



(12)

BREVET DE INVENȚIE

(21) Nr. cerere: **97-01251**

(22) Data de depozit: **06.11.1996**

(45) Data publicării mențiunii acordării brevetului: **28.02.2007** BOPI nr. 2/2007

(30) Prioritate:

06.11.1995 US 60/007.255;

28.02.1996 US 08/608.423;

28.08.1996 US 08/705.484

(86) Cerere internațională PCT:

Nr. **US 96/18003 06.11.1996**

(87) Publicare internațională:

Nr. **WO 97/17432 15.05.1997**

(73) Titular:

• **WISCONSIN ALUMINI RESEARCH
FOUNDATION,
614 NORTH WALNUT STREET, MADISON,
US**

(72) Inventatori:

• **ENSIGN JERALD C.,
1810 NORTH WALNUT STREET,
MADISON, US;**
• **BOWEN DAVID J., 5668 HIGHWAY A,
OREGON, US;**
• **PETELL JAMES, 1427 HUNTERS GLEN,
ZIONSVILLE, US;**
• **FATIG RAYMOND, 30 CLAY COURT,
ZIONSVILLE, US;**
• **SCHOONOVER SUE, 7142 MARSTELLA,
BROWNSBURG, US;**
• **FFRENCH CONSTANT RICHARD H.,
1006 UNIVERSITY BAY DRIVE, MADISON,
US;**

• **ROCHELEAU THOMAS A.,
3100 BUENA VISTA STREET, MADISON,
US;**

• **BLACKBURN MICHAEL B.,
2127 LUANN LANE, MADISON, US;**

• **HEY TIMOTHY D., 1653 CATALINA WAY,
ZIONSVILLE, US;**

• **MERLO DONALD J.,
11845 DUBBIN DRIVE, CARMEL, US;**

• **ORR GREGORY L.,
1028 SARATOGA CIRCLE, INDIANAPOLIS,
US;**

• **ROBERTS JEAN L., 26035 STATE ROAD
NR. 19, ARCADIA, US;**

• **STRICKLAND JAMES A.,
780 MT. ZION ROAD, LEBANON, US;**

• **GUO LINING, 7 NELSON CIRCLE,
BROWNSBURG, US;**

• **CICHE TODD A.,
1609 CHADBOURNE AVENUE, MADISON,
US**

(74) Mandatar:

**ROMINVENT S.A.,
STR. ERMIL PANGRATTI, NR. 35,
SECTOR 1, BUCUREȘTI**

(56) Documente din stadiul tehnicii:

WO 8601509;

US 4837222;

EPO 192713 A;

JP 52116472;

WO 9503695

(54)

TOXINE PROTEICE INSECTICIDE DIN *PHOTORHABDUS*

(57) Rezumat:

Invenția se referă la o metodă de control al unui dăunător insectă, prin hrănirea acestuia cu o proteină recombinată, obținută din bacteria *Photobhabdus* cu activitate toxică, care codifică niște secvențe de aminoacizi selecționați dintre SEQ ID Nr. 12 sau SEQ ID Nr. 47, ce se regă-

sește într-o celulă de plantă transgenică și, respectiv, o plantă transgenică accesibilă dăunătorului- insectă.

Revendicări: 7

Figuri: 6

Examinator: dr. ing. PARASCHIV ADRIANA



Orice persoană interesată are dreptul să formuleze în scris și motivat, la OSIM, o cerere de revocare a hotărârii de acordare a brevetului de invenție, în termen de 6 luni de la publicarea mențiunii acesteia

1 Prezenta invenție se referă la o polinucleotidă care este asociată operațional cu un
 3 promotor heterolog, la o celulă de plantă transgenică ce cuprinde polinucleotida menționată,
 la o proteină recombinantă care are activitate de toxină împotriva unui dăunător insectă, la
 o metodă de control al unui dăunător insectă.

5 Numeroase insecte sunt considerate, în mare măsură, ca dăunătoare pentru dețină-
 torii de locuințe, pentru cei care organizează sau participă la picnicuri, pentru grădinari, pen-
 7 tru fermieri și pentru alții ale căror investiții în agricultură sunt adesea distruse sau diminuate
 ca urmare a distrugerii de către insecte a recoltelor, în special în zonele unde sezonul de
 9 creștere este scurt, pagubele semnificative produse de insecte pot să însemne pierderea
 tuturor profiturilor pentru crescători și o scădere dramatică a recoltei.

11 Împuținarea furnizării de produse agricole particulare are ca urmare, invariabil, costuri
 mai ridicate pentru prelucrătorii de alimente și apoi, în final, pentru consumatorii de plante
 13 alimentare și de produse derivate de la aceste plante.

Prevenirea distrugerii plantelor de cultură și florilor și eliminarea neplăcerii insectelor
 15 dăunătoare s-au bazat de obicei pe pesticide și insecticide organice puternice, cu toxicități
 largi. Aceste produse sintetice au devenit ținta atacului general al populației, ca fiind prea
 17 supărătoare pentru mediu și pentru cei expuși la astfel de agenți. Similar, în așezările neagri-
 cole, proprietarii de locuințe ar fi satisfăcuți să evite insectele în locuințele lor, atunci când
 19 iau masa afară, fără să fie nevoie să omoare insectele.

Utilizarea extinsă a insecticidelor chimice a dat naștere la preocupări privind sănăta-
 21 tea mediului pentru fermieri, companii care produc insecticidele, agenții guvernamentale, gru-
 puri de interes public și pentru public, în general. Dezvoltarea de strategii de dirijare a pesti-
 23 cidelor mai puțin intruzive a fost impulsionată atât de preocuparea societății pentru mediu,
 cât și de dezvoltarea instrumentelor biologice care exploatează mecanismele dirijării inse-
 25 ctelor. Agenții de combatere biologică reprezintă o promisiune alternativă pentru insecticidele
 chimice. La fiecare nivel de dezvoltare evolutivă, organismele sunt prevăzute cu mijloacele
 27 pentru sporirea succesului și supraviețuirii lor. Folosirea moleculelor biologice, ca instrumen-
 te de apărare și agresiune, este cunoscută în regnurile animal și vegetal. Suplimentar, instru-
 29 mentele relativ noi ale ingineriei genetice permit modificări ale insecticidelor biologice pentru
 realizarea de soluții particulare pentru probleme particulare.

31 Un astfel de agent, *Bacillus thuringiensis* (*Bt*), este un agent insecticid eficient și este
 folosit comercial ca atare pe scară largă. De fapt, agentul insecticid al bacteriei *Bt* este o
 33 proteină care, având toxicitate limitată, se poate folosi pe culturile alimentare în ziua
 recoltării. Pentru organismele care nu sunt avute în vedere, toxina *Bt* este o proteină
 35 digestibilă netoxică.

Altă clasă cunoscută de agenți de combatere biologică a insectelor sunt anumite
 37 genuri de nematode cunoscute ca vectori de transmitere pentru simbiozii bacterieni care
 omoară insecte. Nematodele care conțin bacterii insecticide invadează larvele de insecte.
 39 Apoi, bacteriile omoară larvele. Nematodele se reproduc în cadavrul larvar. Urmașii nema-
 todei mănâncă apoi cadavrul din interior. Urmașii nematodei, conținând bacterii astfel pro-
 41 duse, pot apoi să invadeze larve suplimentare.

În trecut, s-au folosit ca agenți de combatere a insectelor nematode, insecticide din
 43 genurile *Steinernema* și *Heterorhabdis*. Aparent, fiecare gen de nematode adăpostește o
 specie anume de bacterie. În nematodele genului *Heterorhabditis*, bacteria simbiotică este
 45 *Photorhabdus luminescens*.

Cu toate că aceste nematode sunt agenți eficienți de combatere a insectelor, în pre-
 47 zent, este dificil, scump și ineficient să se producă, mențină și să se distribuie nematode pen-
 tru combaterea insectelor.

Se cunoaște din domeniu că se poate izola o toxină insecticidă din <i>Photobhabdus luminescens</i> care are activitate numai când se injectează în larvele insectelor <i>Lepidoptere</i> și <i>Coleoptere</i> (Bowen, D.J. și colaboratorii, "Extracellular insecticidal factor produced by <i>Xenorhabdus luminescens</i> ", abstract de la 89 th . Annual meeting of the American Society for Microbiology (ASM) 1989, Ensign, J.C. și colaboratorii, "Crystalline inclusion proteins as an insecticidal toxin of <i>Xenorhabdus luminescens</i> NC-19", abstract de la V th International Colloquium on Invertebrate Pathology and Microbial Control (ICIPMC), August 1990). Aceasta face imposibilă exploatarea efectivă a proprietăților insecticide ale nematodelor sau simbiotului lor bacterian. Ar fi utilă, mai practică și mai puțin consumatoare de muncă o metodă de eliberare pe o arie extinsă a unei toxine insecticide care ar păstra proprietățile sale insecticide după eliberare. Ar fi în mod clar de dorit să se descopere toxine cu activitate orală produsă de către genul <i>Photobhabdus</i> . Izolarea și folosirea acestor toxine sunt dorite datorită motivelor de eficacitate. Până la descoperirile solicitanților, aceste toxine nu au fost izolate sau caracterizate.	1
Toxinele native sunt proteine complexe care sunt produse și secretate de celulele bacteriene în creștere ale genului <i>Photobhabdus</i> . De interes sunt proteinele produse de specia <i>Photobhabdus luminescens</i> . Complexele proteice, cu o mărime moleculară de aproximativ 1 000 kDa, pot fi separate prin analiză pe gel SDS-PAGE, în numeroase componente proteice. Toxinele nu conțin activități hemolizantă, lipază, fosfolipază de tip C sau nuclează. Toxinele prezintă toxicitate semnificativă la administrare prin expunere, pentru numeroase insecte.	15
Problema tehnică pe care o rezolvă invenția este găsirea unei polinucleotide care să conțină informația genetică necesară pentru sinteza unei proteine cu efect toxic asupra unor insecte, care să fie administrată cu ușurință, care să păstreze proprietățile sale insecticide după eliberare și în special toxine cu activitate orală, produse de către genul <i>Photobhabdus</i> :	19
- stabilizarea modificărilor genei care codifică toxina sau a proteinei codificate de genă;	21
- a toxinelor care să inhibe inițial hrănirea larvelor, să mențină activitatea de inhibare a hrănirii, eliminând în același timp activitatea toxicității absolute asupra larvei, schimbare care ar putea permite plantei transformate să supraviețuiască până la recoltare, precum și	23
- dirijarea toxinei spre anumite părți ale larvei insectei, pentru îmbunătățirea eficienței acesteia.	25
Într-un prim aspect, invenția se referă la o polinucleotidă care este asociată operabil cu un promotor heterolog, în care respectiva polinucleotidă codifică o proteină care are activitate de toxină împotriva unui dăunător insectă, în care respectiva proteină cuprinde secvența de aminoacizi aleasă dintre SEQ ID NR:12 sau SEQ ID NR:47.	27
Într-un al doilea aspect, invenția se referă la o celulă de plantă transgenică ce cuprinde polinucleotida menționată și exprimă polinucleotida menționată.	29
Într-un al treilea aspect, invenția se referă la o proteină recombinantă care are activitate de toxină împotriva unui dăunător insectă, în care proteina menționată cuprinde secvența de aminoacizi aleasă dintre SEQ ID NR:12, SEQ ID NR:47.	31
Într-un al patrulea aspect, invenția se referă la o metodă de control al unui dăunător insectă, în care metoda cuprinde hrănirea cu o proteină care cuprinde secvența de aminoacizi aleasă dintre SEQ ID NR:12, SEQ ID NR:47, proteina menționată fiind produsă de și fiind prezentă într-o plantă transgenică, care este accesibilă dăunătorului menționat.	33
Metoda de control al dăunătorului insectă cuprinde alimentarea dăunătorului menționat cu o proteină din <i>Photobhabdus</i> , în care proteina menționată are activitate de toxină (pe cale) orală și în care proteina menționată este produsă de și este prezentă într-o plantă transgenică ce este accesibilă dăunătorului menționat.	35

Prezenta invenție este dirijată spre stabilirea unei clase unice de toxine proteice insecticide din genul *Photorhabdus* care au toxicitate orală împotriva insectelor. O trăsătură unică a *Photorhabdus* este bioluminescența sa. *Photorhabdus* se poate izola dintr-o diversitate de surse. O astfel de sursă o reprezintă nematodele, mai precis nematodele genului *Heterorhabditis*. Altă asemenea sursă o reprezintă probele clinice de la plăgi umane, vezi Farmer și colaboratorii, 1989, *J. Clin. Microbiol.*, pag. 1594-1600. Aceste tulpini saprofite sunt depozitate la Colecția Americană de Tipuri de Culturi (Rockville, MD) ATCC #s 43948, 43949, 43950, 43951 și 43952 și sunt încorporate aici prin referire. Este posibil ca și alte surse să poată adăposti bacterii *Photorhabdus* care produc toxine insecticide. Astfel de surse din mediu ar putea fi atât din mediu terestru, cât și din mediul acvatic.

Genul *Photorhabdus* este definit taxonomic ca un membru al familiei *Enterobacteriaceae*, deși are unele trăsături atipice acestei familii. De exemplu, tulpinile acestui gen nu sunt reducătoare de azotat, produc pigment galben și roșu și sunt bioluminescente. Această ultimă trăsătură este necunoscută la *Enterobacteriaceae*, numai de curând *Photorhabdus* a fost descris ca un gen separat de *Xenorhabdus* (Boemare și colaboratorii, 1993, *Int. J. Syst. Bacteriol.* 43. 249-255). Această diferențiere se bazează pe studii de hibridizare ADN-ADN, diferențe fenotipice (de exemplu, prezența (*Photorhabdus*) sau absența (*Xenorhabdus*) a catalazei și bioluminescenței) și familia de nematode gazdă (*Xenorhabdus*: *Steinemematidae*, *Photorhabdus*; *Heterorhabditidae*). Comparativ, analizele de acid gras celular (Janse și colaboratorii, 1990, *Lett. Appl. Microbiol.* 10. 131-135; Suzuki și colaboratorii, 1990, *J. Gen. Appl. Microbiol.*, 36. 393-401) stau la baza separării *Photorhabdus* de *Xenorhabdus*.

În vederea stabilirii dacă colecția tulpinii dezvăluită aici a cuprins tulpini *Photorhabdus*, tulpinile s-au caracterizat pe baza trăsăturilor care definesc *Photorhabdus* și îl diferențiază de alte *Enterobacteriaceae* și specii *Xenorhabdus*. (Farmer, 1984 *Bergey's Manual of Systemic Bacteriology*, vol. 1, p. 510-511; Akhurst și Boemare, 1988, *J. Gen. Microbiol.* 134, p. 1835-1845; Boemare și colaboratorii, 1993, *Int. J. Syst. Bacteriol.* 43, p. 249-255, care sunt încorporate aici prin referire). Trăsăturile studiate au fost următoarele: ramurile colorate gram negativ, mărimea organismului, pigmentația coloniei, corpi de incluziune, prezența catalazei, capacitatea de a reduce azotatul, bioluminescența, absorbția de colorant, hidroliza gelatină, creștere pe medii selective, temperatura de creștere, supraviețuirea în condiții anaerobe și mobilitatea. Analiza de acid gras s-a folosit pentru a confirma că toate tulpinile de aici aparțin la un singur gen *Photorhabdus*.

Curent, genul bacterian *Photorhabdus* cuprinde o singură specie definită, *Photorhabdus luminescens* (ATCC tulpina #29999, Poinar și colaboratorii, 1977, *Nematologica* 23, 97-102). În literatură, a fost descrisă o diversitate de tulpini înrudite (de exemplu, Akhurst și colaboratorii, 1988, *J. Gen. Microbiol.*, 134, 1835-1845; Boemare și colaboratorii, 1993, *Int. J. Syst. Bacteriol.* 43, p. 249-255; Putz și colaboratorii, 1990, *Appl. environ. Microbiol.*, 56, 181-186). Aici au fost caracterizate numeroase tulpini *Photorhabdus*. Asemenea tulpini sunt listate în tabelul 18 din exemple. Deoarece există curent numai o specie (*luminescens*) definită în genul *Photorhabdus*, trăsăturile speciei *luminescens* s-au folosit aici pentru caracterizarea tulpinilor. Așa cum se poate observa din fig. 5, aceste tulpini sunt în mod clar diverse. Nu se prevede că, în viitor, vor fi alte specii *Photorhabdus* care să aibă unele din atributele speciei *luminescens*, precum și unele caracteristici diferite care, în prezent, nu s-au definit ca o trăsătură a *Photorhabdus luminescens*. Cu toate acestea, invenția de față se referă la orice specie sau tulpini *Photorhabdus* care produc proteine cu activitatea funcțională de agenți de combatere a insectelor, fără a se lua în considerare alte trăsături și caracteristici.

Mai mult, așa cum s-a demonstrat aici, bacteriile genului *Photorhabdus* produc proteine care au activitate funcțională, așa cum s-a definit aici. De interes deosebit sunt proteinele produse de către specia *Photorhabdus luminescens*. Invențiile de aici nu vor fi în nici un mod limitate la tulpinile care sunt descrise aici. Aceste tulpini ilustrează, pentru prima dată, că proteinele produse de către diverse specii izolate ale *Photorhabdus* sunt toxice pentru insecte la expunerea lor la acestea. Astfel, în invențiile descrise aici sunt incluse tulpinile specificate aici și toți mutanții acestora, precum și orice tulpini sau specii ale genului *Photorhabdus* care au activitatea funcțională descrisă aici.

Există câțiva termeni folosiți care au o însemnătate particulară, după cum vom vedea în cele ce urmează.

Prin "activitate funcțională" se înțelege aici că funcția toxinelor proteice, ca agenți de combatere a insectelor, constă în aceea că proteinele sunt active oral sau au un efect toxic, sau sunt capabile să distrugă sau să deterioreze hrănirea, care poate cauza sau nu moartea insectei. Când o insectă vine în contact cu o cantitate eficientă de toxină eliberată prin intermediul expresiei de către o plantă, cu o compoziție(i) formulată de proteină, compoziție(i) pulverizate de proteină, o matrice momeală sau alt sistem de eliberare, urmările constă, de obicei, în moartea insectei sau insectele nu se hrănesc din sursa care face toxinele accesibile pentru insecte. Toxinele proteice discutate aici sunt denumite tipic ca "insecticide". Prin insecticide se înțelege că toxinele proteice au o "activitate funcțională" așa cum s-a definit suplimentar aici și sunt folosite ca agenți de combatere a insectelor.

Prin folosirea termenului "oligonucleotide" se înțelege aici o macromoleculă care constă dintr-o catenă scurtă de nucleotide fie de ADN, fie de ARN. Asemenea lungime ar putea fi de cel puțin o nucleotidă, dar de obicei este în domeniul de circa 10 la circa 12 nucleotide. Determinarea lungimii oligonucleotidei cade în sarcina unui lucrător și nu va fi o limitare aici. Prin urmare, oligonucleotidele pot fi mai mici decât 10 sau mai mari decât 12.

Prin folosirea termenului "toxic" sau "toxicitate", așa cum s-a folosit aici, se înțelege că toxinele produse de către *Photorhabdus* au "activitate funcțională", așa cum s-a definit aici. Prin folosirea termenului "material genetic" se face referire la toate genele, acid nucleic, ADN și ARN.

Bulioanele de fermentare de la tulpinile selectate aici, raportate în tabelul 18, s-au folosit pentru a determina următoarele: mărirea producției toxinei insecticide de către genul *Photorhabdus*, spectrul insecticid al acestor toxine și să asigure sursa materială pentru purificarea complexelor de toxină. Tulpinile caracterizate aici s-au dovedit a avea toxicitate orală împotriva unei diversități de ordine de insecte. Astfel de ordine de insecte includ, dar nu sunt limitate la *Coleoptera*, *Homoptera*, *Lepidoptera*, *Diptera*, *Acarina*, *Hymenoptera* și *Dictyoptera*.

La fel cu alte toxine bacteriene, rata de mutație a bacteriilor dintr-o populație produce numeroase toxine înrudite, ușor diferite în secvență ca să existe. Toxinele de interes aici sunt cele care produc complexe proteice toxice pentru o diversitate de insecte la expunere, așa cum s-a descris aici. De preferință, toxinele sunt active împotriva *Lepidoptera*, *Coleoptera*, *Homoptera*, *Diptera*, *Hymenoptera*, *Dictyoptera* și *Acarina*. Prin invențiile din domeniu se încearcă să se facă referire la toxinele proteice omoloage toxinelor proteice produse de tulpinile de aici și orice derivat al acestor tulpini, precum și toxine proteice produse de către *Photorhabdus*. Aceste proteine omoloage pot să difere ca secvență, dar nu diferă în funcție de acele toxine descrise aici. Toxinele omoloage sunt înțelese ca incluzând complexe de proteine între 300 kDa până la 2 000 kDa și cuprind cel puțin 2 subunități, unde o subunitate este o peptidă care poate sau nu poate să fie aceeași ca cealaltă subunitate. S-au identificat diverse subunități proteice și s-au luat în considerare în exemplele de aici. De obicei, subunitățile proteice sunt între circa 18 kDa la circa 230 kDa; între circa 160 kDa la circa 230 kDa, 100 kDa la 160 kDa; circa 80 kDa la circa 100 kDa și circa 50 kDa la circa 80 kDa.

Prin folosirea termenului "toxină *Photorhabdus*" se înțelege orice proteină produsă de o tulpină a microorganismului *Photorhabdus* care are activitate funcțională împotriva insectelor, unde toxina *Photorhabdus* s-ar putea formula ca o compoziție pulverizabilă, exprimată printr-o plantă transgenică, formulată ca o momeală matrice, eliberată prin intermediul unui *Baculovirus* sau eliberată de către orice altă gazdă aplicabilă sau sistem de eliberare.

Așa cum s-a discutat mai sus, unele tulpini *Photorhabdus* pot fi izolate de la nematode. Unele nematode, viermi cilindrici alungiți, paraziți ai filumului Nematoda, au dezvoltat o capacitate de a exploata larve de insecte, ca un mediu favorabil de creștere. Larvele de insecte asigură o sursă de hrană pentru creșterea nematodelor și un mediu în care se reproduc. Un efect dramatic care urmează invadării larvelor de către anumite nematode este moartea larvară. Moartea larvară rezultă din prezența, în anumite nematode, a bacteriilor care produc o toxină insecticidă care oprește creșterea larvară și inhibă activitatea de hrănire. Interesant e faptul că fiecare gen de nematodă, care parazitează o insectă, găzduiește o specie anume de bacterie, adaptată unic pentru creștere simbiotică cu acea nematodă. Provizoriu, deoarece această cercetare a început, numele genului bacterian *Xenorhabdus* s-a reclassificat în *Xenorhabdus* și *Photorhabdus*. Bacteriile genului *Photorhabdus* sunt caracterizate ca fiind simbiionți ai nematodelor *Heterorhabdus*, în timp ce speciile *Xenorhabdus* sunt simbiionți ai speciei *Steinernema*. Această schimbare în nomenclatură se reflectă în această descriere, dar, în nici un mod, o schimbare în nomenclatură nu va modifica întinderea invențiilor descrise aici.

Peptidele și genele care sunt descrise aici sunt numite conform ghidurilor publicate recent în *Journal of Bacteriology* "Instructions to Authors", p. I-xii (Jan. 1996), care este încorporată aici, prin referire. Următoarele peptide și gene s-au izolat din *Photorhabdus* tulpina W-14.

Nomenclatură peptidă/genă
Complex toxină (Tc)

Nume peptidă	Nume genă	Secvență ID# brevet
regiune genomică tca		
TcaA	tcaA	12
TcaA _{iii}	tcaA	4
TcaB _i	tcaB	3(19, 20)
TcaB _{ii}	tcaB	5
TcaC	tcaC	2
regiune genomică tcb		
TcbA	tcbA	16
TcbA _i	tcbA	(pro-peptidă)
TcbA _{ii}	tcbA	1 (21, 22, 23, 24)
TcbA _{iii}	tcbA	40

Tabel (continuare)

Nume peptidă	Nume genă	Secvență ID# brevet
regiune genomică tcc		
TccA	tccA	8
TccB	tccB	7
regiune genomică tcd		
TcdAi	tcdA	(pro-peptidă)
TcdA _{ii}	tcdA	13, (38, 39, 17, 18)
TcdA _{iii}	tcdA	41, (42, 43)
TcdB	tcdB	14

(Secvența din paranteză arată secvența aminoacidă internă, obținută prin digeste triptice).

Secvențele listate mai sus sunt grupate prin regiunea genomică. Gena *tcbA* s-a exprimat în *E. coli*, ca două fragmente de proteină TcbA și TcbA_{iii}, așa cum s-a ilustrat în exemple. Poate fi benefic să existe o scurtare proteolitică a unor secvențe, pentru a obține activitate mai mare a toxinelor pentru aplicații transgenice comerciale.

Toxinele descrise aici sunt în mod clar unice, prin aceea că toxinele au activitate funcțională care este cheia dezvoltării unei strategii de combatere a insectelor. În dezvoltarea unei strategii pentru combaterea insectelor este posibil să se întârzie sau să se ocolească procesul degradării proteinei prin injectarea unei proteine direct într-un organism, evitând tractul său digestiv. În astfel de cazuri, proteina administrată organismului va reține funcția sa, până când ea este denaturată, degradată nespecific sau eliminată de către sistemul imun din organisme superioare.

Injectarea în insecte a unei toxine insecticide are aplicare potențială numai în laborator și numai pe insecte mari care sunt injectate cu ușurință. Observația că toxinele proteice insecticide descrise aici prezintă activitate toxică, după ingestie orală sau contact cu toxinele, permite dezvoltarea unui plan de control al insectelor, bazat numai pe capacitatea de a încorpora toxinele proteice în hrana insectei. Un astfel de plan ar putea avea ca rezultat producerea de momeli pentru insecte.

Toxinele *Photographus* se pot administra la insecte într-o formă purificată. De asemenea, toxinele pot fi eliberate în cantități de la circa 1 la circa 100 mg/l de mediu. Aceasta poate varia în funcție de condiția formulării, condițiile sursei inoculului, tehnicile pentru izolarea toxinei și altele. Toxinele se pot administra ca o secreție exudată sau proteină celulară, exprimată inițial într-o gazdă heteroloagă procariotă sau eucariotă. De obicei, bacteriile sunt gazdele în care sunt exprimate proteine. Gazdele eucariote pot include, dar nu se limitează la plante, insecte și drojdie. Alternativ, toxinele pot fi produse în bacterii sau plante transgenice, în câmp sau în insectă, printr-un vector baculovirus. Tipic, toxinele vor fi introduse în insectă, prin încorporarea uneia sau mai multor toxine în hrana insectelor.

Letalitatea completă la hrănirea insectelor este utilă, dar nu este necesară pentru atingerea toxicității utile. Dacă insectele evită toxina sau încetează hrănirea, această evitare va fi utilă în unele aplicații, chiar dacă efectele sunt subletale. De exemplu, dacă sunt dorite plante de cultură transgenice, rezistente la insecte, o neplăcere a insectelor să se hrănească pe plante este la fel de utilă ca și toxicitatea letală, deoarece obiectivul final este mai degrabă protecția plantelor, și nu omorârea insectei.

Există numeroase alte căi în care toxinele pot fi încorporate în hrana unei insecte. Ca un exemplu, este posibil să se amestece sursa de hrană a larvei cu proteina toxică, prin stropirea hranei cu o soluție de proteină, cum s-a descris aici. Alternativ, proteina purificată ar putea fi prelucrată prin inginerie genetică într-o bacterie mai puțin dăunătoare, care apoi ar putea fi crescută în cultură și aplicată fie la sursa de hrană, fie s-ar lăsa să stea în sol, într-o zonă în care a fost dorită eradicarea insectei. De asemenea, proteina s-ar putea prelucra prin inginerie genetică într-o sursă de hrană pentru insectă. De exemplu, sursa principală de hrană pentru numeroase larve de insecte o reprezintă materialul vegetal.

Prin încorporarea materialului genetic care codifică proprietățile insecticide ale toxinelor *Photorehabdus* în genomul unei plante consumate de insecta dăunătoare particulară, adultul sau larva vor muri după consumarea ca hrană a plantei. Numeroși membri ai genurilor de monocotiledonate și dicotiledonate au fost transformați. Plante agricole transgenice, precum și fructe și legume sunt de interes comercial. Asemenea plante cultivate includ dar nu se limitează la porumb, orez, soia, rapiță, floarea-soarelui, lucernă, sorg, grâu, bumbac, arahide, tomate, cartof și altele. Există câteva tehnici pentru introducerea materialului genetic străin în celulele plantelor și pentru obținerea plantelor care mențin stabilă și exprimă gena introdusă. Astfel de tehnici includ accelerarea materialului genetic acoperit în microparticule direct în celule (brevete **SUA 4 954 050** pentru Corneli și **5 141 131** pentru DowElanco). Plantele pot fi transformate folosind tehnologia *Agrobacterium*, vezi brevet **SUA 5 177 010** pentru University of Toledo, **5 104 310** pentru Texas A&M, cererea de brevet european **0131624B1**, cererile de brevet european **120516**, **159418B1** și **176 112** pentru Schilperoot, brevete **SUA 5 149 645**, **5 469 976**, **5 464 763** și **4 940 838** și **4 693 976** pentru Scilperoot, cererile de brevet european **116718**, **290799**, **320500** toate pentru MaxPlanck, cererile de brevet european **604662** și **627752** pentru Japan Tobacco, cererile de brevet european **0267159** și **0292435** și brevetul SUA pentru Ciba Geigy, brevetele **SUA 5 463 174** și **4 762 785**, ambele pentru Calgene, și brevetele **SUA 5 004 863** și **5 159 135** ambele pentru Agracetus. Altă transformare include tehnologia contactelor punctiforme, vezi brevetele **SUA 5 302 523** și **5 464 765** ambele pentru Zeneca. Tehnologia electroporării a fost de asemenea folosită pentru transformarea plantelor, vezi **WO 87/06614** pentru Boyce Thomson Institute, **5 472 869** și **5 384 253** ambele pentru Dekalb, **WO 92/09696** și **WO 93/21335** ambele pentru PGS. Toate aceste brevete și publicații asupra transformării sunt încorporate aici, prin referire. Suplimentar, la numeroasele tehnologii pentru transformarea plantelor, tipul țesutului care se contactează cu genele străine poate varia la fel de bine. Astfel de țesut va include, dar nu va fi limitat la țesut embriogenic, țesut calus de tip I și II, hipocotil, meristem și altele. Aproape toate țesuturile plantelor pot fi transformate în timpul diferențierii, folosind tehnici corespunzătoare cuprinse în pregătirea lucrătorului în domeniu.

Altă variabilă este alegerea unui marker selectabil. Preferința pentru un marker particular este la discreția specialistului, dar oricare dintre următorii markeri selectabili se pot folosi împreună cu oricare altă genă nementionată aici, care ar putea să funcționeze ca un marker selectabil. Astfel de markeri selectabili includ dar nu se limitează la gena aminoglicozid fosfotransferază a transpozoniului Tn5 (Aph II), care codifică rezistența la antibioticele kanamicină, neomicină și G418, precum și acele gene care codifică pentru rezistență sau toleranță la glifozat, higromicină, metotrexat, fosfotricină (bialofos), imidazolinone, sulfoniluree și erbicide triazopirimidină, cum ar fi clorosulfuron, bromoxinil, dalapon și altele.

În plus față de un marker selectabil, poate fi dorită folosirea unei gene reporter. În unele situații o genă reporter poate fi folosită fără un marker selectabil. Genele reporter sunt gene care de obicei nu sunt prezente sau exprimate în organismul sau țesutul receptor. În

mod obișnuit, gena reporter codifică pentru o proteină care asigură o schimbare fenotipică sau proprietate enzimatică. Exemple de astfel de gene sunt date în K. Weising și colaboratorii, *Ann. Rev. Genetics*. 22, 421 (1988), care este încorporată aici prin referire. O genă reporter preferată este gena glucuronidază (GUS). Indiferent de tehnica de transformare, gena este încorporată de preferință într-un vector de transfer genă, adaptat pentru a exprima toxinele *Photobacterium* în celula plantei prin includerea în vector a unui promotor de plantă. Suplimentar față de promotorii de plantă, se pot folosi eficient, în celule de plantă, promotori dintr-o diversitate de surse pentru exprimarea genelor străine. De exemplu, pot fi folosiți promotori de origine bacteriană, cum ar fi promotorul octopin sintază, promotorul nopalin sintază, promotorul mannopin sintază, promotori de origine virală, cum ar fi virusul mozaicului conopidei (35S și 19S) și alții. Promotorii de plantă includ, dar nu sunt limitați la ribuloz-1,6-bisfosfat (RUBP) carboxilază subunitatea mică (ssu), promotorul *beta*-conglicinină, promotorul fazeolină, promotorul ADH, promotorii șocului termic și promotori specifici țesutului. De asemenea, promotorii pot să conțină elemente de secvență intensificatoare care poate îmbunătăți eficiența transcripției. Intensificatori tipici includ dar nu se limitează la Adh-intron 1 și Adh-intron 6. Se pot folosi și promotori constitutivi. Promotorii constitutivi dirijează continuu expresia genei în toate tipurile de celule și la toate momentele (de exemplu, actina, ubiquitina, CaMV 35S). Promotorii specifici țesutului sunt responsabili pentru expresia genei în tipuri specifice de celulă sau țesut, cum ar fi frunze sau semințe (de exemplu, zeină, oleozină, napină, ACP) și acești promotori se pot, de asemenea, folosi. De asemenea, promotorii pot să fie activi în timpul unui anumit stadiu al dezvoltării plantelor, precum și activi în țesuturi și organe de plantă. Exemple de astfel de promotori includ dar nu se limitează la specific polenului, specific embrionului, specific mătasei porumbului, specific fibrei de bumbac, specific rădăcinii, promotori specifici endospermului seminței și alții.

În anumite circumstanțe, poate fi de dorit să se folosească un promotor inductiv. Un promotor inductiv este responsabil pentru expresia genelor ca răspuns la un semnal specific, cum ar fi: stimul fizic (gene șoc termic), lumină (carboxilază RUBP), hormon (Em), metaboliti și stres. Se pot folosi alte elemente dorite de transcripție și translație, care funcționează în plante. În domeniu sunt cunoscuți numeroși vectori de transfer genă specifici plantelor.

În plus, se știe că pentru a obține expresia ridicată a genelor bacteriene în plante este de preferat să se refacă prin inginerie genele bacteriene, astfel încât acestea sunt mai eficiente exprimate în citoplasma plantelor. Porumbul este o astfel de plantă la care se preferă să se refacă prin inginerie gena(le) bacteriene, înainte de transformare pentru creșterea nivelului expresiei toxinei în plantă. Un motiv pentru refacere prin inginerie este conținutul G+C foarte scăzut al genei(lor) bacteriene native (și în consecință corectarea față de conținutul ridicat A+T). Aceasta are ca urmare generarea secvențelor care imită sau duplică secvențele de control ale genei plantei, care sunt cunoscute a fi puternic bogate A+T. Prezența unor secvențe bogate A+T în ADN-ul genei(lor) introduse în plante (de exemplu, regiuni box TATA găsite în mod normal în promotorii genei) poate avea ca urmare transcripția aberantă a genei(lor). Pe de altă parte, prezența altor secvențe reglatoare care aparțin mARN-ului transcris (de exemplu, secvențe semnal poliadenilare (AAUAAA), sau secvențe complementare la ARN-uri nucleare mici implicate în splicing pre-mARN) poate duce la instabilitatea ARN-ului. Prin urmare, un scop în modelarea genei(lor) bacteriene prin inginerie, mai preferat cunoscute ca genă(e) de plantă optimizată, este să se genereze o secvență ADN care are un conținut de G+C mai ridicat și de preferință unul apropiat de cel al genelor plantei care codifică pentru enzime metabolice. Alt scop în modelarea genei(lor) optimizate de plantă este generarea unei secvențe ADN care nu numai că are un conținut mai ridicat G+C, dar schimbările care modifică secvența vor fi astfel făcute încât să nu împiedice translația.

Un exemplu de o plantă care are un conținut ridicat G+C este porumbul. Tabelul de mai jos ilustrează cât de ridicat este conținutul G+C în porumb. La fel ca la porumb, se crede că și în alte plante conținutul G+C este de asemenea ridicat.

Tabelul 1

Alcătuirea conținuturilor G+C ale regiunilor genelor de la porumb care codifică proteină

Clasa proteinei ^a	Domeniul G+C%	Medie G+C% ^b
Enzime metabolice (40)	44,4 - 75,3	59,8 (8,0)
Proteine de depozitare		
Grup I (23)	46,0-51,9	48,1 (1,3)
Grup II (13)	60,4-74,3	67,5 (3,2)
Grup I + II (36)	46,0-74,3	55,1 (9,6) ^c
Proteine structurale (18)	48,6-70,5	63,6 (6,7)
Proteine reglatoare (5)	57,2-68,9	62,0 (4,9)
Proteine necaracterizate (9)	41,5-70,3	64,3 (7,2)
Toate proteinele (108)	44,4-75,3	60,8 (5,2)

Numărul genelor din clasă este dat în paranteze.

Deviațiile standard sunt date în paranteze.

Medie grupuri combinate, ignorată în calculul mediei totale.

Pentru datele din tabelul 1, regiunile de codificare ale genelor s-au extras din iritării GenBank (Release 71), iar compozițiile de baze s-au calculat folosind programul MacVectorTM(IBI. New Haven, CT). Secvențele intron s-au ignorat în calcule. Grupul I și II al secvențelor genei proteinei de depozitare s-au remarcat prin diferența lor marcată în compoziția de baze.

Datorită plasticității permise de redundanța codului genetic (adică, unii aminoacizi sunt specificați prin mai mult decât un codon), evoluția genoamelor diferitelor organisme sau clase de organisme a avut drept rezultat folosirea diferențială a codonilor redundanți. Acest "codon bias" se reflectă în media compoziției de bază a regiunilor care codifică proteină. De exemplu, organisme având conținuturi relativ scăzute G+C utilizează codoni care au A sau T în poziția a treia a codonilor redundanți, în timp ce acelea care au conținuturi G+C utilizează codoni care au G sau C în a treia poziție. Se consideră că prezența codonilor "minori" într-o genă a mRNA poate să reducă rata absolută de translație a celui mRNA, în special când abundența relativă a tARN-ului încărcat, care corespunde la codonul minor, este scăzută. O extindere a acesteia este că diminuarea ratei de translație de către codoni minori individuali va fi cel puțin cumulativă pentru codoni minori multipli. Ca urmare, mRNA-uri care au conținuturi relativ ridicate de codoni minori vor avea rate reduse corespunzătoare de translație. Această rată va fi reflectată prin sinteza nivelurilor scăzute ale proteinei codificate.

În scopul refacerii prin inginerie a genei(lor) bacteriene, se determină codonul bias al plantei. Codonul bias este distribuția statistică a codonului pe care îl folosește planta pentru codificarea proteinelor sale. După determinarea bias-ului, se determină procentul frecvenței codonilor din gena(le) de interes. Se vor determina codonii primari preferați de plantă, precum și a doua și a treia opțiune de codoni preferați. Secvența aminoacidă a proteinei de interes este translatată invers, astfel că secvența de acid nucleic rezultată codifică pentru

aceeași proteină ca gena bacteriană nativă, dar secvența de acid nucleic rezultată corespunde primilor codoni preferați ai plantei dorite. Noua secvență se analizează pentru situri enzimice de restricție, care s-au putut crea prin modificare. Siturile identificate sunt modificate în continuare prin înlocuirea codonilor cu a doua sau a treia opțiune de codoni preferați. Alte situri din secvență, care ar putea afecta transcripția sau translația genei de interes, sunt joncțiuni exon: intron 5' sau 3', semnale suplimentare poliA sau semnale de terminare ARN polimerază. În continuare, secvența este analizată și modificată pentru reducerea frecvenței dubletelor TA sau GC. Suplimentar la dublete, secvența G sau C blochează ceea ce are mai mult decât circa 4 resturi, care sunt aceleași care pot afecta transcripția secvenței. Prin urmare, aceste blocaje sunt de asemenea modificate prin înlocuirea codonilor de prima sau a doua opțiune etc., cu codonul de alegere preferat în continuare. Este de preferat ca gena(ele) optimizată de plantă să conțină circa 63% din codonii primei opțiuni, între circa 22% până la circa 37% codoni de a doua opțiune și între 15% și 0% codoni de a treia opțiune, în care procentul total este 100%. Cel mai preferat, gena(ele) optimizate de plantă conțin circa 63% codoni de prima opțiune, cel puțin circa 22% codoni de a doua opțiune, circa 7,5% codon de a treia opțiune și circa 7,5% codoni de a patra opțiune, în care procentul total este 100%. Metoda descrisă mai sus face posibil ca specialistul în domeniu să modifice gena(ele) care sunt străine față de o plantă anume, astfel încât genele sunt exprimate optim în plante. Metoda este explicată suplimentar în cererea SUA aflată în curs de rezolvare, **60/005 405**, depusă în 13 octombrie 1995, care este încorporată aici ca referință.

Astfel, în scopul modelării genei(lor) optimizate de plantă, secvența aminoacidică a toxinelor se translatează invers într-o secvență ADN, utilizând un cod genetic neredundant stabilit de la un tabel de codon bias întocmit pentru secvența ADN pentru planta particulară de transformat. Secvența ADN rezultată, care este complet omogenă în folosirea codonului, este modificată suplimentar pentru a stabili o secvență ADN care pe lângă un grad mai ridicat de diversitate conține, de asemenea, siturile de recunoaștere a enzimelor de restricție, plasate strategic, compoziție dorită de baze și o absență a secvențelor care pot să interfereze cu transcripția genei sau translația produsului mRNA.

Teoretic, se consideră că genele bacteriene pot fi exprimate mai ușor în plante, dacă genele bacteriene sunt exprimate în plastide. Astfel, poate fi posibil să se exprime gene bacteriene în plante, fără optimizarea genelor pentru expresia plantei și să obțină expresie ridicată a proteinei. Vezi, brevete **SUA 4 762 785**, **5 451 513** și **5 545 817**, care sunt încorporate aici ca referință.

Una dintre opiniile luate în considerare în privința exploatarei comerciale a plantelor transgenice este dirijarea rezistenței. Aceasta este o preocupare deosebită cu toxinele *Bacillus thuringiensis*. Există numeroase companii care exploatează comercial *Bacillus thuringiensis* și există preocupare asupra faptului dacă toxinele *Bt* devin rezistente. O strategie pentru dirijarea rezistenței insectelor ar fi să se combine toxinele produse de *Photographus* cu toxine cum ar fi *Bt*, proteine de insecte vegetative (*Ciba Geigy*) sau alte toxine. Combinațiile s-ar putea formula pentru aplicare prin pulverizare sau ar putea fi combinații moleculare. Plantele s-ar putea transforma cu gene *Photographus*, care produc toxine de insecte și alte gene pentru toxine de insecte ca *Bt* cu alte gene toxine de insecte, precum *Bt*.

Cererea de brevet european **0400246A1** descrie transformarea a 2 *Bt* într-o plantă, care ar putea fi oricare 2 gene. Alt mod de a produce o plantă transgenică, care conține mai mult decât o genă de rezistență la insectă, ar fi să se producă 2 plante, fiecare conținând o genă de rezistență la insectă. Aceste plante s-ar putea încrucișa folosind tehnici tradiționale de selecție, pentru a produce o plantă care conține mai mult decât o genă de rezistență la insecte.

1 Suplimentar producerii unei plante transformate care conține genă(e) optimizată de
 3 plantă, există alte sisteme de eliberare unde se poate să se reconstruiască prin inginerie,
 5 genă(e) bacteriană. Urmând aceleași linii, poate fi obținută prin inginerie genetică o toxină
 7 proteică izolată cu ușurință, fuzionată împreună atât la o moleculă care atrage insecte, cât
 9 și la o sursă de hrană, iar activitatea insecticidă a toxinei poate fi prelucrată prin inginerie și
 exprimată în bacterii sau celule eucariote, folosind tehnici standard binecunoscute. După
 purificare în laborator, un asemenea agent toxic cu momeală "încorporată" ar putea fi împachetat în interiorul unei capcane standard pentru insectele din locuințe.

Altă schemă de eliberare este încorporarea materialului genetic al toxinelor într-un
 vector baculovirus. Baculovirusurile infectează insecte gazdă particulare, incluzând pe cele
 avute în vedere cu toxinele *Photorhabdus*. Baculovirus infecțios, care găzduiește un construct de expresie pentru toxinele *Photorhabdus*, s-ar putea introduce în zonele infestării cu
 insecte și prin aceasta ar intoxica sau otrăvi insectele infectate. Transferul proprietăților
 insecticide necesită secvențe de acid nucleic care codifică secvențele aminoacide pentru toxinele *Photorhabdus* integrate într-un vector de expresie proteină, corespunzător pentru gazda
 în care va fi vectorul. O cale pentru a obține o secvență de acid nucleic care codifică o proteină cu proprietăți insecticide este să se izoleze materialul genetic nativ, care produce toxinele
 din *Photorhabdus*, folosind informația dedusă din secvența aminoacidă a toxinei, ale cărei
 porțiuni mari sunt menționate în continuare. Așa cum s-a descris în continuare, sunt dezvăluite de asemenea, metode de purificare a proteinelor responsabile pentru activitatea toxinei.
 Folosind datele secvenței aminoacide N-terminale, cum sunt cele menționate mai jos, cineva
 poate construi oligonucleotide complementare la toate sau la o secțiune a bazelor ADN care
 codifică primii aminoacizi ai toxinei. Aceste oligonucleotide se pot radiomarca și folosi ca
 sonde moleculare pentru izolarea materialului genetic dintr-o bancă genetică genomică, construită de la materialul genetic izolat din tulpini de *Photorhabdus*. Banca genetică poate fi
 donată în vectori plasmid, cosmid, fag sau fagemid. Banca ar putea fi transformată în
Escherichia coli și căutată pentru producerea toxinei de către celulele transformate, folosind anticorpi produși împotriva toxinei sau analize directe pentru toxicitate la insecte.

Această abordare necesită producerea unui set de oligonucleotide, deoarece codul
 genetic degenerat permite ca un aminoacid să poată fi codificat în ADN prin oricare dintre
 cele câteva combinații de 3 nucleotide. De exemplu, aminoacidul arginină poate fi codificat
 prin tripletele de acid nucleic CGA, CGC, CGG, CGT, AGA, și AGG. Deoarece nu se poate
 spune cu precizie care triplet este folosit la acele poziții în gena -toxinei, cineva trebuie să
 prepare oligonucleotide cu fiecare triplet potențial reprezentat. Pot fi necesare mai mult decât
 o moleculă de ADN corespondentă la o subunitate de proteină, pentru construirea unui
 număr suficient de sonde oligonucleotidice pentru recuperarea tuturor subunităților proteinei
 necesare pentru a obține toxicitate orală.

De la secvența aminoacidă a proteinei purificate se poate izola cu ușurință și dona,
 în totalitate sau parțial, materialul genetic responsabil pentru producerea toxinelor într-un
 vector de expresie, folosind oricare dintre cele câteva tehnici binecunoscute pentru specialistul în biologie moleculară. Un vector de expresie tipic este un plasmid ADN, cu toate că
 alte mijloace de transfer includ, dar nu se limitează la, cosmide, fagemide și fagi, care, de
 asemenea, sunt avute în vedere. În plus la trăsăturile necesare sau dorite pentru replicarea
 plasmidului, cum ar fi o origine a replicării și rezistență la antibiotic sau altă formă a unui marker selectabil, cum ar fi gena bar a *Streptomyces hygroscopicus* sau *viridochromogenes*,
 vectori de expresie a proteinei necesită suplimentar, în mod normal, o casetă de expresie
 care încorporează secvențele care acționează *cis*, necesare pentru transcripția și translația
 genei de interes. Secvențele care acționează *cis*, necesare pentru expresie la procariote,
 diferă de cele necesare la eucariote și plante.

O casetă de expresie eucariotă necesită un promotor transcripțional în amonte (5') față de gena de interes, o regiune de terminare transcripțională, cum ar fi un sit de adiție poli-A și un sit de legare ribozom în amonte primului codon al genei de interes. În celule bacteriene, un promotor transcripțional util care ar putea fi inclus în vector este promotorul legare polimerază ARN 17. Promotori, cum sunt cei descriși anterior aici, sunt cunoscuți pentru inițierea eficientă a transcrierii mARN. De asemenea, în amonte de gena de interes, vectorul poate să includă o secvență nucleotidică care codifică o secvență semnal cunoscută pentru dirijarea unei proteine legate covalent la un compartiment particular al celulelor gazdei, cum ar fi suprafața celulei.

Virusi ai insectelor sau baculovirusi sunt cunoscuți ca infectând și afectând în mod dăunător anumite insecte. Efectul virusurilor asupra insectelor este lent, iar virusurile nu opresc hrănirea insectelor. Astfel, virusurile nu sunt avute în vedere ca fiind utile ca agenți de control al insectelor dăunătoare. Combinând genele toxinelor *Photorhabdus* într-un vector baculovirus s-ar putea asigura un mod eficient de transmitere a toxinelor, crescând în același timp letalitatea virusului. Suplimentar, deoarece baculovirusuri diferite sunt specifice pentru diferite insecte, poate fi posibil să se folosească o toxină particulară pentru țintirea selectivă a anumitor insecte dăunătoare. Un vector util în mod deosebit pentru genele toxinelor este virusul poliedrozei nucleare. Vectori de transfer, care folosesc acest virus, au fost descriși și sunt acum vectorii opțiunii pentru transferarea genelor străine în insecte. Gena recombinantă virus-toxină poate fi construită într-o formă transmisibilă oral. De obicei, baculovirusurile infectează insectele victimă, prin mucoasa intestinală a intestinului mediu. Gena toxinei inserată în spatele unui promotor puternic de proteină de acoperire virală va fi exprimat și va omorî rapid insecta infectată.

Suplimentar, la un sistem de eliberare virus de insectă sau baculovirus, sau plantă transgenică pentru toxinele proteice ale prezentei invenții, proteinele pot fi încapsulate folosind tehnologia de încapsulare *Bacillus thuringiensis*, cum ar fi, dar nelimitat la brevetele **SUA 4 695 455, 4 695 462, 4 861 595**, care toate sunt încorporate aici ca referință. Alt sistem de eliberare pentru toxinele proteice ale prezentei invenții este reprezentat de formularea proteinei într-o matrice momeală, care s-ar putea folosi apoi ca momeli în locurile cu insecte, deasupra solului și sub sol. Exemple de astfel de tehnologii includ, dar nu se limitează la, cererea de brevet **PCT WO/93/23998**, care este încorporată aici ca referință.

Așa cum s-a descris mai sus poate deveni necesar să se modifice secvența care codifică proteina, când se exprimă într-o gazdă non-nativă, deoarece preferințele codonului altor gazde pot să difere de cele ale *Photorhabdus*. Într-un astfel de caz, translația poate fi chiar ineficientă într-o nouă gazdă, dacă nu sunt făcute modificările de compensare la secvența de codificare. Suplimentar, pot fi dorite modificări la secvența aminoacidă pentru a evita reactivitatea încrucișată cu proteinele noii gazde sau pentru a rafina proprietățile insecticide ale proteinei în noua gazdă. O genă a toxinei modificată genetic poate să codifice o toxină, care prezintă, de exemplu, toxicitate sporită sau redusă, dezvoltarea rezistenței insectei modificate, stabilitate modificată sau specificitate modificată pentru speciile țintă.

În plus, la genele *Photorhabdus* care codifică toxinele, conținutul prezentei invenții intenționează să includă secvențe înrudite de acid nucleic care codifică biopolimeri aminoacizi omologi la proteinele toxină și care păstrează efectul toxic al proteinelor *Photorhabdus* la speciile de insecte, după ingestie orală.

De exemplu, toxinele folosite în prezenta invenție par să inhibe inițial hrănirea larvară, înaintea de a urma moartea. Prin manipularea secvențelor de acid nucleic ale toxinelor *Photorhabdus* sau secvențelor sale de control, inginerii geneticieni care plasează gena toxină în plante pot să moduleze potența sa sau modul său de acțiune, pentru ca să mențină,

1 de exemplu, activitatea de inhibare a hrănirii, eliminând în același timp activitatea toxicității
absoluate asupra larvei. Această schimbare ar putea permite plantei transformate să supra-
3 viețuiască până la recoltare, fără a avea un efect dramatic inutil asupra ecosistemului, nimi-
cind toate insectele țintă. Toate aceste modificări ale genei care codifică toxina sau ale pro-
5 teinei codificată de genă sunt avute în vedere să intre în cuprinsul prezentei invenții.

Alte modificări avute în vedere ale acidului nucleic se referă la suplimentarea secven-
7 țelor țintă pentru dirijarea toxinei spre părți anume ale larvei insectei, pentru îmbunătățirea
eficienței sale.

9 Tulpinile ATCC 55397, 43948, 43949, 43950, 43951, 43952 s-au depozitat la colecția
americană de culturi tip 12301 Parklawn Drive, Rockville, MD 20852 SUA. Datele secvenței
11 aminoacide și nucleotidice pentru toxina nativă W-14 (ATCC 55397) sunt prezentate mai jos.
De asemenea, aici, este exemplificată izolarea de ADN genomic pentru toxine din gazde
13 bacteriene. Informații suplimentare pot fi găsite în Sambrook J., Fritsch, E.F., și Maniatis.T.
(1989), *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbor Press, care este încor-
15 porată aici ca referință.

Invenția prezintă următoarele avantaje:

- 17 - reducerea influențelor asupra ecosistemului;
- eficacitate superioară.

19 Se dau, în continuare, exemple de realizare a invenției în legătură și cu fig. 1-6 care
reprezintă:

- 21 - fig. 1 este o ilustrare a unei perechi de izolate ADN donate, folosite ca o parte a sec-
venței genelor pentru toxina prezentei invenții;
- 23 - fig. 2 este o hartă a 3 plasmide folosite în procesul secvențierii;
- fig. 3 este o hartă care ilustrează interrelațiile a câteva fragmente ADN parțiale;
- 25 - fig. 4 este o ilustrare a unei analize de omologie dintre secvențele proteinelor TcbA_i
și TcbA_{ii};
- 27 - fig. 5 este o fenogramă a tulpinilor *Photorhabdus*. Relațiile tulpinilor *Photovhabdus*
s-au definit prin rep-PCR.

29 Axa superioară din fig. 5 măsoară procentul similarității tulpinilor pe baza aprecierii
produselor rep-PCR (adică. 0.0 [fără similaritate] la 1.0 [100% similaritate]). La dreapta axei,
31 numerele și literele indică diversele tulpini testate; i4=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,
33 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 la 52=ATCC 43948 la ATCC 43952. Liniile verticale care separă liniile
35 orizontale arată gradul de înrudire (cum s-a citit din intersecția extrapolată a liniei verticale
cu axa superioară) între tulpini sau grupuri de tulpini, la baza liniilor orizontale (de exemplu,
37 tulpina W-14 este similară aproximativ 60% la tulpinile H9 și Hm).

- fig. 6 este o ilustrare a hărților genomice ale tulpinii W-14.

39 În exemple, sunt folosite următoarele abrevieri: Tris = tris (hidroximetil)amino metan;
SDS = dodecilsulfat de sodiu; EDTA = acid etilendiamintetraacetic; IPTG = izopropiltio-B-
41 galactozidă; X-gal = 5-bromo-4-cloro-3-indoil-B-D-galactozidă; CTAB = bromură cetiltrimetil-
amoniu; kbp= kilobaze perechi; dATP, dCTP, dGTP, dTTP, 1 = 2'-deoxinucleozid 5'-trifosfații
43 adeneinei, citozinei, guaninei, timinei și respectiv inozinei; ATP= adenozin 5'-trifosfat.

45 **Exemplul 1. Purificarea toxinei de la *P. luminescens* și demonstrarea toxicității după
eliberarea orală a toxinei purificate**

47 Toxina proteică insecticidă a prezentei invenții s-a purificat de la *P. luminescens* tul-
pina W-14, ATCC număr de acces 55397. Culturile stoc de *P. luminescens* s-au menținut pe
discuri Petri care conțin 2% Proteose Peptone nr. 3 (adică, PP3, Difco Laboratories, Detroit

RO 121280 B1

MI) în 1,5% agar, s-au incubat la 25°C și s-au transferat săptămânal. Colonii de formă primară ale bacteriilor s-au incubat în 200 ml de bulion PP3 suplimentat cu 0,5% polioxietilen sorbitan mono-stearat (Tween 60, Sigma Chemical Company, St. Louis MO) într-un balon de 1 l. Culturile din bulion s-au crescut 72 h, la 30°C, pe un agitator rotativ. Proteinele toxine se pot recupera din culturi crescute în prezență sau absență de Tween; totuși, absența Tween poate afecta forma bacteriei crescute și profilul proteinelor produse de către bacterii. În absență de Tween, se întâlnește o variantă în măsura în care greutatea moleculară a cel puțin unei subunități de toxină identificată se schimbă de la circa 200 kDa la circa 185 kDa.	1
Culturile de 72 h s-au centrifugat la 10.000 x g, timp de 30 min, pentru a îndepărta celulele și resturile celulare. Frațiunea supernatantului, care a conținut activitatea insecticidă, s-a decantat și s-a adus la 50 mM K ₂ HPO ₄ , adăugând un volum corespunzător de 1,0 M K ₂ HPO ₄ . S-a ajustat pH-ul la 8,6 adăugând hidroxid de potasiu. Această fracțiune supernatantă s-a amestecat apoi cu DEAE-Sephacel (Pharmacia LKB Biotechnology), care s-a echilibrat cu 50 mM K ₂ HPO ₄ . Activitatea toxică s-a adsorbit la rășina DEAE. Acest amestec s-a turnat apoi într-o coloană 2,6 X 40 cm și s-a spălat cu 50 mM K ₂ HPO ₄ la temperatura camerei, la o rată de curgere de 30 ml/h, până când efluentul a atins o absorbantă UV la o linie de bază constantă la 280 nm. Apoi, coloana s-a spălat cu KC1 150 mM, până când efluentul a atins din nou o linie de bază constantă la 280 nm. În final, coloana s-a spălat cu KC1 300 mM și fracțiunile s-au colectat.	3
Fracțiunile conținând toxina s-au colectat și s-au sterilizat prin filtrare, folosind o membrană de filtrare cu pori 0,2 μ. Apoi, toxina s-a concentrat și s-a echilibrat la 100 mM KPO ₄ , pH 6,9, folosind o membrană de ultrafiltrare cu separarea unei greutate moleculare de 100 kDa, la 4°C (Centriprep 100, Amicon Division-W.R. Grace and Company). S-a aplicat o probă de 3 ml de concentrat toxină, la partea superioară a unei coloane de filtrare în gel 2,6 x 95 cm Sephacryl S-400 HR (Pharmacia LKB Biotechnology). Tamponul efluent a fost KPO ₄ 100 mM, pH 6,9, care s-a migrat la o rată de curgere 17 ml/h, la 4°C. Efluentul s-a monitorizat la 280 nm.	5
S-au colectat fracțiunile și s-au testat pentru activitate toxică. Toxicitatea fracțiunilor cromatografice a fost examinată într-o analiză biologică, folosind larve de <i>Manduca sexta</i> . Fracțiunile s-au aplicat fie direct pe hrana insectei (Gypsy, hrană granulată din germen de grâu, ICN Biochemicals Division-ICN Biomedicals, Inc.), fie s-au administrat prin injectarea intrahemocelică a unei probe de 5 μl în prima protuberanță abdominală a unei larve în stadiul al 4-lea sau al 5-lea, folosind un ac de mărimea 30. S-a înregistrat greutatea fiecărei larve din grupul de tratament, la intervale de 24 h. Se asumă toxicitatea, dacă insecta a încetat hrănirea și a murit în câteva zile de consum al dietei, sau dacă moartea a avut loc în 24 h după injectarea unei fracțiuni.	7
Fracțiunile toxice s-au colectat și s-au concentrat folosind Centriprep 100 și apoi s-au analizat prin HPLC folosind o coloană cu gel de permeație de 7,5 mm x 60 cm TSK-GEL G-4000 SW cu fosfat de potasiu 100 mM, pH 6,9 tampon eluent care migrează la 0,4 ml/min. Această analiză a arătat toxina proteică ca fiind conținută într-un singur maxim ascuțit care a eluat din coloană cu un timp de retenție de aproximativ 33,6 min. Acest timp de retenție a corespuns la o greutate moleculară estimată de 1000 kDa. Maximul fracțiunilor a fost colectat pentru purificare ulterioară, în timp ce fracțiunile care nu conțin această proteină s-au eliminat. Maximul eluat de la HPLC absoarbe lumina UV la 218 și 280 nm, dar nu absoarbe la 405 nm. Absorbanta la 405 nm s-a dovedit a fi un atribut al compușilor antibiotici xenorhabdină.	9
Electroforeza fracțiunilor maxime colectate într-un gel de agaroză non-denaturată (Metaphor Agarose, FMC BioProducts) a arătat că în maxim sunt prezente 2 complexe	11

proteice. Materialul maximului tamponat în Tris-HCl 50 mM, pH 7,0 s-a separat pe un gel suprapus de agaroză 1,5% tamponat cu Tris-HCl 100 mM la pH 7,0 și gelul care separă agaroză 1,9% s-a tamponat cu Tris-borat 200 mM la pH 8,3, în condiții de tamponare standard (tampon anod Tris-HCl 1M, pH 8,3; tampon catod Tris 0,025 M, glicină 0,92 M). Gelurile au migrat la curent constant 13 niA, la 15°C, până când colorantul de trasare roșu fenol a atins capătul gelului. S-au vizualizat în gelurile de agaroză 2 benzi proteice, folosind colorație *Coomassie brilliant blue*.

Banda cu migrare mai lentă s-a denumit "banda proteică 1" și banda cu migrare mai rapidă s-a denumit "banda proteică 2". Cele două benzi proteice au fost prezente în cantități aproximativ egale. Gelurile de agaroză colorate *Coomassie* s-au folosit ca ghid pentru a exciza precis cele două benzi proteice din porțiunile necolorate ale gelurilor. Bucățile excizate conținând benzile proteice s-au macerat și s-a adăugat o cantitate mică de apă sterilă. Pentru control, s-a excizat o porțiune a gelului care nu a conținut nici o proteină și s-a tratat în aceeași manieră ca bucățile de gel care conțin proteina. S-a recuperat proteina din bucățile de gel prin electroeluție în Tris-borat 100 mM, pH 8,3, la 100 volți (voltaj constant), timp de 2 h. Alternativ, s-a eluat proteina pasiv din bucățile de gel, prin adăugarea unui volum egal de Tris-HCl 50 mM, pH 7,0, la bucățile de gel, apoi incubare la 30°C, 16 h. Aceasta a permis proteinei să difuzeze din gel în tampon care apoi s-a colectat.

Rezultatele testelor de toxicitate la insecte folosind toxină purificată HPLC (33,6 min, maxim) și toxina purificată din gel de agaroză au demonstrat toxicitatea extractelor. Injectarea a 1,5 μg de proteină purificată HPLC omooară în 24 h. Ambele benzi proteice 1 și 2, recuperate din gelul de agaroză prin eluție pasivă sau electroeluție, au fost letale la injectare. Concentrația proteinei estimată pentru aceste probe a fost mai mică de 50 ng/larvă. O comparație a greutateii câștigate și mortalitatea între grupurile de larvă, injectate cu benzile proteice 1 sau 2, arată că proteina banda 1 a fost mai toxică la eliberare prin injectare.

Când toxina purificată HPLC s-a aplicat la hrana larvară, la o concentrație de 7,5 μg/larvă, aceasta a produs o stopare a câștigului greutateii larvare (s-au testat 24 de larve). Larvele au început să se hrănească, dar după consumarea numai unei mici părți din hrana tratată cu toxină, ele au început să prezinte simptomele patologice induse de toxină și larvele au încetat să se hrănească. Insecta devine incoloră și majoritatea larvelor prezintă semne de diaree. Mortalitatea semnificativă a insectei a rezultat când s-au aplicat câteva doze de 5 μg toxină la hrană, pe o perioadă de 7-10 zile.

Banda 1 de proteină separată pe agaroză a inhibat semnificativ câștigul greutateii larvare la o doză 200 ng/larvă. Larvele hrănite cu concentrații similare ale proteinei banda 2 nu s-au inhibat și au câștigat în greutate la aceeași rată ca larvele de control. S-au hrănit 12 larve cu proteină eluată și 45 larve s-au hrănit cu proteina conținută în bucățile de agaroză. Aceste două seturi de date arată că proteina din banda 1 a fost toxică oral pentru *Manduca sexta*. În acest experiment, se pare că proteina din banda 2 nu a fost toxică pentru *Manduca sexta*.

Analizele suplimentare ale proteinei din benzile 1 și 2 prin SDS-PAGE, în condiții de denaturare, a arătat că fiecare bandă a fost compusă din câteva subunități proteice mai mici. Proteinele s-au vizualizat prin colorație *Coomassie brilliant blue*, urmată de colorație argint, pentru a atinge maximul de sensibilitate.

Subunitățile de proteină din cele două benzi au avut o similaritate foarte ridicată. Banda 1 de proteină conține 8 proteine de 25,1, 56,2, 60,8, 65,6, 166, 171, 184 și 208 kDa. Banda 2 de proteină a avut un profil identic, cu excepția că proteinele de 25,1, 60,8 și 65,5 kDa nu au fost prezente. Proteinele 56,2, 60,8, 65,6 și 184 kDa au fost prezente în complexul banda 1 de proteină la concentrații aproximativ egale și reprezintă 80% sau mai mult din conținutul total de proteină a acestui complex.

RO 121280 B1

Toxina nativă purificată-HPLC s-a caracterizat în continuare, după cum urmează. Toxina a fost labilă termic, prin aceea că, după ce a fost încălzită la 60°C, timp de 15 min, aceasta și-a pierdut capacitatea de a omorî sau de inhiba creșterea în greutate, când s-au injectat sau s-au hrănit larve *M.sexata*. Testele au avut ca scop detectarea activităților lipazei, fosfolipazei tip C, nucleazelor sau hemolizei roșii de sânge și s-au realizat cu toxină purificată. Nici una din aceste activități nu a fost prezentată. De asemenea, s-au făcut teste de inhibare antibiotic zonă și toxina purificată nu a reușit să inhibe creșterea bacteriilor gram-negative sau gram-pozitive, drojdii sau fungi filamentoși, arătând că toxicul nu este un antibiotic xenorhabdin.

Toxina nativă purificată - HPLC - s-a testat pentru capacitatea de a omorî alte insecte decât *Manduca sexta*. Tabelul 2 conține denumirile insectelor omorâte de către toxina *P. luminescens* purificată-HPLC în acest studiu.

Tabelul 2

Insecte omorâte de către toxina *P. luminescens*

Nume comun	Ordin	Genuri sau specii	Cale de eliberare
Molia tutunului	<i>Lepidoptera</i>	<i>Manduca sexta</i>	Oral sau injectat
Gândac de făină	<i>Coleptera</i>	<i>Tenebrio molitor</i>	Oral
Pharaoh ant	<i>Hymenoptera</i>	<i>Monomorium pharoanis</i>	Oral
Gândac de bucătărie	<i>Dictyoptera</i>	<i>Pleettella germanica</i>	Oral sau injectat
Țânțar	<i>Diptera</i>	<i>Aedes aegypti</i>	Oral

Exemplul 2. Utilizarea și toxicitatea *Photorhabdus luminescens* s-au caracterizat în continuare. S-a produs bulionul de cultură *Photorhabdus luminescens* (tulpina W-14), așa cum urmează. Mediul de producție a fost 2% Bacto Proteose Peptone® nr. 3 (PP3, Difco Laboratories, Detroit, Michigan) în apă deionizată Milli-Q®. Flacoanele culturi de însămânțare constau din 175 ml mediu plasat într-un flacon cu trei șicane cu un gât Delong, acoperit cu un Kaput și autoclavizat timp de 20 min, T=250°F. Flacoanele de producție constau din 500 ml din 2,8 l 500 ml într-un flacon cu trei șicane cu un gât Delong, acoperit de o spumă silicon Shin-etsu. Acestea s-au autoclavizat timp de 45 min, T=250°F. Cultura de însămânțare s-a incubat la 28°C, la 150 rpm, într-un incubator giratoriu de agitare cu o deplasare de 2 inch. După 16 h de creștere, 1% din cultura de însămânțare s-a plasat într-un flacon de producție care s-a lăsat să crească timp de 24 h înainte de recoltare. Producția toxinei a părut să fie în timpul creșterii fază log. Bulionul microbial s-a transferat la o sticlă de centrifugă de 11 și biomasa celulară s-a peletat (30 min, la 2500 RPM, la 4°C, centrifugă [R.C.F.= ~1600] HG-4L Rotor RC3 Sorval, Dupont, Wilmington, Delaware). Bulionul primar s-a răcit brusc la 4°C, timp de 8-16 h, și s-a recentrifugat cel puțin 2 h (condiții de mai sus), pentru a clarifica suplimentar bulionul prin îndepărtarea unei așa-zise mucopolizaharide care a precipitat la repaus. (O metodă alternativă de producere a combinat ambele etape și a implicat utilizarea unei centrifugări de clarificare timp de 16 h, aceleași condiții de mai sus). Apoi, acest bulion s-a păstrat la 4°C, înainte de biotestare sau filtrare.

