovať použitie indukovateľného promótoru. Indukovateľný promótor zodpovedá za expresiu génov reakciou na špecifický signál ako je: fyzikálna stimulácia (teplotný šok génov), svetlo (RUBP karboxyláza), hormóny (Em), metabolity, a stres. Môžu sa používať aj iné požadované prvky transkripcie a translácie, ktoré sú funkčné v rastlinách. Pre túto techniku je známe množstvo transferových vektorov so špecifitou génov rastlín.

Okrem toho je známe, že na získanie vysokého stupňa expresie bakteriálnych génov v rastlinách sa uprednostňuje opätovná expresia bakteriálnych génov tak, aby boli účinnejšie expresované v cytoplazme rastlín. Kukurica je jedna z takýchto rastlín, ktorá uprednostňuje opätovnú expresiu bakteriálneho génu (-ov) pred transformovaním zvyšovania úrovne expresie toxínu v rastline. Jednou z príčin uprednostňovania opätovnej expresie je veľmi malý obsah G+C v pôvodnom bakteriálnom géne (-och) (a následne neúmerne vysoký obsah A + T). To má za následok, že tvorba sekvencií duplikuje alebo napodobňuje sekvencie regulačných rastlinných génov, ktoré sú známe príliš vysokým obsahom A+T. Prítomnosť niektorých sekvencií bohatých na A+T v DNA génoch fúzovaných do rastlín (napríklad TATA oblasť normálne sa nachádzajúca v génových promótoroch) môže mať za následok aberácie transkribovaného génu (-ov). Na druhej strane prítomnosť iných regulačných sekvencií nachádzajúcich sa v transkribovanej

mRNA (napríklad polyadenylačné signálne sekvencie (AAUAAA) alebo sekvencie komplementárne malým jadrovým RNA vybudovaným do zostrihu pre mRNA) môže viesť ku nestabilite RNA. Z tohto dôvodu je jedným z cieľov v návrhu opätovnej expresie bakteriálneho génu (-ov), ktorá je preferovanejšia pred vylepšeným génom (-mi) rastliny, je vývoj sekvencie DNA, ktorá má vyšší obsah G+C a prednostne ku génom rastliny, ktoré kódujú metabolické enzýmy. Ďalším cieľom v projekte vylepšenia génu (-ov) rastliny (-ov) je vyvinúť takú sekvenciu DNA, ktorá nemá len vyšší obsah G+C, ale aj modifikovanie zmien sekvencií, ktoré by nerobili ťažkosti v translácii.

Príklad rastliny, ktorá má vysoký obsah G+C je kukurica. V nižšie uvedenej tabuľke je znázornené, aký vysoký je obsah G+C v kukurici. Predpokladá sa, že obsah G+C v iných rastlinách je tiež vysoký, podobne ako u kukurice.

Tabuľka 2

Zhrnutie obsahu G+C v proteínoch kódujúcich oblasti génov kukurice

Trieda proteínov <sup>a</sup>	Rozsah % G+C	Priemer % G+C <sup>b</sup>
Metabolické enzýmy (40)	44,4-75,3	59,0 (8,0)
Zásobné proteíny		
Skupina I. (23)	46.0-51,9	48,1 (1,3)
Skupina II. (13)	60.4-74,3	67,5 (3,2)
Skupina I+II. (36)	46.0-74,3	55,1 (9,6) <sup>c</sup>
Štruktúrne proteíny (18)	48.6-70,5	63,6 (6,7)
Regulačné proteíny (5)	57.2-68,9	62,0 (4,9)
Necharakterizované proteíny (9)	41,5-70,3	64,3 (7,2)
Všetky proteíny (108)	44.4-75,3	60,8 (5,2)

- a počet génov v triede daných v zátvorkách
- b štandardné odchýlky dané v zátvorkách
- c priemer kombinovaných skupín, zanedbávaný v kalkulácii celkového priemeru

Údaje o kódovaných oblastiach génov v tabuľke 2 boli vybraté z položiek GenBank (povolenie 71) a základné zloženie bolo vypočítané pomocou programu MacVector<sup>TM</sup> (B, New Haven, CT). V týchto kalkuláciách sa sekvencie intrónov zanedbávali. Skupiny I a II génovej sekvencie zásobných proteínov boli významné z hľadiska zreteľnej rozdielnosti v zložení báz.

Prispôsobivosť násobená nadbytočnosťou genetickým kódom (napríklad niektoré aminokyseliny sú určované viac ako jedným kodónom), evolúcia genómov rozdielnych organizmov alebo tried má za následok rôzne využitia redundantných kodónov. Toto "uprednostňovanie kodónu" sa odráža v priemernom zložení báz oblastí kódujúcich proteíny. Napríklad organizmy s relatívne nízkym obsahom G+C využívajú kodóny obsahujúce A alebo T v tretej pozícii redundantných kodónov, zatiaľ čo tie ktoré majú vyšší obsah G+C využívajú kodóny nesúce G alebo C na tretej pozícii. Týmto sa myslí, že prítomnosť "menej frekventovaných" kodónov v génoch mRNA môže znížiť absolútny počet translácií tejto mRNA, zvlášť, keď relatívna adundancia kontrolnej tRNA, s menej frekventovaným kodónom je nízka. Rozsah znižovania počtu translácií jednotlivých menej frekventovaných kodónov by bol prinajmenšom aditívny k rozmanitým menej frekventovaným kodónom. Preto by mRNA, ktoré majú vysoké relatívne obsahy menej frekventovaných kodónov mali zodpovedať malému počtu translácií. Tento počet by bol zohľadnený pri syntéze nízkych počtov kódovaných proteínov.

Na to, aby sa opätovne fúzoval bakteriálny gén (-y) je determinované uprednostňovanie kodónov rastlín. Uprednostňovanie kodónov je štatisticky vyjadrená distribúcia kodónov, ktoré rastlina využije pre kódovanie svojich proteínov. Následne po

určení uprednostňovania sa určuje percentuálna frekvencia kodónov v géne (-och), ktoré sú predmetom záujmu. Primárne kodóny, ktoré sú uprednostňované rastlinami by sa určovali práve tak, ako aj sekundárne a terciárne vhodné preferované kodóny. Sekvencia aminokyselín proteínu, ktorá je predmetom záujmu, je spätne translatovaná tak, že výsledná sekvencia kódov nukleových kyselín pre rovnaké proteíny je taká, akú má pôvodný bakteriálny gén. Ale táto výsledná sekvencia nukleových kyselín súhlasí s prvými preferovanými kodónmi v žiadanej rastline. Nová sekvencia sa analyzuje kvôli polohe reštrikčných enzýmov preto, aby sa mohla vyvolať modifikácia. Identifikované miesta sa ďalej modifikujú presunom kodónov so sekundárnym alebo terciárnym zložením preferovaných kodónov. Iné miesta v sekvencii génov, ktoré by sa mohli použiť na transláciu alebo transkripciu predmetných génov, sú exon: intron 5' alebo 3' konce, poly A adičné signály alebo RNA polymerizačné terminačné signály. Sekvencia sa ďalej analyzuje a modifikuje tak, aby sa redukovala frekvencia TA alebo GC dvojíc. Okrem dvojíc, sekvenčných blokov G alebo C, môže mať viac ako asi štyri zvyšky, ktoré tak isto môžu spôsobiť transkripciu sekvencií. Preto, tieto bloky taktiež modifikujú substitúciou kodónov prvého alebo sekundárneho výberu atď. s ďalšou preferenciou najvhodnejšieho kodónu. Táto preferencia je preto, že vylepšený gén(-y) rastliny obsahuje (-ú) okolo 63 % prvých najvhodnejších kodónov, medzi okolo 22 % až do asi 37 % kodónov sekundárneho výberu a medzi 15 % až 0 % tretieho výberu kodónov tvoriacich spolu 100 %. Najviac preferovaný vylepšený gén (-y) rastliny obsahuje okolo 63 % kodónov prvého výberu, najmenej asi 22 % kodónov druhého výberu, okolo 7,5 % kodónov tretieho výberu a okolo 7,5 % kodónov štvrtého výberu, pričom celkové percento je 100 %. Vyššie opísaná metóda dáva možnosť odborníkovi modifikovať gén (-y), ktoré sú jednotlivcej rastline cudzie preto, aby sa gény optimálne fúzovali do rastlín. Spôsob sa ďalej opisuje v prihláške PCT/US96/16582, WO 97/13402 zverejnenej 17. apríla 1997.

Na zostavenie vylepšeného rastlinného génu (ov) sa

aminokyselinové sekvencie toxínov upravovanej rastliny reverzne translatovali do sekvencie DNA, s použitím neredundantného genetického kódu, ktorý vznikol z tabuľky kodónových uprednostnení zostavenej pre sekvenciu génov DNA rastliny, ktorá sa má transformovať. Výsledná sekvencia DNA, ktorej použité kodóny sú úplne homogénne, je ďalej modifikovaná tak, aby vznikla taká sekvencia DNA, ktorá okrem toho, aby mala vyšší stupeň diverzity kodónov obsahovala tiež strategicky umiestnené reštrikčné enzýmy na určených miestach, žiadúce zloženie báz, a aby neobsahovala sekvencie, ktoré by mohli interferovať s transkripciou génov, prípadne transláciou produktov mRNA.

Len teoreticky by bakteriálne gény mohli podliehať ľahšie expresii v rastlinách vtedy, keď sú bakteriálne gény expresované do plastidov. Takto by bolo možné expresovať bakteriálne gény do rastliny bez vylepšenia génov pre expresiu do rastlín a získal by sa vysoký výťažok proteínu. Pozri americké patenty 4 762 785, 5 451 513 a 5 545 817, ktoré sú tu včlenené formou odkazu.

Čo sa týka komerčného využívania transgenických rastlín, jedným z problémov je odpor manažmentu. Toto sa zvlášť týka toxínov *Bacillus thuringiensis*. Veľa spoločností komerčne *Bacillus thuringiensis* využíva a tam sú stále vo veľkej neistote vznikajúcej z odporu okolo Bt toxínov. Jednou zo stratégií manažmentu odolnosti hmyzu by mohla byť kombinácia toxínov produkovaných *Photobacterium* s takými toxínmi ako Bt, vegetatívne hmyzie proteíny (Ciba Geigy) alebo s inými toxínami. Kombinácie by sa mohli vytvoriť na aplikáciu rozprašovaním alebo by mohli byť molekulárnymi kombináciami. Rastliny by sa mohli transformovať pomocou tých génov *Photobacterium*, ktoré produkujú hmyzie toxíny a ďalšie gény hmyzích toxínov také ako Bt a aj iné gény hmyzích toxínov, také ako Bt.

Európska patentová prihláška 0400246A1 opisuje transformáciu 2 Bt do rastliny, ktorá by mohla byť aj z 2 génov. Iným spôsobom produkcie transgenickej rastliny obsahujúcej viac ako jeden gén

hmyzej odolnosti, by bola produkcia dvoch rastlín, pričom by každá rastlina obsahovala gén hmyzej odolnosti. Tieto rastliny by sa použili na spätné kríženia tradičnými technikami oploďňovania rastlín na produkciu rastliny obsahujúcej viac ako jeden gén hmyzej odolnosti.

Okrem produkcie transformovanej rastliny obsahujúcej vylepšený rastlinný gén, sú ďalšie spôsoby dodávania, ktoré by mohli byť žiadúce pri novokonštruovaní bakteriálneho génu (-ov). Spolu s touto líniou, genetickým inžinierstvom ľahko izolovateľného proteínového toxínu fúzaného spolu s molekulami zaujímavými pre hmyz - ako zdroja potravy, by sa mohla insekticídna aktivita toxínu konštruovať a expresovať do baktérie alebo do eukaryotických buniek za použitia štandardných, dobre známych techník. Po purifikácii v laboratóriu takýto toxický činiteľ so "zabudovaním" návnady by mohol byť zabalený vnútri v štandardných hmyzích pasciach.

Ďalší spôsob dodávania je zavedenie genetického materiálu toxínov do baculovírusového vektora. Baculovírusy infektujú jedine svojich hmyzích hostiteľov, vrátane svojho žiadúceho pôsobenia s toxínmi *Photorhabdus*. Infekcie baculovírusov sú vhodným prostredím expresie toxínov *Photorhabdus*, ktoré by mohli byť vnesené do oblastí zamorenia hmyzom, čím by tieto intoxikovali alebo jedom infikovali hmyz.

Prenos insekticídnych vlastností vyžaduje sekvencie nukleových kyselín kódujúce aminokyselinové sekvencie toxínu *Photorhabdus* zabudované do proteínového expresného vektora vhodného hostiteľa, v ktorom vektor zostane. Jeden spôsob ako získať sekvenciu nukleových kyselín kódujúcu proteín s insekticídnyimi vlastnosťami, je izolovať pôvodný genetický materiál, ktorý produkuje toxíny z *Photorhabdus*, s použitím informácie odvodzovanej zo sekvencie aminokyselín toxínu veľkých jednotiek, ktoré sú v príkladoch nižšie. Ako je nižšie opísané, metódy purifikácie proteínov zodpovedných za aktivitu toxínov sú

tiež vyriešené.

Použitím údajov N-koncovej sekvencie aminokyselín, tak ako sa uvádza v príkladoch nižšie, môžu sa skonštruovať oleonukleotidy komplementárne k celým alebo k častiam z DNA báz, ktoré kódujú prvé aminokyseliny toxínu. Tieto oligonukleotidy môžu byť radioaktívne značené a použité ako molekulárne sondy na izoláciu genetického materiálu z genómovej genetickej knižnice, ktorá sa zostavila z genetického materiálu izolovaného z kmeňov *Photorhabdus*. Genetická knižnica sa môže klonovať do plazmidov, kozmidov, fágov alebo fágových vektorov. Knižnica by mohla byť transformovaná do *Escherichia coli* a vytriedená pre produkciu toxínu prostredníctvom transformovaných buniek s použitím protilátky kultivovanej oproti toxínu alebo priamych skúšok na hmyziu toxicitu.

Takýto prístup si vyžaduje produkciu batérie oligonukleotidov, pretože degenerácia genetického kódu umožňuje kódovanie aminokyseliny v DNA ktorýmkoľvek z niektorých tripletových kombinácií nukleotidov. Aminokyselina arginín môže byť napríklad kódovaná tripletmi nukleových kyselín CGA, CGC, CGG, CGT, AGA a AGG. Nedá sa predpovedať, ktorý triplet sa použije v danej pozícii v géne toxínu a musia sa pripraviť oligonukleotidy so všetkými potenciálnymi tripletmi predstavujúcimi danú aminokyselinu. Je asi potrebná viac ako jedna DNA molekula, ktorá sa zhoduje s proteínovou subjednotkou, na konštrukciu dostatočného počtu oligonukleotidových sond, ktoré zahŕňajú všetky proteínové subjednotky potrebné na to, aby sa dosiahla orálna toxicita.

Zo sekvencie aminokyselín purifikovaného proteínu sa môže spoľahlivý genetický materiál na produkciu toxínov ľahko izolovať a klonovať vcelku alebo v časti do expresného vektora niekoľkými dobre známymi technikami používanými odborníkmi v molekulárnej biológii. Typickým expresným vektorom je plazmid DNA. Predpokladá a aj iný spôsob prenosu ako zmienený, ktorý má význam, ale

neobmedzuje sa len na kozmidy, fagemidy a fágy, ktoré sa tiež predvídajú. Okrem žiadúcich alebo vyžadovaných vlastností pre replikáciu plazmidu, takými ako sú pôvodná replikácia a odolnosť voči antibiotikám alebo iná forma selektovateľného markera, ako sú *bar* gén *Streptomyces hygrosopicus* alebo *viridochromogenes*, vyžaduje si proteínová expresia vektorov aj prídavnú expresnú kazetu, ktorá včlení cis - aktívne sekvencie potrebné na transkripciu a transláciu génu, ktorý je predmetom záujmu. Cis - aktívne sekvencie, ktoré sa vyžadujú na expresiu v prokaryotoch, sa líšia od tých, ktoré sa vyžadujú pre eukaryoty a rastliny.

Eukaryotická expresná kazeta vyžaduje transkripčný promótor ležiaci proti smeru transkripcie (5') predmetného génu, transkripčnú terminačnú oblasť ako je poly A-aditívne miesto a väzbové miesto na ribozóme proti smeru transkripcie od časti prvého kodónu génu, ktorý je predmetom záujmu. V bakteriálnych bunkách potrebným transkripčným promótorom, ktorý môže byť fúzovaný do vektora, je T7 väzbový promótor RNA polymerázy. Promótory, ako už bolo opísané, sú známe tým, že účinne podporujú transkripciu mRNA. Tiež proti smeru transkripcie od predmetného génu môže vektor vnieť nukleotidové sekvencie kódujúce signálne sekvencie, o ktorých sa vie, že priamo kovalentne viažu proteín do zvláštnych úsekov hostiteľských buniek, takých ako je bunkový povrch.

Vírusy hmyzu alebo baculovírusy sú schopné infikovať a nepriaznivo vplývať na istý hmyz. Tento vplyv vírusov na ne je pomalý a vírusy nezastavia požieraciu aktivitu hmyzu. A tak sa na vírusy nehľadí ako na dôležité insekticidné činitele kontroly škodcov. Fúziou génov toxínov *Photographus* do baculovírusových vektorov by sa mohol uskutočniť účinný spôsob transmitácie toxínov aj v prípade rastu letality vírusov. Pretože rozdielne baculovírusy sú špecifické pre rôznych hmyz, je možné použiť vybraný toxín ako vybraný terč pre zvlášť nebezpečných hmyzích škodcov. Zvlášť potrebným vektorom pre gény toxínov je jadrový



polyhedrálny vírus. Použitie tohto vírusu ako transférových vektorov je opísané a teraz je možné zvoliť vektor na transfer cudzích génov do hmyzu. Vírus-toxínový génový rekombinant sa môže skonštruovať do podoby prenosnej orálne. Baculovírusy obyčajne napadajú hmyzie obete cez strednú vnútornú intestinálnu sliznicu. Gén toxínu inzertovaného za silný plášť vírusového proteínového promotora by sa expresoval a rýchlo zabíjal infektovaný hmyz.

Okrem toho hmyzí vírus, baculovírus alebo transgénová rastlina, ako forma podania proteínových toxínov tohto vynálezu, môžu tiež proteíny opuzdriť použitím technológie *Bacillus thuringiensis*, ktorá je uvedená, ale nie je jediná v amerických patentoch 4 695 455; 4 695 462; 4 861 595, ktoré sa sem včlenili formou odkazu. Iný spôsob prenosu proteínových toxínov v rámci tohto vynálezu, je vloženie proteínu do matrice návnady, ktorá by sa potom mohla použiť nad aj pod základným stanoviskom hmyzích návnad. Príkladom takejto technológie je patentová prihláška PCT WO 93/23998, ktorá však neznamena jediný spôsob, a včlenila sa sem formou odkazu.

Tak ako sa opísalo vyššie, môže byť potrebné modifikovať sekvenciu kódovania proteínu, keď expresia do neživého hostiteľa môže byť odlišná vzhľadom na kódovú preferenciu iných hostiteľov od tých, ktoré platia pre *Photobacterium*. V tomto prípade by mohla byť translácia do iného hostiteľa celkom neúčinná pokiaľ sa neuskutoční kompenzačná modifikácia na pripravenú kódovú sekvenciu. Modifikácia sekvencie aminokyselín by sa mohla vyžadovať preto, aby sa zabránilo inhibičnej krížovej reakcii s proteínmi nového hostiteľa alebo rafinovaním insekticídnych vlastností proteínov v novom hostiteľovi. Geneticky modifikovaný toxín by mohol kódovať prejav toxínu, napríklad zvyšovaním alebo znižovaním toxicity, meniacej sa odolnosti hmyzu v ontogenetickom vývine, zmenami stability alebo modifikáciou záberu druhovej špecifity.

Úmyslom je, v rozsahu prezentovaného vynálezu, pre gény

kódujúce toxíny *Photorhabdus* obsiahnuť sekvencie nukleových kyselín, ktoré kódujú biopolyméry aminokyselín homologické ku proteínovým toxínom, a ktoré si zachovávajú toxické účinky proteínov *Photorhabdus* v jednotlivých druhoch hmyzu po orálnom podaní.

Napríklad sa zdá, že toxíny využívané v tomto vynáleze, ako prvé inhibujú larválnu požíraciu aktivitu pred nastávajúcim exitom. Manipuláciou sekvencie nukleových kyselín toxínov *Photorhabdus* alebo ich regulačných sekvencií, sa pomocou genetického inžinierstva fúzuje toxínový gén do rastliny čo umožní ovplyvňovať jeho činnosť alebo jeho spôsob účinku, napríklad na zachovávanie inhibičnej aktivity požírania, pričom sa eliminuje absolútna toxicita na larvy. Táto zmena by mohla umožniť transformovanej rastline prežiť až do žatvy bez nepotrebných dramatických následkov na zničenom ekosystéme všetkých druhov hmyzu, na ktorý je toxín určený. Do rozsahu vynálezu patria všetky takéto modifikácie génov kódujúcich toxín alebo proteín kódovaný génom.

Ďalšie predvídané modifikácie nukleových kyselín v sebe obsahujú adíciu sekvencií, ktoré sú predmetom záujmu, upriamujúcich toxín do určitých častí lariev hmyzu na to, aby sa zvýšila jeho účinnosť.

Kmeň W-14, ATCC č. 55397 bol uložený v americkej zbierke kultúr - American Type Culture Collection, 12301 Parklawn Drive, Rockville, MD 20852 USA 5. marca 1993. Kmene ATCC č. 43948, 43949, 43950, 43951, 43952 sa získali z americkej zbierky kultúr - American Type Culture Collection, 12301 Parklawn Drive, Rockville, MD 20852 USA a boli znovu uložené v americkej zbierke kultúr - American Type Culture Collection, 12301 Parklawn Drive, Rockville, MD 20852 USA 5. novembra 1996 a druhé uloženie je identifikované ako ATCC č. 55878, 55879, 55880, 55881, a 55882. Údaje o aminokyselinách a sekvencii nukleotidov pôvodného toxínu kmeňa W 14 (ATCC 55397) sú uvádzané nižšie. Izolácia genómovej

DNA pre toxíny z bakteriálnych hostiteľov sa tu tiež vysvetľuje. Ďalšie tu identifikované kmene sú uložené v United States Department of Agriculture, 1815 North University Drive, Peoria, IL 61604. Údaje o aminokyselinách a sekvencii nukleotidov pôvodného toxínu kmeňa W 14 (ATCC 55397) sú uvádzané nižšie. Izolácia genómovej DNA pre toxíny z bakteriálnych hostiteľov sa tu tiež vysvetľuje. Ďalšie tu identifikované kmene sú uložené v United States Department of Agriculture, 1815 North University Drive, Peoria, IL 61604.

V tomto opise špecifikácii sa využívali a skúšali techniky štandardnej a molekulárnej biológie. Ďalšie informácie je možné získať z knihy, ktorá sa sem včlenila ako odkaz: Sambrook, J., Fritsch, E.F., a Maniatis, T. (1989), *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Press, *Current Protocols in Molecular Biology*, F.M. Ausubel a kol., (1997), ktoré sú tu obidve vo forme odkazu.

V príkladoch sa používajú nasledujúce skratky:

Tris = Tris (hydroxymetyl) amínometán;  
SDS = dodecilsíran sodný ;  
EDTA = kyselina etyléndiamíntetraoctová;  
IPTG = izopropyltio-B-galaktozid;  
X-gal = 5-bróm-4-chlór-3-indoyl-B-D-galaktozid;  
CTAB = bromid cetyltrimetylamónny;  
kbp = kilobázové páry;  
dATP, dCTP, dGTP, dTTP, I = 2'-deoxinukleozid-5-trifosfáty adenínu, cytozínu, guanínu, tymínu a prípadne inozínu;  
ATP = adenzín 5'-trifosfát.

### Príklady uskutočnenia vynálezu

#### Príklad 1

Purifikácia toxínu z *P. luminescens* a dôkaz toxicity po orálnom

### dávkovaní purifikovaného toxínu

Insekticídny proteínový toxín vynálezu bol purifikovaný z *P. luminescens* kmeňa W-14, ATCC čísla 55397. Zásobné kultúry *P. luminescens* sa udržiavali na Petriho miskách obsahujúcich 2 % proteósový peptón č. 3 (napríklad PP3, Difco laboratories, Detroit, MI) v 1,5 % agare, inkubované pri 25 °C a preočkovávali sa týždenne. Kolónie pôvodnej formy baktérií sa inokulovali do 200 ml bujónu PP3 doplneného 0,5 % monostearátu sorbitanu polyoxyetylénového (Tween 60, Sigma Chemical Company, St. Louis, MO) v jednom litri. Kultúry rástli v bujóne 72 hodín pri 30 °C na rotačnej trepačke. Toxínové proteíny sa získavali z kultúr rastúcich v prítomnosti alebo neprítomnosti tweenu, aj keď absencia tweenu môže mať vplyv na formu rastu baktérií a profil proteínov, ktoré baktérie produkujú. V neprítomnosti tweenu sa menia rôzne formy do tej miery, že molekulová hmotnosť najmenej jednej identifikovanej subjednotky toxínu sa mení v rozmedzí od 200 kDa až asi 185 kDa.

72 hodinové kultúry sa centrifugovali pri 10 000 x g počas 30 minút preto, aby sa odstránili bunky a poškodené tkanivá. Frakcia supernatantu obsahujúca insekticídne aktívne zložky sa dekantovala a preniesla do 50 mM  $K_2HPO_4$  pridaním príslušného množstva 1,0 M  $K_2HPO_4$ . pH sa upravilo na hodnotu 8,6 prídavkom hydroxidu sodného. Takto upravená supernatantová frakcia sa potom miešala s DEAE-Sephacel živicom (Pharmacia LKB Biotechnology), ktorý sa kondicionoval s 50 mM  $K_2HPO_4$ . Toxická účinná zložka sa adsorbovala na DEAE živicu. Táto zmes sa potom vliala do kolóny 2,6 x 40 cm a premyla 50 mM  $K_2HPO_4$  pri izbovej teplote a pri prietoku 30 ml/hod. až kým efluent nedosiahol stálu hodnotu základnej UV absorbančie pri 280 nm. Kolóna sa potom eluovala 150 mM KCl, až kým efluent opäť nedosiahol stálu hodnotu pri 280 nm. Nakoniec sa kolóna premyla 300 mM KCl a frakcionovanie sa skončilo.

Frakcie obsahujúce toxín sa spájali a sterilizovali na 0,2

mikropórovom membránovom filtri. Toxín sa potom koncentroval a upravil sa 100 mM  $KPO_4$ , s pH = 6,9 a za použitia ultrafiltračnej membrány sa oddelila molekulová hmotnosť 100 kDa pri 4 °C (Centriprep 100, Amicon Division-W.R. Grace a Col). Aplikovali sa 3 ml vzorky toxínového koncentrátu na povrch 2,6 x 95 cm gélovej filtračnej kolóny Sephacryl S-400 HR (Pharmacia LKB Biotechnology). Eluačným pufrom bol 100 mM  $KPO_4$ , s pH = 6,9, ktorý pretekal pri prietoku 17 ml/hod pri 4 °C. Efluent sa monitoroval pri 280 nm.

Frakcie sa zbierali a testoval sa ich toxický účinok. Toxicita chromatografických frakcií sa hodnotila biotestami s použitím lariev *Manduca sexta*. Frakcie sa použili priamo do krmiva hmyzu (mole pšeničnej, živiacej sa klíčkami, ICN Biochemická divízia - ICN Biomedicals, Inc.), alebo sa podali intrahemocelickými injekciami v množstve 5  $\mu$ l vzorky podanej cez prvú panôžku 4. alebo 5. instarového štádia lariev tridsiatkami ihlami. V rámci pokusnej skupiny sa zaznamenávala hmotnosť každej larvy v 24 hodinových intervaloch. Toxicita sa predpokladala vtedy, keď hmyz prestal prijímať potravu a zahynul v priebehu niekoľkých dní následkom konzumácie upravenej hmyzieho krmiva, alebo ak sa úhyn zaznamenal v priebehu 24 hodín po injekčnom podaní frakcie.

Toxické frakcie sa spájali a koncentrovali na Centriprep - 100 a analyzovali sa na HPLC použitím 7,5 mm x 60 cm TSK - GEL G - 4000 SW gélovej prienikovej kolóny so 100 mM fosforečnanu draselného s pH = 6,9 ako eluačného činidla s prietokom 0,4 ml/min. Táto analýza odhalila toxínový proteín, ktorý bol obsiahnutý v jednoduchom ostrom píku, ktorý sa eluoval z kolóny v retenčnom čase približne 33,6 minút. Tento retenčný čas sa zhodoval s odhadovanou molekulovou hmotnosťou 1 000 kDa. Píkové frakcie sa zbierali pre ďalšiu purifikáciu, pričom frakcie neobsahujúce tento proteín sa vyradili. Pík eluovaný na HPLC absorboval UV svetlo pri 218 a 280 nm, ale neabsorboval pri 405 nm. Absorbancia pri 405 nm sa ukázala byť atribútom

xenorabdínových antibiotických zložiek.

Elektroforézou spojených píkových frakcií v nedenaturovanom agarózovom géle (Metaphor Agarose, FMC BioProducts) sa ukázalo, že v piku sú prítomné 2 proteínové komplexy. Píkový materiál, purifikovaný v 50 mM Tris HCl s pH = 7,0 sa oddeľoval na 1,5 % nadstavenom agarózovom géli, ktorý sa pufroval 100 mM Tris HCl s pH 7,0 a 1,9 % rozpusteným agarozovým gélom pufrovaným 200 mM Trisboritanom pri pH = 8,3 za štandardných podmienok pufrovania (anódový pufor 1 M Tris HCl s pH = 8,3 a katódový pufor 0,025 M Tris, 0,192 M glycín). Gély pretekali pri konštantnom prúde 13 mA pri 15 °C až kým fenolová červeň sfarbila tok až do konca gélu. Po ofarbení Coomasiho briliantovou modrou sa ukázali v agarózových géloch 2 proteínové skupiny.

Pomalšia migračná skupina sa zaradila ako "proteínová skupina 1" a rýchlejšia migračná skupina sa zaradila ako "proteínová skupina 2". Dve proteínové skupiny sa vyskytovali približne v rovnakých počtoch. Coomasiho farbou farbené agarózové gély sa použili ako vodítka precízneho odobrania dvoch proteínových skupín nezafarbených častí gélu. Takto odobraté kúsky obsahujúce proteínové skupiny, sa macerovali a pridalo sa k nim malé množstvo sterilnej vody. Podobne kontrola, časť gélu neobsahujúca proteín, sa tiež odobrala a upravila sa rovnakým spôsobom ako kúsky gélu obsahujúce proteíny. Proteín sa získal z gélových kúskov elektroelúciou do 100mM Tris boritanu s pH = 8,3 pri 100 voltoch (konštantná hodnota napätia) počas dvoch hodín. Alternatívne sa proteín eluoval pasívne z kúskov gélu pridaním rovnakého objemu 50 mM Tris HCl pH = 7,0 ku týmto kúskom gélu s následnou 16 hodinovou inkubáciou pri 30 °C. Takto sa proteínom umožnila difúzia z gélu do pufra, ktorý sa zbieral.

Výsledky testov toxicity na hmyze s HPLC purifikovaným toxínom (pík 33,6 min) a toxínom purifikovaným na agarózovom géli dokázali toxicitu extraktov. 1,5 µg injekcie HPLC purifikovaného toxínu zabíjali do 24 hodín. Obe proteínové skupiny 1 a 2 získané

z agarózových gélov pasívnou elúciou alebo elektroelúciou boli letálne po injekčnom podaní. Koncentrácia proteínu, hodnotená v týchto vzorkách bola nižšia ako 50 ng/larva. Porovnanie získanej hmotnosti a mortality medzi skupinami lariev injektovaných proteínovou skupinou 1 alebo 2 indikovalo, že proteínová skupina 1 bola pri podaní injekciou toxickejšia.

Pokiaľ sa do krmiva lariev pridal na HPLC purifikovaný toxín v koncentrácii 7,5 µg/larva zapríčinilo to zastavenie prírastku váhy (24 testovaných lariev). Larvy začnú žrať, ale po konzumácii len veľmi malého množstva toxínom upraveného krmiva sa začali prejavovať patologické symptómy indukované toxínom a larvy prestali potravu prijímať. Výlučky hmyzu sa začali odfarbovať a väčšina lariev mala príznaky hnačky. Významná mortalita hmyzu bola zjavná, keď sa aplikovalo niekoľko 5 µg dávok toxínu do diéty počas 7 až 10 dennej periódy.

Na agarózovom géli separovaná proteínová skupina 1 významne inhibovala prírastok hmotnosti lariev pri dávke 200 ng/larva. Larvy, ktoré sa prikrmovali rovnakými koncentráciami proteínovej skupiny 2 neboli inhibované a prírastok hmotnosti bol rovnaký ako pre kontrolné larvy. Eluovaným proteínom sa krmilo dvanásť lariev a 45 lariev sa krmilo proteínom obsiahnutým v agarózových kúskoch. Tieto dve sady dát indikujú, že proteínová skupina 1 bola orálne toxická na *Manduca sexta*. V tomto experimente sa zdalo, že proteínová skupina 2 nebola toxická na *Manduca sexta*.

Ďalšia analýza proteínových skupín 1 a 2 na SDS-PAGE, v podmienkach denaturácie ukázala, že každá skupina bola zložená z niekoľkých menších proteínových subjednotiek. Proteíny sa objavili po farbení Coomasiho briliantovou modrou, po ktorom nasledovalo farbenie striebrom preto, aby sa dosiahla maximálna citlivosť.

Proteínové podjednotky dvoch skupín boli veľmi podobné. Proteínová skupina 1 obsahuje 8 proteínových subjednotiek 25,1;

56,2 ; 60,8; 65,6; 166; 171; 184 a 208 kDa. Proteínová skupina 2 mala identický profil okrem 25,1 ; 60,8 a 65,6 kDa proteínov, ktoré prítomné neboli. Proteíny 56,2; 60,8; 65,6 a 184 kDa sa nachádzali v komplexe proteínovej skupiny 1 približne v rovnakých koncentráciách a predstavujú 80 % alebo viac z celkového obsahu tohto komplexu.

Pôvodný na HPLC purifikovaný toxín sa ďalej charakterizoval nasledovne. Toxín bol tepelne labilný v tom, že pri dĺžke ohrievania 15 minút pri 60 °C stratil svoju schopnosť zabíjať alebo inhibovať prírastok hmotnosti, pokiaľ bol podaný injekčne alebo bol larvami *M.sexta* požieraný. S purifikovaným toxínom sa robili testy na detekciu lipázy, typu C fosfolipázy, nukleázy alebo hemolytickej aktivity červených krviniek. Žiadna z týchto aktivít prítomná nebola. Robili sa tiež inhibičné testy antibiotickej zóny a purifikovaný toxín nedokázal inhibovať rast gramnegatívnych alebo grampozitívnych baktérií, kvasiniek alebo vláknitých húb. Toto indikovalo, že toxické nie je xenorabdinové antibiotikum.

Pôvodný na HPLC purifikovaný toxín sa testoval aj z hľadiska schopnosti zabíjať hmyz iný ako *Manduca sexta*. V tabuľke 2 je znázornený hmyz zabíjaný purifikovaným toxínom *P.luminescens* toxínom na HPLC v tejto štúdií.



Tabuľka 3

Hmyz usmrcovaný toxínom *P. luminescens*

<u>Bežný názov</u>	<u>Rad</u>	<u>Rod a druh</u>	<u>Cesta podania</u>
	<i>Lepidoptera</i>	<i>Manduca sexta</i>	orálne a injekčne
Múčiar obyčajný	<i>Coleoptera</i>	<i>Tenebrio molitor</i>	orálne
Mravec faraónsky	<i>Hymenoptera</i>	<i>Monomorium pharoanis</i>	orálne
Rus domový	<i>Dictyoptera</i>	<i>Blattella germanica</i>	orálne a injekčne
Komár zimnicový	<i>Diptera</i>	<i>Aedes aegypti</i>	orálne

Ďalšia charakterizácia toxínového komplexu s vysokou molekulovou hmotnosťou.

V ďalších analýzach sa toxínový proteínový komplex podrobil ďalšej charakterizácii z rastového média W-14. Podmienky kultúr a počiatočné kroky purifikácie cez kolónu S-400 HR boli tie isté ako bolo opísané vyššie. Po izolácii toxínového komplexu s vysokou molekulovou hmotnosťou z frakcií kolóny S-400 HR sa toxické frakcie kondicionovali s 10 mM Tris HCl s pH 8,6 a koncentrovali sa v centriplus 100 (Amicon) koncentrátoroch. Proteínový toxínový komplex sa potom aplikoval na slabú anionovú kolónu (WAX), Vydac 301VPH575 (Hesperia, CA), pri rýchlosti 0,5 ml/min. Proteíny sa eluovali s lineárnym gradientom chloridu draselného, 0 až 250 mM KCl, v 10 mM Tris HCl s pH 8,6 počas 50 min. Bolo detegovaných osem proteínových pík absorbancom pri 280 nm.

Bioskúšky s použitím neonátov lariev liskavky uhorkovej

(*Diabrotica undecimpunctata howardi*, SCR) lariev a *Manduca sexta* (THW) sa uskutočnili na všetkých frakciách, ktoré sa eluovali z HPLC kolóny. *Manduca sexta* rástli na potrave mole pšeničnej živiacej sa klíčkami wheat germ (ICN) pri 25 °C v 16 hodinovom svetelnom cykle/8 hodinovom tmavom cykle. SCR rástli na potrave lariev liskavky uhorkovej Insecta-Diet (BioServ) pri 25 °C v 16 hodinovom svetelnom cykle/8 hodinovom tmavom cykle.

Najvyššie mortality sa dosiahli u lariev SCR a THW pre pík 6, ktorý bol eluovaný s asi 112 mM až 132 mM KCl. Analýza SDS-PAGE píku 6 ukázala predovšetkým peptidy 170 kDa, 66 kDa, 63 kDa, 59,5 kDa a 31 kDa. Western blot analýza sa uskutočnila na proteínovej frakcii píku 6 so zmesou polyklonálnych protilátok pripravených proti TcaA<sub>ii</sub>-syn, TcaA<sub>iii</sub>-syn, TcaB<sub>ii</sub>-syn, TcaC-syn a TcbA<sub>ii</sub>-syn peptidom (opísané v príklade 21) a C5F2, monoklonálna protilátka proti peptid TcbA<sub>iii</sub>. Pík 6 obsahoval imunoreaktívne pásy 170 kDa, 90 kDa, 66 kDa, 59,5 kDa a 31 kDa. Sú veľmi blízke k predpovedanej veľkosti pre TcaC (166 kDa), TcaA<sub>ii</sub> + TcaA<sub>iii</sub> (92 kDa), TcaA<sub>iii</sub> (66 kDa), TcaB<sub>ii</sub> (60 kDa) a TcaA<sub>ii</sub> (25 kDa). Pík sa potom ďalej analyzoval elektroforézou na prirodzenom agarózovom géli, ako je tu opísané migroval ako jednotlivý pás s mobilitou, ktorá je podobná pásu 1.

Koncentrácia proteínov v purifikovanom toxínovom proteíne píku 6 sa stanovila s použitím BCA reagentov (Pierce). Roztoky sa urobili v 10 mM Tris, pH 8,6 a aplikovali sa na krmivo v bioskúškach. Po 240 hodinách všetky neonátové larvy na potrave bioskúšok, ktoré prijali viac ako 450 ng alebo viac proteínových frakcií píku 6, boli mŕtve. Skupina lariev, ktoré prijali 90 ng tejto istej frakcie mala 40 % mortalitu. Po 240 hod prežité larvy, ktoré prijali 90 ng a 20 ng proteínovej frakcie píku 6, predstavovali 10 % alebo 70 % kontrolnej hmotnosti.

## Príklad 2

### Užitočnosť insekticídu

Funkčnosť a toxicita *Photorhabdus luminescens* sa ďalej charakterizovala. Bujónová kultúra *Photorhabdus luminescens* (kmeň W-14) sa vyrábala nasledovne. Produkčným médiom bol 2 % baktoproteózový peptón číslo 3 (PP3, Difco Laboratories, Detroit, Michigan) v deionizovanej vode Milli-Q. Kultivačné nádoby obsahovali 175 ml média v 500 ml kmeňových nádobách s Delongovým hrdlom, ktoré boli uzatvorené zátkou Kaput a autoklávovali sa 20 minút pri teplote 250 °F. Produkčné nádoby obsahovali 500 ml v 2,8 litrovej 500 ml zásobnej nádobe s úzkym hrdlom uzatvorenej silikónovým penovým uzáverom Shin-etsu. Tieto sa autoklávovali 45 minút pri teplote 250 °F. Inokulačná kultúra sa inkubovala pri 28 °C v inkubátore pri 150 rpm na rotačnej trepačke s dvojpalcovou výchylkou. Po 16 hodinách rastu sa 1 % inokulačnej kultúry inokulovalo do produkčnej nádoby, čo bolo dostatočné pre 24 hodinový rast pred odberom. Produkcia toxínu sa objavila počas log fázy rastu. Mikrobiálny bujón sa preniesol do 1 l centrifugačnej nádoby a bunková biomasa sa centrifugovala (30 minút pri 2500 RPM pri 4 °C [R.C.F.= okolo 1600] na centrifúge Sorval HG-4L rotor RC3, Dupont, Wilmington, Delaware). Primárny bujón sa 8 až 16 hodín chladil a znova centrifugoval najmenej 2 hodiny (za podmienok rovnakých, ako sú uvedené vyššie), aby sa bujón ďalej purifikoval odbúravaním pravdepodobne mukopolysacharidov, ktoré sa počas centrifugácie usadzovali (alternatívna metóda tohto procesu kombinovala obidva kroky a zahŕňala použité 16 hodinovej prejasňovacej centrifugácie pri rovnakých podmienkach ako je uvedené vyššie). Tento bujón sa pred biotestom alebo filtráciou uskladňoval pri teplote 4 °C.

Bujónová kultúra *Photorhabdus* a proteínový toxín (-y), ktorý sa z bujónu purifikoval, prejavili účinnosť (mortalitu a/alebo inhibíciu rastu, redukovali výlet dospelých jedincov) proti množstvu druhov hmyzu. Podrobnejšie sa aktivita pozorovala proti

liskavkám (*Diabrotica*) (larvy a imága), pásavke zemiakovej a *Popilia japonica*, ktoré patria do radu *Coleoptera*. K ďalším zástupcom *Coleoptera* patria kováčiky, blyskáčiky, skočky, seed beetles, a nosáčky. Účinnosť sa tiež pozorovala proti cikádovitým, ktoré patria k zástupcom radu *Homoptera*. Ďalší zástupcovia *Homoptera* sú kobyľky, méry, méra jabloňová, červcovité, molícovité a peniarkovité práve tak ako množstvo druhov vošiek. Bujón a purifikované frakcie sú tiež účinné proti sivkavcom stavikrvovým, moriam kapustovým, siatici ypsilonovej, *Manduca sexta*, vijačkám kukuričným, *Helicoverpa zea* a obaľovačovi jablčnému, ktoré patria k zástupcom radu *Lepidoptera*. Ďalšími typickými zástupcami tohto radu sú moľa šatná, vijačka papriková, obaľovačovité, mlynárik kapustný, *Heliothis zea*, vreckovcovité, priadkovec americký, vijačka voštinová a sivkavec južný. Účinnosť bola tiež zjavná proti ovocným muškám a larvám komárov, ktoré patria do radu *Diptera*. Ďalšími zástupcami radu *Diptera* sú plodomor hrachový, vrtavka mrkvová, kvetárka kapustová, kvetárka reďkovková, kvetárka cibulová, tipuľa bahenná, mucha domáca a rôzne druhy komárov. Účinnosť je aj proti mravcovi horskému, mravcovi argentínskemu, ktoré patria do rádu, ktorý tiež zahŕňa mravca (*Solenopsis xyloni*), oderous house ants a malé čierne mravce.

Bujón/frakcia je užitočná na to, aby sa redukovala populácia hmyzu a aby sa použili v metódach inhibície populácie hmyzu. Metóda môže obsahovať aplikáciu na lokus hmyzu účinne inaktivujúci hmyz, z množstva opísaných druhov, na ktoré je účinný. Výsledky sú uvedené v tabuľke 4.

Účinnosť proti larvám liskaviek sa testovala nasledovne. Bujónová kultúra *Photobacterium* (sterilizovaného filtráciou, bez buniek) alebo frakcie purifikované na HPLC sa aplikovali priamo na povrch (okolo 1,5 cm<sup>2</sup>) 0,25 ml umelého krmiva v 30 µl podieloch s nasledujúcim zriedením v kontrolnom médiu alebo 10 mM pufri fosforečnanu sodného s pH = 7,0. Misky s krmivom sa vysušili vzduchom v sterilnej prúdiacej atmosfére a jamky sa

infestovali s jednotlivými neonátmi *Diabrotica undecimpunctata howardi* (liskavka uhorková, SCR) vyliahnutými zo sterilizovaných vajíčok s druhým instarovým štádiom SCR, ktoré rástli na umelo pripravenej diéte alebo s druhým instarovým štádiom *Diabrotica virgifera virgifera* (Liskavka sp., WCR) pestovaných na obilných klíčkoch rastúcich na Metromixe. Druhé instarové štádiá lariev sa vážili predtým, ako sa nasadila diéta. Misky sa uzatvorili, vložili do vlhkého rastového kultivátora a udržiavali sa pri 27 °C primeraný čas (4 dni pre neonáty a imága SCR, 2 až 5 dni pre larvy WCR, 7 až 14 dní pre druhé instarové štádium SCR). Mortalita a určenie hmotnosti sa odčítavalo tak, ako sa uviedlo. V tejto štúdií sa použilo vo všeobecnosti 16 jedincov hmyzu na prípravok. Mortality v kontrole boli nasledovné: neonátne larvy < 5 %, dospelých chrobákov 5 %.

Účinnosť proti pásavke zemiakovej sa testovala nasledovne. Bujónová kultúra *Photorhabdus* alebo kontrolné médium sa aplikovalo na povrch (okolo 2,0 cm<sup>2</sup>) 1,5 ml štandardnej umelej výživy, ktorá sa udržiavala v jamkách na 24 jamkovej platni pre tkanivové kultúry. Do každej jamky sa dalo 50 µl prípravku a nasledovalo vysušenie vzduchom. Jednotlivé druhé instarové štádiá lariev pásavky zemiakovej (*Leptinotarsa decemlineata*, CPB) sa potom vložili do tejto diéty a mortalita sa odčítala po 4 dňoch. Na jeden prípravok sa vo všetkých štúdiách použilo 10 lariev. Kontrolná mortalita bola 3,3 %.

Účinnosť proti skarabeusovitým larvám a chrobákom sa testovala nasledovne. Jedince *Popillia japonica* (druhé - tretie instarové štádium) sa zozbierali z infestovaných tráv a udržiavali sa v laboratóriu v pôdno rašelinovej zmesi s kúskami mrkvi, ktoré sa pridávali ako dodatočná výživa. Jedince *Popillia japonica* boli lokálne odchyťované do feromónovej pasce a udržiavali sa v laboratóriu v plastických nádobách s javorovými listami ako potravou. Použil sa nezriedený bujón *Photorhabdus* alebo kontrolné médium na umelo pripravenú diétu liskaviek (30 µl/1,54 cm<sup>2</sup> na chrobáka) alebo kúsky mrkvy (larvy).

Potom sa obidve štádiá samostatne vložili do jamiek s výživou a sledovala sa mortalita a príjem potravy. V obidvoch prípadoch sa pozorovala jasná redukcia v množstve krmenia (a produkcii fekálií).

Účinnosť proti larvám komárov sa testovala nasledovne. Test sa robil na 96-jamkovej mikrotitračnej platni. Každá jamka obsahovala 200  $\mu$ l vodného roztoku (bujónovej kultúry *Photorhabdus*, kontrolného média alebo vody) a približne 20 jeden deň starých lariev (*Aedes aegypti*). 6 jamiek slúžilo na prípravok. Výsledky sa odčítavali po dvoch hodinách po infestácii škodcami a nezmenilo sa počas troch dní pozorovania. V kontrole sa nezistila žiadna mortalita.

Účinnosť proti ovocným muškám sa testovala nasledovne. Zakúpené médium pre *Drosophila melanogaster* sa pripravilo z 50 % suchého média a 50 % tekutiny a to do kontrolného média alebo bujónovej kultúry *Photorhabdus*. Médium sa doprpravovalo prídavkom 8,0 ml sušeného média do každej z 3 kultivačných vialiek pre 1 prípravok a pridaním 8,0 ml príslušnej tekutiny. Neskoré instarové štádiá lariev *Drosophila melanogaster* sa potom umiestnili do každej vialky. Vialky sa inkubovali na laboratórnom stole pri izbovej teplote pod fluorescencnými stropnými svetlami. Počty kukiel alebo dospelých jedincov sa rátali po 3, 7 a 10 dňoch expozície. Zavedenie bujónovej kultúry *Photorhabdus* do živného média pre larvy ovocných mušiek zapríčinilo nepatrnú (17 %), ale významnú redukciu v 10. dni liahnutia dospelých jedincov, v porovnaní s vodou a kontrolným médiom (3 %).

Účinnosť proti cikáde sa testovala nasledovne. Test prijímania potravy cikády (*Macrosteles severini*) sa venoval umožneniu aktívneho príjmu potravy bez ďalšieho vonkajšieho sprostredkovateľa. Zásobáreň na účinný/"krmny" roztok sa urobila tak, že uprostred spodnej časti Petriho misky rozmerov 35 x 10 mm sa urobili 2 jamky. Cez vrch misky sa položil dvojpalcový štvorec Parafilmu M a zabezpečil sa "O" krúžkom. Jednouncový plastický

pohár sa potom infestoval približne 7 cikádami a rezervoár sa položil na vrch pohára parafilmom smerom dole. Testovací roztok sa potom pridal do rezervoára cez jamky. V testoch používaná neriedená bujónová kultúra *Photorhabdus*, bujón a kontrolné médium sa dialyzovali na vodu, aby sa zredukovala kontrolná mortalita. Mortalita sa odčítala v dni 2, kedy sa zistila 26,5 % kontrolná mortalita. V testoch využívajúcich purifikované frakcie (200 mg proteínu/ml) sa použila konečná koncentrácia 5 % sacharózy vo všetkých riedeniach na zlepšenie prežívania cikád. Test sa udržoval v inkubátore pri 28 °C, 70 % relatívnej vlhkosti s fotoperiódou 16/8. Test sa klasifikoval z hľadiska mortality po 72 hodinách. Kontrolná mortalita bola 5,5 %.

Účinnosť proti mravcovi argentínskemu sa testovala nasledovne. 1,5 ml zo 100 % bujónovej kultúry *Photorhabdus*, kontrolného média, alebo vody sa pipetovalo do 2,0 ml purifikovaných sklenených vialiek. Vialky sa uzatvorili kúskom bavlnenej dentálnej nite, ktorá sa navlhčila vhodným prípravkom. Každá vialka sa položila jednotlivo do Petriho misiek rozmerov 60 x 16mm s 8 až 12 dospelých jedincov mravca argentínskeho (*Linepithema humile*). Pre jednu vzorku sa použili tri paralelky. Testovacie platne sa udržovali pri izbovej teplote pod ultrafialovým žiarením na laboratórnom stole. Mortalita sa odčítavala po 5 dňoch expozície. Kontrolná mortalita bola 24 %.

Účinnosť proti mravcovi horskému sa testovala nasledovne. Robotníci čiernych mravcov horských (*Camponotus pennsylvanicus*) sa získali zo stromov DowElancovho majetku v Indianopolise, IN. Testy s tekutou kultúrou *Photorhabdus* sa uskutočnil nasledovne. Do každej plastikovej testovacej nádoby (7 1/8" x 3") sa dalo pätnásť robotníkov, papierové útočiská a 10 ml bujónu alebo kontrolného média do plastikových pohárov. Na vzorku ku mravcom sa dalo bavlnené vlákno cez jamku vo viečku pohára. Všetky prípravky obsahovali 5 % sacharózy. Testy sa udržiavali v tme pri izbovej teplote a hodnotili po 19 dňoch. Kontrolná mortalita bola 9 %. Testy, ktoré obsahovali purifikované frakcie, využívali

umelú mravčiu diétu zmiešanú so vzorkou (purifikovaná frakcia, alebo kontrolný roztok) v množstve 0,2 ml vzorky/2,0 g potravy v plastikovej testovacej skúmavke. Konečná koncentrácia proteínu purifikovanej frakcie bola menšia ako 10 µg/g potravy. 10 mravcov vo vzorke, zdroj vody, útočisko a upravená potrava sa umiestnili do uzavretých plastických nádob a udržiavali sa v tme pri 27 °C vo vlhkom inkubátore. Mortalita sa odčítavala v 10. dni. Kontrolná mortalita nebola žiadna.

Účinnosť proti rôznym larvám *Lepidoptera* sa testovala nasledovne. Tekutá kultúra *Photorhabdus* alebo purifikované frakcie sa aplikovali priamo na povrch (okolo 1,5 cm<sup>2</sup>) 0,25 ml štandardnej umelej výživy v 30 µl podieloch za sebou nasledujúcich riedeniach v kontrolnom médiu alebo prípadne v 10 mM pufru fosforečnanu sodného, pH 7,0. Platne s potravou sa vysušili na vzduchu v sterilnom prietočnom boxe a jamky sa infikovali s jednotlivými neonátnymi larvami. Vajíčka vijačky kukuričnej (*Ostrinia nubilalis*) a *Helicoverpa zea* sa získali z komerčných zdrojov a vyliahli sa v laboratóriu, zatiaľ čo larvy sivkovca stavikrvového (*Spodoptera exigua*), mory kapustovej (*Trichoplusia ni*), sivkavca (*Heliothis virescens*), obaľovača jablčného (*Laspeiresia pomonella*) a siaticie ypsilonovej (*Agrotis ipsilon*) sa získavali interne. Potravinové platne sa uzatvorili po infestácii larvami, vložili do vlhkých rastových komôr a udržiavali v tme pri 27 °C vhodne dlhú periódu. Stanovenie mortality a hmotnosti sa odčítalo v 5 až 7 dňoch pre bujónovú kultúru *Photorhabdus* a 4 až 7 dňoch pre purifikovanú frakciu. Vo všeobecnosti sa použilo vo všetkých štúdiách po 16 jedincov hmyzu na vzorku. Kontrolná mortalita sa pohybovala od 4 do 12,5 % pre kontrolné médium a bola menšia ako 10% v prípade fosfátového pufru.



Tabuľka 4

Účinok bujónovej kultúry *Photorhabdus luminescens* (kmeň W-14) a purifikovanej toxínovej frakcie na mortalitu a rast inhibície rôznych čeľadí/druhov hmyzu.

Čeľaď / druh hmyzu	Bujón	purifikovaná frakcia		
	% mort.	% G.I.	% mort.	R.I.
<b>COLEOPTERA</b>				
<i>Diabrotica</i> (liskavka) <i>undecimpunctata howardi</i> /neonátne larvy	100	na	100	na
/ druhé	na	38,5	nt	nt
/imago	45	nt	nt	nt
<i>diabrotica diabrotica</i> / druhé instarové štádium	na	35	nt	nt
pásavka zemiaková druhé instarové štádium	93	nt	nt	nt
<i>Popillia japonica</i> tretie instarové štád. imago	na na	a.f. a.f.	nt nt	nt nt
<b>DIPTERA</b>				
Ovocná muška (výlet imaga)	17 100	nt na	nt nt	nt nt
larvy komárov				
<b>HOMOPTERA</b>				
<i>Macrosteles</i> <i>severini</i>	96,5	na	100	na
<b>HYMENOPTERA</b>				
mravec argentínsky	75	na	nt	na
mravec horský	71	na	100	na
<b>LEPIDOPTERA</b>				
sivkavec stavikrvový	12,5	36	18,75	41,4
siatica ypsilonová	nt	nt	0	71,2
mora kapustová	nt	nt	21,9	66,8
obaľovač jablčný	nt	nt	6,25	45,9
<i>Helicoverpa zea</i>	56,3	94,2	97,9	na
vijačka kukuričná	96,7	98,4	100	na
<i>Heliothis virescens</i>	13,5	52,5	19,4	85,6

mort. = mortalita, G.I. = inhibícia rastu, na = neaplikovateľné,  
nt = netestované, a.f. = neskŕmiteľné

### Príklad 3

#### Užitočnosť insekticídu v aplikácii na pôdu

Bujónová kultúra *Photorhabdus luminescens* (kmeň W-14) sa ukázala byť účinnou proti *Diabrotica* (liskavkám) keď sa aplikovala priamo do pôdy alebo do pôdnej zmesi (Metromix). Účinok proti neonátom SCR a WCR v metromixe sa testovala nasledovne (tabuľka 5). Test prebiehal na sadenicích obilia (United Agriseeds brand, druh CL614), ktoré klíčili na svetle na vlhkom filtračnom papieri za 6 dní. Po tomto čase boli korene dlhé približne 3 až 6 cm a jednotlivé zrná/sadenice sa zasadili do 591 ml priehľadných plastických pohárov s 50 g suchého Metromixu. 20 neonát SCR alebo WCR sa potom umiestnilo priamo na korene klíčkov a zasypali sa Metromixom. Po infestácii boli klíčky zmáčané 50 ml celkového objemu riedeného roztoku bujónu. Po zmáčaní sa nádoby uzatvorili a uložili pri izbovej teplote na svetle počas 7 dní. Potom sa klíčky premyli preto, aby sa odstránil všetok Metromix a korene sa odobrali a zvážili. Účinnosť sa počítala ako percento zvyšku obilneho koreňa vo vzťahu ku kontrolným rastlinám a ako poškodenie listov indukované spásaním. Poškodenie listov sa odčítavalo vizuálne a počítalo sa buď -, +, ++ alebo +++, s tým, že - reprezentuje nulové poškodenie a +++ reprezentuje poškodení niekoľko.

Účinnosť proti neonátom SCR v pôde sa testovala nasledovne (tabuľka 5). Test prebiehal s obilnými klíčkami (United Agriseeds brand, druh CL614), ktoré klíčili na svetle na vlhkom filtračnom papieri 6 dní. Potom boli korene dlhé približne 3 až 6 cm, jednotlivé zrná/klíčky sa zasadili do 591 ml priehľadných plastických pohárov s 150 g pôdy z poľa v Lebanone, IN, osadenej v predchádzajúcom roku obilím. Táto pôda sa predtým neupravovala insekticídami. Potom sa umiestnilo dvadsať neonát SCR priamo na

korene klíčkov a prikryli sa pôdou. Po infektovaní sa klíčky namočili celkovým objemom 50 ml riedeného roztoku bujónu. Po zaliatí sa neuzatvorené poháre inkubovali v prostredí vysokej vlhkosti v miestnosti (80 %) pri 78°F. Potom sa klíčky premyli, aby sa odstránila všetka pôda a korene sa odobrali a vážili. Aktivita sa počítala ako percento zostatku obilných koreňov v porovnaní ku kontrolným rastlinám a ako poškodenie listov indukované požíeranim. Poškodenie listov sa odčítalo vizuálne a počítalo ako -, +, ++ alebo +++, pričom - znamená bez poškodenia a +++ znamená poškodení niekoľko.

Tabuľka 5

Účinok bujónovej kultúry *Photorhabdus luminescens* (kmeň W-14) na larvy *Diabrotica* po post infestačnom zavlažovaní (Metromix).

Prostriedok	Larva	Poškodenie listu	Hmotnosť koreňa (g)	%
<b><i>Diabrotica undecimpunctata howardi</i></b>				
Voda	-	-	0,4916 ± 0,023	100
Médium (2,0 % obj./obj.)	-	-	0,4416 ± 0,029	100
Bujón (6,25 obj./obj.%)	-	-	0,4641 ± 0,081	100
Voda	+	+++	0,1410 ± 0,006	28,7
Médium (2,0 % obj./obj.)	+	+++	0,1345 ± 0,028	30,4
Bujón (1,56 obj./obj. %)	+	-	0,4830 ± 0,031	104
<b><i>Diabrotica virgifera virgifera</i></b>				
Voda	-	-	0,4446 ± 0,019	100
Bujón (2,0 % obj./obj.)	-	-	0,4069 ± 0,026	100
Voda	+	-	0,2202 ± 0,015	49
bujón (2,0 % obj./obj.)	+	-	0,3879 ± 0,013	95

Tabuľka 6

Účinok bujónovej kultúry *Photorhabdus luminescens* (kmeň W-14) na larvy *Diabrotica undecimpunctata howardi* po post infestačnom zavlažovaní (pôda)

Prostriedok	Larva	Poškodenie listu	Hmotnosť koreňa (g)	%
Voda	-	-	0,2148 ± 0,014	100
Bujón (50 % obj./obj.)	-	-	0,2260 ± 0,016	103
Voda	+	+++	0,0916 ± 0,009	43
bujón (50 % obj./obj.)	+	-	0,2428 ± 0,032	113

Pôsobnosť bujónovej kultúry *Photorhabdus luminescens* (kmeň W-14) proti druhému instarovému štádiu *Popillia japonica* v Metromixe sa sledovalo v testoch, ktoré sa robili nasledovne (tabuľka 7). Približne 50 g suchého Metromixu sa dalo do 591 ml priehľadného plastického pohára. Metromix sa potom zmáčal 50 ml celkového objemu na 50 % (obj./obj.) zriedeného bujónového roztoku. Riedenie surového bujónu sa robilo vodou, pričom 50 % bujón sa pripravil pridaním 25 ml surového bujónu k 25 ml vody tak, aby sa získal celkový objem 50 ml. 1 % (hmotn./obj.) roztok proteózového peptónu č. 3 (PP3), ktorý sa riedil na 50 % koncentrácie normálne používaného média, sa použil ako bujón kontrolný. Po zmáčaní bolo päť jedincov druhého instarového štádia *Popillia japonica* umiestnených na povrch navlhčeného Metromixu. Zdravé larvy *Popillia japonica* sa rýchlo zavrtali do Metromixu. Tie larvy, ktoré sa nezavrtali v priebehu 1 hod. sa odstránili a nahradili sa čerstvými larvami. Poháre sa uzatvorili a umiestnili sa do inkubátora pri 28 °C v tme. Po siedmich dňoch sa larvy z Metromixu odobrali a odčítala sa mortalita. Účinnosť sa vypočítala ako percento mortality v porovnaní s kontrolou.

Tabuľka 7

Účinok bujónovej kultúry *Photorhabdus luminescens* (kmeň W-14) na *Popilla japonica* po predinfestačnom zavlažovaní (Metromix).

Prostriedok	Mortalita*	Mortalita %
Voda	7/15	47
Kontrolné médium (1.0 % obj./obj.)	12/19	63
Bujón (50 % obj./obj.)	17/20	85

\*vyjadrené ako pomer uhynutých/živých lariev

#### Príklad 4

Užitočnosť insekticídu aplikáciou na list

Účinnosť *Photorhabdus* bujónu proti vijačke kukuričnej sa zistila vtedy, keď sa bujón aplikoval priamo na povrch listov (tabuľka 8). V týchto testoch bol *Photorhabdus* bujón 100-násobne riedený kultivačným médiom a normálne sa aplikoval na povrch odobratých listov kukurice v množstve okolo  $6,0 \mu\text{l}/\text{cm}^2$  povrchu listu. Listy sa na vzduchu vysušili a nastrihali sa do pásov veľkých približne 2x2 palce. Listy sa stočili, zabezpečili papierovými svorkami a vložili do 1 uncových plastických pokusných pohárov s 0,25 palcovým 2 % agarom na dne, aby bola zabezpečená dostatočná vlhkosť. Potom sa dovnútra stočeného listu umiestnilo dvanásť neonátov vijačky kukuričnej a pohár sa uzatvoril. Vzorok sa hodnotil po 5 dňovej inkubácii v tme pri  $27^\circ\text{C}$  z hľadiska poškodenia listov požieraním a výťažnosti lariev.

Tabuľka 8

Účinnosť bujónovej kultúry *Photorhabdus luminescens* (kmeň W-14) na larvy vijačky kukuričnej po aplikácii na odobratých listoch kukurice pred infestáciou larvami škodcov.

Prostriedok	Poškodenie listu	Znovuzískané larvy	hmotnosť (mg)
Voda	Rozsiahle	55/120	0,42
Kontrolné médium	Rozsiahle	40/120	0,50
Bujón (1.0 % obj./obj.)	Stopy	3/120	0,15

Účinnosť bujónovej kultúry proti neonátom húseníc sivkovca (*Heliothis virescens*) sa ukázala po použití metódy ponoreného listu. Z rastliny sa obrali čerstvé listy bavlníka nastrihali sa z nich kúsky 18,5 mm korkovrtom. Tieto disky sa jednotlivo namáčali do kontrolného média (PP3) alebo do bujónovej kultúry *Photorhabdus luminescens* (kmeň W-14), ktorá sa prekoncentrovala približne 10-násobne, pričom sa použil tangenciálny filtračný systém Amicon (Beverly, MA) Proflux M12 s filtrom 10 kDa. Nadbytok tekutiny sa odstránil a narovnané papierové svorky sa položili cez stred disku. Papierové svorky sa potom vtlačili do plastického 1,0 uncového pokusného pohára, ktorý obsahoval približne 2,0 ml 1 % agaru. Toto slúžilo na to, aby sa disk listu vznášal nad agarom. Po vysušení listového disku sa jednotlivé neonáty lariev húsenice sivkavca umiestnili na disk a pohár sa prikryl viečkom. Potom sa poháre pevne uzatvorili do plastických sáčkov a uložili v tme pri 27 °C na 5 dní. Po tomto čase sa zostávajúce larvy a list zvažili. Takto sa stanovila miera poškodenia listu (tabuľka 9).

Tabuľka 9

Účinnosť bujónovej kultúry *Photorhabdus luminescens* (kmeň W-14) na neonáty húseníc sivkovca *Heliothis virescens* v teste ponoreného listu bavlníka.

Prostriedok	Listový disk		Konečná hmotnosť lariev (mg)
Kontrolné listy	55,7	1,3	na*
Kontrolné médium	34,0	2,9	4,3 ± 0,91
<i>Photorhabdus</i> bujón	54,3	1,4	0,0**

\* - neaplikovateľné

\*\* - neboli nájdené žijúce larvy

#### Príklad 5, časť A

#### Charakterizácia toxínových peptidových zlúčenín

Subsekventnou analýzou proteínových subjednotiek toxínov skupín, izolovaných spôsobom uvedeným v príklade 1, sa štiepili na 7 % SDS polyakrylamidovom géli pre elektroforézu v pomere 30 : 0,8 (akrylamid : BIS-akrylamid). Táto gélová matrica napomohla lepšiemu rozkladu väčších proteínov. Gélový systém, ktorý sa použil na hodnotenie molekulových hmotností subjednotiek skupiny 1 a skupiny 2 v príklade 1, tvoril 18 % gél v pomere 38 : 0,18 (akrylamid : BIS-akrylamid), umožnil širšie rozpätie veľkosti separácie, ale menší rozklad zlúčenín vyššej molekulovej hmotnosti.

V tejto analýze sa rozložilo skôr 10 ako 8 proteínových skupín. Tabuľka 10 zahŕňa kalkulované molekulové hmotnosti 10 rozložených skupín a priamo porovnáva molekulové hmotnosti odhadované za podmienok v tomto príklade s tými, ktoré boli v príklade predchádzajúcom. Nie je prekvapujúce, že v odlišných separačných podmienkach tohto príkladu, ktoré sa použili tu, sa zistili ďalšie skupiny. Zmeny medzi predchádzajúcim a novým

odhadom molekulovej hmotnosti sa tiež očakávali danými rozdielnosťami v analytických podmienkach. V analýze tohto príkladu sa predpokladá, že odhady vyšších molekulových hmotností sú presnejšie ako v príklade 1 výsledkom zdokonaleného rozkladu. Bohužiaľ sú tieto odhady založené na analýze SDS-PAGE, ktorá je typicky analyticky nepresná a má za následok odhady peptidov, ktoré môžu ďalej byť pozmenené prostredníctvom post- a kotranslačných modifikácií.

Sekvencia aminokyselín sa určila pre N-koncové časti piatich z 10 rozložených peptidov. Tabuľka 10 uvádza vzájomný vzťah molekulovej hmotnosti proteínov a identifikovaných sekvencií. V SEKV. ID. ČÍS. 2 isté analýzy naznačujú, že prolín vo zvyšku 5 by mohol byť asparagín (asn). V SEKV. ID. ČÍS. 3 isté analýzy naznačujú, že zvyšky aminokyseliny na pozíciách 13 a 14 sú obidve arginín (arg). V SEKV. ID. ČÍS.4 isté analýzy naznačujú, že zvyšky aminokyseliny na pozícii 6 by mohli byť tak alanín (ala) ako aj serín (ser). V SEKV. ID. ČÍS.5 isté analýzy naznačujú, že zvyšok aminokyseliny na pozícii 3 by mohla byť kyselina asparágová (asp).

Tabuľka 10

<u>ODHAD</u>	<u>NOVÝ ODHAD*</u>	<u>ZOZNAM SEKV.</u>
208	200,2 kDa	SEKV. ID. ČÍS.1
184	175,0 kDa	SEKV. ID. ČÍS.2
65,6	68,1 kDa	SEKV. ID. ČÍS.3
60,8	65,1 kDa	SEKV. ID. ČÍS.4
56,2	58,3 kDa	SEKV. ID. ČÍS.5
25,1	23,2 kDa	SEKV. ID. ČÍS.15

\* - nové odhady sú založené sa SDS-PAGE a nie sú založené na génových sekvenciách. SDS-PAGE je analyticky nepresná.



#### Príklad 5, časť B

#### Charakterizácia toxínových peptidových zlúčenín

Nová N-koncová sekvencia SEKV. ID. ČÍS.15: Ala Gln Asp Gly Asn Gln Asp Thr Phe Phe Ser Gly Asn Thr sa získala pri ďalšom N-koncovom sekvencovaní peptidov izolovaných z natívnych na HPLC purifikovaných toxínov, ktoré je opísané v príklade 5, v časti A vyššie. Tento peptid pochádza z génu *tcaA*. Peptid označovaný TcaA<sub>ii</sub> začína na pozícii 254 a ide na pozíciu 491, kde začína peptid TcaA<sub>iii</sub>, SEKV. ID. ČÍS. 4. Odhad veľkosti peptidu založený na génovej sekvencii je 25 240 Da.

#### Príklad 6

#### Charakterizácia toxínových peptidových zlúčenín

V ďalších analýzach bol toxínový proteínový komplex znova izolovaný z rastového média *Photobacterium luminescens* (po kultivácii bez prítomnosti tweenu), pričom sa uskutočnilo vyzrážanie 10 % až 80 % síranom amónnym. Nasledoval krok ionovými výmennej chromatografie (Mono Q) a dva kroky chromatografie triediacej veľkosť molekúl. Tieto podmienky boli rovnaké ako sa použili v príklade 1. Počas prvého kroku molekulárneho roztriedenia, sa objavil druhý biologicky aktívny pík pri asi  $100 \pm 10$  kDa. Táto frakcia, na základe merania proteínu, bola 20 až 50 krát menej aktívna ako väčší alebo primárny aktívny pík okolo  $860 \pm 100$  kDa (pôvodný). V tomto experimente sa počas izolácie vyštíepil menší aktívny pík okolo  $325 \pm 50$  kDa, ktorý si udržal značnú časť počiatočnej biologickej aktivity. Predpokladá sa, že pík 325 kDa súvisí alebo je odvodený z píkov 860 kDa.

V tejto analýze sa vyštíepil proteín 56 kDa. N-koncová sekvencia tohto proteínu je tu určená v SEKV. ID. ČÍS. 6. Je pozoruhodné, že tento proteín sa významne podieľa na identite a zachovaní so SEKV. ID. ČÍS. 5 na N-konci naznačujúc, že tieto

dva by mohli byť kódované rôznymi členmi génovej rodina a že proteíny, ktoré každý gén produkuje sú dostatočne podobné na to, aby boli obidva použiteľné v insekticídnom toxínovom komplexe.

Druhý, významný proteín 185 kDa bol rovnako prítomný v množstve prirovnateľnom proteínu 3 z tabuľky 10 a mohol by to byť ten istý proteín alebo proteínový fragment. N-koncová sekvencia tohto proteínu 185 kDa je znázornená v SEKV. ID. ČÍS. 7.

Doplnkové údaje o N-koncovej aminokyselinovej sekvencii sa tiež získali z izolovaných proteínov. Žiaden z určených N-koncových sekvencií sa nejavil identicky k proteínu identifikovanému v tabuľke 10. V izolovaných preparátoch boli prítomné ďalšie proteíny. Jednému z týchto proteínov sa odhadla molekulová hmotnosť 108 kDa a N-koncová sekvencia je znázornená v SEKV. ID. ČÍS. 8. Druhému z týchto proteínov sa odhadla molekulová hmotnosť 80 kDa a N-koncová sekvencia je v SEKV. ID. ČÍS. 9.

Keď sa pík proteínového materiálu, s približne 325 kDa, analyzoval podľa veľkosti, zistili sa skupiny s približne 51, 31, 28 a 22 kDa. Ako vo všetkých prípadoch, v ktorých sa molekulová hmotnosť určila analýzou elektroforetickej pohyblivosti, boli tieto molekulové hmotnosti vystavené falošným efektom, ktoré boli zapríčinené rozdielmi v pufo-iónovej koncentrácii, rozdielnosťami v sile elektroforézy a podobne. Jedným z bežných zistení by bolo pochopenie, že konečné hodnoty molekulovej hmotnosti sa nesmú určovať použitím týchto štandardných metód, a že každá z nich bola predmetom variácií. Odvodilo sa, že proteíny týchto veľkostí sú produktami degradácie väčších druhov proteínov (s veľkosťou približne 200 kDa), ktoré sa zistili vo väčších primárnych toxínových komplexoch.

Nakoniec, niekoľko preparátov obsahovalo proteín, ktorý mal N-koncovú sekvenciu, ktorá je zobrazená v SEKV. ID. ČÍS.10. Táto

sekvencia bola presne homologická so známym chaperoninovým proteínom. Tieto sú známe funkciou zhromažďovania veľkých proteínových komplexov. Aj keď by to prihlasovatelia nemali pripísať takejto zhlukovacej funkcii proteínu, identifikovaného SEKV. ID. ČÍS. 10, táto bola zhodná s existenciou opísaného proteínového toxínového komplexu, a tak by sa mohol chaperoninový proteín zahrnúť do jeho zhluku. Okrem toho, aj keď také proteíny nesvedčia priamo o tom, že majú toxické pôsobenie, mohol by byť tento proteín dôležitý na určenie celkovej štruktúrnej charakteristiky proteínového toxínu a teda môže prispieť k toxickému pôsobeniu alebo k stálosti komplexu in vivo po orálnej aplikácii.

Urobila sa subsekventná analýza stability toxínového proteínového komplexu voči proteínáze K. Určilo sa, že po 24 hodinovej inkubácii komplexu v prítomnosti 10 násobne vyššej koncentrácie proteínázy K sa aktivita prakticky eliminovala (mortalita orálnej aplikácie poklesla na asi 5 %). Tieto údaje súhlasia s proteínovým charakterom toxínu.

Toxickú aktivitu nepripúšťala ani dialyzačná membrána, čo je opäť v súlade so značnou veľkosťou natívneho toxínu.

#### Príklad 7

Izolácia, charakterizácia a čiastková sekvencia aminokyselín toxínov *Photorhabdus*

Izolácia a sekvencovanie N-koncových aminokyselín: V sadách experimentov prevádzaných paralelne s príkladmi 5 a 6 sa vyzrážali proteíny *Photorhabdus* síranom amónnym s upravenou bujónovou kultúrou *Photorhabdus*. Zvyčajne 2 až 3 litre na konečnú koncentráciu buď 10 % alebo 20 % pomalým pridávaním kryštálikov síranu amónneho. Po 1 hodinovom premiešavaní pri 4 °C sa materiál centrifugoval pri 12 000 x g počas 30 minút. Supernatant sa upravil na 80 % síran amónny. Miešal sa pri 4 °C 1 hodinu

a centrifugoval sa pri 12 000 x g počas 60 minút. Zrazenina sa znova rozpustila v jednej desatine objemu 10 mM  $\text{Na}_2\text{PO}_4$  s pH = 7,0 a dialyzovala sa rovnakým fosfátovým pufrom počas noci pri 4 °C. Dialyzovaný materiál sa centrifugoval pri 12 000 x g počas 1 hodiny pred ionovýmennou chromatografiou.

Aminokyselinová kolóna HR 16/50 Q Sepharose (Pharmacia) sa kondicionovala 10 mM  $\text{Na}_2\text{PO}_4$  s pH = 7,0. Centrifugovaná a dialyzovaná zrazenina síranu amónneho sa aplikovala na kolónu Q Sepharose pri prietoku 1,5 ml/min a prepierala sa extenzívne pri 3,0 ml/min. s vyrovnávacím pufrom, až kým optická hustota (vlnová dĺžka 280 nm) nedosiahla hodnotu nižšiu ako 0,100. Ďalej sa použil buď 60 minútový gradient NaCl, ktorý sa menil od 0 po 0,5 M pri prietoku 3 ml/min., alebo séria elučných krokov s použitím 0,1 M; 0,4 M a nakoniec 1,0 M NaCl počas 60 minútovej aplikácie každého z nich na kolónu. Frakcie sa zbierali a koncentrovali na Centriprep 100. Alternatívne by sa mohli proteíny eluovať jednoducho premytím 0,4 M NaCl bez predchádzajúcej elúcie 0,1 M NaCl.

Dva mililitre podielov koncentrovaných vzoriek na Sepharose Q sa dalo pri prietoku 0,5 ml/min. do gélovej filtračnej kolóny HR 16/50 Superose 12 (Pharmacia), ktorá sa nakondicionovala s 10 mM  $\text{Na}_2\text{PO}_4$ , pH = 7,0. Kolóna sa premývala rovnakým pufrom 240 min. pri prietoku 0,5 ml/min. a odoberali sa 2 minútové vzorky. Odobral sa zvyškový objem materiálu a koncentroval sa na Centriprep 100. Dva mililitre podielov vzoriek koncentrovaných na Superose 12 sa dali pri prietoku 0,5 ml/min. na gélovú filtračnú kolónu HR 16/50 Sepharose 4B-CL (Pharmacia), ktorá sa nakondicionovala 10 mM  $\text{Na}_2\text{PO}_4$ , pH = 7,0. Kolóna sa premývala tým istým pufrom 240 min. pri prietoku 0,5 ml/min. a odoberali sa 2 minútové vzorky.

Vylúčený proteínový pík sa podrobil druhej frakcionácii aplikáciou na gélovú filtračnú kolónu so živicom Sepharose CL-4B, ktorý oddeľuje proteíny v rozpätí od okolo 30 kDa až

1 000 kDa. Táto frakcia sa rozštiepila do dvoch pík. Menší pík, pri zvyškovom objeme (> 1 000 kDa) a väčší pík, ktorý eluoval pri zreteľnej molekulovej hmotnosti okolo 860 kDa. Po jednotýždňovej perióde ukázali subsekventné vzorky, získané gélovou filtráciou postupné objavovanie sa tretieho píku asi pri limitovanej proteolýze (približne 325 kDa), ktorý vyzeral, že pochádza z veľkého píku. Biotesty, ktoré sa uskutočnili na troch píkových frakciách ukázali, že zvyškový pík nie je aktívny, aj keď frakcia 860 kDa toxínového komplexu bola veľmi aktívna, ale pík 325 kDa bol menej aktívny i keď bol celkom účinný. SDS-PAGE analýza píkov na Sepharose CL-4B toxínových komplexov a rozdielnych fermentačných produktov odhalila dva jasné peptidové vzory označované "P" a "S". Tieto dva vzory boli charakteristické rozdielnosťami v molekulovej hmotnosti a v koncentrácii peptidových zlúčenín v ich frakciách. "S" vzor, produkovaný najčastejšie, mal 4 peptidy s vysokou molekulovou hmotnosťou (> 150 kDa), zatiaľ čo "P" vzor mal 3 peptidy s vysokou molekulovou hmotnosťou. Okrem toho peptidová frakcia "S" vykazovala 2 až 3 krát vyššiu účinnosť proti vijačke kukuričnej. Táto zmena by mohla súvisieť s obmenami v proteínovej expresii, dôsledkom starnutia inokula a/alebo iných faktorov založených na rastových parametroch starnúcich kultúr.

Miligramové množstvá píkových toxínových komplexov frakcií sa stanovili, že sú peptidové vzory "P" alebo "S" a podrobili sa preparatívnej SDS-PAGE a presali sa Tris glycerínom (Seprabuff<sup>TM</sup>) na membrány PVDF (Problott<sup>TM</sup>, Applied Biosystems) 3 až 4 hodiny. Plaky sa poslali na analýzy aminokyselín a N-koncového sekvencovania aminokyselín na Harvard MicroChem, prípadne na Cambridge ProChem. Tri peptidy vzoru "S" mali jedinečné N-koncové sekvencovanie aminokyselín porovnateľné so sekvenciami, ktoré sa identifikovali v predošlom príklade. Peptidová sada 201 kDa (TcdA<sub>ii</sub>), ďalej uvádzaná ako SEKV. ID Čís. 13, s podielom menším ako 33 % identity aminokyselín a 50 % podobnosti (podobnosť a identita sa vypočítali ručne) so SEKV. ID. Čís. 1 (TcbA<sub>ii</sub>) (tabuľka 10, vertikálne čiary v tabuľke 10 znamenajú

identické aminokyseliny a dvojbodky identifikujú konzervatívne aminokyselinové substitúcie). Druhý peptid 197 kDa, SEKV. ID. Čís. 14 (TcdB), mal 42 % identitu a 58 % podobnosť so SEKV. ID. Čís. 2 (TcaC). Tretí peptid 205 kDa sa označoval TcdA<sub>ii</sub>. Okrem toho obmedzená N-koncová aminokyselinová sekvencia, SEKV. ID. Čís. 16 (TcbA), peptidu s najmenej 235 kDa, bola identická v podobnosti so sekvenciou aminokyselín SEKV. ID. Čís. 12, odvodennej z klonovaného génu (*tcbA*), SEKV. ID. Čís. 11, ktorý obsahuje odvodenú sekvenciu aminokyselín súhlasiacu so SEKV. ID. Čís. 1 (TcbA<sub>ii</sub>). Toto naznačuje, že väčší peptid 235 kDa bol počas fermentácie proteolyticky štiepený na peptid 201 kDa (TcbA<sub>ii</sub>), (SEKV. ID. Čís. 1) s možným následkom v aktivite molekuly. Už v ďalšej sekvencii, sekvencovanie originálne označované ako SEKV. ID. Čís. 5 (TcaB<sub>ii</sub>), o ktorom sa už predtým písalo v príklade 5, sa našiel obsah zvyškov kyseliny aspartámovej (Asp) na tretej pozícii menej často ako glycín (Gly) a dve ďalšie aminokyseliny Gly a Asp na ôsmej alebo deviatej pozícii. Avšak v dvoch ďalších sekvenciách, SEKV. ID. Čís. 2 (TcaC) a SEKV. ID. Čís. 3 (TcaB<sub>i</sub>), sa získala prídavná sekvencia aminokyselín. Denzitometrická kvantifikácia sa uskutočnila s použitím vzorky, ktorá bola identická s "S" preparátom, ktorý sa poslal na analýzu N-konca. Táto analýza ukázala, že peptidy 201 kDa a 197 kDa predstavujú 7,0 %, prípadne 7,2 % z celkového ofarbenia proteínu Coomasiho brilantnou modrou v "S" vzore a sú prítomné v množstve podobnom ako je hojnosť iných peptidov. To evokuje myšlienky, že tieto peptidy môžu reprezentovať proteínové homológy, analogicky k situácii, kedy sa našli s ďalšími bakteriálnymi toxínmi, takými ako rôzne toxíny CryI Bt. Tieto proteíny sa menia od 40 do 90 % homológie na ich N-koncovej sekvencii aminokyselín, ktoré obsahujú toxické fragmenty.

Vnútorne sekvencovanie aminokyselín: Na to, aby sa uľahčilo klonovanie génov peptidového toxínu, získavala sa vnútorná sekvencia aminokyselín vybraných peptidov tak, ako je to ďalej opisované. Miligramové množstvo z píku 2A frakcií, určených ako "P" alebo "S" vzory peptidov, sa podrobilo preparatívnej SDS-

-PAGE a prenášali sa na membrány PVDF Tris glycerínom (Seprabuff™) (ProBlott™, Applied Biosystems) 3 až 4 hodiny. Plaky sa zaslali na Harvard MicroChem alebo na Cambridge ProChem na analýzu aminokyselín a N-koncového sekvencovania aminokyselín. Tri peptidy so vzťahom k TcbA<sub>ii</sub> (obsahujúci SEKV. ID. Čís. 1), TcdA<sub>ii</sub> a TcaB<sub>i</sub> (obsahujúci SEKV. ID. Čís.3) sa podrobili štiepeniu tripsínom v Harvard MicroChem s následnou analýzou na HPLC chromatografe preto, aby sa oddelili jednotlivé peptidy. Analýza aminokyselín N-konca bola založená na selekcii fragmentov triptických peptidov. Dva vnútorné peptidy sa sekvencovali pre peptid TcdA<sub>ii</sub> (205 kDa peptid), ktorý má súvislosť jednak k TcdA<sub>ii</sub>-PT111 (SEKV. ID. Čís. 17) jednak k TcdA<sub>ii</sub>-PT79 (SEKV. ID. Čís. 18). Dva vnútorné peptidy sa sekvencovali pre peptid TcaB<sub>i</sub> (68 kDa peptid), ktorý má súvislosť tak k TcaB<sub>i</sub>-PT158 (SEKV. ID. Čís. 19) ako aj k TcaB<sub>i</sub>-PT108 (SEKV. ID. Čís. 20). Štyri vnútorné peptidy boli sekvencované pre peptid TcbA<sub>ii</sub> (201 kDa peptid) so súvisom jednak k TCBA<sub>ii</sub>-PT103 (SEKV. ID. Čís. 21), jednak k TcbA<sub>ii</sub>-PT56 (SEKV. ID. Čís. 22), TcbA<sub>ii</sub>-PT81(a) (SEKV. ID. Čís. 23) a jednak k TcbA<sub>ii</sub>-PT81(a) (SEKV. ID. Čís. 24).

Tabuľka 11

N-koncové sekvencie aminokyselín.

(podobnosť a identita sa vypočítali ručne)

201 kDa (33 % identita a 50 % podobnosť k SEKV. ID. Čís. 1)

L I G Y N N Q F S G * A	SEKV. ID. Čís. 13
:       :	
F I Q G Y S D L F G N - A	SEKV. ID. Čís. 1

197 kDa (42 % identita a 58 % podobnosť k SEKV. ID. Čís. 2)

M Q N S Q T F S V G E L	SEKV. ID. Čís. 14
:     : :	
M Q D S P E V S I T T L	SEKV. ID. Čís. 2

## Príklad 8

Konštrukcia kozmidovej knižnice genómovej DNA *Photobacterium luminescens* W-14 a opis izolácie génov kódujúcich peptidy obsahujúce preparatívne toxické proteíny

Tak ako sa to vyžaduje pre produkciu *Photobacterium* toxických proteínov pre hmyz v heterológnom hostiteľovi, aj v inom využití, je potrebné izolovať a charakterizovať gény kódujúce tieto peptidy. Tento predmet sa sledoval súbežne. Jeden prístup, opisovaný neskôr, bol založený na použití monoklonálnych a polyklonálnych protilátok, ktoré sa pestovali oproti purifikovanému toxínu. Ten sa potom použil na izoláciu klonov z expresnej knižnice. Ďalší prístup, ktorý sa v tomto príklade opísal, je založený na použití N-koncovej a vnútornej sekvencii aminokyselín preto, aby sa navrhli degenerované oligonukleotidy, ktoré by sa využili na amplifikáciu v PCR. Ďalšia metóda sa môže tiež použiť na identifikáciu klonov DNA, obsahujúcich gény kódujúce peptidy tak, ako to dovoľuje izolácia reštriktívnych génov a determinácia ich bázevej sekvencie DNA.

Izolácia genómovej DNA: Kmeň W-14 druhu *Photobacterium luminescens* (ATCC zbierkové číslo 55 397) rástol na 2 % proteózo-  
vom peptónovom agare č. 3 (Difco Laboratories, Detroit, MI) a insekticídna toxínová schopnosť sa podporovala opakovaným biotestom po pasážovaní, pričom sa používala metóda opísaná vyššie v príklade 1. 50 ml z kultúry, ktorá sa miešala trepaním, sa produkovalo v 175 ml nepriehľadnej nádobe v 2 % proteózo-  
vom peptónovom médiu č. 3. Rástla pri 28 °C a 150 rpm. približne 24 hodín. Z takejto kultúry sa 15 ml odsedimentovalo a zmrazilo v médiu pri -20 °C až pokiaľ sa nerozpustilo na izoláciu DNA. Roztopená kultúra sa centrifugovala (700 x g, 30 min.) a vznášajúci sa pomarančový mukopolysacharidový materiál sa odstránil. Zostávajúci bunkový materiál sa centrifugoval (25 000 x g, 15 min.) preto, aby sa získali odsedimentované bakteriálne bunky a médium sa odstránilo a odhodilo.



Genómová DNA sa izolovala metódou adaptácie CTAB opísanej v časti 2.4.1. v: Bežné postupy molekulárnej biológie (Ausubel a kol. edícia, John Wiley a Sons, 1994) [ako modifikácia sa zaradil salinný šok a všetky použité objemy vzrástli 10 násobne]. Odsedimentované bakteriálne bunky sa resuspendovali v pufri TE (10 mM Tris HCl, 1 mM EDTA, pH = 8,0) na konečný objem 10 ml. Potom sa pridalo 12 ml 5 M NaCl a takáto zmes sa 20 min. centrifugovala pri 15 000 x g. Sediment sa resuspendoval v 5,7 ml TE a 300 ml 10 % SDS a 60 ml 20 mg/ml proteinázy K (Gibco BRL Products, Grand Island, NY; v sterilnej destilovanej vode) sa pridali do suspenzie. Táto zmes sa inkubovala pri 37 °C počas 1 hodiny. Potom sa pridalo približne 10 mg lyzozýmu (Worthington Biochemical Corp., Freehold, NJ). Po ďalších 45 minútach sa pridal 1 ml 5 M NaCl a 800 ml roztoku CTAB/NaCl (10 % hmotn./obj. CTAB, 0,7 M NaCl). Takýto preparát sa inkuboval 10 min. pri 65 °C a potom sa jemne zamiešal a opätovne sa inkuboval a miešal približne 20 minút, aby sa umožnila purifikácia bunkového materiálu. Pridal sa rovnaký objem chloroformu/roztoku izoamylalkoholu (24 : 1, obj./obj.), jemne sa zamiešal a centrifugoval. Po dvoch extrakciách s rovnakým objemom PCI (fenol/chloroform/izoamylalkohol; 50 : 49 : 1, obj./obj./obj.; pufrované s 1 M Tris HCl, pH = 8,0, Intermountain Scientific Corporation, Kaysville, UT) sa DNA vyzrážala s 0,6 objemom izopropanolu. Zrazenina DNA sa jemne odoberala sklenenou tyčinkou, dvakrát sa premyla 70 % etanolom, vysušila sa a rozpustila sa v 2 ml STE (10 mM Tris HCl, pH = 8, 10mM NaCl, 1 mM EDTA). Takýto preparát obsahoval 2,5 mg/ml DNA, čo sa určilo stanovením optickej hustoty pri 260 nm (napr OD<sub>260</sub>).

Rozsah veľkosti molekuly izolovanej genómovej DNA sa hodnotil z dôvodu vhodnosti pre zostavenie knižnice. CHEF gélová analýza sa robila v 1,5 % agarózovom géli (Seakem LE, FMC BioProducts, Rockland, ME) s 0,5 x TBE pufrom (44,5 mM Tris HCl, pH = 8,0; 44,5 mM H<sub>3</sub>BO<sub>3</sub>; 1 mM EDTA) na prístroji BioRad CHEF-DR II s Pulsewave 760 Switcher (Bio-Rad Laboratories Inc., Richmond, CA). Nastavené parametre boli: počiatkový čas A 3 s, konečný čas

A 12 s, 200 voltov, beh teploty 4 až 18 °C, časový priebeh 16,5 hod. Farbenie etídiombromidom a hodnotenie gélu pod ultrafialovým svetlom indikovalo DNA v rozmedzí veľkostí od 30 do 250 kbp.

Zostavenie knižnice: výsledkom preparácie *Photorhabdus* genómovej DNA bola čiastočná zbierka Sau3A. Metóda bola založená na časti 3.1.3. Ausubel (vyššie). Úpravy spočívali v sérii menšieho rozsahu reakcií v rôznych podmienkach, až kým sa nedosiahli výsledky blízke optimálnym. Nasledovalo niekoľko zväčšujúcich reakcií v rôznych podmienkach. Výsledky sa analyzovali na géloch CHEF a len s najobjemnejšou škálou preparátov sa pokračovalo ďalej. V optimálnom prípade sa inkubovalo 200 µg genómovej DNA *Photorhabdusu* s 1,5 jednotkami Sau3A 1 (New England Biolabs, "NEB", Beverly, MA) 15 minút pri 37 °C v celkovom objeme 20ml pufru 1X NEB 4 (dodávaný výrobcom ako 10X). Reakcia sa zastavila pridaním 2 ml PCI a centrifugovala sa pri 8 000 x g 10 minút. Do supernatantu sa pridalo 200 µl 2 M NaCl a 6 ml ľadovo studeného etanolu. Tento preparát sa ochladzoval 30 min. pri -20 °C. Potom sa centrifugoval 15 minút pri 12 000 x g. Supernatant sa odstránil a zrazenina sa vysušila vo vákuovej peci pri 40 °C a potom sa resuspendovala v 400 µl STE. Spektrofotometrické stanovenie indikovalo približne 40 % výťažnosť DNA. Triedenie DNA sa robilo pomocou veľkostnej frakcionácie v sacharózovom gradiente podľa časti 5.3.2. CPMB (citácia). Pripravil sa 10 % až 40 % (hmotn./obj.) lineárny sacharózový gradient s gradientovým rozdeľovaním v skúmavkách Ultra-Clear<sup>TM</sup> (Beckman Instruments, Inc., Palo Alto, CA) a vzorka DNA sa navrsvila navrch. Po centrifugácii (26 000 rpm. 17 hodín, rotor Beckman SW41, 20 °C) sa odoberali frakcie (okolo 750 µl) z povrchu gradientu a analyzovali sa gélovou CHEF elektroforézou (tak, ako bolo opísané skôr). Frakcie, ktoré obsahovali fragmenty Sau3A 1 s veľkosťou v rozmedzí 20 až 40 kpb sa oddelili a DNA sa vyzrážala s použitím modifikovanej metódy v časti 5.3.3. Ausubel (vyššie) (množstvo všetkých roztokov vzrásto približne 6,3 násobne). Po vyzrážavaní, ktoré trvalo celú noc, sa DNA zhromažďovala centrifugáciou (17 000 x g, 15 min.), vysušila,

znova rozpustila v TE. Spojila sa do konečného objemu 80  $\mu$ l a znova sa vyzrážala pridaním 8  $\mu$ l 3 M octanu sodného a 220  $\mu$ l etanolu. Zrazeniny sa zhromažďovali centrifugáciou tak, ako je uvedené vyššie, rozpustili sa v 12  $\mu$ l TE. Koncentrácia sa určila farebnou fluorimetriou na fluorimetri Hoechst 33258 (Polysciences Inc., Warrington, PA), Hoefer TKO 100 (Hoefer Scientific Instruments, San Francisco, CA). Výťažnosť bola približne 2,5  $\mu$ g veľkostne frakcionovanej DNA.

Tridsať  $\mu$ g DNA kozmidu pWE15 (Stratagene, La Jolla, CA) sa štiepil preto, aby sa doplnil 100 jednotkani reštrikčného enzýmu BamH 1 (NEB) vo vyrábanom pufri (konečný objem 200  $\mu$ l, 37 °C, 1 hod.). Reakcia sa extrahovala so 100  $\mu$ l PCI a DNA sa vyzrážala z vodnej fázy pridaním 20  $\mu$ l 3 M octanu sodného a 550  $\mu$ l -20 °C ľadového absolútneho etanolu. Po 20 min. pri -70 °C sa DNA zhromažďovala centrifugáciou (17 000 x g, 15 min.), vysušila sa vo vákuu a rozpustila v 100  $\mu$ l 10 mM Tris HCl s pH = 8,0. K tomu sa pridalo 20  $\mu$ l pufri 10X CIP (100 mM Tris HCl, pH = 8,3; 10 mM ZnCl<sub>2</sub>; 10 mM MgCl<sub>2</sub>) a 1  $\mu$ l (0,25 jednotky) telacej intestinálnej alkalickéj fosfatázy nariedenej 1 : 4 (Boehringer Mannheim Corporation, Indianapolis, IN). Po 30 min. pri 37 °C sa pridávali nasledovné látky: 2  $\mu$ l 0,5 M EDTA s pH = 8,0; 10  $\mu$ l 10 % SDS; 0,5  $\mu$ l z 20 mg/ml proteínázy K (tak, ako sa uvádza vyššie). Nasledovala inkubácia pri 55 °C počas 30 min. V nasledovnej sekvenciálnej extrakcii so 100  $\mu$ l PCI a 100  $\mu$ l fenolu (Intermountain Scientific Corporation, upravený s 1,0 M Tris HCl pH = 8,0) sa defosforylovaná DNA vyzrážala 72  $\mu$ l 7,5 M octanu amónneho a 550  $\mu$ l -20 °C etanolu, inkubovala na ľade 30 min. a centrifugovala tak ako je uvedené vyššie. Scentrifugovaná DNA sa jedenkrát premyla s 500  $\mu$ l -20 °C 70 % etanolu, vysušila sa vo vákuu a rozpustila v 20  $\mu$ l pufri TE.

Ligácia veľkostne frakcionovaných fragmentov Sau3A 1 na BamH-1-štiepený a fosforylovaný vektor pWE15 sa dokončila pomocou T4 ligázy (NEB) modifikáciou postupu v časti 3.3.3. Ausbel (napríklad použitie premiešaného 10X ligačného pufri dodávaného

výrobcom) Ligácia sa uskutočnila počas noci v celkovom objeme 20  $\mu$ l pri 15 °C a potom sa skladovala pri -20 °C.

Štyri  $\mu$ l z ligačnej reakcie kozmidovej DNA, obsahujúce okolo 1  $\mu$ g DNA, sa fúzovali do bakteriofága lambda, pričom sa použil komerčne dostupný fúzovací extrakt (Gigapack III Gold Packing Extract, Stratagene), ktorý sa pripravil podľa predpisu výrobcu. Fúzovaný preparát sa uskladňoval pri 4 °C až dovtedy, kým sa nepoužil. Preparát fúzovaného kozmidu sa použil na infekciu XLI modrých MR buniek *Escherichia coli* (Stratagene) podľa postupov v Gigapack III Gold ("Titrovanie kozmidovej knižnice") tak, ako je to tu opísané. Modré XLI MR bunky rástli na LB médiu (g/l; baktotryptón 10; baktokvasinkový extrakt 5; baktoagar 15; NaCl 5; [Difco laboratories, Detroit, MI]), ktoré obsahovalo 0,2 % (hmotn./obj.) maltózy a 10 mM MgSO<sub>4</sub> pri teplote 37 °C. Po piatich hodinách rastu sa bunky scentrifugovali pri 700 x g (15 min.) a resuspendovali sa do 6 ml 10 mM MgSO<sub>4</sub>. Hustota kultúry sa upravila s 10 mM MgSO<sub>4</sub> na OD<sub>600</sub> = 0,5. Fúzovaná kozmidová knižnica sa zriedila 1 : 10 alebo 1 : 20 sterilným SM médiom (0,1 M NaCl; 10 mM MgSO<sub>4</sub>; 50 mM Tris HCl s pH = 7,0; 0,01 % hmotn./obj. želatíny) a 25  $\mu$ l takto riedeného preparátu sa zmiešalo s 25  $\mu$ l zriedených XLI modrých MR buniek. Zmes sa inkubovala pri 25 °C počas 30 minút (bez pretrepávania). Potom sa pridalo 200  $\mu$ l LB bujónu a inkubácia pokračovala približne hodinu za občasného jemného pretrepávania. Podiely (20 až 40  $\mu$ l) tejto kultúry sa rozotreli na agarové platne s obsahom ampicilínu 100 mg/l (napríklad LB-Amp<sub>100</sub>) a inkubovali sa pri 37 °C počas noci. Na uloženie knižnice bez rozširovania, sa odobrali jednotlivé kolónie a inkubovali sa v jednotlivých jamkách na sterilných 96-jamkových mikrotitračných platniach, pričom každá jamka obsahovala 75  $\mu$ l Terrifikovho bujónu (TB médium: 12 g/l baktotryptónu; 24 g/l baktokvasinkového extraktu; 0,4 % obj./obj. glycerolu; 17 mM KH<sub>2</sub>PO<sub>4</sub>; 72 mM K<sub>2</sub>HPO<sub>4</sub>) a 100 mg/l ampicilínu (napríklad TB-Amp<sub>100</sub>) a inkubovali sa (bez pretrepávania) pri 37 °C počas noci. Po replikácii 96-jamkovej mikrotitračnej platne do svojho duplikátu sa pridalo na platňu

75  $\mu$ l/jamka filtrom sterilizovaného TB: glycerolu (1 : 1 obj./obj.; s alebo bez 100 mg/l ampicilínu). Krátko sa pretrepali pri 100 rpm, pri 37 °C a potom sa uzatvorili parafilmom (American National Can., Greenwich, CT) a uložili sa v mrazničke pri -70 °C. Na duplicitných platniach organizmy rástli a podrobili sa identickým postupom ako primárne platne. Celkom sa pripravilo 40 primárnych mikrotitračných platní (a ich kópií).

Skríning knižnice s rádioaktívne značenou sondou DNA: Na prípravu filtrov s kolóniami s rádioaktívne značenými sondami sa desať 96-jamkových platní z knižnice roztopilo pri 25 °C (horná hranica pre izbovú teplotu). Multikanálové pracovné pomôcky s 96 špičkami sa použili na inokuláciu čerstvej kópie 96-jamkovej platne, ktorá obsahovala 75  $\mu$ l/jamka TB-Amp<sub>100</sub>. Obsah jamiek v kópiách platní rástol cez noc (stacionárne) pri 37 °C. Potom sa pretrepával asi 30 min. pri 100 rpm pri 37 °C. Celkove tieto kópie predstavovali 800 kolónií, pričom tento počet vznikol tým, že niektoré izoláty nenarástli. Použila sa multikanálová pracovná pomôcka na inokuláciu duplikátov impresií z 96-jamkových radov na nylónové membrány (0,45 mikrónu, 220 x 250mm) Magna NT (MSI, Westboro, MA), ktoré sa umiestnili na tuhý LB-Amp<sub>100</sub> (100  $\mu$ l/nádoba) v plastických nádobách pre biotesty (Nunc., 243 x 243 x 18 mm; Curtin Mathison Scientific Inc., Wood Dale, IL). Kolónie na membránach rástli pri 37 °C počas 3 hodín.

Pozitívne kontrolné kolónie (bakteriálny klon obsahujúci sekvenčnú inzerciu GZ4, pozri nižšie) rástli na samostatnej membráne Magna NT (Nunc., 0,45 mikrónov, 82 mm kruh) na médiu LB doplnenom 35 mg/l chloramfenikolu (napríklad LB-Cam<sub>15</sub>) a pripravované boli mimo membrán kolónií knižnice. Bakteriálne kolónie sa na membránach lýzovali a DNA sa denaturovala a neutralizovala podľa postupu prebraného z Genius<sup>TM</sup> System Uživatelská príručka verzia 2.0 (Boehringer Mannheim, Indianapolis, IN). Membrány sa umiestnili stranou s kolóniami na filtračný papier nasiaknutý 0,5 N NaOH a 1,5 M NaCl na 15 min., aby sa denaturovali a neutralizovali sa na filtračnom papieri

nasiaknutom 1 M Tris HCl pH = 8,0; 1,5 M NaCl tiež 15 min. Po UV zosieťovaní použitím stratagénovej UV sady Stratagenlink a samozosieťovaní, sa membrány uložili vysušené pri 25 °C až do použitia. Membrány sa nastrihali na pásiky obsahujúce duplikát odtlačkov z jednotlivých 96-jamkových platní. Potom sa extenzívne premyli podľa metódy uvedenej v CPMB časti 6.4.1. (citované): 3 hod. pri 25 °C v 3X SSC; 0,1 % (hmotn./obj.) SDS. Ďalej nasledovala 1 hod. pri 65 °C v takom istom roztoku a potom sa vypláchli v 2X SSC v príprave pre hybridizačný krok (20X SSC = 3 M NaCl; 0,3 M citrát sodný pH = 7,0).

Amplifikácia špecifického genómového fragmentu génu TcaC. Je založená na určení N-koncovej sekvencie aminokyselín pre purifikáciu peptidovej frakcie TcaC (tu určený ako SEKV. ID. Čís. 2). Reakčná zmes degenerovaných oligonukleotidov (zmes S4Psh) sa syntetizovala štandardnou chemikáliou  $\beta$ -kyanoetyl na aplikovanom biosystémovom syntetizéri Applied BioSystem AB1394 DNA/RNA syntetizér (Perkin Elmer, Foster City, CA). Oligonukleotidy sa zbavovali ochrany 8 hodín pri 55 °C, rozpustili vo vode, kvantifikovali sa spektrofotometrickým meraním a riedili sa na ďalšie použitie. Táto zmes je v súhlase s určenou N-koncovou sekvenciou aminokyselín peptidu TcaC. Určenie sekvencie aminokyselín a zhodnosť degenerovanej DNA sekvencie sa dokázala nižšie, kde A, C, G a T sú štandardné DNA bázy a I predstavuje inozín:

Amino-	Met	Gln	Asp	Ser	Pro	Glu	Val
kyselina							
S4Psh	5'	ATG CA(A/G)	GA(T/C)	(T/A)(C/G)(T/A)	CCI	GA(A/G)	GT3'

Ďalšia sada degenerovaných polynukleotidov sa syntetizovala (zmes P2.3.5.R), pričom reprezentovala doplnok kódovaného reťazca na určenie sekvencie aminokyselín SEKV. ID. Čís. 17:

Amino

kyselina	Ala	Phe	Asn	Ile	Asp	Asp	Val
kodóny	5'GCN	TT(T/C)	AA(T/C)	AT(A/T/C)	GA(T/C)	GA(T/C)	GT 3'
P2.3.5R	3' CG(A/C/G/T)	AA(A/G)	TT(A/G)	TA(T/A/G)	CT(A/G)	CT(A/G)	CA 5'

Tieto oligonukleotidy sa použili ako základy v polymerázových reťazových reakciách (PCR, Roche Molecular Systems, Branchburg, NJ) na amplifikáciu fragmentu špecifického fragmentu DNA z genómovej DNA pripravenej z *Photorhabdus* kmeňa W-14 (pozri vyššie). Typická reakcia (50 µl) obsahovala 125 pmol každej primárnej zmesi P2Psh a P2.3.5R, 253 ng genómovej templátovej DNA, 10 nmol každej dATP, dCTP, dGTP, dTTP, 1X GeneAmp PCR pufor a 2,5 jednotiek AmpliTaq DNA polymerázy (obidve z Roche Molecular Systems; pufor 10X GeneAmp je 100 mM Tris HCl s pH = 8,3; 500mM KCl; 0,01 % hmotn./obj. želatíny). Amplifikácia sa robila na teplotnom cykleri Perkin Elmer Cetus DNA (Perkin Elmer, Foster City, CA) za použitia 35 cyklov 94 °C (1,0 min.), 55 °C (2,0 min.), 72 °C (3,0 min.). Potom nasledovala extenzná perióda 7,0 min. pri 72 °C. Amplifikované produkty sa analyzovali elektroforézou cez 2 % hmotn./obj. NuSieve 3 : 1 agarózu (FMC BioProducts) v pufri TEA (40 mM Tris octan, 2 mM EDTA pH = 8,0). Špecifický produkt odhadnutej veľkosti 250 kbp sa pozoroval oproti množstvu ďalších produktov amplifikácie sfarbením gélu etídium bromidom (0,5 µg/ml) a vyšetrovaním pod ultrafialovým svetlom.

Úsek gélu, ktorý obsahoval produkt približne 250 bp sa odobral a malý plak (0,5 mm priemer) sa premiestnil a použil ako zásoba templátu pre PCR amplifikáciu (40 cyklov). Reakcia (50µl) obsahovala tie isté zlúčeniny, aké sú uvedené vyššie, okrem genómovej templátovej DNA. Nasledujúcou amplifikáciou sa konce fragmentov zarovnali a fosforylovali inkubáciou pri 25 °C počas 20 min. s 1 jednotkou T4 DNA polymerázy (NEB), 1 nmol ATP a 2,15 jednotky T4 kinázy (Pharmacia Biotech Inc., Piscataway, NJ)

DNA fragmenty sa oddeľovali od reziduálnych primérov

elektroforézou cez 1 % hmotn./obj. GTG agarózu (FMC) v TEA. Gélová vrstva obsahujúca fragmenty zjavnej veľkosti 250 bp sa odobrala a DNA sa extrahovala súpravou Qiaex (Qiagen Inc., Chatsworth, CA).

Extrahované fragmenty DNA sa ligátovali k plazmidovému vektoru pBC KS (+) (Stratagene) tak, že sa štiepili na uskutočnenie reštrikčným enzýmom Sma I a extrahovali sa podobne ako sa to opisovalo pre pWE15 DNA vyššie. Typická ligačná reakcia (16,3  $\mu$ l) obsahovala 100 ng vybranej pBC KS (+) DNA, 70 ng fragmentu DNA 250 bp, 1 nmol  $[\text{Co}(\text{NH}_3)_6]\text{Cl}$  a 3,6 Weissových jednotiek T4 DNA ligázy (Collaborative Biomedical Products, Bedford, MA) v 1X ligačnom pufri (50 mM Tris HCl pH = 7,4; 10 mM  $\text{MgCl}_2$ ; 10mM ditiotreitolu; 1 mM spermidinu; 1 mM ATP; 100 mg/ml hovädzieho albumínového séra). Po nočnej inkubácii pri 14 °C sa ligátované produkty fúzovali do zmrazených príslušných DH5 $\alpha$  buniek *Escherichia coli* (Gibco BRL) podľa dodávateľových odporučení a umiestnili sa na LB-Cam<sub>35</sub> platne obsahujúce IPTG (119  $\mu$ g/l) a X-gal (50  $\mu$ g/l). Odobrali sa jednotlivé biele kolónie a pripravil sa plazmid DNA modifikovanou alkalyzačnou / PEG zrážacou metódou (PRISM<sup>TM</sup> Ready Reaction DyeDeoxy<sup>TM</sup>, postupy súprav terminačného cyklu sekvencovania ABI/Perkin Elmer). Nukleotidová sekvencia, obidvoch reťazcov fúzovanej DNA sa určila pomocou primérov T7 [pBC KS (+) bázy 601 až 623: TAAAACGACGGCCAGTGAGCGCG] a primérov LacZ [pBC KS (+) bázy 792 až 816: ATGACCATGATTACGCCAAGCGCGC] a postupu podľa sekvencovacej súpravy PRISM<sup>TM</sup> (ABI/Perkin Elmer). Nefúzované farebné terminačné dideoxyribonukleotidy sa odstránili pasážovaním cez kolóny Centri - Sep 100 (Princeton Separations Inc., Adelphia, NJ) podľa inštrukcií výrobcu. Sekvencia DNA sa získala analýzou vzoriek na sekvenceri ABI modeli 373A DNA (ABI/Perkin Elmer). Objavili sa DNA sekvencie dvoch izolátov, GZ4 a H B14 a sú uvedené na obrázku 1.



Táto sekvencia objasňuje nasledujúce črty:

- i) Bázy 1 až 20 predstavujú jednu zo 64 možných sekvencií S4Psh degenerovaných oligonukleotidov;
- ii) sekvencia aminokyselín 1 až 3 a 6 až 12 sa exaktne zhoduje s determináciou pre N-koniec TcaC (určený ako SEKV. ID. Čís. 2;
- iii) štvrtá kódovaná aminokyselina je zriedkavejšie cysteínový zvyšok ako serínový. Tento rozdiel je kódovaný v rámci degenerácie serínových kodónov (pozri vyššie);
- iv) piata, kódovaná aminokyselina je prolín, súhlasiaci s N-koncovou sekvenciou TcaC, ktorá je určená ako SEKV. ID. Čís. 2;
- v) bázy 257 až 276 kódujú jednu zo 192 možných sekvencií, ktoré boli navrhnuté do degenerovaného spojenia;
- vi) terminačný kodón TGA, ktorý je včlenený pri bázach 268 až 270 je výsledkom komplementarity degeneratívneho zabudovania oligonukleotidového spoja na zodpovedajúcej pozícii a neindukuje skrútený čítací rámec pre zodpovedajúci gén.

Označenie špecifickej sondy peptidového génu TcaC: Fragmenty DNA zhodujúce sa tak ako predtým vyššie 276 bázami sa amplifikovali (35 cyklov) pomocou PCR v 100  $\mu$ l reakčnom objeme, pričom sa použilo 100 pmol z každého priméru P2Psh a P2.3.5R, 10 ng plazmidov GZ4 alebo HB14 ako templátov, 20 nmol z každej dATP, dCTP, dGTP a dTTP, 5 jednotiek polymerázy DNA AmpliTaq a 1X koncentrácia pufru GeneAmp v rovnakých teplotných režimoch ako je opísané vyššie. Produkty amplifikácie sa extrahovali z 1% agarózového gélu GTG súpravou Qiaex a kvantifikovali sa fluorimetricky.

Extrahované produkty amplifikácie z plazmidového templátu HB14 (približne 400 ng) sa rozštiepili na päť častí a označili  $^{32}\text{P}$ -dCTP sondou pomocou veľkej základnej značkovacej zmesi High Prime Labeling Mix (Boehringer Mannheim) podľa návodov výrobcu. Nezavedený rádioizotop sa odstránil pasážovaním cez vzorkovacie purifikačné kolóny NucTrap (Stratagene) podľa inštrukcií dodávateľa. Špecifická aktivita značenej sondy DNA sa určila scintiláciou odčítavanou pri  $3,11 \times 10^8$  dpm/ $\mu\text{g}$ . Takto značená DNA sa použila na membrány so sondami, ktoré sa pripravili z 800 členov genómovej knižnice.

Skríning so špecifickou sondou peptidového génu TcaC: Rádioaktívne značená sonda HB14 sa zohrieva približne 10 minút a potom sa pridala k roztoku "minimal hyb" (Poznámka: metóda "minimal hyb" je prebratá z postupu CERES: Reštrikcia polymorfnej fragmentovej dĺžky, Laboratórny manuál, verzia 4,0 časti 4-40 a 4-47, CERES/NPI, Salt Lake City, UT. NPI v súčasnosti zanikla, so svojimi nástupcami pracujú ako Väzbová Genetika). Roztok "Minimal hyb" obsahuje 10 % hmotn./obj. PEG (polyetylénglykol M.W. pribl. 8 000), 7 % hmotn./obj. SDS, 0,6X SSC, 10 mM pufor fosforečnanu sodného (z 1 M zásobného roztoku obsahujúceho 95 g/l  $\text{NaH}_2\text{PO}_4 \cdot 1\text{H}_2\text{O}$  a 84,5 g/l  $\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$ ), 5 mM EDTA a 100 mg/ml denaturovanej spermálnej DNA lososa. Membrány sa krátko odsali do sucha, potom sa 5 pásov membrány bez prehybridizácie umiestnilo do každej z dvoch plastických krabičiek obsahujúcich 75 ml "Minimal hyb" a 2,6  $\mu\text{g}/\text{ml}$  rádioaktívne značenej sondy HB14. Tieto sa inkubovali počas noci za pomalého trepania (50 rpm.) pri 60 °C. Filtre sa trikrát premyli zakaždým po približne 10 minútach pri 25 °C v premývacom roztoku "minimal hyb" (0,25X SSC, 0,2 % SDS), po ktorom nasledovali dve 30 min. premývania s pomalým trepaním pri 60 °C v tom istom roztoku. Filtre sa umiestnili na papier prekrytý so Saran Wrap (Dow Brands, Indianapolis, IN) do tmavej autorádiografickej kazety a exponovali sa na X-Omat X-lúčový film (Kodak, Rochester, NY) s dvoma zosilňovačmi DuPont Cronex Lighting-Plus C1 (Sigma Chemical Co., St. Louis, MO) počas 4 hodín pri -70 °C. Po

vyvolaní (štandardné fotografické postupy) boli zrejme významné signály v obidvoch replikátoch medzi vysokými zoslabovačmi pozadia, viaceré nepravidelné signály. Filtre sa opäť premyli asi 4 hod. pri 68 °C v premývacom roztoku "minimal hyb" a potom sa znova umiestnili do kaziet a film sa exponoval počas noci pri -70 °C. Identifikovalo sa dvanásť vhodných pozitívov, ktoré spôsobili silné signály na obidvoch duplikátoch 96-jamkových kolóniových impresií. Nijaký signál sa neobjavil na membránach s negatívnou kontrolou (kolónie XLI modré MR bunky obsahujúce pWE15) a veľmi silný signál sa objavil na membránach s pozitívnou kontrolou (DH-5 $\alpha$  bunky, obsahujúce izolát GZ4 produktu PCR), ktorý sa pripravil súčasne s experimentálnymi vzorkami.

Dvanásť zdanlivo pozitívnych kolónií na hybridizáciu sa znova získalo zo zmrazených 96-jamkových platní knižnice a rástli počas noci pri 37 °C na tuhom médiu LB - Amp<sub>100</sub>. Tieto sa potom naočkovali na tuhú pôdu LB - Amp<sub>100</sub> (3/miska plus tri negatívne kontroly: XLI modré MR bunky obsahujúce vektor pWE15). Pripravili sa dve sady membrán (nylón Magna NT, 0,45 mikrónov) na hybridizáciu. Prvá sada sa pripravila umiestnením filtra priamo na kolónie na odočkovanej platni. Potom sa odstránil s adherentnými bakteriálnymi bunkami a podrobila sa postupu, ktorý je uvedený vyššie. Filtre druhej sady sa umiestnili na platne obsahujúce médium LB - Amp<sub>100</sub>, potom sa inokulovali prenesením buniek z odočkovaných platní na filtre. Po nočnom raste pri 37 °C sa filtre odstránili z platní a upravovali sa.

Bakteriálne bunky na filtroch sa podrobili lýze a DNA sa denaturovala umiestnením každého filtra stranou s kolóniami na zmes (1,0 ml) 0,5 N NaOH v plastických miskách 3 min. Filtre sa dosucha odsali na papierovej rúške. Potom sa postup zopakoval s čerstvým 0,5 N NaOH. Po vysušení odsávaním sa filtre neutralizovali tak, že sa každý z nich umiestnil na 1,0 roztoku 1 M Tris HCl s pH = 7,5 počas 3 minút, opäť sa vysušili odsatím a reneutralizovali sa čerstvým pufrom. Toto sa robilo dvoma podobnými spôsobmi namáčania (5 min. každý) na roztok 0,5 M Tris

HCl pH = 7,5 a 1,5 M NaCl. Po vysušení odsatím sa DNA pomocou UV zosieťovala s filtrom (ako vyššie) a filtre sa premyli (25 °C, 100 rpm) v asi 100 ml 3X SSC a 0,1 % (hmotn./obj.) SDS (4 krát, 30 min. každá s čerstvým roztokom pre každé premývanie). Potom sa položili do minimálneho objemu predhybridizačného roztoku [6X SSC plus 1 % hmotn./obj. každého z Ficoll 400 (Pharmacia), polyvinilpyrolidón (priem. M.W. 360 000, Sigma) a hovädzie albumínové sérum frakcia V, (Sigma) 2 hodiny pri 65 °C a 50 rpm. Predhybridizačný roztok sa odstránil a nahradil so <sup>32</sup>P značenou sondou HB14, ktorá sa uchovávala z predchádzajúcej hybridizácie membránovej knižnice, a ktorá bola denaturovaná pri 95 °C 5 min. Hybridizácia sa uskutočnila pri 60 °C 16 hodín za stáleho trepania pri 50 rpm.

Nasledovalo odstránenie roztoku značenej sondy a potom sa membrány premyli 3 krát pri 25 °C (50 rpm, 15 min.) v 3X SSC (okolo 150 ml pre každé premývanie). Potom sa premývali 3 hodiny pri 68 °C (50 rpm) v 0,25X SSC s 0,2 % SDS (premyvací roztok "minimal hyb") a exponovali sa na film s X-lúčmi tak, ako sa opísalo vyššie, 1,5 hodiny pri 25 °C (bez clony zosilňovača transkripcie). Táto expozícia odhalila veľmi silné hybridizačné signály kozmidových izolátov 22612, 25A10, 26A5 a 26B10 a veľmi slabý signál izolátu kozmidu 8B10. Nijaký signál sa neobjavil v kolóniách negatívnej kontroly (pWE15) a veľmi silný signál sa objavil na membránach s pozitívnou kontrolou (DH5α, bunky obsahujúce izolát GZ4 produktu PCR), ktoré sa pripravili súčasne s experimentálnymi vzorkami.

Amplifikácia špecifického genómového fragmentu génu *tcaB*: Je založená na sekvencii N-koncových aminokyselín určených na purifikovanú TcaB<sub>1</sub> peptidovú frakciu (určená ako SEKV. ID. Čís. 3). Zmes degenerovaných oligonukleotidov (zmes P8F) sa syntetizovala tak, ako sa opísalo pre peptid TcaC. Určená sekvencia aminokyselín a zhodujúca sa sekvencia degenerovanej DNA sú uvedené nižšie, pričom A, C, G a T sú štandardné bázy DNA a I predstavuje inozín:

amino

kyselina Leu Phe Thr Gln Thr Leu Lys Glu Ala Arg

P8F 5' TTT ACI CA(A/G) ACI (C/T)TI AAA GAA GCI (A/C)G 3'  
(C/T)TI

Ďalšia sada degenerovaných oligonukleotidov sa syntetizovala (zmes P8.108.3R) reprezentujúc doplnok kódovaných reťazcov na určenie sekvencie aminokyselín vnútorného peptidu TcaB<sub>1</sub>-PT108 (v tejto štúdiu určený ako SEKV. ID. ČÍS. 20):

amino

kyselina Met Tyr Tyr Ile Gln Ala Gln Gln

kodóny ATG TA(T/C) TA(T/C) AT(T/C/A) CA(A/G) GC(A/C/G/T) CA(A/G) CA(A/G)

P8.108.3R 3' AT(A/G) AT(A/G) TA(A/G/T) GT(T/C) CGI GT(T/C) GT 5'

TAC

Tieto oligonukleotidy sa použili ako priméry pre PCR v skúmavkách HotStart 50 (Molecular Bio-Products Inc., San Diego, CA) na amplifikáciu fragmentov špecifickej DNA z genómovej DNA, ktorá bola pripravená z kmeňa W14 druhu *Photorhabdus* (pozri vyššie). Typická reakcia (50 µl) obsahovala (spodná vrstva) 25 pmol každej primérovej zmesi P8F a P8.108.3R s 2 nmol každej z dATP, dCTP, dGTP a dTTP v pufri 1X GeneAmp PCR a (vrchná vrstva) 230 ng génového templátu DNA, 8 nmol každej z dATP, dCTP, dGTP a dTTP a 2,5 jednotky DNA polymerázy AmpliTaq v 1X PCR pufri GeneAmp. Amplifikácia sa robila v 35 cykloch tak, ako sa opísalo pre peptid TcaC. Produkty amplifikácie sa analyzovali na elektroforéze cez 0,7 % hmotn./obj. agarózu SeaKem LE (FMC) v pufri TEA. Pozoroval sa špecifický produkt odhadovaný na veľkosť 1 600 bp.

Štyri takéto reakcie sa spojili a amplifikovaná DNA sa extrahovala z 1,0 % gélu SeaKem LE súpravou Qiaex tak, ako sa opísalo pre peptid TcaC. Extrahovaná DNA sa použila priamo ako

templát pre určovanie sekvencie (sekvencovacia súprava PRISM<sup>TM</sup>) pomocou P8F a P8.108.3R primérových zmesí. Každá reakcia obsahovala okolo 100 ng templátu DNA a 25 pmol jednej primérovej zmesi a upravovala sa podľa štandardných postupov tak, ako sa opísalo pre peptid TcaC. Analýza sekvencií odvodená z extenzie primérov P8F odhalila krátke sekvencie DNA (a kódovanej sekvencie aminokyselín):

GAT GCA TTG NTT GCT  
Asp Ala Leu (Val) Ala,

ktoré sa zhodujú s časťou N-koncovkej peptidovej sekvencie, ktorá je tu určená ako SEKV. ID. ČÍS. 3 (TcaB<sub>i</sub>).

Značenie špecifickej sondy peptidového génu TcaB<sub>i</sub>.

Približne 50 ng gélom purifikovaných fragmentov DNA TcaB<sub>i</sub> sa označilo s <sup>32</sup>P-dCTP tak, ako sa opísalo vyššie a neinkorporované rádioizotopy sa odstránili pasážovaním cez kolónu NICK (Pharmacia). Špecifická aktivita značenej DNA sa určila na 6 x 10<sup>9</sup> dpm/μg. Značená DNA sa použila na prípravu sondových membránových kolónií zo zástupcov genómovej knižnice, ktorá sa hybridizovala na TcaC peptidovú špecifickú sondu.

Membrány obsahujúce 12 kolónií a identifikované v sondovej knižnici TcaC (pozri vyššie) sa zbavili rádioaktívneho TcaC špecifického značkovania dvojnásobným varom približne 30 minút v každom čase v 1 litri 0,1X SSC plus 0,1 % SDS. Odstránenie rádioaktívneho značenia sa overilo 6 hodinovou expozíciou na film. Purifikované membrány sa potom inkubovali s peptid špecifickou sondou TcaB<sub>i</sub>, ktorá sa pripravila vyššie. Značená DNA sa denaturovala varom 10 minút a potom sa pridala k filtrom, ktoré sa inkubovali 1 hod. v 100 ml roztoku "minimal hyb" pri 60 °C. Po celonočnej hybridizácii pri tejto teplote sa odstránil roztok vzorky a filtre sa premyli podľa nasledovného postupu (všetko v 0,3X SSC plus 0,1 % SDS): jedenkrát 5 min. pri 25 °C, jedenkrát 1 hod. pri 60 °C v čerstvom roztoku a jedenkrát 1 hod.

pri 63 °C v čerstvom roztoku. Po 1,5 hod. expozícii na X-lúčový film podľa štandardných postupov sa zaznamenali 4 silne hybridizované kolónie. Tieto tak, ako TcaC-špecifické sondy, boli izolátmi 22612, 25A10, 26A5 a 26B10.

Taký istý roztok sondy TcaB<sub>i</sub> sa riedil rovnakým objemom (okolo 100 ml) roztoku "minimal hyb" a potom sa použil na opis membrán obsahujúcich 800 zástupcov genómovej knižnice. Po hybridizácii, premývaní a expozícii na X-lúčový film tak, ako sa opísalo hore. V tejto vzorke sa ukázali byť silne hybridizované len štyri kozmidové klony 22612, 25A10, 26A5 a 26B10.

Izolácia subklonov obsahujúcich gény kódujúce TcaC a TcaB<sub>i</sub> peptidy a určenie ich sekvencie báz. Tri hybridizačne pozitívne kozmidy kmeňa XL1 modré MR rástli počas noci (200 rpm.) za stáleho trepania pri 30 °C v 100 ml TB-Amp<sub>100</sub>. Po získaní buniek centrifugáciou sa kozmid DNA upravil komerčne dostupnou súpravou (BIGprep<sup>TM</sup>, 5 Prime 3 Prime, Inc., Boulder, CO) podľa návodu na použitie od výrobcu. Týmto postupom sa izoloval len jeden kozmid 26A2. Keď sa štiepil s reštrikčným enzýmom EcoR 1 (NEB) a analyzoval sa gélovou elektroforézou, detegovali sa fragmenty približných veľkostí 14, 10, 8 (vektor), 5; 3,3; 2,9 a 1,5 kbp. Druhý pokus izolovať kozmid DNA z tých istých troch kmeňov (8 ml kultúr TB-Amp<sub>100</sub>, 30 °C) využil metódu minipreparačného varu (Evans G. a Wahl G., 1987, "Kozmidové vektory pre genómové prešmyky a rapídne reštrikčné mapovanie." v návode na techniky molekulárneho klonovania. Meth Enzymology, vol. 152, S. Berger a A. Kimmel, ed., str. 604 - 610). Touto metódou sa izoloval len jeden kozmid 25A10. Keď sa štiepil s reštrikčným enzýmom EcoR 1 (NEB) a analyzoval gélovou elektroforézou, objavila sa identická fragmentačná štruktúra tohto kozmidu, ktorá sa pozorovala predtým u kozmidu 26A5.

Na transformáciu 50 ml buniek DH5α E. coli (Gibco BRL) sa použilo 0,15 μg vzorky DNA kozmidu 26A5 podľa inštrukcií dodávateľa. Jednotlivé izoláty kolónií tohto kmeňa sa inokulovali

do 4 ml TB-Amp<sub>100</sub> a rástli 8 hodín pri 38 °C. Do výslednej koncentrácie 225 µg/ml sa pridal chloramfenikol a inokulácia pokračovala ďalších 24 hodín. Bunky sa potom získali centrifugáciou a zmrazili sa pri -20 °C. Izolácia DNA kozmidu 26A5 sa robila štandardnou metódou alkalickéj miniprep lýzy (Maniatis a kol., op. cit, str. 382), ktorá bola modifikovaná tak, že sa zvýšili všetky objemy o 50 % a uprednostnilo sa miešanie alebo jemné pretrepávanie pred používaním vortexu pri každom kroku metódy. Po premytí centrifugátu DNA v 70 % etanole, sa táto rozpustila v TE s obsahom 25 µg/l ribonukleázy A (Boehringer Mannheim).

Identifikácia fragmentov EcoR 1 hybridizovanými s GZ4 odvođenými a TcaB<sub>1</sub> sondami. Približne 0,4 µg kozmidu 25A10 (z XLI modrých MR buniek) a asi 0,5 µg kozmidu 26A5 (z buniek DH5a chloramfenikolom amplifikované) sa štiepilo s cca 15 jednotkami EcoR 1 (NEB) 85 min., cez noc boli zmrazené, potom sa ohriali na 65 °C päť minút a na 0,7 % agarózovom géli sa podrobili elektroforéze (SeaKem LE, 1X TEA, 80 voltov, 90 min.). Podľa postupu, ktorý je opísaný vyššie sa DNA zafarbila etídium bromidom a fotografovala sa pod ultrafialovým svetlom. EcoR 1 odštiepený z kozmidu 25A10 sa komplexne vyštíepil, ale vzorka kozmidu 26A5 sa vyštepovala v týchto podmienkach len čiastočne. Agarózový gél obsahujúci fragmenty DNA sa podrobil depurinácii, denaturácii a neutralizácii s následným prenosom na nylónovú membránu Magna NT, pričom sa použil vysoko salinný postup (20X SSC) tak, ako je opísané v časti 2.9. Ausubel a kol. (CPMB., op. cit.). Transferovaná DNA sa potom zosieťovala pomocou UV na nylónovú membránu tak, ako predtým.

Špecifický fragment DNA peptidu TcaC, zhodný s inzertom plazmidového izolátu GZ4, sa amplifikoval v PCR v 100 ml reakčnom objeme tak, ako sa to opísalo predtým vyššie. Premyli sa produkty amplifikácie z troch takýchto reakcií a extrahovali sa z 1 % agarózového gélu GTG Qiaexovej súpravy tak, ako je opísané vyššie a kvantifikovalo sa fluorimetriou. Gélom purifikovaná DNA (100



$\mu\text{g}$ ) sa označila s  $^{32}\text{P}$ -dCTP pomocou High Prime Labeling Mix (Boehringer Mannheim) tak, ako sa to opísalo vyššie pri špecifickej aktivite  $6,34 \times 10^8$  dpm/ $\mu\text{g}$ .

Sonda GZ4, označená s  $^{32}\text{P}$  sa varila 10 min. Potom sa pridala do pufru "minimal hyb" ( $1\mu\text{g}/\text{ml}$ ) a pridala sa Southernova prenosová membrána obsahujúca vyzbierané fragmenty kozmidovej DNA a inkubovala sa 4 hod. pri  $60^\circ\text{C}$  s jemným pretrepávaním pri 50 rpm. Potom sa membrána 3 krát premyla pri  $25^\circ\text{C}$  cca 5 min. pre každé jedno premývanie (premyvací roztok "minimal hyb"). Nasledovali dve premývania po 30 min. pri  $60^\circ\text{C}$ . Plak sa exponoval na film (so zosilňovacími clonami) približne 30 min. pri  $-70^\circ\text{C}$ . Vzorka GZ4 silne hybridizovala k 5,0 kbp (zrejma veľkosť) EcoR 1 fragmentu obidvoch týchto kozmidov, 26A5 a 26A10.

Membrána sa zbavila rádioaktivity varom trvajúcim cca 30 min. v 0,1X SSC a 0,1 % SDS. Absencia rádioaktívneho značkovania sa overovala expozíciou na film. Potom sa hybridizovala pri  $60^\circ\text{C}$  3,5 hod. s (denaturovanou) sondou TcaB<sub>i</sub> v pufri "minimal hyb" ktorý sa predtým použil na skrining kolóniových membrán (hore). Premyla sa tak ako sa to opísalo už predtým a exponovala sa na film 40 minút pri  $-70^\circ\text{C}$  s dvoma zosilňovacími clonami. Sonda TcaB<sub>i</sub> s oboma kozmidmi hybridizovala ľahko k cca 5,0 kpb fragmentom EcoR 1 a silne k fragmentom približne 2,9 kpb.

Vzorka DNA kozmidu 26A5, opísaného už predtým (z DH5a buniek) sa použila ako zdroj DNA z ktorej sa subklonovali skupiny, ktoré sú predmetom záujmu. Táto DNA ( $2,5 \mu\text{g}$ ) sa štiepila s cca 3 jednotkami EcoR 1 (NEB) v celkovom objeme  $30 \mu\text{l}$  počas 1,5 hod. preto, aby sa získala čiastočná zbierka tak, ako sa potvrdilo gélovou elektroforézou. Desať  $\mu\text{g}$  pBC KS (+) DNA (Stratagene) sa štiepilo 1,5 hod. s 20 jednotkami EcoR 1 v celkovom objeme  $20 \mu\text{l}$  vedúcom totálnemu štiepeniu, čo sa overilo elektroforézou. Oba preparáty EcoR 1 rozdelenej DNA sa zriedili na  $50 \mu\text{l}$  s vodou, pričom sa v každom preparáte pridal rovnaký objem PCI. Suspenzia sa jemne miešala, centrifugovala

a vodný supernatant sa zbieral. DNA sa vyžrážala so 150  $\mu$ l etanolu a zmes sa uložila pri  $-20\text{ }^{\circ}\text{C}$  počas noci. Po ďalšej centrifugácii a sušení sa zozbierané EcoR 1 pBC KS (+) rozpustili v 100  $\mu$ l TE; čiastočne vyštiepená 26A5 sa rozpustila v 20  $\mu$ l TE. Výťažnosť DNA sa overovala fluorimetricky.

V samostatných reakciách sa ligovalo približne 60 ng EcoR 1 štiepeného spBC KS (+) DNA s približne 180 ng alebo 270 ng čiastočne štiepenej DNA kozmidu 26A5. Ligácia sa uskutočnila v objeme 20  $\mu$ l pri  $15\text{ }^{\circ}\text{C}$  5 hodín ligázou T4 a pufrom z New England BioLabos. Ligačná zmes, riedená sterilným TE na 100  $\mu$ l, sa použila na transformáciu zmrazených, príslušných buniek DH5 $\alpha$  (Gibco BRL) podľa návodu dodávateľa. Rozličné množstvá (25 až 200  $\mu$ l) transformovaných buniek sa dali na čerstvo pripravené tuhé médium LB - Cam<sub>35</sub> s 1 mM IPTG a 50 mg/l X-gal. Misky sa inkubovali pri  $37\text{ }^{\circ}\text{C}$  okolo 20 hod. Potom sa ochladzovali v tme asi 3 hod. preto, aby sa zintenzívnila farba inzertovanej sekvencie. Biele kolónie sa odobrali na misky s rovnakým zložením a inkubovali sa počas noci pri  $37\text{ }^{\circ}\text{C}$ .

Dve kolónie odobrané z každej z vybraných misiek sa preparovali nasledovane. Po odpichnutí bielych kolónií na čerstvé platne, sa vinuté nylónové MT membrány pretlačili do odtlačkových platní, membrána sa odstránila a podrobila sa denaturácii, neutralizácii a UV zosieťovaniu tak, ako je to opísané vyššie pri knižnici membránových kolónií. Kolónie zosieťované na membráne sa odobrali a energicky premyli, vrátane jemného zotretia prebytkov bunkového odpadu s tkanivami. Jedna časť sa hybridizovala s roztokom sondy GZ4 (TcaC), opísaným už skôr a ďalšia časť sa hybridizovala s TcaB<sub>i</sub> roztokom sondy, ktorý bol opísaný skôr, podľa postupu "minimal hyb" s následným premytím a expozíciou na film tak, ako je to opísané pre knižnicu membránových kolónií.

Kolónie vykazujúce hybridizačné signály, boli len vo vzorke GZ4, s obidvomi vzorkami GZ4 a TcaB<sub>i</sub> alebo len so vzorkou TcaB<sub>i</sub>, sa vybrali pre ďalšiu prácu a bunky sa rozotreli na izoláciu

jednotlivých kolónií na LB-Cam<sub>35</sub> médiu s IPTG a X-galom tak, ako predtým. Približne 35 jednotlivých kolónií zo 16 rôznych izolátov sa odpichlo do tekutého média LB - Cam<sub>35</sub> a rástli počas noci pri 37 °C. Bunky sa zbierali pomocou centrifugácie a plazmid DNA sa izoloval štandardne alkalickou lýzou miniprep podľa Maniatisa a kol. ( op. cit. str. 368). Zrazeniny DNA sa rozpustili v TE + 25 µg/ml ribonukleázy A a koncentrácia DNA sa určovala fluorimetricky. Vzorka štiepená EcoR 1 sa analyzovala na gélovej elektroforéze. Odobrali sa ďalšie izoláty tak, ako bolo potrebné. Izolát A17.2 obsahuje len religátovaný pBC KS (+) a používal sa ako (negatívna) kontrola. Izoláty D383 a C44.1, obsahujúce len 2,9 kbp TcaB<sub>i</sub>-hybridizovaného fragmentu EcoR 1 inzertovaného do pBC KS (+). Tieto plazmidy sú znázornené ako pDAB2000, respektíve pDAB2001 na obr.2.

Izolát A35.3 obsahuje len asi 5 kbp, GZ4 - hybridizovaného fragmentu EcoR 1, fúzovaného do pBC KS (+). Tento plazmid sa nazýva pDAB2002 (tiež na obr. 2). Tieto izoláty poskytli templáty pre sekvencovanie DNA.

Plazmidy pDAB2000 a pDAB2001 sa preparovali pomocou súpravy BIGprep<sup>TM</sup> tak, ako predtým. Kultúry (30 ml) rástli počas noci v médiu TB-Cam<sub>35</sub> na OD<sub>600</sub> = 2. Potom sa plazmidy izolovali podľa pokynov výrobcu. Každá zo zrazenín DNA sa znova rozpustila v 100 ml TE a integrita vzorky sa overovala EcoR 1 štiepením a na gélovej elektroforézovej analýze.

Reakcie sekvencovania sa robili duplicitne s jedným replikátom používajúcim templátovú DNA pDAB2000 a ďalšia paralelka používala ako templát DNA pDAB2001. Reakcie sa uskutočňovali pomocou dideoxi farebnej terminačnej cyklickej metódy sekvencovania tak, ako je opísané vyššie pri sekvencovaní GZ4/HB14 DNA. Iniciálne sekvencovanie prebiehalo využívajúc ako priméry LacZ, tak aj priméry T7, ktoré sú opísané vyššie a priméry založené na určovaní sekvencie TcaB<sub>i</sub>, produkty amplifikácie PCR (TH1 = ATTGCAGACTGCCAATCGTTCCG, TH12 =

= GAGAGTATCCAGACCGCGGATGATCTG).

Po zoradení a úprave každého sekvencovaného výťažku, každý sa odstrihol medzi 250 až 350 bázami, v závislosti na integrite chromatografických údajov tak, ako sa interpretovali softvérom Perkin Elmer Applied Biosystems Division SeqEd 675. Subsekventné sekvenčné "krokovanie" sa robilo vybraním vhodnej sekvencie pre nové priméry. S niekoľkými málo výnimkami mali priméry (syntetizované tak ako je opísané vyššie) dĺžku 24 báz s 50 % obsahom G+C. Sekvencovanie touto metódou sa uskutočňovalo na obidvoch reťazcoch asi 2,9 kbp fragmentu EcoR 1.

Na ďalšie použitie ako templátu na sekvencovanie DNA sa plazmid DNA pripravil z izolátu súpravou B16prep<sup>TM</sup>. Sekvenčné reakcie sa uskutočnili a analyzovali tak, ako je opísané vyššie. Na základné sekvenčné reakcie sa použil najprv primér T3 (pBS SK (+) bázy 774 až 796: CGCGCAATTAACCCTCACTAAAG) a primér T7 (pBS KS (+) bázy 621 až 643: GCGCGTAATACGACTCACTATAG) z bočných vektorových sekvencií zakódovaných do inzertovanej DNA. Ďalšia sada primérov (GZ4F: GTATCGATTACAACGCTGTCCTCCC; TH13: GGAAGTG-ACAGCGTTGTAATCGATAC; TH14: ATGTTGGGTGCGTTCGGCTAATGGACATAAC; a LW1-204: GGAAGTGACAGCGTTGTAATCGATAC) sa urobila na základe vnútorných sekvencií, ktoré sa určili predtým degeneráciou oligonukleotidov stredných sekvencií subklonovaných produktov peptidu TcaC PCR. Z údajov tvorených počas počiatkových cyklov sekvencovania sa označili nové sady primérov a použili sa na odmeranie celej dĺžky fragmentu okolo 5 kbp. Celkovo sa použilo 55 oligonukleotidov, kde sa mohlo identifikovať 4832 celkových bp príľahlých sekvencií.

Keď sa sekvencia DNA fragmentu EcoR 1 inzertovaná z pDAB2002 skombinovala s časťou determinovanej sekvencie izolátov pDAB2000, pDAB2001, vytvorila sa celková príľahlá sekvencia 6005 bp (tu určená ako SEKV. ID. Čís. 25). Keď sa dlhé otvorené čítacie rámce translatovali do zodpovedajúcich aminokyselín, sekvencia jasne ukázala N-koncový peptid TcaB<sub>i</sub> (tu

uvedená ako SEKV. ID. ČÍS. 3), ktorý je kódovaný bázami 68 až 124, so zvyškom metionínu nasledujúcim hneď za nimi (začiatok translácie). Vyššie leží potenciálne ribozómové väzbové miesto (bázy 51 až 58) a v smere transkripcie, pri bázach 215 až 277 je kódovaný vnútorný peptid TcaB<sub>i</sub>-PT158 (opísaný ako SEKV. ID. ČÍS. 19). Ďalej v smere transkripcie dole v tom istom čítacom rámci, v bázach 1787 až 1822 je prítomná sekvencia kódujúca vnútorný peptid TcaB<sub>i</sub>-PT108 (určený ako SEKV. ID. ČÍS. 20). Taktiež v tom istom čítacom rámci v bázach 1946 až 1972 je kódovaný N-koncový peptid TcaB<sub>ii</sub> (určený tu ako SEKV. ID. ČÍS. 5) a čítací rámec pokračuje bez prerušenia translácie terminačným kodónom pri nukleotidoch 3632 až 3634.

Neprítomnosť stop kodónu v rámci medzi koncom sekvencie kódujúcej TcaB<sub>i</sub>-PT108 a štart kodónom kódujúcim oblasť TcaB<sub>ii</sub> a neprítomnosť rozpoznávacieho ribozómového väzbového miesta, hneď proti smeru transkripcie od kódovanej oblasti TcaB<sub>ii</sub> indikujú, že peptidy TcaB<sub>ii</sub> a TcaB<sub>i</sub> sú kódované nezávislými otvorenými čítacími rámcami z 3567 bp začínajúceho na bázy pár 65 v SEKV. ID. ČÍS. 25) a sú najpravdepodobnejšie odvodené z nezávislého primárneho produktu génu s 1189 aminokyselinami (131 586 Daltonov, určený ako SEKV. ID. ČÍS. 26) posttranslačným štiepením. Ak aminokyselina hneď predchádzajúca N-koncový peptid TcaB<sub>ii</sub>, predstavuje C-koncovú aminokyselinu peptidu TcaB<sub>i</sub>, potom predpovedané množstvo TcaB<sub>ii</sub> (627 aminokyselín) je 70 814 Daltonov (tu určená ako SEKV. ID. ČÍS. 28) trochu väčšie ako veľkosť zistená SDS-PAGE (68 kDa). Tento peptid by bol kódovaný príslušným rozsahom 1881 párov báz (určený ako SEKV. ID. ČÍS. 27). Predpokladá sa, že pôvodný C-koniec TcaB<sub>i</sub> leží trochu bližšie k C-koncu TcaB<sub>i</sub>-PT108. Molekulová hmotnosť PT108 [3 438 kDa, stanovená N-koncovou aminokyselinovou sekvenčnou analýzou tohto peptidu] predpokladá veľkosť 30 aminokyselín. Ak sa použije veľkosť tohto peptidu na určenie C-konca kódovacej oblasti TcaB<sub>i</sub> (Glu na pozícii 604 SEKV.ID. ČÍS. 28) potom odvodená veľkosť TcaB<sub>i</sub> je určená na 604 aminokyselín alebo 68 463 Daltonov, čo je vo väčšej zhode s experimentálnymi pozorovaniami.

Translácia oblasti kódujúcej peptid TcaB<sub>ii</sub>, pozostávajúcej z 1686 parov báz (určená ako SEKV.ID. ČÍS. 29) vedie k proteínu s 562 aminokyselinami (určený ako SEKV.ID. ČÍS. 30) s predpokladanou hmotnosťou 60 789 Daltonov, čo je v dobrej zhode so zaznamenanými 61 kDa.

Potenciálne ribozómové väzbové miesto (bázy 3633 až 3638) sa našlo 48 bp v smere transkripcie od stop kodónu pre *tcaB* otvorený čítací rámec. V bázach 3694 až 3726 sa objavila sekvencia kódujúca N-koniec peptidu TcaC (určený ako SEKV.ID. ČÍS. 2). Otvorený čítací rámec začínajúci týmto N-koncovým peptidom pokračuje bez prerušenia k báze 6005 (2361 párov báz, tu určená ako prvých 2361 párov báz SEKV.ID. ČÍS.31). Gén (*tcaC*) kódujúci úplný peptid TcaC (so zdanlivou veľkosťou asi 165 kDa, asi 1500 aminokyselín) by obsahoval okolo 4500 bp.

Ďalší izolát obsahujúci klonované fragmenty *EcoR* 1 z kozmidu 26A5, E20.6 sa tiež identifikoval pomocou homológie k predtým zmieneným vzorkám GZ4 a TcaB<sub>i</sub> sondám. Agarózová gélová analýza *EcoR* 1 štiepanej DNA plazmidu fúzovanej z tohto kmeňa (pDAB2004, obr. 2) odhalila fúzované fragmenty odhadovanej veľkosti 2,9; 5 a 3,3 kbp. Sekvenčná analýza DNA iniciovaná priméromi pochádzajúcimi zo sekvencie plazmidu pDAB2002 odhalila, že 3,3 kbp fragment *EcoR* 1 z pDAB2004 leží vedľa 5 kbp fragmentu *EcoR* 1 reprezentujúceho pDAB2002. 2361 pár báz otvoreného čítacieho rámca objaveného v pDAB2002 pokračuje bez prerušenia k ďalším 2094 bázam v pDAB2004 (tu určený ako páry báz 2362 až 4458 SEKV. ID. ČÍS. 31). Sekvenčná analýza DNA využívajúca parentálny kozmid 26A5 DNA ako templát potvrdila kontinuitu otvoreného čítacieho rámca. Celkove otvorený čítací rámec (*tcaC* SEKV. ID. ČÍS. 31) obsahuje 4455 párov báz a kóduje proteín (TcaC) pozostávajúci zo 1485 aminokyselín (tu určená ako SEKV. ID. ČÍS. 32). Vypočítaná molekulová veľkosť 166 214 Daltonov je v súlade s odhadovanou veľkosťou peptidu TcaC (165 kDa) a odvodená sekvencia aminokyselín presne zodpovedá priradeniu pre N-koncovú sekvenciu TcaC (SEKV. ID. ČÍS. 2).

Nepřítomnost sekvencie aminokyselín zhodných so SEKV. ID. ČÍS. 17, ktorá sa použila na označenie degenerovanej oligonukleotidovej zmesi priméru v objavenej sekvencii indikuje, že generácia produktov PCR sa nachádza v izolátoch GZ4 a HB14, ktoré sa použili ako sondy v počiatočnom triedení knižnice a náhodne sa generovali spätným reťazovým narastaním na jednom z primérov v degenerovanom roztoku. Odvodená proteínová sekvencia nezahŕňa klonovaný fragment tu určený ako SEKV. ID. ČÍS. 18. Tieto sekvencie odhaľujú, že plazmid pDAB 2004 obsahuje kompletnú kódovanú oblasť pre peptid TcaC.

Ďalšia analýza SEKV. ID ČÍS. 25 odhalila koniec otvoreného čítacieho rámca (báza 1 až 43), ktorý kóduje posledných 13 aminokyselín TcaA<sub>iii</sub> peptidu, ktorý je tu uvedený ako SEKV. ID. ČÍS. 35. Iba 24 báz oddeľuje koniec TcaA<sub>iii</sub> kódujúcej oblasti a začiatok TcaB<sub>i</sub> kódujúcej oblasti. Medzi 24 bázami sú sekvencie, ktoré môžu slúžiť ako ribozómové viazacie miesto. Hoci je možné, že nie je pravdepodobné, že by génový promótor *Photorhabdus* bol zakódovaný vo vnútri tejto krátkej oblasti. Navrhujeme, aby genómová oblasť tca, ktorá zahŕňa tri dlhé otvorené čítacie rámce [*tcaA* (SEKV. ID. ČÍS. 33), *tcaB* (SEKV. ID. ČÍS. 25, bázy 65 až 36334), a *tcaC* (SEKV. ID. ČÍS. 31), ktorá je oddelená od konca *tcaB* iba 59 bázami] bola regulovaná ako operon s transkripciou začínajúcou proti začiatku *tcaA* génu (SEKV. ID. ČÍS. 33) a výsledkom je polycistronická nosná RNA.

#### Príklad 9

Skríning genómovej knižnice *Photorhabdus* pre gény kódujúce peptid TcbA<sub>ii</sub>

Tento príklad opisuje metódu používanú na identifikáciu klonov DNA, ktoré obsahujú gény kódujúce peptid TcbA<sub>ii</sub>, izoláciu génu a determináciu jeho čiastočnej sekvencie báz DNA.

### Priméry pre reakcie PCR

Polypeptid TcbA<sub>ii</sub> z preparátu účinného na hmyz je okolo 206 kDa. Sekvencia aminokyselín N-konca tohto peptidu je určená ako SEKV. ID. Čís. 1. Štyri zmesi degenerovaných oligonukleotidových primérov ("vodiacich primérov": TH-4, TH-5, TH-6 a TH-7) sa syntetizovali preto, aby kódovali časť tejto aminokyselinovej sekvencie tak, ako je to opísané v príklade 8 a sú zobrazené nižšie.

Tabuľka 12

#### Amino

kyselina	Phe	Ile	Gln	Gly	Tyr	Ser	Asp	Leu	Phe
TH-4	5'-TT(T/C)	ATI	CA(A/G)	CGI	TA(T/C)	TCI	GA(T/C)	CTI	TT-3'
TH-5	5'-TT(T/C)	ATI	CA(A/G)	CGI	TA(T/C)	AG(T/C)	GA(T/C)	CTI	TT-3'
TH-6	5'-TT(T/C)	ATI	CA(A/G)	CGI	TA(T/C)	TCI	GA(T/C)	TT(A/G)	TT-3'
TH-7	5'-TT(T/C)	ATI	CA(A/G)	CGI	TA(T/C)	AG(T/C)	GA(T/C)	TT(A/G)	TT-3'

Tak ako primárna ("a") a sekundárna ("b") sekvencia vnútorného peptidového preparátu (TcbA<sub>ii</sub>-PT81) sa určila a uviedla sa tu ako SEKV. ID. Čís. 23, prípadne ako SEKV. ID. Čís. 24. Štyri zmesi degenerovaných oligonukleotidov ("reverzných primérov": TH-8, TH-9, TH-10 a TH-11) podobne sa odvodili a syntetizovali na kódovanie reverzných komplementov sekvencií, ktoré kódujú časť peptidu SEKV. ID. Čís. 23 tak, ako je znázornené nižšie.

Tabuľka 13

#### Amino

kyselina	Thr	Tyr	Leu	Thr	Ser	Phe	Glu	Gln	Val	Ala	Asn
TH-8	3'-TGI	AT(A/G)	GAI	TGI	AGI	AA(A/G)	CT(T/C)	GT(T/C)	CAI	CGI	TT(G/A)-5'
TH-9	3'-TGI	AT(A/G)	TT(A/G)	TGI	AGI	AA(A/G)	CT(T/C)	GT(T/C)	CAI	CGI	TT(G/A)-5'
TH-10	3'-TGI	AT(A/G)	GAI	TGI	TC(G/A)	AA(A/G)	CT(T/C)	GT(T/C)	CAI	CGI	TT(G/A)-5'
TH-11	3'-TGI	AT(A/G)	TT(A/G)	TGI	TC(G/A)	AA(A/G)	CT(T/C)	GT(T/C)	CAI	CGI	TT(G/A)-5'



Sady týchto primérov sa použili na reakcie PCR na amplifikáciu génových fragmentov kódujúcich  $TcbA_{ii}$  z genómovej DNA *Photorhabdus luminescens* W-14 pripravenej v príklade 6. Všetky reakcie PCR sa robili technikou "horúceho štartu" génmi AmpliWax<sup>TM</sup> a ďalších reagensí a postupov podľa Perkin-Elmera. Obyčajne sa zmes (s celkovým objemom 11  $\mu$ l)  $MgCl_2$ , dNTP, pufrov 10X GeneAmp PCR Buffer II a priméry pridali do skúmaviek obsahujúcich jednotlivé voskové guľičky (10X GenePrep PCR Buffer II je zložený zo 100 mM Tris HCl s pH = 8,3 a 500 mM KCl). Skúmavky sa ohrievali na 80 °C počas 2 minút a mohli sa schladiť. Na povrch parafínovej zátky sa pridal roztok, obsahujúci 10X GeneAmp PCR Buffer II, templát DNA a DNA polymeráza AmpliTaq. Po roztopení voskovej zátky a miešaní zložiek teplotným cyklovaním konečné reakčné podmienky (objem 50  $\mu$ l) boli:  
10mM Tris HCl s pH = 8,3; 50 mM KCl; 2,5 mM  $MgCl_2$ ; 200  $\mu$ M z každej dATP, dCTP, dGTP, dTTP; 1.25 mM zo samostatných zmesí vodiaceho priméru; 1,25  $\mu$ M zo samostatných reverzných primérových zmesí; 1,25 jednotky AmpliTaq DNA polymerázy a 170 ng templátu DNA.

Reaktanty sa vložili do termocyklera (tak, ako je to v príklade 8) a robilo sa podľa nasledovného programu:

Tabuľka 14

Teplota	Čas	Opakovanie cyklu
94 °C	2 minúty	1 x
94 °C	15 sekúnd	
55-65 °C	30 sekúnd	30 x
72 °C	1 minúty	
72 °C	7 minút	3 x
15 °C	stále	

Séria amplifikácií sa robila pri troch rôznych žihacích teplotách (55 °C, 60 °C, 65 °C) degenerovanými primérovými zmesami. Reakcie s temperovaním pri 65 °C nedávali žiadne produkty amplifikácie viditeľné na agarózovej gélovej elektroforéze. Reakcie v temperovacom režime 60 °C a obsahujúce priméry TH-5 + TH-10 tvorili produkt amplifikácie, ktorého mobilita sa zhodovala s 2,9 kbp. Menšie množstvo produktu 2,9 kbp sa tvorilo v týchto podmienkach s priméromi TH-7 + TH-10. Keď sa reakcie temperovali pri 55 °C, tieto uvedené páry primérov tvorili viac 2,9 kbp produktov a tento produkt tvorili aj páry primérov TH-5 + TH-8 a TH-5 + TH-11. Ďalšie veľmi slabé skupiny 2,9 kbp sa objavili v pokusoch obsahujúcich amplifikované produkty z primérových párov TH-7 + TH-8, TH-9, TH-10 alebo TH-11.

Na získanie dostatočného množstva amplifikačného produktu PCR na klonovanie a determináciu sekvencie DNA sa uskutočnilo 10 samostatných reakcií PCR s priméromi TH-5 + TH-10 a robili sa v podmienkach, ktoré sú opísané vyššie s 55 °C žihacou teplotou. Všetky reakcie sa robili v zmesi a produkt 2,9 kbp sa purifikoval

extrakciou na Qiaex z agarózového gélu tak, ako je opísané vyššie

Ďalšie sekvencie, určené pre vnútorné peptidy  $TcbA_{ii}$  sú tu určené ako SEKV. ID. Čís. 21 a SEKV. ID. Čís. 22. Tak, ako predtým sa degenerované oligonukleotidy (reverzné priméry TH-17 a TH-18) pripravovali v súlade s reverzným komplementom sekvencie, ktorá kóduje časť sekvencie aminokyselín týchto peptidov.

Tabuľka 15

Zo SEKV. ID. Čís. 21

Amino

kyselina Met	Glu	Thr	Gln	Asn	Ile	Gln	Glu	Pro
TH-17 3'- TAC	CTT/C	TGI	GTT/C	TTA/G	TAI	GTT/C	GTT/C	GG-5'

Tabuľka 16

Zo SEKV. ID. Čís. 22

Amino

kyselina Asn	Pro	Ile	Asn	Ile	Asn	Thr	Gly	Ile	Asp
TH-18 3'TT(A/G)	GGI	TAI	TT(A/G)	TAI	TT(A?G)	TGI	CCI	TAI	CT(A/G)-5'

Degenerované oligonukleotidy TH-18 a TH-17 sa použili v amplifikačnom experimente s DNA *Photorhabdus luminescens* kmeňa W-14 ako templát a priméry TH-4, TH-5, TH-6 alebo TH-7 ako 5'-priméry. Tieto reakcie amplifikovali produkty s približne 4 kbp a prípadne 4,5 kbp. Tieto DNA sa preniesli z agarózových gélov na nylónové membrány a hybridizovali so značenou sondou  $^{32}P$  (tak, ako je to opísané vyššie), ktorá sa pripravila z 2,9 kbp produktu amplifikovania TH-5 + TH-10 páru primérov. Obidva produkty amplifikácie 4 kbp a 4,5 kbp silne hybridizovali k 2,9 sonde. Tieto výsledky sa použili na zostavenie mapy poradia interných peptidových sekvencií  $TcbA_{ii}$  tak, ako je to znázornené na obr. 3. Približné vzdialenosti medzi primérmami sú znázornené

v nukleotidoch na obr. 3.

Sekvencia DNA 2,9 kbp fragmentu kódujúceho  $TcbA_{ii}$ .

Približne 200 ng purifikovaného fragmentu 2,9 kbp (pripravovaného vyššie) sa vyzrážalo s etanolom a rozpustilo v 17 ml vody. Z toho sa jedna polovica použila ako sekvencovací templát s 25 pmol TH-5 zmesi ako priméry, druhá polovica sa použila ako templát na narastanie TH-10. Sekvencovacie reakcie boli také, ako sú uvedené v príklade 8. Žiadna spoľahlivá sekvencia nevznikala v TH-10 primérovej zmesi. Hoci reakcie s TH-5 primérovou zmesou produkovali sekvencie, ktorých určenie je uvedené nižšie:

```
1  AATCGTGTTG ATCCCTATGC CGNGCCGGGT TCGGTGGAAT CGATGTCCTC ACCGGGGGTT
61  TATTNGAGGG ANTNGTCCCG TGAGGCCAAA AANTGGAATG AAAGAAGTTC AATTTNTTAC
121 CTAGATAAAC GTCGCCCGGN TTTAGAAAGN TTANTGNTCA GCCAGAAAAT TTTGGTTGAG
181 GAAATTCCAC CGNTGGTTCT CTCTATTGAT TNGGGCCTGG CCGGGTTCGA ANNAAAACNA
241 GGAAATNCAC AAGTTGAGGT GATGGNTTTG TNGCNANCTT NTCGTTTAGG TGGGGAGAAA
301 CCTTNTCANC ACGNTTNTGA AACTGTCCGG GAAATCGTCC ATGANCGTGA NCCAGGNTTN
361 CGCCATTGG
```

Na základe týchto sekvencií, sekvencovací primér (TH-21, 5' - CCGGGCGACGTTTATCTAGG - 3') sa určil k reverzným komplementárnym bázam 120 - 139 a inicioval polymerizáciou smerom k 5'-koncu (napríklad TH-5 koniec) gélom purifikovaného 2,9 kbp PCR fragmentu kódujúceho  $TcbA_{ii}$ . Určená sekvencia je uvedená nižšie a je porovnávaná s biochemicky determinovaným N-koncovým peptidom sekvencie  $TcbA_{ii}$  SEKV. ID. ČÍS. 1.

Potvrdenie 2,9 kbp PCR fragmentu sekvencie  $TcbA_{ii}$   
(Podčiarknuté aminokyseliny - kódované degenerovanými oligonukleotidmi).

SEKV.ID.ČÍS.1	F	I	Q	G	Y	S	D	L	F	G	-	-	A	
2,9 kbp sekv.	GC	ATG	CAG	GGG	TAT	AGT	GAC	CTG	TTT	GGT	AAT	CGT	GCT	
			M	Q	G	Y	S	D	L	F	G	N	R	A

Od homológie odvodenej sekvencie aminokyselín k jej biochemickej determinácii je jasné, že 2,9 kbp PCR fragment predstavuje kódovaciu oblasť *TcbA*. Tento fragment 2,9 kbp sa potom použil ako hybridizačná sonda na opis genómovej knižnice *Photorhabdus* W-14, pripravenej v príklade 8 pre kozmidy obsahujúce gén kódujúci *TcbA<sub>ii</sub>*.

#### Skríning *Photorhabdus* kozmidovej knižnice

2,9 kb gélom purifikovaný PCR fragment sa označil <sup>32</sup>P pomocou značkovacej súpravy Boehringer Mannheim High Prime tak, ako sa opísalo v príklade 8. Filtre obsahujúce zvyšky asi 800 kolónií z kozmidovej knižnice sa podrobilo skríningu tak, ako je to opísané predtým (príklad 8) a pozitívne klony sa rozotrelí na izoláciu kolónií a opätovný skríning. Tri klony (8A11, 25G8 a 26D1) dali pozitívne výsledky prostredníctvom niekoľkých skríningových a charakterizačných krokov. Žiadna hybridizácia sa nepozorovala v špecifickej sonde *TcbA<sub>ii</sub>* so štyrmi kozmidmi identifikovanými v príklade 8, ktoré nesú gény *tcaB* a *tcaC*. DNA kozmidov 8A11, 25G8 a 26D1 sa štiepila s reštrikčnými enzýmami *Bgl* 2, *EcoR* 1 alebo *Hind* 3 (buď samostatne alebo v kombinácii jeden s druhým) a fragmenty sa upravovali na agarózovom géli a preniesli sa na nylónovú membránu tak, ako sa opísalo v príklade 8. Membrána sa hybridizovala so značenou sondou <sup>32</sup>P pripravenou z fragmentu 4,5 kbp (vzniknutá amplifikáciou *Photorhabdus* genómovej DNA s primérmí TH-5 + TH-17). Vzorky vzniknuté z kozmidových DNA 8A11 a 26D1 boli identické k tým, ktoré vznikli odstrihnutím genómovej DNA na tej istej membráne. Z toho vyplýva záver, že kozmidy 8A11 a 26D1 sú presné zobrazenia lokusu kódujúceho genóm *TcbA<sub>ii</sub>*. Avšak kozmid 25G8 má jednoduchý fragment *Bgl* II, ktorý je trocha väčší než genómová DNA. Toto môže byť výsledkom polohovania inzertu vnútri vektora.

DNA sekvencia génu kódujúceho *tcba*.

Analýza hybridizácie kozmidu 26D1 na membráne odhalila, že sonda 4,4 kbp hybridizovala k jednoduchému veľkemu fragmentu *EcoR* 1 (väčší ako 9,0 kbp). Tento fragment bol purifikovaný gélom a ligátovaný do *EcoR* I miesta pBC KS (+) tak, ako sa opísalo v príklade 8 pre vznik plazmidu pBC-S1/R1. Sekvencia čiastkovej inzertovanej DNA tohto plazmidu sa determinovala "primérovým krokováním" z bočnej vektorovej sekvencie použitím postupov opísaných v príklade 8. Ďalšie sekvencie vznikali extenziou z nových oligonukleotidov určených z predchádzajúcej determinovanej sekvencie. Keď sa porovnávala determinovaná sekvencia DNA pre gén *tcba* identifikovanej rôznymi metódami (určené ako SEKV. ID. Čís. 11, tak ako sa opísalo v príklade 12 nižšie) úplná homológia sa objavila v nukleotidoch 1-272, 319-826, 2578-3068 a 3068-3540 (celkove 1712 báz). Z tohto je možné vytvoriť záver, že oba spôsoby sa môžu použiť na identifikáciu fragmentov DNA kódujúcich peptid  $TcbA_{ii}$ .

Analýza odvodenej aminokyselinovej sekvencie génu *tcba*.

Sekvencia fragmentu DNA identifikovaného ako SEKV. ID. Čís. 11 kóduje proteín, ktorého odvodená sekvencia aminokyselín je tu určená ako SEKV. ID. Čís. 12. Identitu génu, ktorý kóduje proteín  $TcbA_{ii}$  overuje niekoľko vlastností. N-koncový peptid  $TcbA_{ii}$  (SEKV. ID. Čís. 1; Phe Ile Gln Gly Tyr Ser Asp Leu Phe Gly Asn Arg Ala) je kódovaný tak ako aminokyseliny 88 až 100. Interný peptid  $TcbA_{ii}$   $TcbA_{ii}$ -PT81(a) (SEKV. ID. Čís. 23) je kódovaný tak, ako aminokyseliny 1065 až 1077 a  $TcbA_{ii}$ -PT81(b) (SEKV. ID. Čís. 24) je kódovaný ako aminokyseliny 1571 až 1592. Ďalej, vnútorný peptid  $TcbA_{ii}$ -PT56 (SEKV. ID. Čís. 22) je kódovaný tak ako aminokyseliny 1474 až 1488 a interný peptid  $TcbA_{ii}$ -PT103 (SEKV. ID. Čís. 21) je kódovaný ako aminokyseliny 1614 až 1639. Je bežné, že tento gén je autentickým klonom kódujúcim peptid  $TcbA_{ii}$  tak, ako izolát z preparátov proteínových insekticídov *Photobacterium luminescens* kmeňa W-14.

Proteín izolovaný ako peptid  $TcbA_{ii}$  je odvodený zo štiepenia

dlhšieho peptidu. Dôkazom toho je získaný fakt, že nukleotidy kódujúce N-koncový peptid TcbA<sub>ii</sub> SEKV. ID. ČÍS. 1 idú za 261 bázami (kódujúce 81 N-koncových proximálnych aminokyselín) dlhšieho otvoreného čítacieho rámca (SEKV. ID. ČÍS. 11). Tento čítací rámec začína nukleotidmi kódujúcimi sekvenciu aminokyselín Met, Gln, Asn, Ser, Leu, ktoré sa zhodujú s N-koncovou sekvenciou veľkého peptidu TcbA a je tu určený v SEKV. ID. ČÍS. 16. Predpokladá sa, že TcbA je prekursorový proteín pre TcbA<sub>ii</sub>.

#### Príbuznosť génov *tcbA*, *tcaB* a *tcaC*

Gény *tcaB* a *tcaC* sú viazané blízko a môžu sa transkriptovať tak, ako jednoduchá mRNA (príklad 8). Gén *tcbA* vznikol na kozmidoch, ktoré sa očividne neprekrývajú s tým, ktorý obsahuje *tcaB* a *tcaC* klastre, pretože príslušná genómová knižnica opisuje identifikáciu rozdielnych kozmidov. Avšak porovnanie sekvencie aminokyselín kódovanej génmi *tcaB* a *tcaC* s génom *tcbA* odhaľuje podstatný stupeň homológie. Zachovanie aminokyselín (softvérová analýza sekvencií Protein Alignment Mode of MacVector™, odčítanie matrice pam250, delená hodnota = 2, Oxford Molecular Group, Campbell, CA) je znázornené na obr. 4. V odčítavacom riadku každej skupiny na obr. 4 horné vokáne (˘) indikujú zmeny v homológii alebo zmeny zachovávaní aminokyselín a dolné vokáne (v) indikujú nehomológiu.

Táto analýza ukazuje, že sekvencia aminokyselín peptidu TcbA z rezíduí 1739 - 1894 je vysoko homologická k aminokyselinám 441 - 603 peptidu TcaB<sub>i</sub> (162 z celkove 627 aminokyselín TcaB; SEKV. ID. ČÍS. 28). Tak ako aj sekvencia aminokyselín TcbA 1932 - 2459 je vysoko homologická k aminokyselinám 12 až 531 peptidu TcaB<sub>ii</sub> (520 z celkove 562 aminokyselín; SEKV. ID. ČÍS. 30). Vzhľadom na to, že peptid TcbA (SEKV. ID. ČÍS. 12) obsahuje 2505 aminokyselín a z toho celkove 684 aminokyselín (27 %) pri C-proximálnom konci peptidu, je homologických k TcaB<sub>i</sub> alebo TcaB<sub>ii</sub>, homológie sú usporiadané kolineárne k usporiadaniu myslených proteínov TcaB (SEKV. ID. ČÍS. 26). Veľkostná medzera v homológii TcbA sa zhoduje so spojením medzi TcaB<sub>i</sub> a TcaB<sub>ii</sub>, častí preproteínu TcaB.

Produkty génov TcbA a TcaB sa zjavne evolučne dotýkajú a to predpokladá, že sa podielajú nejakou spoločnou (-ými) funkciou (-ami) v kmeni *Photorhabdus*.

#### Príklad 10

Charakteristika zinkových metaloproteáz v bujóne *Photorhabdus*:  
Inhibícia proteázy, klasifikácia a purifikácia

Inhibícia proteázy a klasifikačné testy: proteázové testy sa uskutočňovali použitím kazeínu FITC ako substrátu rozpusteného vo vode (konečná testovacia koncentrácia 0,08 %). Proteolytické reakcie sa uskutočnili pri 25 °C 1 hod. vo vhodnom pufri s 25 µl bujónu *Photorhabdus* (150 µl celkového objemu reakcie). Vzorky sa tiež testovali v prítomnosti a absencii ditiotreitolu. Po inkubácii sa pridali rovnaké objemy 12 % kyseliny trichlóroctovej na vyzrážanie nerozštiepeného proteínu. Nasledovalo vyzrážavanie 0,5 hod. a potom centrifugácia. 100 µl supernatantu sa umiestnilo na 96-jamkovú mikrotitračnú platňu a pH roztoku sa upravilo pridaním rovnakého objemu 4 N NaOH. Proteolýza sa potom kvantifikovala fluorimetrickým čítačom mikrotitračných platní Fluoroskan II pri excitácii a emisii vlnovej dĺžky 485, prípadne 538 nm. Aktivita proteázy sa testovala v rozpätí pH = 5,0 až 10,0, so zvyšovaním o 0,5 jednotky. Použili sa nasledovné pufre na konečnú koncentráciu octanu sodného 50 mM (pH = 5,0 až 6,5); Tris HCl (pH = 7,0 až 8,0) a bis-Tris propánu (pH = 8,5 až 10,0). Na identifikáciu pozorovaných tried proteáz sa upravil surový bujón s rôznymi proteázovými inhibítormi (0,5 µg/ml konečnej koncentrácie) a potom sa hodnotila aktivita proteázy pri pH = 8,0 na substráte, ktorý je opísaný vyššie. Používali sa proteázové inhibítory vrátane E-64 (L-transepoxysakcinylleucil-amido-[4-guanidino]bután), 3,4-dichlór-izokumarín, leupeptín, pepstatín, amastatín, kyselina etyléndiamíntetraoctová (EDTA) a 1,10-fenantrolín.

Proteázové testy uskutočňované v rozpätí pH odhalili, že



proteáza (-y), ktorá bola skutočne prítomná preukázala (-i) maximálnu aktivitu pri okolo pH = 8,0 (tabuľka 17). Pridanie DTT nemalo na aktivitu proteáz žiaden účinok. Surový bujón sa potom upravil rôznymi proteázovými inhibítormi (tabuľka 18). Styk surového bujónu s inhibítormi, ktoré sú opísané vyššie, odhalil, že 1,10-fenantrolín zapríčinil úplnú inhibíciu celej proteázovej aktivity, pokiaľ sa pridal v konečnej koncentrácii 50  $\mu\text{g}$  s  $\text{IC}_{50} = 5 \mu\text{g}$  v 100  $\mu\text{l}$  z 2 ng/ml surového bujónového roztoku. Tieto dáta poukazujú, že najčastejšie sa vyskytujúca proteáza (-y), nachádzajúca (-e) sa v bujónovej kultúre *Photobacterium luminescens*, je z triedy enzýmov zinkových metaloproteáz.

Tabuľka 17

Účinok pH na aktivitu proteáz v prvý deň produkcie *Photobacterium luminescens* (kmeň W-14).

pH	Flu. jedn <sup>a</sup> aktivita <sup>b</sup>	Percento
5,0	3013 $\pm$ 78	17
5,5	7994 $\pm$ 448	45
6,0	12 965 $\pm$ 483	74
6,5	14 390 $\pm$ 1291	82
7,0	14 386 $\pm$ 1287	82
7,5	14 135 $\pm$ 198	80
8,0	17 582 $\pm$ 831	100
8,5	16 183 $\pm$ 953	92
9,0	16 795 $\pm$ 760	96
9,5	16 279 $\pm$ 1022	93
10,0	15 225 $\pm$ 210	87

a. Flu. jedn. = fluorescenčná jednotka (maximum okolo 28 000; pozadie okolo 2 200).

b. Percento aktivity v pomere k maximu pri pH = 8,0

Tabuľka 18

Účinok rôznych proteázových inhibítorov na aktivitu proteáz pri pH = 8,0, zistené v prvom dni produkcie *Photobacterium luminescens* (kmeňa W-14).

Inhibítor	Korigované Flu. jedn. <sup>a</sup>	per cento inhibície <sup>b</sup>
Kontrola	13 053	0
E - 64	14 259	0
1,10-fenentrolin <sup>c</sup>	15	99
3,4-dichloroizokumarín <sup>d</sup>	7 956	39
Leupeptín	13 074	0
Pepstatín <sup>c</sup>	13 441	0
Amastatín	12 474	4
DMSO kontrola	12 005	8
Metanolová kontrola	12 125	7

a. Korigované Flu.jedn. = Fluorescenčné jednotky - pozadie (2200 flu. jedn.).

b. Per cento inhibície v pomere k proteázovej aktivite pri pH = 8,0.

c. Inhibítory boli rozpustené v metanole.

d. Inhibítory boli rozpustené v DMSO.

Izolácia zinkových metaloproteáz sa uskutočnila aplikáciou dialýzy 10 až 80 % guľičiek síranu amónneho na kolóne Q Sepharose kondiciovanej 50 mM Na<sub>2</sub>PO<sub>4</sub> pH = 7,0 tak, ako je to opísané v príklade 5 pre toxín *Photobacterium luminescens*. Po extenzívnom premytí sa použil 0 až 0,5 M NaCl gradient na elúciu toxínového proteínu. Väčšina biologickej aktivity a proteín sa eluoval z 0,15 až 0,45 M NaCl. Aj tak sa pozorovalo, že väčšina proteolytickej aktivity sa nachádzala vo frakcii 0,25 až 0,35 M NaCl s nejakou aktivitou frakcie 0,15 až 0,25 M NaCl. Analýza SDS-PAGE frakcie 0,25 až 0,35 M NaCl preukázala väčšiu peptidovú skupinu s približne 60

kDa. Frakcia 0,15 až 0,25 M NaCl obsahovala podobnú skupinu 60 kDa, ale pri nižšej relatívnej koncentrácii proteínu. Subsekvantná gélová filtrácia tejto frakcie na kolóne Superose 12 HR 16/50 mala za následok veľký pík migrujúci pri 57,5 kDa, ktorý obsahoval predominantnú (> 90 % celkového farbeného proteínu) skupinu 58,5 kDa zistenú analýzou SDS-PAGE. Doplnková analýza tejto frakcie používajúca rôzne proteázové inhibítory tak, ako je to opísané vyššie, determinovala, že proteáza bola zinková metaloproteáza. Skoro všetky z proteázových aktivít prítomných v bujóne *Photorhabdus* po 1. dni fermentácie sa zhodujú s okolo 58 kDa zinkovej metaloproteázy.

V nasledujúcej druhej izolácii zinkovej metaloproteázy (-áz), kmeň W-14 bujónovej kultúry *Photorhabdus luminescens* rástol tri dni a proteázová aktivita sa zviditeľnila pomocou dodecylsírano-polyakrylamidovou gélovou elektroforézou (SDS-PAGE) obohatenou želatínou tak, ako opísal Schmidt, T.M., Bleakley, B. a Neelson, K.M., 1988. Nosný gél SDS (5,5 x 8 cm) sa pripravil s 12,5 % polyakrylamidu (40 % zásobený roztok akrylamid/bis-akrylamid; Sigma Chemical Co., St. Louis, MO), do ktorého sa zaviedla konečná koncentrácia želatíny 0,1 % (Biorad EIA Grade Reagent, Richmond, CA) pri rozpúšťaní vo vode. Zmesné gély SDS (1,0 x 8 cm) sa pripravili s 5 % polyakrylamidom obohateným s 0,1 % želatínou. Obyčajne sa zriedilo 2,5 µg proteínu, ktorý sa testoval v 0,03 ml nosného pufri SDS-PAGE bez ditiotreitolu (DTT) a naniesol sa na gél. Proteíny sa podrobovali elektroforéze v nosnom pufri SDS (Laemmli, U.K., 1970. Nature 227, 680) pri 0 °C a 8 mA. Po ukončení elektroforézy sa gél premýval 2 hod. v 2,5 % (obj./obj.) tritone X - 100. Gély sa inkubovali 1 hod pri 37 °C v 0,1 M glycíne (pH = 8,0). Po inkubácii sa gély fixovali a farbili počas noci 0,1 % amidovou čerňou vo vode s metanolom a kyselinou octovou (30 : 10 : 60, obj./obj./obj., Sigma Chemical Co). Aktivita proteáz sa zviditeľnila ako svetlá plocha oproti tmavému, amidovou čerňou zafarbenému pozadiu. Bola zapríčinená proteolýzou a subsekvantnou difúziou zo zavedenej želatíny. Nakoniec sa objavili tri odlišné

skupiny produkujúce proteolytickú aktivitu pri 58, 41 a 38 kDa.

Testy sledujúce aktivitu rôznych proteáz v trojdennej bujónovej kultúre W-14 sa uskutočnili pomocou kazeínu FITC rozpusteného vo vode, ako substrátu (konečná testovaná koncentrácia 0,02). Proteolyzačné experimenty sa uskutočnili pri 37 °C 0 až 0,5 hod v 0,1 M Tris HCl (pH = 8,0) s rôznymi proteínovými frakciami v celkovom objeme 0,15 ml. Reakcie sa ukončili pridaním rovnakého objemu 12 % kyseliny trichlóroctovej (TCA) rozpustenej vo vode. Po inkubácii pri izbovej teplote 0,25 hod. sa vzorky centrifugovali pri 10 000 x g počas 0,25 hod. a odoberal sa objem 0,10 ml, ktorý sa rozpipetoval na 96-jamkovú mikrotitračnú platňu. Roztok sa potom neutralizoval pridaním rovnakého objemu 2 N hydroxidu sodného a nasledovala kvantifikácia na fluorimetrickej čítači mikrotitračných platní, Fluoriskane II s excitáciou a emisiou vlnových dĺžok 485, prípadne 538 nm. Meranie aktivity sa uskutočňovalo pomocou kazeínu FITC s rôznymi koncentraciami proteázy pri 37 °C počas 1 až 10 min. Jednotka aktivity sa definovala nezávisle, ako množstvo enzýmu potrebné na produkciu 1 000 fluorescenčných jednotiek za min., a špecifická aktivita sa definovala ako jednotka/mg proteázy.

Štúdie týkajúce sa inhibície sa uskutočňovali s dvoma zinkovými metaloproteázovými inhibítormi: 1,10-fenentrolínom a N-( $\alpha$ -ramnopyranoziloxihydroxyfosfinyl)-Leu-Trp(fosforamidónom). Zásobné roztoky inhibítorov boli rozpustené v 100 % etanole, prípadne vo vode. Koncentrácie zásobných roztokov boli obyčajne 10 mg/ml a 5 mg/ml pre 1,10-fenantrolín, respektíve pre fosforamidón s konečnými koncentraciami inhibítora 0,5 až 1,0 mg/ml na reakciu. Styk trojdňového surového bujónu W-14 s 1,10-fenentrolínom, inhibítorom všetkých zinkových metaloproteáz, zapríčinil úplnú elimináciu celej proteázovej aktivity, zatiaľ čo styk s fosforamidónom, inhibítorom podobným termolýznym proteázam (Weaver, L.H., Kester, W.R a Matthews, B.W., 1977, J. Mol. Biol. 114, 119 - 132), zapríčinil okolo 56 %

redukciu proteázovej aktivity. Zostatková proteolytická aktivita sa nedala ďalej redukovať ďalším pridávaním fosforamidónu.

Proteázy trojdennej bujónovej kultúry *Photorhabdus* kmeňa W-14 sa purifikovali nasledovne: 4,0 litre bujónu sa koncentrovali na špirálovej ultrafiltračnej kazete Amicon typu S1Y100, ktorá bola upevnená na filtračnom zariadení Amicon M-12. Pretekajúci materiál obsahujúci menej ako 100 kDa natívnych proteínov v objeme (3,8 l) sa koncentroval na 0,375 l na špirálovej ultrafiltračnej kazete Amicon typu S1Y100, ktorá bola upevnená na filtračnom zariadení Amicon M-12. Materiál retentátu obsahoval proteíny pohybujúce sa vo veľkostnom rozmedzí 10 až 100 kDa. Tento materiál sa preniesol na kolónu Pharmacia HR 16/10, ktorá bola naplnená silnou anionovýmennou náplňou PerSeptive Biosystem (Framington, MA) Poros<sup>R</sup> 50 HQ, ktorá sa kondicionovala v 10 mM pufri fosforečnanu sodného (pH = 7,0). Proteíny sa naniesli na kolónu pri rýchlosti toku 5 ml/min. Nasledovalo premytie nenaviazaného proteínu pufrom až do  $A_{280} = 0,00$ . Potom sa proteíny eluovali gradientom NaCl od 0 až 1,0 M NaCl do 40 min. pri prietoku 7,5 ml/min. Frakcie sa testovali na proteázovú aktivitu, vyššie, a aktívne frakcie sa zbierali. Proteolyticky aktívne frakcie sa riedili na 50 % (obj./obj.) 10 mM pufrom fosforečnanu sodného (pH = 7,0) a naniesli sa na kolónu Pharmacia HR 10/10 mono Q kondicionovanú 10 mM fosforečnanom sodným. Po premytí kolóny pufrom až do  $A_{280} = 0,00$  sa proteíny eluovali gradientom NaCl od 0 až 0,5 M NaCl 1 hodinu pri prietoku 2 ml/min. Frakcie sa testovali z hľadiska proteázovej aktivity. Tieto proteázy mali najvyššie množstvo fosforamidón-senzitívnej proteázovej aktivity, pričom fosforamidónovo senzitivna aktivita pochádzajúca od 41/38 kDa, proteázy, nižšie, sa zbierala. Zistilo sa, že tieto frakcie sa eluovali v rozmedzí od 0,15 až 0,25 M NaCl. Frakcie obsahujúce predomnanciu na fosforamidón necitlivú proteázovú aktivitu, 58 kDa proteáza, sa tiež zbierali. Zistilo sa, že tieto frakcie sa eluovali v rozmedzí od 0,25 až 0,35 M NaCl. Fosforamidón-senzitívne proteázové frakcie sa potom koncentrovali do konečného objemu 0,75 ml na Millipore Ultrafree-15

centrifugačnom filtračnom zariadení membráne Biomax - 5K NMWL. Tento materiál sa aplikoval pri prietoku 0,5 ml/min. na kolónu Pharmacia MR 10130, ktorá bola naplnená s Pharmacia Sephadex G-50 kondiciovaná v 10 mM pufri fosforečnanu sodného. Frakcie, ktoré majú maximálnu fosforamidón-senzitívnu proteázovú aktivitu boli potom centrifugované v Millipore Ultrafree-15 centrifugačnom filtračnom zariadení na membráne Biomax - 5K NMWL. Analýza proteolytickej aktivity, vyššie, indikovala, že tento materiál má len fosforamidón-senzitívnu proteázovú aktivitu. Zozbieranie fosforamidón-nesenzitívnej proteázy, proteín 58 kDa, pokračoval koncentráciou v Millipore Ultrafree-15 centrifugačnom filtračnom zariadení na membráne Biomax - 5K NMWL a ďalej separáciou na kolóne Pharmacia Superdex - 75. Frakcie obsahujúce proteázu sa zbierali.

Analýzy purifikovaných 58 kDa a 41/38 kDa proteáz odhalili, že hoci obidva typy proteáz boli úplne inhibované 1,10-fenantrolínom, len proteáza 41/38 kDa bola inhibovaná aj fosforamidónom. Ďalšie analýzy surového bujónu indikovali, že aktivita proteázy jednodňovej bujónovej kultúry W-14 mala 23 % z celkovej proteázovej aktivity pochádzajúcej z proteázy 41/38 kDa a narastá až na 44 % v treťom dni bujónovej kultúry W-14.

Štandardná SDS-PAGE analýza pre vyhodnotenie čistoty proteínu a získania amino-koncovej sekvencie sa robila pomocou 4 až 20 % gradientu MiniPlus SeptraGels zakúpeného u Integrated Separation Systems (Natick, MA). Proteíny, ktoré sú amino-koncovo sekvencované, sa farbili na membráne PVDF s následnou purifikáciou, nižšie, (ProBlott Membrány; Applied Biosystems, Forest city, CA), zviditeľnili 0,1 % amidočernou, odobrali sa a poslali na Cambridge Prochem; Cambridge, MA, na sekvencovanie.

Odvodená amino-koncová sekvencia proteáz 58 kDa (SEKV. ID. ČÍS. 45) a 41/38 kDa (SEKV. ID. ČÍS. 44) z tri dni starej bujónovej kultúry W-14 bola DV-GSEKANEKLLK (SEKV. ID. ČÍS. 45)

a prípadne DSGDDDKVTNTDIHR (SEKV. ID. ČÍS. 44).

Sekvencovanie proteázy 41/38 kDa odhalilo niekoľko amino-koncov, z ktorých každý mal ďalšiu aminokyselinu odstránenú v rámci proteolýzy. Hodnotenie primárnej, sekundárnej, terciárnej a kvartérnej sekvencie polypeptidov 38 a 41 kDa umožnilo dedukovať sekvenciu, ktorá je uvedená vyššie a odhalila, že tieto dve proteázy sú homologické.

Príklad 11, Časť A

Skríning genómovej knižnice *Photobacterium* prostredníctvom využitia protilátok pre gény kódujúce TcbA peptid

Tak ako je sekvencovanie opísané vyššie, vhodné sondovanie a sekvencovanie sa uskutočnilo na základe peptidu TcbA<sub>ij</sub> (SEKV. ID. ČÍS. 1). Toto sekvencovanie sa robilo prípravou bujónovej kultúry baktérií a purifikáciou toxínu tak, ako sa opisuje v príkladoch 1 a 2 vyššie.

Genómová DNA sa izolovala z druhu *Photobacterium luminescens*, kmeňa W-14 rastúceho na Graceho médiu hmyzej tkanivovej kultúry. Baktérie rástli v 5 ml kultivačného média v 250 ml Erlenmeyerových bankách pri 28 °C 250 rpm počas približne 24 hodín. Bakteriálne bunky zo 100 ml kultivačného média sa centrifugovali pri 5 000 x g počas 10 minút. Supernatant sa vylial a bunkový sediment sa potom použil na izoláciu genómovej DNA.

Genómová DNA sa izolovala pomocou modifikovanej metódy CTAB opísanej v časti 2.4.3. Ausubel (vyššie). Časť nazvaná "Široká škála CsCl príprav bakteriálnej genómovej DNA" sa používala v kroku 6. V tomto bode sa uskutočňovala extrakcia pridaním chloroform/izoamylalkoholu (24 : 1). Nasledoval krok extrakcie fenol/chloroform/izoamylom (25 : 24 : 1) a nakoniec extrakcia chloroform/izoamyl/alkoholom (24 : 1). DNA sa vyzrážala pridaním

0,6 objemu propanolu. Vyzrážaná DNA sa zachytila a navinula na koniec šikmej sklenenej tyčinky, potom sa krátko ponorila do 70 % etanolu ako posledného premytia a rozpustila sa v 3 ml pufru TE.

Koncentrácia DNA, hodnotená optickou hustotou pri 280/260 nm, bola približne 2mg/ml.

Knižnica sa pripravovala z tejto genómovej DNA. Približne 50 µg genómovej DNA sa čiastočne štiepilo so Sau3 Al. Potom sa použila centrifugácia v hustotnom gradiente NaCl na veľkostnú frakcionáciu čiastočne štiepených fragmentov DNA. Frakcie obsahujúce fragmenty DNA s priemernou veľkosťou 12 kb alebo väčšie tak, ako určila agarová gélová elektroforéza, sa ligátovali do plazmidu BluScript, Stratagene, La Jolla, California a transformovali do *E.coli*, kmeňa DH5α alebo DHB10.

Samostatne sa purifikované podiely proteínu poslali do centra biotechnológie hybridov na Univesity of Wisconsin, Madison, na to, aby sa vyprodukovali monoklonálne protilátky k proteínom. Materiál, ktorý sa poslal, bola purifikovaná frakcia na HPLC obsahujúca natívne skupiny 1 a 2, ktoré boli denaturované pri 65 °C a z ktorých sa 20 µg injektovalo do každej zo štyroch myší. Stabilné hybridy bunkových línií produkujúcich monoklonálne protilátky sa získali po tom, ako bunky zo sleziny neimunizovaných myší fúzovali so stabilnou líniou svalových buniek. Monoklonálne protilátky sa získali z hybridov.

Polyklonálne protilátky sa vytvorili samostatne z natívnej na agarózovom géli purifikovanej skupiny proteínov 1 (pozri príklad 1), ktoré sa potom použili na imunizáciu novozélandských bielych králikov. Proteín sa pripravil odobratím skupiny z natívnych agarózových gélov. Kúsky gélu sa krátko ohriali na 65 °C, aby sa rozpustila agaróza a hneď sa emulgovali s adjuvans. Freundov úplny adjuvans sa použil na primárne imunizácie a Freundov neúplny adjuvans sa použil na 3 doplnkové injekcie



v mesačných intervaloch. Každá injekcia, s približne 0,2 ml emulgovanej skupiny 1, obsahujúcej 50 až 100 mikrogramov proteínov, sa podala mnohonásobnými injekciami do chrbta králika. Sérum sa získalo po 10 dňoch od poslednej injekcie a ďalších odberov krvi v týždňových intervaloch počas troch týždňov. Sérová zložka sa inaktivovala ohrevom na 56 °C počas 15 minút a potom sa uskladnila pri -20 °C.

Monoklonálne a polyklonálne protilátky sa potom použili na opis genómovej knižnice na expresiu antigénov, ktoré by mohli byť detegované epitopom. Pozitívne klony sa detegovali na vzostupných nitrocelulóзовých filtračných kolónach. Urobila sa imunoblotová analýza pozitívnych klonov.

Z analýzy klonov tak, ako sa definovali pri obidvoch imunoblotových reakciách a z Southernovej analýzy vyplynuli v pokusných identifikáciách štyri genómové oblasti.

V prvej oblasti sa kódovanie génov peptidov určilo ako  $TcbA_{ii}$ . Získala sa úplná sekvencia DNA tohto génu (*tcbA*). Je to sada SEKV. ID. ČÍS. 11. Dôkaz, že sekvencia kóduje vnútornú sekvenciu SEKV. ID. ČÍS. 1 demonštruje prítomnosť SEKV. ID. ČÍS. 1 pri aminokyseline číslo 88 z odvodenej sekvencie aminokyselín vzniknutej v otvorenom čítacom rámci SEKV. ID. ČÍS. 11. Toto sa môže potvrdiť odvolaním sa na SEKV. ID. ČÍS. 12, ktorá je odvodenou sekvenciou aminokyselín vytvorenou SEKV. ID. ČÍS. 11.

Druhá oblasť toxínových peptidov obsahuje segmenty súvisiace s vyššie uvedenými  $TcaB_i$ ,  $TcaB_{ii}$  ako aj  $TcaC$ . Nasledujúci skrining knižnice s polyklonálnymi antisérmi, tejto druhej oblasti toxínových génov sa identifikoval niekoľkými klonmi, ktoré produkovali rôzne veľké proteíny, z ktorých všetky podliehali krížovej reakcii s polyklonálnymi protilátkami na imunoblote a tiež sa podieľali na homológii na Southernovom prenose. Porovnanie sekvencie odhalilo, že táto prináleží ku génovému komplexu určovanému vyššie ako *TcaB* a *TcaC*.

Dve ďalšie oblasti protilátok toxínových klonov sa tiež izolovali v polyklonálnom triedení. Tieto oblasti produkovali proteíny, ktoré krížovo reagovali s polyklonálnou protilátkou a tiež sa podieľali na homológii DNA s oblasťami určenými Southernovým prenosom. Toto poukazuje na to, že extracelulárne proteínové gény *P. luminescens* predstavujú rodinu génov, ktoré evolučne spolu súvisia.

Ďalšie sledovania súvislostí, že by mohli byť evolučne súvisiace rozdielnosti v toxínových peptidoch, ktoré obsahujú tieto organizmy, sa vykonávali dvoma krokmi na vyšetrenie ďalších kmeňov *P. luminescens* z hľadiska prítomnosti spomínaných proteínov. Obidva sa robili PCR amplifikáciou genómovej DNA a imunoblotovou analýzou použitím polyklonálnych a monoklonálnych protilátok.

Výsledky ukazujú, že príbuzné proteíny sú produkované *P. luminescens*, kmeňmi WX-2, WX-3, WX-4, WX-5, WX-6, WX-7, WX-8, WX-11, WX-12, WX-15 a W-14.

#### Príklad 11, Časť B

#### Sekvencie a analýza toxínových klonov *tcc*

Ďalšie sekvencovanie DNA sa uskutočnilo na plazmidoch izolovaných z klonov *E. coli*, opísaných v príklade 11, časť A. Nukleotidová sekvencia z tretej oblasti *E. coli* klonov ukázala tri ostré väzbové otvorené čítacie rámce na tomto genómovom lokuse. Tento lokus sa označoval *tcc* s tromi otvorenými čítacími rámcami označovanými *tccA* SEKV. ID. ČÍS. 56, *tccB* SEKV. ID. ČÍS. 58 a *tccC* SEKV. ID. ČÍS. 60. Blízke spojenie medzi týmito otvorenými čítacími rámcami bolo zistené skúškou SEKV. ID. ČÍS. 56, v ktorej 93 bp oddeľuje koncový kodón *tccA* od počiatočného kodónu *tccB* (bázy 2992 až 2994 SEKV. ID. ČÍS. 56) a skúškou SEKV. ID. ČÍS. 58, v ktorej 131 báz oddeľuje koncový kodón *tccB* od počiatočného kodónu *tccC* (bázy 4930 až 4932 SEKV. ID. ČÍS. 58).

Fyzická mapa je ukázaná na obr. 6B.

Odvodzovaná aminokyselina z otvoreného čítacieho rámca *tccA* poukazuje na gény kódujúce proteín 105 459 Da. Tento proteín sa označoval *TccA* (SEQ ID. ČÍS. 57). Prvých 12 aminokyselín tohto proteínu zodpovedá N-koncovkej sekvencii získanej z proteínu 108 kDa, SEKV. ID. ČÍS. 8, identifikovanej predtým ako časť toxínového komplexu.

Odvodzovaná aminokyselina z *tccB* otvoreného čítacieho rámca ukazuje, že gén kóduje proteín s 175 716 Da. Tento proteín sa označoval *TccB* (SEKV. ID. ČÍS. 59). Prvých 11 aminokyselín tohto proteínu zodpovedá N-koncovkej sekvencii získanej z proteínu s odhadovanou molekulovou hmotnosťou 185 kDa, SEKV. ID. ČÍS. 7. Podobnosté analýzy odhalili, že *TccB* proteín je vo vzťahu k proteínom identifikovaným ako *TcbA* SEQ. ID. ČÍS. 12; 37 % podobnosti 27 % identity, *TcdA* SEQ. ID. ČÍS. 47, 35 % podobnosti a 28 % identity, a *TcaB* SEQ. ID. ČÍS. 26; 32 % podobnosti a 26 % identity (s použitím GAP algoritmu Wisconsin Package version 9.0, Genetic Computer Group (GCG) Madison Wisconsin).

Odvodzovaná aminokyselinová sekvencia z *tccc* ukazuje, že tento otvorený čítací rámec kóduje proteín s 111 694 Da a proteínový produkt sa označoval *Tccc* (SEKV. ID. ČÍS. 61).

## Príklad 12

### Charakterizácia kmeňov *Photorhabdus*

Aby sa stanovilo, že kolekcia, ktorá sa tu opisuje, bola zostavená z kmeňov *Photorhabdus*, kmene sa hodnotili v pojmoch uznávaných ako mikrobiologické črty, ktoré sú charakteristické pre *Photorhabdus*, a ktoré ho odlišujú od iných *Enterobacteriaceae* a *Xenorhabdus* spp. (Farmer, J.J., 1984. *Bergey's Manual of Systematic Bacteriology*, Vol 1., str. 510-511 (ed. Kreig N.R. a Holt, J.G.). Williams a Wilkins, Baltimore.; Akhurst a Boemare,

1988, Boemare a kol., 1993). Tieto charakteristické črty sú: Gramovo farbenie negatívnych tyčínok, veľkosť organizmu 0,5 až 2  $\mu\text{m}$  v šírke a 2 až 10  $\mu\text{m}$  v dĺžke, červeno/žltá pigmentácia kolónií, prítomnosť kryštalických inklúzných telies, prítomnosť katalázy, neschopnosť redukovať dusičnany, prítomnosť bioluminiscencie, schopnosť odoberať farbivo z rastového média, pozitívna produkcia proteáz, rastovo teplotné rozmedzie pod 37 °C, prežívanie v anaeróbných podmienkach a pozitívna pohyblivosť (tabuľka 20). Referenčné kmene *Escherichia coli*, *Xenorhabdus* a *Photorhabdus* boli zahrnuté vo všetkých porovnávacích testoch. Výsledky sú v súlade so všetkými kmeňmi, ktoré sú časťou čelade *Enterobacteriaceae* a rodu *Photorhabdus*.

Na stanovenie bioluminiscencie každého kmeňa sa používal luminometer a poskytol kvantitatívne a relatívne meranie produkcie svetla. Na meranie relatívnych svetelných emisných jednotiek, sa bujónové kultúry každého kmeňa (bunky a médiá) merali v troch časových intervaloch po inokulácii v bujónovej kultúre (6, 12 a 24 hod) a porovnávali sa so svietivosťou pozadia (neinokulované médiá a voda). Pred meraním svetelnej emisie z rôznych bujónov, sa stanovila bunková hustota na základe merania svetelnej absorbancie (560 nm) na spektrofotometri Gilford Systems (Oberlin, OH) s použitím nasávania buniek. Urobili sa vhodné riedenia (na normalizovanie optickej hustoty na 1,0 jednotky) pred meraním svietivosti. Podiely riedených bujónov sa potom vložili do kyviet (300  $\mu\text{l}$  do každej) a hodnotli sa v luminometri Bio-Orbit 1251 (Bio-Orbit Oy, Twiku, Finland). Integračná perióda pre každú vzorku bola 45 sekúnd. Vzorky sa kontinuálne premiešavali (odstreďované v prepážkových kyvetách), pričom prebiehalo odčítavanie, či je dostatok prístupného kyslíka. Pozitívny test sa determinoval ako  $\geq 5$ -krát luminiscencia pozadia (okolo 5 až 10 jednotiek). Ďalej sa svietivosť kolónií detegovala na fotografickom filme a vizuálne, po vyvolaní v tmavej komore. Gramové farbiace charakteristiky pre každý kmeň sa stanovili komerčne dostupným zariadením Gramovho farbenia (BBL, Cackeyville, MD) a použili sa v spojení

s Gramomými farebnými kontrolnými diapozitívami (Fischer Scientific, Pittsburgh, PA). Mikroskopické hodnotenie sa potom uskutočnilo na mikroskope Zeiss (Carl Zeiss, Nemecko) pri zväčšení 100 x s olejovo imerzným objektívom (10 x okulárový objektív a 2 x zväčšenie tela mikroskopu). Mikroskopické hodnotenie jednotlivých kmeňov z hľadiska veľkosti organizmu, opisu buniek a inklúzných telies (neskôr, po logaritmickej fáze rastu) sa uskutočňovalo s použitím oporných podložných sklíčok (10 x okulár, 2 x telo mikroskopu a 40 x objektívové zväčšenie) s olejovou imerziou a mikroskopom s fázovým kontrastom a mikrometrom (Akhurst, R.J. a Boemare, N.E. 1990. Hmyzie patogénne Nematoda v biologickej kontrole (vyd. Gaugler, R. a Kaya, H.) str. 75-90. CRC Press, Boca Raton, USA.; Baghdiguian, S., Boyer-Giglio M.H., Thaler, J.O., Bonnot, G., Boemare, N., 1993. Biol Cell 79, 177-185). Pigmentácia kolónií sa zisťovala po inokulácii na živý agar Bacto (Difco Laboratories, Detroit, MI), ktorý sa pripravil podľa priloženého predpisu. Inkubácia prebiehala pri 28 °C a popisy sa uskutočnili po 5 až 7 dňoch. Na test na prítomnosť enzýmu katalázy sa kolónie testovaného organizmu odobrali na malom uzávere zo živnej agarovej platne a položili sa na dno sklenenej testovacej skúmavky. Jeden ml laboratórne pripraveného roztoku peroxidu vodíka sa opatrne pridal po stene skúmavky. Pozitívna reakcia sa zaznamenala vtedy, keď sa bublinky plynu (predpokladaný kyslík) objavili okamžite alebo do 5 sekúnd. Takisto sa hodnotili kontrolné vzorky neinokulovaných živných agarov a roztoku peroxidu vodíka. Na testovanie redukcie dusičnanov sa každá kultúra inokulovala na 10 ml Bacto dusičnanový bujón (Difco Laboratories, Detroit, MI). Po 24 hodinovej inkubácii pri 28 °C sa testovala produkcia dusičnanov pridaním dvoch kvapiek reagentu kyseliny sulfanilínovej a dvoch kvapiek reagentu alfa-naftylamínu (pozri Difco Manual, 10. vydanie, Difco Laboratories, Detroit, MI, 1984). Generácia zreteľnej ružovej alebo červenej farby indikuje tvorbu dusitanov z dusičnanov. Schopnosť každého kmeňa odfarbovať rastové médiá sa testovala na MacConkeyho pôde obsahujúcej farbivo neutrálnu červeň (agar Bacto Tergitol-7 obsahujúci

farbivo - brómtymolovú modrú a EMB agar obsahujúci farbivo eozín-Y (agary z Difco Laboratories, Detroit, MI, všetky sa pripravovali podľa priložených návodov). Po inokulácii na týchto médiách sa odfarbenie médií zaznamenávalo po 5 dňovej inkubácii pri 28 °C. Rast na týchto novších médiách je charakteristický pre zástupcov čelade *Enterobacteriaceae*. Pohyblivosť každého kmeňa sa testovala v roztoku Bacto Mobility Test Medium (Difco Laboratories, Detroit, MI) pripraveného podľa priložených inštrukcií. Cielená stabilná inokulácia sa uskutočnila s každým kmeňom a pohyblivosť sa hodnotila makroskopicky s difúznou zónou rastu, ktorý sa rozvinie od radu inokulácii. V mnohých prípadoch sa pohyblivosť pozorovala tiež mikroskopicky z tekutej kultúry preparátom pod zvlhčenými podložnými sklíčkami. Biochemické výživné hodnotenie pre každý kmeň sa uskutočnil pomocou BBL Enterotube II (Benton, Dickinson, Nemecko). Pracovalo sa podľa návodov s vynechaním toho, že inkubácia trvala 5 dní pri 28 °C. Výsledky boli v súlade s predtým citovanými odkazmi pre *Photorhabdus*. Produkcia proteáz sa testovala pozorovaním hydrolýzy želatíny na Bacto želatíne (Difco Laboratories, Detroit, MI), pričom sa platne pripravili podľa priložených inštrukcií. Kultúry sa inokulovali a platne sa inkubovali 5 dní pri 28 °C. Na hodnotenie rastu pri rôznych teplotách sa na agarové platne [2 % proteózového peptónu č. 3 s dvojpercentným Bacto agarom (Difco, Detroit, MI) v deionizovanej vode] rozotrel z bežného zdroja inokula. Platne sa uzatvorili s Nesco filmom a inkubovali pri 20, 28 a 37 °C tri týždne. Na platniach, ktoré sa inkubovali pri 30 °C nerástlo nič a neukázala sa žiadna bunková aktivita po prenose do inkubátora pri 28 °C počas jedného týždňa. Kyslíkové požiadavky kmeňov *Photorhabdus* sa testovali nasledujúcim spôsobom. Cielená stabilná inokulácia sa robila do tekutého tioglykolátového bujónového média (Difco, Detroit, MI). Skúmavky sa inkubovali jeden týždeň pri izbovej teplote a kultúry sa potom hodnotili podľa typu a rozsahu rastu. Indikátor resazurín určuje úroveň oxidácie média alebo aeróbnej zóny (Difco Manual, 10. edícia, Difco Laboratories, Detroit, MI). Získané výsledky rastových zón testovaných kmeňov *Photorhabdus* boli

v súlade s fakultatívne anaerobnými mikroorganizmami.

Tabuľka 19

Taxonomické znaky kmeňov *Photorhabdus*

Kmeň	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
W-14	= <sup>+</sup>	±	±	<u>rd</u> <u>S</u>	±	=	±	±	±	<u>O</u>	±	±	±	±	±	±	=
WX-1	=	±	±	<u>rd</u> <u>S</u>	±	=	±	±	±	<u>O</u>	±	±	±	±	±	±	=
WX-2	=	±	±	<u>rd</u> <u>S</u>	±	=	±	±	±	<u>O</u>	±	±	±	±	±	±	=
WX-3	=	±	±	<u>rd</u> <u>S</u>	±	=	±	±	±	<u>YT</u>	±	±	±	±	±	±	=
WX-4	=	±	±	<u>rd</u> <u>S</u>	±	=	±	±	±	<u>YT</u>	±	±	±	±	±	±	=
WX-5	=	±	±	<u>rd</u> <u>S</u>	±	=	±	±	±	<u>LO</u>	±	±	±	±	±	±	=
WX-6	=	±	±	<u>rd</u> <u>S</u>	±	=	±	±	±	<u>LY</u>	±	±	±	±	±	±	=
WX-7	=	±	±	<u>rd</u> <u>S</u>	±	=	±	±	±	<u>R</u>	±	±	±	±	±	±	=
WX-8	=	±	±	<u>rd</u> <u>S</u>	±	=	±	±	±	<u>O</u>	±	±	±	±	±	±	=
WX-9	=	±	±	<u>rd</u> <u>S</u>	±	=	±	±	±	<u>YT</u>	±	±	±	±	±	±	=
WX-10	=	±	±	<u>rd</u> <u>S</u>	±	=	±	±	±	<u>Ro</u>	±	±	±	±	±	±	=
WX-11	=	±	±	<u>rd</u> <u>S</u>	±	=	±	±	±	<u>Ro</u>	±	±	±	±	±	±	=
WX-12	=	±	±	<u>rd</u> <u>S</u>	±	=	±	±	±	<u>O</u>	±	±	±	±	±	±	=
WX-14	=	±	±	<u>rd</u> <u>S</u>	±	=	±	±	±	<u>LR</u>	±	±	±	±	±	±	=

Tabuľka 19 - pokračovanie

Kmeň	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
WX-15	=	±	±	<u>rd</u> S	±	=	±	±	±	<u>LR</u>	±	±	±	±	±	±	=
H-9	=	±	=	<u>rd</u> S	±	=	±	±	±	<u>LY</u>	±	±	±	±	±	±	=
Hb	=	±	±	<u>rd</u> S	±	=	±	±	±	<u>YT</u>	±	±	±	±	±	±	=
Hm	=	±	±	<u>rd</u> S	±	=	±	±	±	<u>TY</u>	±	±	±	±	±	±	=
HP88	=	±	±	<u>rd</u> S	±	=	±	±	±	<u>LY</u>	±	±	±	±	±	±	=
NC-1	=	±	±	<u>rd</u> S	±	=	±	±	±	<u>O</u>	±	±	±	±	±	±	=
W30	=	±	±	<u>rd</u> S	±	=	±	±	±	<u>YT</u>	±	±	±	±	±	±	=
WIR	=	±	±	<u>rd</u> S	±	=	±	±	±	<u>RO</u>	±	±	±	±	±	±	=
B2	=	±	±	<u>rd</u> S	±	=	±	±	±	<u>R</u>	±	±	±	±	±	±	=
43948	=	±	±	<u>rd</u> S	±	=	±	±	±	<u>O</u>	±	±	±	±	±	±	=
43949	=	±	±	<u>rd</u> S	±	=	±	±	±	<u>O</u>	±	±	±	±	±	±	=
43950	=	±	±	<u>rd</u> S	±	=	±	±	±	<u>O</u>	±	±	±	±	±	±	=
43951	=	±	±	<u>rd</u> S	±	=	±	±	±	<u>O</u>	±	±	±	±	±	±	=
43952	=	±	±	<u>rd</u> S	±	=	±	±	±	<u>O</u>	±	±	±	±	±	±	=

Hodnotené znaky\*

\* A = Gramovo farbenie, B = kryštalické inklúzne telesá, C = bioluminiscencia, D = tvar bunky, E = pohyblivosť, F = redukcia



dusičnanov, G = prítomnosť katalázy, H = hydrolýza želatíny, I = odfarbenie média, J = pigmentácia, K = rast na EMB agare, L = rast na MacConkeyho agare, M = rast na agare Tergitol- 7, N = fakultatívne anaeróbny, O = rast pri 20 °C, P = rast pri 28 °C, Q = rast pri 37° C, +- +/- = pozitívny alebo negatívny znak, rd = tyčinková baktéria, S = veľkosť v rámci rodových deskriptorov, RO = červeno-oranžová, LR = svetločervená, R = červená O = oranžová, Y = žltá, T = hnedá, LY = svetložltá, YT = žltohnedá a LO = svetlooranžová.

Bunková analýza mastných kyselín je uznávaným prameňom pre charakterizáciu baktérií na rodovej a druhovej úrovni (Tornabene, T.G.1985. Analýza tukov a príbuznosť k chemotaxonómii v mikrobiologických metódach, Vol 18, 209-224.; Goodfellow, M, a O'Donnell, A.G., 1993. Korene systematiky baktérií v príručke nový bakteriálny systém (ed. Goodfellow, M a O'Donnell, A.G.) str.3-54. Londýn: Academic Press Ltd.), tieto odkazy sa sem včlenili ako literatúra a použili sa ako dôkaz, že naša zbierka sa týka rodovej úrovne. Kultúry sa dopravili do externého kontrahovaného laboratória na analýzu metylesterov mastných kyselín (FAME) pomocou Mikrobial ID (MIDI, Newark, De, USA) mikrobiálneho identifikačného systému (MIS). Systém MIS pozostáva z plynového chromatografu Hewlet Packard HP5890A s 25mm x 0,2 mm 5 % metylfenyl silikónom fúzanou silikónovou kapilárnou kolónou. Ako nosný plyn sa použil vodík a plameňovo ionizačné detektorové funkcie v kombinácii s automatickým dávkovačom, integrátor a počítač. Počítač porovnáva vzorky metylesterov mastných kyselín s mikrobiálnou knižnicou mastných kyselín a oproti kalibračnej zmesi známych mastných kyselín. Tak, ako sa vyberali v kontrahovanom laboratóriu, kmene pred analýzou rástli 24 hodín pri 28 °C na tryptikázovom sójovom agare. Extrakcia vzoriek sa uskutočnila kontrahovaným laboratóriom štandardnou metódou FAME. Nebola to priama identifikácia kmeňov k žiadnej skupine luminiscenčných baktérií iných ako *Photorhabdus*. Keď sa uskutočnila klastrová analýza, ktorá porovnávala mastné kyseliny, ktoré prevládali v skupine izolátov, prevládajúce mastné kyseliny

kmeňov sa vzťahli na úroveň rodu.

Evolučná diverzita kmeňov *Photorhabdus* v našej zbierke sa merala analýzou PCR (Polymerázová reťazová reakcia), ktorá sprostredkovala genómové odtlačky prstov pomocou genómovej DNA z každého kmeňa. Tento vynález je založený na rodinách repetitívnych sekvencií DNA prítomných v celom genóme rôznych bakteriálnych druhov (revidované - Versalovič, J., Schneder, M., DE Bruijn, F.J. a Lupski, J.R. 1994. *Methods Mol. Cell. Biol.*, 5, 25-40.). Predpokladá sa, že tri z nich, repetitívna extragenová palindromová sekvencia (REP), enterobakteriálny repetitívny intergénový konsenzus (ERIC) a BOX prvok zohrávajú dôležitú úlohu v organizácii bakteriálneho genómu. Predpokladalo sa, že genómová zložka, ktorá sa vyštiepila selekciou a diferenciálnou disperziou týchto elementov v rámci genómu blízko súvisiaceho s bakteriálnymi kmeňmi, sa môže použiť na rozlišovanie týchto kmeňov (napríklad Louws, F.J., Fulbright, D.W., Stephens, C.T. a DE Bruijn, F.J. 1994. *Appl. Environ. Micro.*, 60, 2286-2295.). Rep-PCR využíva oligonukleotidové priméry komplementárne k týmto repetitívnym sekvenciám na amplifikáciu rôzne veľkých fragmentov DNA ležiacich medzi nimi. Výsledné produkty sa separujú elektroforézou na stanovenie " odtlačkov prstov" DNA pre každý kmeň.

Na izoláciu genómovej DNA z našich kmeňov sa bunkový sediment resuspendoval v pufrí TE (10 mM Tris HCl, 1 mM EDTA, pH = 8,0) na konečný objem 10 ml. Potom sa pridalo 12 ml 5 M NaCl. Táto zmes sa centrifugovala 20 min. pri otáčkach 15 000 x g. Výsledný sediment sa resuspendoval v 5,7 ml TE a pridalo sa 300 µl 10 % SDS a 60 µl 20 mg/ml proteínázy K (Gibco BRL Products, Grand Island, NY). Táto zmes sa inkubovala 1 hod. pri 37 °C, potom sa pridalo približne 10 mg lyzozýmu a zmes sa inkubovala ďalších 45 min. Potom sa pridal jeden mililiter 5 M NaCl a 800 µl roztoku CTAB/NaCl (10 % hmotn./obj. CTAB, 0,7 M NaCl) a zmes sa inkubovala za opatrného trepania 10 min. pri 65 °C. Potom sa inkubovala a trepala ďalších 20 min. na to, aby sa purifikoval

bunkový materiál. Pridal sa rovnaký objem roztoku chloroform/izoamylalkoholu (24 : 1, obj./obj.) jemne sa trepal, a potom sa scentrifugoval. Potom sa dve extrakcie upravili rovnakým objemom fenol/chloroform/izoamylalkoholu (50 : 49 : 1). Genómová DNA sa vyzrážala 0,6 objemom izopropanolu. Vyzrážaná DNA sa odobrala sklenenou tyčinkou, dvakrát sa premyla 70 % etanolom, vysušila sa a rozpustila v 2 ml STE (10 mM Tris HCl, pH = 8,0, 10 mM NaCl, 1mM EDTA). DNA sa potom kvantifikovala pomocou optickej hustoty pri 260 nm. Na to, aby sa vykonala rep-PCR analýza genómovej DNA *Photorhabdus* sa použili nasledovné priméry: REP1R-I, 5'-IIIIICGICGICATCIGGC-3' a REP2-I, 5'-ICGICTTATCIGGCCTAC-3'. PCR sa uskutočnila nasledovnými 25  $\mu$ l reakciami: 7,75  $\mu$ l H<sub>2</sub>O, 2,5  $\mu$ l 10X pufri LA (PanVera Corp., Madison, WI), 16  $\mu$ l zmesi dNTP (2,5 mM z každej), 1  $\mu$ l z každého priméru s 50 pM/ $\mu$ l, 1  $\mu$ l DMSO, 1,5  $\mu$ l genómovej DNA (koncentrácie v rozpätí od 0,075 do 0,480  $\mu$ g/ $\mu$ l) a 0,25  $\mu$ l TaKaRa EX Taq (Pan Vera Corp., Madison. WI). Amplifikácia PCR sa uskutočnila na teplotnom cykleri DNA značky Perkin Elmer (Norwalk, CT) s použitím nasledovných podmienok: 95 °C/7 min., potom 35 cyklov z: 94 °C/ 1 min., 44 °C/1 min., 65 °C/8 min., potom nasledovalo 15 min. pri 65 °C. Po cyklovaní sa pridalo 25  $\mu$ l z tejto reakcie na 5  $\mu$ l 6X gélového nosného pufri (0,25 % brómfenolovej modrej, 40 % hmotn./obj. sacharózy v H<sub>2</sub>O). 15 x 20cm 1 %-agarózového gélu potom dalo na TBE pufor (0,09 M Trisboritanu, 0,002 M EDTA) s použitím 8  $\mu$ l z každej reakcie. Gél tiekol približne 16 hodín pri 45 v. Potom sa gély farbili v 20  $\mu$ g/ml etidium bromidu 1 hodinu a odfarbovali v pufri TBE približne 3 hodiny. Fotografie gélov z polaroidu sa potom dali pod UV žiarenie.

Prítomnosť alebo neprítomnosť skupín špecifických veľkostí pre každý kmeň sa odčítavala z fotografií a zapísala sa ako podobnosť matrice v numerickom taxonomickom softvérovom programe, NTSYS-pc (Exeter Software, Setauket, NY). Kontroly kmeňa *E. coli* HB101 a *Xanthomonas oryzae ssp. oryzae* hodnotené v rovnakom čase produkovali PCR "odtlačky prstov" zhodujúce sa s publikovanými údajmi (Versalovic, J., Koeth, T. a Lupski, J.R. 1991. Nucleic

Acids Res., 19, 6823-6831; Vera Cruz, C.M., Halda-Alija, L., Louws, F., Skinner, D.Z., George, M.L., Nelson, R.J., DE Bruijn, F.J., Rice, C. a Leach, J.E., 1995. Int. Rice Res. Notes, 20, 23-24. Vera Cruz, C.M., Ardales, E.Y., Skinner, D.Z., Talag, J., Nelson, R.J., Louws, F.J., Leung, H., Mew, T.W. a Leach, J.E. 1996. Phytopathology (v tlači). Údaje z kmeňov *Photorhabdus* sa potom analyzovali so sériami programov v NTSYS-pc; SIMQUAL (podobnosť kvalitatívnych údajov) na vytvorenie matrice koeficientov podobnosti (použitý Jaccardov koeficient) a SAHN (Sequential, Agglomerative, Hierarchical and Nested) klastrovania (použitie metódy UPGMA [nevážená metóda párovej skupiny s aritmetickými priemermi]), skupín ktoré sa týkajú kmeňov a mohli byť vyjadrené ako fenogram (obrázok 5). COPH (kofenetické hodnoty) a MXCOMP (porovnanie matric) programy sa použili na vytvorenie kofenetickej hodnoty matrice a porovnanie korelácie medzi touto a pôvodnou maticou, na ktorej je založené klastrovanie. Výsledok normalizovanej Mantelovej štatistiky ( $r$ ) bol vytvorený z merania, ktoré bolo vhodné pre klastrovú analýzu ( $r = 0,8$  až  $0,9$  predstavuje veľmi dobré vlastnosti). V našom prípade  $r = 0,919$ . Preto je naša zbierka porovnaním skupiny rôznych ľahko rozlíšiteľných kmeňov predstavujúcich rod *Photorhabdus*.

### Príklad 13

Insekticídne využitie toxínu (-ov) produkovaného rôznymi kmeňmi *Photorhabdus*

Počiatočné "očkovacie" kultúry rôznych kmeňov *Photorhabdus* sa pripravili inokuláciou 175 ml 2 % proteázového peptónu 3 (PP3) (Difco Laboratoris, Detroit, MI) tekutého média s primárnym variantom subklonu v 500 ml zásobných nádobách s Delongovým hrdlom uzavretým Kaputovou zátkou. Inokulum pre každú očkovaciu kultúru sa získalo z olejom pokrytých šikmých agarov s kultúrami alebo z plochých kultúr. Po inokulácii sa tieto nádoby inkubovali 16 hod. pri 28 °C na rotačnej trepačke pri 150 rpm. Potom sa

očkovací kultúry použili ako jednotné zdroje inokula na fermentáciu každého kmeňa. Dlhodobá ochrana inokula v stave schopnom tvorby toxínu sa uskutočnila zo zostávajúcej očkovacej kultúry v postlog štádiu s pokrytím so sterilným minerálnym olejom a pridaním sterilnej magnetickej miešacej tyčinky pre ďalšie resuspendovanie a kultúry sa uložili v tme pri izbovej teplote. Produkčné bujóny sa inokulovali pridaním 1 % aktívne rastúcej očkovacej kultúry na čerstvé 2 % PP3 médiá (napríklad 1,75 ml na 175 ml čerstvého média). Produkcia bujónov prebiehala buď v 500 ml zásobných bankách (pozri vyššie) alebo v 2800 ml nepriehľadných, naspodu vypuklých bankách (500 ml objemu) uzavretých silikónovou penovou zátkou. Produkčné nádoby sa inkubovali 24 až 48 hod. vo vyššie opísaných podmienkach. Nasledujúcou inkubáciou sa bujóny dali do sterilných jednolitrových polyetylénových nádob, centrifugovali pri 2600 x g 1 hodinu pri 10 °C a dekantovali od sedimentu buniek a poškodeného zvyšku. Tekutý bujón sa potom vákuovo filtroval cez sklenené filtre Whatman GF/D (2,7 µM retencie) a GF/B (1,0 µM retencie) na to, aby sa odstránili zvyšky. Ďalšia čírosť bujónu sa dosiahla na tangenciálnom prietočnom mikrofiltračnom zariadení (Pall Filtron, Northborough, MA) pričom sa použil 0,5 µM filter s otvorenými kanálmi. Ak by bolo potrebné, mohla by sa získať ďalšia čírosť chladením bujónu (na 4 °C) a centrifugovaním počas niekoľkých hodín pri 2600 x g. Po týchto procedúrach sa bujón sterilizoval filtráciou na 0,2 µM nitrocelulóзовom membránovom filtri. Sterilné bujóny sa potom priamo použili na biologické testy, biochemickú analýzu alebo sa koncentrovali (až do 15 násobku) pomocou 10 000 MW rezača, M12 ultrafiltračného zariadenia (Amicon, Beverly, MA) alebo centrifugačnými koncentrátormi (Millipore, Bedford, MA a Pall Filtron, Northborough, MA) s veľkosťou pórov 10 000 MW. V prípade centrifugačných koncentrátorov sa bujón centrifugoval pri 2 000 x g približne 2 hod. Pridalo sa 10 000 MW permeátu k zodpovedajúcemu retentátu na to, aby sa dosiahla želaná koncentrácia zložiek vyššia ako 10 000 MW. Tepelná inaktivácia upravovaných bujónových vzoriek sa dosiahla ohriatím vzoriek na

100 °C v pieskom naplnenom topnom hniezde 10 minút.

Bujón (-y) a toxínový (é) komplex (-y) z rôznych kmeňov *Photorhabdus* boli užitočné na redukcii populácií hmyzu a použil sa v spôsobe inhibície populácie hmyzu, ktorý zahŕňa aplikovanie na miesto hmyzu efektívneho inaktivujúceho insekticídneho množstva opísanej účinnej látky. Demonštrácia spektra insekticídnej účinnosti získanej z bujónov vybranej skupiny kmeňov *Photorhabdus* fermentujúcej tak, ako je opísané vyššie, je zobrazená v tabuľke 20. Je možné, že ďalšie insekticídne účinky sa môžu zaznamenať u týchto kmeňov prostredníctvom vzrastu koncentrácie bujónu alebo s použitím odlišných fermentačných metód. V súlade s účinkom, ktorý sa spája s proteínom, bola insekticídna aktivita všetkých testovaných kmeňov termolabilná (pozri vyššie) .

Bujónová kultúra (-y) rôznych kmeňov *Photorhabdus* preukázali rôznu insekticídnu účinnosť (mortalita a/alebo inhibícia rastu, redukcia výletu dospelých jedincov) oproti množstvu hmyzu. Podrobnejšie, účinnosť je porovnateľná proti larvám liskaviek, a larvám *Anthonomus grandis*, ktoré sú zástupcami hmyzieho radu *Coleoptera*. Ďalší zástupcovia radu *Coleoptera* sú kováčiky obilné, blyskáčiky repkové, skočky chrenové, seed beetles, a pásavka zemiaková. Účinnosť sa tiež zistila proti cikádke *Macrosteles severini* a kobylke obilnej, ktoré sú zástupcami radu *Homoptera*. Ďalším zástupcom *Homoptera* sú ako kobylky, méra hrušková, méra jablňová, červcovité, molicovité, peniarkovité, tak ako aj množstvo druhov vošiek špecifických hostiteľov. Bujóny a purifikovaný toxínový komplex (-y) sú tiež účinné proti sivkavcovi, *Manduca sexta* a vijačke kukuričnej, ktoré sú zástupcami radu *Lepidoptera*. Ďalšími typickými zástupcami tohto radu sú sivkavec stavikrvový, mora kapustová, siatica ypsilonová, *Helicoverpa zea*, obaľovač jablňový, molovitý, vijačka papriková, obaľovačovité, mlynárik kapustný, kvetovka bavlníková, vreckovcovité, priadkovec americký, vijačka voštinová a sivkavec južný. Účinnosť je tiež viditeľná proti

ovocným muškám a larvám komárov, ktoré sú zástupcami radu *Diptera*. Ďalšími zástupcami radu *Diptera* sú plodomor hrachový, vrtavka mrkvová, kvetárka kapustová, kvetárka red'kovkovitá, kvetárka cibuľová, tipuľa bahenná a mucha domová a rôzne druhy komárov. Aktivita bujónu (-ov) a toxínového komplexu (-ov) je tiež zreteľná proti roztočcovi snovaciemu, ktorý je zástupcom radu *Acarina*, ktorý zahŕňa tiež roztočníka jahodového, broad mites, citrus red mite, roztočca ovocného, nahohrdzu hruškovú a tomato russet mite.

Účinnosť proti larvám liskaviek (*Diabrotica* sp.) sa testovala nasledovne. Bujónová kultúra (-y) *Photorhabdus* (0 až 15 násobne koncentrovaná, sterilizovaná filtráciou), 2 % proteázový peptón 3, purifikovaný pepsínový komplex (-y) (0,23 mg/ml) alebo 10 mM pufru fosforečnanu sodného s pH = 7,0 sa aplikovalo priamo na povrch (okolo 1,5 cm<sup>2</sup>) umelo pripraveného krmiva (Rose, R.I. a McCabe, J.M. (1973). J. Econ. Entomol. 66, str. 398-400) v 40 µl podieloch. Toxínový komplex sa riedil v 10 mM pufrí fosforečnanu sodného, pH = 7,0. Misky s krmivom sa vysušili na vzduchu v sterilnom prietokovom boxe a jamky sa infikovali jednotlivými neonátmi *Diabrotica undecimpunctata howardi* (Liskavka uhorková, SCR) vyliahnutými z povrchu sterilizovaných vajíčok. Platne sa pevne uzatvorili, vložili sa do vlhčenej rastovej komory a udržiavali sa pri 27 °C počas vhodne dlhej periódy (3 až 5 dní). Mortalita a hmotnosť lariev sa potom odčítali. Vo všeobecnosti sa použilo 16 jedincov hmyzu na prostriedok vo všetkých štúdiách. Kontrolná mortalita bola všeobecne nižšia ako 5 % .

Účinnosť proti *Anthomonas grandis* sa testovala nasledovne. Koncentrované (1 až 10 násobne) bujóny *Photorhabdus*, kontrolné médium (2 % proteázového peptónu č. 3), purifikovaný toxínový komplex (-y) (0,23mg/ml) alebo 10 mM pufor fosforečnanu sodného, pH = 7,0 sa aplikovalo v 60 µl podieloch na povrch 0,35 g umelej diéty (Stoneville Yellow diéta lepidopter) a usušila sa. Jednotlivé, 12 až 24 hod. larvy *Anthomonas grandis* sa umiestnili

na krmivo a jamky sa pevne uzatvorili a udržovali pri 25 °C, 50 % relatívnej vlhkosti počas 5 dní. Potom sa hodnotila mortalita a hmotnosť lariev. Kontrolná mortalita bola v rozmedzí 0 až 13 %.

Účinnosť proti larvám komárov sa testovala nasledovne. Test sa robil v 96-jamkovej mikrotitračnej platni. Každá jamka obsahovala 200  $\mu$ l vodného roztoku (10 násobne koncentrovanej bujónovej kultúry (-úr) *Photorhabdus*, kontrolné médium (2 % proteázového peptónu č. 3), 10 mM pufru fosforečnanu sodného, toxínový komplex (-y) 0,23 mg/ml alebo voda) a približne 20 1-deň starých lariev (*Aedes aegypti*). 6 jamiek sa použilo na prostriedok. Výsledky sa odčítavali po 3 až 4 dňoch po infestácii. Kontrolná mortalita bola medzi 0 až 20 %.

Účinnosť proti ovocným muškám sa sledovala nasledovne. Kúpené médium pre *Drosophila melanogaster* sa pripravilo použitím 50 % suchého média a 50 % tekutiny alebo vody, kontrolného média (2 % proteázový peptón č. 3), 10 násobne koncentrovanej bujónovej kultúry (-úr) *Photorhabdus*, purifikovaného toxínového komplexu (-ov) (0,23 mg/ml) alebo 10 mM pufru fosforečnanu sodného, pH = 7,0. Toto sa dosiahlo tak, že sa 4,0 ml suchého média vložilo do každej z 3 kultivačných vialiek na prípravok a pridali sa 4,0 ml vhodnej tekutiny. Desať neskorších instarových štádií lariev *Drosophila melanogaster* sa potom vložilo do každej 25 ml vialky. Vialky sa nechali na laboratórnom stole pri izbovej teplote, pod fluorescenčným stropným osvetlením. Kukly alebo dospelé jedince sa odpočítali po 15 dňovej expozícii. Výlet dospelých jedincov sa porovnával k vode a kontrolnému médiu (0 až 16 % redukcia).

Účinnosť proti dospelým jedincom cikádky (*Macrosteles severini*) a poslednému larválnemu štádiu kobyliek (*Peregrinus maidis*) sa testovala testom príjmu potravy, založeným na umožnení aktívneho kŕmenia bez vonkajšieho zásahu. Zásobník pre účinný/"kŕmny" roztok sa urobil dvoma otvormi v strede spodnej



časti 35 x 10 mm Petriho misky. 2 palce štvorca Parafilmu M sa položili cez povrch misky a zabezpečili sa "O" krúžkom. 1 uncový plastický pohár sa potom infikoval približne 7 jedincami skákavého hmyzu a rezervoár sa položil navrch pohára Parafilmom smerom dole. Testovaný roztok sa potom pridal k zásobníku cez otvory. V testoch sa použili 10 násobne koncentrovaná bujónová (-é) kultúra (-y) *Photorhabdus*, bujón a kontrolné médium (2 % proteázový peptón č. 3) a dialyzovali sa oproti 10 mM fosfátovému pufru s pH = 7,0 a sacharóza (do 5 %) sa pridala k výslednému roztoku na redukciu kontrolnej v kontrole. Purifikovaný toxínový komplex (y) (0,23 mg/ml) alebo 10 mM pufru fosforečnanu sodného s pH = 7,0 sa tiež testoval. Mortalita sa zaznamenávala v 3. dni. Test sa udržoval v inkubátore pri 28 °C, 70 % relatívnej vlhkosti s 16/8 svetelnou periódou. Testy sa odčítavali z hľadiska mortality po 72 hod. Kontrolná mortalita bola menšia ako 6 %.

Účinnosť proti larvám *Lepidoptera* sa testovala nasledovne. Koncentrovaná (10 násobne) bujónová kultúra (-y) *Photorhabdus*, kontrolné médium (2 % proteázový peptón), purifikovaný toxínový komplex (-y) (0,23 mg/ml) alebo 10 mM pufru fosforečnanu sodného, pH = 7,0 sa aplikoval priamo na povrch (okolo 1,5 cm<sup>2</sup>) štandardnej výživy lepidopter (Stonewille Yellow výživa) v 40 µl podieloch. Misky s výživou sa nechali vysušiť na vzduchu v sterilnom prietokovom boxe a každá jamka sa infestovala jednotlivými neonátnymi larvami. Vajíčka vijačky kukuričnej (*Ostrinia nubilalis*) a *Manduca sexta* sa získali z komerčných zdrojov a vyľahli v domácej liahni, zatiaľ čo larvy sivkavca (*Heliothis virescens*) sa dodali interne. Po nasledovnej infestácii larvami sa misky s výživou pevne uzatvorili, odložili do vlhčenej rastovej komory a udržovali sa v tme pri 27 °C vhodne dlhý čas. Determinácia mortality a hmotnosti sa odčítavala v 5. dni. Vo všeobecnosti sa používalo 16 jedincov hmyzu pre prípravok vo všetkých štúdiách. Kontrolná mortalita bola vo všeobecnosti v rozsahu od 4 až 12,5 % pre kontrolné médium a menšia ako 10 % pre fosfátový pufru.