Bulionul de cultură *Photorhabdus* și toxina(ele) proteică care s-a purificat din acest bulion au prezentat activitate (mortalitatea și/sau inhibarea creșterii, apariția redusă a adultului) împotriva unui număr de insecte. Mai ales activitatea este împotriva corn root worm (larvă sau adult), gândacului de Colorado al cartofului și turf grub, care sunt membrii ai ordinului

1 de insecte *Coleoptera*. Alți membri ai *Coleoptera* includ specii de wireworm, gândac al pole-
 3 nului, gândaci din genurile *Haltica* și *Epitrix*, gândaci din familia *Bruchidae* și gărgărițe. De
 asemenea, s-a observat activitate împotriva insectelor homeoptere *Cicadellidae*, care sunt
 5 membre ale ordinului *Homoptera*. Alți membri ai *Homoptera* includ planthoppers, pear psylla,
 apple sucker, insecte homeoptere *Ciccoidea*, *Aleyrodidae* și *Cercopidae*, precum și nume-
 7 roase gazde specifice speciilor afide. Bulionul și fracțiunile purificate sunt de asemenea
 active împotriva beet armyworm, cabbage looper, *Noctuidae*. *Pyransta nubilalis* *Heliothis zea*
 și *Crpocapsa pomonella*, care sunt membrii ai ordinului *Lepidoptera*. Alți membri tipici ai
 9 acestui ordin sunt *Tineidae*, Indian mealmoth, leaf roller, genul *Pieris*, cotton bollworm,
 bagworm, Eastern tent Caterpillar, sod webworm și fall armyworm. De asemenea, activitatea
 11 pare să fie împotriva larvelor de țânțari, care sunt membrii ordinului *Diptera*. Alți membri ai
 ordinului *Diptera* sunt *Chironomidae*, carrot fly, cabbage root fly, turnip root fly, onion fly,
 13 *Tipulidae*, *Musca domestica* și diferite specii de țânțari. Activitatea pare să fie împotriva
 insectelor diptere din familia *Camponotus* și *Iridomyrmex humilis*, care sunt membrii ai ordi-
 15 nului care include de asemenea *Lampyridae*, oderous house ant și *Monomorium minimum*.

Bulionul/fracția este folositor pentru reducerea populațiilor de insecte și s-au folosit
 17 într-o metodă a inhibării unei populații de insecte. Metoda poate cuprinde aplicarea, la un lo-
 cus al insectei, a unei cantități eficiente de inactivare a insectei de activ descris. Rezultatele
 19 sunt prezentate în tabelul 3.

Activitatea împotriva lavelor corn root worm s-a testat după cum urmează. Bulion de
 21 cultură *Photorhabdus* (sterilizat pe filtru, celule libere) sau fracțiuni purificate HPLC s-au apli-
 cat direct pe suprafața (~ 1,5 cm²) 0,25 ml de hrană artificială în 30 μ l alicote, după diluția
 23 în mediu de control sau tampon 10 mM fosfat de sodiu, respectiv pH 7,0. Plăcile dietă s-au
 lăsat la uscare în aer, într-o învelitoare-curent sterilă, și godeurile s-au infestat cu o singură
 25 nou apărută *Diabrotica undecimpunctata howardi* (Southern corn rootworm, SCR) incubată
 din ouă sterilizate, cu a doua stare larvară SCR crescută pe dietă artificială sau cu a doua
 27 stare larvară *Diabrotica virgifera virgifera* (Western corn rootworm, WCR), crescute pe răsa-
 duri de cereale în Metromix®. S-au cântărit larve în a doua stare larvară, înainte a adăugarea
 29 la hrană. Plăcile s-au izolat, s-au plasat într-o cameră de creștere umezită și s-au menținut
 la 27°C, pentru o perioadă corespunzătoare (4 zile pentru SCR nou apărut și adult, 2-5 zile
 31 pentru larve WCR, 7-14 zile pentru SCR în a doua stare embrionară). S-au măsurat deter-
 minările de greutate și mortalitate, așa cum s-a indicat. În general, în toate studiile s-au utili-
 33 zat 16 insecte per tratament. Mortalitățile de control au fost următoarele: larve nou apărute,
 < 5%, gândaci adulți, 5%.

Activitatea împotriva gândacului de Colorado al cartofului s-a testat după cum urmea-
 35 ză. Bulion de cultură *Photorhabdus* sau mediul de control s-a aplicat la suprafața (~2,0 cm²)
 37 1,5 ml dietei artificiale standard considerată în godeurile unei plăci de cultură țesut cu 24
 godeuri. Fiecare godeu a primit 50 μ l tratament și s-a lăsat să se usuce la aer. Apoi, larve
 39 individuale de gândac de Colorado al cartofului, în al doilea stadiu larvar (*Leptinotarsa*
decemlineata, CPB), s-au plasat pe hrană și, după 4 zile, s-a înregistrat mortalitatea.
 41 Mortalitatea control a fost 3,3%.

Activitatea împotriva Japanese beetle grubs & beetles s-a testat după cum urmează.
 43 Brazde de pământ cu larve turf grub (*Popillia japonica*, al 2-3-lea stadiu larvar) s-au colectat
 de pe pajiști infestate și s-au menținut în laborator, în amestec sol/turbă cu felii de morcov
 45 adăugat ca hrană suplimentară. Gândacii truf s-au atras local în capcană-feromon și s-au
 menținut în laborator în containere de plastic cu frunze de arțar ca hrană. După aplicare de
 47 bulion de cultură nediluat *Photorhabdus* sau mediu de control la dieta artificială corn root
 worm (30 μ l/1,54 cm², gândaci) sau felii morcov (larve), ambele stadii s-au plasat individual

RO 121280 B1

într-un godeu dietă și s-au observat orice mortalitate și alimentație. În ambele cazuri, a existat o reducere clară în cantitatea hranei (și producerea de excremente) observată.	1
Activitatea împotriva larvelor țânțarilor s-a testat cum urmează. Testul s-a realizat în plăci de microtitrare cu 96-godeuri. Fiecare godeu a conținut 200 μl soluție apoasă (bulion de cultură <i>Photorhabdus</i> , mediu de control sau H ₂ O și aproximativ 20 larve în vârstă de 1 zi (<i>Aedes aegypti</i>). Au fost 6 godeuri per tratament. Rezultatele s-au citit la 2 h după infestare și nu s-au schimbat după perioada de observare de trei zile. Nu a apărut mortalitate control.	3
Activitatea împotriva insectelor diptere din familia <i>Trypetidae</i> s-a testat după cum urmează. Mediul <i>Drosophila melanogaster</i> cumpărat s-a preparat folosind 50% mediu uscat și 50% un lichid, fie apă, fie mediu control, fie bulion cultură <i>Photorhabdus</i> . Aceasta a fost realizată prin plasarea a 8,0 ml mediu uscat în fiecare dintre 3 flacoane ridicate per tratament și adăugând 8,0 ml de lichid corespunzător. Apoi zece larve în stadii larvare târzii de <i>Drosophila melanogaster</i> s-au adăugat la fiecare flacon. Flacoanele s-au ținut pe o masă de laborator, la temperatura camerei, sub lumină fluorescentă din plafon. Numărarea adulți și pupe s-a făcut după 3, 7 și 10 zile de expunere. Încorporarea bulionului de cultură <i>Photorhabdus</i> în mediul de dietă pentru larve de <i>Trypetidae</i> a produs o ușoară (17%) dar semnificativă reducere în apariția de adulți, în ziua -10, cum s-a comparat la apă sau mediu control (reducere 3%).	5
Activitatea împotriva insectelor homeoptere Cicadellidae s-a testat după cum urmează. Testul ingestiei pentru insecte homeoptere <i>Cicadellidae</i> (<i>Macrosteles severini</i>) este destinat a face posibilă ingerarea activului fără alt contact extern. Rezervorul pentru soluția activ/"aliment" este realizat prin 2 lăcașuri în centrul porțiunii de jos a unui vas Petri 35x10 mm. Un pătrat Parafilm M® de 2 inch este plasat peste vârful vasului și se asigură cu un inel "O". Un cub de plastic de 1 oz. s-a infestat apoi cu aproximativ 7 insecte homeoptere <i>Cicadellidae</i> și rezervorul s-a plasat la vârful cupei, sub Parafilm. Soluția test s-a adăugat apoi la rezervor prin lăcașuri. În teste utilizând bulion de cultură <i>Photorhabdus</i> nediluat, bulionul și mediul control s-au dializat contra apă, pentru a reduce mortalitatea control. Mortalitatea s-a raportat la 2 zile, când mortalitatea control a părut a fi 26,5%. În teste utilizând fracțiuni purificate (200 mg proteină/ml), o concentrație finală de sucroză 5% s-a utilizat în toate tratamentele pentru a îmbunătăți supraviețuirea insectelor homeptere <i>Cicadellidae</i> . Testul s-a desfășurat într-un incubator la 28°C, 70% RH cu o fotoperioadă de 16/8. Testul a atins gradul de mortalitate la 72 h. Mortalitatea control a fost 5,5%.	7
Activitatea împotriva insectelor <i>Iridomyrmex humilis</i> s-a testat după cum urmează. Un alicot de 1,5 ml de bulion de cultură <i>Photorhabdus</i> 100%, mediu control sau apă, s-a pipetat în flacoane de sticlă transparentă de 2,0 ml. Flacoanele s-au vârat într-o bucată de fitil dentar de bumbac, care s-a umezit cu tratamentul corespunzător. Fiecare flacon s-a plasat într-un vas Petri separat 60x16 mm cu 8 la 12 insecte adult <i>Iridomyrmex humilis</i> (<i>Linepithema humile</i>). Au fost replicate trei per tratament. Plăcile de biotest s-au ținut pe o masă de laborator, la temperatura camerei, sub lumină fluorescentă din plafon. Citirile pentru mortalitate s-au făcut după 5 zile de expunere. Mortalitatea de control a fost 24%.	9
Activitatea împotriva insectelor din genul <i>Camponotus</i> s-a testat după cum urmează. S-au colectat femele nefertile de insecte (<i>Camponotus pennsylvanicus</i>) din copaci de pe plantația DowElanco în Indianapolis, IN. Testele cu bulion cultură <i>Photorhabdus</i> s-au realizat după cum urmează. Fiecare container biotest de plastic (7 1/8"x3") a ținut 15 femele infertile, un adăpost de hârtie și 10 ml bulion sau mediu control într-un vas mic de plastic. Un fitil de bumbac a eliberat tratamentul, până la un lăcaș în capacul vasului mic. Toate tratamentele au conținut 5% sucroză. Biotestul s-a realizat la întuneric, la temperatura camerei, timp de 19 zile. Mortalitatea control a fost 9%. Testele care eliberează fracțiuni purificate utilizează	11
	13
	15
	17
	19
	21
	23
	25
	27
	29
	31
	33
	35
	37
	39
	41
	43
	45
	47

dietă artificială pentru furnici amestecată cu tratamentul (fracțiune purificată sau soluție control) la o rată de 0,2 ml tratament/2,0 g dietă într-o eprubetă test plastic. Concentrația finală de proteină a fracțiunii purificate a fost mai mică decât 10 µg/g dietă. Zece insecte per tratament, o sursă de apă, adăpost și dieta tratată s-au plasat în containere de plastic etanșe și s-au menținut la întuneric, la 27°C, într-un incubator umezit. S-a înregistrat mortalitatea în ziua 10. Nu a apărut mortalitate control.

Activitatea împotriva diferitelor larve lepidoptere s-a testat după cum urmează. Bulion de cultură *Photorhabdus* sau fracțiuni purificate s-au aplicat direct la suprafața (~1,5 cm²) de 0,25 ml dietei artificiale standard în 30 µl alicote după diluție, în mediu control sau tampon 10 mM fosfat de sodiu, pH 7,0, respectiv. Plăcile de dietă s-au lăsat să se usuce la aer într-o învelitoare-curent sterilă și godeurile s-au infestat cu larve nou ivite individual. Din surse comerciale s-au suplimentat ouă European corn borer (*Ostrinia nubilalis*) și corn earworm (*Helicoverpa zea*) și s-au incubat cu mijloace proprii și s-au suplimentat intern beet armyworm (*Spodoptera exigua*), cabbage looper (*Trichoplusia ni*), tabacco budworm (*Heliothis virescens*), codling moth (*Laspeyresia pomonella*) și black cutworm (*Agrotis ipsilon*). După infestare cu larve, plăcile cu dietă s-au etanșat, s-au plasat într-o cameră de creștere umezită și s-au menținut la întuneric, la 27°C, pentru perioada corespunzătoare. S-au înregistrat determinările de greutate și mortalitate la zilele 5-7 pentru bulion de cultură *Photorhabdus* și 4-7 zile pentru fracția purificată. În tot studiul s-au folosit, în general, 16 insecte per tratament. Mortalitatea control s-a situat între 4-12,5% pentru mediu control și a fost mai mică decât 10% pentru tampon fosfat.

Tabelul 3

Efectul bulionului de cultură *Photorhabdus luminescens* (Tulpina W-14) și fracțiunii toxine purificate, asupra mortalității și inhibării creșterii a diferite ordine/specii de insecte

Ordin/specie de insecte	Bulion		Fracțiune purificată	
	Mort. %	G.I. %	Mort. %	G.I. %
COLEOPTERA				
Vierme rădăcini cereale				
Southern/larvă nou apărută	100	na	100	na
Southern/al 2-lea stadiu larvar	na	38,5	nt	nt
Southern/adult	45	nt	nt	nt
Western/al 2-lea stadiu larvar	na	35	nt	nt
Gândacul de Colorado al cartofului				
al 2-lea stadiu larvar	93	nt	nt	nt
Vierme de gazon				
al 3-lea stadiu larvar	na	a.f	nt	nt
adult	na	A.f	nt	nt
DIPTERA				
Tripedidae (apariție adult)	17	nt	nt	nt
Larvă țânțar	100	na	nt	nt
HOMOPTERA				
Cicadellidae	96,5	na	100	na

Tabelul 3 (continuare)

Ordin/specie de insecte	Bulion		Frațiuni purificate	
	Mort. %	G.I. %	Mort. %	G.I. %
HIMENOPTERA				
Iridomyrmex humilis	75	na	nt	na
Camponotus	71	na	100	na
LEPIDOPTERA				
Beet Armyworm	12,5	36	18,75	41,4
Black Cutworm	nt	nt	0	71,2
Cabbage Looper	nt	nt	21,9	66,8
Codling Moth	nt	nt	6,25	45,0
Corn Earworm	56,3	94,2	97,9	na
European Corn Borer	96,7	98,4	100	na
Tabacco Budworm	13,5	52,5	19,4	85,6

Mort.= mortalitate, G.I.= inhibare creștere, na= neaplicabil,
nt= netestat, a.f.= anti-hrănitor

Exemplul 3. Utilitatea insecticidului pentru aplicare în sol

Bulionul de cultură *Photorhabdus luminescens* (tulpina W-14) s-a arătat a fi activ împotriva corn rootworm, când s-a aplicat direct pe sol sau pe un sol-mix (Metromix®). Activitatea împotriva SCR și WCR nou apărut în Metromix® s-a testat după cum urmează (Tabelul 4). Testul s-a realizat utilizând răsaduri de cereale (United Agriseeds brand CL614), care au germinat la lumină, pe hârtie filtru umed pentru 6 zile. Când rădăcinile au fost aproximativ de 3-6 cm lungime, o singură sămânță/răsad s-a plantat într-o cupă de plastic clar cu 50 mg de Metromix® uscat. Apoi, 20 SCR sau WCR nou apăruti s-au plasat pe rădăcinile răsadurilor și s-au acoperit cu Metromix®. Pe infestare, răsadurile s-au înmuiat apoi cu 50 ml volum total de soluție bulion diluat. După înmuiere, cupele s-au etanșat și s-au lăsat la temperatura camerei, la lumină, timp de 7 zile. După aceea, răsadurile s-au spălat pentru a îndepărta tot Metromix® și rădăcinile s-au excizat și s-au cântărit. Activitatea s-a evaluat ca procentul rădăcinilor de cereale rămase relativ la plantele de control și ca distrugerea frunzei indusă prin hrănire. S-a înregistrat vizual distrugerea frunzei și s-a evaluat fie ca -, +, ++, sau +++, cu "-" reprezentând lipsa distrugerii și "+++" reprezentând distrugere gravă.

Activitatea împotriva SCR nou apărut în sol s-a testat după cum urmează (Tabelul 5). Testul s-a realizat utilizând răsaduri de cereale (United Agriseeds brand CL614) care au germinat la lumină, pe hârtie de filtru umedă, timp de 6 zile. După ce rădăcinile au fost de aproximativ 3-6 cm lungime, o singură sămânță/răsad s-a plantat într-o cupă de plastic clar cu 150 gm sol dintr-un câmp din Lebanon, IN, plantat cu un an înainte cu cereale. Acest sol nu a fost tratat anterior cu insecticide. Apoi, 20 SCR nou apăruti s-au plasat direct pe rădăcinile răsadurilor și s-au acoperit cu sol. După infestare, răsadurile s-au îmbibat cu 50 ml volum total dintr-o soluție de bulion diluat. După înmuiere, cupele neetanșate s-au incubat într-o cameră cu umiditate relativ ridicată (80%) la 78°F. După aceea, răsadurile s-au spălat pentru a îndepărta tot solul și rădăcinile s-au excizat și s-au cântărit. Activitatea s-a evaluat ca procentul rădăcinilor de cereale rămase relativ la plantele control și ca distrugere frunză indusă prin hrănire. S-a înregistrat vizual distrugerea frunzei și s-a evaluat fie ca -, +, ++, sau +++, cu "-" reprezentând fără distrugere și "+++" reprezentând distrugere gravă.

Tabelul 4

Efectul bulionului de cultură *Photorhabdus luminescens* (Tulpina W-14) asupra larvei rootworm după înmuiere postinfestare (Metromix®)

Tratament	Larvă	Distrugere frunză	Greutate rădăcină	%
Southern corn rootworm				
Apă	-	-	0,4916 ± 0,023	100
Mediu (2,0% v/v)	-	-	0,4416 ± 0,029	100
Bulion (6,25% v/v)	-	-	0,4641 ± 0,081	100
Apă	+	+++	0,1410 ± 0,006	28,7
Mediu (2,0% v/v)	+	+++	0,1345 ± 0,028	30,4
Bulion (1,56% v/v)	+	-	0,4830 ± 0,031	104
Western corn rootworm				
Apă	-	-	0,4446 ± 0,019	100
Bulion (2,0% v/v)	-	-	0,4069 ± 0,026	100
Apă	+	-	0,2202 ± 0,015	49
Bulion (2,0% v/v)	+	-	0,3879 ± 0,013	95

Tabelul 5

Efectul bulionului de cultură *Photorhabdus luminescens* (Tulpina W-14) asupra Southern corn root după înmuiere postinfestare (Sol)

Tratament	Larvă	Distrugere frunză	Greutate rădăcină	%
Apă	-		0,2148 ± 0,014	100
Bulion (50% v/v)	-	-	0,2260 ± 0,016	103
Apă	+	+++	0,0916 ± 0,009	43
Bulion (50% v/v)	+	-	0,2428 ± 0,032	113

Activitatea bulionului de cultură *Photorhabdus luminescens* (tulpina W-14) împotriva larvelor turf grub în al doilea stadiu larvar în Metromix® s-a observat în teste realizate după cum urmează (tabelul 6). Aproximativ 50 gm de Metromix® s-a adăugat la o cupă de plastic clar de 591 ml. Metromix®, s-a înmuiat apoi cu 50 ml volum total dintr-o soluție de bulion *Photorhabdus* diluat 50% (v/v). Diluția bulionului brut s-a făcut cu apă, cu 50% bulion, fiind preparat prin adăugarea a 25 ml bulion brut la 25 ml apă pentru un volum total de 50 ml. O soluție 1% (g/v) de proteose peptone #3 (PP3), care este o diluție 50% a concentrației mediului normal, s-a utilizat ca un control bulion. După înmuiere, cinci larve turf grub, în al doilea stadiu larvar, s-au plasat pe vârful Metromix® umezit. O mulțime de larve turf grub în al doilea stadiu au săpat rapid în Metromix®. Acele larve care nu au săpat într-o oră, s-au îndepărtat și s-au înlocuit cu larve noi. Cupele s-au etanșat și s-au plasat într-un incubator la 28°C, la întuneric. După șapte zile, larvele s-au îndepărtat din Metromix® și s-a înregistrat mortalitatea. Activitatea s-a evaluat prin procentul mortalității relativ la control.

Tabelul 6

Efectul bulionului de cultură *Photorhabdus luminescens* (Tulpina W-14) asupra larvei turf grub după înmuiere postinfestare (Metromix®)

Tratament	Mortalitate*	Mortalitate %
Apă	7/15	47
Mediu control (1,0% g/v)	12/19	63
Bulion (50% v/v)	17/20	85

* - exprimată în raportul larvelor moarte pe insectele vii

Exemplul 4. Utilitatea insecticidului la aplicarea pe frunză

Activitatea bulionului *Photorhabdus* împotriva European corn borer s-a observat când bulionul s-a aplicat direct pe suprafața frunzelor de porumb (tabelul 7). În aceste teste, bulionul *Photorhabdus* s-a diluat de 100 ori cu mediu de cultură și s-a aplicat manual la suprafața frunzelor de porumb excizate la o rată de ~6,0 $\mu\text{l}/\text{cm}^2$ pe suprafața frunzei. Frunzele s-au uscat la aer și s-au tăiat fâșii de mărime egală de aproximativ 2x2 inch. Frunzele s-au rulat, s-au asigurat cu agrafe de hârtie și s-au plasat în vase mici de plastic de 1 oz cu 0,25 inch de agar 2% pe suprafața de sus, pentru a preveni umezirea. S-au plasat 20 European corn borer nou apăruiți pe frunzele rulate și cupa s-a sigilat. După incubare pentru 5 zile, la 27°C, la întuneric, s-au înregistrat probele, distrugerea hranei și larvele recuperate.

Tabelul 7

Efectul bulionului de cultură *Photorhabdus luminescens* (Tulpina W-14) asupra larvei European corn borer după aplicarea preinfestării pentru a exciza frunzele de porumb

Tratament	Vătămare frunză	Larve recuperate	Greutate (mg)
Apă	întinsă	55/120	0,42 mg
Mediu control	întinsă	40/120	0,50 mg
Bulion (1,0% v/v)	urme	3/120	0,15 mg

Activitatea bulionului de cultură împotriva *Heliothis virescens* s-a demonstrat folosind o metodologie a frunzei îndoite. Frunze de bumbac proaspete s-au excizat de la plante și discurile frunzei s-au tăiat cu un sfredel-dop de 18,5 mm. Discurile s-au imersat individual în mediul de control (PP3) sau bulion de cultură *Photorhabdus luminescens* / tulpina W-14) care s-a concentrat de aproximativ 10 ori, folosind un sistem de filtrare tangențial Amicon (Beverly, MA), Proflux M 12 cu un filtru de 10 kDa. S-a îndepărtat excesul de lichid și o agrafă de hârtie care întărește s-a plasat către centrul discului. Agrafta de hârtie s-a fixat apoi într-un vas mic de 1,0 oz de plastic, care conține aproximativ 2,0 ml agar 1%. Aceasta servește pentru a suspenda discul frunzei deasupra agarului. După uscarea discului de frunză, s-a plasat pe disc o singură larvă nou apărută și cupa s-a etanșat. Apoi, cupele s-au sigilat într-un sac de plastic și s-au plasat într-un incubator întunecos, 27°C, timp de 5 zile. La această perioadă larvele și materialul din frunze rămas s-au cântărit pentru a stabili o măsură a distrugerii frunzei (tabelul 8).

Tabelul 8

Efectul bulionului de cultură *Photorhabdus luminescens* (tulpina W-14) asupra *Heliothis virescens* nou apărute într-un test al frunzei de bumbac îndoită

Tratament	Discul frunzei	Greutatea finală a larvelor (mg)
Frunze control	55,7 ± 1,3	na*
Mediu control	34,0 ± 2,9	4,3 ± 0,91
Bulion <i>Photorhabdus</i>	54,3 ± 1,4	0,00

* - neaplicabil

** - nu s-au găsit larve în viață

Exemplul 5. Partea A Caracterizarea componentelor peptidei toxine

Într-o analiză ulterioară, subunitățile proteinei toxine ale benzilor izolate ca în exemplul 1 s-au descompus printr-o electroforeză pe un gel de 7% SDS poli(acrilamidă) cu un raport 30:0,8 (acrilamidă: BIS-acrilamidă). Acest gel matrice favorizează mai bine rezoluția proteinelor mai mari. Sistemul de gel, folosit în exemplul 1 pentru a estima greutatea moleculare ale subunităților Banda 1 și Banda 2, a fost un gel 18% cu un raport 38:0,18 (acrilamidă:BIS-acrilamidă), care permite un domeniu larg al separării după mărime, dar rezoluție mai scăzută a componentelor cu greutate moleculară mai ridicată.

În această analiză, s-au descompus 10, mai degrabă decât 8, benzi de proteină. Tabelul 9 prezintă greutatea moleculară calculată ale celor 10 benzi descompuse și compară direct greutatea moleculară estimată în aceste condiții, cu cele din exemplul anterior. Nu este surprinzător că benzile adiționale s-au detectat sub condiții de separare diferite folosite în acest exemplu. Variațiile dintre estimările anterioare și cele noi ale greutății moleculare, de asemenea, sunt de așteptat să dea diferențe în condiții analitice. În analiza acestui exemplu, se crede că greutatea moleculară mai mare estimată este mai precisă decât în exemplul 1, ca un rezultat al rezoluției îmbunătățite. Totuși, acesta s-a estimat pe baza analizei SDS PAGE, care tipic nu are precizie analitică și are ca rezultat estimări de peptide care ar putea fi alterate în continuare, datorită modificărilor post- și cotranslaționale.

Secvențele aminoacide s-au determinat pentru porțiuni N-terminale ale cinci din cele 10 peptide descompuse. Tabelul 9 corelează greutatea moleculară a proteinelor și secvențelor identificate. În SEQ ID NR:2, anumite analize sugerează că prolina la restul 5 poate fi o asparagină (asn). În SEQ ID NR:3, anumite analize sugerează că resturile aminoacide la pozițiile 13 și 14 sunt ambele arginină (arg). În SEQ ID NR: 4, anumite analize sugerează că restul aminoacid la poziția 6 poate să fie alanină (ala, fie serină (ser). În SEQ ID NR:5, anumite analize sugerează că restul aminoacid la poziția 3 poate fi acid aspartic (asp).

Tabelul 9

Estimat Exemplul 1	Estimat nou*	Lista SEQ
208	200,2 kDa	SEQ ID NR:1
184	175,0 kDa	SEQ ID NR:2
65,6	68,1 kDa	SEQ ID NR:3
60,8	65,1 kDa	SEQ ID NR:4
56,2	58,3 kDa	SEQ ID NR:5
25,1	23,2 kDa	SEQ ID NR:15

* Estimările noi se bazează pe SDS PAGE și nu se bazează pe secvențe genice. SDS PAGE nu are precizie analitică.

Exemplul 5. Partea B Caracterizarea componentelor peptidei toxină

Secvența N-terminală nouă, SEQ ID NR:15, Ala Gin Asp Gly Asn Gin Asp Tiu Phe Phe Ser Gly Asn Thr, s-a obținut prin secvențarea suplimentară a N-terminalului peptidelor izolate de la toxină nativă purificată-HPLC, cum s-a descris în exemplul 5, Partea A, de mai sus. Această peptidă vine de la gena *tcaA*. Peptida marcată TcaAii3 începe la poziția 254 și continuă până la poziția 491, unde începe peptida TcaA_{iii}, SEQ ID NR:4. Mărimea estimată a peptidei bazată pe secvența genei este 25,240 Da.

Exemplul 6. Caracterizarea componentelor peptidei toxină

Într-o altă analiză, complexul proteină toxină s-a reizolat din mediul de creștere *Photorhabdus luminescens* (după cultura fără Tween), prin realizarea unei precipitări cu sulfat de amoniu 10-80%, urmată de o etapă de cromatografie schimbătoare de ion (Mono Q) și două etape de cromatografie după mărime moleculară. Aceste condiții au fost asemenea celor din exemplul 1. În timpul primei etape după mărime moleculară, s-a găsit un al doilea peak biologic activ la aproape 100 ± 10 kDa. Pe baza măsurărilor proteinei, această fracțiune a fost de 20-50 ori mai puțin activă decât cea mai mare, sau inițial, peak activ la aproape 860 ± 100 kDa (nativ). De asemenea, în timpul acestui experiment de izolare, s-a desprins un peak activ mai mic de aproape 325 ± 50 kDa care a reținut o parte considerabilă a activității biologice inițiale. Se crede că peak-ul 325 kDa este în legătură cu, sau derivat de la, peak-ul 860 kDa.

O proteină de 56 kDa s-a descompus în această analiză. Secvența N-terminală a acestei proteine este prezentată în SEQ ID NR:6. Este demn de reținut că această proteină prezintă identitate și conservare semnificative cu SEQ ID NR:5 la N-terminus, sugerând că cele două pot fi codificate prin membrii separați ai familiei genei și că proteinele produse de către fiecare genă sunt suficient similare ambelor pentru a fi operabile în complexul toxină insecticidă.

O a doua proteină evidentă 185 kDa a fost prezentă consistent în cantități comparabile cu cele ale proteinei 3 din tabelul 9, și poate fi aceeași proteină sau fragment de proteină. Secvența N-terminală a acestei proteine 185 kDa este prezentată la SEQ ID NR: 7.

De asemenea, s-au obținut date suplimentare despre secvența aminoacidică N-terminală de la proteine izolate. Nici una dintre secvențele N-terminale determinate nu pare identică proteinei identificată în tabelul 9. Alte proteine sunt prezente în preparare izolată. O astfel de proteină are o greutate moleculară estimată de 108 kDa și secvența N-terminală cum se arată în SEQ ID NR:9.

Când materialul proteic s-a analizat după mărime în peak-ul activ de aproximativ 325 kDa, au fost observate benzi de aproximativ 51, 31, 28 și 22 kDa. Așa cum în toate cazurile în care o greutate moleculară s-a determinat prin analiza mobilității electroforetice, aceste greutăți moleculare au fost supuse efectelor de eroare introduse prin diferențe de tărie ionică a tamponului, diferențe de putere electroforetică și altele asemenea. Un specialist în domeniu va înțelege că valorile greutății moleculare definitivă nu pot fi determinate folosind aceste metode standard și că fiecare este supusă variației. S-a presupus că proteinele acestor mărimi sunt produși de degradare ai speciilor de proteine mai mari (mărime de aproximativ 200 kDa), care s-au observat în complexul toxină inițial mai mare.

În final, câteva preparate au inclus o proteină care are secvența N-terminală prezentată în SEQ ID NR:10. Această secvență are omologie puternică la proteine însoțitoare cunoscute, proteine auxiliare cunoscute a funcționa în ansamblul complexelor proteină mare. Cu toate că cererile nu au putut atribui o astfel de funcționare de ansamblu proteinei identificate în SEQ ID NR: 10, aceasta a fost în concordanță cu existența complexului proteină

1 toxină descrisă, așa cum o proteină însoțitoare ar putea fi implicată în ansamblul ei. Mai mult,
 3 cu toate că astfel de proteine nu s-a sugerat direct că ar avea activitate toxică, această pro-
 teină poate fi importantă pentru determinarea întregii naturi structurale a toxinei proteice și,
 astfel, poate contribui la activitatea toxică sau durabilitatea complexului *in vivo* după admi-
 5 nistrare orală.

În continuare, s-a inițiat analiza stabilității complexului toxină proteică la proteinaza K.
 7 S-a determinat că după incubarea 24 h a complexului, în prezența a 10-ori exces molar de
 proteinază K, activitatea s-a eliminat virtual (mortalitatea la aplicare orală în picături la aproa-
 9 pe 5%). Aceste date confirmă natura proteică a toxinei.

Activitatea toxică a fost reținută, de asemenea, printr-o membrană de dializă, confir-
 11 mând din nou mărimea mare a complexului toxină nativă.

Exemplul 7 Izolarea, caracterizarea și secvențarea aminoacidă parțială a toxinelor
 13 *Photorhabdus*

Izolarea și secvențarea aminoacidă N-terminală: într-un ser de experimente desfă-
 15 șurate în paralel cu exemplele 5 și 6, s-a realizat precipitarea cu sulfat de amoniu a proteinelor
Photorhabdus, prin ajustarea bulionului *Photorhabdus*, tipic 2-3 litri, la o concentrație fina-
 17 lă fie de 10%, fie 20%, prin adăugarea lentă a cristalelor de sulfat de amoniu. După agitare
 timp de 1 h, la 4°C, materialul s-a centrifugat la 12000xg, timp de 30 min. Supernatantul s-a
 19 ajustat la 80% sulfat de amoniu, s-a agitat la 4°C, timp de 1 h și s-a centrifugat la 12000xg,
 timp de 60 min. Peletul s-a resuspendat în unu-zece volum de Na₂PO₄ 10 mM pH 7,0 și s-a
 21 dializat împotriva aceluiași tampon fosfat, peste noapte, la 4°C. Materialul dializat s-a centri-
 fugat la 1.2000xg, timp de 1 h, înainte de cromatografie schimbătoare de ion.

O coloană schimb anion HR 16/50 Q Sepharose (Pharmacia) s-a echilibrat cu 10 mM
 Na₂PO₄, pH 7,0. Peletul sulfat de amoniu centrifugat, dializat, s-a aplicat la coloană Q
 25 Sepharose la o rată de 1,5 ml/min și s-a spălat extensiv la 3,0 ml/min cu tampon de echili-
 brare până la densitate optică (O.D. 280), atingând mai puțin de 0,100. În continuare, s-a
 27 aplicat la coloană, fie 60 min, un gradient NaCl în domeniul de la 0 până la 0,5 M la 3 ml/min,
 fie o serie de etape de eluție utilizând 0,1 M, 0,4 M și în final 1,0 NaCl, timp de 60 min, fie-
 29 care. Frațiunile s-au colectat și concentrat folosind un Centriprep 100. Alternativ, proteinele
 ar putea fi eluate printr-o singură spălare 0,4 M NaCl, fără eluție anterioară cu 0,1 M NaCl.

Alicote de 2 ml probe concentrate Q Sepharose s-au încărcat la 0,5 ml/min pe o
 coloană filtrare gel HR 16/50 Sepharose 12 (Pharmacia) echilibrată cu 10 mM Na₂PO₄,
 33 pH 7,0. Coloana s-a spălat cu același tampon pentru 240 min la 0,5 ml/min și 2 min probele
 s-au colectat. Materialul de volum liber s-a colectat și concentrat folosind un Centriprep 100.
 35 Alicote de 2 ml probe concentrate de Superose 12 s-au aplicat la 0,5 ml/min într-o coloană
 de filtrare gel HR 16/50 Sepharose 4B-CL (Pharmacia) cu 10 mM Na₂PO₄, pH 7,0. Coloana
 37 s-a spălat cu același tampon pentru 140 min la 0,5 ml/min și în 2 min s-au colectat probe.

Peak-ul de proteină exclus s-a supus la o a doua fracționare prin aplicarea la o
 39 coloană filtrare gel care folosește o rășină Sepharose CL-4B, care separă proteine de la
 ~ 30 kDa până la 1000 kDa. Această fracțiune s-a descompus în două peak-uri: un peak
 41 minor la volumul liber (>1000 kDa) și un peak major care a eluat la o greutate moleculară
 aparentă, la aproape 860 kDa. După o perioadă de o săptămână, probele supuse în con-
 43 tinuare la filtrare gel au arătat apariția graduală a unui al treilea peak (aproximație 325 kDa),
 care pare să apară din peak-ul major, probabil prin proteoliză limitată. Biotestele realizate
 45 pe cel de-al treilea peak au arătat că peak-ul liber nu are activitate, în timp ce fracția complex
 toxină 869 kDa are activitate ridicată și peak-ul de 325 kDa are activitate scăzută, deși total
 47 puternică. Analiza SDS PAGE ale peak-urilor complex toxină Sepharose CL-4B de la diferite
 produse de fermentație a arătat două modele distincte de peptide, notate "P" și "S". Cele

două modele au înregistrat diferențe în greutatea moleculară și concentrațiile componentei peptidă în fracțiunile lor. Modelul "S", produs mai frecvent, are 4 peptide cu greutatea moleculară ridicată (>150 kDa), în timp ce modelul "P" are 3 peptide cu greutatea moleculară ridicată. În plus, fracția peptidă "S" s-a găsit a fi de 2-3 ori mai activă împotriva European Corn Borer. Această schimbare poate să se refere la variații în expresia proteinei datorate vârstei de inoculare și/sau altor factori bazați pe creșterea parametrilor culturilor îmbătrânite.

Cantități în miligrame de fracțiuni complex toxină peak, determinate a fi modele de peptidă "P" sau "S", s-au supus la SDS PAGE preparativă și s-au transblotat cu TRIS-glicină (Seprabuff™ la membrane PVDF (ProBlott™, Applied Biosystems) pentru 3-4 h. Bloturile s-au trimis pentru analize de aminoacid și aminoacid N-terminal, care secvențează la Harvard MicroChem și Cambridge ProChem, respectiv. Trei peptide în model "S" au secvențe aminoacide unice N-terminale, comparativ cu secvențele identificate în exemplul anterior. O peptidă de 201 kDa (TcdA_{ii}) situată înaintea de SEQ ID NR:13, în cele ce urmează, împarte între 33% identitate aminoacizi și 50% similaritate cu SEQ ID NR:1 (TcbA_{ii}) (Tabelul 10, în tabelul 10 liniile verticale denotă identități aminoacide și coloanele indică substituții aminoacide conservative). O a doua peptidă de 197 kDa, SEQ ID NR: 14 (TcdB) a avut 42% identitate și 58% omologie cu SEQ ID NR:2 (TcaC). O a treia secvență de 205 kDa s-a denumit TcdA_{ii}. În plus, o secvență aminoacidă N-terminală limitată, SEQ ID NR:16 (TcbA) a unei peptide de cel puțin 235 kDa, a fost identică în omologie cu secvența aminoacidă, SEQ ID NR:12, dedusă dintr-o genă donată (tcbA), SEQ ID NR:11, care conține o secvență aminoacidă dedusă corepunzând la SEQ ID NR:1 (TcbA_{ii}). Aceasta arată că peptida mai mare 235+ kDa s-a procesat proteolitic la peptidă 201 kDa, (TcbA_{ii}), (SEQ ID NR: 1), în timpul fermentației, care rezultă probabil în activarea moleculei. Într-o altă secvență, secvența originală raportată ca SEQ ID NR:5 (TcaB_{ii}), raportată în exemplul 5 de mai sus, s-a găsit a conține un rest acid aspartic (Asp) la a treia poziție, mai degrabă decât glicină (Gly) și doi aminoacizi adiționali Gly și Asp la pozițiile a 8-a și respectiv a 9-a. În alte două secvențe, SEQ ID NR:2 (TcaC) și SEQ ID NR: 3 (tCAB_i), s-a obținut secvență aminoacidă adițională. Cantitatea densitometrică s-a realizat folosind o probă care a fost identică preparatului "S", necesitând analiză N-terminală. Această analiză a arătat că peptidele 201 kDa și 197 kDa reprezintă 7,0% și 7,2%, respectiv, din proteina totală colorată Coomassie brilliant blue în modelul "S" și este prezentă în cantități similare altor peptide abundente. S-a apreciat că aceste peptide pot reprezenta omologi de proteină, analog situațiilor găsite cu alte toxine bacteriene, precum diferite toxine CryI *Bt*. Aceste proteine variază de la omologie 40-90% la secvența lor aminoacidă N-terminală, care înconjoară fragmentul toxic.

Secvențarea aminoacidă internă pentru a facilita donarea genelor peptidei toxină, secvențele aminoacide interne ale peptidelor selectate s-au obținut după cum urmează. Cantități în miligrame ale peak-ului fracțiunilor 2A, determinate a fi modele peptidă "P" sau "S", s-au supus la SDS PAGE preparativă și s-a transblotat TRIS-glicină (Seprabuff™ la membrane PVDF (ProBlott™, Applied Biosystems) pentru 3-4 h. Bloturile s-au trimis pentru analize aminoacide și secvențarea aminoacidul N-terminal, la Harvard MicroChem și, respectiv, Cambridge ProChem. Trei peptide menționate ca TcbA_{ii} (care conține SEQ ID NR:1), TcdA_{ii} și TcaB_i (care conține SEQ ID NR:3) s-au supus digestiei tripsină prin Harvard MicroChem, urmată de cromatografia HPLC, pentru a separa peptidele individuale. Analize aminoacide N-terminale s-au realizat pe fragmente de peptidă selectate triptic. Două peptide interne s-au secvențat pentru peptide TcaB_i (peptidă 205 kDa) menționată ca TcaB_i-PT111 (SEQ ID NR:17) și TcaB_i-PT79 (SEQ ID NR:18). S-au secvențat două peptide interne pentru peptidă TcaB_i (peptidă 68 kDa) menționată ca TcaB_i-PT158 (SEQ ID NR:19) și TcaB_i-PT108 (SEQ ID NR:20). S-au secvențat patru peptide interne pentru peptidă TcbA_{ii} (peptidă 201 kDa) menționată ca TCBAII-TP103 (SEQ ID NR:21), TcbA_{ii}PT56 (SEQ ID NR:22), TcbA_{ii}PT81(a) (SEQ ID NR:23) și TcbA_{ii}PT81 (b) (SEQ ID NR:24).

Secvențe aminoacide N-terminale

201 kDa (identitate 33% & similaritate 50% la SEQ ID NR:1)

L I G Y N N Q F S G * A SEQ ID NR:13

: | | | | : |

F I Q G Y S D L F G N - A SEQ ID NR:1

197 kDa (identitate 42% & similaritate 58% SEQ ID NR:2)

M Q N S Q T F S V G E L SEQ ID NR:14

| | : | | : |

M Q D S P E V S I T T L SEQ ID NR:2

Exemplul 8. Construcția unei bănci cosmid a ADN genomic *Photorhabdus luminescens* W-14 și cercetarea acesteia pentru a izola gene care codifică peptide care cuprind prepararea proteinei toxice

Constituind o condiție pentru producerea proteinelor toxice insectă *Photorhabdus* în gazde heteroloage și pentru alte utilizări, este necesară izolarea și caracterizarea genelor care codifică aceste peptide. Acest obiectiv s-a urmărit în paralel. O abordare, descrisă mai târziu, s-a bazat pe utilizarea anticorpilor monoclonali și policlonali apăruți împotriva toxinei purificate, care s-au folosit apoi pentru a izola clone dintr-o bancă de expresie. O altă abordare, descrisă în acest exemplu, se bazează pe utilizarea datelor secvenței aminoacide N-terminale și interne, pentru a desemna oligonucleotide degenerate pentru utilizare în amplificare PCR. Fiecare metodă poate fi utilizată pentru a identifica clone ADN care conțin gene care codifică-peptidă, astfel încât să permită izolarea genelor respective și determinarea secvenței lor ADN bază.

IZOLARE ADN GENOMIC: *Photorhabdus luminescens* tulpina W-14 (ATCC număr de acces 55397) a crescut pe agar 2% proteose peptone #3 (Difco Laboratories, Detroit, MI) și competența toxină insecticidă s-a menținut prin repetarea biotestului după transformare, folosind metoda descrisă în exemplul 1 de mai sus. O cultură de 50 ml agitată s-a produs într-un balon cu șicane de 175 ml în mediu 2% proteose peptone #3, crescut la 28°C și 150 rpm, pentru aproximativ 24 h. Din această cultură, 15 ml s-au peletat și congelat în mediul său, la -20°C, până când s-a dezghețat pentru izolarea ADN-ului. Cultura dezghețată s-a centrifugat (700xg, 30 min) și s-a îndepărtat materialul mucopolizaharid portocaliu care plutește. Materialul celular rămas s-a centrifugat (25000xg, 15 min) la pelet, celulele bacteriale și mediul s-au îndepărtat și înlăturat.

S-a izolat ADN genomic printr-o adaptare a metodei CTAB descrisă în secțiunea 2.4.1 din *Current Protocols in Molecular Biology* (Ausubel et al., eds. John Wiley & Sons, 1994) (modificată pentru a include o sare șoc și cu toate volumele crescute de 10 ori). Celule bacteriene peletate s-au resuspendat în tampon TE (10 mM Tris-HCl, 1 mM EDTA, pH 8,0) la un volum final de 10 ml, apoi s-au adăugat 12 ml de NaCl 5M; acest amestec s-a centrifugat 20 min, la 15000xg. Peletul s-a resuspendat în 5,7 ml TE și 300 ml SDS 10% și 60 ml proteinază K 20 mg/ml (Gibco BRL Products, Grand Island, NY; în apă distilată sterilă) s-au adăugat la suspensie. Acest amestec s-a incubat la 37°C, pentru o oră; apoi s-au adăugat aproximativ 10 mg lizozimă (Worthington Biocchemical Corp., Freehold, NJ). După încă 45 min, s-au adăugat 1 ml de NaCl 5M și 800 ml de soluție CTAB/NaCl (10% g/v CTAB,

RO 121280 B1

0,7 MnaCl). Acest preparat s-a incubat 10 min, la 65°C, apoi s-a agitat blând și s-a incubat suplimentar și agitat pentru aproximativ 20 min, pentru a sprijini clarificarea materialului celular. S-a adăugat un volum egal de soluție cloroform/alcool izoamilic (24:1, v/v), s-a amestecat blând și s-a centrifugat. După două extracții cu un volum egal de PCI (fenol/cloroform/alcool izoamilic; 50:49:1, v/v/v; s-a echilibrat cu IM Tris-HCl, pH 8,0; Intermountain Scientific Corporation, Kaysville, UT), a precipitat ADN-ul cu 0.6 volume de izopropanol, ADN-ul precipitat s-a îndepărtat blând cu o baghetă de sticlă, s-a spălat de două ori cu 70% etanol, s-a uscat și s-a dizolvat în 2 ml STE (10 mM Tris-HCl pH 8,0, 10 mM NaCl, 1 mM EDTA). Acest preparat a conținut 2,5 mg/ml ADN, cum s-a determinat prin densitate optică la 260 nm (de exemplu, OD₂₆₀).

Domeniul mărimii moleculare al ADN genomic izolat s-a evaluat pentru potrivire la construirea băncii. Analiza gel CHSF s-a realizat în geluri de agaroză 1,5% (Seakem® LE, FMC BioProducts, Rockland, ME) cu tampon 0,5xTBE (44,5 mM Tris-HCl pH 8,0, 44,5 mM H₃BO₃, 1 mM EDTA) pe un aparat BioRad CHEF-DRii cu un Tulseware 760 Switcher (Bio-Rad Laboratories, Inc., Richmond, CA). Parametrii de funcționare sunt: timp inițial A, 3 s; timp final A, 12 s; 200 volți; temperatura de funcționare, 4-18°C; perioada de funcționare, 16,5 h. Colorarea bronură de etidiu și examinarea gelului sub lumină ultravioletă au arătat că ADN-ul s-a întins de la 30-250 kbp în mărime.

CONSTRUCȚIA BĂNCII: O digestie Sau3 A parțială s-a făcut acestui preparat ADN genomic *Photothabdus*. Metoda s-a bazat pe secțiunea 3.1.3 a Ausubel (supra). Adaptările au inclus reacții de funcționare la scară mică, sub diferite condiții, până când s-au obținut rezultate aproape optime. S-au realizat câteva reacții la scară mare, cu diferite condiții, rezultatele s-au analizat pe geluri CHEF și numai prepararea la scară mare, cea mai bună, s-a continuat. În cazul optim, 200 μg ADN genomic *Photothabdus* s-au incubat eu 1,5 unități de Sau3A 1 (New England Biolabs, "NEB", Beverly, MA) pentru 15 min, în volum total de 2 ml tampon 1 x NEB 4 (suplimentat ca 10x de către producător). Reacția s-a stopat prin adăugarea a 2 ml PCZ și centrifugare la 8000xg, timp 10 min. S-au adăugat la supernatant 200 μl de NaCl 5M, plus 6 ml gheață-etanol rece. Acest preparat s-a răcit timp de 30 min, la -20°C, apoi s-a centrifugat la 12000xg timp de 15 min. Supernatantul s-a îndepărtat și precipitatul s-a uscat într-un cuptor cu vid, la 40°C, apoi s-a resuspendat în 400 μl STE. Testul spectrofotometric a indicat aproximativ 40% recuperare a ADN-ului absorbit. ADN-ul digerat s-a fracționat după mărime pe un gradient de sucroză, conform secțiunii 5.3.2 a CPMB (op.cit.). S-a preparat un gradient linear sucroză 10% la 40% (g/v) cu un gradient făcut în tuburi Ultra-Clear™ (Beckman Instruments, Inc., Palo Alto, CA) și proba ADN s-a stratificat la vârf. După centrifugare, (26ppp rpm, 17 h, rotor Beckman SW41,20°C), fracțiunile (aproape 759 μl) s-au tras în jos de la vârful gradientului și s-au analizat prin electroforeză pe gel CHEF (cum s-a descris mai devreme). S-au selectat fracțiunile conținând fragmente Sau3A 1 în domeniul de mărime 20-40 kbp și s-au precipitat printr-o modificare (cantități ale tuturor soluțiilor crescute aproximativ de 6,3-ori) ale metodei descrise în secțiunea 5.5. a lui Ausubel (supra.). După precipitare, peste noapte, s-a colectat ADN-ul prin centrifugare (17000xg, 15 min), s-au uscat, s-au redizolvat în TE, s-au depozitat într-un volum total de 80 μl și reprecipitat cu încă 8 μl acetat de sodiu 3M și 220 μl etanol. Peletul colectat prin centrifugare ca mai sus s-a resuspendat în 12 μl TE. S-a determinat concentrația ADN-ului prin fluorimetrie Hoechst 33258 dye (Polysciences, Inc., Warrington, PA) într-un fluorimetru Hoefer TKO100 (Hoefer Scientific Instruments, San Francisco, CA). S-au recuperat aproximativ 2,5 μg ADN fracționat după mărime.

S-au digerat 20 μg cosmid pWe 15 ADN (Stratagene, La Jolla, CA) la completare cu 100 unități enzimă de restricție BamHI (NEB) în tamponul fabricantului (volum final 200 μl,

1 37°C, o oră). Reacția s-a extras cu 100 μ l PCI și s-a precipitat ADN din faza apoasă, prin
 3 adiția a 20 μ l acetat de sodiu 3M și 550 μ l etanol absolut -20°C. După 20 min, la -70°C, s-a
 colectat ADN prin centrifugare (17000xg, 15 min), s-a uscat sub vid și s-a dizolvat în 180 μ l
 5 Tris-HCl 10 mM, pH 8,0. La aceasta s-au adăugat 20 μ l tampon 10x CIP (100 mM Tris-HCl,
 pH 8,3; 10 mM ZnCl₂; 10 mM MgCl₂) și 1 μ l (0,25 unități) de fosfatază alcalină intestinală de
 7 vițel 1:4 (Boehringer mannheim Corporation, Indianapolis, IN). După 30 min, la 370°C, s-au
 făcut următoarele adiții: 2 μ l 0,5 M EDTA, pH 8,0; 10 μ l 10% SDS; 0,5 μ l de proteinază K
 9 20 mg/ml (ca mai sus), urmat de incubare la 55°C, timp de 30 min. După extracțiile secven-
 țiale cu 100 μ l de PCI și 100 μ l fenol (Intermountain Scientific Corporation, echilibrat cu
 11 Tris-HCl 1 M, pH 8,0), s-a precipitat ADN-ul defosforilat prin adiția a 72 μ l acetat de amoniu
 7,5 M și 550 μ l etanol -20°C, incubare pe gheață pentru 30 min și centrifugare ca mai sus.
 13 ADN-ul peletat s-a spălat o dată cu 500 μ l -20°C 70%, s-a uscat sub vid și s-a dizolvat în
 20 μ l de tampon TE.

15 Ligarea fragmentelor Sau3A 1 fracționate după mărime la vectorul pWE15 digerat
 BamH 1 și fosfatizat s-a realizat folosind ligaza T4 (NEB) prin modificarea (adică, folosirea
 17 unui tampon de ligare 10X premixat furnizat de producător) protocolului în secțiunea 3,33 a
 lui Ausubel. Ligarea s-a realizat peste noapte, într-un volum total de 20 μ l la 15°C, urmată
 de depozitare la -20°C.

19 4 μ l din reacția de ligare ADN cosmid, conținând aproximativ 1 μ g ADN, s-au împachet-
 21 etat în bacteriofag lambda, folosind un extract de împachetare comercial (Gigapack® III
 Gold Packaging Extract, Startagene), urmând instrucțiunile fabricantului. Preparatul de împachetare
 23 s-a depozitat la 4°C, până la folosire. Preparatul cosmid împachetat s-a utilizat pen-
 tru a infecta celule *Escherichia coli* XLI Blue MR (Stratagene) conform protoalelor
 25 Gigapack® III Gold ("Titering the Cosmid Library"), după cum urmează. Celule XLI Blue MR
 s-au crescut în mediu LB (g/L: Bacto-triptonă, 10; extract Bacto-drojdie, 5; agar-Bacto, 15;
 27 NaCl, 5; [Difco Laboratories, detroit, MI]) care conține 0,2% (g/v) maltoză plus 10 mM
 MgSO₄, la 37°C. După 5 h de creștere, celulele s-au peletat la 700xg (15 min) și s-au resus-
 29 pendat în 6ml 10 mM MgSO₄. Densitatea de cultură s-a ajustat cu 10 mM MgSO₄ la
 OD₆₀₀=0,5. Banca cosmid împachetată s-a diluat 1:10 sau 1:20 cu mediu steril SM (0,1 M
 31 NaCl, 10 mM MgSO₄, 50 mM Tris-HCl pH 7,5, 0,01% g/v gelatină) și 25 μ l preparat diluat s-a
 amestecat cu 25 μ l celule XL1 Blue MR diluate. Amestecul s-a incubat la 25°C, pentru 30 min
 (fără agitare), apoi s-au adăugat 200 μ l bulion LB și incubarea a continuat aproximativ o oră,
 33 agitând blând, ocazional. Alicotele acestei culturi (20-40 μ l) s-au împrăștiat pe plăci de agar
 LB care conțin 100 mg/l ampicilină (adică, LB-Amp₁₀₀) și s-au incubat peste noapte, la 37°C.
 35 Pentru a depozita banca fără amplificare, coloniile individuale s-au adunat și s-au inoculat
 în godeurile individuale ale plăcilor microgodeuri sterile cu 96-godeuri; fiecare godeu care
 37 conține 75 μ l Terrific Broth (mediu TB: 12 g/l Bacto-triptonă, 24 g/l extract de Bacto-drojdie,
 0,4% v/v glicerol, 17 mM KH₂PO₄, 72 mM K₂HPO₄) plus 100 mg/l ampicilină (adică,
 39 TB-Amp₁₀₀) și s-au incubat (fără agitare) peste noapte, la 37°C. După replicarea plăcii de
 96-godeuri într-o placă copie, s-au adăugat la placă 75 μ l/godeu TB sterilizat prin
 41 filtrare: glicerol (1:1, v/v; cu sau fără 100 mg/l ampicilină), s-a scuturat sumar la 100 rpm și
 s-a închis cu Parafilm® (American National Can, Greenwich, CT) și s-au plasat la -70°C,
 43 într-un congelator pentru depozitare. Plăcile copie s-au crescut și s-au procesat identic cu
 placa mașter. S-a preparat un total de 40 de asemenea plăci mașter (și copiile lor).

45 CERCETAREA BĂNCII CU PROBE ADN RADIOMARCAT: Pentru a prepara filtre
 colonie pentru testarea cu probe radioactive marcate, zece plăci de 96 godeuri ale băncii
 47 s-au dezghețat la 25°C (reper vârf la temperatura camerei). Un instrument de acoperire șablon
 cu 96 vârfuri s-a folosit pentru a inocula o placă copie cu 96 godeuri, nouă, care conține

RO 121280 B1

75 µl/godeu de TB-Amp₁₀₀. Placa copie s-a crescut peste noapte (staționar) la 37°C, apoi s-a scuturat timp de 30 min, la 100 rpm la 37°C. Un total de 800 colonii a fost reprezentat în aceste plăci copii, datorită noncreșterii unor izolate. Instrumentul șablon s-a folosit pentru a inocula amprente ale matricii de 96 godeuri pe membrane de naylon (0,45 µ, 200x250 mm) Magna NT (MSI, Westboro, MA) care fuseseră plasate pe LB-Amp¹⁰⁰ solid (100 ml/vas) în vase de plastic Bio-test (Nune, 243x243x18 mm; Curtin Mathison Scientific, Inc., Wood Dale, IL). Coloniile s-au crescut pe membrane la 37°C, timp de aproximativ 3h.

O colonie de control pozitivă (o clonă bacterială care conține un insert secvență GZ4, vezi mai jos) s-a crescut pe o membrană Magna NT (Nune, 0,45 µ, cerc 82 mm) separată pe mediu LB suplimentat cu 35 mg/l cloramfenicol (adică, LB-Cam₃₅) și s-au procesat de-a lungul membranelor coloniei băncii. Coloniile bacteriale de pe membrană s-au lizat și ADN-ul s-a denaturat și s-a neutralizat în conformitate cu protocolul luat de la Genius™ System User's Guide Versiunea 2.0 (Boehringer Mannheim, Indianapolis, IN). Membranele s-au plasat cu partea cu colonia deasupra pe o hârtie de filtru îmbibată cu 0,5 N NaOH plus 1,5 M NaCl, timp de 15 min, pentru a denatura, și s-au neutralizat pe o hârtie de filtru îmbibată cu 1M Tris-HCl pH 8,0 1,5 M NaCl, timp de 15 min. După crosslinking-UV folosind un set Stratagene UV Stratalinker pe auto crosslink, membranele s-au depozitat la 25°C, până la utilizare. Membranele s-au netezit în fâșii care conțin amprente duplicat ale unei singure plăci de 96-godeuri, apoi s-a spălat extensiv prin metoda secțiunii 6.4.1 în CPMB (op.cit.): 3 h, la 25°C, în 3x SSC, 0,1% (g/v) SDS, urmat de o oră, la 65°C, în aceeași soluție, apoi s-a clătit în 2x SSC în prepararea pentru etapa de hibridizare (20x SSC= 3M NaCl, 0,3 M citrat de sodiu, pH 7,0).

Amplificarea unui fragment genomic specific al genei *tcaC*

Bazat pe secvența aminoacidă N-terminală, s-a determinat pentru fracția peptidă TcaC purificată [descriș aici ca SEQ ID NR:2], un depozit de oligonucleotide degenerate (rezervă S4Psh) s-a sintetizat prin chimie standard *p*-cianoetil pe un Applied BioSystem ABI394 DNA/RNA Synthesizer (Perkin Elmer, Foster City, CA). Oligonucleotidele s-au deprotejat 8 h, la 55°C, s-au dizolvat în apă, s-au determinat cantitativ prin măsurători de spectrofotometrie și s-au diluat pentru utilizare. Acest depozit corespunde secvenței aminoacide N-terminale determinată peptidei TcaC.

Secvența aminoacidă determinată și secvența ADN degenerată sunt date mai jos, unde A, C, G și T sunt baze ADN standard și I reprezintă inozină:

Aminoacid	Met	Gin	Asp	Ser	Pro	Glu	Val
S4Psh	5'	ATG	CA(A/G)	GA(T/C)	(T/A)(C/G)(T/A)	CCI	GA(A/G) GT 3'

S-a sintetizat alt set de oligonucleotide degenerate (depozit P2.3.5R), care reprezintă complementul catenei care codifică pentru secvența aminoacidă determinată în SEQ ID NR:17:

Amino-							
acid	Ala	Phe	Asp	Ile	Asp	Asp	Val
Codonii 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'

Aceste oligonucleotide s-au folosit ca primeri în Polymerase Chain Reactions (PCR®, Roche Molecular Systems, Branchburg, NJ), pentru a amplifica fragmentul ADN specific din ADN genomic preparat din tulpila W-14 *Photorhabdus* (vezi mai sus). O reacție tipică (50 µl) care conține 125 pmol din fiecare depozit primer P2Psh și P2.3.5R, 255 ng șablon ADN genomic, 10 nmol din fiecare din dATP, dCTP, dGTP și dTTP, tampon PCR IX GeneAmp® și 2,5 unități de ADN polimerază AmpliTaq® (ambele de la Roche Molecular Systems; tampon

1 10X GeneAmp® este 100 mM Tris-HCl pH 8,3, 500 mM KCl, 0,01% g/v gelatină). Amplificările
 3 s-au realizat într-un Perkin Elmer Cetus DNA Thermal Cyclor(Perkin Elmer, Foster City, Ca),
 folosind 35 cicluri de 94°C (1,0 min), 55°C (2,0 min), 72°C (3,0 min), urmat de o perioadă de
 5 extinsie de 7,0 min, la 72°C. Produsele de amplificare s-au analizat prin electroforeză prin
 7 2% g/v NuSieve® 3:1 agaroză (FMC Bio Products) în tampon TEA (40 mM Tris-acetat, 2mM
 EDTA, pH 8,0). Un produs specific, de mărime estimată 250 bp, s-a observat la numeroase
 alte produse amplificate prin bromură de etidiu (0,5 µg/ml), colorarea gelului și examinarea
 sub lumină ultraviolet.

9 Regiunea gelului care conține un produs de aproximativ 250 bp s-a excizat și un dop
 mic (0,5 mm dia.) s-a îndepărtat și s-a utilizat pentru a alimenta matrița pentru amplificarea
 11 PCR (40 cicluri). Reacția (50 µl) a conținut aceleași componente ca mai sus, minus matrița
 ADN genomic. După amplificare, capetele fragmentelor s-au teșit și s-au fosforilat prin incu-
 13 bare la 25°C, pentru 20 min, cu o unitate ADN polimerază T4 (NEB), 1 nmol ATP și 2,15
 unități de T4 kinază (Pharmacia Biotech Inc., Piscataway, NJ).

15 Fragmentele ADN s-au separat din primeri reziduali prin electroforeză prin agaroză
 1% g/v GTG® (FMC) în TEA. S-a excizat o felie de gel care conține fragmente de mărime
 17 aparentă 250 bp și s-a extras ADN-ul folosind un kit Qiaex (Qiagen Inc., Chatsworth, CA).

Fragmentele ADN extrase s-au ligat la vectorul plasmid pBC KS(+) (Stratagene) care
 19 a fost digerat la completare cu enzima de restricție Sma 1 și s-a extras într-un mod similar
 celui descris pentru ADN pW15, de mai sus. O reacție de ligare tipică (16,3 µl) a conținut
 21 100 ng digest ADN pBC KS(+), 70 ng ADN fragment de 250 bp, 1 nmol [Co(NH₃)₆]Cl₃ și 3,9
 unități Weiss AND ligază T4 (Collaborative Biomedical Products, Bedford, MA), în 1X tampon
 23 de ligare (50 mM Tris-HCl), pH 7,4; 10 mM MgCl₂; 10 mM ditiotritol; 1 mM spermidină, 1 mM
 ATP, 100 mg/ml albumină serică bovină). După incubare peste noapte, la 14°C, produsele
 25 ligate s-au transformat la îngheț, celule competente *Escherichia coli* DH5α (Gico BRL)
 conform recomandărilor furnizorilor și s-au acoperit pe plăci LB-Cam₃₅, conținând IPTG
 27 (119 µg/ml) și X-gal (50 µg/ml). S-au străpuns colonii albe independente și s-a preparat ADN
 plasmidul, printr-o metodă modificată alcalinizare/precipitare PEG (Prism™ Ready Reaction
 29 Dyedexy™ Terminator Cycle Sequencing Kit Protocols; ABI/Perkin Elmer). Secvența
 nucleică a ambelor catene ale ADN insert s-au determinat folosind primeri T7 [pBC KS(+)
 31 baze 601-623:TAAAACGACGGCCAGTGAGCGCG]și LacZ[pBCKS(+)]baze 792-816:
 ATGACCATGATTACGCCAAGCGCGC) și protocoale de alimentare cu kit de secvențare
 33 PRISM™ (ABI/Perkin Elmer). Dideoxiribonucleotide colorant-terminator nonîncorporate s-au
 îndepărtat prin trecere prin coloane Centri-Sep 100 (Princeton Separations, Inc., Adelphia,
 35 NJ), conform instrucțiunilor fabricantului. Secvența ADN s-a obținut prin analiza probelor pe
 un ABI Model 373 A DNA Sequencer (ABI/Perkin Elmer). Secvențele ADN ale celor două
 37 izolate, GZ4 și HB14, au rezultat așa cum sunt prezentate în fig. 1.

Această secvență ilustrează caracteristicile următoare: 1) bazele 1-20 reprezintă una
 39 din cele 64 secvențe posibile ale oligonucleotidelor degenerate S4Psh, ii) secvența amino-
 acizilor 1-3 și 6-12 corespunde exact celei determinate pentru N-terminus ale TcaC (descrișă
 41 ca SEQ ID NR:2), iii) al patrulea aminoacid codificat este un rest cisteină mai degrabă decât
 serină. Această diferență este codificată în degenerare pentru codonii serină (vezi mai sus),
 43 iv) al cincelea aminoacid codificat este prolină, corespunzător secvenței TcaC N-terminal dat
 ca SEQ ID NR:2, v) bazele 257-276 codifică una dintre cele 192 secvențe posibile desem-
 45 nate în depozitul degenerat, vi) codonul terminare TGA introdus la bazele 268-270 este
 rezultatul complementarității la constructul degenerat în depozitul oligonucleotidic la poziția
 47 corespunzătoare și nu indică un cadru de citire micșorat pentru gena corespunzătoare.

RO 121280 B1

<i>Marcarea probei peptidă tcaC specific-genă.</i>	1
Fragmente AND care corespund celor 276 date de mai sus s-au amplificat (35 cicluri) prin PCR® într-un volum de reacție de 100 µl, folosind 100 pmol fiecare dintre primerii P2Psh și P2.3.5T, 10 ng de plasmizi GZ4 și HB14 ca matrițe, 20 nmol fiecare dintre dATP, dCTP, dGTOP și dTTP, 5 unități ADN polimerază AmpliTaq® și 1X concentrație tampon GeneAmp®, sub aceleași regimuri de temperatură cum s-a descris mai sus. Produsele de amplificare s-au extras dintr-un gel agaroză 1% GTG® prin kit Qiaex și s-a determinat cantitativ prin fluorometrie.	3 5 7
Produsele de amplificare extrase din matrița plasmid HB14 (aproximativ 400 ng) s-au divizat în cinci alicote și s-au marcat cu ³² P-dCTP, folosind High Prime Labeling Mix (Boehringer Mannheim), conform instrucțiunilor producătorului. Radioizotopul neîncorporat s-a îndepărtat prin trecere prin NucTrap® Probe Purification Columns (Stratagene), conform instrucțiunilor furnizorului. Activitatea specifică a produsului ADN marcat s-a determinat, prin numărare scintilație, a fi 3,11 x 108 dpm/µg. Acest ADN marcat s-a utilizat pentru a proba membrane preparate dintre cei 800 membrii ai băncii genomice.	9 11 13 15
CERCETAREA CU O PROBĂ PEPTIDĂ-TCAC SPECIFIC GENĂ. Proba HB14 radiomarcată s-a fiert aproximativ 10 min, apoi s-a adăugat soluție "hib minimal". [Notă: Metoda "minimal hyb" este luată de la un protocol CERES; <i>Restriction Fragment Length Polymorphism Laboratory Manual version 4,0</i> , secțiuni 4-40 și 4.47; CERES/NPI, Salt Lake City, UT. NPI, care acum este scos din uz, cu succesorii săi care funcționează ca Linkage Gebetics], soluția "minimal hyb" conține 10% g/v PEG (polietilen glicol, greutate moleculară aproximativ 8000), 7% g/v SDS; 0,6X SSC, tampon 10 mM fosfat de sodiu (de la un stoc IX care conține 95 g/l NaH ₂ PO ₄ •1H ₂ O și 84,5 g/l Na ₂ HPO ₄ •7H ₂ O), 5 mM EDTA și 100 mg/ml ADN denaturat din spermă de somon. Membranele s-au șters, s-au uscat scurt, apoi, fără prehibridizare, 5 fâșii ale membranei s-au plasat în fiecare din cele două cutii de plastic care conțin 75 ml de "minimal hyb" și 2,6 ng/ml probă HB14 radiomarcată. Acesta s-a incubat peste noapte cu agitare ușoară (50 rpm), la 60°C, în "soluție de spălare minimal hyb" (0,25X SSC, 0,2% SDS), urmată de două spălări de 30 min, cu scuturare ușoară la 60°C în aceeași soluție. Filtrele s-au plasat pe hârtie acoperită cu Saran Wrap® (Dow Brands, Indianapolis, IN), într-o casetă autoradiografică cu lumină densă și s-a expus la film raze X X-Omat (Kodak, Rochester, NY) cu doi amplificatori DuPont Cronex Lightning-Plus CI (Sigma Chemical Co., St. Louis, MO), pentru 4 h, la -70°C. Pe dezvoltare (proceduri fotografice standard), s-au evidențiat semnale semnificative în ambele replicare, pe un fundal intens al semnalelor slabe mai neregulate. Filtrele s-au spălat din nou pentru aproape 4 h la 68GC în "soluție de spălare minimal hyb" și s-au plasat apoi, din nou, în casete, iar filmul s-a expus peste noapte la -70°C. S-au identificat 12 pozitive posibile datorate semnalelor puternice pe ambele amprente colonii duplicat 96-godeuri. Nu s-a văzut nici un semnal cu membrane negative de control (colonii ale celulelor XLI Blue MR care conțin pWE 15), dar s-a observat un foarte puternic semnal cu membrane pozitive de control (celule DH5α care conțin GZ4 izolat al produsului PCR), care s-a procesat în mod curent cu probe experimentale. Cele 12 colonii putative hibridizare-pozitive s-au recuperat de la plăcile înghețate 96-godeuri ale băncii și s-au crescut peste noapte la 37°C, pe mediu solid LB-Amp100. Pentru hibridizare, s-au preparat două seturi de membrane (Magna NT nylon, 0,45 µ). Primul set s-a preparat prin plasarea unui filtru direct în coloane pe o bucată placă rămasă, apoi s-a îndepărtat cu celule bacteriene aderente și s-a procedat ca mai jos. Filtrele celui de-al doilea set s-au plasat pe plăci care conțin mediu LB-Amp ₁₀₀ , apoi s-au inoculat cu celule de transferare de la bucata placă rămasă la filtre. După creșterea peste noapte, la 37°C, filtrele s-au îndepărtat de pe plăci și s-au procesat.	17 19 21 23 25 27 29 31 33 35 37 39 41 43 45 47

1 Celulele bacteriene de pe filtre s-au uzat cu ADN denaturat prin plasarea fiecărui filtru
 3 cu colonia în sus la un depozit (1,0 ml) de 0,5 N NaOH, într-o placă de plastic, pentru 3 min.
 5 Filtrele s-au șters uscat pe o hârtie prosop, apoi procesul s-a repetat cu NaOH 0,5 N proas-
 7 păt. După ștergere uscată, filtrele s-au neutralizat prin plasarea fiecăreia pe un depozit 1,0 ml
 9 de Tris-HCl 1M, pH 7,5, timp de 3 min, s-au șters uscat și s-au reneutralizat cu tampon
 11 proaspăt. Aceasta a fost urmată de două îmbibări similare (5 min, fiecare) pe depozite cu
 13 0,5 M Tris-HCl, pH 7,5 plus 1,5 M NaCl. După ștergere uscată, ADN-ul s-a crosslink-at UV
 15 pe filtru (ca mai sus) și filtrele s-au spălat (25°C, 100 rpm) în aproape 100 ml de 3X SSC plus
 0,1% (g/v) SDS (4 ori, 30 min, fiecare cu soluție proaspătă pentru fiecare spălare). Apoi, s-au
 plasat într-un volum minim de soluție de prehibridizare [6X SSC plus 1% g/v fiecare de Ficoll
 400 (Pharmacia), polivinilpirolidonă (greutate moleculară medie 360000; Sigma) și albumină
 serică bovină Fracția V; (Sigma)] pentru 2 h, la 65°C, 50 rpm. Soluția de prehibridizare s-a
 îndepărtat și s-a plasat cu proba BH14 ³²P-marcată care a fost salvat de la hibridizarea anteri-
 oară a membranelor băncii și care s-au denaturat la 95°C, timp de 5 min. Hibridizarea s-a
 realizat la 60°, timp de 16 h, cu scuturare la 50 rpm.

După îndepărtarea soluției probei marcate, membranele s-au spălat de 2 ori la 25°C
 (50 rpm, 15 min) în 3X SSC (aproape 150 ml fiecare spălat). Apoi s-au spălat timp de 3 h,
 la 68°C (50 rpm) în 0,25X SSC plus 0,2% SDS (soluție de spălare "minimal hyb") și s-au
 expus la film raze X, așa cum s-a descris mai sus, pentru 1,5 la 25°C (cercetări neaplicate).
 Această expunere a demonstrat semnale foarte puternice de hibridizare la izolații cosmid
 22G12, 25A10, 16A5 și 16B10 și un semnal foarte slab cu izolatul cosmid 8B10. Nu s-a
 văzut nici un semnal la coloniile (pWE15) control negativ și s-a văzut un semnal foarte puter-
 nic cu membrane control pozitive (celule DH5α care conțin izolatul GZ4 al produsului PCR),
 care a fost procesat în mod curent cu probe experimentale.

AMPLIFICAREA UNUI FRAGMENT GENOMIC SPECIFIC AL GENEI TCAB. Pe baza
 unei secvențe aminoacide N-terminale determinată pentru fracția peptidă TcaB_i purificată
 (descrisă aici ca SEQ ID NR:3), s-a sintetizat un depozit al oligonucleotidelor degenerate
 (depozit P8F), așa cum s-a descris pentru peptida TcaC. Secvența aminoacidă determinată
 și secvența ADN degenerată corespunzătoare sunt date mai jos, unde A, C, G și T sunt baze
 ADN standard și I reprezintă inozină:

Amino

acid	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)G3'
			(C/T)TI							

Un alt set de oligonucleotide degenerate s-au sintetizat (depozit P8.108.3R), repre-
 zentând complementul catenei care codifică pentru secvența aminoacidă a peptidei interne
 TcaB_iPT108 (descrisă aici ca SEQ ID NR:20):

Amino

acid	Met	Tyr	Tyr	Oe	Gin	Ala	Gin	Gin
Codon	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				

Aceste oligonucleotide s-au folosit ca primeri pentru PCR® folosind HotStart 50
 Tubes™ (molecular Bio-products, Inc., San Diego. CA) pentru a amplifica un fragment ADN
 specific de la ADN genomic preparat de la *Photorhabdus tulpina* W-14 (vezi mai sus). O
 reacție tipică (50 μl) a conținut (linia de jos) 25 pmol din fiecare primer depozit P8F și
 P8.108.3R cu 2 nmol fiecare dintre dATP, dCTP, dGTP, și dTTP, în tampon IX GeneAmp®
 PCR și (linie vârf) 230 ng ADN matriță genomic, 8 nmol fiecare dintre dATP, dGTP și dTTP

RO 121280 B1

și 2,5 unități ADN polimerază AmpliTaq®, în tampon IX GeneAmp®PCR. Amplificarea s-a realizat prin 35 cicluri, așa cum s-a descris pentru peptida TcaC. Produsele de amplificare s-au analizat prin electroforeză prin agaroză 0,7% g/v SeaKem® LE (FMC) în tampon TE. S-a observat un produs specific de mărime estimată de 1600 bp.	1
S-au depozitat patru astfel de reacții și ADN amplificat s-a extras dintr-un gel 1,0% SeaKem® LE prin kit Qiaex, așa cum s-a descris pentru peptida TcaC. ADN-ul extras s-a utilizat direct ca matriță pentru determinarea secvenței (PRISM™ Sequencing Kit), folosind depozitele de primeri P8F și P8.108.3R. Fiecare reacție a conținut aproape 100 ng ADN matriță și 25 pmol dintr-un depozit primer și s-a procesat conform protocoalelor standard, așa cum s-a descris pentru peptida TcaC. O analiză a secvenței derivate de la extinsia primerilor P8F demonstrează că secvența ADN este scurtă (și secvența aminoacidă codificată):	3
GAT GCA TTG CTT GCT	5
Asp Ala Leu (Val) Ala	7
care corespunde la o porțiune a secvenței peptidei N-terminale descrisă ca SEQ ID NR:3 (TcaB _i).	9
<i>Marcarea probei peptidă-tcaB_i specifică genă</i>	11
Aproximativ 50 ng ADN fragment TcaB _i purificat pe gel s-a marcat cu ³² P-dCTP cum s-a descris mai sus și radioizotopii neîncorporați s-au îndepărtat prin trecere prin NICK Column® (Pharmacia). Activitatea specifică a ADN s-a marcat de 6 x 10 ⁹ dpm/μg. Acest ADN marcat s-a folosit pentru a testa membranele colonii preparate de la membranele băncii genomice care s-au hibridizat la proba specifică peptidă-TcaC.	13
Membranele care conțin 12 colonii identificate în cercetarea băncii probă-TcaC (vezi mai sus) s-au îndepărtat de marcajul specific-TcaC radioactiv, prin fierbere de două ori timp de aproximativ 30 min, de fiecare dată într-un litru de 0,1 X SSC plus 0,1% SDS. Îndepărtarea radiomarcajului s-a verificat cu o expunere pe film de 6 h. Membranele îndepărtate s-au incubat apoi cu proba specific-peptidă TcaB _t preparată mai sus. ADN marcat s-a denaturat prin fierbere timp de 10 min și apoi s-a adăugat la filtratele care fuseseră incubate timp de o oră, în 100 ml soluție "minimal hyb", la 60°C. După hibridizarea peste noapte, la această temperatură, soluția probă s-a îndepărtat și filtratele s-au spălat așa cum urmează (toate în 0,3X SSC plus 0,1% SDS): o dată timp de 10 min, la 25°C, o dată timp de o oră, la 60°C în soluție proaspătă și o dată timp de oră la 63°C în soluție proaspătă. După 1,5 h de expunere la film raze X, după procedurile standard, s-au observat 4 colonii puternic hibridizate. Acestea au fost, așa ca și în cazul probei specifice TcaC, izolatele 22G12, 25A10, 26A5 și 26B10.	15
Aceeași soluție probă TcaB _i s-a diluat cu un volum egal (aproximativ 100 ml) soluție "minimal hyb" și apoi s-au folosit pentru a cerceta membranele care conțin 800 membrii ale băncii genomice. După hibridizare, spălare și expunere la film raze X, așa cum s-a descris mai sus, s-au găsit doar patru clone cosmid 22G12, 25A10, 26A5 și 26b10, care să hibridizeze puternic la această probă.	17
Izolarea sublonelor care conțin gene care codifică peptidele TcaC și TcaB _i și determinarea secvenței de bază ADN a acestora:	19
S-au crescut cu agitare peste noapte (200 rpm) la 30°C, în 100 ml TB-Amp ₁₀₀ 3 cosmizi hibridizați pozitiv în tulpina XL1 Blue MR. După recoltarea celulelor prin centrifugare ADN cosmid s-a preparat folosind un kit accesibil comercial (BIGprep™, 5 Prime 3 Prime, Inc., Boulder, CO), folosind protocolul producătorului. Doar un cosmid, 26A5, s-a izolat cu succes prin această procedură. Când s-a digerat cu enzima de restricție EcoR 1 (NEB) și s-a analizat prin electroforeză pe gel, s-au detectat fragmente de mărimi aproximative 14, 10, 8 (vector), 5, 3,3, 2,9 și 1,5 kbp. O a doua încercare de a izola ADN cosmid din aceleași trei tulpini (8 ml cultură; TB-Amp ₁₀₀ , 30°C, folosind metoda miniprep fierbere (Evans G. și	21

G.Wahl., 1987, *Cosmid vectors for genomic walking and rapid restriction mapping* în *Guide to Molecular Cloning Techniques. Meth. Enzymology*, vol. 152, S. Berger și A. Kimmel, eds., pg. 604-610). Doar un cosmid, 25A10, s-a izolat cu succes prin această metodă. Când s-a digerat cu enzima de restricție EcoR 1 (NEB) și s-a analizat prin electroforeză pe gel, acest cosmid a arătat un model de fragmentare identic cu cel observat mai devreme la cosmidul 26 A5.

O probă de 0,15 µg ADN cosmid 26A5 s-a utilizat pentru a transforma 50 ml de celule DH5α E.coli (Gibco BRL) prin protocolul furnizorului. Un izolat colonie individual al acestei tulpini s-a inoculat în 4 ml de TB-Amp₁₀₀ și s-a crescut 8 h, la 37°C. S-a adăugat cloramfenicol până la concentrația finală de 225 µg/ml, incubarea s-a continuat pentru încă 24 h, apoi celulele s-au recoltat prin centrifugare și s-au înghețat la -20° C. Izolarea ADN-ului cosmid 26 A5 s-a făcut printr-o miniprep liză alcalină standard (Maniatis et al., op.cit, p.382), modificată prin creșterea tuturor volumelor prin 50% și cu agitare sau amestecare blândă, mai degrabă decât prin învârtire, la fiecare etapă. După spălarea peletului ADN în 70% etanol, s-a dizolvat în TE care conține 25 (ag/ml ribonuclează A (Boehringer Mannheim).

Identificarea fragmentelor EcoR 1 care hibridizează la probele GZ4-derivată și TcaB_i. Aproximativ 0,4 µg de 25A10 cosmid (din celule XL1 Blue MR) și aproximativ 0,5 µg de 26A5 cosmid (din celule DH5α amplificate cloramfenicate) s-au digerat fiecare cu aproape 15 unități de EcoR 1 (NEB), timp de 85 min, s-au înghețat peste noapte, apoi s-au încălzit la 65°C, pentru 5 min, și s-au supus electroforezei în gel de agaroză 0,7% (Seakem®, 1X TEA, 80 volți, 90 min). ADN-ul s-a colorat cu bromură de etidiu, așa cum s-a descris mai sus, și s-a fotografiat sub lumină ultravioletă. Digerarea cu EcoR 1 a cosmidului 25A10 a fost o digereare completă, dar proba cosmidului 26A5 a fost digerată parțial în aceste condiții. Gelul de agaroză care conține fragmentele ADN s-a supus la depurare, denaturare și neutralizare, urmate de Southern blot pe membrane de nylon Magna NT, utilizând un protocol sare ridicat (20X SSC), toate așa cum s-a descris în secțiunea 2.9 a lui Ausubel et al. (CPMB, op.cit). ADN-ul transfectat s-a crosslinkat UV la membrana de nylon ca mai înainte.

Un fragment ADN specific peptidei TcaC care corespunde insertului izolatului GZ4 plasmid, s-a amplificat prin PCR® într-un volum de reacție de 100 ml așa cum s-a descris mai sus mai devreme. Produsele amplificate de la 3 astfel de reacții s-au depozitat și s-au extras de la gel de agaroză 1% GTG® prin kit Qiaex, așa cum s-a descris mai sus, și s-a determinat cantitativ prin fluorometrie. ADN-ul purificat-gel (100 ng) s-a marcat cu ³²P-dCTP, utilizând High Prime Labeling Mix (Boehringer Mannheim) așa cum s-a descris mai sus, la o activitate specifică de 6,3 4x10⁸ dpm/µg.

Proba GZ4 marcată ³²P s-a fiert 10 min, apoi s-a adăugat la tampon "minimal hyb" (la 1 ng/ml) și s-a adăugat membrana Southern blot care conține fragmentele ADN cosmid digerate și s-a incubat timp de 4 h, la 60°C, cu scuturare blândă la 50 rpm. Apoi, membrana s-a spălat de 3 ori la 25°C, aproape 5 min de fiecare dată (soluție de spălare "minimal hyb"), urmată de două spălări timp de 30 min, fiecare la 60°C. Blot-ul s-a expus la film (cu cercetare amplificată) timp de aproape 30 min, la -70°C. Proba GZ4 a hibridizat puternic la 5,0 kbp (mărime aparentă) fragmentul EcoR 1 al ambilor doi cosmizi, 26A5 și 25A10.

S-a îndepărtat radioactivitatea de pe membrană, prin fierbere timp de aproximativ 30 min în 0,1X SSC plus 0,1% SDS și absența radiomarcajului s-a verificat prin expunere la film. Apoi, s-a hibridizat la 60°C, timp de 3,5 h cu proba TcaB_i (denaturată) în tampon "minimal hyb" folosit anterior pentru cercetarea membranelor colonii (mai sus), s-a spălat așa cum s-a descris mai devreme și s-a expus la film timp de 40 min, la -70°C, cu doi amplificatori de cercetare. Proba TcaB_i a hibridizat, cu ambii cosmizi, ușor cu un fragment EcoR 1 de aproximativ 5,0 kbp și puternic, cu un fragment de aproximativ 2,9 kbp.

RO 121280 B1

Proba ADN cosmid 26A5 descrisă mai sus (din celule DH5α) s-a utilizat ca sursă a ADN-ului de la care s-au subclonat benzile de interes. Acest ADN (2,5 μg) s-a digerat cu aproximativ 3 unități de EcoR 1 (NEB) într-un volum total de 30 μl pentru aproximativ 1,5 h, pentru a obține o digestie parțială, așa cum s-a confirmat prin electroforeza pe gel. S-au digerat 10 μg ADN pBC KS (+) (Stratagene) timp de 1,5 h cu 20 unități EcoR 1 într-un volum total de 20 μl, conducând la digestia totală, așa cum s-a confirmat prin electroforeză. Ambele preparate ADN EcoR 1 - tăiat sau diluat până la 50 μl cu apă, la fiecare s-a adăugat un volum egal de PCI, suspensia s-a amestecat blând, s-a filat într-o microcentrifugă și s-a colectat supernatantul apos. S-a precipitat ADN cu 150 μl etanol și amestecul s-a plasat la -20°C, peste noapte. Urmând centrifugarea și uscarea, pBC KS (+) digerat EcoR 1 s-a dizolvat în 100 μl TE; 26A5 digerat parțial s-a dizolvat în 20 μl TE. Recuperarea ADN-ului s-a verificat prin fluorometrie.

În reacții separate, ADN pBC KS (+) digerat EcoR 1 s-a ligat cu aproximativ 180 ng sau 270 ng de ADN cosmid 26A5 digerat parțial. Ligările s-au realizat într-un volum de 20 μl, la 15°C, timp de 5 h, folosind ligază T4 și tampon de la New England BioLabs. Amestecul de ligare, diluat la 100 μl cu steril TE, s-a utilizat pentru a transforma celule DH5α competente, înghețate (Gibco BRL) în conformitate cu instrucțiunile furnizorului. Cantități variabile (25-200 μl) de celule transformate s-au depus pe mediu LB-Cam₃₅ solid proaspăt preparat cu 1 mM IPTG și 50 mg/l X-gal. Plăcile s-au incubat la 37°C, aproape 20 h, apoi s-au răcit în întuneric timp de aproximativ 3 h, pentru intensificarea culorii, în vederea selectării inser-tului. Coloniile albe s-au cules pe restul plăcii aceleiași compoziții și s-au incubat peste noapte la 37°C.

Două proeminente de colonii, de pe fiecare rest de placă selectată, s-au preparat după cum urmează. După ridicarea coloniilor albe la plăci proaspete, membrane de nylon Magna NT rotunde s-au presat pe resturile plăcilor, membrana s-a desprins și s-a supus denaturării, neutralizării și crosslink-ării UV, așa cum s-a descris mai sus pentru membranele colonie ale băncii. Proeminența coloniei crosslink-ate s-a spălat viguros incluzând eliminarea prin ștergere blândă a reziduurilor celulare în exces cu un țesut. Un set s-a hibridizat cu soluție probă GZ4 (TcaC) descrisă mai devreme și celălalt set s-a hibridizat cu soluția probă TcaB, descrisă mai devreme, în conformitate cu protocolul "minimal hyb", urmând spălarea și expunerea la film, așa cum s-a descris pentru membranele colonie ale băncii.

Coloniile care au arătat semnale de hibridizare, fie numai cu proba GZ4, fie cu ambele probe GZ4 și TcaB, sau numai cu proba TcaB, s-au selectat pentru prelucrare ulterioară și celulele s-au striat pentru izolarea coloniei individuale pe mediu LB-Cam₃₅ cu IPTG și X-gal, ca mai înainte. Aproximativ 35 colonii individuale din 16 izolate diferite au fost sepa-rate în mediul LB-Cam₃₅ lichid și s-au crescut peste noapte la 37°C; celulele s-au centrifugat prin centrifugare și s-a izolat ADN plasmid prin miniprep liză alcalină standard în conformitate cu Maniatis et al., (ap.cit.p.368). Peletele ADN s-au dizolvat în TE + 25 μg/ml ribonuclează A și concentrația ADN s-a determinat prin fluorometrie. Modelul de digerare EcoR 1 s-a ana-lizat prin electroforeză pe gel. Următoarele izolate s-au selectat ca folositoare. Izolatul A 17.2 conține doar pBC KS (+) religat și s-a folosit pentru un control (negativ). Izolatele D38.3 și C44.1 fiecare conținând doar 2,9 kbp, fragmentul TcaB, care hibridizează EcoR 1, s-au inserat în pBC KS (+). Acești plasmizi denumiți pDAB2000 și, respectiv, pDAB2001 sunt ilustrați în fig. 2.

Izolatul A35.3 conține doar aproximativ 5 kbp, fragmentul GZ4- care hibridizează EcoR 1 s-au inserat în pBC KS (+). Acest plasmid s-a denumit pDAB2002 (de asemenea, în fig. 2). Aceste izolate furnizează matrițe pentru secvențarea ADN.

1 Plasmidii pDAB2000 și pDAB2001 s-au preparat folosind kitul BIGprep™, ca mai sus.
 2 Culturile (30 ml) s-au crescut peste noapte în TB-Cam₃₅, la un OD₆₀₀ din 2, apoi plasmidul
 3 s-a izolat conform cu îndrumările producătorului. Peletele ADN s-au dizolvat fiecare în 100 µl
 4 TE și integritatea probei s-a verificat prin digestie cu EcoR 1 și prin analiză electroforetică
 5 pe gel.

6 Reacțiile de secvențare s-au efectuat în duplicat, cu un replicat folosind ca matriță
 7 ADN pDAB2000 și celălalt replicat folosind ca matriță ADN pDAB2001. Reacțiile s-au realizat
 8 folosind metoda de secvențare ciclu de terminare colorant dideoxi, așa cum s-a descris mai
 9 sus pentru secvențarea ADN-urilor GZ4/HB14. Secvențarea inițială s-a desfășurat utilizând,
 10 ca primeri, primerii LacZ și T7 descriși mai sus, plus primeri bazați pe secvența determinată
 11 a produsului amplificat PCR TcaB_i (TH1 = ATTGCAGACTGCCAATCGCTTCGG, TH12 =
 12 GAGAGTATCCAGACGCCGGATGATCTG).

13 După alinierea și editarea fiecărui produs de secvențare, fiecare s-a trunctat la între
 14 250 la 350 baze, în funcție de integritatea datelor cromatografice, așa cum s-a interpretat
 15 prin software-ul Perkin Elmer Applied Biosystem Division SeqEd 675. "Etapetele" de secven-
 16 țare care au urmat s-au realizat prin selectarea secvenței corespunzătoare pentru noii pri-
 17 meri. Cu câteva excepții, primerii (sintetizați cum s-a descris mai sus) au fost lungi de 24
 18 baze cu o compoziție G+C 50%. Secvențarea prin această metodă s-a realizat la ambele
 19 catene ale fragmentului EcoR 1 de aproximativ 2,9 kbp.

20 Pentru a servi în continuare ca matriță pentru secvențarea ADN, ADN plasmid din
 21 izolatul pDAB2002 s-a preparat prin kitul BIGperp™. Reacțiile de secvențare s-au realizat
 22 și analizat așa cum s-a descris mai sus. Inițial, un primer T3 (pBS SK (+) bazele
 23 774-796:CGCGCAATTAACCTCACTAAAG) și un primer T7 (pBS KS (+) bazele 621-643:
 24 GCGCGTAATACGACTCACTATAG) s-au utilizat pentru a prima reacțiile de secvențare de
 25 la secvențele vectorului de flancare citite în ADN inserat. Un alt set de primeri, (GZ4F:
 26 GTATCGATTACAACGCTGTCACTTCCC; TH13: GGGAAGTGACAGCGTTGTAA
 27 TCGATAC; TH14: ATGTTGGGTGCGTCCGCTAATGGACATAAC; și LW1-204:
 28 GGGAAGTGACAGC GTTGTAAATCGATAC) s-a construit pentru a prima de la secvențele
 29 interne, care s-au determinat anterior prin secvențarea mediată-oligonucleotidă degenerată
 30 a produselor PCR peptide TcaC. De la datele generate în timpul rundelor inițiale de secven-
 31 țare, s-a desemnat și s-a utilizat un nou set de primeri, pentru a străbate întreaga lungime
 32 a fragmentului de ~5 kbp. S-a utilizat un total de 55 oligoprimeri, care au permis identificarea
 33 tuturor celor 8432 bp ale secvenței contigue.

34 Când secvența ADN a fragmentului insert EcoR 1 a pDAB2002 se combină cu partea
 35 secvenței determinată a izolatelor pDAB2000 și pDAB2001, s-a generat un total de 6005 bp
 36 de secvență contiguă (descrisă aici ca SEQ ID NR: 25). Când cadrele de citire lungi, deschi-
 37 se, s-au translat în aminoacizii corespunzători, secvența arată clar peptida TcaB; N-terminală
 38 (descrisă ca SEQ ID NR: 3) codificată de bazele 19-75, urmând imediat restul metionihă. În
 39 sus, se află un sit de legare ribozom potențial (bazele 1-9) și în jos, la bazele 166-288, este
 40 codificată peptida internă TcaB_i-PT158 (descrisă aici ca SEQ ID NR: 19). În continuare, în
 41 jos, în același cadru de citire, la bazele 1738-1773, există o secvență care codifică peptida
 42 internă TcaB_j-PTIOS (descrisă aici ca SEQ ID NR: 20). De asemenea, în același cadru de
 43 citire, la bazele 1897-1923 este codificată peptida TcaB_n N-terminală (descrisă aici ca SEQ
 44 ID NR: 5) și cadrul de citire continuă neîntrerupt, până la codonul de terminare translație la
 45 nucleotidele 3586-3588.

46 Lipsa unui codon stop în cadru, între sfârșitul secvenței care codifică TcaB_i-PT108
 47 și începutul regiunii care codifică TcaB_{ii}, și lipsa unui sit de legare ribozom imediat mai sus
 de regiunea care codifică TcaB_{ii}, sugerează că peptidele TcaB_i și TcaB_{ii} sunt codificate de

RO 121280 B1

un singur cadru de citire de 3567 bp, care începe la baza pereche 16 în SEQ ID NR:25, și, foarte probabil, derivă dintr-un produs genă primar individual de 1198 aminoacizi (131.586 Daltoni; descris aici ca SEQ ID NR:26), printr-un clivaj posttranslațional. Dacă aminoacidul care precedă imediat peptida TcaB_i N-terminală reprezintă aminoacidul C-terminal al peptidei TcaB_i, atunci masa previzionată TcaB_i (627 aminoacizi) este 70.814 Daltoni (descrisă aici ca SEQ ID NR: 28), ceva mai mare decât mărimea observată prin SDS-PAGE (68 kDa). Această peptidă ar fi codificată printr-o întindere contiguă de 1881 baze perechi (descrisă aici ca SEQ ID NR: 27). S-a considerat că C-terminusul nativ al TcaB_i este cumva mai apropiat de C-terminusul TcaB_i-PT108. Masa moleculară a PT108 (3,438 kDa; determinată în timpul analizei secvenței aminoacide N-terminală a acestei peptide) prevede o mărime de 30 aminoacizi. Utilizând mărimea acestei peptide pentru a desemna C-terminusul regiunii care codifică TcaB_i [Glu la poziția 604 a SEQ ID NR: 28], mărimea derivată a TcaB_i este determinată a fi de 604 aminoacizi sau 68.463 Daltoni, în conformitate cu observațiile experimentale.

Translarea regiunii de 1686 baze perechi, care codifică peptida TcaB_i (descrisă aici ca SEQ ID NR: 29), a dat o proteină de 562 aminoacizi (descrisă aici ca SEQ ID NR: 30), cu masa previzionată de 60.789 Daltoni, care corespunde bine cu cea observată de 61 kDa.

Un sit potențial de legare ribozom (bazele 3633-3638) este găsit 48 bp mai jos de codonul stop pentru cadrul deschis de citire tcaB. La bazele 3645-3677, s-a găsit o secvență care codifică N-terminusul peptidei TcaC, (descrisă aici ca SEQ ID NR:2). Cadrul deschis de citire inițiat de peptida N-terminală continuă neîntrerupt până la baza 6005 (2361 baze perechi, descrise aici ca primele 2361 baze perechi ale SEQ ID NR: 31). O genă (tcaC) care codifică întreaga peptidă TcaC, (mărime aparentă -165 kDa; -1500 aminoacizi) ar cuprinde aproximativ 4500 bp.

Alt izolat care conține fragmentele EcoR 1 donate ale cosmidului 26A5, E20.6 s-a identificat, de asemenea, prin omologie cu probele GZ4 și TcaB_i menționate anterior. Analiza pe gel de agaroză, a digeratelor EcoR 1 ale ADN-ului plasmidului adăpostit de această tulpină (pDAB2004, fig. 2), a arătat fragmentele inserate ale mărimilor estimate 2.9, 5 și 3,3 kbp. Analiza secvenței ADN inițiată de la primerii desemnați de la secvența plasmidului pDAB2002 a arătat că fragmentul de 3,3 kbp EcoR 1 al pDAB2004 este adiacent fragmentului de 5 kbp EcoR 1 reprezentat în pDAB2002. Cadrul deschis de citire de 2361 baze perechi descoperit în pDAB2002 continuă neîntrerupt pentru încă 2094 baze în pDAB2004 [descrise aici ca bazele perechi de la 2362 la 4458 ale SEQ ID NR:31]. Analiza secvenței ADN utilizând ADN cosmid 26A5 primar ca matriță a confirmat continuitatea cadrului deschis de citire. Împreună, cadrul deschis de citire (TcaC SEQ ID NR:31) cuprinde 4455 baze perechi și codifică o proteină (TcaC) de 1485 aminoacizi [descrisă aici ca SEQ ID NR:32]. Mărimea moleculară calculată de 166.214 daltoni este conformă cu mărimea estimată a peptidei TcaC (165 kDa) și secvența aminoacidă derivată se potrivește exact cu cea descrisă pentru secvența TcaC N-terminală [SEQ ID NR:2].

Lipsa unei secvențe aminoacide corespunzătoare SEQ ID NR:17, folosită pentru a descrie depozitul primer oligonucleotidă degenerată în secvența descoperită, arată că generarea produselor PCR[®] găsită în izolatele GZ4 și HB14, care s-au utilizat ca probe în cercetarea băncii inițiale, au fost generate întâmplător prin primarea catenei-inverse de unul din primeri în depozitul degenerat. În continuare, secvența proteinei derivate nu include fragmentul intern descris aici ca SEQ ID NR:18. Aceste secvențe arată că plasmidul pDAB2004 conține întreaga regiune care codifică pentru peptida TcaC.

Exemplul 9. Cercetarea băncii genomice *Photorhabdus* pentru genele care codifică peptida TcbA_{ii}

Acest exemplu descrie o metodă utilizată pentru a identifica clonele ADN care conțin genele care codifică peptida TcbA_{ii}, izolarea acestor gene și determinarea secvenței lor parțiale de baze ADN.

Poli-peptida TcbA_{ii} preparării active insectă este - 602 kDa. Secvența aminoacidică a N-terminusului acestei peptide este descrisă ca SEQ ID NR:1. Patru depozite de primeri oligonucleotide degenerate ("primeri forward": TH-4, TH-5, TH-6 și TH-7) s-au sintetizat pentru a codifica o porțiune a acestei secvențe aminoacide, așa cum s-a descris în exemplul 8 și sunt descrise și mai jos.

Tabelul 11

Primerii și reacțiile PCR

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

În plus, o secvență primară ("a") și una secundară ("b") ale preparării peptidei interne (TcbA_{ii}-PT81) s-au determinat și s-au descris aici ca SEQ ID NR:23 și, respectiv, SEQ ID NR:24. Patru depozite de oligonucleotide degenerate ("primeri reverse": TH-8, TH-9, TH-10 și TH-11) s-au construit și s-au sintetizat pentru a codifica complementul invers al secvențelor care codifică o porțiune a peptidei SEQ ID NR:23, așa cum se arată mai jos.

Tabelul 12

Amino acid	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'

Seturi din acești primeri s-au utilizat în reacții PCR®, pentru a amplifica fragmentele de genă care codifică TcbA_{ii} de la ADN genomic *Photorhabdus luminescens* W-14 preparat în exemplul 6. Toate reacțiile PCR s-au realizat cu tehnica "Hot Star" folosind muguri AmpliWax™ și alți reactivi și protocoale Perkin Elmer, de obicei, un amestec (volum total 11 μl) de MgCl₂, dNTP's, tampon PCR II10X GeneAmp® și s-au adăugat primerii la eprubete care conțin un singur bob de ceară, [tamponul PCR II10X GeneAmp® este compus din 10 mM Tris-HCl, pH 8,3 și 500 mM KC1]. Eprubetele s-au încălzit 2 min, la 80°C, și s-au lăsat să se răcească. S-a adăugat o soluție care conține tampon PCR II10X GeneAmp®, matriță ADN și polimerază ADN, la vârful dopului etanș de ceară. După topirea dopului etanș

RO 121280 B1

de ceară și amestecarea componentelor prin ciclare termică, condițiile de reacție finale (volum de 50 μ l) au fost: 10 mM Tris-HCl, pH 8,3; 50 mM KCl; 2,5 mM $MgCl_2$; 200 μ M fiecare din dATP, dCTP, dGTP, dTTP; 1,25 mM într-un singur depozit primer forward; 1,25 μ M într-un singur depozit primer revers, 1,25 unități polimerază ADN AmpliTaq® și 170 ng matriță AND.

Reacțiile s-au plasat într-un termocicler (ca în exemplul 8) și s-au desfășurat cu următorul program:

Tabelul 13

Temperatură	Temp	Repere ciclu
94°C	2 min	1X
94°C 55-65°C 72°C	15 s 30 s 1 min	30X
72°C	7 min	1X
15°C	Constant	

O serie de amplificări s-a desfășurat la aceste trei temperaturi normalizatoare (55°, 60°, 65°C), utilizând depozitele primeri degenerați. Reacțiile cu normalizare la 65°C nu au avut produse de amplificare vizibile, după electroforeza pe gel de agaroză. Reacțiile având regim de normalizare de 60°C și care conțin primerii TH-5+TH-10 au avut drept rezultat un produs de amplificare care a avut o mobilitate corespunzătoare la 2,9 kbp. O cantitate de produs mai mică de 2,9 kbp s-a produs în aceste condiții cu primerii TH-7+TH-10. Când reacțiile s-au normalizat la 55°C, aceste perechi de primeri au produs mai mult decât un produs de 2,9 kbp și acest produs s-a produs, de asemenea, cu perechile de primeri TH-5+TH-8 și TH-5+TH-11. În plus, benzi de 2,9 kbp foarte pale s-au observat în pasajele care conțin produse amplificate de la perechile de primeri TH-7 plus TH-8, TH-10 sau TH-11.

Pentru a obține un produs amplificat PCR suficient pentru donarea și determinarea secvenței ADN, 10 reacții PCR separate s-au inițiat folosind primerii TH-5+TH-10 și s-au desfășurat utilizând condițiile de mai sus cu normalizare la 55°C. Toate reacțiile s-au depozitat și produsul de 2,9 kbp s-a purificat prin extracție Qiaex de la gel de agaroză, așa cum s-a descris mai sus.

Secvențe adiționale determinate pentru peptida internă TcbAj sunt descrise aici ca SEQ ID NR: 21 și SEQ ID NR:22. Ca mai înainte, oligonucleotidele degenerare (primerii reverse TH-17 și TH-18) s-au realizat în conformitate cu complementul invers care codifică o porțiune a secvenței aminoacide a acestor peptide.

Tabelul 14

De la SEQ ID NR: 21

Amino acid	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'

De la SEQ ID NR:22

Amino acid	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'

Oligonucleotidele degenerate TH-18 și TH-17 s-au utilizat în experimentul de amplificare cu ADN *Photorhabdus hammescens* W-14 ca matriță și primerii TH-4, TH-5, TH-6 sau TH-7 ca primeri 5' (forward). Aceste reacții au amplificat produși de aproximativ 4 kbp și, respectiv, 4,5 kbp. Aceste ADN-uri s-au transferat de la gel de agaroză pe membrane nylon și s-au hibridizat cu proba marcată 32P (așa cum s-a descris mai sus), preparată de la produsul de 2,9 kbp amplificat prin perechea primer TH-5+TH-10. Atât produsul de amplificare de 4 kbp, cât și cel de 4,5 kbp au hibridizat puternic cu proba de 2,9 kbp. Aceste rezultate s-au utilizat pentru a construi o hartă pentru secvența peptidei interne TcbA_{ii}, așa cum s-a arătat în fig. 3. Distanța aproximativă dintre primeri este prezentată în fig. 3.

Secvența ADN a fragmentului de 2,9 kbp care codifică TcbA_{ii}

S-au precipitat cu etanol aproximativ 200 ng din fragmentul de 2,9 kbp (preparat mai sus) și s-au dizolvat în 17 ml apă. O jumătate din acesta s-a utilizat ca matriță de secvențare cu 25 pmol din depozitul TH-5 ca primer și cealaltă jumătate s-a utilizat ca matriță pentru primarea TH-10. Reacțiile de secvențare au fost așa cum s-a arătat în exemplul 8. Nu s-a produs nici o secvență de bază folosind depozitul primer TH-10; totuși, reacțiile cu depozit primer TH-5 au produs secvența descrisă mai jos:

```

1   AATCGTGTTG ATCCCTATGC CGNGCCGGGT TCGGTGGAAT CGATGTCCTC
23  ACCGGGGGTT
25  61   TATTNGAGGG ANTNGTCCCG TGAGGCCAAA AANTGGAATG AAAGAAGTTC
    AAATTTNTTAC
27  121  CTAGATAAAC GTAGCCCGGN TTTAGAAAGN TTANTGNTCA
    GCCAGAAAATTTTGGTTGAG
29  181  GAAATTCAC CGNTGGTCT CTCTATTGAT TNGGGCCTGG CCGGGTTCGA
    ANNAAAACNA
31  241  GAAATNCAC AAGTTGAGGTGATGGNTTGTNGCNANCTTNTCGTTTAGGT
    GGGGAGAAA
33  301  CCTTNTCANCACGNTTNTGAACTGACCGGGAAATCGTCCATGANCGTGANC
    CAGGNTTN
35  361  CGCCATTGG

```

Bazat pe această secvență, un primer de secvențare (TH-21, 5'-CCGGGCGACGTTTATC TAGG-3') s-a proiectat pentru a inversa bazele complement 120-139 și a iniția polimerizarea spre capătul 5' (adică, TH-5) al fragmentului PCR de 2,9 kbp purificat-gel care codifică TcbA_{ii}.

Secvența determinată este prezentată mai jos și este comparată cu secvența peptidei N-terminală determinată biochimic a TcbA_{fi} SEQ ID NR: 1.

Confirmarea secvenței fragmentului PCR 2,9 kbp TcbA_{ii} [Aminoacizii subliniați=codificați de oligonucleotidele degenerate].

```

45  SEQ  FIQ GYS D L F G      -      -      A

```

```

47  ID NR:1      |      || || || || || ||

```

```

      GCATG CAG GGG TATAGTGAC CTGTTT GGT AATCGTGCT

```

```

      M Q      GY      S D L F G N R A

```

RO 121280 B1

De la omologia secvenței de aminoacizi, derivată cu cea determinată biochimic, a fost clar că fragmentul PCR de 2,9 kbp reprezintă regiunea care codifică TcbA. Acest fragment de 2,9 kbp s-a utilizat apoi ca probă de hibridizare pentru cercetarea băncii genomice *Photorhabdus* W-14 preparată în exemplul 8, pentru cosmizii care conțin gena care codifică TcbA_{ii}.

Cercetarea băncii cosmid *Pphotorhabdus*

Fragmentul PCR purificat-gel de 2,9 kbp s-a marcat cu ³²P folosind kitul de marcare Boehringer Mannheim High Prime, așa cum s-a descris în exemplul 8. Filtrele care conțin remanenți de la aproximativ 800 de colonii din banca cosmidului s-au cercetat așa cum s-a descris mai sus (exemplul 8) și clonele pozitive s-au dungat pentru izolarea coloniilor și s-au recercetat. 3 clone (8A11, 25G8 și 26D1) au dat rezultate pozitive pe parcursul mai multor etape de cercetare și caracterizare. Nu s-a observat hibridizarea probei specifice TcbA_{ii} cu nici unul dintre cei 4 cosmizi identificați în exemplul 8 și care conțin genele tcaB și tcaC. ADN din cosmizii 8A11, 25G8 și 26D1 s-au digerat cu enzimele de restricție Bgl 2, EcoR 1 sau Hind 3 (fiecare singur sau în combinație cu un altul) și fragmentele s-au separat pe un gel de agaroză și s-au transferat pe o membrană nylon, așa cum s-a descris în exemplul 8. Membrana s-a hibridizat cu proba marcată ³²P preparată de la fragmentul de 4,5 kbp (generat prin amplificarea ADN genomic de la *Photorhabdus* cu primerii TH-5+TH-17). Modelele generate de la ADN cosmizi 8A11 și 26D1 au fost identice cu acelea generate cu ADN genomic tăiat-similar pe aceeași membrană. S-a conchis că respectivii cosmizi 8A11 și 26D1 sunt reprezentările precise ale TcbA_{ii} genomic care codifică locus. Totuși, cosmidul 25G8 are un fragment individual Bg 1 care este puțin mai mare decât ADN-ul genomic. Aceasta poate rezulta din poziționarea insertului în vector.

Secvența ADN a genei care codifică tcbA

Analiza hibridizării membranei cosmidului 26D1 a relevat ca proba de 4,5 kbp a hibridizat la un singur fragment EcoR 1 mare (mai mare decât 9 kbp). Acest fragment s-a purificat-gel și s-a ligat în situl EcoR 1 al pBC KS (+), așa cum s-a descris în exemplul 8, pentru a genera plasmidul pBC-SI/RI. Secvența ADN parțială a insertului ADN a acestui plasmid s-a determinat prin "plimbarea primerului" de la secvența vectorului de flancare, utilizând procedura descrisă în exemplul 8. În continuare, secvența s-a generat prin extinderea de la noile oligonucleotide proiectate de la secvența determinată mai devreme. Când s-a comparat cu secvența ADN determinată pentru gena tcbA identificată prin alte metode (descrisă aici ca SEQ ID NR:11, așa cum s-a descris în exemplul 12 mai jos), s-a găsit omologie completă la nucleotidele 1-272, 319-862, 2578-3036 și 3068-3540 (total de baze=1712). S-a conchis că ambele abordări pot fi utilizate pentru a identifica fragmentele ADN care codifică peptida TcbA_{ii}.

Analiza secvenței aminoacide derivate a genei tcbA

Secvența fragmentului ADN identificat ca SEQ ID NR:11 codifică o proteină a cărei secvență aminoacidică derivată este descrisă aici ca SEQ ID NR:12. Câteva trăsături verifică identitatea genei cu cea care codifică proteina TcbA_{ii}. Peptida N-terminală TcbA_{ii} (SEQ ID NR: 1; Phe Ile Gin Gly Tyr Ser Asp Leu Phe Gly Asn Arg Ala) este codificată de aminoacizii 88-100. TcbA_{ii}-PT81 (a) peptida internă TcbA_{ii} (SEQ ID NR:23) este codificată de aminoacizii 1065-1077 și TcbA_{ii}-PT81 (b) (SEQ ID NR:24) este codificată ca aminoacizii 1571-1592. Mai departe, peptida internă TcbA_{ii}-PT56 (SEQ ID NR: 22) este codificată ca aminoacizii 1474-1488 și peptida internă TcbA_{ii}-PT103 (SEQ ID NR:24) este codificată ca aminoacizii 1614-1639. Este evident că această genă este o clonă autentică care codifică peptida TcbA_{ii}, așa cum s-a izolat de la prepararea proteinei insecticide a *Photorhabdus luminescens* tulpina W-14.

Proteina izolată ca peptida TcbA_{ii} este derivată de la clivajul unei peptide mai lungi. O dovadă pentru aceasta este furnizată de faptul că nucleotidele care codifică peptida TcbA_{ii} N-terminală SEQ ID NR:1 sunt precedate de 261 baze (care codifică 87 aminoacizii N-terminali-proximali) ale unui cadru deschis de citire mai lung (SEQ ID NR:11). Acest cadru de citire începe cu oligonucleotidele care codifică secvența aminoacidă Met Gin Asn Ser Leu, care corespunde cu secvența N-terminală a peptidei TcbA mai mari și descrisă aici ca SEQ ED NR: 16. S-a considerat că TcbA este proteina precursoră pentru TcbA_{ii}.

Relațiile genelor tcbA, tcaB și tcaC

Genele tcaB și tcaC sunt linkate aproape și pot fi transcrise ca un singur mRNA (exemplul 8). Gena tcbA este apărută pe cosmidul care aparent nu îi înfășoară pe cei care adăpostesc clusterul tcaB și tcaC, de vreme ce cercetarea băncii genomice respective a identificat cosmizi diferiți. Totuși, compararea secvențelor amino codificate de genele tcaB și tcaC cu gena tcbA a relevat un grad substanțial de omologie. Conservarea aminoacidului (Protein Alignment Mode of MacVector™ Sequence Analysis Software, care marchează matricea pam250, valoare tăiată=2; Kodak Scientific Imaging Systems, Rochester, NY) este prezentată în fig. 4. Pe linia de marcat a fiecărui tablou din fig. 4, semnele "A" indică omologia sau conservarea schimbărilor aminoacide și semnele "V" indică nonomologia.

Această analiză arată că secvența aminoacidă a peptidei TcbA din resturile 1739 la 1894 este omoloagă ridicat cu aminoacizii 441 la 603 ai peptidei TcaB_i (162 dintr-un total de 627 aminoacizi ai P8; SEQ ID NR:28). În plus, secvența aminoacizilor TcbA 1932 la 2459 este omoloagă înalt cu aminoacizii 12 la 531 ai peptidei TcaB_{ii} (520 din totalul de 562 aminoacizi; SEQ ID NR: 30). Considerând că peptida TcbA (SEQ ID NR:12) cuprinde 2595 aminoacizi, un total de 684 aminoacizi (27%), la capătul C-proximal al acesteia, este omolog cu peptidele TcaB_i și TcaB_{ii} și omoloagele sunt aranjate colinier pe aranjamentul preproteinei putative TcaB (SEQ ID NR:26). Un gol măsurabil în omologia TcaB_{ii} coincide cu joncțiunea dintre porțiunile TcaB_i și TcaB_{ii} ale preproteinei TcaB. Evident, produsele genei TcbA și TcaB sunt legate evolutiv și s-a propus că ele au funcție (i) comună în *Photorhabdus*.

Exemplul 10. *Caracterizarea zinc-metaloproteazelor în bulion Photorhabdus: inhibarea proteazei, clasificare și purificare*

Inhibarea proteazei și testele de clasificare: testele protează s-au realizat folosind cazeină-FITC dizolvată în apă ca substrat (concentrația finală a testului 0,08%). Reacțiile de proteoliză s-au realizat la 25°C, timp de o oră, în tamponul potrivit cu bulion *Photorhabdus* (volumul total al reacției 150 μl). Probele s-au testat, de asemenea, în prezența și absența ditiotritolului. După incubare s-a adăugat un volum egal de 12% acid tricloracetic pentru a precipita proteina nedigerată. După precipitare timp de 0,5 h și centrifugarea care a urmat, s-au plasat 100 μg supernatant într-o placă de microtitrare cu 96-godeuri și pH-ul soluției s-a ajustat prin adăugarea unui volum egal de NaOH 4N. Proteoliza s-a cuantificat apoi utilizând un cititor de placă fluorometric Fluoroskan II la excitarea și emisia lungimii de undă de 485 și, respectiv 538 nm. Activitatea proteazei s-a testat într-un domeniu de la pH 5,0-10,0 în 5 unități crescute. S-au folosit următoarele tamponuri la concentrația finală de 50 mM: acetat de sodiu (pH 5,0-6,5); Tris-HCl (pH 7,0-8,0) și bis-Tris propan (pH 8,5-10,0). Pentru a identifica clasa proteazei(lor) observate, bulionul brut s-a tratat cu o varietate de inhibitori protează (concentrație finală 0,5 μg/μg) și apoi s-a examinat pentru activitatea proteazei pH 8,0, folosind substratul descris mai sus. Inhibitorii proteazei utilizați includ E-64 (L-trans-expoxisaccinilleucilamino[4-, -guamdino]-butan), 3,4-dicloroizocumarin, Leupeptin, pepstatin, amastatin, acid etilendiamintetraacetic (EDTA) și 1,10 fenantrolin. Testele protează realizate peste domeniul pH relevă că proteaza(ele) au fost într-adevăr cele care prezintă activitatea maximă la ~pH 8,0 (tabelul 16). Adiția de DTT nu a avut nici un efect asupra

RO 121280 B1

activității proteazei. Bulionul brut s-a tratat apoi cu o varietate de inhibitori de protează (tabelul 17). Tratatul bulionului brut cu inhibitorii descriși mai sus arată că 1,10 fenantroin a cauzat inhibarea completă a întregii activități a proteazei, când s-a adăugat la o concentrație finală de 50 μg cu $\text{IC}_{50} = 5 \mu\text{g}$ în 100 μl la o soluție de bulion brut 2 mg/ml. Aceste date arată că proteaza(ele) cea mai abundentă găsită în bulionul *Photorhabdus* este de forma clasei zinc-metaloproteazei a enzimelor.

Tabelul 16

Efectul pH-ului asupra activității din ziua 1 a *Photorhabdus*

pH	Activitate ^a unit.flu. ^b	Procent
5,0	3013 $\pm$ 78	17
5,5	7994 $\pm$ 448	45
6,0	12965 $\pm$ 483	74
6,5	14390 $\pm$ 1291	82
7,0	14386 $\pm$ 1287	82
7,5	14135 $\pm$ 198	80
8,0	17582 $\pm$ 831	100
8,5	16183 $\pm$ 953	92
9,0	16795 $\pm$ 760	96
9,5	16279 $\pm$ 1022	93
10,0	15225 $\pm$ 210	87

a. Activitate=procent activitate relativă la maximum pH 8,0.

b. Unit. flu=unități fluorescente (maximum = -28,000; fundal = -2200).

Tabelul 17

Efectul diferiților inhibitori de protează asupra activității proteazei la pH 8 găsită în producția din ziua 1 a *Photorhabdus luminescens* (tulpina W-14)

Inhibitor	Unit. Flu. corectate ^a	Procent de inhibare
Control	13053	0
E-64	14259	0
1-10 Fenantrolin ^c	15	99
3,4 Diclorizocoumarin ^d	7956	39
Leupeptin	13074	0
Pepstatin ^c	13441	0
Amastatin	12474	4
Control DMSO	12005	8
Control Metanol	12125	7

a. Unit. flu. corectate = unități fluorescente - fundal (2200 unit. flu).

b. Procentul inhibării relative la activitatea proteazei la pH 8,0.

c. Inhibitorii s-au dizolvat în metanol.

d. Inhibitorii s-au dizolvat în DMSO.

Izolarea zinc-metalproteazei s-a realizat prin aplicarea peletului de sulfat de amoniu 10-80% dializat la o coloană Sepharose Q echilibrată la 50 mM Na₂PO₄, pH 7,0 așa cum s-a descris în exemplul 5 pentru toxina *Photorhabdus*. După spălare intensivă, un gradient 0 la 0,5 M NaCl s-a utilizat pentru a elua proteina toxinei. Majoritatea activității biologice și proteina s-au eluat de la 0,15-0,45 M NaCl. Totuși, s-a observat că majoritatea activității proteolitice a fost prezentă în fracția 0,25-0,35 M NaCl cu ceva activitate în fracția 0,15-0,25 M NaCl. Analiza SDS PAGE a fracțiunii 0,25-0,35 M NaCl a arătat o bandă importantă a peptidei de aproximativ 69 kDa. Fracția 0,15-0,25 M NaCl a conținut o bandă similară de 60 kDa, dar o concentrație relativă a proteinei mai scăzută. Filtrarea subsecventă pe gel a acestor fracțiuni folosind o coloană Superose 12 HR 16/50 a rezultat într-un peak major care migrează la 57,5 kDa, care a conținut banda de 58,5 kDa predominantă (>90% din total proteinei colorate) prin analiza SDS PAGE. Analiza ulterioară a acestui fragment, folosind diferiți inhibitori de protează, așa cum s-a descris mai sus, a determinat că proteaza a fost o zinc-metalprotează. Aproape toată activitatea proteazei prezentă în bulion *Photorhabdus* în ziua 1 de fermentare corespunde cu ~58 kDa zinc-metalproteazei.

Într-oaltă izolare secundară a zinc-metalproteazei(ele), bulionul *Photorhabdus* crescut trei zile s-a colectat, iar activitatea proteazei s-a vizualizat cu electroforeză pe gel de sodiu dodecil sulfat-poliacrilamidă (SDS-PAGE), legat cu gelatină, așa cum s-a descris în Schmidt, T.M., Bleakley, B. și Neelson, K.M. 1988. Gelurile de migrare SDS (5,5x8 cm) s-au realizat cu 12,5% poliacrilamidă (40% soluție stoc de acrilamidă/bis-acrilamidă; Sigma Chemical Co., St. Louis, MO), în care concentrația finală de gelatină 0,1% (reactiv grad Biorad EIA; Richmond CA) s-a încorporat pe dizolvare în apă. Gelurile SDS de stivuire (1,0x8 cm) au fost construite cu 5% poliacrilamidă, au fost legate de asemenea cu 0,1% gelatină. În mod tipic, 2,5 μg de proteină de testat s-au diluat în 0,03 ml tampon de încărcare SDS-PAGE fără ditiotreitol (DTT) și s-au încărcat pe gel. Proteinele s-au supus electroforezei în tampon de migrare SDS (Laemmli, U.K. 1970, Nature 227,680) la 0°C și la 8 mA. După ce electroforeza a fost completă, gelul s-a spălat 2 h în 2,5% (v/v) Triton X-100. Gelurile s-au incubat apoi o oră, la 37°C, în 0,1 M glicină (pH 8,0). După incubare, gelurile s-au fixat și s-au colorat peste noapte 0,1% amido black în metanol-acid acetic-apă (30:10:60, vol./vol./vol.; Sigma Chemical Co.). Activitatea proteazei s-a vizualizat ca zone luminoase pe negru, amido black a colorat fundalul, datorită proteolizei și difuziei subsecvente a gelatinei încorporate. S-au observat cel puțin 3 benzi distincte produse de activitatea proteolitică la 58-, 41- și 38 kDa.

Teste de activitate ale diferitelor proteaze în bulion de cultură W-14, ziua a 3-a, s-au realizat utilizând cazeină FITC dizolvată în apă ca substrat (concentrația finală a testului 0,02%). Experimentele proteolizei s-au realizat la 37°C, timp de 0-0,5 h, în 0,1 M Tris-HCl (pH 8,0) cu diferite fracțiuni de proteină într-un volum total 0.4-5 ml. Reacțiile s-au terminat prin adăugarea unui volum egal de acid tricloracetic (TCA) 12% dizolvat în apă. După incubare la temperatura camerei, timp de 0,25 h, probele s-au centrifugat la 10000xg 0,25 h și s-au îndepărtat 10 ml alicote și s-au plasat pe plăci de microtitrare cu 96 godeuri. Soluția s-a neutralizat apoi prin adăugarea unui volum egal de hidroxid de sodiu 2N, urmată de cuantificare folosind un cititor de placă fluorometric Fluoroskan II cu excitare și emisie de lungimi de undă de 485 și, respectiv, 538 nm. Măsurarea activității s-a realizat folosind cazeină FITC cu diferite concentrații de protează la 37°C, 0-10 min. O unitate de activitate s-a definit arbitrar, precum cantitatea de enzimă necesară pentru a produce 1000 unități fluorescente/min și activitatea specifică s-a definit ca unități/mg de protează.

Studiile de inhibare s-au realizat folosind 2 inhibitori zinc-metalprotează; 1,10 fenantrolin și N-(a-ramnopiranosiloxihidroxifosfinil)-Leu-Trp(fosforamidon) cu soluții stoc ale

RO 121280 B1

inhibitorilor dizolvați în 100% etanol și, respectiv, apă. Concentrațiile stoc au fost de obicei 10 mg/ml și 5 mg/ml pentru 1,10 fenantrolin și respectiv fosforamidon, concentrația finală a inhibitorului de 0,5-1,0 mg/ml per reacție. Tratamentul de 3 zile al bulionului brut W-14 cu 1,10 fenantrolin, un inhibitor al tuturor zinc-metaloproteazelor, a rezultat în eliminarea completă a întregii activități a proteazei, în timp ce tratamentul cu fosforaminodin, un inhibitor al proteazelor ca termolizina (Waver, L.H., Kester, W.R., și Matthews, B.W., 1977, *J. Mol. Biol.*, 114, 119-132) a rezultat în reducerea cu ~56% a activității proteazei. Activitatea proteolitică reziduală nu a putut fi redusă în continuare cu fosforamidon adițional.

Proteaza de 3 zile a bulionului *Photorhabdus* W-14 s-a purificat așa cum urmează: 4,0 l bulion s-au concentrat folosind un cartuș de ultrafiltrare spirală Amicon tip SI Y100, atașat la un dispozitiv de filtrare Amicon M-12. Materialul care curge, având proteine native mai mici în lungime decât 100 kDa (3,8 L), s-a concentrat la 0,375 L folosind un cartuș de ultrafiltrare spirală Amicon tip S1 Y10, atașat la un dispozitiv de filtrare Amicon M-12. Materialul reținut a conținut proteine variind în mărime de la 10-100 kDa. Acest material s-a încărcat pe o coloană Pharmacia HR 16/10 care s-a împachetat cu PerSeptive Biosystem (Framington, MA) Poros®50 HQ, împachetare puternic schimbătoare anion, care s-a echilibrat într-un tampon de fosfat de sodiu 10 mM (pH 7,0). Proteinele s-au încărcat pe coloană la o rată de curgere de 5 ml/min, urmată de spălarea proteinelor nelegate cu tampon, până când $A_{280} = 0,00$. După aceea, proteinele s-au eluat folosind gradient NaCl la 0-1,0 M NaCl, în 40 min, la o rată de curgere de 7,5 ml/min. Frațiunile s-au testat pentru o activitate a proteazei, supra., iar frațiunile active s-au depozitat. Frațiunile active proteolitic s-au diluat cu 50% (v/v) tampon fosfat de sodiu 10 mM (pH 7,0) și s-au încărcat pe o coloană Pharmacia HR 10/10 Mono Q, echilibrată în 10 mM fosfat de sodiu. După spălarea coloanei cu tampon, până când $A_{280} = 0,00$, proteinele s-au eluat folosind un gradient NaCl de 0-0,5 M NaCl, o oră, la o rată de curgere de 2,0 ml/min. Frațiunile s-au testat pentru activitatea proteazei. Acele frațiuni care au avut cea mai mare cantitate de activitate a proteazei sensibilă la fosforamidon, activitatea sensibilă la fosforamidon fiind datorată proteazei 41/38 kDa, infra., s-au depozitat. Aceste frațiuni s-au găsit a elua la un domeniu de 0,15-0,25 M NaCl. Frațiunile care conțin activitatea proteazei insensibilă la fosforamidon predominantă, proteaza 58 kDa, de asemenea s-au depozitat. Aceste frațiuni s-au găsit a elua la un domeniu de 0,25-0,35M NaCl. Frațiunile proteaza sensibilă la fosforamidon s-au concentrat apoi la volumul final de 0,75 ml, folosind un dispozitiv centrifugal de filtrare Millipore Ultrafree®-15 membrană Biomax-5K NMWL. Acest material s-a aplicat la o rată de curgere de 0,5 ml/min pe o coloană Pharmacia HR 10/30, care fusese împachetată cu Pharmacia Sephadex G-50 echilibrată într-un tampon de 10 mM sulfat de sodiu (pH 7,0)/0,1 M NaCl. Frațiunile care au avut activitatea proteazei sensibilă la fosforamidon, maximă, s-au depozitat apoi și s-au centrifugat pe un dispozitiv centrifugal de filtrare Millipore Ultrafree®-15 membrană Biomax-5K NMWL. Analiza activității proteolitice, supra., indică acest material a avea doar activitate a proteazei sensibilă la fosforamidon. Depozitarea proteazei insensibile la fosforamidon, proteina 58 kDa a fost urmată de concentrarea într-un dispozitiv centrifugal de filtrare Millipore Ultrafree®-15 membrană Biomax-5K NMWL și de separarea, în continuare, pe o coloană Pharmacia Supradex-75. Frațiunile care conțin protează s-au depozitat.

Analiza proteazelor purificate 58- și 41/38 kDa a relevat că, în timp ce ambele tipuri de proteaze au fost complet inhibitate cu 1,10 fenantrolin, doar protează 41/38 kDa a fost inhibată cu fosforamidon. Analiza ulterioară a bulionului brut a arătat că activitatea proteazei bulionului W-14, din ziua 1, a fost 23% din totalul activității proteazei datorată proteazei 41/38 kDa, care a crescut la 44% în ziua 3-a a bulionului W-14.

Analiza SDS-PAGE standard pentru examinarea purității proteinei și obținerea secvenței terminale amino s-au realizat folosind 4-20% gradient MiniPlus SeptraGels achiziționat de la Integrated Separation Systems (Natick, MA). Proteinele de secvențat amino-terminal s-au blott-at pe membrane PVDF, urmând purificarea, infra., (ProBlott™ Membranes; Applied Biosystems, Foster City, CA), s-au vizualizat cu 0,1 amido black, s-au excizat și s-au trimis la Cambridge Prochem., Cambridge, MA, pentru secvențare.

Secvențele terminale amino deduse ale proteazelor 58- (SEQ ID NR:45) și 41/38 kDa (SEQ ID NR:44) din bulion W-14, vechi de 3 zile, au fost DV-GSEKANEKLLK (SEQ ID NR:45) și, respectiv, DSGDDDKVTNTDIHR (SEQ ID NR: 44).

Secvențarea proteazei 41/38 kDa a relevat mai mulți terminali amino, fiecare având un aminoacid adițional îndepărtat prin proteoliză. Examinarea secvențelor primară, secundară, terțiară și cuaternară pentru polipeptidele 38 și 41 kDa a permis deducerea secvenței arătată mai sus și a relevat că aceste două proteaze sunt omoloage.

Exemplul 11. Partea A. Cercetarea băncii genomice *Photorhabdus* via folosirea anticorpilor pentru genele care codifică peptida TcbA

În paralel cu secvențarea descrisă mai sus, un experiment corespunzător și secvențare s-au făcut și pe baza peptidei TcbA_{ii} (SEQ ID NR:1). Această secvențare s-a realizat prin prepararea bulionului de cultură bacterian și purificarea toxinei, cum s-a descris în exemplele 1 și 2, de mai sus.

S-a izolat ADN genomic de lei. *Photorhabdus luminescens* tulpina W-14 crescut în mediu de cultură țesut insectă al lui Grace. Bacteria s-a crescut în 5 ml mediu de cultură într-un pahar Erlenmeyer, la 28°C și 250 rpm, pentru aproximativ 24 h. Celulele bacteriene de la 100 ml mediul de cultură s-au peletat la 5000xg, timp de 10 min. Supernatantul s-a îndepărtat și peletele celulare s-au folosit apoi pentru izolarea ADN-ului genomic.

S-au izolat ADN genomic folosind o modificare a metodei CTAB descrisă în Secțiunea 2.4.3. a lui Ausubel (supra.), secțiune intitulată *Large Scale CsCl prep of bacterial genomic DNA*, s-a urmat prin etapa 6. La acest punct, s-a realizat o extracție suplimentară cloroform/alcool izoamilic (24:1), urmată de o etapă de extracție fenol/cloroform/izoamil (25:24:1) și o extracție finală cloroform/alcool izoamilic (24:1). S-a precipitat ADN-ul prin adăugarea unui volum 0,6 de izopropanol. ADN-ul precipitat s-a suspendat și a atins vârful unei baghete de sticlă întinsă, s-a afundat scurt în etanol 70% ca o spălare finală și s-a dizolvat în 3 ml tampon TE.

Concentrarea ADN-ului, exprimată prin densitate optică la 280/260 nm, s-a aproximat la 2 mg/ml.

Folosind acest ADN genomic s-a preparat o bancă. Aproximativ 50 μg de ADN genomic s-au digerat parțial cu Sau3 A. Apoi, centrifugarea gradientului de densitate NaCl s-a folosit pentru a măsura fragmentele ADN digerate, parțial fracționate. Frațiunile care conțineau fragmente ADN, cu o mărime medie de 12 kb sau mai mare, așa cum s-a determinat prin electroforeză pe gel de agaroză, s-au ligat în plasmidul BluScript, Stratagene, La Jolla, California și s-au transformat în tulpini DH5α sau DHB10 E.coli.

Separat, alicotele purificate ale proteinei s-au trimis la Centrul de biotehnologie hibridoma al Universității din Wisconsin, Madison, pentru producerea anticorpilor monoclonali la proteine. Materialul care a fost trimis s-a purificat HPLC, iar fracțiunile care conțin benzile native 1 și 2, care fuseseră denaturate la 65°C și din care 20 μg s-au injectat în fiecare dintre cei 4 șoareci. Liniile de celule hibridoma, producătoare de anticorpi monoclonali stabili, s-au recuperat de pe celule de splină de la șoareci imunizați, s-au topit cu o linie celulară mieloma stabilă. Anticorpii monoclonali s-au recuperat de la hibridoame.

RO 121280 B1

Separat, anticorpii policlonali s-au creat prin luarea benzii 1 native, purificată pe gel de agaroză (vezi exemplul 1) proteină, care s-a folosit apoi pentru a imuniza un iepure alb de Noua Zeelandă. Proteina s-a preparat prin excizarea benzii de la gelurile de agaroză native, încălzirea sumară a bucăților de gel la 65°C, pentru a topi agaroză și emulsionarea imediată cu adjuvant. Adjuvantul Freund complet s-a folosit pentru imunizarea primară, iar Freund incomplet s-a folosit pentru 3 injectări adiționale la intervale de 1 lună. Pentru fiecare injectare s-au administrat subcutanat, prin injectări multiple în spatele iepurelui, aproximativ 0,2 ml de bandă 1 emulsionată care conține 50 la 100 µg de proteină. Serul s-a obținut la 10 zile după ce ultima injecție și sângerările adiționale s-au realizat la intervale de 1 săptămână, timp de 3 săptămâni. Complementul serului s-a inactivat prin încălzire la 56°C, timp de 15 min, și s-a stocat la -20°C.

Anticorpii monoclonali și policlonali s-au utilizat apoi pentru a cerceta banca genomică, pentru exprimarea antigenilor care ar putea fi depistați prin epitopi. Clonele pozitive s-au detectat pe proeminențele coloniei de pe filtrul de nitroceluloză. S-a întreprins o analiză imunoblot a clonelor pozitive.

O analiză a clonelor definită prin imunoblot și analiză Southern a avut ca rezultat tentativa de identificare a cinci clase de clone.

În prima clasă de clone a fost o genă care codifică peptida desemnată aici ca TcbA_i. S-a obținut întreaga secvență ADN a acestei gene (TcbA). Aceasta s-a prezentat ca SEQ ID NR: 11. Confirmarea că această secvență codifică secvența internă a SEQ ID NR: 1 a fost demonstrată prin prezența SEQ ID NR:1 la aminoacidul numărul 88 din secvența aminoacidă dedusă, creată prin cadrul deschis de citire al SEQ ID NR: 11. Aceasta poate fi confirmată prin referire la SEQ ID NR: 12, care este secvența aminoacidă dedusă creată prin SEQ ID NR: 11.

A doua clasă de peptide toxină conține segmentele la care se face referință mai sus ca TcaB_i, TcaB_{ii} și TcaC. Urmând cercetarea băncii cu antiseruri policlonale, această a doua clasă de genă toxină s-a identificat prin mai multe clone care au produs proteine de diferite mărimi, dintre care toate au reacționat încrucișat cu anticorpul policlonal pe un imunoblot și s-au găsit, de asemenea, că au omologie ADN pe un Southern Blot. Compararea secvențelor relevă că ele aparțin aceluiași complex de gene desemnat TcaB și TcaC, mai sus.

Trei alte clase de clone toxină anticorp s-au izolat, de asemenea, în cercetarea policlonală. Aceste clase au produs proteine care reacționează încrucișat cu un anticorp policlonal și, de asemenea, au omologie ADN cu clasele, așa cum s-a determinat prin Southern Blot. Clasele au fost desemnate Clasa III, Clasa IV și Clasa V. A fost de asemenea posibil să se identifice monoclonalii care reacționează încrucișat cu Clasele I, II, III și IV. Aceasta sugerează că toate au regiuni de înaltă omologie proteină. Astfel, se pare că genele proteinei extracelulare *P. luminescens* nu prezintă o familie de gene care sunt înrudite în evoluție.

Pentru a putea continua conceptul că ar putea fi variații înrudite evolutiv în peptidele toxine conținute în acest organism, au fost întreprinse două abordări pentru a examina alte tulpini *P. luminescens* pentru prezența proteinelor înrudite. Aceasta s-a realizat atât prin amplificarea PCR a ADN-ului genomic, cât și prin analiza imunoblot folosind anticorpi monoclonali și policlonali.

Rezultatele arată că proteinele înrudite sunt produse de tulpinile *P. luminescens* WX-2, WX-3, WX-4, WX-5, WX-6, WX-7, WX-8, WX-11, WX-12, WX-15 și W-14.

Exemplul 11. Partea B. Secvența și analiza clonelor toxină din Clasa III - tcc

Secvențarea ADN ulterioară s-a realizat pe plasmizi izolați din clone E.coli Clasa III, descriși în exemplul 11, partea A. Secvența de nucleotide s-a arătat a fi de trei ori linkată, aproape cadru deschis de citire la acest locus genomic. Acest locus s-a denumit tec cu trei

1 cadre deschise de citire denumite tccA SEQ ID NR: 56, tccB SEQ ID NR:58 și tccC SEQ ID
NR:60 (Fig. 6B).

3 Aminoacidul dedus de la cadrul deschis de citire tccA arată că gena codifică o pro-
teină de 105.459 Da. Această proteină s-a denumit TccA. Primii 12 aminoacizi ai acestei pro-
5 teine potriviți cu secvența N-terminală s-au obținut de la o proteină de 108 kDa, SEQ ID
NR:7, identificată mai devreme ca parte a complexului toxină.

7 Aminoacidul dedus de la cadrul deschis de citire TccB arată că gena codifică o pro-
teină de 157.716 Da. Această proteină potrivită cu secvența N-terminală s-a obținut de la o
9 proteină cu greutatea moleculară estimată de 185 kDa, SEQ ID NR:8.

11 Secvența aminoacidică dedusă din tccC arată că acest cadru deschis de citire codifică
o proteină de 111.694 Da și produsul proteină s-a denumit TccC.

Exemplul 12. Caracterizarea tulpinilor *Photorhabdus*

13 Pentru a stabili dacă respectiva colecție descrisă aici a fost cuprinsă de tulpinile
Photorhabdus, tulpinile respective s-au evaluat pentru recunoașterea trăsăturilor micro-
15 biologice care sunt caracteristice la *Photordabds* și care o diferențiază de alte
Enterobacteriacee și *Xenorhabdus spp.* (Farmer, J.J. 1984, *Bergey's Manual of Systemic*
17 *Bacteriology*, vol 1, pp. 510-511. (Ed. Kreig N.R. and Hoit, J. G.). Wiffiams & Wilkins,
Baltimore.; Akhurst and Boemare, 1988, Boemare et al., 1993). Aceste trăsături caracteris-
19 tice sunt după cum urmează: tijele negative colorate ale lui Gram, mărirea organismului de
0,5-2 μm în lățime și de 2-10 μm în lungime, pigmentarea coloniei roșu/galbenă, prezența
21 corpurilor de incluziune cristalin, prezența catalazei, inabilitatea de a reduce nitrații, prezența
bioluminescenței, abilitatea de a relua colorant din mediul de cultură, pozitiv pentru produ-
23 cerea proteazei, intervalul temperaturii de creștere sub 37°C, supraviețuire în condiții anae-
robe și capacitate de mișcare pozitivă. (tabelul 18). Tulpini de referință *Escherichia coli*,
25 *Xenorhabdus* și *Photorhabdus* s-au inclus în toate testele pentru comparare. Rezultatele de
suprafață sunt consistente pentru toate tulpinile ce sunt parte a familiei *Enterobacteriacee*
27 și a genului *Photorhabdus*.