Účinnosť proti roztočovi snovaciemu (*Tetranychus urticae*) sa určovala nasledovne. Mladé ovocné rastliny sa upravili na jednoduchý kotyledon a postriekali sa s 10 násobne koncentrovaným bujónom (-i), kontrolným médiom (2 % protázový peptón), purifikovaným toxínovým komplexom (-mi) (0,23 mg/ml) alebo 10 mM pufor fosforečnanu sodného, pH = 7,0. Po vysušení sa rastliny infestovali zmesnou populáciom roztočov a nechali sa pri laboratórnej teplote a vlhkosti 72 hod. Žijúce roztoče sa odčítavali na stanovenie úrovne kontroly.

Tabuľka 20

Sledované insekticídne spektrum bujónov rôznych kmeňov *Photorhabdus*.

<i>Photorhabdus</i> kmeň	Senzitívne* druhy hmyzu
WX-1	3**, 4, 5, 6, 7, 8
WX-2	2, 4
WX-3	1, 4
WX-4	1, 4
WX-5	4
WX-6	4
WX-7	3, 4, 5, 6, 7, 8,
WX-8	1, 2, 4
WX-9	1, 2, 4
WX-10	4
WX-11	1, 2, 4
WX-12	2, 4, 5, 6, 7, 8
WX-14	1, 2, 4
WX-15	1, 2, 4
W30	3, 4, 5, 8
NC-1	1, 2, 3, 4, 5, 6, 7, 8, 9
WIR	2, 3, 5, 6, 7, 8,

Tabuľka 20 - pokračovanie

<i>Photorhabdus</i> kmeň	Senzitívne* druhy hmyzu
HP88	1, 3, 4, 5, 7, 8
Hb	3, 4, 5, 7, 8
HM	1, 2, 3, 4, 5, 7, 8
H9	1, 2, 3, 4, 5, 6, 7, 8
W-14	1, 2, 3, 4, 5, 6, 7, 8, 10
ATCC 43948	4
ATCC 43949	4
ATCC 43950	4
ATCC 43951	4
ATCC 43952	4

\* =  $\geq 25$  % mortalita a/alebo inhibícia rastu oproti kontrole

\*\* = 1 - Sivkavec (*Heliothis virescens*), 2 - vijačka kukuričná,

3 - *Manduca sexta*, 4 - Liskavka uhorková (SCR, *Diabrotica undecimpunctata howardi*), 5 - *Anthonomus grandis*, 6 - komár, 7 - ovocná muška, 8 - *Macrostesles severini*, 9 - kobylka obilná, 10 - roztočec snovací.

#### Príklad 14

*Photorhabdus*, kmene iné ako W-14: purifikácia, charakterizácia a spektrum účinnosti

#### Purifikácia

Postup, ktorý nasleduje, je podobný tomu, ktorý bol vyvinutý na purifikáciu W-14 a je založený na purifikácii tých frakcií, ktoré majú najsilnejšiu účinnosť proti liskavke uhorkovej (SCR), ktorá sa determinovala v biotestoch (pozri príklad 13). Obyčajne

sa 4 až 20 l bujónu filtrovalo tak, ako je opísané v príklade 13, zozbierali sa a koncentrovali na špirálovej ultrafiltračnej náplni Amicon typ SLY100 upevnenej na filtračnom zariadení Amicon M-12. Retentát obsahoval pôvodné proteíny skladajúce sa z molekulových veľkostí väčších ako 100 kDa, zatiaľ čo filtrát, ktorý vznikol prietokom cez materiál obsahoval pôvodné proteíny menšej veľkosti ako 100 kDa. Väčšina účinnosti proti SCR bola obsiahnutá v 100 kDa retentáte. Retentát sa potom kontinuálne diafiltroval s 10mM fosforečnanom sodným (pH = 7,0), až kým filtrát dosiahol  $A_{280} < 0,100$ . Ak neboli inakšie podmienky, všetky postupy od tohto bodu sa uskutočnili v pufri, tak ako je definovaný - 10mM fosforečnan sodný (pH = 7,0). Retentát sa koncentroval na konečný objem približne 0,20 l a filtroval pomocou 0,45 mm Nalgene<sup>TM</sup> - filtračnej jednotky pre sterilnú filtráciu. Filtrovaný materiál sa pri prietoku 7,5 ml/min. vložil do kolóny Pharmacia HR16/10, ktorá bola naplnená silnou anionovýmennou matricou kondicionovanou pufrom použitím PerSeptive Biosystem Sprint HPLC systém. Po nanesení sa kolóna premývala s pufrom, až kým sa dosiahla  $A_{280} < 0,100$ . Potom sa proteíny eluovali z kolóny pufrom s 0,4 M NaCl, pri prietoku 2,5 ml/min. počas 20 min. do celkového objemu 50 ml. Potom sa kolóna premývala pufrom, ktorý obsahoval 1,0 M NaCl pri tom istom prietoku ďalších 20 min (konečný objem = 50 ml). Proteíny eluované 0,4 M a 1,0 M NaCl sa dali do separačného dialýzového vaku (Spectra/Por Membrane MWCO: 2000) a dialyzovali sa cez noc pri 4 °C v 12 l pufri. Väčšina účinnosti proti SCR bola obsiahnutá v 0,4 M frakcii. 0,4 M frakcia sa ďalej purifikovala aplikovaním 20 ml na kolónu Pharmacia XK 26/100, ktorá bola znovu naplnená sorbentom Sepharose CL4B (Pharmacia), pri prietoku 0,75 ml/min. Potom sa frakcie spájali na základe  $A_{280}$  píkoveho profilu a koncentrovali sa na konečný objem 0,75 ml na centrifugačnom zariadení Millipore Ultrafree s membránou Biomax-50K NMWL. Koncentrácie proteínov sa určili pomocou súpravy Biorad Protein Assay s hovädzím gamaglobulínom ako štandardom.

## Charakterizácia

Pôvodná molekulová hmotnosť SCR toxínového komplexu sa určila na kolóne Pharmacia HR 16/50, ktorá bola znovu naplnená Sepharose CL4B v pufri. Potom sa kolóna kalibrovala proteínmi so známou molekulovou veľkosťou, ktorá bola dostatočná na výpočet toxínovej približnej pôvodnej molekulovej veľkosti. Tak, ako je uvedené v tabuľke 21, molekulová veľkosť toxínového komplexu sa pohybovala v rozmedzí od 777 kDa kmeňa Hb až 1900 kDa kmeňa WX-14. Výťažok toxínového komplexu bol tiež rôzny, od produkcie 0,8 mg/l kmeňa WX-12 až do 7,0 mg/l, ktorý produkoval kmeň Hb .

Proteíny, nachádzajúce sa v toxínovom komplexe, sa hodnotili z hľadiska individuálnej polypeptidovej veľkosti SDS-PAGE analýzou. Obyčajne sa nanieslo 20 mg proteínu toxínového komplexu z každého kmeňa na 2 až 15 % polyakrylamidový gél (Integrated separation systems) a podrobilo sa elektroforéze pri 20 mA v pufri Biorad SDS-PAGE. Po ukončení elektroforézy sa gély farbili v Biorad Coomasieho modrej R-250 počas noci (0,2 % v metanole: kyselina octová: voda = 40 : 10 : 40 obj./obj./obj.). Potom sa gély odfarbovali v zmesi metanol : kyselina octová : voda = 40 : 10 : 40 (obj./obj./obj.). Gély sa potom oplachovali vodou 15 min. a skenovali na Molecular Dynamics Personal Laser Densitometer. Pruhy sa kvantifikovali a molekulová hmotnosť sa vypočítala ako pomer k vysokej molekulovej hmotnosti štandardov Biorad, ktoré sa pohybovali od 200 do 245 kDa.

Veľkosti jednotlivých polypeptidov obsahujúce SCR toxínový komplex z každého kmeňa sú prehľadne uvedené v tabuľke 22. Veľkosti jednotlivých polypeptidov boli v rozmedzí od 230 kDa kmeňa WX-1 do veľkosti 16 kDa ako je vidieť pre kmeň WX-7. Každý kmeň, okrem kmeňa Hb, mal polypeptidy obsahujúce toxínový komplex, ktorý bol v rozmedzí 160 až 230 kDa, v rozmedzí 100 až 160 kDa a v rozmedzí 50 až 80 kDa. Tieto údaje indikujú že toxínový komplex sa môže meniť v zložení peptidu a komponentov

peptidu od kmeňa ku kmeňu, predsa len vo všetkých prípadoch sa javia toxínové vlastnosti ako konzistentný, široký, oligomérený, proteínový komplex.

Tabuľka 21

Charakterizácia toxínového komplexu *Photorhabdus*, kmeňov mimo W-14

Kmeň	Približná pôvodná molekulová hmotnosť <sup>a</sup>	Výťažnosť aktívnej frakcie (mg/l) <sup>b</sup>
H9	972 000	1,8
HB	777 000	7,0
Hm	1 400 000	1,1
HP88	813 000	2,5
NCl	1 092 000	3,3
WIR	979 000	1,0
WX-1	973 000	0,8
WX-2	951 000	2,2
WX-7	1 000 000	1,5
WX-12	898 000	0,4
WX-14	1 900 000	1,9
W-14	860 000	7,5

a pôvodná molekulová hmotnosť určená na kolóne Pharmacia HR 16/50 naplnenej so Sepharose CL4B  
b množstvo toxínového komplexu získaného z bujónovej kultúry.

### Spektrum aktivity

Tak ako je zobrazené v tabuľke 23, toxínové komplexy purifikované z kmeňov Hm a H9 sa testovali z hľadiska účinnosti proti veľkému počtu hmyzu s porovnaním s toxínovým komplexom z kmeňa W-14. Testy sa uskutočňovali tak, ako je opísané v príklade 13. Toxínové komplexy zo všetkých troch kmeňov preukázali účinnosť proti sivkavcovi (*Heliothis virescens*), *Manduca sexta*, vijačke kukuričnej, liskavke uhorkovej a cikádke *Macrosteles severini*. Ďalej toxínový komplex kmeňov Hm a W-14 tiež preukázali účinnosť proti roztočcovi snovaciemu. Toxínový



Tabuľka 23

Pozorované insekticídne spektrum purifikovaného toxínového komplexu z kmeňov *Photorhabdus*

Kmeň <i>Photorhabdus</i>	Citlivé* druhy hmyzu
Toxínový komplex Hm	1**, 2, 3, 5, 6, 7, 8
Toxínový komplex H9	1, 2, 3, 6, 7, 8
Toxínový komplex W-14	1, 2, 3, 4, 5, 6, 7, 8,

\* = > 25 % mortalita alebo inhibícia rastu oproti kontrole  
\* = > 25 % mortalita alebo inhibícia rastu oproti kontrole  
\*\* = 1 - sivkavec (*Heliothis virescens*), 2 - vijačka kukuričná  
3 - liskavka uhorková, 4 - komár, 5 - roztočec snováci,  
6 - cikádka (*Macrosteles severini*), 7- ovocná muška,  
8 - *Anthomonas grandis*

Príklad 15

Subfrakcionácia *Photorhabdus* proteínového toxínového komplexu

Proteínový toxínový komplex *Photorhabdus* sa izoloval tak, ako je opísané v príklade 14. Ďalej sa 10 mg toxínu aplikovalo na kolónu MonoQ 5/5 kondicionovanú 20 mM Tris HCl s pH = 7,0 pri prietoku 1 ml/min. Kolóna sa premývala 20 mM Tris HCl s pH = 7,0 až kým sa optická hustota pri 280 nm vrátila na základnú absorbanciu. Proteíny viazané v kolóne sa eluovali lineárnym gradientom NaCl od 0 do 1,0 M NaCl v 20 mM Tris HCl s pH 7,0 pri prietoku 1 ml/min počas 30 min. Jeden ml frakcií sa zbral a podrobil biotestu na liskavke uhorkovej (SCR) (pozri príklad 13). Účinné píky sa určili v sérii riedení každej frakcie v biotestoch s SCR. Objavili sa dva účinné píky proti SCR a pomenovali sa A (eluát v okolo 0,2 až 0,3 M NaCl) a B (eluované v 0,3 až 0,4 M NaCl). Účinné píky A a B sa zbierali jednotlivo a oba píky sa ďalej purifikovali 3-krokovým postupom, ktorý je opísaný nižšie.



Tuhý  $(\text{NH}_4)_2\text{SO}_4$  sa pridal do vyššie opísanej proteínovej frakcie na konečnú koncentráciu 1,7 M. Proteíny sa potom aplikovali na kolónu fenyl-Superose 5/5, kondiciovanú 1,7 M  $(\text{NH}_4)_2\text{SO}_4$  v 50 mM pufri fosforečnanu draselného s pH = 7,0 pri 1 ml/min. Proteíny, zachytené na kolóne sa eluovali lineárnym gradientom 1,7 M  $(\text{NH}_4)_2\text{SO}_4$ , 0 % etylénglykolu, 50 mM fosforečnanu draselného, pH = 7,0 až po 25 % etylénglykolu, 25 mM fosforečnanu sodného pH = 7,0 (žiaden  $(\text{NH}_4)_2\text{SO}_4$ ) pri prietoku 0,5 ml/min. Frakcie sa počas noci dialyzovali oproti 10 mM pufru fosforečnanu sodného s pH = 7,0. Účinnosť proti SCR sa určovali v každej frakcii biotestom.

Frakcie s najvyššou účinnosťou sa zbierali a aplikovali na kolónu MonoQ 5/5, ktorá bola kondiciovaná 20 mM Tris HCl s pH = 7,0, 1 ml/min. Proteíny zachytené na kolóne sa eluovali 1 ml/min lineárnym gradientom 0 až 1 M NaCl v 20 mM Tris HCl s pH = 7,0.

V konečnom kroku purifikácie sa najúčinnšie frakcie, získané predtým (určené SCR biotestom) spájali a dali sa na druhú fenyl-Superose 5/5 kolónu. Tuhý  $(\text{NH}_4)_2\text{SO}_4$  sa pridal na konečnú koncentráciu 1,7 M. Roztok sa potom nanášal do kolóny kondiciovanej 1,7 M  $(\text{NH}_4)_2\text{SO}_4$  v 50 mM pufri fosforečnanu draselného s pH = 7,0, 1 ml/min. Proteíny zachytené na kolóne sa eluovali lineárnym gradientom 1,7 M  $(\text{NH}_4)_2\text{SO}_4$ , 50 mM fosforečnanu draselného s pH = 7,0 až do 10 mM fosforečnanu sodného s pH = 7,0, 0,5 ml/min. Frakcie sa dialyzovali počas noci oproti 10 mM pufru fosforečnanu sodného s pH = 7,0. Účinky v každej frakcii proti SCR sa určili biotestom.

Konečný purifikovaný proteín, vzniknutý opísaným trojkrokovým postupom z píku A, sa nazval toxín A. Konečný purifikovaný proteín z píku B sa nazval toxín B.

## Charakterizácia a sekvencovanie aminokyselín toxínu A a toxínu B

Na SDS-PAGE oba toxíny A aj B obsahovali 2 veľké peptidy (> 90 % celkového Commassie zafarbeného proteínu): 192 kDa (nazvané A1 prípadne B1) a 58 kDa (nazvané A2 a B2). V obidvoch toxínoch sa A aj B odhalila len jedna veľká skupina v natívnom PAGE, čo indikovalo, že A1 a A2 boli subjednotky jedného proteínového komplexu a B1 a B2 boli subjednotkami jedného proteínového komplexu. Ďalej sa natívna molekulová hmotnosť obidvoch toxínov A aj B určila na 860 kDa gélovou filtračnou chromatografiou. Zistilo sa, že relatívne molárne koncentrácie A1 k A2 sú ekvivalentné 1 k 1 tak, ako boli determinované denzitometrickou analýzou SDS-PAGE gélov. Podobne B1 a B2 peptidy boli prítomné v rovnakej molárnej koncentrácii.

Toxín A a toxín B sa podrobili elektroforéze v 10 % SDS-PAGE a presali sa na membrány PVDF. Plaky sa poslali na analýzu aminokyselín a N-terminálneho sekvencovania aminokyselín na Harward MicroChem a prípadne na Cambridge ProChem. N-koncová aminosekvencia B1 sa určila, že je identická so SEKV. ID. ČÍS. 1, TcbA<sub>ii</sub> oblasť génu tcbA (SEKV. ID. ČÍS. 12, pozícia 87 až 99). Jedinečná N-koncová sekvencia sa získala pre peptid B2 (SEKV. ID. ČÍS. 40). N-koncová sekvencia aminokyselín peptidu B2 bola identická s oblasťou TcbA<sub>iii</sub> odvodenej sekvencie aminokyselín pre gén tcbA (SEKV. ID. ČÍS. 12, pozícia 1935 až 1945). Preto toxín B obsahoval prednostne dva peptidy, TcbA<sub>ii</sub> a TcbA<sub>iii</sub>, o ktorých sa zistilo, že sú odvodené z rovnakého génového produktu TcbA.

N-koncová sekvencia A2 (SEKV. ID. ČÍS. 41) bola jedinečná v porovnaní s peptidom TcbA<sub>iii</sub> a ďalších peptidov. Peptid A2 sa označoval TcdA<sub>iii</sub> (pozri príklad 17). SEKV. ID. ČÍS. 6 sa určila ako zmes aminokyselinových sekvencií SEKV. ID. ČÍS. 40 a 41

Peptidy A1 a A2 sa ďalej podrobili vnútornému aminokyselinovému sekvencovaniu. Na vnútorné sekvencovanie

aminokyselín sa podrobilo 10 µg toxínu A elektroforéze v 10 % SDS-PAGE a presalo na membránu PVDF. Potom sa plaky farbili amidočernou. Peptidy A1 a A2, označované TcdA<sub>ii</sub> a TcdA<sub>iii</sub>, boli odobrané z plaku a poslané na Harvard MicroChem a Cambridge ProChem. Peptidy boli podrobené štiepeniu tripsínom, a potom nasledovala HPLC chromatografia na separáciu jednotlivých peptidov. Analýza N-koncovej aminokyseliny sa robila na vybraných triptických peptidových fragmentoch. Dve vnútorné sekvencie aminokyselín peptidu A1 (TcdA<sub>ii</sub>-PK71, SEKV. ID. ČÍS. 38 a TcdA<sub>ii</sub>-PK44, SEKV. ID. ČÍS. 39) sa preukázali, že majú signifikantné homológie z odvodenými aminokyselinovými sekvenciami oblasti TcbA<sub>ii</sub> génu *tcbA* (SEKV. ID. ČÍS. 12). Podobne N-koncová sekvencia (SEKV. ID. ČÍS. 41) a dve vnútorné sekvencie peptidov A2 (TcdA<sub>iii</sub>-PK57, SEKV. ID. ČÍS. 42 a TcdA<sub>iii</sub>-PK20, SEKV. ID. ČÍS. 43) sa tiež ukázali ako signifikantne homologické s odvodenými aminokyselinovými sekvenciami oblasti TcbA<sub>iii</sub> génu *tcbA* (SEKV. ID. ČÍS. 12).

Záver vyššie uvedených výsledkov je, že toxínový komplex má najmenej dva proteínové toxínové komplexy účinné proti SCR, a to toxín A a toxín B. Toxín A a toxín B sú podobné vo svojich pôvodných a subjednotkových molekulových hmotnostiach, avšak zloženie ich peptidov je rozdielne. Toxín A obsahuje peptidy TcdA<sub>ii</sub> a TcdA<sub>iii</sub> ako prevládajúce peptidy a toxín B obsahuje TcbA<sub>ii</sub> a TcbA<sub>iii</sub> ako prevládajúce peptidy.

Purifikácia a charakterizácia toxínu C Tca peptidov.

Komplex proteínového toxínu *Photorhabdus* sa izoloval tak, ako bolo opísané vyššie. Potom sa asi 50 mg toxínu aplikovalo na MonoQ 10/10 kolónu zrovnováženú s 20 mM Tris-HCl, pH 7,0 pri rýchlosti toku 2 ml/min. Kolóna sa premyla s 20mM Tris HCl, pH 7,0, pokiaľ sa optická hustota pri 280 nm nevrátila na základnú úroveň. Proteíny naviazané v kolóne sa eluovali s lineárnym gradientom od 0 do 1M NaCl v 20 mM Tris HCl, pH 7,0 pri 2 ml/60 min. Zozbierali sa 2 ml frakcií a podrobili sa Western

analýze s použitím pAb TcaB<sub>ii</sub>-syn protilátky (pozri príklad 21) ako primárnej protilátky. Frakcie, ktoré reagovali s pAb TcaB<sub>ii</sub>-syn protilátkou, sa kombinovali a pridal sa tuhý (NH<sub>4</sub>)SO<sub>4</sub> do konečnej koncentrácie 1,7 M. Proteíny sa potom aplikovali na kolónu fenyl-Superose 10/10 zrovnováženú s 1,7 M (NH<sub>4</sub>)SO<sub>4</sub> v pufrí fosforečnanu draselného, pH 7 pri 1 ml/min. Proteíny naviazané v kolóne sa eluovali s lineárnym gradientom 1,7 M (NH<sub>4</sub>)SO<sub>4</sub>, 50 ml fosforečnanu draselného, pH 7,0 do 10 ml fosforečnanu draselného, pH 7,0 pri 1 ml/min počas 120 min. Zbierali sa frakcie po 2 ml, dialyzovali sa cez noc oproti pufru 10 mM fosforečnanu draselného, pH 7,0 a analyzovali sa Western analýzou s použitím pAb TcaB<sub>ii</sub>-syn protilátok ako primárnych protilátok.

Frakcie, ktoré navzájom reagovali s protilátkou sa vylúhovali a aplikovali na Mono Q 5/5 kolónu zrovnováženú s 20 mM Tris-HCl, pH 7,0 pri 1 ml/min. Proteíny viazané v kolóne sa eluovali s 1 ml/min s lineárnym gradientom od 0 do 1 M NaCl v 20 mM Tris-HCl, pH 7,0 počas 30 min.

Vyššie zreagované frakcie s protilátkou pAb TcaB<sub>ii</sub>-syn sa vylúhovali a vložili sa do kolóny fenyl-Superose 5/5. Do konečnej koncentrácie 1,7 M sa dodal tuhý (NH<sub>4</sub>)SO<sub>4</sub>. Roztok sa potom aplikoval do kolóny zrovnováženej s 1,7 (NH<sub>4</sub>)SO<sub>4</sub> v 50 mM fosforečnanu draselného, pH 7,0 do 10 ml fosforečnanu draselného, pH 7,0 pri 0,5 ml/min počas 60 min. Frakcie sa dialyzovali cez noc proti pufru 10 mM fosforečnanu draselného, pH 7,0.

Pre posledný purifikačný krok sa vyššie stanovené frakcie pomocou Western analýzy, ktoré reagovali s protilátkou pAb TcaB<sub>ii</sub>-syn, kombinovali a aplikovali na kolónu Mono Q 5/5 zrovnováženú s 20 mM Tris-HCl, pH 7,0 pri 1 ml/min. Proteíny viazané v kolóne sa eluovali pri 1 ml/min lineárnym gradientom od 0 do 1 M NaCl v 20 mM Tris-HCl, pH 7,0 počas 30 min.

Posledná purifikovaná proteínová frakcia obsahovala 6

hlavných peptidov vyhodnotených SDS-PAGE: 165 kDa, 90 kDa, 64 kDa, 62 kDa, 58 kDa a 22 kDa. LD50 insekticídnej aktivity tejto purifikovanej frakcie sa stanovilo na 100 ng a 500 ng oproti SCR alebo ECB.

Vyššie uvedené peptidy sa naniesli na membrány PVDF a plaky sa odoslali na analýzu aminokyselín a päť aminokyselín dlhého N-koncového sekvencovania na Harvard MicroChem alebo na Cambridge Prochem. Sekvencia N-koncovej aminokyseliny peptidu 165 kDa sa stanovila, že je identická s peptidom TcaC (SEKV. ID. Čís. 2, pozícia 1 až 5). Sekvencia N-koncovej aminokyseliny peptidu 90 kDa sa stanovila, že je TcaA<sub>ii</sub> oblasťou odvodenej sekvencie aminokyseliny pre *tcaA* gén (SEKV. ID. Čís. 33, pozícia 254 až 258). Sekvencia N-koncovej aminokyseliny peptidu 64 kDa sa stanovila, že je identická s peptidom TcaB<sub>i</sub> (SEKV. ID. Čís. 3, pozícia 1 až 5). Sekvencia N-koncovej aminokyseliny peptidu 62 kDa sa stanovila, že je TcaA<sub>ii</sub> oblasťou odvodenej sekvencie aminokyseliny pre *tcaA* gén (SEKV. ID. Čís. 33, pozícia 489 až 493). Sekvencia N-koncovej aminokyseliny peptidu 58 kDa sa stanovila, že je identická s peptidom TcaB<sub>ii</sub> (SEKV. ID. Čís. 5, pozícia 1 až 5). Sekvencia N-koncovej aminokyseliny peptidu 22 kDa (SEKV. ID. Čís. 62) sa stanovila, že je TcaA<sub>i</sub> oblasťou, označenou ako TcaA<sub>iv</sub>, odvodenej sekvencie aminokyseliny pre *tcaA* gén (SEKV. ID. Čís. 34, pozícia 98 až 102). Je potrebné poznamenať, že *tcaA*, *tcaB* a *tcaC* gény pochádzajú z toho istého *tca* operónu (obr. 6A).

Päť µg purifikovanej Tca frakcie, purifikovaný toxín A a purifikovaný toxín B sa analyzovali pomocou Western blot s použitím nasledovných protilátok individuálne ako primárnej protilátky: pAb TcaB<sub>ii</sub>-syn protilátka, mAb CF52 protilátka, pAb TcdA<sub>ii</sub>-syn protilátka a pAb Tcd<sub>iii</sub>-syn protilátka (Príklad 21). S TcaB<sub>ii</sub>-syn protilátkou reagovali iba purifikované frakcie Tca peptidov, ale nie toxín A alebo B. S protilátkou mAb CF52 reagoval iba toxín B, ale nie frakcia Tca peptidov alebo toxín A. Buď s pAb TcdA<sub>ii</sub>-syn protilátkou alebo pAb Tcd<sub>iii</sub>-syn protilátkou

reagoval iba toxín A, ale nie toxín B alebo frakcia Tca peptidov. To ukazuje, že insekticídna aktivita pozorovaná v purifikovanej frakcii Tca peptidov je nezávislá na toxíne A a toxíne B. Purifikovaná frakcia Tca peptidov je tretím špecifickým proteínovým toxínom, ktorý je označený ako toxín C.

#### Príklad 16

#### Štiepenie a aktivácia TcbA peptidu

V toxínovom komplexe B, vznikajú peptidy TcbA<sub>ii</sub> a TcbA<sub>iii</sub> z jednoduchého génového produktu TcbA (príklad 15). Upravovanie peptidu TcbA na TcbA<sub>ii</sub> a TcbA<sub>iii</sub> je prevažne činnosťou proteázy (-áz) *Photorhabdus* a najpravdepodobnejšie metaloproteázami opísanými v príklade 10. V niektorých prípadoch, keď bol bujón *Photorhabdus* W-14 upravovaný, sa zaznamenalo, že peptid TcbA bol prítomný v toxínovom komplexe ako hlavná zložka, okrem peptidov TcbA<sub>ii</sub> a TcbA<sub>iii</sub>. Rovnaké postupy, ktoré sa opísali pre purifikáciu toxínového komplexu B (príklad 15), sa použili na obohatenie peptidu TcbA z toxínovej komplexovej frakcie bujónu z W-14. Konečný purifikovaný materiál sa analyzoval v 4 až 20 % gradiente SDS-PAGE a hlavné peptidy sa kvantifikovali denzitometricky. Určilo sa, že TcbA, TcbA<sub>ii</sub> a TcbA<sub>iii</sub> obsahovali 58 %, 36 % a prípadne 6 % z celkového proteínu. Identita týchto proteínov sa potvrdila ich príslušnými molekulovými veľkosťami na SDS-PAGE a analýzou Western blot s použitím monošpecifických protilátok. Pôvodné molekulové hmotnosti tejto frakcie sa určili na 860 kDa.

Štiepenie TcbA sa hodnotilo po reakcii vyššie purifikovaného materiálu s purifikovanými 38 kDa a 54 kDa *Photorhabdus* metaloproteázami W-14 (príklad 10), a trypsínom ako kontrolným enzýmom (Sigma, MO). Štandardná reakcia pozostávala z 17,5 µg vyššie purifikovanej frakcie, 1,5 jednotky proteázy a 0,1 M Tris pufru s pH = 8,0 v celkovom objeme 100 µl. V kontrolnej reakcii sa vynechala proteáza. Reakčná zmes sa

inkubovala pri 37 °C počas 90 min. Na konci reakcie sa zobralo 20 µl, ktoré sa varili s pufrom vzorky SDS-PAGE, hneď na elektroforetickú analýzu v 4 až 20 % gradiente SDS-PAGE. Určilo sa z SDS-PAGE, že v obidvoch proteázových prípravkoch 38 kDa a 58 kDa, vzrástlo množstvo peptidov TcbA<sub>ii</sub> a TcbA<sub>iii</sub> asi 3-násobne, zatiaľ čo množstvo peptidu TcbA proporcionálne pokleslo (tabuľka 24). Relatívna redukcia a zvýšenie vybraných peptidov sa potvrdila analýzou Western blot. Gélová filtrácia štiepeného materiálu ďalej odhalila, že pôvodná molekulová hmotnosť komplexu zostala rovnaká. V trypsínovej úprave boli peptidy TcbA a TcbA<sub>ii</sub> boli nešpecificky rozložené na malé peptidy. Toto indikuje, že proteázy *Photorhabdus* 38 kDa a 58 kDa môžu špecificky premeniť peptid TcbA na peptidy TcbA<sub>ii</sub> a TcbA<sub>iii</sub>. Proteázy upravovanej a neupravovanej kontroly zostatkovej 80 µl reakčnej zmesi sa sériovo riedili 10 mM pufrom fosforečnanu sodného s pH = 7,0 a analyzovali biotestom na SCR. Porovnaním účinnosti niekoľkých riedení sa určilo, že 38 kDa proteázový prípravok zvýšil insekticídnu aktivitu proti SCR približne 3 až 4 násobne. Rast inhibície ostatného hmyzu v proteázovom prípravku bol tiež prudší ako v kontrole (tabuľka 24).

Tabuľka 24

Konverzia a aktivácia TcbA peptidu do peptidov TcbA<sub>ii</sub> a TcbA<sub>iii</sub> proteázovým prípravkom.

	Kontrola	proteázový prípravok 38 kDa
TcbA (% celkového proteínu)	58	18
TcbA <sub>ii</sub> (% celkového proteínu)	36	64
TcbA <sub>iii</sub> (% celkového proteínu)	6	18
LD50 (µg proteínu)	2,1	0,52
Hmotnosť SCR (mg/hmyz)*	0,2	0,1

\*: indikácia inhibície rastu meraním priemernej hmotnosti živého hmyzu po 5 dňoch skrmovania diéty v teste.

#### Aktivácia a spracovanie toxínu B SCR črevnou proteázou

Už v druhej demonštrácii proteolytickej aktivácie sa skúšalo, či sa toxíny W-14 hmyzom spracovávajú. Toxín B purifikovaný z bujónu *Photorhabdus* W-14 (pozri príklad 15) bol prítomný hlavne v intaktných TcbA peptidoch, ako sa zistilo pomocou SDS-PAGE a Western blot analýzou s použitím monoklonálnej protilátky. LD50 tejto frakcie oproti SCR sa stanovilo na okolo 700 ng.

SCR larvy rástli na potrave pre coleoptera do dosiahnutia štvrtého vývojového štádia (okolo 100 až 125 mg celkovej hmotnosti každého jedinca hmyzu). Obsah čriev SCR sa zisťoval nasledovne: črevá sa odstránili s použitím odstraňovacích nožníc a pinziet. Po odstránení zvyšku tukového materiálu, ktorý pokrýval steny čriev, sa asi 40 čriev homogenizovalo v mikrocentrifugovej trubici s obsahom 100  $\mu$ l sterilnej vody. Trubica sa potom centrifugovala pri 14 000 rpm počas 10 min a peleta sa oddelila. Supernatant sa uskladňoval pri  $-70$  °C v mrazničke do použitia.

Spracovanie toxínu B hmyzím črevom sa zistilo spracovaním vyššie purifikovaného toxínu B so zozbieraným obsahom čriev SCR. Reakcia pozostávala zo 40  $\mu$ g toxínu B (1 mg/ml), 50  $\mu$ l obsahu čriev SCR a 0,1 M Tris pufra, pH 8,0 v celkovom objeme 100  $\mu$ l. Pri kontrolnej skúške sa vynechal obsah čriev SCR. Reakčná zmes sa inkubovala cez noc pri 37 °C. Na konci reakcie sa vybral 1  $\mu$ l a varil sa s rovnakým objemom vzorky 2 x SDS-PAGE pufra pre SDS-PAGE analýzu. Zvyšných 90  $\mu$ l reakčnej zmesi sa postupne rozpustilo s 10 mM pufra fosforečnanu sodného a analyzovalo sa SCR bioskúškou. SDS-PAGE analýza ukázala spracovanie v SCR obsahu čriev, peptid TcbA bol kompletne rozštiepený na menšie peptidy. Analýza nedenaturovaných toxínových frakcií ukázala, že natívna-



veľkosť okolo 860 kDa, zostala rovnaká, aj keď dlhšie peptidy boli fragmentované. V SCR bioanalýzach sa zistilo, že LD50 toxínu B smerovaného SCR čriev bol okolo 70 ng; čo predstavuje 10 násobný nárast. V oddelenom experimente spracovanie proteázou K kompletne eliminovalo aktivitu toxínu.

#### Príklad 17

#### Skríning knižnice pre génové kódovanie peptidu TcdAii

Dokončilo sa klonovanie a charakteristika kódovania génu peptidu TcdA<sub>ii</sub>, ktorý je opísaný ako SEKV. ID. Čís. 17 (vnútorný peptid TcdA<sub>ii</sub>-PT111 N-koncovej sekvencie) a SEKV. ID. Čís. 18 (vnútorný peptid TcdA<sub>ii</sub>-PT79 N-koncová sekvencia). Dve zmesi degenerovaných oligonukleotidov, na označenie kódovania aminokyselinových sekvencií SEKV. ID. Čís. 17 (tabuľka 25) a SEKV. ID. Čís. 18 (tabuľka 26) a reverzné zložky týchto sekvencií sa syntetizovali a opísali v príklade 8. DNA sekvencia oligonukleotidov je uvedená nižšie:

#### Tabuľka 25

#### Degenerované oligonukleotidy pre SEKV. ID. Čís. 17

P2-PT111	1	2	3	4	5	6	7	8
Aminokyselina	Ala	Phe	Asn	Ile	Asp	Asp	Val	Ser
Kodóny	5' GCN	TT(T/C)	AA(T/C)	AT(T/C/A)	GA(T/C)	GA(T/C)	GTN 3'	
P2.3.6.CB	5' GC(A[C[G{T)	TT(T/C)	AAT	ATT	GAT	GAT	GT 3'	
P2.3.5	5' GC(A[C[G{T)	TT(T/C)	AA(T/C)	AT(T/C/A)	GA(T/C)	GA(T/C)	GT 3'	
P2.3.5R	5' AC	(G/A)TC	(G/A)TC	(T/G/A)AT	(G/A)TT	(G/A) AA	(A/C/G/T) GC 3'	
P2.3.5RI	5' ACI	TCI	TCI	ATI	TTI	AAI	GC 3'	
P2.3R.CB	5'CAG	(A/G)CT	(A/C)AC	ATC	ATC	AAT	ATT	AAA 3'

Tabuľka 26

Degenerované oligonukleotidy pre SEKV. ID. ČÍS. 18

P2-PT79	1	2	3	4	5	6	7	8	9	10	11	12	13
Aminokyselina	Phe	Ile	Val	Tyr	Thr	Ser	Leu	Gly	Val	Asn	Pro	Asn	Asn
Kodóny*	5'TTY	ATH	GTN	TAY	ACN	6	6	GGN	GTN	AAV	CCN	AAV	AAV 3'
P2.79.2	5'TTY	ATV	GTK	TAT	ACV	TCI	YTR	GGV	GTK	AAT	CCR	AAT	AAT 3'
P2.79.3	5'TTT	ATT	GTK	TAT	ACV	AGV	YTR	GGV	GTK	AAT	CCR	AAT	AAT 3'
P2.79.R.1	5'ATT	ATT	YGG	ATT	MAC	RCC	YAR	RCT	RGT	ATA	MAC	AAT	AAA 3'
P2.79R.CB	5'ATT	ATT	YGG	ATT	MAC	ACC	CAG	RCT	GGT	ATA	MAC	AAT	AAA 3'

\* Podľa IUPAC - IUB kódov nukleotidov: Y = C alebo T, H = A, C alebo T, N = A, C, G alebo T, K = G alebo T, R = A alebo G a M = A alebo C

Polymerázové reťazové reakcie (PCR) sa uskutočňovali v podstate tak, ako je to opísané v príklade 8, s použitím pôvodných primérov P2.3.6.CB alebo P2.3.5. a ako reverzné priméry P2.79.R.1 alebo P2.79R.CB vo všetkých pokračujúcich/reverzných kombináciách, použitím genómovej DNA *Photorhabdus* W-14 ako templátu. V ďalšej sade reakcií sa použili priméry P2.79.2 alebo P2.79.3 ako postačujúce priméry a P2.3.5R, P2.3.5RI a P2.3R.CB ako reverzné priméry vo všetkých pokračujúcich/reverzných kombináciách. Len v reakciách obsahujúcich P2.3.6.CB ako pokračujúce priméry kombinované s P2.79.R.1 alebo s P2.79R.CB ako reverznými priméromi sa objavili neartefaktické amplifikované produkty, s odhadovanou veľkosťou (pohyblivosť na agarózovom géle) 2500 párov báz. Poradie primérov použitých na zistenie tohto produktu amplifikácie indikuje, že peptidový fragment TcdA<sub>ii</sub>-PT111 leží amino-proximálne k peptidovému fragmentu TcdA<sub>ii</sub>-PT79.

Produkty PCR 2500 bp sa ligátovali na plazmidový vektor pCR<sup>TM</sup>II (Invitrogen, San Diego, CA) podľa dodávateľových inštrukcií, a DNA sekvencie sa determinovali cez konce fúzovaných fragmentov dvoch izolátov (HS24 a HS27) použitím dodávateľom doporučených primérov a sekvencovacích metód, ktoré sa opísali predtým. Sekvencie obidvoch izolátov boli rovnaké. Nové priméry sa syntetizovali na základe určenia sekvencie a použili sa na hlavné doplnkové sekvencovacie reakcie na to, aby sa získalo celkovo 2557 báz inzertu (SEKV. ID. ČÍS. 36). Translácia čiastkových peptidov kódovaných ako SEKV. ID. ČÍS. 36 poskytuje 845 aminokyselinových sekvencií opísaných ako SEKV. ID. ČÍS. 37. Analýza proteínovej homológie tejto časti peptidového fragmentu TcdA<sub>ii</sub> odhaľuje substanciálnu aminokyselinovú homológiu (68 % podobnosti; 53 % identity) k zvyšku 542 až 1390 proteínu TcbA (SEKV. ID. ČÍS. 12). Je preto zrejmé, že gén predstavovaný v časti SEKV. ID. ČÍS. 36 produkuje proteín podobný, ale nie identický aminokyselinovej sekvencii ako je u TcbA proteínu, a ktorý má pravdepodobne podobnú, ale nie identickú biologickú účinnosť ako proteín TcbA.

V inom príklade sa izoloval gén kódujúci peptidy TcdA<sub>ii</sub>-PK44 a TcdA<sub>iii</sub> N-koncového 58 kDa peptidu, opísané ako SEKV. ID. ČÍS. 39 (vnútorný peptid TcdA<sub>ii</sub>-PK44 sekvencie) a SEKV. ID. ČÍS. 41 (TcdA<sub>iii</sub> 58 kDa N-koncovej peptidovej sekvencie). Dve zmesi degenerovaných oligonukleotidov na označenie kódovania aminokyselinových sekvencií, opísaných ako SEKV. ID. ČÍS. 39 (tabuľka 28) a SEKV. ID. ČÍS. 41 (tabuľka 27) a reverzné doplnky týchto sekvencií sa syntetizovali tak, ako je opísané v príklade 3, ich DNA sekvencie.

Tabuľka 27

Degenerované oligonukleotidy pre SEKV.ID.ČÍS.41

Kodón č.	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Aminokyselina	Leu	Arg	Ser	Ala	Asn	Thr	Leu	Thr	Asp	Leu	Phe	Leu	Pro	Gln
A2.1	5' YTR	CGY	AGY	GCI	AAT	ACY	YTR	ACY	GAT	YTR	TTT	YTR	CCR	CA 3'
A2.2				GCI	AAT	ACI	YTR	ACI	GAY	YTR	TTY	YTR	CCI	CA 3'
A2.3.R		5' TG	YGG	YAR	AAA	YAR	RTC	RGT	YAR	RGT	RTT	IGC	RCT	RCG 3'
A2.4.R				5' TG	IGG	CAG	AAA	CAG	RTC	IGT	CAG	IGT	ATT	ICG 3'

Tabuľka 28

Degenerované oligonukleotidy pre SEKV.ID.ČÍS.39

Aminokyselina č.	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Kodón č.	1	2	3	4	5	6	7	8	9
Aminokyselina	Gly	Pro	Val	Glu	Ile	Asn	Thr	Ala	Ile
A1.44.1	5' GGY	CCR	GTK	GAA	ATT	AAT	ACC	GCI	AT 3'
A1.44.1R	5' ATI	GCG	GTA	TTA	ATT	TCM	ACY	GGR	CC 3'
A1.44.2	5'GGI	CCI	GTI	GAR	ATY	AAV	ACI	GCI	AT 3'
A1.44.2R	5' ATI	GCI	GTR	TTR	ATY	TCI	ACI	GGI	CC 3'

Polymerázové reťazové reakcie (PCR) sa pripravili v podstate tak, ako sa opísalo v príklade 8, použitím ako pokračujúceho priméru A1.44.1 alebo A1.44.2 a ako reverzných primérov A2.3R alebo A2.4R vo všetkých pokračujúcich/reverzných kombináciách s použitím genómovej DNA *Photorhabdus* W-14 ako

templátu. V inej sade reakcií sa použili ako pôvodné priméry A2.1 alebo A2.2 a A1.44.1R a A1.44.2R sa použili ako revezné priméry vo všetkých pokračujúcich/reverzných kombináciách. Len v reakciách obsahujúcich A1.44.1 alebo A1.44.2 ako pokračujúce priméry kombinované s A2.3R ako spätného priméru sa objavil neartefaktický amplifikačný produkt, s odhadovanou veľkosťou (mobilita na agarových géloch) 1400 párov báz. Poradie primérov použitých na zistenie tohto amplifikačného produktu indikuje, že peptidový fragment TcdA<sub>ii</sub>-PK44 leží amino-proximálne k 58 kDa peptidovému fragmentu TcdA<sub>iii</sub>.

PCR produkty 1400 bp sa ligátovali k plazmidovému vektoru pCR<sup>TM</sup> podľa dodávateľových inštrukcií. Určili sa sekvencie DNA cez koniec inzertovaných fragmentov štyroch izolátov pomocou primérov podobných v sekvencii k dodávateľom odporučených primérov a použitím metód sekvencovania, ktoré sa opísali predtým. Sekvencia nukleových kyselín všetkých izolátov sa líšila tak, ako sa očakávalo, v oblastiach zhodujúcich sa s degenerovanými primérovými sekvenciami, ale aminokyselinová sekvencia odvodená z týchto údajov bola rovnaká ako skutočné sekvencie aminokyselín peptidov, ktoré sa determinovali predtým (SEKV. ID. ČÍS. 41 a 39).

Skríning genómovej kozmidovej knižnice W-14 tak, ako je opísaná v príklade 8, s rádioaktívne značenou vzorkou, ktorá obsahovala DNA pripravenú vyššie (SEKV. ID. ČÍS. 36) identifikoval päť hybridizovaných kozmidových izolátov, menovite 17D9, 20B10, 21D2, 27B10 a 26D1. Tieto kozmidy boli odlišné od tých, ktoré boli predtým identifikované so vzorkami zhodujúcimi sa s génmi opísanými ako SEKV. ID. ČÍS 11 alebo SEKV. ID. ČÍS. 25. Analýzy reštrikčného enzýmu a DNA plakových hybridizácií identifikovali tri *EcoR I* fragmenty s približnou veľkosťou 3,7; 3,7 a 1,1 kbp, ktoré zasahovali oblasti obsahujúce DNA SEKV. ID. ČÍS. 36. Skríning genómovej kozmidovej knižnice W-14 s použitím ako rádioaktívne značenej sondy 1,4 kbp DNA fragmentu pripraveného v tomto príklade identifikoval päť rovnakých

kozmidov (17D9, 20B10, 21D2, 27B10 a 26D1). DNA blot hybridizácia k *EcoR I* - štiepených kozmidových DNA tiež preukázala hybridizáciu na rovnaké podmnožiny *EcoR I* fragmentov tak, ako je to vidieť s 2,5 kbp  $TcdA_{ii}$  génovou sondou, ktorá ukazuje, že obidva fragmenty sú kódované na genómovej DNA.

Determinácia sekvencie DNA klonovaných *EcoR I* fragmentov odhalila neprerušný čítací rámec z 7551 párov báz (SEKV. ID. ČÍS. 46) kódujúci 282,9 kDa proteín z 2516 aminokyselín (SEKV. ID. ČÍS. 47). Analýza aminokyselinovej sekvencie tohto proteínu odhalila všetky očakávané vnútorné fragmenty peptidov  $TcdA_{ii}$  (SEKV. ID. ČÍS. 17, 18, 37, 38 a 39) a N-koniec  $TcdA_{iii}$  peptidu (SEKV. ID. ČÍS. 41) a všetky  $TcdA_{iii}$  vnútorné peptidy (SEKV. ID. ČÍS. 42 a 43). Peptidy izolované a identifikované ako  $TcdA_{ii}$  a  $TcdA_{iii}$  sú každý produktami otvoreného čítacieho rámca, označeného *tcdA*, opísané ako SEKV. ID. ČÍS. 46. SEKV. ID. ČÍS. 47 ďalej ukazuje štart na pozícii 89, sekvencie opísanej ako SEKV. ID. ČÍS. 13, ktorá je N-koncovou sekvenciou peptidu s veľkosťou približne 201 kDa, pričom indikuje, že iniciačný proteín produkovaný SEKV. ID. ČÍS. 46 je upravovaný podobne, ako predchádzajúci opísaný pre SEKV. ID. ČÍS. 12. Proteín sa ďalej štiepil, aby vznikol produkt s veľkosťou 209,2 kDa, kódovaný v SEKV. ID. ČÍS. 48 a opísaný ako SEKV. ID. ČÍS. 49 ( $TcdA_{ii}$  peptid) a produkt s veľkosťou 63,6 kDa, kódovaný v SEKV. ID. ČÍS. 50 a opísaný ako SEKV. ID. ČÍS. 51 ( $TcdA_{iii}$  peptid). Uvažovalo sa o tom, že insekticídna aktivita identifikovaná ako toxín A (príklad 15) pochádzala z produktov SEKV. ID. ČÍS. 46. Ako príklad sa uvádza proteín s plnou dĺžkou 282,9 kDa opísaný ako SEKV. ID. ČÍS. 47, sa upravoval na produkciu peptidov opísaných ako SEKV. ID. ČÍS. 49 a 51. Uvažovalo sa o tom, že insekticídna aktivita identifikovaná ako toxín B (príklad č. 15) pochádza z produktov SEKV. ID. ČÍS. 11, ako príklad sa uviedol 280,6 kDa proteín opísaný ako SEKV. ID. ČÍS. 12. Tento proteín je proteolyticky spracovaný, čo viedlo k 207,6 kDa peptidu opísanému ako SEKV. ID. ČÍS. 53, ktorý je kódovaný SEKV. ID. ČÍS. 52 a 62,9 kDa peptidu s N-koncovou sekvenciou opísanou ako SEKV. ID.

Čís. 40 a ďalej opísanou ako SEKV. ID. Čís. 55, ktorá je kódovaná v SEKV. ID. Čís. 54.

Porovnanie aminokyselinovej sekvencie medzi proteínmi opísanými ako SEKV. ID. Čís. 12 a SEKV. ID. Čís. 47 odhalilo, že majú 69 % podobnosť a 54 % identitu s použitím Wisconsin Package Version 8.0, Genetics Computer Group (GCG), Madison, WI, alebo 60 % podobnosť a 54 % identitu s použitím programu verzie 9.0. Tento vysoký stupeň vývojového príbuzenstva nie je rovnaký po celej dĺžke aminokyselinovej sekvencie týchto peptidov, ale je vyšší smerom ku karboxy-zakončenému koncu proteínov, zatiaľ čo peptidy opísané ako SEKV.ID.Čís. 51 (pochádzajúca zo SEKV. ID. Čís. 47) a SEKV. ID. Čís. 55 (pochádzajúca zo SEKV.ID.Čís. 12) majú 76 % podobnosť a 64 % identitu s použitím Wisconsin Package Version 8.0, Genetics Computer Group (GCG), Madison, WI alebo 71 % podobnosti a 64 % identity s použitím programu verzie 9.0.

#### Príklad 18

Eliminácia vijačky kukuričnej, ktorá indukuje listové poškodenie na rastlinách kukurice aplikáciou postreku bujónu z *Photorhabdus* (kmeň W-14).

Schopnosť *Photorhabdus* toxínu (-ov) redukovať poškodenie zapríčinené larvami hmyzu sa demonštrovalo meraním poškodenia listu zapríčineného vijačkou kukuričnou (*Ostrinia nubilalis*) infestovanou do rastlín kukurice ošetrovaných bujónom *Photorhabdus*. Urobila sa fermentácia bujónu *Photorhabdus*, kmeňa W-14 a koncentrovala sa približne 10 krát použitím ultrafiltrácie (10 000 MW veľkosť pórov) tak, ako je to opísané v príklade 13. Výsledný koncentrovaný bujón sa potom sterilizoval filtráciou na 0,2 mikrónomovom nitrocelulóзовom membránovom filtri. Podobne pripravená vzorka neinokulovaného 2 % proteózového peptónu č. 3 sa použila ako kontrola. Rastliny kukurice (DowElanco patentová odroda) rástli zo semena do vegetatívneho štádia 7 alebo 8 v hrncoch obsahujúcich zmes pôdy v skleníku (27 °C deň, 22 °C

noc, okolo 50 % relatívna vlhkosť, 14 hod. dĺžka dňa, zalievanie/hnojenie podľa potreby). Testované rastliny sa upravili náhodne do úplných označených blokov (3 paralelky/prípravok, 6 rastlín/prípravok) v skleníku s teplotou okolo 22 °C cez deň, 18 °C v noci, bez umelého osvetlenia a s čiastočným zatienením, okolo 50 % relatívna vlhkosť a zalievanie/hnojenie podľa potreby. Prostriedky (neinokulované médiá a koncentrovaný bujón *Photorhabdus*) sa aplikovali rozstrekovacou injekciou, 2 ml sa aplikovali priamo (okolo 2 palcov) nad stonkou a ďalšie 2,0 ml sa aplikovali cirkulárnym pohybom z približne jednej stopy nad stonkou. Jedna skupina rastlín nedostala žiaden prípravok. Po aplikácii prípravkov sa vysušili (asi 30 minút) a dvanásť neonátov vijačky kukuričnej (vajíčka sa získali z komerčných zdrojov a vyliahli sa v laboratóriu) sa aplikovalo priamo na stonku. Po jednom týždni sa odčítalo na rastlinách poškodenie listov použitím modifikovanej Guthriho škály (Koziel, M.G., Beland, G.L., Bowman, C., Carozzi, N.B., Crenshaw, R., Crossland, L., Dawson, J., Desai, N., Hill, M., Kadwell, S., Launis, K., Lewis, K., Maddox, D., McPherson, K., Meghji, M.Z., Merlin, E., Rhodes, R., Waren, G.W., Wright, N. a Evola, S.V., 1993).

Bio/technology, 11, 194-195 a údaje sa porovnávali štatisticky [T-test (LSD)  $p < 0,05$  a Tukeyho Studentovo rozdelenie (HSD) test  $p < 0,1$ ]. Výsledky sú uvedené v tabuľke 29. Skóre 1 nepredstavuje žiadne poškodenie, skóre 2 predstavuje jemné poškodenie tzv. "okenná tabuľa" na rozvinutom liste bez dierkovej penetrácie a skóre 5 predstavuje penetráciu listu s predĺženým poškodením a/alebo požieranie stredného rebra listu evidentného na viac ako troch listoch (poškodenie < 1 palec). Tieto údaje indikujú, že bujón alebo iné proteíny obsahujúce frakcie ochrániť proti špecifickým hmyzím škodcom, keď sa nanesú formou spraya, alebo keď gény alebo ich deriváty, ktoré kódujú proteín alebo jeho časti, sú dodané prostredníctvom transgéennej rastliny alebo mikróbom.



Tabuľka 29

Účinok bujónovej kultúry *Photorhabdus* na vijačku kukuričnú - indukujúcu poškodenie listu kukurice.

Prípravok	Priemer Goutriho skóre
Bez prípravku	5,02 <sup>a</sup>
Neinokulované médium	5,15 <sup>a</sup>
<i>Photorhabdus</i> bujón	2,24 <sup>b</sup>

Priemery s rozdielnymi číslicami sú štatisticky významné ( $p < 0,05$  alebo  $p < 0,1$ ).

#### Príklad 19

Genetické inžinierstvo génov na expresiu do *E. coli*

#### Zhrnutie konštrukcií

Zostrojovala sa séria plazmidov na expresiu génu *tcbA* *Photorhabdus* kmeňa W-14 do *Escherichia coli*. Zoznam plazmidov je uvedený v tabuľke 30. Nasleduje krátky opis každej konštrukcie ako aj zhrnutie získaných údajov z expresie *E. coli*.

Tabuľka 30

Expresné plazmidy pre gén *tcba*

Plazmid	Gén	Vektor/Selekcia	Oddelenie
pDAB2025	<i>tcba</i>	pBC/Chl	Intracelulárna
pDAB2026	<i>tcba</i>	pAcGP67B/Amp	Baculovírus sekretovaný
pDAB2027	<i>tcba</i>	pET27b/Kan	Periplazma
pDAB2028	<i>tcba</i>	pET15- <i>tcba</i>	Intracelulárna

Skratky: Kan = Kanamycín, Chl = chloramfenikol, Amp = Ampicilín

#### Konštrukcia pDAB2025

V príklade 9 sa opisuje veľký fragment EcoR I, ktorý hybridizuje k sonde  $TcbA_{ii}$ . Tento fragment sa subklonoval do pBC (Stratagene, LaJolla CA) na vytvorenie pDAB2025. Sekvenčná analýza indikuje, že tento fragment má 8816 párov báz. Fragment kóduje gén *tcba*, s iniciáciou ATG na pozícii 571 a termináciou TAA na pozícii 8086. Fragment preto nesie 570 párov báz DNA *Photorhabdus* proti smeru transkripcie od ATG a 730 párov báz v smere transkripcie od TAA.

#### Konštrukcia plazmidu pDAB2026

Gén *tcba* sa amplifikoval v PCR z plazmidu pDAB2025 s nasledujúcimi primérmami: 5' primér (S1Ac51) 5' TTT AAA CCA TGG GAA ACT CAT TAT CAA GCA CTA TC 3' a 3' primér (S1Ac31) TTT AAA GCG GCC GCT TAA CGG ATG GTA TAA CGA ATA TG 3'. PCR sa pripravila zo súpravy TaKaRa LA PCR z PanVera (Madison, WI) v nasledujúcej reakcii: 57,5 ml vody, 10 ml 10X pufu LA, 16 ml dNTP (2,5 mM z každého zásobného roztoku), 20 ml z každého priméru

s 10 pmolov/ml, 300 ng plazmidu pDAB2025 obsahujúceho gén *tcbA* W-14 a jeden ml polymerázy TaKaRa LA Taq. Podmienky cyklovania boli 98 °C/20 s, 68 °C/5 min., 72 °C/10 min. v 30 cykloch. Očakávaný produkt PCR okolo 7526 bp sa izoloval na 0,8 % agarózovom géli v pufri TBE (100 mM Tris, 90 mM kyseliny bórovej, 1 mM EDTA) a purifikoval sa súpravou Qiaex II (Chatsworth, CA). Purifikovaný gén *tcbA* sa štiepil s *Nco I* a *Not I* a ligátoval do baculovírusového transferového vektora pAcGP67B (PharMingen (San Diego, CA) a transformovali sa do DH5α *E. coli*. Výsledný rekombinant sa nazval pDAB2026. Gén *tcbA* sa potom odstrihol z pDAB2026 a transferoval do pET27b na vytvorenie plazmidu pDAB2027. Mutácia meniaci zmysel kodónu v géne *tcbA* sa opravila v pDAB2027.

Opravený gén *tcbA* obsahuje dve zmeny oproti sekvencii, ktorá je znázornená v sekvencii SEKV. ID. Čís. 11. Výmenu A > G na 212, pričom sa mení asparagín 71 na serín 71 a výmenu G > A na 229, pričom sa mení alanín 77 na treonín 77. Obe tieto zmeny sú proti smeru transkripcie navrhnutého N-konca TcbA<sub>ii</sub>.

#### Konštrukcia pDAB2028

*tcbA* kódujúci oblasť pDAB2027 sa transferoval do vektora pET15b. Tento proces sa vykonal vynútenými ligáciami, pričom DNA bola odstrihnutá reštrikčnými enzýmami *Nco I* a *Xho I*. Výsledný rekombinant sa nazval pDAB2028.

#### Expresia TcbA do *E. coli* z plazmidu pDAB2028

Expresia *tcbA* do *E. coli* sa získala modifikáciou metód, ktoré sa opísali predtým v Studier a kol. (Studier, F.W., Rosenberg, A., Dunn, J. a Dubendorff, J., (1990). Využitie T7 RNA polymerázy na priamu expresiu klonovaných génov. *Methods. Enzymol.*, 185, str. 60-89.). Príslušné bunky *E. coli* kmeňa BL21 (DE3) sa transformovali s plazmidom pDAB2028 a naočkovali na LB agar obsahujúci 100 µg/ml ampicilínu a 40 mM glukózy.

Transformované bunky sa rozotreli na miskú s hustotou niekoľko sto izolovaných kolónií na miskú. Po celonočnej inkubácii pri 37 °C sa bunky odobrali z misiek a suspendovali v médiu LB, ktoré obsahovalo 100 µg/ml ampicilínu. Typickými objemami kultúry boli od 200 do 500 ml. V čase nula boli hustoty kultúr (OD600) 0,05 až 0,15, v závislosti od experimentu. Kultúry sa trepali pri jednej z troch teplôt (22 °C, 30 °C alebo 37 °C) až kým sa nedosiahla hustota od 0,15 do 0,5. V tomto čase sa indukovali s 1 mM izopropiltio-β-galaktozidom (IPTG). Kultúry sa inkubovali pri určenej teplote počas 4 až 5 hodín a potom sa preniesli do 4 °C až pokiaľ sa nezačali spracovávať (12 až 72) hod.

Purifikácia a charakterizácia TcbA expresovaného do *E.coli* z plazmidu pDAB2028

Kultúry *E. coli* expresované peptidmi TcbA sa spracovávali nasledovne. Bunky sa odobrali centrifugáciou pri 17 000 x g, médiá sa dekantovali a uchovávali sa v oddelených nádobách.

Médiá sa koncentrovali asi 8 x na filtračnom systéme M12 (Amicon, Beverly MA) a molekulová hmotnosť 100 kDa sa oddelila filtračne. Koncentrované médiá sa nanášali na anionovú kolónu a zachytené proteíny sa eluovali s 1,0 M NaCl. Zistilo sa, že elučný pík 1,0 M NaCl zapríčiňoval mortalitu lariev liskavky uhorkovej (SCR) larvy (tabuľka 30). Frakcia 1,0 M NaCl sa dialyzovala oproti 10 mM pufru fosforečnanu sodnému s pH = 7,0. Koncentrovala sa a podrobila sa gélovej filtrácii na Sepharose CL-4B (Pharmacia, Piscataway, New Jersey). Oblasť elučného profilu CL-4B, ktorá bola s súhlase s vypočítanou molekulovou hmotnosťou (okolo 900 kDa) takou, akú má natívny toxínový komplex W-14, sa zozbierala, koncentrovala a hodnotila sa v biotestoch proti larvám. Zistilo sa, že táto zozbieraná frakcia 900 kDa je insekticídne účinná (pozri tabuľku 31 nižšie), s podobnými symptómami aké zapríčiňoval natívny toxínový komplex W-14. Táto frakcia sa vystavila pôsobeniu proteínázy K a tepelnej úprave a účinnosť v obidvoch prípadoch bola buď eliminovaná alebo

redukovaná, poskytujúc tak dôkaz, že účinný je natívny proteín. Okrem toho účinná frakcia bola testovaná imunologicky pozitívne pre peptidy TcbA a TcbA<sub>iii</sub> v imunoblotovej analýze, keď sa testovala s anti-TcbA<sub>iii</sub> monoklonálnym antidiobdy (tabuľka 31).

Tabuľka 31

Výsledky imunoblotu a biotestov na SCR

Frakcia	Účinok na SCR		Imunoblot	Natívna veľkosť
	% mortality	% inhibície rastu		
Médiá TcbA 1,0 M	+++	+++	TcbA	
Ionomenič				
Médiá TcbA CL-4B	+++	+++	TcbA, TcbA <sub>iii</sub>	okolo 900 kDa
Médiá TcbA CL-4B + proteínáza K	++	+++	NT	
Médiá TcbA CL-4B + tepelná úprava	-	-	NT	
Bunkový sup.CL-4B TcbA	-	+++	NT	okolo 900 900 kD

PK = úprava s proteínázou K počas 2 hodín; tepelná úprava = 100 °C počas 10 minút; ND = nedetegované; NT = netestovaná. Spôsob hodnotenia mortality a inhibície rastu sa hodnotil ako porovnanie s kontrolnými vzorkami; 5 až 24 % = "+", 25 až 49 % = "++", 50 až 100 % = "+++".

Bunkový sediment sa znova suspendoval v 10 mM pufrí fosforečnanu sodného s pH = 7,0 a lýzoval sa pasážovaním cez bunkový rozprašovač Bio-Neb<sup>TM</sup> (Glas-Col Inc., Terra Haute, IN). Sedimenty sa upravili s DNázou na odstránenie DNA

a centrifugovali sa pri 17 000 x g na to, aby sa oddelili bunkové pelety od bunkového supernatantu. Supernatantová frakcia sa dekantovala a filtrovala cez filter s 0,2 mikrónmi preto, aby sa odstránili veľké častice a podrobila sa anionovýmennej chromatografii. Naviazané proteíny sa eluovali s 1,0 M NaCl, dialyzovali a koncentrovali sa na koncentrátore Biomax<sup>TM</sup> (Millipore Corp., Bedford, MA) s oddeľovanou molekulovou hmotnosťou 50 000 Daltonov. Koncentrované frakcie sa podrobili gélovej filtračnej chromatografii na granulovanej matrici Sepharose CL-4B. Údaje biotestov pre materiál pripravovaný týmto spôsobom sú v tabuľke 30 a sú označené ako " bunkový sup. TcbA ".

V ďalších použitých metódach sa spracovalo veľké množstvo materiálu. Bunkové sedimenty sa resuspendovali v 10 mM pufrí fosforečnanu sodného s pH = 7,0 a úplne sa zhomogenizovali na 40 ml tkanivových mlynčekoch Kontes Glass Company (Vineland, NJ). Bunkový odpad sa odsedimentoval centrifugáciou pri 25 000 x g a bunkový supernatant sa dekantoval, prefiltraval cez 0,2 mikrónový filter a podrobil sa anionovýmennej chromatografii na kolóne Pharmacia 10/10, ktorá bola naplnená granulami HQ 50. Naviazané proteíny sa eluovali gradientom NaCl od 0,0 do 1,0 M. Frakcie obsahujúce proteín TcbA sa spájali a koncentrovali s 50 kDa koncentrátorom a podrobili sa gélovej filtračnej chromatografii na granulovej matrici Pharmacia CL-4B. Frakcie, ktoré obsahovali oligomér TcbA s molekulovou hmotnosťou približne 900 kDa, sa zbierali a podrobili sa anionovýmennej chromatografii na kolóne Pharmacia Mono Q 10/10 kondiciovanej s pufrom 20 mM Tris s pH = 7,3. Gradient od 0,0 do 1,0 M NaCl sa použil na elúciu rekombinantného proteínu TcbA. Rekombinantný TcbA eluoval z kolóny pri koncentrácii soli približne 0,3 až 0,4 M NaCl s rovnakou molaritou, pri ktorej sa eluoval natívny oligomér TcbA z kolóny Mono Q 10/10. Zistilo sa, že frakcia rekombinantu TcbA zapríčiňovala podobnú mortalitu SCR v experimentoch s biotestami, aká je v tabuľke 31.

Druhá sada konštrukcií expresií sa pripravila a testovala na expresiu *Tcaba* proteínového toxínu.

Konštrukcia pDAB2030: expresný plazmid pre *tcbA* kódujúcu oblasť

Plazmid pDAB2028 (tu pozri) obsahuje *tcbA* kódujúcu oblasť komerčného vektora pET15 (Novagen, Madison, WI), kóduje ampicilínový selekčný marker. Plazmid pDAB2030 sa vytvoril na expresiu *tcbA* kódujúcej oblasti z plazmidu, ktorý kóduje kanamycínový selekčný marker. Toto sa uskutočnilo delením pET27 (Novagen, Madison, WI) kanamycínového plazmidu a pDAB2028 s *Xba I* a *Xho I*. Toto uvoľnilo celé viacnásobné klonovacie miesto vrátane *tcbA* kódujúcej oblasti od plazmidu pDAB2028. Dva oddelené plazmidy sa zmiešali a ligátovali sa. Rekombinantné plazmidy sa rozdelili na kanamycíne a tie, ktoré obsahovali pDAB2028 fragmenty sa identifikovali reštrikčnou analýzou. Nový rekombinantný plazmid sa nazval pDAB2030.

Konštrukcia plazmidu pDAB2031: korekcia mutácií v *tcbA*<sub>1</sub>

Dve mutácie v N-koncovej *tcbA* kódujúcej oblasti ako bolo opísané v príklade 19 (SEKV. ID. ČÍS. 11; A > G na 212 zmena asparagínu 71 na serín 71; G > A na 229 zmena alanínu 77 na treonín 77) sa korigovali nasledovne: PCR produkt sa generoval s použitím primérov TH50 (5' ACC GTC TTC TTT ACG ATC AGT G 3') a S1Ac51 (5' TTT AAA CCA TGG GAA ACT CAT TAT CAA GCA CTA TC 3') a pDAB2025 ako templátu na generáciu 1778 bp produktu. Tento PCR produkt sa klonoval na plazmid pCR2.1 (Invitrogen, San Diego, CA) a klon sa izoloval a sekvencoval. Klon sa štiepil s *Nco I* a *Pin AI* a fragment 1670 bp sa purifikoval z 1 % agarózového gélu. Plazmid s obsahom mutovanej *tcbA* kódujúcej oblasti (pDAB 2030) sa štiepil s *Nco I* a *Not I* a purifikoval sa od fragmentu 1670 bp v 0,8 % agaróze s Qiaex II (Qiagen, Chatsworth, CA). Korigovaný *Nco I*/*Pin AI* fragment sa potom rozpustil v pDAB2030. Ligátovaná DNA sa potom transformovala do DH5α *E. coli*. Klon sa izoloval, sekvencoval a zistilo sa, že je potrebná korekcia.

Tento plazmid s obsahom korigovanej *tcbA* kódujúcej oblasti sa nazýva pDAB2031.

Konštrukcia pDAB2033 a pDAB2034: expresia plazmidov pre *tcbA*

Expresia plazmidov pDAB2025 a pDAB2027 až 2031 sa celá uskutočnila na expresnom systéme Bactriophage T7. Ďalší vektorový systém sa použil na expresiu *tcbA* génu a jeho derivátov. Expresný vektor Trc99a (Pharmacia Biotech, Piscataway, NJ) obsahuje silný trc promótor proti smeru viacnásobného klonovacieho miesta s 5' *Nco I* miestom, ktoré je porovnateľné s *tcbA* kódujúcou oblasťou od pDAB2030 a pDAB2031. Avšak plazmid nebol kompatibilný s 3' miestom. Preto sa miesto *Hind III* od Trc99a oddelilo a otupilo sa spracovaním s T4 DNA polymerázou (Boehringer Mannheim, Indianapolis, In). Vektorový plazmid sa potom rozdelil *Nco I* a potom nasledovalo spracovanie s alkalickou fosfatázou. Plazmidy pDAB2030 a pDAB2031 sa každý delil s *Xho I* (rezy 3'-konca *tcbA* kódujúcej oblasti) a potom nasledovalo spracovanie s T4 DNA polymerázou na zaoblenie koncov. Plazmidy sa potom delili s *Nco I*, DNA sa potom extrahovali s fenolom etanolom sa vyzrážali a znovu rozpustili v pufri. Plazmidy pDAB2030 a pDAB2031 a Trc99a sa zmiešali oddelene, ligátovali sa a transformovali do buniek DH5 $\alpha$  a uložili sa na LB médium s obsahom ampicilínu a 50 mM glukózy. Rekombinantné plazmidy sa identifikovali reštrikčným štiepením. Nové plazmidy sa nazvali pDAB2033 (obsahuje *tcbA* kódujúcu sekvenciu s dvomi mutáciami v *tcbA*<sub>i</sub>) a pDAB2034 (obsahuje korigovanú verziu *tcbA* z pDAB2031).

Konštrukcia plazmidu pDAB2032: expresný plazmid pre *tcbA*<sub>ii</sub>*A*<sub>iii</sub>

Plazmid, ktorý kóduje *TcbA*<sub>ii</sub>*A*<sub>iii</sub> časť *TcbA* sa vytvoril podobným spôsobom ako plazmid pDAB2031. PCR produkt sa generoval s použitím TH42 (5' TAG GTC TCC ATG GCT TTT ATA CAA GGT TAT AGT GAT CTG 3') a TH50 (5' ACC GTC TTC TTT ACG ATC AGT G 3') primérov a plazmidu pDAB2025 ako templátu. To viedlo k produktu 1521 bp s počiatočným kodónom na začiatku kódujúcej sekvencie *tcbA*<sub>ii</sub>.



Tento PCR produkt sa izoloval v 1 % agarózovom géli a purifikoval sa. Purifikovaný produkt sa klonoval na pCR2.1 ako vyššie a korigovaný klon sa identifikoval DNA sekvenčnou analýzou. Tento klon sa štiepil s *Nco I* a *Pin AI*, fragment 1414 bp sa izoloval v 1 % agarózovom géli a ligátoval sa na *Nco I* a *Pin AI* miesta plazmidu pDAB2030 a transformoval sa do DH5 $\alpha$  *E. coli*. Tento nový plazmid skonštruovaný na expresiu  $TcbA_{ii}A_{iii}$  *E. coli*, sa nazval pDAB2032.

Expresia *tcbA* a  $tcbA_{ii}A_{iii}$  z plazmidov pDAB2030, pDAB2031 a pDAB2032

Tu opísaná expresia *tcbA* v *E. coli* z plazmidov pDAB2030, pDAB2031 a pDAB2032, okrem expresie  $tcbA_{ii}A_{iii}$ , sa uskutočnila v *E. coli* kmeňa HMS174 (DE3) (Novagen, Madison, WI).

Expresia *tcbA* z plazmidu pDAB2033

Plazmid pDAB2033 sa transformoval do buniek BL21 (Novagen, Madison, WI) a uložil sa na LB s obsahom 100 mikroorganizmov/ml ampicilínu a 50 mM glukózy. Platne so rozložili tak, aby na každej platni bolo niekoľko sto kolónií, potom nasledovala inkubácia buď pri 30 °C alebo pri 37 °C cez noc. Kolónie sa zozbierali z platní a suspendovali sa v LB s obsahom 100 mikroorganizmov/ml ampicilínu, ale bez glukózy. Typický obsah kultúry bol 250 ml v jednej 1 l fľaši s deleným dnom. Kultúry sa indukovali, keď dosiahli hustotu 0,3 až 0,6 OD600 nm. Najčastejšie sa táto hustota dosiahla bezprostredne po suspendovaní buniek z platní a nevyžadovala rastovú periódu v kvapalnom médiu. Použili sa dve indukčné metódy. Metóda 1: bunky sa indukovali s 1 mM IPTG pri 37 °C. Kultúry sa trepali pri 200 rpm na plochej trepačke počas 5 hodín a zozbierali sa. Metóda 2: Kultúry sa indukovali s 25 mikromolárnym IPTG pri 30 °C a trepali sa pri 200 rpm počas 15 hodín buď pri 20 °C alebo pri 30 °C. Kultúry sa nechali stáť pri 4 °C pokiaľ sa nepoužili na purifikáciu.

Purifikácia TcbA od *E. coli*

Purifikácia, bioskúška a imunoblot analýza TcbA a TcbA<sub>ii</sub>A<sub>iii</sub> sa uskutočnili tak, ako tu bolo opísané. Výsledky niektorých reprezentačných pokusov expresie *E. coli* sú ukázané v tabuľke 32. Všetky materiály, ktoré sú ukázané v tabuľke 32, sa purifikovali od médií frakcie kultúr. Predpovedaná natívna molekulová hmotnosť je približne 900 kD ako je tu opísané. Čistota vzoriek a množstvo TcbA vo vzťahu k znečisťujúcim proteínom sa menia s každou prípravou.

Tabuľka 32

Bioskúška aktivity a imunoblotová analýza TcbA a derivátov produkovaných v *E. coli* a purifikovaných z média kultúry

Plazmid	Kódujúca oblasť	E. coli reťazec	Southern Corn Rootworm Bioassay activity		Peptidy detegované imunoblotom	Mikrogramy proteínu aplikované v diéte
			% narastania inhibície	% usmrtených		
pDAB2030	<i>tcbA</i>	BL21 (DE3)	-	+++	TcbA + TcbA <sub>iii</sub>	1-8
pDAB2031	<i>tcbA</i>	BL21 (DE3)	-	+++	TcbA + TcbA <sub>iii</sub>	1-10
pDAB2033	<i>tcbA</i>	BL21	-	+++	TcbA + TcbA <sub>iii</sub>	1-2
pDAB2032	<i>tcbA</i> <sub>ii</sub> - <i>A</i> <sub>iii</sub>	HMS 174 (DE3)	+++	+	TcbA <sub>ii</sub> A <sub>iii</sub> + TcbA <sub>iii</sub>	13-27

Hodnotiaci systém pre mortalitu a inhibíciu rastu na liskavke uhorkovej v porovnaní s kontrolnou vzorkou; 5 - 24 % "+", 25 - 49 % "++", 50 - 100 % "+++".

## Príklad 20

Charakterizácia toxínových peptidov s matricovo asistovanou laserovou desorpčnou ionizačnou time-of-flight hmotnostnou spektroskopiou

Toxíny, ktoré sa izolovali z bujónu W-14 sa purifikovali spôsobom ako je opísané v príklade 15. V niektorých prípadoch sa TcaB proteínový toxín znovu spracoval s proteázou (príklad 16), ktorá sa izolovala z bujónu W-14 ako je predtým opísané (príklad 15). Molekulová hmotnosť proteínu sa určila s použitím matricovo asistovanej laserovej desorpčnej ionizačnej time-of-flight hmotnostnej spektroskopie, tu nazývanej MALDI-TOF, na pracovnej stanici VOYAGER BIOSPECTROMETRY s technológiou DELAYED EXTRACTION (PerSeptive Biosystems, Framingham, MA). Typicky sa proteín, o ktorý bol záujem (100 až 500 pmol v 5  $\mu$ l), zmiešal s 1  $\mu$ l acetonitrilu a dialyzoval sa počas 0,5 až 1 hod. na filtri Millipore VS s veľkosťou pórov 0,025  $\mu$ M (Millipore Corp. Bedford, MA). Dialýza sa uskutočnila plávaním filtra na vode (shiny side up), potom nasledovalo dodanie zmesi proteín-acetonitril po kvapkách na povrch filtra. Po dialýze sa dialyzovaný proteín odstránil s použitím pipety a zmiešal sa s matricou, ktorá obsahovala kyselinu sinapínovú a kyselinu trifluóroctovú podľa výrobných inštrukcií. Proteín a matrica mali schopnosť spoločne kryštalizovať na okolo zlatom potiahnutej platni s veľkosťou 3 cm<sup>2</sup> (PerSeptive Corp.). Excitácia kryštálov a následná hmotnostná analýza sa uskutočnili s použitím nasledujúcich podmienok: nastavenie lasera 3050, tlak 4,55e-07, vstupná nízka hmotnosť 1500,0, negatívne ióny - vypnuté, akceleračné napätie 25 000, mriežkové napätie 90,0 %, napätie vodiča 0,010 %, lineárny mód a čas pulzového oneskorenia 350 ns.

Údaje z hmotnostnej analýzy proteínu sú ukázané v tabuľke 33. Údaje získané z MALDI-TOF sa porovnali s hypotetickými výsledkami z informácie o génovej sekvencii a predtým stanovenými s SDS-PAGE.

Tabuľka 33

Molekulárna analýza peptidov pomocou MALDI-TOF, SDS-PAGE a predpokladané stanovenie založené na génovej sekvencii

Peptid	Predpoklad (gén)	SDS-PAGE	MALDI-TOF
TcbA	280,634 Da	240,000 Da	281,040 Da
TcbA <sub>i/ii</sub>	217,710 Da	nerozložený	216,812 Da
TcbA <sub>ii</sub>	207,698 Da	201,000 Da	206,473 Da
TcbA <sub>iii</sub>	62,943 Da	58,000 Da	63,520 Da
-----			
TcdA <sub>ii</sub>	209,218 Da	188,000 Da	208,186 Da
TcdA <sub>iii</sub>	63,520 Da	56,000 Da	63,544 Da
-----			
TcbA <sub>ii</sub>	generovaná proteáza	201,000 Da	216,614 Da <sup>^</sup> 215,123 Da <sup>^</sup> 210,391 Da <sup>^</sup> 208,680 Da <sup>^</sup>
TcbA <sub>iii</sub>	generovaná proteáza	56,000 Da	64,111 Da

<sup>^</sup> Údaje normalizovanej TcbA, mnohonásobné fragmenty pozorované v TcbA<sub>ii/ii</sub>

#### Príklad 21

#### Produkcia peptidových špecifických polyklonálnych protilátok

Deväť zložiek peptidu toxínového komplexu W-14, menovite TcaA, TcaA<sub>iii</sub>, TcaB<sub>ii</sub>, TcaC, TcbA<sub>ii</sub>, TcbA<sub>iii</sub>, TcdA<sub>ii</sub> a TcdA<sub>iii</sub> sa vybrali ako ciele, voči ktorým sa budú tvoriť protilátky. Celková DNA a údaje od odvodených aminokyselinových sekvencií pre tieto peptidy ukazujú, že sekvenčná homológia medzi niektorými peptidmi bola podstatná. Ak sa použil celý peptid ako imunogén na indukciu tvorby protilátky, potom výsledné protilátky mohli viazať

mnohonásobné peptidy pri príprave toxínov. Na predídenie tohto problému sa protilátky generovali tak, aby sa mohli viazať špecificky k jednej oblasti každého peptidu, ktorý je predmetom záujmu. Jediná oblasť (subpeptid) každého cieľového peptidu sa vybrala na základe analýzy opísanej nižšie.

Každá celá peptidová sekvencia sa analyzovala s použitím MacVector<sup>TM</sup> Protein Analysis Tool (IBI Sequence Analysis Software, International Biotechnologies, Inc., P. O. Box 9558, New Haven, CT 06535) na stanovenie jej antigénového indexu. Tento program sa vybral na lokalizáciu novej vonku umiestnenej aminokyselinovej sekvencie, t.j. oblastí, ktoré by mohli mať antigénové miesta. Tento spôsob kombinoval informácie z hydrofobicity, pravdepodobnosti povrchu a predpovedí bočnej ohybnosti pozdĺž sekundárnej štruktúrnej predpovedi na produkciu kompozitnej predpovedi povrchových črt proteínu. Ohodnotenie každej analýzy sa normalizovalo na hodnotu medzi -1,0 a + 1,0 (MacVector<sup>TM</sup> Manual). Hodnoty antigénového indexu sa získali pre celú sekvenciu cieľového peptidu. U každého peptidu sa plocha pokrývajúca 19 alebo viac aminokyselín, ktorá ukázala vysoký antigénový index od pôvodnej sekvencie, znovu analyzovala na stanovenie antigénového indexu subpeptidu bez bočných zvyškov. Táto opakovaná analýza bola potrebná, pretože antigénový index peptidu by mohol byť ovplyvnený bočnými aminokyselinovými zvyškami. Ak izolovaná subpeptidová sekvencia neudrží vysoký antigénový index, potom sa musí vybrať nová oblasť a analýza sa musí opakovať.

Každá oddelená subpeptidová sekvencia sa vyrovnala a porovnala sa so všetkými siedmymi cieľovými peptidovými sekvenciami s použitím MacVector<sup>TM</sup> linearizačného programu. Ak oddelená subpeptidová sekvencia ukázala identitu (väčšiu ako 20 %) k inému cieľovému peptidu, izolovalo sa a znovu sa analyzovalo sa 19 alebo viac nových aminokyselinových oblastí. Od všetkých cieľových peptidov sa oddelili jediné subpeptidové sekvencie pokrývajúce 19 alebo viac aminokyselín, ktoré ukázali

vysoký antigénový index.

Sekvencie siedmych subpeptidov sa poslali do Genemed Biotechnology Inc. Posledný aminokyselinový zvyšok každého subpeptidu sa vymazal, pretože neukázal významný efekt na antigénový index. Cysteínový zvyšok sa dodal k N-koncovému zvyšku každej subpeptidovej sekvencie, okrem TcaB<sub>i</sub>-syn, ktorý obsahuje vnútorný cysteínový zvyšok. Prítomnosť cysteínových zvyškov uľahčuje konjugáciu nosného proteínu (KLH). Výsledné peptidové produkty korešpondujúce s určenými toxínovými peptidmi a sekvenciami SEKV. ID. Čís. sú ukázané v tabuľke 34.

Tabuľka 34

Aminokyselinové sekvencie pre syntetické peptidy

SEQ ID. Čís. Peptidová aminokyseliná sekvencia

---

63	TcaA <sub>ii</sub> -syn	NH2-	(C) L R G N S P T N P D K D G I F A Q V A
64	TcaA <sub>iii</sub> -syn	NH2-	(C) Y T P D Q T P S F Y E T A F R S A D G
65	TcaB <sub>i</sub> -syn	NH2-	H G Q S Y N D N N Y C N F T L S I N T
66	TcaB <sub>iii</sub> -syn	NH2-	(C) V D P K T L Q R Q Q A G G D G T G S S
67	TcaC-syn	NH2-	(C) Y K A P Q R Q E D G D S N A V T Y D K
68	TcbA <sub>ii</sub> -syn	NH2-	(C) Y N E N P S S E D K K W Y F S S K D D
69	TcbA <sub>iii</sub> -syn	NH2-	(C) F D S Y S Q L Y E E N I N A G E Q R A
70	TcdA <sub>ii</sub> -syn	NH2-	(C) N P N N S S N K L M F Y P V Y Q Y S G N T
71	TcdA <sub>iii</sub> -syn	NH2-	(C) V S Q G S G S A G S G N N N L A F G A G

Každý konjugovaný syntetický peptid sa injektoval dvom králikom podľa Genemedovho akceleračného programu. Predimúnne a postimúnne séra boli k dispozícii na testovanie po jednom mesiaci.

Predbežný test obidvoch pred- a postimúnnych sér od každého králika sa uskutočnil v Genemed Biotechnologies Inc. Genemed oznámil, že s použitím obidvoch techník ELISA a Western blot sa

zistila reakcia postimúnneho séra k príslušným syntetickým peptidom. Následne sa testovali séra s celým cieľovým peptidom pomocou Western blot analýzy. Dve dávky čiastočne purifikovaného W-14 toxínového komplexu rodu *Photorhabdus* sa použil ako antigén. Dve vzorky ukázali aktivitu proti liskavke uhorkovej. Ich peptidové vzory na géli SDS-PAGE sa mierne odlišovali.

Použili sa predliate SDS-polyakrylamidové gély s gradientom 4 až 20 % (Integrated Separation Systems, Natick, MA 01760). Na každú gélovú sondu sa nanieslo 1 až 8 µg proteínu. Uskutočnila sa elektroforéza a proteíny sa rozložili na nitrocelulózovej membráne Hybond-ECL<sup>TM</sup> (Amersham International). Membrána sa blokovala s 10 % mliekom v TBST (25 mM Tris HCl, pH 7,4, 136 mM NaCl, 2,7 mM KCl, 0,1 % Tween 20) počas jednej hodiny pri izbovej teplote. Každé králičie sérum sa rozpustilo v 10 % mlieku/TBST na pomer 1 : 500. Použili sa aj ostatné riedenia medzi 1 : 50 až 1 : 1000. Sérum sa dodalo do membrány a umiestnilo sa na rovinnú trepačku počas najmenej jednej hodiny. Membrána sa dôkladne vyprala s blokovacím roztokom alebo TBST. Zriedenie 1 : 2000 sekundárnych protilátok (kozí anti-myší IgG konjugovaný s chrenovou peroxidázou, BioRad Laboratories) v 10 % mlieku/TBST sa aplikovalo na membránu umiestnenú rovinnou trepačke počas jednej hodiny. Membrána sa následne vyprala so zvýšeným množstvom TBST. Detekcia proteínu sa uskutočnila s použitím ECL (Enhanced Chemiluminescence) detekčného zariadenia (Amersham International).

Analýza Western blot sa uskutočnila kvôli identifikácii viazacej špecificity každej anti-syntetickej peptidovej protilátky. Všetky syntetické polyklonálne protilátky ukázali špecificitu smerom k spracovaným, a ak sú použiteľné, smerom k nespracovaným, cieľovým peptidom z proteínových frakcií, ktoré sú odvodené od bujónu kultúry *Photorhabdus*. Zistili sa rôzne protilátky na uznanie či nespracované alebo spracované rekombinantné proteíny odvodené od heterologických expresných systémov, ako sú bakteriálne alebo hmyzie bunky, s použitím

baculovírusových expresných systémov. V prvom prípade protilátka anti-TcbA<sub>iii</sub>-syn ukázala nejakú skríženú reaktivitu k peptidu anti-TcbA<sub>iii</sub>. V druhom prípade protilátka anti-TcaC-syn uznala neidentifikovaný 190 kDa peptid vo frakciách W-14 toxínového komplexu.

## Príklad 22

### Charakterizácia kmeňov *Photorhabdus*

Na stanovenie, že tu opísaná zbierka bola obsiahnutá v kmeňoch *Photorhabdus*, sa kmene tu opísané stanovili v zmysle uznaných mikrobiologických črt, ktoré sú charakteristické pre bakteriálny rod *Photorhabdus* a ktorými sa odlišuje od ostatných druhov *Enterobacteriaceae* a *Xenorhabdus* (Farmer, J.J. 1984, Bergey's Manual of Systemic Bacteriology, Vol 1., str. 510 - 511. (ed. Kreig N.R. a Holt, J.G.) Williams & Wilkins, Baltimore; Akhurst and Boemare, 1988, J. Gen. Microbiol. 134, 1835 - 1845; Forst and Nealon, 1996. Microbiol. Rev. 60, 21 - 43). Tieto charakteristické črty sú nasledujúce: Gramovo negatívne farbenie tyčiniek, veľkosť organizmu 0,3 až 2 µm v šírke a 2 až 10 µm v dĺžke [s príležitostnými vláknami (15 až 50 µm) a sféroplastom], pigmentácia kolónie na živnom agare od žltej do oranžovej/červenej, prítomnosť kryštalických inklúzných telies, prítomnosť katalázy, neschopnosť redukovať dusičnany, prítomnosť bioluminiscencie, schopnosť prijímať farbivo z rastového média, pozitívne na tvorbu proteázy, rast pri teplote menšej ako 37 °C, úhyn pri anaeróbných podmienkach a pozitívna pohyblivosť (tabuľka 33). Testovacie metódy sa kontrolovali s použitím referenčných rodov *Escherichia coli*, *Xenorhabdus* a *Photorhabdus*. Celkové výsledky sú konzistentné so všetkými kmeňmi, ktoré sú súčasťou čelade *Enterobacteriaceae* a druhu *Photorhabdus*. Treba poznamenať, že DEP1, DEP2 a DEP3 majú vzťah ku kmeňom *Photorhabdus* získaných z americkej zbierky kultúr - American Type Culture Collection, 12301 Parklawn Drive, Rockville, MD 20852 USA (ATCC č. 29304, 29999 a 51583) a boli



znovu uložené v zbierke kultúr - Agricultural Research Culture Collection (NRRL), International Depository Authority, 1815 N. University Street, Peoria, IL 61604, USA 29. apríla 1997 a druhé uloženie je identifikované ako NRRL č. B-21707, 21708 a 21709.

Na stanovenie bioluminiscencie týchto kmeňov *Photorhabdus* sa použil luminometer. Pri meraní prítomnosti alebo neprítomnosti relatívnych svetelných emitujúcich jednotiek sa merali bujóny z každého kmeňa (bunky a médium) trikrát po očkovaní v kvapalnej kultúre (24, 48, 72 hodín) a porovnali sa so svietivosťou pozadia (neočkované médium). Niekoľko kmeňov *Xenorhabdus* sa testovalo ako negatívne kontroly na svetelnosť. Pred meraním svetelnej emisie z rôznych bujónov sa stanovila bunková hustota meraním svetelnej absorbancie (560 nm) v spektrometri Gilford Systems (Oberlin, OH) s použitím slamkovej bunky. Výsledné svetlom emitované jednotky sa potom normalizovali na hustotu buniek. Alikvótné časti bujónov s umiestnili do 96 sônd na mikrotitračných platniach (každá po 100  $\mu$ l) a merali sa na luminometri Packard Lumicount<sup>TM</sup> (Packard Instrument Co., Meriden, CT). Meracia perióda pre každú vzorku bola od 0,1 do 1,0 s. Vzorky sa budili v luminometri počas 10 s pred ich odčítaním. Pozitívny test sa stanovil, že je 5 násobkom luminiscencie pozadia (okolo 1 až 15 relatívnych svetelných jednotiek). Navyše stupeň svetelnosti kolónie sa potvrdil prekryvaním fotografických filmov a pohľadom po vizuálnej adaptácii v tmavej komore. Gramove kmeňové charakteristiky každého kmeňa sa stanovili s komerčným Gramovým kmeňovým prístrojom (BBL, Cockeysville, MD) použitým v zapojení s Gramovými kmeňovými kontrolnými diapozitívmi (Fisher Scientific, Pittsburgh, PA). Mikroskopické vyhodnotenie sa uskutočnilo s použitím mikroskopu Zeiss (Carl Zeiss, Germany) so sošovkou objektívu 100X s olejovou imerziou (s okulárom 10X a tubusovým zväčšením 2X). Mikroskopické pozorovanie jednotlivých kmeňov kvôli veľkosti organizmu, opisu buniek a inklúzných telies (tie posledné dve pozorovania po logaritmickom raste) sa uskutočnili s použitím vlhkých podporných sklíčok (okulár 10X, tubus 2X a zväčšenie objektívu 40X) a s fázovou kontrastnou

mikroskopiou s mikrometrom (Akhurst, R.J. a Boemare, N.E. 1990. Entomopathogenetic Nematodes in Biological Control (ed. Gaugler, R. and Kaya, H.), str. 75 - 90. CRC Press, Boca Raton, USA.; Baghdiguian S., Boyer-Giglio M. H., Thaler, J. O., Bonnot G., Boemare N. 1993. Biol. Cell 79, 177 - 185.). Pigmentácia kolónie sa uskutočnila po očkovaní na živný agar Bacto (Difco Laboratories, Detroit, MI), pripravený podľa inštrukcií. Inkubácia prebehla pri 28 °C a opisy sa vypracovali po piatich dňoch. Na testovanie prítomnosti enzýmovej katalázy sa premiestnili kolónie testovaného organizmu na malý uzáver z platne živného agaru a umiestnili sa na dno sklenenej testovacej tuby. Jeden ml roztoku úžitkového peroxidu vodíka sa pomaly dodal dovnútra tuby. Pozitívna reakcia sa zistila, keď sa bubliny plynu (uvoľnený kyslík) zjavili bezprostredne alebo do 5 sekúnd. Tiež sa uskutočnili kontroly nenačkovaného živného agaru a roztoku peroxidu vodíka. Na testovanie dusičnanovej redukcie sa každá kultúra naočkovala do 10 ml dusičnanového bujónu Bacto (Difco Laboratories, Detroit, MI). Po 24 hodinách inkubácie s miernym miešaním pri 28 °C sa tvorba dusitanu testovala dodaním dvoch kvapiek reagentu kyseliny sulfanilovej a dvoch kvapiek alfa-naftylamínového reagentu (pozri Difco Manual, 10. vydanie, Difco Laboratories, Detroit, MI, 1984). Tvorba zreteľnej ružovej alebo červenej farby naznačuje tvorbu dusitanu z dusičnanu, zatiaľ čo neprítomnosť tvorby farby naznačuje, že kmeň je negatívny v redukcii dusičnanov. Neskôr sa dodal jemne zomletý zinok na ďalšie potvrdenie prítomnosti neredukovaného dusičnanu; stanovený tvorbou dusičnanu a výslednou červenou farbou. Schopnosť každého kmeňa prijať farbivo z rastového média sa testovala na agare Bacto MacConkey s obsahom farbiva neutrálneho červeného; agar Bacto Tergitol-7 obsahoval farbivo bromotymolová červená a agar Bacto EMB obsahoval farbivo eozín-Y (rastové agary z Difco Laboratories, Detroit, MI, všetky pripravené podľa inštrukcií). Po naočkovaní na tieto média sa prijímanie farieb zaznamenalo po inkubácii pri 28 °C počas 5 dní. Rast na týchto posledných médiách je charakteristický pre členov čelade *Enterobacteriaceae*. Pohyblivosť každého kmeňa sa testovala

s použitím roztoku Bacto Motility Test Medium (Difco Laboratories, Detroit, MI) pripraveného podľa inštrukcií. Uskutočnilo sa butt-stab očkovanie s každým kmeňom a pohyblivosť sa sledovala mikroskopicky pomocou difúznej zóny rastového rozširovania od línie očkovania. Tvorba proteázy sa testovala pozorovaním hydrolýzy rôsolu s použitím rôsolu Bacto (Difco Laboratories, Detroit, MI) pripraveného podľa inštrukcií. Kultúry sa naočkovali a tuby alebo platne sa inkubovali pri 28 °C počas 5 dní. Hydrolýza rôsolu sa potom kontrolovala pri izbovej teplote t.j. menšej ako 22 °C. Na stanovenie rastu pri rôznych teplotách sa agarové platne [2 % proteózového peptónu #3 s 2 % agaru Bacto-Agar (Difco, Detroit, MI) v deionizovanej vode] sa rozstrihali z bežného zdroja očkovania. Platne sa inkubovali pri 20, 28 a 37 °C počas troch týždňov. Hladiny teplôt inkubátora sa kontrolovali s elektronickým termočlánkom a meradlom na zabezpečenie platných teplotných zadání. Požiadavky na kyslík pre kmene *Photorhabdus* sa testovali nasledujúcim spôsobom. Uskutočnilo sa očkovanie butt-stab do fluidného tioglyklátového bujónového média (Difco, Detroit, MI). Tuby sa inkubovali pri izbovej teplote počas jedného týždňa a kultúry sa potom skúšali na typ a rozsah rastu. Indikátor resazurín demonštruje prítomnosť oxidácie média alebo aeróbnu zónu (Difco Manual, 10. vydanie, Difco Laboratories, Detroit, MI). Rastová zóna získaná pre kmene *Photorhabdus* bola konzistentná s tou, ktorá sa získala od fakultatívnych anaeróbných mikroorganizmov. V prípade nejasných výsledkov sa finálne agarové koncentrácie fluidného tioglykolátového bujónového média zvýšili na 0,75 % a rastové charakteristiky sa znovu skontrolovali.

Tabuľka 35

Taxonomické črty kmeňov *Photorhabdus*

Retazec	A*	B	C	D	E	F	G	H	I	J <sup>§</sup>	K	L	M	N	O	P	Q
P. zealandica	-#	+	+	rd S	+	-	+	+	+	PO	+	+	+	+	+	+	-
P. hepialus	-	+	+	rd S	+	-	+	+	+	Y	+	+	+	+	+	+	-
HB-Arg	-	+	+	rd S	+	-	+	+	+	W	+	+	+	+	+	+	-
HB Oswedo	-	+	+	rd S	+	-	+	+	+	W	+	+	+	+	+	+	-
HB Lewiston	-	+	+	rd S	+	-	+	+	+	T	+	+	+	+	+	+	-
K-122	-	+	+	rd S	+	-	+	+	+	Y	+	+	+	+	+	+	-
HMGD	-	+	+	rd S	+	-	+	+	+	Rd	+	+	+	+	+	+	-
Indicus	-	+	+	rd S	+	-	+	+	+	W	+	+	+	+	+	+	-
GD	-	+	+	rd S	+	-	+	+	+	YT	+	+	+	+	+	+	-
PWH-5	-	+	+	rd S	+	-	+	+	+	Y	+	+	+	+	+	+	-
Megidis	-	+	+	rd S	+	-	+	+	+	R	+	+	+	+	+	+	-
HF-85	-	+	+	rd S	+	-	+	+	+	R	+	+	+	+	+	+	-
A. Cows	-	+	+	rd S	+	-	+	+	+	PR	+	+	+	+	+	+	-
MP1	-	+	+	rd S	+	-	+	+	+	T	+	+	+	+	+	+	-
MP2	-	+	+	rd S	+	-	+	+	+	T	+	+	+	+	+	+	-
MP3	-	+	+	rd S	+	-	+	+	+	R	+	+	+	+	+	+	-
MP4	-	+	+	rd S	+	-	+	+	+	Y	+	+	+	+	+	+	-
MP5	-	+	+	rd S	+	-	+	+	+	PR	+	+	+	+	+	+	-
GL98	-	+	+	rd S	+	-	+	+	+	W	+	+	+	+	+	+	-
GL101	-	+	+	rd S	+	-	+	+	+	W	+	+	+	+	+	+	-
GL138	-	+	+	rd S	+	-	+	+	+	W	+	+	+	+	+	+	-
GL155	-	+	+	rd S	+	-	+	+	+	W	+	+	+	+	+	+	-

Tabuľka 35 - pokračovanie

Reťazec	A*	B	C	D	E	F	G	H	I	J <sup>§</sup>	K	L	M	N	O	P	Q
GL217	-	+	+	rd S	+	-	+	+	+	Y	+	+	+	+	+	+	-
GL257	-	+	+	rd S	+	-	+	+	+	O	+	+	+	+	+	+	-
DEP1	-	+	+	rd S	+	-	+	+	+	W	+	+	+	+	+	+	-
DEP2	-	+	+	rd S	+	-	+	+	+	PR	+	+	+	+	+	+	-
DEP3	-	+	+	rd S	+	-	+	+	+	CR	+	+	+	+	+	+	-

\*: A = Gramovo zafarbenie, B = kryštalické inklúzne telesá, C = bioluminiscencia, D = forma buniek, E = pohyblivosť, F = redukcia dusičnanov, G = prítomnosť katalázy, H = hydrolýza rôsolu, I = prijímanie farbiva, J = pigmentácia na živnom agare (niektoré farebné posuny po piatom dni), K = rast na EMB agare, l = rast na agare MacConkey, M = rast na agare Tergitol - 7, N = fakultatívny anaeróbny rast, O = rast pri 20 °C, P = rast pri 28 °C, Q = rast pri 37 °C.

#: + = pozitívny pre črtu, - = negatívny pre črtu; rd = tyčinka, S = triedený v rámci rodového opisu.

§ W = biela, CR = smotanová, Y = žltá, YT = žltohnedá, T = hnedá, PO = svetlo oranžová, O = oranžová, PR = svetlo červená, R = červená.

Evolučná diverzita kmeňov *Photorhabdus* v tejto kolekcii sa merala analýzou PCR (polymerázová reťazová reakcia, Polymerase Chain Reaction) médiovaným genómovým odtlačkom prsta s pozitívom genómovej DNA z každého kmeňa. Táto technika je založená na rodinách opakovaných DNA sekvencií, ktoré sú prítomné v genóme rôznych bakteriálnych druhov (Versalovic, J., Schneider, M., DE Bruijn, F. J. a Lupski, J. R. 1994. *Methods Mol. Cell. Biol.*, 5, 25 - 40). Tri z nich, opakovaná mimogénová palindromová sekvencia

(REP), enterobakteriálny opakovaný vnútrogénový konsenzus (ERIC) a BOX prvok sa uvažujú, že hrajú významnú úlohu v organizácii bakteriálneho genómu. Predpokladá sa, že genómová organizácia je usporiadaná prostredníctvom výberu a rozdielna disperzia týchto prvkov vo vnútri genómu blízkych príbuzných kmeňov sa môže použiť na rozlišovanie týchto kmeňov (napr. Louws, F. J., Fulbright, D. W., Stephens, C. T. a DE Bruijn, F. J. 1994 Appl. Environ. Micro. 60, 2286 - 2295). REP-PCR používa oligonukleotidové priméry ako doplnok k týmto opakovaným sekvenciám na amplifikáciu rôzne veľkých DNA fragmentov, ktoré ležia medzi nimi. Výsledné produkty sa oddeľujú elektroforézou na stanovenie DNA "odtlačkov prstov" pre každý reťazec.

Na izoláciu genómovej DNA z našich kmeňov sa bunkové pelety resuspendovali v TE pufri (10 mM Tris-HCl, 1 mM EDTA, pH 8,0) na konečný objem 10 ml a potom sa dodalo 12 ml 5 M NaCl. Táto zmes sa centrifugovala 20 min pri 15 000 x g. Výsledná peleta sa resuspendovala v 5,7 ml TE, potom sa dodalo 300 µl 10 % SDS a 60 µl 20 mg/ml proteínázy K (Gibco BRL Products, Grand Island, NY). Táto zmes sa inkubovala pri 37 °C počas 1 hod, potom sa dodalo približne 10 mg lyzozýmu a zmes sa inkubovala počas ďalších 45 min. Potom sa dodal 1 ml 5 M NaCl a 800 µl CTAB/NaCl roztoku (10 % hmotn./obj. CTAB, 0,7 M NaCl) a zmes sa inkubovala 10 min pri 65 °C, jemne sa pretrepávala, potom sa inkubovala a jemne pretrepávala počas ďalších 20 min. na podporu vyčistenia bunkového materiálu. Dodal sa rovnaký objem roztoku chloroform/izoamylalkohol (24 : 1, obj./obj.), jemne sa miešalo a potom sa centrifugovalo. Potom sa uskutočnili dve extrakcie s rovnakým objemom roztoku fenol/chloroform/izoamylalkohol (50 : 49 : 1). Genómová DNA sa zrážala s 0,6 objemu izopropanolu. Vyzrážaná DNA sa odstránila so sklenenou tyčinkou, dvakrát sa premyla so 70 % etanolom, sušila sa a rozpustila v 2 ml STE (10 mM Tris-HCl, pH 8,0, 10 mM NaCl, 1 mM EDTA). DNA sa potom kvantifikovala pomocou optickej hustoty pri 260 nm. Na uskutočnenie REP-PCR analýzy genómovej *Photorhabdus* DNA sa použili nasledujúce priméry: REP1-I, 5'-IIIICGICGICATCIGGC-3'

a REP2-I, 5'-ICGICTTATCIGGCCTAC-3'. PCR sa uskutočnila s použitím nasledujúcich 25  $\mu$ l reakcie: 7,75  $\mu$ l H<sub>2</sub>O, 2,5  $\mu$ l 10X LA pufru (PanVera Corp., Madison, WI), 16  $\mu$ l dNTP zmesi (2,5 mM z každého), 1  $\mu$ l každého priméru pri 50 pM/ $\mu$ l, 1  $\mu$ l DMSO, 1,5  $\mu$ l genómovej DNA (koncentrácie v rozsahu od 0,075 do 0,480  $\mu$ l/ $\mu$ l) a 0,25  $\mu$ l TaKaRa EX Taq (PanVera Corp., Madison, WI). PCR amplifikácia sa uskutočnila v Perkin Elmer DNA termálnom cyklovači (Norwalk, CT) s použitím nasledovných podmienok: 95 °C/7 min., potom 35 cyklov mimo, 94 °C/1 min., 44 °C/1 min., 65 °C/8 min., potom nasledovalo 15 min. pri 65 °C. Po cyklovaní sa 25  $\mu$ l reakcie dodalo do 5  $\mu$ l 6X gélového nakladacieho pufru (0,25 % brómfenolová modrá, 40 % hmotn./obj. sacharózy vo vode). 15 x 20 cm 1 % agarózového gélu sa potom dalo na TBE pufor s použitím 8  $\mu$ l z každej reakcie. Gél tiekol približne počas 16 hodín pri 45v. Gély sa potom zafarbili v 20  $\mu$ l/ml etídium bromidu počas 1 hodiny a odfarbovali v TBE pufri počas približne 3 hodín. Polaroidové fotografie gélov sa získali pod UV osvetlením.

Prítomnosť alebo neprítomnosť pásov na špecifických veľkostiach pre každý kmeň sa zisťovala z fotografií a zadávala sa ako matica podobnosti v numerickom taxonomickom softvérovom programe NTSYS-pc (Exeter Software, Setauket, NY). Kontroly kmeňov *E. coli* HB101 a *Xanthomonas oryzae* pv. *oryzae* uskutočnené za tých istých podmienok ako produkcia PCR otlačkov prstov korešpondujú s publikovanými správami (Versalovic, J., Koeth, T. a Lupski, J. R. 1991. Nucleic Acids Res. 19, 6823 - 6831; Vera Cruz, C. M., Halda-Alija, L., Louws, F., Skinner D.Z., George, M. L., Nelson, R. J., DE Bruijn, F. J., Rice, C. a Leach, J. E. 1995. Int. Rice Res. Notes, 20, 23-24; Vera Cruz, C. M., Ardales, E. Y, Skinner, D. Z., Talag, J., Nelson, R.J., Louws, F. J., Leung, H., Mew, T. W. a Leach, J. E. 1996. Phytopathology 86, 1352-1359). Údaje z kmeňov *Photorhabdus* sa analyzovali so sériou programov NTSYS-pc; SIMQUAL (Similarity for Qualitative data) na generovanie matice koeficientov podobnosti (s použitím Jaccardovho koeficientu) a SAHN (Sequential, Agglomerative, Hierarchial and Nested) zoskupovaním [s použitím UPGMA

(Unweighted Pair-Group Method with Arithmetic Averages) method], pomocou ktorého sa zistilo, ktoré skupiny patria ku kmeňom a môžu byť vyjadrené ako fenogram (obr. 7). Programy COPH (kofenetické hodnoty) a MXCOMP (porovnanie matíc) sa použili na generovanie kofenetických hodnôt matíc a porovnanie korelácií medzi nimi a pôvodnej matice, na základe ktorej sa uskutočnilo zoskupovanie. Výsledná normalizovaná Mantelova štatistika ( $r$ ) sa generovala a je mierou vhodnosti prispôsobenia pre zoskupovaciú analýzu ( $r = 0,8$  až  $0,9$  predstavuje veľmi dobré prispôsobenie). V našom prípade  $r$  bol  $0,924$ . Preto sa kolekcia porovnávala s rôznymi skupinami ľahko odlišiteľných kmeňov predstavujúcich rod *Photorhabdus*.

#### Príklad 23

Insekticídna použiteľnosť toxínu(ov) produkovaného rôznymi kmeňmi *Photorhabdus*

Počiatkové "zásobné" kultúry rôznych kmeňov *Photorhabdus* sa produkovali naočkováním 175 ml 2 % proteózového peptónu #3 (PP3) (Difco Laboratories, Detroit, MI) kvapalného média s primárnym koloniálnym variantom v 500 ml trojhrdlej banke s Delongovým hrdlom, uzavretým s Kaputovým uzáverom. Po očkovaní sa banka inkubovala počas 24 až 72 hodín pri  $28\text{ }^{\circ}\text{C}$  v rotačnej trepačke pri 150 rpm pokiaľ sa nedosiahla stacionárna fáza. Kultúry sa preniesli do sterilných magnetických miešacích valcov a prekryli sa so sterilným minerálnym olejom kvôli obmedzeniu prístupu vzduchu. Skladované kultúry sa uchovávali v tme pri izbovej teplote. Tieto kultúry sa potom použili ako očkovacie zdroje na fermentáciu každého kmeňa.

"Násadové" banky alebo kultúry sa produkovali buď očkovaním 2 ml olejom prekrytých kultúr alebo prenesením primárneho koloniálneho variantu do 175 ml sterilného média v 500 ml trojhrdlovej banke uzatvorenej s Kaputovým uzáverom. (Je tiež možné použitie iných očkovacích zdrojov). Typicky nasledovalo 16



hodín inkubácie pri 28 °C v rotačnej trepačke pri 150 rpm, potom sa násadové kultúry preniesli do produkčných baniek. Produkčné banky sa zvyčajne očkovali dodaním asi 1 % aktívne rastúcej násadovej kultúry k sterilnému 2 % PP3 médiu (napríklad 2,0 ml na 175 ml sterilného média). Produkcia bujónov prebehla v 500 ml trojhrdlých bankách s Kaputovým uzáverom. Produkčné banky sa iniciovali pri 28 °C na rotačnej trepačke pri 150 rpm. Produkčné fermentácie sa ukončili po 24 až 72 hodinách, hoci úspešná fermentácia nebola obmedzená týmto časom trvania. Po určenej inkubácii sa bujóny umiestnili do sterilných 1,0 l polyetylénových fliaš, otáčali sa pri 260 x g počas 1 hod. pri 10 °C a dekantovali sa od buniek a zvyškových peliet. Ďalšie purifikácia bujónov sa dosiahla s tangenciálnym prietokovým mikrofiltračným zariadením (Pall Filtron, Northborough, MA) s použitím membránového polyétersulfónového filtra s 0,5 µM otvorenými kanálmi. Výsledné bujóny sa potom koncentrovali (nad 10 násobok) s použitím 10 000 alebo 100 000 MW cut-off membrány, ultrafiltračného zariadenia M12 (Amicon, Beverly MA) alebo centrifugálnym koncentrátorom (Millipore, Bedford, MA a Pall Filtron, Northborough, MA) s veľkosťou pórov 10 000 alebo 100 000 MW. V prípade centrifúgových koncentrátorov sa bujón otáčal pri 2 000 x g približne 2 hodiny. Membránový permeát sa dodal do korešpondujúceho retentátu na dosiahnutie želanej koncentrácie zložiek väčšej ako použitá veľkosť pórov. Po týchto procedúrach sa bujón použil na biochemické analýzy alebo na sterilizovaný filter s použitím 0,2 µM membrány z nitrátu celulózy na biologické ohodnotenie. Tepelná inaktivácia spracovaných vzoriek sa dosiahla zahriatím vzoriek na 100 °C v blokoch plnených pieskom počas 10 minút.

Bujón(y) a toxínový(é) komplex(y) z rôznych kmeňov *Photorhabdus* bol užitočný na redukciu populácií hmyzu a použil sa v spôsobe inhibície hmyzej populácie, ktorý zahŕňal aplikovanie na miesto hmyzu efektívneho inaktivujúceho insekticídneho množstva opísanej účinnej látky. Demonštrácia šírky insekticídnej aktivity pozorovanej z bujónov vybraných skupín kmeňov

*Photorhabdus*, fermentovaných ako je opísané vyššie, je ukázaná v tabuľke 36. Je možné, že zdokonalené alebo dodatočné insekticídne aktivity by mohli byť zistené s týmito kmeňmi cez narastanie koncentrácie bujónov alebo s použitím rôznych fermentačných metód. V súlade s aktivitou býva združený proteín, insekticídna aktivita všetkých testovaných kmeňov bola teplotne labilná.

Bujón(y) kultúr z rôznych kmeňov *Photorhabdus* ukazuje rôznu insekticídnu aktivitu (mortalita a/alebo rast inhibície) oproti množstvu hmyzu. Špecifickejšie sa aktivita pozoruje oproti liskavke uhorkovej, ktorá je členom hmyzu rádu *Coleoptera*. Iné členy *Coleoptera* zahŕňajú *Anthrenus grandis*, kováčiky obilné, blyskáčiky repkové, skočky chrenové, osivové chrobáky a pásavku zemiakovú. Bujóny a purifikovaný toxínový komplex(y) sú tiež aktívne voči sivkavcovi, *Manduca sexta* a vijačke kukuričnej, ktoré sú členmi rádu *Lepidoptera*. Iné typické členy tohto rádu sú sivkavec stavikrvový, mora kapustová, siatica ypsilonová, *Helicoverpa zea*, obaľovač jablčný, moľovitý, vijačka papriková, obaľovačovité, mlynárík kapustný, kvetovka bavlníková, vreckovcovité, priadkovec americký, vijačka voštinová a sivkavec južný. Aktivita sa tiež pozoruje oproti nemeckým švábom, ktoré sú členom rádu *Dictyoptera* (alebo *Blattodea*). Iné členy tohto rádu sú orientálne šváby a americké šváby.

Aktivita proti larvám liskaviek uhorkových sa testovala nasledovne. Bujón(y) kultúry *Photorhabdus* (10 krát koncentrované, sterilizovaný filter), 2 % proteózový peptón #3 (10 krát koncentrovaný), purifikovaný toxínový komplex(y), 10 mM pufofosforečnanu sodného, pH 7,0 sa priamo aplikovali na povrch (okolo 1,5 cm<sup>2</sup>) umelej diéty (Rose, R. I. a McCabe, J. M. 1973. J. Econ. Entomol. 66, 398 - 400) v 40 µl podieloch. Toxínový komplex sa zriedil v 10 mM pufofosforečnanu sodného, pH 7,0. Diétne platne sa mohli vysušiť na vzduchu v sterilnej prietokovej peci a sondy sa zamorili s jedným neonátom *Diabrotica undecimpunctata howardi* (liskavka uhorková, SCR) vyliahnutým

z povrchovo sterilizovaných vajíčok. Platne sa zaizolovali a umiestnili do vlhkej rastovej komory a nechali sa tam pri 27 °C počas určeného času (3 až 5 dní). Potom sa stanovila mortalita a hmotnosť lariev. Vo všeobecnosti sa použilo 16 kusov hmyzu na spracovanie počas celého štúdia. Kontrolná mortalita bola vo všeobecnosti menšia ako 5 %.

Aktivita proti larvám lepidoptera sa testovala nasledovne. Bujón(y) kultúry *Photorhabdus* (10 krát koncentrované, sterilizovaný filter), kontrolné médium (2 % proteózový peptón #3), purifikovaný toxínový komplex(y), 10 mM pufor fosforečnanu sodného, pH 7,0 sa priamo aplikovali na povrch (okolo 1,5 cm<sup>2</sup>) štandardnej umelej lepidopterovej diéty (Stoneville Yellow diét) 40 µl podieloch. Diétne platne sa mohli vysušiť na vzduchu v sterilnej prietokovej peci a sondy sa zamorili s jedným neonátom. Vajíčka eur *Ostrinia nubilalis* a *Manduca sexta* sa získali z komerčných zdrojov a vyliahli sa v liahni, zatiaľ čo larvy *Heliothis virescens* sa dodávali interne. Nasledovalo zamorenie larvami a diétne platne sa zaizolovali a umiestnili do vlhkej rastovej komory a nechali sa v tme pri 27 °C počas určeného času. Mortalita a hmotnosť lariev sa určili na piaty deň. Kontrolná mortalita bola vo všeobecnosti v rozsahu od 0 do okolo 12,5 % pre kontrolné médium a menej ako 10 % pre fosforečnanový pufor.

Aktivita proti švábom sa testovala nasledovne. Koncentrované bujóny (10 krát) kultúr *Photorhabdus* a kontrolné médium (2 % proteózový peptón #3) sa priamo aplikovali na povrch (okolo 1,5 cm<sup>2</sup>) štandardnej lepidopterovej diéty (Stoneville Yellow diét) 40 µl podieloch. Diétne platne sa mohli vysušiť na vzduchu v sterilnej prietokovej peci a sondy sa zamorili s jedným CO<sub>2</sub> anestetizovaným prvým vývojovým štádiom *Blatella germanica*. Nasledovalo zamorenie a diétne platne sa zaizolovali a umiestnili do vlhkej rastovej komory a nechali sa v tme pri 27 °C počas určeného času. Mortalita a hmotnosť lariev sa určili na piaty deň. Kontrolná mortalita bola menej ako 10 %.