29 S-a utilizat un luminometru pentru a stabili bioluminiscenta fiecărei tulpini și pentru
a furniza o măsurare cantitativă și relativă a producției de lumină. Pentru măsurarea unităților
de emisie de lumină relative, bulioane din fiecare tulpină (celule și mediu) s-au măsurat la
31 trei intervale de timp, după inoculare în lichid de cultură (6, 12 și 24 h) și s-au comparat cu
luminozitatea fundalului (mediu neinoculat și apă). Înainte de a măsura emisiunea de lumină
33 din diferite bulioane, densitatea celulelor s-a stabilit prin măsurarea absorbantei luminoase
(560 nm) într-un spectrometru Gilford Systems (Oberlin, OH), folosind o celulă de sorbție.
35 S-au relizat apoi diluțiile adecvate (pentru a normaliza densitatea optică la 1,0 unități), înainte
de a măsura luminozitatea. Alicotele bulioanelor diluate s-au plasat apoi în cuve (300 μl
37 fiecare) și s-au citit în Bio Orbit 1251 Luminometer (Bio-Orbit Oy, Twiku, Finland). Perioada
de integrare pentru fiecare probă a fost de 45 s. Probele s-au amestecat continuu (filat în
39 cuve cu praguri), în timp ce a fost citit pentru a asigura disponibilitatea de oxigen. Un test
pozitiv s-a determinat ca fiind ≥ 5 ori luminescența fundalului (~5-10 unități). În plus, lumino-
41 zitatea coloniei s-a detectat cu film de acoperire fotografic și vizual, după adaptarea la came-
ra obscură. Caracteristicile de colorare ale lui Gram, ale fiecărei tulpini, s-au stabilit cu un kit
43 de culoare a lui Gram comercial (BBL, Cockeysville, MD), utilizate cu diapozitive de control
culoare a lui Gram (Fisher Scientific, Pittsburgh, PA). Evaluarea microscopică s-a realizat
45 apoi utilizând un microscop Zeiss (Carl Zeiss, Germania), obiective de imersie 100X (cu
mărire 10X ocular și 2X corp). Examinarea microscopică a tulpinilor individuale pentru mări-
47 mea organismului, descrierea celulară și corpurile de incluziune (mai târziu după creșterea
logaritmică) s-au realizat utilizând diapozitive cu montura umedă (mărire 10X ocular, 2X corp

RO 121280 B1

și 40X obiectiv) cu imersiune în ulei și microscopia contrast fază cu un micrometru (Akhurst, R. J. and Boemare, N. E. 1990, <i>Entomopathogenic Nematodes in Biological Control</i> (ed. Gaugler, R. and Kaya, H.), pp 75-90, CRC Press, Boca Raton, USA.; Baghdiguian S., Boyer-Giglio M. H., Thaler, J. O., Bonnot G., Boemare N., 1993, <i>Biol. Cell</i> 79,177-185.).	1
Pigmentarea coloniei s-a observat după inocularea pe agar nutrient Bacto (Difco Laboratories, Detroit, MI), preparat ca în instrucțiunile de pe etichetă. Incubarea a apărut la 28°C și caracterizarea s-a produs după 5-7 zile. Pentru a testa prezența catalazei enzimei, o colonie din organismul test s-a îndepărtat pe mic tampon de la placa nutrientă de agar și s-a plasat pe fondul eprubetei de sticlă test. S-a adăugat blând, pe marginea eprubetei, 1 ml de soluție de apă oxigenată obișnuită. S-a înregistrat o reacție pozitivă când au apărut, imediat sau în 5 min, bule de gaz (presupus oxigen). S-au examinat, de asemenea, controalele agarului nutrient neinoculat și ale soluției de apă oxigenată. Pentru a testa reducerea nitratului, fiecare cultură s-a inoculat cu 10 ml Bacto Nitrate Broth (Difco Laboratories, Detroit, MI). După 24 h de incubare la 28°C, producerea de nitrit s-a testat prin adăugarea a două picături de reactiv acid sulfanilic și 2 picături de reactiv <i>alfa</i> -naftilamină (vezi <i>Difco Manual</i> , 10 th edition, Difco Laboratories, Detroit, MI, 1984). Generarea unei culori roz sau roșu distincte indică formarea nitritului din nitrat. Capacitatea fiecărei tulpini de absorbi colorant din mediul de cultură s-a testat cu agar Bacto MacConkey, care conține colorant natural roșu; agar Bacto Tergitol-7, care conține colorantul bromtimol albastru și agar Bacto EMB care conține colorantul eosin-y (agaruri de la Difco Laboratories, Detroit, MI, toate preparate în conformitate cu instrucțiunile de pe etichetă). După incubare pe aceste medii, colorantul absorbit s-a înregistrat după incubare la 28°C, timp de 5 zile. Creșterea pe aceste medii din urmă este caracteristică pentru membrii familiei <i>Enterobacteriaceae</i> . Mobilitatea fiecărei tulpini s-a testat folosind o soluție de Bacto Motility Test Medium (Difco Laboratories, Detroit, MI) preparată ca în instrucțiunile de pe etichetă. O inoculare punct de joncțiune-străpungere s-a realizat cu fiecare tulpină și mobilitatea s-a apreciat macroscopic prin zona difuză a împrăștierei creșterii de la linia de inoculare. În multe cazuri, mobilitatea s-a observat microscopic din lichidul de cultură sub diapozitive cu montura umedă. Evaluarea nutrientului biochimic pentru fiecare tulpină s-a realizat folosind BBL Enterotube II (Benton, Dickinson, Germania). S-au urmat instrucțiunile produsului, cu excepția că incubarea s-a realizat la 28°C, timp de 5 zile. Rezultatele au fost asemănătoare cu raportul menționat anterior despre <i>Photorhabdus</i> . Producerea de protează s-a testat prin observarea hidrolizei gelatinei, utilizând plăci gelatină Bacto (Difco Laboratories, Detroit, MI), preparate ca în instrucțiunile de pe etichetă. Culturile s-au inoculat și plăcile s-au incubat la 28°C, timp de 5 zile. Pentru a evalua creșterea la diferite temperaturi, plăcile de agar [2% proteose peptone #3 cu 2% Bacto-Agar (Difco, Detroit, MI) în apă deionizată] s-au striat de la sursa comună a inoculării. Plăcile s-au sigilat cu film Nesco® și s-au incubat la 20, 28 și 37°C, timp de 3 săptămâni. Plăcile care nu au arătat creștere la 37°C nu au demonstrat viabilitatea celulelor după transfer într-un incubator la 28°C, timp de o săptămână. Necesarul de oxigen al tulpinilor <i>Photorhabdus</i> s-a testat în următorul mod. S-a realizat o inoculare punct de joncțiune-străpungere în mediu de bulion tioglicolat fluid (Difco, Detroit, MI). Eprubetele s-au incubat la temperatura camerei pentru o săptămână și culturile s-au examinat pentru tipul și extinderea creșterii. Indicatorul resazurin demonstrează nivelul oxidării medii a zonelor aerobe (<i>Difco Manual</i> , 10 th edition, Difco Laboratories, Detroit, MI). Rezultatele zonele de creștere obținute pentru tulpinile <i>Photorhabdus</i> testate au fost asemănătoare cu cele ale microorganismelor anaerobe facultative.	3 5 7 9 11 13 15 17 19 21 23 25 27 29 31 33 35 37 39 41 43 45

Trăsăturile taxonomice ale tulpinilor *Photorhabdus* - Evaluarea trăsăturilor*

1

3

5

7

9

11

13

15

17

19

21

23

25

27

29

31

33

35

37

Tabelul 10

Trăsăturile taxonomice ale tulpinilor Photorhabdus - Evaluarea trăsăturilor*

Tulpina	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
W-14	<u>-1</u>	±	±	<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-12	=	±	±	<u>rd</u> <u>S</u>	±	=	±	±	±	<u>RO</u>	±	±	±	±	±	±	=
WX-13	=	±	±	<u>rd</u> <u>S</u>	±	=	±	±	±	<u>O</u>	±	±	±	±	±	±	=
WX-14	=	±	±	<u>rd</u> <u>S</u>	±	=	±	±	±	<u>LR</u>	±	±	±	±	±	±	=
WX-15	=	±	±	<u>rd</u> <u>S</u>	±	=	±	±	±	<u>LR</u>	±	±	±	±	±	±	=
H9	=	±	±	<u>rd</u> <u>S</u>	±	=	±	±	±	<u>LY</u>	±	±	±	±	±	±	=
Hb	=	±	±	<u>rd</u> <u>S</u>	±	=	±	±	±	<u>YT</u>	±	±	±	±	±	±	=

Tabelul 18 (continuare)

Tulpina	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	
Hm	=	±	±	<u>rd</u> <u>S</u>	±	=	±	±	±	<u>TY</u>	±	±	±	±	±	±	=	3
HP88	=	±	±	<u>rd</u> <u>S</u>	±	=	±	±	±	<u>LY</u>	±	±	±	±	±	±	=	5
NC-1	=	±	±	<u>rd</u> <u>S</u>	±	=	±	±	±	<u>O</u>	±	±	±	±	±	±	=	7
W30	=	±	±	<u>rd</u> <u>S</u>	±	=	±	±	±	<u>YT</u>	±	±	±	±	±	±	=	9
WIR	=	±	±	<u>rd</u> <u>S</u>	±	=	±	±	±	<u>RO</u>	±	±	±	±	±	±	=	11
B2	=	±	±	<u>rd</u> <u>S</u>	±	=	±	±	±	<u>R</u>	±	±	±	±	±	±	=	13
43948	=	±	±	<u>rd</u> <u>S</u>	±	=	±	±	±	<u>O</u>	±	±	±	±	±	±	=	15
43949	=	±	±	<u>rd</u> <u>S</u>	±	=	±	±	±	<u>O</u>	±	±	±	±	±	±	=	17
43950	=	±	±	<u>rd</u> <u>S</u>	±	=	±	±	±	<u>O</u>	±	±	±	±	±	±	=	19
43851	=	±	±	<u>rd</u> <u>S</u>	±	=	±	±	±	<u>O</u>	±	±	±	±	±	±	=	21
42952	=	±	±	<u>rd</u> <u>S</u>	±	=	±	±	±	<u>O</u>	±	±	±	±	±	±	=	23

A= colorantul lui Gram, B= corpuri cristaline de incluziune, C=bioluminescență, D= forma celulei, E=mobilitate, F= reducerea nitrului, G= prezența catalazei, H= hidroliza gelatinei, I= absorbția colorantului, J= pigmentarea, K= creșterea pe agar EMB, L= creșterea pe agar MacConkey, M= creșterea pe agar Tergitol-7, N= anaerobe facultative, O= creșterea la 28°C, Q= creșterea la 37°C, I - +/- = pentru trăsătură pozitivă sau negativă, rd= baghetă, S= măsurat cu descriptori Genus, RO= roșu-portocaliu, LR= roșu deschis, R= roșu, O= portocaliu, Y= galben, T= maroniu, LY= galben deschis, YT= galben maroniu și LO= portocaliu deschis.

Analiza acizilor grași celulari este recunoscută ca un instrument pentru caracterizarea bacterială la nivel de genuri și specii (Tornabene, T.G., 1985, *Lipid Analysis and the Relationship to Chemotaxonomy in Methods in Microbiology*, vol. 18, 209-234; Goodfellow, M. and O'Donnell, A.G., 1993, *Roots of Bacterial Systematics in Handbook of New Bacterial Systematics* (ed. Goodfellow, M. & O'Donnell, A.G.), pp. 3-54, London: Academic Press Ltd), aceste referințe sunt încorporate aici prin referire și s-au utilizat pentru a confirma că, colecția noastră s-a înrudit la nivel de gen. Culturile s-au transportat la un laborator extern cu contract pentru analiza metil ester acid gras (FAME), utilizând un Microbial ID (Midi, Newark, DE, USA) *Microbial Identification System* (MIS). Sistemul MIS constă dintr-un cromatograf gaz Hewlett Packard HP5890A cu o coloană capilară 5% metilfenil silicon topit silica de 25 mm x 0,2 mm. Se folosește hidrogen, ca gaz purtător, și un detector ionizare cu flacără, funcționează împreună cu un aparat pentru luarea probelor automat, un integrator și un

1 computer. Computerul compară probele de metil ester acid gras cu o bancă de acizi grași
 3 microbiană și cu un amestec de calibrare de acizi grași cunoscuți. Așa cum au fost selectate
 de laboratorul contractant, tulpinile au crescut 24 h, la 28°C, pe un agar tripticază soia,
 înainte de analiză. Extragerea probelor s-a realizat de către laboratorul contractant în confor-
 5 mitate cu metodologia standard FAME. Nu s-au identificat direct tulpini din alt grup bacterial
 luminiscent, decât *Photorhabdus*. Când s-a realizat analiza pe grupuri care compară profi-
 7 lurile acizilor grași ai unui grup de izolate, profilurile acizilor grași ai tulpinei erau înrudite la
 nivel de gen.

9 Diversitatea evolutivă a tulpinilor *Photorhabdus* din colecția noastră s-a măsurat prin
 analiza PCR (Polymerase Chain Reaction) a amprentelor genomice mediate, folosind ADN
 11 genomic de la fiecare tulpină. Această tehnică este bazată pe familiile de secvențe ADN re-
 petitive, prezente de-a lungul genomilor diverselor specii bacteriene (revăzută de Versalovic,
 13 J., Schneider, N., DE Bruijn, F.J. și Lupski, J.R., 1994, *Methods Mol. Cell. Biol.*, 5,25-40.).
 Trei din aceste secvențe palindrom extragenice repetitive (REP), consens intergenic repe-
 15 titiv enterobacterial (ERIC) și elementul BOX sunt considerate a juca un rol important în orga-
 nizarea genomului bacterian. Organizarea genomică este considerată a fi formată prin selec-
 17 tarea și dispersia diferențială a acestor elemente în genomul tulpinilor bacteriene strâns înru-
 dite, poate fi folosită pentru a diferenția aceste tulpini (de exemplu, Louws, F.J., Fulbright,
 19 D.W., Stephens, C.T. și DE Bruijn, F.J., 1994, *Appl. Environ. Micro.* 60, 2286-2295).
 REP-PCR utilizează primeri oligonucleotide complementari la aceste secvențe repetitive,
 21 pentru a amplifica fragmentele ADN de mărime variabilă, care se găsesc între ele. Produsele
 rezultate sunt separate prin electroforeză, pentru a stabili "amprenta" ADN pentru fiecare
 23 tulpină.

25 Pentru a izola ADN genomic pentru tulpinile noastre, peletele celulare s-au resuspen-
 dat în tampon TE (10 mM Tris- HCl, 1 mM EDTA, pH 8,0) până la un volum final de 10 ml
 și s-au adăugat 10 ml NaCl 5 M. Acest amestec s-a centrifugat 20 min, la 15000xg. Peletul
 27 rezultat s-a resuspendat în 5,7 ml TE și s-au adăugat 300 μ l SDS și 60 μ l 20 mg/ml prote-
 inază K (Gibco BRL Products, Grand Island, NY). Acest amestec s-a incubat la 37°C, timp
 29 de o oră, apoi s-au adăugat aproximativ 10 mg lizozimă și amestecul s-a incubat pentru încă
 45 min. S-a adăugat apoi 1 ml NaCl 5M și 800 μ l soluție CTAB/NaCl (10% g/v CTAB, NaCl
 31 0,7 M) și amestecul s-a incubat 10 min, la 65°C, agitând blând, apoi s-a incubat și s-a agitat
 pentru încă 20 min, pentru a ajuta limpezirea materialului celular. S-a adăugat un volum egal
 33 de soluție cloroform/alcool izoamilic (24:1, v/v), s-a amestecat blând și apoi s-a centrifugat.
 S-au realizat apoi 2 extracții cu un volum egal de fenol/cloroform/aleool izoamilic (50:49:1).
 35 ADN genomic s-a precipitat cu 0,6 volume de izopropanol. ADN precipitat s-a îndepărtat cu
 o baghetă de sticlă, s-a spălat de 2 ori cu 70% etanol, s-a uscat și s-a dizolvat în 2 ml STE
 37 (10 mM Tris-HCl pH 8,0, 10 mM NaCl, 1 mM EDTA). Apoi, ADN-ul s-a cuantificat prin
 densitate optică 260 iun. Pentru a realiza analiza rep-PCR a ADN-ului genomic
 39 *Photorhabdus* s-au utilizat următorii primeri: REP1R-I; 5'-IIICGICGICATCIGGC-3' și REP2-I;
 5'-ICGICTTATCIGGCCTAC-3'. PCR s-a realizat folosind următorii 25 μ l de reacție: 7,75 μ l
 41 apă, 2,5 μ l tampon 10X LA (PanVera Corp., Madison, WI), 16 μ l amestec dNTP (2,5 mM fie-
 care), 1 μ l din fiecare primer la 50 pM/ μ l, 1 μ l DMSO, 1,5 μ l ADN genomic (domeniul con-
 43 concentrațiilor de la 0,075-0,480 μ g/ μ l) și 0,25 μ l TaKaRa EX Taq (PanVera Corp., Madison,
 WI). Amplificarea PCR s-a realizat într-un Perkin Elmer DNA Thermal Cycler (Norwalk, CT),
 45 care a utilizat următoarele condiții: 97°C/7 min, apoi 35 cicluri la 94°C/1 min, 44°C/1 min,
 65°C/8 min; urmate de 15 min, la 65°C. După ciclare, cei 25 μ l de reacție s-au adăugat la 6x
 47 tampon de încărcare gel (0,25% albastru de bromfenol, 40% g/v sucroză în apă). Un gel de
 agaroză 1% 15x20 cm s-a trecut apoi în tampon TBE (0,09 M Tris-borat, 0,02 M EDTA),

folosind 8 µl din fiecare reacție. Gelul s-a trecut timp de aproximativ 16 h, la 45 v. Gelurile s-au colorat apoi în 20 µg/ml bromură de etidiu, timp de o oră, și s-au decolorat în tampon TBE timp de aproximativ 3 h. S-au făcut apoi fotografii Polaroid® sub lumină UV ale gelurilor.	1 3
S-au măsurat de pe fotografii prezența și absența benzilor la mărime specifice pentru fiecare tulpină și s-a introdus ca o matrice de similaritate într-un program software de taxonomie numerică, NTSYS-pc (Exeter Software, Setauket, NY). Controale din tulpina HB 101 <i>E.coli</i> și <i>Xanthomonas oryzae</i> pv. <i>Oryzae</i> s-au testat în același timp ca "amprente" PCR produse, care corespund rapoartelor publicate (Versalovic, J., Koeuth, T. și Lupsky, J.R., 1991, <i>Nucleic Acids Res.</i> 19, 6823-6831; Vera Cruz, CM. Halda-Alija, L. Louws, F., Skinner, D.Z., George, M.L., Nelson, R.J., DE Bruijn, F.J., Rice, C și Leach, J.E., 1995, <i>Int. Rice Res. Notes</i> , 20, 23-24; Vera Cruz, CM., Ardales, E.Y., Skinner, D.Z., Talag, J., Nelson, R.J., Louws, F.J., Leung, H., Mew, T.W. și Leach, J.E., 1996, <i>Phytopathology</i> (respectiv, în presă). Apoi, s-au analizat date de la tulpini de <i>Photorhabdus</i> cu o serie de programe în NTSYS-pc; SIMQUAL (Similarity for Qualitative data) pentru a genera o matrice a coeficienților de similaritate (folosind coeficientul Jaccard) și SAHN (Sequential, Agglomerative, heirarchical and Nested) pentru a realiza grupări folosind metoda [UPGMA (Unweighted Pair-Group Method with Arithmetic Averages)] care grupează tulpini înrudite și pot fi exprimate ca o fenogramă (fig. 5). Programele CPH (valori cofenetice) și MXCOMP (matrice comparare) s-au folosit pentru a genera o matrice valoare cogenetică și compară corelația dintre acesta și matricea originală pe care s-a bazat gruparea. S-a generat o statistică Mantel (r) normalizată rezultată, care este o măsură a valorii armonizării pentru o analiză de grup (4=0,8-0,9 reprezintă o valoare foarte bună). În cazul nostru, r= 0,919. De aceea, colecția noastră este compusă dintr-un grup diferit al tulpinilor ușor de evidențiat, ale genului <i>Photorhabdus</i> .	5 7 9 11 13 15 17 19 21 23
Exemplul 13. Utilitatea insecticidă a toxineilor produsă de tulpini diferite <i>Photorhabdus</i>	25
Culturile "semințe" inițiale ale tulpinilor diferite <i>Photorhabdus</i> s-au produs prin inocularea a 175 ml de 2% Proteose Peptone #3 (PP3) (Difco Laboratories, Detroit, MI) mediu lichid cu o subclonă variată primară într-un flacon de 500 ml cu trei șicane cu un gât Delong, acoperit cu un Kaput. Inoculul pentru fiecare cultură de semințe a derivat de la culturi înclinate agar acoperire-ulei sau culturi pe placă. După inoculare, aceste flacoane s-au incubat pentru 16 h, la 28°C, pentru agitator rotativ la 150 rpm. Aceste culturi de semințe s-au folosit apoi ca surse de inocul uniform, pentru a da fermentație fiecărei tulpini. În plus, s-a acoperit cultura de semințe post-log cu ulei mineral steril, adăugând un agitator magnetic steril pentru resuspendare ulterioară și stocând cultura la întuneric, la temperatura camerei, și asigurând păstrare pe termen lung a inoculului într-o stare competent-toxină. Bulioanele produse s-au inoculat prin adăugarea a 1% de cultură de creștere activă a semințelor la mediu 2% PP3 proaspăt (de exemplu, 1,75 ml per 175 ml mediu proaspăt). Producția bulioanelor s-a realizat fie în flacoane de 500 ml cu trei șicane (vezi mai sus), fie în flacoane cu fund convex de 2800 ml șicane (volum 500 ml) acoperite printr-o spumă silicon. Flacoanele de producție s-au incubat pentru 24-48 h, în condițiile menționate mai sus. După incubare, bulioanele s-au distribuit în sticle sterile de polietilenă de 11, filat la 2600xg pentru o oră, la 10°C, și s-a decantat de celule și resturi de pelet. Apoi, bulionul lichid s-a filtrat la vid, prin filtre de sticlă Whatman GF/D (2,7 uM reținere) și GF/B (1,0 fM reținere), pentru a îndepărta resturile. Clarificarea bulionului, suplimentară, s-a realizat cu un dispozitiv de microfiltrare cu curgere tangențială (Pali Fiitron, Northborough, MA), folosind un filtru cu canal-deschis de 0,5 µM. Când este necesar, clarificarea suplimentară poate fi obținută prin răcirea bulionului (la 4°C) și centrifugare pentru câteva ore la 2600xg. După aceste proceduri, bulionul s-a sterilizat prin filtrare folosind un filtru cu membrană de nitroceluloză	27 29 31 33 35 37 39 41 43 45 47

1 0,2 μ M. Bulioanele sterile s-au folosit apoi direct pentru teste biologice, analize biochimice
sau s-au concentrat (până la 15-ori) folosind o deconectare de 10000 MW, dispozitiv de
3 ultra-filtrare M12 (Amicon, Beverly, MA) sau concentratori centrifugali (Millipore, Bedford, MA
și Pali Fltron, Northborough, MA) cu o mărime de por de 10000 MW. În cazul
5 concentratoarelor centrifugale, bulionul s-a filat la 2000xg pentru aproximativ 2 h. Pătrunsul
10000 MW s-a adăugat la reținerea corespunzătoare, pentru a obține concentrația dorită a
7 componentelor mai mari de 10000 MW. Inactivarea la cald a probelor de bulion procesat s-a
realizat prin încălzirea probelor la 100°C, într-un bloc de încălzire umplut cu nisip, pentru
9 10 min.

Bulionul(le) și complexul(le) toxină de la diferite tulpini de *Photorhabdus* sunt utilizate
11 pentru reducerea populațiilor de insecte și s-au folosit într-o metodă de inhibare a unei popu-
lații de insecte, care cuprinde aplicarea la un loc al insectei, a unei cantități eficiente de
13 inactivare insectă din activul descris. O demonstrație a întinderii activității insecticide obser-
vată de la bulionul unui grup selectat al tulpinilor *Photorhabdus* fermentate, cum s-a descris
15 mai sus, este arătată în tabelul 19. Este posibil ca activități insecticide suplimentare să poată
fi detectate cu aceste tulpini, prin concentrația crescută a bulionului sau prin folosirea diferi-
17 telor metode de fermentare. Compatibil cu activitatea asociată cu o proteină, activitatea
insecticidă a tuturor tulpinilor testate a fost termic labilă (vezi mai sus).

Bulionul(le) de cultură din diferite tulpini de *Photorhabdus* arată activitate insecticidă
19 diferențială (mortalitatea și/sau inhibarea creșterii a redus emergența adultului) împotriva
unui număr de insecte. Mai specific, activitatea se observă împotriva larvelor corn rootworm
21 și boli weevil, care sunt membri ai ordinului *Coleoptera*. Alți membri ai *Coleoptera* includ
23 wireworm, pollen beetles, flea beetles și gândacul cartofului de Colorado. Activitatea s-a
observat de asemenea și împotriva insectelor homeoptere din familia *Cicadellidae* și corn
25 plant hopper, care sunt membri ai familiei *Homoptera*. Alți membri ai *Homoptera* includ
planthopper, pear psylla, apple sucker, scale insects, whitefly, spittle bug, la fel de bine ca
27 numeroase gazde caracteristice speciilor afide. Bulioanele și complexe toxină purificată
sunt de asemenea active împotriva tobacco budworm, tobacco hornworm și European corn
29 borer, care sunt membri ai ordinului *Lepidoptera*. Alți membri tipici ai acestui ordin sunt beet
armyworm, cabbage looper, black cutworm, corn earworm, codling moth, clothes moth,
31 Indian mealmoth, leaf roller, cabbage worm, cotton bollworm, bagworm, Eastern tent
caterpillar, sod webworm și fall armyworm. Activitatea s-a observat de asemenea împotriva
33 larvelor fruitfly și țânțari care sunt membri ai ordinului *Diptera*. Alți membri ai ordinului
Diptera sunt: pea midge, carrot fly, cabbage root fly, turnip root fly, onion fly, crane fly și
35 musca de casă și diferite specii de țânțari. Activitatea cu bulion(e) și complex(e) toxină este
de asemenea observată împotriva two-spotted spider mite, care este un membru al ordinului
37 *Acarina*, care include strawberry spider mite, broad mite, citrus red mite, European red mite,
pear rust mite și tomato russet mite.

39 Activitatea împotriva larvelor corn rootworm s-a testat cum urmează: bulion(e) de
cultură *Photorhabdus* (concentrat de 0-15 ori, sterilizat prin filtrare), 2% Proteose Peptone
41 #3, complex(e) toxină purificat [0,23 mg/ml] sau tampon 10 mM fosfat de sodiu, pH 7,0 s-au
aplicat direct la suprafața (aproape 1,5 cm²) dietei artificiale (Rose, R.I. li McCabe, J.M.
43 (1973), *J. Econ. Entomol.* 66, (398-400) în 40 ui alicote. Complexul toxină s-a diluat în tam-
pon 10 mM fosfat de sodiu, pH 7,0. Plăcile cu dietă s-au lăsat la aer-uscat într-o nișă-curgere
45 sterilă și godeurile s-au infestat cu individ nou apărut *Diabrotica undecimpunctata howardi*
(Southern corn rootworm, SCR) crestat de pe suprafață de ouă sterilizate. Plăcile s-au etan-
47 șat, s-au plasat într-o cameră de creștere umezită și s-a menținut la 27°C, pentru perioada
corespunzătoare (3-5 zile). Apoi, s-au înregistrat procente de mortalitate și determinări de
49 greutate larvală. În general, în toate studiile s-au folosit 16 insecte per tratament. Mortalitatea
control a fost în general mai puțin de 5%.

RO 121280 B1

Activitatea împotriva boll weevil (<i>Anthonomus grandis</i>) s-a testat așa cum urmează.	1
Bulioane <i>Photorhabdus</i> concentrate (1-10 ori), mediu de control (2% Proteose Peptone #3), complex(e) toxină purificat [0,23 mg/ml] sau tampon 10 mM fosfat de sodiu, pH 7,0 s-au aplicat în 60 μ l alicote la suprafața 0,35 g dietă artificială (Stoneville Yellow Lepidoptera dietă) și s-au lăsat să se usuce. O singură larvă boli weevil de 12-24 h s-a plasat pe dietă și godeurile s-au etanșat și s-au ținut la 25°C, 50% RH, timp de 5 zile. Apoi s-au evaluat mortalitatea și greutatea larvale. Mortalitatea control s-a situat între 0-13%.	3
Activitatea împotriva larvelor de țânțari s-a testat după cum urmează. Testul s-a condus într-o placă microtitroare cu 96-godeuri. Fiecare godeu a conținut 200 μ l soluție apoasă (concentrată de 10 ori) de bulion(e) de cultură <i>Photorhabdus</i> , mediu control (2% Proteose Peptone #3), tampon 10 mM fosfat de sodiu, complex(e) toxină 0,23 mg/ml sau apă și aproximativ 20 larve în vârstă de o zi (<i>Aedes aegypti</i>). Au fost 6 godeuri per tratament. Rezultatele s-au citit la 3-4 zile după infestare. Mortalitatea control a fost între 0-20%.	5
Activitatea împotriva fruitflies s-a testat după cum urmează. Mediu <i>Drosophila melanogaster</i> cumpărat s-a preparat folosind 50% mediu uscat și 50% lichid sau apă, mediu control (2% Proteose Peptone #3), bulion(e) de cultură <i>Photorhabdus</i> concentrat de 10-ori, complex(e) toxină purificat [0,23 mg/ml] sau tampon 10 mM fosfat de sodiu, pH 7,0. Acesta s-a acoperit prin plasarea a 4,0 ml mediu uscat în fiecare din 3 fiole de creștere per tratament și adăugare de 4,0 ml de lichid corespunzător. 10 larve de <i>Drosophila melanogaster</i> în stadii larvare târzii s-au adăugat la fiecare fiolă de 25 ml. Fiolele s-au ținut pe o masă de laborator, la temperatura camerei, sub un plafon cu lumină fluorescentă. După 15 zile de expunere, s-au numărat adulții și pupele. Apariția adulților s-a comparat cu mediul control și apos (0-16% reducere).	7
Activitatea împotriva adulților aster leafhopper (<i>Macrostelus severini</i>) și nimfelor corn planthoppre (<i>Peregrinus maidis</i>) s-a testat cu un test de ingerare menit să permită ingerarea activ fără alt contact extern. Rezervorul pentru soluția activ/"hrană" este construit prin realizarea a două găuri în centrul porțiunii de jos a unui vas Petri 35x10 mm. S-a plasat pe vârful vasului un pătrat de Parafilm M® de 2 inch și s-a asigurat cu un inel "O". Apoi s-a infestat o cană de plastic de 1 oz, cu aproximativ 7 indivizi și rezervorul s-a plasat deasupra vasului Parafilm M®, în jos. Soluția test s-a adăugat apoi la rezervor, prin găuri. În testele care au folosit bulion(e) de cultură <i>Photorhabdus</i> concentrat de 10 ori, bulionul și mediu de control (2% proteoză peptonă #3) s-au dializat pe un tampon de 10 mM fosfat de sodiu, pH 7,0 și la soluția rezultantă s-a adăugat sucroză (la 5%), pentru a reduce mortalitatea controlului. Complexul(ele) toxină purificate [0,23 mg/ml] sau tampon fosfat de sodiu 10 mM, pH 7,0 s-au testat de asemenea. Mortalitatea s-a raportat în ziua a 3-a. Testele s-au ținut într-un incubator la 28°C, 70% RH cu o fotoperioadă 16/8. Testele s-au gradat pentru mortalitate la 72 h. Mortalitatea controlului a fost mai mică de 6%.	9
Activitatea împotriva larvelor lepidoteran s-a testat așa cum urmează. Bulion(e) de cultură <i>Photorhabdus</i> concentrat(e) (10 ori), mediu de control (2% proteoză peptonă #3), complex(e) toxină purificate [0,23 mg/ml] sau tampon fosfat de sodiu 10 mM, pH 7,0 s-au aplicat direct pe suprafața (~1,5 cm ²) regimului standard artificial lepidoteran (regimul Stoneville Yellow) în 40 μ l alicote. Plăcile cu hrană s-au lăsat să se usuce în aer, într-un curs steril acoperit, și fiecare godeu s-a infestat cu o singură larvă nou-apărută. Ouă European corn borer (<i>Ostrinia nubilalis</i>) și tobacco hornworm (<i>Manduca sexta</i>) s-au obținut din surse comerciale și s-au incubat local, în timp ce larve de tobacco budworm (<i>Heliothis virescens</i>) s-au furnizat intern. După infestarea cu larve, plăcile cu hrană s-au sigilat, s-au plasat într-o cameră de creștere umedă și s-au menținut la întuneric la 27°C, pentru o perioadă adecvată. Determinările mortalității și greutății s-au făcut în ziua a 5-a. În general, s-au folosit 16 insecte per tratament în toate studiile. Mortalitatea controlului a fost în general în domeniul 4-12,5% pentru mediul de control și mai puțin de 10% pentru tamponul fosfat.	11
Activitatea împotriva two-spotted spider mite (<i>Tetranychus urticae</i>) s-a determinat așa cum urmează. Plante tinere de dovlecel s-au tăiat la un singur cotiledon și s-au pulverizat	13

RO 121280 B1

1 pentru golire cu bulion(e) concentrat, mediu de control (2% proteoză peptonă #3),
 3 complex(e) toxină purificate [0,23 mg/ml] sau tampon fosfat de sodiu 10 mM, pH 7,0. După
 5 uscare, plantele s-au infestat cu o populație mixtă de spider mite și s-au ținut la temperatura
 și umiditatea laboratorului, timp de 72 h. Insectele vii s-au numărat pentru a determina nivelul
 de combatere.

Tabelul 19

Spectrul insecticid observat la bulioane de la diferite tulpini Photorhabdus

Tulpina <i>Photorhabdus</i>	Speciile de insecte sensibile*
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, 4, 5, 6, 7, 8
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, 9, 10
ATCC 43948	4
ATCC 43949	4
ATCC 43950	4
ATCC 43951	4
ATCC 43952	4

* = ≥25% mortalitate și/sau inhibare a creșterii vs. control

** = 1; tobacco budworm, 2; European corn borer, 3; tobacco hornworm, 4; Southern corn rootworm, 5; boll weevil, 6; țânțar, 7; fruit fly, 8; aster leafhopper, 9; hom planthopper, 10; two-spotted spider mite.

RO 121280 B1

Exemplul 14. *Tulpini non W-14 Photordabdus: purificare, caracterizare și spectru de activitate. Purificare.* 1

Protocolul, așa cum urmează, este similar cu cel dezvoltat pentru purificarea W-14 și s-a stabilit pe baza purificării acelor fracțiuni care au cea mai mare parte a activității împotriva Southern corn root worm (SCR), așa cum s-a determinat în bioteste (vezi exemplul 13). De obicei, s-au primit și s-au concentrat 4-20 L bulion care a fost filtrat, așa cum s-a descris în exemplul 13, folosind un cartuș de ultrafiltrare spirală Amicon tip S1 Y100, atașat la un dispozitiv de filtrare Amicon M-12. Retenentul a conținut proteine native care au greutatea moleculară mai mare de 100 kDa, în timp ce materialul de curgere a conținut proteine native mai mici de 100 kDa în mărime. Majoritatea activității împotriva SCR a fost conținută în retenentul de 100kDa. Retenentul s-a diafiltrat, apoi, în continuu cu fosfat de sodiu 10 MM (pH 7,0), până când filtratul a atins un $A_{280} < 0,100$. Dacă nu s-ar fi început astfel, din acest punct, toate procedurile s-au desfășurat în tampon definit prin fosfat de sodiu 10 mM (pH 7,0). Retenentul s-a concentrat apoi la un volum final de aproximativ 0,20 l și s-a filtrat utilizând o unitate de filtrare sterilă Nalgene™ Filterware 0,45 mm. Materialul filtrat s-a încărcat 7,5 ml/min. pe o coloană Pharmacia HR16/10 care fusese împachetată cu PerSeptive Biosystem Poros® 50 HQ matrice de schimbare puternică anioni echilibrată în tampon, utilizând un sistem HPLC PerSeptive Biosystem Sprint®. După încărcare, coloana s-a spălat cu tampon, până când s-a atins $A_{280} 0,100$. Proteinele s-au eluat apoi de pe coloană la 2,5 ml/min, timp de 20 min, până s-a atins un volum de 50 ml folosind tampon cu 0,4 M NaCl. Coloana s-a spălat apoi folosind tampon cu 1,0 M NaCl la aceeași rată de curgere, pentru încă 20 min (volum final = 50 ml). Proteinele eluate ce 0,4 M și 1,0 M NaCl s-au plasat apoi în truse de dializă separate (Spectra/Por® Membrane MWCO: 2.000) și s-au lăsat să dializeze peste noapte, la 4°C, în 12 L tampon. Majoritatea activității împotriva SCR a fost conținută în fracțiunea 0,4 M. Fracțiunea 0,4 M s-a purificat în continuare prin aplicarea unei coloane Pharmacia XK 26/100 care fusese preîmpachetată cu Sepharose CL4B (Pharmacia), folosind o rată de curgere de 0,75 ml/min. Fracțiunile s-au depozitat pe baza profilului peak-ului A_{280} și s-au concentrat la un volum final de 0,75 ml, utilizând un dispozitiv centrifugal de filtrare Millipore Ultrafree® -15 membrană Biomax-50 KNMWL. Concentrația proteinelor s-a determinat folosind un kit Biorad Protein Assay cu globulină standard bovină.

Caracterizarea

Greutatea moleculară nativă a complexului toxină SCR s-a determinat folosind o Pharmacia HR 16/50 care fusese preîmpachetată cu un tampon Sepharose CL4B. Coloana s-a calibrat apoi utilizând proteine cu greutatea moleculară cunoscută, permițând astfel calcularea mărimii moleculare native, aproximativă, a toxinei. Așa cum se arată în tabelul 20, mărimea moleculară a complexului toxină a fost în domeniul de la 777 kDa, cu tulpina Hb, până la 1.900 kDa cu tulpina WX-14. Randamentul complexului de toxine a variat, de asemenea, de la tulpina WX-12 care a produs 0,8 mg/L, la tulpina Hb care a produs 7,0 mg/L.

Proteinele găsite în complexul toxină s-au examinat pentru mărimea polipeptidei individuale, utilizând analiza SDS-PAGE. Tipic, 20 mg proteină a complexului toxină de la fiecare tulpină s-au încărcat pe un gel poliacrilamidă 2-15% (Integrated Separation Systems) și s-au supus electroforezei la 20 mA în tampon Biorad SDS-PAGE. După terminarea electroforezei, gelurile s-au colorat peste noapte în albastru Biorad Coomassie R-250 (0,2% în metanokacid acetic:apă; 40:10:40 v/v/v). În continuare, gelurile s-au decolorat în metanokacid acetic:apă; 40:10:40 (v/v/v). Gelurile s-au limpezit apoi cu apă, timp de 15 min, și s-au scanat folosind un Molecular Dynamics Personal Laser Densitometer®. Benzile s-au cuantificat și mărimea moleculară s-a calculat așa cum s-a comparat cu standardele de greutate moleculară înaltă Biorad, care au variat de la 200-45 kDa.

Mărimile polipeptidelor individuale care conțin complexul toxină SCR din fiecare tulpină sunt listate în tabelul 21. Mărimile polipeptidelor individuale au variat de la 230 kDa cu tulpina WX-1, la 16 kDa, așa cum s-a observat la tulpina WX-7. Fiecare tulpină, cu excepția tulpinii Hb, a avut polipeptide care cuprind complexul toxină care sunt în domeniul 160-230 kDa, domeniul 100-160 kDa și domeniul 50-80 kDa. Aceste date arată că respectivul complex toxină poate varia în compoziția peptidei și componentele de la tulpină la tulpină, totuși, în toate cazurile, atributele toxinei par să consistă într-un complex proteină oligamERIC.

Tabelul 20

Caracterizarea complexului toxină din tulpini *Photorhabdus non W-14*

Tulpina	Gr. molec. nativă aprox. ^a	Producere fracțiune activă (mg/L) ^b
Hb	972.000	1,8
Hb	777.000	7,0
Hm	1.400.000	1,1
HP88	813.000	2,5
NCI	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 Greutatea moleculară nativă determinată folosind o coloană Pharmacia HR 16/50 împachetată cu Sepharose CL4B;

^b Cantitatea de complex toxină recuperată de la bulionul de cultură.

Spectrul de activitate

Așa cum se prezintă în tabelul 21, complexe toxină purificate de la tulpini Hm și H9 s-au testat pentru activitatea împotriva unei varietăți de insecte, pentru comparare cu complexul toxină de la tulpina W-14. Testele s-au realizat așa cum s-a descris în exemplul 13. Complexul toxină de la toate trei tulpinile a arătat activitate împotriva tobacco budworm, European corn borer, Southern corn rootworm și aster leafhopper. În plus, complexul toxină de la tulpinile Hm și W-14 a arătat, de asemenea, activitate împotriva two-spotted spider mite. Mai mult, complexul toxină de la W-14 a arătat activitate împotriva larvelor de țânțar. Aceste date arată că, pe măsură ce ele au similarități în activitățile între diferite ordine de insecte, complexul toxină poate, de asemenea, arăta activități diferite împotriva altor ordine de insecte.

RO 121280 B1

Tabelul 21

Mărimile aproximative (kDa) ale peptidelor într-un complex toxină purificat de la non W-14 *Photorhabdus*

H9	Hb	Hm	HP	NC-1	WIR	WX-1	WX-2	WX-7	WX-12	WX-14	W-14
			88								
180	150	170	170	180	170	230	200	200	180	210	190
170	140	140	160	170	160	190	170	180	160	180	180
160	139	100	140	140	120	170	150	110	140	160	170
140	130	81	130	110	110	160	120	87	139	120	160
120	120	72	129	44	89	110	110	75	130	110	150
98	100	68	110	16	79	98	82	43	110	100	130
87	98	49	100		74	76	64	33	92	95	120
84	88	46	86		62	58	37	28	87	80	110
79	81	30	81		51	53	30	26	80	69	93
72	75	22	77		40	41		23	73	49	90
68	69	20	73		39	35		22	59	41	77
60	60	19	60		37	31		21	56	33	69
57	57		58		33	28		19	51		65
52	54		45		30	24		18	37		63
46	49		39		28	22		16	33		60
40	44		35		27				32		51
37	39				25				26		46
	37				23						40
	35										39
											29

Tabelul 22

Spectrul insecticid observat la un complex toxină purificat de la tulpini *Photorhabdus*

Tulplina <i>Photorhabdus</i>	Speciile de insecte sensibile*
Complex toxină Hm	1**, 2, 3, 5, 6, 7, 8
Complex toxină H9	1, 2, 3, 6, 7, 8
Complex toxină W-14	1, 2, 3, 4, 5, 6, 7, 8

* = >25% mortalitate sau inhibarea creșterii;

** = 1; Tobacco bud worm, 2; European corn borer, 3; Southern corn root worm, 4) Mosquito, 5; Two-spotted spider mite, 6; Aster leafhopper, 7; fruit Fly, 8; Boli Weevil.

Exemplul 15. Subfracționarea complexului toxină proteină *Photorhabdus*

Complexul toxină proteină *Photorhabdus* s-a izolat așa cum s-a descris în exemplul 14. În continuare, s-au aplicat aproximativ 10 mg toxină la o coloană MonoQ 5/5 echilibrată cu 20 mM Tris-HCl, pH 7,0 la o rată de curgere de 1 ml/min. Coloana s-a spălat cu 20 mM Tris-HCl, pH 7,0, până când densitatea optică la 280 nm s-a întors la absorbanta liniei de bază. Proteinele legate la coloană s-au eluat cu un gradient linear de 0 la 1,0 M NaCl în 20 mM Tris-HCl, pH 7,0 la 1 ml/min pentru 30 min. S-au colectat fracțiuni de 1 ml și s-au supus la biotest Southern corn rootworm (SCR) (vezi exemplul 13). Peak-urile activității s-au determinat printr-o serie de diluții ale fiecărei fracțiuni în bioteste SCR. Două peak-uri de activitate împotriva SCR s-au observat și s-au numit A (eluat la aproape 0,2-0,3 M NaCl) și B (eluat la 0,3-0,4 M NaCl). Peak-urile de activitate A și B s-au depozitat separat și ambele peak-uri s-au purificat suplimentar folosind o procedură în 3 etape, descrisă mai jos.

$(\text{NH}_4)_2\text{SO}_4$ solid s-a adăugat la fracțiunea de proteină de mai sus, la o concentrație finală de 1,7 M. Apoi, proteinele s-au aplicat la o coloană fenil-Superose 5/5 echilibrată cu 1,7 M $(\text{NH}_4)_2\text{SO}_4$ în tampon 5 mM fosfat de potasiu, pH 7 la 1 ml/min. Proteinele legate la coloană s-au eluat cu un gradient liniar de 1,7 M $(\text{NH}_4)_2\text{SO}_4$, 0% etilen glicol, 50 mM fosfat de potasiu, pH 7,0, la 25% etilen glicol, 25 mM fosfat de potasiu, pH 7,0 (nu $(\text{NH}_4)_2\text{SO}_4$ la 0,5 ml/min. Fracțiunile s-au dializat peste noapte peste un tampon 10 mM fosfat de sodiu, pH 7,0. Activitățile în fiecare fracțiune împotriva SCR s-au determinat prin biotest.

Fracțiunile cu activitate mai ridicată s-au depozitat și s-au aplicat la o coloană MonoQ 5/5, care s-a echilibrat cu 20 mM Tris-HCl, pH 7,0 la 1 ml/min. Proteinele legate la coloană s-au eluat la 1 ml/min printr-un gradient linear de la 0 până la 1 M NaCl în 20 mM Tris-HCl, pH 7,0.

Pentru etapa finală de purificare, fracțiunile mai active, de mai sus, (determinate prin biotest SCR) s-au depozitat și s-au supus la o a doua coloană fenil-Superose 5/5. $(\text{NH}_4)_2\text{SO}_4$ solid s-a adăugat la o concentrație finală de 1,7 M. Apoi, soluția s-a încărcat la coloana echilibrată cu 1,7 M $(\text{NH}_4)_2\text{SO}_4$ în tampon 50 mM fosfat de potasiu, pH 7 la 1 ml/min. Proteinele legate la coloană s-au eluat cu un gradient linear de 1,7 M $(\text{NH}_4)_2\text{SO}_4$ mM fosfat de potasiu, pH 7,0, la 10 mM fosfat de potasiu, pH 7,0 la 0,5 ml/min. Fracțiunile s-au dializat peste noapte împotriva unui tampon 10 mM fosfat de sodiu, pH 7,0. Activitățile în fiecare fracțiune împotriva SCR s-au determinat prin biomasă.

Proteina finală purificată prin procedura în 3 etape de mai sus, de la peak-ul A, s-a numit toxina A, iar proteina finală purificată de la peak-ul B s-a numit toxina B.

Caracterizarea și secvențarea aminoacidă a toxinei A și toxinei B în SDS-PAGE

Ambele toxine A și B au conținut două peptide majore (>90% din proteina totală colorată Commaassie): 192 kDa (numită A1 și, respectiv, B1) și 58 kDa (numite A2 și, respectiv, B2). Ambele toxine A și B relevă numai o bandă majoră în PAGE nativă, indicând că A1 și A2 sunt subunități ale unui complex proteină, iar B1 și B2 sunt subunități ale unui complex proteină. Suplimentar, greutatea moleculară nativă a ambelor toxine A și B s-a determinat, prin cromatografie filtrare gel, a fi 860 kDa. Concentrațiile molare relative ale A1 și A2 s-au apreciat a fi o echivalență 1 la 1, cum s-a determinat prin analiza densitometrică a gelurilor SDS-PAGE. În mod similar, peptidele B1 și B2 au prezentat aceeași concentrație molară.

Toxina A și toxina B s-au supus electroforezei în 10% SDS-PAGE și s-au transblottat la membrane PVDF. Blott-urile s-au transmis pentru analiza aminoacidă și secvențarea aminoacidă N-terminală la Harvard MicroChem, și respectiv, Cambridge ProChem. Amino secvența N-terminală a lui B1 s-a determinat a fi identică la SEQ ID NR:1, regiunea TcbA₁ a genei tcbA (SEQ ID NR: 12, poziția 87 la 99). O secvență N-terminală unică s-a obținut pentru peptida B2 (SEQ ID NR:40). Secvența aminoacidă N-terminală a peptidei B2 a fost

identică cu regiunea TcbA_{iii} a secvenței aminoacide derivate de la gena tcbA (SEQ ID NR:12, poziția 1935 la 1945). De aceea, toxina B conținea predominant 2 peptide, TcbA_i și TcbA_{iii}, care s-a observat că sunt derivate de la același produs genă, TcbA.

Secvența N-terminală a A2 (SEQ ID NR:41) a fost unică în comparație cu peptida TcbA_{iii} și alte peptide. Peptida A2 s-a denumit TcdA_{iii} (vezi exemplul 17). SEQ ID NR:6 s-a determinat a fi un amestec al secvențelor aminoacide SEQ ID NR: 40 și 41.

Peptidele A1 și A2 s-au supus în continuare secvențării aminoacide interne. Pentru secvențarea aminoacidă internă s-au supus electroforezei 10 μg de toxină A în 10% SDS-PAGE și s-au transblottat la o membrană PVDF. După ce blott-ul s-a colorat cu amido black, peptidele A1 și A2, denumite TcdA_i și respectiv TcdA_{iii}, s-au excizat de pe blot și s-au trimis la Harvard MicroChem și la Cambridge ProChem. Peptidele s-au supus digerației tripsin, urmată de cromatografia HPLC, pentru a separa peptidele individuale. Analiza aminoacidului N-terminal s-a realizat pe fragmentele de peptidă triptic selectate. Două secvențe aminoacide interne ale peptidei A (TcdA_i-PK71, SEQ ID NR:38 și TcdA_i-PK44, SEQ ID NR:39) s-au găsit a avea omologii semnificative cu secvențele aminoacide deduse ale regiunii Tcb_i a genei tcbA (SEQ ID NR:12). Similar, secvența N-terminală (SEQ ID NR:41) și 2 secvențe interne ale peptidelor A2 (TcdA_{iii}-PK57, SEQ ID NR:42 și TcdA_{iii}-PK20, SEQ ID NR:43) au prezentat, de asemenea, omologie semnificativă cu secvențele aminoacide deduse ale regiunii TcbA_{iii} ale genei tcbA (SEQ ID NR:12).

În rezumatul rezultatelor de mai sus, complexul toxină are cel puțin 2 complexe toxină proteină active împotriva SCR: toxina A și toxina B. Toxina A și toxina B sunt similare în greutatea lor moleculară native și în subunități, totuși, compozițiile lor de peptide sunt diferite. Toxina A conține peptidele TcdA_i și TcdA_{iii} ca peptide majoritare și toxina B conține TcbA_i și TcbA_{iii} ca peptide majoritare.

Exemplul 16. Clivarea și activarea peptidelor TcbA

În complexul toxină B, peptidele TcbA_i și TcbA_{iii} creează de la un singur produs genă TcbA (exemplul 15). Procesarea peptidei TcbA la TcbA_i și TcbA_{iii} este prezumată prin acțiunea proteazei(lor) *Photorhabdus* și, cel mai probabil, a metaloproteazelor descrise în exemplul 10. În unele cazuri, s-a notat că, atunci când bulionul W-14 *Photorhabdus* s-a procesat, peptida TcbA s-a prezentat în complexul B toxină, ca o componentă majoră, în adăugare cu peptidele TcbA_i și TcbA_{iii}. Proceduri identice, descrise pentru purificarea complexului toxină B (exemplul 15), s-au folosit pentru a îmbogăți peptida TcbA de la fracțiunea complex toxină a bulionului W-14. Materialul final purificat s-a analizat într-un gradient 4-20% SDS-PAGE și peptidele majore s-au cuantificat prin densitometrie. S-a determinat că TcbA, TcbA_i și TcbA_{iii} au cuprins 58%, 36% și respectiv 6% din proteina totală. Identitățile acestor peptide s-au confirmat prin măsurile lor moleculare în analizele SDS-PAGE și Western Blot, folosind anticorpi specifici. Greutatea moleculară nativă a acestei fracțiuni s-a determinat a fi de 860 kDa.

Clivajul TcbA s-a evaluat prin tratarea materialului purificat de mai sus cu 38 kDa și 58 kDa metaloprotează W-14 *Photorhabdus* (exemplul 10) și Tripsin ca enzimă de control (Sigma, MO). Reacția standard a constatat în 17,5 μg din fracția purificată de mai sus, 1,5 unități protează și 0,1 M tampon Tris, pH 8,0, într-un volum total de 100 μl. Pentru reacția de control s-a omis protează. Amestecul de reacție s-a incubat la 37°C, timp de 90 min. La sfârșitul reacției, s-au luat 20 μl și s-au fiert cu tampon probă SDS-PAGE pentru analiza electroforetică în SDS-PAGE gradient 4-20%. Din SDS-PAGE s-a determinat că, atât în tratamentul cu protează 38 kDa, cât și în cel cu protează 58 kDa, cantitatea de peptide tcbA_i și TcbA_{iii} a crescut de 3 ori, în timp ce cantitatea de peptidă TcbA a scăzut proporțional (tabelul 23). Reducerea relativă și creșterea peptidelor selectate s-au confirmat prin analiză Western Blot.

Mai mult, filtrarea gelului materialului clivat a relevat că mărimea moleculară nativă a complexului a rămas aceeași. Prin tratament cu Trypsină, peptidele TcbA_{ii} și TcbA_{iii} au fost digerate nonspecific în peptide mici. Aceasta arată că proteazele *Photorhabdus* 38 kDa și 58 kDa pot procesa specific peptida TcbA, în peptidele TcbA_{ii} și TcbA_{iii}. Controlul protează tratat și netratat al celor 80 μl rămași din amestecul de reacție s-a diluat în serie cu 10 mM tompo fosfat de sodiu, pH 7,0 și s-a analizat prin biotest SCR. Prin compararea activității la diferite diluții, s-a determinat că tratamentul protează 38 kDa a crescut activitatea insecticidă SCR de aproximativ 3 până la 4 ori. Inhibarea creșterii insectelor rămase în tratament protează este, de asemenea, mai gravă decât cea a controlului (tabelul 23).

Tabelul 23

Conversia și activarea peptidei TcbA în peptidele TcbA_{ii} și TcbA_{iii} prin tratament protează

	Control	Tratament protează 38 kDa
S0 (% din total proteină)	58	18
S1 (% din total proteină)	36	64
S9 (% din total proteină)	6	18
LD 50 (μg proteină)	2,1	0,52
Greutate SCR (mg/insectă)*	0,2	0,1

* - un indicator al inhibării creșterii prin măsurarea greutății medii a insectelor vii, după 5 zile de regim în test.

Exemplul 17. Cercetarea băncii pentru gene care codifică peptida TcdA_{ii}

S-au finalizat donarea și caracterizarea genei care codifică peptida TcdA_{ii}, descrisă ca SEQ ID NR:17 (secvența peptidei interne TcdA_{ii}-PT111 N-terminală) și SEQ ID NR:18 (secvența peptidei interne TcdA_{ii}-PT79 N-terminală). S-au sintetizat așa cum s-a descris în exemplul 8, 2 depozite de oligonucleotide degenerate proiectate să codifice secvențele aminoacide a SEQ ID NR:17 (tabelul 24) și SEQ ID NR:18 (tabelul 25) și complementele inverse ale acestora două. Secvența ADN a oligonucleotidelor este dată mai jos:

RO 121280 B1

Tabelul 24

Oligonucleotida degenerată pentru SEQ ID NR:17

P2-PT11	1	2	3	4	5	6	7	8
Amilacid	Ala	Phe	Asn	Ile	Asp	Asp	Val	Ser
Codonii	5' GGN	TTT(TTC)	AA(TTC)	ATT(TTCA)	GAT(TAC)	GAT(TAC)	GTN(3')	
P2.3.6CB	5' GC(A/C/G/T)	TT(TTC)	AAT	ATT	GAT	GAT	GT(3')	
P2.3.5	5' OC(A/C/G/T)	TT(TTC)	AA(TTC)	ATT(TTCA)	GAT(TAC)	GAT(TAC)	GT(3')	
P2.3.5R	5' AC	(GAA)TC	(GAA)TC	(TGA)AT	(GATT)	(GAA)AA	(A/C/G/T)CC(3')	
P2.3.5RI	5' ACI	TCI	TCI	ATI	TTI	AAI	OC(3')	
P2.3R/CB	5' CAO	(A/C)AC	(A/C)AC	ATC	ATC	AAT	ATT	AAA(3')

Tabelul 25

Oligonucleotida degenerată pentru SEQ ID NR:18

P2-PT9	1	2	3	4	5	6	7	8	9	10	11	12	13
Amilacidizi	Phe	Ile	Val	Tyr	Thr	Ser	Leu	Gly	Val	Asn	Pro	Asn	Asn
Codonii	5' TTY	ATH	GTN	TAY	ACN	6	6	GON	GTN	AA(3')	CCN	AA(3')	AA(3')
P2.79.2	5' TTY	ATY	GTK	TAT	ACY	TCI	YTR	GOY	GTK	AAT	CCR	AAT	AAT(3')
P2.79.3	5' TTY	ATT	GTK	TAT	ACY	AGY	YTR	GOY	GTK	AAT	CCR	AAT	AAT(3')
P2.79.R.1	5' ATT	ATT	YGG	ATT	MAC	RCC	YAR	RCT	RCT	ATA	MAC	AAT	AAA(3')
P2.79R/CB	5' ATT	ATT	YGG	ATT	MAC	ACC	CAO	RCT	RCT	ATA	MAC	AAT	AAA(3')

* codurile pentru nucleotide în conformitate cu IUPAC-IUB, Y=C sau T, H=A, C sau T, N=A, C, G sau T, K=G sau T, R= A sau G și M=A sau C.

1 Reacții polimerazice de lanț (PCR) s-au realizat în principal cum s-a descris în
 3 exemplul 8, utilizând ca primeri forward P2.3.6.CB sau P2.3.5 și ca primeri reverse
 P2.79.R.1. sau P2.79R.CB, în toate combinațiile forward/reverse, folosind ADN genomic
 5 *Photobabdus* W-14 ca matrită. În alt set de reacții, primerii P2.79.2 sau P2.79.3 s-au folosit
 ca primeri înainte și P.2.3.5R, P2.3.5RJ și P2.3R.CB s-au utilizat ca primeri inversi în toate
 7 combinațiile înainte/invers. Doar în reacțiile conținând P2.3.6.CB ca primeri înainte combinat
 cu P2.79.R. 1. sau P2.79R.CB ca primeri invers s-a observat un produs amplificat
 9 non-artefactual cu mărime estimată (mobilitate pe geluri agaroză) de 2500 perechi baze.
 Ordinul primerilor folosiți pentru a obține acest produs de amplificare arată că fragmentul
 peptidă TcdA_{PT111} este situat amino-proximal la fragmentul peptidă TcdA_{PT79}.

11 Produsele PCR 2500 bp s-au ligat la vectorul plasmid pCM™II (Invitrogen, San
 Diego, CA), conform instrucțiunilor furnizorilor, iar secvențele ADN situate de-a lungul
 13 capetelor fragmentelor insert ale celor două izolate (HS24 și HS27) s-au determinat utilizând
 primerii recomandați de furnizori și metodele de secvențare descrise anterior. Secvența
 15 ambelor izolate a fost aceeași. Primerii noi s-au sintetizat pe baza secvenței determinate și
 s-au folosit la reacțiile de secvențare primer adițional, pentru a obține un total de 2557 baze
 17 ale insertului [SEQ ID NR:36]. Translația peptidei parțiale codificată prin SEQ ID NR:36 a dat
 secvența 845 aminoacidică descrisă ca SEQ ID NR:37. Analiza omologiei de proteină a
 19 acestei porțiuni a fragmentului de peptidă TcdA[^] arată o omologie aminoacidă substanțială
 (68% similaritate; 53% identitate) la resturile 542 la 1390 ale proteinei TcbA [SEQ ID NR: 12].
 21 Cu toate acestea, se pare că gena reprezentată în parte prin SEQ ID NR:36 produce o
 proteină a secvenței aminoacide similare, dar nu identică, precum proteina TcbA și a cărei
 23 asemănare a avut activitate similară, dar nu identică cu proteina TcbA.

Într-o altă realizare, o genă care codifică peptidele TcdA_{PK44} și peptidă N-terminală
 25 TcdA_{III} de 58 kDa, descrisă ca SEQ ID NR:9 (secvența TcdA_{PK44} peptidă internă), s-a
 izolat SEQ ID NR:41 (secvența peptidei N-terminale TcdA_{III} de 58 kDa). Două depozite de
 27 oligonucleotide degenerate, desemnate a codifica secvențele aminoacide, descrise precum
 SEQ ID NR:39 (tabelul 27) și SEQ ID NR:41 (tabelul 26) și complemente reverse ale acestor
 29 secvențe, s-au sintetizat în exemplul 8 și secvențele lor AND.

RO 121280 B1

1

Tabelul 26

3

Oligonucleotida degenerată pentru SEQ ID NR:41

Codon #	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Aminoacid	Leu	Arg	Ser	Ala	Asn	Thr	Leu	Thr	Asp	Leu	Phe	Leu	Pro	Gln
A2.1	5' YTR	COY	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	RTG	YAR	RTG	RJT	IGC	RCT	RCG 3'
A2.4.R				5' TG	IGG	CAG	AAA	CAG	RTC	IGT	CAG	IGT	ATT	IGC 3'

11

13

15

17

Tabelul 27

Oligonucleotida degenerată pentru SEQ ID NR:39

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

19

21

23

25

27

S-au realizat Polymerase Chain Reactions (PCR) în esență, așa cum s-a descris în exemplul 8, folosind ca primeri înainte Al.44.1 sau Al.44.2 și ca primeri inverși A2.3R sau A2.4R în toate combinațiile înainte/invers, folosind ca matriță ADN genomic *Photothabdus* W-14. Într-un alt set de reacții, primerii A2.1 sau A2.2 s-au folosit ca primeri înainte și A1.44.1R și Al.44.2R s-au folosit ca primeri inverși în toate combinațiile înainte/invers. Doar în reacțiile care au conținut Al.44.1 sau Al.44.2 ca primeri înainte combinați cu A2.3R ca primer invers, s-a observat un produs amplificat non-artefactual, de mărime estimată (mobilitate pe geluri de agaroză) de 1400 baze perechi. Ordinea primerilor utilizați pentru a obține acest produs de amplificare arată că fragmentul de peptidă TcdA_{ii} TK44 este situat amino-proximal la fragmentul peptidă de 58 kDa al TcdA_{iii}.

Produsele PCR de 1400 bp s-au ligat la vectorul plasmid pCR™II conform instrucțiunilor furnizorului. Secvențele ADN de-a lungul capetelor fragmentelor insert a patru izolate s-au determinat utilizând primeri similari în secvență la primerii recomandați de furnizor și utilizând metode de secvențare descrise anterior. Secvența de acid nucleic a tuturor izolatelor diferă, așa cum s-a așteptat, în regiuni care corespund secvențelor primer degenerate, dar secvențele aminoacide deduse din aceste date au fost aceleași precum secvențele aminoacide actuale pentru peptide determinate anterior (SEQ ID NR:41 și 39).

Cercetarea băncii cosmid genomic W-1.4, așa cum s-a descris în exemplul 8 cu o probă radiomarcată, a cuprins ADN-ul preparat mai sus (SEQ ID NR:36), a identificat cinci cosmizi izolați care hibridizează 19D9, 20B10, 21D2, 27B10 și 26D1. Acești compuși au fost diferiți de cei identificați anterior cu probe care corespund genelor descrise ca SEQ ID NR:11 sau SEQ ID NR:25. Analizele enzimă de restricție și hibridizări blot ADN au identificat fragmente EcoR I, de mărimi aproximative 3,7, 3,7 și 1,1 kbp, care cuprind regiunea cu ADN-ul SEQ ID NR:36. Cercetarea băncii cosmid genomic W-14, utilizând ca probă fragmentul ADN radiomarcant de 1,4 kbp preparat în acest exemplu, a identificat aceiași cinci cosmizi (17D9, 20B10, 21D2, 27B10 și 26D1). Hibridizarea blot ADN la EcoR I-digerat ADN-uri cosmid, de asemenea, a arătat hibridizare la același subset al fragmentelor EcoR I, cum s-a observat cu proba genă TcdA_{ii} de 2,5 kbp, care arată că ambele fragmente sunt codificate pe ADN genomic.

Determinarea secvenței ADN a fragmentelor EcoR I donat a arătat un cadru de citire neîntrerupt de 7551 baze perechi (SEQ ID NR:46), care codifică o proteină de 282,9 kDa de 2516 aminoacizi (SEQ ID NR:47). Analiza secvenței aminoacide a acestei proteine a revelat toate fragmentele interne ale peptidelor TcdA_{ii} SEQ ID NR:17, 18, 37, 38 și 39 și N-terminusul peptidă TcdA_{iii} (SEQ ID NR:41) și toate peptidele interne TcdA_{iii} (SEQ ID NR:42 și 43) așteptate. Peptidele izolate și identificate ca TcdA_{ii} și TcdA_{iii} sunt fiecare produs al cadrului deschis de citire, denumit tcdA, descris ca SEQ ID NR:46. În continuare, SEQ ID NR:47 arată, începând de la poziția 89, secvența descrisă ca SEQ ID NR:13 care este secvența N-terminală a unei peptide de mărime aproximativă 201 kDa, care arată că proteina inițială produsă de la SEQ ID NR:46 este procesată într-o manieră similară cu cea descrisă mai devreme pentru SEQ ID NR:12. În plus, proteina este mai departe clivată pentru a genera un produs cu mărimea 209,2 kDa, codificat de SEQ ID NR:48 și descris ca SEQ ID NR:49 (peptida TcdA_{ii}) și un produs cu mărimea de 63,6 kDa codificat de SEQ ID NR:50 și descris ca SEQ ID NR:51 (peptida TcdA_{iii}). Astfel, se consideră că activitatea insecticidă identificată ca toxina A (exemplul 15) derivată de la produsele SEQ ID NR:46 exemplificate prin proteina de lungime întreagă de 282,9 kDa, descrisă ca SEQ ID NR:47, este procesată pentru a produce peptidele descrise ca SEQ ID NR:49 și 51. Se consideră că activitatea insecticidă identificată ca toxina B (exemplul 15) derivă de la produsele SEQ ID NR:11, așa cum s-a exemplificat prin proteina de 280,6 kDa descrisă ca SEQ ID NR:12. Această proteină este procesată proteolitic pentru a obține peptida de 207,6 kDa descrisă ca SEQ ID NR:53, care este codificată de SEQ ID NR:52 și peptida de 62,9 kDa care are secvența

N-terminală descrisă ca SEQ ID NR:40 și mai departe descrisă ca SEQ ID NR:55, care este codificată de SEQ ID NR:54.

Comparațiile secvenței aminoacide, între proteinele descrise ca SEQ ID NR:12 și SEQ ID NR:47, au arătat că ele au similaritate 69% și identitate 54%. Acest grad înalt de relații în evoluție nu este uniform de-a lungul întregii secvențe aminoacide a acestor peptide, dar este mai mare spre capătul carboxi-terminal al proteinelor, pe măsură ce peptidele descrise ca SEQ ID NR:51 (derivată de la SEQ ID NR:47) și SEQ ID NR:55 (derivată de la SEQ ID NR:12) au similaritate 76% și identitate 64%.

Exemplul 18. Combaterea distrugerii de porumb

Capacitatea toxinei(lor) *Photorhabdus* de a reduce distrugerea plantei cauzată de larvele insectelor s-a demonstrat prin măsurarea distrugerii frunzei cauzată de European cornborer (*Ostrinia nubilalis*) infestată pe plante de porumb tratate cu bulion *Photorhabdus*. Fermentarea bulionului de la *Photorhabdus* tulpina W-14 s-a produs și s-a concentrat de aproximativ 10 ori prin ultrafiltrare (10000 MW mărime por), așa cum s-a descris în exemplul 13. Bulionul concentrat rezultat s-a sterilizat apoi prin filtrare, folosind ca filtre membrane de nitroceluloză 0,2 μ . O probă preparată similar, de 2% proteose peptone #3 neinoculată, s-a folosit în scop de control. Plantele de porumb (o linie încrucișată proprietatea DowElanco) s-au crescut din semințe la stadiul vegetativ 7 sau 8, în vase care conțin un amestec fără pământ, într-o seră (27°C ziua; 22°C noaptea, aproape 50% RH, 14 h zi-lumină, udate/fertilizate după nevoie). Plantele test s-au aranjat într-o schemă bloc complet randomizat (3 repetiții/tratament, 6 plante/tratament), într-o seră cu temperatura de aproape 22°C ziua; 18°C noaptea, fără lumină artificială și cu umbră parțială, aproape 50% RH și udate/fertilizate după nevoie. S-au aplicat tratamente (mediu neinoculat și bulion concentrat *Photorhabdus*) cu o seringă spray, 2,0 ml aplicat direct (aproape 6 inch) deasupra verticilului și încă 2,0 ml aplicați într-o mișcare circulară de la aproximativ un picior deasupra verticilului. În plus, un grup de plante nu a primit tratament. După ce tratamentele s-au uscat (aproximativ 30 min), 12 larve nou apărute, European corn borer (obținute din surse comerciale și incubate cu mijloace proprii), s-au aplicat direct pe verticil. După 1 săptămână, plantele s-au cercetat pentru distrugerea frunzelor, folosind o scală Guthrie modificată (Koziel, M.G., Belnad, G.L., Bowman, C., Carozzi, N.B., Crenshaw, R., Crossland L., Dawson, 1, Desai, N., Hill, M., Kadwell, S., Launis, K., Lewis, K., Maddox, D., McPherson, K., Meghji, M.Z., Merlin, E., Rhodes, R., Warren, G.W., Wright, M. și Evola, S.V., 1993, *Bio/Technology*, 11, 194-195) și măsurătorile s-au comparat statistic [T-test (LSD) $p < 0,05$ și Tukey's Studentized Range (HDS) Test $p < 0,1$]. Rezultatele sunt prezentate în tabelul 28. Ca referință, 1 înseamnă nici o distrugere, 2 reprezintă distrugere fină "fereastră" pe frunza nedesfăcută fără penetrație por și 5 reprezintă penetrația frunzei cu leziuni alungite și/sau nervura de mijloc mâncată, evidentă la mai mult de 3 frunze (leziuni < 1 inch). Aceste date arată că bulionul sau alte proteine care conțin fracțiunile pot conferi protecție împotriva anumitor insecte parazite, când se administrează în formula spray sau când gena sau derivații acesteia care codifică proteina, sau o parte a acesteia este administrată *via* o plantă transgenică sau un microb.

Tabelul 28

Efectul bulionului de cultură *Photorhabdus* pe distrugerea frunzei induse

Tratament	Măsurare Guthrie medie
Fără tratament	5,02 ^a
Mediu neinoculat	5,15 ^a
Bulion <i>Photorhabdus</i>	2,24 ^b

Mediile cu litere diferite sunt diferite statistic ($p < 0,05$ sau $p < 0,1$).

Exemplul 19. Ingineria genetică a genelor pentru exprimarea în *E.coli*

Rezumatul construcțiilor

S-au construit o serie de plasmizi pentru a exprima gena *Photorhabdus W-14* în *Escherichia coli*. În tabelul 29, este prezentată o listă a plasmizilor. S-au obținut date pentru o descriere pe scurt a fiecărui construct ca un rezultat al expresiei *E.coli*.

Tabelul 29

Plasmizi de expresie pentru gena *tcbA*

Plasmid	Genă	Vector /Selecție	Comportament
pDAB634	<i>tcbA</i>	pBC/Chl	Intracelular
pAcGP67/ <i>tcbA</i>	<i>tcbA</i>	pAcG67B/Amp	Baculovirus secretat
pDAB635	<i>tcbA</i>	pET27b/Kan	Periplasm
pET15- <i>tcbA</i>	<i>tcbA</i>	pET15- <i>tcbA</i>	Intracelular

Abrevieri: Kan=kanamicină, Chl=cloramfenicol, Amp=ampicilină

Construcția pDAB634

În exemplul 9, este descris un fragment mare *EcoR* I care hibridizează la proba *TcbA_{ii}*. Acest fragment s-a subclonat în pBC (Stratagene, La Jolla CA). Analiza de secvență a arătat că acest fragment este format din 8816 perechi baze. Fragmentul codifică gena *tcbA* cu ATG care inițiază la poziția 571 și TAA care termină la poziția 8086. Prin urmare, fragmentul poartă 570 perechi baze de ADN *Photorhabdus* în amonte de ATG și 730 perechi baze în avalul TAA.

Construcția plasmidului pAcGP67B/*tcbA*

Gena *tcbA* s-a amplificat PCR folosind primerii următori: 5' primer (S1Ac51) 5' TTT AAA CCA TGG GAA ACT CAT TAT CAA GCA CTA TC 3' primer (S1Ac31) 5' TTT AAA GCG GCC GCT TAA CGG ATG GTA-TAA CGA TAT TG 3'. S-a realizat PCR folosind un kit TaKaRa LA PCR de la PanVera (Madison, Wisconsin) în reacția următoare: 57,5 ml apă, 10 ml tampon 10X LA, 16 ml dNTPs (2mM fiecare soluție stoc), 20 ml fiecare primer la 10 pmoli/ml, 300 ng plasmid pDAB634 care conține gena *tcbA* și 1 ml polimerază TaKaRa LA Taq. Condițiile de ciclare au fost 98°C/20 s, 68°C/5 min, 72°C/10 min pentru 30 cicluri. Un produs PCR așteptat de aproape 7526 bp s-a izolat într-un gel de 0,8% agaroză în tampon TBE (100 mM Tris, 90 mM acid boric, 1 mM EDTA) și s-a purificat utilizând un kit Qiaex II de la Qiagen (Chatsworth, California). Gena *tcbA* purificată s-a digerat cu *Nco* I și *Not* I și s-a ligat în vectorul transfer baculovirus pAcGP67B (PharMingen (San Diego, California)) și s-a transformat în *E.coli* DH5a. Gena *tcbA* s-a tăiat apoi de la pAcGP67B și s-a transferat la pET27b pentru a crea plasmidul pDAB635. S-a reparat în pDAB635 o mutație în sens greșit, în gena *tcbA*.

Gena *tcbA* reparată conține 2 modificări de la secvența arătată în SEQID NR: 11; un A>G la schimbarea 212, o aspargină 71 la serină 71 și un G>A la schimbarea 229, o alanină 77 la treonină 77. Aceste schimbări sunt amândouă mai sus de N-terminusul *TcbA_{ii}* propus.

Construcția pET15-*tcbA*

Regiunea care codifică *tcbA* a pDAB635 s-a transfectat la vectorul pET15b. Aceasta s-a realizat utilizând ligări shotgun, ADN-urile s-au tăiat cu enzimele de restricție *Nco* I și *Xho* I. Recombinantul rezultat s-a denumit pET15-*tcbA*.

Exprimarea *TcbA* în *E.coli* de la plasmidul pET15-*tcbA*

RO 121280 B1

Exprimarea *tcbA* în *E.coli* s-a obținut prin modificarea metodelor anterioare descrise de Studier et al. (Studier, F.W., Rosenberg, A., Dunn, J. și Dubendorff, J., (1990) *Use of T7 RNA polymerase to direct expression of cloned genes. Methods Enzymol.*, 185, 60-89). Celule corespunzătoare 2s.col/tulpina BL21(DE3) s-au transformat cu plasmidul pET15-*tcbA* și s-au depus pe agar LB care conține 100 μg ampicilina și 40 mM glucoza. Celulele transformate s-au depus la o densitate de câteva sute de colonii izolate/placă. După incubare peste noapte, la 37°C, celulele s-au curățat de pe plăci și s-au suspendat în bulion LB care conține 100 μg ampicilina. Volumele tipice de cultură au fost de 200-500 ml. La momentul zero, densitățile culturii (OD₆₀₀) erau de la 0,15-0,5, în funcție de experiment. Culturile s-au scuturat la una din trei temperaturi (22°C, 30°C sau 37°C), până când densitatea de 0,15-0,5 s-a obținut la momentul în care ele s-au indus cu 1 mM izopropiltio-β-galactozidă (IPTG). Culturile s-au incubat la temperatura desemnată, timp de 4-5 h, și apoi s-au transfectat la 4°C până la procesare (12-72 h).

Purificarea și caracterizarea *TcbA* exprimată în *E.coli* de la plasmidul pET15-*tcbA*

Culturile *E.coli* care exprimă peptide *TcbA* s-au procesat așa cum urmează. Celulele s-au recoltat prin centrifugare la 17.000 x G și mediul s-a decantat și s-a păstrat într-un container separat.

Mediul s-a concentrat de aproape 8x folosind sistemul de filtrare M12 (Amicon, Beverly MA) și un filtru de separare cu masa moleculară 100 kDa. Mediul concentrat s-a încărcat pe o coloană schimb anion și proteinele legate s-au eluat cu 1,0 M NaCl. S-a constatat că peak-ul de eluție 1,0 M NaCl a cauzat mortalitatea larvelor Southern corn rootworm (SCR) (tabelul 30). Frațiunea 1,0 M NaCl s-a dializat pe tampon 10 mM fosfat de sodiu pH 7,0, concentrat și s-a supus la filtrare pe gel pe Sepharose CL-4B (Pharmacia, Piscataway, New Jersey). Regiunea profilului de eluție CL-4B care corespunde masei moleculare calculate (aproape 900 kDa), precum complexul toxină W-14 nativă, s-a colectat, concentrat și biotestat împotriva larvelor. Frațiunea colectată de 900 kDa s-a găsit a avea activitate insecticidă (vezi tabelul 30, mai jos) cu simptome similare cu cea cauzată de complexul toxină W-14 nativ. Această fracțiune s-a supus la Proteinase K și tratament termic, activitatea în ambele cazuri fie s-a redus, fie a fost eliminată, furnizând dovada că activitatea este de natură proteică în natură. În plus, fracțiunea activă s-a testat imunologic pozitiv pentru peptidele *TcbA* și *TcbA_{III}* în analize imunoblot, când s-a testat un anticorp monoclonal anti-*TcbA_{III}* (tabelul 30).

Tabelul 30

Rezultatele biotestelor imunoblot și SCR

Fracțiunea	Activitate SCR		Imunoblot	Mărime nativă
	Mortalitate %	Inhibarea creșterii %		Mărime estimată CL-4B
Mediu Tcb 1,0 M	+++	+++	TcbA	
Schimb de ioni				
Mediu TcbA CL-4B	+++	+++	Tcb, TcbA _{III}	-900 kDa
Mediu TcbA CL-4B + Proteinsase K	++	+++	NT	
Mediu TcbA CL-4B + tratament termic	-	-	NT	
Celule TcbA Sup CL-4B	-	+++	NT	~900 kD

PK= Tratament cu Proteinase K, 2 h; Tratament termic= 100°C, timp de 10 min; ND= nedetectat; NT= netestat. Sistemul de măsurare pentru mortalitate și inhibarea creșterii, așa cum s-a comparat cu probele de control; 5-24%= "+", 25-49%= "++", 50-100%= "+++".

Peletul de celule s-a resuspendat în tampon 10 mM fosfat de sodiu, pH 7,0 și s-a uzat prin trecere printr-un dispozitiv de pulverizare celule Bio-Neb™ (Glas-Col Inc., Terra Haute, IN). Peletele s-au tratat cu DNază pentru a îndepărta ADN-ul și s-au centrifugat la 17.000xg, pentru a separa celulele pelet de supernatantul celule. Frațiunea supernatant s-a decantat și s-a filtrat printr-un filtru 0,2μ, pentru a îndepărta particulele mari și s-a supus la cromatografie schimb de anioni. Proteinele legate s-au eluat cu 1,0 M NaCl, s-au dializat și s-au concentrat utilizând concentratori Biomax™ (Millipore Corp, Bedford, MA) cu greutate moleculară îndepărtată de 50000Daltoni. Frațiunea concentrată s-a supus cromatografiei filtrare gel, utilizând matriță granulată Sepharose CL-4B. Datele biotestului pentru materialul preparat în acest mod sunt prezentate în tabelul 30 și sunt denumite "celule TcbA sup".

Într-o altă metodă, pentru a mări cantități mari de material, peletele celule s-au resuspendat în tampon fosfat de sodiu 10 mM, pH 7,0 și s-au omogenizat complet, utilizând un mecanism de lustruit țesut de 40 ml Kontes Glass Company (Vineland, NJ). Restul celular s-a peletat prin centrifugare la 25000xg și supernatantul celule s-a decantat, s-a trecut printr-un filtru 0,2 μ și s-a supus cromatografiei schimb de anioni, utilizând o coloană Pharmacia 10/10 împachetată cu granule Poros HQ 50. Proteinele legate s-au eluat prin realizarea unui gradient NaCl de 0,0 la 1,0 M. Frațiunile care conțin proteina TcbA s-au combinat și s-au concentrat utilizând un concentrator 50 kDa și s-au supus la cromatografie filtrare pe gel, utilizând o matrice granulată Pharmacia CL-4B. Frațiunile care conțin oligomer TcbA, masă moleculară aproximativ 900 kDa, s-au colectat și s-au supus la cromatografie schimb de anioni, utilizând o coloană Pharmacia Mono Q 10/10 echilibrată cu tampon Tris 20 mM, pH 7,3. S-a folosit un gradient de 0,0 la 1,0 M NaCl pentru a elua proteina TcbA recombinantă. TcbA recombinantă s-a eluat de la coloană la o concentrație salină de aproximativ 0,3-0,4 M NaCl, aceeași molaritate la care oligomerul TcbA s-a eluat de pe coloana Mono Q 10/10. Frațiunea recombinantă TcbA s-a găsit a cauza mortalitatea SCR în experimentele biotest, similare cu acelea prezentate în tabelul 30.

Revendicări

1. Polinucleotidă care este asociată operabil cu un promotor heterolog, în care respectiva polinucleotidă codifică o proteină care are activitate de toxină împotriva unui dăunător insectă, în care respectiva proteină cuprinde secvența de aminoacizi aleasă dintre SEQ ID NR:12 și SEQ ID NR:47.

2. Celulă de plantă transgenică, care cuprinde o polinucleotidă conform revendicării 1.

3. Celulă de plantă transgenică, care exprimă o polinucleotidă conform revendicării 1.

4. Proteină recombinantă care are activitate de toxină împotriva unui dăunător insectă, în care respectiva proteină cuprinde secvența de aminoacizi aleasă dintre SEQ ID NR:12 sau SEQ ID NR:47.

5. Metodă de control al unui dăunător insectă, în care metoda respectivă cuprinde hrănirea cu o proteină conform revendicării 4, a dăunătorului menționat.

6. Metodă conform revendicării 5, în care proteina menționată este produsă de o polinucleotidă și este prezentă într-o plantă transgenică, care este accesibilă dăunătorului menționat.

7. Metodă de control al unui dăunător insectă, în care metoda respectivă cuprinde alimentarea respectivului dăunător cu o proteină din *Photographus*, în care proteina menționată are activitate de toxină, pe cale orală, și în care proteina menționată este produsă de și este prezentă într-o plantă transgenică, care este accesibilă dăunătorului menționat.

RO 121280 B1

LISTA DE SECVENȚE	1
(1) INFORMAȚIE GENERALĂ:	3
(i) SOLICITANT:	
Ensign, Jerald C	
Bowen, David J	5
Petell, James	
Fatig, Raymond	7
Schoonover, Sue	
French-Constant, Richard	9
Orr, Gregory L	
Merlo, Donald J	11
Roberts, Jean L	
Rocheleau, Thomas A	13
Blackburn, Michael B	
Hey, Tomothy D	15
Strickland, James A	
(ii) TITLUL INVENȚIEI: Toxine proteine insecticide din <i>Photorhabdus</i>	17
(iii) NUMĂR SECVENȚE: 61	
(iv) ADRESA PENTRU CORESPONDENȚĂ:	19
(A) ADRESANT: Quarly & Brady	
(B) STRADA: 1 South Pinckney Street	21
(C) ORAȘ: Madison	
(D) STAT: WI	23
(E) ȚARA: US	
(F) ZIP: 53703	25
(v) FORMA CITIBILĂ PE COMPUTER:	
(A) TIP MEDIU: Disc floppy	27
(B) COMPUTER: compatibil PC IBM	
(C) SISTEM DE OPERARE: PC-DOS/MS-DOS	29
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30	
(vi) DATE CURENTE ALE CERERII:	31
(A) NUMĂR CERERE:	
(B) DATA DEPUNERII:	33
(C) CLASIFICARE:	
(vii) DATA DE PRIORITATE A CERERII:	35
(A) NUMĂR CERERE: US 08/063 615	
(B) DATA DEPUNERII: 18 mai 1993	37
(vii) DATA DE PRIORITATE A CERERII:	
(A) NUMĂR CERERE: US 08/395 497	39
(B) DATA DEPUNERII: 28 februarie 1995	
(vii) DATA DE PRIORITATE A CERERII:	41
(A) NUMĂR CERERE: US 60/007 255	
(B) DATA DEPUNERII: 06-nov-1995	43
(vii) DATA DE PRIORITATE A CERERII:	
(A) NUMĂR CERERE: US 08/608 423	45
(B) DATA DEPUNERII: 28-febr-1996	
(vii) DATA DE PRIORITATE A CERERII:	47
(A) NUMĂR CERERE: US 08/705 484	
(B) DATA DEPUNERII: 28-aug-1996	49

- 1 (viii) INFORMAȚIE MANDATAR/AGENT:
 3 (A) NUME: Seay, Nicholas J
 (B) NUMĂR DE ÎNREGISTRARE: 27386
 (C) NUMĂR REFERINȚĂ/DOSAR: 960296.93804
- 5 (ix) INFORMAȚIE TELECOMUNICAȚII:
 (A) TELEFON: 608-251-5000
 7 (B) TELEFAX: 608-251-9166
- (2) INFORMAȚIA PENTRU SECV. ID. NR. 1:
 9 (i) CARACTERISTICI DE SECVENȚĂ:
 (A) LUNGIME: 11 aminoacizi
 11 (B) TIP: aminoacid
 (C) ÎMPLETIRE:
 13 (D) TOPOLOGIE: lineară
 (ii) TIPUL MOLECULEI: proteină
- 15 (v) TIPUL FRAGMENTULUI: N-terminal
- 17 (xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 1:
 19 Phe Ile Gln Gly Tyr Ser Asp Leu Phe Gly Asn
 1 5 10
- 21 (2) INFORMAȚIA PENTRU SECV. ID. NR. 2:
 (i) CARACTERISTICI DE SECVENȚĂ:
 23 (A) LUNGIME: 12 aminoacizi
 (B) TIP: aminoacid
 25 (C) ÎMPLETIRE:
 (D) TOPOLOGIE: lineară
 27 (ii) TIPUL MOLECULEI: proteină
 (v) TIPUL FRAGMENTULUI: N-terminal
- 29 (xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 2:
 31 Met Gln Asp Ser Pro Glu Val Ser Ile Thr Thr Trp
 1 5 10
- 33
- INFORMAȚIA PENTRU SECV. ID. NR. 3:
 35 (i) CARACTERISTICI DE SECVENȚĂ:
 (A) LUNGIME: 19 aminoacizi
 37 (B-) TIP: aminoacid
 (C) ÎMPLETIRE:
 39 (D) TOPOLOGIE: lineară
 (ii) TIPUL MOLECULEI: proteină
- 41 (v) TIPUL FRAGMENTULUI: N-terminal
- (xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. :3:
 43 (2) INFORMAȚIA PENTRU SECV. ID. NR. 4:
 (i) CARACTERISTICI DE SECVENȚĂ:
 45 (A) LUNGIME: 14 aminoacizi
 (B) TIP: aminoacid
 47 (C) ÎMPLETIRE:
 (D) TOPOLOGIE: lineară

RO 121280 B1

(ii) TIPUL MOLECULEI: proteină	1
(v) TIPUL FRAGMENTULUI: N-terminal	
(xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 4	3
Ala Ser Pro Leu Ser Thr Ser Glu Leu Thr Ser Lys Leu Asn 1 5 10	5
(2) INFORMAȚIA PENTRU SECV. ID. NR. 5:	7
(i) CARACTERISTICI DE SECVENȚĂ:	
(A) LUNGIME: 9 aminoacizi	9
(B) TIP: aminoacid	
(C) ÎMPLETIRE:	11
(D) TOPOLOGIE: lineară	
(ii) TIPUL MOLECULEI: proteină	13
(v) TIPUL FRAGMENTULUI: N-terminal	
(xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 5:	15
Ala Gly Asp Thr Ala Asn Ile Gly Asp 1 5	17
(2) INFORMAȚIA PENTRU SECV. ID. NR. 6:	19
(i) CARACTERISTICI DE SECVENȚĂ:	
(A) LUNGIME: 15 aminoacizi	21
(B) TIP: aminoacid	
(C) ÎMPLETIRE:	23
(D) TOPOLOGIE: lineară	
(ii) TIPUL MOLECULEI: proteină	25
(v) TIPUL FRAGMENTULUI: N-terminal	
(xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. :6:	27
Leu Gly Gly Ala Ala Thr Leu Leu Asp Leu Leu Pro Gln Ile 1 5 10 15	29
(2) INFORMAȚIA PENTRU SECV. ID. NR. 7:	31
(i) CARACTERISTICI DE SECVENȚĂ:	
(A) LUNGIME: 11 aminoacizi	33
(B) TIP: aminoacid	
(C) ÎMPLETIRE:	35
(D) TOPOLOGIE: lineară	
(ii) TIPUL MOLECULEI: proteină	37
(v) TIPUL FRAGMENTULUI: N-terminal	
(xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 7	39
Met Leu Ser Thr Met Glu Lys Gln Leu Asn Glu 1 5 10	41
(2) INFORMAȚIA PENTRU SECV. ID. NR. 8:	43
(i) CARACTERISTICI DE SECVENȚĂ:	
(A) LUNGIME: 9 aminoacizi	45
(B) TIP: aminoacid	
(C) ÎMPLETIRE:	47
(D) TOPOLOGIE: lineară	
(ii) TIPUL MOLECULEI: proteină	49
(v) TIPUL FRAGMENTULUI: N-terminal	
(xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 8:	51

RO 121280 B1

1 Met Asn Leu Ala Ser Pro Leu Ile Ser
1 5

3

5 (2) INFORMAȚIA PENTRU SECV. ID. NR. 9:
 (i) CARACTERISTICI DE SECVENȚĂ:
 (A) LUNGIME: 16 aminoacizi
 (B) TIP: aminoacid
 (C) ÎMPLETIRE:
 (D) TOPOLOGIE: lineară
 (ii) TIPUL MOLECULEI: proteină
 (v) TIPUL FRAGMENTULUI: N-terminal
 (xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 9:

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

15

17 (2) INFORMAȚIE PENTRU SECV. ID NR 10:
 (i) CARACTERISTICI DE SECVENȚĂ:
 (A) LUNGIME: 20 aminoacizi
 (B) TIP: aminoacid
 (C) ÎMPLETIRE:
 (D) TOPOLOGIE: lineară
 (ii) TIP MOLECULĂ: proteină
 (v) TIP FRAGMENT: N-terminal
 (xi) DESCRIEREA SECVENȚEI: SECV. ID NR: 10:

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

27 Arg Gly Val Asn
20

29 (2) INFORMAȚIA PENTRU SECV. ID. NR. 11:
 (i) CARACTERISTICI DE SECVENȚĂ:
 (A) LUNGIME: 7515 baze perechi
 (B) TIP: acid nucleic
 (C) ÎMPLETIRE: dublă
 (D) TOPOLOGIE: lineară
 (ii) TIPUL MOLECULEI: proteină
 (v) TIPUL FRAGMENTULUI: ADN (genomic)
 (ix) TRĂSĂTURA:
 (A) NUME/COD: CDS
 (B) LOCALIZARE: 1..7515

39

41 (xi) DESCRIEREA SECVENȚEI: SECV. ID. NR.: 11:
 ATG CAA AAC TCA TTA TCA AGC ACT ATC GAT ACT ATT TGT CAG AAA CTG 43
 Met Gln Asn Ser Leu Ser Ser Thr Ile Asp Thr Ile Cys Gln Lys Leu
1 5 10 15

43

45 CCA 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

47

49 CCG GAA AAA ACT CCG GGA ATG GTT AAT TGG GGG GAA GCA AAA CCG ATT 144
 Arg Glu Lys Thr Arg Gly Met Val Asn Trp Gly Glu Ala Lys Arg Ile
35 40 45

51

53 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 GGC TAT GCT AAT CCG CTG CTG AAA AAC GCT GTT CCG TTC 240
 Arg Ile Phe Ala Tyr Ala Asn Pro Leu Leu Lys Asn Ala Val Arg Leu

RO 121280 B1

65	70	75	80		1
GGT ACC CGG CAA ATG TTG GGT TTT ATA CAA GGT TAT AGT GAT CTG TTT	288				3
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				5
Gly Asn Arg Ala Asp Asn Tyr Ala Ala Pro Gly Ser Val Ala Ser Met					
100	105	110			7
TTC TCA CCG GCG GCT TAT TTG ACG GAA TTG TAC CGT GAA GCC AAA AAC	384				9
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				11
Leu His Asp Ser Ser Ser Ile Tyr Tyr Leu Asp Lys Arg Arg Pro Asp					13
130	135	140			
TTA GCA AGC TTA ATG CTC AGC CAG AAA AAT ATG GAT GAG GAA ATT TCA	480				15
Leu Ala Ser Leu Met Leu Ser Gln Lys Asn Met Asp Glu Glu Ile Ser					
145	150	155	150		
ACG CTG GCT CTC TCT AAT GAA TTG TGC CTT GCC GCG ATC GAA ACA AAA	528				17
Thr Leu Ala Leu Ser Asn Glu Leu Cys Leu Ala Gly Ile Glu Thr Lys					19
165	170	175			
ACA GGA AAA TCA CAA GAT GAA GTG ATG GAT ATG TTG TCA ACT TAT CGT	576				21
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				23
Leu Ser Gly Glu Thr Pro Tyr His His Ala Tyr Glu Thr Val Arg Glu					25
195	200	205			
ATC GTT CAT GAA CGT GAT CCA GGA TTT CGT CAT TTG TCA CAG GCA CCC	672				27
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				29
Ile Val Ala Ala Lys Leu Asp Pro Val Thr Leu Leu Gly Ile Ser Ser					31
225	230	235	240		
CAT ATT TCG CCA GAA CTG TAT AAC TTG CTG ATT GAG GAG ATC CCG GAA	768				33
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				35
Lys Asp Glu Ala Ala Leu Asp Thr Leu Tyr Lys Thr Asn Phe Gly Asp					37
260	265	270			
ATT ACT ACT GCT CAG TTA ATG TCC CCA AGT TAT CTG GCC CGG TAT TAT	864				39
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				41
Gly Val Ser Pro Glu Asp Ile Ala Tyr Val Thr Thr Ser Leu Ser His					43
290	295	300			
GTT GGA TAT AGC AGT GAT ATT CTG GTT ATT CCG TTG GTC GAT GGT GTG	960				45
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				47
Gly Lys Met Glu Val Val Arg Val Thr Arg Thr Pro Ser Asp Asn Tyr					
325	330	335			

RO 121280 B1

1	ACC AGT CAG ACG AAT TAT ATT GAG CTG TAT CCA CAG GGT GGC GAC AAT	1056
3	Thr Ser Gln Thr Asn Tyr Ile Glu Leu Tyr Pro Gln Gly Gly Asp Asn	
	340 345 350	
5	TAT TTC ATC AAA TAC AAT CTA AGC AAT AGT TTT GGT TTC GAT GAT TTT	1104
	Tyr Leu Ile Lys Tyr Asn Leu Ser Asn Ser Phe Gly Leu Asp Asp Phe	
	355 360 365	
7	TAT CTG CAA TAT AAA GAT GGT TCC GCT GAT TGG ACT GAG ATT GCC CAT	1152
9	Tyr Leu Gln Tyr Lys Asp Gly Ser Ala Asp Trp Thr Glu Ile Ala His	
	370 375 380	
11	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	
13	ACA ATC AAA CGT AGT GAC TCT GAC AAT ATA CTC AGT ATA GGG TTA CAA	1248
15	Thr Ile Lys Arg Ser Asp Ser Asp Asn Ile Leu Ser Ile Gly Leu Gln	
	405 410 415	
17	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	
19	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	
21	CGG TTG CTC AAA GCT ACC GGC CTC TCT TTT GCT ACG TTG GAG CGT ATT	1392
23	Arg Leu Leu Lys Ala Thr Gly Leu Ser Phe Ala Thr Leu Glu Arg Ile	
	450 455 460	
25	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	
27	AAG GTT TAT CGG GTA AAA TTC TAT ATT GAT CGT TAT GGC ATC AGT GAA	1488
29	Lys Val Tyr Arg Val Lys Phe Tyr Ile Asp Arg Tyr Gly Ile Ser Glu	
	485 490 495	
31	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	
33	GGC AAT CAG CTT ACC CAG TTT GAG CAA CTA TTT AAT CAC CCG CCG CTC	1584
35	Gly Asn Gln Leu Ser Gln Phe Glu Gln Leu Phe Asn His Pro Pro Leu	
	515 520 525	
37	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	
39	AAT CCT GAT CTC AAC CTT AAA CCA GAC AGT ACC GGT GAT GAT CAA GGC	1680
	Asn Pro Asp Leu Asn Leu Lys Pro Asp Ser Thr Gly Asp Asp Gln Arg	
	545 550 555 560	
41	AAG GCG GTT TTA AAA CGC GCG TTT CAG GTT AAC GCC AGT GAG TTG TAT	1728
43	Lys Ala Val Leu Lys Arg Ala Phe Gln Val Asn Ala Ser Glu Leu Tyr	
	565 570 575	
45	CAG ATC 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	
47	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	

RO 121280 B1

595	600	605	1
ATT CAT AAC CTG ACT ATT GCT GAA TTG AAC ATT TTG TTG GTG ATT TGT 1872			3
Ile His Asn Leu Thr Ile Ala Glu Leu Asn Ile Leu Leu Val Ile Cys			
610	615	620	5
GGC TAT GGC GAC ACC AAC ATT TAT CAG ATT ACC GAC GAT AAT TTA GCC 1920			7
Gly Tyr Gly Asp Thr Asn Ile Tyr Gln Ile Thr Asp Asp Asn Leu Ala			
625	630	635	9
AAA ATA GTG GAA ACA TTG TTG TGG ATC ACT CAA TGG TTG AAG ACC CAA 1958			11
Lys Ile Val Glu Thr Leu Leu Trp Ile Thr Gln Trp Leu Lys Thr Gln			
645	650	655	13
AAA TGG ACA GTT ACC GAC CTG TTT CTG ATG ACC ACG GCC ACT TAC AGC 2016			15
Lys Trp Thr Val Thr Asp Leu Phe Leu Met Thr Thr Ala Thr Tyr Ser			
660	665	670	17
ACC ACT TTA ACG CCA GAA ATT AGC AAT CTG ACG GCT ACG TTG TCT TCA 2064			19
Thr Thr Leu Thr Pro Glu Ile Ser Asn Leu Thr Ala Thr Leu Ser Ser			
575	680	685	21
ACT TTG CAT GGC AAA GAG AGT CTG ATT GGG GAA GAT CTG AAA AGA GCA 2112			23
Thr Leu His Gly Lys Glu Ser Leu Ile Gly Glu Asp Leu Lys Arg Ala			
690	695	700	25
ATG GCG CCT TGC TTC ACT TCG GCT TTG CAT TTG ACT TCT CAA GAA GTT 2150			27
Met Ala Pro Cys Phe Thr Ser Ala Leu His Leu Thr Ser Gln Glu Val			
705	710	715	29
GGC TAT GAC CTG CTG TTG TGG ATA GAC CAG ATT CAA CCG CCA CAA ATA 2208			31
Ala Tyr Asp Leu Leu Leu Trp Ile Asp Gln Ile Gln Pro Ala Gln Ile			
725	730	735	33
ACT GTT GAT GGG TTT TGG GAA GAA GTG CAA ACA ACA CCA ACC AGC TTG 2256			35
Thr Val Asp Gly Phe Trp Glu Glu Val Gln Thr Thr Pro Thr Ser Leu			
740	745	750	37
AAG GTG ATT ACC TTT GCT CAG GTG CTG GCA CAA TTG AGC CTG ATC TAT 2304			39
Lys Val Ile Thr Phe Ala Gln Val Leu Ala Gln Leu Ser Leu Ile Tyr			
755	760	765	41
CGT CGT ATT GGG TTA AGT GAA ACG GAA CTG TCA CTG ATC GTG ACT CAA 2352			43
Arg Arg Ile Gly Leu Ser Glu Thr Glu Leu Ser Leu Ile Val Thr Gln			
770	775	780	45
TCT TCT CTG CTA GTG GCA GGC AAA AGC ATA CTG GAT CAC GGT CTG TTA 2400			47
Ser Ser Leu Leu Val Ala Gly Lys Ser Ile Leu Asp His Gly Leu Leu			
785	790	795	
ACC CTG ATG GCC TTG GAA GGT TTT CAT ACC TGG GTT AAT GGC TTG GGC 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	

RO 121280 B1

1	ACA	CAG	ATT	GAC	GCT	ATT	CTG	CAA	TGG	TTA	CAG	ATG	TCT	TCG	GCC	TTG	2540
3	Thr	Gln	Ile	Asp	Ala	Ile	Leu	Gln	Trp	Leu	Gln	Met	Ser	Ser	Ala	Leu	865 870 875 880
5	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
7	ATA	GAT	CAT	AAC	TAT	GCT	GCC	TGG	CAA	GCT	GCG	GCG	GCT	GCG	CTG	ATG	2736
9	Ile	Asp	His	Asn	Tyr	Ala	Ala	Trp	Gln	Ala	Ala	Ala	Ala	Ala	Leu	Met	900 905 910
11	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
13	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
15	GGA	GTA	CGT	GAT	CGT	AAC	GCT	TTA	TAT	ACC	TAT	TTG	CTG	ATT	GAT	AAT	2880
17	Gly	Val	Arg	Asp	Arg	Asn	Gly	Leu	Tyr	Thr	Tyr	Leu	Leu	Ile	Asp	Asn	945 950 955 960
19	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
21	GGT	ATT	CAA	CTG	TAC	GTT	AAC	CGG	GCT	TTA	AAC	CGA	GAT	GAA	GGT	CAG	2976
23	Gly	Ile	Gln	Leu	Tyr	Val	Asn	Arg	Ala	Leu	Asn	Arg	Asp	Glu	Gly	Gln	980 985 990
25	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
27	TAC	AAT	AAA	CGT	TAC	AGT	ACT	TGG	GCT	GGT	GTC	TCT	GAA	CTG	GTC	TAT	3072
29	Tyr	Asn	Lys	Arg	Tyr	Ser	Thr	Trp	Ala	Gly	Val	Ser	Glu	Leu	Val	Tyr	1010 1015 1020
31	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
33	ATC	ATC	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
35	GAT	ACG	GTG	GAA	GAT	GCT	TTC	AAA	ACT	TAT	TTG	ACC	ACC	TTT	CAG	CAG	3216
37	Asp	Thr	Val	Glu	Asp	Ala	Phe	Lys	Thr	Tyr	Leu	Thr	Ser	Phe	Glu	Gln	1060 1065 1070
39	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
41	GAT	CAA	GGA	TTA	ACT	TAT	TTT	ATC	GGT	ATC	GAC	CAA	GCA	GCT	CCG	GGT	3312
43	Asp	Gln	Gly	Leu	Thr	Tyr	Phe	Ile	Gly	Ile	Asp	Gln	Ala	Ala	Pro	Gly	1090 1095 1100
45	ACG	TAT	TAC	TGG	CGT	AGT	GTT	GAT	CAC	AGC	AAA	TGT	GAA	AAT	GCG	AGG	3350
	Thr	Tyr	Tyr	Trp	Arg	Ser	Val	Asp	His	Ser	Lys	Cys	Glu	Asn	Gly	Lys	1105 1110 1115 1120
47	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	

RO 121280 B1

	1125	1130	1135	1
AAT CCT TGG AAA AAT ATC ATC CGT CCG GTT GTT TAT ATG TCC CGC TTA 3456				3
Asn Pro Trp Lys Asn Ile Ile Arg Pro Val Val Tyr Met Ser Arg Leu	1140	1145	1150	
				5
TAT CTG CTA TGG CTG GAG CAG CAA TCA AAG AAA AGT GAT GAT GGT AAA 3504				7
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				9
Thr Thr Ile Tyr Gln Tyr Asn Leu Lys Leu Ala His Ile Arg Tyr Asp	1170	1175	1180	
				11
GGT AGT TGG AAT ACA CCA TTT ACT TTT GAT GTG ACA GAA AAG GTA AAA 3600				13
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				15
Asn Tyr Thr Ser Ser Thr Asp Ala Ala Glu Ser Leu Gly Leu Tyr Cys	1205	1210	1215	
				17
ACT GGT TAT CAA GGG GAA GAC ACT CTA TTA GTT ATG TTC TAT TCG ATG 3696				19
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				21
Gln Ser Ser Tyr Ser Ser Tyr Thr Asp Asn Asn Ala Pro Val Thr Gly	1235	1240	1245	
				23
CTA TAT ATT TTC GCT GAT ATG TCA TCA GAC AAT ATG ACG AAT GCA CAA 3792				25
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				27
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				29
Ala Asp Pro Asp Ser Asp Asn Lys Lys Val Ile Thr Arg Arg Val Asn	1285	1290	1295	
				31
AAC CGT TAT CCG GAG GAT TAT GAA ATT CCT TCC TCT GTG ACA AGT AAC 3936				33
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				35
Ser Asn Tyr Ser Trp Gly Asp His Ser Leu Thr Met Leu Tyr Gly Gly	1315	1320	1325	
				37
AGT GTT CCT AAT ATT ACT TTT GAA TCG CCG GCA CAA GAT TTA AGG CTA 4032				39
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 ACT ATT ATT CAT AAT GGA TAT GCG GGA ACC 4080				41
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 CCT TCA TTA GGT GAT 4128				43
Arg Arg Ile Gln Cys Asn Leu Met Lys Gln Tyr Ala Ser Leu Gly Asp	1365	1370	1375	
				45
AAA TTT ATA ATT TAT GAT TCA TCA TTT GAT GAT GCA AAC CGT TTT AAT 4176				47
Lys Phe Ile Ile Tyr Asp Ser Ser Phe Asp Asp Ala Asn Arg Phe Asn	1380	1385	1390	

RO 121280 B1

1
CTG GTG CCA TTG TTT AAA TTC GGA AAA GAC GAG AAC TCA GAT GAT AGT 4284
Leu Val Pro Leu Phe Lys Phe Gly Lys Asp Glu Asn Ser Asp Asp Ser
1395 1400 1405

3
ATT TGT ATA TAT AAT GAA AAC CCT TCC TCT GAA GAT AAG AAG TCG TAT 428
Ile Cys Ile Tyr Asn Glu Asn Pro Ser Ser Glu Asp Lys Lys Trp Tyr
1410 1415 1420

5
TTT TCT TCG AAA GAT GAC AAT AAA ACA GCG GAT TAT AAT CGT GGA ACT 4324
Phe Ser Ser Lys Asp Asp Asn Lys Thr Ala Asp Tyr Asn Gly Gly Thr
1425 1430 1435 1440

7
CAA TGT ATA GAT GCT GGA ACC AGT AAC AAA GAT TTT TAT TAT AAT CTC 436
Gln Cys Ile Asp Ala Gly Thr Ser Asn Lys Asp Phe Tyr Tyr Asn Leu
1445 1450 1455

9
CAG GAG ATT GAA GTA ATT AGT GTT ACT GGT GGG TAT TGG TCG AGT TAT 4414
Gln Glu Ile Glu Val Ile Ser Val Thr Gly Gly Tyr Trp Ser Ser Tyr
1460 1465 1470

11
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

13
GTA AAA GTC ACC GTA AAA CCG GGT GGT GAC GAT CAA ATC TTT ACT GCT 4514
Val Lys Val Thr Val Lys Ala Gly Gly Asp Asp Gln Ile Phe Thr Ala
1490 1495 1500

15
GAT AAT AGT ACC TAT GTT CCT CAG CAA CCG GCA CCC AGT TTT GAG GAG 4564
Asp Asn Ser Thr Tyr Val Pro Gln Gln Pro Ala Pro Ser Phe Glu Glu
1505 1510 1515 1520

17
ATG ATT TAT CAG TTC AAT AAC CTG ACA ATA GAT TGT AAG AAT TTA AAT 4604
Met Ile Tyr Gln Phe Asn Asn Leu Thr Ile Asp Cys Lys Asn Leu Asn
1525 1530 1535

19
TTC ATC GAC AAT CAG GCA CAT ATT GAG ATT GAT TTC ACC GCT ACG GCA 4654
Phe Ile Asp Asn Gln Ala His Ile Glu Ile Asp Phe Thr Ala Thr Ala
1540 1545 1550

21
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

23
AAA AAA GTT CTC GGT ACT GAG AAC GTG ATT GCG TTA TAT AGC GAA AAT 4752
Lys Lys Val Leu Gly Thr Glu Asn Val Ile Ala Leu Tyr Ser Glu Asn
1570 1575 1580

25
AAC CGT 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

27
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

29
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

31
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

33
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

RO 121280 B1

1650	1655	1660	1
GAG AAC GAT AGT TTT GTG ATT TAT CAA GGA GAA CTT ACC GAA ACA AGT 5010			3
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			5
Gln Thr Val Val Lys Val Phe Leu Ser Tyr Phe Ile Glu Ala Thr Gly			
1685	1690	1695	7
AAT AAG AAC CAC TTA TGG GTA CGT GCT AAA TAC CAA AAG GAA ACG ACT 5136			9
Asn Lys Asn His Leu Trp Val Arg Ala Lys Tyr Gln Lys Glu Thr Thr			
1700	1705	1710	
GAT AAG ATC TTC TTC GAC CGT ACT GAT GAG AAA GAT CCG CAC GGT TGG 5184			11
Asp Lys Ile Leu Phe Asp Arg Thr Asp Glu Lys Asp Pro His Gly Trp			
1715	1720	1725	13
TTT CTC AGC GAC GAT CAC AAG ACC TTT AGT GGT CTC TCT TCC GCA CAG 5232			15
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 GCG GCC AAT CCT 5280			17
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			19
Leu Tyr Phe Trp Glu Leu Phe Tyr Tyr Thr Pro Met Met Met Ala His			
1765	1770	1775	21
CGT TTG TTG CAG GAA CAG AAT TTT GAT GCG GCG AAC CAT TGG TTC CGT 5376			23
Arg Leu Leu Gln Glu Gln Asn Phe Asp Ala Ala Asn His Trp Phe Arg			
1780	1785	1790	25
TAT GTC TGG AGT CCA TCC GGT TAT ATC GTT GAT GGT AAA ATT GCT ATC 5424			27
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			29
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			31
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 5563			33
Met His Tyr Lys Val Ala Thr Phe Met Ala Thr Leu Asp Leu Leu Met			
1845	1850	1855	35
GCC CGT GGT GAT GCT GCT TAC CGC CAG TTA GAG CGT GAT ACG TTG GCT 5615			37
Ala Arg Gly Asp Ala Ala Tyr Arg Gln Leu Glu Arg Asp Thr Leu Ala			
1860	1865	1870	39
GAA GCT AAA ATG TGG TAT ACA CAG GCG CTT AAT CTG TTG GGT GAT GAG 5664			41
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			43
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			45
Ala Ala Ser Lys Thr Thr Gln Gln Val Arg Gln Gln Val Leu Thr Gln			
1905	1910	1915	1920
			47

RO 121280 B1

1	TTG	CGT	CTC	AAT	AGC	AGG	GTA	AAA	ACC	CCG	TTG	CTA	GGA	ACA	GCC	AAT	5303
3	Leu	Arg	Leu	Asn	Ser	Arg	Val	Lys	Thr	Pro	Leu	Leu	Gly	Thr	Ala	Asn	
				1925						1930						1935	
5	TCC	CTG	ACC	GCT	TTA	TTT	CTG	CCG	CAG	CAA	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	
7	TAC	TGG	CCG	ACA	CTG	CGC	CAG	CGT	ATG	TTT	AAT	TTA	CGT	CAT	AAT	CTG	5904
9	Tyr	Trp	Arg	Thr	Leu	Ala	Gln	Arg	Met	Phe	Asn	Leu	Arg	His	Asn	Leu	
				1955					1960							1965	
11	TCG	ATT	GAC	GCC	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	
13	GAT	CCA	AAA	GCT	TTA	CTG	AGT	CGC	CCG	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							2000	
15	GCC	GAC	TTG	CCG	AAG	GCG	CCG	CTG	ACT	ATT	CAC	CGC	TTC	CCT	CAA	ATG	6048
17	Ala	Asp	Leu	Pro	Lys	Ala	Pro	Leu	Thr	Ile	His	Arg	Phe	Pro	Gln	Met	
					2005						2010					2015	
19	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	
21	TCA	CTA	TTG	GGG	TAC	AGT	GAG	COT	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	
23	CTA	CTG	CAA	ACC	CAA	GCC	AGC	GAG	TTA	ATA	CTG	ACC	AGT	ATT	CGT	ATG	6192
25	Leu	Leu	Gln	Thr	Gln	Ala	Ser	Glu	Leu	Ile	Leu	Thr	Ser	Ile	Arg	Met	
				2050					2055							2060	
27	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	
				2055				2070					2075			2080	
29	GTC	TCT	TTA	GCT	GGA	GTG	CAA	CAA	CGG	TTT	GAC	AGC	TAT	AGC	CAA	CTG	6288
31	Val	Ser	Leu	Ala	Gly	Val	Gln	Gln	Arg	Phe	Asp	Ser	Tyr	Ser	Gln	Leu	
					2085						2090					2095	
33	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	
35	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	
37	GGC	GCG	GGT	GTT	GAT	ATG	GCA	CCA	AAT	ATC	TTC	GCG	CTG	GCT	GAT	GCC	6432
39	Gly	Ala	Gly	Val	Asp	Met	Ala	Pro	Asn	Ile	Phe	Gly	Leu	Ala	Asp	Gly	
				2130					2135							2140	
41	GCC	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							2160	
43	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	
45	GAA	ATA	TAT	CGC	CGT	CGC	CGT	CAA	GAA	TGG	AAA	ATT	CAG	CGT	GAC	AAC	5576
47	Glu	Ile	Tyr	Arg	Arg	Arg	Arg	Gln	Glu	Trp	Lys	Ile	Gln	Arg	Asp	Asn	

RO 121280 B1

	2180	2185	2190	1
GCA CAA GCG GAG ATT AAC CAG TTA AAC GCG CAA CTG GAA TCA CTG TCT 6624				3
Ala Gln Ala Glu Ile Asn Gln Leu Asn Ala Gln Leu Glu Ser Leu Ser	2195	2200	2205	5
ATT CGC CGT GAA GCC GCT GAA ATG CAA AAA GAG TAC CTG AAA ACC CAG 6672				7
Ile Arg Arg Glu Ala Ala Glu Met Gln Lys Glu Tyr Leu Lys Thr Gln	2210	2215	2220	9
CAA GCT CAG GCG CAG GCA CAA CTT ACT TTC TTA AGA AGC AAA TTC AGT 6720				11
Gln Ala Gln Ala Gln Ala Gln Leu Thr Phe Leu Arg Ser Lys Phe Ser	2225	2230	2235	13
AAT CAA GCG TTA TAT AGT TCG TTA CGA GCG CGT TTG TCA GGT ATT TAT 6768				15
Asn Gln Ala Leu Tyr Ser Trp Leu Arg Gly Arg Leu Ser Gly Ile Tyr	2245	2250	2255	17
TTC CAG TTC TAT GAC TTG GCC GTA TCA CGT TCC CTG ATG GCA GAG CAA 6816				19
Phe Gln Phe Tyr Asp Leu Ala Val Ser Arg Cys Leu Met Ala Glu Gln	2260	2265	2270	21
TCC TAT CAA TGG GAA GCT AAT GAT AAT TCC ATT AGC TTT GTC AAA CCG 6864				23
Ser Tyr Gln Trp Glu Ala Asn Asp Asn Ser Ile Ser Phe Val Lys Pro	2275	2280	2285	25
GGT GCA TGG CAA GGA ACT TAC GCC GGC TTA TTG TGT GGA GAA GCT TTG 6912				27
Gly Ala Trp Gln Gly Thr Tyr Ala Gly Leu Leu Cys Gly Glu Ala Leu	2290	2295	2300	29
ATA CAA AAT CTG GCA CAA ATG GAA GAG GCA TAT CTG AAA TGG CAA TCT 6960				31
Ile Gln Asn Leu Ala Gln Met Glu Glu Ala Tyr Leu Lys Trp Glu Ser	2305	2310	2315	33
CGC GCT TTG GAA GTA GAA CGC ACG GTT TCA TTG GCA GTG GTT TAT GAT 7008				35
Arg Ala Leu Glu Val Glu Arg Thr Val Ser Leu Ala Val Val Tyr Asp	2325	2330	2335	37
TCA CTG GAA GGT AAT GAT CGT TTT AAT TTA GCG GAA CAA ATA CCT GCA 7056				39
Ser Leu Glu Gly Asn Asp Arg Phe Asn Leu Ala Glu Gln Ile Pro Ala	2340	2345	2350	41
TTA TTG GAT AAG GCG GAG GGA ACA GCA GCA ACT AAA GAA AAT GGG TTA 7104				43
Leu Leu Asp Lys Gly Glu Gly Thr Ala Gly Thr Lys Glu Asn Gly Leu	2355	2360	2365	45
TCA TTG GCT AAT GCT ATC CTG TCA GCT TCG GTC AAA TTG TCC GAC TTG 7152				47
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	2395	2390	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 ACC 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	

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

5 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

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

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

(2) INFORMAȚIA PENTRU SECV. ID. NR. 12:

(i) CARACTERISTICI DE SECVENȚĂ:

(A) LUNGIME: 2505 aminoacizi

(B) TIP: aminoacid

(D) TOPOLOGIE: lineară

(ii) TIPUL MOLECULEI: proteină

(xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 12:

21 Met Gln Asn Ser Leu Ser Ser Thr Ile Asp Thr Ile Cys Gln Lys Leu
 1 5 10 15

23 Gln Leu Thr Cys Pro Ala Glu Ile Ala Leu Tyr Pro Phe Asp Thr Phe
 20 25 30

25 Arg Glu Lys Thr Arg Gly Met Val Asn Trp Gly Glu Ala Lys Arg Ile
 35 40 45

27 Tyr Glu Ile Ala Gln Ala Glu Gln Asp Arg Asn Leu Leu His Glu Lys
 50 55 60

29 Arg Ile Phe Ala Tyr Ala Asn Pro Leu Leu Lys Asn Ala Val Arg Leu
 65 70 75 80

31 Gly Thr Arg Gln Met Leu Gly Phe Ile Gln Gly Tyr Ser Asp Leu Phe
 85 90 95

33 Gly Asn Arg Ala Asp Asn Tyr Ala Ala Pro Gly Ser Val Ala Ser Met
 100 105 110

35 Phe Ser Pro Ala Ala Tyr Leu Thr Glu Leu Tyr Arg Glu Ala Lys Asn
 115 120 125

37 Leu His Asp Ser Ser Ser Ile Tyr Tyr Leu Asp Lys Arg Arg Pro Asp
 130 135 140

39 Leu Ala Ser Leu Met Leu Ser Gln Lys Asn Met Asp Glu Glu Ile Ser
 145 150 155 160

43 Thr Leu Ala Leu Ser Asn Glu Leu Cys Leu Ala Gly Ile Glu Thr Lys
 165 170 175

45 Thr Gly Lys Ser Gln Asp Glu Val Met Asp Met Leu Ser Thr Tyr Arg
 180 185 190

RO 121280 B1

Leu Ser Gly Glu Thr Pro Tyr His His Ala Tyr Glu Thr Val Arg Glu	1
195 200 205	
Ile Val His Glu Arg Asp Pro Gly Phe Arg His Leu Ser Gln Ala Pro	3
210 215 220	
Ile Val Ala Ala Lys Leu Asp Pro Val Thr Leu Leu Gly Ile Ser Ser	5
225 230 235 240	
His Ile Ser Pro Glu Leu Tyr Asn Leu Leu Ile Glu Glu Ile Pro Glu	7
245 250 255	
Lys Asp Glu Ala Ala Leu Asp Thr Leu Tyr Lys Thr Asn Phe Gly Asp	9
260 265 270	
Ile Thr Thr Ala Gln Leu Met Ser Pro Ser Tyr Leu Ala Arg Tyr Tyr	11
275 280 285	
Gly Val Ser Pro Glu Asp Ile Ala Tyr Val Thr Thr Ser Leu Ser His	13
290 295 300	
Val Gly Tyr Ser Ser Asp Ile Leu Val Ile Pro Leu Val Asp Gly Val	15
305 310 315 320	
Gly Lys Met Glu Val Val Arg Val Thr Arg Thr Pro Ser Asp Asn Tyr	17
325 330 335	
Thr Ser Gln Thr Asn Tyr Ile Glu Leu Tyr Pro Gln Gly Gly Asp Asn	19
340 345 350	
Tyr Leu Ile Lys Tyr Asn Leu Ser Asn Ser Phe Gly Leu Asp Asp Phe	21
355 360 365	
Tyr Leu Gln Tyr Lys Asp Gly Ser Ala Asp Trp Thr Glu Ile Ala His	23
370 375 380	
Asn Pro Tyr Pro Asp Met Val Ile Asn Gln Lys Tyr Glu Ser Gln Ala	25
385 390 395 400	
Thr Ile Lys Arg Ser Asp Ser Asp Asn Ile Leu Ser Ile Gly Leu Gln	27
405 410 415	
Arg Trp His Ser Gly Ser Tyr Asn Phe Ala Ala Ala Asn Phe Lys Ile	29
420 425 430	
Asp Gln Tyr Ser Pro Lys Ala Phe Leu Leu Lys Met Asn Lys Ala Ile	31
435 440 445	
Arg Leu Leu Lys Ala Thr Gly Leu Ser Phe Ala Thr Leu Glu Arg Ile	33
450 455 460	
Val Asp Ser Val Asn Ser Thr Lys Ser Ile Thr Val Glu Val Leu Asn	35
465 470 475 480	
Lys Val Tyr Arg Val Lys Phe Tyr Ile Asp Arg Tyr Gly Ile Ser Glu	37
485 490 495	
Glu Thr Ala Ala Ile Leu Ala Asn Ile Asn Ile Ser Gln Gln Ala Val	39
500 505 510	
Gly Asn Gln Leu Ser Gln Phe Glu Gln Leu Phe Asn His Pro Pro Leu	41
515 520 525	
Asn Gly Ile Arg Tyr Glu Ile Ser Glu Asp Asn Ser Lys His Leu Pro	43
530 535 540	

RO 121280 B1

1	Asn Pro Asp Leu Asn Leu Lys Pro Asp Ser Thr Gly Asp Asp Gln Arg	545	550	555	560
3	Lys Ala Val Leu Lys Arg Ala Phe Gln Val Asn Ala Ser Glu Leu Tyr	565	570	575	
5	Gln Met Leu Leu Ile Thr Asp Arg Lys Glu Asp Gly Val Ile Lys Asn	580	585	590	
7	Asn Leu Glu Asn Leu Ser Asp Leu Tyr Leu Val Ser Leu Leu Ala Gln	595	600	605	
9	Ile His Asn Leu Thr Ile Ala Glu Leu Asn Ile Leu Leu Val Ile Cys	610	615	620	
11	Gly Tyr Gly Asp Thr Asn Ile Tyr Gln Ile Thr Asp Asp Asn Leu Ala	625	630	635	640
13	Lys Ile Val Glu Thr Leu Leu Trp Ile Thr Gln Trp Leu Lys Thr Gln	645	650	655	
15	Lys Trp Thr Val Thr Asp Leu Phe Leu Met Thr Thr Ala Thr Tyr Ser	660	665	670	
17	Thr Thr Leu Thr Pro Glu Ile Ser Asn Leu Thr Ala Thr Leu Ser Ser	675	680	685	
19	Thr Leu His Gly Lys Glu Ser Leu Ile Gly Glu Asp Leu Lys Arg Ala	690	695	700	
21	Met Ala Pro Cys Phe Thr Ser Ala Leu His Leu Thr Ser Gln Glu Val	705	710	715	720
23	Ala Tyr Asp Leu Leu Leu Trp Ile Asp Gln Ile Gln Pro Ala Gln Ile	725	730	735	
25	Thr Val Asp Gly Phe Trp Glu Glu Val Gln Thr Thr Pro Thr Ser Leu	740	745	750	
27	Lys Val Ile Thr Phe Ala Gln Val Leu Ala Gln Leu Ser Leu Ile Tyr	755	760	765	
29	Arg Arg Ile Gly Leu Ser Glu Thr Glu Leu Ser Leu Ile Val Thr Gln	770	775	780	
31	Ser Ser Leu Leu Val Ala Gly Lys Ser Ile Leu Asp His Gly Leu Leu	785	790	795	800
33	Thr Leu Met Ala Leu Glu Gly Phe His Thr Trp Val Asn Gly Leu Gly	805	810	815	
35	Gln His Ala Ser Leu Ile Leu Ala Ala Leu Lys Asp Gly Ala Leu Thr	820	825	830	
37	Val Thr Asp Val Ala Gln Ala Met Asn Lys Glu Glu Ser Leu Leu Gln	835	840	845	
39	Met Ala Ala Asn Gln Val Glu Lys Asp Leu Thr Lys Leu Thr Ser Trp	850	855	860	
41	Thr Gln Ile Asp Ala Ile Leu Gln Trp Leu Gln Met Ser Ser Ala Leu	865	870	875	880
43	Ala Val Ser Pro Leu Asp Leu Ala Gly Met Met Ala Leu Lys Tyr Gly	885	890	895	

RO 121280 B1

Ile Asp His Asn Tyr Ala Ala Trp Gln Ala Ala Ala Ala Ala Leu Met	1
900 905 910	
Ala Asp His Ala Asn Gln Ala Gln Lys Lys Leu Asp Glu Thr Phe Ser	3
915 920 925	
Lys Ala Leu Cys Asn Tyr Tyr Ile Asn Ala Val Val Asp Ser Ala Ala	5
930 935 940	
Gly Val Arg Asp Arg Asn Gly Leu Tyr Thr Tyr Leu Leu Ile Asp Asn	7
945 950 955 960	
Gln Val Ser Ala Asp Val Ile Thr Ser Arg Ile Ala Glu Ala Ile Ala	9
965 970 975	
Gly Ile Gln Leu Tyr Val Asn Arg Ala Leu Asn Arg Asp Glu Gly Gln	11
980 985 990	
Leu Ala Ser Asp Val Ser Thr Arg Gln Phe Phe Thr Asp Trp Glu Arg	13
995 1000 1005	
Tyr Asn Lys Arg Tyr Ser Thr Trp Ala Gly Val Ser Glu Leu Val Tyr	15
1010 1015 1020	
Tyr Pro Glu Asn Tyr Val Asp Pro Thr Gln Arg Ile Gly Gln Thr Lys	17
1025 1030 1035 1040	
Met Met Asp Ala Leu Leu Gln Ser Ile Asn Gln Ser Gln Leu Asn Ala	19
1045 1050 1055	
Asp Thr Val Glu Asp Ala Phe Lys Thr Tyr Leu Thr Ser Phe Glu Gln	21
1060 1065 1070	
Val Ala Asn Leu Lys Val Ile Ser Ala Tyr His Asp Asn Val Asn Val	23
1075 1080 1085	
Asp Gln Gly Leu Thr Tyr Phe Ile Gly Ile Asp Gln Ala Ala Pro Gly	25
1090 1095 1100	
Thr Tyr Tyr Trp Arg Ser Val Asp His Ser Lys Cys Glu Asn Gly Lys	27
1105 1110 1115 1120	
Phe Ala Ala Asn Ala Trp Gly Glu Trp Asn Lys Ile Thr Cys Ala Val	29
1125 1130 1135	
Asn Pro Trp Lys Asn Ile Ile Arg Pro Val Val Tyr Met Ser Arg Leu	31
1140 1145 1150	
Tyr Leu Leu Trp Leu Glu Gln Gln Ser Lys Lys Ser Asp Asp Gly Lys	33
1155 1160 1165	
Thr Thr Ile Tyr Gln Tyr Asn Leu Lys Leu Ala His Ile Arg Tyr Asp	35
1170 1175 1180	
Gly Ser Trp Asn Thr Pro Phe Thr Phe Asp Val Thr Glu Lys Val Lys	37
1185 1190 1195 1200	
Asn Tyr Thr Ser Ser Thr Asp Ala Ala Glu Ser Leu Gly Leu Tyr Cys	39
1205 1210 1215	
Thr Gly Tyr Gln Gly Glu Asp Thr Leu Leu Val Met Phe Tyr Ser Met	41
1220 1225 1230	
Gln Ser Ser Tyr Ser Ser Tyr Thr Asp Asn Asn Ala Pro Val Thr Gly	43
1235 1240 1245	

RO 121280 B1

1 Leu Tyr Ile Phe Ala Asp Met Ser Ser Asp Asn Met Thr Asn Ala Gln
 1250 1255 1260
 3 Ala Thr Asn Tyr Trp Asn Asn Ser Tyr Pro Gln Phe Asp Thr Val Met
 1265 1270 1275 1280
 5 Ala Asp Pro Asp Ser Asp Asn Lys Lys Val Ile Thr Arg Arg Val Asn
 1285 1290 1295
 7 Asn Arg Tyr Ala Glu Asp Tyr Glu Ile Pro Ser Ser Val Thr Ser Asn
 1300 1305 1310
 9 Ser Asn Tyr Ser Trp Gly Asp His Ser Leu Thr Met Leu Tyr Gly Gly
 1315 1320 1325
 11 Ser Val Pro Asn Ile Thr Phe Glu Ser Ala Ala Glu Asp Leu Arg Leu
 1330 1335 1340
 13 Ser Thr Asn Met Ala Leu Ser Ile Ile His Asn Gly Tyr Ala Gly Thr
 1345 1350 1355 1360
 15 Arg Arg Ile Gln Cys Asn Leu Met Lys Gln Tyr Ala Ser Leu Gly Asp
 1365 1370 1375
 17 Lys Phe Ile Ile Tyr Asp Ser Ser Phe Asp Asp Ala Asn Arg Phe Asn
 1380 1385 1390
 19 Leu Val Pro Leu Phe Lys Phe Gly Lys Asp Glu Asn Ser Asp Asp Ser
 1395 1400 1405
 21 Ile Cys Ile Tyr Asn Glu Asn Pro Ser Ser Glu Asp Lys Lys Trp Tyr
 1410 1415 1420
 23 Phe Ser Ser Lys Asp Asp Asn Lys Thr Ala Asp Tyr Asn Gly Gly Thr
 1425 1430 1435 1440
 25 Gln Cys Ile Asp Ala Gly Thr Ser Asn Lys Asp Phe Tyr Tyr Asn Leu
 1445 1450 1455
 27 Gln Glu Ile Glu Val Ile Ser Val Thr Gly Gly Tyr Trp Ser Ser Tyr
 1460 1465 1470
 29 Lys Ile Ser Asn Pro Ile Asn Ile Asn Thr Gly Ile Asp Ser Ala Lys
 1475 1480 1485
 31 Val Lys Val Thr Val Lys Ala Gly Gly Asp Asp Gln Ile Phe Thr Ala
 1490 1495 1500
 33 Asp Asn Ser Thr Tyr Val Pro Gln Gln Pro Ala Pro Ser Phe Glu Glu
 1505 1510 1515 1520
 35 Met Ile Tyr Gln Phe Asn Asn Leu Thr Ile Asp Cys Lys Asn Leu Asn
 1525 1530 1535
 37 Phe Ile Asp Asn Gln Ala His Ile Glu Ile Asp Phe Thr Ala Thr Ala
 1540 1545 1550
 39 Gln Asp Gly Arg Phe Leu Gly Ala Glu Thr Phe Ile Ile Pro Val Thr
 1555 1560 1565
 41 Lys Lys Val Leu Gly Thr Glu Asn Val Ile Ala Leu Tyr Ser Glu Asn
 1570 1575 1580
 43 Asn Gly Val Gln Tyr Met Gln Ile Gly Ala Tyr Arg Thr Arg Leu Asn
 1585 1590 1595 1600

RO 121280 B1

Thr Leu Phe Ala Gln Gln Leu Val Ser Arg Ala Asn Arg Gly Ile Asp 1605 1610 1615	1
Ala Val Leu Ser Met Glu Thr Gln Asn Ile Gln Glu Pro Gln Leu Gly 1620 1625 1630	3
Ala Gly Thr Tyr Val Gln Leu Val Leu Asp Lys Tyr Asp Glu Ser Ile 1635 1640 1645	5
His Gly Thr Asn Lys Ser Phe Ala Ile Glu Tyr Val Asp Ile Phe Lys 1650 1655 1660	7
Glu Asn Asp Ser Phe Val Ile Tyr Gln Gly Glu Leu Ser Glu Thr Ser 1665 1670 1675 1680	9
Gln Thr Val Val Lys Val Phe Leu Ser Tyr Phe Ile Glu Ala Thr Gly 1685 1690 1695	11
Asn Lys Asn His Leu Trp Val Arg Ala Lys Tyr Gln Lys Glu Thr Thr 1700 1705 1710	13
Asp Lys Ile Leu Phe Asp Arg Thr Asp Glu Lys Asp Pro His Gly Trp 1715 1720 1725	15
Phe Leu Ser Asp Asp His Lys Thr Phe Ser Gly Leu Ser Ser Ala Gln 1730 1735 1740	17
Ala Leu Lys Asn Asp Ser Glu Pro Met Asp Phe Ser Gly Ala Asn Ala 1745 1750 1755 1760	19
Leu Tyr Phe Trp Glu Leu Phe Tyr Tyr Thr Pro Met Met Met Ala His 1765 1770 1775	21
Arg Leu Leu Gln Glu Gln Asn Phe Asp Ala Ala Asn His Trp Phe Arg 1780 1785 1790	23
Tyr Val Trp Ser Pro Ser Gly Tyr Ile Val Asp Gly Lys Ile Ala Ile 1795 1800 1805	25
Tyr His Trp Asn Val Arg Pro Leu Glu Glu Asp Thr Ser Trp Asn Ala 1810 1815 1820	27
Gln Gln Leu Asp Ser Thr Asp Pro Asp Ala Val Ala Gln Asp Asp Pro 1825 1830 1835 1840	29
Met His Tyr Lys Val Ala Thr Phe Met Ala Thr Leu Asp Leu Leu Met 1845 1850 1855	31
Ala Arg Gly Asp Ala Ala Tyr Arg Gln Leu Glu Arg Asp Thr Leu Ala 1860 1865 1870	33
Glu Ala Lys Met Trp Tyr Thr Gln Ala Leu Asn Leu Leu Gly Asp Glu 1875 1880 1885	35
Pro Gln Val Met Leu Ser Thr Thr Trp Ala Asn Pro Thr Leu Gly Asn 1890 1895 1900	37
Ala Ala Ser Lys Thr Thr Gln Gln Val Arg Gln Gln Val Leu Thr Gln 1905 1910 1915 1920	39
Leu Arg Leu Asn Ser Arg Val Lys Thr Pro Leu Leu Gly Thr Ala Asn 1925 1930 1935	41
Ser Leu Thr Ala Leu Phe Leu Pro Gln Glu Asn Ser Lys Leu Lys Gly 1940 1945 1950	43
	45
	47

RO 121280 B1

1 Tyr Trp Arg Thr Leu Ala Gln Arg Met Phe Asn Leu Arg His Asn Leu
1955 1960 1965

3 Ser Ile Asp Gly Gln Pro Leu Ser Leu Pro Leu Tyr Ala Lys Pro Ala
1970 1975 1980

5 Asp Pro Lys Ala Leu Leu Ser Ala Ala Val Ser Ala Ser Gln Gly Gly
1985 1990 1995 2000

7 Ala Asp Leu Pro Lys Ala Pro Leu Thr Ile His Arg Phe Pro Gln Met
2005 2010 2015

9 Leu Glu Gly Ala Arg Gly Leu Val Asn Gln Leu Ile Gln Phe Gly Ser
2020 2025 2030

11 Ser Leu Leu Gly Tyr Ser Glu Arg Gln Asp Ala Glu Ala Met Ser Gln
2035 2040 2045

13 Leu Leu Gln Thr Gln Ala Ser Glu Leu Ile Leu Thr Ser Ile Arg Met
2050 2055 2060

15 Gln Asp Asn Gln Leu Ala Glu Leu Asp Ser Glu Lys Thr Ala Leu Gln
2065 2070 2075 2080

17 Val Ser Leu Ala Gly Val Gln Gln Arg Phe Asp Ser Tyr Ser Gln Leu
2085 2090 2095

19 Tyr Glu Glu Asn Ile Asn Ala Gly Glu Gln Arg Ala Leu Ala Leu Arg
2100 2105 2110

21 Ser Glu Ser Ala Ile Glu Ser Gln Gly Ala Gln Ile Ser Arg Met Ala
2115 2120 2125

23 Gly Ala Gly Val Asp Met Ala Pro Asn Ile Phe Gly Leu Ala Asp Gly
2130 2135 2140

25 Gly Met His Tyr Gly Ala Ile Ala Tyr Ala Ile Ala Asp Gly Ile Glu
2145 2150 2155 2160

27 Leu Ser Ala Ser Ala Lys Met Val Asp Ala Glu Lys Val Ala Gln Ser
2165 2170 2175

29 Glu Ile Tyr Arg Arg Arg Arg Gln Glu Trp Lys Ile Gln Arg Asp Asn
2180 2185 2190

31 Ala Gln Ala Glu Ile Asn Gln Leu Asn Ala Gln Leu Glu Ser Leu Ser
2195 2200 2205

33 Ile Arg Arg Glu Ala Ala Glu Met Gln Lys Glu Tyr Leu Lys Thr Gln
2210 2215 2220

35 Gln Ala Gln Ala Gln Ala Gln Leu Thr Phe Leu Arg Ser Lys Phe Ser
2225 2230 2235 2240

37 Asn Gln Ala Leu Tyr Ser Trp Leu Arg Gly Arg Leu Ser Gly Ile Tyr
2245 2250 2255

39 Phe Gln Phe Tyr Asp Leu Ala Val Ser Arg Cys Leu Met Ala Glu Gln
2260 2265 2270

41 Ser Tyr Gln Trp Glu Ala Asn Asp Asn Ser Ile Ser Phe Val Lys Pro
2275 2280 2285

43 Gly Ala Trp Gln Gly Thr Tyr Ala Gly Leu Leu Cys Gly Glu Ala Leu
2290 2295 2300

RO 121280 B1

Ile Gln Asn Leu Ala Gln Met Glu Glu Ala Tyr Leu Lys Trp Glu Ser 2305 2310 2315 2320	1
Arg Ala Leu Glu Val Glu Arg Thr Val Ser Leu Ala Val Val Tyr Asp 2325 2330 2335	3
Ser Leu Glu Gly Asn Asp Arg Phe Asn Leu Ala Glu Gln Ile Pro Ala 2340 2345 2350	5
Leu Leu Asp Lys Gly Glu Gly Thr Ala Gly Thr Lys Glu Asn Gly Leu 2355 2360 2365	7
Ser Leu Ala Asn Ala Ile Leu Ser Ala Ser Val Lys Leu Ser Asp Leu 2370 2375 2380	9
Lys Leu Gly Thr Asp Tyr Pro Asp Ser Ile Val Gly Ser Asn Lys Val 2385 2390 2395 2400	11
Arg Arg Ile Lys Gln Ile Ser Val Ser Leu Pro Ala Leu Val Gly Pro 2405 2410 2415	13
Tyr Gln Asp Val Gln Ala Met Leu Ser Tyr Gly Gly Ser Thr Gln Leu 2420 2425 2430	15
Pro Lys Gly Cys Ser Ala Leu Ala Val Ser His Gly Thr Asn Asp Ser 2435 2440 2445	17
Gly Gln Phe Gln Leu Asp Phe Asn Asp Gly Lys Tyr Leu Pro Phe Glu 2450 2455 2460	19
Gly Ile Ala Leu Asp Asp Gln Gly Thr Leu Asn Leu Gln Phe Pro Asn 2465 2470 2475 2480	21
Ala Thr Asp Lys Gln Lys Ala Ile Leu Gln Thr Met Ser Asp Ile Ile 2485 2490 2495	23
Leu His Ile Arg Tyr Thr Ile Arg * 2500 2505	25
(2) INFORMAȚIA PENTRU SECV. ID. NR. 13:	27
(i) CARACTERISTICI DE SECVENȚĂ:	29
(A) LUNGIME: 12 aminoacizi	31
(B) TIP: aminoacid	33
(C) ÎMPLETIRE: simplă	35
(D) TOPOLOGIE: lineară	37
(ii) TIPUL MOLECULEI: peptidă	39
(xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 13:	41
Leu Ile Gly Tyr Asn Asn Gln Phe Ser Gly Xaa Ala 1 5 10	43
(2) INFORMAȚIA PENTRU SECV. ID. NR. :14	45
(i) CARACTERISTICI DE SECVENȚĂ:	47
(A) LUNGIME: 12 aminoacizi	49
(B) TIP: aminoacid	
(C) ÎMPLETIRE: simplă	
(D) TOPOLOGIE: lineară	
(ii) TIPUL MOLECULEI: peptidă	
(xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 14:	