Tabuľka 36

Sledované insekticídne spektrum bujónov z rôznych kmeňov  
*Photorhabdus*

Kmene <i>Photorhabdus</i>	Citlivé* hmyzie druhy
<i>P. zealandica</i>	1**, 2, 4
<i>P. hepialus</i>	1, 2, 4
HB-Arg	1, 2, 4
HB Oswedo	1, 2, 4
HB Lewiston	1, 2, 4
K-122	1, 4
HMGD	1, 4
Indicus	1, 2, 4
GD	2, 4
PWH-5	1, 2, 4
Megidis	1, 2, 4
HF-85	1, 2, 4
A. Cows	1, 4
MP1	1, 2, 4
MP2	1, 2, 4
MP3	4
MP4	1, 4
MP5	4
GL98	1, 4
GL101	1, 4, 5
GL138	1, 2, 4
GL155	1, 4
GL217	1, 2, 4
GL257	1, 4
DEP1	1, 4
DEP2	1, 2, 3, 4
DEP3	4

- \* = <sup>3</sup> 25 % mortality a/alebo ihibície rastu oproti kontrole  
\*\* = 1. sivkavec, 2. vijačka kukuričná, 3. *Manduca sexta*,  
4. liskavka uhorková, 5. šváb

#### Príklad 24

Souther analýza kmeňov *Photorhabdus* nie W-14 s použitím W-14 génovej sondy

Kmene *Photorhabdus* rástli na agare 2 % proteózového peptónu #3 (Difco Laboratories, Detroit, MI) a insekticídna toxínová spôsobilosť sa udržiavala opakovaním bioskúšky po prestupe. 50 ml trepanej kultúry sa vytvorilo v 175 ml nádobách v médiu 2 % proteózového peptónu #3, pri raste pri 28 °C a 150 rpm približne 24 hodín. 15 ml tejto kultúry sa centrifugovalo (700 x g, 30 min) a mrazilo sa vo svojom médiu pri -20 °C a znovu sa rozmrazovalo, pokiaľ sa nerozmrazilo (pomaly v ľadovej vode) na izoláciu DNA. Rozmrazené kultúry W-14 sa centrifugovali (900 x g, 15 min, 4 °C) a plávajúci bunkový oranžový mukopolysacharidový materiál sa odstránil. Zostávajúci bunkový materiál sa centrifugoval (25 000 x g, 4 °C) na peletovanie bakteriálnych buniek a médium sa odstránilo a vyradilo.

Celková DNA sa izolovala adaptáciou CTAB metódy, ktorá je opísaná v časti 2.4.1 od Ausubela a kol. (1994). Modifikácia zahŕňala vysokoslaný šok, a všetky objemy vzrástli 10 krát nad "miniprep" odporúčané objemy. Všetky centrifugácie sa uskutočňovali pri 4 °C, ak nebolo špecifikované inak. Peletované bakteriálne bunky sa resuspendovali v TE pufri (10 mM Tris-HCl, 1 mM EDTA, pH 8,0) na konečný objem 10 ml, potom sa pridalo 12 ml 5 M NaCl; táto zmes sa centrifugovala 20 min pri 15 000 x g. Peleta sa resuspendovala v 5,7 ml TE a 300 µl 10 % SDS a do suspenzie sa dodalo 60 µl 20 mg/ml proteínázy K (v sterilnej destilovanej vode, Gibco BRL Products, Grand Island, NY). Zmes sa inkubovala pri 37 °C počas 1 hod; potom sa pridalo približne 10 mg lyzozýmu (Worthington Biochemical Corp.,

Freehold, NJ). Po ďalšej 45 min inkubácii sa dodal 1 ml 5 M NaCl a 800 µl roztoku CTAB/NaCl (10 % hmotn./obj. CTAB, 0,7 M NaCl). Táto zmes sa inkubovala 10 min pri 65 °C potom sa zľahka premiešavala a ďalej inkubovala a premiešavala približne počas 20 min na purifikáciu bunkového materiálu. Dodal sa ekvivalentný objem roztoku chloroform/izoamylalkohol (24 : 1, obj. : obj.), zmes sa veľmi jemne miešala a fázy sa separovali centrifugáciou pri 12 000 x g počas 15 min. Horná (vodná) fáza sa jemne odstránila s pipetou so širokým otvorom a dvakrát sa extrahovala ako je vyššie uvedené s ekvivalentným objemom PCI (fenol/chloroform/izoamylalkohol; 50 : 49 : 1, obj. : obj. : obj.; zrovnovážené s 1 M Tris-HCl, pH 8,0; Intermountain Scientific Corporation, Kaysville, UT). DNA vyzrážaná s 0,6 objemu izopropanolu sa jemne preniesla na sklenenú tyčinku, dvakrát sa premyla so 70 % etanolom, vysušila sa a rozpustila v 2 ml STE (10 mM Tris-HCl, 10 mM NaCl, 1 mM EDTA, pH 8). Táto zmes podľa stanovenia optickej hustoty pri 260 nm obsahovala 2,5 mg/ml DNA.

Identifikácia *Bgl II/Hind III* hybridizujúcich fragmentov k *tc*-génovým špecifickým skúškam

Približne 10 µg genómovej DNA sa úplne rozštiepilo s okolo 30 jednotkami každého *Bgl II* a *Hind III* (NEB) počas 180 min, cez noc sa mrazilo, potom zohrievalo pri 65 °C počas 5 min a podrobilo sa elektroforéze v 0,8 % agarózovom géli (Seakem LE, 1X TEA, 80 volt, 90 min). DNA sa zafarbila s etídiom bromidom (50 µg/ml) ako je opísané vyššie a fotografovala pri ultrafialovom svetle. DNA fragmenty v agarózovom géli sa vystavili depurinácii (5 min v 0,2 M HCl), denaturácii (15 min v 0,5 M NaOH, 1,5 M NaCl) a neutralizácii (15 min v 0,5 M Tris-HCl, pH 8,0, 1,5 M NaCl), s tromi premytiami destilovanou vodou medzi každým krokom DNA sa preniesla pomocou Southern blotting z gélu na nylónovú membránu (Amersham, Arlington Heights, IL) s použitím vysoko slaného protokolu (20X SSC) ako je opísané v časti od 2.9 Ausubela a kol. (CPMB, op. cit). Prenesená

DNA sa potom zosieťovala s nylónovou membránou ultrafialovým svetlom s použitím sady Stratagene UV Stratalinker na samozosieťovanie. Membrána sa uskladňovala suchá pri 25 °C do použitia.

Hybridizácia sa uskutočnila s použitím systému priameho označovacieho a detekčného systému ECL<sup>TM</sup> (Amersham, Arlington Heights, IL), nasledovaná protokolmi poskytnutými výrobcom. V stručnosti sa sondy pripravili kovalentným spojením denaturovanej DNA k enzýmu chrenovej peroxidázy. Raz označená sonda sa použila v hybridizačných podmienkach, ktoré udržali enzymatickú aktivitu. Nehybridizovaná sonda sa odstránila dvomi jemnými vypratiami vždy po 20 min pri 42 °C v 0,5 x SSC, 0,4 % SDS a 6 M močovine. Po tomto nasledovali dve vypratia vždy po 5 min pri izbovej teplote v 2 x SSC. Ako je nariadené výrobcom reagenty ECL<sup>TM</sup> sa použili na detekciu hybridizujúcich DNA pásov. Zistili sa niektoré faktory, ktoré vplyvajú na schopnosť detegovať gény bez vzťahu medzi rôznymi kmeňmi *Photobacterium* a kmeňom W-14. Po prvé sa neuplatnila v týchto hybridizáciách vysoká "stringency". Zo stavu techniky je známe, že meniaci sa "stringency" hybridizácie a pracie podmienky budú mať vplyv na štruktúru a intenzitu hybridizujúcich pásov. Po druhé Southern blotové "blot to blot" variácie budú vplyvať na mobilitu hybridizujúcich pásov a odhady molekulovej hmotnosti. Preto sa W-14 zahrnul ako štandard na všetky Southern bloty.

Génové špecifické sondy odvodené od W-14 toxínových génov sa použili v týchto hybridizáciách. Nasledujúci text je zoznam špecifických koordinát v každej génovej sekvencii, ku ktorej korešponduje sonda. Sonda špecifická pre *tcaB<sub>i</sub>/B<sub>ii</sub>*: 1174 až 3642 Sekv. ID #25, sonda špecifická pre *tcaC*: 3637 až 6005 Sekv. ID #25, sonda špecifická pre *tcbA*: 2097 až 4964 Sekv. ID #11, sonda špecifická pre *tcdA*: 1660 až 4191 Sekvencovanie ID #46. Nasledujúce tabuľky sumarizujú analýzy Southern blot kmeňov *Photobacterium*. V prípade, že sa vyskytujú hybridizačné sondy, hybridizované fragmenty sa označili buď identické alebo rozdielne od štruktúry pozorovanej v kmeni W-14.

Tabuľka 37

Southern analýza kmeňov *Photorhabdus*

Kmene	tcdA	tcbA	tcaC	tcaB <sub>i</sub> /ii
WX-1	D	D	D	D
WX-2	D	D	-	D
WX-3	D	D	D	D
WX-4	D	D	ND	D
WX-5	D	D	D	D
WX-6	D	D	D	D
WX-7	D	D	ND	D
WX-8	D	D	D	D
WX-9	ND	D	D	D
WX-10	ND	D	D	D
WX-11	ND	D	D	D
WX-12	D	D	D	D
WX-13	D	D	D	D
WX-14	D	D	D	D
WX-15	D	D	D	D
HP88	D	-	D	D
Hm	D	-	D	D
Hb	D	-	D	-
H9	D	-	I	D
B2	D	-	D	-
NC-1	D	-	D	D
WIR	D	-	D	D
W30	D	D	D	D
W-14	I	I	I	I



ND = nestanovené; - = nedetegovateľný hybridizačný produkt;  
I = identická fragmentová štruktúra; D = rôzna fragmentová štruktúra.

Tabuľka 38

Southern analýza kmeňov *Photorhabdus*

Kmene	tcdA	tcbA	tcaC	tcaB <sub>i/ii</sub>
K-122	3.3, 2.8	D	-	ND
PWH-5	+	D	D	-
Indicus	D	D	3.0	I
Megidis	D	D	D	-
GD	D	D	D	-
HF-85	D	D	D	-
MP 3	D	-	D	-
MP 1	D	+	D	-
A. Cows	D	+	D	-
HB-Arg	D	ND	D	-
HMGD	D	D	D	-
HB Lewiston	D	D	D	-
HB Oswego	D	D	D	-
W-14	I	I	I	I

ND = nestanovené; - = nedetegovateľný hybridizačný produkt;  
I = identická fragmentová štruktúra; D = rôzna fragmentová štruktúra; + = nestanovená hybridizačná fragmentová štruktúra.



ND = nestanovené; - = nedetegovateľný hybridizačný produkt;  
I = identická fragmentová štruktúra; D = rôzna fragmentová štruktúra; + = nestanovená hybridizačná fragmentová štruktúra.

Z týchto analýz je zrejmé, že homológy génov W-14 sú dispergované cez rôzne kmene *Photobhabdus* ako je evidentné z rozdielov vo veľkosti génových fragmentov medzi W-14 a inými kmeňmi.

#### Príklad 25

N-koncové aminokyselinové sekvencie toxínových komplexov peptidov z rôznych kmeňov *Photobhabdus*

Vzťahy peptidov izolovaných z rôznych kmeňov *Photobhabdus*, ako bolo opísané v príklade 14, sa podrobili sekvencovaniu N-koncových aminokyselín. N-koncové aminokyselinové sekvencie toxínových peptidov v rôznych kmeňoch sa porovnali s toxínovými peptidmi W-14. V tabuľke 40 porovnanie toxínových peptidov porovnaných s údajmi ukázalo, že identické alebo homologické (aspoň 40 % podobnosti k W-14 génu/peptidom) toxínové peptidy boli prítomné vo všetkých kmeňoch. Napríklad N-koncová aminokyselinová sekvencia TcaC, SEKV. ID. Čís. 2 sa zistila ako identická s 160 kDa peptidom v HP88, ale tiež homológy boli prítomné v kmeňoch WIR, H9, Hb, WX-1 a Hm. Niektoré peptidy alebo homológy sa pozorovali v iných kmeňoch; avšak nie všetky peptidy boli sekvencované pre toxínové komplexy z iných kmeňov vďaka blokovaniu N-zakončenia alebo nízkemu nadbytku. Navyše mnohé iné N-koncové aminokyselinové sekvencie (SEKV. ID. Čís. 82 až 88) sa získali pre toxínové komplexy peptidov z iných kmeňov, ktoré mali podobnosť k peptidom W-14 a v niektorých prípadoch boli identické jeden s druhým. Napríklad identická aminokyselinová sekvencia SEKV. ID. Čís. 82 sa získala pre 64 kDa peptid prítomný v oboch kmeňoch HP88 a Hb a homologické sekvencie pre peptid 70 kDa v kmeni NC-1 (SEKV. ID. Čís. 83).

Tabuľka 40

Porovnanie homológie N-koncových sekvencií medzi proteínmi izolovanými z kmeňov nie W-14

W-14 Peptid	W-14 Gén	W-14 SEQ ID	SEQ ID Čís:	Kmeň	Iden-tický	Homoló-gia
TcaA <sub>ii</sub>	tcaA	15				
TcaA <sub>iii</sub>	tcaA	4				
TcbB <sub>i</sub>	tcaB	3	76	H9	-	74 kDa
			76	Hm	-	71 kDa
TcbB <sub>ii</sub>	tcaB	5		H9	61 kDa	-
				Hm	61 kDa	-
TcaC	tcaA	2	72	Hb	-	160 kDa
				HP88	160 kDa	-
			73	WIR	-	170 kDa
			74	H9	-	180 kDa
			75	Hm	-	170 kDa
			80	WX-1	-	170 kDa
TcbA <sub>ii</sub>	tcbA	1				
TcbA <sub>iii</sub>	tcbA	40				
TccA	tcbA	8	77	Hb	-	81 kDa
TccB	tccB	7		WX-1	170 kDa	-
				WX-2	180 kDa	-
				WX-14	180 kDa	-
				WIR	170 kDa	-
			78	H9	-	170 kDa
				NC-1	140 kDa	-
			79	Hm	-	190 kDa
TcdA <sub>ii</sub>	tcca					
TcdA <sub>iii</sub>	tcca	41	81	Hb	57 kDa	-
				H9	-	69 kDa
?	?	9		Hb	86 kDa	-
				HP88	86 kDa	-

Homológia sa vzťahuje k aminokyselinovým sekvenciám, ktoré mali podobnosť aspoň 40 % k W-14 génu/peptidom. Podobné zvyšky sa identifikovali ako členy v jednej z nasledujúcich skupín: (P, A, G, S, T); (Q, N, E, B, D, Z); (H, K, R); L, I, V, M) a (F, Y, W).

Príklad 26

Imunologická analýza kmeňov *Photorhabdus*

Bujóny kultúr kmeňov *Photorhabdus* sa koncentrovali 10 až

15 krát s použitím ultrafiltračného zariadenia Centriprep-10 (Amicon, Inc. Beverly, MA 01915). Koncentrácia proteínov bola v rozsahu od 0,3 do 3,0 mg na ml. 10 až 20  $\mu$ g celkového proteínu sa umiestnilo do každej sondy a zalialo sa 4 až 20 % polyakrylamidovým gélom (Integrated Separation Systems, Natick, MA 01760). Gélová elektroforéza sa uskutočňovala 1,25 hod. s použitím konštantného prúdového nastavenia pri 25 mA na gél. Gél sa elektroblotoval na nitrocelulóзовú membránu Hybond-ECL<sup>TM</sup> (Amersham Corporation, Arlington Hts, IL 60005) s použitím polosuchého elektroblottera (Pharmacia Biotech Inc., Piscataway, NJ 08854). Konštantný prúd sa aplikoval pri 0,75 mA na cm počas 2,5 hod. Membrána sa blokovala s 10 % mliekom TBST (25 mM Tris-HCl, pH 7,4, 136 mM NaCl, 2,7 mM KCl, 0,1 % Tween 20) počas jednej hodiny pri izbovej teplote. Každá primárna protilátka sa rozpustila v 10 % mlieku/TBST na 1 : 500. Tiež sa použilo iné zriedenie od 1 : 50 až 1 : 1000. Membrána sa inkubovala v primárnej protilátke najmenej počas jednej hodiny. Potom sa dokonale vyprala s blokujúcim roztokom alebo TBST. Zriedenie sekundárnych protilátok 1 : 2000 (kozí antimyší IgG alebo kozí antikráličí TgG sa spojil s chrenovou peroxidázou; BioRoad Laboratories, Hercules, CA 94547) v 10 % mlieku/TBST sa aplikovalo na membránu, ktorá sa umiestnila na plochú trepačku počas jednej hodiny. Detekcia proteínu sa uskutočnila s použitím ECL detekčným vybavením (Enhanced Chemiluminescence) (Amersham International).

Panel peptidových špecifických protilátok generovaný proti W-14 peptidom sa použil na charakterizáciu proteínového zloženia bujónov z deviatich kmeňov *Photorhabdus* nie W-14 s použitím analýzy Western blot. Navyše sa použila jedna monoklonálna protilátka (MAb-C5F2), ktorá uznala proteín TcbA<sub>iii</sub> vo W-14 derivovanom toxínovom komplexe. Výsledky (tabuľka 39) ukázali vzájomné uznanie protilátok k niektorým proteínom v týchto bujónoch. V niektorých prípadoch proteíny, ktoré boli uznané protilátkami mali tú istú veľkosť ako cieľové peptidy W-14. V iných prípadoch proteíny, ktoré boli uznané protilátkami boli

menšie ako cieľové peptidy W-14. Tieto údaje ukazujú, že niektoré kmeňe *Photobacterium* nie W-14 môžu produkovať podobné proteíny ako kmeň W-14. Rozdiel by mohol byť vďaka delícii alebo spracovaniu proteínov alebo degradačnému procesu. Niektoré z kmeňov neobsahovali proteín(y), ktoré by mohli byť uznané niektorými protilátkami, avšak je možné, že koncentrácia je výrazne nižšia ako tá, ktorá bola pozorovaná pre peptidy W-14. Pri porovnaní s rôznymi homológmi toxínového peptidu, tieto výsledky ukázali roznosť medzi kmeňmi *Photobacterium*.

Tabuľka 41

Vzájomné rozlíšenie monoklonálnymi protilátkami alebo polyklonálnymi protilátkami generovanými proti peptidom W-14 na proteín(y) v bujónoch vybraných z kmeňov *Photobacterium* nie W-14

Photo- rhabdus kmeň	MAD C5F2	PAb TcdA ii- syn	PAb TcdA iii- syn	PAb TcaC -syn	PAb TcaB ii- syn	PAb TcbA iii- syn	PAb TcaB i- syn	PAb TcaA ii- syn	PAb TcaA iii- syn
MP1	-	+	+	+	-	+	+	+	+
MP2	+	+	+	+	-	+	+	+	+
MP3	-	+	+	+	-	NT	+	+	-
A.cows	-	+	+	+	-	NT	+	+	+
Hb-osw	-	-	NT	+	+	NT	+	+	+
H-Arg	-	+	+	+	-	NT	+	+	+
Hb-leu	-	+	+	+	-	NT	+	+	+
Indicus	+	+	+	+	+	NT	+	+	+
HF85	-	+	+	+	-	+	+	+	+
W-14	+	+	+	+	+	+	+	+	+

+: pozitívna reakcia; -: negatívna reakcia; NT: netestované.

Zvyšné kmeňe *Photorhabdus* nie W-14 sa charakterizovali analýzou Western blot s použitím bujónu kultúry a/alebo čiastočne purifikovaných proteínových frakcií ako antigénu. Panel protilátok zahŕňa MAb-C5F2, MAb-DE1 (rozlišujúci TcdA<sub>II</sub>), PAbDE2 (rozlišujúci TcaB), PAb-TcbA<sub>ii</sub>-syn, PAb-TcaC-syn, PAb-TcaB<sub>ii</sub>-syn, PAb-TcbA<sub>iii</sub>-syn, PAb-TcaB<sub>i</sub>-syn. Tieto protilátky ukázali vzájomnú reaktivitu s proteínmi v bujóne a v čiastočne purifikovaných frakciách kmeňov nie W-14.

Údaje ukazujú, že protilátky by sa mohli použiť na identifikáciu proteínov v bujóne ako aj v čiastočne purifikovaných proteínových frakciách.

Tabuľka 42

Vzájomné rozlíšenie monoklonálnymi protilátkami alebo polyklonálnymi protilátkami generovanými proti peptidom W-14 na proteín(y) v bujónoch alebo čiastočne purifikovaných proteínových frakciách vybraných z kmeňov *Photorhabdus* nie W-14

Photo- rhabdus kmeň	Monoklonálne protiilátky		Polyklonálne protlilátky					
	Mab C5F2	Mab- DE1	PAb- DE2	PAb TcbA <sub>ii</sub> -syn	PAb TcaC- syn	PAb TcaB <sub>ii</sub> -syn	PAb- TcbA <sub>iii</sub> -syn	PAb TcaB <sub>i</sub> -syn
WX-1	+	+	+	+	+	+	+	+
WX-2	+	+	+	+	+	+	NT	+
WX-3	+	NT	+	NT	NT	NT	NT	NT
WX-5	+	NT	+	NT	NT	NT	NT	NT
WX-6	+	NT	NT	NT	NT	NT	NT	NT
WX-7	+	+	+	+	+	+	NT	+
WX-8	+	NT	NT	NT	NT	NT	NT	NT

Tabuľka 42 - pokračovanie

Photo- rhabdus kmeň	Monoklonálne protilátky		Polyklonálne protilátky					
	Mab C5F2	Mab- DE1	PAb- DE2	PAb TcbA <sub>ii</sub> -syn	PAb TcaC- syn	PAb TcaB <sub>ii</sub> -syn	PAb- TcbA <sub>iii</sub> -syn	PAb TcaB <sub>i</sub> -syn
WX-9	+	NT	NT	NT	NT	NT	NT	NT
WX-10	-	NT	NT	NT	NT	NT	NT	NT
WX-12	+	+	+	+	+	+	+	+
WX-14	+	+	+	+	NT	+	NT	+
WX-15	+	NT	NT	NT	NT	NT	NT	NT
W30	+	+	+	NT	NT	NT	NT	NT
Hb	-	NT	+	NT	+	NT	-	+
H9	-	-	+	NT	+	+	NT	NT
Hm	-	NT	+	+	+	+	NT	++
HP88	-	NT	+	-	+	-	-	+
NC-1	+	-	+	+	+	+	NT	+
WIR	-	NT	+	+	+	+	+	+
W-14	+	+	+	+	+	+	+	+

+: pozitívna reakcia; -: negatívna reakcia; NT: netestované.

#### Príklad 27

Bakteriálna expresia kódujúcej oblasti *tcdA*

Spracovanie génu *tcdA* pre bakteriálnu expresiu

Konce 5' a 3' kódujúcej oblasti *tcdA* (SEKV. ID ČÍS. 46) sa modifikovali dodaním užitočného klonujúceho miesta na inzerciu



segmentu do heterologických expresných vektorov. Konce sa modifikovali s použitím špecifických primérov v polymerázových reťazových reakciách (PCR), ktoré sa v podstate uskutočnili ako je opísané v príklade 8. Sada primérov ako je opísané nižšie sa použila v konjunkcii s kozmidom 21D2.4 ako templátom na vytvorenie produktov s určito modifikovanými koncami.

Prvá sada primérov sa použila na modifikáciu konca 5' génu, na inzerciu špecifického miesta *Nco I* ako kodónového iniciátora s použitím pokračujúceho priméru A0F1 (5' GAT CGA TCG ATC CAT GGC CAA CGA GTC TGT AAA AGA GAT ACC TGA TG TAT TAA AAA GCC AGT GTG 3') a dodaním špecifických miest *Bgl II*, *Sal I* *Not I* na uľahčenie inzercie zvyšku génu s použitím reverzného priméru A0R1 (5' GAT CGA TCG TAC GCG GCC GCT CGA TCG ATC GTC GAC CCA TTG ATT TGA GAT CTG GGC GGC GGG TAT CCA GAT AAT AAA CGG AGT CAC 3').

Iná PCR reakcia sa zvolila na modifikáciu konca 3' génu dodaním ďalšieho stop kodónu a vhodných reštrikčných miest na klonovanie. Pokračujúci primér A0F2 (3' ACT GGC TGC GTG GTC GAC TGG CGG CGA TTT ACT 3') sa použil na amplifikáciu cez špecifické miesto *Sal I* v géne, neskôr sa použil na klonovanie modifikovaného konca 3'. Reverzný primér A0R2 (5' CGA TGC ATG CTG CGG CCG CAG GCC TTC CTC GAG TCA TTA TTT AAT GGT GTA GCG AAT ATG CAA AAT 3') sa použil na inzerciu druhého stop kodónu (TGA) a klonujúcich miest *Xho I*, *Stu I* a *Not I*. Bakteriálny expresný vektor pET27b (Novagen, Madison, WI) sa modifikoval na vymazanie miesta *Bgl II* na pozícii 446 podľa štandardných techník molekulárnej biológie.

Produkty 497 bp z prvej amplifikačnej reakcie (A0F1 + A0R1) na modifikáciu konca 5' génu sa ligátovali k modifikovanému vektoru pET27b podľa dodávateľových inštrukcií. DNA sekvencie amplifikovanej časti troch izolátov sa stanovili s použitím dodávateľom odporúčaných primérov a sekvencovacích metód, ktoré sú opísané v predchádzajúcom texte. Sekvencia všetkých izolátov bola tá istá.

Jeden izolát sa potom použil ako klonujúci vektor na inzerciu strednej časti génu *tcdA* na fragment 6341 bp *bgl II* až *Sal I*. Výsledný klon sa nazval MC4 a obsahoval všetky ale najviac 3' časti *tcdA* kódujúcej sekvencie. Nakoniec sa na skompletovanie plnej dĺžky kódujúcej oblasti ligátoval PCR produkt 832 bp, ktorý pochádza z druhej PCR amplifikácie (A0F2 + A0R2) na modifikáciu konca 3' génu, na izoláciu MC4 na fragment *Sal I* až *Not I* podľa štandardných techník molekulárnej biológie. Kódujúca oblasť *tcdA* sa sekvencovala a zistilo sa, že je úplná. Výsledný plazmid sa nazval pDAB2035.

Konštrukcia plazmidov pDAB2036, pDAB2037 a pDAB2038 na bakteriálnu expresiu *tcdA*

Kódujúca oblasť *tcdA* sa oddelila od plazmidu pDAB2035 s reštrikčnými enzýmami *Nco I* a *Xho I* a gél sa purifikoval. Fragментy sa ligátovali na miesta *Nco I* a *Xho I* expresného vektora pET15 na vytvorenie plazmidu pDAB2036. Ďalej sa pDAB2035 delil s *Nco I* a *Not I* na uvoľnenie *tcdA* kódujúcej oblasti, ktorá sa ligátovala na miesta *Nco I* a *Not I* expresného vektora pET28b na vytvorenie plazmidu pDAB2037. Nakoniec sa plazmid pDAB2035 delil s *Nco I* a *Stu I* na uvoľnenie kódujúcej oblasti *tcdA*. Tento fragment sa ligátoval na expresný vektor Trc99a, ktorý sa delil s *Hind III*, potom nasledovalo spracovanie s T4 DNA polymerázou na zaoblenie koncov. Vektor sa potom delil s *Nco I* a ligátoval sa s *Nco I*/*Stu I* oddelený *tcdA* fragment. Výsledný plazmid sa nazval pDAB2038.

Expresia *tcdA* od plazmidu pDAB2038

Plazmid pDAB2038 sa transformoval na bunky *BL21* a bol expresovaný ako bolo opísané vyššie pre plazmid pDAB2033 v príklade 19.

Purifikácia *tcdA* od *E. coli*

Expresná kultúra sa centrifugovala pri 10 300 x g počas 30 min a zozbierala sa v supernatante. Zriedila sa dvomi objemami vody a aplikovala sa na pórovitú kolónu 50 HQ (Perspective Systems, MA) pri rýchlosti toku 7,5 ml/min (1,6 x 10 cm), ktorá sa zrovnovážnila s pufrom 10 mM fosforečnanu sodného, pH 7,0 (pufor A). Kolóna sa premyla s pufrom A, pokiaľ sa optická hustota pri 280 nm nevrátila na základnú úroveň. Proteínové väzby ku kolóne sa potom eluovali s 1 M NaCl v pufri A.

Frakcia sa uložila v 20 ml podieloch do kolóny gélovej filtrácie Sepharose CL-4B (2,6 x 100 cm), ktorá bola zrovnovážená s pufrom A. Proteín sa eluoval v pufri A pri rýchlosti toku 0,75 ml/min. Frakcie s retenčným časom medzi 260 a 460 minútami sa združili a aplikovali sa do kolóny Mono Q 5/5 pri 1 ml/min, ktorá bola zrovnovážená s 20 mM Tris-HCl, pH 7,0 (pufor B). Kolóna sa premyla s pufrom B, pokiaľ sa optická hustota pri 280 nm nevrátila na základnú úroveň. Proteínové väzby ku kolóne sa eluovali s lineárnym gradientom od 0 do 1 M NaCl v pufri B pri 1 ml/min počas 30 min. Zozbieral sa 1 mililiter frakcií, sériovo sa zriedili a podrobili sa SCR bioskúške. Zistilo sa že, frakcie eluované medzi 0,1 a 0,3 M NaCl majú najvyššiu insekticídnu aktivitu. Western analýzy aktívnych frakcií s použitím protilátky pAb TcdA<sub>ii</sub>-syn a protilátky pAb TcdA<sub>iii</sub>-syn indikujú prítomnosť peptidov TcdA<sub>ii</sub> a TcdA<sub>iii</sub>.

ZOZNAM SEKVENCÍ

(1) VŠEOBECNÉ INFORMÁCIE

(i) ŽIADATEĽ: Jerald C. Ensign

David J. Bowen

James Petell

Raymond Fatig

Sue Schoonover

Richard Ffrench - Constant

Gregory L. Orr

Donald J. Merlo

Jean L. Roberts

Thomas A. Rocheleau

(ii) NÁZOV VYNÁLEZU: Insekticídne proteínové toxíny  
z *Photorhabdus*

(iii) POČET SEKVENCÍ: 88

(iv) KONTAKTNÁ ADRESA:

(A) ADRESÁT: DowElanco

(B) ULICA : 9330 Zionsville Road

(C) MESTO : Indianapolis

(D) ŠTÁT : IN

(E) KRAJINA: USA

(F) PSČ: : 46268

(v) SPÔSOB ČITATEĽNOSTI NA POČÍTAČI:

(A) TYP MÉDIA: Disketa

(B) POČÍTAČ : Kompatibilný IBM PC

(C) OPERAČNÝ SYSTÉM: PC-DOS/MS-DOS

(D) SOFTWARE : Patent Release # 1.0,  
verzia č. 1.30

(vi) SÚČASNÉ ÚDAJE PRIHLÁŠKY:

(A) ČÍSLO PRIHLÁŠKY:

(B) DÁTUM REGISTRÁCIE:

(C) ZARADENIE:

(vii) PREDOSLÉ ÚDAJE PRIHLÁŠKY:

(A) ČÍSLO PRIHLÁŠKY: US 08/063 615

(B) DÁTUM REGISTRÁCIE: 18. máj 1993

(vii) PREDOSLÉ ÚDAJE PRIHLÁŠKY:

(A) ČÍSLO PRIHLÁŠKY: US 08/395 497

(B) DÁTUM REGISTRÁCIE: 28. február 1995

(vii) PREDOSLÉ ÚDAJE PRIHLÁŠKY:

(A) ČÍSLO PRIHLÁŠKY: US 60/007 255

(B) DÁTUM REGISTRÁCIE: 6. november 1995

(vii) PREDOSLÉ ÚDAJE PRIHLÁŠKY:

(A) ČÍSLO PRIHLÁŠKY: US 08/608 423

(B) DÁTUM REGISTRÁCIE: 28. február 1996

(vii) PREDOSLÉ ÚDAJE PRIHLÁŠKY:

(A) ČÍSLO PRIHLÁŠKY: US 08/705 484

(B) DÁTUM REGISTRÁCIE: 28. august 1996

(vii) PREDOSLÉ ÚDAJE PRIHLÁŠKY:

(A) ČÍSLO PRIHLÁŠKY: US 08/743 699

(B) DÁTUM REGISTRÁCIE: 6. november 1996

(viii) INFORMÁCIA O PRÁVNIKOVÍ/ZÁSTUPCOVI:

(A) MENO: Andrea T. Borucki

(B) REGISTRAČNÉ ČÍSLO: 33651

(C) REFERENČNÉ/ZOZNAMOVÉ ČÍSLO: 50301E

(ix) TELEKOMUNIKAČNÉ INFORMÁCIE

(A) TELEFÓN: 317 - 337 - 4846

(B) TELEFAX: 317 - 337 - 4847

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 1:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĹŽKA: 11 aminokyselín
- (B) TYP: aminokyselina
- (C) REĹAZEC:
- (D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: proteín

(v) TYP FRAGMENTU: N-koncový

(xi) OPIS SEKVENCIE: SEKV. ID ČÍS. 1 (TcbA<sub>ii</sub> N-koniec):

Phe	Ile	Gln	Gly	Tyr	Ser	Asp	Leu	Phe	Gly	Asn
1			5						10	

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 2:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĹŽKA: 12 aminokyselín
- (B) TYP: aminokyselina
- (C) REĹAZEC:
- (D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: proteín

(v) TYP FRAGMENTU: N-koncový

(xi) OPIS SEKVENCIE: SEKV. ID ČÍS. 2 (TcaC N-koniec):

Met	Gln	Asp	Ser	Pro	Glu	Val	Ser	Ile	Thr	Thr	Trp
1				5							10

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 3:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĹŽKA: 19 aminokyselín
- (B) TYP: aminokyselina
- (C) REŤAZEC:
- (D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: proteín

(v) TYP FRAGMENTU: N-koncový

(xi) OPIS SEKVENCIE: SEKV. ID ČÍS. 3 (TcaB<sub>i</sub> N-koniec):

Ser Glu Ser Leu Phe Thr Gln Thr Leu Lys Glu Ala Arg Arg  
1                      5                      10

Asp Ala Leu Val Ala  
15

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 4:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĹŽKA: 14 aminokyselín
- (B) TYP: aminokyselina
- (C) REŤAZEC:
- (D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: proteín

(v) TYP FRAGMENTU: N-koncový

(xi) OPIS SEKVENCIE: SEKV. ID ČÍS. 4 (TcaA<sub>iii</sub> N-koniec):

Ala Ser Pro Leu Ser Thr Ser Glu Leu Thr Ser Lys Leu Asn  
1                      5                      10







(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 9:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĚLKA: 16 aminokyselín
- (B) TYP: aminokyselina
- (C) REĚAZEC:
- (D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: proteín

8

(v) TYP FRAGMENTU: N-koncový

(xi) OPIS SEKVENCIE: SEKV. ID ČÍS. 9:

Met	Ile	Asn	Leu	Asp	Ile	Asn	Glu	Gln	Asn	Lys	Ile	Met	Val	Val
1				5					10					15
Ser														

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 10:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĚLKA: 20 aminokyselín
- (B) TYP: aminokyselina
- (C) REĚAZEC:
- (D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: proteín

(v) TYP FRAGMENTU: N-koncový

(xi) OPIS SEKVENCIE: SEKV. ID ČÍS. 10:

Ala	Ala	Lys	Asp	Val	Lys	Phe	Gly	Ser	Asp	Ala	Arg	Val	Lys	Met
1				5					10					15

Leu Arg Gly Val Asn

20

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 11:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĚLKA: 7515 párov báz
- (B) TYP: nukleová kyselina
- (C) REŤAZEC: dvojitý
- (D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: DNA (genómová)

(xi) CHARAKTERISTICKÝ ZNAK:

- (A) MENO/ZNAČKA: CDS
- (B) POLOHA: 1..7515

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 11 (tcbA gén):

ATG	CAA	AAC	TCA	TTA	TCA	AGC	ACT	ATC	GAT	ACT	ATT	TGT	CAG	AAA	CTG	48
Met	Gln	Asn	Ser	Leu	Ser	Ser	Thr	Ile	Asp	Thr	Ile	Cys	Gln	Lys	Leu	
1				5					10					15		
CAA	TTA	ACT	TGT	CCG	GCG	GAA	ATT	GCT	TTG	TAT	CCC	TTT	GAT	ACT	TTC	96
Gln	Leu	Thr	Cys	Pro	Ala	Glu	Ile	Ala	Leu	Tyr	Pro	Phe	Asp	Thr	Phe	
			20					25					30			
CGG	GAA	AAA	ACT	CGG	GGA	ATG	GTT	AAT	TGG	GGG	GAA	GCA	AAA	CGG	ATT	144
Arg	Glu	Lys	Thr	Arg	Gly	Met	Val	Asn	Trp	Gly	Glu	Ala	Lys	Arg	Ile	
		35					40					45				
TAT	GAA	ATT	GCA	CAA	GCG	GAA	CAG	GAT	AGA	AAC	CTA	CTT	CAT	GAA	AAA	192
Tyr	Glu	Ile	Ala	Gln	Ala	Glu	Gln	Asp	Arg	Asn	Leu	Leu	His	Glu	Lys	
	50					55					60					
CGT	ATT	TTT	GCC	TAT	GCT	AAT	CCG	CTG	CTG	AAA	AAC	GCT	GTT	CGG	TTG	240
Arg	Ile	Phe	Ala	Tyr	Ala	Asn	Pro	Leu	Leu	Lys	Asn	Ala	Val	Arg	Leu	
	65				70					75					80	
GGT	ACC	CGG	CAA	ATG	TTG	GGT	TTT	ATA	CAA	GGT	TAT	AGT	GAT	CTG	TTT	288
Gly	Thr	Arg	Gln	Met	Leu	Gly	Phe	Ile	Gln	Gly	Tyr	Ser	Asp	Leu	Phe	
				85					90					95		
GGT	AAT	CGT	GCT	GAT	AAC	TAT	GCC	GCG	CCG	GGC	TCG	GTT	GCA	TCG	ATG	336
Gly	Asn	Arg	Ala	Asp	Asn	Tyr	Ala	Ala	Pro	Gly	Ser	Val	Ala	Ser	Met	
			100				105						110			
TTC	TCA	CCG	GCG	GCT	TAT	TTG	ACG	GAA	TTG	TAC	CGT	GAA	GCC	AAA	AAC	384
Phe	Ser	Pro	Ala	Ala	Tyr	Leu	Thr	Glu	Leu	Tyr	Arg	Glu	Ala	Lys	Asn	
		115					120					125				
TTG	CAT	GAC	AGC	AGC	TCA	ATT	TAT	TAC	CTA	GAT	AAA	CGT	CGC	CCG	GAT	432
Leu	His	Asp	Ser	Ser	Ser	Ile	Tyr	Tyr	Leu	Asp	Lys	Arg	Arg	Pro	Asp	
	130					135					140					

TTA	GCA	AGC	TTA	ATG	CTC	AGC	CAG	AAA	AAT	ATG	GAT	GAG	GAA	ATT	TCA	480
Leu	Ala	Ser	Leu	Met	Leu	Ser	Gln	Lys	Asn	Met	Asp	Glu	Glu	Ile	Ser	
145					150					155					160	
ACG	CTG	GCT	CTC	TCT	AAT	GAA	TTG	TGC	CTT	GCC	GGG	ATC	GAA	ACA	AAA	528
Thr	Leu	Ala	Leu	Ser	Asn	Glu	Leu	Cys	Leu	Ala	Gly	Ile	Glu	Thr	Lys	
				165					170						175	
ACA	GGA	AAA	TCA	CAA	GAT	GAA	GTG	ATG	GAT	ATG	TTG	TCA	ACT	TAT	CGT	576
Thr	Gly	Lys	Ser	Gln	Asp	Glu	Val	Met	Asp	Met	Leu	Ser	Thr	Tyr	Arg	
			180					185						190		
TTA	AGT	GGA	GAG	ACA	CCT	TAT	CAT	CAC	GCT	TAT	GAA	ACT	GTT	CGT	GAA	624
Leu	Ser	Gly	Glu	Thr	Pro	Tyr	His	His	Ala	Tyr	Glu	Thr	Val	Arg	Glu	
		195					200						205			
ATC	GTT	CAT	GAA	CGT	GAT	CCA	GGA	TTT	CGT	CAT	TTG	TCA	CAG	GCA	CCC	672
Ile	Val	His	Glu	Arg	Asp	Pro	Gly	Phe	Arg	His	Leu	Ser	Gln	Ala	Pro	
	210					215					220					
ATT	GTT	GCT	GCT	AAG	CTC	GAT	CCT	GTG	ACT	TTG	TTG	GGT	ATT	AGC	TCC	720
Ile	Val	Ala	Ala	Lys	Leu	Asp	Pro	Val	Thr	Leu	Leu	Gly	Ile	Ser	Ser	
225					230					235					240	
CAT	ATT	TCG	CCA	GAA	CTG	TAT	AAC	TTG	CTG	ATT	GAG	GAG	ATC	CCG	GAA	768
His	Ile	Ser	Pro	Glu	Leu	Tyr	Asn	Leu	Leu	Ile	Glu	Glu	Ile	Pro	Glu	
				245					250					255		
AAA	GAT	GAA	GCC	GCG	CTT	GAT	ACG	CTT	TAT	AAA	ACA	AAC	TTT	GGC	GAT	816
Lys	Asp	Glu	Ala	Ala	Leu	Asp	Thr	Leu	Tyr	Lys	Thr	Asn	Phe	Gly	Asp	
			260					265					270			
ATT	ACT	ACT	GCT	CAG	TTA	ATG	TCC	CCA	AGT	TAT	CTG	GCC	CGG	TAT	TAT	864
	Ile	Thr	Thr	Ala	Gln	Leu	Met	Ser	Pro	Ser	Tyr	Leu	Ala	Arg	Tyr	Tyr
			275					280					285			
GGC	GTC	TCA	CCG	GAA	GAT	ATT	GCC	TAC	GTG	ACG	ACT	TCA	TTA	TCA	CAT	912
Gly	Val	Ser	Pro	Glu	Asp	Ile	Ala	Tyr	Val	Thr	Thr	Ser	Leu	Ser	His	
	290					295					300					
GTT	GGA	TAT	AGC	AGT	GAT	ATT	CTG	GTT	ATT	CCG	TTG	GTC	GAT	GGT	GTG	960
Val	Gly	Tyr	Ser	Ser	Asp	Ile	Leu	Val	Ile	Pro	Leu	Val	Asp	Gly	Val	
305					310					315					320	
GGT	AAG	ATG	GAA	GTA	GTT	CGT	GTT	ACC	CGA	ACA	CCA	TCG	GAT	AAT	TAT	1008
Gly	Lys	Met	Glu	Val	Val	Arg	Val	Thr	Arg	Thr	Pro	Ser	Asp	Asn	Tyr	
				325					330					335		
ACC	AGT	CAG	ACG	AAT	TAT	ATT	GAG	CTG	TAT	CCA	CAG	GGT	GGC	GAC	AAT	1056
Thr	Ser	Gln	Thr	Asn	Tyr	Ile	Glu	Leu	Tyr	Pro	Gln	Gly	Gly	Asp	Asn	
			340					345					350			
TAT	TTG	ATC	AAA	TAC	AAT	CTA	AGC	AAT	AGT	TTT	GGT	TTG	GAT	GAT	TTT	1104
Tyr	Leu	Ile	Lys	Tyr	Asn	Leu	Ser	Asn	Ser	Phe	Gly	Leu	Asp	Asp	Phe	
		355					360					365				
TAT	CTG	CAA	TAT	AAA	GAT	GGT	TCC	GCT	GAT	TGG	ACT	GAG	ATT	GCC	CAT	1152
Tyr	Leu	Gln	Tyr	Lys	Asp	Gly	Ser	Ala	Asp	Trp	Thr	Glu	Ile	Ala	His	
	370					375					380					
AAT	CCC	TAT	CCT	GAT	ATG	GTC	ATA	AAT	CAA	AAG	TAT	GAA	TCA	CAG	GCG	1200
Asn	Pro	Tyr	Pro	Asp	Met	Val	Ile	Asn	Gln	Lys	Tyr	Glu	Ser	Gln	Ala	
385					390					395					400	
ACA	ATC	AAA	CGT	AGT	GAC	TCT	GAC	AAT	ATA	CTC	AGT	ATA	GGG	TTA	CAA	1248
Thr	Ile	Lys	Arg	Ser	Asp	Ser	Asp	Asn	Ile	Leu	Ser	Ile	Gly	Leu	Gln	
				405					410					415		

AGA	TGG	CAT	AGC	GGT	AGT	TAT	AAT	TTT	GCC	GCC	GCC	AAT	TTT	AAA	ATT	1296
Arg	Trp	His	Ser	Gly	Ser	Tyr	Asn	Phe	Ala	Ala	Ala	Asn	Phe	Lys	Ile	
			420					425					430			
GAC	CAA	TAC	TCC	CCG	AAA	GCT	TTC	CTG	CTT	AAA	ATG	AAT	AAG	GCT	ATT	1344
Asp	Gln	Tyr	Ser	Pro	Lys	Ala	Phe	Leu	Leu	Lys	Met	Asn	Lys	Ala	Ile	
		435				440						445				
CGG	TTG	CTC	AAA	GCT	ACC	GGC	CTC	TCT	TTT	GCT	ACG	TTG	GAG	CGT	ATT	1392
Arg	Leu	Leu	Lys	Ala	Thr	Gly	Leu	Ser	Phe	Ala	Thr	Leu	Glu	Arg	Ile	
	450					455					460					
GTT	GAT	AGT	GTT	AAT	AGC	ACC	AAA	TCC	ATC	ACG	GTT	GAG	GTA	TTA	AAC	1440
Val	Asp	Ser	Val	Asn	Ser	Thr	Lys	Ser	Ile	Thr	Val	Glu	Val	Leu	Asn	
465				470						475					480	
AAG	GTT	TAT	CGG	GTA	AAA	TTC	TAT	ATT	GAT	CGT	TAT	GGC	ATC	AGT	GAA	1488
Lys	Val	Tyr	Arg	Val	Lys	Phe	Tyr	Ile	Asp	Arg	Tyr	Gly	Ile	Ser	Glu	
			485					490						495		
GAG	ACA	GCC	GCT	ATT	TTG	GCT	AAT	ATT	AAT	ATC	TCT	CAG	CAA	GCT	GTT	1536
Glu	Thr	Ala	Ala	Ile	Leu	Ala	Asn	Ile	Asn	Ile	Ser	Gln	Gln	Ala	Val	
			500					505					510			
GGC	AAT	CAG	CTT	AGC	CAG	TTT	GAG	CAA	CTA	TTT	AAT	CAC	CCG	CCG	CTC	1584
Gly	Asn	Gln	Leu	Ser	Gln	Phe	Glu	Gln	Leu	Phe	Asn	His	Pro	Pro	Leu	
		515					520					525				
AAT	GGT	ATT	CGC	TAT	GAA	ATC	AGT	GAG	GAC	AAC	TCC	AAA	CAT	CTT	CCT	1632
Asn	Gly	Ile	Arg	Tyr	Glu	Ile	Ser	Glu	Asp	Asn	Ser	Lys	His	Leu	Pro	
	530					535					540					
AAT	CCT	GAT	CTG	AAC	CTT	AAA	CCA	GAC	AGT	ACC	GGT	GAT	GAT	CAA	CGC	1680
Asn	Pro	Asp	Leu	Asn	Leu	Lys	Pro	Asp	Ser	Thr	Gly	Asp	Asp	Gln	Arg	
545					550					555					560	
AAG	GCG	GTT	TTA	AAA	CGC	GCG	TTT	CAG	GTT	AAC	GCC	AGT	GAG	TTG	TAT	1728
Lys	Ala	Val	Leu	Lys	Arg	Ala	Phe	Gln	Val	Asn	Ala	Ser	Glu	Leu	Tyr	
			565					570						575		
CAG	ATG	TTA	TTG	ATC	ACT	GAT	CGT	AAA	GAA	GAC	GGT	GTT	ATC	AAA	AAT	1776
Gln	Met	Leu	Leu	Ile	Thr	Asp	Arg	Lys	Glu	Asp	Gly	Val	Ile	Lys	Asn	
			580					585						590		
AAC	TTA	GAG	AAT	TTG	TCT	GAT	CTG	TAT	TTG	GTT	AGT	TTG	CTG	GCC	CAG	1824
Asn	Leu	Glu	Asn	Leu	Ser	Asp	Leu	Tyr	Leu	Val	Ser	Leu	Leu	Ala	Gln	
		595					600					605				
ATT	CAT	AAC	CTG	ACT	ATT	GCT	GAA	TTG	AAC	ATT	TTG	TTG	GTG	ATT	TGT	1872
Ile	His	Asn	Leu	Thr	Ile	Ala	Glu	Leu	Asn	Ile	Leu	Leu	Val	Ile	Cys	
	610					615					620					
GGC	TAT	GGC	GAC	ACC	AAC	ATT	TAT	CAG	ATT	ACC	GAC	GAT	AAT	TTA	GCC	1920
Gly	Tyr	Gly	Asp	Thr	Asn	Ile	Tyr	Gln	Ile	Thr	Asp	Asp	Asn	Leu	Ala	
625					630					635					640	
AAA	ATA	GTG	GAA	ACA	TTG	TTG	TGG	ATC	ACT	CAA	TGG	TTG	AAG	ACC	CAA	1968
Lys	Ile	Val	Glu	Thr	Leu	Leu	Trp	Ile	Thr	Gln	Trp	Leu	Lys	Thr	Gln	
			645					650						655		
AAA	TGG	ACA	GTT	ACC	GAC	CTG	TTT	CTG	ATG	ACC	ACG	GCC	ACT	TAC	AGC	2016
Lys	Trp	Thr	Val	Thr	Asp	Leu	Phe	Leu	Met	Thr	Thr	Ala	Thr	Tyr	Ser	
			660					665					670			
ACC	ACT	TTA	ACG	CCA	GAA	ATT	AGC	AAT	CTG	ACG	GCT	ACG	TTG	TCT	TCA	2064
Thr	Thr	Leu	Thr	Pro	Glu	Ile	Ser	Asn	Leu	Thr	Ala	Thr	Leu	Ser	Ser	
		675					680					685				

ACT TTG CAT GGC AAA GAG AGT CTG ATT GGG GAA GAT CTG AAA AGA GCA 2112  
 Thr Leu His Gly Lys Glu Ser Leu Ile Gly Glu Asp Leu Lys Arg Ala  
 690 695 700

ATG GCG CCT TGC TTC ACT TCG GCT TTG CAT TTG ACT TCT CAA GAA GTT 2160  
 Met Ala Pro Cys Phe Thr Ser Ala Leu His Leu Thr Ser Gln Glu Val  
 705 710 715 720

GCG TAT GAC CTG CTG TTG TGG ATA GAC CAG ATT CAA CCG GCA CAA ATA 2208  
 Ala Tyr Asp Leu Leu Leu Trp Ile Asp Gln Ile Gln Pro Ala Gln Ile  
 725 730 735

ACT GTT GAT GGG TTT TGG GAA GAA GTG CAA ACA ACA CCA ACC AGC TTG 2256  
 Thr Val Asp Gly Phe Trp Glu Glu Val Gln Thr Thr Pro Thr Ser Leu  
 740 745 750

AAG GTG ATT ACC TTT GCT CAG GTG CTG GCA CAA TTG AGC CTG ATC TAT 2304  
 Lys Val Ile Thr Phe Ala Gln Val Leu Ala Gln Leu Ser Leu Ile Tyr  
 755 760 765

CGT CGT ATT GGG TTA AGT GAA ACG GAA CTG TCA CTG ATC GTG ACT CAA 2352  
 Arg Arg Ile Gly Leu Ser Glu Thr Glu Leu Ser Leu Ile Val Thr Gln  
 770 775 780

TCT TCT CTG CTA GTG GCA GGC AAA AGC ATA CTG GAT CAC GGT CTG TTA 2400  
 Ser Ser Leu Leu Val Ala Gly Lys Ser Ile Leu Asp His Gly Leu Leu  
 785 790 795 800

ACC CTG ATG GCC TTG GAA GGT TTT CAT ACC TGG GTT AAT GGC TTG GGG 2448  
 Thr Leu Met Ala Leu Glu Gly Phe His Thr Trp Val Asn Gly Leu Gly  
 805 810 815

CAA CAT GCC TCC TTG ATA TTG GCG GCG TTG AAA GAC GGA GCC TTG ACA 2496  
 Gln His Ala Ser Leu Ile Leu Ala Ala Leu Lys Asp Gly Ala Leu Thr  
 820 825 830

GTT ACC GAT GTA GCA CAA GCT ATG AAT AAG GAG GAA TCT CTC CTA CAA 2544  
 Val Thr Asp Val Ala Gln Ala Met Asn Lys Glu Glu Ser Leu Leu Gln  
 835 840 845

ATG GCA GCT AAT CAG GTG GAG AAG GAT CTA ACA AAA CTG ACC AGT TGG 2592  
 Met Ala Ala Asn Gln Val Glu Lys Asp Leu Thr Lys Leu Thr Ser Trp  
 850 855 860

ACA CAG ATT GAC GCT ATT CTG CAA TGG TTA CAG ATG TCT TCG GCC TTG 2640  
 Thr Gln Ile Asp Ala Ile Leu Gln Trp Leu Gln Met Ser Ser Ala Leu  
 865 870 875 880

GCG GTT TCT CCA CTG GAT CTG GCA GGG ATG ATG GCC CTG AAA TAT GGG 2688  
 Ala Val Ser Pro Leu Asp Leu Ala Gly Met Met Ala Leu Lys Tyr Gly  
 885 890 895

ATA GAT CAT AAC TAT GCT GCC TGG CAA GCT GCG GCG GCT GCG CTG ATG 2736  
 Ile Asp His Asn Tyr Ala Ala Trp Gln Ala Ala Ala Ala Ala Leu Met  
 900 905 910

GCT GAT CAT GCT AAT CAG GCA CAG AAA AAA CTG GAT GAG ACG TTC AGT 2784  
 Ala Asp His Ala Asn Gln Ala Gln Lys Lys Leu Asp Glu Thr Phe Ser  
 915 920 925

AAG GCA TTA TGT AAC TAT TAT ATT AAT GCT GTT GTC GAT AGT GCT GCT 2832  
 Lys Ala Leu Cys Asn Tyr Tyr Ile Asn Ala Val Val Asp Ser Ala Ala  
 930 935 940

GGA GTA CGT GAT CGT AAC GGT TTA TAT ACC TAT TTG CTG ATT GAT AAT 2880  
 Gly Val Arg Asp Arg Asn Gly Leu Tyr Thr Tyr Leu Leu Ile Asp Asn  
 945 950 955 960

CAG GTT TCT GCC GAT GTG ATC ACT TCA CGT ATT GCA GAA GCT ATC GCC 2928  
 Gln Val Ser Ala Asp Val Ile Thr Ser Arg Ile Ala Glu Ala Ile Ala  
 965 970 975

GGT ATT CAA CTG TAC GTT AAC CGG GCT TTA AAC CGA GAT GAA GGT CAG 2976  
Gly Ile Gln Leu Tyr Val Asn Arg Ala Leu Asn Arg Asp Glu Gly Gln  
980 985 990

CTT GCA TCG GAC GTT AGT ACC CGT CAG TTC TTC ACT GAC TGG GAA CGT 3024  
Leu Ala Ser Asp Val Ser Thr Arg Gln Phe Phe Thr Asp Trp Glu Arg  
995 1000 1005

TAC AAT AAA CGT TAC AGT ACT TGG GCT GGT GTC TCT GAA CTG GTC TAT 3072  
Tyr Asn Lys Arg Tyr Ser Thr Trp Ala Gly Val Ser Glu Leu Val Tyr  
1010 1015 1020

TAT CCA GAA AAC TAT GTT GAT CCC ACT CAG CGC ATT GGG CAA ACC AAA 3120  
Tyr Pro Glu Asn Tyr Val Asp Pro Thr Gln Arg Ile Gly Gln Thr Lys  
1025 1030 1035 1040

ATG ATG GAT GCG CTG TTG CAA TCC ATC AAC CAG AGC CAG CTA AAT GCG 3168  
Met Met Asp Ala Leu Leu Gln Ser Ile Asn Gln Ser Gln Leu Asn Ala  
1045 1050 1055

GAT ACG GTG GAA GAT GCT TTC AAA ACT TAT TTG ACC AGC TTT GAG CAG 3216  
Asp Thr Val Glu Asp Ala Phe Lys Thr Tyr Leu Thr Ser Phe Glu Gln  
1060 1065 1070

GTA GCA AAT CTG AAA GTA ATT AGT GCT TAC CAC GAT AAT GTG AAT GTG 3264  
Val Ala Asn Leu Lys Val Ile Ser Ala Tyr His Asp Asn Val Asn Val  
1075 1080 1085

GAT CAA GGA TTA ACT TAT TTT ATC GGT ATC GAC CAA GCA GCT CCG GGT 3312  
Asp Gln Gly Leu Thr Tyr Phe Ile Gly Ile Asp Gln Ala Ala Pro Gly  
1090 1095 1100

ACG TAT TAC TGG CGT AGT GTT GAT CAC AGC AAA TGT GAA AAT GGC AAG 3360  
Thr Tyr Tyr Trp Arg Ser Val Asp His Ser Lys Cys Glu Asn Gly Lys  
1105 1110 1115 1120

TTT GCC GCT AAT GCT TGG GGT GAG TGG AAT AAA ATT ACC TGT GCT GTC 3408  
Phe Ala Ala Asn Ala Trp Gly Glu Trp Asn Lys Ile Thr Cys Ala Val  
1125 1130 1135

AAT CCT TGG AAA AAT ATC ATC CGT CCG GTT GTT TAT ATG TCC CGC TTA 3456  
Asn Pro Trp Lys Asn Ile Ile Arg Pro Val Val Tyr Met Ser Arg Leu  
1140 1145 1150

TAT CTG CTA TGG CTG GAG CAG CAA TCA AAG AAA AGT GAT GAT GGT AAA 3504  
Tyr Leu Leu Trp Leu Glu Gln Gln Ser Lys Lys Ser Asp Asp Gly Lys  
1155 1160 1165

ACC ACG ATT TAT CAA TAT AAC TTA AAA CTG GCT CAT ATT CGT TAC GAC 3552  
Thr Thr Ile Tyr Gln Tyr Asn Leu Lys Leu Ala His Ile Arg Tyr Asp  
1170 1175 1180

GGT AGT TGG AAT ACA CCA TTT ACT TTT GAT GTG ACA GAA AAG GTA AAA 3600  
Gly Ser Trp Asn Thr Pro Phe Thr Phe Asp Val Thr Glu Lys Val Lys  
1185 1190 1195 1200

AAT TAC ACG TCG AGT ACT GAT GCT GCT GAA TCT TTA GGG TTG TAT TGT 3648  
Asn Tyr Thr Ser Ser Thr Asp Ala Ala Glu Ser Leu Gly Leu Tyr Cys  
1205 1210 1215

ACT GGT TAT CAA GGG GAA GAC ACT CTA TTA GTT ATG TTC TAT TCG ATG 3696  
Thr Gly Tyr Gln Gly Glu Asp Thr Leu Leu Val Met Phe Tyr Ser Met  
1220 1225 1230

CAG AGT AGT TAT AGC TCC TAT ACC GAT AAT AAT GCG CCG GTC ACT GGG 3744  
Gln Ser Ser Tyr Ser Ser Tyr Thr Asp Asn Asn Ala Pro Val Thr Gly  
1235 1240 1245

CTA TAT ATT TTC GCT GAT ATG TCA TCA GAC AAT ATG ACG AAT GCA CAA 3792  
Leu Tyr Ile Phe Ala Asp Met Ser Ser Asp Asn Met Thr Asn Ala Gln  
1250 1255 1260

GCA ACT AAC TAT TGG AAT AAC AGT TAT CCG CAA TTT GAT ACT GTG ATG 3840  
Ala Thr Asn Tyr Trp Asn Asn Ser Tyr Pro Gln Phe Asp Thr Val Met  
1265 1270 1275 1280

GCA GAT CCG GAT AGC GAC AAT AAA AAA GTC ATA ACC AGA AGA GTT AAT 3888  
Ala Asp Pro Asp Ser Asp Asn Lys Lys Val Ile Thr Arg Arg Val Asn  
1285 1290 1295

AAC CGT TAT GCG GAG GAT TAT GAA ATT CCT TCC TCT GTG ACA AGT AAC 3936  
Asn Arg Tyr Ala Glu Asp Tyr Glu Ile Pro Ser Ser Val Thr Ser Asn  
1300 1305 1310

AGT AAT TAT TCT TGG GGT GAT CAC AGT TTA ACC ATG CTT TAT GGT GGT 3984  
Ser Asn Tyr Ser Trp Gly Asp His Ser Leu Thr Met Leu Tyr Gly Gly  
1315 1320 1325

AGT GTT CCT AAT ATT ACT TTT GAA TCG GCG GCA GAA GAT TTA AGG CTA 4032  
Ser Val Pro Asn Ile Thr Phe Glu Ser Ala Ala Glu Asp Leu Arg Leu  
1330 1335 1340

TCT ACC AAT ATG GCA TTG AGT ATT ATT CAT AAT GGA TAT GCG GGA ACC 4080  
Ser Thr Asn Met Ala Leu Ser Ile Ile His Asn Gly Tyr Ala Gly Thr  
1345 1350 1355 1360

CGC CGT ATA CAA TGT AAT CTT ATG AAA CAA TAC GCT TCA TTA GGT GAT 4128  
Arg Arg Ile Gln Cys Asn Leu Met Lys Gln Tyr Ala Ser Leu Gly Asp  
1365 1370 1375

AAA TTT ATA ATT TAT GAT TCA TCA TTT GAT GAT GCA AAC CGT TTT AAT 4176  
Lys Phe Ile Ile Tyr Asp Ser Ser Phe Asp Asp Ala Asn Arg Phe Asn  
1380 1385 1390

CTG GTG CCA TTG TTT AAA TTC GGA AAA GAC GAG AAC TCA GAT GAT AGT 4224  
Leu Val Pro Leu Phe Lys Phe Gly Lys Asp Glu Asn Ser Asp Asp Ser  
1395 1400 1405

ATT TGT ATA TAT AAT GAA AAC CCT TCC TCT GAA GAT AAG AAG TGG TAT 4272

Ile Cys Ile Tyr Asn Glu Asn Pro Ser Ser Glu Asp Lys Lys Trp Tyr  
1410 1415 1420

TTT TCT TCG AAA GAT GAC AAT AAA ACA GCG GAT TAT AAT GGT GGA ACT 4320  
Phe Ser Ser Lys Asp Asp Asn Lys Thr Ala Asp Tyr Asn Gly Gly Thr  
1425 1430 1435 1440

CAA TGT ATA GAT GCT GGA ACC AGT AAC AAA GAT TTT TAT TAT AAT CTC 4368  
Gln Cys Ile Asp Ala Gly Thr Ser Asn Lys Asp Phe Tyr Tyr Asn Leu  
1445 1450 1455

CAG GAG ATT GAA GTA ATT AGT GTT ACT GGT GGG TAT TGG TCG AGT TAT 4416  
Gln Glu Ile Glu Val Ile Ser Val Thr Gly Gly Tyr Trp Ser Ser Tyr  
1460 1465 1470

AAA ATA TCC AAC CCG ATT AAT ATC AAT ACG GGC ATT GAT AGT GCT AAA 4464  
Lys Ile Ser Asn Pro Ile Asn Ile Asn Thr Gly Ile Asp Ser Ala Lys  
1475 1480 1485

GTA AAA GTC ACC GTA AAA GCG GGT GGT GAC GAT CAA ATC TTT ACT GCT 4512  
Val Lys Val Thr Val Lys Ala Gly Gly Asp Asp Gln Ile Phe Thr Ala  
1490 1495 1500

GAT AAT AGT ACC TAT GTT CCT CAG CAA CCG GCA CCC AGT TTT GAG GAG 4560  
Asp Asn Ser Thr Tyr Val Pro Gln Gln Pro Ala Pro Ser Phe Glu Glu  
1505 1510 1515 1520



ATG ATT TAT CAG TTC AAT AAC CTG ACA ATA GAT TGT AAG AAT TTA AAT 4608  
Met Ile Tyr Gln Phe Asn Asn Leu Thr Ile Asp Cys Lys Asn Leu Asn  
1525 1530 1535

TTC ATC GAC AAT CAG GCA CAT ATT GAG ATT GAT TTC ACC GCT ACG GCA 4656  
Phe Ile Asp Asn Gln Ala His Ile Glu Ile Asp Phe Thr Ala Thr Ala  
1540 1545 1550

CAA GAT GGC CGA TTC TTG GGT GCA GAA ACT TTT ATT ATC CCG GTA ACT 4704  
Gln Asp Gly Arg Phe Leu Gly Ala Glu Thr Phe Ile Ile Pro Val Thr  
1555 1560 1565

AAA AAA GTT CTC GGT ACT GAG AAC GTG ATT GCG TTA TAT AGC GAA AAT 4752  
Lys Lys Val Leu Leu Gly Thr Glu Asn Val Ile Ala Leu Tyr Ser Glu Asn  
1570 1575 1580

AAC GGT GTT CAA TAT ATG CAA ATT GGC GCA TAT CGT ACC CGT TTG AAT 4800  
Asn Gly Val Gln Tyr Met Gln Ile Gly Ala Tyr Arg Thr Arg Leu Asn  
1585 1590 1595 1600

ACG TTA TTC GCT CAA CAG TTG GTT AGC CGT GCT AAT CGT GGC ATT GAT 4848  
Thr Leu Phe Ala Gln Gln Leu Val Ser Arg Ala Asn Arg Gly Ile Asp  
1605 1610 1615

GCA GTG CTC AGT ATG GAA ACT CAG AAT ATT CAG GAA CCG CAA TTA GGA 4896  
Ala Val Leu Ser Met Glu Thr Gln Asn Ile Gln Glu Pro Gln Leu Gly  
1620 1625 1630

GCG GGC ACA TAT GTG CAG CTT GTG TTG GAT AAA TAT GAT GAG TCT ATT 4944  
Ala Gly Thr Tyr Val Gln Leu Val Leu Asp Lys Tyr Asp Glu Ser Ile  
1635 1640 1645

CAT GGC ACT AAT AAA AGC TTT GCT ATT GAA TAT GTT GAT ATA TTT AAA 4992  
His Gly Thr Asn Lys Ser Phe Ala Ile Glu Tyr Val Asp Ile Phe Lys  
1650 1655 1660

GAG AAC GAT AGT TTT GTG ATT TAT CAA GGA GAA CTT AGC GAA ACA ACT 5040  
Glu Asn Asp Ser Phe Val Ile Tyr Gln Gly Glu Leu Ser Glu Thr Ser  
1665 1670 1675 1680

CAA ACT GTT GTG AAA GTT TTC TTA TCC TAT TTT ATA GAG GCG ACT GGA 5088  
Gln Thr Val Val Lys Val Phe Leu Ser Tyr Phe Ile Glu Ala Thr Gly  
1685 1690 1695

AAT AAG AAC CAC TTA TGG GTA CGT GCT AAA TAC CAA AAG GAA ACG ACT 5136  
Asn Lys Asn His Leu Trp Val Arg Ala Lys Tyr Gln Lys Glu Thr Thr  
1700 1705 1710

GAT AAG ATC TTG TTC GAC CGT ACT GAT GAG AAA GAT CCG CAC GGT TGG 5184  
Asp Lys Ile Leu Phe Asp Arg Thr Asp Glu Lys Asp Pro His Gly Trp  
1715 1720 1725

TTT CTC AGC GAC GAT CAC AAG ACC TTT AGT GGT CTC TCT TCC GCA CAG 5232  
Phe Leu Ser Asp Asp His Lys Thr Phe Ser Gly Leu Ser Ser Ala Gln  
1730 1735 1740

GCA TTA AAG AAC GAC AGT GAA CCG ATG GAT TTC TCT GGC GCC AAT GCT 5280  
Ala Leu Lys Asn Asp Ser Glu Pro Met Asp Phe Ser Gly Ala Asn Ala  
1745 1750 1755 1760

CTC TAT TTC TGG GAA CTG TTC TAT TAC ACG CCG ATG ATG ATG GCT CAT 5328  
Leu Tyr Phe Trp Glu Leu Phe Tyr Tyr Thr Pro Met Met Met Ala His  
1765 1770 1775

CGT TTG TTG CAG GAA CAG AAT TTT GAT GCG GCG AAC CAT TGG TTC CGT 5376  
Arg Leu Leu Gln Glu Gln Asn Phe Asp Ala Ala Asn His Trp Phe Arg  
1780 1785 1790

TAT GTC TGG AGT CCA TCC GGT TAT ATC GTT GAT GGT AAA ATT GCT ATC 5424  
Tyr Val Trp Ser Pro Ser Gly Tyr Ile Val Asp Gly Lys Ile Ala Ile  
1795 1800 1805

TAC CAC TGG AAC GTG CGA CCG CTG GAA GAA GAC ACC AGT TGG AAT GCA 5472  
Tyr His Trp Asn Val Arg Pro Leu Glu Glu Asp Thr Ser Trp Asn Ala  
1810 1815 1820

CAA CAA CTG GAC TCC ACC GAT CCA GAT GCT GTA GCC CAA GAT GAT CCG 5520  
Gln Gln Leu Asp Ser Thr Asp Pro Asp Ala Val Ala Gln Asp Asp Pro  
1825 1830 1835 1840

ATG CAC TAC AAG GTG GCT ACC TTT ATG GCG ACG TTG GAT CTG CTA ATG 5568  
Met His Tyr Lys Val Ala Thr Phe Met Ala Thr Leu Asp Leu Leu Met  
1845 1850 1855

GCC CGT GGT GAT GCT GCT TAC CGC CAG TTA GAG CGT GAT ACG TTG GCT 5616  
Ala Arg Gly Asp Ala Ala Tyr Arg Gln Leu Glu Arg Asp Thr Leu Ala  
1860 1865 1870

GAA GCT AAA ATG TGG TAT ACA CAG GCG CTT AAT CTG TTG GGT GAT GAG 5664  
Glu Ala Lys Met Trp Tyr Thr Gln Ala Leu Asn Leu Leu Gly Asp Glu  
1875 1880 1885

CCA CAA GTG ATG CTG AGT ACG ACT TGG GCT AAT CCA ACA TTG GGT AAT 5712  
Pro Gln Val Met Leu Ser Thr Thr Trp Ala Asn Pro Thr Leu Gly Asn  
1890 1895 1900

GCT GCT TCA AAA ACC ACA CAG CAG GTT CGT CAG CAA GTG CTT ACC CAG 5760  
Ala Ala Ser Lys Thr Thr Gln Gln Val Arg Gln Gln Val Leu Thr Gln  
1905 1910 1915 1920

TTG CGT CTC AAT AGC AGG GTA AAA ACC CCG TTG CTA GGA ACA GCC AAT 5808  
Leu Arg Leu Asn Ser Arg Val Lys Thr Pro Leu Leu Gly Thr Ala Asn  
1925 1930 1935

TCC CTG ACC GCT TTA TTC CTG CCG CAG GAA AAT AGC AAG CTC AAA GGC 5856  
Ser Leu Thr Ala Leu Phe Leu Pro Gln Glu Asn Ser Lys Leu Lys Gly  
1940 1945 1950

TAC TGG CGG ACA CTG GCG CAG CGT ATG TTT AAT TTA CGT CAT AAT CTG 5904  
Tyr Trp Arg Thr Leu Ala Gln Arg Met Phe Asn Leu Arg His Asn Leu  
1955 1960 1965

TCG ATT GAC GGC CAG CCG CTC TCC TTG CCG CTG TAT GCT AAA CCG GCT 5952  
Ser Ile Asp Gly Gln Pro Leu Ser Leu Pro Leu Tyr Ala Lys Pro Ala  
1970 1975 1980

GAT CCA AAA GCT TTA CTG AGT GCG GCG GTT TCA GCT TCT CAA GGG GGA 6000  
Asp Pro Lys Ala Leu Leu Ser Ala Ala Val Ser Ala Ser Gln Gly Gly  
1985 1990 1995 2000

GCC GAC TTG CCG AAG GCG CCG CTG ACT ATT CAC CGC TTC CCT CAA ATG 6048  
Ala Asp Leu Pro Lys Ala Pro Leu Thr Ile His Arg Phe Pro Gln Met  
2005 2010 2015

CTA GAA GGG GCA CGG GGC TTG GTT AAC CAG CTT ATA CAG TTC GGT AGT 6096  
Leu Glu Gly Ala Arg Gly Leu Val Asn Gln Leu Ile Gln Phe Gly Ser  
2020 2025 2030

TCA CTA TTG GGG TAC AGT GAG CGT CAG GAT GCG GAA GCT ATG AGT CAA 6144  
Ser Leu Leu Gly Tyr Ser Glu Arg Gln Asp Ala Glu Ala Met Ser Gln  
2035 2040 2045

CTA CTG CAA ACC CAA GCC AGC GAG TTA ATA CTG ACC AGT ATT CGT ATG 6192  
Leu Leu Gln Thr Gln Ala Ser Glu Leu Ile Leu Thr Ser Ile Arg Met  
2050 2055 2060

CAG GAT AAC CAA TTG GCA GAG CTG GAT TCG GAA AAA ACC GCC TTG CAA 6240  
Gln Asp Asn Gln Leu Ala Glu Leu Asp Ser Glu Lys Thr Ala Leu Gln  
2065 2070 2075 2080

GTC TCT TTA GCT GGA GTG CAA CAA CGG TTT GAC AGC TAT AGC CAA CTG 6288  
Val Ser Leu Ala Gly Val Gln Gln Arg Phe Asp Ser Tyr Ser Gln Leu  
2085 2090 2095

TAT GAG GAG AAC ATC AAC GCA GGT GAG CAG CGA GCG CTG GCG TTA CGC 6336  
Tyr Glu Glu Asn Ile Asn Ala Gly Glu Gln Arg Ala Leu Ala Leu Arg  
2100 2105 2110

TCA GAA TCT GCT ATT GAG TCT CAG GGA GCG CAG ATT TCC CGT ATG GCA 6384  
Ser Glu Ser Ala Ile Glu Ser Gln Gly Ala Gln Ile Ser Arg Met Ala  
2115 2120 2125

GGC GCG GGT GTT GAT ATG GCA CCA AAT ATC TTC GGC CTG GCT GAT GGC 6432  
Gly Ala Gly Val Asp Met Ala Pro Asn Ile Phe Gly Leu Ala Asp Gly  
2130 2135 2140

GGC ATG CAT TAT GGT GCT ATT GCC TAT GCC ATC GCT GAC GGT ATT GAG 6480  
Gly Met His Tyr Gly Ala Ile Ala Tyr Ala Ile Ala Asp Gly Ile Glu  
2145 2150 2155 2160

TTG AGT GCT TCT GCC AAG ATG GTT GAT GCG GAG AAA GTT GCT CAG TCG 6528  
Leu Ser Ala Ser Ala Lys Met Val Asp Ala Glu Lys Val Ala Gln Ser  
2165 2170 2175

GAA ATA TAT CGC CGT CGC CGT CAA GAA TGG AAA ATT CAG CGT GAC AAC 6576  
Glu Ile Tyr Arg Arg Arg Arg Gln Glu Trp Lys Ile Gln Arg Asp Asn  
2180 2185 2190

GCA CAA GCG GAG ATT AAC CAG TTA AAC GCG CAA CTG GAA TCA CTG TCT 6624  
Ala Gln Ala Glu Ile Asn Gln Leu Asn Ala Gln Leu Glu Ser Leu Ser  
2195 2200 2205

ATT CGC CGT GAA GCC GCT GAA ATG CAA AAA GAG TAC CTG AAA ACC CAG 6672  
Ile Arg Arg Glu Ala Ala Glu Met Gln Lys Glu Tyr Leu Lys Thr Gln  
2210 2215 2220

CAA GCT CAG GCG CAG GCA CAA CTT ACT TTC TTA AGA AGC AAA TTC AGT 6720  
Gln Ala Gln Ala Gln Ala Gln Leu Thr Phe Leu Arg Ser Lys Phe Ser  
2225 2230 2235 2240

AAT CAA GCG TTA TAT AGT TGG TTA CGA GGG CGT TTG TCA GGT ATT TAT 6768  
Asn Gln Ala Leu Tyr Ser Trp Leu Arg Gly Arg Leu Ser Gly Ile Tyr  
2245 2250 2255

TTC CAG TTC TAT GAC TTG GCC GTA TCA CGT TGC CTG ATG GCA GAG CAA 6816  
Phe Gln Phe Tyr Asp Leu Ala Val Ser Arg Cys Leu Met Ala Glu Gln  
2260 2265 2270

TCC TAT CAA TGG GAA GCT AAT GAT AAT TCC ATT AGC TTT GTC AAA CCG 6864  
Ser Tyr Gln Trp Glu Ala Asn Asp Asn Ser Ile Ser Phe Val Lys Pro  
2275 2280 2285

GGT GCA TGG CAA GGA ACT TAC GCC GGC TTA TTG TGT GGA GAA GCT TTG 6912  
Gly Ala Trp Gln Gly Thr Tyr Ala Gly Leu Leu Cys Gly Glu Ala Leu  
2290 2295 2300

ATA CAA AAT CTG GCA CAA ATG GAA GAG GCA TAT CTG AAA TGG GAA TCT 6960  
Ile Gln Asn Leu Ala Gln Met Glu Glu Ala Tyr Leu Lys Trp Glu Ser  
2305 2310 2315 2320

CGC GCT TTG GAA GTA GAA CGC ACG GTT TCA TTG GCA GTG GTT TAT GAT 7008  
Arg Ala Leu Glu Val Glu Arg Thr Val Ser Leu Ala Val Val Tyr Asp  
2325 2330 2335

TCA CTG GAA GGT AAT GAT CGT TTT AAT TTA GCG GAA CAA ATA CCT GCA 7056  
Ser Leu Glu Gly Asn Asp Arg Phe Asn Leu Ala Glu Gln Ile Pro Ala  
2340 2345 2350

TTA TTG GAT AAG GGG GAG GGA ACA GCA GGA ACT AAA GAA AAT GGG TTA 7104  
Leu Leu Asp Lys Gly Glu Gly Thr Ala Gly Thr Lys Glu Asn Gly Leu  
2355 2360 2365

TCA TTG GCT AAT GCT ATC CTG TCA GCT TCG GTC AAA TTG TCC GAC TTG 7152  
Ser Leu Ala Asn Ala Ile Leu Ser Ala Ser Val Lys Leu Ser Asp Leu  
2370 2375 2380

AAA CTG GGA ACG GAT TAT CCA GAC AGT ATC GTT GGT AGC AAC AAG GTT 7200  
Lys Leu Gly Thr Asp Tyr Pro Asp Ser Ile Val Gly Ser Asn Lys Val  
2385 2390 2395 2400

CGT CGT ATT AAG CAA ATC AGT GTT TCG CTA CCT GCA TTG GTT GGG CCT 7248  
Arg Arg Ile Lys Gln Ile Ser Val Ser Leu Pro Ala Leu Val Gly Pro  
2405 2410 2415

TAT CAG GAT GTT CAG GCT ATG CTC AGC TAT GGT GGC AGT ACT CAA TTG 7296  
Tyr Gln Asp Val Gln Ala Met Leu Ser Tyr Gly Gly Ser Thr Gln Leu  
2420 2425 2430

CCG AAA GGT TGT TCA GCG TTG GCT GTG TCT CAT GGT ACC AAT GAT AGT 7344  
Pro Lys Gly Cys Ser Ala Leu Ala Val Ser His Gly Thr Asn Asp Ser  
2435 2440 2445

GGT CAG TTC CAG TTG GAT TTC AAT GAC GGC AAA TAC CTG CCA TTT GAA 7392  
Gly Gln Phe Gln Leu Asp Phe Asn Asp Gly Lys Tyr Leu Pro Phe Glu  
2450 2455 2460

GGT ATT GCT CTT GAT GAT CAG GGT ACA CTG AAT CTT CAA TTT CCG AAT 7440  
Gly Ile Ala Leu Asp Asp Gln Gly Thr Leu Asn Leu Gln Phe Pro Asn  
2465 2470 2475 2480

GCT ACC GAC AAG CAG AAA GCA ATA TTG CAA ACT ATG AGC GAT ATT ATT 7488  
Ala Thr Asp Lys Gln Lys Ala Ile Leu Gln Thr Met Ser Asp Ile Ile  
2485 2490 2495

TTG CAT ATT CGT TAT ACC ATC CGT TAA 7515  
Leu His Ile Arg Tyr Thr Ile Arg \*  
2500 2505

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 12:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĹŽKA: 2504 aminokyselín
- (B) TYP: aminokyselina
- (D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: proteín

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 12 (TcBA proteín):

Met Gln Asn Ser Leu Ser Ser Thr Ile Asp Thr Ile Cys Gln Lys Leu  
1 5 10 15  
Gln Leu Thr Cys Pro Ala Glu Ile Ala Leu Tyr Pro Phe Asp Thr Phe  
20 25 30  
Arg Glu Lys Thr Arg Gly Met Val Asn Trp Gly Glu Ala Lys Arg Ile  
35 40 45  
Tyr Glu Ile Ala Gln Ala Glu Gln Asp Arg Asn Leu Leu His Glu Lys  
50 55 60  
Arg Ile Phe Ala Tyr Ala Asn Pro Leu Leu Lys Asn Ala Val Arg Leu  
65 70 75 80  
Gly Thr Arg Gln Met Leu Gly Phe Ile Gln Gly Tyr Ser Asp Leu Phe  
85 90 95  
Gly Asn Arg Ala Asp Asn Tyr Ala Ala Pro Gly Ser Val Ala Ser Met  
100 105 110  
Phe Ser Pro Ala Ala Tyr Leu Thr Glu Leu Tyr Arg Glu Ala Lys Asn  
115 120 125  
Leu His Asp Ser Ser Ser Ile Tyr Tyr Leu Asp Lys Arg Arg Pro Asp  
130 135 140  
Leu Ala Ser Leu Met Leu Ser Gln Lys Asn Met Asp Glu Glu Ile Ser  
145 150 155 160  
Thr Leu Ala Leu Ser Asn Glu Leu Cys Leu Ala Gly Ile Glu Thr Lys  
165 170 175  
Thr Gly Lys Ser Gln Asp Glu Val Met Asp Met Leu Ser Thr Tyr Arg  
180 185 190  
Leu ~~Ser Gly~~ Glu Thr Pro Tyr His His Ala Tyr Glu Thr Val Arg Glu  
195 200 205  
Ile Val His Glu Arg Asp Pro Gly Phe Arg His Leu Ser Gln Ala Pro  
210 215 220  
Ile Val Ala Ala Lys Leu Asp Pro Val Thr Leu Leu Gly Ile Ser Ser  
225 230 235 240  
His Ile Ser Pro Glu Leu Tyr Asn Leu Leu Ile Glu Glu Ile Pro Glu  
245 250 255  
Lys Asp Glu Ala Ala Leu Asp Thr Leu Tyr Lys Thr Asn Phe Gly Asp  
260 265 270  
Ile Thr Thr Ala Gln Leu Met Ser Pro Ser Tyr Leu Ala Arg Tyr Tyr  
275 280 285  
Gly Val Ser Pro Glu Asp Ile Ala Tyr Val Thr Thr Ser Leu Ser His  
290 295 300  
Val Gly Tyr Ser Ser Asp Ile Leu Val Ile Pro Leu Val Asp Gly Val  
305 310 315 320  
Gly Lys Met Glu Val Val Arg Val Thr Arg Thr Pro Ser Asp Asn Tyr  
325 330 335  
Thr Ser Gln Thr Asn Tyr Ile Glu Leu Tyr Pro Gln Gly Gly Asp Asn  
340 345 350

Tyr Leu Ile Lys Tyr Asn Leu Ser Asn Ser Phe Gly Leu Asp Asp Phe  
355 360 365

Tyr Leu Gln Tyr Lys Asp Gly Ser Ala Asp Trp Thr Glu Ile Ala His  
370 375 380

Asn Pro Tyr Pro Asp Met Val Ile Asn Gln Lys Tyr Glu Ser Gln Ala  
385 390 395 400

Thr Ile Lys Arg Ser Asp Ser Asp Asn Ile Leu Ser Ile Gly Leu Gln  
405 410 415

Arg Trp His Ser Gly Ser Tyr Asn Phe Ala Ala Ala Asn Phe Lys Ile  
420 425 430

Asp Gln Tyr Ser Pro Lys Ala Phe Leu Leu Lys Met Asn Lys Ala Ile  
435 440 445

Arg Leu Leu Lys Ala Thr Gly Leu Ser Phe Ala Thr Leu Glu Arg Ile  
450 455 460

Val Asp Ser Val Asn Ser Thr Lys Ser Ile Thr Val Glu Val Leu Asn  
465 470 475 480

Lys Val Tyr Arg Val Lys Phe Tyr Ile Asp Arg Tyr Gly Ile Ser Glu  
485 490 495

Glu Thr Ala Ala Ile Leu Ala Asn Ile Asn Ile Ser Gln Gln Ala Val  
500 505 510

Gly Asn Gln Leu Ser Gln Phe Glu Gln Leu Phe Asn His Pro Pro Leu  
515 520 525

Asn Gly Ile Arg Tyr Glu Ile Ser Glu Asp Asn Ser Lys His Leu Pro  
530 535 540

Asn Pro Asp Leu Asn Leu Lys Pro Asp Ser Thr Gly Asp Asp Gln Arg  
545 550 555 560

Lys Ala Val Leu Lys Arg Ala Phe Gln Val Asn Ala Ser Glu Leu Tyr  
565 570 575

Gln Met Leu Leu Ile Thr Asp Arg Lys Glu Asp Gly Val Ile Lys Asn  
580 585 590

Asn Leu Glu Asn Leu Ser Asp Leu Tyr Leu Val Ser Leu Leu Ala Gln  
595 600 605

Ile His Asn Leu Thr Ile Ala Glu Leu Asn Ile Leu Leu Val Ile Cys  
610 615 620

Gly Tyr Gly Asp Thr Asn Ile Tyr Gln Ile Thr Asp Asp Asn Leu Ala  
625 630 635 640

Lys Ile Val Glu Thr Leu Leu Trp Ile Thr Gln Trp Leu Lys Thr Gln  
645 650 655

Lys Trp Thr Val Thr Asp Leu Phe Leu Met Thr Thr Ala Thr Tyr Ser  
660 665 670

Thr Thr Leu Thr Pro Glu Ile Ser Asn Leu Thr Ala Thr Leu Ser Ser  
675 680 685

Thr Leu His Gly Lys Glu Ser Leu Ile Gly Glu Asp Leu Lys Arg Ala  
690 695 700

Met Ala Pro Cys Phe Thr Ser Ala Leu His Leu Thr Ser Gln Glu Val  
705 710 715 720

Ala Tyr Asp Leu Leu Leu Trp Ile Asp Gln Ile Gln Pro Ala Gln Ile  
725 730 735

Thr Val Asp Gly Phe Trp Glu Glu Val Gln Thr Thr Pro Thr Ser Leu  
740 745 750

Lys Val Ile Thr Phe Ala Gln Val Leu Ala Gln Leu Ser Leu Ile Tyr -  
755 760 765

Arg Arg Ile Gly Leu Ser Glu Thr Glu Leu Ser Leu Ile Val Thr Gln  
770 775 780

Ser Ser Leu Leu Val Ala Gly Lys Ser Ile Leu Asp His Gly Leu Leu  
785 790 795 800

Thr Leu Met Ala Leu Glu Gly Phe His Thr Trp Val Asn Gly Leu Gly  
805 810 815

Gln His Ala Ser Leu Ile Leu Ala Ala Leu Lys Asp Gly Ala Leu Thr  
820 825 830

Val Thr Asp Val Ala Gln Ala Met Asn Lys Glu Glu Ser Leu Leu Gln  
835 840 845

Met Ala Ala Asn Gln Val Glu Lys Asp Leu Thr Lys Leu Thr Ser Trp  
850 855 860

Thr Gln Ile Asp Ala Ile Leu Gln Trp Leu Gln Met Ser Ser Ala Leu  
865 870 875 880

Ala Val Ser Pro Leu Asp Leu Ala Gly Met Met Ala Leu Lys Tyr Gly  
885 890 895

Ile Asp His Asn Tyr Ala Ala Trp Gln Ala Ala Ala Ala Ala Leu Met  
900 905 910

Ala Asp His Ala Asn Gln Ala Gln Lys Lys Leu Asp Glu Thr Phe Ser  
915 920 925

Lys Ala Leu Cys Asn Tyr Tyr Ile Asn Ala Val Val Asp Ser Ala Ala  
930 935 940

Gly Val Arg Asp Arg Asn Gly Leu Tyr Thr Tyr Leu Leu Ile Asp Asn  
945 950 955 960

Gln Val Ser Ala Asp Val Ile Thr Ser Arg Ile Ala Glu Ala Ile Ala  
965 970 975

Gly Ile Gln Leu Tyr Val Asn Arg Ala Leu Asn Arg Asp Glu Gly Gln  
980 985 990

Leu Ala Ser Asp Val Ser Thr Arg Gln Phe Phe Thr Asp Trp Glu Arg  
995 1000 1005

Tyr Asn Lys Arg Tyr Ser Thr Trp Ala Gly Val Ser Glu Leu Val Tyr  
1010 1015 1020

Tyr Pro Glu Asn Tyr Val Asp Pro Thr Gln Arg Ile Gly Gln Thr Lys  
1025 1030 1035 1040

Met Met Asp Ala Leu Leu Gln Ser Ile Asn Gln Ser Gln Leu Asn Ala  
1045 1050 1055

Asp Thr Val Glu Asp Ala Phe Lys Thr Tyr Leu Thr Ser Phe Glu Gln  
1060 1065 1070

Val Ala Asn Leu Lys Val Ile Ser Ala Tyr His Asp Asn Val Asn Val

1075				1080				1085							
Asp	Gln	Gly	Leu	Thr	Tyr	Phe	Ile	Gly	Ile	Asp	Gln	Ala	Ala	Pro	Gly
1090						1095					1100				
Thr	Tyr	Tyr	Trp	Arg	Ser	Val	Asp	His	Ser	Lys	Cys	Glu	Asn	Gly	Lys
1105					1110					1115					1120
Phe	Ala	Ala	Asn	Ala	Trp	Gly	Glu	Trp	Asn	Lys	Ile	Thr	Cys	Ala	Val
				1125					1130						1135
Asn	Pro	Trp	Lys	Asn	Ile	Ile	Arg	Pro	Val	Val	Tyr	Met	Ser	Arg	Leu
			1140					1145						1150	
Tyr	Leu	Leu	Trp	Leu	Glu	Gln	Gln	Ser	Lys	Lys	Ser	Asp	Asp	Gly	Lys
		1155					1160						1165		
Thr	Thr	Ile	Tyr	Gln	Tyr	Asn	Leu	Lys	Leu	Ala	His	Ile	Arg	Tyr	Asp
	1170					1175						1180			
Gly	Ser	Trp	Asn	Thr	Pro	Phe	Thr	Phe	Asp	Val	Thr	Glu	Lys	Val	Lys
1185					1190					1195					1200
Asn	Tyr	Thr	Ser	Ser	Thr	Asp	Ala	Ala	Glu	Ser	Leu	Gly	Leu	Tyr	Cys
				1205					1210						1215
Thr	Gly	Tyr	Gln	Gly	Glu	Asp	Thr	Leu	Leu	Val	Met	Phe	Tyr	Ser	Met
			1220						1225					1230	
Gln	Ser	Ser	Tyr	Ser	Ser	Tyr	Thr	Asp	Asn	Asn	Ala	Pro	Val	Thr	Gly
		1235					1240						1245		
Leu	Tyr	Ile	Phe	Ala	Asp	Met	Ser	Ser	Asp	Asn	Met	Thr	Asn	Ala	Gln
	1250					1255						1260			
Ala	Thr	Asn	Tyr	Trp	Asn	Asn	Ser	Tyr	Pro	Gln	Phe	Asp	Thr	Val	Met
1265					1270					1275					1280
Ala	Asp	Pro	Asp	Ser	Asp	Asn	Lys	Lys	Val	Ile	Thr	Arg	Arg	Val	Asn
				1285					1290						1295
Asn	Arg	Tyr	Ala	Glu	Asp	Tyr	Glu	Ile	Pro	Ser	Ser	Val	Thr	Ser	Asn
			1300						1305						1310
Ser	Asn	Tyr	Ser	Trp	Gly	Asp	His	Ser	Leu	Thr	Met	Leu	Tyr	Gly	Gly
		1315					1320						1325		
Ser	Val	Pro	Asn	Ile	Thr	Phe	Glu	Ser	Ala	Ala	Glu	Asp	Leu	Arg	Leu
	1330					1335						1340			
Ser	Thr	Asn	Met	Ala	Leu	Ser	Ile	Ile	His	Asn	Gly	Tyr	Ala	Gly	Thr
1345					1350					1355					1360
Arg	Arg	Ile	Gln	Cys	Asn	Leu	Met	Lys	Gln	Tyr	Ala	Ser	Leu	Gly	Asp
				1365					1370						1375
Lys	Phe	Ile	Ile	Tyr	Asp	Ser	Ser	Phe	Asp	Asp	Ala	Asn	Arg	Phe	Asn
			1380						1385					1390	
Leu	Val	Pro	Leu	Phe	Lys	Phe	Gly	Lys	Asp	Glu	Asn	Ser	Asp	Asp	Ser
		1395					1400						1405		
Ile	Cys	Ile	Tyr	Asn	Glu	Asn	Pro	Ser	Ser	Glu	Asp	Lys	Lys	Trp	Tyr
	1410					1415							1420		
Phe	Ser	Ser	Lys	Asp	Asp	Asn	Lys	Thr	Ala	Asp	Tyr	Asn	Gly	Gly	Thr
1425					1430					1435					1440
Gln	Cys	Ile	Asp	Ala	Gly	Thr	Ser	Asn	Lys	Asp	Phe	Tyr	Tyr	Asn	Leu
				1445					1450						1455



Gln Glu Ile Glu Val Ile Ser Val Thr Gly Gly Tyr Trp Ser Ser Tyr  
1460 1465 1470

Lys Ile Ser Asn Pro Ile Asn Ile Asn Thr Gly Ile Asp Ser Ala Lys  
1475 1480 1485

Val Lys Val Thr Val Lys Ala Gly Gly Asp Asp Gln Ile Phe Thr Ala  
1490 1495 1500

Asp Asn Ser Thr Tyr Val Pro Gln Gln Pro Ala Pro Ser Phe Glu Glu  
1505 1510 1515 1520

Met Ile Tyr Gln Phe Asn Asn Leu Thr Ile Asp Cys Lys Asn Leu Asn  
1525 1530 1535

Phe Ile Asp Asn Gln Ala His Ile Glu Ile Asp Phe Thr Ala Thr Ala  
1540 1545 1550

Gln Asp Gly Arg Phe Leu Gly Ala Glu Thr Phe Ile Ile Pro Val Thr  
1555 1560 1565

Lys Lys Val Leu Gly Thr Glu Asn Val Ile Ala Leu Tyr Ser Glu Asn  
1570 1575 1580

Asn Gly Val Gln Tyr Met Gln Ile Gly Ala Tyr Arg Thr Arg Leu Asn  
1585 1590 1595 1600

Thr Leu Phe Ala Gln Gln Leu Val Ser Arg Ala Asn Arg Gly Ile Asp  
1605 1610 1615

Ala Val Leu Ser Met Glu Thr Gln Asn Ile Gln Glu Pro Gln Leu Gly  
1620 1625 1630

Ala Gly Thr Tyr Val Gln Leu Val Leu Asp Lys Tyr Asp Glu Ser Ile  
1635 1640 1645

His Gly Thr Asn Lys Ser Phe Ala Ile Glu Tyr Val Asp Ile Phe Lys  
1650 1655 1660

Glu Asn Asp Ser Phe Val Ile Tyr Gln Gly Glu Leu Ser Glu Thr Ser  
1665 1670 1675 1680

Gln Thr Val Val Lys Val Phe Leu Ser Tyr Phe Ile Glu Ala Thr Gly  
1685 1690 1695

Asn Lys Asn His Leu Trp Val Arg Ala Lys Tyr Gln Lys Glu Thr Thr  
1700 1705 1710

Asp Lys Ile Leu Phe Asp Arg Thr Asp Glu Lys Asp Pro His Gly Trp  
1715 1720 1725

Phe Leu Ser Asp Asp His Lys Thr Phe Ser Gly Leu Ser Ser Ala Gln  
1730 1735 1740

Ala Leu Lys Asn Asp Ser Glu Pro Met Asp Phe Ser Gly Ala Asn Ala  
1745 1750 1755 1760

Leu Tyr Phe Trp Glu Leu Phe Tyr Tyr Thr Pro Met Met Met Ala His  
1765 1770 1775

Arg Leu Leu Gln Glu Gln Asn Phe Asp Ala Ala Asn His Trp Phe Arg  
1780 1785 1790

Tyr Val Trp Ser Pro Ser Gly Tyr Ile Val Asp Gly Lys Ile Ala Ile  
1795 1800 1805

Tyr His Trp Asn Val Arg Pro Leu Glu Glu Asp Thr Ser Trp Asn Ala  
1810 1815 1820

Gln Gln Leu Asp Ser Thr Asp Pro Asp Ala Val Ala Gln Asp Asp Pro  
1825 1830 1835 1840

Met His Tyr Lys Val Ala Thr Phe Met Ala Thr Leu Asp Leu Leu Met  
1845 1850 1855

Ala Arg Gly Asp Ala Ala Tyr Arg Gln Leu Glu Arg Asp Thr Leu Ala  
1860 1865 1870

Glu Ala Lys Met Trp Tyr Thr Gln Ala Leu Asn Leu Leu Gly Asp Glu  
1875 1880 1885

Pro Gln Val Met Leu Ser Thr Thr Trp Ala Asn Pro Thr Leu Gly Asn  
1890 1895 1900

Ala Ala Ser Lys Thr Thr Gln Gln Val Arg Gln Gln Val Leu Thr Gln  
1905 1910 1915 1920

Leu Arg Leu Asn Ser Arg Val Lys Thr Pro Leu Leu Gly Thr Ala Asn  
1925 1930 1935

Ser Leu Thr Ala Leu Phe Leu Pro Gln Glu Asn Ser Lys Leu Lys Gly  
1940 1945 1950

Tyr Trp Arg Thr Leu Ala Gln Arg Met Phe Asn Leu Arg His Asn Leu  
1955 1960 1965

Ser Ile Asp Gly Gln Pro Leu Ser Leu Pro Leu Tyr Ala Lys Pro Ala  
1970 1975 1980

Asp Pro Lys Ala Leu Leu Ser Ala Ala Val Ser Ala Ser Gln Gly Gly  
1985 1990 1995 2000

Ala Asp Leu Pro Lys Ala Pro Leu Thr Ile His Arg Phe Pro Gln Met  
2005 2010 2015

Leu Glu Gly Ala Arg Gly Leu Val Asn Gln Leu Ile Gln Phe Gly Ser  
2020 2025 2030

Ser Leu Leu Gly Tyr Ser Glu Arg Gln Asp Ala Glu Ala Met Ser Gln  
2035 2040 2045

Leu Leu Gln Thr Gln Ala Ser Glu Leu Ile Leu Thr Ser Ile Arg Met  
2050 2055 2060

Gln Asp Asn Gln Leu Ala Glu Leu Asp Ser Glu Lys Thr Ala Leu Gln  
2065 2070 2075 2080

Val Ser Leu Ala Gly Val Gln Gln Arg Phe Asp Ser Tyr Ser Gln Leu  
2085 2090 2095

Tyr Glu Glu Asn Ile Asn Ala Gly Glu Gln Arg Ala Leu Ala Leu Arg  
2100 2105 2110

Ser Glu Ser Ala Ile Glu Ser Gln Gly Ala Gln Ile Ser Arg Met Ala  
2115 2120 2125

Gly Ala Gly Val Asp Met Ala Pro Asn Ile Phe Gly Leu Ala Asp Gly  
2130 2135 2140

Gly Met His Tyr Gly Ala Ile Ala Tyr Ala Ile Ala Asp Gly Ile Glu  
2145 2150 2155 2160

Leu Ser Ala Ser Ala Lys Met Val Asp Ala Glu Lys Val Ala Gln Ser  
2165 2170 2175

Glu Ile Tyr Arg Arg Arg Arg Gln Glu Trp Lys Ile Gln Arg Asp Asn  
2180 2185 2190

Ala Gln Ala Glu Ile Asn Gln Leu Asn Ala Gln Leu Glu Ser Leu Ser  
2195 2200 2205

Ile Arg Arg Glu Ala Ala Glu Met Gln Lys Glu Tyr Leu Lys Thr Gln

2210	2215	2220
Gln Ala Gln Ala Gln Ala Gln Leu Thr Phe Leu Arg Ser Lys Phe Ser 2225 2230 2235 2240		
Asn Gln Ala Leu Tyr Ser Trp Leu Arg Gly Arg Leu Ser Gly Ile Tyr 2245 2250 2255		
Phe Gln Phe Tyr Asp Leu Ala Val Ser Arg Cys Leu Met Ala Glu Gln 2260 2265 2270		
Ser Tyr Gln Trp Glu Ala Asn Asp Asn Ser Ile Ser Phe Val Lys Pro 2275 2280 2285		
Gly Ala Trp Gln Gly Thr Tyr Ala Gly Leu Leu Cys Gly Glu Ala Leu 2290 2295 2300		
Ile Gln Asn Leu Ala Gln Met Glu Glu Ala Tyr Leu Lys Trp Glu Ser 2305 2310 2315 2320		
Arg Ala Leu Glu Val Glu Arg Thr Val Ser Leu Ala Val Val Tyr Asp 2325 2330 2335		
Ser Leu Glu Gly Asn Asp Arg Phe Asn Leu Ala Glu Gln Ile Pro Ala 2340 2345 2350		
Leu Leu Asp Lys Gly Glu Gly Thr Ala Gly Thr Lys Glu Asn Gly Leu 2355 2360 2365		
Ser Leu Ala Asn Ala Ile Leu Ser Ala Ser Val Lys Leu Ser Asp Leu 2370 2375 2380		
Lys Leu Gly Thr Asp Tyr Pro Asp Ser Ile Val Gly Ser Asn Lys Val 2385 2390 2395 2400		
Arg Arg Ile Lys Gln Ile Ser Val Ser Leu Pro Ala Leu Val Gly Pro 2405 2410 2415		
Tyr Gln Asp Val Gln Ala Met Leu Ser Tyr Gly Gly Ser Thr Gln Leu 2420 2425 2430		
Pro Lys Gly Cys Ser Ala Leu Ala Val Ser His Gly Thr Asn Asp Ser 2435 2440 2445		
Gly Gln Phe Gln Leu Asp Phe Asn Asp Gly Lys Tyr Leu Pro Phe Glu 2450 2455 2460		
Gly Ile Ala Leu Asp Asp Gln Gly Thr Leu Asn Leu Gln Phe Pro Asn 2465 2470 2475 2480		
Ala Thr Asp Lys Gln Lys Ala Ile Leu Gln Thr Met Ser Asp Ile Ile 2485 2490 2495		
Leu His Ile Arg Tyr Thr Ile Arg * 2500 2505		

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 13:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĹŽKA: 12 aminokyselín
- (B) TYP: aminokyselina
- (C) REĹAZEC: jednoduchý
- (D) TOPOLOGIA: lineárna





(ii) TYP MOLEKULY: peptid

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 18 (TcdA<sub>ii</sub> - PT79  
interný peptid):

Phe Ile Val Tyr Thr Ser Leu Gly Val Asn Pro Asn Asn Ser  
1 5 10

Ser Asn  
15

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 19:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĚLKA: 21 aminokyselín
- (B) TYP: aminokyselina
- (C) REĚAZEC: jednoduchý
- (D) TOPOLÓGIA: lineárna

(ii) TYP MOLEKULY: peptid

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 19 (TcaB<sub>i</sub> - PT158  
interný peptid):

Ile Ser Asp Leu Val Thr Thr Ser Pro Leu Ser Glu Ala Ile Gly  
1 5 10 15

Ser Leu Gln Leu Phe Ile  
20

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 20:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĚLKA: 12 aminokyselín
- (B) TYP: aminokyselina
- (C) REĚAZEC: jednoduchý
- (D) TOPOLÓGIA: lineárna







(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 25:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĹŽKA: 6054 párov báz
- (B) TYP: nukleová kyselina
- (C) REĹAZEC: dvojitý
- (D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: DNA (genómová)

(ix) CHARAKTERISTICKÝ ZNAK:

- (A) MENO/ZNAČKA: CDS
- (B) POLOHA: 1..43
- (D) ĎALŠIE INFORMÁCIE: /produkt= "koniec TcaA<sub>iii</sub>"

(ix) CHARAKTERISTICKÝ ZNAK:

- (A) MENO/ZNAČKA: RBS
- (B) POLOHA: 51 58

(ix) CHARAKTERISTICKÝ ZNAK:

- (A) MENO/ZNAČKA: CDS
- (B) POLOHA: 65 3634
- (D) ĎALŠIE INFORMÁCIE: /produkt= "TcaB<sub>i</sub>"

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 25:

```
A GTA GCC CAA AAC TTA AGT GCC GCA ATC AGC AAT CGT CAG TAACCGGATA 50
Val Ala Gln Asn Leu Ser Ala Ala Ile Ser Asn Arg Gln ...

AAGAAGGAAT TGATT ATG TCT GAA TCT TTA TTT ACA CAA ACG TTG AAA GAA 100
Met Ser Glu Ser Leu Phe Thr Gln Thr Leu Lys Glu
1 5 10

GCG CGC CGT GAT GCA TTG GTT GCT CAT TAT ATT GCT ACT CAG GTG CCC 148
Ala Arg Arg Asp Ala Leu Val Ala His Tyr Ile Ala Thr Gln Val Pro
15 20 25

GCA GAT TTA AAA GAG AGT ATC CAG ACC GCG GAT GAT CTG TAC GAA TAT 196
Ala Asp Leu Lys Glu Ser Ile Gln Thr Ala Asp Asp Leu Tyr Glu Tyr
30 35 40

CTG TTG CTG GAT ACC AAA ATT AGC GAT CTG GTT ACT ACT TCA CCG CTG 244
Leu Leu Leu Asp Thr Lys Ile Ser Asp Leu Val Thr Thr Ser Pro Leu
45 50 55 60

TCC GAA GCG ATT GGC AGT CTG CAA TTG TTT ATT CAT CGT GCG ATA GAG 292
Ser Glu Ala Ile Gly Ser Leu Gln Leu Phe Ile His Arg Ala Ile Glu
65 70 75
```

GGC TAT GAC GGC ACG CTG GCA GAC TCA GCA AAA CCC TAT TTT GCC GAT 340  
Gly Tyr Asp Gly Thr Leu Ala Asp Ser Ala Lys Pro Tyr Phe Ala Asp  
80 85 90

GAA CAG TTT TTA TAT AAC TGG GAT AGT TTT AAC CAC CGT TAT AGC ACT 388  
Glu Gln Phe Leu Tyr Asn Trp Asp Ser Phe Asn His Arg Tyr Ser Thr  
95 100 105

TGG GCT GGC AAG GAA CGG TTG AAA TTC TAT GCC GGG GAT TAT ATT GAT 436  
Trp Ala Gly Lys Glu Arg Leu Lys Phe Tyr Ala Gly Asp Tyr Ile Asp  
110 115 120

CCA ACA TTG CGA TTG AAT AAG ACC GAG ATA TTT ACC GCA TTT GAA CAA 484  
Pro Thr Leu Arg Leu Asn Lys Thr Glu Ile Phe Thr Ala Phe Glu Gln  
125 130 135 140

GGT ATT TCT CAA GGG AAA TTA AAA AGT GAA TTA GTC GAA TCT AAA TTA 532  
Gly Ile Ser Gln Gly Lys Leu Lys Ser Glu Leu Val Glu Ser Lys Leu  
145 150 155

CGT GAT TAT CTA ATT AGT TAT GAC ACT TTA GCC ACC CTT GAT TAT ATT 580  
Arg Asp Tyr Leu Ile Ser Tyr Asp Thr Leu Ala Thr Leu Asp Tyr Ile  
160 165 170

ACT GCC TGC CAA GGC AAA GAT AAT AAA ACC ATC TTC TTT ATT GGC CGT 628  
Thr Ala Cys Gln Gly Lys Asp Asn Lys Thr Ile Phe Phe Ile Gly Arg  
175 180 185

ACA CAG AAT GCA CCC TAT GCA TTT TAT TGG CGA AAA TTA ACT TTA GTC 676  
Thr Gln Asn Ala Pro Tyr Ala Phe Tyr Trp Arg Lys Leu Thr Leu Val  
190 195 200

ACT GAT GGC GGT AAG TTG AAA CCA GAT CAA TGG TCA GAG TGG CGA GCA 724  
Thr Asp Gly Gly Lys Leu Lys Pro Asp Gln Trp Ser Glu Trp Arg Ala  
205 210 215 220

ATT AAT GCC GGG ATT AGT GAG GCA TAT TCA GGG CAT GTC GAG CCT TTC 772  
Ile Asn Ala Gly Ile Ser Glu Ala Tyr Ser Gly His Val Glu Pro Phe  
225 230 235

TGG GAA AAT AAC AAG CTG CAC ATC CGT TGG TTT ACT ATC TCG AAA GAA 820  
Trp Glu Asn Asn Lys Leu His Ile Arg Trp Phe Thr Ile Ser Lys Glu  
240 245 250

GAT AAA ATA GAT TTT GTT TAT AAA AAC ATC TGG GTG ATG AGT AGC GAT 868  
Asp Lys Ile Asp Phe Val Tyr Lys Asn Ile Trp Val Met Ser Ser Asp  
255 260 265

TAT AGC TGG GCA TCA AAG AAA AAA ATC TTG GAA CTT TCT TTT ACT GAC 916  
Tyr Ser Trp Ala Ser Lys Lys Lys Ile Leu Glu Leu Ser Phe Thr Asp  
270 275 280

TAC AAT AGA GTT GGA GCA ACA GGA TCA TCA AGC CCG ACT GAA GTA GCT 964  
Tyr Asn Arg Val Gly Ala Thr Gly Ser Ser Ser Pro Thr Glu Val Ala  
285 290 295 300

TCA CAA TAT GGT TCT GAT GCT CAG ATG AAT ATT TCT GAT GAT GGG ACT 1012  
Ser Gln Tyr Gly Ser Asp Ala Gln Met Asn Ile Ser Asp Asp Gly Thr  
305 310 315

GTA CTT ATT TTT CAG AAT GCC GGC GGA GCT ACT CCC AGT ACT GGA GTG 1060  
Val Leu Ile Phe Gln Asn Ala Gly Gly Ala Thr Pro Ser Thr Gly Val  
320 325 330

ACG TTA TGT TAT GAC TCT GGC AAC GTG ATT AAG AAC CTA TCT AGT ACA 1108  
Thr Leu Cys Tyr Asp Ser Gly Asn Val Ile Lys Asn Leu Ser Ser Thr  
335 340 345

GGA AGT GCA AAT TTA TCG TCA AAG GAT TAT GCC ACA ACT AAA TTA CGC 1156  
Gly Ser Ala Asn Leu Ser Ser Lys Asp Tyr Ala Thr Thr Lys Leu Arg  
350 355 360

ATG TGT CAT GGA CAA AGT TAC AAT GAT AAT AAC TAC TGC AAT TTT ACA 1204  
Met Cys His Gly Gln Ser Tyr Asn Asp Asn Asn Tyr Cys Asn Phe Thr  
365 370 375 380

CTC TCT ATT AAT ACA ATA GAA TTC ACC TCC TAC GGC ACA TTC TCA TCA 1252  
Leu Ser Ile Asn Thr Ile Glu Phe Thr Ser Tyr Gly Thr Phe Ser Ser  
385 390 395

GAT GGA AAA CAA TTT ACA CCA CCT TCT GGT TCT GCC ATT GAT TTA CAC 1300  
Asp Gly Lys Gln Phe Thr Pro Pro Ser Gly Ser Ala Ile Asp Leu His  
400 405 410

CTC CCT AAT TAT GTA GAT CTC AAC GCG CTA TTA GAT ATT AGC CTC GAT 1348  
Leu Pro Asn Tyr Val Asp Leu Asn Ala Leu Leu Asp Ile Ser Leu Asp  
415 420 425

TCA CTA CTT AAT TAT GAC GTT CAG GGG CAG TTT GGC GGA TCT AAT CCG 1396  
Ser Leu Leu Asn Tyr Asp Val Gln Gly Gln Phe Gly Gly Ser Asn Pro  
430 435 440

GTT GAT AAT TTC AGT GGT CCC TAT GGT ATT TAT CTA TGG GAA ATC TTC 1444  
Val Asp Asn Phe Ser Gly Pro Tyr Gly Ile Tyr Leu Trp Glu Ile Phe  
445 450 455 460

TTC CAT ATT CCG TTC CTT GTT ACG GTC CGT ATG CAA ACC GAA CAA CGT 1492  
Phe His Ile Pro Phe Leu Val Thr Val Arg Met Gln Thr Glu Gln Arg  
465 470 475

TAC GAA GAC GCG GAC ACT TGG TAC AAA TAT ATT TTC CGC AGC GCC GGT 1540  
Tyr Glu Asp Ala Asp Thr Trp Tyr Lys Tyr Ile Phe Arg Ser Ala Gly  
480 485 490

TAT CGC GAT GCT AAT GGC CAG CTC ATT ATG GAT GGC AGT AAA CCA CGT 1588  
Tyr Arg Asp Ala Asn Gly Gln Leu Ile Met Asp Gly Ser Lys Pro Arg  
495 500 505

TAT TGG AAT GTG ATG CCA TTG CAA CTG GAT ACC GCA TGG GAT ACC ACA 1636  
Tyr Trp Asn Val Met Pro Leu Gln Leu Asp Thr Ala Trp Asp Thr Thr  
510 515 520

CAG CCC GCC ACC ACT GAT CCA GAT GTG ATC GCT ATG GCG GAC CCG ATG 1684  
Gln Pro Ala Thr Thr Asp Pro Asp Val Ile Ala Met Ala Asp Pro Met  
525 530 535 540

CAT TAC AAG CTG GCG ATA TTC CTG CAT ACC CTT GAT CTA TTG ATT GCC 1732  
His Tyr Lys Leu Ala Ile Phe Leu His Thr Leu Asp Leu Leu Ile Ala  
545 550 555

CGA GGC GAC AGC GCT TAC CGT CAA CTT GAA CGC GAT ACT CTA GTC GAA 1780  
Arg Gly Asp Ser Ala Tyr Arg Gln Leu Glu Arg Asp Thr Leu Val Glu  
560 565 570

GCC AAA ATG TAC TAC ATT CAG GCA CAA CAG CTA CTG GGA CCG CGC CCT 1828  
Ala Lys Met Tyr Tyr Ile Gln Ala Gln Gln Leu Leu Gly Pro Arg Pro  
575 580 585

GAT ATC CAT ACC ACC AAT ACT TGG CCA AAT CCC ACC TTG AGT AAA GAA 1876  
Asp Ile His Thr Thr Asn Thr Trp Pro Asn Pro Thr Leu Ser Lys Glu  
590 595 600

GCT GGC GCT ATT GCC ACA CCG ACA TTC CTC AGT TCA CCG GAG GTG ATG 1924  
Ala Gly Ala Ile Ala Thr Pro Thr Phe Leu Ser Ser Pro Glu Val Met  
605 610 615 620

ACG TTC GCT GCC TGG CTA AGC GCA GGC GAT ACC GCA AAT ATT GGC GAC 1972  
Thr Phe Ala Ala Trp Leu Ser Ala Gly Asp Thr Ala Asn Ile Gly Asp  
625 630 635

GGT ~~GAT~~ TTC TTG CCA CCG TAC AAC GAT GTA CTA CTC GGT TAC TGG GAT 2020  
 Gly Asp Phe Leu Pro Pro Tyr Asn Asp Val Leu Leu Gly Tyr Trp Asp  
 640 645 650

AAA CTT GAG TTA CGC CTA TAC AAC CTG CGC CAC AAT CTG AGT CTG GAT 2068  
 Lys Leu Glu Leu Arg Leu Tyr Asn Leu Arg His Asn Leu Ser Leu Asp  
 655 660 665

GGT CAA CCG CTA AAT CTG CCA CTG TAT GCC ACG CCG GTA GAC CCG AAA 2116  
 Gly Gln Pro Leu Asn Leu Pro Leu Tyr Ala Thr Pro Val Asp Pro Lys  
 670 675 680

ACC CTG CAA CGC CAG CAA GCC GGA GGG GAC GGT ACA GGC AGT AGT CCG 2164  
 Thr Leu Gln Arg Gln Gln Ala Gly Gly Asp Gly Thr Gly Ser Ser Pro  
 685 690 695 700

GCT GGT GGT CAA GGC AGT GTT CAG GGC TGG CGC TAT CCG TTA TTG GTA 2212  
 Ala Gly Gly Gln Gly Ser Val Gln Gly Trp Arg Tyr Pro Leu Leu Val  
 705 710 715

GAA CGC GCC CGC TCT GCC GTG AGT TTG TTG ACT CAG TTC GGC AAC AGC 2260  
 Glu Arg Ala Arg Ser Ala Val Ser Leu Leu Thr Gln Phe Gly Asn Ser  
 720 725 730

TTA CAA ACA ACG TTA GAA CAT CAG GAT AAT GAA AAA ATG ACG ATA CTG 2308  
 Leu Gln Thr Thr Leu Glu His Gln Asp Asn Glu Lys Met Thr Ile Leu  
 735 740 745

TTG CAG ACT CAA CAG GAA GCC ATC CTG AAA CAT CAG CAC GAT ATA CAA 2356  
 Leu Gln Thr Gln Gln Glu Ala Ile Leu Lys His Gln His Asp Ile Gln  
 750 755 760

CAA AAT AAT CTA AAA GGA TTA CAA CAC AGC CTG ACC GCA TTA CAG GCT 2404  
 Gln Asn Asn Leu Lys Gly Leu Gln His Ser Leu Thr Ala Leu Gln Ala  
 765 770 775 780

AGC CGT GAT GGC GAC ACA TTG CGG CAA AAA CAT TAC AGC GAC CTG ATT 2452  
 Ser Arg Asp Gly Asp Thr Leu Arg Gln Lys His Tyr Ser Asp Leu Ile  
 785 790 795

AAC GGT GGT CTA TCT GCG GCA GAA ATC GCC GGT CTG ACA CTA CGC AGC 2500  
 Asn Gly Gly Leu Ser Ala Ala Glu Ile Ala Gly Leu Thr Leu Arg Ser  
 800 805 810