1

Met Gln Asn Ser Gln Thr Phe Ser Val Gly Glu Leu
1 5 10

3

(2) INFORMAȚIA PENTRU SECV. ID. NR. 15:

5

(i) CARACTERISTICI DE SECVENȚĂ:

7

(A) LUNGIME: 9 aminoacizi

(B) TIP: aminoacid

(C) ÎMPLETIRE: simplă

9

(D) TOPOLOGIE: lineară

(ii) TIPUL MOLECULEI: peptidă

11

(xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 15

13

Ala Gln Asp Gly Asn Gln Asp Thr Phe Phe Ser Gly Asn Thr
1 5 10

15

(2) INFORMAȚIA PENTRU SECV. ID. NR. 16:

17

(i) CARACTERISTICI DE SECVENȚĂ:

(A) LUNGIME: 5 aminoacizi

(B) TIP: aminoacid

19

(C) ÎMPLETIRE: simplă

(D) TOPOLOGIE: lineară

21

(ii) TIPUL MOLECULEI: peptidă

(xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 16

23

Met Gln Asn Ser Leu
1 5

25

(2) INFORMAȚIA PENTRU SECV. ID. NR. 17:

27

(i) CARACTERISTICI DE SECVENȚĂ:

(A) LUNGIME: 10 aminoacizi

(B) TIP: aminoacid

29

(C) ÎMPLETIRE: simplă

31

(D) TOPOLOGIE: lineară

(ii) TIPUL MOLECULEI: peptidă

33

(xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. :17:

35

Ala Phe Asn Ile Asp Asp Val Ser Leu Phe
1 5 10

37

SECV. ID-. NR. 18:

39

(i) CARACTERISTICI DE SECVENȚĂ:

(A) LUNGIME: 16 aminoacizi

(B) TIP: aminoacid

41

(C) ÎMPLETIRE: simplă

(D) TOPOLOGIE: lineară

43

(ii) TIPUL MOLECULEI: peptidă

(xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 18:

45

Phe Ile Val Tyr Thr Ser Leu Gly Val Asn Pro Asn Asn Ser Ser Asn
1 5 10 15

47

RO 121280 B1

(2) INFORMAȚIA PENTRU SECV. ID. NR. 19:	1
(i) CARACTERISTICI DE SECVENȚĂ:	
(A) LUNGIME: 21 aminoacizi	3
(B) TIP: aminoacid	
(C) ÎMPLETIRE: simplă	5
(D) TOPOLOGIE: lineară.	
(ii) TIPUL MOLECULEI: peptidă	7
(xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 19:	9
Ile Ser Asp Leu Val Thr Thr Ser Pro Leu Ser Glu Ala Ile Gly Ser 1 5 10 15 Leu Gln Leu Phe Ile 20	11
	13
(2) INFORMAȚIA PENTRU SECV. ID. NR. 20:	
(i) CARACTERISTICI DE SECVENȚĂ:	15
(A) LUNGIME: 12 aminoacizi	
(B) TIP: aminoacid	17
(C) ÎMPLETIRE: simplă	
(D) TOPOLOGIE: lineară	19
(ii) TIPUL MOLECULEI: peptidă	
(xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 20:	21
Met Tyr Tyr Ile Gln Ala Gln Gln Leu Leu Gly Pro 1 5 10	23
(2) INFORMAȚIA PENTRU SECV. ID. NR. 21:	25
(i) CARACTERISTICI DE SECVENȚĂ:	
(A) LUNGIME: 26 aminoacizi	27
(B) TIP: aminoacid	
(C) ÎMPLETIRE: simplă	29
(D) TOPOLOGIE: lineară	
(ii) TIPUL MOLECULEI: peptidă	31
(xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 21:	33
Gly Ile Asp Ala Val Leu Ser Met Glu Thr Gln Asn Ile Gln Glu Pro 1 5 10 15 Gln Leu Gly Ala Gly Thr Tyr Val Gln Leu 20 25	35
	37
(2) INFORMAȚIA PENTRU SECV. ID. NR. 22:	
(i) CARACTERISTICI DE SECVENȚĂ:	39
(A) LUNGIME: 15 aminoacizi	
(B) TIP: aminoacid	41
(C) ÎMPLETIRE: simplă	
(D) TOPOLOGIE: lineară	43

RO 121280 B1

- 1 (ii) TIPUL MOLECULEI: peptidă
 (xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 22:
- 3
- 5 **Ile Ser Asn Pro Ile Asn Ile Asn Thr Gly Ile Asp Ser Ala Lys**
 1 5 10 15
- 7 (2) INFORMAȚIA PENTRU SECV. ID. NR. 23:
 (i) CARACTERISTICI DE SECVENȚĂ:
 (A) LUNGIME: 13 aminoacizi
 (B) TIP: aminoacid
 (C) ÎMPLETIRE: simplă
 (D) TOPOLOGIE: lineară
 (ii) TIPUL MOLECULEI: peptidă
 (xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 23:
- 13
- 15 **Thr Tyr Leu Thr Ser Phe Glu Gln Val Ala Asn Leu Lys**
 1 5 10
- 17 (2) INFORMAȚIA PENTRU SECV. ID. NR. 24:
 (i) CARACTERISTICI DE SECVENȚĂ:
 (A) LUNGIME: 22 aminoacizi
 (B) TIP: aminoacid
 (C) ÎMPLETIRE: simplă
 (D) TOPOLOGIE: lineară
 (ii) TIPUL MOLECULEI: peptidă
 (xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 24:
- 19
- 21
- 23
- 25 **Val Leu Gly Thr Glu Asn Val Ile Ala Leu Tyr Ser Glu Asn Asn Gly**
 1 5 10 15
- 27
- 29 **Val Gln Tyr Met Gln Ile**
 20
- 31 (2) INFORMAȚIA PENTRU SECV. ID. NR. 25:
 (i) CARACTERISTICI DE SECVENȚĂ:
 (A) LUNGIME: 6005 baze perechi
 (B) TIP: acid nucleic
 (C) ÎMPLETIRE: dublă
 (D) TOPOLOGIE: lineară
 (ii) TIPUL MOLECULEI: ADN (genomic)
 (ix) TRĂSĂTURA:
 (A) NUME/COD: RBS
 (B) LOCALIZARE: 1..9
 (ix) TRĂSĂTURA:
 (A) NUME/COD: CDS
 (B) LOCALIZARE: 16...3585
 (D) ALTĂ INFORMAȚIE: /produs= "p8"
 (xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 25:
- 33
- 35
- 37
- 39
- 41
- 43
- 45

RO 121280 B1

AAGAAGGAAT	CGATT	ATG	TCT	GAA	TCT	TTA	TTT	ACA	CAA	ACG	TTG	AAA	GAA	51	3		
		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	99	5
Ala	Arg	Arg	Asp	Ala	Leu	Val	Ala	His	Tyr	Ile	Ala	Thr	Gln	Val	Pro		7
		15				20					25						
GCA	GAT	TTA	AAA	GAG	AGT	ATC	CAG	ACC	GCG	GAT	GAT	CTG	TAC	GAA	TAT	147	9
Ala	Asp	Leu	Lys	Glu	Ser	Ile	Gln	Thr	Ala	Asp	Asp	Leu	Tyr	Glu	Tyr		11
		30				35					40						
CTG	TTG	CTG	GAT	ACC	AAA	ATT	AGC	GAT	CTG	GTT	ACT	ACT	TCA	CCG	CTG	195	13
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	243	15
Ser	Glu	Ala	Ile	Gly	Ser	Leu	Gln	Leu	Phe	Ile	His	Arg	Ala	Ile	Glu		17
				65					70					75			
GGC	TAT	GAC	GGC	ACG	CTG	GCA	GAC	TCA	GCA	AAA	CCC	TAT	TTT	GCC	GAT	291	19
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	339	21
Glu	Gln	Phe	Leu	Tyr	Asn	Trp	Asp	Ser	Phe	Asn	His	Arg	Tyr	Ser	Thr		23
		95				100						105					
TGG	GCT	GGC	AAG	GAA	CGG	TTG	AAA	TTC	TAT	GCC	GGG	GAT	TAT	ATT	GAT	387	25
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	435	27
Pro	Thr	Leu	Arg	Leu	Asn	Lys	Thr	Glu	Ile	Phe	Thr	Ala	Phe	Glu	Gln		29
		125				130				135					140		

RO 121280 B1

1	GGT	ATT	TCT	CAA	GGG	AAA	TTA	AAA	AGT	GAA	TTA	GTC	GAA	TCT	AAA	TTA	433
3	Gly	Ile	Ser	Gln	Gly	Lys	Leu	Lys	Ser	Glu	Leu	Val	Glu	Ser	Lys	Leu	
					145					150					155		
5	CGT	GAT	TAT	CTA	ATT	AGT	TAT	GAC	ACT	TTA	GCC	ACC	CTT	GAT	TAT	ATT	531
	Arg	Asp	Tyr	Leu	Ile	Ser	Tyr	Asp	Thr	Leu	Ala	Thr	Leu	Asp	Tyr	Ile	
					160				165						170		
7	ACT	GCC	TGC	CAA	GGC	AAA	GAT	AAT	AAA	ACC	ATC	TTC	TTT	ATT	GGC	CGT	579
9	Thr	Ala	Cys	Gln	Gly	Lys	Asp	Asn	Lys	Thr	Ile	Phe	Phe	Ile	Gly	Arg	
					175			180						185			
11	ACA	CAG	AAT	GCA	CCC	TAT	GCA	TTT	TAT	TGG	CGA	AAA	TTA	ACT	TTA	GTC	627
	Thr	Gln	Asn	Ala	Pro	Tyr	Ala	Phe	Tyr	Trp	Arg	Lys	Leu	Thr	Leu	Val	
					190			195				200					
13	ACT	GAT	GGC	GGT	AAG	TTG	AAA	CCA	GAT	CAA	TGG	TCA	GAG	TGG	CGA	GCA	675
	Thr	Asp	Gly	Gly	Lys	Leu	Lys	Pro	Asp	Gln	Trp	Ser	Glu	Trp	Arg	Ala	
						210				215					220		
15	ATT	AAT	GCC	GGG	ATT	AGT	GAG	GCA	TAT	TCA	GGG	CAT	GTC	GAG	CCT	TTC	723
17	Ile	Asn	Ala	Gly	Ile	Ser	Glu	Ala	Tyr	Ser	Gly	His	Val	Glu	Pro	Phe	
					225					230					235		
19	TGG	GAA	AAT	AAC	AAG	CTG	CAC	ATC	CGT	TGG	TTT	ACT	ATC	TCG	AAA	GAA	771
	Trp	Glu	Asn	Asn	Lys	Leu	His	Ile	Arg	Trp	Phe	Thr	Ile	Ser	Lys	Glu	
					240				245					250			
21	GAT	AAA	ATA	GAT	TTT	GTT	TAT	AAA	AAC	ATC	TGG	GTG	ATG	AGT	AGC	GAT	819
23	Asp	Lys	Ile	Asp	Phe	Val	Tyr	Lys	Asn	Ile	Trp	Val	Met	Ser	Ser	Asp	
					255			260					265				
25	TAT	ACC	TGG	GCA	TCA	AAG	AAA	AAA	ATC	TTG	GAA	CTT	TCT	TTT	ACT	GAC	867
	Tyr	Ser	Trp	Ala	Ser	Lys	Lys	Lys	Ile	Leu	Glu	Leu	Ser	Phe	Thr	Asp	
					270			275				280					
27	TAC	AAT	AGA	GTT	GGA	GCA	ACA	GGA	TCA	TCA	AGC	CCG	ACT	GAA	GTA	GCT	915
	Tyr	Asn	Arg	Val	Gly	Ala	Thr	Gly	Ser	Ser	Ser	Pro	Thr	Glu	Val	Ala	
					285			290			295					300	
29	TCA	CAA	TAT	GGT	TCT	GAT	GCT	CAG	ATG	AAT	ATT	TCT	GAT	GAT	GGG	ACT	963
31	Ser	Gln	Tyr	Gly	Ser	Asp	Ala	Gln	Met	Asn	Ile	Ser	Asp	Asp	Gly	Thr	
					305				310						315		
33	GTA	CTT	ATT	TTT	CAG	AAT	GCC	GGC	GGA	GCT	ACT	CCC	AGT	ACT	GGA	GTG	1011
	Val	Leu	Ile	Phe	Gln	Asn	Ala	Gly	Gly	Ala	Thr	Pro	Ser	Thr	Gly	Val	
					320				325					330			
35	ACG	TTA	TGT	TAT	GAC	TCT	GGC	AAC	GTG	ATT	AAG	AAC	CTA	TCT	AGT	ACA	1059
37	Thr	Leu	Cys	Tyr	Asp	Ser	Gly	Asn	Val	Ile	Lys	Asn	Leu	Ser	Ser	Thr	
					335			340				345					
39	GGA	AGT	GCA	AAT	TTA	TCG	TCA	AAG	GAT	TAT	GCC	ACA	ACT	AAA	TTA	CGC	1107
	Gly	Ser	Ala	Asn	Leu	Ser	Ser	Lys	Asp	Tyr	Ala	Thr	Thr	Lys	Leu	Arg	
					350			355				360					
41	ATG	TGT	CAT	GGA	CAA	AGT	TAC	AAT	GAT	AAT	AAC	TAC	TGC	AAT	TTT	ACA	1155
	Met	Cys	His	Gly	Gln	Ser	Tyr	Asn	Asp	Asn	Asn	Tyr	Cys	Asn	Phe	Thr	
					365			370			375					380	
43	CTC	TCT	ATT	AAT	ACA	ATA	GAA	TTC	ACC	TCC	TAC	GGC	ACA	TTC	TCA	TCA	1203
45	Leu	Ser	Ile	Asn	Thr	Ile	Glu	Phe	Thr	Ser	Tyr	Gly	Thr	Phe	Ser	Ser	
					385					390					395		
47	GAT	GGA	AAA	CAA	TTT	ACA	CCA	CCT	TCT	GGT	TCT	GCC	ATT	GAT	TTA	CAC	1251
	Asp	Gly	Lys	Gln	Phe	Thr	Pro	Pro	Ser	Gly	Ser	Ala	Ile	Asp	Leu	His	

RO 121280 B1

	400	405	410		1
CTC CCT AAT TAT GTA GAT CTC AAC GCG CTA TTA GAT ATT AGC CTC GAT 1293					3
Leu Pro Asn Tyr Val Asp Leu Asn Ala Leu Leu Asp Ile Ser Leu Asp	415	420	425		
					5
TCA CTA CTT AAT TAT GAC GTT CAG GGG CAG TTT GGC GGA TCT AAT CCG 1347					7
Ser Leu Leu Asn Tyr Asp Val Gln Gly Gln Phe Gly Gly Ser Asn Pro	430	435	440		
					9
GTT GAT AAT TTC AGT GGT CCC TAT GGT ATT TAT CTA TGG GAA ATC TTC 1395					11
Val Asp Asn Phe Ser Gly Pro Tyr Gly Ile Tyr Leu Trp Glu Ile Phe	445	450	455	460	
					13
TTC CAT ATT CCG TTC CTT GTT ACG GTC CGT ATG CAA ACC GAA CAA CGT 1443					15
Phe His Ile Pro Phe Leu Val Thr Val Arg Met Gln Thr Glu Gln Arg	465	470	475		
					17
TAC GAA GAC GCG GAC ACT TGG TAC AAA TAT ATT TTC CGC AGC GCC GGT 1491					19
Tyr Glu Asp Ala Asp Thr Trp Tyr Lys Tyr Ile Phe Arg Ser Ala Gly	480	485	490		
					21
TAT CGC GAT GCT AAT GGC CAG CTC ATT ATG GAT GGC AGT AAA CCA CGT 1539					23
Tyr Arg Asp Ala Asn Gly Gln Leu Ile Met Asp Gly Ser Lys Pro Arg	495	500	505		
					25
TAT TGG AAT CTG ATG CCA TTG CAA CTC GAT ACC GCA TGG GAT ACC ACA 1587					27
Tyr Trp Asn Val Met Pro Leu Gln Leu Asp Thr Ala Trp Asp Thr Thr	510	515	520		
					29
CAG CCC GCC ACC ACT GAT CCA GAT GTG ATC GCT ATG GCG GAC CCG ATG 1635					31
Gln Pro Ala Thr Thr Asp Pro Asp Val Ile Ala Met Ala Asp Pro Met	525	530	535	540	
					33
CAT TAC AAG CTG GCG ATA TTC CTG CAT ACC CTT GAT CTA TTG ATT GCC 1683					35
His Tyr Lys Leu Ala Ile Phe Leu His Thr Leu Asp Leu Leu Ile Ala	545	550	555		
					37
CGA GGC GAC AGC GCT TAC CGT CAA CTT GAA GCG GAT ACT CTA GTC GAA 1731					39
Arg Gly Asp Ser Ala Tyr Arg Gln Leu Glu Arg Asp Thr Leu Val Glu	560	565	570		
					41
GCC AAA ATG TAC TAC ATT CAG GCA CAA CAG CTA CTG GGA CCG CGC CCT 1779					43
Ala Lys Met Tyr Tyr Ile Gln Ala Gln Gln Leu Leu Gly Pro Arg Pro	575	580	585		
					45
GAT ATC CAT ACC ACC AAT ACT TGG CCA AAT CCC ACC TTG AGT AAA GAA 1827					47
Asp Ile His Thr Thr Asn Thr Trp Pro Asn Pro Thr Leu Ser Lys Glu	590	595	600		
					39
GCT GGC GCT ATT GCC ACA CCG ACA TTC CTC AGT TCA CCG GAG GTG ATG 1875					41
Ala Gly Ala Ile Ala Thr Pro Thr Phe Leu Ser Ser Pro Glu Val Met	605	610	615	620	
					43
ACG TTC GCT GCC TGG CTA AGC GCA GGC GAT ACC GCA AAT ATT GGC GAC 1923					45
Thr Phe Ala Ala Trp Leu Ser Ala Gly Asp Thr Ala Asn Ile Gly Asp	625	630	635		
					47
GGT GAT TTC TTG CCA CCG TAC AAC GAT GTA CTA CTC GGT TAC TGG GAT 1971					49
Gly Asp Phe Leu Pro Pro Tyr Asn Asp Val Leu Leu Gly Tyr Trp Asp	640	645	650		
					51
AAA CTT GAG TTA CGC CTA TAC AAC CTG CGC CAC AAT CTG AGT CTG GAT 2019					53
Lys Leu Glu Leu Arg Leu Tyr Asn Leu Arg His Asn Leu Ser Leu Asp	655	660	665		

RO 121280 B1

1
3
5
7
9
11
13
15
17
19
21
23
25
27
29
31
33
35
37
39
41
43
45
47

GGT CAA CCG CTA AAT CTG CCA CTG TAT GCC ACG CCG GTA GAC CCG AAA 2057
Gly Gln Pro Leu Asn Leu Pro Leu Tyr Ala Thr Pro Val Asp Pro Lys
570 675 680

ACC CTG CAA CGC CAG CAA GCC GGA GGG GAC GGT ACA GGC AGT AGT CCG 2115
Thr Leu Gln Arg Gln Ala Gly Gly Asp Gly Thr Gly Ser Ser Pro
685 690 695 700

GCT GGT GGT CAA CGC AGT GTT CAG GCC TGG CGC TAT CCG TTA TTG GTA 2153
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 2211
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 2259
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 2307
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 2355
Gln Asn Asn Leu Lys Gly Leu Gln His Ser Leu Thr Ala Leu Gln Ala
765 770 775 780

AGC CGT GAT GCC GAC ACA TTG CGG CAA AAA CAT TAC AGC GAC CTG ATT 2403
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 2451
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 GCC 2499
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 CCT AAC GGT GGA 2547
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 2595
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 GCG ATT TCA GAA GTG ACA GCA 2643
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 2691
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 CCC CAG ATA CAA AGC CTG CAA GAG 2739
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 2787
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 2835
Ala Asn Ala Gln Ala Ile Tyr Asp Leu Gln Thr Thr Arg Phe Thr Gly

RO 121280 B1

925	930	935	940	1
CAG GCA CTG TAT AAC TGG ATG GCC GGT CGT CTC TCC GCG CTC TAT TAC 2383				3
Gln Ala Leu Tyr Asn Trp Met Ala Gly Arg Leu Ser Ala Leu Tyr Tyr	945	950	955	5
CAA ATG TAT GAT TCC ACT CTG CCA ATC TGT CTC CAG CCA AAA GCC GCA 2931				7
Gln Met Tyr Asp Ser Thr Leu Pro Ile Cys Leu Gln Pro Lys Ala Ala	960	965	970	9
TTA GTA CAG GAA TTA GGC GAG AAA GAG AGC GAC AGT CTT TTC CAG GTT 2979				11
Leu Val Gln Glu Leu Gly Glu Lys Glu Ser Asp Ser Leu Phe Gln Val	975	980	985	13
CCG GTG TGG AAT GAT CTG TGG CAA GGG CTG TTA GCA GGA GAA GGT TTA 3027				15
Pro Val Trp Asn Asp Leu Trp Gln Gly Leu Leu Ala Gly Glu Gly Leu	990	995	1000	17
AGT TCA GAG CTA CAG AAA CTG GAT GCC ATC TGG CTT GCA CGT GGT GGT 3075				19
Ser Ser Glu Leu Gln Lys Leu Asp Ala Ile Trp Leu Ala Arg Gly Gly	1005	1010	1015	21
ATT GGG CTA GAA GCC ATC CGC ACC GTG TCG CTG GAT ACC CTG TTT GGC 3123				23
Ile Gly Leu Glu Ala Ile Arg Thr Val Ser Leu Asp Thr Leu Phe Gly	1025	1030	1035	25
ACA GGG ACG TTA AGT GAA AAT ATC AAT AAA GTG CTT AAC GGG GAA ACG 3171				27
Thr Gly Thr Leu Ser Glu Asn Ile Asn Lys Val Leu Asn Gly Glu Thr	1040	1045	1050	29
GTA TCT CCA TCC GGT GGC GTC ACT CTG GCG CTG ACA GGG GAT ATC TTC 3219				31
Val Ser Pro Ser Gly Gly Val Thr Leu Ala Leu Thr Gly Asp Ile Phe	1055	1060	1065	33
CAA GCA ACA CTG GAT TTG AGT CAG CTA GGT TTG GAT AAC TCT TAC AAC 3267				35
Gln Ala Thr Leu Asp Leu Ser Gln Leu Gly Leu Asp Asn Ser Tyr Asn	1070	1075	1080	37
TTG GGT AAC GAG AAG AAA CGT CGT ATT AAA CGT ATC GCC GTC ACC CTG 3315				39
Leu Gly Asn Glu Lys Lys Arg Arg Ile Lys Arg Ile Ala Val Thr Leu	1085	1090	1095	41
CCA ACA CTT CTG GGG CCA TAT CAA GAT CTT GAA GCC ACA CTG GTA ATG 3363				43
Pro Thr Leu Leu Gly Pro Tyr Gln Asp Leu Glu Ala Thr Leu Val Met	1105	1110	1115	45
GGT GCG GAA ATC GCC GCC TTA TCA CAC GGT GTG AAT GAC GGA GCG CCG 3411				
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 3459				
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 3507				
Asp Ala Thr Thr Gly Thr Leu Glu Leu Asn Ile Phe His Ala Gly Lys	1150	1155	1160	
GAG GGA ACC CAA CAC GAG TTG GTC GCG AAT CTG AGT GAC ATC ATT CTC 3555				
Glu Gly Thr Gln His Glu Leu Val Ala Asn Leu Ser Asp Ile Ile Val	1165	1170	1175	
CAT CTG AAT TAC ATC ATT CGA GAC GCG TAA ATTTCTTTTC TTGTGCGATT 3605				
His Leu Asn Tyr Ile Ile Arg Asp Ala *	1185	1190		

RO 121280 B1

1
3 ACAGGTCCCT ATCAGGGGCC TGTATTAAAG GAGTACTTTA TGCAGGATTC ACCAGAAGTA 3655
5 TCGATTACAA CGCTGTCACT TCCCAAAGGT GCGGGTGCTA TCAATGGCAT GGGAGAAGCA 3725
7 CTGAATGCTG CCGGCCCTGA TGGAAATGGCC TCCCTATCTC TGCCATTACC CCTTTTCGACC 3785
9 GGCAGAGGGA CGCCTCCTCG ATTATCGCTG ATTTACAGCA ACACTGCCAG TAATGGGCCT 3845
11 TTGGGCATCG GCTGGCAATG CCGTGTTATG TCCATTAGCC GACGCACCCA ACATGGCATT 3905
13 CCACAATACG GTAATGACGA CACGTTCCCTA TCCCCACAAG GCGAGGTCAT GAATATCGCC 3965
15 CTGAATGACC AAGGGCAACC TGATATCCGT CAAGACGTTA AAACGCTGCA AGGCGTTACC 4025
17 TTGCCAATTT CCTATACCGT GACCCGCTAT CAAGCCCGCC AGATCCTGGA TTTCAGTAAA 4085
19 ATCGAATACT GGCAACCTGC CTCCGGTCAA GAAGGACCGG CTTTCTGGCT GATATCGACA 4145
21 CCGGACGGGC ATCTACACAT CTTAGCGAAA ACCCGGCAGG CTTGTCTGGC AAATCCGCAA 4205
23 AATGACCAAC AAATCGCCCA GTGTTTCTG GAAGAACTG TGACGCCAGC CGGTGAACAT 4265
25 GTCAGTATC AATATCGAGC CGAAGATGAA GCCCATTGTG ACGACAATGA AAAAACCGCT 4325
27 CATCCCAATG TTACCGCACA GCGCTATCTG GTACAGGTGA ACTACAGGCA ACATCAAACC 4385
29 ACAAGCCAGC CTGTTCTGAC TGCATAACGC ACCTCCCGCA CCGGAAGAGT GGCTGTTTCA 4445
31 TCTGGTCTTT GACCACGGTG AGCGCGTACC TCACTTCATA CCCTGCCAAC ATGGGATGCA 4505
33 GGTACAGCGC AATGGTCTGT ACGCCCGGAT ATCTTCTCTC GCTATGAATA TGGTTTIGAA 4565
35 GTGCGTACTC GCGGCTTATG TCAACAAGTG CTGATGTTTC ACCGCACCGC GCTCATGGCC 4625
37 GGAGAAGCCA GTACCAATGA CGCCCCGGAA CTGGTTGGAC GCTTAATACT GGAATATGAC 4685
39 AAAAACGCCA GCGTCACCAC GTTGATTACC ATCCGTCAAT TAAGCCATGA ATCGGACGGG 4745
41 AGGCCAGTCA CCCAGCCACC ACTAGAACTA GCCTGGCAAC GGTTCGATCT GGAGAAAATC 4805
43 CCGACATGGC AACGCTTTGA CGCACTAGAT AATTTTAACT CGCAGCAACG TTATCAACTG 4865
45 GTTGATCTGC GGGGAGAAGG GTTGCCAGGT ATGCTGTATC AAGATCGAGG CGCTTGGTGG 4925
47 TATAAAGCTC CGCAACGTCA GGAAGACGGA GACAGCAATG CCCTCACTTA CGACAAAATC 4985
GCCCCACTGC CTACCCTACC CAATTTGCAG GATAATGCCT CATTGATGGA TATCAACGOA 5045
GACGGCCAAC TGGATTGGGT TGTTACCGCC TCCGGTATTC GCGGATACCA TAGTCAGCAA 5105
CCCGATGCAA AGTGGACGCA CTTTACGCCA ATCAATGCCT TGCCCGTCCA ATATTTTCAT 5165
CCAAGCATCC AGTTCGCTGA CCTTACCGGG GCAGGCTTAT CTGATTTAGT GTTGATCGGG 5225
CCGAAAAGCG TCGGTCTATA TGCCAACCAG CGAAACGGCT GCGGTAAAGG AGAAGATGTC 5285
CCCCAATCCA CAGGTATCAC CCTGCCTGTC ACAGGGACCG ATGCCCCGAA ACTGGTGGCT 5345
TTCAGTGATA TGCTCGGTTT CCGTCAACAA CATCTGGTGG AAATCAAGCG TAATCGCGTC 5405
ACCTGTTGGC CGAATCTAGG GCATGGCCGT TTCCGTCAAC CACTAACTCT GTCAGGATTT 5465
AGCCAGCCCG AAAATAGCTT CAATCCCGAA CCGCTGTTTC TGGCGGATAT CGACGGCTCC 5525
GGCACCACCG ACCTTATCTA TCGCAATCC GCCTCTTTGC TCATTTATCT CAACCAAAGT 5585

RO 121280 B1

GGTAATCAGT TTGATGCCCC GTTGACATTA GCGTTGCCAG AAGGCGTACA ATTTGACAAAC 5645
 ACTTGCCCAAC TTCAAGTCGC CGATATTCAG GGATTAGGGA TAGCCAGCTT GATTCTGACT 5705
 GTGCCACATA TCGCGCCACA TCACTGGCGT TGTGACCTGT CACTGACCAA ACCCTGGTTG 5765
 TTGAATGTAA TGAACAATAA CCGGGGCGCA CATCACACGC TACATTATCG TAGTTCCCGG 5825
 CAATTCTGGT TGGATGAAAA ATTACAGCTC ACCAAAGCAG GCAAATCTCC GGCTTGTTAT 5885
 CTGCCGTTTC CAATGCATTT GCTATGGTAT ACCGAAATTC AGGATGAAAT CAGCGGCAAC 5945
 CCGCTCACCA GTGAAGTCAA CTACAGCCAC GCGCTCTGGG ATGGTAAAGA GCGGGAATTC 6005

2) INFORMAȚIA PENTRU SECV. ID. NR. 26:

(i) CARACTERISTICI DE SECVENȚĂ:

(A) LUNGIME: 1190 aminoacizi

(B) TIP: aminoacid

(D) TOPOLOGIE: lineară

(ii) TIPUL MOLECULEI: proteină

(xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 26:

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	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	

RO 121280 B1

1 Lys Leu Lys Pro Asp Gln Trp Ser Glu Trp Arg Ala Ile Asn Ala Gly
 210 215 220
 3 Ile Ser Glu Ala Tyr Ser Gly His Val Glu Pro Phe Trp Glu Asn Asn
 225 230 235 240
 5 Lys Leu His Ile Arg Trp Phe Thr Ile Ser Lys Glu Asp Lys Ile Asp
 245 250 255
 7 Phe Val Tyr Lys Asn Ile Trp Val Met Ser Ser Asp Tyr Ser Trp Ala
 260 265 270
 9 Ser Lys Lys Lys Ile Leu Glu Leu Ser Phe Thr Asp Tyr Asn Arg Val
 275 280 285
 11 Gly Ala Thr Gly Ser Ser Ser Pro Thr Glu Val Ala Ser Gln Tyr Gly
 290 295 300
 13 Ser Asp Ala Gln Met Asn Ile Ser Asp Asp Gly Thr Val Leu Ile Phe
 305 310 315 320
 15 Gln Asn Ala Gly Gly Ala Thr Pro Ser Thr Gly Val Thr Leu Cys Tyr
 325 330 335
 17 Asp Ser Gly Asn Val Ile Lys Asn Leu Ser Ser Thr Gly Ser Ala Asn
 340 345 350
 19 Leu Ser Ser Lys Asp Tyr Ala Thr Thr Lys Leu Arg Met Cys His Gly
 355 360 365
 21 Gln Ser Tyr Asn Asp Asn Asn Tyr Cys Asn Phe Thr Leu Ser Ile Asn
 370 375 380
 23 Thr Ile Glu Phe Thr Ser Tyr Gly Thr Phe Ser Ser Asp Gly Lys Gln
 385 390 395 400
 25 Phe Thr Pro Pro Ser Gly Ser Ala Ile Asp Leu His Leu Pro Asn Tyr
 405 410 415
 27 Val Asp Leu Asn Ala Leu Leu Asp Ile Ser Leu Asp Ser Leu Leu Asn
 420 425 430
 29 Tyr Asp Val Gln Gly Gln Phe Gly Gly Ser Asn Pro Val Asp Asn Phe
 435 440 445
 31 Ser Gly Pro Tyr Gly Ile Tyr Leu Trp Glu Ile Phe Phe His Ile Pro
 450 455 460
 33 Phe Leu Val Thr Val Arg Met Gln Thr Glu Gln Arg Tyr Glu Asp Ala
 465 470 475 480
 35 Asp Thr Trp Tyr Lys Tyr Ile Phe Arg Ser Ala Gly Tyr Arg Asp Ala
 485 490 495
 37 Asn Gly Gln Leu Ile Met Asp Gly Ser Lys Pro Arg Tyr Trp Asn Val
 500 505 510
 39 Met Pro Leu Gln Leu Asp Thr Ala Trp Asp Thr Thr Gln Pro Ala Thr
 515 520 525
 41 Thr Asp Pro Asp Val Ile Ala Met Ala Asp Pro Met His Tyr Lys Leu
 530 535 540
 43 Ala Ile Phe Leu His Thr Leu Asp Leu Leu Ile Ala Arg Gly Asp Ser
 545 550 555 560

RO 121280 B1

Ala Tyr Arg Gln Leu Glu Arg Asp Thr Leu Val Glu Ala Lys Met Tyr	1
565 570 575	3
Tyr Ile Gln Ala Gln Gln Leu Leu Gly Pro Arg Pro Asp Ile His Thr	5
580 585 590	7
Thr Asn Thr Trp Pro Asn Pro Thr Leu Ser Lys Glu Ala Gly Ala Ile	9
595 600 605	11
Ala Thr Pro Thr Phe Leu Ser Ser Pro Glu Val Met Thr Phe Ala Ala	13
610 615 620	15
Trp Leu Ser Ala Gly Asp Thr Ala Asn Ile Gly Asp Gly Asp Phe Leu	17
625 630 635 640	19
Pro Pro Tyr Asn Asp Val Leu Leu Gly Tyr Trp Asp Lys Leu Glu Leu	21
645 650 655	23
Arg Leu Tyr Asn Leu Arg His Asn Leu Ser Leu Asp Gly Gln Pro Leu	25
660 665 670	27
Asn Leu Pro Leu Tyr Ala Thr Pro Val Asp Pro Lys Thr Leu Gln Arg	29
675 680 685	31
Gln Gln Ala Gly Gly Asp Gly Thr Gly Ser Ser Pro Ala Gly Gly Gln	33
690 695 700	35
Gly Ser Val Gln Gly Trp Arg Tyr Pro Leu Leu Val Glu Arg Ala Arg	37
705 710 715 720	39
Ser Ala Val Ser Leu Leu Thr Gln Phe Gly Asn Ser Leu Gln Thr Thr	41
725 730 735	43
Leu Glu His Gln Asp Asn Glu Lys Met Thr Ile Leu Leu Gln Thr Gln	45
740 745 750	47
Gln Glu Ala Ile Leu Lys His Gln His Asp Ile Gln Gln Asn Asn Leu	
755 760 765	
Lys 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	

RO 121280 B1

1 Ala Gln Lys Gln Ile Thr Leu Ser Glu Thr Glu Gln Ala Asn Ala Gln
915 920 925

3 Ala Ile Tyr Asp Leu Gln Thr Thr Arg Phe Thr Gly Gln Ala Leu Tyr
930 935 940

5 Asn Trp Met Ala Gly Arg Leu Ser Ala Leu Tyr Tyr Gln Met Tyr Asp
945 950 955 960

7 Ser Thr Leu Pro Ile Cys Leu Gln Pro Lys Ala Ala Leu Val Gln Glu
965 970 975

9 Leu Gly Glu Lys Glu Ser Asp Ser Leu Phe Gln Val Pro Val Trp Asn
980 985 990

11 Asp Leu Trp Gln Gly Leu Leu Ala Gly Glu Gly Leu Ser Ser Glu Leu
995 1000 1005

13 Gln Lys Leu Asp Ala Ile Trp Leu Ala Arg Gly Gly Ile Gly Leu Glu
1010 1015 1020

15 Ala Ile Arg Thr Val Ser Leu Asp Thr Leu Phe Gly Thr Gly Thr Leu
1025 1030 1035 1040

17 Ser Glu Asn Ile Asn Lys Val Leu Asn Gly Glu Thr Val Ser Pro Ser
1045 1050 1055

19 Gly Gly Val Thr Leu Ala Leu Thr Gly Asp Ile Phe Gln Ala Thr Leu
1060 1065 1070

21 Asp Leu Ser Gln Leu Gly Leu Asp Asn Ser Tyr Asn Leu Gly Asn Glu
1075 1080 1085

23 Lys Lys Arg Arg Ile Lys Arg Ile Ala Val Thr Leu Pro Thr Leu Leu
1090 1095 1100

25 Gly Pro Tyr Gln Asp Leu Glu Ala Thr Leu Val Met Gly Ala Glu Ile
1105 1110 1115 1120

27 Ala Ala Leu Ser His Gly Val Asn Asp Gly Gly Arg Phe Val Thr Asp
1125 1130 1135

29 Phe Asn Asp Ser Arg Phe Leu Pro Phe Glu Gly Arg Asp Ala Thr Thr
1140 1145 1150

31 Gly Thr Leu Glu Leu Asn Ile Phe His Ala Gly Lys Glu Gly Thr Gln
1155 1160 1165

33 His Glu Leu Val Ala Asn Leu Ser Asp Ile Ile Val His Leu Asn Tyr
1170 1175 1180

35 Ile Ile Arg Asp Ala *
1185 1190

(2) INFORMAȚIA PENTRU SECV. ID. NR. 27:

(i) CARACTERISTICI DE SECVENȚĂ:

(A) LUNGIME: 1881 baze perechi

(B) TIP: acid nucleic

(C) ÎMPLETIRE: dublă

(D) TOPOLOGIE: lineară

(ii) TIPUL MOLECULEI: ADN (genomic)

(ix) TRĂSĂTURA:

RO 121280 B1

(A) NUME/COD: CDS
 (B) LOCALIZARE: 1..1881
 (D) ALTĂ INFORMAȚIE: /produs= "p8"
 (xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 27:

ATG TCT GAA TCT TTA TTT ACA CAA ACG TTG AAA GAA GCG CGC CGT GAT	48	1
Met Ser Glu Ser Leu Phe Thr Gln Thr Leu Lys Glu Ala Arg Arg Asp		7
1 5 10 15		
GCA TTG GTT GCT CAT TAT ATT GCT ACT CAG GTG CCC GCA GAT TTA AAA	96	9
Ala Leu Val Ala His Tyr Ile Ala Thr Gln Val Pro Ala Asp Leu Lys		
20 25 30		11
GAG AGT ATC CAG ACC GCG GAT GAT CTG TAC GAA TAT CTG TTG CTG GAT	144	13
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	15
Thr Lys Ile Ser Asp Leu Val Thr Thr Ser Pro Leu Ser Glu Ala Ile		
50 55 60		17
GGC AGT CTG CAA TTG TTT ATT CAT CGT GCG ATA GAG GGC TAT GAC GGC	240	19
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	21
Thr Leu Ala Asp Ser Ala Lys Pro Tyr Phe Ala Asp Glu Gln Phe Leu		
85 90 95		23
TAT AAC TGG GAT AGT TTT AAC CAC CGT TAT AGC ACT TCG GCT GGC AAG	336	25
Tyr Asn Trp Asp Ser Phe Asn His Arg Tyr Ser Thr Trp Ala Gly Lys		
100 105 110		
GAA CCG TTG AAA TTC TAT GCC GGG GAT TAT ATT GAT CCA ACA TTG CGA	384	27
Glu Arg Leu Lys Phe Tyr Ala Gly Asp Tyr Ile Asp Pro Thr Leu Arg		
115 120 125		29
TTG AAT AAG ACC GAG ATA TTT ACC GCA TTT GAA CAA GGT ATT TCT CAA	432	31
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	33
Gly Lys Leu Lys Ser Glu Leu Val Glu Ser Lys Leu Arg Asp Tyr Leu		
145 150 155 160		35
ATT AGT TAT GAC ACT TTA GCC ACC CTT GAT TAT ATT ACT GCC TGC CAA	528	37
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	39
Gly Lys Asp Asn Lys Thr Ile Phe Phe Ile Gly Arg Thr Gln Asn Ala		
180 185 190		41
CCC TAT GCA TTT TAT TGG CGA AAA TTA ACT TTA GTC ACT GAT GGC GGT	624	43
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 GGC	672	45
Lys Leu Lys Pro Asp Gln Trp Ser Glu Trp Arg Ala Ile Asn Ala Gly		
210 215 220		47

RO 121280 B1

1
3 ATT AGT GAG GCA TAT TCA GGG CAT CTC GAG CCT TTC TGG GAA AAT AAC 717
Ile Ser Glu Ala Tyr Ser Gly His Val Glu Pro Phe Trp Glu Asn Asn
225 230 235 240

5 AAG CTG CAC ATC CGT TGG TTT ACT ATC TCG AAA GAA GAT AAA ATA GAT 768
Lys Leu His Ile Arg Trp Phe Thr Ile Ser Lys Glu Asp Lys Ile Asp
245 250 255

7 TTT GTT TAT AAA AAC ATC TCG GTG ATG AGT AGC GAT TAT AGC TGG GCA 816
Phe Val Tyr Lys Asn Ile Trp Val Met Ser Ser Asp Tyr Ser Trp Ala
260 265 270

9 TCA AAG AAA AAA ATC TTG GAA CTT TCT TTT ACT GAC TAC AAT AGA GTT 864
Ser Lys Lys Lys Ile Leu Glu Leu Ser Phe Thr Asp Tyr Asn Arg Val
275 280 285

11 CGA GCA ACA GGA TCA TCA AGC CCG ACT GAA CTA CCT TCA CAA TAT GGT 912
Gly Ala Thr Gly Ser Ser Ser Pro Thr Glu Val Ala Ser Gln Tyr Gly
290 295 300

13 TCT GAT GCT CAG ATG AAT ATT TCT GAT GAT GGG ACT GTA CTT ATT TTT 960
Ser Asp Ala Gln Met Asn Ile Ser Asp Asp Gly Thr Val Leu Ile Phe
305 310 315 320

15 CAG AAT GCC GGC GGA GCT ACT CCC AGT ACT GGA GTG ACG TTA TGT TAT 1008
Gln Asn Ala Gly Gly Ala Thr Pro Ser Thr Gly Val Thr Leu Cys Tyr
325 330 335

17 GAC TCT GGC AAC GTG ATT AAG AAC CTA TCT AGT ACA GGA AGT GCA AAT 1056
Asp Ser Gly Asn Val Ile Lys Asn Leu Ser Ser Thr Gly Ser Ala Asn
340 345 350

19 TTA TCG TCA AAG GAT TAT GCC ACA ACT AAA TTA CGC ATG TGT CAT GGA 1104
Leu Ser Ser Lys Asp Tyr Ala Thr Thr Lys Leu Arg Met Cys His Gly
355 360 365

21 CAA AGT TAC AAT GAT AAT AAC TAC TGC AAT TTT ACA CTC TCT ATT AAT 1152
Gln Ser Tyr Asn Asp Asn Asn Tyr Cys Asn Phe Thr Leu Ser Ile Asn
370 375 380

23 ACA ATA GAA TTC ACC TCC TAC GGC ACA TTC TCA TCA GAT GGA AAA CAA 1200
Thr Ile Glu Phe Thr Ser Tyr Gly Thr Phe Ser Ser Asp Gly Lys Gln
385 390 395 400

25 TTT ACA CCA CCT TCT GGT TCT GCC ATT GAT TTA CAC CTC CCT AAT TAT 1248
Phe Thr Pro Pro Ser Gly Ser Ala Ile Asp Leu His Leu Pro Asn Tyr
405 410 415

27 GTA GAT CTC AAC GCG CTA TTA GAT ATT AGC CTC GAT TCA CTA CTT AAT 1296
Val Asp Leu Asn Ala Leu Leu Asp Ile Ser Leu Asp Ser Leu Leu Asn
420 425 430

29 TAT GAC GTT CAG GGG CAG TTT GGC GGA TCT AAT CCG GTT GAT AAT TTC 1344
Tyr Asp Val Gln Gly Gln Phe Gly Gly Ser Asn Pro Val Asp Asn Phe
435 440 445

31 AGT GGT CCC TAT GGT ATT TAT CTA TCG GAA ATC TTC TTC CAT ATT CCG 1392
Ser Gly Pro Tyr Gly Ile Tyr Leu Trp Glu Ile Phe Phe His Ile Pro
450 455 460

33 TTC CTT GTT ACG GTC CGT ATG CAA ACC GAA CAA CGT TAC GAA GAC GCG 1440
Phe Leu Val Thr Val Arg Met Gln Thr Glu Gln Arg Tyr Glu Asp Ala
465 470 475 480

35 GAC ACT TGG TAC AAA TAT ATT TTC CCG AGC GCC GGT TAT CCG GAT GCT 1488

37
39
41
43
45
47

RO 121280 B1

Asp Thr Trp Tyr Lys Tyr Ile Phe Arg Ser Ala Gly Tyr Arg Asp Ala	1
435 490 495	
AAT GGC CAG CTC ATT ATG GAT GGC AGT AAA CCA CGT TAT TGG AAT GTG 1535	3
Asn Gly Gln Leu Ile Met Asp Gly Ser Lys Pro Arg Tyr Trp Asn Val	
500 505 510	5
ATG CCA TTG CAA CTG GAT ACC GCA TGG GAT ACC ACA CAG CCC GCC ACC 1584	7
Met Pro Leu Gln Leu Asp Thr Ala Trp Asp Thr Thr Gln Pro Ala Thr	
515 520 525	9
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	11
530 535 540	
GCG ATA TTC CTG CAT ACC CTT GAT CTA TTG ATT GCC CGA GGC GAC AGC 1680	13
Ala Ile Phe Leu His Thr Leu Asp Leu Leu Ile Ala Arg Gly Asp Ser	
545 550 555 560	15
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	17
565 570 575	
TAC ATT CAG GCA CAA CAG CTA CTG GGA CCG CGC CCT GAT ATC CAT ACC 1776	19
Tyr Ile Gln Ala Gln Gln Leu Leu Gly Pro Arg Pro Asp Ile His Thr	
580 585 590	21
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	23
595 600 605	
GCC ACA CCG ACA TTC CTC AGT TCA CCG GAG GTG ATG ACG TTC GCT GCC 1872	25
Ala Thr Pro Thr Phe Leu Ser Ser Pro Glu Val Met Thr Phe Ala Ala	
610 615 620	27
TGG CTA AGC 1881	
Trp Leu Ser 625	29
(2) INFORMAȚIA PENTRU SECV. ID. NR. 28:	31
(i) CARACTERISTICI DE SECVENȚĂ:	
(A) LUNGIME: 627 aminoacizi	33
(B) TIP: amirioacid	
(D) TOPOLOGIE: lineară	35
(ii) TIPUL MOLECULEI: proteină	
(xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 28:	37
Met Ser Glu Ser Leu Phe Thr Gln Thr Leu Lys Glu Ala Arg Arg Asp	39
1 5 10 15	
Ala Leu Val Ala His Tyr Ile Ala Thr Gln Val Pro Ala Asp Leu Lys	41
20 25 30	
Glu Ser Ile Gln Thr Ala Asp Asp Leu Tyr Glu Tyr Leu Leu Asp	43
35 40 45	
Thr Lys Ile Ser Asp Leu Val Thr Thr Ser Pro Leu Ser Glu Ala Ile	45
50 55 60	
Gly Ser Leu Gln Leu Phe Ile His Arg Ala Ile Glu Gly Tyr Asp Gly	47

RO 121280 B1

1		65				70				75					80	
3	Thr	Leu	Ala	Asp	Ser	Ala	Lys	Pro	Tyr	Phe	Ala	Asp	Glu	Gln	Phe	Leu
					85					90					95	
5	Tyr	Asn	Trp	Asp	Ser	Phe	Asn	His	Arg	Tyr	Ser	Thr	Trp	Ala	Gly	Lys
				100					105					110		
7	Glu	Arg	Leu	Lys	Phe	Tyr	Ala	Gly	Asp	Tyr	Ile	Asp	Pro	Thr	Leu	Arg
			115					120				125				
9	Leu	Asn	Lys	Thr	Glu	Ile	Phe	Thr	Ala	Phe	Glu	Gln	Gly	Ile	Ser	Gln
		130					135					140				
11	Gly	Lys	Leu	Lys	Ser	Glu	Leu	Val	Glu	Ser	Lys	Leu	Arg	Asp	Tyr	Leu
	145					150					155					160
13	Ile	Ser	Tyr	Asp	Thr	Leu	Ala	Thr	Leu	Asp	Tyr	Ile	Thr	Ala	Cys	Gln
				165						170					175	
15	Gly	Lys	Asp	Asn	Lys	Thr	Ile	Phe	Phe	Ile	Gly	Arg	Thr	Gln	Asn	Ala
			180						185					190		
17	Pro	Tyr	Ala	Phe	Tyr	Trp	Arg	Lys	Leu	Thr	Leu	Val	Thr	Asp	Gly	Gly
		195					200						205			
19	Lys	Leu	Lys	Pro	Asp	Gln	Trp	Ser	Glu	Trp	Arg	Ala	Ile	Asn	Ala	Gly
	210						215					220				
21	Ile	Ser	Glu	Ala	Tyr	Ser	Gly	His	Val	Glu	Pro	Phe	Trp	Glu	Asn	Asn
	225					230					235					240
23	Lys	Leu	His	Ile	Arg	Trp	Phe	Thr	Ile	Ser	Lys	Glu	Asp	Lys	Ile	Asp
				245						250				255		
25	Phe	Val	Tyr	Lys	Asn	Ile	Trp	Val	Met	Ser	Ser	Asp	Tyr	Ser	Trp	Ala
			260						265					270		
27	Ser	Lys	Lys	Lys	Ile	Leu	Glu	Leu	Ser	Phe	Thr	Asp	Tyr	Asn	Arg	Val
		275						280					285			
29	Gly	Ala	Thr	Gly	Ser	Ser	Ser	Pro	Thr	Glu	Val	Ala	Ser	Gln	Tyr	Gly
	290						295					300				
31	Ser	Asp	Ala	Gln	Met	Asn	Ile	Ser	Asp	Asp	Gly	Thr	Val	Leu	Ile	Phe
	305					310					315					320
33	Gln	Asn	Ala	Gly	Gly	Ala	Thr	Pro	Ser	Thr	Gly	Val	Thr	Leu	Cys	Tyr
				325						330					335	
35	Asp	Ser	Gly	Asn	Val	Ile	Lys	Asn	Leu	Ser	Ser	Thr	Gly	Ser	Ala	Asn
			340						345					350		
37	Leu	Ser	Ser	Lys	Asp	Tyr	Ala	Thr	Thr	Lys	Leu	Arg	Met	Cys	His	Gly
		355						360					365			
39	Gln	Ser	Tyr	Asn	Asp	Asn	Asn	Tyr	Cys	Asn	Phe	Thr	Leu	Ser	Ile	Asn
	370						375					380				
41	Thr	Ile	Glu	Phe	Thr	Ser	Tyr	Gly	Thr	Phe	Ser	Ser	Asp	Gly	Lys	Gln
	385					390					395					400
43	Phe	Thr	Pro	Pro	Ser	Gly	Ser	Ala	Ile	Asp	Leu	His	Leu	Pro	Asn	Tyr
				405						410				415		
45	Val	Asp	Leu	Asn	Ala	Leu	Leu	Asp	Ile	Ser	Leu	Asp	Ser	Leu	Leu	Asn

RO 121280 B1

	420	425	430	1
Tyr Asp Val Gln Gly Gln Phe Gly Gly Ser Asn Pro Val Asp Asn Phe	435	440	445	3
Ser Gly Pro Tyr Gly Ile Tyr Leu Trp Glu Ile Phe Phe His Ile Pro	450	455	460	5
Phe Leu Val Thr Val Arg Met Gln Thr Glu Gln Arg Tyr Glu Asp Ala	465	470	475	7
Asp Thr Trp Tyr Lys Tyr Ile Phe Arg Ser Ala Gly Tyr Arg Asp Ala	485	490	495	9
Asn Gly Gln Leu Ile Met Asp Gly Ser Lys Pro Arg Tyr Trp Asn Val	500	505	510	11
Met Pro Leu Gln Leu Asp Thr Ala Trp Asp Thr Thr Gln Pro Ala Thr	515	520	525	13
Thr Asp Pro Asp Val Ile Ala Met Ala Asp Pro Met His Tyr Lys Leu	530	535	540	15
Ala Ile Phe Leu His Thr Leu Asp Leu Leu Ile Ala Arg Gly Asp Ser	545	550	555	17
Ala Tyr Arg Gln Leu Glu Arg Asp Thr Leu Val Glu Ala Lys Met Tyr	565	570	575	19
Tyr Ile Gln Ala Gln Gln Leu Leu Gly Pro Arg Pro Asp Ile His Thr	580	585	590	21
Thr Asn Thr Trp Pro Asn Pro Thr Leu Ser Lys Glu Ala Gly Ala Ile	595	600	605	23
Ala Thr Pro Thr Phe Leu Ser Ser Pro Glu Val Met Thr Phe Ala Ala	610	615	620	25
Trp Leu Ser	625			27
(2) INFORMAȚIA PENTRU SECV. ID. NR. 29:				29
(i) CARACTERISTICI DE SECVENȚĂ:				31
(A) LUNGIME: 1689 baze perechi				33
(B) TIP: acid nucleic				35
(C) ÎMPLETIRE: dublă				37
(D) TOPOLOGIE: lineară				39
(ii) TIPUL MOLECULEI: ADN (genomic)				41
(ix) TRĂSĂTURA:				43
(A) NUME/COD: CDS				45
(B) LOCALIZARE: 1..1689				47
(D) ALTĂ INFORMAȚIE: /produs= "s8"				
(xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 29:				
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

RO 121280 B1

1																	
3	AAC	GAT	GTA	CTA	CTC	GGT	TAC	TGG	GAT	AAA	CTT	GAG	TTA	CGC	CTA	TAC	36
	Asn	Asp	Val	Leu	Leu	Gly	Tyr	Trp	Asp	Lys	Leu	Glu	Leu	Arg	Leu	Tyr	
			20						25					30			
5	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				
7	CTG	TAT	CCC	ACG	CCG	GTA	GAC	CCG	AAA	ACC	CTG	CAA	CCG	CAG	CAA	CCC	192
	Leu	Tyr	Ala	Thr	Pro	Val	Asp	Pro	Lys	Thr	Leu	Gln	Arg	Gln	Gln	Ala	
		50					55					60					
9	CGA	GGC	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	
11	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		
13	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			
15	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				
17	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					
19	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	
21	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		
23	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			
25	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				
27	AAC	CTC	TTC	GGG	CTG	GCT	AAC	GCT	GCA	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					
29	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		
31	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		
33	GAA	TGG	GCA	TTG	CAA	CGG	GAT	ATT	GCT	GAT	AAC	GAA	ATA	ACC	CAA	CTG	816
	Glu	Trp	Ala	Gln	Gln	Arg	Asp	Ile	Ala	Asp	Asn	Glu	Ile	Thr	Gln	Leu	
			260						265					270			
35	GAT	GCC	CAG	ATA	CAA	AGC	CTG	CAA	GAG	CAA	ATC	ACG	ATG	GCA	CAA	AAA	864

RO 121280 B1

Asp	Ala	Gln	Ile	Gln	Ser	Leu	Gln	Glu	Gln	Ile	Thr	Met	Ala	Gln	Lys	1
275							280					285				3
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	5
290						295					300					7
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	1
305					310					315					320	3
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	5
					325				330					335		7
CCA	ATC	TGT	CTC	CAG	CCA	AAA	GCC	GCA	TTA	GTA	CAG	GAA	TTA	GCC	GAG	1056
Pro	Ile	Cys	Leu	Gln	Pro	Lys	Ala	Ala	Leu	Val	Gln	Glu	Leu	Gly	Glu	11
			340					345					350			13
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	15
		355				360						365				17
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	19
		370				375					380					21
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	23
385					390					395					400	25
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	27
				405				410						415		29
ATC	AAT	AAA	GTG	CTT	AAC	GGG	GAA	ACG	GTA	TCT	CCA	TCC	GGT	GGC	CTC	1296
Ile	Asn	Lys	Val	Leu	Asn	Gly	Glu	Thr	Val	Ser	Pro	Ser	Gly	Gly	Val	31
			420				425						430			33
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	35
			435			440					445					37
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	39
		450			455						460					41
CGT	ATT	AAA	CGT	ATC	GCC	GTG	ACC	CTG	CCA	ACA	CTT	CTG	GCG	CCA	TAT	1440
Arg	Ile	Lys	Arg	Ile	Ala	Val	Thr	Leu	Pro	Thr	Leu	Leu	Gly	Pro	Tyr	43
465					470				475					480		45
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	37
				485				490					495			39
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	41
			500				505						510			43
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	45
		515					520					525				47
GAG	CTC	AAT	ATT	TTC	GAT	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	49
		530				535					540					51
GTC	GCG	AAT	CTG	AGT	GAC	ATC	ATT	GTG	GAT	CTG	AAT	TAC	ATC	ATT	GGA	1680
Val	Ala	Asn	Leu	Ser	Asp	Ile	Ile	Val	His	Leu	Asn	Tyr	Ile	Ile	Arg	53
		545				550				555					560	55
GAC	GCG	TAA														1639
Asp	Ala	*														57

RO 121280 B1

(2) INFORMAȚIA PENTRU SECV. ID. NR. 30:

(i) CARACTERISTICI DE SECVENȚĂ:

(A) LUNGIME: 563 aminoacizi

(B) TIP: aminoacid

(D) TOPOLOGIE: lineară

(ii) TIPUL MOLECULEI: proteină 20

(xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 30:

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				

RO 121280 B1

225	230	235	240	1
Ser Ala Gly Ile	Ser Glu Val Thr	Ala Gly Tyr Gln Arg Arg	Gln Glu	3
	245	250	255	
Glu Trp Ala Leu	Gln Arg Asp Ile	Ala Asp Asn Glu Ile	Thr Gln Leu	5
	260	265	270	
Asp Ala Gln Ile	Gln Ser Leu Gln	Glu Gln Ile Thr Met	Ala Gln Lys	7
	275	280	285	
Gln Ile Thr Leu	Ser Glu Thr Glu	Gln Ala Asn Ala	Gln Ala Ile Tyr	9
	290	295	300	
Asp Leu Gln Thr	Thr Arg Phe Thr	Gly Gln Ala Leu	Tyr Asn Trp Met	11
305	310	315	320	
Ala Gly Arg Leu	Ser Ala Leu Tyr	Tyr Gln Met Tyr	Asp Ser Thr Leu	13
	325	330	335	
Pro Ile Cys Leu	Gln Pro Lys Ala	Ala Leu Val Gln	Glu Leu Gly Glu	15
	340	345	350	
Lys Glu Ser Asp	Ser Leu Phe Gln	Val Pro Val Trp	Asn Asp Leu Trp	17
	355	360	365	
Gln Gly Leu Leu	Ala Gly Glu Gly	Leu Ser Ser Glu	Leu Gln Lys Leu	19
	370	375	380	
Asp Ala Ile Trp	Leu Ala Arg Gly	Gly Ile Gly Leu	Glu Ala Ile Arg	21
385	390	395	400	
Thr Val Ser Leu	Asp Thr Leu Phe	Gly Thr Gly Thr	Leu Ser Glu Asn	23
	405	410	415	
Ile Asn Lys Val	Leu Asn Gly Glu	Thr Val Ser Pro	Ser Gly Gly Val	25
	420	425	430	
Thr Leu Ala Leu	Thr Gly Asp Ile	Phe Gln Ala Thr	Leu Asp Leu Ser	27
	435	440	445	
Gln Leu Gly Leu	Asp Asn Ser Tyr	Asn Leu Gly Asn	Glu Lys Lys Arg	29
	450	455	460	
Arg Ile Lys Arg	Ile Ala Val Thr	Leu Pro Thr Leu	Leu Gly Pro Tyr	31
465	470	475	480	
Gln Asp Leu Glu	Ala Thr Leu Val	Met Gly Ala Glu	Ile Ala Ala Leu	33
	485	490	495	
Ser His Gly Val	Asn Asp Gly Gly	Arg Phe Val Thr	Asp Phe Asn Asp	35
	500	505	510	
Ser Arg Phe Leu	Pro Phe Glu Gly	Arg Asp Ala Thr	Thr Gly Thr Leu	37
	515	520	525	
Glu Leu Asn Ile	Phe His Ala Gly	Lys Glu Gly Thr	Gln His Glu Leu	39
	530	535	540	
Val Ala Asn Leu	Ser Asp Ile Ile	Val His Leu Asn	Tyr Ile Ile Arg	41
545	550	555	560	
Asp Ala *				43

(2) INFORMAȚIA PENTRU SECV. ID. NR. 31:

(i) CARACTERISTICI DE SECVENȚĂ:

RO 121280 B1

1 (A) LUNGIME: 4458 baze perechi
 (B) TIP: acid nucleic
 3 (C) ÎMPLETIRE: dublă
 (D) TOPOLOGIE: lineară
 5 (ii) TIPUL MOLECULEI.: ADN (genomic)
 (ix) TRĂSĂTURA:
 7 (A) NUME/COD: CDS
 (B) LOCALIZARE: 1..4458
 9 (xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 31:

11 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
 13 1 5 10 15

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

17 CCT GAT GCA 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
 19 35 40 45

21 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

23 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
 25 65 70 75 80

27 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
 29 85 90 95

31 CTA TCC CCA CAA GGC GAG GTC ATG AAT ATC GCC CTG AAT GAC CAA GGC 336
 Leu Ser Pro Gln Gly Glu Val Met Asn Ile Ala Leu Asn Asp Gln Gly
 100 105 110

33 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
 35 115 120 125

37 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

39 TTC AGT AAA ATC GAA TAC TCG 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
 41 145 150 155 160

43 GCT TTC TGG CTG ATA TCG ACA CCG GAC GGG CAT CTA CAC ATC TTA GGC 528
 Ala Phe Trp Leu Ile Ser Thr Pro Asp Gly His Leu His Ile Leu Gly
 165 170 175

45 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

47

RO 121280 B1

180	185	190		1
GCC CAG TGG TTG CTG GAA GAA ACT GTG ACG CCA GCC GGT GAA CAT CTC 624				3
Ala Gln Trp Leu Leu Glu Glu Thr Val Thr Pro Ala Gly Glu His Val				
195	200	205		5
AGC TAT CAA TAT CGA GCC GAA GAT GAA GCC CAT TGT GAC GAC AAT GAA 672				7
Ser Tyr Gln Tyr Arg Ala Glu Asp Glu Ala His Cys Asp Asp Asn Glu				
210	215	220		
AAA ACC GCT CAT CCC AAT GTT ACC GCA CAG CGC TAT CTG GTA CAG GTG 720				9
Lys Thr Ala His Pro Asn Val Thr Ala Gln Arg Tyr Leu Val Gln Val				
225	230	235	240	
AAC TAC GGC AAC ATC AAA CCA CAA GCC AGC CTG TTC GTA CTG GAT AAC 768				11
Asn Tyr Gly Asn Ile Lys Pro Gln Ala Ser Leu Phe Val Leu Asp Asn				
245	250	255		13
GCA CCT CCC GCA CCG GAA GAG TGG CTG TTT CAT CTG GTC TTT GAC CAC 816				15
Ala Pro Pro Ala Pro Glu Glu Trp Leu Phe His Leu Val Phe Asp His				
260	265	270		
GGT GAG CGC GAT ACC TCA CTT CAT ACC GTG CCA ACA TGG GAT GCA GGT 864				17
Gly Glu Arg Asp Thr Ser Leu His Thr Val Pro Thr Trp Asp Ala Gly				
275	280	285		19
ACA GCG CAA TGG TCT GTA CGC CCG GAT ATC TTC TCT CGC TAT GAA TAT 912				21
Thr Ala Gln Trp Ser Val Arg Pro Asp Ile Phe Ser Arg Tyr Glu Tyr				
290	295	300		
GGT TTT GAA GTG CGT ACT CGC CGC TTA TGT CAA CAA GTG CTG ATG TTT 960				23
Gly Phe Glu Val Arg Thr Arg Arg Leu Cys Gln Gln Val Leu Met Phe				
305	310	315	320	25
CAC CGC ACC GCG CTC ATG GCC CGA GAA GCC AGT ACC AAT GAC GCC CCG 1008				27
His Arg Thr Ala Leu Met Ala Gly Glu Ala Ser Thr Asn Asp Ala Pro				
325	330	335		
GAA CTG GTT GGA CGC TTA ATA CTG GAA TAT GAC AAA AAC GCC AGC GTC 1056				29
Glu Leu Val Gly Arg Leu Ile Leu Glu Tyr Asp Lys Asn Ala Ser Val				
340	345	350		
ACC ACG TTG ATT ACC ATC CGT CAA TTA AGC CAT GAA TCG GAC GGG AGG 1104				31
Thr Thr Leu Ile Thr Ile Arg Gln Leu Ser His Glu Ser Asp Gly Arg				
355	360	365		33
CCA GTC ACC CAG CCA CCA CTA GAA CTA GCC TGG CAA CCG TTT GAT CTG 1152				35
Pro Val Thr Gln Pro Pro Leu Glu Leu Ala Trp Gln Arg Phe Asp Leu				
370	375	380		
GAG AAA ATC CCG ACA TGG CAA CGC TTT GAC GCA CTA GAT AAT TTT AAC 1200				37
Glu Lys Ile Pro Thr Trp Gln Arg Phe Asp Ala Leu Asp Asn Phe Asn				
385	390	395	400	39
TCG CAG CAA CGT TAT CAA CTG GTT GAT CTG CGG GGA GAA GGG TTG CCA 1248				41
Ser Gln Gln Arg Tyr Gln Leu Val Asp Leu Arg Gly Glu Gly Leu Pro				
405	410	415		
GGT ATG CTG TAT CAA GAT CGA GGC GCT TGG TGG TAT AAA GCT CCG CAA 1296				43
Gly Met Leu Tyr Gln Asp Arg Gly Ala Trp Trp Tyr Lys Ala Pro Gln				
420	425	430		
CGT CAG GAA GAC GGA GAC AGC AAT GCC GTC ACT TAC GAC AAA ATC GCC 1344				45
Arg Gln Glu Asp Gly Asp Ser Asn Ala Val Thr Tyr Asp Lys Ile Ala				
435	440	445		47