ACC GCC ATG ATT ACC AAT GGC GTT GCA ACG GGA TTG CTG ATT GCC GGC 2548  
 Thr Ala Met Ile Thr Asn Gly Val Ala Thr Gly Leu Leu Ile Ala Gly  
 815 820 825

GGA ATC GCC AAC GCG GTA CCT AAC GTC TTC GGG CTG GCT AAC GGT GGA 2596  
 Gly Ile Ala Asn Ala Val Pro Asn Val Phe Gly Leu Ala Asn Gly Gly  
 830 835 840

TCG GAA TGG GGA GCG CCA TTA ATT GGC TCC GGG CAA GCA ACC CAA GTT 2644  
 Ser Glu Trp Gly Ala Pro Leu Ile Gly Ser Gly Gln Ala Thr Gln Val  
 845 850 855 860

GGC GCC GGC ATC CAG GAT CAG AGC GCG GGC ATT TCA GAA GTG ACA GCA 2692  
 Gly Ala Gly Ile Gln Asp Gln Ser Ala Gly Ile Ser Glu Val Thr Ala  
 865 870 875

GGC TAT CAG CGT CGT CAG GAA GAA TGG GCA TTG CAA CGG GAT ATT GCT 2740  
 Gly Tyr Gln Arg Arg Gln Glu Glu Trp Ala Leu Gln Arg Asp Ile Ala  
 880 885 890

GAT AAC GAA ATA ACC CAA CTG GAT GCC CAG ATA CAA AGC CTG CAA GAG 2788  
 Asp Asn Glu Ile Thr Gln Leu Asp Ala Gln Ile Gln Ser Leu Gln Glu  
 895 900 905

CAA ATC ACG ATG GCA CAA AAA CAG ATC ACG CTC TCT GAA ACC GAA CAA 2836  
 Gln Ile Thr Met Ala Gln Lys Gln Ile Thr Leu Ser Glu Thr Glu Gln  
 910 915 920

GCG AAT GCC CAA GCG ATT TAT GAC CTG CAA ACC ACT CGT TTT ACC GGG 2884  
Ala Asn Ala Gln Ala Ile Tyr Asp Leu Gln Thr Thr Arg Phe Thr Gly  
925 930 935 940

CAG GCA CTG TAT AAC TGG ATG GCC GGT CGT CTC TCC GCG CTC TAT TAC 2932  
Gln Ala Leu Tyr Asn Trp Met Ala Gly Arg Leu Ser Ala Leu Tyr Tyr  
945 950 955

CAA ATG TAT GAT TCC ACT CTG CCA ATC TGT CTC CAG CCA AAA GCC GCA 2980  
Gln Met Tyr Asp Ser Thr Leu Pro Ile Cys Leu Gln Pro Lys Ala Ala  
960 965 970

TTA GTA CAG GAA TTA GGC GAG AAA GAG AGC GAC AGT CTT TTC CAG GTT 3028  
Leu Val Gln Glu Leu Gly Glu Lys Glu Ser Asp Ser Leu Phe Gln Val  
975 980 985

CCG GTG TGG AAT GAT CTG TGG CAA GGG CTG TTA GCA GGA GAA GGT TTA 3076  
Pro Val Trp Asn Asp Leu Trp Gln Gly Leu Leu Ala Gly Glu Gly Leu  
990 995 1000

AGT TCA GAG CTA CAG AAA CTG GAT GCC ATC TGG CTT GCA CGT GGT GGT 3124  
Ser Ser Glu Leu Gln Lys Leu Asp Ala Ile Trp Leu Ala Arg Gly Gly  
1005 1010 1015 1020

ATT GGG CTA GAA GCC ATC CGC ACC GTG TCG CTG GAT ACC CTG TTT GGC 3172  
Ile Gly Leu Glu Ala Ile Arg Thr Val Ser Leu Asp Thr Leu Phe Gly  
1025 1030 1035

ACA GGG ACG TTA AGT GAA AAT ATC AAT AAA GTG CTT AAC GGG GAA ACG 3220  
Thr Gly Thr Leu Ser Glu Asn Ile Asn Lys Val Leu Asn Gly Glu Thr  
1040 1045 1050

GTA TCT CCA TCC GGT GGC GTC ACT CTG GCG CTG ACA GGG GAT ATC TTC 3268  
Val Ser Pro Ser Gly Gly Val Thr Leu Ala Leu Thr Gly Asp Ile Phe  
1055 1060 1065

CAA GCA ACA CTG GAT TTG AGT CAG CTA GGT TTG GAT AAC TCT TAC AAC 3316  
Gln Ala Thr Leu Asp Leu Ser Gln Leu Gly Leu Asp Asn Ser Tyr Asn  
1070 1075 1080

TTG GGT AAC GAG AAG AAA CGT CGT ATT AAA CGT ATC GCC GTC ACC CTG 3364  
Leu Gly Asn Glu Lys Lys Arg Arg Ile Lys Arg Ile Ala Val Thr Leu  
1085 1090 1095 1100

CCA ACA CTT CTG GGG CCA TAT CAA GAT CTT GAA GCC ACA CTG GTA ATG 3412  
Pro Thr Leu Leu Gly Pro Tyr Gln Asp Leu Glu Ala Thr Leu Val Met  
1105 1110 1115

GGT GCG GAA ATC GCC GCC TTA TCA CAC GGT GTG AAT GAC GGA GGC CGG 3460  
Gly Ala Glu Ile Ala Ala Leu Ser His Gly Val Asn Asp Gly Gly Arg  
1120 1125 1130

TTT GTT ACC GAC TTT AAC GAC AGC CGT TTT CTG CCT TTT GAA GGT CGA 3508  
Phe Val Thr Asp Phe Asn Asp Ser Arg Phe Leu Pro Phe Glu Gly Arg  
1135 1140 1145

GAT GCA ACA ACC GGC ACA CTG GAG CTC AAT ATT TTC CAT GCG GGT AAA 3556  
Asp Ala Thr Thr Gly Thr Leu Glu Leu Asn Ile Phe His Ala Gly Lys  
1150 1155 1160

GAG GGA ACG CAA CAC GAG TTG GTC GCG AAT CTG AGT GAC ATC ATT GTG 3604  
Glu Gly Thr Gln His Glu Leu Val Ala Asn Leu Ser Asp Ile Ile Val  
1165 1170 1175 1180

CAT CTG AAT TAC ATC ATT CGA GAC GCG TAA ATTTCTTTTC TTTGTCGATT 3654  
His Leu Asn Tyr Ile Ile Arg Asp Ala \*  
1185 1190

ACAGGTCCCT ATCAGGGGCC TGTTATTAAG GAGTACTTTA TGCAGGATTC ACCAGAAGTA 3714  
TCGATTACAA CGCTGTCACT TCCCAAAGGT GCGGGTGCTA TCAATGGCAT GGGAGAAGCA 3774  
CTGAATGCTG CCGGCCCTGA TGGAATGGCC TCCCTATCTC TGCCATTACC CCTTTCGACC 3834  
GGCAGAGGGA CGGCTCCTGG ATTATCGCTG ATTTACAGCA ACAGTGCAGG TAATGGGCCT 3894  
TTCGGCATCG GCTGGCAATG CGGTGTATG TCCATTAGCC GACGCACCCA ACATGGCATT 3954  
CCACAATACG GTAATGACGA CACGTTCCTA TCCCCACAAG GCGAGGTCAT GAATATCGCC 4014  
CTGAATGACC AAGGGCAACC TGATATCCGT CAAGACGTTA AAACGCTGCA AGGCGTTACC 4074  
TTGCCAATTT CCTATACCGT GACCCGCTAT CAAGCCCGCC AGATCCTGGA TTTTCAGTAA 4134  
ATCGAATACT GGCAACCTGC CTCCGGTCAA GAAGGACGCG CTTTCTGGCT GATATCGACA 4194  
CCGGACGGGC ATCTACACAT CTTAGGGAAA ACCGCGCAGG CTTGTCTGGC AAATCCGCAA 4254  
AATGACCAAC AAATCGCCCA GTGGTTGCTG GAAGAACTG TGACGCCAGC CCGTGAACAT 4314  
GTCAGCTATC AATATCGAGC CGAAGATGAA GCCCATTGTG ACGACAATGA AAAAACCGCT 4374  
CATCCCAATG TTACCGCACA GCGCTATCTG GTACAGGTGA ACTACAGGCA ACATCAAACC 4434  
  
ACAAGCCAGC CTGTTTCGTAC TGGATAACGC ACCTCCCGCA CCGGAAGAGT GGCTGTTTTCA 4494  
TCTGGTCTTT GACCACGGTG AGCGCGTACC TCACTTCATA CCGTGCCAAC ATGGGATGCA 4554  
GGTACAGCGC AATGGTCTGT ACGCCCGGAT ATCTTCTCTC GCTATGAATA TGGTTTTGAA 4614  
GTGCGTACTC GCCGCTTATG TCAACAAGTG CTGATGTTTC ACCGCACCGC GCTCATGGCC 4674  
GGAGAAGCCA GTACCAATGA CGCCCCGAA CTGGTTGGAC GCTTAATACT GGAATATGAC 4634  
AAAAACGCCA GCGTCACCAC GTTGATTACC ATCCGTCAAT TAAGCCATGA ATCGGACGGG 4794  
AGGCCAGTCA CCCAGCCACC ACTAGAACTA GCCTGGCAAC GGTTTGATCT GGAGAAAATC 4854  
CCGACATGGC AACGCTTGA CGCACTAGAT AATTTTAACT CGCAGCAACG TTATCAACTG 4865  
GTTGATCTGC GGGGAGAAGG GTTGCCAGGT ATGCTGTATC AAGATCGAGG CGCTTGGTGG 4914  
TATAAAGCTC CGCAACGTCA GGAAGACGGA GACAGCAATG CCGTCACTTA CGACAAAATC 4974  
GCCCCACTGC CTACCCTACC CAATTTGCAG GATAATGCCT CATTGATGGA TATCAACGGA 5034  
GACGGCCAAC TGGATTGGGT TGTTACCGCC TCCGGTATTC GCGGATACCA TAGTCAGCAA 5094  
CCCAGTGGAA AGTGGACGCA CTTTACGCCA ATCAATGCCT TGCCCGTGGG ATATTTTCAT 5214  
CCAAGCATCC AGTTCGCTGA CTTACCGGG GCAGGCTTAT CTGATTTAGT GTTGATCGGG 5274  
CCGAAAAGCG TGCGTCTATA TGCCAACCAG CGAAACGGCT GCGGTAAAGG AGAAGATGTC 5334  
CCCCAATCCA CAGGTATCAC CCTGCCTGTC ACAGGGACCG ATGCCCGCAA ACTGGTGGCT 5394  
TTCAGTGATA TGCTCGGTTT CCGTCAACAA CATCTGGTGG AAATCAAGGG TAATCGCGTC 5454  
ACCTGTTGGC CGAATCTAGG GCATGGCCGT TTCGGTCAAC CACTAACTCT GTCAGGATTT 5514  
AGCCAGCCCC AAAATAGCTT CAATCCCGAA CCGCTGTTTC TGGCGGATAT CGACGGCTCC 5574  
GGCACCACCG ACCTTATCTA TGCGCAATCC GGCTCTTTGC TCATTTATCT CAACCAAAGT 5634  
GGTAATCAGT TTGATGCCCC GTTGACATTA GCGTTGCCAG AAGGCGTACA ATTTGACAAC 5694  
ACTTGCCAAC TTCAAGTCGC CGATATTCAG GGATTAGGGA TAGCCAGCTT GATTCTGACT 5754

GTGCCACATA TCGCGCCACA TCACTGGCGT TGTGACCTGT CACTGACCAA ACCCTGGTTG 5814  
TTGAATGTAA TGAACAATAA CCGGGGCGCA CATCACACGC TACATTATCG TAGTTCCGCG 5874  
CAATTCTGGT TGGATGAAAA ATTACAGCTC ACCAAAGCAG GCAAATCTCC GGCTTGTTAT 5934  
CTGCCGTTTC CAATGCATTT GCTATGGTAT ACCGAAATTC AGGATGAAAT CAGCGGCAAC 5994  
CGGCTCACCA GTGAAGTCAA CTACAGCCAC GCGTCTGGG ATGGTAAAGA GCGGGAATTC 6054

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 26:

(i) CHARAKTERISTIKA SEKVENCIE:

(A) DĹŽKA: 1189 aminokyselín

(B) TYP: aminokyselina

(D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: proteín

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 26 (TcaB proteín):

Ala Leu Val Ala His Tyr Ile Ala Thr Gln Val Pro Ala Asp Leu Lys  
20 25 30  
Glu Ser Ile Gln Thr Ala Asp Asp Leu Tyr Glu Tyr Leu Leu Leu Asp  
35 40 45  
Thr Lys Ile Ser Asp Leu Val Thr Thr Ser Pro Leu Ser Glu Ala Ile  
50 55 60  
Gly Ser Leu Gln Leu Phe Ile His Arg Ala Ile Glu Gly Tyr Asp Gly  
65 70 75 80  
Thr Leu Ala Asp Ser Ala Lys Pro Tyr Phe Ala Asp Glu Gln Phe Leu  
85 90 95  
Tyr Asn Trp Asp Ser Phe Asn His Arg Tyr Ser Thr Trp Ala Gly Lys  
100 105 110  
Glu Arg Leu Lys Phe Tyr Ala Gly Asp Tyr Ile Asp Pro Thr Leu Arg  
115 120 125  
Leu Asn Lys Thr Glu Ile Phe Thr Ala Phe Glu Gln Gly Ile Ser Gln  
130 135 140  
Gly Lys Leu Lys Ser Glu Leu Val Glu Ser Lys Leu Arg Asp Tyr Leu  
145 150 155 160  
Ile Ser Tyr Asp Thr Leu Ala Thr Leu Asp Tyr Ile Thr Ala Cys Gln  
165 170 175  
Gly Lys Asp Asn Lys Thr Ile Phe Phe Ile Gly Arg Thr Gln Asn Ala  
180 185 190  
Pro Tyr Ala Phe Tyr Trp Arg Lys Leu Thr Leu Val Thr Asp Gly Gly  
195 200 205

Lys Leu Lys Pro Asp Gln Trp Ser Glu Trp Arg Ala Ile Asn Ala Gly  
210 215 220

Ile Ser Glu Ala Tyr Ser Gly His Val Glu Pro Phe Trp Glu Asn Asn  
225 230 235 240

Lys Leu His Ile Arg Trp Phe Thr Ile Ser Lys Glu Asp Lys Ile Asp  
245 250 255

Phe Val Tyr Lys Asn Ile Trp Val Met Ser Ser Asp Tyr Ser Trp Ala  
260 265 270

Ser Lys Lys Lys Ile Leu Glu Leu Ser Phe Thr Asp Tyr Asn Arg Val  
275 280 285

Gly Ala Thr Gly Ser Ser Ser Pro Thr Glu Val Ala Ser Gln Tyr Gly  
290 295 300

Ser Asp Ala Gln Met Asn Ile Ser Asp Asp Gly Thr Val Leu Ile Phe  
305 310 315 320

Gln Asn Ala Gly Gly Ala Thr Pro Ser Thr Gly Val Thr Leu Cys Tyr  
325 330 335

Asp Ser Gly Asn Val Ile Lys Asn Leu Ser Ser Thr Gly Ser Ala Asn  
340 345 350

Leu Ser Ser Lys Asp Tyr Ala Thr Thr Lys Leu Arg Met Cys His Gly  
355 360 365

Gln Ser Tyr Asn Asp Asn Asn Tyr Cys Asn Phe Thr Leu Ser Ile Asn  
370 375 380

Thr Ile Glu Phe Thr Ser Tyr Gly Thr Phe Ser Ser Asp Gly Lys Gln  
385 390 395 400

Phe Thr Pro Pro Ser Gly Ser Ala Ile Asp Leu His Leu Pro Asn Tyr  
405 410 415

Val Asp Leu Asn Ala Leu Leu Asp Ile Ser Leu Asp Ser Leu Leu Asn  
420 425 430

Tyr Asp Val Gln Gly Gln Phe Gly Gly Ser Asn Pro Val Asp Asn Phe  
435 440 445

Ser Gly Pro Tyr Gly Ile Tyr Leu Trp Glu Ile Phe Phe His Ile Pro  
450 455 460

Phe Leu Val Thr Val Arg Met Gln Thr Glu Gln Arg Tyr Glu Asp Ala  
465 470 475 480

Asp Thr Trp Tyr Lys Tyr Ile Phe Arg Ser Ala Gly Tyr Arg Asp Ala  
485 490 495

Asn Gly Gln Leu Ile Met Asp Gly Ser Lys Pro Arg Tyr Trp Asn Val  
500 505 510

Met Pro Leu Gln Leu Asp Thr Ala Trp Asp Thr Thr Gln Pro Ala Thr  
515 520 525

Thr Asp Pro Asp Val Ile Ala Met Ala Asp Pro Met His Tyr Lys Leu  
530 535 540

Ala Ile Phe Leu His Thr Leu Asp Leu Leu Ile Ala Arg Gly Asp Ser  
545 550 555 560

Ala Tyr Arg Gln Leu Glu Arg Asp Thr Leu Val Glu Ala Lys Met Tyr  
565 570 575



Tyr Ile Gln Ala Gln Gln Leu Leu Gly Pro Arg Pro Asp Ile His Thr  
580 585 590

Thr Asn Thr Trp Pro Asn Pro Thr Leu Ser Lys Glu Ala Gly Ala Ile  
595 600 605

Ala Thr Pro Thr Phe Leu Ser Ser Pro Glu Val Met Thr Phe Ala Ala  
610 615 620

Trp Leu Ser Ala Gly Asp Thr Ala Asn Ile Gly Asp Gly Asp Phe Leu  
625 630 635 640

Pro Pro Tyr Asn Asp Val Leu Leu Gly Tyr Trp Asp Lys Leu Glu Leu  
645 650 655

Arg Leu Tyr Asn Leu Arg His Asn Leu Ser Leu Asp Gly Gln Pro Leu  
660 665 670

Asn Leu Pro Leu Tyr Ala Thr Pro Val Asp Pro Lys Thr Leu Gln Arg  
675 680 685

Gln Gln Ala Gly Gly Asp Gly Thr Gly Ser Ser Pro Ala Gly Gly Gln  
690 695 700

Gly Ser Val Gln Gly Trp Arg Tyr Pro Leu Leu Val Glu Arg Ala Arg  
705 710 715 720

Ser Ala Val Ser Leu Leu Thr Gln Phe Gly Asn Ser Leu Gln Thr Thr  
725 730 735

Leu Glu His Gln Asp Asn Glu Lys Met Thr Ile Leu Leu Gln Thr Gln  
740 745 750

Gln Glu Ala Ile Leu Lys His Gln His Asp Ile Gln Gln Asn Asn Leu  
755 760 765

L. : Gly Leu Gln His Ser Leu Thr Ala Leu Gln Ala Ser Arg Asp Gly  
770 775 780

Asp Thr Leu Arg Gln Lys His Tyr Ser Asp Leu Ile Asn Gly Gly Leu  
785 790 795 800

Ser Ala Ala Glu Ile Ala Gly Leu Thr Leu Arg Ser Thr Ala Met Ile  
805 810 815

Thr Asn Gly Val Ala Thr Gly Leu Leu Ile Ala Gly Gly Ile Ala Asn  
820 825 830

Ala Val Pro Asn Val Phe Gly Leu Ala Asn Gly Gly Ser Glu Trp Gly  
835 840 845

Ala Pro Leu Ile Gly Ser Gly Gln Ala Thr Gln Val Gly Ala Gly Ile  
850 855 860

Gln Asp Gln Ser Ala Gly Ile Ser Glu Val Thr Ala Gly Tyr Gln Arg  
865 870 875 880

Arg Gln Glu Glu Trp Ala Leu Gln Arg Asp Ile Ala Asp Asn Glu Ile  
885 890 895

Thr Gln Leu Asp Ala Gln Ile Gln Ser Leu Gln Glu Gln Ile Thr Met  
900 905 910

Ala Gln Lys Gln Ile Thr Leu Ser Glu Thr Glu Gln Ala Asn Ala Gln  
915 920 925

Ala Ile Tyr Asp Leu Gln Thr Arg Phe Thr Gly Gln Ala Leu Tyr  
930 935 940

Asn Trp Met Ala Gly Arg Leu Ser Ala Leu Tyr Tyr Gln Met Tyr Asp  
945 950 955 960  
Ser Thr Leu Pro Ile Cys Leu Gln Pro Lys Ala Ala Leu Val Gln Glu  
965 970 975  
Leu Gly Glu Lys Glu Ser Asp Ser Leu Phe Gln Val Pro Val Trp Asn  
980 985 990  
Asp Leu Trp Gln Gly Leu Leu Ala Gly Glu Gly Leu Ser Ser Glu Leu  
995 1000 1005  
Gln Lys Leu Asp Ala Ile Trp Leu Ala Arg Gly Gly Ile Gly Leu Glu  
1010 1015 1020  
Ala Ile Arg Thr Val Ser Leu Asp Thr Leu Phe Gly Thr Gly Thr Leu  
1025 1030 1035 1040  
Ser Glu Asn Ile Asn Lys Val Leu Asn Gly Glu Thr Val Ser Pro Ser  
1045 1050 1055  
Gly Gly Val Thr Leu Ala Leu Thr Gly Asp Ile Phe Gln Ala Thr Leu  
1060 1065 1070  
Asp Leu Ser Gln Leu Gly Leu Asp Asn Ser Tyr Asn Leu Gly Asn Glu  
1075 1080 1085  
Lys Lys Arg Arg Ile Lys Arg Ile Ala Val Thr Leu Pro Thr Leu Leu  
1090 1095 1100  
Gly Pro Tyr Gln Asp Leu Glu Ala Thr Leu Val Met Gly Ala Glu Ile  
1105 1110 1115 1120  
Ala Ala Leu Ser His Gly Val Asn Asp Gly Gly Arg Phe Val Thr Asp  
1125 1130 1135  
Phe Asn Asp Ser Arg Phe Leu Pro Phe Glu Gly Arg Asp Ala Thr Thr  
1140 1145 1150  
Gly Thr Leu Glu Leu Asn Ile Phe His Ala Gly Lys Glu Gly Thr Gln  
1155 1160 1165  
His Glu Leu Val Ala Asn Leu Ser Asp Ile Ile Val His Leu Asn Tyr  
1170 1175 1180  
Ile Ile Arg Asp Ala \*  
1185 1190

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 27:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĹŽKA: 1881 párov báz
- (B) TYP: nukleová kyselina
- (C) REĹAZEC: dvojitý
- (D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: DNA (genómová)

(ix) CHARAKTERISTICKÝ ZNAK:

(A) MENO/ZNAČKA: CDS

(B) POLOHA: 1..1881

(C) ĎALŠIE INFORMÁCIE: tcaB<sub>i</sub>

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 27 (kódovaná oblasť  
tcaB<sub>i</sub>):

ATG	TCT	GAA	TCT	TTA	TTT	ACA	CAA	ACG	TTG	AAA	GAA	GCG	CGC	CGT	GAT	48
Met	Ser	Glu	Ser	Leu	Phe	Thr	Gln	Thr	Leu	Lys	Glu	Ala	Arg	Arg	Asp	
1				5					10					15		
GCA	TTG	GTT	GCT	CAT	TAT	ATT	GCT	ACT	CAG	GTG	CCC	GCA	GAT	TTA	AAA	96
Ala	Leu	Val	Ala	His	Tyr	Ile	Ala	Thr	Gln	Val	Pro	Ala	Asp	Leu	Lys	
			20					25					30			
GAG	AGT	ATC	CAG	ACC	GCG	GAT	GAT	CTG	TAC	GAA	TAT	CTG	TTG	CTG	GAT	144
Glu	Ser	Ile	Gln	Thr	Ala	Asp	Asp	Leu	Tyr	Glu	Tyr	Leu	Leu	Leu	Asp	
		35					40					45				
ACC	AAA	ATT	AGC	GAT	CTG	GTT	ACT	ACT	TCA	CCG	CTG	TCC	GAA	GCG	ATT	192
Thr	Lys	Ile	Ser	Asp	Leu	Val	Thr	Thr	Ser	Pro	Leu	Ser	Glu	Ala	Ile	
	50					55					60					
GGC	AGT	CTG	CAA	TTG	TTT	ATT	CAT	CGT	GCG	ATA	GAG	GGC	TAT	GAC	GGC	240
Gly	Ser	Leu	Gln	Leu	Phe	Ile	His	Arg	Ala	Ile	Glu	Gly	Tyr	Asp	Gly	
65					70				75						80	
ACG	CTG	GCA	GAC	TCA	GCA	AAA	CCC	TAT	TTT	GCC	GAT	GAA	CAG	TTT	TTA	288
Thr	Leu	Ala	Asp	Ser	Ala	Lys	Pro	Tyr	Phe	Ala	Asp	Glu	Gln	Phe	Leu	
				85					90					95		
TAT	AAC	TGG	GAT	AGT	TTT	AAC	CAC	CGT	TAT	AGC	ACT	TGG	GCT	GGC	AAG	336
Tyr	Asn	Trp	Asp	Ser	Phe	Asn	His	Arg	Tyr	Ser	Thr	Trp	Ala	Gly	Lys	
			100					105					110			
GAA	CGG	TTG	AAA	TTC	TAT	GCC	GGG	GAT	TAT	ATT	GAT	CCA	ACA	TTG	CGA	384
Glu	Arg	Leu	Lys	Phe	Tyr	Ala	Gly	Asp	Tyr	Ile	Asp	Pro	Thr	Leu	Arg	
		115					120					125				
TTG	AAT	AAG	ACC	GAG	ATA	TTT	ACC	GCA	TTT	GAA	CAA	GGT	ATT	TCT	CAA	432
Leu	Asn	Lys	Thr	Glu	Ile	Phe	Thr	Ala	Phe	Glu	Gln	Gly	Ile	Ser	Gln	
	130					135					140					
GGG	AAA	TTA	AAA	AGT	GAA	TTA	GTC	GAA	TCT	AAA	TTA	CGT	GAT	TAT	CTA	480
Gly	Lys	Leu	Lys	Ser	Glu	Leu	Val	Glu	Ser	Lys	Leu	Arg	Asp	Tyr	Leu	
145					150					155					160	
ATT	AGT	TAT	GAC	ACT	TTA	GCC	ACC	CTT	GAT	TAT	ATT	ACT	GCC	TGC	CAA	528
Ile	Ser	Tyr	Asp	Thr	Leu	Ala	Thr	Leu	Asp	Tyr	Ile	Thr	Ala	Cys	Gln	
				165					170					175		
GGC	AAA	GAT	AAT	AAA	ACC	ATC	TTC	TTT	ATT	GGC	CGT	ACA	CAG	AAT	GCA	576
Gly	Lys	Asp	Asn	Lys	Thr	Ile	Phe	Phe	Ile	Gly	Arg	Thr	Gln	Asn	Ala	
			180					185					190			
CCC	TAT	GCA	TTT	TAT	TGG	CGA	AAA	TTA	ACT	TTA	GTC	ACT	GAT	GGC	GGT	624
Pro	Tyr	Ala	Phe	Tyr	Trp	Arg	Lys	Leu	Thr	Leu	Val	Thr	Asp	Gly	Gly	
		195					200					205				
AAG	TTG	AAA	CCA	GAT	CAA	TGG	TCA	GAG	TGG	CGA	GCA	ATT	AAT	GCC	GGG	672
Lys	Leu	Lys	Pro	Asp	Gln	Trp	Ser	Glu	Trp	Arg	Ala	Ile	Asn	Ala	Gly	
	210					215						220				

ATT Ile 225	AGT Ser	GAG Glu	GCA Ala	TAT Tyr	TCA Ser 230	GGG Gly	CAT His	GTC Val	GAG Glu	CCT Pro 235	TTC Phe	TGG Trp	GAA Glu	AAT Asn	AAC Asn 240	720
AAG Lys	CTG Leu	CAC His	ATC Ile	CGT Arg 245	TGG Trp	TTT Phe	ACT Thr	ATC Ile	TCG Ser 250	AAA Lys	GAA Glu	GAT Asp	AAA Lys	ATA Ile	GAT Asp 255	768
TTT Phe	GTT Val	TAT Tyr	AAA Lys 260	AAC Asn	ATC Ile	TGG Trp	GTG Val	ATG Met 265	AGT Ser	AGC Ser	GAT Asp	TAT Tyr	AGC Ser 270	TGG Trp	GCA Ala	816
TCA Ser	AAG Lys	AAA Lys 275	AAA Lys	ATC Ile	TTG Leu	GAA Glu	CTT Leu 280	TCT Ser	TTT Phe	ACT Thr	GAC Asp 285	TAC Tyr	AAT Asn	AGA Arg	GTT Val	864
GGA Gly 290	GCA Ala	ACA Thr	GGA Gly	TCA Ser	TCA Ser	AGC Ser 295	CCG Pro	ACT Thr	GAA Glu	GTA Val	GCT Ala 300	TCA Ser	CAA Gln	TAT Tyr	GGT Gly	912
TCT Ser 305	GAT Asp	GCT Ala	CAG Gln	ATG Met	AAT Asn 310	ATT Ile	TCT Ser	GAT Asp	GAT Asp	GGG Gly 315	ACT Thr	GTA Val	CTT Leu	ATT Ile	TTT Phe 320	960
CAG Gln	AAT Asn	GCC Ala	GGC Gly 325	GGA Gly	GCT Ala	ACT Thr	CCC Pro	AGT Ser	ACT Thr	GGA Gly 330	GTG Val	ACG Thr	TTA Leu	TGT Cys	TAT Tyr 335	1008
GAC Asp	TCT Ser	GGC Gly	AAC Asn 340	GTG Val	ATT Ile	AAG Lys	AAC Asn	CTA Leu	TCT Ser	AGT Ser	ACA Thr	GGA Gly	AGT Ser	GCA Ala	AAT Asn 350	1056
TTA Leu	TCG Ser	TCA Ser 355	AAG Lys	GAT Asp	TAT Tyr	GCC Ala	ACA Thr	ACT Thr	AAA Lys	TTA Leu	CGC Arg	ATG Met	TGT Cys	CAT His	GGA Gly	1104
CAA Gln 370	AGT Ser	TAC Tyr	AAT Asn	GAT Asp	AAT Asn	AAC Asn 375	TAC Tyr	TGC Cys	AAT Asn	TTT Phe	ACA Thr 380	CTC Leu	TCT Ser	ATT Ile	AAT Asn	1152
ACA Thr 385	ATA Ile	GAA Glu	TTC Phe	ACC Thr	TCC Ser	TAC Tyr 390	GGC Gly	ACA Thr	TTC Phe	TCA Ser 395	TCA Ser	GAT Asp	GGA Gly	AAA Lys	CAA Gln 400	1200
TTT Phe	ACA Thr	CCA Pro	CCT Pro	TCT Ser 405	GGT Gly	TCT Ser	GCC Ala	ATT Ile	GAT Asp 410	TTA Leu	CAC His	CTC Leu	CCT Pro	AAT Asn 415	TAT Tyr	1248
GTA Val	GAT Asp	CTC Leu	AAC Asn 420	GCG Ala	CTA Leu	TTA Leu	GAT Asp	ATT Ile 425	AGC Ser	CTC Leu	GAT Asp	TCA Ser	CTA Leu 430	CTT Leu	AAT Asn	1296
TAT Tyr	GAC Asp	GTT Val	CAG Gln	GGG Gly	CAG Gln	TTT Phe	GGC Gly	GGA Gly	TCT Ser	AAT Asn	CCG Pro	GTT Val	GAT Asp	AAT Asn	TTC Phe	1344
AGT Ser	GGT Gly	CCC Pro	TAT Tyr	GGT Gly	ATT Ile	TAT Tyr	CTA Leu	TGG Trp	GAA Glu	ATC Ile	TTC Phe	TTC Phe	CAT His	ATT Ile	CCG Pro	1392
	450					455				460						
TTC Phe 465	CTT Leu	GTT Val	ACG Thr	GTC Val	CGT Arg 470	ATG Met	CAA Gln	ACC Thr	GAA Glu	CAA Gln 475	CGT Arg	TAC Tyr	GAA Glu	GAC Asp	GCG Ala 480	1440
GAC Asp	ACT Thr	TGG Trp	TAC Tyr	AAA Lys 485	TAT Tyr	ATT Ile	TTC Phe	CGC Arg	AGC Ser	GCC Ala 490	GGT Gly	TAT Tyr	CGC Arg	GAT Asp	GCT Ala 495	1488

AAT GGC CAG CTC ATT ATG GAT GGC AGT AAA CCA CGT TAT TGG AAT GTG 1536  
 Asn Gly Gln Leu Ile Met Asp Gly Ser Lys Pro Arg Tyr Trp Asn Val  
 500 505 510

ATG CCA TTG CAA CTG GAT ACC GCA TGG GAT ACC ACA CAG CCC GCC ACC 1584  
 Met Pro Leu Gln Leu Asp Thr Ala Trp Asp Thr Thr Gln Pro Ala Thr  
 515 520 525

ACT GAT CCA GAT GTG ATC GCT ATG GCG GAC CCG ATG CAT TAC AAG CTG 1632  
 Thr Asp Pro Asp Val Ile Ala Met Ala Asp Pro Met His Tyr Lys Leu  
 530 535 540

GCG ATA TTC CTG CAT ACC CTT GAT CTA TTG ATT GCC CGA GGC GAC AGC 1680  
 Ala Ile Phe Leu His Thr Leu Asp Leu Leu Ile Ala Arg Gly Asp Ser  
 545 550 555 560

GCT TAC CGT CAA CTT GAA CGC GAT ACT CTA GTC GAA GCC AAA ATG TAC 1728  
 Ala Tyr Arg Gln Leu Glu Arg Asp Thr Leu Val Glu Ala Lys Met Tyr  
 565 570 575

TAC ATT CAG GCA CAA CAG CTA CTG GGA CCG CGC CCT GAT ATC CAT ACC 1776  
 Tyr Ile Gln Ala Gln Gln Leu Leu Gly Pro Arg Pro Asp Ile His Thr  
 580 585 590

ACC AAT ACT TGG CCA AAT CCC ACC TTG AGT AAA GAA GCT GGC GCT ATT 1824  
 Thr Asn Thr Trp Pro Asn Pro Thr Leu Ser Lys Glu Ala Gly Ala Ile  
 595 600 605

GCC ACA CCG ACA TTC CTC AGT TCA CCG GAG GTG ATG ACG TTC GCT GCC 1872  
 Ala Thr Pro Thr Phe Leu Ser Ser Pro Glu Val Met Thr Phe Ala Ala  
 610 615 620

TGG CTA AGC 1881  
 Trp Leu Ser  
 625

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 28:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĹŽKA: 627 aminokyselín
- (B) TYP: aminokyselina
- (D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: proteín

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 28 (TcaB<sub>i</sub> proteín):

Met Ser Glu Ser Leu Phe Thr Gln Thr Leu Lys Glu Ala Arg Arg Asp  
 1 5 10 15

Ala Leu Val Ala His Tyr Ile Ala Thr Gln Val Pro Ala Asp Leu Lys  
 20 25 30

Glu Ser Ile Gln Thr Ala Asp Asp Leu Tyr Glu Tyr Leu Leu Asp  
 35 40 45

Thr Lys Ile Ser Asp Leu Val Thr Thr Ser Pro Leu Ser Glu Ala Ile  
 50 55 60

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

Ser Gly Pro Tyr Gly Ile Tyr Leu Trp Glu Ile Phe Phe His Ile Pro  
450 455 460  
Phe Leu Val Thr Val Arg Met Gln Thr Glu Gln Arg Tyr Glu Asp Ala  
465 470 475 480  
Asp Thr Trp Tyr Lys Tyr Ile Phe Arg Ser Ala Gly Tyr Arg Asp Ala  
485 490 495  
Asn Gly Gln Leu Ile Met Asp Gly Ser Lys Pro Arg Tyr Trp Asn Val  
500 505 510  
Met Pro Leu Gln Leu Asp Thr Ala Trp Asp Thr Thr Gln Pro Ala Thr  
515 520 525  
Thr Asp Pro Asp Val Ile Ala Met Ala Asp Pro Met His Tyr Lys Leu  
530 535 540  
Ala Ile Phe Leu His Thr Leu Asp Leu Leu Ile Ala Arg Gly Asp Ser  
545 550 555 560  
Ala Tyr Arg Gln Leu Glu Arg Asp Thr Leu Val Glu Ala Lys Met Tyr  
565 570 575  
Tyr Ile Gln Ala Gln Gln Leu Leu Gly Pro Arg Pro Asp Ile His Thr  
580 585 590  
Thr Asn Thr Trp Pro Asn Pro Thr Leu Ser Lys Glu Ala Gly Ala Ile  
595 600 605  
Ala Thr Pro Thr Phe Leu Ser Ser Pro Glu Val Met Thr Phe Ala Ala  
610 615 620  
Trp Leu Ser  
625

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 29:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĹŽKA: 1689 párov báz
- (B) TYP: nukleová kyselina
- (C) REŤAZEC: dvojitý
- (D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: DNA (genómová)

(ix) CHARAKTERISTICKÝ ZNAK:

- (A) MENO/ZNAČKA: CDS
- (B) POLOHA: 1..1689
- (C) ĎALŠIE INFORMÁCIE: tcaB<sub>ii</sub>

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 29 (kódovaná oblasť tcaB<sub>ii</sub>):

GCA GGC GAT ACC GCA AAT ATT GGC GAC GGT GAT TTC TTG CCA CCG TAC	48
Ala Gly Asp Thr Ala Asn Ile Gly Asp Gly Asp Phe Leu Pro Pro Tyr	
1 5 10 15	
AAC GAT GTA CTA CTC GGT TAC TGG GAT AAA CTT GAG TTA CGC CTA TAC	96
Asn Asp Val Leu Leu Gly Tyr Trp Asp Lys Leu Glu Leu Arg Leu Tyr	
20 25 30	
AAC CTG CGC CAC AAT CTG AGT CTG GAT GGT CAA CCG CTA AAT CTG CCA	144
Asn Leu Arg His Asn Leu Ser Leu Asp Gly Gln Pro Leu Asn Leu Pro	
35 40 45	
CTG TAT GCC ACG CCG GTA GAC CCG AAA ACC CTG CAA CGC CAG CAA GCC	192
Leu Tyr Ala Thr Pro Val Asp Pro Lys Thr Leu Gln Arg Gln Gln Ala	
50 55 60	
GGA GGG GAC GGT ACA GGC AGT AGT CCG GCT GGT GGT CAA GGC AGT GTT	240
Gly Gly Asp Gly Thr Gly Ser Ser Pro Ala Gly Gly Gln Gly Ser Val	
65 70 75 80	
CAG GGC TGG CGC TAT CCG TTA TTG GTA GAA CGC GCC CGC TCT GCC GTG	288
Gln Gly Trp Arg Tyr Pro Leu Leu Val Glu Arg Ala Arg Ser Ala Val	
85 90 95	
AGT TTG TTG ACT CAG TTC GGC AAC AGC TTA CAA ACA ACG TTA GAA CAT	336
Ser Leu Leu Thr Gln Phe Gly Asn Ser Leu Gln Thr Thr Leu Glu His	
100 105 110	
CAG GAT AAT GAA AAA ATG ACG ATA CTG TTG CAG ACT CAA CAG GAA GCC	384
Gln Asp Asn Glu Lys Met Thr Ile Leu Leu Gln Thr Gln Gln Glu Ala	
115 120 125	
ATC CTG AAA CAT CAG CAC GAT ATA CAA CAA AAT AAT CTA AAA GGA TTA	432
Ile Leu Lys His Gln His Asp Ile Gln Gln Asn Asn Leu Lys Gly Leu	
130 135 140	
CAA CAC AGC CTG ACC GCA TTA CAG GCT AGC CGT GAT GGC GAC ACA TTG	480
Gln His Ser Leu Thr Ala Leu Gln Ala Ser Arg Asp Gly Asp Thr Leu	
145 150 155 160	
CGG CAA AAA CAT TAC AGC GAC CTG ATT AAC GGT GGT CTA TCT GCG GCA	528
Arg Gln Lys His Tyr Ser Asp Leu Ile Asn Gly Gly Leu Ser Ala Ala	
165 170 175	
GAA ATC GCC GGT CTG ACA CTA CGC AGC ACC GCC ATG ATT ACC AAT GGC	576
Glu Ile Ala Gly Leu Thr Leu Arg Ser Thr Ala Met Ile Thr Asn Gly	
180 185 190	
GTT GCA ACG GGA TTG CTG ATT GCC GGC GGA ATC GCC AAC GCG GTA CCT	624
Val Ala Thr Gly Leu Leu Ile Ala Gly Gly Ile Ala Asn Ala Val Pro	
195 200 205	
AAC GTC TTC GGG CTG GCT AAC GGT GGA TCG GAA TGG GGA GCG CCA TTA	672
Asn Val Phe Gly Leu Ala Asn Gly Gly Ser Glu Trp Gly Ala Pro Leu	
210 215 220	
ATT GGC TCC GGG CAA GCA ACC CAA GTT GGC GCC GGC ATC CAG GAT CAG	720
Ile Gly Ser Gly Gln Ala Thr Gln Val Gly Ala Gly Ile Gln Asp Gln	
225 230 235 240	
AGC GCG GGC ATT TCA GAA GTG ACA GCA GGC TAT CAG CGT CGT CAG GAA	768
Ser Ala Gly Ile Ser Glu Val Thr Ala Gly Tyr Gln Arg Arg Gln Glu	
245 250 255	
GAA TGG GCA TTG CAA CGG GAT ATT GCT GAT AAC GAA ATA ACC CAA CTG	816
Glu Trp Ala Leu Gln Arg Asp Ile Ala Asp Asn Glu Ile Thr Gln Leu	
260 265 270	
GAT GCC CAG ATA CAA AGC CTG CAA GAG CAA ATC ACG ATG GCA CAA AAA	864
Asp Ala Gln Ile Gln Ser Leu Gln Glu Gln Ile Thr Met Ala Gln Lys	
275 280 285	



CAG ATC ACG CTC TCT GAA ACC GAA CAA GCG AAT GCC CAA GCG ATT TAT 912  
Gln Ile Thr Leu Ser Glu Thr Glu Gln Ala Asn Ala Gln Ala Ile Tyr  
290 295 300

GAC CTG CAA ACC ACT CGT TTT ACC GGG CAG GCA CTG TAT AAC TGG ATG 960  
Asp Leu Gln Thr Thr Arg Phe Thr Gly Gln Ala Leu Tyr Asn Trp Met  
305 310 315 320

GCC GGT CGT CTC TCC GCG CTC TAT TAC CAA ATG TAT GAT TCC ACT CTG 1008  
Ala Gly Arg Leu Ser Ala Leu Tyr Tyr Gln Met Tyr Asp Ser Thr Leu  
325 330 335

CCA ATC TGT CTC CAG CCA AAA GCC GCA TTA GTA CAG GAA TTA GGC GAG 1056  
Pro Ile Cys Leu Gln Pro Lys Ala Ala Leu Val Gln Glu Leu Gly Glu  
340 345 350

AAA GAG AGC GAC AGT CTT TTC CAG GTT CCG GTG TGG AAT GAT CTG TGG 1104  
Lys Glu Ser Asp Ser Leu Phe Gln Val Pro Val Trp Asn Asp Leu Trp  
355 360 365

CAA GGG CTG TTA GCA GGA GAA GGT TTA AGT TCA GAG CTA CAG AAA CTG 1152  
Gln Gly Leu Leu Ala Gly Glu Gly Leu Ser Ser Glu Leu Gln Lys Leu  
370 375 380

GAT GCC ATC TGG CTT GCA CGT GGT GGT ATT GGG CTA GAA GCC ATC CGC 1200  
Asp Ala Ile Trp Leu Ala Arg Gly Gly Ile Gly Leu Glu Ala Ile Arg  
385 390 395 400

ACC GTG TCG CTG GAT ACC CTG TTT GGC ACA GGG ACG TTA AGT GAA AAT 1248  
Thr Val Ser Leu Asp Thr Leu Phe Gly Thr Gly Thr Leu Ser Glu Asn  
405 410 415

ATC AAT AAA GTG CTT AAC GGG GAA ACG GTA TCT CCA TCC GGT GGC GTC 1296  
Ile Asn Lys Val Leu Asn Gly Glu Thr Val Ser Pro Ser Gly Gly Val  
420 425 430

ACT CTG GCG CTG ACA GGG GAT ATC TTC CAA GCA ACA CTG GAT TTG AGT 1344  
Thr Leu Ala Leu Thr Gly Asp Ile Phe Gln Ala Thr Leu Asp Leu Ser  
435 440 445

CAG CTA GGT TTG GAT AAC TCT TAC AAC TTG GGT AAC GAG AAG AAA CGT 1392  
Gln Leu Gly Leu Asp Asn Ser Tyr Asn Leu Gly Asn Glu Lys Lys Arg  
450 455 460

CGT ATT AAA CGT ATC GCC GTC ACC CTG CCA ACA CTT CTG GGG CCA TAT 1440  
Arg Ile Lys Arg Ile Ala Val Thr Leu Pro Thr Leu Gly Pro Tyr  
465 470 475 480

CAA GAT CTT GAA GCC ACA CTG GTA ATG GGT GCG GAA ATC GCC GCC TTA 1488  
Gln Asp Leu Glu Ala Thr Leu Val Met Gly Ala Glu Ile Ala Ala Leu  
485 490 495

TCA CAC GGT GTG AAT GAC GGA GGC CGG TTT GTT ACC GAC TTT AAC GAC 1536  
Ser His Gly Val Asn Asp Gly Gly Arg Phe Val Thr Asp Phe Asn Asp  
500 505 510

AGC CGT TTT CTG CCT TTT GAA GGT CGA GAT GCA ACA ACC GGC ACA CTG 1584  
Ser Arg Phe Leu Pro Phe Glu Gly Arg Asp Ala Thr Thr Gly Thr Leu  
515 520 525

GAG CTC AAT ATT TTC CAT GCG GGT AAA GAG GGA ACG CAA CAC GAG TTG 1632  
Glu Leu Asn Ile Phe His Ala Gly Lys Glu Gly Thr Gln His Glu Leu  
530 535 540

GTC GCG AAT CTG AGT GAC ATC ATT GTG CAT CTG AAT TAC ATC ATT CGA 1680  
Val Ala Asn Leu Ser Asp Ile Ile Val His Leu Asn Tyr Ile Ile Arg  
545 550 555 560

GAC GCG TAA  
Asp Ala \* 1689

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 30:

(i) CHARAKTERISTIKA SEKVENCIE:

(A) DĹŽKA: 562 aminokyselín

(B) TYP: aminokyselina

(D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: proteín

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 30 (TcaB<sub>ii</sub> proteín):

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

Asp Ala Gln Ile Gln Ser Leu Gln Glu Gln Ile Thr Met Ala Gln Lys  
 275 280 285  
 Gln Ile Thr Leu Ser Glu Thr Glu Gln Ala Asn Ala Gln Ala Ile Tyr  
 290 295 300  
 Asp Leu Gln Thr Thr Arg Phe Thr Gly Gln Ala Leu Tyr Asn Trp Met  
 305 310 315 320  
 Ala Gly Arg Leu Ser Ala Leu Tyr Tyr Gln Met Tyr Asp Ser Thr Leu  
 325 330 335  
 Pro Ile Cys Leu Gln Pro Lys Ala Ala Leu Val Gln Glu Leu Gly Glu  
 340 345 350  
 Lys Glu Ser Asp Ser Leu Phe Gln Val Pro Val Trp Asn Asp Leu Trp  
 355 360 365  
 Gln Gly Leu Leu Ala Gly Glu Gly Leu Ser Ser Glu Leu Gln Lys Leu  
 370 375 380  
 Asp Ala Ile Trp Leu Ala Arg Gly Gly Ile Gly Leu Glu Ala Ile Arg  
 385 390 395 400  
 Thr Val Ser Leu Asp Thr Leu Phe Gly Thr Gly Thr Leu Ser Glu Asn  
 405 410 415  
 Ile Asn Lys Val Leu Asn Gly Glu Thr Val Ser Pro Ser Gly Gly Val  
 420 425 430  
 Thr Leu Ala Leu Thr Gly Asp Ile Phe Gln Ala Thr Leu Asp Leu Ser  
 435 440 445  
 Gln Leu Gly Leu Asp Asn Ser Tyr Asn Leu Gly Asn Glu Lys Lys Arg  
 450 455 460  
 Arg Ile Lys Arg Ile Ala Val Thr Leu Pro Thr Leu Leu Gly Pro Tyr  
 465 470 475 480  
 Gln Asp Leu Glu Ala Thr Leu Val Met Gly Ala Glu Ile Ala Ala Leu  
 485 490 495  
 Ser His Gly Val Asn Asp Gly Gly Arg Phe Val Thr Asp Phe Asn Asp  
 500 505 510  
 Ser Arg Phe Leu Pro Phe Glu Gly Arg Asp Ala Thr Thr Gly Thr Leu  
 515 520 525  
 Glu Leu Asn Ile Phe His Ala Gly Lys Glu Gly Thr Gln His Glu Leu  
 530 535 540  
 Val Ala Asn Leu Ser Asp Ile Ile Val His Leu Asn Tyr Ile Ile Arg  
 545 550 555 560  
 Asp Ala \*

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 31:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĹŽKA: 4458 párov báz
- (B) TYP: nukleová kyselina
- (C) REŤAZEC: dvojité
- (D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY:DNA (genómová)

(ix) CHARAKTERISTICKÝ ZNAK:

(A) MENO/ZNAČKA: CDS

(B) POLOHA: 1..4458

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 31 (tcac gén):

ATG	CAG	GAT	TCA	CCA	GAA	GTA	TCG	ATT	ACA	ACG	CTG	TCA	CTT	CCC	AAA	48
Met	Gln	Asp	Ser	Pro	Glu	Val	Ser	Ile	Thr	Thr	Leu	Ser	Leu	Pro	Lys	
1				5					10					15		
GGT	GGC	GGT	GCT	ATC	AAT	GGC	ATG	GGA	GAA	GCA	CTG	AAT	GCT	GCC	GGC	96
Gly	Gly	Gly	Ala	Ile	Asn	Gly	Met	Gly	Glu	Ala	Leu	Asn	Ala	Ala	Gly	
			20					25					30			
CCT	GAT	GGA	ATG	GCC	TCC	CTA	TCT	CTG	CCA	TTA	CCC	CTT	TCG	ACC	GGC	144
Pro	Asp	Gly	Met	Ala	Ser	Leu	Ser	Leu	Pro	Leu	Pro	Leu	Ser	Thr	Gly	
		35					40					45				
AGA	GGG	ACG	GCT	CCT	GGA	TTA	TCG	CTG	ATT	TAC	AGC	AAC	AGT	GCA	GGT	192
Arg	Gly	Thr	Ala	Pro	Gly	Leu	Ser	Leu	Ile	Tyr	Ser	Asn	Ser	Ala	Gly	
	50					55					60					
AAT	GGG	CCT	TTC	GGC	ATC	GGC	TGG	CAA	TGC	GGT	GTT	ATG	TCC	ATT	AGC	240
Asn	Gly	Pro	Phe	Gly	Ile	Gly	Trp	Gln	Cys	Gly	Val	Met	Ser	Ile	Ser	
65					70					75					80	
CGA	CGC	ACC	CAA	CAT	GGC	ATT	CCA	CAA	TAC	GGT	AAT	GAC	GAC	ACG	TTC	288
Arg	Arg	Thr	Gln	His	Gly	Ile	Pro	Gln	Tyr	Gly	Asn	Asp	Asp	Thr	Phe	
				85					90					95		
CTA	TCC	CCA	CAA	GGC	GAG	GTC	ATG	AAT	ATC	GCC	CTG	AAT	GAC	CAA	GGG	336
Leu	Ser	Pro	Gln	Gly	Glu	Val	Met	Asn	Ile	Ala	Leu	Asn	Asp	Gln	Gly	
			100					105					110			
CAA	CCT	GAT	ATC	CGT	CAA	GAC	GTT	AAA	ACG	CTG	CAA	GGC	GTT	ACC	TTG	384
Gln	Pro	Asp	Ile	Arg	Gln	Asp	Val	Lys	Thr	Leu	Gln	Gly	Val	Thr	Leu	
		115					120					125				
CCA	ATT	TCC	TAT	ACC	GTG	ACC	CGC	TAT	CAA	GCC	CGC	CAG	ATC	CTG	GAT	432
Pro	Ile	Ser	Tyr	Thr	Val	Thr	Arg	Tyr	Gln	Ala	Arg	Gln	Ile	Leu	Asp	
		130				135					140					
TTC	AGT	AAA	ATC	GAA	TAC	TGG	CAA	CCT	GCC	TCC	GGT	CAA	GAA	GGA	CGC	480
Phe	Ser	Lys	Ile	Glu	Tyr	Trp	Gln	Pro	Ala	Ser	Gly	Gln	Glu	Gly	Arg	
145					150					155					160	
GCT	TTC	TGG	CTG	ATA	TCG	ACA	CCG	GAC	GGG	CAT	CTA	CAC	ATC	TTA	GGG	528
Ala	Phe	Trp	Leu	Ile	Ser	Thr	Pro	Asp	Gly	His	Leu	His	Ile	Leu	Gly	
				165					170					175		
AAA	ACC	GCG	CAG	GCT	TGT	CTG	GCA	AAT	CCG	CAA	AAT	GAC	CAA	CAA	ATC	576
Lys	Thr	Ala	Gln	Ala	Cys	Leu	Ala	Asn	Pro	Gln	Asn	Asp	Gln	Gln	Ile	
			180					185					190			
GCC	CAG	TGG	TTG	CTG	GAA	GAA	ACT	GTG	ACG	CCA	GCC	GGT	GAA	CAT	GTC	624
Ala	Gln	Trp	Leu	Leu	Glu	Glu	Thr	Val	Thr	Pro	Ala	Gly	Glu	His	Val	
		195					200					205				
AGC	TAT	CAA	TAT	CGA	GCC	GAA	GAT	GAA	GCC	CAT	TGT	GAC	GAC	AAT	GAA	672
Ser	Tyr	Gln	Tyr	Arg	Ala	Glu	Asp	Glu	Ala	His	Cys	Asp	Asp	Asn	Glu	
		210				215					220					

AAA Lys 225	ACC Thr	GCT Ala	CAT His	CCC Pro	AAT Asn 230	GTT Val	ACC Thr	GCA Ala	CAG Gln	CGC Arg 235	TAT Tyr	CTG Leu	GTA Val	CAG Gln	GTG Val 240	720
AAC Asn	TAC Tyr	GGC Gly	AAC Asn	ATC Ile 245	AAA Lys	CCA Pro	CAA Gln	GCC Ala	AGC Ser 250	CTG Leu	TTC Phe	GTA Val	CTG Leu	GAT Asp 255	AAC Asn	768
GCA Ala	CCT Pro	CCC Pro	GCA Ala 260	CCG Pro	GAA Glu	GAG Glu	TGG Trp	CTG Leu 265	TTT Phe	CAT His	CTG Leu	GTC Val	TTT Phe 270	GAC Asp	CAC His	816
GGT Gly	GAG Glu	CGC Arg 275	GAT Asp	ACC Thr	TCA Ser	CTT Leu	CAT His 280	ACC Thr	GTG Val	CCA Pro	ACA Thr	TGG Trp 285	GAT Asp	GCA Ala	GGT Gly	864
ACA Thr 290	GCG Ala	CAA Gln	TGG Trp	TCT Ser	GTA Val	CGC Arg 295	CCG Pro	GAT Asp	ATC Ile	TTC Phe 300	TCT Ser	CGC Arg	TAT Tyr	GAA Glu	TAT Tyr	912
GGT Gly 305	TTT Phe	GAA Glu	GTG Val	CGT Arg 310	ACT Thr	CGC Arg	CGC Arg	TTA Leu	TGT Cys	CAA Gln 315	CAA Gln	GTG Val	CTG Leu	ATG Met	TTT Phe 320	960
CAC His	CGC Arg	ACC Thr	GCG Ala	CTC Leu 325	ATG Met	GCC Ala	GGA Gly	GAA Glu	GCC Ala 330	AGT Ser	ACC Thr	AAT Asn	GAC Asp	GCC Ala 335	CCG Pro	1008
GAA Glu	CTG Leu	GTT Val	GGA Gly 340	CGC Arg	TTA Leu	ATA Ile	CTG Leu	GAA Glu 345	TAT Tyr	GAC Asp	AAA Lys	AAC Asn	GCC Ala 350	AGC Ser	GTC Val	1056
ACC Thr	ACG Thr	TTG Leu 355	ATT Ile	ACC Thr	ATC Ile	CGT Arg	CAA Gln 360	TTA Leu	AGC Ser	CAT His	GAA Glu 365	TCG Ser	GAC Asp	GGG Gly	AGG Arg	1104
CCA Pro 370	GTC Val	ACC Thr	CAG Gln	CCA Pro	CCA Pro	CTA Leu 375	GAA Glu	CTA Leu	GCC Ala	TGG Trp 380	CAA Gln	CGG Arg	TTT Phe	GAT Asp	CTG Leu	1152
GAG Glu 385	AAA Lys	ATC Ile	CCG Pro	ACA Thr	TGG Trp 390	CAA Gln	CGC Arg	TTT Phe	GAC Asp 395	GCA Ala	CTA Leu	GAT Asp	AAT Asn	TTT Phe 400	AAC Asn	1200
TCG Ser	CAG Gln	CAA Gln	CGT Arg 405	TAT Tyr	CAA Gln	CTG Leu	GTT Val	GAT Asp 410	CTG Leu	CGG Arg	GGA Gly	GAA Glu	GGG Gly	TTG Leu 415	CCA Pro	1248
GGT Gly	ATG Met	CTG Leu 420	TAT Tyr	CAA Gln	GAT Asp	CGA Arg	GGC Gly 425	GCT Ala	TGG Trp	TGG Trp	TAT Tyr	AAA Lys 430	GCT Ala	CCG Pro	CAA Gln	1296
CGT Arg	CAG Gln	GAA Glu 435	GAC Asp	GGA Gly	GAC Asp	AGC Ser	AAT Asn 440	GCC Ala	GTC Val	ACT Thr	TAC Tyr	GAC Asp 445	AAA Lys	ATC Ile	GCC Ala	1344
CCA Pro 450	CTG Leu	CCT Pro	ACC Thr	CTA Leu	CCC Pro	AAT Asn 455	TTG Leu	CAG Gln	GAT Asp	AAT Asn	GCC Ala 460	TCA Ser	TTG Leu	ATG Met	GAT Asp	1392
ATC Ile 465	AAC Asn	GGA Gly	GAC Asp	GGC Gly	CAA Gln 470	CTG Leu	GAT Asp	TGG Trp	GTT Val 475	GTT Val	ACC Thr	GCC Ala	TCC Ser	GGT Gly	ATT Ile 480	1440
CGC Arg	GGA Gly	TAC Tyr	CAT His	AGT Ser 485	CAG Gln	CAA Gln	CCC Pro	GAT Asp	GGA Gly 490	AAG Lys	TGG Trp	ACG Thr	CAC His	TTT Phe 495	ACG Thr	1488

CCA ATC AAT GCC TTG CCC GTG GAA TAT TTT CAT CCA AGC ATC CAG TTC 1536  
Pro Ile Asn Ala Leu Pro Val Glu Tyr Phe His Pro Ser Ile Gln Phe  
500 505 510

GCT GAC CTT ACC GGG GCA GGC TTA TCT GAT TTA GTG TTG ATC GGG CCG 1584  
Ala Asp Leu Thr Gly Ala Gly Leu Ser Asp Leu Val Leu Ile Gly Pro  
515 520 525

AAA AGC GTG CGT CTA TAT GCC AAC CAG CGA AAC GGC TGG CGT AAA GGA 1632  
Lys Ser Val Arg Leu Tyr Ala Asn Gln Arg Asn Gly Trp Arg Lys Gly  
530 535 540

GAA GAT GTC CCC CAA TCC ACA GGT ATC ACC CTG CCT GTC ACA GGG ACC 1680  
Glu Asp Val Pro Gln Ser Thr Gly Ile Thr Leu Pro Val Thr Gly Thr  
545 550 555 560

GAT GCC CGC AAA CTG GTG GCT TTC AGT GAT ATG CTC GGT TCC GGT CAA 1728  
Asp Ala Arg Lys Leu Val Ala Phe Ser Asp Met Leu Gly Ser Gly Gln  
565 570 575

CAA CAT CTG GTG GAA ATC AAG GGT AAT CGC GTC ACC TGT TGG CCG AAT 1776  
Gln His Leu Val Glu Ile Lys Gly Asn Arg Val Thr Cys Trp Pro Asn  
580 585 590

CTA GGG CAT GGC CGT TTC GGT CAA CCA CTA ACT CTG TCA GGA TTT AGC 1824  
Leu Gly His Gly Arg Phe Gly Gln Pro Leu Thr Leu Ser Gly Phe Ser  
595 600 605

CAG CCC GAA AAT AGC TTC AAT CCC GAA CGG CTG TTT CTG GCG GAT ATC 1872  
Gln Pro Glu Asn Ser Phe Asn Pro Glu Arg Leu Phe Leu Ala Asp Ile  
610 615 620

GAC GGC TCC GGC ACC ACC GAC CTT ATC TAT GCG CAA TCC GGC TCT TTG 1920  
Asp Gly Ser Gly Thr Thr Asp Leu Ile Tyr Ala Gln Ser Gly Ser Leu  
625 630 635 640

CTC ATT TAT CTC AAC CAA AGT GGT AAT CAG TTT GAT GCC CCG TTG ACA 1968  
Leu Ile Tyr Leu Asn Gln Ser Gly Asn Gln Phe Asp Ala Pro Leu Thr  
645 650 655

TTA GCG TTG CCA GAA GGC GTA CAA TTT GAC AAC ACT TGC CAA CTT CAA 2016  
Leu Ala Leu Pro Glu Gly Val Gln Phe Asp Asn Thr Cys Gln Leu Gln  
660 665 670

GTC GCC GAT ATT CAG GGA TTA GGG ATA GCC AGC TTG ATT CTG ACT GTG 2064  
Val Ala Asp Ile Gln Gly Leu Gly Ile Ala Ser Leu Ile Leu Thr Val  
675 680 685

CCA CAT ATC GCG CCA CAT CAC TGG CGT TGT GAC CTG TCA CTG ACC AAA 2112  
Pro His Ile Ala Pro His His Trp Arg Cys Asp Leu Ser Leu Thr Lys  
690 695 700

CCC TGG TTG TTG AAT GTA ATG AAC AAT AAC CGG GGC GCA CAT CAC ACG 2160  
Pro Trp Leu Leu Asn Val Met Asn Asn Asn Arg Gly Ala His His Thr  
705 710 715 720

CTA CAT TAT CGT AGT TCC GCG CAA TTC TGG TTG GAT GAA AAA TTA CAG 2208  
Leu His Tyr Arg Ser Ser Ala Gln Phe Trp Leu Asp Glu Lys Leu Gln  
725 730 735

CTC ACC AAA GCA GGC AAA TCT CCG GCT TGT TAT CTG CCG TTT CCA ATG 2256  
Leu Thr Lys Ala Gly Lys Ser Pro Ala Cys Tyr Leu Pro Phe Pro Met  
740 745 750

CAT TTG CTA TGG TAT ACC GAA ATT CAG GAT GAA ATC AGC GGC AAC CGG 2304  
His Leu Leu Trp Tyr Thr Glu Ile Gln Asp Glu Ile Ser Gly Asn Arg  
755 760 765

CTC ACC AGT GAA GTC AAC TAC AGC CAC GGC GTC TGG GAT GGT AAA GAG 2352  
 Leu Thr Ser Glu Val Asn Tyr Ser His Gly Val Trp Asp Gly Lys Glu  
 770 775 780

CGG GAA TTC AGA GGA TTT GGC TGC ATC AAA CAG ACA GAT ACC ACA ACG 2400  
 Arg Glu Phe Arg Gly Phe Gly Cys Ile Lys Gln Thr Asp Thr Thr Thr  
 785 790 795 800

TTT TCT CAC GGC ACC GCC CCC GAA CAG GCG GCA CCG TCG CTG AGT ATT 2448  
 Phe Ser His Gly Thr Ala Pro Glu Gln Ala Ala Pro Ser Leu Ser Ile  
 805 810 815

AGC TGG TTT GCC ACC GCC ATG GAT GAA GTA GAC AGC CAA TTA GCT ACG 2496  
 Ser Trp Phe Ala Thr Gly Met Asp Glu Val Asp Ser Gln Leu Ala Thr  
 820 825 830

GAA TAT TGG CAG GCA GAC ACG CAA GCT TAT AGC GGA TTT GAA ACC CGT 2544  
 Glu Tyr Trp Gln Ala Asp Thr Gln Ala Tyr Ser Gly Phe Glu Thr Arg  
 835 840 845

TAT ACC GTC TGG GAT CAC ACC AAC CAG ACA GAC CAA GCA TTT ACC CCC 2592  
 Tyr Thr Val Trp Asp His Thr Asn Gln Thr Asp Gln Ala Phe Thr Pro  
 850 855 860

AAT GAG ACA CAA CGT AAC TGG CTG ACG CGA GCG CTT AAA GGC CAA CTG 2640  
 Asn Glu Thr Gln Arg Asn Trp Leu Thr Arg Ala Leu Lys Gly Gln Leu  
 865 870 875 880

CTA CGC ACT GAG CTC TAC GGT CTG GAC GGA ACA GAT AAG CAA ACA GTG 2688  
 Leu Arg Thr Glu Leu Tyr Gly Leu Asp Gly Thr Asp Lys Gln Thr Val  
 885 890 895

CCT TAT ACC GTC AGT GAA TCG CGC TAT CAG GTA CGC TCT ATT CCC GTA 2736  
 Pro Tyr Thr Val Ser Glu Ser Arg Tyr Gln Val Arg Ser Ile Pro Val  
 900 905 910

AAT AAA GAA ACT GAA TTA TCT GCC TGG GTG ACT GCT ATT GAA AAT CGC 2784  
 Asn Lys Glu Thr Glu Leu Ser Ala Trp Val Thr Ala Ile Glu Asn Arg  
 915 920 925

AGC TAC CAC TAT GAA CGT ATC ATC ACT GAC CCA CAG TTC AGC CAG AGT 2832  
 Ser Tyr His Tyr Glu Arg Ile Ile Thr Asp Pro Gln Phe Ser Gln Ser  
 930 935 940

ATC AAG TTG CAA CAC GAT ATC TTT GGT CAA TCA CTG CAA AGT GTC GAT 2880  
 Ile Lys Leu Gln His Asp Ile Phe Gly Gln Ser Leu Gln Ser Val Asp ---  
 945 950 955 960

ATT GCC TGG CCG CGC CGC GAA AAA CCA GCA GTG AAT CCC TAC CCG CCT 2928  
 Ile Ala Trp Pro Arg Arg Glu Lys Pro Ala Val Asn Pro Tyr Pro Pro  
 965 970 975

ACC CTG CCG GAA ACG CTA TTT GAC AGC AGC TAT GAT GAT CAA CAA CAA 2976  
 Thr Leu Pro Glu Thr Leu Phe Asp Ser Ser Tyr Asp Asp Gln Gln Gln  
 980 985 990

CTA TTA CGT CTG GTG AGA CAA AAA AAT AGC TGG CAT CAC CTG ACT GAT 3024  
 Leu Leu Arg Leu Val Arg Gln Lys Asn Ser Trp His His Leu Thr Asp  
 995 1000 1005

GGG GAA AAC TGG CGA TTA GGT TTA CCG AAT GCA CAA CGC CGT GAT GTT 3072  
 Gly Glu Asn Trp Arg Leu Gly Leu Pro Asn Ala Gln Arg Arg Asp Val  
 1010 1015 1020

TAT ACT TAT GAC CGG AGC AAA ATT CCA ACC GAA GGG ATT TCC CTT GAA 3120  
 Tyr Thr Tyr Asp Arg Ser Lys Ile Pro Thr Glu Gly Ile Ser Leu Glu  
 1025 1030 1035 1040

ATC TTG CTG AAA GAT GAT GGC CTG CTA GCA GAT GAA AAA GCG GCC GTT 3168  
 Ile Leu Leu Lys Asp Asp Gly Leu Leu Ala Asp Glu Lys Ala Ala Val  
 1045 1050 1055

TAT CTG GGA CAA CAA CAG ACG TTT TAC ACC GCC GGT CAA GCG GAA GTC 3216  
 Tyr Leu Gly Gln Gln Gln Thr Phe Tyr Thr Ala Gly Gln Ala Glu Val  
 1060 1065 1070

ACT CTA GAA AAA CCC ACG TTA CAA GCA CTG GTC GCG TTC CAA GAA ACC 3264  
 Thr Leu Glu Lys Pro Thr Leu Gln Ala Leu Val Ala Phe Gln Glu Thr  
 1075 1080 1085

GCC ATG ATG GAC GAT ACC TCA TTA CAG GCG TAT GAA GGC GTG ATT GAA 3312  
 Ala Met Met Asp Asp Thr Ser Leu Gln Ala Tyr Glu Gly Val Ile Glu  
 1090 1095 1100

GAG CAA GAG TTG AAT ACC GCG CTG ACA CAG GCC GGT TAT CAG CAA GTC 3360  
 Glu Gln Glu Leu Asn Thr Ala Leu Thr Gln Ala Gly Tyr Gln Gln Val  
 1105 1110 1115 1120

GCG CGG TTG TTT AAT ACC AGA TCA GAA AGC CCG GTA TGG GCG GCA CGG 3408  
 Ala Arg Leu Phe Asn Thr Arg Ser Glu Ser Pro Val Trp Ala Ala Arg  
 1125 1130 1135

CAA GGT TAT ACC GAT TAC GGT GAC GCC GCA CAG TTC TGG CGG CCT CAG 3456  
 Gln Gly Tyr Thr Asp Tyr Gly Asp Ala Ala Gln Phe Trp Arg Pro Gln  
 1140 1145 1150

GCT CAG CGT AAC TCG TTG CTG ACA GGG AAA ACC ACA CTG ACC TGG GAT 3504  
 Ala Gln Arg Asn Ser Leu Leu Thr Gly Lys Thr Thr Leu Thr Trp Asp  
 1155 1160 1165

ACC CAT CAT TGT GTA ATA ATA CAG ACT CAA GAT GCC GCT GGA TTA ACG 3552  
 Thr His His Cys Val Ile Ile Gln Thr Gln Asp Ala Ala Gly Leu Thr  
 1170 1175 1180

ACG CAA GCC CAT TAC GAT TAT CGT TTC CTT ACA CCG GTA CAA CTG ACA 3600  
 Thr Gln Ala His Tyr Asp Tyr Arg Phe Leu Thr Pro Val Gln Leu Thr  
 1185 1190 1195 1200

GAT ATT AAT GAT AAT CAA CAT ATT GTG ACT CTG GAC GCG CTA GGT CGC 3648  
 Asp Ile Asn Asp Asn Gln His Ile Val Thr Leu Asp Ala Leu Gly Arg  
 1205 1210 1215

GTA ACC ACC AGC CGG TTC TGG GGC ACA GAG GCA GGA CAA GCC GCA GGC 3696  
 Val Thr Thr Ser Arg Phe Trp Gly Thr Glu Ala Gly Gln Ala Ala Gly  
 1220 1225 1230

TAT TCC AAC CAG CCC TTC ACA CCA CCG GAC TCC GTA GAT AAA GCG CTG 3744  
 Tyr Ser Asn Gln Pro Phe Thr Pro Pro Asp Ser Val Asp Lys Ala Leu  
 1235 1240 1245

GCA TTA ACC GGC GCA CTC CCT GTT GCC CAA TGT TTA GTC TAT GCC GTT 3792  
 Ala Leu Thr Gly Ala Leu Pro Val Ala Gln Cys Leu Val Tyr Ala Val  
 1250 1255 1260

GAT AGC TGG ATG CCG TCG TTA TCT TTG TCT CAG CTT TCT CAG TCA CAA 3840  
 Asp Ser Trp Met Pro Ser Leu Ser Leu Ser Gln Leu Ser Gln Ser Gln  
 1265 1270 1275 1280

GAA GAG GCA GAA GCG CTA TGG GCG CAA CTG CGT GCC GCT CAT ATG ATT 3888  
 Glu Glu Ala Glu Ala Leu Trp Ala Gln Leu Arg Ala Ala His Met Ile  
 1285 1290 1295

ACC GAA GAT GGG AAA GTG TGT GCG TTA AGC GGG AAA CGA GGA ACA AGC 3936  
 Thr Glu Asp Gly Lys Val Cys Ala Leu Ser Gly Lys Arg Gly Thr Ser  
 1300 1305 1310



CAT CAG AAC CTG ACG ATT CAA CTT ATT TCG CTA TTG GCA AGT ATT CCC 3984  
His Gln Asn Leu Thr Ile Gln Leu Ile Ser Leu Leu Ala Ser Ile Pro  
1315 1320 1325

CGT TTA CCG CCA CAT GTA CTG GGG ATC ACC ACT GAT CGC TAT GAT AGC 4032  
Arg Leu Pro Pro His Val Leu Gly Ile Thr Thr Asp Arg Tyr Asp Ser  
1330 1335 1340

GAT CCG CAA CAG CAG CAC CAA CAG ACG GTG AGC TTT AGT GAC GGT TTT 4080  
Asp Pro Gln Gln Gln His Gln Gln Thr Val Ser Phe Ser Asp Gly Phe  
1345 1350 1355 1360

GGC CGG TTA CTC CAG AGT TCA GCT CGT CAT GAG TCA GGT GAT GCC TGG 4128  
Gly Arg Leu Leu Gln Ser Ser Ala Arg His Glu Ser Gly Asp Ala Trp  
1365 1370 1375

CAA CGT AAA GAG GAT GGC GGG CTG GTC GTG GAT GCA AAT GGC GTT CTG 4176  
Gln Arg Lys Glu Asp Gly Gly Leu Val Val Asp Ala Asn Gly Val Leu  
1380 1385 1390

GTC AGT GCC CCT ACA GAC ACC CGA TGG GCC GTT TCC GGT CGC ACA GAA 4224  
Val Ser Ala Pro Thr Asp Thr Arg Trp Ala Val Ser Gly Arg Thr Glu  
1395 1400 1405

TAT GAC GAC AAA GGC CAA CCT GTG CGT ACT TAT CAA CCC TAT TTT CTA 4272  
Tyr Asp Asp Lys Gly Gln Pro Val Arg Thr Tyr Gln Pro Tyr Phe Leu  
1410 1415 1420

AAT GAC TGG CGT TAC GTT AGT GAT GAC AGC GCA CGA GAT GAC CTG TTT 4320  
Asn Asp Trp Arg Tyr Val Ser Asp Asp Ser Ala Arg Asp Asp Leu Phe  
1425 1430 1435 1440

GCC GAT ACC CAC CTT TAT GAT CCA TTG GGA CGG GAA TAC AAA GTC ATC 4368  
Ala Asp Thr His Leu Tyr Asp Pro Leu Gly Arg Glu Tyr Lys Val Ile  
1445 1450 1455

ACT GCT AAG AAA TAT TTG CGA GAA AAG CTG TAC ACC CCG TGG TTT ATT 4416  
Thr Ala Lys Lys Tyr Leu Arg Glu Lys Leu Tyr Thr Pro Trp Phe Ile  
1460 1465 1470

GTC AGT GAG GAT GAA AAC GAT ACA GCA TCA AGA ACC CCA TAG 4458  
Val Ser Glu Asp Glu Asn Asp Thr Ala Ser Arg Thr Pro \*

(2) INFORMÁCIA PRE SEKVENCIU SEKV. ID. ČÍS. 32:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĹŽKA: 1485 aminokyselín
- (B) TYP: aminokyselina
- (D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: proteín

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 32 (TcaC protein):

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

290	295	300
Gly Phe Glu Val Arg Thr Arg Arg Leu Cys Gln Gln Val Leu Met Phe 305 310 315 320		
His Arg Thr Ala Leu Met Ala Gly Glu Ala Ser Thr Asn Asp Ala Pro 325 330 335		
Glu Leu Val Gly Arg Leu Ile Leu Glu Tyr Asp Lys Asn Ala Ser Val 340 345 350		
Thr Thr Leu Ile Thr Ile Arg Gln Leu Ser His Glu Ser Asp Gly Arg 355 360 365		
Pro Val Thr Gln Pro Pro Leu Glu Leu Ala Trp Gln Arg Phe Asp Leu 370 375 380		
Glu Lys Ile Pro Thr Trp Gln Arg Phe Asp Ala Leu Asp Asn Phe Asn 385 390 395 400		
Ser Gln Gln Arg Tyr Gln Leu Val Asp Leu Arg Gly Glu Gly Leu Pro 405 410 415		
Gly Met Leu Tyr Gln Asp Arg Gly Ala Trp Trp Tyr Lys Ala Pro Gln 420 425 430		
Arg Gln Glu Asp Gly Asp Ser Asn Ala Val Thr Tyr Asp Lys Ile Ala 435 440 445		
Pro Leu Pro Thr Leu Pro Asn Leu Gln Asp Asn Ala Ser Leu Met Asp 450 455 460		
Ile Asn Gly Asp Gly Gln Leu Asp Trp Val Val Thr Ala Ser Gly Ile 465 470 475 480		
Arg Gly Tyr His Ser Gln Gln Pro Asp Gly Lys Trp Thr His Phe Thr 485 490 495		
Pro Ile Asn Ala Leu Pro Val Glu Tyr Phe His Pro Ser Ile Gln Phe 500 505 510		
Ala Asp Leu Thr Gly Ala Gly Leu Ser Asp Leu Val Leu Ile Gly Pro 515 520 525		
Lys Ser Val Arg Leu Tyr Ala Asn Gln Arg Asn Gly Trp Arg Lys Gly 530 535 540		
Glu Asp Val Pro Gln Ser Thr Gly Ile Thr Leu Pro Val Thr Gly Thr 545 550 555 560		
Asp Ala Arg Lys Leu Val Ala Phe Ser Asp Met Leu Gly Ser Gly Gln 565 570 575		
Gln His Leu Val Glu Ile Lys Gly Asn Arg Val Thr Cys Trp Pro Asn 580 585 590		
Leu Gly His Gly Arg Phe Gly Gln Pro Leu Thr Leu Ser Gly Phe Ser 595 600 605		
Gln Pro Glu Asn Ser Phe Asn Pro Glu Arg Leu Phe Leu Ala Asp Ile 610 615 620		
Asp Gly Ser Gly Thr Thr Asp Leu Ile Tyr Ala Gln Ser Gly Ser Leu 625 630 635 640		
Leu Ile Tyr Leu Asn Gln Ser Gly Asn Gln Phe Asp Ala Pro Leu Thr 645 650 655		
Leu Ala Leu Pro Glu Gly Val Gln Phe Asp Asn Thr Cys Gln Leu Gln 660 665 670		

Val Ala Asp Ile Gln Gly Leu Gly Ile Ala Ser Leu Ile Leu Thr Val  
675 680 685

Pro His Ile Ala Pro His His Trp Arg Cys Asp Leu Ser Leu Thr Lys  
690 695 700

Pro Trp Leu Leu Asn Val Met Asn Asn Asn Arg Gly Ala His His Thr  
705 710 715 720

Leu His Tyr Arg Ser Ser Ala Gln Phe Trp Leu Asp Glu Lys Leu Gln  
725 730 735

Leu Thr Lys Ala Gly Lys Ser Pro Ala Cys Tyr Leu Pro Phe Pro Met  
740 745 750

His Leu Leu Trp Tyr Thr Glu Ile Gln Asp Glu Ile Ser Gly Asn Arg  
755 760 765

Leu Thr Ser Glu Val Asn Tyr Ser His Gly Val Trp Asp Gly Lys Glu  
770 775 780

Arg Glu Phe Arg Gly Phe Gly Cys Ile Lys Gln Thr Asp Thr Thr Thr  
785 790 795 800

Phe Ser His Gly Thr Ala Pro Glu Gln Ala Ala Pro Ser Leu Ser Ile  
805 810 815

Ser Trp Phe Ala Thr Gly Met Asp Glu Val Asp Ser Gln Leu Ala Thr  
820 825 830

Glu Tyr Trp Gln Ala Asp Thr Gln Ala Tyr Ser Gly Phe Glu Thr Arg  
835 840 845

Tyr Thr Val Trp Asp His Thr Asn Gln Thr Asp Gln Ala Phe Thr Pro  
850 855 860

Asn Glu Thr Gln Arg Asn Trp Leu Thr Arg Ala Leu Lys Gly Gln Leu  
865 870 875 880

Leu Arg Thr Glu Leu Tyr Gly Leu Asp Gly Thr Asp Lys Gln Thr Val  
885 890 895

Pro Tyr Thr Val Ser Glu Ser Arg Tyr Gln Val Arg Ser Ile Pro Val  
900 905 910

Asn Lys Glu Thr Glu Leu Ser Ala Trp Val Thr Ala Ile Glu Asn Arg  
915 920 925

Ser Tyr His Tyr Glu Arg Ile Ile Thr Asp Pro Gln Phe Ser Gln Ser  
930 935 940

Ile Lys Leu Gln His Asp Ile Phe Gly Gln Ser Leu Gln Ser Val Asp  
945 950 955 960

Ile Ala Trp Pro Arg Arg Glu Lys Pro Ala Val Asn Pro Tyr Pro Pro  
965 970 975

Thr Leu Pro Glu Thr Leu Phe Asp Ser Ser Tyr Asp Asp Gln Gln Gln  
980 985 990

Leu Leu Arg Leu Val Arg Gln Lys Asn Ser Trp His His Leu Thr Asp  
995 1000 1005

Gly Glu Asn Trp Arg Leu Gly Leu Pro Asn Ala Gln Arg Arg Asp Val  
1010 1015 1020

Tyr Thr Tyr Asp Arg Ser Lys Ile Pro Thr Glu Gly Ile Ser Leu Glu  
1025 1030 1035 1040

Ile Leu Leu Lys Asp Asp Gly Leu Leu Ala Asp Glu Lys Ala Ala Val  
1045 1050 1055

Tyr Leu Gly Gln Gln Gln Thr Phe Tyr Thr Ala Gly Gln Ala Glu Val  
1060 1065 1070

Thr Leu Glu Lys Pro Thr Leu Gln Ala Leu Val Ala Phe Gln Glu Thr  
1075 1080 1085

Ala Met Met Asp Asp Thr Ser Leu Gln Ala Tyr Glu Gly Val Ile Glu  
1090 1095 1100

Glu Gln Glu Leu Asn Thr Ala Leu Thr Gln Ala Gly Tyr Gln Gln Val  
1105 1110 1115 1120

Ala Arg Leu Phe Asn Thr Arg Ser Glu Ser Pro Val Trp Ala Ala Arg  
1125 1130 1135

Gln Gly Tyr Thr Asp Tyr Gly Asp Ala Ala Gln Phe Trp Arg Pro Gln  
1140 1145 1150

Ala Gln Arg Asn Ser Leu Leu Thr Gly Lys Thr Thr Leu Thr Trp Asp  
1155 1160 1165

Thr His His Cys Val Ile Ile Gln Thr Gln Asp Ala Ala Gly Leu Thr  
1170 1175 1180

Thr Gln Ala His Tyr Asp Tyr Arg Phe Leu Thr Pro Val Gln Leu Thr  
1185 1190 1195 1200

Asp Ile Asn Asp Asn Gln His Ile Val Thr Leu Asp Ala Leu Gly Arg  
1205 1210 1215

Val Thr Thr Ser Arg Phe Trp Gly Thr Glu Ala Gly Gln Ala Ala Gly  
1220 1225 1230

Tyr Ser Asn Gln Pro Phe Thr Pro Pro Asp Ser Val Asp Lys Ala Leu  
1235 1240 1245

Ala Leu Thr Gly Ala Leu Pro Val Ala Gln Cys Leu Val Tyr Ala Val  
1250 1255 1260

Asp Ser Trp Met Pro Ser Leu Ser Leu Ser Gln Leu Ser Gln Ser Gln  
1265 1270 1275 1280

Glu Glu Ala Glu Ala Leu Trp Ala Gln Leu Arg Ala Ala His Met Ile  
1285 1290 1295

Thr Glu Asp Gly Lys Val Cys Ala Leu Ser Gly Lys Arg Gly Thr Ser  
1300 1305 1310

His Gln Asn Leu Thr Ile Gln Leu Ile Ser Leu Leu Ala Ser Ile Pro  
1315 1320 1325

Arg Leu Pro Pro His Val Leu Gly Ile Thr Thr Asp Arg Tyr Asp Ser  
1330 1335 1340

Asp Pro Gln Gln Gln His Gln Gln Thr Val Ser Phe Ser Asp Gly Phe  
1345 1350 1355 1360

Gly Arg Leu Leu Gln Ser Ser Ala Arg His Glu Ser Gly Asp Ala Trp  
1365 1370 1375

Gln Arg Lys Glu Asp Gly Gly Leu Val Val Asp Ala Asn Gly Val Leu  
1380 1385 1390

Val Ser Ala Pro Thr Asp Thr Arg Trp Ala Val Ser Gly Arg Thr Glu  
1395 1400 1405

Tyr Asp Asp Lys Gly Gln Pro Val Arg Thr Tyr Gln Pro Tyr Phe Leu  
1410 1415 1420

Asn Asp Trp Arg Tyr Val Ser Asp Asp Ser Ala Arg Asp Asp Leu Phe

1425		1430		1435		1440
Ala Asp Thr His Leu Tyr Asp Pro Leu Gly Arg Glu Tyr Lys Val Ile						
		1445		1450		1455
Thr Ala Lys Lys Tyr Leu Arg Glu Lys Leu Tyr Thr Pro Trp Phe Ile						
		1460		1465		1470
Val Ser Glu Asp Glu Asn Asp Thr Ala Ser Arg Thr Pro *						
		1475		1480		1485

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 33:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĹŽKA: 3288 párov báz
- (B) TYP: nukleová kyselina
- (C) REŤAZEC: dvojitý
- (D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: DNA (genómová)

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 33 (tcaA gén):

ATG GTG ACT GTT ATG CAA AAT AAA ATA TCA TTT TTA TCA GGT ACA TCC	48
Met Val Thr Val Met Gln Asn Lys Ile Ser Phe Leu Ser Gly Thr Ser	
1 5 10 15	
GAA CAG CCC CTG CTT GAC GCC GGT TAT CAA AAC GTA TTT GAT ATC GCA	96
Glu Gln Pro Leu Leu Asp Ala Gly Tyr Gln Asn Val Phe Asp Ile Ala	
20 25 30	
TCA ATC AGC CGG GCT ACT TTC GTT CAA TCC GTT CCC ACC CTG CCC GTT	144
Ser Ile Ser Arg Ala Thr Phe Val Gln Ser Val Pro Thr Leu Pro Val	
35 40 45	
AAA GAG GCT CAT ACC GTC TAT CGT CAG GCG CGG CAA CGT GCG GAA AAT	192
Lys Glu Ala His Thr Val Tyr Arg Gln Ala Arg Gln Arg Ala Glu Asn	
50 55 60	
CTG AAA TCC CTC TAC CGA GCC TGG CAA TTG CGT CAG GAG CCG GTT ATT	240
Leu_Lys Ser Leu Tyr Arg Ala Trp Gln Leu Arg Gln Glu Pro Val Ile	
65 70 75 80	
AAA GGG CTG GCT AAA CTT AAC CTA CAA TCC AAC GTT TCT GTG CTT CAA	288
Lys Gly Leu Ala Lys Leu Asn Leu Gln Ser Asn Val Ser Val Leu Gln	
85 90 95	
GAT GCT TTG GTA GAG AAT ATT GGC GGT GAT GGG GAT TTC AGC GAT TTA	336
Asp Ala Leu Val Glu Asn Ile Gly Gly Asp Gly Asp Phe Ser Asp Leu	
100 105 110	
ATG AAC CGT GCC AGT CAA TAT GCT GAC GCT GCC TCT ATT CAA TCC CTA	384
Met Asn Arg Ala Ser Gln Tyr Ala Asp Ala Ala Ser Ile Gln Ser Leu	
115 120 125	
TTT TCA CCG GGC CGT TAT GCT TCC GCA CTC TAC AGA GTT GCT AAA GAT	432
Phe Ser Pro Gly Arg Tyr Ala Ser Ala Leu Tyr Arg Val Ala Lys Asp	
130 135 140	

CTG CAT AAA TCA GAT TCC AGT TTG CAT ATT GAT AAT CGC CGC GCT GAT 480  
Leu His Lys Ser Asp Ser Ser Leu His Ile Asp Asn Arg Arg Ala Asp  
145 150 155 160

CTG AAG GAT CTG ATA TTA AGC GAA ACG ACG ATG AAT AAA GAG GTC ACT 528  
Leu Lys Asp Leu Ile Leu Ser Glu Thr Thr Met Asn Lys Glu Val Thr  
165 170 175

TCC CTT GAT ATC TTG TTG GAT GTG CTA CAA AAA GGC GGT AAA GAT ATT 576  
Ser Leu Asp Ile Leu Leu Asp Val Leu Gln Lys Gly Gly Lys Asp Ile

180 185 190

ACT GAG CTG TCC GGC GCA TTC TTC CCA ATG ACG TTA CCT TAT GAC GAT 624  
Thr Glu Leu Ser Gly Ala Phe Phe Pro Met Thr Leu Pro Tyr Asp Asp  
195 200 205

CAT CTG TCG CAA ATC GAT TCC GCT TTA TCG GCA CAA GCC AGA ACG CTG 672  
His Leu Ser Gln Ile Asp Ser Ala Leu Ser Ala Gln Ala Arg Thr Leu  
210 215 220

AAC GGT GTG TGG AAT ACT TTG ACA GAT ACC ACG GCA CAA GCG GTT TCA 720  
Asn Gly Val Trp Asn Thr Leu Thr Asp Thr Thr Ala Gln Ala Val Ser  
225 230 235 240

GAA CAA ACC AGT AAT ACG AAT ACA CGC AAA CTG TTC GCT GCC CAA GAT 768  
Glu Gln Thr Ser Asn Thr Asn Thr Arg Lys Leu Phe Ala Ala Gln Asp  
245 250 255

GGT AAT CAA GAT ACA TTT TTT TCC GGA AAC ACT TTT TAT TTC AAA GCG 816  
Gly Asn Gln Asp Thr Phe Phe Ser Gly Asn Thr Phe Tyr Phe Lys Ala  
260 265 270

GTG GGA TTC AGC GGG CAA CCT ATG GTT TAC CTG TCA CAG TAC ACC AGC 864  
Val Gly Phe Ser Gly Gln Pro Met Val Tyr Leu Ser Gln Tyr Thr Ser  
275 280 285

GGG AAC GGC ATT GTC GGC GCA CAA TTG ATT GCA GGT AAT CCA GAC CAA 912  
Gly Asn Gly Ile Val Gly Ala Gln Leu Ile Ala Gly Asn Pro Asp Gln  
290 295 300

GCC GCC GCC GCA ATA GTC GCA CCG TTG AAA CTC ACT TGG TCA ATG GCA 960  
Ala Ala Ala Ala Ile Val Ala Pro Leu Lys Leu Thr Trp Ser Met Ala  
305 310 315 320

AAA CAG TGT TAC TAC CTC GTC GCT CCC GAT GGT ACA ACG ATG GGA GAC 1008  
Lys Gln Cys Tyr Tyr Leu Val Ala Pro Asp Gly Thr Thr Met Gly Asp  
325 330 335

GGT AAT GTT CTG ACC GGC TGT TTC TTA AGA GGC AAC AGC CCA ACT AAC 1056  
Gly Asn Val Leu Thr Gly Cys Phe Leu Arg Gly Asn Ser Pro Thr Asn  
340 345 350

CCG GAT AAA GAC GGT ATT TTT GCT CAG GTA GCC AAC AAA TCA GGC AGT 1104  
Pro Asp Lys Asp Gly Ile Phe Ala Gln Val Ala Asn Lys Ser Gly Ser  
355 360 365

ACT CAG CCT TTG CCA AGC TTC CAT CTG CCG GTC ACA CTG GAA CAC AGC 1152  
Thr Gln Pro Leu Pro Ser Phe His Leu Pro Val Thr Leu Glu His Ser  
370 375 380

GAG AAT AAA GAT CAG TAC TAT CTG AAA ACA GAG CAG GGT TAT ATC ACG 1200  
Glu Asn Lys Asp Gln Tyr Tyr Leu Lys Thr Glu Gln Gly Tyr Ile Thr  
385 390 395 400

GTA GAT AGT TCC GGA CAG TCA AAT TGG AAA AAC GCG CTG GTT ATC AAT 1248  
Val Asp Ser Ser Gly Gln Ser Asn Trp Lys Asn Ala Leu Val Ile Asn  
405 410 415

GGG ACA AAA GAC AAG GGG CTG TTA TTA ACC TTT TGC AGC GAT AGC TCA 1296  
Gly Thr Lys Asp Lys Gly Leu Leu Thr Phe Cys Ser Asp Ser Ser  
420 425 430

GGC ACT CCG ACA AAC CCT GAT GAT GTG ATT CCT CCC GCT ATC AAT GAT 1344  
Gly Thr Pro Thr Asn Pro Asp Asp Val Ile Pro Pro Ala Ile Asn Asp  
435 440 445

ATT CCA TCG CCG CCA GCC CGC GAA ACA CTG TCA CTG ACG CCG GTC AGT 1392  
Ile Pro Ser Pro Pro Ala Arg Glu Thr Leu Ser Leu Thr Pro Val Ser  
450 455 460

TAT CAA TTG ATG ACC AAT CCG GCA CCG ACA GAA GAT GAT ATT ACC AAC 1440

Tyr Gln Leu Met Thr Asn Pro Ala Pro Thr Glu Asp Asp Ile Thr Asn  
465 470 475 480

CAT TAT GGT TTT AAC GGC GCT AGC TTA CGG GCT TCT CCA TTG TCA ACC 1488  
His Tyr Gly Phe Asn Gly Ala Ser Leu Arg Ala Ser Pro Leu Ser Thr  
485 490 495

AGC GAG TTG ACC AGC AAA CTG AAT TCT ATC GAT ACT TTC TGT GAG AAG 1536  
Ser Glu Leu Thr Ser Lys Leu Asn Ser Ile Asp Thr Phe Cys Glu Lys  
500 505 510

ACC CGG TTA AGC TTC AAT CAG TTA ATG GAT TTG ACC GCT CAG CAA TCT 1584  
Thr Arg Leu Ser Phe Asn Gln Leu Met Asp Leu Thr Ala Gln Gln Ser  
515 520 525

TAC AGT CAA AGC AGC ATT GAT GCG AAA GCA GCC AGC CGC TAT GTT CGT 1632  
Tyr Ser Gln Ser Ser Ile Asp Ala Lys Ala Ala Ser Arg Tyr Val Arg  
530 535 540

TTT GGG GAA ACC ACC CCA ACC CGC GTC AAT GTC TAC GGT GCC GCT TAT 1680  
Phe Gly Glu Thr Thr Pro Thr Arg Val Asn Val Tyr Gly Ala Ala Tyr  
545 550 555 560

CTG AAC AGC ACA CTG GCA GAC GCG GCT GAT GGT CAA TAT CTG TGG ATT 1728  
Leu Asn Ser Thr Leu Ala Asp Ala Ala Asp Gly Gln Tyr Leu Trp Ile  
565 570 575

CAG ACT GAT GGC AAG AGC CTA AAT TTC ACT GAC GAT ACG GTA GTC GCC 1776  
Gln Thr Asp Gly Lys Ser Leu Asn Phe Thr Asp Asp Thr Val Val Ala  
580 585 590

TTA GCC GGT CGC GCT GAA AAG CTG GTA CGT TTA TCA TCC CAG ACC GGG 1824  
Leu Ala Gly Arg Ala Glu Lys Leu Val Arg Leu Ser Ser Gln Thr Gly  
595 600 605

CTA TCA TTT GAA GAA TTG GAC TGG CTG ATT GCC AAT GCC AGT CGT AGT 1872  
Leu Ser Phe Glu Glu Leu Asp Trp Leu Ile Ala Asn Ala Ser Arg Ser  
610 615 620

GTG CCG GAC CAC CAC GAC AAA ATT GTG CTG GAT AAG CCG GTC CTT GAA 1920  
Val Pro Asp His His Asp Lys Ile Val Leu Asp Lys Pro Val Leu Glu  
625 630 635 640

GCA CTG GCA GAG TAT GTC AGC CTA AAA CAG CGC TAT GGG CTT GAT GCC 1968  
Ala Leu Ala Glu Tyr Val Ser Leu Lys Gln Arg Tyr Gly Leu Asp Ala  
645 650 655

AAT ACC TTT GCG ACC TTC ATT AGT GCA GTA AAT CCT TAT ACG CCA GAT 2016  
Asn Thr Phe Ala Thr Phe Ile Ser Ala Val Asn Pro Tyr Thr Pro Asp  
660 665 670

CAG ACA CCC AGT TTC TAT GAA ACC GCT TTC CGC TCT GCC GAC GGT AAT 2064  
Gln Thr Pro Ser Phe Tyr Glu Thr Ala Phe Arg Ser Ala Asp Gly Asn  
675 680 685



CAT His 690	GTC Val	ATT Ile	GCG Ala	CTA Leu	GGT Gly	ACA Thr	GAG Glu	GTG Val	AAA Lys	TAT Tyr	GCA Ala	GAA Glu	AAT Asn	GAG Glu	CAG Gln	2112
GAT Asp 705	GAG Glu	TTA Leu	GCC Ala	GCC Ala	ATA Ile	TGC Cys	TGC Cys	AAA Lys	GCA Ala	TTG Leu	GGT Gly	GTC Val	ACC Thr	AGT Ser	GAT Asp 720	2160
GAA Glu	CTG Leu	CTC Leu	CGT Arg	ATT Ile	GGT Gly	CGC Arg	TAT Tyr	TGC Cys	TTC Phe	GGT Gly	AAT Asn	GCA Ala	GGC Gly	AGT Ser	TTT Phe 735	2208
ACC Thr	TTG Leu	GAT Asp	GAA Glu	TAT Tyr	ACC Thr	GCC Ala	AGT Ser	CAG Gln	TTG Leu	TAT Tyr	CGC Arg	TTC Phe	GGC Gly	GCC Ala	ATT Ile	2256
CCC Pro	CGT Arg	TTG Leu	TTT Phe	GGG Gly	CTG Leu	ACA Thr	TTT Phe	GCC Ala	CAA Gln	GCC Ala	GAA Glu	ATT Ile	TTA Leu	TGG Trp	CGT Arg	2304
CTG Leu	ATG Met	GAA Glu	GGC Gly	GGA Gly	AAA Lys	GAT Asp	ATC Ile	TTA Leu	TTG Leu	CAA Gln	CAG Gln	TTA Leu	GGT Gly	CAG Gln	GCA Ala	2352
AAA Lys 785	TCC Ser	CTG Leu	CAA Gln	CCA Pro	CTG Leu	GCT Ala	ATT Ile	TTA Leu	CGC Arg	CGT Arg	ACC Thr	GAG Glu	CAG Gln	GTG Val	CTG Leu 800	2400
GAT Asp	TGG Trp	ATG Met	TCG Ser	TCC Ser	GTA Val	AAT Asn	CTA Leu	AGT Ser	CTG Leu	ACT Thr	TAT Tyr	CTG Leu	CAA Gln	GGG Gly	ATG Met 815	2448
GTA Val	AGT Ser	ACG Thr	CAA Gln	TGG Trp	AGC Ser	GGT Gly	ACC Thr	GCC Ala	ACC Thr	GCT Ala	GAG Glu	ATG Met	TTC Phe	AAT Asn	TTC Phe	2496
TTG Leu	GAA Glu	AAC Asn	GTT Val	TGT Cys	GAC Asp	AGC Ser	GTG Val	AAT Asn	AGT Ser	CAA Gln	GCT Ala	GCC Ala	ACT Thr	AAA Lys	GAA Glu	2544
ACA Thr	ATG Met	GAT Asp	TCG Ser	GCG Ala	TTA Leu	CAG Gln	CAG Gln	AAA Lys	GTG Val	CTG Leu	CGG Arg	GCG Ala	CTA Leu	AGC Ser	GCC Ala	2592
GGT Gly 865	TTC Phe	GGC Gly	ATT Ile	AAG Lys	AGC Ser	AAT Asn	GTG Val	ATG Met	GGT Gly	ATC Ile	GTC Val	ACC Thr	TTC Phe	TGG Trp	CTG Leu 880	2640
GAG Glu	AAA Lys	ATC Ile	ACA Thr	ATC Ile	GGT Gly	AGT Ser	GAT Asp	AAT Asn	CCT Pro	TTT Phe	ACA Thr	TTG Leu	GCA Ala	AAC Asn	TAC Tyr 895	2688
TGG Trp	CAT His	GAT Asp	ATT Ile	CAA Gln	ACC Thr	CTG Leu	TTT Phe	AGC Ser	CAT His	GAC Asp	AAT Asn	GCC Ala	ACG Thr	TTA Leu	GAG Glu	2736
TCC Ser	TTA Leu	CAA Gln	ACC Thr	GAC Asp	ACT Thr	TCT Ser	CTG Leu	GTA Val	ATT Ile	GCT Ala	ACT Thr	CAG Gln	CAA Gln	CTT Leu	AGC Ser	2784
CAG Gln	CTA Leu	GTG Val	TTA Leu	ATT Ile	GTG Val	AAA Lys	TGG Trp	CTG Leu	AGC Ser	CTG Leu	ACC Thr	GAG Glu	CAG Gln	GAT Asp	CTG Leu	2832
CAA Gln	TTA Leu	CTG Leu	ACA Thr	ACC Thr	TAT Tyr	CCC Pro	GAA Glu	CGT Arg	TTA Leu	ATC Ile	AAC Asn	GGC Gly	ATC Ile	ACG Thr	AAT Asn 960	2880

GTT CCT GTA CCC AAT CCG GAG CTA TTA CTC ACG CTA TCA CGT TTT AAG 2928  
 Val Pro Val Pro Asn Pro Glu Leu Leu Leu Thr Leu Ser Arg Phe Lys  
 965 970 975

CAG TGG GAA ACT CAA GTC ACC GTT TCC CGT GAT GAA GCG ATG CGC TGT 2976  
 Gln Trp Glu Thr Gln Val Thr Val Ser Arg Asp Glu Ala Met Arg Cys  
 980 985 990

TTC GAT CAA TTA AAT GCC AAT GAT ATG ACG ACT GAA AAT GCA GGT TCA 3024  
 Phe Asp Gln Leu Asn Ala Asn Asp Met Thr Thr Glu Asn Ala Gly Ser  
 995 1000 1005

CTG ATC GCC ACA TTG TAT GAG ATG GAT AAA GGT ACG GGA GCG CAA GTT 3072  
 Leu Ile Ala Thr Leu Tyr Glu Met Asp Lys Gly Thr Gly Ala Gln Val  
 1010 1015 1020

AAT ACC TTG CTA TTA GGT GAA AAT AAC TGG CCG AAA AGT TTT ACC TCT 3120  
 Asn Thr Leu Leu Leu Gly Glu Asn Asn Trp Pro Lys Ser Phe Thr Ser  
 1025 1030 1035 1040

CTC TGG CAA CTT CTG ACC TGG TTA CGC GTC GGG CAA AGA CTG AAT GTC 3168  
 Leu Trp Gln Leu Leu Thr Trp Leu Arg Val Gly Gln Arg Leu Asn Val  
 1045 1050 1055

GGT AGT ACC ACT CTG GGC AAT CTG TTG TCC ATG ATG CAA GCA GAC CCT 3216  
 Gly Ser Thr Thr Leu Gly Asn Leu Leu Ser Met Met Gln Ala Asp Pro  
 1060 1065 1070

GCT GCC GAG AGT AGC GCT TTA TTG GCA TCA GTA GCC CAA AAC TTA AGT 3264  
 Ala Ala Glu Ser Ser Ala Leu Leu Ala Ser Val Ala Gln Asn Leu Ser  
 1075 1080 1085

GCC GCA ATC AGC AAT CGT CAG TAA 3288  
 Ala Ala Ile Ser Asn Arg Gln \*\*\*  
 1090 1095

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 34:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĹŽKA: 1095 aminokyselín
- (B) TYP: aminokyselina
- (D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: proteín

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 34 (TcaA proteín):

Charakteristický znak	od	do	Opis
	254	267	SEKV. ID. ČÍS. 15
	254	492	TcaA <sub>ii</sub> peptid

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

Thr Gln Pro Leu Pro Ser Phe His Leu Pro Val Thr Leu Glu His Ser  
370 375 380

Glu Asn Lys Asp Gln Tyr Tyr Leu Lys Thr Glu Gln Gly Tyr Ile Thr  
385 390 395 400

Val Asp Ser Ser Gly Gln Ser Asn Trp Lys Asn Ala Leu Val Ile Asn  
405 410 415

Gly Thr Lys Asp Lys Gly Leu Leu Leu Thr Phe Cys Ser Asp Ser Ser  
420 425 430

Gly Thr Pro Thr Asn Pro Asp Asp Val Ile Pro Pro Ala Ile Asn Asp  
435 440 445

Ile Pro Ser Pro Pro Ala Arg Glu Thr Leu Ser Leu Thr Pro Val Ser  
450 455 460

Tyr Gln Leu Met Thr Asn Pro Ala Pro Thr Glu Asp Asp Ile Thr Asn  
465 470 475 480

His Tyr Gly Phe Asn Gly Ala Ser Leu Arg Ala Ser Pro Leu Ser Thr  
485 490 W4 » 495

Ser Glu Leu Thr Ser Lys Leu Asn Ser Ile Asp Thr Phe Cys Glu Lys  
500 505 510

Thr Arg Leu Ser Phe Asn Gln Leu Met Asp Leu Thr Ala Gln Gln Ser  
515 520 525

Tyr Ser Gln Ser Ser Ile Asp Ala Lys Ala Ala Ser Arg Tyr Val Arg  
530 535 540

Phe Gly Glu Thr Thr Pro Thr Arg Val Asn Val Tyr Gly Ala Ala Tyr  
545 550 555 560

Leu Asn Ser Thr Leu Ala Asp Ala Ala Asp Gly Gln Tyr Leu Trp Ile  
565 570 575

Gln Thr Asp Gly Lys Ser Leu Asn Phe Thr Asp Asp Thr Val Val Ala  
580 585 590

Leu Ala Gly Arg Ala Glu Lys Leu Val Arg Leu Ser Ser Gln Thr Gly  
595 600 605

Leu Ser Phe Glu Glu Leu Asp Trp Leu Ile Ala Asn Ala Ser Arg Ser  
610 615 620

Val Pro Asp His His Asp Lys Ile Val Leu Asp Lys Pro Val Leu Glu  
625 630 635 640

Ala Leu Ala Glu Tyr Val Ser Leu Lys Gln Arg Tyr Gly Leu Asp Ala  
645 650 655

Asn Thr Phe Ala Thr Phe Ile Ser Ala Val Asn Pro Tyr Thr Pro Asp  
660 665 670

Gln Thr Pro Ser Phe Tyr Glu Thr Ala Phe Arg Ser Ala Asp Gly Asn  
675 680 685

His Val Ile Ala Leu Gly Thr Glu Val Lys Tyr Ala Glu Asn Glu Gln  
690 695 700

Asp Glu Leu Ala Ala Ile Cys Cys Lys Ala Leu Gly Val Thr Ser Asp  
705 710 715 720

Glu Leu Leu Arg Ile Gly Arg Tyr Cys Phe Gly Asn Ala Gly Ser Phe  
725 730 735

Thr Leu Asp Glu Tyr Thr Ala Ser Gln Leu Tyr Arg Phe Gly Ala Ile  
740 745 750

Pro Arg Leu Phe Gly Leu Thr Phe Ala Gln Ala Glu Ile Leu Trp Arg  
755 760 765

Leu Met Glu Gly Gly Lys Asp Ile Leu Leu Gln Gln Leu Gly Gln Ala  
770 775 780

Lys Ser Leu Gln Pro Leu Ala Ile Leu Arg Arg Thr Glu Gln Val Leu  
785 790 795 800

Asp Trp Met Ser Ser Val Asn Leu Ser Leu Thr Tyr Leu Gln Gly Met  
805 810 815

Val Ser Thr Gln Trp Ser Gly Thr Ala Thr Ala Glu Met Phe Asn Phe  
820 825 830

Leu Glu Asn Val Cys Asp Ser Val Asn Ser Gln Ala Ala Thr Lys Glu  
835 840 845

Thr Met Asp Ser Ala Leu Gln Gln Lys Val Leu Arg Ala Leu Ser Ala  
850 855 860

Gly Phe Gly Ile Lys Ser Asn Val Met Gly Ile Val Thr Phe Trp Leu  
865 870 875 880

Glu Lys Ile-Thr Ile Gly Ser Asp Asn Pro Phe Thr Leu Ala Asn Tyr  
885 890 895

Trp His Asp Ile Gln Thr Leu Phe Ser His Asp Asn Ala Thr Leu Glu  
900 905 910

Ser Leu Gln Thr Asp Thr Ser Leu Val Ile Ala Thr Gln Gln Leu Ser  
915 920 925

Gln Leu Val Leu Ile Val Lys Trp Leu Ser Leu Thr Glu Gln Asp Leu  
930 935 940

Gln Leu Leu Thr Thr Tyr Pro Glu Arg Leu Ile Asn Gly Ile Thr Asn  
945 950 955 960

Val Pro Val Pro Asn Pro Glu Leu Leu Leu Thr Leu Ser Arg Phe Lys  
965 970 975

Gln Trp Glu Thr Gln Val Thr Val Ser Arg Asp Glu Ala Met Arg Cys  
980 985 990

Phe Asp Gln Leu Asn Ala Asn Asp Met Thr Thr Glu Asn Ala Gly Ser  
995 1000 1005

Leu Ile Ala Thr Leu Tyr Glu Met Asp Lys Gly Thr Gly Ala Gln Val  
1010 1015 1020

Asn Thr Leu Leu Leu Gly Glu Asn Asn Trp Pro Lys Ser Phe Thr Ser  
1025 1030 1035 1040

Leu Trp Gln Leu Leu Thr Trp Leu Arg Val Gly Gln Arg Leu Asn Val  
1045 1050 1055

Gly Ser Thr Thr Leu Gly Asn Leu Leu Ser Met Met Gln Ala Asp Pro  
1060 1065 1070

Ala Ala Glu Ser Ser Ala Leu Leu Ala Ser Val Ala Gln Asn Leu Ser  
1075 1080 1085

Ala Ala Ile Ser Asn Arg Gln ...  
1090 1095

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 35:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĹŽKA: 603 aminokyselín
- (B) TYP: aminokyselina
- (D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: proteín

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 35 (TcaA<sub>iii</sub> proteín):

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

Phe Gly Ala Ile Pro Arg Leu Phe Gly Leu Thr Phe Ala Gln Ala Glu  
260 265 270

Ile Leu Trp Arg Leu Met Glu Gly Gly Lys Asp Ile Leu Leu Gln Gln  
275 280 285

Xxx Gly Gln Ala Lys Ser Leu Gln Pro Leu Ala Ile Leu Arg Arg Thr  
290 295 300

Glu Gln Val Leu Asp Trp Met Ser Pro Val Asn Leu Ser Leu Thr Tyr  
305 310 315 320

Leu Gln Gly Met Val Ser Thr Gln Trp Ser Gly Thr Ala Thr Ala Glu  
325 330 335

Met Phe Asn Phe Leu Glu Asn Val Cys Asp Ser Val Asn Ser Gln Ala  
340 345 350

Xxx Thr Lys Glu Thr Met Asp Ser Ala Leu Gln Gln Lys Val Leu Arg  
355 360 365

Ala Leu Ser Ala Gly Phe Gly Ile Lys Ser Asn Val Met Gly Ile Val  
370 375 380

Thr Phe Trp Leu Glu Lys Ile Thr Ile Gly Arg Asp Asn Pro Phe Thr  
385 390 395 400

Leu Ala Asn Tyr Trp His Asp Ile Gln Thr Leu Phe Ser His Asp Asn  
405 410 415

Ala Thr Leu Glu Ser Leu Gln Thr Asp Thr Ser Leu Val Ile Ala Thr  
420 425 430

Gln Gln Leu Ser Gln Leu Val Leu Ile Val Lys Trp Val Ser Leu Thr  
435 440 445

Glu Gln Asp Leu Gln Leu Leu Thr Thr Tyr Pro Glu Arg Leu Ile Asn  
450 455 460

Gly Ile Thr Asn Val Pro Val Pro Asn Pro Glu Leu Leu Leu Thr Leu  
465 470 475 480

Ser Arg Phe Lys Gln Trp Glu Thr Gln Val Thr Val Ser Arg Asp Glu  
485 490 495

Ala Met Arg Cys Phe Asp Gln Leu Asn Ala Asn Asp Met Thr Thr Glu  
500 505 510

Asn Ala Gly Ser Leu Ile Ala Thr Leu Tyr Glu Met Asp Lys Gly Thr  
515 520 525

Gly Ala Gln Val Asn Thr Leu Leu Leu Gly Glu Asn Asn Trp Pro Lys  
530 535 540

Ser Phe Thr Ser Leu Trp Gln Leu Leu Thr Trp Leu Arg Val Gly Gln  
545 550 555 560

Arg Leu Asn Val Gly Ser Thr Thr Leu Gly Asn Leu Leu Ser Met Met  
565 570 575

Gln Ala Asp Pro Ala Ala Glu Ser Ser Ala Leu Leu Ala Ser Val Ala  
580 585 590

Gln Asn Leu Ser Ala Ala Ile Ser Asn Arg Gln \*  
595 600

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 36:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĚLKA: 2557 párov báz
- (B) TYP: nukleová kyselina
- (D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: DNA (genómová)

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 36 (tcdA interný fragment):

GAATTCGGCT	TGCGTTTAAT	ATTGATGATG	TCTCGCTCTT	CCGCCTGCTT	AAAATTACCG	60
ACCATGATAA	TAAAGATGGA	AAAATTAATA	ATAACCTAAA	GAATCTTTCC	AATTTATATA	120
TTGGAAAATT	ACTGGCAGAT	ATTCATCAAT	TAACCATTGA	TGAACTGGAT	TTATTACTGA	180
TTGCCGTAGG	TGAAGGAAAA	ACTAATTTAT	CCGCTATCAG	TGATAAGCAA	TTGGCTACCC	240
TGATCAGAAA	ACTCAATACT	ATTACCAGCT	GGCTACATAC	ACAGAAGTGG	AGTGTATTCC	300
AGCTATTTAT	CATGACCTCC	ACCAGCTATA	ACAAAACGCT	AACGCCTGAA	ATTAAGAATT	360
TGCTGGATAC	CGTCTACCAC	GGTTTACAAG	GTTTTGATAA	AGACAAAGCA	GATTTGCTAC	420
ATGTCATGGC	GCCCTATAT	GCGGCCACCT	TGCAATTATC	ATCGGAAAAT	GTCGCCCACT	480
CGGTACTCCT	TTGGGCAGAT	AAGTTACAGC	CCGGCGACGG	CGCAATGACA	GCAGAGGGAN	540
TCTGGGACTG	GTTGAATACT	AAGTATACGC	CGGGTTCATC	GGAAGCCGTA	GAAACGCAGG	600
AACATATCGT	TCAGTATTGT	CAGGCTCTGG	CACAATTGGA	AATGGTTTAC	CATTCCACCG	660
GCATCAACGA	AAACGCCTTC	CGTCTATTTG	TGACAAAACC	AGAGATGTTT	GGCGCTGCAA	720
CTGGAGCAGC	GCCCCGCGAT	GATGCCCTTT	CACTGATTAT	GCTGACACGT	TTTGCGGATT	780
GGGTGAACGC	ACTAGGCGAA	AAAGCGTCTT	CGGTGCTAGC	GGCATTTGAA	GCTAACTCGT	840
TAACGGCAGA	ACAACCTGGCT	GATGCCATGA	ATCTTGATGC	TAATTTGCTG	TTGCAAGCCA	900
GTATTCAAGC	ACAAAATCAT	CAACATCTTC	CCCCAGTAAC	TCCAGAAAAT	GCGTTCCTCT	960
GTTGGACATC	TATCAATACT	ATCCTGCAAT	GGGTTAATGT	CGCACAAACA	TTGAAATGTC	1020
GCCCCACAGG	GCGTTTCCGC	TTTGGTCGGG	CTGGATTATA	TTCAATCAAT	GAAAGAGACA	1080
CCGACCTATG	CCCAGTGGGA	AAACGCGGCA	GGCGTATTAA	CCGCCGGGTT	GAATTCAACA	1140
ACAGGCTAAT	ACATTACAAC	GCTTTTCTGG	ATGAATCTCG	CAGTGCCGCA	TTAAGCACCT	1200
ACTATATCCG	TCAAGTCGCC	AAGGCAGCGG	CGGCTATTAA	AAGCCGTGAT	GACTTGTATC	1260
AATACTTACT	GATTGATAAT	CAGGTTTCTG	CGGCAATAAA	AACCACCCGG	ATCGCCGAAG	1320
CCATTGCCAG	TATTCAACTG	TACGTCAACC	GGGCATTGGA	AAATGTGGAA	GAAAATGCCA	1380
ATTCCGGGGT	TATCAGCCGC	CAATTCTTTA	TCGACTGGGA	CAAATACAAT	AAACGCTACA	1440
GCACTTGGGC	GGGTGTTTCT	CAATTAGTTT	ACTACCCGGA	AACTATATT	GATCCGACCA	1500
TGCGTATCGG	ACAAACCAAA	ATGATGGACG	CATTACTGCA	ATCCGTCAGC	CAAAGCCAAT	1560
TAAACGCCGA	TACCGTCGAA	GATGCCTTTA	TGTCTTATCT	GACATCGTTT	GAACAAGTGG	1620
CTAATCTTAA	AGTTATTAGC	GCATATCACG	ATAATATTAA	TAACGATCAA	GGGCTGACCT	1680
ATTTTATCGG	ACTCAGTGAA	ACTGATGCCG	GTGAATATTA	TTGGCGCAGT	GTCGATCACA	1740
GTAAATTCAA	CGACGGTAAA	TTCGCGGCTA	ATGCCTGGAG	TGAATGGCAT	AAAATTGATT	1800