RO 121280 B1

1	CCA	CTG	CGT	ACC	CTA	CCC	AAT	TTG	CAG	GAT	AAT	GCC	TCA	TTG	ATG	GAT	1392
3	Pro	Leu	Pro	Thr	Leu	Pro	Asn	Leu	Gln	Asp	Asn	Ala	Ser	Leu	Met	Asp	
	450						455					460					
5	ATC	AAC	GGA	GAC	GGC	CAA	CTG	GAT	TGG	GTT	GTT	ACC	GCC	TCC	GGT	ATT	1440
	Ile	Asn	Gly	Asp	Gly	Gln	Leu	Asp	Trp	Val	Val	Thr	Ala	Ser	Gly	Ile	
	465					470					475					480	
7	CCC	GGA	TAC	CAT	AGT	CAG	CAA	CCC	GAT	GGA	AAG	TGG	ACG	CAC	TTT	ACG	1488
9	Arg	Gly	Tyr	His	Ser	Gln	Gln	Pro	Asp	Gly	Lys	Trp	Thr	His	Phe	Thr	
					485				490						495		
11	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		
13	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				
15	AAA	AGC	GTG	CGT	CTA	TAT	GCC	AAC	CAG	CGA	AAC	GGC	TGG	CGT	AAA	GGA	1632
17	Lys	Ser	Val	Arg	Leu	Tyr	Ala	Asn	Gln	Arg	Asn	Gly	Trp	Arg	Lys	Gly	
	530						535					540					
19	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	
21	GAT	GCC	CGC	AAA	CTG	GTG	GCT	TTC	AGT	GAT	ATG	CTC	GGT	TCC	GGT	CAA	1728
23	Asp	Ala	Arg	Lys	Leu	Val	Ala	Phe	Ser	Asp	Met	Leu	Gly	Ser	Gly	Gln	
					565				570						575		
25	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			
27	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					
29	CAG	CCC	GAA	AAT	AGC	TTC	AAT	CCC	GAA	CGG	CTG	TTT	CTG	GCG	GAT	ATC	1872
31	Gln	Pro	Glu	Asn	Ser	Phe	Asn	Pro	Glu	Arg	Leu	Phe	Leu	Ala	Asp	Ile	
	610						615					620					
33	GAC	GGC	TCC	GGC	ACC	ACC	GAC	CTT	ATC	TAT	GGG	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	
35	CTC	ATT	TAT	CTC	AAC	CAA	AGT	GGT	AAT	CAG	TTT	GAT	GCC	CCG	TTG	ACA	1968
37	Leu	Ile	Tyr	Leu	Asn	Gln	Ser	Gly	Asn	Gln	Phe	Asp	Ala	Pro	Leu	Thr	
					645				650						655		
39	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			
41	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					
43	CCA	CAT	ATC	GCG	CCA	CAT	CAC	TGG	CGT	TGT	GAC	CTG	TCA	CTG	ACC	AAA	2112
45	Pro	His	Ile	Ala	Pro	His	His	Trp	Arg	Cys	Asp	Leu	Ser	Leu	Thr	Lys	
	690						695					700					
47	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	

RO 121280 B1

705	710	715	720	1
CTA CAT TAT CGT ACT TCC GCG CAA TTC TCG TTC GAT GAA AAA TTA CAG 2208				3
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				5
Leu Thr Lys Ala Gly Lys Ser Pro Ala Cys Tyr Leu Pro Phe Pro Met	740	745	750	7
CAT TTG CTA TGG TAT ACC GAA ATT CAG GAT GAA ATC AGC GGC AAC CGG 2304				9
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				11
Leu Thr Ser Glu Val Asn Tyr Ser His Gly Val Trp Asp Gly Lys Glu	770	775	780	13
CGG GAA TTC ACA GGA TTT GGC TGC ATC AAA CAG ACA GAT ACC ACA ACC 2400				15
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				17
Phe Ser His Gly Thr Ala Pro Glu Gln Ala Ala Pro Ser Leu Ser Ile	805	810	815	
AGC TGG TTT GCC ACC GGC ATG GAT GAA GTA GAC AGC CAA TTA GCT ACG 2496				19
Ser Trp Phe Ala Thr Gly Met Asp Glu Val Asp Ser Gln Leu Ala Thr	820	825	830	21
GAA TAT TGG CAG GCA GAC ACG CAA GCT TAT AGC GGA TTT GAA ACC CGT 2544				23
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				25
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				27
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				31
Leu Arg Thr Glu Leu Tyr Gly Leu Asp Gly Thr Asp Lys Gln Thr Val	885	890	895	
CCT TAT ACC GTC ACT GAA TCG CCG TAT CAG GTA CCG TCT ATT CCC GTA 2736				33
Pro Tyr Thr Val Ser Glu Ser Arg Tyr Gln Val Arg Ser Ile Pro Val	900	905	910	35
AAT AAA GAA ACT GAA TTA TCT GCC TGG CTG ACT GCT ATT GAA AAT CGC 2784				37
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				39
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				41
Ile Lys Leu Gln His Asp Ile Phe Gly Gln Ser Leu Gln Ser Val Asp	945	950	955	960
ATT GCC TGG CCG CCG CCG GAA AAA CCA GCA GTG AAT CCC TAC CCG COT 2928				43
Ile Ala Trp Pro Arg Arg Glu Lys Pro Ala Val Asn Pro Tyr Pro Pro	965	970	975	45
				47

RO 121280 B1

1 ACT TTG CCG GAA ACG CTA TTT GAC AGC AGC TAT GAT GAT CAA CAA CAA 3275
Thr Leu Pro Glu Thr Leu Phe Asp Ser Ser Tyr Asp Asp Gln Gln Gln 980 985 990

3 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

5 GCG GAA AAC TGG CGA TTA GGT TTA CCG AAT GCA CAA CCG CGT GAT GTT 3072
Gly Glu Asn Trp Arg Leu Gly Leu Pro Asn Ala Gln Arg Arg Asp Val 1010 1015 1020

7 TAT ACT TAT GAC CCG 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

9 ATC TTG CTG AAA GAT GAT GGC CTG CTA CCA GAT GAA AAA GCG GCC CTT 3168
Ile Leu Leu Lys Asp Asp Gly Leu Leu Ala Asp Glu Lys Ala Ala Val 1045 1050 1055

11 TAT CTG CGA CAA CAA CAG ACG TTT TAC ACC GCC CGT CAA GCC CAA GTC 3216
Tyr Leu Gly Gln Gln Thr Phe Tyr Thr Ala Gly Gln Ala Glu Val 1060 1065 1070

13 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

15 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

17 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

19 GCG CCG TTG TTT AAT ACC AGA TCA GAA AGC CCG GTA TGG GCG GCA CCG 3408
Ala Arg Leu Phe Asn Thr Arg Ser Glu Ser Pro Val Trp Ala Ala Arg 1125 1130 1135

21 CAA GGT TAT ACC GAT TAC GGT GAC GCC GCA CAG TTC TGG CCG CCT CAG 3456
Gln Gly Tyr Thr Asp Tyr Gly Asp Ala Ala Gln Phe Trp Arg Pro Gln 1140 1145 1150

23 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

25 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

27 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

29 GAT ATT AAT GAT AAT CAA CAT ATT GTG ACT CTG GAC GCG CTA GGT CCG 3648
Asp Ile Asn Asp Asn Gln His Ile Val Thr Leu Asp Ala Leu Gly Arg 1205 1210 1215

31 GTA ACC ACC AGC CCG 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

33 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

RO 121280 B1

1235	1240	1245	1
GCA TTA ACC GGC GCA CTC CCT GTT GCC CAA TGT TTA GTC TAT GCC GTT 3722			3
Ala Leu Thr Gly Ala Leu Pro Val Ala Gln Cys Leu Val Tyr Ala Val			
1250	1255	1260	5
GAT AGC TGG ATG CCG TCG TTA TCT TTG TCT CAG CTT TCT CAG TCA CAA 3840			7
Asp Ser Trp Met Pro Ser Leu Ser Leu Ser Gln Leu Ser Gln			
1265	1270	1275	9
GAA GAG GCA GAA GCG CTA TGG GCG CAA CTG CGT GCC GCT CAT ATG ATT 3888			11
Glu Glu Ala Glu Ala Leu Trp Ala Gln Leu Arg Ala Ala His Met Ile			
1285	1290	1295	13
ACC GAA GAT GGG AAA GTG TGT GCG TTA AGC GGG AAA CGA GGA ACA AGC 3936			15
Thr Glu Asp Gly Lys Val Cys Ala Leu Ser Gly Lys Arg Gly Thr Ser			
1300	1305	1310	17
CAT CAG AAC CTG ACG ATT CAA CTT ATT TCG CTA TTG GCA AGT ATT CCC 3984			19
His Gln Asn Leu Thr Ile Gln Leu Ile Ser Leu Leu Ala Ser Ile Pro			
1315	1320	1325	21
CGT TTA CCG CCA CAT GTA CTG GGG ATC ACC ACT GAT CGC TAT GAT AGC 4032			23
Arg Leu Pro Pro His Val Leu Gly Ile Thr Thr Asp Arg Tyr Asp Ser			
1330	1335	1340	25
GAT CCG CAA CAG CAG CAC CAA CAG ACG GTG AGC TTT AGT GAC GGT TTT 4080			27
Asp Pro Gln Gln Gln His Gln Gln Thr Val Ser Phe Ser Asp Gly Phe			
1345	1350	1355	29
GGC CGG TTA CTC CAG AGT TCA GCT CGT CAT GAG TCA GGT GAT GCC TGG 4128			31
Gly Arg Leu Leu Gln Ser Ser Ala Arg His Glu Ser Gly Asp Ala Trp			
1365	1370	1375	33
CAA CGT AAA GAG GAT GGC GGG CTG GTC GTG GAT GCA AAT GGC GTT CTG 4176			35
Gln Arg Lys Glu Asp Gly Gly Leu Val Val Asp Ala Asn Gly Val Leu			
1380	1385	1390	37
GTC AGT GCC CCT ACA GAC ACC CGA TGG GCC GTT TCC GGT CGC ACA GAA 4224			39
Val Ser Ala Pro Thr Asp Thr Arg Trp Ala Val Ser Gly Arg Thr Glu			
1395	1400	1405	41
TAT GAC GAC AAA GGC CAA CCT GTG CGT ACT TAT CAA CCC TAT TTT CTA 4272			43
Tyr Asp Asp Lys Gly Gln Pro Val Arg Thr Tyr Gln Pro Tyr Phe Leu			
1410	1415	1420	45
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	
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 *			
1475	1480	1485	

(2) INFORMAȚIA PENTRU SECV. ID. NR. 32:

(i) CARACTERISTICI DE SECVENȚĂ:

RO 121280 B1

1 (A) LUNGIME: 1486 aminoacizi
 (B) TIP: aminoacid
 3 (D) TOPOLOGIE: lineară 5
 (ii) TIPUL MOLECULEI: proteină
 5 (xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 32:

7 Met Gln Asp Ser Pro Glu Val Ser Ile Thr Thr Leu Ser Leu Pro Lys
 1 5 10 15

9 Gly Gly Gly Ala Ile Asn Gly Met Gly Glu Ala Leu Asn Ala Ala Gly
 20 25 30

11 Pro Asp Gly Met Ala Ser Leu Ser Leu Pro Leu Pro Leu Ser Thr Gly
 35 40 45

13 Arg Gly Thr Ala Pro Gly Leu Ser Leu Ile Tyr Ser Asn Ser Ala Gly
 50 55 60

15 Asn Gly Pro Phe Gly Ile Gly Trp Gln Cys Gly Val Met Ser Ile Ser
 65 70 75 80

17 Arg Arg Thr Gln His Gly Ile Pro Gln Tyr Gly Asn Asp Asp Thr Phe
 85 90 95

19 Leu Ser Pro Gln Gly Glu Val Met Asn Ile Ala Leu Asn Asp Gln Gly
 100 105 110

21 Gln Pro Asp Ile Arg Gln Asp Val Lys Thr Leu Gln Gly Val Thr Leu
 115 120 125

23 Pro Ile Ser Tyr Thr Val Thr Arg Tyr Gln Ala Arg Gln Ile Leu Asp
 130 135 140

25 Phe Ser Lys Ile Glu Tyr Trp Gln Pro Ala Ser Gly Gln Glu Gly Arg
 145 150 155 160

27 Ala Phe Trp Leu Ile Ser Thr Pro Asp Gly His Leu His Ile Leu Gly
 165 170 175

29 Lys Thr Ala Gln Ala Cys Leu Ala Asn Pro Gln Asn Asp Gln Gln Ile
 180 185 190

31 Ala Gln Trp Leu Leu Glu Glu Thr Val Thr Pro Ala Gly Glu His Val
 195 200 205

33 Ser Tyr Gln Tyr Arg Ala Glu Asp Glu Ala His Cys Asp Asp Asn Glu
 210 215 220

35 Lys Thr Ala His Pro Asn Val Thr Ala Gln Arg Tyr Leu Val Gln Val
 225 230 235 240

37 Asn Tyr Gly Asn Ile Lys Pro Gln Ala Ser Leu Phe Val Leu Asp Asn
 245 250 255

39 Ala Pro Pro Ala Pro Glu Glu Trp Leu Phe His Leu Val Phe Asp His
 260 265 270

41 Gly Glu Arg Asp Thr Ser Leu His Thr Val Pro Thr Trp Asp Ala Gly
 275 280 285

43 Thr Ala Gln Trp Ser Val Arg Pro Asp Ile Phe Ser Arg Tyr Glu Tyr

RO 121280 B1

290	295	300	1
Gly Phe Glu Val Arg Thr Arg Arg Leu Cys Gln Gln Val Leu Met Phe 305 310 315 320			3
His Arg Thr Ala Leu Met Ala Gly Glu Ala Ser Thr Asn Asp Ala Pro 325 330 335			5
Glu Leu Val Gly Arg Leu Ile Leu Glu Tyr Asp Lys Asn Ala Ser Val 340 345 350			7
Thr Thr Leu Ile Thr Ile Arg Gln Leu Ser His Glu Ser Asp Gly Arg 355 360 365			9
Pro Val Thr Gln Pro Pro Leu Glu Leu Ala Trp Gln Arg Phe Asp Leu 370 375 380			11
Glu Lys Ile Pro Thr Trp Gln Arg Phe Asp Ala Leu Asp Asn Phe Asn 385 390 395 400			13
Ser Gln Gln Arg Tyr Gln Leu Val Asp Leu Arg Gly Glu Gly Leu Pro 405 410 415			15
Gly Met Leu Tyr Gln Asp Arg Gly Ala Trp Trp Tyr Lys Ala Pro Gln 420 425 430			17
Arg Gln Glu Asp Gly Asp Ser Asn Ala Val Thr Tyr Asp Lys Ile Ala 435 440 445			19
Pro Leu Pro Thr Leu Pro Asn Leu Gln Asp Asn Ala Ser Leu Met Asp 450 455 460			21
Ile Asn Gly Asp Gly Gln Leu Asp Trp Val Val Thr Ala Ser Gly Ile 465 470 475 480			23
Arg Gly Tyr His Ser Gln Gln Pro Asp Gly Lys Trp Thr His Phe Thr 485 490 495			25
Pro Ile Asn Ala Leu Pro Val Glu Tyr Phe His Pro Ser Ile Gln Phe 500 505 510			27
Ala Asp Leu Thr Gly Ala Gly Leu Ser Asp Leu Val Leu Ile Gly Pro 515 520 525			29
Lys Ser Val Arg Leu Tyr Ala Asn Gln Arg Asn Gly Trp Arg Lys Gly 530 535 540			31
Glu Asp Val Pro Gln Ser Thr Gly Ile Thr Leu Pro Val Thr Gly Thr 545 550 555 560			33
Asp Ala Arg Lys Leu Val Ala Phe Ser Asp Met Leu Gly Ser Gly Gln 565 570 575			35
Gln His Leu Val Glu Ile Lys Gly Asn Arg Val Thr Cys Trp Pro Asn 580 585 590			37
Leu Gly His Gly Arg Phe Gly Gln Pro Leu Thr Leu Ser Gly Phe Ser 595 600 605			39
Gln Pro Glu Asn Ser Phe Asn Pro Glu Arg Leu Phe Leu Ala Asp Ile 610 615 620			41
Asp Gly Ser Gly Thr Thr Asp Leu Ile Tyr Ala Gln Ser Gly Ser Leu 625 630 635 640			43
Leu Ile Tyr Leu Asn Gln Ser Gly Asn Gln Phe Asp Ala Pro Leu Thr			45
			47

RO 121280 B1

[illegible]

RO 121280 B1

995	1000	1005	1
Gly Glu Asn Trp Arg Leu Gly Leu Pro Asn Ala Gln Arg Arg Asp Val 1010 1015 1020			3
Tyr Thr Tyr Asp Arg Ser Lys Ile Pro Thr Glu Gly Ile Ser Leu Glu 1025 1030 1035 1040			5
Ile Leu Leu Lys Asp Asp Gly Leu Leu Ala Asp Glu Lys Ala Ala Val 1045 1050 1055			7
Tyr Leu Gly Gln Gln Gln Thr Phe Tyr Thr Ala Gly Gln Ala Glu Val 1060 1065 1070			9
Thr Leu Glu Lys Pro Thr Leu Gln Ala Leu Val Ala Phe Gln Glu Thr 1075 1080 1085			11
Ala Met Met Asp Asp Thr Ser Leu Gln Ala Tyr Glu Gly Val Ile Glu 1090 1095 1100			13
Glu Gln Glu Leu Asn Thr Ala Leu Thr Gln Ala Gly Tyr Gln Gln Val 1105 1110 1115 1120			15
Ala Arg Leu Phe Asn Thr Arg Ser Glu Ser Pro Val Trp Ala Ala Arg 1125 1130 1135			17
Gln Gly Tyr Thr Asp Tyr Gly Asp Ala Ala Gln Phe Trp Arg Pro Gln 1140 1145 1150			19
Ala Gln Arg Asn Ser Leu Leu Thr Gly Lys Thr Thr Leu Thr Trp Asp 1155 1160 1165			21
Thr His His Cys Val Ile Ile Gln Thr Gln Asp Ala Ala Gly Leu Thr 1170 1175 1180			23
Thr Gln Ala His Tyr Asp Tyr Arg Phe Leu Thr Pro Val Gln Leu Thr 1185 1190 1195 1200			25
Asp Ile Asn Asp Asn Gln His Ile Val Thr Leu Asp Ala Leu Gly Arg 1205 1210 1215			27
Val Thr Thr Ser Arg Phe Trp Gly Thr Glu Ala Gly Gln Ala Ala Gly 1220 1225 1230			29
Tyr Ser Asn Gln Pro Phe Thr Pro Pro Asp Ser Val Asp Lys Ala Leu 1235 1240 1245			31
Ala Leu Thr Gly Ala Leu Pro Val Ala Gln Cys Leu Val Tyr Ala Val 1250 1255 1260			33
Asp Ser Trp Met Pro Ser Leu Ser Leu Ser Gln Leu Ser Gln Ser Gln 1265 1270 1275 1280			35
Glu Glu Ala Glu Ala Leu Trp Ala Gln Leu Arg Ala Ala His Met Ile 1285 1290 1295			37
Thr Glu Asp Gly Lys Val Cys Ala Leu Ser Gly Lys Arg Gly Thr Ser 1300 1305 1310			39
His Gln Asn Leu Thr Ile Gln Leu Ile Ser Leu Leu Ala Ser Ile Pro 1315 1320 1325			41
Arg Leu Pro Pro His Val Leu Gly Ile Thr Thr Asp Arg Tyr Asp Ser 1330 1335 1340			43
Asp Pro Gln Gln Gln His Gln Gln Thr Val Ser Phe Ser Asp Gly Phe			45
			47

RO 121280 B1

1	1345	1350	1355	1360
3	Gly Arg Leu Leu Gln Ser Ser Ala Arg His Glu Ser Gly Asp Ala Trp	1365	1370	1375
5	Gln Arg Lys Glu Asp Gly Gly Leu Val Val Asp Ala Asn Gly Val Leu	1380	1385	1390
7	Val Ser Ala Pro Thr Asp Thr Arg Trp Ala Val Ser Gly Arg Thr Glu	1395	1400	1405
9	Tyr Asp Asp Lys Gly Gln Pro Val Arg Thr Tyr Gln Pro Tyr Phe Leu	1410	1415	1420
11	Asn Asp Trp Arg Tyr Val Ser Asp Asp Ser Ala Arg Asp Asp Leu Phe	1425	1430	1435
13	Ala Asp Thr His Leu Tyr Asp Pro Leu Gly Arg Glu Tyr Lys Val Ile	1445	1450	1455
15	Thr Ala Lys Lys Tyr Leu Arg Glu Lys Leu Tyr Thr Pro Trp Phe Ile	1460	1465	1470
17	Val Ser Glu Asp Glu Asn Asp Thr Ala Ser Arg Thr Pro *	1475	1480	1485
21	(2) INFORMAȚIA PENTRU SECV. ID. NR. 33:			
	(i) CARACTERISTICI DE SECVENȚĂ:			
23	(A) LUNGIME: 3288 baze perechi			
	(B) TIP: acid. nucleic			
25	(C) ÎMPLETIRE: dublă			
	(D) TOPOLOGIE: lineară			
27	(ii) TIPUL MOLECULEI: ADN (genomic)			
	(xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 33:			
29	ATG GTG ACT GTT ATG CAA AAT AAA ATA TCA TTT TTA TCA GGT ACA TCC 48			
31	Met Val Thr Val Met Gln Asn Lys Ile Ser Phe Leu Ser Gly Thr Ser	5	10	15
33	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
35	TCA ATC AGC CGG GCT ACT TTC GTT CAA TCC GTT CCC ACC CTG CCC GTT 144			
37	Ser Ile Ser Arg Ala Thr Phe Val Gln Ser Val Pro Thr Leu Pro Val	35	40	45
39	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
41	CTG AAA TCC CTC TAC CGA GCC TGG CAA TTG CGT CAG GAG CCG GTT ATT 240			
43	Leu Lys Ser Leu Tyr Arg Ala Trp Gln Leu Arg Gln Glu Pro Val Ile	65	70	80
45	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
47	GAT GCT TTC 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			

RO 121280 B1

100	105	110		1
ATG AAC CGT GCC AGT CAA TAT GCT GAC GCT GCC TCT ATT CAA TCC CTA 384				3
Met Asn Arg Ala Ser Gln Tyr Ala Asp Ala Ala Ser Ile Gln Ser Leu	115	120	125	
				5
TTT TCA CCG GCC CGT TAT GCT TCC GCA CTC TAC AGA GTT GCT AAA GAT 432				7
Phe Ser Pro Gly Arg Tyr Ala Ser Ala Leu Tyr Arg Val Ala Lys Asp	130	135	140	
				9
CTG CAT AAA TCA GAT TCC AGT TTG CAT ATT GAT AAT CGC CGC GCT GAT 480				11
Leu His Lys Ser Asp Ser Ser Leu His Ile Asp Asn Arg Arg Ala Asp	145	150	155	160
				13
CTG AAG GAT CTG ATA TTA AGC GAA ACG ACG ATG AAT AAA GAG GTC ACT 528				15
Leu Lys Asp Leu Ile Leu Ser Glu Thr Thr Met Asn Lys Glu Val Thr	165	170	175	
				17
TCC CTT GAT ATC TTG TTG GAT GTO CTA CAA AAA GGC GGT AAA GAT ATT 576				19
Ser Leu Asp Ile Leu Leu Asp Val Leu Gln Lys Gly Gly Lys Asp Ile	180	185	190	
				21
ACT GAG CTG TCC GGC GCA TTC TTC CCA ATG ACG TTA CCT TAT GAC GAT 624				23
Thr Glu Leu Ser Gly Ala Phe Phe Pro Met Thr Leu Pro Tyr Asp Asp	195	200	205	
				25
CAT CTG TCG CAA ATC GAT TCC GCT TTA TCG GCA CAA GCC AGA ACG CTG 672				27
His Leu Ser Gln Ile Asp Ser Ala Leu Ser Ala Gln Ala Arg Thr Leu	210	215	220	
				29
AAC GGT GTG TGG AAT ACT TTG ACA GAT ACC ACG GCA CAA GCG GTT TCA 720				31
Asn Gly Val Trp Asn Thr Leu Thr Asp Thr Thr Ala Gln Ala Val Ser	225	230	235	240
				33
GAA CAA ACC AGT AAT ACG AAT ACA CGC AAA CTG TTC GCT GCC CAA GAT 768				35
Glu Gln Thr Ser Asn Thr Asn Thr Arg Lys Leu Phe Ala Ala Gln Asp	245	250	255	
				37
GGT AAT CAA GAT ACA TTT TTT TCC GGA AAC ACT TTT TAT TTC AAA GCG 816				39
Gly Asn Gln Asp Thr Phe Phe Ser Gly Asn Thr Phe Tyr Phe Lys Ala	260	265	270	
				41
GTG GGA TTC AGC GGG CAA CCT ATG GTT TAC CTG TCA CAG TAC ACC AGC 864				43
Val Gly Phe Ser Gly Gln Pro Met Val Tyr Leu Ser Gln Tyr Thr Ser	275	280	285	
				45
GGG AAC GGC ATT GTC GGC GCA CAA TTG ATT GCA GGT AAT CCA GAC CAA 912				47
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	

RO 121280 B1

1	ACT	CAG	CCT	TTG	CCA	AGC	TTC	CAT	CTG	CCG	GTC	ACA	CTG	GAA	CAC	AGC	1152
3	Thr	Gln	Pro	Leu	Pro	Ser	Phe	His	Leu	Pro	Val	Thr	Leu	Glu	His	Ser	
	370						375					380					
5	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	
7	GTA	GAT	AGT	TCC	GGA	CAG	TCA	AAT	TCG	AAA	AAC	GCG	CTG	GTT	ATC	AAT	1248
	Val	Asp	Ser	Ser		Gln	Ser	Asn	Trp	Lys	Asn	Ala	Leu	Val	Ile	Asn	
					405					410					415		
9	GGG	ACA	AAA	GAC	AAG	GGG	CTG	TTA	TTA	ACC	TTT	TGC	AGC	GAT	AGC	TCA	1296
11	Gly	Thr	Lys	Asp	Lys	Gly	Leu	Leu	Leu	Thr	Phe	Cys	Ser	Asp	Ser	Ser	
				420					425					430			
13	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				
15	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	
17				450			455					460					
19	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	
21	CAT	TAT	GGT	TTT	AAC	GGC	GCT	AGC	TTA	CCG	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		
23	AGC	GAG	TTG	ACC	AGC	AAA	CTG	AAT	TCT	ATC	GAT	ACT	TTC	TGT	GAG	AAG	1536
25	Ser	Glu	Leu	Thr	Ser	Lys	Leu	Asn	Ser	Ile	Asp	Thr	Phe	Cys	Glu	Lys	
				500					505					510			
27	ACC	CCG	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				
29	TAC	AGT	CAA	AGC	AGC	ATT	GAT	GCG	AAA	GCA	GCC	AGC	CGC	TAT	GTT	CGT	1632
31	Tyr	Ser	Gln	Ser	Ser	Ile	Asp	Ala	Lys	Ala	Ala	Ser	Arg	Tyr	Val	Arg	
		530					535					540					
33	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	
35	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	
37					565					570					575		
39	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			
41	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				
43	CTA	TCA	TTT	GAA	GAA	TTC	GAC	TCG	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	
45			610				615					620					
47	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	

RO 121280 B1

525											535											540	1
GCA	CTG	GCA	GAG	TAT	GTC	AGC	CTA	AAA	CAG	CGC	TAT	GGG	CTT	GAT	GCC	1353	3						
Ala	Leu	Ala	Glu	Tyr	Val	Ser	Leu	Lys	Gln	Arg	Tyr	Gly	Leu	Asp	Ala								
				645					650					655				5					
AAT	ACC	TTT	GCG	ACC	TTC	ATT	AGT	GCA	GTA	AAT	CCT	TAT	ACG	CCA	GAT	2015	7						
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	9						
Gln	Thr	Pro	Ser	Phe	Tyr	Glu	Thr	Ala	Phe	Arg	Ser	Ala	Asp	Gly	Asn								
				675					680					685									
CAT	GTC	ATT	GCG	CTA	GGT	ACA	GAG	GTG	AAA	TAT	GCA	GAA	AAT	GAG	CAG	2112	11						
His	Val	Ile	Ala	Leu	Gly	Thr	Glu	Val	Lys	Tyr	Ala	Glu	Asn	Glu	Gln								
				690					695					700				13					
GAT	GAG	TTA	GCC	GCC	ATA	TGC	TGC	AAA	GCA	TTG	GGT	GTC	ACC	AGT	GAT	2160	15						
Asp	Glu	Leu	Ala	Ala	Ile	Cys	Cys	Lys	Ala	Leu	Gly	Val	Thr	Ser	Asp								
				705					710					715									
GAA	CTG	CTC	CGT	ATT	GGT	CGC	TAT	TGC	TTC	GGT	AAT	GCA	GGC	AGT	TTT	2208	17						
Glu	Leu	Leu	Arg	Ile	Gly	Arg	Tyr	Cys	Phe	Gly	Asn	Ala	Gly	Ser	Phe								
				725					730					735				19					
ACC	TTG	GAT	GAA	TAT	ACC	GCC	AGT	CAG	TTG	TAT	CGC	TTC	GCC	GCC	ATT	2256	21						
Thr	Leu	Asp	Glu	Tyr	Thr	Ala	Ser	Gln	Leu	Tyr	Arg	Phe	Gly	Ala	Ile								
				740					745					750									
CCC	CGT	TTG	TTT	GCG	CTG	ACA	TTT	GCC	CAA	GCC	GAA	ATT	TTA	TGG	CGT	2304	23						
Pro	Arg	Leu	Phe	Gly	Leu	Thr	Phe	Ala	Gln	Ala	Glu	Ile	Leu	Trp	Arg								
				755					760					765				25					
CTG	ATG	GAA	GGC	GGA	AAA	GAT	ATC	TTA	TTG	CAA	CAG	TTA	GGT	CAG	GCA	2352	27						
Leu	Met	Glu	Gly	Gly	Lys	Asp	Ile	Leu	Leu	Gln	Gln	Leu	Gly	Gln	Ala								
				770					775					780									
AAA	TCC	CTG	CAA	CCA	CTG	GCT	ATT	TTA	CGC	CGT	ACC	GAG	CAG	CTC	CTG	2400	29						
Lys	Ser	Leu	Gln	Pro	Leu	Ala	Ile	Leu	Arg	Arg	Thr	Glu	Gln	Val	Leu								
				785					790					795									
GAT	TGG	ATG	TCG	TCC	GTA	AAT	CTA	AGT	CTG	ACT	TAT	CTG	CAA	GGG	ATG	2448	31						
Asp	Trp	Met	Ser	Ser	Val	Asn	Leu	Ser	Leu	Thr	Tyr	Leu	Gln	Gly	Met								
				805					810					815				33					
GTA	AGT	ACG	CAA	TGG	AGC	GGT	ACC	GCC	ACC	GCT	GAG	ATG	TTC	AAT	TTC	2496	35						
Val	Ser	Thr	Gln	Trp	Ser	Gly	Thr	Ala	Thr	Ala	Glu	Met	Phe	Asn	Phe								
				820					825					830									
TTG	GAA	AAC	GTT	TGT	GAC	AGC	GTG	AAT	AGT	CAA	GCT	GCC	ACT	AAA	GAA	2544	37						
Leu	Glu	Asn	Val	Cys	Asp	Ser	Val	Asn	Ser	Gln	Ala	Ala	Thr	Lys	Glu								
				835					840					845				39					
ACA	ATG	GAT	TCG	GCG	TTA	CAG	CAG	AAA	GTG	CTG	CGG	GCG	CTA	AGC	GCC	2592	41						
Thr	Met	Asp	Ser	Ala	Leu	Gln	Gln	Lys	Val	Leu	Arg	Ala	Leu	Ser	Ala								
				850					855					860									
GGT	TTC	GCG	ATT	AAG	AGC	AAT	GTG	ATG	GGT	ATC	GTC	ACC	TTC	TGG	CTG	2640	43						
Gly	Phe	Gly	Ile	Lys	Ser	Asn	Val	Met	Gly	Ile	Val	Thr	Phe	Trp	Leu								
				865					870					875				45					
GAG	AAA	ATC	ACA	ATC	GGT	AGT	GAT	AAT	CCT	TTT	ACA	TTG	GCA	AAC	TAC	2688	47						
Glu	Lys	Ile	Thr	Ile	Gly	Ser	Asp	Asn	Pro	Phe	Thr	Leu	Ala	Asn	Tyr								
				885					890					895									

RO 121280 B1

1	TGG GAT GAT ATT TAA ACC CTG TTT AGC CAT GAC AAT GCC ACG TTA GAG	2716
3	Trp His Asp Ile Gln Thr Leu Phe Ser His Asp Asn Ala Thr Leu Glu	
	900 905 910	
5	TGC TTA CAA ACC GAC ACT TCT CTG GTA ATT GCT ACT CAG CAA CTT AGC	2784
	Ser Leu Gln Thr Asp Thr Ser Leu Val Ile Ala Thr Gln Gln Leu Ser	
	915 920 925	
7	CAG CTA GTG TTA ATT GTG AAA TGG CTG AGC CTG ACC GAG CAG GAT CTG	2832
9	Gln Leu Val Leu Ile Val Lys Trp Leu Ser Leu Thr Glu Gln Asp Leu	
	930 935 940	
11	CAA TTA CTG ACA ACC TAT CCC GAA CGT TTA ATC AAC GGC ATC ACG AAT	2880
	Gln Leu Leu Thr Thr Tyr Pro Glu Arg Leu Ile Asn Gly Ile Thr Asn	
	945 950 955 960	
13	GTT CCT GTA CCC AAT CCG GAG CTA TTA CTC ACG CTA TCA CGT TTT AAG	2928
15	Val Pro Val Pro Asn Pro Glu Leu Leu Leu Thr Leu Ser Arg Phe Lys	
	965 970 975	
17	CAG TGG GAA ACT CAA GTC ACC GTT TCC CGT GAT GAA GCG ATG CCG TGT	2976
	Gln Trp Glu Thr Gln Val Thr Val Ser Arg Asp Glu Ala Met Arg Cys	
	980 985 990	
19	TTC GAT CAA TTA AAT GCC AAT GAT ATG ACG ACT GAA AAT GCA GGT TCA	3024
21	Phe Asp Gln Leu Asn Ala Asn Asp Met Thr Thr Glu Asn Ala Gly Ser	
	995 1000 1005	
23	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	
25	AAT ACC TTG CTA TTA GGT GAA AAT AAC TGG CCG AAA AGT TTT ACC TCT	3120
27	Asn Thr Leu Leu Leu Gly Glu Asn Asn Trp Pro Lys Ser Phe Thr Ser	
	1025 1030 1035 1040	
29	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	
31	GGT AGT ACC ACT CTG GGC AAT CTG TTG TCC ATG ATG CAA GCA GAC CCT	3216
33	Gly Ser Thr Thr Leu Gly Asn Leu Leu Ser Met Met Gln Ala Asp Pro	
	1060 1065 1070	
35	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	
37	GCC GCA ATC AGC AAT CGT CAG TAA	3285
39	Ala Ala Ile Ser Asn Arg Gln ...	
	1090 1095	

- 41 (2) INFORMAȚIA PENTRU SECV. ID. NR. 34:
- 43 (i) CARACTERISTICI DE SECVENȚĂ:
- 45 (A) LUNGIME: 1095 aminoacizi
- 45 (B) TIP: aminoacid
- 45 (C) TOPOLOGIE: lineară
- 47 (ii) TIPUL MOLECULEI: proteină
- 47 (xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 34:

RO 121280 B1

Trăsătura	De la	Până la	Descrierea		
	254	267	SECV. ID. NR. 15	1	
	254	492	peptida Tca _{ii}	3	
Met Val Thr Val Met Gln Asn Lys Ile Ser Phe Leu Ser Gly Thr Ser	1	5	10	15	5
Glu Gln Pro Leu Leu Asp Ala Gly Tyr Gln Asn Val Phe Asp Ile Ala	20	25	30	35	7
Ser Ile Ser Arg Ala Thr Phe Val Gln Ser Val Pro Thr Leu Pro Val	35	40	45	50	9
Lys Glu Ala His Thr Val Tyr Arg Gln Ala Arg Gln Arg Ala Glu Asn	50	55	60	65	11
Leu Lys Ser Leu Tyr Arg Ala Trp Gln Leu Arg Gln Glu Pro Val Ile	65	70	75	80	13
Lys Gly Leu Ala Lys Leu Asn Leu Gln Ser Asn Val Ser Val Leu Gln	85	90	95	100	15
Asp Ala Leu Val Glu Asn Ile Gly Gly Asp Gly Asp Phe Ser Asp Leu	100	105	110	115	17
Met Asn Arg Ala Ser Gln Tyr Ala Asp Ala Ala Ser Ile Gln Ser Leu	115	120	125	130	19
Phe Ser Pro Gly Arg Tyr Ala Ser Ala Leu Tyr Arg Val Ala Lys Asp	130	135	140	145	21
Leu His Lys Ser Asp Ser Ser Leu His Ile Asp Asn Arg Arg Ala Asp	145	150	155	160	23
Leu Lys Asp Leu Ile Leu Ser Glu Thr Thr Met Asn Lys Glu Val Thr	165	170	175	180	25
Ser Leu Asp Ile Leu Leu Asp Val Leu Gln Lys Gly Gly Lys Asp Ile	180	185	190	195	27
Thr Glu Leu Ser Gly Ala Phe Phe Pro Met Thr Leu Pro Tyr Asp Asp	195	200	205	210	29
His Leu Ser Gln Ile Asp Ser Ala Leu Ser Ala Gln Ala Arg Thr Leu	210	215	220	225	31
Asn Gly Val Trp Asn Thr Leu Thr Asp Thr Thr Ala Gln Ala Val Ser	225	230	235	240	33
Glu Gln Thr Ser Asn Thr Asn Thr Arg Lys Leu Phe Ala Ala Gln Asp	245	250	255	260	35
Gly Asn Gln Asp Thr Phe Phe Ser Gly Asn Thr Phe Tyr Phe Lys Ala	260	265	270	275	37
Val Gly Phe Ser Gly Gln Pro Met Val Tyr Leu Ser Gln Tyr Thr Ser	275	280	285	290	39
Gly Asn Gly Ile Val Gly Ala Gln Leu Ile Ala Gly Asn Pro Asp Gln	290	295	300	305	41
Ala Ala Ala Ala Ile Val Ala Pro Leu Lys Leu Thr Trp Ser Met Ala	305	310	315	320	43
Lys Gln Cys Tyr Tyr Leu Val Ala Pro Asp Gly Thr Thr Met Gly Asp	325	330	335	340	45
Gly Asn Val Leu Thr Gly Cys Phe Leu Arg Gly Asn Ser Pro Thr Asn					47

RO 121280 B1

1		340		345		350	
3	Pro Asp Lys Asp Gly Ile Phe	Ala Gln Val Ala Asn Lys Ser Gly Ser	355	360	365		
5	Thr Gln Pro Leu Pro Ser Phe His Leu Pro Val Thr Leu Glu His Ser		370	375	380		
7	Glu Asn Lys Asp Gln Tyr Tyr Leu Lys Thr Glu Gln Gly Tyr Ile Thr		385	390	395	400	
9	Val Asp Ser Ser Gly Gln Ser Asn Trp Lys Asn Ala Leu Val Ile Asn		405	410	415		
11	Gly Thr Lys Asp Lys Gly Leu Leu Leu Thr Phe Cys Ser Asp Ser Ser		420	425	430		
13	Gly Thr Pro Thr Asn Pro Asp Asp Val Ile Pro Pro Ala Ile Asn Asp		435	440	445		
15	Ile Pro Ser Pro Pro Ala Arg Glu Thr Leu Ser Leu Thr Pro Val Ser		450	455	460		
17	Tyr Gln Leu Met Thr Asn Pro Ala Pro Thr Glu Asp Asp Ile Thr Asn		465	470	475	480	
19	His Tyr Gly Phe Asn Gly Ala Ser Leu Arg Ala Ser Pro Leu Ser Thr		485	490	495		
21	Ser Glu Leu Thr Ser Lys Leu Asn Ser Ile Asp Thr Phe Cys Glu Lys		500	505	510		
23	Thr Arg Leu Ser Phe Asn Gln Leu Met Asp Leu Thr Ala Gln Gln Ser		515	520	525		
25	Tyr Ser Gln Ser Ser Ile Asp Ala Lys Ala Ala Ser Arg Tyr Val Arg		530	535	540		
27	Phe Gly Glu Thr Thr Pro Thr Arg Val Asn Val Tyr Gly Ala Ala Tyr		545	550	555	560	
29	Leu Asn Ser Thr Leu Ala Asp Ala Ala Asp Gly Gln Tyr Leu Trp Ile		565	570	575		
31	Gln Thr Asp Gly Lys Ser Leu Asn Phe Thr Asp Asp Thr Val Val Ala		580	585	590		
33	Leu Ala Gly Arg Ala Glu Lys Leu Val Arg Leu Ser Ser Gln Thr Gly		595	600	605		
35	Leu Ser Phe Glu Glu Leu Asp Trp Leu Ile Ala Asn Ala Ser Arg Ser		610	615	620		
37	Val Pro Asp His His Asp Lys Ile Val Leu Asp Lys Pro Val Leu Glu		625	630	635	640	
39	Ala Leu Ala Glu Tyr Val Ser Leu Lys Gln Arg Tyr Gly Leu Asp Ala		645	650	655		
41	Asn Thr Phe Ala Thr Phe Ile Ser Ala Val Asn Pro Tyr Thr Pro Asp		660	665	670		
43	Gln Thr Pro Ser Phe Tyr Glu Thr Ala Phe Arg Ser Ala Asp Gly Asn		675	680	685		
45	His Val Ile Ala Leu Gly Thr Glu Val Lys Tyr Ala Glu Asn Glu Gln						

RO 121280 B1

690	695	700	1
Asp Glu Leu Ala Ala Ile Cys Cys Lys Ala Leu Gly Val Thr Ser Asp			3
705	710	715	720
Glu Leu Leu Arg Ile Gly Arg Tyr Cys Phe Gly Asn Ala Gly Ser Phe			5
	725	730	735
Thr Leu Asp Glu Tyr Thr Ala Ser Gln Leu Tyr Arg Phe Gly Ala Ile			7
	740	745	750
Pro Arg Leu Phe Gly Leu Thr Phe Ala Gln Ala Glu Ile Leu Trp Arg			9
	755	760	765
Leu Met Glu Gly Gly Lys Asp Ile Leu Leu Gln Gln Leu Gly Gln Ala			11
	770	775	780
Lys Ser Leu Gln Pro Leu Ala Ile Leu Arg Arg Thr Glu Gln Val Leu			13
	785	790	800
Asp Trp Met Ser Ser Val Asn Leu Ser Leu Thr Tyr Leu Gln Gly Met			15
	805	810	815
Val Ser Thr Gln Trp Ser Gly Thr Ala Thr Ala Glu Met Phe Asn Phe			17
	820	825	830
Leu Glu Asn Val Cys Asp Ser Val Asn Ser Gln Ala Ala Thr Lys Glu			19
	835	840	845
Thr Met Asp Ser Ala Leu Gln Gln Lys Val Leu Arg Ala Leu Ser Ala			21
	850	855	860
Gly Phe Gly Ile Lys Ser Asn Val Met Gly Ile Val Thr Phe Trp Leu			23
	865	870	875
Glu Lys Ile Thr Ile Gly Ser Asp Asn Pro Phe Thr Leu Ala Asn Tyr			25
	885	890	895
Trp His Asp Ile Gln Thr Leu Phe Ser His Asp Asn Ala Thr Leu Glu			27
	900	905	910
Ser Leu Gln Thr Asp Thr Ser Leu Val Ile Ala Thr Gln Gln Leu Ser			29
	915	920	925
Gln Leu Val Leu Ile Val Lys Trp Leu Ser Leu Thr Glu Gln Asp Leu			31
	930	935	940
Gln Leu Leu Thr Thr Tyr Pro Glu Arg Leu Ile Asn Gly Ile Thr Asn			33
	945	950	955
Val Pro Val Pro Asn Pro Glu Leu Leu Leu Thr Leu Ser Arg Phe Lys			35
	965	970	975
Gln Trp Glu Thr Gln Val Thr Val Ser Arg Asp Glu Ala Met Arg Cys			37
	980	985	990
Phe Asp Gln Leu Asn Ala Asn Asp Met Thr Thr Glu Asn Ala Gly Ser			39
	995	1000	1005
Leu Ile Ala Thr Leu Tyr Glu Met Asp Lys Gly Thr Gly Ala Gln Val			41
	1010	1015	1020
Asn Thr Leu Leu Leu Gly Glu Asn Asn Trp Pro Lys Ser Phe Thr Ser			43
	1025	1030	1035
Leu Trp Gln Leu Leu Thr Trp Leu Arg Val Gly Gln Arg Leu Asn Val			45
			47

RO 121280 B1

1
3
5
7
9
11
13
15
17
19
21
23
25
27
29
31
33
35
37
39
41
43
45
47

1045
1050
1055
1060
1065
1070
1075
1080
1085
1090
1095

Gly Ser Thr Thr Leu Gly Asn Leu Leu Ser Met Met Gln Ala Asp Pro
Ala Ala Glu Ser Ser Ala Leu Leu Ala Ser Val Ala Gln Asn Leu Ser
Ala Ala Ile Ser Asn Arg Gln ...

(2) INFORMAȚIA PENTRU SECV. ID. NR. 35:
(i) CARACTERISTICI DE SECVENȚĂ:
(A) LUNGIME: 603 aminoacizi
(B) TIP: aminoacid
(C) TOPOLOGIE: lineară
(ii) TIPUL MOLECULEI: proteină
(xi) DESCRIEREA SECVENȚEI SECV ID. NR. 35:

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

RO 121280 B1

Val Thr Ser Asp Glu Leu Leu Arg Ile Gly Arg Tyr Cys Phe Gly Asn 225 230 235 240	1
Ala Gly Arg Phe Thr Leu Asp Glu Tyr Thr Ala Ser Gln Leu Tyr Arg 245 250 255	3
Phe Gly Ala Ile Pro Arg Leu Phe Gly Leu Thr Phe Ala Gln Ala Glu 260 265 270	5
Ile Leu Trp Arg Leu Met Glu Gly Gly Lys Asp Ile Leu Leu Gln Gln 275 280 285	7
Xxx Gly Gln Ala Lys Ser Leu Gln Pro Leu Ala Ile Leu Arg Arg Thr 290 295 300	9
Glu Gln Val Leu Asp Trp Met Ser Pro Val Asn Leu Ser Leu Thr Tyr 305 310 315 320	11
Leu Gln Gly Met Val Ser Thr Gln Trp Ser Gly Thr Ala Thr Ala Glu 325 330 335	13
Met Phe Asn Phe Leu Glu Asn Val Cys Asp Ser Val Asn Ser Gln Ala 340 345 350	15
Xxx Thr Lys Glu Thr Met Asp Ser Ala Leu Gln Gln Lys Val Leu Arg 355 360 365	17
Ala Leu Ser Ala Gly Phe Gly Ile Lys Ser Asn Val Met Gly Ile Val 370 375 380	19
Thr Phe Trp Leu Glu Lys Ile Thr Ile Gly Arg Asp Asn Pro Phe Thr 385 390 395 400	21
Leu Ala Asn Tyr Trp His Asp Ile Gln Thr Leu Phe Ser His Asp Asn 405 410 415	23
Ala Thr Leu Glu Ser Leu Gln Thr Asp Thr Ser Leu Val Ile Ala Thr 420 425 430	25
Gln Gln Leu Ser Gln Leu Val Leu Ile Val Lys Trp Val Ser Leu Thr 435 440 445	27
Glu Gln Asp Leu Gln Leu Leu Thr Thr Tyr Pro Glu Arg Leu Ile Asn 450 455 460	29
Gly Ile Thr Asn Val Pro Val Pro Asn Pro Glu Leu Leu Leu Thr Leu 465 470 475 480	31
Ser Arg Phe Lys Gln Trp Glu Thr Gln Val Thr Val Ser Arg Asp Glu 485 490 495	33
Ala Met Arg Cys Phe Asp Gln Leu Asn Ala Asn Asp Met Thr Thr Glu 500 505 510	35
Asn Ala Gly Ser Leu Ile Ala Thr Leu Tyr Glu Met Asp Lys Gly Thr 515 520 525	37
Gly Ala Gln Val Asn Thr Leu Leu Leu Gly Glu Asn Asn Trp Pro Lys 530 535 540	39
Ser Phe Thr Ser Leu Trp Gln Leu Leu Thr Trp Leu Arg Val Gly Gln 545 550 555 560	41
Arg Leu Asn Val Gly Ser Thr Thr Leu Gly Asn Leu Leu Ser Met Met 565 570 575	43
	45
	47

RO 121280 B1

Gln Ala Asp Pro Ala Ala Glu Ser Ser Ala Leu Leu Ala Ser Val Ala
530 585 590

Gln Asn Leu Ser Ala Ala Ile Ser Asn Arg Gln •
595 600

(2) INFORMAȚIA PENTRU SECV. ID. NR. 36:

(i) CARACTERISTICI DE SECVENȚĂ:

(A) LUNGIME: 2557 baze perechi

(B) TIP: acid nucleic

(C) TOPOLOGIE: lineară

(ii) TIPUL MOLECULEI: ADN (genomic)

(xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 36:

13	GAATTCGGCT TCGTTTAAT ATTGATGATG TCTCGCTCTT CCGCCTGCTT AAAATTACCG	60
15	ACCATGATAA TAAAGATGGA AAAATTAAAA ATAACCTAAA GAATCTTTCC AATTTATATA	120
17	TTGGAAAAAT ACTGCCAGAT ATTCATCAAT TAACCATTGA TGAAGTGGAT TTATTACTGA	180
19	TTGCCGTAGG TGAAGGAAAA ACTAATTTAT CCGCTATCAG TGATAAGCAA TTGGCTACCC	240
21	TGATCAGAAA ACTCAATACT ATTACCAGCT GGCTACATAC ACAGAAGTGG AGTGTATTC	300
23	AGCTATTTAT CATGACCTCC ACCAGCTATA ACAAACCGCT AACGCCTGAA ATTAAGAATT	360
25	TGCTGGATAC CGTCTACCAC GGTTTACAAG GTTTTGATAA AGACAAAGCA GATTTGCTAC	420
27	ATGTCATGGC GCCCTATATT GCGGCCACCT TGCAATATC ATCGGAAAAAT GTCGCCCACT	480
29	CGTACTCCT TTGGGCAGAT AAGTTACAGC CCGCGGACGG CGCAATGACA GCAGAGGGAN	540
31	TCTGGGACTG GTTGAATACT AAGTATACGC CCGGTTTCATC GGAAGCCGTA GAAACGCAGG	600
33	AACATATCGT TCAGTATTGT CAGGCTCTGG CACAATTGGA AATGGTTTAC CATTCCACCG	660
35	GCATCAACGA AAACGCCTTC CGTCTATTTG TGACAAAACC AGAGATGTTT GCGCTGCAA	720
37	CTGGAGCAGC GCCCGCGCAT GATGCCCTTT CACTGATTAT GCTGACACGT TTTGCGGATT	780
39	GGGTGAACGC ACTAGGCGAA AAAGCGTCTT CCGTGCTAGC GGCATTTGAA GCTAACTCGT	840
41	TAACGGCAGA ACAACTGGCT GATGCCATGA ATCTTGATGC TAATTTGCTG TTGCAAGCCA	900
43	GTATTCAAGC ACAAATCAT CAACATCTTC CCCAGTAAC TCCAGAAAAT GCGTTCTCCT	960
45	GTTGGACATC TATCAATACT ATCCTGCAAT GGGTTAATGT CGCACAACAA TTGAAATGTC	1020
47	GCCCCACAGG GCGTTTCCGC TTTGGTCCGG CTGGATTATA TTCAATCAAT GAAAGAGACA	1080
49	CCGACCTATG CCCAGTGGGA AAACCGCGCA GCGTATTAA CCGCCGGGT GAATTCACAA	1140
51	ACAGGCTAAT ACATTACAAC GCTTTCTGGA ATGAATCTCG CAGTGCCGCA TTAAGCACCT	1200
53	ACTATATCCG TCAAGTCGCC AAGGCAGCGG CCGCTATTAA AAGCCGTGAT GACTTGTATC	1260
55	AATACTTACT GATTGATAAT CAGGTTTCTG CCGCAATAAA AACCACCCGG ATCGCCGAAG	1320
57	CCATTGCCAG TATTCAACTG TACGTCAACC GGGCATTGGA AAATGTGGAA GAAAATGCCA	1380
59	ATTCGGGGGT TATCAGCCGC CAATTCTTTA TCGACTGGGA CAAATACAAT AAACGCTACA	1440
61	GCACTTGGGC GGGTGTTCCT CAATTAGTTT ACTACCCGGA AAATATATAT GATCCGACCA	1500
63	TGCGTATCGG ACAAACCAA ATGATGGACG CATTACTGCA ATCCGTCAGC CAAAGCCAA	1560
65	TAAACGCCGA TACCGTCGAA GATGCCCTTA TGTCTTATCT GACATCGTTT GAACAAGTGG	1620
67	CTAATCTTAA AGTTATTAGC GCATATCAGC ATAATATTAA TAACGATCAA GCGCTGACCT	1680
69	ATTTTATCGG ACTCAGTGAA ACTGATGCCG GTGAATATTA TTGGCGCAGT GTCGATCACA	1740
71	GTAAATTCAA CGACGGTAAA TTCCGGGCTA ATGCCGAGG TGAATGGCAT AAAATGATT	1800
73	GTCCAATTAA CCCTTATAAA AGCACTATCC GTCCAGTGAT ATATAAATCC CCGCTGTATC	1860

RO 121280 B1

TGCTCTGGTT GGAACAAAAG GAGATCACCA AACAGACAGG AAATAGTAA GATGGCTATC	1920	1
AACTGAAAC GGATTATCGT TATGAACTAA AATTGGCGCA TATCCGCTAT GATGGCACTT	1980	3
CGAATACGCC AATCACCTTT GATGTCAATA AAAAAATATC CGAGCTAAAA CTGGAaaaaa	2040	
ATAGAGCGCC CGGACTCTAT TGTGCCGGTT ATCAAGGTGA AGATACGTTG CTGGTGATGT	2100	5
TTTATAACCA ACAAGACACA CTAGATAGTT ATAAAAACGC TTCAATGCAA GGAATATATA	2160	
TCTTTGCTGA TATGGCATCC AAAGATATGA CCCCAGAACA GAGCAATGTT TATCGGGATA	2220	7
ATAGCTATCA ACAATTTGAT ACCAATAATG TCAGAAGAGT GAATAACCGC TATGCAGAGG	2280	
ATTATGAGAT TCCTTCTTCG GTAAGTAGCC GTAAAGACTA TGGTTGGGGA GATTATTACC	2340	9
TCAGCATGGT ATATAACGGA GATATTCCAA CTATCAATTA CAAAGCCGCA TCAAGTGATT	2400	
TAAAAATTTA TATTTCACCA AAATTAAGAA TTATTCATAA TGGATATGAA GGACAGAAGC	2460	11
GCAATCAATG CAATTTGATG AATAAATATG GCAAACCTAGG TGATAAATTT ATTGTGTATA	2520	13
CCAGCCTGGG CGTTAATCCG AATAATAAGC CGAATTC	2557	

(2) INFORMAȚIA PENTRU SECV. ID. NR. 37:

(i) CARACTERISTICI DE SECVENȚĂ:

(A) LUNGIME: 845 aminoacizi

(B) TIP: aminoacid

(C) TOPOLOGIE: lineară

(ii) TIPUL MOLECULEI: proteină (parțial)

(xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 37:

Ala Phe Asn Ile Asp Asp Val Ser Leu Phe Arg Leu Leu Lys Ile Thr	1 5 10 15	23
Asp His Asp Asn Lys Asp Gly Lys Ile Lys Asn Asn Leu Lys Asn Leu	20 25 30	25
Ser Asn Leu Tyr Ile Gly Lys Leu Leu Ala Asp Ile His Gln Leu Thr	35 40 45	27
Ile Asp Glu Leu Asp Leu Leu Leu Ile Ala Val Gly Glu Gly Lys Thr	50 55 60	29
Asn Leu Ser Ala Ile Ser Asp Lys Gln Leu Ala Thr Leu Ile Arg Lys	65 70 75 80	31
Leu Asn Thr Ile Thr Ser Trp Leu His Thr Gln Lys Trp Ser Val Phe	85 90 95	33
Gln Leu Phe Ile Met Thr Ser Thr Ser Tyr Asn Lys Thr Leu Thr Pro	100 105 110	35
Glu Ile Lys Asn Leu Leu Asp Thr Val Tyr His Gly Leu Gln Gly Phe	115 120 125	37
Asp Lys Asp Lys Ala Asp Leu Leu His Val Met Ala Pro Tyr Ile Ala	130 135 140	39
Ala Thr Leu Gln Leu Ser Ser Glu Asn Val Ala His Ser Val Leu Leu	145 150 155 160	41
Trp Ala Asp Lys Leu Gln Pro Gly Asp Gly Ala Met Thr Ala Glu Gly	165 170 175	43
Phe Trp Asp Trp Leu Asn Thr Lys Tyr Thr Pro Gly Ser Ser Glu Ala		45
		47

RO 121280 B1

	180	185	190
1			
3	Val Glu Thr Gln Glu His Ile Val Gln Tyr Cys Gln Ala Leu Ala Gln 195 200 205		
5	Leu Glu Met Val Tyr His Ser Thr Gly Ile Asn Glu Asn Ala Phe Arg 210 215 220		
7	Leu Phe Val Thr Lys Pro Glu Met Phe Gly Ala Ala Thr Gly Ala Ala 225 230 235 240		
9	Pro Ala His Asp Ala Leu Ser Leu Ile Met Leu Thr Arg Phe Ala Asp 245 250 255		
11	Trp Val Asn Ala Leu Gly Glu Lys Ala Ser Ser Val Leu Ala Ala Phe 260 265 270		
13	Glu Ala Asn Ser Leu Thr Ala Glu Gln Leu Ala Asp Ala Met Asn Leu 275 280 285		
15	Asp Ala Asn Leu Leu Leu Gln Ala Ser Ile Gln Ala Gln Asn His Gln 290 295 300		
17	His Leu Pro Pro Val Thr Pro Glu Asn Ala Phe Ser Cys Trp Thr Ser 305 310 315 320		
19	Ile Asn Thr Ile Leu Gln Trp Val Asn Val Ala Gln Gln Leu Lys Cys 325 330 335		
21	Arg Pro Thr Gly Arg Phe Arg Phe Gly Arg Ala Gly Leu Tyr Ser Ile 340 345 350		
23	Asn Glu Arg Asp Thr Asp Leu Cys Pro Val Gly Lys Arg Gly Arg Arg 355 360 365		
25	Ile Asn Arg Arg Val Glu Phe Asn Asn Arg Leu Ile His Tyr Asn Ala 370 375 380		
27	Phe Leu Asp Glu Ser Arg Ser Ala Ala Leu Ser Thr Tyr Tyr Ile Arg 385 390 395 400		
29	Gln Val Ala Lys Ala Ala Ala Ala Ile Lys Ser Arg Asp Asp Leu Tyr 405 410 415		
31	Gln Tyr Leu Leu Ile Asp Asn Gln Val Ser Ala Ala Ile Lys Thr Thr 420 425 430		
33	Arg Ile Ala Glu Ala Ile Ala Ser Ile Gln Leu Tyr Val Asn Arg Ala 435 440 445		
35	Leu Glu Asn Val Glu Glu Asn Ala Asn Ser Gly Val Ile Ser Arg Gln 450 455 460		
37	Phe Phe Ile Asp Trp Asp Lys Tyr Asn Lys Arg Tyr Ser Thr Trp Ala 465 470 475 480		
39	Gly Val Ser Gln Leu Val Tyr Tyr Pro Glu Asn Tyr Ile Asp Pro Thr 485 490 495		
41	Met Arg Ile Gly Gln Thr Lys Met Met Asp Ala Leu Leu Gln Ser Val 500 505 510		
43	Ser Gln Ser Gln Leu Asn Ala Asp Thr Val Glu Asp Ala Phe Met Ser 515 520 525		
45	Tyr Leu Thr Ser Phe Glu Gln Val Ala Asn Leu Lys Val Ile Ser Ala		
47			

RO 121280 B1

530	535	540	1
Tyr His Asp Asn Ile Asn Asn Asp Gln Gly Leu Thr Tyr Phe Ile Gly 545 550 555 560			3
Leu Ser Glu Thr Asp Ala Gly Glu Tyr Tyr Trp Arg Ser Val Asp His 565 570 575			5
Ser Lys Phe Asn Asp Gly Lys Phe Ala Ala Asn Ala Trp Ser Glu Trp 580 585 590			7
His Lys Ile Asp Cys Pro Ile Asn Pro Tyr Lys Ser Thr Ile Arg Pro 595 600 605			9
Val Ile Tyr Lys Ser Arg Leu Tyr Leu Leu Trp Leu Glu Gln Lys Glu 610 615 620			11
Ile Thr Lys Gln Thr Gly Asn Ser Lys Asp Gly Tyr Gln Thr Glu Thr 625 630 635 640			13
Asp Tyr Arg Tyr Glu Leu Lys Leu Ala His Ile Arg Tyr Asp Gly Thr 645 650 655			15
Trp Asn Thr Pro Ile Thr Phe Asp Val Asn Lys Lys Ile Ser Glu Leu 660 665 670			17
Lys Leu Glu Lys Asn Arg Ala Pro Gly Leu Tyr Cys Ala Gly Tyr Gln 675 680 685			19
Gly Glu Asp Thr Leu Leu Val Met Phe Tyr Asn Gln Gln Asp Thr Leu 690 695 700			21
Asp Ser Tyr Lys Asn Ala Ser Met Gln Gly Leu Tyr Ile Phe Ala Asp 705 710 715 720			23
Met Ala Ser Lys Asp Met Thr Pro Glu Gln Ser Asn Val Tyr Arg Asp 725 730 735			25
Asn Ser Tyr Gln Gln Phe Asp Thr Asn Asn Val Arg Arg Val Asn Asn 740 745 750			27
Arg Tyr Ala Glu Asp Tyr Glu Ile Pro Ser Ser Val Ser Ser Arg Lys 755 760 765			29
Asp Tyr Gly Trp Gly Asp Tyr Tyr Leu Ser Met Val Tyr Asn Gly Asp 770 775 780			31
Ile Pro Thr Ile Asn Tyr Lys Ala Ala Ser Ser Asp Leu Lys Ile Tyr 785 790 795 800			33
Ile Ser Pro Lys Leu Arg Ile Ile His Asn Gly Tyr Glu Gly Gln Lys 805 810 815			35
Arg Asn Gln Cys Asn Leu Met Asn Lys Tyr Gly Lys Leu Gly Asp Lys 820 825 830			37
Phe Ile Val Tyr Thr Ser Leu Gly Val Asn Pro Asn Asn 835 840 845			39
			41
			43
			45
			47

- 1 (2) INFORMAȚIA PENTRU SECV. ID. NR. 38:
 (i) CARACTERISTICI DE SECVENȚĂ:
 3 (A) LUNGIME: 16 aminoacizi
 (B) TIP: aminoacid
 5 (C) ÎMPLETIRE: simplă
 (D) TOPOLOGIE: lineară
 7 (ii) TIPUL MOLECULEI: proteină
 (v) TIPUL FRAGMENTULUI: N-terminal
 9 (xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 38:

11 Arg Tyr Tyr Asn Leu Ser Asp Glu Glu Leu Ser Gln Phe Ile Gly
 1 5 10 15
 Lys

- 13 (2) INFORMAȚIA PENTRU SECV. ID. NR. 39:
 (i) CARACTERISTICI DE SECVENȚĂ:
 15 (A) LUNGIME: 20 aminoacizi
 17 (B) TIP: aminoacid
 (C) ÎMPLETIRE: simplă
 19 (D) TOPOLOGIE: lineară
 (ii) TIPUL MOLECULEI: proteină
 21 (v) TIPUL FRAGMENTULUI: N-terminal
 (xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 39:

23 Gly Thr Ala Thr Asp Val Ser Gly Pro Val Glu Ile Asn Thr Ala
 1 5 10 15
 25 Ile Ser Pro Ala Lys
 20

- 27 (2) INFORMAȚIA PENTRU SECV. ID. NR. 40:
 (i) CARACTERISTICI DE SECVENȚĂ:
 29 (A) LUNGIME: 11 aminoacizi
 31 (B) TIP: aminoacid
 (C) ÎMPLETIRE: simplă
 33 (D) TOPOLOGIE: lineară
 (ii) TIPUL MOLECULEI: proteină
 35 (v) TIPUL FRAGMENTULUI: N-terminal
 (xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 40:

37 Ala Asn Ser Leu Tyr Ala Leu Phe Leu Pro Gln
 1 5 10

- 39 (2) INFORMAȚIA PENTRU SECV. ID. NR. 41:
 (i) CARACTERISTICI DE SECVENȚĂ:
 41 (A) LUNGIME: 14 aminoacizi
 43 (B) TIP: aminoacid
 (C) ÎMPLETIRE: simplă
 45 (D) TOPOLOGIE: lineară

RO 121280 B1

(ii) TIPUL MOLECULEI: proteină	1
(v) TIPUL FRAGMENTULUI: N-terminal	
(xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 41:	3
Leu Arg Ser Ala Asn Thr Leu Thr Asp Leu Phe Leu Pro Gln 1 5 10	5
(2) INFORMAȚIA PENTRU SECV. ID. NR. 42:	7
(i) CARACTERISTICI DE SECVENȚĂ:	
(A) LUNGIME: 19 aminoacizi	9
(B) TIP: aminoacid	
(C) ÎMPLETIRE: simplă	11
(D) TOPOLOGIE: lineară	
(ii) TIPUL MOLECULEI: proteină	13
(v) TIPUL FRAGMENTULUI: N-terminal	
(xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 42:	15
Arg Ala Leu Glu Val Glu Arg Thr Val Ser Leu Ala Glu Val Tyr 1 5 10 15	17
Ala Gly Leu Glu	19
(2) INFORMAȚIA PENTRU SECV. ID. NR. 43:	21
(i) CARACTERISTICI DE SECVENȚĂ:	
(A) LUNGIME: 11 aminoacizi	23
(B) TIP: aminoacid	
(C) ÎMPLETIRE: simplă	25
(D) TOPOLOGIE: lineară	
(ii) TIPUL MOLECULEI: proteină	27
(v) TIPUL FRAGMENTULUI: N-terminal	
(xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 43:	29
Ile Arg Glu Asp Tyr Pro Ala Ser Leu Gly Lys 1 5 10	31
(2) INFORMAȚIA PENTRU SECV. ID. NR. 44:	33
(i) CARACTERISTICI DE SECVENȚĂ:	
(A) LUNGIME: 16 aminoacizi	35
(B) TIP: aminoacid	
(C) ÎMPLETIRE: simplă	37
(D) TOPOLOGIE: lineară	
(ii) TIPUL MOLECULEI: proteină	39
(v) TIPUL FRAGMENTULUI: N-terminal	
(xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 44:	41
Asp Asp Ser Gly Asp Asp Asp Lys Val Thr Asn Thr Asp Ile His 1 5 10 15	43
Arg	45

RO 121280 B1

- 1 (2) INFORMAȚIA PENTRU SECV. ID. NR. 45:
 (i) CARACTERISTICI DE SECVENȚĂ:
 3 (A) LUNGIME: 13 aminoacizi
 (B) TIP: aminoacid
 5 (C) ÎMPLETIRE: simplă
 (D) TOPOLOGIE: lineară
 7 (ii) TIPUL MOLECULEI: proteină
 (v) TIPUL FRAGMENTULUI: N-terminal
 9 (xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 45:

11 Asp Val Xaa Gly Ser Glu Lys Ala Asn Glu Lys Leu Lys
 1 5 10

- 13 (2) INFORMAȚIA PENTRU SECV. ID. NR. 46:
 (i) CARACTERISTICI DE SECVENȚĂ:
 15 (A) LUNGIME: 7551 baze perechi
 (B) TIP: acid nucleic
 17 (C) ÎMPLETIRE: dublă
 (D) TOPOLOGIE: lineară
 19 (ii) TIPUL MOLECULEI: ADN (genomic)
 (xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 46: (tcdA)

23 ATG AAC GAG TCT GTA AAA GAG ATA CCT GAT GTA TTA AAA AGC CAG TCT 48
 Met Asn Glu Ser Val Lys Glu Ile Pro Asp Val Leu Lys Ser Gln Cys
 1 5 10 15
 25 GGT TTT AAT TGT CTG ACA GAT ATT ACC 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
 27 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
 31 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
 33 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

RO 121280 B1

GCC ATT CTC GCT GCC AAT GGT GAA CTG ATA GGC TAT AAC AAT GAA TTT	288	1
Ala Ile Leu Ala Pro Asn Ala Glu Leu Ile Gly Tyr Asn Asn Gln Phe		
85 90 95		3
AGC GGT AGA GCC AGT CAA TAT GTT GCG CCG GGT ACC GTT TCT TCC ATG	336	5
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	7
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	9
Leu His Ala Ser Asp Ser Val Tyr Tyr Leu Asp Thr Arg Arg Pro Asp		
130 135 140		11
CTC AAA TCA ATG GCG CTC AGT CAG CAA AAT ATG GAT ATA GAA TTA TCC	480	13
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	15
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	17
Ser Lys Leu Glu Asn Tyr Thr Lys Val Met Glu Met Leu Ser Thr Phe		
180 185 190		19
CGT CCT TCC GCC GCA ACG CCT TAT CAT GAT GCT TAT GAA AAT GTG CGT	624	21
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	23
Glu Val Ile Gln Leu Gln Asp Pro Gly Leu Glu Gln Leu Asn Ala Ser		
210 215 220		
CCG GCA ATT GCC GCG TTG ATG CAT CAA GCC TCC CTA TTG GGT ATT AAC	720	25
Pro Ala Ile Ala Gly Leu Met His Gln Ala Ser Leu Leu Gly Ile Asn		
225 230 235 240		27
GCT TCA ATC TCG CCT GAG CTA TTT AAT ATT CTG ACG GAG GAG ATT ACC	768	29
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	31
Gly 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	33
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	35
Ser Asp Glu Glu Leu Ser Gln Phe Ile Gly Lys Ala Ser Asn Phe Gly		
290 295 300		37
CAA CAG GAA TAT AGT AAT AAC CAA CTT ATT ACT CCG GTA GTC AAC AGC	960	39
Gln Gln Glu Tyr Ser Asn Asn Gln Leu Ile Thr Pro Val Val Asn Ser		
305 310 315 320		
AGT GAT GCG ACG GTT AAG GTA TAT CCG ATC ACC CCG CAA TAT ACA ACC	1008	41
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	43
Asn Ala Tyr Gln Met Asp Val Glu Leu Phe Pro Phe Gly Gly Glu Asn		
340 345 350		45
TAT CCG TTA GAT TAT AAA TTC AAA AAT TTT TAT AAT GCC TCT TAT TTA	1104	47
Tyr Arg Leu Asp Tyr Lys Phe Lys Asn Phe Tyr Asn Ala Ser Tyr Leu		
355 360 365		