GTCCAATTAA CCCTTATAAA AGCACTATCC GTCCAGTGAT ATATAAATCC CGCCTGTATC 1860  
 TGCTCTGGTT GGAACAAAAG GAGATCACCA AACAGACAGG AAATAGTAAA GATGGCTATC 1920  
 AAAGTAAAAC GGATTATCGT TATGAACTAA AATTGGCGCA TATCCGCTAT GATGGCACTT 1980  
 GGAATACGCC AATCACCTTT GATGTCAATA AAAAAATATC CGAGCTAAAA CTGAAAAAAA 2040  
 5 ATAGAGCGCC CGGACTCTAT TGTGCCGGTT ATCAAGGTGA AGATACGTTG CTGGTGATGT 2100  
 TTTATAACCA ACAAGACACA CTAGATAGTT ATAAAAACGC TTCAATGCAA GGACTATATA 2160  
 TCTTTGCTGA TATGGCATCC AAAGATATGA CCCCAGAACA GAGCAATGTT TATCGGGATA 2220  
 ATAGCTATCA ACAATTTGAT ACCAATAATG TCAGAAGAGT GAATAACCGC TATGCAGAGG 2280  
 ATTATGAGAT TCCTTCTTCG GTAAGTAGCC GTAAAGACTA TGGTTGGGGA GATTATTACC 2340  
 10 TCAGCATGGT ATATAACGGA GATATTCCAA CTATCAATTA CAAAGCCGCA TCAAGTGATT 2400  
 TAAAAATTTA TATTTACCA AAATTAAGAA TTATTCATAA TGGATATGAA GGACAGAAGC 2460  
 GCAATCAATG CAATTTGATG AATAAATATG GCAAACCTAGG TGATAAATTT ATTGTGTATA 2520  
 CCAGCCTGGG CGTTAATCCG AATAATAAGC CGAATTC 2557

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 37:

(i) CHARAKTERISTIKA SEKVENCIE:

(A) DĹŽKA: 845 aminokyselín

(B) TYP: aminokyselina

(D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: proteín (čiastkový)

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 37 (TcdA interný peptid):

Ala	Phe	Asn	Ile	Asp	Asp	Val	Ser	Leu	Phe	Arg	Leu	Leu	Lys	Ile	Thr
1				5					10					15	
Asp	His	Asp	Asn	Lys	Asp	Gly	Lys	Ile	Lys	Asn	Asn	Leu	Lys	Asn	Leu
			20					25					30		
Ser	Asn	Leu	Tyr	Ile	Gly	Lys	Leu	Leu	Ala	Asp	Ile	His	Gln	Leu	Thr
		35					40					45			
Ile	Asp	Glu	Leu	Asp	Leu	Leu	Leu	Ile	Ala	Val	Gly	Glu	Gly	Lys	Thr
	50					55					60				
Asn	Leu	Ser	Ala	Ile	Ser	Asp	Lys	Gln	Leu	Ala	Thr	Leu	Ile	Arg	Lys
65					70					75					80
Leu	Asn	Thr	Ile	Thr	Ser	Trp	Leu	His	Thr	Gln	Lys	Trp	Ser	Val	Phe
				85					90					95	
Gln	Leu	Phe	Ile	Met	Thr	Ser	Thr	Ser	Tyr	Asn	Lys	Thr	Leu	Thr	Pro
				100				105						110	
Glu	Ile	Lys	Asn	Leu	Leu	Asp	Thr	Val	Tyr	His	Gly	Leu	Gln	Gly	Phe
		115					120					125			
Asp	Lys	Asp	Lys	Ala	Asp	Leu	Leu	His	Val	Met	Ala	Pro	Tyr	Ile	Ala
	130					135					140				

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

Trp Ala Asp Lys Leu Gln Pro Gly Asp Gly Ala Met Thr Ala Glu Gly  
165 170 175

Phe Trp Asp Trp Leu Asn Thr Lys Tyr Thr Pro Gly Ser Ser Glu Ala  
180 185 190

Val Glu Thr Gln Glu His Ile Val Gln Tyr Cys Gln Ala Leu Ala Gln  
195 200 205

Leu Glu Met Val Tyr His Ser Thr Gly Ile Asn Glu Asn Ala Phe Arg  
210 215 220

Leu Phe Val Thr Lys Pro Glu Met Phe Gly Ala Ala Thr Gly Ala Ala  
225 230 235 240

Pro Ala His Asp Ala Leu Ser Leu Ile Met Leu Thr Arg Phe Ala Asp  
245 250 255

Trp Val Asn Ala Leu Gly Glu Lys Ala Ser Ser Val Leu Ala Ala Phe  
260 265 270

Glu Ala Asn Ser Leu Thr Ala Glu Gln Leu Ala Asp Ala Met Asn Leu  
275 280 285

Asp Ala Asn Leu Leu Leu Gln Ala Ser Ile Gln Ala Gln Asn His Gln  
290 295 300

His Leu Pro Pro Val Thr Pro Glu Asn Ala Phe Ser Cys Trp Thr Ser  
305 310 315 320

Ile Asn Thr Ile Leu Gln Trp Val Asn Val Ala Gln Gln Leu Lys Cys  
325 330 335

Arg Pro Thr Gly Arg Phe Arg Phe Gly Arg Ala Gly Leu Tyr Ser Ile  
340 345 350

Asn Glu Arg Asp Thr Asp Leu Cys Pro Val Gly Lys Arg Gly Arg Arg  
355 360 365

Ile Asn Arg Arg Val Glu Phe Asn Asn Arg Leu Ile His Tyr Asn Ala  
370 375 380

Phe Leu Asp Glu Ser Arg Ser Ala Ala Leu Ser Thr Tyr Tyr Ile Arg  
385 390 395 400

Gln Val Ala Lys Ala Ala Ala Ala Ile Lys Ser Arg Asp Asp Leu Tyr  
405 410 415

Gln Tyr Leu Leu Ile Asp Asn Gln Val Ser Ala Ala Ile Lys Thr Thr  
420 425 430

Arg Ile Ala Glu Ala Ile Ala Ser Ile Gln Leu Tyr Val Asn Arg Ala  
435 440 445

Leu Glu Asn Val Glu Glu Asn Ala Asn Ser Gly Val Ile Ser Arg Gln  
450 455 460

Phe Phe Ile Asp Trp Asp Lys Tyr Asn Lys Arg Tyr Ser Thr Trp Ala  
465 470 475 480

Gly Val Ser Gln Leu Val Tyr Tyr Pro Glu Asn Tyr Ile Asp Pro Thr  
485 490 495

Met Arg Ile Gly Gln Thr Lys Met Met Asp Ala Leu Leu Gln Ser Val  
500 505 510

Ser Gln Ser Gln Leu Asn Ala Asp Thr Val Glu Asp Ala Phe Met Ser  
515 520 525

Tyr Leu Thr Ser Phe Glu Gln Val Ala Asn Leu Lys Val Ile Ser Ala  
530 535 540

Tyr His Asp Asn Ile Asn Asn Asp Gln Gly Leu Thr Tyr Phe Ile Gly  
545 550 555 560

Leu Ser Glu Thr Asp Ala Gly Glu Tyr Tyr Trp Arg Ser Val Asp His  
565 570 575

Ser Lys Phe Asn Asp Gly Lys Phe Ala Ala Asn Ala Trp Ser Glu Trp  
580 585 590

His Lys Ile Asp Cys Pro Ile Asn Pro Tyr Lys Ser Thr Ile Arg Pro  
595 600 605

Val Ile Tyr Lys Ser Arg Leu Tyr Leu Leu Trp Leu Glu Gln Lys Glu  
610 615 620

Ile Thr Lys Gln Thr Gly Asn Ser Lys Asp Gly Tyr Gln Thr Glu Thr  
625 630 635 640

Asp Tyr Arg Tyr Glu Leu Lys Leu Ala His Ile Arg Tyr Asp Gly Thr  
645 650 655

Trp Asn Thr Pro Ile Thr Phe Asp Val Asn Lys Lys Ile Ser Glu Leu  
660 665 670

Lys Leu Glu Lys Asn Arg Ala Pro Gly Leu Tyr Cys Ala Gly Tyr Gln  
675 680 685

Gly Glu Asp Thr Leu Leu Val Met Phe Tyr Asn Gln Gln Asp Thr Leu  
690 695 700

Asp Ser Tyr Lys Asn Ala Ser Met Gln Gly Leu Tyr Ile Phe Ala Asp  
705 710 715 720

Met Ala Ser Lys Asp Met Thr Pro Glu Gln Ser Asn Val Tyr Arg Asp  
725 730 735

Asn Ser Tyr Gln Gln Phe Asp Thr Asn Asn Val Arg Arg Val Asn Asn  
740 745 750

Arg Tyr Ala Glu Asp Tyr Glu Ile Pro Ser Ser Val Ser Ser Arg Lys  
755 760 765

Asp Tyr Gly Trp Gly Asp Tyr Tyr Leu Ser Met Val Tyr Asn Gly Asp  
770 775 780

Ile Pro Thr Ile Asn Tyr Lys Ala Ala Ser Ser Asp Leu Lys Ile Tyr  
785 790 795 800

Ile Ser Pro Lys Leu Arg Ile Ile His Asn Gly Tyr Glu Gly Gln Lys  
805 810 815

Arg Asn Gln Cys Asn Leu Met Asn Lys Tyr Gly Lys Leu Gly Asp Lys  
820 825 830

Phe Ile Val Tyr Thr Ser Leu Gly Val Asn Pro Asn Asn  
835 840 845

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 38:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĚLKA: 16 aminokyselín
- (B) TYP: aminokyselina
- (C) REŤAZEC: jednoduchý
- (D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: proteín

(v) TYP FRAGMENTU: N-koncový

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 38 (TcdA<sub>ii</sub> - pk71  
interný peptid):

Arg	Tyr	Tyr	Asn	Leu	Ser	Asp	Glu	Glu	Leu	Ser	Gln	Phe	Ile	Gly
1			5						10					15

Lys

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 39:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĚLKA: 20 aminokyselín
- (B) TYP: aminokyselina
- (C) REŤAZEC: jednoduchý
- (D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: proteín

(v) TYP FRAGMENTU: N-koncový

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 39 (TcdA<sub>ii</sub> - pk44  
interný peptid):

Gly Thr Ala Thr Asp Val Ser Gly Pro Val Glu Ile Asn Thr Ala

1	5	10	15
---	---	----	----

Ile Ser Pro Ala Lys  
20

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 40:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĹŽKA: 11 aminokyselín
- (B) TYP: aminokyselina
- (C) REĹAZEC: jednoduchý
- (D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: proteín

(v) TYP FRAGMENTU: N-koncový

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 40 (TcbA<sub>iii</sub> N-koniec):

Ala	Asn	Ser	Leu	Thr	Ala	Leu	Phe	Leu	Pro	Gln
1			5					10		

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 41:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĹŽKA: 14 aminokyselín
- (B) TYP: aminokyselina
- (C) REĹAZEC: jednoduchý
- (D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: proteín

(v) TYP FRAGMENTU: N-koncový

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 41 (TcdA<sub>iii</sub> N-koniec):

Leu	Arg	Ser	Ala	Asn	Thr	Leu	Thr	Asp	Leu	Phe	Leu	Pro	Gln
1				5					10				

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 42:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĚLKA: 19 aminokyselín
- (B) TYP: aminokyselina
- (C) REŤAZEC: jednoduchý
- (D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: proteín

(v) TYP FRAGMENTU: N-koncový

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 42 (TcdA - pk57  
interný peptid):

Arg	Ala	Leu	Glu	Val	Glu	Arg	Thr	Val	Ser	Leu	Ala	Glu	Val	Tyr
1				5					10					15

Ala Gly Leu Glu

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 43:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĚLKA: 11 aminokyselín
- (B) TYP: aminokyselina
- (C) REŤAZEC: jednoduchý
- (D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: proteín

(v) TYP FRAGMENTU: N-koncový

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 43 (TcdA<sub>iii</sub> - pk20  
interný peptid):

Ile	Arg	Glu	Asp	Tyr	Pro	Ala	Ser	Leu	Gly	Lys
1				5					10	

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 44:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĚŽKA: 16 aminokyselín
- (B) TYP: aminokyselina
- (C) REĚAZEC: jednoduchý
- (D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: proteín

(v) TYP FRAGMENTU: N-koncový

(xi) OPIS SEKVENCIE: SEKV. ID. ĆÍS. 44:

Asp	Asp	Ser	Gly	Asp	Asp	Asp	Lys	Val	Thr	Asn	Thr	Asp	Ile	His
1				5					10					15

Arg

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ĆÍS. 45:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĚŽKA: 13 aminokyselín
- (B) TYP: aminokyselina
- (C) REĚAZEC: jednoduchý
- (D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: proteín

(v) TYP FRAGMENTU: N-koncový

(xi) OPIS SEKVENCIE: SEKV. ID. ĆÍS. 45:

Asp	Val	Xaa	Gly	Ser	Glu	Lys	Ala	Asn	Glu	Lys	Leu	Lys
1				5					10			

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ĆÍS. 46:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĚŽKA: 7551 párov báz
- (B) TYP: nukleová kyselina
- (C) REĚAZEC: dvojité
- (D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY:DNA (genómová)

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 46 (tcdA):

ATG	AAC	GAG	TCT	GTA	AAA	GAG	ATA	CCT	GAT	GTA	TTA	AAA	AGC	CAG	TGT	48
Met	Asn	Glu	Ser	Val	Lys	Glu	Ile	Pro	Asp	Val	Leu	Lys	Ser	Gln	Cys	
1				5					10					15		
GGT	TTT	AAT	TGT	CTG	ACA	GAT	ATT	AGC	CAC	AGC	TCT	TTT	AAT	GAA	TTT	96
Gly	Phe	Asn	Cys	Leu	Thr	Asp	Ile	Ser	His	Ser	Ser	Phe	Asn	Glu	Phe	
			20					25					30			
CGC	CAG	CAA	GTA	TCT	GAG	CAC	CTC	TCC	TGG	TCC	GAA	ACA	CAC	GAC	TTA	144
Arg	Gln	Gln	Val	Ser	Glu	His	Leu	Ser	Trp	Ser	Glu	Thr	His	Asp	Leu	
		35					40					45				
TAT	CAT	GAT	GCA	CAA	CAG	GCA	CAA	AAG	GAT	AAT	CGC	CTG	TAT	GAA	GCG	192
Tyr	His	Asp	Ala	Gln	Gln	Ala	Gln	Lys	Asp	Asn	Arg	Leu	Tyr	Glu	Ala	
	50					55					60					
CGT	ATT	CTC	AAA	CGC	GCC	AAT	CCC	CAA	TTA	CAA	AAT	GCG	GTG	CAT	CTT	240
Arg	Ile	Leu	Lys	Arg	Ala	Asn	Pro	Gln	Leu	Gln	Asn	Ala	Val	His	Leu	
65					70					75					80	
GCC	ATT	CTC	GCT	CCC	AAT	GCT	GAA	CTG	ATA	GGC	TAT	AAC	AAT	CAA	TTT	288
Ala	Ile	Leu	Ala	Pro	Asn	Ala	Glu	Leu	Ile	Gly	Tyr	Asn	Asn	Gln	Phe	
				85					90					95		
AGC	GGT	AGA	GCC	AGT	CAA	TAT	GTT	GCG	CCG	GGT	ACC	GTT	TCT	TCC	ATG	336
Ser	Gly	Arg	Ala	Ser	Gln	Tyr	Val	Ala	Pro	Gly	Thr	Val	Ser	Ser	Met	
			100					105					110			
TTC	TCC	CCC	GCC	GCT	TAT	TTG	ACT	GAA	CTT	TAT	CGT	GAA	GCA	CGC	AAT	384
Phe	Ser	Pro	Ala	Ala	Tyr	Leu	Thr	Glu	Leu	Tyr	Arg	Glu	Ala	Arg	Asn	
		115					120					125				
TTA	CAC	GCA	AGT	GAC	TCC	GTT	TAT	TAT	CTG	GAT	ACC	CGC	CGC	CCA	GAT	432
Leu	His	Ala	Ser	Asp	Ser	Val	Tyr	Tyr	Leu	Asp	Thr	Arg	Arg	Pro	Asp	
	130					135					140					
CTC	AAA	TCA	ATG	GCG	CTC	AGT	CAG	CAA	AAT	ATG	GAT	ATA	GAA	TTA	TCC	480
Leu	Lys	Ser	Met	Ala	Leu	Ser	Gln	Gln	Asn	Met	Asp	Ile	Glu	Leu	Ser	
145					150					155					160	
ACA	CTC	TCT	TTG	TCC	AAT	GAG	CTG	TTA	TTG	GAA	AGC	ATT	AAA	ACT	GAA	528
Thr	Leu	Ser	Leu	Ser	Asn	Glu	Leu	Leu	Leu	Glu	Ser	Ile	Lys	Thr	Glu	
				165					170					175		
TCT	AAA	CTG	GAA	AAC	TAT	ACT	AAA	GTG	ATG	GAA	ATG	CTC	TCC	ACT	TTC	576
Ser	Lys	Leu	Glu	Asn	Tyr	Thr	Lys	Val	Met	Glu	Met	Leu	Ser	Thr	Phe	
			180					185					190			
CGT	CCT	TCC	GGC	GCA	ACG	CCT	TAT	CAT	GAT	GCT	TAT	GAA	AAT	GTG	CGT	624
Arg	Pro	Ser	Gly	Ala	Thr	Pro	Tyr	His	Asp	Ala	Tyr	Glu	Asn	Val	Arg	
		195					200					205				
GAA	GTT	ATC	CAG	CTA	CAA	GAT	CCT	GGA	CTT	GAG	CAA	CTC	AAT	GCA	TCA	672
Glu	Val	Ile	Gln	Leu	Gln	Asp	Pro	Gly	Leu	Glu	Gln	Leu	Asn	Ala	Ser	



210	215	220	
CCG GCA ATT GCC GGG TTG ATG CAT CAA GCC TCC CTA TTG GGT ATT AAC 720 Pro Ala Ile Ala Gly Leu Met His Gln Ala Ser Leu Leu Gly Ile Asn 225 230 235 240			
GCT TCA ATC TCG CCT GAG CTA TTT AAT ATT CTG ACG GAG GAG ATT ACC 768 Ala Ser Ile Ser Pro Glu Leu Phe Asn Ile Leu Thr Glu Glu Ile Thr 245 250 255			
GAA GGT AAT GCT GAG GAA CTT TAT AAG AAA AAT TTT GGT AAT ATC GAA 816 Glu Gly Asn Ala Glu Glu Leu Tyr Lys Lys Asn Phe Gly Asn Ile Glu 260 265 270			
CCG GCC TCA TTG GCT ATG CCG GAA TAC CTT AAA CGT TAT TAT AAT TTA 864 Pro Ala Ser Leu Ala Met Pro Glu Tyr Leu Lys Arg Tyr Tyr Asn Leu 275 280 285			
AGC GAT GAA GAA CTT AGT CAG TTT ATT GGT AAA GCC AGC AAT TTT GGT 912 Ser Asp Glu Glu Leu Ser Gln Phe Ile Gly Lys Ala Ser Asn Phe Gly 290 295 300			
CAA CAG GAA TAT AGT AAT AAC CAA CTT ATT ACT CCG GTA GTC AAC AGC 960 Gln Gln Glu Tyr Ser Asn Asn Gln Leu Ile Thr Pro Val Val Asn Ser 305 310 315 320			
AGT GAT GGC ACG GTT AAG GTA TAT CGG ATC ACC CGC GAA TAT ACA ACC 1008 Ser Asp Gly Thr Val Lys Val Tyr Arg Ile Thr Arg Glu Tyr Thr Thr 325 330 335			
AAT GCT TAT CAA ATG GAT GTG GAG CTA TTT CCC TTC GGT GGT GAG AAT 1056 Asn Ala Tyr Gln Met Asp Val Glu Leu Phe Pro Phe Gly Gly Glu Asn 340 345 350			
TAT CGG TTA GAT TAT AAA TTC AAA AAT TTT TAT AAT GCC TCT TAT TTA 1104 Tyr Arg Leu Asp Tyr Lys Phe Lys Asn Phe Tyr Asn Ala Ser Tyr Leu 355 360 365			
TCC ATC AAG TTA AAT GAT AAA AGA GAA CTT GTT CGA ACT GAA GGC GCT 1152 Ser Ile Lys Leu Asn Asp Lys Arg Glu Leu Val Arg Thr Glu Gly Ala 370 375 380			
CCT CAA GTC AAT ATA GAA TAC TCC GCA AAT ATC ACA TTA AAT ACC GCT 1200 Pro Gln Val Asn Ile Glu Tyr Ser Ala Asn Ile Thr Leu Asn Thr Ala 385 390 395 400			
GAT ATC AGT CAA CCT TTT GAA ATT GGC CTG ACA CGA GTA CTT CCT TCC 1248 Asp Ile Ser Gln Pro Phe Glu Ile Gly Leu Thr Arg Val Leu Pro Ser 405 410 415			
GGT TCT TGG GCA TAT GCC GCC GCA AAA TTT ACC GTT GAA GAG TAT AAC 1296 Gly Ser Trp Ala Tyr Ala Ala Ala Lys Phe Thr Val Glu Glu Tyr Asn 420 425 430			
CAA TAC TCT TTT CTG CTA AAA CTT AAC AAG GCT ATT CGT CTA TCA CGT 1344 Gln Tyr Ser Phe Leu Leu Lys Leu Asn Lys Ala Ile Arg Leu Ser Arg 435 440 445			
GCG ACA GAA TTG TCA CCC ACG ATT CTG GAA GGC ATT GTG CGC AGT GTT 1392 Ala Thr Glu Leu Ser Pro Thr Ile Leu Glu Gly Ile Val Arg Ser Val 450 455 460			
AAT CTA CAA CTG GAT ATC AAC ACA GAC GTA TTA GGT AAA GTT TTT CTG 1440 Asn Leu Gln Leu Asp Ile Asn Thr Asp Val Leu Gly Lys Val Phe Leu 465 470 475 480			
ACT AAA TAT TAT ATG CAG CGT TAT GCT ATT CAT GCT GAA ACT GCC CTG 1488 Thr Lys Tyr Tyr Met Gln Arg Tyr Ala Ile His Ala Glu Thr Ala Leu 485 490 495			
ATA CTA TGC AAC GCG CCT ATT TCA CAA CGT TCA TAT GAT AAT CAA CCT 1536			

Ile	Leu	Cys	Asn	Ala	Pro	Ile	Ser	Gln	Arg	Ser	Tyr	Asp	Asn	Gln	Pro	
			500					505					510			
AGC	CAA	TTT	GAT	CGC	CTG	TTT	AAT	ACG	CCA	TTA	CTG	AAC	GGA	CAA	TAT	1584
Ser	Gln	Phe	Asp	Arg	Leu	Phe	Asn	Thr	Pro	Leu	Leu	Asn	Gly	Gln	Tyr	
		515					520					525				
TTT	TCT	ACC	GGC	GAT	GAG	GAG	ATT	GAT	TTA	AAT	TCA	GGT	AGC	ACC	GGC	1632
Phe	Ser	Thr	Gly	Asp	Glu	Glu	Ile	Asp	Leu	Asn	Ser	Gly	Ser	Thr	Gly	
	530					535					540					
GAT	TGG	CGA	AAA	ACC	ATA	CTT	AAG	CGT	GCA	TTT	AAT	ATT	GAT	GAT	GTC	1680
Asp	Trp	Arg	Lys	Thr	Ile	Leu	Lys	Arg	Ala	Phe	Asn	Ile	Asp	Asp	Val	
545					550					555					560	
TCG	CTC	TTC	CGC	CTG	CTT	AAA	ATT	ACC	GAC	CAT	GAT	AAT	AAA	GAT	GGA	1728
Ser	Leu	Phe	Arg	Leu	Leu	Lys	Ile	Thr	Asp	His	Asp	Asn	Lys	Asp	Gly	
				565					570					575		
AAA	ATT	AAA	AAT	AAC	CTA	AAG	AAT	CTT	TCC	AAT	TTA	TAT	ATT	GGA	AAA	1776
Lys	Ile	Lys	Asn	Asn	Leu	Lys	Asn	Leu	Ser	Asn	Leu	Tyr	Ile	Gly	Lys	
			580					585					590			
TTA	CTG	GCA	GAT	ATT	CAT	CAA	TTA	ACC	ATT	GAT	GAA	CTG	GAT	TTA	TTA	1824
Leu	Leu	Ala	Asp	Ile	His	Gln	Leu	Thr	Ile	Asp	Glu	Leu	Asp	Leu	Leu	
		595				600						605				
CTG	ATT	GCC	GTA	GGT	GAA	GGA	AAA	ACT	AAT	TTA	TCC	GCT	ATC	AGT	GAT	1872
Leu	Ile	Ala	Val	Gly	Glu	Gly	Lys	Thr	Asn	Leu	Ser	Ala	Ile	Ser	Asp	
	610					615					620					
AAG	CAA	TTG	GCT	ACC	CTG	ATC	AGA	AAA	CTC	AAT	ACT	ATT	ACC	AGC	TGG	1920
Lys	Gln	Leu	Ala	Thr	Leu	Ile	Arg	Lys	Leu	Asn	Thr	Ile	Thr	Ser	Trp	
625					630					635					640	
CTA	CAT	ACA	CAG	AAG	TGG	AGT	GTA	TTC	CAG	CTA	TTT	ATC	ATG	ACC	TCC	1968
Leu	His	Thr	Gln	Lys	Trp	Ser	Val	Phe	Gln	Leu	Phe	Ile	Met	Thr	Ser	
			645					650					655			
ACC	AGC	TAT	AAC	AAA	ACG	CTA	ACG	CCT	GAA	ATT	AAG	AAT	TTG	CTG	GAT	2016
Thr	Ser	Tyr	Asn	Lys	Thr	Leu	Thr	Pro	Glu	Ile	Lys	Asn	Leu	Leu	Asp	
			660					665					670			
ACC	GTC	TAC	CAC	GGT	TTA	CAA	GGT	TTT	GAT	AAA	GAC	AAA	GCA	GAT	TTG	2064
Thr	Val	Tyr	His	Gly	Leu	Gln	Gly	Phe	Asp	Lys	Asp	Lys	Ala	Asp	Leu	
		675					680					685				
CTA	CAT	GTC	ATG	GCG	CCC	TAT	ATT	GCG	GCC	ACC	TTG	CAA	TTA	TCA	TCG	2112
Leu	His	Val	Met	Ala	Pro	Tyr	Ile	Ala	Ala	Thr	Leu	Gln	Leu	Ser	Ser	
		690				695					700					
GAA	AAT	GTC	GCC	CAC	TCG	GTA	CTC	CTT	TGG	GCA	GAT	AAG	TTA	CAG	CCC	2160
Glu	Asn	Val	Ala	His	Ser	Val	Leu	Leu	Trp	Ala	Asp	Lys	Leu	Gln	Pro	
705					710					715					720	
GGC	GAC	GGC	GCA	ATG	ACA	GCA	GAA	AAA	TTC	TGG	GAC	TGG	TTG	AAT	ACT	2208
Gly	Asp	Gly	Ala	Met	Thr	Ala	Glu	Lys	Phe	Trp	Asp	Trp	Leu	Asn	Thr	
				725					730					735		
AAG	TAT	ACG	CCG	GGT	TCA	TCG	GAA	GCC	GTA	GAA	ACG	CAG	GAA	CAT	ATC	2256
Lys	Tyr	Thr	Pro	Gly	Ser	Ser	Glu	Ala	Val	Glu	Thr	Gln	Glu	His	Ile	
			740					745					750			
GTT	CAG	TAT	TGT	CAG	GCT	CTG	GCA	CAA	TTG	GAA	ATG	GTT	TAC	CAT	TCC	2304
Val	Gln	Tyr	Cys	Gln	Ala	Leu	Ala	Gln	Leu	Glu	Met	Val	Tyr	His	Ser	
		755				760						765				
ACC	GGC	ATC	AAC	GAA	AAC	GCC	TTC	CGT	CTA	TTT	GTG	ACA	AAA	CCA	GAG	2352
Thr	Gly	Ile	Asn	Glu	Asn	Ala	Phe	Arg	Leu	Phe	Val	Thr	Lys	Pro	Glu	
	770					775					780					

ATG Met 785	TTT Phe	GGC Gly	GCT Ala	GCA Ala	ACT Thr	GGA Gly	GCA Ala	GCG Ala	CCC Pro	GCG Ala	CAT His	GAT Asp	GCC Ala	CTT Leu	TCA Ser	2400
CTG Leu	ATT Ile	ATG Met	CTG Leu	ACA Thr	CGT Arg	TTT Phe	GCG Ala	GAT Asp	TGG Trp	GTG Val	AAC Asn	GCA Ala	CTA Leu	GGC Gly	GAA Glu	2448
AAA Lys	GCG Ala	TCC Ser	TCG Ser	GTG Val	CTA Leu	GCG Ala	GCA Ala	TTT Phe	GAA Glu	GCT Ala	AAC Asn	TCG Ser	TTA Leu	ACG Thr	GCA Ala	2496
GAA Glu	CAA Gln	CTG Leu	GCT Ala	GAT Asp	GCC Ala	ATG Met	AAT Asn	CTT Leu	GAT Asp	GCT Ala	AAT Asn	TTG Leu	CTG Leu	TTG Leu	CAA Gln	2544
GCC Ala	AGT Ser	ATT Ile	CAA Gln	GCA Ala	CAA Gln	AAT Asn	CAT His	CAA Gln	CAT His	CTT Leu	CCC Pro	CCA Pro	GTA Val	ACT Thr	CCA Pro	2592
GAA Glu	AAT Asn	GCG Ala	TTC Phe	TCC Ser	TGT Cys	TGG Trp	ACA Thr	TCT Ser	ATC Ile	AAT Asn	ACT Thr	ATC Ile	CTG Leu	CAA Gln	TGG Trp	2640
GTT Val	AAT Asn	GTC Val	GCA Ala	CAA Gln	CAA Gln	TTG Leu	AAT Asn	GTC Val	GCC Ala	CCA Pro	CAG Gln	GGC Gly	GTT Val	TCC Ser	GCT Ala	2688
TTG Leu	GTC Val	GGG Gly	CTG Leu	GAT Asp	TAT Tyr	ATT Ile	CAA Gln	TCA Ser	ATG Met	AAA Lys	GAG Glu	ACA Thr	CCG Pro	ACC Thr	TAT Tyr	2736
GCC Ala	CAG Gln	TGG Trp	GAA Glu	AAC Asn	GCG Ala	GCA Ala	GGC Gly	GTA Val	TTA Leu	ACC Thr	GCC Ala	GGG Gly	TTG Leu	AAT Asn	TCA Ser	2784
CAA Gln	CAG Gln	GCT Ala	AAT Asn	ACA Thr	TTA Leu	CAC His	GCT Ala	TTT Phe	CTG Leu	GAT Asp	GAA Glu	TCT Ser	CGC Arg	AGT Ser	GCC Ala	2832
GCA Ala	TTA Leu	AGC Ser	ACC Thr	TAC Tyr	TAT Tyr	ATC Ile	CGT Arg	CAA Gln	GTC Val	GCC Ala	AAG Lys	GCA Ala	GCG Ala	GCG Ala	GCT Ala	2880
ATT Ile	AAA Lys	AGC Ser	CGT Arg	GAT Asp	GAC Asp	TTG Leu	TAT Tyr	CAA Gln	TAC Tyr	TTA Leu	CTG Leu	ATT Ile	GAT Asp	AAT Asn	CAG Gln	2928
GTT Val	TCT Ser	GCG Ala	GCA Ala	ATA Ile	AAA Lys	ACC Thr	ACC Thr	CGG Arg	ATC Ile	GCC Ala	GAA Glu	GCC Ala	ATT Ile	GCC Ala	AGT Ser	2976
ATT Ile	CAA Gln	CTG Leu	TAC Tyr	GTC Val	AAC Asn	CGG Arg	GCA Ala	TTG Leu	GAA Glu	AAT Asn	GTG Val	GAA Glu	GAA Glu	AAT Asn	GCC Ala	3024
AAT Asn	TCG Ser	GGG Gly	GTT Val	ATC Ile	AGC Ser	CGC Arg	CAA Gln	TTC Phe	TTT Phe	ATC Ile	GAC Asp	TGG Trp	GAC Asp	AAA Lys	TAC Tyr	3072
AAT Asn	AAA Lys	CGC Arg	TAC Tyr	AGC Ser	ACT Thr	TGG Trp	GCG Ala	GGT Gly	GTT Val	TCT Ser	CAA Gln	TTA Leu	GTT Val	TAC Tyr	TAC Tyr	3120
CCG Pro	GAA Glu	AAC Asn	TAT Tyr	ATT Ile	GAT Asp	CCG Pro	ACC Thr	ATG Met	CGT Arg	ATC Ile	GGA Gly	CAA Gln	ACC Thr	AAA Lys	ATG Met	3168
ATG Met	GAC Asp	GCA Ala	TTA Leu	CTG Leu	CAA Gln	TCC Ser	GTC Val	AGC Ser	CAA Gln	AGC Ser	CAA Gln	TTA Leu	AAC Asn	GCC Ala	GAT Asp	3216

ACC	GTC	GAA	GAT	GCC	TTT	ATG	TCT	TAT	CTG	ACA	TCG	TTT	GAA	CAA	GTG	3264
Thr	Val	Glu	Asp	Ala	Phe	Met	Ser	Tyr	Leu	Thr	Ser	Phe	Glu	Gln	Val	
		1075						1080				1085				
GCT	AAT	CTT	AAA	GTT	ATT	AGC	GCA	TAT	CAC	GAT	AAT	ATT	AAT	AAC	GAT	3312
Ala	Asn	Leu	Lys	Val	Ile	Ser	Ala	Tyr	His	Asp	Asn	Ile	Asn	Asn	Asp	
	1090					1095					1100					
CAA	GGG	CTG	ACC	TAT	TTT	ATC	GGA	CTC	AGT	GAA	ACT	GAT	GCC	GGT	GAA	3360
Gln	Gly	Leu	Thr	Tyr	Phe	Ile	Gly	Leu	Ser	Glu	Thr	Asp	Ala	Gly	Glu	
	1105				1110					1115					1120	
TAT	TAT	TGG	CGC	AGT	GTC	GAT	CAC	AGT	AAA	TTC	AAC	GAC	GGT	AAA	TTC	3408
Tyr	Tyr	Trp	Arg	Ser	Val	Asp	His	Ser	Lys	Phe	Asn	Asp	Gly	Lys	Phe	
				1125					1130					1135		
GCG	GCT	AAT	GCC	TGG	AGT	GAA	TGG	CAT	AAA	ATT	GAT	TGT	CCA	ATT	AAC	3456
Ala	Ala	Asn	Ala	Trp	Ser	Glu	Trp	His	Lys	Ile	Asp	Cys	Pro	Ile	Asn	
			1140					1145					1150			
CCT	TAT	AAA	AGC	ACT	ATC	CGT	CCA	GTG	ATA	TAT	AAA	TCC	CGC	CTG	TAT	3504
Pro	Tyr	Lys	Ser	Thr	Ile	Arg	Pro	Val	Ile	Tyr	Lys	Ser	Arg	Leu	Tyr	
		1155					1160					1165				
CTG	CTC	TGG	TTG	GAA	CAA	AAG	GAG	ATC	ACC	AAA	CAG	ACA	GGA	AAT	AGT	3552
Leu	Leu	Trp	Leu	Glu	Gln	Lys	Glu	Ile	Thr	Lys	Gln	Thr	Gly	Asn	Ser	
	1170					1175					1180					
AAA	GAT	GGC	TAT	CAA	ACT	GAA	ACG	GAT	TAT	CGT	TAT	GAA	CTA	AAA	TTG	3600
Lys	Asp	Gly	Tyr	Gln	Thr	Glu	Thr	Asp	Tyr	Arg	Tyr	Glu	Leu	Lys	Leu	
	1185				1190					1195					1200	
GCG	CAT	ATC	CGC	TAT	GAT	GGC	ACT	TGG	AAT	ACG	CCA	ATC	ACC	TTT	GAT	3648
Ala	His	Ile	Arg	Tyr	Asp	Gly	Thr	Trp	Asn	Thr	Pro	Ile	Thr	Phe	Asp	
				1205					1210					1215		
GTC	AAT	AAA	AAA	ATA	TCC	GAG	CTA	AAA	CTG	GAA	AAA	AAT	AGA	GCG	CCC	3696
Val	Asn	Lys	Lys	Ile	Ser	Glu	Leu	Lys	Leu	Glu	Lys	Asn	Arg	Ala	Pro	
			1220					1225					1230			
GGA	CTC	TAT	TGT	GCC	GGT	TAT	CAA	GGT	GAA	GAT	ACG	TTG	CTG	GTG	ATG	3744
Gly	Leu	Tyr	Cys	Ala	Gly	Tyr	Gln	Gly	Glu	Asp	Thr	Leu	Leu	Val	Met	
		1235					1240					1245				
TTT	TAT	AAC	CAA	CAA	GAC	ACA	CTA	GAT	AGT	TAT	AAA	AAC	GCT	TCA	ATG	3792
Phe	Tyr	Asn	Gln	Gln	Asp	Thr	Leu	Asp	Ser	Tyr	Lys	Asn	Ala	Ser	Met	
	1250					1255					1260					
CAA	GGA	CTA	TAT	ATC	TTT	GCT	GAT	ATG	GCA	TCC	AAA	GAT	ATG	ACC	CCA	3840
Gln	Gly	Leu	Tyr	Ile	Phe	Ala	Asp	Met	Ala	Ser	Lys	Asp	Met	Thr	Pro	
	1265				1270					1275					1280	
GAA	CAG	AGC	AAT	GTT	TAT	CGG	GAT	AAT	AGC	TAT	CAA	CAA	TTT	GAT	ACC	3888
Glu	Gln	Ser	Asn	Val	Tyr	Arg	Asp	Asn	Ser	Tyr	Gln	Gln	Phe	Asp	Thr	
				1285					1290					1295		
AAT	AAT	GTC	AGA	AGA	GTG	AAT	AAC	CGC	TAT	GCA	GAG	GAT	TAT	GAG	ATT	3936
Asn	Asn	Val	Arg	Arg	Val	Asn	Asn	Arg	Tyr	Ala	Glu	Asp	Tyr	Glu	Ile	
			1300					1305					1310			
CCT	TCC	TCG	GTA	AGT	AGC	CGT	AAA	GAC	TAT	GGT	TGG	GGA	GAT	TAT	TAC	3984
Pro	Ser	Ser	Val	Ser	Ser	Arg	Lys	Asp	Tyr	Gly	Trp	Gly	Asp	Tyr	Tyr	
		1315					1320					1325				
CTC	AGC	ATG	GTA	TAT	AAC	GGA	GAT	ATT	CCA	ACT	ATC	AAT	TAC	AAA	GCC	4032
Leu	Ser	Met	Val	Tyr	Asn	Gly	Asp	Ile	Pro	Thr	Ile	Asn	Tyr	Lys	Ala	
	1330					1335					1340					
GCA	TCA	AGT	GAT	TTA	AAA	ATC	TAT	ATC	TCA	CCA	AAA	TTA	AGA	ATT	ATT	4080
Ala	Ser	Ser	Asp	Leu	Lys	Ile	Tyr	Ile	Ser	Pro	Lys	Leu	Arg	Ile	Ile	

1345				1350					1355				1360				
CAT	AAT	GGA	TAT	GAA	GGA	CAG	AAG	CGC	AAT	CAA	TGC	AAT	CTG	ATG	AAT	4128	
His	Asn	Gly	Tyr	Glu	Gly	Gln	Lys	Arg	Asn	Gln	Cys	Asn	Leu	Met	Asn		
				1365					1370				1375				
AAA	TAT	GGC	AAA	CTA	GGT	GAT	AAA	TTT	ATT	GTT	TAT	ACT	AGC	TTG	GGG	4176	
Lys	Tyr	Gly	Lys	Leu	Gly	Asp	Lys	Phe	Ile	Val	Tyr	Thr	Ser	Leu	Gly		
			1380					1385					1390				
GTC	AAT	CCA	AAT	AAC	TCG	TCA	AAT	AAG	CTC	ATG	TTT	TAC	CCC	GTC	TAT	4224	
Val	Asn	Pro	Asn	Asn	Ser	Ser	Asn	Lys	Leu	Met	Phe	Tyr	Pro	Val	Tyr		
		1395					1400					1405					
CAA	TAT	AGC	GGA	AAC	ACC	AGT	GGA	CTC	AAT	CAA	GGG	AGA	CTA	CTA	TTC	4272	
Gln	Tyr	Ser	Gly	Asn	Thr	Ser	Gly	Leu	Asn	Gln	Gly	Arg	Leu	Leu	Phe		
	1410					1415					1420						
CAC	CGT	GAC	ACC	ACT	TAT	CCA	TCT	AAA	GTA	GAA	GCT	TGG	ATT	CCT	GGA	4320	
His	Arg	Asp	Thr	Thr	Tyr	Pro	Ser	Lys	Val	Glu	Ala	Trp	Ile	Pro	Gly		
	1425				1430					1435					1440		
GCA	AAA	CGT	TCT	CTA	ACC	AAC	CAA	AAT	GCC	GCC	ATT	GGT	GAT	GAT	TAT	4368	
Ala	Lys	Arg	Ser	Leu	Thr	Asn	Gln	Asn	Ala	Ala	Ile	Gly	Asp	Asp	Tyr		
				1445					1450					1455			
GCT	ACA	GAC	TCT	CTG	AAT	AAA	CCG	GAT	GAT	CTT	AAG	CAA	TAT	ATC	TTT	4416	
Ala	Thr	Asp	Ser	Leu	Asn	Lys	Pro	Asp	Asp	Leu	Lys	Gln	Tyr	Ile	Phe		
			1460					1465					1470				
ATG	ACT	GAC	AGT	AAA	GGG	ACT	GCT	ACT	GAT	GTC	TCA	GGC	CCA	GTA	GAG	4464	
Met	Thr	Asp	Ser	Lys	Gly	Thr	Ala	Thr	Asp	Val	Ser	Gly	Pro	Val	Glu		
		1475					1480					1485					
ATT	AAT	ACT	GCA	ATT	TCT	CCA	GCA	AAA	GTT	CAG	ATA	ATA	GTC	AAA	GCG	4512	
Ile	Asn	Thr	Ala	Ile	Ser	Pro	Ala	Lys	Val	Gln	Ile	Ile	Val	Lys	Ala		
	1490					1495					1500						
GGT	GGC	AAG	GAG	CAA	ACT	TTT	ACC	GCA	GAT	AAA	GAT	GTC	TCC	ATT	CAG	4560	
Gly	Gly	Lys	Glu	Gln	Thr	Phe	Thr	Ala	Asp	Lys	Asp	Val	Ser	Ile	Gln		
	1505				1510					1515					1520		
CCA	TCA	CCT	AGC	TTT	GAT	GAA	ATG	AAT	TAT	CAA	TTT	AAT	GCC	CTT	GAA	4608	
Pro	Ser	Pro	Ser	Phe	Asp	Glu	Met	Asn	Tyr	Gln	Phe	Asn	Ala	Leu	Glu		
			1525						1530					1535			
ATA	GAC	GGT	TCT	GGT	CTG	AAT	TTT	ATT	AAC	AAC	TCA	GCC	AGT	ATT	GAT	4656	
Ile	Asp	Gly	Ser	Gly	Leu	Asn	Phe	Ile	Asn	Asn	Ser	Ala	Ser	Ile	Asp		
			1540					1545					1550				
GTT	ACT	TTT	ACC	GCA	TTT	GCG	GAG	GAT	GGC	CGC	AAA	CTG	GGT	TAT	GAA	4704	
Val	Thr	Phe	Thr	Ala	Phe	Ala	Glu	Asp	Gly	Arg	Lys	Leu	Gly	Tyr	Glu		
		1555					1560					1565					
AGT	TTC	AGT	ATT	CCT	GTT	ACC	CTC	AAG	GTA	AGT	ACC	GAT	AAT	GCC	CTG	4752	
Ser	Phe	Ser	Ile	Pro	Val	Thr	Leu	Lys	Val	Ser	Thr	Asp	Asn	Ala	Leu		
	1570					1575					1580						
ACC	CTG	CAC	CAT	AAT	GAA	AAT	GGT	GCG	CAA	TAT	ATG	CAA	TGG	CAA	TCC	4800	
Thr	Leu	His	His	Asn	Glu	Asn	Gly	Ala	Gln	Tyr	Met	Gln	Trp	Gln	Ser		
	1585				1590				1595						1600		
TAT	CGT	ACC	CGC	CTG	AAT	ACT	CTA	TTT	GCC	CGC	CAG	TTG	GTT	GCA	CGC	4848	
Tyr	Arg	Thr	Arg	Leu	Asn	Thr	Leu	Phe	Ala	Arg	Gln	Leu	Val	Ala	Arg		
			1605					1610						1615			
GCC	ACC	ACC	GGA	ATC	GAT	ACA	ATT	CTG	AGT	ATG	GAA	ACT	CAG	AAT	ATT	4896	
Ala	Thr	Thr	Gly	Ile	Asp	Thr	Ile	Leu	Ser	Met	Glu	Thr	Gln	Asn	Ile		
			1620					1625					1630				
CAG	GAA	CCG	CAG	TTA	GGC	AAA	GGT	TTC	TAT	GCT	ACG	TTC	GTG	ATA	CCT	4944	

Gln	Glu	Pro	Gln	Leu	Gly	Lys	Gly	Phe	Tyr	Ala	Thr	Phe	Val	Ile	Pro	
		1635					1640					1645				
CCC	TAT	AAC	CTA	TCA	ACT	CAT	GGT	GAT	GAA	CGT	TGG	TTT	AAG	CTT	TAT	4992
Pro	Tyr	Asn	Leu	Ser	Thr	His	Gly	Asp	Glu	Arg	Trp	Phe	Lys	Leu	Tyr	
	1650				1655					1660						
ATC	AAA	CAT	GTT	GTT	GAT	AAT	AAT	TCA	CAT	ATT	ATC	TAT	TCA	GGC	CAG	5040
Ile	Lys	His	Val	Val	Asp	Asn	Asn	Ser	His	Ile	Ile	Tyr	Ser	Gly	Gln	
1665				1670					1675						1680	
CTA	ACA	GAT	ACA	AAT	ATA	AAC	ATC	ACA	TTA	TTT	ATT	CCT	CTT	GAT	GAT	5088
Leu	Thr	Asp	Thr	Asn	Ile	Asn	Ile	Thr	Leu	Phe	Ile	Pro	Leu	Asp	Asp	
				1685					1690					1695		
GTC	CCA	TTG	AAT	CAA	GAT	TAT	CAC	GCC	AAG	GTT	TAT	ATG	ACC	TTC	AAG	5136
Val	Pro	Leu	Asn	Gln	Asp	Tyr	His	Ala	Lys	Val	Tyr	Met	Thr	Phe	Lys	
			1700					1705					1710			
AAA	TCA	CCA	TCA	GAT	GGT	ACC	TGG	TGG	GGC	CCT	CAC	TTT	GTT	AGA	GAT	5184
Lys	Ser	Pro	Ser	Asp	Gly	Thr	Trp	Trp	Gly	Pro	His	Phe	Val	Arg	Asp	
		1715					1720					1725				
GAT	AAA	GGA	ATA	GTA	ACA	ATA	AAC	CCT	AAA	TCC	ATT	TTG	ACC	CAT	TTT	5232
Asp	Lys	Gly	Ile	Val	Thr	Ile	Asn	Pro	Lys	Ser	Ile	Leu	Thr	His	Phe	
	1730					1735					1740					
GAG	AGC	GTC	AAT	GTC	CTG	AAT	AAT	ATT	AGT	AGC	GAA	CCA	ATG	GAT	TTC	5280
Glu	Ser	Val	Asn	Val	Leu	Asn	Asn	Ile	Ser	Ser	Glu	Pro	Met	Asp	Phe	
1745					1750					1755					1760	
AGC	GGC	GCT	AAC	AGC	CTC	TAT	TTC	TGG	GAA	CTG	TTC	TAC	TAT	ACC	CCG	5328
Ser	Gly	Ala	Asn	Ser	Leu	Tyr	Phe	Trp	Glu	Leu	Phe	Tyr	Tyr	Thr	Pro	
			1765						1770					1775		
ATG	CTG	GTT	GCT	CAA	CGT	TTG	CTG	CAT	GAA	CAG	AAC	TTC	GAT	GAA	GCC	5376
Met	Leu	Val	Ala	Gln	Arg	Leu	Leu	His	Glu	Gln	Asn	Phe	Asp	Glu	Ala	
			1780					1785					1790			
AAC	CGT	TGG	CTG	AAA	TAT	GTC	TGG	AGT	CCA	TCC	GGT	TAT	ATT	GTC	CAC	5424
Asn	Arg	Trp	Leu	Lys	Tyr	Val	Trp	Ser	Pro	Ser	Gly	Tyr	Ile	Val	His	
		1795					1800					1805				
GGC	CAG	ATT	CAG	AAC	TAC	CAG	TGG	AAC	GTC	CGC	CCG	TTA	CTG	GAA	GAC	5472
Gly	Gln	Ile	Gln	Asn	Tyr	Gln	Trp	Asn	Val	Arg	Pro	Leu	Leu	Glu	Asp	
	1810					1815					1820					
ACC	AGT	TGG	AAC	AGT	GAT	CCT	TTG	GAT	TCC	GTC	GAT	CCT	GAC	GCG	GTA	5520
Thr	Ser	Trp	Asn	Ser	Asp	Pro	Leu	Asp	Ser	Val	Asp	Pro	Asp	Ala	Val	
1825					1830					1835					1840	
GCA	CAG	CAC	GAT	CCA	ATG	CAC	TAC	AAA	GTT	TCA	ACT	TTT	ATG	CGT	ACC	5568
Ala	Gln	His	Asp	Pro	Met	His	Tyr	Lys	Val	Ser	Thr	Phe	Met	Arg	Thr	
				1845					1850					1855		
TTG	GAT	CTA	TTG	ATA	GCA	CGC	GGC	GAC	CAT	GCT	TAT	CGC	CAA	CTG	GAA	5616
Leu	Asp	Leu	Leu	Ile	Ala	Arg	Gly	Asp	His	Ala	Tyr	Arg	Gln	Leu	Glu	
			1860					1865					1870			
CGA	GAT	ACA	CTC	AAC	GAA	GCG	AAG	ATG	TGG	TAT	ATG	CAA	GCG	CTG	CAT	5664
Arg	Asp	Thr	Leu	Asn	Glu	Ala	Lys	Met	Trp	Tyr	Met	Gln	Ala	Leu	His	
		1875					1880					1885				
CTA	TTA	GGT	GAC	AAA	CCT	TAT	CTA	CCG	CTG	AGT	ACG	ACA	TGG	AGT	GAT	5712
Leu	Leu	Gly	Asp	Lys	Pro	Tyr	Leu	Pro	Leu	Ser	Thr	Thr	Trp	Ser	Asp	
	1890					1895					1900					
CGA	CGA	CTA	GAC	AGA	GCC	GCG	GAT	ATC	ACT	ACC	CAA	AAT	GCT	CAC	GAC	5760
Pro	Arg	Leu	Asp	Arg	Ala	Ala	Asp	Ile	Thr	Thr	Gln	Asn	Ala	His	Asp	
	1905				1910					1915					1920	

AGC	GCA	ATA	GTC	GCT	CTG	CGG	CAG	AAT	ATA	CCT	ACA	CCG	GCA	CCT	TTA	5808
Ser	Ala	Ile	Val	Ala	Leu	Arg	Gln	Asn	Ile	Pro	Thr	Pro	Ala	Pro	Leu	
				1925					1930					1935		
TCA	TTG	CGC	AGC	GCT	AAT	ACC	CTG	ACT	GAT	CTC	TTC	CTG	CCG	CAA	ATC	5856
Ser	Leu	Arg	Ser	Ala	Asn	Thr	Leu	Thr	Asp	Leu	Phe	Leu	Pro	Gln	Ile	
			1940					1945					1950			
AAT	GAA	GTG	ATG	ATG	AAT	TAC	TGG	CAG	ACA	TTA	GCT	CAG	AGA	GTA	TAC	5904
Asn	Glu	Val	Met	Met	Asn	Tyr	Trp	Gln	Thr	Leu	Ala	Gln	Arg	Val	Tyr	
		1955					1960					1965				
AAT	CTG	CGT	CAT	AAC	CTC	TCT	ATC	GAC	GGC	CAG	CCG	TTA	TAT	CTG	CCA	5952
Asn	Leu	Arg	His	Asn	Leu	Ser	Ile	Asp	Gly	Gln	Pro	Leu	Tyr	Leu	Pro	
	1970					1975					1980					
ATC	TAT	GCC	ACA	CCG	GCC	GAT	CCG	AAA	GCG	TTA	CTC	AGC	GCC	GCC	GTT	6000
Ile	Tyr	Ala	Thr	Pro	Ala	Asp	Pro	Lys	Ala	Leu	Leu	Ser	Ala	Ala	Val	
	1985				1990						1995				2000	
GCC	ACT	TCT	CAA	GGT	GGA	GGC	AAG	CTA	CCG	GAA	TCA	TTT	ATG	TCC	CTG	6048
Ala	Thr	Ser	Gln	Gly	Gly	Gly	Lys	Leu	Pro	Glu	Ser	Phe	Met	Ser	Leu	
			2005						2010					2015		
TGG	CGT	TTC	CCG	CAC	ATG	CTG	GAA	AAT	GCG	CGC	GGC	ATG	GTT	AGC	CAG	6096
Trp	Arg	Phe	Pro	His	Met	Leu	Glu	Asn	Ala	Arg	Gly	Met	Val	Ser	Glr.	
			2020					2025					2030			
CTC	ACC	CAG	TTC	GGC	TCC	ACG	TTA	CAA	AAT	ATT	ATC	GAA	CGT	CAG	GAC	6144
Leu	Thr	Gln	Phe	Gly	Ser	Thr	Leu	Gln	Asn	Ile	Ile	Glu	Arg	Gln	Asp	
		2035					2040					2045				
GCG	GAA	GCG	CTC	AAT	GCG	TTA	TTA	CAA	AAT	CAG	GCC	GCC	GAG	CTG	ATA	6192
Ala	Glu	Ala	Leu	Asn	Ala	Leu	Leu	Gln	Asn	Gln	Ala	Ala	Glu	Leu	Ile	
	2050					2055					2060					
TTG	ACT	AAC	CTG	AGC	ATT	CAG	GAC	AAA	ACC	ATT	GAA	GAA	TTG	GAT	GCC	6240
Leu	Thr	Asn	Leu	Ser	Ile	Gln	Asp	Lys	Thr	Ile	Glu	Glu	Leu	Asp	Ala	
	2065				2070					2075				2080		
GAG	AAA	ACG	GTG	TTG	GAA	AAA	TCC	AAA	GCG	GGA	GCA	CAA	TCG	CGC	TTT	6288
Glu	Lys	Thr	Val	Leu	Glu	Lys	Ser	Lys	Ala	Gly	Ala	Gln	Ser	Arg	Phe	
			2085					2090						2095		
GAT	AGC	TAC	GGC	AAA	CTG	TAC	GAT	GAG	AAT	ATC	AAC	GCC	GGT	GAA	AAC	6336
Asp	Ser	Tyr	Gly	Lys	Leu	Tyr	Asp	Glu	Asn	Ile	Asn	Ala	Gly	Glu	Asn	
			2100					2105					2110			
CAA	GCC	ATG	ACG	CTA	CGA	GCG	TCC	GCC	GCC	GGG	CTT	ACC	ACG	GCA	GTT	6384
Gln	Ala	Met	Thr	Leu	Arg	Ala	Ser	Ala	Ala	Gly	Leu	Thr	Thr	Ala	Val	
		2115					2120					2125				
CAG	GCA	TCC	CGT	CTG	GCC	GGT	GCG	GCG	GCT	GAT	CTG	GTG	CCT	AAC	ATC	6432
Gln	Ala	Ser	Arg	Leu	Ala	Gly	Ala	Ala	Ala	Asp	Leu	Val	Pro	Asn	Ile	
	2130					2135					2140					
TTC	GGC	TTT	GCC	GGT	GGC	GGC	AGC	CGT	TGG	GGG	GCT	ATC	GCT	GAG	GCC	6480
Phe	Gly	Phe	Ala	Gly	Gly	Gly	Ser	Arg	Trp	Gly	Ala	Ile	Ala	Glu	Ala	
	2145				2150					2155				2160		
ACA	GGT	TAT	GTG	ATG	GAA	TTC	TCC	GCG	AAT	GTT	ATG	AAC	ACC	GAA	GCG	6528
Thr	Gly	Tyr	Val	Met	Glu	Phe	Ser	Ala	Asn	Val	Met	Asn	Thr	Glu	Ala	
			2165					2170					2175			
GAT	AAA	ATT	AGC	CAA	TCT	GAA	ACC	TAC	CGT	CGT	CGC	CGT	CAG	GAG	TGG	6576
Asp	Lys	Ile	Ser	Gln	Ser	Glu	Thr	Tyr	Arg	Arg	Arg	Arg	Gln	Glu	Trp	
			2180					2185					2190			
GAG	ATC	CAG	CGG	AAT	AAT	GCC	GAA	GCS	GAA	TTG	AAG	CAA	ATC	GAT	GCT	6624
Glu	Ile	Gln	Arg	Asn	Asn	Ala	Glu	Ala	Glu	Leu	Lys	Gln	Ile	Asp	Ala	
		2195				2200						2205				

CAG Gln	CTC Leu	AAA Lys	TCA Ser	CTC Leu	GCT Ala	GTA Val	CGC Arg	CGC Arg	GAA Glu	GCC Ala	GCC Ala	GTA Val	TTG Leu	CAG Gln	AAA Lys	6672
	2210					2215				2220						
ACC Thr	AGT Ser	CTG Leu	AAA Lys	ACC Thr	CAA Gln	CAA Gln	GAA Glu	CAG Gln	ACC Thr	CAA Gln	TCT Ser	CAA Gln	TTG Leu	GCC Ala	TTC Phe	6720
	2225				2230					2235					2240	
CTG Leu	CAA Gln	CGT Arg	AAG Lys	TTC Phe	AGC Ser	AAT Asn	CAG Gln	GCG Ala	TTA Leu	TAC Tyr	AAC Asn	TGG Trp	CTG Leu	CGT Arg	GGT Gly	6768
				2245					2250						2255	
CGA Arg	CTG Leu	GCG Ala	GCG Ala	ATT Ile	TAC Tyr	TTC Phe	CAG Gln	TTC Phe	TAC Tyr	GAT Asp	TTG Leu	GCC Ala	GTC Val	GCG Ala	CGT Arg	6816
			2260					2265					2270			
TGC Cys	CTG Leu	ATG Met	GCA Ala	GAA Glu	CAA Gln	GCT Ala	TAC Tyr	CGT Arg	TGG Trp	GAA Glu	CTC Leu	AAT Asn	GAT Asp	GAC Asp	TCT Ser	6864
		2275					2280					2285				
GCC Ala	CGC Arg	TTC Phe	ATT Ile	AAA Lys	CCG Pro	GGC Gly	GCC Ala	TGG Trp	CAG Gln	GGA Gly	ACC Thr	TAT Tyr	GCC Ala	GGT Gly	CTG Leu	6912
	2290					2295					2300					
CTT Leu	GCA Ala	GGT Gly	GAA Glu	ACC Thr	TTG Leu	ATG Met	CTG Leu	AGT Ser	CTG Leu	GCA Ala	CAA Gln	ATG Met	GAA Glu	GAC Asp	GCT Ala	6960
	2305				2310					2315					2320	
CAT His	CTG Leu	AAA Lys	CGC Arg	GAT Asp	AAA Lys	CGC Arg	GCA Ala	TTA Leu	GAG Glu	GTT Val	GAA Glu	CGC Arg	ACA Thr	GTA Val	TCG Ser	7008
				2325					2330						2335	
CTG Leu	GCC Ala	GAA Glu	GTT Val	TAT Tyr	GCA Ala	GGA Gly	TTA Leu	CCA Pro	AAA Lys	GAT Asp	AAC Asn	GGT Gly	CCA Pro	TTT Phe	TCC Ser	7056
			2340					2345					2350			
CTG Leu	GCT Ala	CAG Gln	GAA Glu	ATT Ile	GAC Asp	AAG Lys	CTG Leu	GTG Val	AGT Ser	CAA Gln	GGT Gly	TCA Ser	GGC Gly	AGT Ser	GCC Ala	7104
		2355				2360						2365				
GGC Gly	AGT Ser	GGT Gly	AAT Asn	AAT Asn	AAT Asn	TTG Leu	GCG Ala	TTC Phe	GGC Gly	GCC Ala	GGC Gly	ACG Thr	GAC Asp	ACT Thr	AAA Lys	7152
	2370					2375					2380					
ACC Thr	TCT Ser	TTG Leu	CAG Gln	GCA Ala	TCA Ser	GTT Val	TCA Ser	TTC Phe	GCT Ala	GAT Asp	TTG Leu	AAA Lys	ATT Ile	CGT Arg	GAA Glu	7200
	2385				2390					2395					2400	
GAT Asp	TAC Tyr	CCG Pro	GCA Ala	TCG Ser	CTT Leu	GGC Gly	AAA Lys	ATT Ile	CGA Arg	CGT Arg	ATC Ile	AAA Lys	CAG Gln	ATC Ile	AGC Ser	7248
			2405						2410					2415		
GTC Val	ACT Thr	TTG Leu	CCC Pro	GCG Ala	CTA Leu	CTG Leu	GGA Gly	CCG Pro	TAT Tyr	CAG Gln	GAT Asp	GTA Val	CAG Gln	GCA Ala	ATA Ile	7296
			2420					2425					2430			
TTG Leu	TCT Ser	TAC Tyr	GGC Gly	GAT Asp	AAA Lys	GCC Ala	GGA Gly	TTA Leu	GCT Ala	AAC Asn	GGC Gly	TGT Cys	GAA Glu	GCG Ala	CTG Leu	7344
		2435				2440						2445				
GCA Ala	GTT Val	TCT Ser	CAC His	GGT Gly	ATG Met	AAT Asn	GAC Asp	AGC Ser	GGC Gly	CAA Gln	TTC Phe	CAG Gln	CTC Leu	GAT Asp	TTC Phe	7392
	2450				2455						2460					
AAC Asn	GAT Asp	GGC Gly	AAA Lys	TTC Phe	CTG Leu	CCA Pro	TTC Phe	GAA Glu	GGC Gly	ATC Ile	GCC Ala	ATT Ile	GAT Asp	CAA Gln	GGC Gly	7440
	2465				2470				2475						2480	
ACG Thr	CTG Leu	ACA Thr	CTG Leu	AGC Ser	TTC Phe	CCA Pro	AAT Asn	GCA Ala	TCT Ser	ATG Met	CCG Pro	GAG Glu	AAA Lys	GGT Gly	AAA Lys	7488



	2485		2490		2495	
CAA GCC ACT ATG TTA AAA ACC CTG AAC GAT ATC ATT TTG CAT ATT CGC						7536
Gln Ala Thr Met Leu Lys Thr Leu Asn Asp Ile Ile Leu His Ile Arg						
	2500		2505		2510	
TAC ACC ATT AAA TAA		7551				
Tyr Thr Ile Lys ...						
	2516					

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 47:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĹŽKA: 2516 aminokyselín
- (B) TYP: aminokyselina
- (C) REĹAZEC: jednoduchý
- (D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: proteín

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 47 (TcdA):

Charakt. znak	od	do	Opis
Peptid	1	2516	TcdA proteíny
Peptid	89	1937	TcdA <sub>ii</sub> peptid
Fragment	89	100	TcdA <sub>ii</sub> N-koniec (SEKV. ID. ČÍS. 13)
Fragment	284	299	(SEKV. ID. ČÍS. 38)
Fragment	554	563	(SEKV. ID. ČÍS. 17)
Fragment	1080	1092	(SEKV. ID. ČÍS. 23;12/13)
Fragment	1385	1400	(SEKV. ID. ČÍS. 18)
Fragment	1478	1497	(SEKV. ID. ČÍS. 39)
Fragment	1620	1642	(SEKV. ID. ČÍS. 21;19/23)
Fragment	1938	1948	(SEKV. ID. ČÍS. 41)
Peptid	1938	2516	TcdA <sub>iii</sub> peptid
Fragment	2327	2345	(SEKV. ID. ČÍS. 42)
Fragment	2398	2408	(SEKV. ID. ČÍS. 43)

Met Asn Glu Ser Val Lys Glu Ile Pro Asp Val Leu Lys Ser Gln Cys  
1 5 10 15  
Gly Phe Asn Cys Leu Thr Asp Ile Ser His Ser Ser Phe Asn Glu Phe  
20 25 30  
Arg Gln Gln Val Ser Glu His Leu Ser Trp Ser Glu Thr His Asp Leu  
35 40 45  
Tyr His Asp Ala Gln Gln Ala Gln Lys Asp Asn Arg Leu Tyr Glu Ala  
50 55 60  
Arg Ile Leu Lys Arg Ala Asn Pro Gln Leu Gln Asn Ala Val His Leu  
65 70 75 80  
Ala Ile Leu Ala Pro Asn Ala Glu Leu Ile Gly Tyr Asn Asn Gln Phe  
85 90 95  
Ser Gly Arg Ala Ser Gln Tyr Val Ala Pro Gly Thr Val Ser Ser Met  
100 105 110  
Phe Ser Pro Ala Ala Tyr Leu Thr Glu Leu Tyr Arg Glu Ala Arg Asn  
115 120 125  
Leu His Ala Ser Asp Ser Val Tyr Tyr Leu Asp Thr Arg Arg Pro Asp  
130 135 140  
Leu Lys Ser Met Ala Leu Ser Gln Gln Asn Met Asp Ile Glu Leu Ser  
145 150 155 160  
Thr Leu Ser Leu Ser Asn Glu Leu Leu Leu Glu Ser Ile Lys Thr Glu  
165 170 175  
Ser Lys Leu Glu Asn Tyr Thr Lys Val Met Glu Met Leu Ser Thr Phe  
180 185 190  
Arg Pro Ser Gly Ala Thr Pro Tyr His Asp Ala Tyr Glu Asn Val Arg  
195 200 205  
Glu Val Ile Gln Leu Gln Asp Pro Gly Leu Glu Gln Leu Asn Ala Ser  
210 215 220  
Pro Ala Ile Ala Gly Leu Met His Gln Ala Ser Leu Leu Gly Ile Asn  
225 230 235 240  
Ala Ser Ile Ser Pro Glu Leu Phe Asn Ile Leu Thr Glu Glu Ile Thr  
245 250 255  
Glu Gly Asn Ala Glu Glu Leu Tyr Lys Lys Asn Phe Gly Asn Ile Glu  
260 265 270  
Pro Ala Ser Leu Ala Met Pro Glu Tyr Leu Lys Arg Tyr Tyr Asn Leu  
275 280 285  
Ser Asp Glu Glu Leu Ser Gln Phe Ile Gly Lys Ala Ser Asn Phe Gly  
290 295 300  
Gln Gln Glu Tyr Ser Asn Asn Gln Leu Ile Thr Pro Val Val Asn Ser  
305 310 315 320  
Ser Asp Gly Thr Val Lys Val Tyr Arg Ile Thr Arg Glu Tyr Thr Thr  
325 330 335  
Asn Ala Tyr Gln Met Asp Val Glu Leu Phe Pro Phe Gly Gly Glu Asn  
340 345 350

Tyr Arg Leu Asp Tyr Lys Phe Lys Asn Phe Tyr Asn Ala Ser Tyr Leu  
355 360 365

Ser Ile Lys Leu Asn Asp Lys Arg Glu Leu Val Arg Thr Glu Gly Ala  
370 375 380

Pro Gln Val Asn Ile Glu Tyr Ser Ala Asn Ile Thr Leu Asn Thr Ala  
385 390 395 400

Asp Ile Ser Gln Pro Phe Glu Ile Gly Leu Thr Arg Val Leu Pro Ser  
405 410 415

Gly Ser Trp Ala Tyr Ala Ala Ala Lys Phe Thr Val Glu Glu Tyr Asn  
420 425 430

Gln Tyr Ser Phe Leu Leu Lys Leu Asn Lys Ala Ile Arg Leu Ser Arg  
435 440 445

Ala Thr Glu Leu Ser Pro Thr Ile Leu Glu Gly Ile Val Arg Ser Val  
450 455 460

Asn Leu Gln Leu Asp Ile Asn Thr Asp Val Leu Gly Lys Val Phe Leu  
465 470 475 480

Thr Lys Tyr Tyr Met Gln Arg Tyr Ala Ile His Ala Glu Thr Ala Leu  
485 490 495

Ile Leu Cys Asn Ala Pro Ile Ser Gln Arg Ser Tyr Asp Asn Gln Pro  
500 505 510

Ser Gln Phe Asp Arg Leu Phe Asn Thr Pro Leu Leu Asn Gly Gln Tyr  
515 520 525

Phe Ser Thr Gly Asp Glu Glu Ile Asp Leu Asn Ser Gly Ser Thr Gly  
530 535 540

Asp Trp Arg Lys Thr Ile Leu Lys Arg Ala Phe Asn Ile Asp Asp Val  
545 550 555 560

Ser Leu Phe Arg Leu Leu Lys Ile Thr Asp His Asp Asn Lys Asp Gly  
565 570 575

Lys Ile Lys Asn Asn Leu Lys Asn Leu Ser Asn Leu Tyr Ile Gly Lys  
580 585 590

Leu Leu Ala Asp Ile His Gln Leu Thr Ile Asp Glu Leu Asp Leu Leu  
595 600 605

Leu Ile Ala Val Gly Glu Gly Lys Thr Asn Leu Ser Ala Ile Ser Asp  
610 615 620

Lys Gln Leu Ala Thr Leu Ile Arg Lys Leu Asn Thr Ile Thr Ser Trp  
625 630 635 640

Leu His Thr Gln Lys Trp Ser Val Phe Gln Leu Phe Ile Met Thr Ser  
645 650 655

Thr Ser Tyr Asn Lys Thr Leu Thr Pro Glu Ile Lys Asn Leu Leu Asp  
660 665 670

Thr Val Tyr His Gly Leu Gln Gly Phe Asp Lys Asp Lys Ala Asp Leu  
675 680 685

Leu His Val Met Ala Pro Tyr Ile Ala Ala Thr Leu Gln Leu Ser Ser  
690 695 700

Glu Asn Val Ala His Ser Val Leu Leu Trp Ala Asp Lys Leu Gln Pro  
705 710 715 720

Gly Asp Gly Ala Met Thr Ala Glu Lys Phe Trp Asp Trp Leu Asn Thr  
725 730 735

Lys Tyr Thr Pro Gly Ser Ser Glu Ala Val Glu Thr Gln Glu His Ile  
740 745 750

Val Gln Tyr Cys Gln Ala Leu Ala Gln Leu Glu Met Val Tyr His Ser  
755 760 765

Thr Gly Ile Asn Glu Asn Ala Phe Arg Leu Phe Val Thr Lys Pro Glu  
770 775 780

Met Phe Gly Ala Ala Thr Gly Ala Ala Pro Ala His Asp Ala Leu Ser  
785 790 795 800

Leu Ile Met Leu Thr Arg Phe Ala Asp Trp Val Asn Ala Leu Gly Glu  
805 810 815

Lys Ala Ser Ser Val Leu Ala Ala Phe Glu Ala Asn Ser Leu Thr Ala  
820 825 830

Glu Gln Leu Ala Asp Ala Met Asn Leu Asp Ala Asn Leu Leu Leu Gln  
835 840 845

Ala Ser Ile Gln Ala Gln Asn His Gln His Leu Pro Pro Val Thr Pro  
850 855 860

Glu Asn Ala Phe Ser Cys Trp Thr Ser Ile Asn Thr Ile Leu Gln Trp  
865 870 875 880

Val Asn Val Ala Gln Gln Leu Asn Val Ala Pro Gln Gly Val Ser Ala  
885 890 895

Leu Val Gly Leu Asp Tyr Ile Gln Ser Met Lys Glu Thr Pro Thr Tyr  
900 905 910

Ala Gln Trp Glu Asn Ala Ala Gly Val Leu Thr Ala Gly Leu Asn Ser  
915 920 925

Gln Gln Ala Asn Thr Leu His Ala Phe Leu Asp Glu Ser Arg Ser Ala  
930 935 940

Ala Leu Ser Thr Tyr Tyr Ile Arg Gln Val Ala Lys Ala Ala Ala Ala  
945 950 955 960

Ile Lys Ser Arg Asp Asp Leu Tyr Gln Tyr Leu Leu Ile Asp Asn Gln  
965 970 975

Val Ser Ala Ala Ile Lys Thr Thr Arg Ile Ala Glu Ala Ile Ala Ser  
980 985 990

Ile Gln Leu Tyr Val Asn Arg Ala Leu Glu Asn Val Glu Glu Asn Ala  
995 1000 1005

Asn Ser Gly Val Ile Ser Arg Gln Phe Phe Ile Asp Trp Asp Lys Tyr  
1010 1015 1020

Asn Lys Arg Tyr Ser Thr Trp Ala Gly Val Ser Gln Leu Val Tyr Tyr  
1025 1030 1035 1040

Pro Glu Asn Tyr Ile Asp Pro Thr Met Arg Ile Gly Gln Thr Lys Met  
1045 1050 1055

Met Asp Ala Leu Leu Gln Ser Val Ser Gln Ser Gln Leu Asn Ala Asp  
1060 1065 1070

Thr Val Glu Asp Ala Phe Met Ser Tyr Leu Thr Ser Phe Glu Gln Val  
1075 1080 1085

Ala Asn Leu Lys Val Ile Ser Ala Tyr His Asp Asn Ile Asn Asn Asp  
1090 1095 1100

Gln Gly Leu Thr Tyr Phe Ile Gly Leu Ser Glu Thr Asp Ala Gly Glu  
1105 1110 1115 1120

Tyr Tyr Trp Arg Ser Val Asp His Ser Lys Phe Asn Asp Gly Lys Phe  
1125 1130 1135

Ala Ala Asn Ala Trp Ser Glu Trp His Lys Ile Asp Cys Pro Ile Asn  
1140 1145 1150

Pro Tyr Lys Ser Thr Ile Arg Pro Val Ile Tyr Lys Ser Arg Leu Tyr  
1155 1160 1165

Leu Leu Trp Leu Glu Gln Lys Glu Ile Thr Lys Gln Thr Gly Asn Ser  
1170 1175 1180

Lys Asp Gly Tyr Gln Thr Glu Thr Asp Tyr Arg Tyr Glu Leu Lys Leu  
1185 1190 1195 1200

Ala His Ile Arg Tyr Asp Gly Thr Trp Asn Thr Pro Ile Thr Phe Asp  
1205 1210 1215

Val Asn Lys Lys Ile Ser Glu Leu Lys Leu Glu Lys Asn Arg Ala Pro  
1220 1225 1230

Gly Leu Tyr Cys Ala Gly Tyr Gln Gly Glu Asp Thr Leu Leu Val Met  
1235 1240 1245

Phe Tyr Asn Gln Gln Asp Thr Leu Asp Ser Tyr Lys Asn Ala Ser Met  
1250 1255 1260

Gln Gly Leu Tyr Ile Phe Ala Asp Met Ala Ser Lys Asp Met Thr Pro  
1265 1270 1275 1280

Glu Gln Ser Asn Val Tyr Arg Asp Asn Ser Tyr Gln Gln Phe Asp Thr  
1285 1290 1295

Asn Asn Val Arg Arg Val Asn Asn Arg Tyr Ala Glu Asp Tyr Glu Ile  
1300 1305 1310

Pro Ser Ser Val Ser Ser Arg Lys Asp Tyr Gly Trp Gly Asp Tyr Tyr  
1315 1320 1325

Leu Ser Met Val Tyr Asn Gly Asp Ile Pro Thr Ile Asn Tyr Lys Ala  
1330 1335 1340

Ala Ser Ser Asp Leu Lys Ile Tyr Ile Ser Pro Lys Leu Arg Ile Ile  
1345 1350 1355 1360

His Asn Gly Tyr Glu Gly Gln Lys Arg Asn Gln Cys Asn Leu Met Asn  
1365 1370 1375

Lys Tyr Gly Lys Leu Gly Asp Lys Phe Ile Val Tyr Thr Ser Leu Gly  
1380 1385 1390

Val Asn Pro Asn Asn Ser Ser Asn Lys Leu Met Phe Tyr Pro Val Tyr  
1395 1400 1405

Gln Tyr Ser Gly Asn Thr Ser Gly Leu Asn Gln Gly Arg Leu Leu Phe  
1410 1415 1420

His Arg Asp Thr Thr Tyr Pro Ser Lys Val Glu Ala Trp Ile Pro Gly  
1425 1430 1435 1440

Ala Lys Arg Ser Leu Thr Asn Gln Asn Ala Ala Ile Gly Asp Asp Tyr  
1445 1450 1455

Ala Thr Asp Ser Leu Asn Lys Pro Asp Asp Leu Lys Gln Tyr Ile Phe  
1460 1465 1470

Met Thr Asp Ser Lys Gly Thr Ala Thr Asp Val Ser Gly Pro Val Glu  
1475 1480 1485

Ile Asn Thr Ala Ile Ser Pro Ala Lys Val Gln Ile Ile Val Lys Ala  
1490 1495 1500

Gly Gly Lys Glu Gln Thr Phe Thr Ala Asp Lys Asp Val Ser Ile Gln  
1505 1510 1515 1520

Pro Ser Pro Ser Phe Asp Glu Met Asn Tyr Gln Phe Asn Ala Leu Glu  
1525 1530 1535

Ile Asp Gly Ser Gly Leu Asn Phe Ile Asn Asn Ser Ala Ser Ile Asp  
1540 1545 1550

Val Thr Phe Thr Ala Phe Ala Glu Asp Gly Arg Lys Leu Gly Tyr Glu  
~~1555~~ 1560 1565

Ser Phe Ser Ile Pro Val Thr Leu Lys Val Ser Thr Asp Asn Ala Leu  
1570 1575 1580

Thr Leu His His Asn Glu Asn Gly Ala Gln Tyr Met Gln Trp Gln Ser  
1585 1590 1595 1600

Tyr Arg Thr Arg Leu Asn Thr Leu Phe Ala Arg Gln Leu Val Ala Arg  
1605 1610 1615

Ala Thr Thr Gly Ile Asp Thr Ile Leu Ser Met Glu Thr Gln Asn Ile  
1620 1625 1630

Gln Glu Pro Gln Leu Gly Lys Gly Phe Tyr Ala Thr Phe Val Ile Pro  
1635 1640 1645

Pro Tyr Asn Leu Ser Thr His Gly Asp Glu Arg Trp Phe Lys Leu Tyr  
1650 1655 1660

Ile Lys His Val Val Asp Asn Asn Ser His Ile Ile Tyr Ser Gly Gln  
1665 1670 1675 1680

Leu Thr Asp Thr Asn Ile Asn Ile Thr Leu Phe Ile Pro Leu Asp Asp  
1685 1690 1695

Val Pro Leu Asn Gln Asp Tyr His Ala Lys Val Tyr Met Thr Phe Lys  
1700 1705 1710

Lys Ser Pro Ser Asp Gly Thr Trp Trp Gly Pro His Phe Val Arg Asp  
1715 1720 1725

Asp Lys Gly Ile Val Thr Ile Asn Pro Lys Ser Ile Leu Thr His Phe  
1730 1735 1740

Glu Ser Val Asn Val Leu Asn Asn Ile Ser Ser Glu Pro Met Asp Phe  
1745 1750 1755 1760

Ser Gly Ala Asn Ser Leu Tyr Phe Trp Glu Leu Phe Tyr Tyr Thr Pro  
1765 1770 1775

Met Leu Val Ala Gln Arg Leu Leu His Glu Gln Asn Phe Asp Glu Ala  
1780 1785 1790

Asn Arg Trp Leu Lys Tyr Val Trp Ser Pro Ser Gly Tyr Ile Val His  
1795 1800 1805

Gly Gln Ile Gln Asn Tyr Gln Trp Asn Val Arg Pro Leu Leu Glu Asp  
1810 1815 1820

Thr Ser Trp Asn Ser Asp Pro Leu Asp Ser Val Asp Pro Asp Ala Val  
1825 1830 1835 1840

Ala Gln His Asp Pro Met His Tyr Lys Val Ser Thr Phe Met Arg Thr  
1845 1850 1855

Leu Asp Leu Leu Ile Ala Arg Gly Asp His Ala Tyr Arg Gln Leu Glu  
1860 1865 1870

Arg Asp Thr Leu Asn Glu Ala Lys Met Trp Tyr Met Gln Ala Leu His  
1875 1880 1885

Leu Leu Gly Asp Lys Pro Tyr Leu Pro Leu Ser Thr Thr Trp Ser Asp  
1890 1895 1900

Pro Arg Leu Asp Arg Ala Ala Asp Ile Thr Thr Gln Asn Ala His Asp  
1905 1910 1915 1920

Ser Ala Ile Val Ala Leu Arg Gln Asn Ile Pro Thr Pro Ala Pro Leu  
1925 1930 1935

Ser Leu Arg Ser Ala Asn Thr Leu Thr Asp Leu Phe Leu Pro Gln Ile  
1940 1945 1950

Asn Glu Val Met Met Asn Tyr Trp Gln Thr Leu Ala Gln Arg Val Tyr  
1955 1960 1965

Asn Leu Arg His Asn Leu Ser Ile Asp Gly Gln Pro Leu Tyr Leu Pro  
1970 1975 1980

Ile Tyr Ala Thr Pro Ala Asp Pro Lys Ala Leu Leu Ser Ala Ala Val  
1985 1990 1995 2000

Ala Thr Ser Gln Gly Gly Gly Lys Leu Pro Glu Ser Phe Met Ser Leu  
2005 2010 2015

Trp Arg Phe Pro His Met Leu Glu Asn Ala Arg Gly Met Val Ser Gln  
2020 2025 2030

Leu Thr Gln Phe Gly Ser Thr Leu Gln Asn Ile Ile Glu Arg Gln Asp  
2035 2040 2045

Ala Glu Ala Leu Asn Ala Leu Leu Gln Asn Gln Ala Ala Glu Leu Ile  
2050 2055 2060

Leu Thr Asn Leu Ser Ile Gln Asp Lys Thr Ile Glu Glu Leu Asp Ala  
2065 2070 2075 2080

Glu Lys Thr Val Leu Glu Lys Ser Lys Ala Gly Ala Gln Ser Arg Phe  
2085 2090 2095

Asp Ser Tyr Gly Lys Leu Tyr Asp Glu Asn Ile Asn Ala Gly Glu Asn  
2100 2105 2110

Gln Ala Met Thr Leu Arg Ala Ser Ala Ala Gly Leu Thr Thr Ala Val  
2115 2120 2125

Gln Ala Ser Arg Leu Ala Gly Ala Ala Ala Asp Leu Val Pro Asn Ile  
2130 2135 2140

Phe Gly Phe Ala Gly Gly Gly Ser Arg Trp Gly Ala Ile Ala Glu Ala  
2145 2150 2155 2160

Thr Gly Tyr Val Met Glu Phe Ser Ala Asn Val Met Asn Thr Glu Ala  
2165 2170 2175

Asp Lys Ile Ser Gln Ser Glu Thr Tyr Arg Arg Arg Arg Gln Glu Trp  
2180 2185 2190

Glu Ile Gln Arg Asn Asn Ala Glu Ala Glu Leu Lys Gln Ile Asp Ala  
2195 2200 2205

Gln Leu Lys Ser Leu Ala Val Arg Arg Glu Ala Ala Val Leu Gln Lys  
2210 2215 2220

Thr Ser Leu Lys Thr Gln Gln Glu Gln Thr Gln Ser Gln Leu Ala Phe  
2225 2230 2235 2240

Leu Gln Arg Lys Phe Ser Asn Gln Ala Leu Tyr Asn Trp Leu Arg Gly  
2245 2250 2255

Arg Leu Ala Ala Ile Tyr Phe Gln Phe Tyr Asp Leu Ala Val Ala Arg  
2260 2265 2270

Cys Leu Met Ala Glu Gln Ala Tyr Arg Trp Glu Leu Asn Asp Asp Ser  
2275 2280 2285

Ala Arg Phe Ile Lys Pro Gly Ala Trp Gln Gly Thr Tyr Ala Gly Leu  
2290 2295 2300

Leu Ala Gly Glu Thr Leu Met Leu Ser Leu Ala Gln Met Glu Asp Ala  
2305 2310 2315 2320

His Leu Lys Arg Asp Lys Arg Ala Leu Glu Val Glu Arg Thr Val Ser  
2325 2330 2335

Leu Ala Glu Val Tyr Ala Gly Leu Pro Lys Asp Asn Gly Pro Phe Ser  
2340 2345 2350

Leu Ala Gln Glu Ile Asp Lys Leu Val Ser Gln Gly Ser Gly Ser Ala  
2355 2360 2365

Gly Ser Gly Asn Asn Asn Leu Ala Phe Gly Ala Gly Thr Asp Thr Lys  
2370 2375 2380

Thr Ser Leu Gln Ala Ser Val Ser Phe Ala Asp Leu Lys Ile Arg Glu  
2385 2390 2395 2400

Asp Tyr Pro Ala Ser Leu Gly Lys Ile Arg Arg Ile Lys Gln Ile Ser  
2405 2410 2415

Val Thr Leu Pro Ala Leu Leu Gly Pro Tyr Gln Asp Val Gln Ala Ile  
2420 2425 2430

Leu Ser Tyr Gly Asp Lys Ala Gly Leu Ala Asn Gly Cys Glu Ala Leu  
2435 2440 2445

Ala Val Ser His Gly Met Asn Asp Ser Gly Gln Phe Gln Leu Asp Phe  
2450 2455 2460

Asn Asp Gly Lys Phe Leu Pro Phe Glu Gly Ile Ala Ile Asp Gln Gly  
2465 2470 2475 2480

Thr Leu Thr Leu Ser Phe Pro Asn Ala Ser Met Pro Glu Lys Gly Lys  
2485 2490 2495

Gln Ala Thr Met Leu Lys Thr Leu Asn Asp Ile Ile Leu His Ile Arg  
2500 2505 2510

Tyr Thr Ile Lys  
2516



(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 48:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĚLKA: 5547 párov báz
- (B) TYP: nukleová kyselina
- (C) REŤAZEC: dvojitý
- (D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: DNA (genómová)

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 46 (kódovaná oblasť  
tcdA<sub>i</sub>):

CTG	ATA	GGC	TAT	AAC	AAT	CAA	TTT	AGC	GGT	AGA	GCC	AGT	CAA	TAT	GTT	48
Leu	Ile	Gly	Tyr	Asn	Asn	Gln	Phe	Ser	Gly	Arg	Ala	Ser	Gln	Tyr	Val	
1				5					10					15		
GCG	CCG	GGT	ACC	GTT	TCT	TCC	ATG	TTC	TCC	CCC	GCC	GCT	TAT	TTG	ACT	96
Ala	Pro	Gly	Thr	Val	Ser	Ser	Met	Phe	Ser	Pro	Ala	Ala	Tyr	Leu	Thr	
			20					25					30			
GAA	CTT	TAT	CGT	GAA	GCA	CGC	AAT	TTA	CAC	GCA	AGT	GAC	TCC	GTT	TAT	144
Glu	Leu	Tyr	Arg	Glu	Ala	Arg	Asn	Leu	His	Ala	Ser	Asp	Ser	Val	Tyr	
		35					40					45				
TAT	CTG	GAT	ACC	CGC	CGC	CCA	GAT	CTC	AAA	TCA	ATG	GCG	CTC	AGT	CAG	192
Tyr	Leu	Asp	Thr	Arg	Arg	Pro	Asp	Leu	Lys	Ser	Met	Ala	Leu	Ser	Gln	
	50					55					60					
CAA	AAT	ATG	GAT	ATA	GAA	TTA	TCC	ACA	CTC	TCT	TTG	TCC	AAT	GAG	CTG	240
Gln	Asn	Met	Asp	Ile	Glu	Leu	Ser	Thr	Leu	Ser	Leu	Ser	Asn	Glu	Leu	
65					70					75					80	
TTA	TTG	GAA	AGC	ATT	AAA	ACT	GAA	TCT	AAA	CTG	GAA	AAC	TAT	ACT	AAA	288
Leu	Leu	Glu	Ser	Ile	Lys	Thr	Glu	Ser	Lys	Leu	Glu	Asn	Tyr	Thr	Lys	
				85					90					95		
GTG	ATG	GAA	ATG	CTC	TCC	ACT	TTC	CGT	CCT	TCC	GGC	GCA	ACG	CCT	TAT	336
Val	Met	Glu	Met	Leu	Ser	Thr	Phe	Arg	Pro	Ser	Gly	Ala	Thr	Pro	Tyr	
			100					105					110			
CAT	GAT	GCT	TAT	GAA	AAT	GTG	CGT	GAA	GTT	ATC	CAG	CTA	CAA	GAT	CCT	384
His	Asp	Ala	Tyr	Glu	Asn	Val	Arg	Glu	Val	Ile	Gln	Leu	Gln	Asp	Pro	
		115					120					125				
GGA	CTT	GAG	CAA	CTC	AAT	GCA	TCA	CCG	GCA	ATT	GCC	GGG	TTG	ATG	CAT	432
Gly	Leu	Glu	Gln	Leu	Asn	Ala	Ser	Pro	Ala	Ile	Ala	Gly	Leu	Met	His	
	130					135					140					
CAA	GCC	TCC	CTA	TTG	GGT	ATT	AAC	GCT	TCA	ATC	TCG	CCT	GAG	CTA	TTT	480
Gln	Ala	Ser	Leu	Leu	Gly	Ile	Asn	Ala	Ser	Ile	Ser	Pro	Glu	Leu	Phe	
145					150					155					160	

AAT ATT CTG ACG GAG GAG ATT ACC GAA GGT AAT GCT GAG GAA CTT TAT 528  
 Asn Ile Leu Thr Glu Glu Ile Thr Glu Gly Asn Ala Glu Glu Leu Tyr  
 165 170 175

AAG AAA AAT TTT GGT AAT ATC GAA CCG GCC TCA TTG GCT ATG CCG GAA 576  
 Lys Lys Asn Phe Gly Asn Ile Glu Pro Ala Ser Leu Ala Met Pro Glu  
 180 185 190

TAC CTT AAA CGT TAT TAT AAT TTA AGC GAT GAA GAA CTT AGT CAG TTT 624  
 Tyr Leu Lys Arg Tyr Tyr Asn Leu Ser Asp Glu Glu Leu Ser Gln Phe  
 195 200 205

ATT GGT AAA GCC AGC AAT TTT GGT CAA CAG GAA TAT AGT AAT AAC CAA 672  
 Ile Gly Lys Ala Ser Asn Phe Gly Gln Gln Glu Tyr Ser Asn Asn Gln  
 210 215 220

CTT ATT ACT CCG GTA GTC AAC AGC AGT GAT GGC ACG GTT AAG GTA TAT 720  
 Leu Ile Thr Pro Val Val Asn Ser Ser Asp Gly Thr Val Lys Val Tyr  
 225 230 235 240

CGG ATC ACC CGC GAA TAT ACA ACC AAT GCT TAT CAA ATG GAT GTG GAG 768  
 Arg Ile Thr Arg Glu Tyr Thr Thr Asn Ala Tyr Gln Met Asp Val Glu  
 245 250 255

CTA TTT CCC TTC GGT GGT GAG AAT TAT CGG TTA GAT TAT AAA TTC AAA 816  
 Leu Phe Pro Phe Gly Gly Glu Asn Tyr Arg Leu Asp Tyr Lys Phe Lys  
 260 265 270

AAT TTT TAT AAT GCC TCT TAT TTA TCC ATC AAG TTA AAT GAT AAA AGA 864  
 Asn Phe Tyr Asn Ala Ser Tyr Leu Ser Ile Lys Leu Asn Asp Lys Arg  
 275 280 285

GAA CTT GTT CGA ACT GAA GGC GCT CCT CAA GTC AAT ATA GAA TAC TCC 912  
 Glu Leu Val Arg Thr Glu Gly Ala Pro Gln Val Asn Ile Glu Tyr Ser  
 290 295 300

GCA AAT ATC ACA TTA AAT ACC GCT GAT ATC AGT CAA CCT TTT GAA ATT 960  
 Ala Asn Ile Thr Leu Asn Thr Ala Asp Ile Ser Gln Pro Phe Glu Ile  
 305 310 315 320

GGC CTG ACA CGA GTA CTT CCT TCC GGT TCT TGG GCA TAT GCC GCC GCA 1008  
 Gly Leu Thr Arg Val Leu Pro Ser Gly Ser Trp Ala Tyr Ala Ala Ala  
 325 330 335

AAA TTT ACC GTT GAA GAG TAT AAC CAA TAC TCT TTT CTG CTA AAA CTT 1056  
 Lys Phe Thr Val Glu Glu Tyr Asn Gln Tyr Ser Phe Leu Leu Lys Leu  
 340 345 350

AAC AAG GCT ATT CGT CTA TCA CGT GCG ACA GAA TTG TCA CCC ACG ATT 1104  
 Asn Lys Ala Ile Arg Leu Ser Arg Ala Thr Glu Leu Ser Pro Thr Ile  
 355 360 365

CTG GAA GGC ATT GTG CGC AGT GTT AAT CTA CAA CTG GAT ATC AAC ACA 1152  
 Leu Glu Gly Ile Val Arg Ser Val Asn Leu Gln Leu Asp Ile Asn Thr  
 370 375 380

GAC GTA TTA GGT AAA GTT TTT CTG ACT AAA TAT TAT ATG CAG CGT TAT 1200  
 Asp Val Leu Gly Lys Val Phe Leu Thr Lys Tyr Tyr Met Gln Arg Tyr  
 385 390 395 400

GCT ATT CAT GCT GAA ACT GCC CTG ATA CTA TGC AAC GCG CCT ATT TCA 1248  
 Ala Ile His Ala Glu Thr Ala Leu Ile Leu Cys Asn Ala Pro Ile Ser  
 405 410 415

CAA CGT TCA TAT GAT AAT CAA CCT AGC CAA TTT GAT GCG CTG TTT AAT 1296  
 Gln Arg Ser Tyr Asp Asn Gln Pro Ser Gln Phe Asp Arg Leu Phe Asn  
 420 425 430

ACG CCA TTA CTG AAC GGA CAA TAT TTT TCT ACC GGC GAT GAG GAG ATT 1344  
 Thr Pro Leu Leu Asn Gly Gln Tyr Phe Ser Thr Gly Asp Glu Glu Ile

	435				440				445							
GAT	TTA	AAT	TCA	GGT	AGC	ACC	GGC	GAT	TGG	CGA	AAA	ACC	ATA	CTT	AAG	1392
Asp	Leu	Asn	Ser	Gly	Ser	Thr	Gly	Asp	Trp	Arg	Lys	Thr	Ile	Leu	Lys	
	450					455					460					
CGT	GCA	TTT	AAT	ATT	GAT	GAT	GTC	TCG	CTC	TTC	CGC	CTG	CTT	AAA	ATT	1440
Arg	Ala	Phe	Asn	Ile	Asp	Asp	Val	Ser	Leu	Phe	Arg	Leu	Leu	Lys	Ile	
	465				470					475					480	
ACC	GAC	CAT	GAT	AAT	AAA	GAT	GGA	AAA	ATT	AAA	AAT	AAC	CTA	AAG	AAT	1488
Thr	Asp	His	Asp	Asn	Lys	Asp	Gly	Lys	Ile	Lys	Asn	Asn	Leu	Lys	Asn	
				485					490					495		
CTT	TCC	AAT	TTA	TAT	ATT	GGA	AAA	TTA	CTG	GCA	GAT	ATT	CAT	CAA	TTA	1536
Leu	Ser	Asn	Leu	Tyr	Ile	Gly	Lys	Leu	Leu	Ala	Asp	Ile	His	Gln	Leu	
			500					505					510			
ACC	ATT	GAT	GAA	CTG	GAT	TTA	TTA	CTG	ATT	GCC	GTA	GGT	GAA	GGA	AAA	1584
Thr	Ile	Asp	Glu	Leu	Asp	Leu	Leu	Leu	Ile	Ala	Val	Gly	Glu	Gly	Lys	
		515					520					525				
ACT	AAT	TTA	TCC	GCT	ATC	AGT	GAT	AAG	CAA	TTG	GCT	ACC	CTG	ATC	AGA	1632
Thr	Asn	Leu	Ser	Ala	Ile	Ser	Asp	Lys	Gln	Leu	Ala	Thr	Leu	Ile	Arg	
	530					535					540					
AAA	CTC	AAT	ACT	ATT	ACC	AGC	TGG	CTA	CAT	ACA	CAG	AAG	TGG	AGT	GTA	1680
Lys	Leu	Asn	Thr	Ile	Thr	Ser	Trp	Leu	His	Thr	Gln	Lys	Trp	Ser	Val	
	545				550					555					560	
TTC	CAG	CTA	TTT	ATC	ATG	ACC	TCC	ACC	AGC	TAT	AAC	AAA	ACG	CTA	ACG	1728
Phe	Gln	Leu	Phe	Ile	Met	Thr	Ser	Thr	Ser	Tyr	Asn	Lys	Thr	Leu	Thr	
			565						570					575		
CCT	GAA	ATT	AAG	AAT	TTG	CTG	GAT	ACC	GTC	TAC	CAC	GGT	TTA	CAA	GGT	1776
Pro	Glu	Ile	Lys	Asn	Leu	Leu	Asp	Thr	Val	Tyr	His	Gly	Leu	Gln	Gly	
			580					585					590			
TTT	GAT	AAA	GAC	AAA	GCA	GAT	TTG	CTA	CAT	GTC	ATG	GCG	CCC	TAT	ATT	1824
Phe	Asp	Lys	Asp	Lys	Ala	Asp	Leu	Leu	His	Val	Met	Ala	Pro	Tyr	Ile	
		595				600						605				
GCG	GCC	ACC	TTG	CAA	TTA	TCA	TCG	GAA	AAT	GTC	GCC	CAC	TCG	GTA	CTC	1872
Ala	Ala	Thr	Leu	Gln	Leu	Ser	Ser	Glu	Asn	Val	Ala	His	Ser	Val	Leu	
	610					615					620					
CTT	TGG	GCA	GAT	AAG	TTA	CAG	CCC	GGC	GAC	GGC	GCA	ATG	ACA	GCA	GAA	1920
Leu	Trp	Ala	Asp	Lys	Leu	Gln	Pro	Gly	Asp	Gly	Ala	Met	Thr	Ala	Glu	
	625				630					635					640	
AAA	TTC	TGG	GAC	TGG	TTG	AAT	ACT	AAG	TAT	ACG	CCG	GGT	TCA	TCG	GAA	1968
Lys	Phe	Trp	Asp	Trp	Leu	Asn	Thr	Lys	Tyr	Thr	Pro	Gly	Ser	Ser	Glu	
			645						650					655		
GCC	GTA	GAA	ACG	CAG	GAA	CAT	ATC	GTT	CAG	TAT	TGT	CAG	GCT	CTG	GCA	2016
Ala	Val	Glu	Thr	Gln	Glu	His	Ile	Val	Gln	Tyr	Cys	Gln	Ala	Leu	Ala	
			660					665					670			
CAA	TTG	GAA	ATG	GTT	TAC	CAT	TCC	ACC	GGC	ATC	AAC	GAA	AAC	GCC	TTC	2064
Gln	Leu	Glu	Met	Val	Tyr	His	Ser	Thr	Gly	Ile	Asn	Glu	Asn	Ala	Phe	
		675					680					685				
CGT	CTA	TTT	GTG	ACA	AAA	CCA	GAG	ATG	TTT	GGC	GCT	GCA	ACT	GGA	GCA	2112
Arg	Leu	Phe	Val	Thr	Lys	Pro	Glu	Met	Phe	Gly	Ala	Ala	Thr	Gly	Ala	
	690					695					700					
GCG	CCC	GCG	CAT	GAT	GCC	GTT	TCA	CTG	ATT	ATG	CTG	ACA	CGT	TTT	GCG	2160
Ala	Pro	Ala	His	Asp	Ala	Leu	Ser	Leu	Ile	Met	Leu	Thr	Arg	Phe	Ala	
	705				710					715					720	
GAT	TGG	GTG	AAC	GCA	CTA	GCC	GAA	AAA	GCS	TCC	TCG	GTG	CTA	GCG	GCA	2208

Asp	Trp	Val	Asn	Ala	Leu	Gly	Glu	Lys	Ala	Ser	Ser	Val	Leu	Ala	Ala		
				725					730			---		735			
TTT	GAA	GCT	AAC	TCG	TTA	ACG	GCA	GAA	CAA	CTG	GCT	GAT	GCC	ATG	AAT	2256	
Phe	Glu	Ala	Asn	Ser	Leu	Thr	Ala	Glu	Gln	Leu	Ala	Asp	Ala	Met	Asn		
			740					745					750				
CTT	GAT	GCT	AAT	TTG	CTG	TTG	CAA	GCC	AGT	ATT	CAA	GCA	CAA	AAT	CAT	2304	
Leu	Asp	Ala	Asn	Leu	Leu	Leu	Gln	Ala	Ser	Ile	Gln	Ala	Gln	Asn	His		
			755				760					765					
CAA	CAT	CTT	CCC	CCA	GTA	ACT	CCA	GAA	AAT	GCG	TTC	TCC	TGT	TGG	ACA	2352	
Gln	His	Leu	Pro	Pro	Val	Thr	Pro	Glu	Asn	Ala	Phe	Ser	Cys	Trp	Thr		
			770			775					780						
TCT	ATC	AAT	ACT	ATC	CTG	CAA	TGG	GTT	AAT	GTC	GCA	CAA	CAA	TTG	AAT	2400	
Ser	Ile	Asn	Thr	Ile	Leu	Gln	Trp	Val	Asn	Val	Ala	Gln	Gln	Leu	Asn		
					790					795					800		
GTC	GCC	CCA	CAG	GGC	GTT	TCC	GCT	TTG	GTC	GGG	CTG	GAT	TAT	ATT	CAA	2448	
Val	Ala	Pro	Gln	Gly	Val	Ser	Ala	Leu	Val	Gly	Leu	Asp	Tyr	Ile	Gln		
				- 805					810					815			
TCA	ATG	AAA	GAG	ACA	CCG	ACC	TAT	GCC	CAG	TGG	GAA	AAC	GCG	GCA	GGC	2496	
Ser	Met	Lys	Glu	Thr	Pro	Thr	Tyr	Ala	Gln	Trp	Glu	Asn	Ala	Ala	Gly		
			820					825					830				
GTA	TTA	ACC	GCC	GGG	TTG	AAT	TCA	CAA	CAG	GCT	AAT	ACA	TTA	CAC	GCT	2544	
Val	Leu	Thr	Ala	Gly	Leu	Asn	Ser	Gln	Gln	Ala	Asn	Thr	Leu	His	Ala		
			835				840					845					
TTT	CTG	GAT	GAA	TCT	CGC	AGT	GCC	GCA	TTA	AGC	ACC	TAC	TAT	ATC	CGT	2592	
Phe	Leu	Asp	Glu	Ser	Arg	Ser	Ala	Ala	Leu	Ser	Thr	Tyr	Tyr	Ile	Arg		
			850			855					860						
CAA	GTC	GCC	AAG	GCA	GCG	GCG	GCT	ATT	AAA	AGC	CGT	GAT	GAC	TTG	TAT	2640	
Gln	Val	Ala	Lys	Ala	Ala	Ala	Ala	Ile	Lys	Ser	Arg	Asp	Asp	Leu	Tyr		
					870					875				880			
CAA	TAC	TTA	CTG	ATT	GAT	AAT	CAG	GTT	TCT	GCG	GCA	ATA	AAA	ACC	ACC	2688	
Gln	Tyr	Leu	Leu	Ile	Asp	Asn	Gln	Val	Ser	Ala	Ala	Ile	Lys	Thr	Thr		
				885				890						895			
CGG	ATC	GCC	GAA	GCC	ATT	GCC	AGT	ATT	CAA	CTG	TAC	GTC	AAC	CGG	GCA	2736	
Arg	Ile	Ala	Glu	Ala	Ile	Ala	Ser	Ile	Gln	Leu	Tyr	Val	Asn	Arg	Ala		
			900					905					910				
TTG	GAA	AAT	GTG	GAA	GAA	AAT	GCC	AAT	TCG	GGG	GTT	ATC	AGC	CGC	CAA	2784	
Leu	Glu	Asn	Val	Glu	Glu	Asn	Ala	Asn	Ser	Gly	Val	Ile	Ser	Arg	Gln		
			915				920					925					
TTC	TTT	ATC	GAC	TGG	GAC	AAA	TAC	AAT	AAA	CGC	TAC	AGC	ACT	TGG	GCG	2832	
Phe	Phe	Ile	Asp	Trp	Asp	Lys	Tyr	Asn	Lys	Arg	Tyr	Ser	Thr	Trp	Ala		
			930			935					940						
GGT	GTT	TCT	CAA	TTA	GTT	TAC	TAC	CCG	GAA	AAC	TAT	ATT	GAT	CCG	ACC	2880	
Gly	Val	Ser	Gln	Leu	Val	Tyr	Tyr	Pro	Glu	Asn	Tyr	Ile	Asp	Pro	Thr		
					950					955				960			
ATG	CGT	ATC	GGA	CAA	ACC	AAA	ATG	ATG	GAC	GCA	TTA	CTG	CAA	TCC	GTC	2928	
Met	Arg	Ile	Gly	Gln	Thr	Lys	Met	Met	Asp	Ala	Leu	Leu	Gln	Ser	Val		
				965					970					975			
AGC	CAA	AGC	CAA	TTA	AAC	GCC	GAT	ACC	GTC	GAA	GAT	GCC	TTT	ATG	TCT	2976	
Ser	Gln	Ser	Gln	Leu	Asn	Ala	Asp	Thr	Val	Glu	Asp	Ala	Phe	Met	Ser		
				980				985					990				
TAT	CTG	ACA	TCG	TTT	GAA	CAA	GTG	GCT	AAT	CTT	AAA	GTT	ATT	AGC	GCA	3024	
Tyr	Leu	Thr	Ser	Phe	Glu	Gln	Val	Ala	Asn	Leu	Lys	Val	Ile	Ser	Ala		
				995			1000					1005					

TAT CAC GAT AAT ATT AAT AAC GAT CAA GGG CTG ACC TAT TTT ATC GGA 3072  
Tyr His Asp Asn Ile Asn Asn Asp Gln Gly Leu Thr Tyr Phe Ile Gly  
1010 1015 1020

CTC AGT GAA ACT GAT GCC GGT GAA TAT TAT TGG CGC AGT GTC GAT CAC 3120  
Leu Ser Glu Thr Asp Ala Gly Glu Tyr Tyr Trp Arg Ser Val Asp His  
1025 1030 1035 1040

AGT AAA TTC AAC GAC GGT AAA TTC GCG GCT AAT GCC TGG AGT GAA TGG 3168  
Ser Lys Phe Asn Asp Gly Lys Phe Ala Ala Asn Ala Trp Ser Glu Trp  
1045 1050 1055

CAT AAA ATT GAT TGT CCA ATT AAC CCT TAT AAA AGC ACT ATC CGT CCA 3216  
His Lys Ile Asp Cys Pro Ile Asn Pro Tyr Lys Ser Thr Ile Arg Pro  
1060 1065 1070

GTG ATA TAT AAA TCC CGC CTG TAT CTG CTC TGG TTG GAA CAA AAG GAG 3264  
Val Ile Tyr Lys Ser Arg Leu Tyr Leu Leu Trp Leu Glu Gln Lys Glu  
1075 1080 1085

ATC ACC AAA CAG ACA GGA AAT AGT AAA GAT GGC TAT CAA ACT GAA ACG 3312  
Ile Thr Lys Gln Thr Gly Asn Ser Lys Asp Gly Tyr Gln Thr Glu Thr  
1090 1095 1100

GAT TAT CGT TAT GAA CTA AAA TTG GCG CAT ATC CGC TAT GAT GGC ACT 3360  
Asp Tyr Arg Tyr Glu Leu Lys Leu Ala His Ile Arg Tyr Asp Gly Thr  
1105 1110 1115 1120

TGG AAT ACG CCA ATC ACC TTT GAT GTC AAT AAA AAA ATA TCC GAG CTA 3408  
Trp Asn Thr Pro Ile Thr Phe Asp Val Asn Lys Lys Ile Ser Glu Leu  
1125 1130 1135

AAA CTG GAA AAA AAT AGA GCG CCC GGA CTC TAT TGT GCC GGT TAT CAA 3456  
Lys Leu Glu Lys Asn Arg Ala Pro Gly Leu Tyr Cys Ala Gly Tyr Gln  
1140 1145 1150

GGT GAA GAT ACG TTG CTG GTG ATG TTT TAT AAC CAA CAA GAC ACA CTA 3504  
Gly Glu Asp Thr Leu Leu Val Met Phe Tyr Asn Gln Gln Asp Thr Leu  
1155 1160 1165

GAT AGT TAT AAA AAC GCT TCA ATG CAA GGA CTA TAT ATC TTT GCT GAT 3552  
Asp Ser Tyr Lys Asn Ala Ser Met Gln Gly Leu Tyr Ile Phe Ala Asp  
1170 1175 1180

ATG GCA TCC AAA GAT ATG ACC CCA GAA CAG AGC AAT GTT TAT CGG GAT 3600  
Met Ala Ser Lys Asp Met Thr Pro Glu Gln Ser Asn Val Tyr Arg Asp  
1185 1190 1195 1200

AAT AGC TAT CAA CAA TTT GAT ACC AAT AAT GTC AGA AGA GTG AAT AAC 3648  
Asn Ser Tyr Gln Gln Phe Asp Thr Asn Asn Val Arg Arg Val Asn Asn  
1205 1210 1215

CGC TAT GCA GAG GAT TAT GAG ATT CCT TCC TCG GTA AGT AGC CGT AAA 3696  
Arg Tyr Ala Glu Asp Tyr Glu Ile Pro Ser Ser Val Ser Ser Arg Lys  
1220 1225 1230

GAC TAT GGT TGG GGA GAT TAT TAC CTC AGC ATG GTA TAT AAC GGA GAT 3744  
Asp Tyr Gly Trp Gly Asp Tyr Tyr Leu Ser Met Val Tyr Asn Gly Asp  
1235 1240 1245

ATT GCA ACT ATC AAT TAC AAA GCC GCA TCA AGT GAT TTA AAA ATC TAT 3792  
Ile Pro Thr Ile Asn Tyr Lys Ala Ala Ser Ser Asp Leu Lys Ile Tyr  
1250 1255 1260

ATC TCA GCA AAA TTA AGA ATT ATT CAT AAT GGA TAT GAA GGA CAG AAG 3840  
Ile Ser Pro Lys Leu Arg Ile Ile His Asn Gly Tyr Glu Gly Gln Lys  
1265 1270 1275 1280

CGC AAT CAA TGC AAT CTG ATG AAT AAA TAT GGC AAA CTA GGT GAT AAA 3998  
Arg Asn Gln Cys Asn Leu Met Asn Lys Tyr Gly Lys Leu Gly Asp Lys  
1285 1290 1295

TTT ATT GTT TAT ACT AGC TTG GGG GTC AAT CCA AAT AAC TCG TCA AAT 3936  
Phe Ile Val Tyr Thr Ser Leu Gly Val Asn Pro Asn Asn Ser Ser Asn  
1300 1305 1310

AAG CTC ATG TTT TAC CCC GTC TAT CAA TAT AGC GGA AAC ACC AGT GGA 3984  
Lys Leu Met Phe Tyr Pro Val Tyr Gln Tyr Ser Gly Asn Thr Ser Gly  
1315 1320 1325

CTC AAT CAA GGG AGA CTA CTA TTC CAC CGT GAC ACC ACT TAT CCA TCT 4032  
Leu Asn Gln Gly Arg Leu Leu Phe His Arg Asp Thr Thr Tyr Pro Ser  
1330 1335 1340

AAA GTA GAA GCT TGG ATT CCT GGA GCA AAA CGT TCT CTA ACC AAC CAA 4080  
Lys Val Glu Ala Trp Ile Pro Gly Ala Lys Arg Ser Leu Thr Asn Gln  
1345 1350 1355 1360

AAT GCC GCC ATT GGT GAT GAT TAT GCT ACA GAC TCT CTG AAT AAA CCG 4128  
Asn Ala Ala Ile Gly Asp Asp Tyr Ala Thr Asp Ser Leu Asn Lys Pro  
1365 1370 1375

GAT GAT CTT AAG CAA TAT ATC TTT ATG ACT GAC AGT AAA GGG ACT GCT 4176  
Asp Asp Leu Lys Gln Tyr Ile Phe Met Thr Asp Ser Lys Gly Thr Ala  
1380 1385 1390

ACT GAT GTC TCA GGC CCA GTA GAG ATT AAT ACT GCA ATT TCT CCA GCA 4224  
Thr Asp Val Ser Gly Pro Val Glu Ile Asn Thr Ala Ile Ser Pro Ala  
1395 1400 1405

AAA GTT CAG ATA ATA GTC AAA GCG GGT GGC AAG GAG CAA ACT TTT ACC 4272  
Lys Val Gln Ile Ile Val Lys Ala Gly Gly Lys Glu Gln Thr Phe Thr  
1410 1415 1420

GCA GAT AAA GAT GTC TCC ATT CAG CCA TCA CCT AGC TTT GAT GAA ATG 4320  
Ala Asp Lys Asp Val Ser Ile Gln Pro Ser Pro Ser Phe Asp Glu Met  
1425 1430 1435 1440

AAT TAT CAA TTT AAT GCC CTT GAA ATA GAC GGT TCT GGT CTG AAT TTT 4368  
Asn Tyr Gln Phe Asn Ala Leu Glu Ile Asp Gly Ser Gly Leu Asn Phe  
1445 1450 1455

ATT AAC AAC TCA GCC AGT ATT GAT GTT ACT TTT ACC GCA TTT GCG GAG 4416  
Ile Asn Asn Ser Ala Ser Ile Asp Val Thr Phe Thr Ala Phe Ala Glu  
1460 1465 1470

GAT GGC CGC AAA CTG GGT TAT GAA AGT TTC AGT ATT CCT GTT ACC CTC 4464  
Asp Gly Arg Lys Leu Gly Tyr Glu Ser Phe Ser Ile Pro Val Thr Leu  
1475 1480 1485

AAG GTA AGT ACC GAT AAT GCC CTG ACC CTG CAC CAT AAT GAA AAT GGT 4512  
Lys Val Ser Thr Asp Asn Ala Leu Thr Leu His His Asn Glu Asn Gly  
1490 1495 1500

GCG CAA TAT ATG CAA TGG CAA TCC TAT CGT ACC CGC CTG AAT ACT CTA 4560  
Ala Gln Tyr Met Gln Trp Gln Ser Tyr Arg Thr Arg Leu Asn Thr Leu  
1505 1510 1515 1520

TTT GCC CGC CAG TTG GTT GCA CGC GCC ACC ACC GGA ATC GAT ACA ATT 4608  
Phe Ala Arg Gln Leu Val Ala Arg Ala Thr Thr Gly Ile Asp Thr Ile  
1525 1530 1535

CTG AGT ATG GAA ACT CAG AAT ATT CAG GAA CCG CAG TTA GGC AAA GGT 4656  
Leu Ser Met Glu Thr Gln Asn Ile Gln Glu Pro Gln Leu Gly Lys Gly  
1540 1545 1550

TTC TAT GCT ACG TTC GTG ATA CCT GTC TAT AAC CTA TCA ACT CAT GGT 4704  
Phe Tyr Ala Thr Phe Val Ile Pro Pro Tyr Asn Leu Ser Thr His Gly  
1555 1560 1565

GAT GAA CGT TGG TTT AAG CTT TAT ATC AAA CAT GTT GTT GAT AAT AAT 4752  
Asp Glu Arg Trp Phe Lys Leu Tyr Ile Lys His Val Val Asp Asn Asn

1570		1575		1580	
TCA CAT ATT ATC TAT	TCA GGC CAG CTA ACA	GAT ACA AAT ATA AAC	ATC 4800		
Ser His Ile Ile Tyr	Ser Gly Gln Leu Thr	Asp Thr Asn Ile Asn	Ile 1600		
1585	1590	1595	1600		
ACA TTA TTT ATT CCT	CTT GAT GAT GTC CCA	TTG AAT CAA GAT TAT	CAC 4848		
Thr Leu Phe Ile Ile	Pro Leu Asp Asp Val	Pro Leu Asn Gln Asp	Tyr His 1615		
1605	1610				
GCC AAG GTT TAT ATG	ACC TTC AAG AAA TCA	CCA TCA GAT GGT ACC	TGG 4896		
Ala Lys Val Tyr Met	Thr Phe Lys Lys Ser	Pro Ser Asp Gly Thr	Trp 1630		
1620	1625	1630			
TGG GGC CCT CAC TTT	GTT AGA GAT GAT AAA	GGA ATA GTA ACA ATA	AAC 4944		
Trp Gly Pro His Phe	Val Arg Asp Asp Lys	Gly Ile Val Thr Ile	Asn 1645		
1635	1640	1645			
CCT AAA TCC ATT TTG	ACC CAT TTT GAG AGC	GTC AAT GTC CTG AAT	AAT 4992		
Pro Lys Ser Ile Leu	Thr His Phe Glu Ser	Val Asn Val Leu Asn	Asn 1660		
1650	1655	1660			
ATT AGT AGC GAA CCA	ATG GAT TTC AGC GGC	GCT AAC AGC CTC TAT	TTC 5040		
Ile Ser Ser Glu Pro	Met Asp Phe Ser Gly	Ala Asn Ser Leu Tyr	Phe 1680		
1665	1670	1675	1680		
TGG GAA CTG TTC TAC	TAT ACC CCG ATG CTG	GTT GCT CAA CGT TTG	CTG 5088		
Trp Glu Leu Phe Tyr	Tyr Thr Pro Met Leu	Val Ala Gln Arg Leu	Leu 1695		
1685	1690	1695			
CAT GAA CAG AAC TTC	GAT GAA GCC AAC CGT	TGG CTG AAA TAT GTC	TGG 5136		
His Glu Gln Asn Phe	Asp Glu Ala Asn Arg	Trp Leu Lys Tyr Val	Trp 1710		
1700	1705	1710			
AGT CCA TCC GGT TAT	ATT GTC CAC GGC CAG	ATT CAG AAC TAC CAG	TGG 5184		
Ser Pro Ser Gly Tyr	Ile Val His Gly Gln	Ile Gln Asn Tyr Gln	Trp 1725		
1715	1720	1725			
AAC GTC CGC CCG TTA	CTG GAA GAC ACC AGT	TGG AAC AGT GAT CCT	TTG 5232		
Asn Val Arg Pro Leu	Leu Glu Asp Thr Ser	Trp Asn Ser Asp Pro	Leu 1740		
1730	1735	1740			
GAT TCC GTC GAT CCT	GAC GCG GTA GCA CAG	CAC GAT CCA ATG CAC	TAC 5280		
Asp Ser Val Asp Pro	Asp Ala Val Ala Gln	His Asp Pro Met His	Tyr 1760		
1745	1750	1755	1760		
AAA GTT TCA ACT TTT	ATG CGT ACC TTG GAT	CTA TTG ATA GCA CGC	GGC 5328		
Lys Val Ser Thr Phe	Met Arg Thr Leu Asp	Leu Leu Ile Ala Arg	Gly 1775		
1765	1770	1775			
GAC CAT GCT TAT CGC	CAA CTG GAA CGA GAT	ACA CTC AAC GAA GCG	AAG 5376		
Asp His Ala Tyr Arg	Gln Leu Glu Arg Asp	Thr Leu Asn Glu Ala	Lys 1790		
1780	1785	1790			
ATG TGG TAT ATG CAA	GCG CTG CAT CTA TTA	GGT GAC AAA CCT TAT	CTA 5424		
Met Trp Tyr Met Gln	Ala Leu His Leu Leu	Gly Asp Lys Pro Tyr	Leu 1805		
1795	1800	1805			
CCG CTG AGT ACG ACA	TGG AGT GAT CCA CGA	CTA GAC AGA GCC GCG	GAT 5472		
Pro Leu Ser Thr Thr	Trp Ser Asp Pro Arg	Leu Asp Arg Ala Ala	Asp 1820		
1810	1815	1820			
ATC ACT ACC CAA AAT	GCT CAC GAC AGC GCA	ATA GTC GCT CTG CGG	CAG 5520		
Ile Thr Thr Gln Asn	Ala His Asp Ser Ala	Ile Val Ala Leu Arg	Gln 1840		
1825	1830	1835	1840		
AAT ATA CCT ACA CCG	GCA CCT TTA TCA		5547		
Asn Ile Pro Thr Pro	Ala Pro Leu Ser		1849		
1845	1849				

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 49:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĹŽKA: 1849 aminokyselín
- (B) TYP: aminokyseliny
- (C) REŤAZEC: jednoduchý
- (D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: proteín

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 49 (TcdA<sub>ii</sub>):

Charakt. znak	od	do	Opis
Peptid	1	1849	TcdA <sub>ii</sub> peptid
Fragment	1	12	TcdA <sub>ii</sub> N-koniec (SEKV. ID. ČÍS. 13)
Fragment	196	211	(SEKV. ID. ČÍS. 38)
Fragment	466	475	(SEKV. ID. ČÍS. 17)
Fragment	993	1004	(SEKV. ID. ČÍS. 23;12/13)
Fragment	1297	1312	(SEKV. ID. ČÍS. 18)
Fragment	1390	1409	(SEKV. ID. ČÍS. 39)
Fragment	1532	1554	(SEKV. ID. ČÍS. 21;19/23)

Leu Ile Gly Tyr Asn Asn Gln Phe Ser Gly Arg Ala Ser Gln Tyr Val  
 1 5 10 15

Ala Pro Gly Thr Val Ser Ser Met Phe Ser Pro Ala Ala Tyr Leu Thr  
 20 25 30

Glu Leu Tyr Arg Glu Ala Arg Asn Leu His Ala Ser Asp Ser Val Tyr  
 35 40 45

Tyr Leu Asp Thr Arg Arg Pro Asp Leu Lys Ser Met Ala Leu Ser Gln  
 50 55 60

Gln Asn Met Asp Ile Glu Leu Ser Thr Leu Ser Leu Ser Asn Glu Leu  
 65 70 75 80

Leu Leu Glu Ser Ile Lys Thr Glu Ser Lys Leu Glu Asn Tyr Thr Lys  
 85 90 95

Val Met Glu Met Leu Ser Thr Phe Arg Pro Ser Gly Ala Thr Pro Tyr  
 100 105 110

His Asp Ala Tyr Glu Asn Val Arg Glu Val Ile Gln Leu Gln Asp Pro  
 115 120 125

Gly Leu Glu Gln Leu Asn Ala Ser Pro Ala Ile Ala Gly Leu Met His  
 130 135 140

Gln Ala Ser Leu Leu Gly Ile Asn Ala Ser Ile Ser Pro Glu Leu Phe  
 145 150 155 160



Asn Ile Leu Thr Glu Glu Ile Thr Glu Gly Asn Ala Glu Glu Leu Tyr  
165 170 175

Lys Lys Asn Phe Gly Asn Ile Glu Pro Ala Ser Leu Ala Met Pro Glu  
180 185 190

Tyr Leu Lys Arg Tyr Tyr Asn Leu Ser Asp Glu Glu Leu Ser Gln Phe  
195 200 205

Ile Gly Lys Ala Ser Asn Phe Gly Gln Gln Glu Tyr Ser Asn Asn Gln  
210 215 220

Leu Ile Thr Pro Val Val Asn Ser Ser Asp Gly Thr Val Lys Val Tyr  
225 230 235 240

Arg Ile Thr Arg Glu Tyr Thr Thr Asn Ala Tyr Gln Met Asp Val Glu  
245 250 255

Leu Phe Pro Phe Gly Gly Glu Asn Tyr Arg Leu Asp Tyr Lys Phe Lys  
260 265 270

Asn Phe Tyr Asn Ala Ser Tyr Leu Ser Ile Lys Leu Asn Asp Lys Arg  
275 285

Glu Leu Val Arg Thr Glu Gly Ala Pro Gln Val Asn Ile Glu Tyr Ser  
290 295 300

Ala Asn Ile Thr Leu Asn Thr Ala Asp Ile Ser Gln Pro Phe Glu Ile  
305 310 315 320

Gly Leu Thr Arg Val Leu Pro Ser Gly Ser Trp Ala Tyr Ala Ala Ala  
325 330 335

Lys Phe Thr Val Glu Glu Tyr Asn Gln Tyr Ser Phe Leu Leu Lys Leu  
340 345 350

Asn Lys Ala Ile Arg Leu Ser Arg Ala Thr Glu Leu Ser Pro Thr Ile  
355 360 365

Leu Glu Gly Ile Val Arg Ser Val Asn Leu Gln Leu Asp Ile Asn Thr  
370 375 380

Asp Val Leu Gly Lys Val Phe Leu Thr Lys Tyr Tyr Met Gln Arg Tyr  
385 390 395 400

Ala Ile His Ala Glu Thr Ala Leu Ile Leu Cys Asn Ala Pro Ile Ser  
405 410 415

Gln Arg Ser Tyr Asp Asn Gln Pro Ser Gln Phe Asp Arg Leu Phe Asn  
420 425 430

Thr Pro Leu Leu Asn Gly Gln Tyr Phe Ser Thr Gly Asp Glu Glu Ile  
435 440 445

Asp Leu Asn Ser Gly Ser Thr Gly Asp Trp Arg Lys Thr Ile Leu Lys  
450 455 460

Arg Ala Phe Asn Ile Asp Asp Val Ser Leu Phe Arg Leu Leu Lys Ile  
465 470 475 480

Thr Asp His Asp Asn Lys Asp Gly Lys Ile Lys Asn Asn Leu Lys Asn  
485 490 495

Leu Ser Asn Leu Tyr Ile Gly Lys Leu Leu Ala Asp Ile His Gln Leu  
500 505 510

Thr Ile Asp Glu Leu Asp Leu Leu Leu Ile Ala Val Gly Glu Gly Lys  
515 520 525

Thr Asn Leu Ser Ala Ile Ser Asp Lys Gln Leu Ala Thr Leu Ile Arg  
530 535 540

Lys Leu Asn Thr Ile Thr Ser Trp Leu His Thr Gln Lys Trp Ser Val  
545 550 555 560

Phe Gln Leu Phe Ile Met Thr Ser Thr Ser Tyr Asn Lys Thr Leu Thr  
565 570 575

Pro Glu Ile Lys Asn Leu Leu Asp Thr Val Tyr His Gly Leu Gln Gly  
580 585 590

Phe Asp Lys Asp Lys Ala Asp Leu Leu His Val Met Ala Pro Tyr Ile  
595 600 605

Ala Ala Thr Leu Gln Leu Ser Ser Glu Asn Val Ala His Ser Val Leu  
610 615 620

Leu Trp Ala Asp Lys Leu Gln Pro Gly Asp Gly Ala Met Thr Ala Glu  
625 630 635 640

Lys Phe Trp Asp Trp Leu Asn Thr Lys Tyr Thr Pro Gly Ser Ser Glu  
645 650 655

Ala Val Glu Thr Gln Glu His Ile Val Gln Tyr Cys Gln Ala Leu Ala  
660 665 670

Gln Leu Glu Met Val Tyr His Ser Thr Gly Ile Asn Glu Asn Ala Phe  
675 680 685

Arg Leu Phe Val Thr Lys Pro Glu Met Phe Gly Ala Ala Thr Gly Ala  
690 695 700

Ala Pro Ala His Asp Ala Leu Ser Leu Ile Met Leu Thr Arg Phe Ala  
705 710 715 720

Asp Trp Val Asn Ala Leu Gly Glu Lys Ala Ser Ser Val Leu Ala Ala  
725 730 735

Phe Glu Ala Asn Ser Leu Thr Ala Glu Gln Leu Ala Asp Ala Met Asn  
740 745 750

Leu Asp Ala Asn Leu Leu Leu Gln Ala Ser Ile Gln Ala Gln Asn His  
755 760 765

Gln His Leu Pro Pro Val Thr Pro Glu Asn Ala Phe Ser Cys Trp Thr  
770 775 780

Ser Ile Asn Thr Ile Leu Gln Trp Val Asn Val Ala Gln Gln Leu Asn  
785 790 795 800

Val Ala Pro Gln Gly Val Ser Ala Leu Val Gly Leu Asp Tyr Ile Gln  
805 810 815

Ser Met Lys Glu Thr Pro Thr Tyr Ala Gln Trp Glu Asn Ala Ala Gly  
820 825 830

Val Leu Thr Ala Gly Leu Asn Ser Gln Gln Ala Asn Thr Leu His Ala  
835 840 845

Phe Leu Asp Glu Ser Arg Ser Ala Ala Leu Ser Thr Tyr Tyr Ile Arg  
850 855 860

Gln Val Ala Lys Ala Ala Ala Ala Ile Lys Ser Arg Asp Asp Leu Tyr  
865 870 875 880

Gln Tyr Leu Leu Ile Asp Asn Gln Val Ser Ala Ala Ile Lys Thr Thr  
885 890 895

Arg ~~Trp~~ Ala Glu Ala Ile Ala Ser Ile Gln Leu Tyr Val Asn Arg Ala  
900 905 910  
Leu Glu Asn Val Glu Glu Asn Ala Asn Ser Gly Val Ile Ser Arg Gln  
915 920 925  
Phe Phe Ile Asp Trp Asp Lys Tyr Asn Lys Arg Tyr Ser Thr Trp Ala  
930 935 940  
Gly Val Ser Gln Leu Val Tyr Tyr Pro Glu Asn Tyr Ile Asp Pro Thr  
945 950 955 960  
Met Arg Ile Gly Gln Thr Lys Met Met Asp Ala Leu Leu Gln Ser Val  
965 970 975  
Ser Gln Ser Gln Leu Asn Ala Asp Thr Val Glu Asp Ala Phe Met Ser  
980 985 990  
Tyr Leu Thr Ser Phe Glu Gln Val Ala Asn Leu Lys Val Ile Ser Ala  
995 1000 1005  
Tyr His Asp Asn Ile Asn Asn Asp Gln Gly Leu Thr Tyr Phe Ile Gly  
1010 1015 1020  
Leu Ser Glu Thr Asp Ala Gly Glu Tyr Tyr Trp Arg Ser Val Asp His  
1025 1030 1035 1040  
Ser Lys Phe Asn Asp Gly Lys Phe Ala Ala Asn Ala Trp Ser Glu Trp  
1045 1050 1055  
His Lys Ile Asp Cys Pro Ile Asn Pro Tyr Lys Ser Thr Ile Arg Pro  
1060 1065 1070  
Val Ile Tyr Lys Ser Arg Leu Tyr Leu Leu Trp Leu Glu Gln Lys Glu ---  
1075 1080 1085  
Ile Thr Lys Gln Thr Gly Asn Ser Lys Asp Gly Tyr Gln Thr Glu Thr  
1090 1095 1100  
Asp Tyr Arg Tyr Glu Leu Lys Leu Ala His Ile Arg Tyr Asp Gly Thr  
1105 1110 1115 1120  
Trp Asn Thr Pro Ile Thr Phe Asp Val Asn Lys Lys Ile Ser Glu Leu  
1125 1130 1135  
Lys Leu Glu Lys Asn Arg Ala Pro Gly Leu Tyr Cys Ala Gly Tyr Gln  
1140 1145 1150  
Gly Glu Asp Thr Leu Leu Val Met Phe Tyr Asn Gln Gln Asp Thr Leu  
1155 1160 1165  
Asp Ser Tyr Lys Asn Ala Ser Met Gln Gly Leu Tyr Ile Phe Ala Asp  
1170 1175 1180  
Met Ala Ser Lys Asp Met Thr Pro Glu Gln Ser Asn Val Tyr Arg Asp  
1185 1190 1195 1200  
Asn Ser Tyr Gln Gln Phe Asp Thr Asn Asn Val Arg Arg Val Asn Asn  
1205 1210 1215  
Arg Tyr Ala Glu Asp Tyr Glu Ile Pro Ser Ser Val Ser Ser Arg Lys  
1220 1225 1230  
Asp Tyr Gly Trp Gly Asp Tyr Tyr Leu Ser Met Val Tyr Asn Gly Asp  
1235 1240 1245  
Ile Pro Thr Ile Asn Tyr Lys Ala Ala Ser Ser Asp Leu Lys Ile Tyr  
1250 1255 1260

Ile Ser Pro Lys Leu Arg Ile Ile His Asn Gly Tyr Glu Gly Gln Lys  
1265 1270 1275 1280

Arg Asn Gln Cys Asn Leu Met Asn Lys Tyr Gly Lys Leu Gly Asp Lys  
1285 1290 1295

Phe Ile Val Tyr Thr Ser Leu Gly Val Asn Pro Asn Asn Ser Ser Asn  
1300 1305 1310

Lys Leu Met Phe Tyr Pro Val Tyr Gln Tyr Ser Gly Asn Thr Ser Gly  
1315 1320 1325

Leu Asn Gln Gly Arg Leu Leu Phe His Arg Asp Thr Thr Tyr Pro Ser  
1330 1335 1340

Lys Val Glu Ala Trp Ile Pro Gly Ala Lys Arg Ser Leu Thr Asn Gln  
1345 1350 1355 1360

Asn Ala Ala Ile Gly Asp Asp Tyr Ala Thr Asp Ser Leu Asn Lys Pro  
1365 1370 1375

Asp Asp Leu Lys Gln Tyr Ile Phe Met Thr Asp Ser Lys Gly Thr Ala  
1380 1385 1390

Thr Asp Val Ser Gly Pro Val Glu Ile Asn Thr Ala Ile Ser Pro Ala  
1395 1400 1405

Lys Val Gln Ile Ile Val Lys Ala Gly Gly Lys Glu Gln Thr Phe Thr  
1410 1415 1420

Ala Asp Lys Asp Val Ser Ile Gln Pro Ser Pro Ser Phe Asp Glu Met  
1425 1430 1435 1440

Asn Tyr Gln Phe Asn Ala Leu Glu Ile Asp Gly Ser Gly Leu Asn Phe  
1445 1450 1455

Ile Asn Asn Ser Ala Ser Ile Asp Val Thr Phe Thr Ala Phe Ala Glu  
1460 1465 1470

Asp Gly Arg Lys Leu Gly Tyr Glu Ser Phe Ser Ile Pro Val Thr Leu  
1475 1480 1485

Lys Val Ser Thr Asp Asn Ala Leu Thr Leu His His Asn Glu Asn Gly  
1490 1495 1500

Ala Gln Tyr Met Gln Trp Gln Ser Tyr Arg Thr Arg Leu Asn Thr Leu  
1505 1510 1515 1520

Phe Ala Arg Gln Leu Val Ala Arg Ala Thr Thr Gly Ile Asp Thr Ile  
1525 1530 1535

Leu Ser Met Glu Thr Gln Asn Ile Gln Glu Pro Gln Leu Gly Lys Gly  
1540 1545 1550

Phe Tyr Ala Thr Phe Val Ile Pro Pro Tyr Asn Leu Ser Thr His Gly  
1555 1560 1565

Asp Glu Arg Trp Phe Lys Leu Tyr Ile Lys His Val Val Asp Asn Asn  
1570 1575 1580

Ser His Ile Ile Tyr Ser Gly Gln Leu Thr Asp Thr Asn Ile Asn Ile  
1585 1590 1595 1600

Thr Leu Phe Ile Pro Leu Asp Asp Val Pro Leu Asn Gln Asp Tyr His  
1605 1610 1615

Ala Lys Val Tyr Met Thr Phe Lys Lys Ser Pro Ser Asp Gly Thr Trp  
1620 1625 1630

Trp Gly Pro His Phe Val Arg Asp Asp Lys Gly Ile Val Thr Ile Asn  
1635 1640 1645

Pro Lys Ser Ile Leu Thr His Phe Glu Ser Val Asn Val Leu Asn Asn  
1650 1655 1660

Ile Ser Ser Glu Pro Met Asp Phe Ser Gly Ala Asn Ser Leu Tyr Phe  
1665 1670 1675 1680

Trp Glu Leu Phe Tyr Tyr Thr Pro Met Leu Val Ala Gln Arg Leu Leu  
1685 1690 1695

His Glu Gln Asn Phe Asp Glu Ala Asn Arg Trp Leu Lys Tyr Val Trp  
1700 1705 1710

Ser Pro Ser Gly Tyr Ile Val His Gly Gln Ile Gln Asn Tyr Gln Trp  
1715 1720 1725

Asn Val Arg Pro Leu Leu Glu Asp Thr Ser Trp Asn Ser Asp Pro Leu  
1730 1735 1740

Asp Ser Val Asp Pro Asp Ala Val Ala Gln His Asp Pro Met His Tyr  
1745 1750 1755 1760

Lys Val Ser Thr Phe Met Arg Thr Leu Asp Leu Leu Ile Ala Arg Gly  
1765 1770 1775

Asp His Ala Tyr Arg Gln Leu Glu Arg Asp Thr Leu Asn Glu Ala Lys  
1780 1785 1790

Met Trp Tyr Met Gln Ala Leu His Leu Leu Gly Asp Lys Pro Tyr Leu  
1795 1800 1805

Pro Leu Ser Thr Thr Trp Ser Asp Pro Arg Leu Asp Arg Ala Ala Asp  
1810 1815 1820

Ile Thr Thr Gln Asn Ala His Asp Ser Ala Ile Val Ala Leu Arg Gln  
1825 1830 1835 1840

Asn Ile Pro Thr Pro Ala Pro Leu Ser  
1845 1849

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 50:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĹŽKA: 1740 párov báz
- (B) TYP: nukleová kyselina
- (C) REŤAZEC: dvojitý
- (D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: DNA (genómová)

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 50 (kódovaná oblasť  
TcdA<sub>iii</sub>):

TTG	CGC	AGC	GCT	AAT	ACC	CTG	ACT	GAT	CTC	TTC	CTG	CCG	CAA	ATC	AAT	48
Leu	Arg	Ser	Ala	Asn	Thr	Leu	Thr	Asp	Leu	Phe	Leu	Pro	Gln	Ile	Asn	
1				5					10					15		
GAA	GTG	ATG	ATG	AAT	TAC	TGG	CAG	ACA	TTA	GCT	CAG	AGA	GTA	TAC	AAT	96
Glu	Val	Met	Met	Asn	Tyr	Trp	Gln	Thr	Leu	Ala	Gln	Arg	Val	Tyr	Asn	
			20				25						30			
CTG	CGT	CAT	AAC	CTC	TCT	ATC	GAC	GGC	CAG	CCG	TTA	TAT	CTG	CCA	ATC	144
Leu	Arg	His	Asn	Leu	Ser	Ile	Asp	Gly	Gln	Pro	Leu	Tyr	Leu	Pro	Ile	
		35					40					45				
TAT	GCC	ACA	CCG	GCC	GAT	CCG	AAA	GCG	TTA	CTC	AGC	GCC	GCC	GTT	GCC	192
Tyr	Ala	Thr	Pro	Ala	Asp	Pro	Lys	Ala	Leu	Leu	Ser	Ala	Ala	Val	Ala	
	50					55					60					
ACT	TCT	CAA	GGT	GGA	GGC	AAG	CTA	CCG	GAA	TCA	TTT	ATG	TCC	CTG	TGG	240
Thr	Ser	Gln	Gly	Gly	Gly	Lys	Leu	Pro	Glu	Ser	Phe	Met	Ser	Leu	Trp	
65					70					75					80	
CGT	TTC	CCG	CAC	ATG	CTG	GAA	AAT	GCG	CGC	GGC	ATG	GTT	AGC	CAG	CTC	288
Arg	Phe	Pro	His	Met	Leu	Glu	Asn	Ala	Arg	Gly	Met	Val	Ser	Gln	Leu	
				85					90					95		
ACC	CAG	TTC	GGC	TCC	ACG	TTA	CAA	AAT	ATT	ATC	GAA	CGT	CAG	GAC	GCG	336
Thr	Gln	Phe	Gly	Ser	Thr	Leu	Gln	Asn	Ile	Ile	Glu	Arg	Gln	Asp	Ala	
			100					105					110			
GAA	GCG	CTC	AAT	GCG	TTA	TTA	CAA	AAT	CAG	GCC	GCC	GAG	CTG	ATA	TTG	384
Glu	Ala	Leu	Asn	Ala	Leu	Leu	Gln	Asn	Gln	Ala	Ala	Glu	Leu	Ile	Leu	
			115				120					125				
ACT	AAC	CTG	AGC	ATT	CAG	GAC	AAA	ACC	ATT	GAA	GAA	TTG	GAT	GCC	GAG	432
Thr	Asn	Leu	Ser	Ile	Gln	Asp	Lys	Thr	Ile	Glu	Glu	Leu	Asp	Ala	Glu	
	130					135					140					
AAA	ACG	GTG	TTG	GAA	AAA	TCC	AAA	GCG	GGA	GCA	CAA	TCG	CGC	TTT	GAT	480
Lys	Thr	Val	Leu	Glu	Lys	Ser	Lys	Ala	Gly	Ala	Gln	Ser	Arg	Phe	Asp	
145					150					155					160	
AGC	TAC	GGC	AAA	CTG	TAC	GAT	GAG	AAT	ATC	AAC	GCC	GGT	GAA	AAC	CAA	528
Ser	Tyr	Gly	Lys	Leu	Tyr	Asp	Glu	Asn	Ile	Asn	Ala	Gly	Glu	Asn	Gln	
				165					170					175		
GCC	ATG	ACG	CTA	CGA	GCG	TCC	GCC	GCC	GGG	CTT	ACC	ACG	GCA	GTT	CAG	576
Ala	Met	Thr	Leu	Arg	Ala	Ser	Ala	Ala	Gly	Leu	Thr	Thr	Ala	Val	Gln	
			180					185					190			
GCA	TCC	CGT	CTG	GCC	GGT	GCG	GCG	GCT	GAT	CTG	GTG	CCT	AAC	ATC	TTC	624
Ala	Ser	Arg	Leu	Ala	Gly	Ala	Ala	Ala	Asp	Leu	Val	Pro	Asn	Ile	Phe	
		195				200						205				
GGC	TTT	GCC	GGT	GGC	GGC	AGC	CGT	TGG	GGG	GCT	ATC	GCT	GAG	GCG	ACA	672
Gly	Phe	Ala	Gly	Gly	Gly	Ser	Arg	Trp	Gly	Ala	Ile	Ala	Glu	Ala	Thr	
	210					215					220					
GGT	TAT	GTG	ATG	GAA	TTC	TCC	GCG	AAT	GTT	ATG	AAC	ACC	GAA	GCG	GAT	720
Gly	Tyr	Val	Met	Glu	Phe	Ser	Ala	Asn	Val	Met	Asn	Thr	Glu	Ala	Asp	
225					230					235					240	
AAA	ATT	AGC	CAA	TCT	GAA	ACC	TAC	CGT	CGT	CGC	CGT	CAG	GAG	TGG	GAG	768
Lys	Ile	Ser	Gln	Ser	Glu	Thr	Tyr	Arg	Arg	Arg	Arg	Gln	Glu	Trp	Glu	
				245				250						255		
ATC	CAG	CGG	AAT	AAT	GCC	GAA	GCG	GAA	TTG	AAG	CAA	ATC	GAT	GCT	CAG	816
Ile	Gln	Arg	Asn	Asn	Ala	Glu	Ala	Glu	Leu	Lys	Gln	Ile	Asp	Ala	Gln	
			260					265					270			

CTC AAA TCA CTC GCT GTA CGC CGC GAA GCC GCC GTA TTG CAG AAA ACC 864  
Leu Lys Ser Leu Ala Val Arg Arg Glu Ala Ala Val Leu Gln Lys Thr  
275 280 285

AGT CTG AAA ACC CAA CAA GAA CAG ACC CAA TCT CAA TTG GCC TTC CTG 912  
Ser Leu Lys Thr Gln Gln Glu Gln Thr Gln Ser Gln Leu Ala Phe Leu  
290 295 300

CAA CGT AAG TTC AGC AAT CAG GCG TTA TAC AAC TGG CTG CGT GGT CGA 960  
Gln Arg Lys Phe Ser Asn Gln Ala Leu Tyr Asn Trp Leu Arg Gly Arg  
305 310 315 320

CTG GCG GCG ATT TAC TTC CAG TTC TAC GAT TTG GCC GTC GCG CGT TGC 1008  
Leu Ala Ala Ile Tyr Phe Gln Phe Tyr Asp Leu Ala Val Ala Arg Cys  
325 330 335

CTG ATG GCA GAA CAA GCT TAC CGT TGG GAA CTC AAT GAT GAC TCT GCC 1056  
Leu Met Ala Glu Gln Ala Tyr Arg Trp Glu Leu Asn Asp Asp Ser Ala  
340 345 350

CGC TTC ATT AAA CCG GGC GCC TGG CAG GGA ACC TAT GCC GGT CTG CTT 1104  
Arg Phe Ile Lys Pro Gly Ala Trp Gln Gly Thr Tyr Ala Gly Leu Leu  
355 360 365

GCA GGT GAA ACC TTG ATG CTG AGT CTG GCA CAA ATG GAA GAC GCT CAT 1152  
Ala Gly Glu Thr Leu Met Leu Ser Leu Ala Gln Met Glu Asp Ala His  
370 375 380

CTG AAA CGC GAT AAA CGC GCA TTA GAG GTT GAA CGC ACA GTA TCG CTG 1200  
Leu Lys Arg Asp Lys Arg Ala Leu Glu Val Glu Arg Thr Val Ser Leu  
385 390 395 400

GCC GAA GTT TAT GCA GGA TTA CCA AAA GAT AAC GGT CCA TTT TCC CTG 1248  
Ala Glu Val Tyr Ala Gly Leu Pro Lys Asp Asn Gly Pro Phe Ser Leu  
405 410 415

GCT CAG GAA ATT GAC AAG CTG GTG AGT CAA GGT TCA GGC AGT GCC GGC 1296  
Ala Gln Glu Ile Asp Lys Leu Val Ser Gln Gly Ser Gly Ser Ala Gly  
420 425 430

AGT GGT AAT AAT AAT TTG GCG TTC GGC GCC GGC ACG GAC ACT AAA ACC 1344  
Ser Gly Asn Asn Asn Leu Ala Phe Gly Ala Gly Thr Asp Thr Lys Thr  
435 440 445

TCT TTG CAG GCA TCA GTT TCA TTC GCT GAT TTG AAA ATT CGT GAA GAT 1392  
Ser Leu Gln Ala Ser Val Ser Phe Ala Asp Leu Lys Ile Arg Glu Asp  
450 455 460

TAC CCG GCA TCG CTT GGC AAA ATT CGA CGT ATC AAA CAG ATC AGC GTC 1440  
Tyr Pro Ala Ser Leu Gly Lys Ile Arg Arg Ile Lys Gln Ile Ser Val  
465 470 475 480

ACT TTG CCC GCG CTA CTG GGA CCG TAT CAG GAT GTA CAG GCA ATA TTG 1488  
Thr Leu Pro Ala Leu Leu Gly Pro Tyr Gln Asp Val Gln Ala Ile Leu  
485 490 495

TCT TAC GGC GAT AAA GCC GGA TTA GCT AAC GGC TGT GAA GCG CTG GCA 1536  
Ser Tyr Gly Asp Lys Ala Gly Leu Ala Asn Gly Cys Glu Ala Leu Ala  
500 505 510

GTT TCT CAC GGT ATG AAT GAC AGC GGC CAA TTC CAG CTC GAT TTC AAC 1584  
Val Ser His Gly Met Asn Asp Ser Gly Gln Phe Gln Leu Asp Phe Asn  
515 520 525

GAT GGC AAA TTC CTG CCA TTC GAA GGC ATC GCC ATT GAT CAA GGC ACG 1632  
Asp Gly Lys Phe Leu Pro Phe Glu Gly Ile Ala Ile Asp Gln Gly Thr  
530 535 540

CTG ACA CTG AGC TTC CCA AAT GCA TCT ATG CCG GAG AAA GGT AAA CAA 1680  
 Leu Thr Leu Ser Phe Pro Asn Ala Ser Met Pro Glu Lys Gly Lys Gln  
 545 550 555 560

GCC ACT ATG TTA AAA ACC CTG AAC GAT ATC ATT TTG CAT ATT CGC TAC 1728  
 Ala Thr Met Leu Lys Thr Leu Asn Asp Ile Ile Leu His Ile Arg Tyr  
 565 570 575

ACC ATT AAA TAA 1740  
 Thr Ile Lys ...  
 579

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍŠ. 51:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĚLKA: 579 aminokyselín
- (B) TYP: aminokyseliny
- (C) REŤAZEC: jednoduchý
- (D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: proteín

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍŠ. 51 (TcdA<sub>iii</sub>):

Leu Arg Ser Ala Asn Thr Leu Thr Asp Leu Phe Leu Pro Gln Ile Asn  
 1 5 10 15

Glu Val Met Met Asn Tyr Trp Gln Thr Leu Ala Gln Arg Val Tyr Asn  
 20 25 30

Leu Arg His Asn Leu Ser Ile Asp Gly Gln Pro Leu Tyr Leu Pro Ile  
 35 40 45

Tyr Ala Thr Pro Ala Asp Pro Lys Ala Leu Leu Ser Ala Ala Val Ala  
 50 55 60

Thr Ser Gln Gly Gly Gly Lys Leu Pro Glu Ser Phe Met Ser Leu Trp  
 65 70 75 80

Arg Phe Pro His Met Leu Glu Asn Ala Arg Gly Met Val Ser Gln Leu  
 85 90 95

Thr Gln Phe Gly Ser Thr Leu Gln Asn Ile Ile Glu Arg Gln Asp Ala  
 100 105 110

Glu Ala Leu Asn Ala Leu Leu Gln Asn Gln Ala Ala Glu Leu Ile Leu  
 115 120 125

Thr Asn Leu Ser Ile Gln Asp Lys Thr Ile Glu Glu Leu Asp Ala Glu  
 130 135 140

Lys Thr Val Leu Glu Lys Ser Lys Ala Gly Ala Gln Ser Arg Phe Asp  
 145 150 155 160

Ser Tyr Gly Lys Leu Tyr Asp Glu Asn Ile Asn Ala Gly Glu Asn Gln  
 165 170 175



Ala Met Thr Leu Arg Ala Ser Ala Ala Gly Leu Thr Thr Ala Val Gln  
180 185 190

Ala Ser Arg Leu Ala Gly Ala Ala Ala Asp Leu Val Pro Asn Ile Phe  
195 200 205

Gly Phe Ala Gly Gly Gly Ser Arg Trp Gly Ala Ile Ala Glu Ala Thr  
210 215 220

Gly Tyr Val Met Glu Phe Ser Ala Asn Val Met Asn Thr Glu Ala Asp  
225 230 235 240

Lys Ile Ser Gln Ser Glu Thr Tyr Arg Arg Arg Arg Gln Glu Trp Glu  
245 250 255

Ile Gln Arg Asn Asn Ala Glu Ala Glu Leu Lys Gln Ile Asp Ala Gln  
260 265 270

Leu Lys Ser Leu Ala Val Arg Arg Glu Ala Ala Val Leu Gln Lys Thr  
275 280 285

Ser Leu Lys Thr Gln Gln Glu Gln Thr Gln Ser Gln Leu Ala Phe Leu  
290 295 300

Gln Arg Lys Phe Ser Asn Gln Ala Leu Tyr Asn Trp Leu Arg Gly Arg  
305 310 315 320

Leu Ala Ala Ile Tyr Phe Gln Phe Tyr Asp Leu Ala Val Ala Arg Cys  
325 330 335

Leu Met Ala Glu Gln Ala Tyr Arg Trp Glu Leu Asn Asp Asp Ser Ala  
340 345 350

Arg Phe Ile Lys Pro Gly Ala Trp Gln Gly Thr Tyr Ala Gly Leu Leu  
355 360 365

Ala Gly Glu Thr Leu Met Leu Ser Leu Ala Gln Met Glu Asp Ala His  
370 375 380

Leu Lys Arg Asp Lys Arg Ala Leu Glu Val Glu Arg Thr Val Ser Leu  
385 390 395 400

Ala Glu Val Tyr Ala Gly Leu Pro Lys Asp Asn Gly Pro Phe Ser Leu  
405 410 415

Ala Gln Glu Ile Asp Lys Leu Val Ser Gln Gly Ser Gly Ser Ala Gly  
420 425 430

Ser Gly Asn Asn Asn Leu Ala Phe Gly Ala Gly Thr Asp Thr Lys Thr  
435 440 445

Ser Leu Gln Ala Ser Val Ser Phe Ala Asp Leu Lys Ile Arg Glu Asp  
450 455 460

Tyr Pro Ala Ser Leu Gly Lys Ile Arg Arg Ile Lys Gln Ile Ser Val  
465 470 475 480

Thr Leu Pro Ala Leu Leu Gly Pro Tyr Gln Asp Val Gln Ala Ile Leu  
485 490 495

Ser Tyr Gly Asp Lys Ala Gly Leu Ala Asn Gly Cys Glu Ala Leu Ala  
500 505 510

Val Ser His Gly Met Asn Asp Ser Gly Gln Phe Gln Leu Asp Phe Asn  
515 520 525

Asp Gly Lys Phe Leu Pro Phe Glu Gly Ile Ala Ile Asp Gln Gly Thr  
530 535 540

Leu Thr Leu Ser Phe Pro Asn Ala Ser Met Pro Glu Lys Gly Lys Gln  
 545 550 555 560  
 Ala Thr Met Leu Lys Thr Leu Asn Asp Ile Ile Leu His Ile Arg Tyr  
 565 570 575  
 Thr Ile Lys ...  
 579

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 52:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĚLKA: 5532 párov báz
- (B) TYP: nukleová kyselina
- (C) REŤAZEC: dvojitý
- (D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: DNA (genómová)

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 52 (kódovaná oblasť tcbA<sub>ii</sub>):

TTT	ATA	CAA	GGT	TAT	AGT	GAT	CTG	TTT	GGT	AAT	CGT	GCT	GAT	AAC	TAT	48
Phe	Ile	Gln	Gly	Tyr	Ser	Asp	Leu	Phe	Gly	Asn	Arg	Ala	Asp	Asn	Tyr	
1				5				10						15		
GCC	GCG	CCG	GGC	TCG	GTT	GCA	TCG	ATG	TTC	TCA	CCG	GCG	GCT	TAT	TTG	96
Ala	Ala	Pro	Gly	Ser	Val	Ala	Ser	Met	Phe	Ser	Pro	Ala	Ala	Tyr	Leu	
			20					25						30		
ACG	GAA	TTG	TAC	CGT	GAA	GCC	AAA	AAC	TTG	CAT	GAC	AGC	AGC	TCA	ATT	144
Thr	Glu	Leu	Tyr	Arg	Glu	Ala	Lys	Asn	Leu	His	Asp	Ser	Ser	Ser	Ile	
		35					40					45				
TAT	TAC	CTA	GAT	AAA	CGT	CGC	CCG	GAT	TTA	GCA	AGC	TTA	ATG	CTC	AGC	192
Tyr	Tyr	Leu	Asp	Lys	Arg	Arg	Pro	Asp	Leu	Ala	Ser	Leu	Met	Leu	Ser	
		50				55					60					
CAG	AAA	AAT	ATG	GAT	GAG	GAA	ATT	TCA	ACG	CTG	GCT	CTC	TCT	AAT	GAA	240
Gln	Lys	Asn	Met	Asp	Glu	Glu	Ile	Ser	Thr	Leu	Ala	Leu	Ser	Asn	Glu	
65					70					75					80	
TTG	TGC	CTT	GCC	GGG	ATC	GAA	ACA	AAA	ACA	GGA	AAA	TCA	CAA	GAT	GAA	288
Leu	Cys	Leu	Ala	Gly	Ile	Glu	Thr	Lys	Thr	Gly	Lys	Ser	Gln	Asp	Glu	
				85					90					95		
GTG	ATG	GAT	ATG	TTG	TCA	ACT	TAT	CGT	TTA	AGT	GGA	GAG	ACA	CCT	TAT	336
Val	Met	Asp	Met	Leu	Ser	Thr	Tyr	Arg	Leu	Ser	Gly	Glu	Thr	Pro	Tyr	
			100					105					110			
CAT	CAC	GCT	TAT	GAA	ACT	GTT	CGT	GAA	ATC	GTT	CAT	GAA	CGT	GAT	CCA	384
His	His	Ala	Tyr	Glu	Thr	Val	Arg	Glu	Ile	Val	His	Glu	Arg	Asp	Pro	
			115				120						125			
GGA	TTT	CGT	CAT	TTG	TCA	CAG	GCA	CCC	ATT	GTT	GCT	GCT	AAG	CTC	GAT	432
Gly	Phe	Arg	His	Leu	Ser	Gln	Ala	Pro	Ile	Val	Ala	Ala	Lys	Leu	Asp	
		130				135						140				
CCT	GTG	ACT	TTG	TTG	GGT	ATT	AGC	TCC	CAT	ATT	TCG	CCA	GAA	CTG	TAT	480

Pro	Val	Thr	Leu	Leu	Gly	Ile	Ser	Ser	His	Ile	Ser	Pro	Glu	Leu	Tyr	
145					150					155					160	
AAC	TTG	CTG	ATT	GAG	GAG	ATC	CCG	GAA	AAA	GAT	GAA	GCC	GCG	CTT	GAT	528
Asn	Leu	Leu	Ile	Glu	Glu	Ile	Pro	Glu	Lys	Asp	Glu	Ala	Ala	Leu	Asp	
				165					170					175		
ACG	CTT	TAT	AAA	ACA	AAC	TTT	GGC	GAT	ATT	ACT	ACT	GCT	CAG	TTA	ATG	576
Thr	Leu	Tyr	Lys	Thr	Asn	Phe	Gly	Asp	Ile	Thr	Thr	Ala	Gln	Leu	Met	
			180					185					190			
TCC	CCA	AGT	TAT	CTG	GCC	CGG	TAT	TAT	GGC	GTC	TCA	CCG	GAA	GAT	ATT	624
Ser	Pro	Ser	Tyr	Leu	Ala	Arg	Tyr	Tyr	Gly	Val	Ser	Pro	Glu	Asp	Ile	
		195					200					205				
GCC	TAC	GTG	ACG	ACT	TCA	TTA	TCA	CAT	GTT	GGA	TAT	AGC	AGT	GAT	ATT	672
Ala	Tyr	Val	Thr	Thr	Ser	Leu	Ser	His	Val	Gly	Tyr	Ser	Ser	Asp	Ile	
	210					215					220					
CTG	GTT	ATT	CCG	TTG	GTC	GAT	GGT	GTG	GGT	AAG	ATG	GAA	GTA	GTT	CGT	720
Leu	Val	Ile	Pro	Leu	Val	Asp	Gly	Val	Gly	Lys	Met	Glu	Val	Val	Arg	
225					230					235					240	
GTT	ACC	CGA	ACA	CCA	TCG	GAT	AAT	TAT	ACC	AGT	CAG	ACG	AAT	TAT	ATT	768
Val	Thr	Arg	Thr	Pro	Ser	Asp	Asn	Tyr	Thr	Ser	Gln	Thr	Asn	Tyr	Ile	
				245					250					255		
GAG	CTG	TAT	CCA	CAG	GGT	GGC	GAC	AAT	TAT	TTG	ATC	AAA	TAC	AAT	CTA	816
Glu	Leu	Tyr	Pro	Gln	Gly	Gly	Asp	Asn	Tyr	Leu	Ile	Lys	Tyr	Asn	Leu	
			260					265					270			
AGC	AAT	AGT	TTT	GGT	TTG	GAT	GAT	TTT	TAT	CTG	CAA	TAT	AAA	GAT	GGT	864
Ser	Asn	Ser	Phe	Gly	Leu	Asp	Asp	Phe	Tyr	Leu	Gln	Tyr	Lys	Asp	Gly	
		275				280						285				
TCC	GCT	GAT	TGG	ACT	GAG	ATT	GCC	CAT	AAT	CCC	TAT	CCT	GAT	ATG	GTC	912
Ser	Ala	Asp	Trp	Thr	Glu	Ile	Ala	His	Asn	Pro	Tyr	Pro	Asp	Met	Val	
	290					295					300					
ATA	AAT	CAA	AAG	TAT	GAA	TCA	CAG	GCG	ACA	ATC	AAA	CGT	AGT	GAC	TCT	960
Ile	Asn	Gln	Lys	Tyr	Glu	Ser	Gln	Ala	Thr	Ile	Lys	Arg	Ser	Asp	Ser	
305					310					315					320	
GAC	AAT	ATA	CTC	AGT	ATA	GGG	TTA	CAA	AGA	TGG	CAT	AGC	GGT	AGT	TAT	1008
Asp	Asn	Ile	Leu	Ser	Ile	Gly	Leu	Gln	Arg	Trp	His	Ser	Gly	Ser	Tyr	
				325					330					335		
AAT	TTT	GCC	GCC	GCC	AAT	TTT	AAA	ATT	GAC	CAA	TAC	TCC	CCG	AAA	GCT	1056
Asn	Phe	Ala	Ala	Ala	Asn	Phe	Lys	Ile	Asp	Gln	Tyr	Ser	Pro	Lys	Ala	
			340				345						350			
TTC	CTG	CTT	AAA	ATG	AAT	AAG	GCT	ATT	CGG	TTG	CTC	AAA	GCT	ACC	GGC	1104
Phe	Leu	Leu	Lys	Met	Asn	Lys	Ala	Ile	Arg	Leu	Leu	Lys	Ala	Thr	Gly	
		355				360						365				
CTC	TCT	TTT	GCT	ACG	TTG	GAG	CGT	ATT	GTT	GAT	AGT	GTT	AAT	AGC	ACC	1152
Leu	Ser	Phe	Ala	Thr	Leu	Glu	Arg	Ile	Val	Asp	Ser	Val	Asn	Ser	Thr	
	370					375					380					
AAA	TCC	ATC	ACG	GTT	GAG	GTA	TTA	AAC	AAG	GTT	TAT	CGG	GTA	AAA	TTC	1200
Lys	Ser	Ile	Thr	Val	Glu	Val	Leu	Asn	Lys	Val	Tyr	Arg	Val	Lys	Phe	
385				390						395					400	
TAT	ATT	GAT	CGT	TAT	GGC	ATC	AGT	GAA	GAG	ACA	GCC	GCT	ATT	TTG	GCT	1248
Tyr	Ile	Asp	Arg	Tyr	Gly	Ile	Ser	Glu	Glu	Thr	Ala	Ala	Ile	Leu	Ala	
				405					410					415		
AAT	ATT	AAT	ATC	TCT	CAG	CAA	GCT	GTT	GGC	AAT	CAG	CTT	AGC	CAG	TTT	1296
Asn	Ile	Asn	Ile	Ser	Gln	Gln	Ala	Val	Gly	Asn	Gln	Leu	Ser	Gln	Phe	
			420					425					430			

GAG CAA CTA TTT AAT CAC CCG CCG CTC AAT GGT ATT CGC TAT GAA ATC 1344  
Glu Gln Leu Phe Asn His Pro Pro Leu Asn Gly Ile Arg Tyr Glu Ile  
435 440 445

AGT GAG GAC AAC TCC AAA CAT CTT CCT AAT CCT GAT CTG AAC CTT AAA 1392  
Ser Glu Asp Asn Ser Lys His Leu Pro Asn Pro Asp Leu Asn Leu Lys  
450 455 460

CCA GAC AGT ACC GGT GAT GAT CAA CGC AAG GCG GTT TTA AAA CGC GCG 1440  
Pro Asp Ser Thr Gly Asp Asp Gln Arg Lys Ala Val Leu Lys Arg Ala  
465 470 475 480

TTT CAG GTT AAC GCC AGT GAG TTG TAT CAG ATG TTA TTG ATC ACT GAT 1488  
Phe Gln Val Asn Ala Ser Glu Leu Tyr Gln Met Leu Leu Ile Thr Asp  
485 490 495

CGT AAA GAA GAC GGT GTT ATC AAA AAT AAC TTA GAG AAT TTG TCT GAT 1536  
Arg Lys Glu Asp Gly Val Ile Lys Asn Asn Leu Glu Asn Leu Ser Asp  
500 505 510

CTG TAT TTG GTT AGT TTG CTG GCC CAG ATT CAT AAC CTG ACT ATT GCT 1584  
Leu Tyr Leu Val Ser Leu Leu Ala Gln Ile His Asn Leu Thr Ile Ala  
515 520 525

GAA TTG AAC ATT TTG TTG GTG ATT TGT GGC TAT GGC GAC ACC AAC ATT 1632  
Glu Leu Asn Ile Leu Leu Val Ile Cys Gly Tyr Gly Asp Thr Asn Ile  
530 535 540

TAT CAG ATT ACC GAC GAT AAT TTA GCC AAA ATA GTG GAA ACA TTG TTG 1680  
Tyr Gln Ile Thr Asp Asp Asn Leu Ala Lys Ile Val Glu Thr Leu Leu  
545 550 555 560

TGG ATC ACT CAA TGG TTG AAG ACC CAA AAA TGG ACA GTT ACC GAC CTG 1728  
Trp Ile Thr Gln Trp Leu Lys Thr Gln Lys Trp Thr Val Thr Asp Leu  
565 570 575

TTT CTG ATG ACC ACG GCC ACT TAC AGC ACC ACT TTA ACG CCA GAA ATT 1776  
Phe Leu Met Thr Thr Ala Thr Tyr Ser Thr Thr Leu Thr Pro Glu Ile  
580 585 590

AGC AAT CTG ACG GCT ACG TTG TCT TCA ACT TTG CAT GGC AAA GAG AGT 1824  
Ser Asn Leu Thr Ala Thr Leu Ser Ser Thr Leu His Gly Lys Glu Ser  
595 600 605

CTG ATT GGG GAA GAT CTG AAA AGA GCA ATG GCG CCT TGC TTC ACT TCG 1872  
Leu Ile Gly Glu Asp Leu Lys Arg Ala Met Ala Pro Cys Phe Thr Ser  
610 615 620

GCT ~~TTG~~ ~~CAT~~ TTG ACT TCT CAA GAA GTT GCG TAT GAC CTG CTG TTG TGG 1920  
Ala Leu His Leu Thr Ser Gln Glu Val Ala Tyr Asp Leu Leu Leu Trp  
625 630 635 640

ATA GAC CAG ATT CAA CCG GCA CAA ATA ACT GTT GAT GGG TTT TGG GAA 1968  
Ile Asp Gln Ile Gln Pro Ala Gln Ile Thr Val Asp Gly Phe Trp Glu  
645 650 655

GAA GTG CAA ACA ACA CCA ACC AGC TTG AAG GTG ATT ACC TTT GCT CAG 2016  
Glu Val Gln Thr Thr Pro Thr Ser Leu Lys Val Ile Thr Phe Ala Gln  
660 665 670

GTG CTG GCA CAA TTG AGC CTG ATC TAT CGT CGT ATT GGG TTA AGT GAA 2064  
Val Leu Ala Gln Leu Ser Leu Ile Tyr Arg Arg Ile Gly Leu Ser Glu  
675 680 685

ACG GAA CTG TCA CTG ATC GTG ACT CAA TCT TCT CTG CTA GTG GCA GGC 2112  
Thr Glu Leu Ser Leu Ile Val Thr Gln Ser Ser Leu Leu Val Ala Gly  
690 695 700

AAA AGC ATA CTG GAT CAC GGT CTG TTA ACC CTG ATG GCC TTG GAA GGT 2160  
Lys Ser Ile Leu Asp His Gly Leu Leu Thr Thr Leu Met Ala Leu Glu Gly  
705 710 715 720

TTT	CAT	ACC	TGG	GTT	AAT	GGC	TTG	GGG	CAA	CAT	GCC	TCC	TTG	ATA	TTG	2208
Phe	His	Thr	Trp	Val	Asn	Gly	Leu	Gly	Gln	His	Ala	Ser	Leu	Ile	Leu	
				725					730					735		
GCG	GCG	TTG	AAA	GAC	GGA	GCC	TTG	ACA	GTT	ACC	GAT	GTA	GCA	CAA	GCT	2256
Ala	Ala	Leu	Lys	Asp	Gly	Ala	Leu	Thr	Val	Thr	Asp	Val	Ala	Gln	Ala	
			740					745					750			
ATG	AAT	AAG	GAG	GAA	TCT	CTC	CTA	CAA	ATG	GCA	GCT	AAT	CAG	GTG	GAG	2304
Met	Asn	Lys	Glu	Glu	Ser	Leu	Leu	Gln	Met	Ala	Ala	Asn	Gln	Val	Glu	
		755					760					765				
AAG	GAT	CTA	ACA	AAA	CTG	ACC	AGT	TGG	ACA	CAG	ATT	GAC	GCT	ATT	CTG	2352
Lys	Asp	Leu	Thr	Lys	Leu	Thr	Ser	Trp	Thr	Gln	Ile	Asp	Ala	Ile	Leu	
	770					775					780					
CAA	TGG	TTA	CAG	ATG	TCT	TCG	GCC	TTG	GCG	GTT	TCT	CCA	CTG	GAT	CTG	2400
Gln	Trp	Leu	Gln	Met	Ser	Ser	Ala	Leu	Ala	Val	Ser	Pro	Leu	Asp	Leu	
785					790					795					800	
GCA	GGG	ATG	ATG	GCC	CTG	AAA	TAT	GGG	ATA	GAT	CAT	AAC	TAT	GCT	GCC	2448
Ala	Gly	Met	Met	Ala	Leu	Lys	Tyr	Gly	Ile	Asp	His	Asn	Tyr	Ala	Ala	
				805					810					815		
TGG	CAA	GCT	GCG	GCG	GCT	GCG	CTG	ATG	GCT	GAT	CAT	GCT	AAT	CAG	GCA	2496
Trp	Gln	Ala	Ala	Ala	Ala	Ala	Leu	Met	Ala	Asp	His	Ala	Asn	Gln	Ala	
			820					825					830			
CAG	AAA	AAA	CTG	GAT	GAG	ACG	TTC	AGT	AAG	GCA	TTA	TGT	AAC	TAT	TAT	2544
Gln	Lys	Lys	Leu	Asp	Glu	Thr	Phe	Ser	Lys	Ala	Leu	Cys	Asn	Tyr	Tyr	
	835						840					845				
ATT	AAT	GCT	GTT	GTC	GAT	AGT	GCT	GCT	GGA	GTA	CGT	GAT	CGT	AAC	GGT	2592
Ile	Asn	Ala	Val	Val	Asp	Ser	Ala	Ala	Gly	Val	Arg	Asp	Arg	Asn	Gly	
	850					855					860					
TTA	TAT	ACC	TAT	TTG	CTG	ATT	GAT	AAT	CAG	GTT	TCT	GCC	GAT	GTG	ATC	2640
Leu	Tyr	Thr	Tyr	Leu	Leu	Ile	Asp	Asn	Gln	Val	Ser	Ala	Asp	Val	Ile	
865					870					875					880	
ACT	TCA	CGT	ATT	GCA	GAA	GCT	ATC	GCC	GGT	ATT	CAA	CTG	TAC	GTT	AAC	2688
Thr	Ser	Arg	Ile	Ala	Glu	Ala	Ile	Ala	Gly	Ile	Gln	Leu	Tyr	Val	Asn	
				885					890					895		
CGG	GCT	TTA	AAC	CGA	GAT	GAA	GGT	CAG	CTT	GCA	TCG	GAC	GTT	AGT	ACC	2736
Arg	Ala	Leu	Asn	Arg	Asp	Glu	Gly	Gln	Leu	Ala	Ser	Asp	Val	Ser	Thr	
			900					905					910			
CGT	CAG	TTC	TTC	ACT	GAC	TGG	GAA	CGT	TAC	AAT	AAA	CGT	TAC	AGT	ACT	2784
Arg	Gln	Phe	Phe	Thr	Asp	Trp	Glu	Arg	Tyr	Asn	Lys	Arg	Tyr	Ser	Thr	
		915					920					925				
TGG	GCT	GGT	GTC	TCT	GAA	CTG	GTC	TAT	TAT	CCA	GAA	AAC	TAT	GTT	GAT	2832
Trp	Ala	Gly	Val	Ser	Glu	Leu	Val	Tyr	Tyr	Pro	Glu	Asn	Tyr	Val	Asp	
	930					935					940					
CCC	ACT	CAG	CGC	ATT	GGG	CAA	ACC	AAA	ATG	ATG	GAT	GCG	CTG	TTG	CAA	2880
Pro	Thr	Gln	Arg	Ile	Gly	Gln	Thr	Lys	Met	Met	Asp	Ala	Leu	Leu	Gln	
945					950					955					960	
TCC	ATC	AAC	CAG	AGC	CAG	CTA	AAT	GCG	GAT	ACG	GTG	GAA	GAT	GCT	TTC	2928
Ser	Ile	Asn	Gln	Ser	Gln	Leu	Asn	Ala	Asp	Thr	Val	Glu	Asp	Ala	Phe	
				965					970					975		
AAA	ACT	TAT	TTG	ACC	AGC	TTT	GAG	CAG	GTA	GCA	AAT	CTG	AAA	GTA	ATT	2976
Lys	Thr	Tyr	Leu	Thr	Ser	Phe	Glu	Gln	Val	Ala	Asn	Leu	Lys	Val	Ile	
			980					985					990			
AGT	GCT	TAC	CAC	GAT	AAT	GTG	AAT	GTG	GAT	CAA	GGA	TTA	ACT	TAT	TTT	3024
Ser	Ala	Tyr	His	Asp	Asn	Val	Asn	Val	Asp	Gln	Gly	Leu	Thr	Tyr	Phe	

995					1000					1005						
ATC	GGT	ATC	GAC	CAA	GCA	GCT	CCG	GGT	ACG	TAT	TAC	TGG	CGT	AGT	GTT	3072
Ile	Gly	Ile	Asp	Gln	Ala	Ala	Pro	Gly	Thr	Tyr	Tyr	Trp	Arg	Ser	Val	
	1010						1015					1020				
GAT	CAC	AGC	AAA	TGT	GAA	AAT	GGC	AAG	TTT	GCC	GCT	AAT	GCT	TGG	GGT	3120
Asp	His	Ser	Lys	Cys	Glu	Asn	Gly	Lys	Phe	Ala	Ala	Asn	Ala	Trp	Gly	
	1025					1030				1035					1040	
GAG	TGG	AAT	AAA	ATT	ACC	TGT	GCT	GTC	AAT	CCT	TGG	AAA	AAT	ATC	ATC	3168
Glu	Trp	Asn	Lys	Ile	Thr	Cys	Ala	Val	Asn	Pro	Trp	Lys	Asn	Ile	Ile	
					1045					1050					1055	
CGT	CCG	GTT	GTT	TAT	ATG	TCC	CGC	TTA	TAT	CTG	CTA	TGG	CTG	GAG	CAG	3216
Arg	Pro	Val	Val	Tyr	Met	Ser	Arg	Leu	Tyr	Leu	Leu	Trp	Leu	Glu	Gln	
					1060					1065					1070	
CAA	TCA	AAG	AAA	AGT	GAT	GAT	GGT	AAA	ACC	ACG	ATT	TAT	CAA	TAT	AAC	3264
Gln	Ser	Lys	Lys	Ser	Asp	Asp	Gly	Lys	Thr	Thr	Ile	Tyr	Gln	Tyr	Asn	
			1075					1080							1085	
TTA	AAA	CTG	GCT	CAT	ATT	CGT	TAC	GAC	GGT	AGT	TGG	AAT	ACA	CCA	TTT	3312
Leu	Lys	Leu	Ala	His	Ile	Arg	Tyr	Asp	Gly	Ser	Trp	Asn	Thr	Pro	Phe	
	1090						1095								1100	
ACT	TTT	GAT	GTG	ACA	GAA	AAG	GTA	AAA	AAT	TAC	ACG	TCG	AGT	ACT	GAT	3360
Thr	Phe	Asp	Val	Thr	Glu	Lys	Val	Lys	Asn	Tyr	Thr	Ser	Ser	Thr	Asp	
	1105					1110					1115				1120	
GCT	GCT	GAA	TCT	TTA	GGG	TTG	TAT	TGT	ACT	GGT	TAT	CAA	GGG	GAA	GAC	3408
Ala	Ala	Glu	Ser	Leu	Gly	Leu	Tyr	Cys	Thr	Gly	Tyr	Gln	Gly	Glu	Asp	
					1125					1130					1135	
ACT	CTA	TTA	GTT	ATG	TTC	TAT	TCG	ATG	CAG	AGT	AGT	TAT	AGC	TCC	TAT	3456
Thr	Leu	Leu	Val	Met	Phe	Tyr	Ser	Met	Gln	Ser	Ser	Tyr	Ser	Ser	Tyr	
					1140					1145					1150	
ACC	GAT	AAT	AAT	GCG	CCG	GTC	ACT	GGG	CTA	TAT	ATT	TTC	GCT	GAT	ATG	3504
Thr	Asp	Asn	Asn	Ala	Pro	Val	Thr	Gly	Leu	Tyr	Ile	Phe	Ala	Asp	Met	
					1155				1160						1165	
TCA	TCA	GAC	AAT	ATG	ACG	AAT	GCA	CAA	GCA	ACT	AAC	TAT	TGG	AAT	AAC	3552
Ser	Ser	Asp	Asn	Met	Thr	Asn	Ala	Gln	Ala	Thr	Asn	Tyr	Trp	Asn	Asn	
					1170										1180	
AGT	TAT	CCG	CAA	TTT	GAT	ACT	GTG	ATG	GCA	GAT	CCG	GAT	AGC	GAC	AAT	3600
Ser	Tyr	Pro	Gln	Phe	Asp	Thr	Val	Met	Ala	Asp	Pro	Asp	Ser	Asp	Asn	
	1185					1190					1195				1200	
AAA	AAA	GTC	ATA	ACC	AGA	AGA	GTT	AAT	AAC	CGT	TAT	GCG	GAG	GAT	TAT	3648
Lys	Lys	Val	Ile	Thr	Arg	Arg	Val	Asn	Arg	Tyr	Ala	Glu	Asp	Tyr		
					1205					1210					1215	
GAA	ATT	CCT	TCC	TCT	GTG	ACA	AGT	AAC	AGT	AAT	TAT	TCT	TGG	GGT	GAT	3696
Glu	Ile	Pro	Ser	Ser	Val	Thr	Ser	Asn	Ser	Asn	Tyr	Ser	Trp	Gly	Asp	
					1220					1225					1230	
CAC	AGT	TTA	ACC	ATG	CTT	TAT	GGT	GGT	AGT	GTT	CCT	AAT	ATT	ACT	TTT	3744
His	Ser	Leu	Thr	Met	Leu	Tyr	Gly	Gly	Ser	Val	Pro	Asn	Ile	Thr	Phe	
					1235					1240					1245	
GAA	TCG	GCG	GCA	GAA	GAT	TTA	AGG	CTA	TCT	ACC	AAT	ATG	GCA	TTG	AGT	3792
Glu	Ser	Ala	Ala	Glu	Asp	Leu	Arg	Leu	Ser	Thr	Asn	Met	Ala	Leu	Ser	
					1250										1260	
ATT	ATT	CAT	AAT	GGA	TAT	GCG	GGA	ACC	CGC	CGT	ATA	CAA	TGT	AAT	CTT	3840
Ile	Ile	His	Asn	Gly	Tyr	Ala	Gly	Thr	Arg	Arg	Ile	Gln	Cys	Asn	Leu	
					1265						1270				1280	
ATG	AAA	CAA	TAC	GCT	TCA	TTA	GGT	GAT	AAA	TTT	ATA	ATT	TAT	GAT	TCA	3888

Met Lys Gln Tyr Ala Ser Leu Gly Asp Lys Phe Ile Ile Tyr Asp Ser  
1285 1290 1295

TCA TTT GAT GAT GCA AAC CGT TTT AAT CTG GTG CCA TTG TTT AAA TTC 3936  
Ser Phe Asp Asp Ala Asn Arg Phe Asn Leu Val Pro Leu Phe Lys Phe  
1300 1305 1310

GGA AAA GAC GAG AAC TCA GAT GAT AGT ATT TGT ATA TAT AAT GAA AAC 3984  
Gly Lys Asp Glu Asn Ser Asp Asp Ser Ile Cys Ile Tyr Asn Glu Asn  
1315 1320 1325

CCT TCC TCT GAA GAT AAG AAG TGG TAT TTT TCT TCG AAA GAT GAC AAT 4032  
Pro Ser Ser Glu Asp Lys Lys Trp Tyr Phe Ser Ser Lys Asp Asp Asn  
1330 1335 1340

AAA ACA GCG GAT TAT AAT GGT GGA ACT CAA TGT ATA GAT GCT GGA ACC 4080  
Lys Thr Ala Asp Tyr Asn Gly Gly Thr Gln Cys Ile Asp Ala Gly Thr  
1345 1350 1355 1360

AGT AAC AAA GAT TTT TAT TAT AAT CTC CAG GAG ATT GAA GTA ATT AGT 4128  
Ser Asn Lys Asp Phe Tyr Tyr Asn Leu Gln Glu Ile Glu Val Ile Ser  
1365 1370 1375

GTT ACT GGT GGG TAT TGG TCG AGT TAT AAA ATA TCC AAC CCG ATT AAT 4176  
Val Thr Gly Gly Tyr Trp Ser Ser Tyr Lys Ile Ser Asn Pro Ile Asn  
1380 1385 1390

ATC AAT ACG GGC ATT GAT AGT GCT AAA GTA AAA GTC ACC GTA AAA GCG 4224  
Ile Asn Thr Gly Ile Asp Ser Ala Lys Val Lys Val Thr Val Lys Ala  
1395 1400 1405

GGT GGT GAC GAT CAA ATC TTT ACT GCT GAT AAT AGT ACC TAT GTT CCT 4272  
Gly Gly Asp Asp Gln Ile Phe Thr Ala Asp Asn Ser Thr Tyr Val Pro  
1410 1415 1420

CAG CAA CCG GCA CCC AGT TTT GAG GAG ATG ATT TAT CAG TTC AAT AAC 4320  
Gln Gln Pro Ala Pro Ser Phe Glu Glu Met Ile Tyr Gln Phe Asn Asn  
1425 1430 1435 1440

CTG ACA ATA GAT TGT AAG AAT TTA AAT TTC ATC GAC AAT CAG GCA CAT 4368  
Leu Thr Ile Asp Cys Lys Asn Leu Asn Phe Ile Asp Asn Gln Ala His  
1445 1450 1455

ATT GAG ATT GAT TTC ACC GCT ACG GCA CAA GAT GGC CGA TTC TTG GGT 4416  
Ile Glu Ile Asp Phe Thr Ala Thr Ala Gln Asp Gly Arg Phe Leu Gly  
1460 1465 1470

GCA GAA ACT TTT ATT ATC CCG GTA ACT AAA AAA GTT CTC GGT ACT GAG 4464  
Ala ~~Thr~~ Thr Phe Ile Ile Pro Val Thr Lys Lys Val Leu Gly Thr Glu  
1475 1480 1485

AAC GTG ATT GCG TTA TAT AGC GAA AAT AAC GGT GTT CAA TAT ATG CAA 4512  
Asn Val Ile Ala Leu Tyr Ser Glu Asn Asn Gly Val Gln Tyr Met Gln  
1490 1495 1500

ATT GGC GCA TAT CGT ACC CGT TTG AAT ACG TTA TTC GCT CAA CAG TTG 4560  
Ile Gly Ala Tyr Arg Thr Arg Leu Asn Thr Leu Phe Ala Gln Gln Leu  
1505 1510 1515 1520

GTT AGC CGT GCT AAT CGT GGC ATT GAT GCA GTG CTC AGT ATG GAA ACT 4608  
Val Ser Arg Ala Asn Arg Gly Ile Asp Ala Val Leu Ser Met Glu Thr  
1525 1530 1535

CAG AAT ATT CAG GAA CCG CAA TTA GGA GCG GGC ACA TAT GTG CAG CTT 4656  
Gln Asn Ile Gln Glu Pro Gln Leu Gly Ala Gly Thr Tyr Val Gln Leu  
1540 1545 1550

GTG TTG GAT AAA TAT GAT GAG TCT ATT CAT GGC ACT AAT AAA AGC TTT 4704  
Val Leu Asp Lys Tyr Asp Glu Ser Ile His Gly Thr Asn Lys Ser Phe  
1555 1560 1565

GCT ATT GAA TAT GTT GAT ATA TTT AAA GAG AAC GAT AGT TTT GTG ATT 4752  
Ala Ile Glu Tyr Val Asp Ile Phe Lys Glu Asn Asp Ser Phe Val Ile  
1570 1575 1580

TAT CAA GGA GAA CTT AGC GAA ACA AGT CAA ACT GTT GTG AAA GTT TTC 4800  
Tyr Gln Gly Glu Leu Ser Glu Thr Ser Gln Thr Val Val Lys Val Phe  
1585 1590 1595 1600

TTA TCC TAT TTT ATA GAG GCG ACT GGA AAT AAG AAC CAC TTA TGG GTA 4848  
Leu Ser Tyr Phe Ile Glu Ala Thr Gly Asn Lys Asn His Leu Trp Val  
1605 1610 1615 -

CGT GCT AAA TAC CAA AAG GAA ACG ACT GAT AAG ATC TTG TTC GAC CGT 4896  
Arg Ala Lys Tyr Gln Lys Glu Thr Thr Asp Lys Ile Leu Phe Asp Arg  
1620 1625 1630

ACT GAT GAG AAA GAT CCG CAC GGT TGG TTT CTC AGC GAC GAT CAC AAG 4944  
Thr Asp Glu Lys Asp Pro His Gly Trp Phe Leu Ser Asp Asp His Lys  
1635 1640 1645

ACC TTT AGT GGT CTC TCT TCC GCA CAG GCA TTA AAG AAC GAC AGT GAA 4992  
Thr Phe Ser Gly Leu Ser Ser Ala Gln Ala Leu Lys Asn Asp Ser Glu  
1650 1655 1660

CCG ATG GAT TTC TCT GGC GCC AAT GCT CTC TAT TTC TGG GAA CTG TTC 5040  
Pro Met Asp Phe Ser Gly Ala Asn Ala Leu Tyr Phe Trp Glu Leu Phe  
1665 1670 1675 1680

TAT TAC ACG CCG ATG ATG ATG GCT CAT CGT TTG TTG CAG GAA CAG AAT 5088  
Tyr Tyr Thr Pro Met Met Met Ala His Arg Leu Leu Gln Glu Gln Asn  
1685 1690 1695

TTT GAT GCG GCG AAC CAT TGG TTC CGT TAT GTC TGG AGT CCA TCC GGT 5136  
Phe Asp Ala Ala Asn His Trp Phe Arg Tyr Val Trp Ser Pro Ser Gly  
1700 1705 1710

TAT ATC GTT GAT GGT AAA ATT GCT ATC TAC CAC TGG AAC GTG CGA CCG 5184  
Tyr Ile Val Asp Gly Lys Ile Ala Ile Tyr His Trp Asn Val Arg Pro  
1715 1720 1725

CTG GAA GAA GAC ACC AGT TGG AAT GCA CAA CAA CTG GAC TCC ACC GAT 5232  
Leu Glu Glu Asp Thr Ser Trp Asn Ala Gln Gln Leu Asp Ser Thr Asp  
1730 1735 1740

CCA GAT GCT GTA GCC CAA GAT GAT CCG ATG CAC TAC AAG GTG GCT ACC 5280  
Pro Asp Ala Val Ala Gln Asp Asp Pro Met His Tyr Lys Val Ala Thr  
1745 1750 1755 1760

TTT ATG GCG ACG TTG GAT CTG CTA ATG GCC CGT GGT GAT GCT GCT TAC 5328  
Phe Met Ala Thr Leu Asp Leu Leu Met Ala Arg Gly Asp Ala Ala Tyr  
1765 1770 1775

CGC CAG TTA GAG CGT GAT ACG TTG GCT GAA GCT AAA ATG TGG TAT ACA 5376  
Arg Gln Leu Glu Arg Asp Thr Leu Ala Glu Ala Lys Met Trp Tyr Thr  
1780 1785 1790

CAG GCG CTT AAT CTG TTG GGT GAT GAG CCA CAA GTG ATG CTG AGT ACG 5424  
Gln Ala Leu Asn Leu Leu Gly Asp Glu Pro Gln Val Met Leu Ser Thr  
1795 1800 1805

ACT TGG GCT AAT CCA ACA TTG GGT AAT GCT GCT TCA AAA ACC ACA CAG 5472  
Thr Trp Ala Asn Pro Thr Leu Gly Asn Ala Ala Ser Lys Thr Thr Gln  
1810 1815 1820

CAG GTT CGT CAG CAA GTG CTT ACC CAG TTG CGT CTC AAT AGC AGG GTA 5520  
Gln Val Arg Gln Gln Val Leu Thr Gln Leu Arg Leu Asn Ser Arg Val  
1825 1830 1835 1840

AAA ACC CCG TTG 5532  
Lys Thr Pro Leu  
1844



(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 53:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĹŽKA: 1844 aminokyselín
- (B) TYP: aminokyseliny
- (C) REĹAZEC: jednoduchý
- (D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: proteín

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 53 (TcbA<sub>ii</sub>):

Charakt. znak	od	do	Opis
Peptid	1	1844	TcbA <sub>ii</sub> peptid
Fragment	1	11	(SEKV. ID. ČÍS. 1)
Fragment	978	990	(SEKV. ID. ČÍS. 23)
Fragment	1387	1401	(SEKV. ID. ČÍS. 22)
Fragment	1484	1505	(SEKV. ID. ČÍS. 24)
Fragment	1527	1552	(SEKV. ID. ČÍS. 21)

Phe Ile Gln Gly Tyr Ser Asp Leu Phe Gly Asn Arg Ala Asp Asn Tyr  
1 5 10 15  
Ala Ala Pro Gly Ser Val Ala Ser Met Phe Ser Pro Ala Ala Tyr Leu  
20 25 30  
Thr Glu Leu Tyr Arg Glu Ala Lys Asn Leu His Asp Ser Ser Ser Ile  
35 40 45  
Tyr Tyr Leu Asp Lys Arg Arg Pro Asp Leu Ala Ser Leu Met Leu Ser  
50 55 60  
Gln Lys Asn Met Asp Glu Glu Ile Ser Thr Leu Ala Leu Ser Asn Glu  
65 70 75 80  
Leu Cys Leu Ala Gly Ile Glu Thr Lys Thr Gly Lys Ser Gln Asp Glu  
85 90 95  
Val Met Asp Met Leu Ser Thr Tyr Arg Leu Ser Gly Glu Thr Pro Tyr  
100 105 110  
His His Ala Tyr Glu Thr Val Arg Glu Ile Val His Glu Arg Asp Pro  
115 120 125  
Gly Phe Arg His Leu Ser Gln Ala Pro Ile Val Ala Ala Lys Leu Asp  
130 135 140  
Pro Val Thr Leu Leu Gly Ile Ser Ser His Ile Ser Pro Glu Leu Tyr  
145 150 155 160  
Asn Leu Leu Ile Glu Glu Ile Pro Glu Lys Asp Glu Ala Ala Leu Asp  
165 170 175

Thr Leu Tyr Lys Thr Asn Phe Gly Asp Ile Thr Thr Ala Gln Leu Met  
180 185 190

Ser Pro Ser Tyr Leu Ala Arg Tyr Tyr Gly Val Ser Pro Glu Asp Ile  
195 200 205

Ala Tyr Val Thr Thr Ser Leu Ser His Val Gly Tyr Ser Ser Asp Ile  
210 215 220

Leu Val Ile Pro Leu Val Asp Gly Val Gly Lys Met Glu Val Val Arg  
225 230 235 240

Val Thr Arg Thr Pro Ser Asp Asn Tyr Thr Ser Gln Thr Asn Tyr Ile  
245 250 255

Glu Leu Tyr Pro Gln Gly Gly Asp Asn Tyr Leu Ile Lys Tyr Asn Leu  
260 265 270

Ser Asn Ser Phe Gly Leu Asp Asp Phe Tyr Leu Gln Tyr Lys Asp Gly  
275 280 285

Ser Ala Asp Trp Thr Glu Ile Ala His Asn Pro Tyr Pro Asp Met Val  
290 295 300

Ile Asn Gln Lys Tyr Glu Ser Gln Ala Thr Ile Lys Arg Ser Asp Ser  
305 310 315 320

Asp Asn Ile Leu Ser Ile Gly Leu Gln Arg Trp His Ser Gly Ser Tyr  
325 330 335

Asn Phe Ala Ala Ala Asn Phe Lys Ile Asp Gln Tyr Ser Pro Lys Ala  
340 345 350

Phe Leu Leu Lys Met Asn Lys Ala Ile Arg Leu Leu Lys Ala Thr Gly  
355 360 365

Leu Ser Phe Ala Thr Leu Glu Arg Ile Val Asp Ser Val Asn Ser Thr  
370 375 380

Lys Ser Ile Thr Val Glu Val Leu Asn Lys Val Tyr Arg Val Lys Phe  
385 390 395 400

Tyr Ile Asp Arg Tyr Gly Ile Ser Glu Glu Thr Ala Ala Ile Leu Ala  
405 410 415

Asn Ile Asn Ile Ser Gln Gln Ala Val Gly Asn Gln Leu Ser Gln Phe  
420 425 430

Glu Gln Leu Phe Asn His Pro Pro Leu Asn Gly Ile Arg Tyr Glu Ile  
435 440 445

Ser Glu Asp Asn Ser Lys His Leu Pro Asn Pro Asp Leu Asn Leu Lys  
450 455 460

Pro Asp Ser Thr Gly Asp Asp Gln Arg Lys Ala Val Leu Lys Arg Ala  
465 470 475 480

Phe Gln Val Asn Ala Ser Glu Leu Tyr Gln Met Leu Leu Ile Thr Asp  
485 490 495

Arg Lys Glu Asp Gly Val Ile Lys Asn Asn Leu Glu Asn Leu Ser Asp  
500 505 510

Leu Tyr Leu Val Ser Leu Leu Ala Gln Ile His Asn Leu Thr Ile Ala  
515 520 525

Glu Leu Asn Ile Leu Leu Val Ile Cys Gly Tyr Gly Asp Thr Asn Ile  
530 535 540

Tyr Gln Ile Thr Asp Asp Asn Leu Ala Lys Ile Val Glu Thr Leu Leu  
545 550 560

Trp Ile Thr Gln Trp Leu Lys Thr Gln Lys Trp Thr Val Thr Asp Leu  
565 570 575

Phe Leu Met Thr Thr Ala Thr Tyr Ser Thr Thr Leu Thr Pro Glu Ile  
580 585 590

Ser Asn Leu Thr Ala Thr Leu Ser Ser Thr Leu His Gly Lys Glu Ser  
595 600 605

Leu Ile Gly Glu Asp Leu Lys Arg Ala Met Ala Pro Cys Phe Thr Ser  
610 615 620

Ala Leu His Leu Thr Ser Gln Glu Val Ala Tyr Asp Leu Leu Leu Trp  
625 630 635 640

Ile Asp Gln Ile Gln Pro Ala Gln Ile Thr Val Asp Gly Phe Trp Glu  
645 650 655

Glu Val Gln Thr Thr Pro Thr Ser Leu Lys Val Ile Thr Phe Ala Gln  
660 665 670

Val Leu Ala Gln Leu Ser Leu Ile Tyr Arg Arg Ile Gly Leu Ser Glu  
675 680 685

Thr Glu Leu Ser Leu Ile Val Thr Gln Ser Ser Leu Leu Val Ala Gly  
690 695 700

Lys Ser Ile Leu Asp His Gly Leu Leu Thr Leu Met Ala Leu Glu Gly  
705 710 715 720

Phe His Thr Trp Val Asn Gly Leu Gly Gln His Ala Ser Leu Ile Leu  
725 730 735

Ala Ala Leu Lys Asp Gly Ala Leu Thr Val Thr Asp Val Ala Gln Ala  
740 745 750

Met Asn Lys Glu Glu Ser Leu Leu Gln Met Ala Ala Asn Gln Val Glu  
755 760 765

Lys Asp Leu Thr Lys Leu Thr Ser Trp Thr Gln Ile Asp Ala Ile Leu  
770 775 780

Gln Trp Leu Gln Met Ser Ser Ala Leu Ala Val Ser Pro Leu Asp Leu  
785 790 795 800

Ala Gly Met Met Ala Leu Lys Tyr Gly Ile Asp His Asn Tyr Ala Ala  
805 810 815

Trp Gln Ala Ala Ala Ala Ala Leu Met Ala Asp His Ala Asn Gln Ala  
820 825 830

Gln Lys Lys Leu Asp Glu Thr Phe Ser Lys Ala Leu Cys Asn Tyr Tyr  
835 840 845

Ile Asn Ala Val Val Asp Ser Ala Ala Gly Val Arg Asp Arg Asn Gly  
850 855 860

Leu Tyr Thr Tyr Leu Leu Ile Asp Asn Gln Val Ser Ala Asp Val Ile  
865 870 875 880

Thr Ser Arg Ile Ala Glu Ala Ile Ala Gly Ile Gln Leu Tyr Val Asn  
885 890 895

Arg Ala Leu Asn Arg Asp Glu Gly Gln Leu Ala Ser Asp Val Ser Thr  
900 905 910

Arg Gln Phe Phe Thr Asp Trp Glu Arg Tyr Asn Lys Arg Tyr Ser Thr  
915 920 925

Trp Ala Gly Val Ser Glu Leu Val Tyr Tyr Pro Glu Asn Tyr Val Asp  
930 935 940

Pro Thr Gln Arg Ile Gly Gln Thr Lys Met Met Asp Ala Leu Leu Gln  
945 950 955 960

Ser Ile Asn Gln Ser Gln Leu Asn Ala Asp Thr Val Glu Asp Ala Phe  
965 970 975

Lys Thr Tyr Leu Thr Ser Phe Glu Gln Val Ala Asn Leu Lys Val Ile  
980 985 990

Ser Ala Tyr His Asp Asn Val Asn Val Asp Gln Gly Leu Thr Tyr Phe  
995 1000 1005

Ile Gly Ile Asp Gln Ala Ala Pro Gly Thr Tyr Tyr Trp Arg Ser Val  
1010 1015 1020

Asp His Ser Lys Cys Glu Asn Gly Lys Phe Ala Ala Asn Ala Trp Gly  
1025 1030 1035 1040

Glu Trp Asn Lys Ile Thr Cys Ala Val Asn Pro Trp Lys Asn Ile Ile  
1045 1050 1055

Arg Pro Val Val Tyr Met Ser Arg Leu Tyr Leu Leu Trp Leu Glu Gln  
1060 1065 1070

Gln Ser Lys Lys Ser Asp Asp Gly Lys Thr Thr Ile Tyr Gln Tyr Asn  
1075 1080 1085

Leu Lys Leu Ala His Ile Arg Tyr Asp Gly Ser Trp Asn Thr Pro Phe  
1090 1095 1100

Thr Phe Asp Val Thr Glu Lys Val Lys Asn Tyr Thr Ser Ser Thr Asp  
1105 1110 1115 1120

Ala Ala Glu Ser Leu Gly Leu Tyr Cys Thr Gly Tyr Gln Gly Glu Asp  
- 1125 1130 1135

Thr Leu Leu Val Met Phe Tyr Ser Met Gln Ser Ser Tyr Ser Ser Tyr  
1140 1145 1150

Thr Asp Asn Asn Ala Pro Val Thr Gly Leu Tyr Ile Phe Ala Asp Met  
1155 1160 1165

Ser Ser Asp Asn Met Thr Asn Ala Gln Ala Thr Asn Tyr Trp Asn Asn  
1170 1175 1180

Ser Tyr Pro Gln Phe Asp Thr Val Met Ala Asp Pro Asp Ser Asp Asn  
1185 1190 1195 1200

Lys Lys Val Ile Thr Arg Arg Val Asn Asn Arg Tyr Ala Glu Asp Tyr  
1205 1210 1215

Glu Ile Pro Ser Ser Val Thr Ser Asn Ser Asn Tyr Ser Trp Gly Asp  
1220 1225 1230

His Ser Leu Thr Met Leu Tyr Gly Gly Ser Val Pro Asn Ile Thr Phe  
1235 1240 1245

Glu Ser Ala Ala Glu Asp Leu Arg Leu Ser Thr Asn Met Ala Leu Ser  
1250 1255 1260

Ile Ile His Asn Gly Tyr Ala Gly Thr Arg Arg Ile Gln Cys Asn Leu  
1265 1270 1275 1280

Met Lys Gln Tyr Ala Ser Leu Gly Asp Lys Phe Ile Ile Tyr Asp Ser  
1285 1290 1295

Ser Phe Asp Asp Ala Asn Arg Phe Asn Leu Val Pro Leu Phe Lys Phe  
1300 1305 1310

Gly Lys Asp Glu Asn Ser Asp Asp Ser Ile Cys Ile Tyr Asn Glu Asn  
1315 1320 1325

Pro Ser Ser Glu Asp Lys Lys Trp Tyr Phe Ser Ser Lys Asp Asp Asn  
1330 1335 1340

Lys Thr Ala Asp Tyr Asn Gly Gly Thr Gln Cys Ile Asp Ala Gly Thr  
1345 1350 1355 1360

Ser Asn Lys Asp Phe Tyr Tyr Asn Leu Gln Glu Ile Glu Val Ile Ser  
1365 1370 1375

Val Thr Gly Gly Tyr Trp Ser Ser Tyr Lys Ile Ser Asn Pro Ile Asn  
1380 1385 1390

Ile Asn Thr Gly Ile Asp Ser Ala Lys Val Lys Val Thr Val Lys Ala  
1395 1400 1405

Gly Gly Asp Asp Gln Ile Phe Thr Ala Asp Asn Ser Thr Tyr Val Pro  
1410 1415 1420

Gln Gln Pro Ala Pro Ser Phe Glu Glu Met Ile Tyr Gln Phe Asn Asn  
1425 1430 1435 1440

Leu Thr Ile Asp Cys Lys Asn Leu Asn Phe Ile Asp Asn Gln Ala His  
1445 1450 1455

Ile Glu Ile Asp Phe Thr Ala Thr Ala Gln Asp Gly Arg Phe Leu Gly  
1460 1465 1470

Ala Glu Thr Phe Ile Ile Pro Val Thr Lys Lys Val Leu Gly Thr Glu  
1475 1480 1485

Asn Val Ile Ala Leu Tyr Ser Glu Asn Asn Gly Val Gln Tyr Met Gln  
1490 1495 1500

Ile Gly Ala Tyr Arg Thr Arg Leu Asn Thr Leu Phe Ala Gln Gln Leu  
1505 1510 1515 1520

Val Ser Arg Ala Asn Arg Gly Ile Asp Ala Val Leu Ser Met Glu Thr  
1525 1530 1535

Gln Asn Ile Gln Glu Pro Gln Leu Gly Ala Gly Thr Tyr Val Gln Leu  
1540 1545 1550

Val Leu Asp Lys Tyr Asp Glu Ser Ile His Gly Thr Asn Lys Ser Phe  
1555 1560 1565

Ala Ile Glu Tyr Val Asp Ile Phe Lys Glu Asn Asp Ser Phe Val Ile  
1570 1575 1580

Tyr Gln Gly Glu Leu Ser Glu Thr Ser Gln Thr Val Val Lys Val Phe  
1585 1590 1595 1600

Leu Ser Tyr Phe Ile Glu Ala Thr Gly Asn Lys Asn His Leu Trp Val  
1605 1610 1615

Arg Ala Lys Tyr Gln Lys Glu Thr Thr Asp Lys Ile Leu Phe Asp Arg  
1620 1625 1630

Thr Asp Glu Lys Asp Pro His Gly Trp Phe Leu Ser Asp Asp His Lys  
1635 1640 1645

Thr Phe Ser Gly Leu Ser Ser Ala Gln Ala Leu Lys Asn Asp Ser Glu  
1650 1655 1660

Pro Met Asp Phe Ser Gly Ala Asn Ala Leu Tyr Phe Trp Glu Leu Phe  
1665 1670 1675 1680

Tyr Tyr Thr Pro Met Met Met Ala His Arg Leu Leu Gln Glu Gln Asn  
1685 1690 1695

Phe Asp Ala Ala Asn His Trp Phe Arg Tyr Val Trp Ser Pro Ser Gly  
1700 1705 1710

Tyr Ile Val Asp Gly Lys Ile Ala Ile Tyr His Trp Asn Val Arg Pro  
1715 1720 1725

Leu Glu Glu Asp Thr Ser Trp Asn Ala Gln Gln Leu Asp Ser Thr Asp  
1730 1735 1740

Pro Asp Ala Val Ala Gln Asp Asp Pro Met His Tyr Lys Val Ala Thr  
1745 1750 1755 1760

Phe Met Ala Thr Leu Asp Leu Leu Met Ala Arg Gly Asp Ala Ala Tyr  
1765 1770 1775

Arg Gln Leu Glu Arg Asp Thr Leu Ala Glu Ala Lys Met Trp Tyr Thr  
1780 1785 1790

Gln Ala Leu Asn Leu Leu Gly Asp Glu Pro Gln Val Met Leu Ser Thr  
1795 1800 1805

Thr Trp Ala Asn Pro Thr Leu Gly Asn Ala Ala Ser Lys Thr Thr Gln  
1810 1815 1820

Gln Val Arg Gln Gln Val Leu Thr Gln Leu Arg Leu Asn Ser Arg Val  
1825 1830 1835 1840

Lys Thr Pro Leu  
1844

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 54:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĚLKA: 1722 párov báz
- (B) TYP: nukleová kyselina
- (C) REŤAZEC: dvojitý
- (D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: DNA (genómová)

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 54 (kódovaná oblasť  
tcbA<sub>iii</sub>):

CTA GGA ACA GCC AAT TCC CTG ACC GCT TTA TTC CTG CCG CAG GAA AAT 48  
Leu Gly Thr Ala Asn Ser Leu Thr Ala Leu Phe Leu Pro Gln Glu Asn  
1 5 10 15

AGC AAG CTC AAA GGC TAC TGG CGG ACA CTG GCG CAG CGT ATG TTT AAT 96  
Ser Lys Leu Lys Gly Tyr Trp Arg Thr Leu Ala Gln Arg Met Phe Asn  
20 25 30

TTA CGT CAT AAT CTG TCG ATT GAC GGC CAG CCG CTC TCC TTG CCG CTG 144  
Leu Arg His Asn Leu Ser Ile Asp Gly Gln Pro Leu Ser Leu Pro Leu  
35 40 45

TAT GCT AAA CCG GCT GAT CCA AAA GCT TTA CTG AGT GCG GCG GTT TCA 192  
Tyr Ala Lys Pro Ala Asp Pro Lys Ala Leu Leu Ser Ala Ala Val Ser  
50 55 60

GCT TCT CAA GGG GGA GCC GAC TTG CCG AAG GCG CCG CTG ACT ATT CAC 240  
Ala Ser Gln Gly Gly Ala Asp Leu Pro Lys Ala Pro Leu Thr Ile His  
65 70 75 80

CGC TTC CCT CAA ATG CTA GAA GGG GCA CGG GGC TTG GTT AAC CAG CTT 288  
Arg Phe Pro Gln Met Leu Glu Gly Ala Arg Gly Leu Val Asn Gln Leu  
85 90 95

ATA CAG TTC GGT AGT TCA CTA TTG GGG TAC AGT GAG CGT CAG GAT GCG 336  
Ile Gln Phe Gly Ser Ser Leu Leu Gly Tyr Ser Glu Arg Gln Asp Ala  
100 105 110

GAA GCT ATG AGT CAA CTA CTG CAA ACC CAA GCC AGC GAG TTA ATA CTG 384  
Glu Ala Met Ser Gln Leu Leu Gln Thr Gln Ala Ser Glu Leu Ile Leu  
115 120 125

ACC AGT ATT CGT ATG CAG GAT AAC CAA TTG GCA GAG CTG GAT TCG GAA 432  
Thr Ser Ile Arg Met Gln Asp Asn Gln Leu Ala Glu Leu Asp Ser Glu  
130 135 140

AAA ACC GCC TTG CAA GTC TCT TTA GCT GGA GTG CAA CAA CGG TTT GAC 480  
Lys Thr Ala Leu Gln Val Ser Leu Ala Gly Val Gln Gln Arg Phe Asp  
145 150 155 160

AGC TAT AGC CAA CTG TAT GAG GAG AAC ATC AAC GCA GGT GAG CAG CGA 528  
Ser Tyr Ser Gln Leu Tyr Glu Glu Asn Ile Asn Ala Gly Glu Gln Arg  
165 170 175

GCG CTG GCG TTA CGC TCA GAA TCT GCT ATT GAG TCT CAG GGA GCG CAG 576  
Ala Leu Ala Leu Arg Ser Glu Ser Ala Ile Glu Ser Gln Gly Ala Gln  
180 185 190

ATT TCC CGT ATG GCA GGC GCG GGT GTT GAT ATG GCA CCA AAT ATC TTC 624  
Ile Ser Arg Met Ala Gly Ala Gly Val Asp Met Ala Pro Asn Ile Phe  
195 200 205

GGC CTG GCT GAT GGC GGC ATG CAT TAT GGT GCT ATT GCC TAT GCC ATC 672  
Gly Leu Ala Asp Gly Gly Met His Tyr Gly Ala Ile Ala Tyr Ala Ile  
210 215 220

GCT GAC GGT ATT GAG TTG AGT GCT TCT GCC AAG ATG GTT GAT GCG GAG 720  
Ala Asp Gly Ile Glu Leu Ser Ala Ser Ala Lys Met Val Asp Ala Glu  
225 230 235 240

AAA GTT GCT CAG TCG GAA ATA TAT CGC CGT CGC CGT CAA GAA TGG AAA 768  
Lys Val Ala Gln Ser Glu Ile Tyr Arg Arg Arg Arg Gln Glu Trp Lys  
245 250 255

ATT CAG CGT GAC AAC GCA CAA GCG GAG ATT AAC CAG TTA AAC GCG CAA 816  
Ile Gln Arg Asp Asn Ala Gln Ala Glu Ile Asn Gln Leu Asn Ala Gln  
260 265 270

CTG GAA TCA CTG TCT ATT CGC CGT GAA GCC GCT GAA ATG CAA AAA GAG 864  
Leu Glu Ser Leu Ser Ile Arg Arg Glu Ala Ala Glu Met Gln Lys Glu  
275 280 285

TAC CTG AAA ACC CAG CAA GCT CAG GCG CAG GCA CAA CTT ACT TTC TTA 912  
Tyr Leu Lys Thr Gln Gln Ala Gln Ala Gln Ala Gln Leu Thr Phe Leu  
290 295 300

AGA AGC AAA TTC AGT AAT CAA GCG TTA TAT AGT TGG TTA CGA GGG CGT 960  
Arg Ser Lys Phe Ser Asn Gln Ala Leu Tyr Ser Trp Leu Arg Gly Arg  
305 310 315 320

TTG TCA GGT ATT TAT TTC CAG TTC TAT GAC TTG GCC GTA TCA CGT TGC 1008  
Leu Ser Gly Ile Tyr Phe Gln Phe Tyr Asp Leu Ala Val Ser Arg Cys  
325 330 335

CTG ATG GCA GAG CAA TCC TAT CAA TGG GAA GCT AAT GAT AAT TCC ATT 1056  
Leu Met Ala Glu Gln Ser Tyr Gln Trp Glu Ala Asn Asp Asn Ser Ile  
340 345 350

AGC TTT GTC AAA CCG GGT GCA TGG CAA GGA ACT TAC GCC GGC TTA TTG 1104  
Ser Phe Val Lys Pro Gly Ala Trp Gln Gly Thr Tyr Ala Gly Leu Leu  
355 360 365

TGT GGA GAA GCT TTG ATA CAA AAT CTG GCA CAA ATG GAA GAG GCA TAT 1152  
Cys Gly Glu Ala Leu Ile Gln Asn Leu Ala Gln Met Glu Glu Ala Tyr  
370 375 380

CTG AAA TGG GAA TCT CGC GCT TTG GAA GTA GAA CGC ACG GTT TCA TTG 1200  
Leu Lys Trp Glu Ser Arg Ala Leu Glu Val Glu Arg Thr Val Ser Leu  
385 390 395 400

GCA GTG GTT TAT GAT TCA CTG GAA GGT AAT GAT CGT TTT AAT TTA GCG 1248  
Ala Val Val Tyr Asp Ser Leu Glu Gly Asn Asp Arg Phe Asn Leu Ala  
405 410 415

GAA CAA ATA CCT GCA TTA TTG GAT AAG GGG GAG GGA ACA GCA GGA ACT 1296  
Glu Gln Ile Pro Ala Leu Leu Asp Lys Gly Glu Gly Thr Ala Gly Thr  
420 425 430

AAA GAA AAT GGG TTA TCA TTG GCT AAT GCT ATC CTG TCA GCT TCG GTC 1344  
Lys Glu Asn Gly Leu Ser Leu Ala Asn Ala Ile Leu Ser Ala Ser Val  
435 440 445

AAA TTG TCC GAC TTG AAA CTG GGA ACG GAT TAT CCA GAC AGT ATC GTT 1392  
Lys Leu Ser Asp Leu Lys Leu Gly Thr Asp Tyr Pro Asp Ser Ile Val  
450 455 460

GGT AGC AAC AAG GTT CGT CGT ATT AAG CAA ATC AGT GTT TCG CTA CCT 1440  
Gly Ser Asn Lys Val Arg Arg Ile Lys Gln Ile Ser Val Ser Leu Pro  
465 470 475 480

GCA TTG GTT GGG CCT TAT CAG GAT GTT CAG GCT ATG CTC AGC TAT GGT 1488  
Ala Leu Val Gly Pro Tyr Gln Asp Val Gln Ala Met Leu Ser Tyr Gly  
485 490 495

GGC AGT ACT CAA TTG CCG AAA GGT TGT TCA GCG TTG GCT GTG TCT CAT 1536  
Gly Ser Thr Gln Leu Pro Lys Gly Cys Ser Ala Leu Ala Val Ser His  
500 505 510

GGT ACC AAT GAT AGT GGT CAG TTC CAG TTG GAT TTC AAT GAC GGC AAA 1584  
Gly Thr Asn Asp Ser Gly Gln Phe Gln Leu Asp Phe Asn Asp Gly Lys  
515 520 525

TAC CTG CCA TTT GAA GGT ATT GCT CTT GAT GAT CAG GGT ACA CTG AAT 1632  
Tyr Leu Pro Phe Glu Gly Ile Ala Leu Asp Asp Gln Gly Thr Leu Asn  
530 535 540



CTT CAA TTT CCG AAT GCT ACC GAC AAG CAG AAA GCA ATA TTG CAA ACT 1680  
 Leu Gln Phe Pro Asn Ala Thr Asp Lys Gln Lys Ala Ile Leu Gln Thr  
 545 550 555 560

ATG AGC GAT ATT ATT TTG CAT ATT CGT TAT ACC ATC CGT TAA 1722  
 Met Ser Asp Ile Ile Leu His Ile Arg Tyr Thr Ile Arg ...  
 565 570 573

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 55:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĹŽKA: 573 aminokyselín
- (B) TYP: aminokyseliny
- (C) REĹAZEC: jednoduchý
- (D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: proteín

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 55 (Tcba<sub>iii</sub>):

Leu Gly Thr Ala Asn Ser Leu Thr Ala Leu Phe Leu Pro Gln Glu Asn  
 1 5 10 15

Ser Lys Leu Lys Gly Tyr Trp Arg Thr Leu Ala Gln Arg ~~Met Phe~~ Asn  
 20 25 30

Leu Arg His Asn Leu Ser Ile Asp Gly Gln Pro Leu Ser Leu Pro Leu  
 35 40 45

Tyr Ala Lys Pro Ala Asp Pro Lys Ala Leu Leu Ser Ala Ala Val Ser  
 50 55 60

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

Arg Phe Pro Gln Met Leu Glu Gly Ala Arg Gly Leu Val Asn Gln Leu  
 85 90 95

Ile Gln Phe Gly Ser Ser Leu Leu Gly Tyr Ser Glu Arg Gln Asp Ala  
 100 105 110

Glu Ala Met Ser Gln Leu Leu Gln Thr Gln Ala Ser Glu Leu Ile Leu  
 115 120 125

Thr Ser Ile Arg Met Gln Asp Asn Gln Leu Ala Glu Leu Asp Ser Glu  
 130 135 140

Lys Thr Ala Leu Gln Val Ser Leu Ala Gly Val Gln Gln Arg Phe Asp  
 145 150 155 160

Ser Tyr Ser Gln Leu Tyr Glu Glu Asn Ile Asn Ala Gly Glu Gln Arg  
 165 170 175

Ala Leu Ala Leu Arg Ser Glu Ser Ala Ile Glu Ser Gln Gly Ala Gln  
 180 185 190

Ile Ser Arg Met Ala Gly Ala Gly Val Asp Met Ala Pro Asn Ile Phe  
195 200 205

Gly Leu Ala Asp Gly Gly Met His Tyr Gly Ala Ile Ala Tyr Ala Ile  
210 215 220

Ala Asp Gly Ile Glu Leu Ser Ala Ser Ala Lys Met Val Asp Ala Glu  
225 230 235 240

Lys Val Ala Gln Ser Glu Ile Tyr Arg Arg Arg Arg Gln Glu Trp Lys  
245 250 255

Ile Gln Arg Asp Asn Ala Gln Ala Glu Ile Asn Gln Leu Asn Ala Gln  
260 265 270

Leu Glu Ser Leu Ser Ile Arg Arg Glu Ala Ala Glu Met Gln Lys Glu  
275 280 285

Tyr Leu Lys Thr Gln Gln Ala Gln Ala Gln Ala Gln Leu Thr Phe Leu  
290 295 300

Arg Ser Lys Phe Ser Asn Gln Ala Leu Tyr Ser Trp Leu Arg Gly Arg  
305 310 315 320

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

Leu Met Ala Glu Gln Ser Tyr Gln Trp Glu Ala Asn Asp Asn Ser Ile  
340 345 350

Ser Phe Val Lys Pro Gly Ala Trp Gln Gly Thr Tyr Ala Gly Leu Leu  
355 360 365

Cys Gly Glu Ala Leu Ile Gln Asn Leu Ala Gln Met Glu Glu Ala Tyr  
370 375 380

Leu Lys Trp Glu Ser Arg Ala Leu Glu Val Glu Arg Thr Val Ser Leu  
385 390 395 400

Ala Val Val Tyr Asp Ser Leu Glu Gly Asn Asp Arg Phe Asn Leu Ala  
405 410 415

Glu Gln Ile Pro Ala Leu Leu Asp Lys Gly Glu Gly Thr Ala Gly Thr  
420 425 430

Lys Glu Asn Gly Leu Ser Leu Ala Asn Ala Ile Leu Ser Ala Ser Val  
435 440 445

Lys Leu Ser Asp Leu Lys Leu Gly Thr Asp Tyr Pro Asp Ser Ile Val  
450 455 460

Gly Ser Asn Lys Val Arg Arg Ile Lys Gln Ile Ser Val Ser Leu Pro  
465 470 475 480

Ala Leu Val Gly Pro Tyr Gln Asp Val Gln Ala Met Leu Ser Tyr Gly  
485 490 495

Gly Ser Thr Gln Leu Pro Lys Gly Cys Ser Ala Leu Ala Val Ser His  
500 505 510

Gly Thr Asn Asp Ser Gly Gln Phe Gln Leu Asp Phe Asn Asp Gly Lys  
515 520 525

Tyr Leu Pro Phe Glu Gly Ile Ala Leu Asp Asp Gln Gly Thr Leu Asn  
530 535 540

Leu Gln Phe Pro Asn Ala Thr Asp Lys Gln Lys Ala Ile Leu Gln Thr  
545 550 555 560

Met Ser Asp Ile Ile Leu His Ile Arg Tyr Thr Ile Arg ...  
565 570 573

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 56:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĹŽKA: 2994 párov báz
- (B) TYP: nukleová kyselina
- (C) REĹAZEC: dvojitý
- (D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: DNA (genómová)

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 56 (tcca):

1	ATG AAT CAA CTC GCC AGT CCC CTG ATT TCC CGC ACC GAA GAG ATC CAC	48
1	Met Asn Gln Leu Ala Ser Pro Leu Ile Ser Arg Thr Glu Glu Ile His	16
49	AAC TTA CCC GGT AAA TTG ACC GAT CTT GGT TAT ACC TCA GTG TTT GAT	96
17	Asn Leu Pro Gly Lys Leu Thr Asp Leu Gly Tyr Thr Ser Val Phe Asp	32
97	GTG GTA CGT ATG CCG CGT GAG CGT TTT ATT CGT GAG CAT CGT GCT GAT	144
33	Val Val Arg Met Pro Arg Glu Arg Phe Ile Arg Glu His Arg Ala Asp	48
145	CTC GGG CGC AGT GCT GAA AAA ATG TAT GAC CTG GCA GTG GGC TAT GCT	192
49	Leu Gly Arg Ser Ala Glu Lys Met Tyr Asp Leu Ala Val Gly Tyr Ala	64
193	CAT CAG GTG TTA CAC CAT TTT CGC CGT AAT TCT CTT AGT GAA GCT GTT	240
65	His Gln Val Leu His His Phe Arg Arg Asn Ser Leu Ser Glu Ala Val	80
241	CAG TTT GGC TTG AGA AGT CCG TTC TCC GTA TCA GGC CCG GAT TAC GCC	288
81	Gln Phe Gly Leu Arg Ser Pro Phe Ser Val Ser Gly Pro Asp Tyr Ala	96
289	AAT CAG TTT CTT GAT GCA AAC ACG GGT TGG AAA GAT AAA GCA CCA AGT	336
97	Asn Gln Phe Leu Asp Ala Asn Thr Gly Trp Lys Asp Lys Ala Pro Ser	112
337	GGA TCA CCG GAA GCC AAT GAT GCG CCG GTA GCC TAT CTG ACT CAT ATT	384
113	Gly Ser Pro Glu Ala Asn Asp Ala Pro Val Ala Tyr Leu Thr His Ile	128
385	TAT CAA TTG GCC CTT GAA CAG GAA AAG AAT GGC GCC ACT ACC ATT ATG	432
129	Tyr Gln Leu Ala Leu Glu Gln Glu Lys Asn Gly Ala Thr Thr Ile Met	144
433	AAT ACG CTG GCG GAG CGT CGC CCC GAT CTG GGT GCT TTG TTA ATT AAT	480
145	Asn Thr Leu Ala Glu Arg Arg Pro Asp Leu Gly Ala Leu Leu Ile Asn	160
481	GAT AAA GCA ATC AAT GAG GTG ATA CCG CAA TTG CAG TTG GTC AAT GAA	528
161	Asp Lys Ala Ile Asn Glu Val Ile Pro Gln Leu Gln Leu Val Asn Glu	176

529 ATT CTG TCC AAA GCT ATT CAG AAG AAA CTG AGT TTG ACT GAT CTG GAA 576  
177 Ile Leu Ser Lys Ala Ile Gln Lys Lys Leu Ser Leu Thr Asp Leu Glu 192

577 GCG GTA AAC GCC AGA CTT TCC ACT ACC CGT TAC CCG AAT AAT CTG CCG 624  
193 Ala Val Asn Ala Arg Leu Ser Thr Thr Arg Tyr Pro Asn Asn Leu Pro 208

625 TAT CAT TAT GGT CAT CAG CAG ATT CAG ACA GCT CAA TCG GTA TTG GGT 672  
209 Tyr His Tyr Gly His Gln Gln Ile Gln Thr Ala Gln Ser Val Leu Gly 224

673 ACT ACG TTG CAA GAT ATC ACT TTG CCA CAG ACG CTG GAT CTG CCG CAA 720  
225 Thr Thr Leu Gln Asp Ile Thr Leu Pro Gln Thr Leu Asp Leu Pro Gln 240

721 AAC TTC TGG GCA ACA GCA AAA GGA AAA CTG AGC GAT ACG ACT GCC AGT 768  
241 Asn Phe Trp Ala Thr Ala Lys Gly Lys Leu Ser Asp Thr Thr Ala Ser 256

769 GCT TTG ACC CGA CTG CAA ATC ATG GCG AGT CAG TTT TCG CCA GAG CAG 816  
257 Ala Leu Thr Arg Leu Gln Ile Met Ala Ser Gln Phe Ser Pro Glu Gln 272

817 CAG AAA ATC ATT ACG GAG ACT GTC GGT CAG GAT TTC TAT CAG CTT AAC 864  
273 Gln Lys Ile Ile Thr Glu Thr Val Gly Gln Asp Phe Tyr Gln Leu Asn 288

865 TAT GGT GAC AGT TCG CTT ACT GTG AAT AGT TTC AGC GAC ATG ACC ATA 912  
289 Tyr Gly Asp Ser Ser Leu Thr Val Asn Ser Phe Ser Asp Met Thr Ile 304

913 ATG ACT GAT CGA ACA AGT TTG ACT GTA CCC CAG GTA GAA CTG ATG TTG 960  
305 Met Thr Asp Arg Thr Ser Leu Thr Val Pro Gln Val Glu Leu Met Leu 320

961 TGT TCA ACT GTC GGA GGT TCT ACG GTT GTT AAG TCT GAT AAT GTG AGT  
1008  
321 Cys Ser Thr Val Gly Gly Ser Thr Val Val Lys Ser Asp Asn Val Ser 336

1009 TCT GGT GAC ACG ACA GCG ACG CCA TTT GCG TAT GGC GCC CGC TTT ATT  
1056  
337 Ser Gly Asp Thr Thr Ala Thr Pro Phe Ala Tyr Gly Ala Arg Phe Ile 352

1057 CAT GCC GGT AAG CCG GAG GCG ATT ACC CTG AGT CGC AGT GGT GCG GAG  
1104  
353 ~~His Ala~~ Gly Lys Pro Glu Ala Ile Thr Leu Ser Arg Ser Gly Ala Glu 368

1105 GCG CAT TTT GCT CTG ACG GTT AAC AAT CTG ACA GAT GAC AAG TTG GAC  
1152  
369 Ala His Phe Ala Leu Thr Val Asn Asn Leu Thr Asp Asp Lys Leu Asp 384

1153 CGT ATT AAC CGC ACA GTG CGC CTG CAA AAA TGG CTG AAT CTG CCT TAT  
1200  
385 Arg Ile Asn Arg Thr Val Arg Leu Gln Lys Trp Leu Asn Leu Pro Tyr 400

1201 GAG GAT ATT GAC CTG TTA GTG ACT TCT GCT ATG GAT GCG GAA ACA GGA  
1248  
401 Glu Asp Ile Asp Leu Leu Val Thr Ser Ala Met Asp Ala Glu Thr Gly 416

1249 AAT ACC GCG CTG TCG ATG AAC GAC AAT ACG CTG CGT ATG TTG GGA GTG  
1296  
417 Asn Thr Ala Leu Ser Met Asn Asp Asn Thr Leu Arg Met Leu Gly Val 432

1297 TTC AAA CAT TAT CAG GCG AAG TAT GGT GTT AGC GCT AAA CAA TTT GCT  
1344  
433 Phe Lys His Tyr Gln Ala Lys Tyr Gly Val Ser Ala Lys Gln Phe Ala 448

1345 GGC TGG CTG CGC GTA GTG GCC CCG TTT GCC ATT ACA CCG GCA ACG CCG  
1392  
449 Gly Trp Leu Arg Val Val Ala Pro Phe Ala Ile Thr Pro Ala Thr Pro 464

1393 TTT TTA GAC CAA GTG TTT AAC TCC GTC GGC ACC TTT GAT ACA CCG-TTT  
1440  
465 Phe Leu Asp Gln Val Phe Asn Ser Val Gly Thr Phe Asp Thr Pro Phe 480

1441 GTG ATA GAT AAT CAG GAT TTT GTC TAT ACA TTG ACC ACC GGG GGC GAT  
1488  
481 Val Ile Asp Asn Gln Asp Phe Val Tyr Thr Leu Thr Thr Gly Gly Asp 496

1489 GGG GCG CGT GTT AAG CAT ATC AGC ACG GCA CTG GGC CTC AAT CAT CGT  
1536  
497 Gly Ala Arg Val Lys His Ile Ser Thr Ala Leu Gly Leu Asn His Arg 512

1537 CAG TTC CTG TTA TTG GCG GAT AAT ATT GCC CGT CAA CAG GGG AAT GTC  
1584  
513 Gln Phe Leu Leu Leu Ala Asp Asn Ile Ala Arg Gln Gln Gly Asn Val 528

1585 ACG CAA AGC ACA CTC AAC TGT AAT CTG TTT GTG GTG TCA GCT TTC TAC  
1632  
529 Thr Gln Ser Thr Leu Asn Cys Asn Leu Phe Val Val Ser Ala Phe Tyr 544

1633 CGT CTG GCT AAT TTG GCG CGC ACA TTG GGG ATA AAT CCA GAG TCT TTC  
1680  
545 Arg Leu Ala Asn Leu Ala Arg Thr Leu Gly Ile Asn Pro Glu Ser Phe 560

1681 TGT GCC TTG GTT GAT CGA TTA GAT GCA GGT ACA GGC ATC GTC TGG CAG  
1728  
561 Cys Ala Leu Val Asp Arg Leu Asp Ala Gly Thr Gly Ile Val Trp Gln 576

1729 CAA TTG GCA GGG AAA CCC ACA ATC ACG GTA CCA CAA AAA GAT TCC CCG  
1776  
577 Gln Leu Ala Gly Lys Pro Thr Ile Thr Val Pro Gln Lys Asp Ser Pro 592

1777 CTG GCG GCG GAT ATT CTG AGT TTG CTG CAA GCG CTA AGT GCG ATT GCT  
1824  
593 Leu Ala Ala Asp Ile Leu Ser Leu Leu Gln Ala Leu Ser Ala Ile Ala 608

1825 CAA TGG CAA CAA CAG CAC GAT TTA GAA TTT TCA GCA CTG CTT TTG CTG  
1872  
609 Gln Trp Gln Gln Gln His Asp Leu Glu Phe Ser Ala Leu Leu Leu Leu 624

1873 TTG AGT GAC AAC CCT ATT TCT ACC TCG CAG GGC ACT GAC GAT CAA TTG  
1920  
625 Leu Ser Asp Asn Pro Ile Ser Thr Ser Gln Gly Thr Asp Asp Gln Leu 640

1921 AAC TTT ATC CGT CAA GTG TGG CAG AAC CTA GGC AGT ACG TTT GTG GGT  
1968  
641 Asn Phe Ile Arg Gln Val Trp Gln Asn Leu Gly Ser Thr Phe Val Gly 656

1969 GCA ACA TTG TTG TCC CGC AGT GGG GCA CCA TTA GTC GAT ACC AAC GGC  
2016  
657 Ala Thr Leu Leu Ser Arg Ser Gly Ala Pro Leu Val Asp Thr Asn Gly 672

2017 CAC GCT ATT GAC TGG TTT GCT CTG CTC TCA GCA GGT AAT AGT CCG CTT  
2064  
673 His Ala Ile Asp Trp Phe Ala Leu Leu Ser Ala Gly Asn Ser Pro Leu 688

2065 ATC GAT AAG GTT GGT CTG GTG ACT GAT GCT GGC ATA CAA AGT GTT ATA  
2112  
689 Ile Asp Lys Val Gly Leu Val Thr Asp Ala Gly Ile Gln Ser Val Ile 704

2113 GCA ACG GTG GTC AAT ACA CAA AGC TTA TCT GAT GAA GAT AAG AAG CTG  
2160  
705 Ala Thr Val Val Asn Thr Gln Ser Leu Ser Asp Glu Asp Lys Lys Leu 720

2161 GCA ATC ACT ACT CTG ACT AAT ACG TTG AAT CAG GTA CAG AAA ACT CAA  
2208  
721 Ala Ile Thr Thr Leu Thr Asn Thr Leu Asn Gln Val Gln Lys Thr Gln 736

2209 CAG GGC GTG GCC GTC AGT CTG TTG GCG CAG ACT CTG AAC GTG AGT CAG  
2256  
737 Gln Gly Val Ala Val Ser Leu Leu Ala Gln Thr Leu Asn Val Ser Gln 752

2257 TCA CTG CCT GCG TTA TTG TTG CGC TGG AGT GGA CAA ACA ACC TAC CAG  
2304  
753 Ser Leu Pro Ala Leu Leu Leu Arg Trp Ser Gly Gln Thr Thr Tyr Gln 768

2305 TGG TTG AGT GCG ACT TGG GCA TTG AAG GAT GCC GTT AAG ACT GCC GCC  
2352  
769 Trp Leu Ser Ala Thr Trp Ala Leu Lys Asp Ala Val Lys Thr Ala Ala 784

2353 GAT ATT CCC GCT GAC TAT CTG CGT CAA TTA CGT GAA GTG GTA CGC CGC  
2400  
785 Asp Ile Pro Ala Asp Tyr Leu Arg Gln Leu Arg Glu Val Val Arg Arg 800

2401 TCC TTG TTG ACC CAA CAA TTC ACG CTG AGT CCT GCA ATG GTG CAA ACC  
2448  
801 Ser Leu Leu Thr Gln Gln Phe Thr Leu Ser Pro Ala Met Val Gln Thr 816

2449 TTG CTG GAC TAT CCA GCC TAT TTT GGC GCT TCC GCA GAA ACA GTG ACC  
2496  
817 Leu Leu Asp Tyr Pro Ala Tyr Phe Gly Ala Ser Ala Glu Thr Val Thr 832

2497 GAT ATC AGT TTG TGG ATG CTT TAT ACC CTG AGC TGT TAT AGC GAT TTA  
2544  
833 Asp Ile Ser Leu Trp Met Leu Tyr Thr Leu Ser Cys Tyr Ser Asp Leu 848

2545 TTG CTC CAA ATG GGT GAA GCT GGT GGT ACC GAA GAT GAT GTA CTG GCC  
2592  
849 Leu Leu Gln Met Gly Glu Ala Gly Gly Thr Glu Asp Asp Val Leu Ala 864

2593 TAC TTA CGC ACA GCT AAT GCT ACC ACA CCG TTG AGC CAA TCT GAT GCT  
2640  
865 Tyr Leu Arg Thr Ala Asn Ala Thr Thr Pro Leu Ser Gln Ser Asp Ala 880

2641 GCA CAG ACG TTG GCA ACG CTA TTG GGT TGG GAG GTT AAC GAG TTG CAA  
2688  
881 Ala Gln Thr Leu Ala Thr Leu Leu Gly Trp Glu Val Asn Glu Leu Gln 896

2689 GCC GCT TGG TCG GTA TTG GGC GGG ATT GCC AAA ACC ACA CCG CAA CTG  
2736  
897 Ala Ala Trp Ser Val Leu Gly Gly Ile Ala Lys Thr Thr Pro Gln Leu 912

2737 GAT GCG CTT CTG CGT TTG CAA CAG GCA CAG AAC CAA ACT GGT CTT GGC  
2784  
913 Asp Ala Leu Leu Arg Leu Gln Gln Ala Gln Asn Gln Thr Gly Leu Gly 928

2785 GTT ACA CAG CAA CAG CAA GGC TAT CTC CTG AGT CGT GAC AGT GAT TAT  
2832  
929 Val Thr Gln Gln Gln Gln Gly Tyr Leu Leu Ser Arg Asp Ser Asp Tyr 944

2833 ACC CTT TGG CAA AGC ACC GGT CAG GCG CTG GTG GCT GGC GTA TCC CAT  
2880  
945 Thr Leu Trp Gln Ser Thr Gly Gln Ala Leu Val Ala Gly Val Ser His 960

2881 GTC AAG GGC AGT AAC TGA GCATGGCAGA GCTCACTACC TGAGTGGATT TGATTT  
2934  
961 Val Lys Gly Ser Asn End 965

2935 TTCCGTATGG CCTAATGAGG CTATTTCTAA ACCGCCATTT AAGTAAGGCA GATAATTATG  
2994

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 57:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĹŽKA: 965 aminokyselín
- (B) TYP: aminokyselina
- (C) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: proteín

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 57 (TccA peptid):

Charakt. znak	od	do	Opis
	1	10	(SEKV. ID. ČÍS. 8)

1 Met Asn Gln Leu Ala Ser Pro Leu Ile Ser Arg Thr Glu Glu Ile His 16  
17 Asn Leu Pro Gly Lys Leu Thr Asp Leu Gly Tyr Thr Ser Val Phe Asp 32  
33 Val Val Arg Met Pro Arg Glu Arg Phe Ile Arg Glu His Arg Ala Asp 48  
49 Leu Gly Arg Ser Ala Glu Lys Met Tyr Asp Leu Ala Val Gly Tyr Ala 64  
65 His Gln Val Leu His His Phe Arg Arg Asn Ser Leu Ser Glu Ala Val 80  
81 Gln Phe Gly Leu Arg Ser Pro Phe Ser Val Ser Gly Pro Asp Tyr Ala 96

97 Asn Gln Phe Leu Asp Ala Asn Thr Gly Trp Lys Asp Lys Ala Pro Ser 112  
113 Gly Ser Pro Glu Ala Asn Asp Ala Pro Val Ala Tyr Leu Thr His Ile 128  
129 Tyr Gln Leu Ala Leu Glu Gln Glu Lys Asn Gly Ala Thr Thr Ile Met 144  
145 Asn Thr Leu Ala Glu Arg Arg Pro Asp Leu Gly Ala Leu Leu Ile Asn 160  
161 Asp Lys Ala Ile Asn Glu Val Ile Pro Gln Leu Gln Leu Val Asn Glu 176  
177 Ile Leu Ser Lys Ala Ile Gln Lys Lys Leu Ser Leu Thr Asp Leu Glu 192

193 Ala Val Asn Ala Arg Leu Ser Thr Thr Arg Tyr Pro Asn Asn Leu Pro 208  
209 Tyr His Tyr Gly His Gln Gln Ile Gln Thr Ala Gln Ser Val Leu Gly 224  
225 Thr Thr Leu Gln Asp Ile Thr Leu Pro Gln Thr Leu Asp Leu Pro Gln 240  
241 Asn Phe Trp Ala Thr Ala Lys Gly Lys Leu Ser Asp Thr Thr Ala Ser 256  
257 Ala Leu Thr Arg Leu Gln Ile Met Ala Ser Gln Phe Ser Pro Glu Gln 272  
273 Gln Lys Ile Ile Thr Glu Thr Val Gly Gln Asp Phe Tyr Gln Leu Asn 288  
289 Tyr Gly Asp Ser Ser Leu Thr Val Asn Ser Phe Ser Asp Met Thr Ile 304  
305 Met Thr Asp Arg Thr Ser Leu Thr Val Pro Gln Val Glu Leu Met Leu 320  
321 Cys Ser Thr Val Gly Gly Ser Thr Val Val Lys Ser Asp Asn Val Ser 336  
337 Ser Gly Asp Thr Thr Ala Thr Pro Phe Ala Tyr Gly Ala Arg Phe Ile 352  
353 His Ala Gly Lys Pro Glu Ala Ile Thr Leu Ser Arg Ser Gly Ala Glu 368  
369 Ala His Phe Ala Leu Thr Val Asn Asn Leu Thr Asp Asp Lys Leu Asp 384  
385 Arg Ile Asn Arg Thr Val Arg Leu Gln Lys Trp Leu Asn Leu Pro Tyr 400  
401 Glu Asp Ile Asp Leu Leu Val Thr Ser Ala Met Asp Ala Glu Thr Gly 416  
417 Asn Thr Ala Leu Ser Met Asn Asp Asn Thr Leu Arg Met Leu Gly Val 432  
433 Phe Lys His Tyr Gln Ala Lys Tyr Gly Val Ser Ala Lys Gln Phe Ala 448  
449 Gly Trp Leu Arg Val Val Ala Pro Phe Ala Ile Thr Pro Ala Thr Pro 464  
465 Phe Leu Asp Gln Val Phe Asn Ser Val Gly Thr Phe Asp Thr Pro Phe 480  
481 Val Ile Asp Asn Gln Asp Phe Val Tyr Thr Leu Thr Thr Gly Gly Asp 496  
497 Gly Ala Arg Val Lys His Ile Ser Thr Ala Leu Gly Leu Asn His Arg 512  
513 Gln Phe Leu Leu Leu Ala Asp Asn Ile Ala Arg Gln Gln Gly Asn Val 528  
529 Thr Gln Ser Thr Leu Asn Cys Asn Leu Phe Val Val Ser Ala Phe Tyr 544  
545 Arg Leu Ala Asn Leu Ala Arg Thr Leu Gly Ile Asn Pro Glu Ser Phe 560  
561 Cys Ala Leu Val Asp Arg Leu Asp Ala Gly Thr Gly Ile Val Trp Gln 576  
577 Gln Leu Ala Gly Lys Pro Thr Ile Thr Val Pro Gln Lys Asp Ser Pro 592  
593 Leu Ala Ala Asp Ile Leu Ser Leu Leu Gln Ala Leu Ser Ala Ile Ala 608  
609 Gln Trp Gln Gln Gln His Asp Leu Glu Phe Ser Ala Leu Leu Leu Leu 624  
625 Leu Ser Asp Asn Pro Ile Ser Thr Ser Gln Gly Thr Asp Asp Gln Leu 640  
641 Asn Phe Ile Arg Gln Val Trp Gln Asn Leu Gly Ser Thr Phe Val Gly 656



657 Ala Thr Leu Leu Ser Arg Ser Gly Ala Pro Leu Val Asp Thr Asn Gly 672  
673 His Ala Ile Asp Trp Phe Ala Leu Leu Ser Ala Gly Asn Ser Pro Leu 688  
689 Ile Asp Lys Val Gly Leu Val Thr Asp Ala Gly Ile Gln Ser Val Ile 704  
705 Ala Thr Val Val Asn Thr Gln Ser Leu Ser Asp Glu Asp Lys Lys Leu 720  
721 Ala Ile Thr Thr Leu Thr Asn Thr Leu Asn Gln Val Gln Lys Thr Gln 736  
737 Gln Gly Val Ala Val Ser Leu Leu Ala Gln Thr Leu Asn Val Ser Gln 752  
  
753 Ser Leu Pro Ala Leu Leu Leu Arg Trp Ser Gly Gln Thr Thr Tyr Gln 768  
769 Trp Leu Ser Ala Thr Trp Ala Leu Lys Asp Ala Val Lys Thr Ala Ala 784  
785 Asp Ile Pro Ala Asp Tyr Leu Arg Gln Leu Arg Glu Val Val Arg Arg 800  
801 Ser Leu Leu Thr Gln Gln Phe Thr Leu Ser Pro Ala Met Val Gln Thr 816  
817 Leu Leu Asp Tyr Pro Ala Tyr Phe Gly Ala Ser Ala Glu Thr Val Thr 832  
833 Asp Ile Ser Leu Trp Met Leu Tyr Thr Leu Ser Cys Tyr Ser Asp Leu 848  
849 Leu Leu Gln Met Gly Glu Ala Gly Gly Thr Glu Asp Asp Val Leu Ala 864  
865 Tyr Leu Arg Thr Ala Asn Ala Thr Thr Pro Leu Ser Gln Ser Asp Ala 880  
881 Ala Gln Thr Leu Ala Thr Leu Leu Gly Trp Glu Val Asn Glu Leu Gln 896  
897 Ala Ala Trp Ser Val Leu Gly Gly Ile Ala Lys Thr Thr Pro Gln Leu 912  
913 Asp Ala Leu Leu Arg Leu Gln Gln Ala Gln Asn Gln Thr Gly Leu Gly 928  
929 Val Thr Gln Gln Gln Gln Gly Tyr Leu Leu Ser Arg Asp Ser Asp Tyr 944  
945 Thr Leu Trp Gln Ser Thr Gly Gln Ala Leu Val Ala Gly Val Ser His 960  
961 Val Lys Gly Ser Asn 965

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 58:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĚLKA: 4932 párov báz
- (B) TYP: nukleová kyselina
- (C) REŤAZEC: dvojité
- (D) TOPOLÓGIA: lineárna

(ii) TYP MOLEKULY: DNA (genómová)

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 58 (tccb):

1	ATG	TTA	TCG	ACA	ATG	GAA	AAA	CAA	CTG	AAT	GAA	TCC	CAG	CGT	GAT	GCG	48
1	Met	Leu	Ser	Thr	Met	Glu	Lys	Gln	Leu	Asn	Glu	Ser	Gln	Arg	Asp	Ala	16
49	TTG	GTG	ACT	GGC	TAT	ATG	AAT	TTT	GTG	GCG	CCG	ACG	TTG	AAA	GGC	GTC	96
17	Leu	Val	Thr	Gly	Tyr	Met	Asn	Phe	Val	Ala	Pro	Thr	Leu	Lys	Gly	Val	32
97	AGT	GGT	CAG	CCG	GTG	ACG	GTG	GAA	GAT	TTA	TAC	GAA	TAT	TTG	CTG	ATT	144
33	Ser	Gly	Gln	Pro	Val	Thr	Val	Glu	Asp	Leu	Tyr	Glu	Tyr	Leu	Leu	Ile	48
145	GAC	CCG	GAA	GTG	GCT	GAT	GAG	GTT	GAG	ACG	AGT	CGG	GTA	GCA	CAA	GCG	192
49	Asp	Pro	Glu	Val	Ala	Asp	Glu	Val	Glu	Thr	Ser	Arg	Val	Ala	Gln	Ala	64
193	ATT	GCC	AGC	ATA	CAG	CAA	TAT	ATG	ACT	CGT	CTG	GTC	AAC	GGC	TCT	GAA	240
65	Ile	Ala	Ser	Ile	Gln	Gln	Tyr	Met	Thr	Arg	Leu	Val	Asn	Gly	Ser	Glu	80
241	CCG	GGG	CGT	CAG	GCG	ATG	GAG	CCT	TCT	ACA	GCT	AAC	GAA	TGG	CGT	GAT	288
81	Pro	Gly	Arg	Gln	Ala	Met	Glu	Pro	Ser	Thr	Ala	Asn	Glu	Trp	Arg	Asp	96
289	AAT	GAT	AAC	CAA	TAT	GCT	ATC	TGG	GCT	GCG	GGG	GCT	GAG	GTT	CGA	AAT	336
97	Asn	Asp	Asn	Gln	Tyr	Ala	Ile	Trp	Ala	Ala	Gly	Ala	Glu	Val	Arg	Asn	112
337	TAC	GCT	GAA	AAC	TAT	ATT	TCA	CCC	ATC	ACC	CGG	CAG	GAA	AAA	AGC	CAT	384
113	Tyr	Ala	Glu	Asn	Tyr	Ile	Ser	Pro	Ile	Thr	Arg	Gln	Glu	Lys	Ser	His	128
385	TAT	TTC	TCG	GAG	CTG	GAG	ACG	ACT	TTA	AAT	CAG	AAT	CGA	CTC	GAT	CCG	432
129	Tyr	Phe	Ser	Glu	Leu	Glu	Thr	Thr	Leu	Asn	Gln	Asn	Arg	Leu	Asp	Pro	144
433	GAT	CGT	GTG	CAG	GAT	GCT	GTT	TTG	GCG	TAT	CTC	AAT	GAG	TTT	GAG	GCA	480
145	Asp	Arg	Val	Gln	Asp	Ala	Val	Leu	Ala	Tyr	Leu	Asn	Glu	Phe	Glu	Ala	160
481	GTG	AGT	AAT	CTA	TAT	GTG	CTC	AGT	GGT	TAT	ATT	AAT	CAG	GAT	AAA	TTT	528
161	Val	Ser	Asn	Leu	Tyr	Val	Leu	Ser	Gly	Tyr	Ile	Asn	Gln	Asp	Lys	Phe	176
529	GAC	CAA	GCT	ATC	TAC	TAC	TTT	ATT	GGT	CGC	ACT	ACC	ACT	AAA	CCG	TAT	576
177	Asp	Gln	Ala	Ile	Tyr	Tyr	Phe	Ile	Gly	Arg	Thr	Thr	Thr	Lys	Pro	Tyr	192
577	CGC	TAC	TAC	TGG	CGT	CAG	ATG	GAT	TTG	AGT	AAG	AAC	CGT	CAA	GAT	CCG	624
193	Arg	Tyr	Tyr	Trp	Arg	Gln	Met	Asp	Leu	Ser	Lys	Asn	Arg	Gln	Asp	Pro	208
625	GCA	GGG	AAT	CCG	GTG	ACG	CCA	AAT	TGC	TGG	AAT	GAT	TGG	CAG	GAA	ATC	672
209	Ala	Gly	Asn	Pro	Val	Thr	Pro	Asn	Cys	Trp	Asn	Asp	Trp	Gln	Glu	Ile	224
673	ACT	TTG	CCG	CTG	TCT	GGT	GAT	ACG	GTG	CTG	GAG	CAT	ACA	GTT	CGC	CCG	720
225	Thr	Leu	Pro	Leu	Ser	Gly	Asp	Thr	Val	Leu	Glu	His	Thr	Val	Arg	Pro	240
721	GTA	TTT	TAT	AAT	GAT	CGA	CTA	TAT	GTG	GCT	TGG	GTT	GAG	CGT	GAC	CCG	768
241	Val	Phe	Tyr	Asn	Asp	Arg	Leu	Tyr	Val	Ala	Trp	Val	Glu	Arg	Asp	Pro	256
769	GCA	GTA	CAG	AAG	GAT	GCT	GAC	GGT	AAA	AAC	ATC	GGT	AAA	ACC	CAT	GCC	816
257	Ala	Val	Gln	Lys	Asp	Ala	Asp	Gly	Lys	Asn	Ile	Gly	Lys	Thr	His	Ala	272
817	TAC	AAC	ATA	AAG	TTT	GGT	TAT	AAA	CGT	TAT	GAT	GAT	ACT	TGG	ACA	GCG	864
273	Tyr	Asn	Ile	Lys	Phe	Gly	Tyr	Lys	Arg	Tyr	Asp	Asp	Thr	Trp	Thr	Ala	288

865	CCG	AAT	ACG	ACC	ACG	TTA	ATG	ACA	CAA	CAA	GCA	GGG	GAA	AGT	TCA	GAA	912
289	Pro	Asn	Thr	Thr	Thr	Leu	Met	Thr	Gln	Gln	Ala	Gly	Glu	Ser	Ser	Glu	304
913	ACA	CAG	CGA	TCC	AGC	CTG	CTG	ATT	GAT	GAA	TCT	AGC	ACC	ACA	TTG	CGC	960
305	Thr	Gln	Arg	Ser	Ser	Leu	Leu	Ile	Asp	Glu	Ser	Ser	Thr	Thr	Leu	Arg	320
961	CAA	GTT	AAT	CTG	TTG	GCT	ACC	ACC	GAT	TTT	AGT	ATC	GAT	CCG	ACG	GAG	
1008																	
321	Gln	Val	Asn	Leu	Leu	Ala	Thr	Thr	Asp	Phe	Ser	Ile	Asp	Pro	Thr	Glu	336
1009	GAA	ACG	GAC	AGT	AAC	CCG	TAT	GGC	CGC	CTA	ATG	TTG	GGG	GTG	TTT	GTC	
1056																	
337	Glu	Thr	Asp	Ser	Asn	Pro	Tyr	Gly	Arg	Leu	Met	Leu	Gly	Val	Phe	Val	352
1057	CGT	CAA	TTT	GAA	GGT	GAT	GGG	GCC	AAT	AGA	AAA	AAT	AAA	CCC	GTT	GTT	
1104																	
353	Arg	Gln	Phe	Glu	Gly	Asp	Gly	Ala	Asn	Arg	Lys	Asn	Lys	Pro	Val	Val	368
1105	TAT	GGT	TAT	CTC	TAT	TGT	GAC	TCA	GCT	TTC	AAT	CGT	CAT	GTT	CTC	AGG	
1152																	
369	Tyr	Gly	Tyr	Leu	Tyr	Cys	Asp	Ser	Ala	Phe	Asn	Arg	His	Val	Leu	Arg	384
1153	CCG	TTA	AGT	AAG	AAC	TTT	TTG	TTC	AGT	ACT	TAC	CGT	GAT	GAA	ACG	GAT	
1200																	
385	Pro	Leu	Ser	Lys	Asn	Phe	Leu	Phe	Ser	Thr	Tyr	Arg	Asp	Glu	Thr	Asp	400
1201	GGT	CAA	AAC	AGC	TTG	CAA	TTT	GCG	GTA	TAC	GAT	AAA	AAG	TAT	GTA	ATT	
1248																	
401	Gly	Gln	Asn	Ser	Leu	Gln	Phe	Ala	Val	Tyr	Asp	Lys	Lys	Tyr	Val	Ile	416
1249	ACT	AAG	GTT	GTT	ACA	GGT	GCA	ACG	GAA	GAT	CCC	GAA	AAT	ACA	GGA	TGG	
1296																	
417	Thr	Lys	Val	Val	Thr	Gly	Ala	Thr	Glu	Asp	Pro	Glu	Asn	Thr	Gly	Trp	432
1297	GTA	AGT	AAA	GTT	GAT	GAC	TTG	AAA	CAA	GGC	ACT	ACT	GGG	GCC	TAT	GTG	
1344																	
433	Val	Ser	Lys	Val	Asp	Asp	Leu	Lys	Gln	Gly	Thr	Thr	Gly	Ala	Tyr	Val	448
1345	TAT	ATC	GAT	CAA	GAT	GGC	CTG	ACG	CTT	CAT	ATA	CAA	ACC	ACA	ACT	AAT	
1392																	
449	Tyr	Ile	Asp	Gln	Asp	Gly	Leu	Thr	Leu	His	Ile	Gln	Thr	Thr	Thr	Asn	464
1393	GGG	GAT	TTT	ATT	AAC	CGT	CAT	ACG	TTT	GGA	TAT	AAC	GAT	CTT	GTA	TAT	
1440																	
465	Gly	Asp	Phe	Ile	Asn	Arg	His	Thr	Phe	Gly	Tyr	Asn	Asp	Leu	Val	Tyr	480
1441	GAT	TCT	AAG	TCT	GGT	TAT	GGT	TTC	ACG	TGG	TCA	GGA	AAT	GAA	GGT	TTT	
1488																	
481	Asp	Ser	Lys	Ser	Gly	Tyr	Gly	Phe	Thr	Trp	Ser	Gly	Asn	Glu	Gly	Phe	496
1489	TAT	CTG	GAT	TAC	CAT	GAT	GGA	AAT	TAT	TAC	ACC	TTT	CAT	AAT	GCA	ATA	
1536																	
497	Tyr	Leu	Asp	Tyr	His	Asp	Gly	Asn	Tyr	Tyr	Thr	Phe	His	Asn	Ala	Ile	512

1537 ATC AAC TAC TAT CCG TCT GGA TAT GGT GGT GGA TCT GTT CCT AAT GGA  
1584  
513 Ile Asn Tyr Tyr Pro Ser Gly Tyr Gly Gly Gly Ser Val Pro Asn Gly 528

1585 ACG TGG GCG TTA GAG CAA AGG ATT AAT GAG GGA TGG GCT ATT GCT CCC  
1632  
529 Thr Trp Ala Leu Glu Gln Arg Ile Asn Glu Gly Trp Ala Ile Ala Pro 544

1633 CTG CTT GAT ACT CTC CAT ACT GTT ACT GTG AAG GGC AGT TAT ATC GCT  
1680  
545 Leu Leu Asp Thr Leu His Thr Val Thr Val Lys Gly Ser Tyr Ile Ala 560

1681 TGG GAA GGG GAA ACA CCT ACC GGT TAT AAT CTG TAT ATT CCA GAT GGT  
1728  
561 Trp Glu Gly Glu Thr Pro Thr Gly Tyr Asn Leu Tyr Ile Pro Asp Gly 576

1729 ACC GTG TTG CTA GAT TGG TTT GAT AAA ATA AAT TTT GCT ATT GGT CTT  
1776  
577 Thr Val Leu Leu Asp Trp Phe Asp Lys Ile Asn Phe Ala Ile Gly Leu 592

1777 AAT AAG CTT GAG TCT GTA TTT ACG TCG CCA GAT TGG CCA ACA CTA ACC  
1824  
593 Asn Lys Leu Glu Ser Val Phe Thr Ser Pro Asp Trp Pro Thr Leu Thr 608

1825 ACT ATC AAA AAT TTC AGT AAA ATC GCC GAT AAC CGC AAA TTC TAT CAG  
1872  
609 Thr Ile Lys Asn Phe Ser Lys Ile Ala Asp Asn Arg Lys Phe Tyr Gln 624

1873 GAA ATC AAT GCT GAG ACG GCG GAT GGA CGC AAC CTG TTT AAA CGT TAC  
1920  
625 Glu Ile Asn Ala Glu Thr Ala Asp Gly Arg Asn Leu Phe Lys Arg Tyr 640

1921 AGT ACT CAA ACT TTC GGA CTT ACC AGC GGT GCG ACT TAT TCT ACA ACT  
1968  
641 Ser Thr Gln Thr Phe Gly Leu Thr Ser Gly Ala Thr Tyr Ser Thr Thr 656

1969 TAT ACT TTG TCT GAG GCG GAT TTC TCC ACT GAT CCG GAC AAA AAC TAC  
2016  
657 Tyr Thr Leu Ser Glu Ala Asp Phe Ser Thr Asp Pro Asp Lys Asn Tyr 672

2017 CTA CAG GTT TGT TTG AAT GTC GTG TGG GAT CAT TAT GAC CGC CCG TCA  
2064  
673 Leu Gln Val Cys Leu Asn Val Val Trp Asp His Tyr Asp Arg Pro Ser 688

2065 GGG AAA AAA GGG GCT TAT TCT TGG GTC AGT AAG TGG TTT AAC GTC TAT  
2112  
689 Gly Lys Lys Gly Ala Tyr Ser Trp Val Ser Lys Trp Phe Asn Val Tyr 704

2113 GTT GCG TTG CAA GAT AGC AAA GCT CCG GAT GCC ATT CCT CGA TTA GTT  
2160  
705 Val Ala Leu Gln Asp Ser Lys Ala Pro Asp Ala Ile Pro Arg Leu Val 720

2161 TCC CGT TAC GAT AGT AAA CGT GGT CTG GTG CAA TAT CTG GAC TTC TGG  
2208  
721 Ser Arg Tyr Asp Ser Lys Arg Gly Leu Val Gln Tyr Leu Asp Phe Trp 736

2209 ACC TCA TCA TTA CCC GCG AAA ACC CGT CTT AAC ACC ACC TTT GTG CGT  
2256  
737 Thr Ser Ser Leu Pro Ala Lys Thr Arg Leu Asn Thr Thr Phe Val Arg 752

2257 ACT TTG ATT GAG AAG GCT AAT CTG GGG CTG GAT AGT TTG CTG GAT TAC  
2304  
753 Thr Leu Ile Glu Lys Ala Asn Leu Gly Leu Asp Ser Leu Leu Asp Tyr 768

2305 ACC TTG CAG GCA GAT CCT TCT CTG GAA GCA GAT TTA GTG ACT GAC GGC  
2352  
769 Thr Leu Gln Ala Asp Pro Ser Leu Glu Ala Asp Leu Val Thr Asp Gly 784

2353 AAA AGC GAA CCA ATG GAC TTT AAT GGT TCA AAC GGT CTC TAT TTC TGG  
2400  
785 Lys Ser Glu Pro Met Asp Phe Asn Gly Ser Asn Gly Leu Tyr Phe Trp 800

2401 GAA TTG TTC TTT CAC CTG CCG TTT TTG GTT GCT ACA CGC TTT GCC AAC  
2448  
801 Glu Leu Phe Phe His Leu Pro Phe Leu Val Ala Thr Arg Phe Ala Asn 816

2449 GAA CAG CAA TTT TCG CCG GCA CAA AAG AGT TTG CAT TAC ATC TTT GAC  
2496  
817 Glu Gln Gln Phe Ser Pro Ala Gln Lys Ser Leu His Tyr Ile Phe Asp 832

2497 CCG GCG ATG AAA AAC AAG CCA CAC AAT GCC CCG GCT TAT TGG AAT GTA  
2544  
833 Pro Ala Met Lys Asn Lys Pro His Asn Ala Pro Ala Tyr Trp Asn Val 848

2545 CGT CCG TTG GTT GAA GGA AAC AGC GAT TTG TCA CGT CAT TTG GAC GAT  
2592  
849 Arg Pro Leu Val Glu Gly Asn Ser Asp Leu Ser Arg His Leu Asp Asp 864

2593 TCT ATA GAC CCA GAT ACT CAA GCT TAT GCT CAT CCG GTG ATA TAC CAG  
2640  
865 Ser Ile Asp Pro Asp Thr Gln Ala Tyr Ala His Pro Val Ile Tyr Gln 880

2641 AAA GCG GTG TTT ATT GCC TAT GTC AGT AAC CTG ATT GCT CAG GGA GAT  
2688  
881 Lys Ala Val Phe Ile Ala Tyr Val Ser Asn Leu Ile Ala Gln Gly Asp 896

2689 ATG TGG TAT CGC CAA TTG ACT CGT GAC GGT CTG ACT CAG GCC CGT GTC  
2736  
897 Met Trp Tyr Arg Gln Leu Thr Arg Asp Gly Leu Thr Gln Ala Arg Val 912

2737 TAT TAC AAT CTG GCC GCT GAA TTG CTA GGG CCT CGT CCG GAT GTA TCG  
2784  
913 Tyr Tyr Asn Leu Ala Ala Glu Leu Leu Gly Pro Arg Pro Asp Val Ser 928

2785 CTG AGT AGC ATT TGG ACG CCG CAA ACC CTG GAT ACC TTA GCA GCC GGG  
2832  
929 Leu Ser Ser Ile Trp Thr Pro Gln Thr Leu Asp Thr Leu Ala Ala Gly 944

2833 CAA AAA GCG GTT TTA CGT GAT TTT GAG CAC CAG TTG GCT AAT AGT GAT  
2880  
945 Gln Lys Ala Val Leu Arg Asp Phe Glu His Gln Leu Ala Asn Ser Asp 960

2881 ACC GCT TTA CCC GCA TTG CCG GGC CGC AAT GTC AGC TAC TTG AAA CTG  
2928  
961 Thr Ala Leu Pro Ala Leu Pro Gly Arg Asn Val Ser Tyr Leu Lys Leu 976

2929 GCA GAT AAT GGC TAC TTT AAT GAA CCG CTC AAT GTT CTG ATG TTG TCT  
2976  
977 Ala Asp Asn Gly Tyr Phe Asn Glu Pro Leu Asn Val Leu Met Leu Ser 992

2977 CAC TGG GAT ACG TTG GAT GCA CCG TTA TAC AAT CTG CGT CAT AAC CTG  
3024  
993 His Trp Asp Thr Leu Asp Ala Arg Leu Tyr Asn Leu Arg His Asn Leu  
1008

3025 ACC GTT GAT GGC AAG CCG CTT TCG CTG CCG CTG TAT GCT GCG CCT GTT  
3072  
1009 Thr Val Asp Gly Lys Pro Leu Ser Leu Pro Leu Tyr Ala Ala Pro Val  
1024

3073 GAT CCG GTA GCG TTG TTG GCT CAG CGT GCT CAG TCC GGC ACG TTG ACG  
3120  
1025 Asp Pro Val Ala Leu Leu Ala Gln Arg Ala Gln Ser Gly Thr Leu Thr  
1040

3121 AAT GGC GTC AGT GGC GCC ATG TTG ACG GTG CCG CCA TAC CGT TTC AGC  
3168  
1041 Asn Gly Val Ser Gly Ala Met Leu Thr Val Pro Pro Tyr Arg Phe Ser  
1056

3169 GCT ATG TTG CCG CGA GCT TAC AGC GCC GTG GGT ACG TTG ACC AGT TTT  
3216  
1057 Ala Met Leu Pro Arg Ala Tyr Ser Ala Val Gly Thr Leu Thr Ser Phe  
1072

3217 GGT CAG AAC CTG CTT AGT TTG TTG GAA CGT AGC GAA CGA GCC TGT CAA  
3264  
1073 Gly Gln Asn Leu Leu Ser Leu Leu Glu Arg Ser Glu Arg Ala Cys Gln  
1088

3265 GAA GAG TTG GCG CAA CAG CAA CTG TTG GAT ATG TCC AGC TAT GCC ATC  
3312  
1089 Glu Glu Leu Ala Gln Gln Gln Leu Leu Asp Met Ser Ser Tyr Ala Ile  
1104

3313 ACG TTG CAA CAA CAG GCG CTG GAT GGA TTG GCG GCA GAT CGT CTG GCG  
3360  
1105 Thr Leu Gln Gln Gln Ala Leu Asp Gly Leu Ala Ala Asp Arg Leu Ala  
1120

3361 CTG CTA GCT AGT CAG GCT ACG GCA CAA CAG CGT CAT GAC CAT TAT TAC  
3408  
1121 Leu Leu Ala Ser Gln Ala Thr Ala Gln Gln Arg His Asp His Tyr Tyr  
1136

3409 ACT CTG TAT CAG AAC AAC ATC TCC AGT GCG GAA CAA CTG GTG ATG GAC  
3456  
1137 Thr Leu Tyr Gln Asn Asn Ile Ser Ser Ala Glu Gln Leu Val Met Asp  
1152

3457 ACC CAA AGG TCA GCA CAA TCC CTG ATT TCT TCT TCC ACT GGT GTA CAA  
3504  
1153 Thr Gln Thr Ser Ala Gln Ser Leu Ile Ser Ser Ser Thr Gly Val Gln  
1168

3505 ACT GCC AGT GGG GCA CTG AAA GTG ATC CCG AAT ATC TTT GGT TTG GCT  
3552  
1169 Thr Ala Ser Gly Ala Leu Lys Val Ile Pro Asn Ile Phe Gly Leu Ala  
1184

3553 GAT GGC GGC TCG CGC TAT GAA GGA GTA ACG GAA GCG ATT GCC ATC GGG  
3600  
1185 Asp Gly Gly Ser Arg Tyr Glu Gly Val Thr Glu Ala Ile Ala Ile Gly  
1200

3601 TTA ATG GCT GCC GGA CAA GCC ACC AGC GTG GTG GCC GAG CGT CTG GCA  
3648  
1201 Leu Met Ala Ala Gly Gln Ala Thr Ser Val Val Ala Glu Arg Leu Ala  
1216

3649 ACC ACG GAG AAT TAC CGC CGC CGC CGT GAA GAG TGG CAA ATC CAA TAC  
3696  
1217 Thr Thr Glu Asn Tyr Arg Arg Arg Arg Glu Glu Trp Gln Ile Gln Tyr  
1232

3697 CAG CAG GCA CAG TCT GAG GTC GAC GCA TTA CAG AAA CAG TTG GAT GCG  
3744  
1233 Gln Gln Ala Gln Ser Glu Val Asp Ala Leu Gln Lys Gln Leu Asp Ala  
1248

3745 CTG GCA GTG CGC GAG AAA GCA GCT CAA ACT TCC CTG CAA CAG GCG AAG  
3792  
1249 Leu Ala Val Arg Glu Lys Ala Ala Gln Thr Ser Leu Gln Gln Ala Lys  
1264

3793 GCA CAG CAG GTA CAA ATT CGG ACC ATG CTG ACT TAC TTA ACT ACT CGT  
3840  
1265 Ala Gln Gln Val Gln Ile Arg Thr Met Leu Thr Tyr Leu Thr Thr Arg  
1280

3841 TTC ACC CAG GCG ACT CTG TAC CAG TGG CTG AGT GGT CAA TTA TCC GCG  
3888  
1281 Phe Thr Gln Ala Thr Leu Tyr Gln Trp Leu Ser Gly Gln Leu Ser Ala  
1296

3889 TTG TAT TAT CAA GCG TAT GAT GCC GTG GTT GCT CTC TGC CTC TCC GCC  
3936  
1297 Leu Tyr Tyr Gln Ala Tyr Asp Ala Val Val Ala Leu Cys Leu Ser Ala  
1312

3937 CAA GCT TGC TGG CAG TAT GAA TTG GGT GAT TAC GCT ACC ACT TTT ATC  
3984  
1313 Gln Ala Cys Trp Gln Tyr Glu Leu Gly Asp Tyr Ala Thr Thr Phe Ile  
1328

3985 CAG ACC GGT ACC TGG AAC GAC CAT TAC CGT GGT TTG CAA GTG GGG GAG  
4032  
1329 Gln Thr Gly Thr Trp Asn Asp His Tyr Arg Gly Leu Gln Val Gly Glu  
1344

4033 ACA CTG CAA CTC AAT TTG CAT CAG ATG GAA GCG GCC TAT TTA GTT CGT  
4080  
1345 Thr Leu Gln Leu Asn Leu His Gln Met Glu Ala Ala Tyr Leu Val Arg  
1360

4081 CAC GAA CGC CGT CTT AAT GTG ATC CGT ACT GTG TCG CTC AAA AGC CTA  
4128  
1361 His Glu Arg Arg Leu Asn Val Ile Arg Thr Val Ser Leu Lys Ser Leu  
1376

~~4129 TTG GGT~~ GAT GAT GGT TTT GGT AAG TTA AAA ACC GAA GGC AAA GTC GAC  
4176  
1377 Leu Gly Asp Asp Gly Phe Gly Lys Leu Lys Thr Glu Gly Lys Val Asp  
1392

4177 TTT CCA TTA AGC GAA AAG CTG TTT GAC AAC GAC TAT CCG GGG CAC TAT  
4224  
1393 Phe Pro Leu Ser Glu Lys Leu Phe Asp Asn Asp Tyr Pro Gly His Tyr  
1408

4225 TTG CGC CAG ATT AAA ACT GTG TCA GTG ACG TTG CCG ACG TTA GTC GGG  
4272  
1409 Leu Arg Gln Ile Lys Thr Val Ser Val Thr Leu Pro Thr Leu Val Gly  
1424

4273 CCG TAT CAA AAC GTG AAG GCA ACG CTC ACT CAG ACC AGC AGC AGT ATA  
4320  
1425 Pro Tyr Gln Asn Val Lys Ala Thr Leu Thr Gln Thr Ser Ser Ser Ile  
1440

4321 TTG TTA GCA GCA GAT ATC AAT GGT GTT AAA CGT CTC AAT GAT CCG ACA  
4368  
1441 Leu Leu Ala Ala Asp Ile Asn Gly Val Lys Arg Leu Asn Asp Pro Thr  
1456

4369 GGT AAA GAG GGT GAT GCG ACG CAT ATT GTC ACC AAT CTG CGT GCC AGC  
4416  
1457 Gly Lys Glu Gly Asp Ala Thr His Ile Val Thr Asn Leu Arg Ala Ser  
1472

4417 CAG CAG GTG GCG CTC TCT TCT GGC ATT AAT GAT GCC GGT AGC TTT GAG  
4464  
1473 Gln Gln Val Ala Leu Ser Ser Gly Ile Asn Asp Ala Gly Ser Phe Glu  
1488

4465 TTG CGT TTG GAA GAT GAG CGC TAT CTA TCA TTT GAG GGG ACT GGA GCT  
4512  
1489 Leu Arg Leu Glu Asp Glu Arg Tyr Leu Ser Phe Glu Gly Thr Gly Ala  
1504



4513 GTT TCC AAA TGG ACT CTT AAC TTC CCG CGT TCT GTG GAT GAG CAT ATT  
4560  
1505 Val Ser Lys Trp Thr Leu Asn Phe Pro Arg Ser Val Asp Glu His Ile  
1520

4561 GAC GAT AAG ACA TTG AAA GCG GAT GAG ATG CAG GCC GCA CTG TTG GCG  
4608  
1521 Asp Asp Lys Thr Leu Lys Ala Asp Glu Met Gln Ala Ala Leu Leu Ala  
1536

4609 AAT ATG GAT GAT GTG CTG GTG CAG GTG CAT TAT ACC GCC TGC GAC GGC  
4656  
1537 Asn Met Asp Asp Val Leu Val Gln Val His Tyr Thr Ala Cys Asp Gly  
1552

4657 GGC GCC AGT TTC GCA AAC CAG GTC AAG AAA ACA CTC TCT TAA CATTAACTTT 4708  
1553 Gly Ala Ser Phe Ala Asn Gln Val Lys Lys Thr Leu Ser End 1565

4709 TAACTAATCC CTCCCACTCT GTTCGCCAGA GTGGGAGAAG GTTTGTTCATA TCTAAAATCA 4768

4770 ATCTTGGGAT CTTTCTCCAT TTCATTGGAA GGGAAGCTGT AAAACAAATA AGGAATATGA 4828

4829 TATG 4932

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 59:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĹŽKA: 1565 aminokyselín
- (B) TYP: aminokyselina
- (C) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: proteín

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 59 (Tccb peptid):

Charakt. znak	od	do	Opis
1	Met	Leu	Ser Thr Met Glu Lys Gln Leu Asn Glu Ser Gln Arg Asp Ala
16			
17	Leu	Val	Thr Gly Tyr Met Asn Phe Val Ala Pro Thr Leu Lys Gly Val
32			
33	Ser	Gly	Gln Pro Val Thr Val Glu Asp Leu Tyr Glu Tyr Leu Leu Ile
48			
49	Asp	Pro	Glu Val Ala Asp Glu Val Glu Thr Ser Arg Val Ala Gln Ala
64			
65	Ile	Ala	Ser Ile Gln Gln Tyr Met Thr Arg Leu Val Asn Gly Ser-Glu
80			

81 Pro Gly Arg Gln Ala Met Glu Pro Ser Thr Ala Asn Glu Trp Arg Asp  
96

97 Asn Asp Asn Gln Tyr Ala Ile Trp Ala Ala Gly Ala Glu Val Arg Asn  
112

113 Tyr Ala Glu Asn Tyr Ile Ser Pro Ile Thr Arg Gln Glu Lys Ser His  
128

129 Tyr Phe Ser Glu Leu Glu Thr Thr Leu Asn Gln Asn Arg Leu Asp Pro  
144

145 Asp Arg Val Gln Asp Ala Val Leu Ala Tyr Leu Asn Glu Phe Glu Ala  
160

161 Val Ser Asn Leu Tyr Val Leu Ser Gly Tyr Ile Asn Gln Asp Lys Phe  
176

177 Asp Gln Ala Ile Tyr Tyr Phe Ile Gly Arg Thr Thr Thr Lys Pro Tyr  
192

193 Arg Tyr Tyr Trp Arg Gln Met Asp Leu Ser Lys Asn Arg Gln Asp Pro  
208

209 Ala Gly Asn Pro Val Thr Pro Asn Cys Trp Asn Asp Trp Gln Glu Ile  
224

225 Thr Leu Pro Leu Ser Gly Asp Thr Val Leu Glu His Thr Val Arg Pro  
240

241 Val Phe Tyr Asn Asp Arg Leu Tyr Val Ala Trp Val Glu Arg Asp Pro  
256

257 Ala Val Gln Lys Asp Ala Asp Gly Lys Asn Ile Gly Lys Thr His Ala  
272

273 Tyr Asn Ile Lys Phe Gly Tyr Lys Arg Tyr Asp Asp Thr Trp Thr Ala  
288

289 Pro Asn Thr Thr Thr Leu Met Thr Gln Gln Ala Gly Glu Ser Ser Glu  
304

305 Thr Gln Arg Ser Ser Leu Leu Ile Asp Glu Ser Ser Thr Thr Leu Arg  
320

321 Gln Val Asn Leu Leu Ala Thr Thr Asp Phe Ser Ile Asp Pro Thr Glu  
336

337 Glu Thr Asp Ser Asn Pro Tyr Gly Arg Leu Met Leu Gly Val Phe Val  
352

353 Arg Gln Phe Glu Gly Asp Gly Ala Asn Arg Lys Asn Lys Pro Val Val  
368

369 Tyr Gly Tyr Leu Tyr Cys Asp Ser Ala Phe Asn Arg His Val Leu Arg  
384

385 Pro Leu Ser Lys Asn Phe Leu Phe Ser Thr Tyr Arg Asp Glu Thr Asp  
400

401 Gly Gln Asn Ser Leu Gln Phe Ala Val Tyr Asp Lys Lys Tyr Val Ile  
416

417 Thr Lys Val Val Thr Gly Ala Thr Glu Asp Pro Glu Asn Thr Gly Trp  
432

433 Val Ser Lys Val Asp Asp Leu Lys Gln Gly Thr Thr Gly Ala Tyr Val  
448

449 Tyr Ile Asp Gln Asp Gly Leu Thr Leu His Ile Gln Thr Thr Thr Asn  
464

465 Gly Asp Phe Ile Asn Arg His Thr Phe Gly Tyr Asn Asp Leu Val Tyr  
480

481 Asp Ser Lys Ser Gly Tyr Gly Phe Thr Trp Ser Gly Asn Glu Gly Phe  
496

497 Tyr Leu Asp Tyr His Asp Gly Asn Tyr Tyr Thr Phe His Asn Ala Ile  
512

513 Ile Asn Tyr Tyr Pro Ser Gly Tyr Gly Gly Gly Ser Val Pro Asn Gly  
528

529 Thr Trp Ala Leu Glu Gln Arg Ile Asn Glu Gly Trp Ala Ile Ala Pro  
544

545 Leu Leu Asp Thr Leu His Thr Val Thr Val Lys Gly Ser Tyr Ile Ala  
560

561 Trp Glu Gly Glu Thr Pro Thr Gly Tyr Asn Leu Tyr Ile Pro Asp Gly  
576

577 Thr Val Leu Leu Asp Trp Phe Asp Lys Ile Asn Phe Ala Ile Gly Leu  
592

593 Asn Lys Leu Glu Ser Val Phe Thr Ser Pro Asp Trp Pro Thr Leu Thr  
608

609 Thr Ile Lys Asn Phe Ser Lys Ile Ala Asp Asn Arg Lys Phe Tyr Gln  
624

625 Glu Ile Asn Ala Glu Thr Ala Asp Gly Arg Asn Leu Phe Lys Arg Tyr  
640

641 Ser Thr Gln Thr Phe Gly Leu Thr Ser Gly Ala Thr Tyr Ser Thr Thr  
656

657 Tyr Thr Leu Ser Glu Ala Asp Phe Ser Thr Asp Pro Asp Lys Asn Tyr  
672

673 Leu Gln Val Cys Leu Asn Val Val Trp Asp His Tyr Asp Arg Pro Ser  
688

689 Gly Lys Lys Gly Ala Tyr Ser Trp Val Ser Lys Trp Phe Asn Val Tyr  
704

705 Val Ala Leu Gln Asp Ser Lys Ala Pro Asp Ala Ile Pro Arg Leu Val  
720

721 Ser Arg Tyr Asp Ser Lys Arg Gly Leu Val Gln Tyr Leu Asp Phe Trp  
736

737 Thr Ser Ser Leu Pro Ala Lys Thr Arg Leu Asn Thr Thr Phe Val Arg  
752

753 Thr Leu Ile Glu Lys Ala Asn Leu Gly Leu Asp Ser Leu Leu Asp Tyr  
768

769 Thr Leu Gln Ala Asp Pro Ser Leu Glu Ala Asp Leu Val Thr Asp Gly  
784

785 Lys Ser Glu Pro Met Asp Phe Asn Gly Ser Asn Gly Leu Tyr Phe Trp  
800

801 Glu Leu Phe Phe His Leu Pro Phe Leu Val Ala Thr Arg Phe Ala Asn  
816

817 Glu Gln Gln Phe Ser Pro Ala Gln Lys Ser Leu His Tyr Ile Phe Asp  
832

833 Pro Ala Met Lys Asn Lys Pro His Asn Ala Pro Ala Tyr Trp Asn Val  
848

849 Arg Pro Leu Val Glu Gly Asn Ser Asp Leu Ser Arg His Leu Asp Asp  
864

865 Ser Ile Asp Pro Asp Thr Gln Ala Tyr Ala His Pro Val Ile Tyr Gln  
880

881 Lys Ala Val Phe Ile Ala Tyr Val Ser Asn Leu Ile Ala Gln Gly Asp  
896

897 Met Trp Tyr Arg Gln Leu Thr Arg Asp Gly Leu Thr Gln Ala Arg Val  
912

913 Tyr Tyr Asn Leu Ala Ala Glu Leu Leu Gly Pro Arg Pro Asp Val Ser  
928

929 Leu Ser Ser Ile Trp Thr Pro Gln Thr Leu Asp Thr Leu Ala Ala Gly  
944

945 Gln Lys Ala Val Leu Arg Asp Phe Glu His Gln Leu Ala Asn Ser Asp  
960

961 Thr Ala Leu Pro Ala Leu Pro Gly Arg Asn Val Ser Tyr Leu Lys Leu  
976

977 Ala Asp Asn Gly Tyr Phe Asn Glu Pro Leu Asn Val Leu Met Leu Ser  
992

993 His Trp Asp Thr Leu Asp Ala Arg Leu Tyr Asn Leu Arg His Asn Leu  
1008

1009 Thr Val Asp Gly Lys Pro Leu Ser Leu Pro Leu Tyr Ala Ala Pro Val  
1024

1025 Asp Pro Val Ala Leu Leu Ala Gln Arg Ala Gln Ser Gly Thr Leu Thr  
1040

1041 Asn Gly Val Ser Gly Ala Met Leu Thr Val Pro Pro Tyr Arg Phe Ser  
1056

1057 Ala Met Leu Pro Arg Ala Tyr Ser Ala Val Gly Thr Leu Thr Ser Phe  
1072

1073 Gly Gln Asn Leu Leu Ser Leu Leu Glu Arg Ser Glu Arg Ala Cys Gln  
1088

1089 Glu Glu Leu Ala Gln Gln Gln Leu Leu Asp Met Ser Ser Tyr Ala Ile  
1104

1105 Thr Leu Gln Gln Gln Ala Leu Asp Gly Leu Ala Ala Asp Arg Leu Ala  
1120

1121 Leu Leu Ala Ser Gln Ala Thr Ala Gln Gln Arg His Asp His Tyr Tyr  
1136

1137 Thr Leu Tyr Gln Asn Asn Ile Ser Ser Ala Glu Gln Leu Val Met Asp  
1152

1153 Thr Gln Thr Ser Ala Gln Ser Leu Ile Ser Ser Ser Thr Gly Val Gln  
1168

1169 Thr Ala Ser Gly Ala Leu Lys Val Ile Pro Asn Ile Phe Gly Leu Ala  
1184

1185 Asp Gly Gly Ser Arg Tyr Glu Gly Val Thr Glu Ala Ile Ala Ile Gly  
1200  
1201 Leu Met Ala Ala Gly Gln Ala Thr Ser Val Val Ala Glu Arg Leu Ala  
1216  
1217 Thr Thr Glu Asn Tyr Arg Arg Arg Arg Glu Glu Trp Gln Ile Gln Tyr  
1232  
1233 Gln Gln Ala Gln Ser Glu Val Asp Ala Leu Gln Lys Gln Leu Asp Ala  
1248  
1249 Leu Ala Val Arg Glu Lys Ala Ala Gln Thr Ser Leu Gln Gln Ala Lys  
1264  
1265 Ala Gln Gln Val Gln Ile Arg Thr Met Leu Thr Tyr Leu Thr Thr Arg  
1280  
1281 Phe Thr Gln Ala Thr Leu Tyr Gln Trp Leu Ser Gly Gln Leu Ser Ala  
1296  
1297 Leu Tyr Tyr Gln Ala Tyr Asp Ala Val Val Ala Leu Cys Leu Ser Ala  
1312  
1313 Gln Ala Cys Trp Gln Tyr Glu Leu Gly Asp Tyr Ala Thr Thr Phe Ile  
1328  
1329 Gln Thr Gly Thr Trp Asn Asp His Tyr Arg Gly Leu Gln Val Gly Glu  
1344  
1345 Thr Leu Gln Leu Asn Leu His Gln Met Glu Ala Ala Tyr Leu Val Arg  
1360  
1361 His Glu Arg Arg Leu Asn Val Ile Arg Thr Val Ser Leu Lys Ser Leu  
1376  
1377 Leu Gly Asp Asp Gly Phe Gly Lys Leu Lys Thr Glu Gly Lys Val Asp  
1392  
1393 Phe Pro Leu Ser Glu Lys Leu Phe Asp Asn Asp Tyr Pro Gly His Tyr  
1408  
1409 Leu Arg Gln Ile Lys Thr Val Ser Val Thr Leu Pro Thr Leu Val Gly  
1424  
1425 Pro Tyr Gln Asn Val Lys Ala Thr Leu Thr Gln Thr Ser Ser Ser Ile  
1440  
1441 Leu Leu Ala Ala Asp Ile Asn Gly Val Lys Arg Leu Asn Asp Pro Thr  
1456  
1457 Gly Lys Glu Gly Asp Ala Thr His Ile Val Thr Asn Leu Arg Ala Ser  
1472  
1473 Gln Gln Val Ala Leu Ser Ser Gly Ile Asn Asp Ala Gly Ser Phe Glu  
1488  
1489 Leu Arg Leu Glu Asp Glu Arg Tyr Leu Ser Phe Glu Gly Thr Gly Ala  
1504  
1505 Val Ser Lys Trp Thr Leu Asn Phe Pro Arg Ser Val Asp Glu His Ile  
1520  
1521 Asp Asp Lys Thr Leu Lys Ala Asp Glu Met Gln Ala Ala Leu Leu Ala  
1536  
1537 Asn Met Asp Asp Val Leu Val Gln Val His Tyr Thr Ala Cys Asp Gly  
1552  
1553 Gly Ala Ser Phe Ala Asn Gln Val Lys Lys Thr Leu Ser 1565

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 60:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĚLKA: 3132 párov báz
- (B) TYP: nukleová kyselina
- (C) REĤAZEC: dvojitý
- (D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: DNA (genómová)

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 60 (tccc):

1	ATG AGT CCG TCT GAG ACT ACT CTT TAT ACT CAA ACC CCA ACA GTC AGC	48
1	Met Ser Pro Ser Glu Thr Thr Leu Tyr Thr Gln Thr Pro Thr Val Ser	16
49	GTG TTA GAT AAT CGC GGT CTG TCC ATT CGT GAT ATT GGT TTT CAC CGT	96
17	Val Leu Asp Asn Arg Gly Leu Ser Ile Arg Asp Ile Gly Phe His Arg	32
97	ATT GTA ATC GGG GGG GAT ACT GAC ACC CGC GTC ACC CGT CAC CAG TAT	
144		
33	Ile Val Ile Gly Gly Asp Thr Asp Thr Arg Val Thr Arg His Gln Tyr	48
145	GAT GCC CGT GGA CAC CTG AAC TAC AGT ATT GAC CCA CGC TTG TAT GAT	
192		
49	Asp Ala Arg Gly His Leu Asn Tyr Ser Ile Asp Pro Arg Leu Tyr Asp	64
193	GCA AAG CAG GCT GAT AAC TCA GTA AAG CCT AAT TTT GTC TGG CAG CAT	
240		
65	Ala Lys Gln Ala Asp Asn Ser Val Lys Pro Asn Phe Val Trp Gln His	80
241	GAT CTG GCC GGT CAT GCC CTG CGG ACA GAG AGT GTC GAT GCT GGT CGT	
288		
81	Asp Leu Ala Gly His Ala Leu Arg Thr Glu Ser Val Asp Ala Gly Arg	96
289	ACT GTT GCA TTG AAT GAT ATT GAA GGT CGT TCG GTA ATG ACA ATG AAT	
336		
97	Thr Val Ala Leu Asn Asp Ile Glu Gly Arg Ser Val Met Thr Met Asn	
112		
337	GCG ACC GGT GTT CGT CAG ACC CGT CGC TAT GAA GGC AAC ACC TTG CCC	
384		
113	Ala Thr Gly Val Arg Gln Thr Arg Arg Tyr Glu Gly Asn Thr Leu Pro	
128		
385	GGT CGC TTG TTA TCT GTG AGC GAG CAA GTT TTC AAC CAA GAG AGT GCT	
432		
129	Gly Arg Leu Leu Ser Val Ser Glu Gln Val Phe Asn Gln Glu Ser Ala	
144		
433	AAA GTG ACA GAG CGC TTT ATC TGG GCT GGG AAT ACA ACC TCG GAG AAA	
480		

145 Lys Val Thr Glu Arg Phe Ile Trp Ala Gly Asn Thr Thr Ser Glu Lys  
160

481 GAG TAT AAC CTC TCC GGT CTG TGT ATA CGC CAC TAC GAC ACA GCG GGA  
528

161 Glu Tyr Asn Leu Ser Gly Leu Cys Ile Arg His Tyr Asp Thr Ala Gly  
176

529 GTG ACC CGG TTG ATG AGT CAG TCA CTG GCG GGC GCC ATG CTA TCC CAA  
576

177 Val Thr Arg Leu Met Ser Gln Ser Leu Ala Gly Ala Met Leu Ser Gln  
192

577 TCT CAC CAA TTG CTG GCG GAA GGG CAG GAG GCT AAC TGG AGC GGT GAC  
624

193 Ser His Gln Leu Leu Ala Glu Gly Gln Glu Ala Asn Trp Ser Gly Asp  
208

625 GAC GAA ACT GTC TGG CAG GGA ATG CTG GCA AGT GAG GTC TAT ACG ACA  
672

209 Asp Glu Thr Val Trp Gln Gly Met Leu Ala Ser Glu Val Tyr Thr Thr  
224

673 CAA AGT ACC ACT AAT GCC ATC GGG GCT TTA CTG ACC CAA ACC GAT GCG  
720

225 Gln Ser Thr Thr Asn Ala Ile Gly Ala Leu Leu Thr Gln Thr Asp Ala  
240

721 AAA GGC AAT ATT CAG CGT CTG GCT TAT GAC ATT GCC GGT CAG TTA AAA  
768

241 Lys Gly Asn Ile Gln Arg Leu Ala Tyr Asp Ile Ala Gly Gln Leu Lys  
256

769 GGG AGT TGG TTG ACG GTG AAA GGC CAG AGT GAA CAG GTG ATT GTT AAG  
816

257 Gly Ser Trp Leu Thr Val Lys Gly Gln Ser Glu Gln Val Ile Val Lys  
272

817 TCC CTG AGC TGG TCA GCC GCA GGT CAT AAA TTG CGT GAA GAG CAC GGT  
864

273 Ser Leu Ser Trp Ser Ala Ala Gly His Lys Leu Arg Glu Glu His Gly  
288

865 AAC GGC GTG GTT ACG GAG TAC AGT TAT GAG CCG GAA ACT CAA CGT CTG  
912

289 Asn Gly Val Val Thr Glu Tyr Ser Tyr Glu Pro Glu Thr Gln Arg Leu  
304

913 ATA GGT ATC ACC ACC CGG CGT GCC GAA GGG AGT CAA TCA GGA GCC AGA  
960

305 Ile Gly Ile Thr Thr Arg Arg Ala Glu Gly Ser Gln Ser Gly Ala Arg  
320

961 GTA TTG CAG GAT CTA CGC TAT AAG TAT GAT CCG GTG GGG AAT GTT ATC  
1008

321 Val Leu Gln Asp Leu Arg Tyr Lys Tyr Asp Pro Val Gly Asn Val Ile 336

1009 AGT ATC CAT AAT GAT GCC GAA GCT ACC CGC TTT TGG CGT AAT CAG AAA  
1056

337 Ser Ile His Asn Asp Ala Glu Ala Thr Arg Phe Trp Arg Asn Gln Lys 352  
1057 GTG GAG CCG GAG AAT CGC TAT GTT TAT GAT TCT CTG TAT CAG CTT ATG  
1104  
353 Val Glu Pro Glu Asn Arg Tyr Val Tyr Asp Ser Leu Tyr Gln Leu Met 368  
1105 AGT GCG ACA GGG CGT GAA ATG GCT AAT ATC GGT CAG CAA AGC AAC CAA  
1152  
369 Ser Ala Thr Gly Arg Glu Met Ala Asn Ile Gly Gln Gln Ser Asn Gln  
384  
1153 CTT CCC TCA CCC GTT ATA CCT GTT CCT ACT GAC GAC AGC ACT TAT ACC  
1200  
385 Leu Pro Ser Pro Val Ile Pro Val Pro Thr Asp Asp Ser Thr Tyr Thr 400  
1201 AAT TAC CTT CGT ACC TAT ACT TAT GAC CGT GGC GGT AAT TTG GTT CAA  
1248  
401 Asn Tyr Leu Arg Thr Tyr Thr Tyr Asp Arg Gly Gly Asn Leu Val Gln 416  
1249 ATC CGA CAC AGT TCA CCC GCG ACT CAA AAT AGT TAC ACC ACA GAT ATC  
1296  
417 Ile Arg His Ser Ser Pro Ala Thr Gln Asn Ser Tyr Thr Thr Asp Ile 432  
1297 ACC GTT TCA AGC CGC AGT AAC CGG GCG GTA TTG AGT ACA TTA ACG ACA  
1344  
433 Thr Val Ser Ser Arg Ser Asn Arg Ala Val Leu Ser Thr Leu Thr Thr 448  
1345 GAT CCA ACC CGA GTG GAT GCG CTA TTT GAT TCC GGC GGT CAT CAG AAG  
1392  
449 Asp Pro Thr Arg Val Asp Ala Leu Phe Asp Ser Gly Gly His Gln Lys 464  
1393 ATG TTA ATA CCG GGG CAA AAT CTG GAT TGG AAT ATT CGG GGT GAA TTG  
1440  
465 Met Leu Ile Pro Gly Gln Asn Leu Asp Trp Asn Ile Arg Gly Glu Leu 480  
1441 CAA CGA GTC ACA CCG GTG AGC CGT GAA AAT AGC AGT GAC AGT GAA TGG  
1488  
481 Gln Arg Val Thr Pro Val Ser Arg Glu Asn Ser Ser Asp Ser Glu Trp 496  
1489 TAT CGC TAT AGC AGT GAT GGC ATG CGG CTG CTA AAA GTG AGT GAA CAG  
1536  
497 Tyr Arg Tyr Ser Ser Asp Gly Met Arg Leu Leu Lys Val Ser Glu Gln 512  
1537 CAG ACG GGC AAC AGT ACT CAA GTA CAA CGG GTG ACT TAT CTG CCG GGA  
1584  
513 Gln Thr Gly Asn Ser Thr Gln Val Gln Arg Val Thr Tyr Leu Pro Gly 528  
1585 TTA GAG CTA CGG ACA ACT GGG GTT GCA GAT AAA ACA ACC GAA GAT TTG  
1632  
529 Leu Glu Leu Arg Thr Thr Gly Val Ala Asp Lys Thr Thr Glu Asp Leu 544  
1633 CAG GTG ATT ACG GTA GGT GAA GCG GGT CGC GCA CAG GTA AGG GTA TTG  
1680  
545 Gln Val Ile Thr Val Gly Glu Ala Gly Arg Ala Gln Val Arg Val Leu 560  
1681 CAC TGG GAA AGT GGT AAG CCG ACA GAT ATT GAC AAC AAT CAG GTG CGC  
1728



561 His Trp Glu Ser Gly Lys Pro Thr Asp Ile Asp Asn Asn Gln Val Arg 576  
---  
1729 TAC AGC TAC GAT AAT CTG CTT GGC TCC AGC CAG CTT GAA CTG GAT AGC  
1776  
577 Tyr Ser Tyr Asp Asn Leu Leu Gly Ser Ser Gln Leu Glu Leu Asp Ser 592  
  
1777 GAA GGG CAG ATT CTC AGT CAG GAA GAG TAT TAT CCG TAT GGC GGT ACG  
1824  
593 Glu Gly Gln Ile Leu Ser Gln Glu Glu Tyr Tyr Pro Tyr Gly Gly Thr 608  
  
1825 GCG ATA TGG GCG GCG AGA AAT CAG ACA GAA GCC AGC TAC AAA TTT ATT  
1872  
609 Ala Ile Trp Ala Ala Arg Asn Gln Thr Glu Ala Ser Tyr Lys Phe Ile 624  
  
1873 CGT TAC TCC GGT AAA GAG CGG GAT GCC ACT GGA TTG TAT TAT TAC GGC  
1920  
625 Arg Tyr Ser Gly Lys Glu Arg Asp Ala Thr Gly Leu Tyr Tyr Tyr Gly 640  
-  
  
1921 TAC CGT TAT TAT CAA CCT TGG GTG GGT CGA TGG TTG AGT GCT GAT CCG  
1968  
641 Tyr Arg Tyr Tyr Gln Pro Trp Val Gly Arg Trp Leu Ser Ala Asp Pro 656  
  
1969 GCG GGA ACC GTG GAT GGG CTG AAT TTG TAC CGA ATG GTG AGG AAT AAC  
2016  
657 Ala Gly Thr Val Asp Gly Leu Asn Leu Tyr Arg Met Val Arg Asn Asn 672  
  
2017 CCC ATC ACA TTG ACT GAC CAT GAC GGA TTA GCA CCG TCT CCA AAT AGA  
2064  
673 Pro Ile Thr Leu Thr Asp His Asp Gly Leu Ala Pro Ser Pro Asn Arg 688  
  
2065 AAT CGA AAT ACA TTT TGG TTT GCT TCA TTT TTG TTT CGT AAA CCT GAT  
2112  
689 Asn Arg Asn Thr Phe Trp Phe Ala Ser Phe Leu Phe Arg Lys Pro Asp 704  
  
2113 GAG GGA ATG TCC GCG TCA ATG AGA CGG GGA CAA AAA ATT GGC AGA GCC  
2160  
705 Glu Gly Met Ser Ala Ser Met Arg Arg Gly Gln Lys Ile Gly Arg Ala 720  
  
2161 ATT GCC GGC GGG ATT GCG ATT GGC GGT CTT GCG GCT ACC ATT GCC GCT  
2208  
721 Ile Ala Gly Gly Ile Ala Ile Gly Gly Leu Ala Ala Thr Ile Ala Ala 736  
  
2209 ACG GCT GGC GCG GCT ATC CCC GTC ATT CTG GGG GTT GCG GCC GTA GGC  
2256  
737 Thr Ala Gly Ala Ala Ile Pro Val Ile Leu Gly Val Ala Ala Val Gly 752  
  
2257 GCG GGG ATT GGC GCG TTG ATG GGA TAT AAC GTC GGT AGC CTG CTG GAA  
2304  
753 Ala Gly Ile Gly Ala Leu Met Gly Tyr Asn Val Gly Ser Leu Leu Glu 768  
  
2305 AAA GGC GGG GCA TTA CTT GCT CGA CTC GTA CAG GGG AAA TCG ACG TTA  
2352  
769 Lys Gly Gly Ala Leu Leu Ala Arg Leu Val Gln Gly Lys Ser Thr Leu 784  
  
2353 GTA CAG TCG GCG GCT GGC GCG GCT GCC GGA GCG AGT TCA GCC GCG GCT  
2400  
785 Val Gln Ser Ala Ala Gly Ala Ala Ala Gly Ala Ser Ser Ala Ala Ala 800

2401 TAT GGC GCA CGG GCA CAA GGT GTC GGT GTT GCA TCA GCC GCC GGG GCG  
2448  
801 Tyr Gly Ala Arg Ala Gln Gly Val Gly Val Ala Ser Ala Ala Gly Ala 816

2449 GTA ACA GGG GCT GTG GGA TCA TGG ATA AAT AAT GCT GAT CGG GGG ATT  
2496  
817 Val Thr Gly Ala Val Gly Ser Trp Ile Asn Asn Ala Asp Arg Gly Ile 832

2497 GGC GGC GCT ATT GGG GCC GGG AGT GCG GTA GGC ACC ATT GAT ACT ATG  
2544  
833 Gly Gly Ala Ile Gly Ala Gly Ser Ala Val Gly Thr Ile Asp Thr Met 848

2545 TTA GGG ACT GCC TCT ACC CTT ACC CAT GAA GTC GGG GCA GCG GCG GGT  
2592  
849 Leu Gly Thr Ala Ser Thr Leu Thr His Glu Val Gly Ala Ala Ala Gly 864

2593 GGG GCG GCG GGT GGG ATG ATC ACC GGT ACG CAA GGG AGT ACT CGG GCA  
2640  
865 Gly Ala Ala Gly Gly Met Ile Thr Gly Thr Gln Gly Ser Thr Arg Ala 880

2641 GGT ATC CAT GCC GGT ATT GGC ACC TAT TAT GGC TCC TGG ATT GGT TTT  
2688  
881 Gly Ile His Ala Gly Ile Gly Thr Tyr Tyr Gly Ser Trp Ile Gly Phe 896

2689 GGT TTA GAT GTC GCT AGT AAC CCC GCC GGA CAT TTA GCG AAT TAC GCA  
2736  
897 Gly Leu Asp Val Ala Ser Asn Pro Ala Gly His Leu Ala Asn Tyr Ala 912

2737 GTG GGT TAT GCC GCT GGT TTG GGT GCT GAA ATG GCT GTC AAC AGA ATA  
2784  
913 Val Gly Tyr Ala Ala Gly Leu Gly Ala Glu Met Ala Val Asn Arg Ile 928

2785- ATG GGT GGT GGA TTT TTG AGT AGG CTC TTA GGC CCG GTT GTC AGC CCA  
2832  
929 Met Gly Gly Gly Phe Leu Ser Arg Leu Leu Gly Arg Val Val Ser Pro 944

2833 TAT GCC GCC GGT TTA GCC AGA CAA TTA GTA CAT TTC AGT GTC GCC AGA  
2880  
945 Tyr Ala Ala Gly Leu Ala Arg Gln Leu Val His Phe Ser Val Ala Arg 960

2881 CCT GTC TTT GAG CCG ATA TTT AGT GTT CTC GGC GGG CTT GTC GGT GGT  
2928  
961 Pro Val Phe Glu Pro Ile Phe Ser Val Leu Gly Gly Leu Val Gly Gly 976

2929 ATT GGA ACT GGC CTG CAC AGA GTG ATG GGA AGA GAG AGT TGG ATT TCC  
2976  
977 Ile Gly Thr Gly Leu His Arg Val Met Gly Arg Glu Ser Trp Ile Ser 992

2977 AGA GCG TTA AGT GCT GCC GGT AGT GGT ATA GAT CAT GTC GCT GGC ATG  
3024  
993 Arg Ala Leu Ser Ala Ala Gly Ser Gly Ile Asp His Val Ala Gly Met  
1008

3025 ATT GGT AAT CAG ATC AGA GGC AGG GTC TTG ACC ACA ACC GGG ATC GCT  
3072

1009 Ile Gly Asn Gln Ile Arg Gly Arg Val Leu Thr Thr Thr Gly Ile Ala  
1024

3073 AAT GCG ATA GAC TAT GGC ACC AGT GCT GTG GGA GCC GCA CGA CGA GTT  
3120  
1025 Asn Ala Ile Asp Tyr Gly Thr Ser Ala Val Gly Ala Ala Arg Arg Val  
1040

3121 TTT TCT TTG TAA 3132  
1041 Phe Ser Leu End 1043

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 61:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĹŽKA: 1043 aminokyselín
- (B) TYP: aminokyselina
- (C) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: proteín

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 61 (Tccc peptid):

1	Met Ser Pro Ser Glu Thr Thr Leu Tyr Thr Gln Thr Pro Thr Val Ser	16
17	Val Leu Asp Asn Arg Gly Leu Ser Ile Arg Asp Ile Gly Phe His Arg	32
33	Ile Val Ile Gly Gly Asp Thr Asp Thr Arg Val Thr Arg His Gln Tyr	48
49	Asp Ala Arg Gly His Leu Asn Tyr Ser Ile Asp Pro Arg Leu Tyr Asp	64
65	Ala Lys Gln Ala Asp Asn Ser Val Lys Pro Asn Phe Val Trp Gln His	80
81	Asp Leu Ala Gly His Ala Leu Arg Thr Glu Ser Val Asp Ala Gly Arg	96
97	Thr Val Ala Leu Asn Asp Ile Glu Gly Arg Ser Val Met Thr Met Asn	112
113	Ala Thr Gly Val Arg Gln Thr Arg Arg Tyr Glu Gly Asn Thr Leu Pro	128
129	Gly Arg Leu Leu Ser Val Ser Glu Gln Val Phe Asn Gln Glu Ser Ala	144
145	Lys Val Thr Glu Arg Phe Ile Trp Ala Gly Asn Thr Thr Ser Glu Lys	160
161	Glu Tyr Asn Leu Ser Gly Leu Cys Ile Arg His Tyr Asp Thr Ala Gly	176
177	Val Thr Arg Leu Met Ser Gln Ser Leu Ala Gly Ala Met Leu Ser Gln	192
193	Ser His Gln Leu Leu Ala Glu Gly Gln Glu Ala Asn Trp Ser Gly Asp	208
209	Asp Glu Thr Val Trp Gln Gly Met Leu Ala Ser Glu Val Tyr Thr Thr	224
225	Gln Ser Thr Thr Asn Ala Ile Gly Ala Leu Leu Thr Gln Thr Asp Ala	240
241	Lys Gly Asn Ile Gln Arg Leu Ala Tyr Asp Ile Ala Gly Gln Leu Lys	256

257 Gly Ser Trp Leu Thr Val Lys Gly Gln Ser Glu Gln Val Ile Val Lys 272  
273 Ser Leu Ser Trp Ser Ala Ala Gly His Lys Leu Arg Glu Glu His Gly 288  
289 Asn Gly Val Val Thr Glu Tyr Ser Tyr Glu Pro Glu Thr Gln Arg Leu 304  
305 Ile Gly Ile Thr Thr Arg Arg Ala Glu Gly Ser Gln Ser Gly Ala Arg 320  
321 Val Leu Gln Asp Leu Arg Tyr Lys Tyr Asp Pro Val Gly Asn Val Ile 336  
337 Ser Ile His Asn Asp Ala Glu Ala Thr Arg Phe Trp Arg Asn Gln Lys 352

353 Val Glu Pro Glu Asn Arg Tyr Val Tyr Asp Ser Leu Tyr Gln Leu Met 368  
369 Ser Ala Thr Gly Arg Glu Met Ala Asn Ile Gly Gln Gln Ser Asn Gln 384  
385 Leu Pro Ser Pro Val Ile Pro Val Pro Thr Asp Asp Ser Thr Tyr Thr 400  
401 Asn Tyr Leu Arg Thr Tyr Thr Tyr Asp Arg Gly Gly Asn Leu Val Gln 416  
417 Ile Arg His Ser Ser Pro Ala Thr Gln Asn Ser Tyr Thr Thr Asp Ile 432  
433 Thr Val Ser Ser Arg Ser Asn Arg Ala Val Leu Ser Thr Leu Thr Thr 448  
449 Asp Pro Thr Arg Val Asp Ala Leu Phe Asp Ser Gly Gly His Gln Lys 464  
465 Met Leu Ile Pro Gly Gln Asn Leu Asp Trp Asn Ile Arg Gly Glu Leu 480  
481 Gln Arg Val Thr Pro Val Ser Arg Glu Asn Ser Ser Asp Ser Glu Trp 496  
497 Tyr Arg Tyr Ser Ser Asp Gly Met Arg Leu Leu Lys Val Ser Glu Gln 512  
513 Gln Thr Gly Asn Ser Thr Gln Val Gln Arg Val Thr Tyr Leu Pro Gly 528  
529 Leu Glu Leu Arg Thr Thr Gly Val Ala Asp Lys Thr Thr Glu Asp Leu 544  
545 Gln Val Ile Thr Val Gly Glu Ala Gly Arg Ala Gln Val Arg Val Leu 560  
561 His Trp Glu Ser Gly Lys Pro Thr Asp Ile Asp Asn Asn Gln Val Arg 576  
577 Tyr Ser Tyr Asp Asn Leu Leu Gly Ser Ser Gln Leu Glu Leu Asp Ser 592  
593 Glu Gly Gln Ile Leu Ser Gln Glu Glu Tyr Tyr Pro Tyr Gly Gly Thr 608  
609 Ala Ile Trp Ala Ala Arg Asn Gln Thr Glu Ala Ser Tyr Lys Phe Ile 624  
625 Arg Tyr Ser Gly Lys Glu Arg Asp Ala Thr Gly Leu Tyr Tyr Tyr Gly 640  
641 Tyr Arg Tyr Tyr Gln Pro Trp Val Gly Arg Trp Leu Ser Ala Asp Pro 656  
657 Ala Gly Thr Val Asp Gly Leu Asn Leu Tyr Arg Met Val Arg Asn Asn 672  
673 Pro Ile Thr Leu Thr Asp His Asp Gly Leu Ala Pro Ser Pro Asn Arg 688  
689 Asn Arg Asn Thr Phe Trp Phe Ala Ser Phe Leu Phe Arg Lys Pro Asp 704  
705 Glu Gly Met Ser Ala Ser Met Arg Arg Gly Gln Lys Ile Gly Arg Ala 720  
721 Ile Ala Gly Gly Ile Ala Ile Gly Gly Leu Ala Ala Thr Ile Ala Ala 736  
737 ~~Thr Ala~~ Gly Ala Ala Ile Pro Val Ile Leu Gly Val Ala Ala Val Gly 752  
753 Ala Gly Ile Gly Ala Leu Met Gly Tyr Asn Val Gly Ser Leu Leu Glu 768

769 Lys Gly Gly Ala Leu Leu Ala Arg Leu Val Gln Gly Lys Ser Thr Leu 784  
785 Val Gln Ser Ala Ala Gly Ala Ala Ala Gly Ala Ser Ser Ala Ala Ala 800  
801 Tyr Gly Ala Arg Ala Gln Gly Val Gly Val Ala Ser Ala Ala Gly Ala 816  
817 Val Thr Gly Ala Val Gly Ser Trp Ile Asn Asn Ala Asp Arg Gly Ile 832  
833 Gly Gly Ala Ile Gly Ala Gly Ser Ala Val Gly Thr Ile Asp Thr Met 848  
849 Leu Gly Thr Ala Ser Thr Leu Thr His Glu Val Gly Ala Ala Ala Gly 864  
865 Gly Ala Ala Gly Gly Met Ile Thr Gly Thr Gln Gly Ser Thr Arg Ala 880  
881 Gly Ile His Ala Gly Ile Gly Thr Tyr Tyr Gly Ser Trp Ile Gly Phe 896  
897 Gly Leu Asp Val Ala Ser Asn Pro Ala Gly His Leu Ala Asn Tyr Ala 912  
913 Val Gly Tyr Ala Ala Gly Leu Gly Ala Glu Met Ala Val Asn Arg Ile 928  
  
929 Met Gly Gly Gly Phe Leu Ser Arg Leu Leu Gly Arg Val Val Ser Pro 944  
945 Tyr Ala Ala Gly Leu Ala Arg Gln Leu Val His Phe Ser Val Ala Arg 960  
961 Pro Val Phe Glu Pro Ile Phe Ser Val Leu Gly Gly Leu Val Gly Gly 976  
977 Ile Gly Thr Gly Leu His Arg Val Met Gly Arg Glu Ser Trp Ile Ser 992  
993 Arg Ala Leu Ser Ala Ala Gly Ser Gly Ile Asp His Val Ala Gly Met 1008  
1009 Ile Gly Asn Gln Ile Arg Gly Arg Val Leu Thr Thr Thr Gly Ile Ala 1024  
1025 Asn Ala Ile Asp Tyr Gly Thr Ser Ala Val Gly Ala Ala Arg Arg Val 1040  
1041 Phe Ser Leu 1043

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 62: TcaA<sub>iv</sub>

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĚŽKA: 5 aminokyselín
- (B) TYP: aminokyselina
- (C) REĤAZEC: jednoduchý
- (D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: proteín

(v) TYP FRAGMENTU: interný

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 62: TcaA<sub>iv</sub>

Asn Ile Gly Gly Asp

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 63: TcaA<sub>ii</sub> - syn

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĚŽKA: 20 aminokyselín
- (B) TYP: aminokyselina
- (C) REĚAZEC: jednoduchý
- (D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: proteín

(v) TYP FRAGMENTU: interný

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 63: TcaA<sub>ii</sub> - syn

Cys	Leu	Arg	Gly	Asn	Ser	Pro	Thr	Asn	Pro	Asp	Lys	Asp	Gly	Ile
1				5					10					15

Phe	Ala	Gln	Val	Ala
				20

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 64: TcaA<sub>iii</sub> - syn

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĚŽKA: 20 aminokyselín
- (B) TYP: aminokyselina
- (C) REĚAZEC: jednoduchý
- (D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: proteín

(v) TYP FRAGMENTU: interný

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 64: TcaA<sub>iii</sub> - syn

Cys	Tyr	Thr	Pro	Asp	Gln	Thr	Pro	Ser	Phe	Tyr	Glu	Thr	Ala	Phe
1				5					10					15

Arg Ser Ala Asp Gly

20

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 65: TcaB<sub>i</sub> - syn

(i) CHARAKTERISTIKA SEKVENCIE:

(A) DĹŽKA: 19 aminokyselín

(B) TYP: aminokyselina

(C) REĹAZEC: jednoduchý

(D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: proteín

(v) TYP FRAGMENTU: interný

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 65: TcaB<sub>i</sub> - syn

His	Gly	Gln	Ser	Tyr	Asn	Asp	Asn	Asn	Tyr	Cys	Asn	Phe	Thr	Leu
1			5					10						15

Ser Ile Asn Thr  
19

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 66: TcaB<sub>ii</sub> - syn

(i) CHARAKTERISTIKA SEKVENCIE:

(A) DĹŽKA: 20 aminokyselín

(B) TYP: aminokyselina

(C) REĹAZEC: jednoduchý

(D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: proteín

(v) TYP FRAGMENTU: interný

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 66: TcaB<sub>ii</sub> - syn





(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 68: TcbA<sub>ii</sub> - syn

Cys	Tyr	Asn	Glu	Asn	Pro	Ser	Ser	Glu	Asp	Lys	Lys	Trp	Tyr	Phe
1				5					10					15
Ser Ser Lys Asp Asp														
20														

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 69: TcbA<sub>iii</sub> - syn

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĹŽKA: 20 aminokyselín
- (B) TYP: aminokyselina
- (C) REĹAZEC: jednoduchý
- (D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: proteín

(v) TYP FRAGMENTU: interný

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 69: TcbA<sub>iii</sub> - syn

Cys	Phe	Asp	Ser	Tyr	Ser	Gln	Leu	Tyr	Glu	Glu	Asn	Ile	Asn	Ala
1				5					10					15
Gly Glu Gln Arg Ala														
20														

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 70: TcdA<sub>i</sub> - syn

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĹŽKA: 22 aminokyselín
- (B) TYP: aminokyselina
- (C) REĹAZEC: jednoduchý
- (D) TOPOLOGIA: lineárna

(ii) TYP MOLEKULY: proteín









(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 79:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĚLKA: 15 aminokyselín
- (B) TYP: aminokyselina
- (C) REŤAZEC: jednoduchý
- (D) TOPOLÓGIA: lineárna

(ii) TYP MOLEKULY: proteín

(v) TYP FRAGMENTU: N-koncový

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 79: 109 kDa - Hm

Met Leu Asp Ile Met Glu Lys Gln Leu Asn Glu Ser Glu Arg Asp  
1                                    5                                    10                                    15

(2) INFORMÁCIA O SEKVENCII SEKV. ID. ČÍS. 80:

(i) CHARAKTERISTIKA SEKVENCIE:

- (A) DĚLKA: 8 aminokyselín
- (B) TYP: aminokyselina
- (C) REŤAZEC: jednoduchý
- (D) TOPOLÓGIA: lineárna

(ii) TYP MOLEKULY: proteín

(v) TYP FRAGMENTU: N-koncový

(xi) OPIS SEKVENCIE: SEKV. ID. ČÍS. 80: 170 kDa - WX - 1

Met Gln Asp Ser Arg Glu Val Ser  
1                                    5











## P A T E N T O V É   N Á R O K Y

1. V podstate čistá kultúra mikroorganizmov kmeňa *Photorhabdus* označená B2, DEP1, DEP2, DEP3, P. zealandrica, P. hepialus, HB-Arg, HB Oswego, HB Oswego, HB Lewiston, K-122, HMGD, Indicus, GD, PWH-5, Megidis, HF-85, A. Cows, MP1, MP2, MP3, MP4, MP5, GL98, GL101, GL138, GL55, GL217 alebo GL257.
2. Kompozícia obsahujúca účinné množstvo *Photorhabdus* proteínového toxínu, ktorý má funkčnú aktivitu proti hmyzu, vyznačujúca sa tým, že toxín je produkovaný purifikovanou kultúrou kmeňa *Photorhabdus* označenou B2, DEP1, DEP2, DEP3, P. zealandrica, P. hepialus, HB-Arg, HB Oswego, HB Oswego, HB Lewiston, K-122, HMGD, Indicus, GD, PWH-5, Megidis, HF-85, A. Cows, MP1, MP2, MP3, MP4, MP5, GL98, GL101, GL138, GL55, GL217 alebo GL257.
3. Purifikovaný prípravok obsahujúci proteín s N-koncovou aminokyselinovou sekvenciou vybranou zo skupiny pozostávajúcej zo SEKV. ID. ČÍS. 62, SEKV. ID. ČÍS. 72, SEKV. ID. ČÍS. 73, SEKV. ID. ČÍS. 74, SEKV. ID. ČÍS. 75, SEKV. ID. ČÍS. 76, SEKV. ID. ČÍS. 77, SEKV. ID. ČÍS. 78, SEKV. ID. ČÍS. 79, SEKV. ID. ČÍS. 80, SEKV. ID. ČÍS. 81, SEKV. ID. ČÍS. 82, SEKV. ID. ČÍS. 83, SEKV. ID. ČÍS. 84, SEKV. ID. ČÍS. 85, SEKV. ID. ČÍS. 86, SEKV. ID. ČÍS. 87 a SEKV. ID. ČÍS. 88.
4. Chimerický DNA konštrukt, ktorý je adaptovaný na expresiu v prokaryotickom alebo eukaryotickom hostiteľovi, a ktorý obsahuje 5' až 3' transkripčný promótor aktívny v hostiteľovi, DNA sekvenciu kódujúcu proteín *Photorhabdus*, ktorý má funkčnú aktivitu proti hmyzu, a transkripčný terminátor, vyznačujúci sa tým, že proteín kódovaný DNA sekvenciou má N-koncovú sekvenciu vybranú zo skupiny pozostávajúcej z SEKV. ID. ČÍS. 62, SEKV. ID. ČÍS. 72, SEKV. ID. ČÍS. 73, SEKV. ID. ČÍS. 74, SEKV. ID. ČÍS. 75, SEKV. ID. ČÍS. 76, SEKV. ID. ČÍS. 77, SEKV.

ID. Čís. 78, SEKV. ID. Čís. 79, SEKV. ID. Čís. 80, SEKV. ID. Čís. 81, SEKV. ID. Čís. 82, SEKV. ID. Čís. 83, SEKV. ID. Čís. 84, SEKV. ID. Čís. 85, SEKV. ID. Čís. 86, SEKV. ID. Čís. 87 a SEKV. ID. Čís. 88.

5. Spôsob kontroly hmyzu, ktorý zahŕňa orálne podanie hmyzu účinného množstva proteínového toxínu s funkčnou aktivitou proti hmyzu, vyznačujúci sa tým, že proteín je produkovaný purifikovanou kultúrou kmeňa *Photobacterium* označenou B2, DEP1, DEP2, DEP3, *P. zealandrica*, *P. hepialus*, HB-Arg, HB Oswego, HB Oswego, HB Lewiston, K-122, HMGD, Indicus, GD, PWH-5, Megidis, HF-85, A. Cows, MP1, MP2, MP3, MP4, MP5, GL98, GL101, GL138, GL55, GL217 alebo GL257.
6. Spôsob podľa nároku 5, vyznačujúci sa tým, že hmyz je rádu *Lepidoptera*, *Coleoptera*, *Hymenoptera*, *Diptera*, *Dictyoptera*, *Acarina* alebo *Homoptera*.
7. Spôsob podľa nároku 6, vyznačujúci sa tým, že druh hmyzu je z rádu *Coleoptera* a je to liskavka uhorková, pásavka zemiaková, múčny červ, *Anthonomus grandis*, turf grub alebo druh hmyzu je z rádu *Lepidoptera* a je to sivkavec stavikrvový, siatica ypsilonová, mora kapustová, obaľovač jablčný, *Helicoverpa Zea*, vijačka kukuričná, *Manduca sexta* a sivkavec.
8. DNA alebo RNA oligonukleotid kódujúci aminokyselinovú sekvenciu vybranú zo SEKV. ID. Čís. 62, SEKV. ID. Čís. 72, SEKV. ID. Čís. 73, SEKV. ID. Čís. 74, SEKV. ID. Čís. 75, SEKV. ID. Čís. 76, SEKV. ID. Čís. 77, SEKV. ID. Čís. 78, SEKV. ID. Čís. 79, SEKV. ID. Čís. 80, SEKV. ID. Čís. 81, SEKV. ID. Čís. 82, SEKV. ID. Čís. 83, SEKV. ID. Čís. 84, SEKV. ID. Čís. 85, SEKV. ID. Čís. 86, SEKV. ID. Čís. 87 a SEKV. ID. Čís. 88.
9. Použitie oligonukleotidu podľa nároku 8 ako sondy na izolovanie genetického materiálu z organizmov *Enterobacteraceae*, *Photobacterium* alebo *Photobacterium*

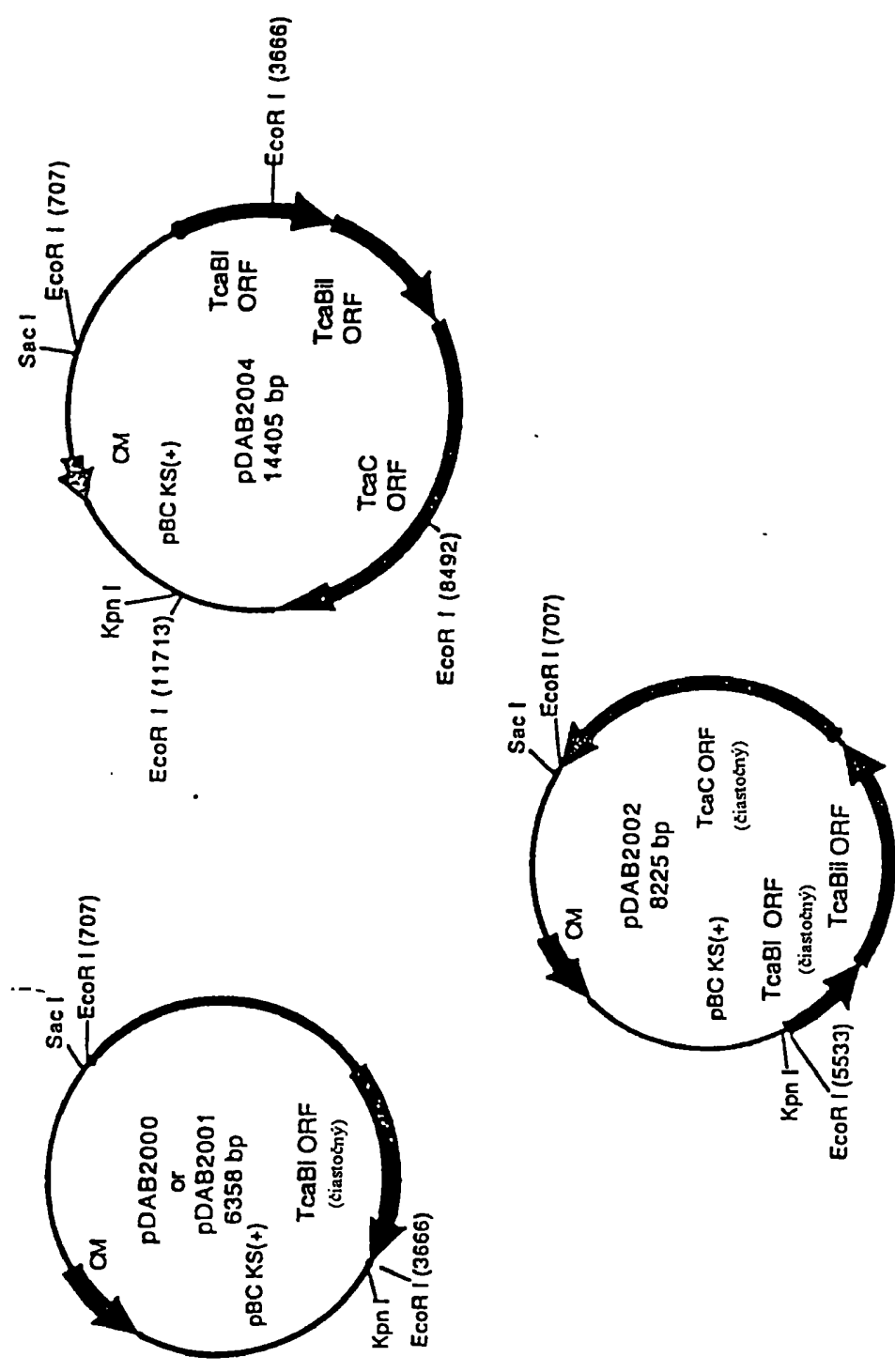
*luminescence.*

10. Spôsob expresie proteínu produkovaného purifikovanou bakteriálnou kultúrou rodu *Photobacterium* v prokaryotickom alebo eukaryotickom hostiteľovi v účinnom množstve tak, že proteín má funkčnú aktivitu proti hmyzu, vyznačujúci sa tým, že spôsob zahŕňa konštrukciu chimerických DNA konštruktov s 5' a 3' promótormi, DNA sekvenčný kódujúci proteín, transkripčný terminátor a potom transferovanie chimerického DNA konštruktú do hostiteľa, pričom proteínová kódujúca DNA sekvencia má N-koncovú aminokyselinovú sekvenciu vybranú zo skupiny pozostávajúcej zo SEKV. ID. Čís. 62, SEKV. ID. Čís. 72, SEKV. ID. Čís. 73, SEKV. ID. Čís. 74, SEKV. ID. Čís. 75, SEKV. ID. Čís. 76, SEKV. ID. Čís. 77, SEKV. ID. Čís. 78, SEKV. ID. Čís. 79, SEKV. ID. Čís. 80, SEKV. ID. Čís. 81, SEKV. ID. Čís. 82, SEKV. ID. Čís. 83, SEKV. ID. Čís. 84, SEKV. ID. Čís. 85, SEKV. ID. Čís. 86, SEKV. ID. Čís. 87 a SEKV. ID. Čís. 88.
  
11. Spôsob prípravy protilátky proti fragmentu proteínu, ktorý je časťou proteínu s funkčnou aktivitou, kde proteín je produkovaný baktériou z čelade *Enterobacteraceae*, vyznačujúci sa tým, že spôsob zahŕňa:
  - a) izoláciu fragmentu proteínu, kde proteínový fragment tvorí najmenej šesť aminokyselín,
  - b) imunizáciu cicavčieho druhu s proteínovým fragmentom,
  - c) zozbieranie séra s obsahom protilátky alebo protilátky zo sleziny cicavčieho druhu, kde zozbieraná protilátka je protilátka k proteínovému fragmentu s funkčnou aktivitou.
  
12. Spôsob podľa nároku 11, vyznačujúci sa tým, že proteínový fragment je vybraný zo skupiny pozostávajúcej z SEKV. ID. Čís. 63, SEKV. ID. Čís. 64, SEKV. ID. Čís. 65, SEKV. ID. Čís.

66, SEKV. ID. ČÍS. 67, SEKV. ID. ČÍS. 68, SEKV. ID. ČÍS. 69,  
SEKV. ID. ČÍS. 70 a SEKV. ID. ČÍS. 71.

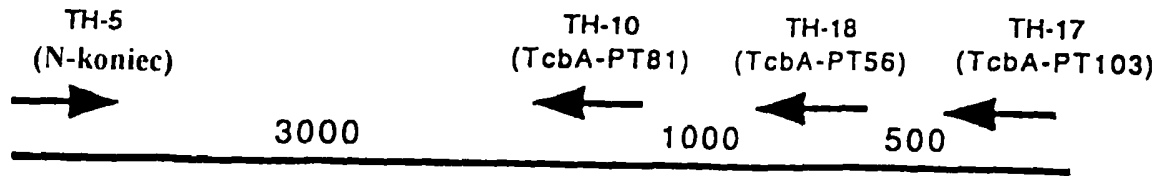
1	ATG CAG GAT TGT COG GAA GTA TCG ATT ACA ACG CTG TCA CTT COC AAA GGT GGC GGT	GGT GGC GGT
	TAC GTC CTA ACA GGC CTT CAT ACG TAA TGT TCC GAC AGT GAA GGG TTT OCA COG OCA	CCA COG OCA
1	Met Gln Asp Cys Pro Glu Val Ser Ile Thr Thr Leu Ser Leu Pro Lys Gly Gly Gly	Gly
	P2Psh	
58	GCT ATC AAT GGC ATG GGA GAA GCA CTG AAT GCT GGC GGC CCT GAT GGA ATG GGC TOC	TOC
	CGA TAG TTA CCG TAC OCT CTT CGT GAC TTA CGA CCG GGA CTA OCT TAC CCG AGG	AGG
20	Ala Ile Asn Gly Met Gly Glu Ala Leu Asn Ala Ala Gly Pro Asp Gly Met Ala Ser	Met Ala Ser
115	CTA TCT CTG CCA TTA CCG CTT TOG ACC GGC AGA GGG ACG GCT CCT GGA TTA TOG CTG	CTG
	GAT AGA GAC GGT AAT GGG GAA AGC TGG CCG TCT OOC TOC CGA GGA OCT AAT AGC GAC	GAC
39	Leu Ser Leu Pro Leu Pro Leu Ser Thr Gly Arg Gly Thr Ala Pro Gly Leu Ser Leu	Leu Ser Leu
172	ATT TAC AGC AAC AGT GCA GGT AAT GGG OCT TTC GGC ATC GGC TGG CAA TGC GGT GTT	GTT
	TAA ATG TCG TTG TCA OGT CCA TTA CCG GGA AAG CCG TAG CCG ACC GTT ACG CCA CAA	CAA
58	Ile Tyr Ser Asn Ser Ala Gly Asn Gly Pro Phe Gly Ile Gly Trp Gln Cys Gly Val	Val
229	ATG TOC ATT AGC CGA CCG ACC CAA CAT GGC CTT CAA CAT TGA CGA GGT	GGT
	TAC AGG TAA TCG GCT CCG TGG GTT GTA CCG GAA GTT GTA ACT GCT GCA	GCA
77	Met Ser Ile Ser Arg Arg Thr Gln His Gly Leu Gln His ... Arg Arg	Arg
	P2.3.5R	

Obr. 1



Obr. 2 Plazmidy použité v sekvencovaní *tca* lokusu. CM = gén rezistentný na chloramfenikol. ORF = otvorený čítací rámec.





Obr. 3 Fyzická mapa DNA fragmentov *tcb* lokusu. Stanovená vzdialenosť medzi fragmentami je daná v nukleotidoch.

Tcba	1740	1750	1760	1770	1780
	SSAQALKNDS	EPMDFSGANA	LYFWELFYVT	PMMMAHRLLO	EQNFDAAANHW
		n			
		450	460	470	480
TcaBi	gS	nPvDFSGpyg	iYlWEiFfhi	PflvtvRmqt	EQryedAdtW>
	^^	^^^^^^^^v^	^^^^^^^^^^	^-----v^v	^^-----v^v

Tcba	1790	1800	1810	1820	1830
	FRYVWSPSGY	IVDGKIAIYH	WNVRLPEEDT	SWNAQQLDST	DPDAVAQDDP
	rdangql				
	490	510	520	530	
TcaBi	ykYifrsaGY	ImDGskprY-	WNVmPLqLDT	awdttQpatT	DPDviAmaDP>
	^^^^-^^^^	^^^^-v^v^	^^^^-v^v^	^^^^v^v-^^	^^-^^v-^^

Tcba	1840	1850	1860	1870	1880
	MHYKVATFMA	TLDLLMARGD	AAYRQLERDT	LAEAKMWTQ	ALNLLGDEPQ
	540	550	560	570	580
TcaBi	MHYKlAiFlh	TLDLLiARGD	sAYRQLERDT	LvEAKMyYiQ	AqgLLGprPd>
	^^^^^^-v^v	^^^^^^^^^^	^^^^^^^^^^	^-----v^	^v^^^^v^v^

Tcba	1890	1900	1910	1920	1930
	VMLSTTWANP	TLGNAASKTT	QQVRQOVLTO	LRLNSRVKTP	LLGTANSLTA
		600			
TcaBi	ihctnTWpNP	TLsk>			
	^v^-----	^^^^			

Tcba	1940	1950	1960	1970	1980
	LFLPQENSKL	KGWRTLAQR	MFNLRHNLSI	DGQPLSLPLY	AKPADPKALL
	20	30	40	50	60
TcaBii	_FLPpyNdvL	lGVWdkLeLR	lyNLRHNLSI	DGQPLnLPLY	AtPvDPKtLq>
	^^-v^v-^v	v^v^v-^v	^^^^^^^^^^	^^^^^^^^^^	^-----v^v

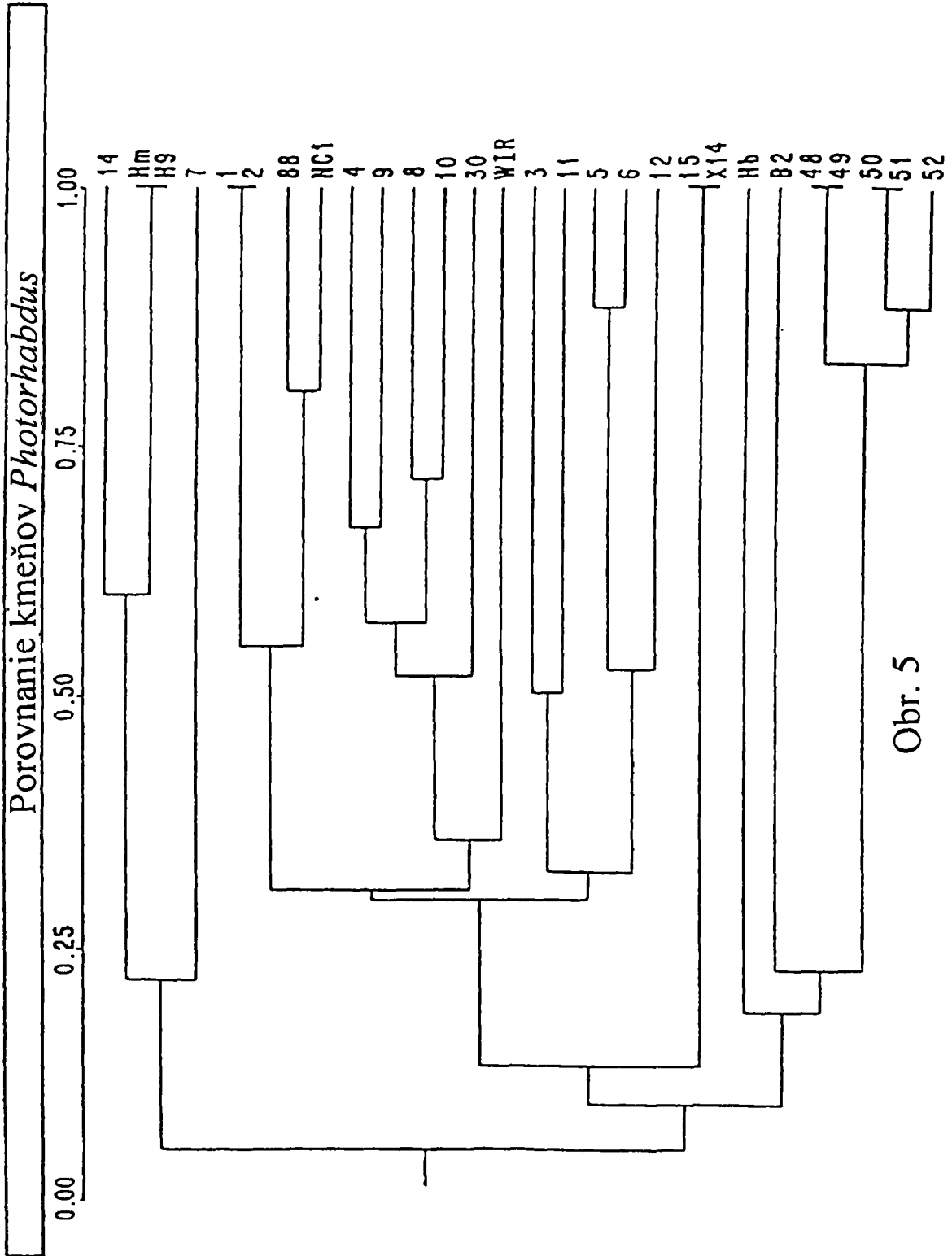
Tcba	1990	2000	2010	2020	2030
	SAAVSASQGG	ADLPKAPLTI	HRFPQMLEGA	RGLVNQLIQF	GSSLLGYSER
		gw			
	70	80	90	100	110
TcaBii	rqqaggdgtG	sspaggqgsV	qRyPllvErA	RsaVslLtQF	GnSLqtTlEh>
	-----v-^	^v^v-^v	^^^^v^v^v	^^v^v-^^	^^^^v-vv^

Tcba	2040	2050	2060	2070	2080
	QDAEAMSOLL	QTQASELIIT	SIRMQDNOLA	ELDSEKTALQ	VSLAGVQORF
	120	130	140	150	160
TcaBii	QDnEkMtill	QTQqeailkh	qhdiQqNnLk	gLqhsLTALQ	aSrdGdtlRq>
	^^-v^v^v	^^-----v^v	v^v^v^v^v	^^v-v^v^	^v-^v^v^v

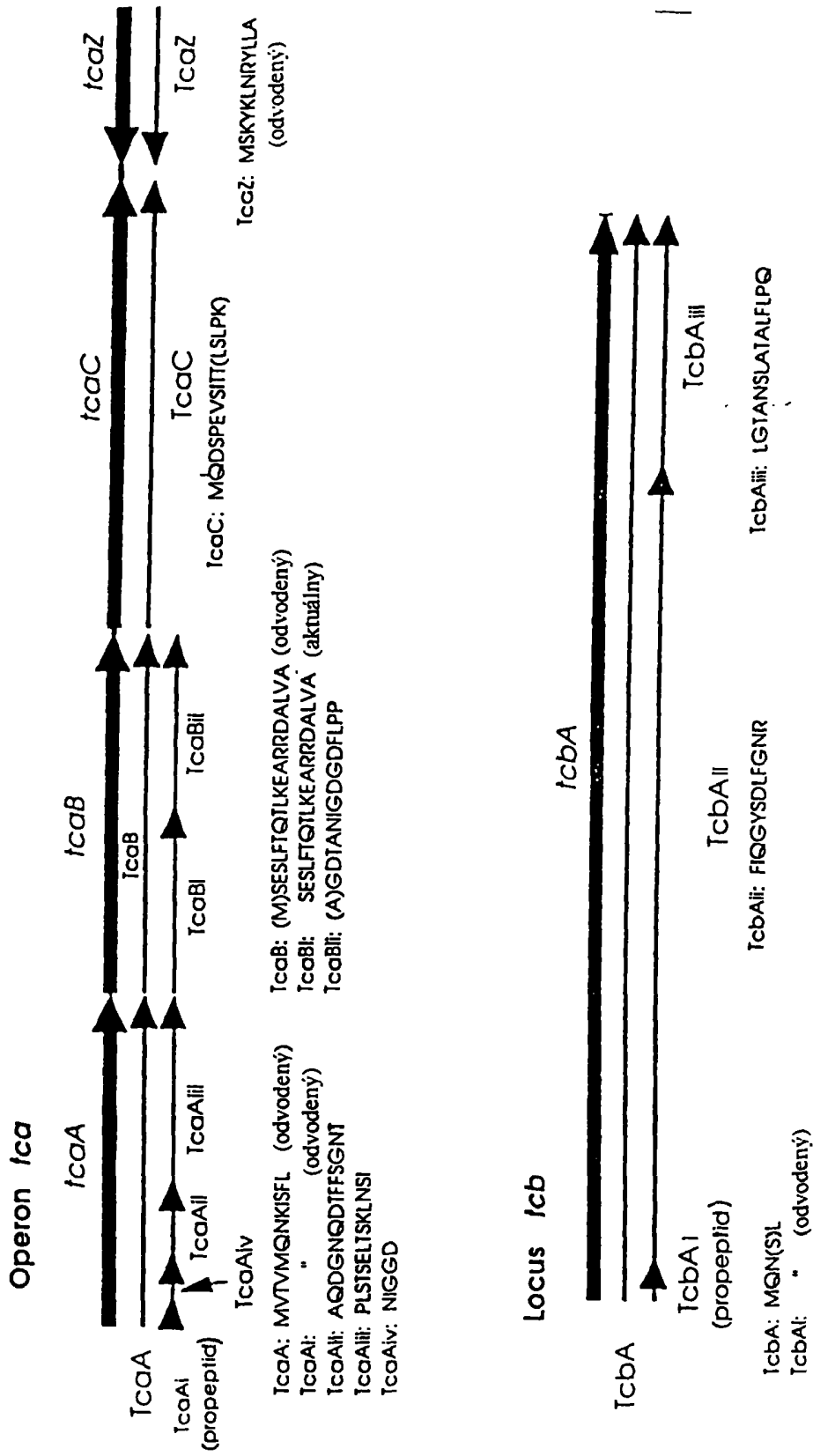
Obr. 4A

TcbA	2090	2100	2110	2120	2130
	DSYSQLYEEN	INAGEQRALA	LRSESAIESQ	GAQISRMAGA	GVDMAPNIFG
TcaBii	170	180	190	200	210
	khYSdLingg	lsAaEiagLt	LRStamI-tn	GvatglliaG	GinavPNvFG>
	~v~v~v~v~v~v	~v~v~v~v~v~v	~v~v~v~v~v~v	~v~v~v~v~v~v	~v~v~v~v~v~v
TcbA	2140	2150	2160	2170	2180
	LADGGMHYGA	IAYAIADGIE	LSASAKMVDA	EKVAQSEIYR	RRRQEWKIQR
TcaBii	220	230	240	250	260
	LANGGsewGA	pligsgqatq	vgAgiqdqSA	gisevtagYq	RRqeEWalQR>
	~v~v~v~v~v~v	~v~v~v~v~v~v	~v~v~v~v~v~v	~v~v~v~v~v~v	~v~v~v~v~v~v
TcbA	2190	2200	2210	2220	2230
	DNAQAEINQL	NAQLESLSIR	REAAEMQKEY	LKTQQAQAQA	QLTFLRSKFS
TcaBii	270	280	290	300	310
	DiAdnEItQL	daQiqSLqeq	itmAqkQitl	seTeQAnAQA	iydlqttrFt>
	~v~v~v~v~v~v	~v~v~v~v~v~v	~v~v~v~v~v~v	~v~v~v~v~v~v	~v~v~v~v~v~v
TcbA	2240	2250	2260	2270	2280
	NQALYSWLRG	RLSGIYQFY	DLAVSRCLMA	EQSYQWEAND	NSISFVKPGA
TcaBii	320	330	340	360	
	gQALYnWmaG	RLSalyYQmY	DstlpicLqP	kaalvqEgek	eSdSlfqvPv>
	~v~v~v~v~v~v	~v~v~v~v~v~v	~v~v~v~v~v~v	~v~v~v~v~v~v	~v~v~v~v~v~v
TcbA	2290	2300	2310	2320	2330
	WQGTYAGLLC	GEALIQNLAQ	MEEAYLKWES	RALEVERTVS	LAVVYDSLEG
TcaBii	370	380	390	400	410
	WndlWqGLLa	GEgLSseLqk	ldaiwLargg	igLEaiRTVS	Ldtlfgt--G>
	~v~v~v~v~v~v	~v~v~v~v~v~v	~v~v~v~v~v~v	~v~v~v~v~v~v	~v~v~v~v~v~v
TcbA	2340	2350	2360	2370	2380
	NDRFNLAEQI	PALLDKGEGT	AGTKKNGLSL	ANAILSASVK	LSDLKLGTDY
TcaBii	420	430	440	450	
	----tLsEnI	nkvLn-GEtv	spsggvtLaL	tgdlfqAtld	LSqLgLdnsY>
	~v~v~v~v~v~v	~v~v~v~v~v~v	~v~v~v~v~v~v	~v~v~v~v~v~v	~v~v~v~v~v~v
TcbA	2390	2400	2410	2420	2430
	PDSIVGSNKV	RRIKQISVSL	PALVGPYQDV	QAMLSYGGST	QLPKGCSALA
TcaBii	460	470	480	490	500
	-n--lGneKk	RRIKrIaVtL	PtLlGPYQDl	eAtLvmGaea	aLshGvndgg>
	~v~v~v~v~v~v	~v~v~v~v~v~v	~v~v~v~v~v~v	~v~v~v~v~v~v	~v~v~v~v~v~v
TcbA	2440	2450	2460	2470	
	VSHGTNDSGQ	FQLDFNDGKY	LPFEGIALDD	QGTNLNLQFPN	
TcaBii	510	520	530		
	rfvtdfndsr	F-LpF-eGrd	attgtleLn>		
	~v~v~v~v~v~v	~v~v~v~v~v~v	~v~v~v~v~v~v		

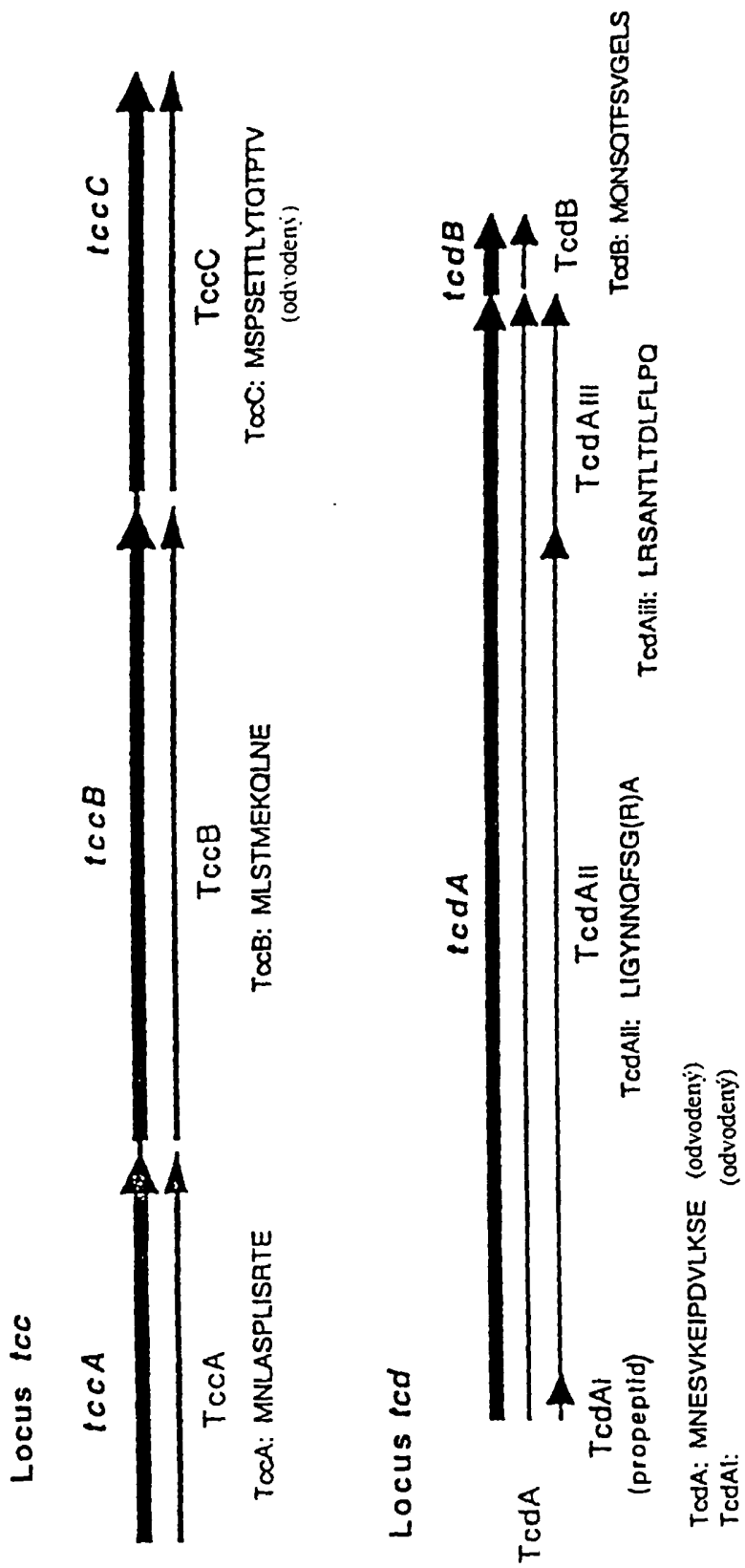
Obr. 4B



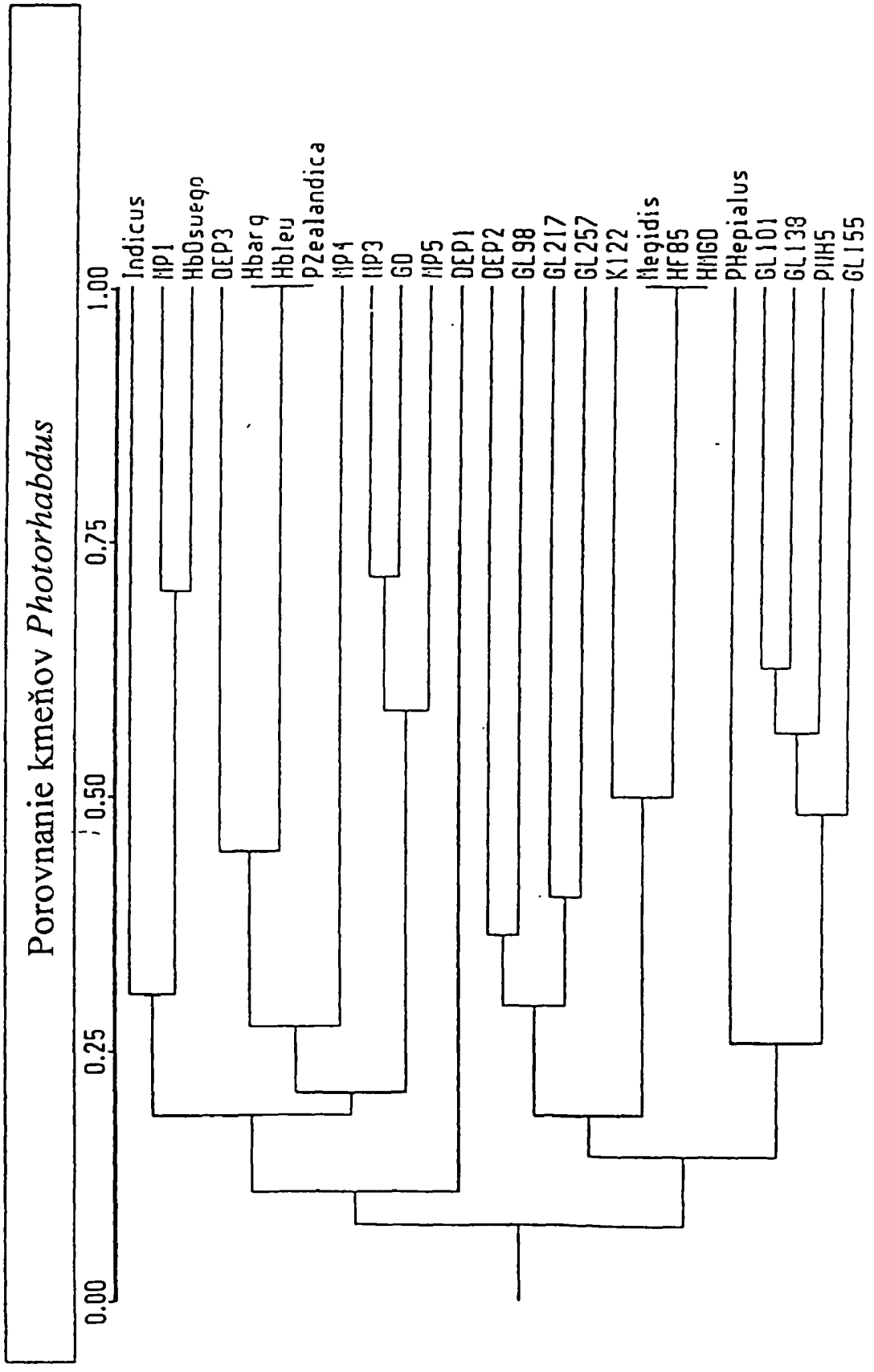
Obr. 5



Obr. 6A Loci *tca* a *tcb*, primárne génové produkty a odvodené peptidy



Obr. 6B Loci *tcc* a *tcd*, primárne génové produkty a odvodené peptidy



Obr. 7