RO 121280 B1

3	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					
5	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	
7	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		
9	GGT	TCT	TCG	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			
11	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				
13	GGC	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					
15	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	
17	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		
19	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			
21	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				
23	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					
25	GAT	TCG	CGA	AAA	ACC	ATA	CTT	AAG	CGT	GCA	TTT	AAT	ATT	CAT	CAT	CTC	1680
	Asp	Trp	Arg	Lys	Thr	Ile	Leu	Lys	Arg	Ala	Phe	Asn	Ile	Asp	Asp	Val	
	545					550					555					560	
27	TCG	CTC	TTC	CGC	CTG	CTT	AAA	ATT	ACC	GAC	CAT	GAT	AAT	AAA	GAT	GGA	1728
	Ser																

RO 121280 B1

645															650															655																
ACC	AGC	TAT	AAC	AAA	ACG	CTA	ACG	CCT	GAA	ATT	AAG	AAT	TTG	CTG	GAT	2016	3																													
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	5																													
Thr	Val	Tyr	His	Gly	Leu	Gln	Gly	Phe	Asp	Lys	Asp	Lys	Ala	Asp	Leu		7																													
			675				680					685																																		
CTA	CAT	GTC	ATC	CGC	CCC	TAT	ATT	CCG	GCC	ACC	TTG	CAA	TTA	TCA	TCG	2112	9																													
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	11																													
Glu	Asn	Val	Ala	His	Ser	Val	Leu	Leu	Trp	Ala	Asp	Lys	Leu	Gln	Pro		13																													
					710					715					720																															
GGC	GAC	GGC	GCA	ATG	ACA	GCA	GAA	AAA	TTC	TGG	GAC	TGG	TTG	AAT	ACT	2208	15																													
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	17																													
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	19																													
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	21																													
Thr	Gly	Ile	Asn	Glu	Asn	Ala	Phe	Arg	Leu	Phe	Val	Thr	Lys	Pro	Glu		23																													
						775					780																																			
ATG	TTT	GCC	GCT	GCA	ACT	GCA	GCA	CCG	CCC	CCG	CAT	GAT	GCC	CTT	TCA	2400	25																													
Met	Phe	Gly	Ala	Ala	Thr	Gly	Ala	Ala	Pro	Ala	His	Asp	Ala	Leu	Ser																															
					790				795						800																															
CTG	ATT	ATG	CTG	ACA	CGT	TTT	GCG	GAT	TCG	GTG	AAC	GCA	CTA	GCC	GAA	2448	27																													
Leu	Ile	Met	Leu	Thr	Arg	Phe	Ala	Asp	Trp	Val	Asn	Ala	Leu	Gly	Glu																															
				805					810					815																																
AAA	GCG	TCC	TCG	GTG	CTA	GCG	GCA	TTT	GAA	GCT	AAC	TCG	TTA	ACG	GCA	2496	29																													
Lys	Ala	Ser	Ser	Val	Leu	Ala	Ala	Phe	Glu	Ala	Asn	Ser	Leu	Thr	Ala																															
			820					825					830																																	
GAA	CAA	CTG	GCT	GAT	GCC	ATG	AAT	CTT	GAT	GCT	AAT	TTG	CTG	TTG	CAA	2544	31																													
Glu	Gln	Leu	Ala	Asp	Ala	Met	Asn	Leu	Asp	Ala	Asn	Leu	Leu	Leu	Gln		33																													
			835				840					845																																		
GCC	AGT	ATT	CAA	GCA	CAA	AAT	CAT	CAA	CAT	CTT	CCC	CCA	GTA	ACT	CCA	2592	35																													
Ala	Ser	Ile	Gln	Ala	Gln	Asn	His	Gln	His	Leu	Pro	Pro	Val	Thr	Pro																															
			850			855					860																																			
GAA	AAT	GCG	TTC	TCC	TGT	TGG	ACA	TCT	ATC	AAT	ACT	ATC	CTG	CAA	TGG	2640	37																													
Glu	Asn	Ala	Phe	Ser	Cys	Trp	Thr	Ser	Ile	Asn	Thr	Ile	Leu	Gln	Trp																															
					870				875						880																															
GTT	AAT	GTC	GCA	CAA	CAA	TTG	AAT	GTC	GCC	CCA	CAG	GGC	GTT	TCC	GCT	2688	39																													
Val	Asn	Val	Ala	Gln	Gln	Leu	Asn	Val	Ala	Pro	Gln	Gly	Val	Ser	Ala		41																													
				885				890						895																																
TTG	GTC	GGG	CTG	GAT	TAT	ATT	CAA	TCA	ATG	AAA	GAG	ACA	CCG	ACC	TAT	2736	43																													
Leu	Val	Gly	Leu	Asp	Tyr	Ile	Gln	Ser	Met	Lys	Glu	Thr	Pro	Thr	Tyr																															
			900				905						910																																	
GCC	CAG	TGG	GAA	AAC	GCG	GCA	GGC	GTA	TTA	ACC	GCC	GGG	TTG	AAT	TCA	2784	45																													
Ala	Gln	Trp	Glu	Asn	Ala	Ala	Gly	Val	Leu	Thr	Ala	Gly	Leu	Asn	Ser																															
			915				920					925																																		
CAA	CAG	GCT	AAT	ACA	TTA	CAC	GCT	TTT	CTG	GAT	GAA	TCT	CGC	AGT	GCC	2832	47																													

RO 121280 B1

1	Gln Gln Ala Asn Thr Leu His Ala Phe Leu Asp Glu Ser Arg Ser Ala	
3	930 935 940	
5	GCA TTA AGC ACC TAC TAT ATC CGT CAA GTC GCC AAG GCA GCG GCG GCT Ala Leu Ser Thr Tyr Tyr Ile Arg Gln Val Ala Lys Ala Ala Ala	2880 945 950 955 960
7	ATT AAA AGC CGT GAT GAC TTG TAT CAA TAC TTA CTG ATT GAT AAT CAG Ile Lys Ser Arg Asp Asp Leu Tyr Gln Tyr Leu Leu Ile Asp Asn Gln	2928 965 970 975
9	GTT TCT GCG GCA ATA AAA ACC ACC CGG ATC GCC GAA GCC ATT GCC AGT Val Ser Ala Ala Ile Lys Thr Thr Arg Ile Ala Glu Ala Ile Ala Ser	2976 980 985 990
11	ATT CAA CTG TAC GTC AAC CGG GCA TTG GAA AAT GTG GAA GAA AAT GCC Ile Gln Leu Tyr Val Asn Arg Ala Leu Glu Asn Val Glu Glu Asn Ala	3024 995 1000 1005
13	AAT TCG GGG GTT ATC AGC CGC CAA TTC TTT ATC GAC TGG GAC AAA TAC Asn Ser Gly Val Ile Ser Arg Gln Phe Phe Ile Asp Trp Asp Lys Tyr	3072 1010 1015 1020
15	AAT AAA CGC TAC AGC ACT TGG GCG GGT GTT TCT CAA TTA GTT TAC TAC Asn Lys Arg Tyr Ser Thr Trp Ala Gly Val Ser Gln Leu Val Tyr Tyr	3120 1025 1030 1035 1040
17	CCG GAA AAC TAT ATT GAT CCG ACC ATG CGT ATC GGA CAA ACC AAA ATG Pro Glu Asn Tyr Ile Asp Pro Thr Met Arg Ile Gly Gln Thr Lys Met	3168 1045 1050 1055
19	ATG GAC GCA TTA CTG CAA TCC GTC AGC CAA AGC CAA TTA AAC GCC GAT Met Asp Ala Leu Leu Gln Ser Val Ser Gln Ser Gln Leu Asn Ala Asp	3216 1060 1065 1070
21	ACC GTC GAA GAT GCC TTT ATG TCT TAT CTG ACA TCG TTT GAA CAA CTG Thr Val Glu Asp Ala Phe Met Ser Tyr Leu Thr Ser Phe Glu Gln Val	3264 1075 1080 1085
23	GCT AAT CTT AAA GTT ATT AGC GCA TAT CAC GAT AAT ATT AAT AAC GAT Ala Asn Leu Lys Val Ile Ser Ala Tyr His Asp Asn Ile Asn Asn Asp	3312 1090 1095 1100
25	CAA GGG CTG ACC TAT TTT ATC GGA CTC AGT GAA ACT GAT GCC GGT GAA Gln Gly Leu Thr Tyr Phe Ile Gly Leu Ser Glu Thr Asp Ala Gly Glu	3360 1105 1110 1115 1120
27	TAT TAT TGG CGC AGT GTC GAT CAC ACT AAA TTC AAC GAC GGT AAA TTC Tyr Tyr Trp Arg Ser Val Asp His Ser Lys Phe Asn Asp Gly Lys Phe	3408 1125 1130 1135
29	GCG GCT AAT GCC TGG AGT GAA TGG CAT AAA ATT GAT TGT CCA ATT AAC Ala Ala Asn Ala Trp Ser Glu Trp His Lys Ile Asp Cys Pro Ile Asn	3456 1140 1145 1150
31	CCT TAT AAA ACC ACT ATC COT CCA CTG ATA TAT AAA TCC CGC CTG TAT Pro Tyr Lys Ser Thr Ile Arg Pro Val Ile Tyr Lys Ser Arg Leu Tyr	3504 1155 1160 1165
33	CTG CTC TGG TTG GAA CAA AAG GAG ATC ACC AAA CAG ACA GGA AAT AGT Leu Leu Trp Leu Glu Gln Lys Glu Ile Thr Lys Gln Thr Gly Asn Ser	3552 1170 1175 1180
35	AAA GAT GGC TAT CAA ACT GAA ACG GAT TAT COT TAT GAA CTA AAA TTG Lys Asp Gly Tyr Gln Thr Glu Thr Asp Tyr Arg Tyr Glu Leu Lys Leu	3600 1185 1190 1195 1200
37	GCG CAT ATC CGC TAT GAT GCC ACT TCG AAT ACG CCA ATC ACC TTT GAT Ala His Ile Arg Tyr Asp Gly Thr Trp Asn Thr Pro Ile Thr Phe Asp	3648 1205 1210 1215
39		
41		
43		
45		
47		

RO 121280 B1

GTC AAT AAA AAA ATA TCC GAG CTA AAA CTG GAA AAA AAT ACA GCG CCC	3636	1
Val Asn Lys Lys Ile Ser Glu Leu Lys Leu Glu Lys Asn Arg Ala Pro		
1220 1225 1230		3
GGA CTC TAT TGT GCC GGT TAT CAA GGT GAA GAT ACG TTG CTG GTG ATG	3744	5
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	7
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	9
Gln Gly Leu Tyr Ile Phe Ala Asp Met Ala Ser Lys Asp Met Thr Pro		
1265 1270 1275 1280		11
GAA CAG AGC AAT GTT TAT CGG GAT AAT AGC TAT CAA CAA TTT GAT ACC	3888	13
Glu Gln Ser Asn Val Tyr Arg Asp Asn Ser Tyr Gln Gln Phe Asp Thr		
1285 1290 1295		
AAT AAT CTC AGA AGA GTG AAT AAC CGC TAT GCA GAG GAT TAT GAG ATT	3936	15
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	17
Pro Ser Ser Val Ser Ser Arg Lys Asp Tyr Gly Trp Gly Asp Tyr Tyr		
1315 1320 1325		19
CTC AGC ATG GTA TAT AAC GGA GAT ATT CCA ACT ATC AAT TAC AAA GCC	4032	21
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	23
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	25
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	27
Lys Tyr Gly Lys Leu Gly Asp Lys Phe Ile Val Tyr Thr Ser Leu Gly		
1380 1385 1390		29
GTC AAT CCA AAT AAC TCG TCA AAT AAG CTC ATG TTT TAC CCC GTC TAT	4224	31
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 GCG AGA CTA CTA TTC	4272	33
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	35
His Arg Asp Thr Thr Tyr Pro Ser Lys Val Glu Ala Trp Ile Pro Gly		
1425 1430 1435 1440		37
GCA AAA CGT TCT CTA ACC AAC CAA AAT GCC GCC ATT GGT GAT GAT TAT	4368	39
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	41
Ala Thr Asp Ser Leu Asn Lys Pro Asp Asp Leu Lys Gln Tyr Ile Phe		
1460 1465 1470		
ATG ACT GAC AGT AAA GCG ACT GCT ACT GAT CTC TCA GCG CCA GTA GAG	4464	43
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	45
Ile Asn Thr Ala Ile Ser Pro Ala Lys Val Gln Ile Ile Val Lys Ala		
1490 1495 1500		47

RO 121280 B1

1	GGT GGC AAG GAG CAA ACT TTT ACC GCA GAT AAA GAT GTC TCC ATT CAG	1560
3	Gly Gly Lys Glu Gln Thr Phe Thr Ala Asp Lys Asp Val Ser Ile Gln	1505 1510 1515 1520
5	CCA TCA CCT AGC TTT GAT GAA ATG AAT TAT CAA TTT AAT GCC CTT GAA	4603
	Pro Ser Pro Ser Phe Asp Glu Met Asn Tyr Gln Phe Asn Ala Leu Glu	1525 1530 1535
7	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
9	GTT ACT TTT ACC GCA TTT GCG GAG GAT GGC CGC AAA CTG GGT TAT GAA	4704
11	Val Thr Phe Thr Ala Phe Ala Glu Asp Gly Arg Lys Leu Gly Tyr Glu	1555 1560 1565
13	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
15	ACC CTG CAC CAT AAT GAA AAT GGT GCG CAA TAT ATG CAA TCG CAA TCC	4800
	Thr Leu His His Asn Glu Asn Gly Ala Gln Tyr Met Gln Trp Gln Ser	1585 1590 1595 1600
17	TAT CGT ACC CGC CTG AAT ACT CTA TTT GCC CGC CAG TTG GTT GCA CGC	4848
19	Tyr Arg Thr Arg Leu Asn Thr Leu Phe Ala Arg Gln Leu Val Ala Arg	1605 1610 1615
21	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
23	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
25	CCC TAT AAC CTA TCA ACT CAT GGT GAT GAA COT TGG TTT AAG CTT TAT	4992
27	Pro Tyr Asn Leu Ser Thr His Gly Asp Glu Arg Trp Phe Lys Leu Tyr	1650 1655 1660
29	ATC AAA CAT GTT GTT GAT AAT AAT TCA CAT ATT ATC TAT TCA GCC CAG	5040
	Ile Lys His Val Val Asp Asn Asn Ser His Ile Ile Tyr Ser Gly Gln	1665 1670 1675 1680
31	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
33	GTC CCA TTG AAT CAA GAT TAT CAC GCC AAG GTT TAT ATG ACC TTC AAG	5136
35	Val Pro Leu Asn Gln Asp Tyr His Ala Lys Val Tyr Met Thr Phe Lys	1700 1705 1710
37	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
39	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
41	GAG AGC GTC AAT GTC CTG AAT AAT ATT AGT ACC GAA CCA ATG GAT TTC	5280
43	Glu Ser Val Asn Val Leu Asn Asn Ile Ser Ser Glu Pro Met Asp Phe	1745 1750 1755 1760
45	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
47	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	

RO 121280 B1

	1730	1785	1790		1
AAC CGT TGG CTC AAA TAT GTC TGG ACT CCA TCC GGT TAT ATT GTC CAC	5424				3
Asn Arg Trp Leu Lys Tyr Val Trp Ser Pro Ser Gly Tyr Ile Val His	1795	1800	1805		5
GGC CAG ATT CAG AAC TAC CAG TGG AAC GTC CGC CCG TTA CTG GAA GAC	5472				7
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 CAT TCC CTC GAT CCT GAC GCG GTA	5520				9
Thr Ser Trp Asn Ser Asp Pro Leu Asp Ser Val Asp Pro Asp Ala Val	1825	1830	1835	1840	11
GCA CAG CAC GAT CCA ATG CAC TAC AAA GTT TCA ACT TTT ATG CGT ACC	5568				13
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				15
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				17
Arg Asp Thr Leu Asn Glu Ala Lys Met Trp Tyr Met Gln Ala Leu His	1875	1880	1885		19
CTA TTA GGT GAC AAA CCT TAT CTA CCG CTG AGT ACG ACA TGG AGT GAT	5712				21
Leu Leu Gly Asp Lys Pro Tyr Leu Pro Leu Ser Thr Thr Trp Ser Asp	1890	1895	1900		
CCA CGA CTA GAC AGA GCC GCG GAT ATC ACT ACC CAA AAT GCT CAC GAC	5760				23
Pro Arg Leu Asp Arg Ala Ala Asp Ile Thr Thr Gln Asn Ala His Asp	1905	1910	1915	1920	
AGC CCA ATA GTC GCT CTG CCG CAG AAT ATA CCT ACA CCG GCA CCT TTA	5808				25
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				27
Ser Leu Arg Ser Ala Asn Thr Leu Thr Asp Leu Phe Leu Pro Gln Ile	1940	1945	1950		29
AAT GAA GTG ATG ATG AAT TAC TGG CAG ACA TTA GCT CAG AGA GTA TAC	5904				31
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				33
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				35
Ile Tyr Ala Thr Pro Ala Asp Pro Lys Ala Leu Leu Ser Ala Ala Val	1985	1990	1995	2000	37
GCC ACT TCT CAA GGT GGA GGC AAG CTA CCG GAA TCA TTT ATG TCC CTG	6048				39
Ala Thr Ser Gln 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				41
Trp Arg Phe Pro His Met Leu Glu Asn Ala Arg Gly Met Val Ser Gln	2020	2025	2030		
CTC ACC CAG TTC GGC TCC ACG TTA CAA AAT ATT ATC GAA CGT CAG GAC	6144				43
Leu Thr Gln Phe Gly Ser Thr Leu Gln Asn Ile Ile Glu Arg Gln Asp	2035	2040	2045		45
GCG GAA GCG CTC AAT GCG TTA TTA CAA AAT CAG GCC GCC GAG CTG ATA	6192				47
Ala Glu Ala Leu Asn Ala Leu Leu Gln Asn Gln Ala Ala Glu Leu Ile	2050	2055	2060		
TTG ACT AAC CTG ACC ATT CAG GAC AAA ACC ATT GAA GAA TTG GAT GCC	6240				

RO 121280 B1

1	Leu	Thr	Asn	Leu	Ser	Ile	Gln	Asp	Lys	Thr	Ile	Glu	Glu	Leu	Asp	Ala	
	2065					2070					2075					2080	
3	GAG	AAA	ACG	GTG	TTG	GAA	AAA	TCC	AAA	GCG	GGA	GCA	CAA	TCG	CGC	TTT	6283
5	Glu	Lys	Thr	Val		2085	Glu	Lys	Ser	Lys	Ala	Gly	Ala	Gln	Ser	Arg	Phe
										2090						2095	
7	GAT	AGC	TAC	GGC	AAA	CTG	TAC	GAT	GAG	AAT	ATC	AAC	GCC	GCT	GAA	AAC	6336
	Asp	Ser	Tyr			2100	Gly	Lys	Leu	Tyr	Asp	Glu	Asn	Ile	Asn	Ala	Gly
										2105						2110	Asn
9	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	
11	CAG	GCA	TCC	CGT	CTG	GCC	GGT	GCG	GCG	GCT	GAT	CTG	GTG	CCT	AAC	ATC	6432
13	Gln	Ala	Ser	Arg	Leu	Ala	Gly	Ala	Ala	Ala	Asp	Leu	Val	Pro	Asn	Ile	
			2130					2135					2140				
15	TTC	GCC	TTT	GCC	GGT	GCC	GCG	AGC	CGT	TGG	GGG	GCT	ATC	GCT	GAG	GCG	6480
	Phe	Gly	Phe	Ala	Gly	Gly	Gly	Ser	Arg	Trp	Gly	Ala	Ile	Ala	Glu	Ala	
						2150						2155				2160	
17	ACA	GCT	TAT	GTG	ATG	GAA	TTC	TCC	GCG	AAT	GTT	ATC	AAC	ACC	GAA	GCG	6528
	Thr	Gly	Tyr	Val	Met	Glu	Phe	Ser	Ala	Asn	Val	Met	Asn	Thr	Glu	Ala	
						2165					2170					2175	
19	GAT	AAA	ATT	AGC	CAA	TCT	GAA	ACC	TAC	CGT	CGT	CGC	CGT	CAG	GAG	TGG	6576
21	Asp	Lys	Ile	Ser	Gln	Ser	Glu	Thr	Tyr	Arg	Arg	Arg	Arg	CAG	Gln	Glu	Trp
					2180					2185					2190		
23	GAG	ATC	CAG	CGG	AAT	AAT	GCC	GAA	GCG	GAA	TTC	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			
25	CAG	CTC	AAA	TCA	CTC	GCT	GTA	GCG	GCG	GAA	GCC	GCC	GTA	TTG	CAG	AAA	6672
	Gln	Leu	Lys	Ser	Leu	Ala	Val	Arg	Arg	Glu	Ala	Ala	Val	Leu	Gln	Lys	
					2210			2215					2220				
27	ACC	AGT	CTG	AAA	ACC	CAA	CAA	GAA	CAG	ACC	CAA	TCT	CAA	TTG	GCC	TTC	6720
	Thr	Ser	Leu	Lys	Thr	Gln	Gln	Glu	Gln	Thr	Gln	Ser	Gln	Leu	Ala	Phe	
29						2230						2235				2240	
31	CTG	CAA	CGT	AAG	TTC	AGC	AAT	CAG	GCG	TTA	TAC	AAC	TGG	CTG	CGT	GGT	6768
	Leu	Gln	Arg	Lys	Phe	Ser	Asn	Gln	Ala	Leu	Tyr	Asn	Trp	Leu	Arg	Gly	
						2245				2250					2255		
33	CCA	CTG	GCG	GCG	ATT	TAC	TTC	CAG	TTC	TAC	GAT	TTG	GCC	GTC	GCG	CGT	6816
	Arg	Leu	Ala	Ala	Ile	Tyr	Phe	Gln	Phe	Tyr	Asp	Leu	Ala	Val	Ala	Arg	
						2260				2265					2270		
35	TGC	CTG	ATG	GCA	GAA	CAA	GCT	TAC	CGT	TGG	GAA	CTC	AAT	GAT	GAC	TCT	6864
	Cys	Leu	Met	Ala	Glu	Gln	Ala	Tyr	Arg	Trp	Glu	Leu	Asn	Asp	Asp	Ser	
37					2275				2280				2285				
39	GCC	GCG	TTC	ATT	AAA	CCG	GCG	GCC	TGG	CAG	GGA	ACC	TAT	GCC	CGT	CTG	6912
	Ala	Arg	Phe	Ile	Lys	Pro	Gly	Ala	Trp	Gln	Gly	Thr	Tyr	Ala	Gly	Leu	
					2290			2295				2300					
41	CTT	GCA	GCT	GAA	ACC	TTG	ATG	CTG	AGT	CTG	GCA	CAA	ATG	GAA	GAC	GCT	6960
	Leu	Ala	Gly	Glu	Thr	Leu	Met	Leu	Ser	Leu	Ala	Gln	Met	Glu	Asp	Ala	
						2310						2315				2320	
43	CAT	CTG	AAA	CGC	GAT	AAA	CGC	GCA	TTA	GAG	GTT	GAA	CGC	ACA	GTA	TCG	7008
	His	Leu	Lys	Arg	Asp	Lys	Arg	Ala	Leu	Glu	Val	Glu	Arg	Thr	Val	Ser	
						2325					2330					2335	
45	CTG	GCC	GAA	GTT	TAT	GCA	GGA	TTA	CCA	AAA	GAT	AAC	GGT	CCA	TTT	TCC	7056
47	Leu	Ala	Glu	Val	Tyr	Ala	Gly	Leu	Pro	Lys	Asp	Asn	Gly	Pro	Phe	Ser	
					2340					2345					2350		

RO 121280 B1

CTG GCT CAG GAA ATT GAC AAG CTG GTG AGT CAA GGT TCA GGC AGT GCC	7174	1
Leu Ala Gln Glu Ile Asp Lys Leu Val Ser Gln Gly Ser Gly Ser Ala		
2355 2360 2365		3
GGC AGT GGT AAT AAT AAT TTG GCG TTC GGC GCC GGC ACG GAC ACT AAA	7152	5
Gly Ser Gly Asn Asn Asn Leu Ala Phe Gly Ala Gly Thr Asp Thr Lys		
2370 2375 2380		
ACC TCT TTG CAG GCA TCA GTT TCA TTC GCT GAT TTG AAA ATT CGT GAA	7200	7
Thr Ser Leu Gln Ala Ser Val Ser Phe Ala Asp Leu Lys Ile Arg Glu		
2385 2390 2395 2400		9
GAT TAC CCG GCA TCG CTT GGC AAA ATT CGA CGT ATC AAA CAG ATC AGC	7248	11
Asp Tyr Pro Ala Ser Leu Gly Lys Ile Arg Arg Ile Lys Gln Ile Ser		
2405 2410 2415		
GTC ACT TTG CCC GCG CTA CTG GGA CCG TAT CAG GAT GTA CAG GCA ATA	7296	13
Val Thr Leu Pro Ala Leu Leu Gly Pro Tyr Gln Asp Val Gln Ala Ile		
2420 2425 2430		
TTG TCT TAC GGC GAT AAA GCC GGA TTA GCT AAC GGC TGT GAA GCG CTG	7344	15
Leu Ser Tyr Gly Asp Lys Ala Gly Leu Ala Asn Gly Cys Glu Ala Leu		
2435 2440 2445		17
GCA GTT TCT CAC GGT ATG AAT GAC AGC GGC CAA TTC CAG CTC GAT TTC	7392	19
Ala Val Ser His Gly Met Asn Asp Ser Gly Gln Phe Gln Leu Asp Phe		
2450 2455 2460		
AAC GAT GGC AAA TTC CTG CCA TTC GAA GGC ATC GCC ATT GAT CAA GCC	7440	21
Asn Asp Gly Lys Phe Leu Pro Phe Glu Gly Ile Ala Ile Asp Gln Gly		
2465 2470 2475 2480		23
ACG CTG ACA CTG AGC TTC CCA AAT GCA TCT ATG CCG GAG AAA GGT AAA	7488	25
Thr Leu Thr Leu Ser Phe Pro Asn Ala Ser Met Pro Glu Lys Gly Lys		
2485 2490 2495		
CAA GCC ACT ATG TTA AAA ACC CTG AAC GAT ATC ATT TTG CAT ATT CGC	7536	27
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	29
Tyr Thr Ile Lys ...		
2516		31

RO 121280 B1

(2) INFORMAȚIA PENTRU SECV. ID. NR. 47:

(i) CARACTERISTICI DE SECVENȚĂ:

(A) LUNGIME: 2516 baze perechi

(B) TIP: acid nucleic

(C) ÎMPLETIRE: simplă

(D) TOPOLOGIE: lineară

(ii) TIPUL MOLECULEI: proteină

(xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 47 (TcdA)

Țesătura	De la	Până la	Descrierea
Peptidă	1	2516	Proteină TcdA
Peptidă	89	1937	Peptidă TcdA _{ii}
Fragment	89	100	S2 N-terminus (SECV. ID NR. 13)
Fragment	284	299	(SECV. ID NR. 38)
Fragment	554	563	(SECV. ID NR. 17)
Fragment	1080	1092	(SECV. ID NR. 23; 12/13)
Fragment	1385	1400	(SECV. ID NR. 18)
Țesătura	De la	Până la	Descrierea
Fragment	1478	1479	(SECV. ID NR. 39)
Fragment	1620	1642	(SECV. ID NR. 21; 19/23)
Fragment	1938	1948	(SECV. ID NR. 41)
Peptidă	1938	2516	Peptidă TcdA _{iii}
Fragment	2327	2345	(SECV. ID NR. 42)
Fragment	2398	2408	(SECV. ID NR. 43)

RO 121280 B1

Met	Asn	Glu	Ser	Val	Lys	Glu	Ile	Pro	Asp	Val	Leu	Lys	Ser	Gln	Cys	1
1			5						10					15		
Gly	Phe	Asn	Cys	Leu	Thr	Asp	Ile	Ser	His	Ser	Ser	Phe	Asn	Glu	Phe	3
		20					25						30			
Arg	Gln	Gln	Val	Ser	Glu	His	Leu	Ser	Trp	Ser	Glu	Thr	His	Asp	Leu	5
		35					40					45				
Tyr	His	Asp	Ala	Gln	Gln	Ala	Gln	Lys	Asp	Asn	Arg	Leu	Tyr	Glu	Ala	7
	50					55					60					
Arg	Ile	Leu	Lys	Arg	Ala	Asn	Pro	Gln	Leu	Gln	Asn	Ala	Val	His	Leu	9
65					70					75					80	
Ala	Ile	Leu	Ala	Pro	Asn	Ala	Glu	Leu	Ile	Gly	Tyr	Asn	Asn	Gln	Phe	11
				85					90					95		
Ser	Gly	Arg	Ala	Ser	Gln	Tyr	Val	Ala	Pro	Gly	Thr	Val	Ser	Ser	Met	13
			100					105					110			
Phe	Ser	Pro	Ala	Ala	Tyr	Leu	Thr	Glu	Leu	Tyr	Arg	Glu	Ala	Arg	Asn	15
		115					120					125				
Leu	His	Ala	Ser	Asp	Ser	Val	Tyr	Tyr	Leu	Asp	Thr	Arg	Arg	Pro	Asp	17
	130					135					140					
Leu	Lys	Ser	Met	Ala	Leu	Ser	Gln	Gln	Asn	Met	Asp	Ile	Glu	Leu	Ser	19
145					150					155					160	
Thr	Leu	Ser	Leu	Ser	Asn	Glu	Leu	Leu	Leu	Glu	Ser	Ile	Lys	Thr	Glu	21
				165					170					175		
Ser	Lys	Leu	Glu	Asn	Tyr	Thr	Lys	Val	Met	Glu	Met	Leu	Ser	Thr	Phe	23
			180					185					190			
Arg	Pro	Ser	Gly	Ala	Thr	Pro	Tyr	His	Asp	Ala	Tyr	Glu	Asn	Val	Arg	25
		195					200					205				
Glu	Val	Ile	Gln	Leu	Gln	Asp	Pro	Gly	Leu	Glu	Gln	Leu	Asn	Ala	Ser	27
	210					215					220					
Pro	Ala	Ile	Ala	Gly	Leu	Met	His	Gln	Ala	Ser	Leu	Leu	Gly	Ile	Asn	29
225					230					235					240	
Ala	Ser	Ile	Ser	Pro	Glu	Leu	Phe	Asn	Ile	Leu	Thr	Glu	Glu	Ile	Thr	31
				245					250					255		
Glu	Gly	Asn	Ala	Glu	Glu	Leu	Tyr	Lys	Lys	Asn	Phe	Gly	Asn	Ile	Glu	33
			260					265					270			
Pro	Ala	Ser	Leu	Ala	Met	Pro	Glu	Tyr	Leu	Lys	Arg	Tyr	Tyr	Asn	Leu	35
		275					280					285				
Ser	Asp	Glu	Glu	Leu	Ser	Gln	Phe	Ile	Gly	Lys	Ala	Ser	Asn	Phe	Gly	37
	290					295					300					
Gln	Gln	Glu	Tyr	Ser	Asn	Asn	Gln	Leu	Ile	Thr	Pro	Val	Val	Asn	Ser	39
305					310					315					320	
Ser	Asp	Gly	Thr	Val	Lys	Val	Tyr	Arg	Ile	Thr	Arg	Glu	Tyr	Thr	Thr	41
				325					330					335		
Asn	Ala	Tyr	Gln	Met	Asp	Val	Glu	Leu	Phe	Pro	Phe	Gly	Gly	Glu	Asn	43
			340					345					350			
Tyr	Arg	Leu	Asp	Tyr	Lys	Phe	Lys	Asn	Phe	Tyr	Asn	Ala	Ser	Tyr	Leu	45
		355					360					365				
Ser	Ile	Lys	Leu	Asn	Asp	Lys	Arg	Glu	Leu	Val	Arg	Thr	Glu	Gly	Ala	47

RO 121280 B1

1		370		375		380													
3	Pro	Gln	Val	Asn	Ile	Glu	Tyr	Ser	Ala	Asn	Ile	Thr	Leu	Asn	Thr	Ala			
	385					390					395					400			
5	Asp	Ile	Ser	Gln	Pro	Phe	Glu	Ile	Gly	Leu	Thr	Arg	Val	Leu	Pro	Ser			
					405					410					415				
7	Gly	Ser	Trp	Ala	Tyr	Ala	Ala	Ala	Lys	Phe	Thr	Val	Glu	Glu	Tyr	Asn			
				420					425					430					
9	Gln	Tyr	Ser	Phe	Leu	Leu	Lys	Leu	Asn	Lys	Ala	Ile	Arg	Leu	Ser	Arg			
			435					440					445						
11	Ala	Thr	Glu	Leu	Ser	Pro	Thr	Ile	Leu	Glu	Gly	Ile	Val	Arg	Ser	Val			
	450						455						460						
13	Asn	Leu	Gln	Leu	Asp	Ile	Asn	Thr	Asp	Val	Leu	Gly	Lys	Val	Phe	Leu			
	465					470					475					480			
15	Thr	Lys	Tyr	Tyr	Met	Gln	Arg	Tyr	Ala	Ile	His	Ala	Glu	Thr	Ala	Leu			
					485					490					495				
17	Ile	Leu	Cys	Asn	Ala	Pro	Ile	Ser	Gln	Arg	Ser	Tyr	Asp	Asn	Gln	Pro			
				500					505					510					
19	Ser	Gln	Phe	Asp	Arg	Leu	Phe	Asn	Thr	Pro	Leu	Leu	Asn	Gly	Gln	Tyr			
			515					520					525						
21	Phe	Ser	Thr	Gly	Asp	Glu	Glu	Ile	Asp	Leu	Asn	Ser	Gly	Ser	Thr	Gly			
	530					535						540							
23	Asp	Trp	Arg	Lys	Thr	Ile	Leu	Lys	Arg	Ala	Phe	Asn	Ile	Asp	Asp	Val			
	545					550					555					560			
25	Ser	Leu	Phe	Arg	Leu	Leu	Lys	Ile	Thr	Asp	His	Asp	Asn	Lys	Asp	Gly			
				565					570						575				
27	Lys	Ile	Lys	Asn	Asn	Leu	Lys	Asn	Leu	Ser	Asn	Leu	Tyr	Ile	Gly	Lys			
				580					585					590					
29	Leu	Leu	Ala	Asp	Ile	His	Gln	Leu	Thr	Ile	Asp	Glu	Leu	Asp	Leu	Leu			
			595					600					605						
31	Leu	Ile	Ala	Val	Gly	Glu	Gly	Lys	Thr	Asn	Leu	Ser	Ala	Ile	Ser	Asp			
	610					615						620							
33	Lys	Gln	Leu	Ala	Thr	Leu	Ile	Arg	Lys	Leu	Asn	Thr	Ile	Thr	Ser	Trp			
	625					630					635					640			
35	Leu	His	Thr	Gln	Lys	Trp	Ser	Val	Phe	Gln	Leu	Phe	Ile	Met	Thr	Ser			
				645					650						655				
37	Thr	Ser	Tyr	Asn	Lys	Thr	Leu	Thr	Pro	Glu	Ile	Lys	Asn	Leu	Leu	Asp			
				660					665					670					
39	Thr	Val	Tyr	His	Gly	Leu	Gln	Gly	Phe	Asp	Lys	Asp	Lys	Ala	Asp	Leu			
			675					680					685						
41	Leu	His	Val	Met	Ala	Pro	Tyr	Ile	Ala	Ala	Thr	Leu	Gln	Leu	Ser	Ser			
	690						695					700							
43	Glu	Asn	Val	Ala	His	Ser	Val	Leu	Leu	Trp	Ala	Asp	Lys	Leu	Gln	Pro			
	705					710					715					720			
45	Gly	Asp	Gly	Ala	Met	Thr	Ala	Glu	Lys	Phe	Trp	Asp	Trp	Leu	Asn	Thr			
				725						730				735					
47	Lys	Tyr	Thr	Pro	Gly	Ser	Ser	Glu	Ala	Val	Glu	Thr	Gln	Glu	His	Ile			
				740					745					750					

RO 121280 B1

Val Gln Tyr Cys Gln Ala Leu Ala Gln Leu Glu Met Val Tyr His Ser	1
755 760 765	
Thr Gly Ile Asn Glu Asn Ala Phe Arg Leu Phe Val Thr Lys Pro Glu	3
770 775 780	
Met Phe Gly Ala Ala Thr Gly Ala Ala Pro Ala His Asp Ala Leu Ser	5
785 790 795 800	
Leu Ile Met Leu Thr Arg Phe Ala Asp Trp Val Asn Ala Leu Gly Glu	7
805 810 815	
Lys Ala Ser Ser Val Leu Ala Ala Phe Glu Ala Asn Ser Leu Thr Ala	9
820 825 830	
Glu Gln Leu Ala Asp Ala Met Asn Leu Asp Ala Asn Leu Leu Leu Gln	11
835 840 845	
Ala Ser Ile Gln Ala Gln Asn His Gln His Leu Pro Pro Val Thr Pro	13
850 855 860	
Glu Asn Ala Phe Ser Cys Trp Thr Ser Ile Asn Thr Ile Leu Gln Trp	15
865 870 875 880	
Val Asn Val Ala Gln Gln Leu Asn Val Ala Pro Gln Gly Val Ser Ala	17
885 890 895	
Leu Val Gly Leu Asp Tyr Ile Gln Ser Met Lys Glu Thr Pro Thr Tyr	19
900 905 910	
Ala Gln Trp Glu Asn Ala Ala Gly Val Leu Thr Ala Gly Leu Asn Ser	21
915 920 925	
Gln Gln Ala Asn Thr Leu His Ala Phe Leu Asp Glu Ser Arg Ser Ala	23
930 935 940	
Ala Leu Ser Thr Tyr Tyr Ile Arg Gln Val Ala Lys Ala Ala Ala Ala	25
945 950 955 960	
Ile Lys Ser Arg Asp Asp Leu Tyr Gln Tyr Leu Leu Ile Asp Asn Gln	27
965 970 975	
Val Ser Ala Ala Ile Lys Thr Thr Arg Ile Ala Glu Ala Ile Ala Ser	29
980 985 990	
Ile Gln Leu Tyr Val Asn Arg Ala Leu Glu Asn Val Glu Glu Asn Ala	31
995 1000 1005	
Asn Ser Gly Val Ile Ser Arg Gln Phe Phe Ile Asp Trp Asp Lys Tyr	33
1010 1015 1020	
Asn Lys Arg Tyr Ser Thr Trp Ala Gly Val Ser Gln Leu Val Tyr Tyr	35
1025 1030 1035 1040	
Pro Glu Asn Tyr Ile Asp Pro Thr Met Arg Ile Gly Gln Thr Lys Met	37
1045 1050 1055	
Met Asp Ala Leu Leu Gln Ser Val Ser Gln Ser Gln Leu Asn Ala Asp	39
1060 1065 1070	
Thr Val Glu Asp Ala Phe Met Ser Tyr Leu Thr Ser Phe Glu Gln Val	41
1075 1080 1085	
Ala Asn Leu Lys Val Ile Ser Ala Tyr His Asp Asn Ile Asn Asn Asp	43
1090 1095 1100	
Gln Gly Leu Thr Tyr Phe Ile Gly Leu Ser Glu Thr Asp Ala Gly Glu	45
1105 1110 1115 1120	
Tyr Tyr Trp Arg Ser Val Asp His Ser Lys Phe Asn Asp Gly Lys Phe	47
1125 1130 1135	

RO 121280 B1

1 Ala Ala Asn Ala Trp Ser Glu Trp His Lys Ile Asp Cys Pro Ile Asn
1140 1145 1150

3 Pro Tyr Lys Ser Thr Ile Arg Pro Val Ile Tyr Lys Ser Arg Leu Tyr
1155 1160 1165

5 Leu Leu Trp Leu Glu Gln Lys Glu Ile Thr Lys Gln Thr Gly Asn Ser
1170 1175 1180

7 Lys Asp Gly Tyr Gln Thr Glu Thr Asp Tyr Arg Tyr Glu Leu Lys Leu
1185 1190 1195 1200

9 Ala His Ile Arg Tyr Asp Gly Thr Trp Asn Thr Pro Ile Thr Phe Asp
1205 1210 1215

11 Val Asn Lys Lys Ile Ser Glu Leu Lys Leu Glu Lys Asn Arg Ala Pro
1220 1225 1230

13 Gly Leu Tyr Cys Ala Gly Tyr Gln Gly Glu Asp Thr Leu Leu Val Met
1235 1240 1245

15 Phe Tyr Asn Gln Gln Asp Thr Leu Asp Ser Tyr Lys Asn Ala Ser Met
1250 1255 1260

17 Gln Gly Leu Tyr Ile Phe Ala Asp Met Ala Ser Lys Asp Met Thr Pro
1265 1270 1275 1280

19 Glu Gln Ser Asn Val Tyr Arg Asp Asn Ser Tyr Gln Gln Phe Asp Thr
1285 1290 1295

21 Asn Asn Val Arg Arg Val Asn Asn Arg Tyr Ala Glu Asp Tyr Glu Ile
1300 1305 1310

23 Pro Ser Ser Val Ser Ser Arg Lys Asp Tyr Gly Trp Gly Asp Tyr Tyr
1315 1320 1325

25 Leu Ser Met Val Tyr Asn Gly Asp Ile Pro Thr Ile Asn Tyr Lys Ala
1330 1335 1340

27 Ala Ser Ser Asp Leu Lys Ile Tyr Ile Ser Pro Lys Leu Arg Ile Ile
1345 1350 1355 1360

29 His Asn Gly Tyr Glu Gly Gln Lys Arg Asn Gln Cys Asn Leu Met Asn
1365 1370 1375

31 Lys Tyr Gly Lys Leu Gly Asp Lys Phe Ile Val Tyr Thr Ser Leu Gly
1380 1385 1390

33 Val Asn Pro Asn Asn Ser Ser Asn Lys Leu Met Phe Tyr Pro Val Tyr
1395 1400 1405

35 Gln Tyr Ser Gly Asn Thr Ser Gly Leu Asn Gln Gly Arg Leu Leu Phe
1410 1415 1420

37 His Arg Asp Thr Thr Tyr Pro Ser Lys Val Glu Ala Trp Ile Pro Gly
1425 1430 1435 1440

39 Ala Lys Arg Ser Leu Thr Asn Gln Asn Ala Ala Ile Gly Asp Asp Tyr
1445 1450 1455

41 Ala Thr Asp Ser Leu Asn Lys Pro Asp Asp Leu Lys Gln Tyr Ile Phe
1460 1465 1470

43 Met Thr Asp Ser Lys Gly Thr Ala Thr Asp Val Ser Gly Pro Val Glu
1475 1480 1485

45 Ile Asn Thr Ala Ile Ser Pro Ala Lys Val Gln Ile Ile Val Lys Ala
1490 1495 1500

47 Gly Gly Lys Glu Gln Thr Phe Thr Ala Asp Lys Asp Val Ser Ile Gln

RO 121280 B1

1505	1510	1515	1520	1
Pro Ser Pro Ser Phe Asp Glu Met Asn Tyr Gln Phe Asn Ala Leu Glu	1525	1530	1535	3
Ile Asp Gly Ser Gly Leu Asn Phe Ile Asn Asn Ser Ala Ser Ile Asp	1540	1545	1550	5
Val Thr Phe Thr Ala Phe Ala Glu Asp Gly Arg Lys Leu Gly Tyr Glu	1555	1560	1565	7
Ser Phe Ser Ile Pro Val Thr Leu Lys Val Ser Thr Asp Asn Ala Leu	1570	1575	1580	9
Thr Leu His His Asn Glu Asn Gly Ala Gln Tyr Met Gln Trp Gln Ser	1585	1590	1595	11
Tyr Arg Thr Arg Leu Asn Thr Leu Phe Ala Arg Gln Leu Val Ala Arg	1605	1610	1615	13
Ala Thr Thr Gly Ile Asp Thr Ile Leu Ser Met Glu Thr Gln Asn Ile	1620	1625	1630	15
Gln Glu Pro Gln Leu Gly Lys Gly Phe Tyr Ala Thr Phe Val Ile Pro	1635	1640	1645	17
Pro Tyr Asn Leu Ser Thr His Gly Asp Glu Arg Trp Phe Lys Leu Tyr	1650	1655	1660	19
Ile Lys His Val Val Asp Asn Asn Ser His Ile Ile Tyr Ser Gly Gln	1665	1670	1675	21
Leu Thr Asp Thr Asn Ile Asn Ile Thr Leu Phe Ile Pro Leu Asp Asp	1685	1690	1695	23
Val Pro Leu Asn Gln Asp Tyr His Ala Lys Val Tyr Met Thr Phe Lys	1700	1705	1710	25
Lys Ser Pro Ser Asp Gly Thr Trp Trp Gly Pro His Phe Val Arg Asp	1715	1720	1725	27
Asp Lys Gly Ile Val Thr Ile Asn Pro Lys Ser Ile Leu Thr His Phe	1730	1735	1740	29
Glu Ser Val Asn Val Leu Asn Asn Ile Ser Ser Glu Pro Met Asp Phe	1745	1750	1755	31
Ser Gly Ala Asn Ser Leu Tyr Phe Trp Glu Leu Phe Tyr Tyr Thr Pro	1765	1770	1775	33
Met Leu Val Ala Gln Arg Leu Leu His Glu Gln Asn Phe Asp Glu Ala	1780	1785	1790	35
Asn Arg Trp Leu Lys Tyr Val Trp Ser Pro Ser Gly Tyr Ile Val His	1795	1800	1805	37
Gly Gln Ile Gln Asn Tyr Gln Trp Asn Val Arg Pro Leu Leu Glu Asp	1810	1815	1820	39
Thr Ser Trp Asn Ser Asp Pro Leu Asp Ser Val Asp Pro Asp Ala Val	1825	1830	1835	41
Ala Gln His Asp Pro Met His Tyr Lys Val Ser Thr Phe Met Arg Thr	1845	1850	1855	43
Leu Asp Leu Leu Ile Ala Arg Gly Asp His Ala Tyr Arg Gln Leu Glu	1860	1865	1870	45
Arg Asp Thr Leu Asn Glu Ala Lys Met Trp Tyr Met Gln Ala Leu His	1875	1880	1885	47

RO 121280 B1

1 Leu Leu Gly Asp Lys Pro Tyr Leu Pro Leu Ser Thr Thr Trp Ser Asp
 1390 1895 1900
 3 Pro Arg Leu Asp Arg Ala Ala Asp Ile Thr Thr Gln Asn Ala His Asp
 1905 1910 1915 1920
 5 Ser Ala Ile Val Ala Leu Arg Gln Asn Ile Pro Thr Pro Ala Pro Leu
 1925 1930 1935
 7 Ser Leu Arg Ser Ala Asn Thr Leu Thr Asp Leu Phe Leu Pro Gln Ile
 1940 1945 1950
 9 Asn Glu Val Met Met Asn Tyr Trp Gln Thr Leu Ala Gln Arg Val Tyr
 1955 1960 1965
 11 Asn Leu Arg His Asn Leu Ser Ile Asp Gly Gln Pro Leu Tyr Leu Pro
 1970 1975 1980
 13 Ile Tyr Ala Thr Pro Ala Asp Pro Lys Ala Leu Leu Ser Ala Ala Val
 1985 1990 1995 2000
 15 Ala Thr Ser Gln Gly Gly Gly Lys Leu Pro Glu Ser Phe Met Ser Leu
 2005 2010 2015
 17 Trp Arg Phe Pro His Met Leu Glu Asn Ala Arg Gly Met Val Ser Gln
 2020 2025 2030
 19 Leu Thr Gln Phe Gly Ser Thr Leu Gln Asn Ile Ile Glu Arg Gln Asp
 2035 2040 2045
 21 Ala Glu Ala Leu Asn Ala Leu Leu Gln Asn Gln Ala Ala Glu Leu Ile
 2050 2055 2060
 23 Leu Thr Asn Leu Ser Ile Gln Asp Lys Thr Ile Glu Glu Leu Asp Ala
 2065 2070 2075 2080
 25 Glu Lys Thr Val Leu Glu Lys Ser Lys Ala Gly Ala Gln Ser Arg Phe
 2085 2090 2095
 27 Asp Ser Tyr Gly Lys Leu Tyr Asp Glu Asn Ile Asn Ala Gly Glu Asn
 2100 2105 2110
 29 Gln Ala Met Thr Leu Arg Ala Ser Ala Ala Gly Leu Thr Thr Ala Val
 2115 2120 2125
 31 Gln Ala Ser Arg Leu Ala Gly Ala Ala Ala Asp Leu Val Pro Asn Ile
 2130 2135 2140
 33 Phe Gly Phe Ala Gly Gly Gly Ser Arg Trp Gly Ala Ile Ala Glu Ala
 2145 2150 2155 2160
 35 Thr Gly Tyr Val Met Glu Phe Ser Ala Asn Val Met Asn Thr Glu Ala
 2165 2170 2175
 37 Asp Lys Ile Ser Gln Ser Glu Thr Tyr Arg Arg Arg Arg Gln Glu Trp
 2180 2185 2190
 39 Glu Ile Gln Arg Asn Asn Ala Glu Ala Glu Leu Lys Gln Ile Asp Ala
 2195 2200 2205
 41 Gln Leu Lys Ser Leu Ala Val Arg Arg Glu Ala Ala Val Leu Gln Lys
 2210 2215 2220
 43 Thr Ser Leu Lys Thr Gln Gln Glu Gln Thr Gln Ser Gln Leu Ala Phe
 2225 2230 2235 2240
 45 Leu Gln Arg Lys Phe Ser Asn Gln Ala Leu Tyr Asn Trp Leu Arg Gly
 2245 2250 2255
 47 Arg Leu Ala Ala Ile Tyr Phe Gln Phe Tyr Asp Leu Ala Val Ala Arg
 2260 2265 2270

RO 121280 B1

Cys Leu Met Ala Glu Gln Ala Tyr Arg Trp Glu Leu Asn Asp Asp Ser 2275 2280 2285	1
Ala Arg Phe Ile Lys Pro Gly Ala Trp Gln Gly Thr Tyr Ala Gly Leu 2290 2295 2300	3
Leu Ala Gly Glu Thr Leu Met Leu Ser Leu Ala Gln Met Glu Asp Ala 2305 2310 2315 2320	5
His Leu Lys Arg Asp Lys Arg Ala Leu Glu Val Glu Arg Thr Val Ser 2325 2330 2335	7
Leu Ala Glu Val Tyr Ala Gly Leu Pro Lys Asp Asn Gly Pro Phe Ser 2340 2345 2350	9
Leu Ala Gln Glu Ile Asp Lys Leu Val Ser Gln Gly Ser Gly Ser Ala 2355 2360 2365	11
Gly Ser Gly Asn Asn Asn Leu Ala Phe Gly Ala Gly Thr Asp Thr Lys 2370 2375 2380	13
Thr Ser Leu Gln Ala Ser Val Ser Phe Ala Asp Leu Lys Ile Arg Glu 2385 2390 2395 2400	15
Asp Tyr Pro Ala Ser Leu Gly Lys Ile Arg Arg Ile Lys Gln Ile Ser 2405 2410 2415	17
Val Thr Leu Pro Ala Leu Leu Gly Pro Tyr Gln Asp Val Gln Ala Ile 2420 2425 2430	19
Leu Ser Tyr Gly Asp Lys Ala Gly Leu Ala Asn Gly Cys Glu Ala Leu 2435 2440 2445	21
Ala Val Ser His Gly Met Asn Asp Ser Gly Gln Phe Gln Leu Asp Phe 2450 2455 2460	23
Asn Asp Gly Lys Phe Leu Pro Phe Glu Gly Ile Ala Ile Asp Gln Gly 2465 2470 2475 2480	25
Thr Leu Thr Leu Ser Phe Pro Asn Ala Ser Met Pro Glu Lys Gly Lys 2485 2490 2495	27
Gln Ala Thr Met Leu Lys Thr Leu Asn Asp Ile Ile Leu His Ile Arg 2500 2505 2510	29
Tyr Thr Ile Lys 2516	31

(2) INFORMAȚIA PENTRU SECV. ID. NR. 48:

(i) CARACTERISTICI DE SECVENȚĂ:

(A) LUNGIME: 5547 baze perechi

(B) TIP: acid nucleic

(C) ÎMPLETIRE: dublă

(D) TOPOLOGIE: lineară

(ii) TIPUL MOLECULEI: ADN (genomic)

(xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 48

(regiune de codificare tcdA_{II}):

CTG ATA GGC TAT AAC AAT CAA TTT AGC GGT AGA GCC ACT CAA TAT GTT 48	
Leu Ile Gly Tyr Asn Asn Gln Phe Ser Gly Arg Ala Ser Gln Tyr Val	
1 5 10 15	45

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	47

RO 121280 B1

1	GAA	CTT	TAT	CGT	GAA	GCA	CGC	AAT	TTA	CAC	GCA	AGT	GAC	TCC	GTT	TAT	144
3	Glu	Leu	Tyr	Arg	Glu	Ala	Arg	Asn	Leu	His	Ala	Ser	Asp	Ser	Val	Tyr	
			35					40					45				
5	TAT	CTG	GAT	ACC	CGC	CGC	CCA	GAT	CTC	AAA	TCA	ATG	CCG	CTC	AGT	CAG	192
	Tyr	Leu	Asp	Thr	Arg	Arg	Pro	Asp	Leu	Lys	Ser	Met	Ala	Leu	Ser	Gln	
		50					55					60					
7	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		
9	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		
11	CTG	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			
13	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				
15	GGA	CTT	GAG	CAA	CTC	AAT	GCA	TCA	CCG	GCA	ATT	CCC	GGG	TTG	ATG	CAT	432
	Gly	Leu	Glu	Gln	Leu	Asn	Ala	Ser	Pro	Ala	Ile	Ala	Gly	Leu	Met	His	
		130					135					140					
17	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		
19	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		
21	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			
23	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				
25	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					
27	CTT	ATT	ACT	CCG	GTA	GTC	AAC	AGC	AGT	GAT	GGC	ACG	GTT	AAG	GTA	TAT	720
	Leu	Ile	Thr	Pro	Val</												

RO 121280 B1

305	310	315	320	
GGC CTG ACA CGA GTA CTT CCT TCC GGT TCT TCG GCA TAT GCC GCC GCA 1008				1
Gly Leu Thr Arg Val Leu Pro Ser Gly Ser Trp Ala Tyr Ala Ala Ala				3
325		330	335	5
AAA TTT ACC GTT GAA GAG TAT AAC CAA TAC TCT TTT CTG CTA AAA CTT 1056				7
Lys Phe Thr Val Glu Glu Tyr Asn Gln Tyr Ser Phe Leu Leu Lys Leu				9
340		345	350	11
AAC AAG GCT ATT CCT CTA TCA CGT GCG ACA GAA TTG TCA CCC ACG ATT 1104				13
Asn Lys Ala Ile Arg Leu Ser Arg Ala Thr Glu Leu Ser Pro Thr Ile				15
355		360	365	17
CTG GAA GGC ATT GTG GCG AGT GTT AAT CTA CAA CTG GAT ATC AAC ACA 1152				19
Leu Glu Gly Ile Val Arg Ser Val Asn Leu Gln Leu Asp Ile Asn Thr				21
370		375	380	23
GAC GTA TTA GGT AAA GTT TTT CTG ACT AAA TAT TAT ATG CAG CGT TAT 1200				25
Asp Val Leu Gly Lys Val Phe Leu Thr Lys Tyr Tyr Met Gln Arg Tyr				27
385		390	395	29
GCT ATT CAT GCT GAA ACT GCC CTG ATA CTA TCC AAC CCG CCT ATT TCA 1248				31
Ala Ile His Ala Glu Thr Ala Leu Ile Leu Cys Asn Ala Pro Ile Ser				33
405		410	415	35
CAA CGT TCA TAT GAT AAT CAA CCT AGC CAA TTT GAT CCG CTG TTT AAT 1296				37
Gln Arg Ser Tyr Asp Asn Gln Pro Ser Gln Phe Asp Arg Leu Phe Asn				39
420		425	430	41
ACG CCA TTA CTG AAC GGA CAA TAT TTT TCT ACC GGC GAT GAG GAG ATT 1344				43
Thr Pro Leu Leu Asn Gly Gln Tyr Phe Ser Thr Gly Asp Glu Glu Ile				45
435		440	445	47
GAT TTA AAT TCA GGT AGC ACC GCG GAT TGG CGA AAA ACC ATA CTT AAG 1392				49
Asp Leu Asn Ser Gly Ser Thr Gly Asp Trp Arg Lys Thr Ile Leu Lys				51
450		455	460	53
CGT GCA TTT AAT ATT GAT GAT GTC TCG CTC TTC CCG CTG CTT AAA ATT 1440				55
Arg Ala Phe Asn Ile Asp Asp Val Ser Leu Phe Arg Leu Leu Lys Ile				57
465		470	475	59
ACC GAC CAT GAT AAT AAA GAT GGA AAA ATT AAA AAT AAC CTA AAG AAT 1488				61
Thr Asp His Asp Asn Lys Asp Gly Lys Ile Lys Asn Asn Leu Lys Asn				63
485		490	495	65
CTT TCC AAT TTA TAT ATT GGA AAA TTA CTG GCA GAT ATT CAT CAA TTA 1536				67
Leu Ser Asn Leu Tyr Ile Gly Lys Leu Leu Ala Asp Ile His Gln Leu				69
500		505	510	71
ACC ATT GAT GAA CTG GAT TTA TTA CTG ATT GCC GTA GGT GAA GGA AAA 1584				73
Thr Ile Asp Glu Leu Asp Leu Leu Leu Ile Ala Val Gly Glu Gly Lys				75
515		520	525	77
ACT AAT TTA TCC GCT ATC AGT GAT AAG CAA TTG GCT ACC CTG ATC AGA 1632				79
Thr Asn Leu Ser Ala Ile Ser Asp Lys Gln Leu Ala Thr Leu Ile Arg				81
530		535	540	83
AAA CTC AAT ACT ATT ACC AGC TGG CTA CAT ACA CAG AAG TGG AGT GTA 1680				85
Lys Leu Asn Thr Ile Thr Ser Trp Leu His Thr Gln Lys Trp Ser Val				87
545		550	555	89
TTC CAG CTA TTT ATC ATG ACC TCC ACC AGC TAT AAC AAA ACG CTA ACG 1728				91
Phe Gln Leu Phe Ile Met Thr Ser Thr Ser Tyr Asn Lys Thr Leu Thr				93
565		570	575	95
CCT GAA ATT AAG AAT TTG CTG GAT ACC GTC TAC CAC GGT TTA CAA GGT 1776				97
Pro Glu Ile Lys Asn Leu Leu Asp Thr Val Tyr His Gly Leu Gln Gly				99
580		585	590	101
TTT GAT AAA GAC AAA GCA GAT TTG CTA CAT GTC ATG GCG CCC TAT ATT 1824				103

RO 121280 B1

1	Phe	Asp	Lys	Asp	Lys	Ala	Asp	Leu	Leu	His	Val	Met	Ala	Pro	Tyr	Ile	
3			595					600					605				
5	GCG	GCC	ACC	TTG	CAA	TTA	TCA	TCC	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					
7	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	
9	AAA	TTC	TGG	GAC	TGG	TTG	AAT	ACT	AAG	TAT	ACG	CCG	GGT	TCA	TCG	GAA	1963
	Lys	Phe	Trp	Asp	Trp	Leu	Asn	Thr	Lys	Tyr	Thr	Pro	Gly	Ser	Ser	Glu	
					645					650						655	
11	GCC	CTA	GAA	ACG	CAG	GAA	CAT	ATC	CTT	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			
13	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				
15	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					
17	GCG	CCC	GCG	CAT	GAT	GCC	CTT	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	
19	GAT	TGG	GTG	AAC	GCA	CTA	GCC	GAA	AAA	GCG	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		
21	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			
23	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					
25	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					
27	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	
	785				790						795					800	
29	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		
31	TCA	ATG	AAA	GAG	ACA	CCG	ACC	TAT	GCC	CAG	TGG	GAA	AAC	GCG	GCA	GCC	2496
	Ser	Met	Lys	Glu	Thr	Pro	Thr	Tyr	Ala	Gln	Trp	Glu	Asn	Ala	Ala	Gly	
				820				825						830			
33	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					
35	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					
37	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	
	865				870					875						880	

RO 121280 B1

CAA TAC TTA CTG ATT GAT AAT CAG GTT TCT GCG GCA ATA AAA ACC ACC 2635	1
Gln Tyr Leu Leu Ile Asp Asn Gln Val Ser Ala Ala Ile Lys Thr Thr 885 890 895	3
CGG ATC GCC GAA GCC ATT CCC AGT ATT CAA CTG TAC CTC AAC CCG GCA 2735	5
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 ACC CCG CAA 2781	7
Leu Glu Asn Val Glu Glu Asn Ala Asn Ser Gly Val Ile Ser Arg Gln 915 920 925	
TTT TTT ATC GAC TGG GAC AAA TAC AAT AAA CCG TAC AGC ACT TCG CCG 2832	9
Phe Phe Ile Asp Trp Asp Lys Tyr Asn Lys Arg Tyr Ser Thr Trp Ala 930 935 940	11
GGT GTT TCT CAA TTA GTT TAC TAC CCG GAA AAC TAT ATT GAT CCG ACC 2880	13
Gly Val Ser Gln Leu Val Tyr Tyr Pro Glu Asn Tyr Ile Asp Pro Thr 945 950 955 960	
ATG CGT ATC GGA CAA ACC AAA ATG ATG GAC GCA TTA CTG CAA TCC GTC 2928	15
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 ATC TCT 2976	17
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	19
Tyr Leu Thr Ser Phe Glu Gln Val Ala Asn Leu Lys Val Ile Ser Ala 995 1000 1005	21
TAT CAC GAT AAT ATT AAT AAC GAT CAA CCG CTG ACC TAT TTT ATC CGA 3072	23
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 TCG CCG AGT CTC GAT CAC 3120	25
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 TCG 3158	27
Ser Lys Phe Asn Asp Gly Lys Phe Ala Ala Asn Ala Trp Ser Glu Trp 1045 1050 1055	29
CAT AAA ATT GAT TGT CCA ATT AAC CCT TAT AAA AGC ACT ATC CGT CCA 3216	31
His Lys Ile Asp Cys Pro Ile Asn Pro Tyr Lys Ser Thr Ile Arg Pro 1060 1065 1070	
GTG ATA TAT AAA TCC CCG CTG TAT CTG CTC TGG TTG GAA CAA AAG GAG 3264	33
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 ACC 3312	35
Ile Thr Lys Gln Thr Gly Asn Ser Lys Asp Gly Tyr Gln Thr Glu Thr 1090 1095 1100	37
GAT TAT CGT TAT GAA CTA AAA TTG GCG CAT ATC CCG TAT GAT GGC ACT 3360	39
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	41
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	43
Lys Leu Glu Lys Asn Arg Ala Pro Gly Leu Tyr Cys Ala Gly Tyr Gln 1140 1145 1150	
GGT GAA GAT ACC TTG CTG GTG ATG TTT TAT AAC CAA CAA GAC ACA CTA 3501	45
Gly Glu Asp Thr Leu Leu Val Met Phe Tyr Asn Gln Gln Asp Thr Leu 1155 1160 1165	47

RO 121280 B1

1
3 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

5 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

7 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 Val Arg Arg Val Asn Asn
1205 1210 1215

9 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

11 GAC TAT GGT TCG GCA 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

13 ATT CCA 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

15 ATC TCA CCA AAA TTA AGA ATT ATT CAT AAT GCA TAT GAA GCA CAG AAG 3840
Ile Ser Pro Lys Leu Arg Ile Ile His Asn Gly Tyr Glu Gly Gln Lys
1265 1270 1275 1280

17 CGC AAT CAA TGC AAT CTG ATG AAT AAA TAT GGC AAA CTA GGT GAT AAA 3888
Arg Asn Gln Cys Asn Leu Met Asn Lys Tyr Gly Lys Leu Gly Asp Lys
1285 1290 1295

19 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

21 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

23 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

25 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

27 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

29 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

31 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

33 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

35 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

37 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

RO 121280 B1

1445	1450	1455	
ATT AAC AAC TCA GCC AGT ATT GAT GTT ACT TTT ACC GCA TTT GCG GAG 1446			1
Ile Asn Asn Ser Ala Ser Ile Asp Val Thr Phe Thr Ala Phe Ala Glu			3
1460	1465	1470	
GAT GGC CGC AAA CTG GGT TAT GAA AGT TTC AGT ATT CCT GTT ACC CTC 1464			5
Asp Gly Arg Lys Leu Gly Tyr Glu Ser Phe Ser Ile Pro Val Thr Leu			7
1475	1480	1485	
AAG GTA AGT ACC GAT AAT GCC CTG ACC CTG CAC CAT AAT GAA AAT GGT 4512			9
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			11
Ala Gln Tyr Met Gln Trp Gln Ser Tyr Arg Thr Arg Leu Asn Thr Leu			
1505	1510	1515	
TTT GCC CGC CAG TTG GTT GCA CGC GCC ACC ACC GGA ATC GAT ACA ATT 4608			13
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			15
Leu Ser Met Glu Thr Gln Asn Ile Gln Glu Pro Gln Leu Gly Lys Gly			17
1540	1545	1550	
TTC TAT GCT ACG TTC GTG ATA CCT CCC TAT AAC CTA TCA ACT CAT GGT 4704			19
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			21
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			23
Ser His Ile Ile Tyr Ser Gly Gln Leu Thr Asp Thr Asn Ile Asn Ile			
1585	1590	1595	
ACA TTA TTT ATT CCT CTT GAT GAT GTC CCA TTG AAT CAA GAT TAT CAC 4848			25
Thr Leu Phe Ile Pro Leu Asp Asp Val Pro Leu Asn Gln Asp Tyr His			
1605	1610	1615	
GCC AAG GTT TAT ATG ACC TTC AAG AAA TCA CCA TCA GAT GGT ACC TGG 4896			27
Ala Lys Val Tyr Met Thr Phe Lys Lys Ser Pro Ser Asp Gly Thr Trp			29
1620	1625	1630	
TGG GGC CCT CAC TTT GTT AGA GAT GAT AAA GGA ATA GTA ACA ATA AAC 4944			31
Trp Gly Pro His Phe Val Arg Asp Asp Lys Gly Ile Val Thr Ile Asn			
1635	1640	1645	
CCT AAA TCC ATT TTG ACC CAT TTT GAG AGC GTC AAT GTC CTG AAT AAT 4992			33
Pro Lys Ser Ile Leu Thr His Phe Glu Ser Val Asn Val Leu Asn Asn			
1650	1655	1660	
ATT AGT AGC GAA CCA ATG GAT TTC AGC GGC GCT AAC AGC CTC TAT TTC 5040			35
Ile Ser Ser Glu Pro Met Asp Phe Ser Gly Ala Asn Ser Leu Tyr Phe			37
1665	1670	1675	
TGG GAA CTG TTC TAC TAT ACC CCG ATG CTG GTT GCT CAA CGT TTG CTG 5088			39
Trp Glu Leu Phe Tyr Tyr Thr Pro Met Leu Val Ala Gln Arg Leu Leu			
1685	1690	1695	
CAT GAA CAG AAC TTC GAT GAA GCC AAC CGT TGG CTG AAA TAT GTC TGG 5136			41
His Glu Gln Asn Phe Asp Glu Ala Asn Arg Trp Leu Lys Tyr Val Trp			
1700	1705	1710	
AGT CCA TCC GGT TAT ATT GTC CAC GGC CAG ATT CAG AAC TAC CAG TGG 5184			43
Ser Pro Ser Gly Tyr Ile Val His Gly Gln Ile Gln Asn Tyr Gln Trp			
1715	1720	1725	
AAC GTC CGC CCG TTA CTG GAA GAC ACC AGT TGG AAC AGT GAT CCT TTG 5232			45
			47

1 Asn Val Arg Pro Leu Leu Glu Asp Thr Ser Trp Asn Ser Asp Pro Leu
1730 1735 1740

3 GAT TCC GTC GAT CCT GAC GCG GTA GCA CAG CAC GAT CCA ATG CAC TAC 5230
Asp Ser Val Asp Pro Asp Ala Val Ala Gln His Asp Pro Met His Tyr
1745 1750 1755 1760

5 AAA GTT TCA ACT TTT ATG CGT ACC TTG GAT CTA TTG ATA GCA CCG GGC 5328
Lys Val Ser Thr Phe Met Arg Thr Leu Asp Leu Leu Ile Ala Arg Gly
1765 1770 1775

7 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
1780 1785 1790

9 ATC 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
1795 1800 1805

11 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
1810 1815 1820

13 ATC ACT ACC CAA AAT GCT CAC GAC AGC GCA ATA GTC GCT CTG CCG CAG 5520
Ile Thr Thr Gln Asn Ala His Asp Ser Ala Ile Val Ala Leu Arg Gln
1825 1830 1835 1840

15 AAT ATA CCT ACA CCG GCA CCT TTA TCA 5547
Asn Ile Pro Thr Pro Ala Pro Leu Ser
1845 1849

(2) INFORMAȚIA PENTRU SECV. ID. NR. 49:

(i) CARACTERISTICI DE SECVENȚĂ:

(A) LUNGIME: 1849 baze perechi

(B) TIP: acid nucleic

(C) ÎMPLETIRE: simplă

(D) TOPOLOGIE: lineară

(ii) TIPUL MOLECULEI: proteină

(xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 49: (TcdA_{ii})

Țesătura	De la	Până la	Descrierea
Peptidă	1	1849	Peptidă TcdA _{ii}
Fragment	1	12	S2 N-terminus (SECV. ID NR. 13)
Fragment	196	211	(SECV. ID NR. 38)
Fragment	466	475	(SECV. ID NR. 17)
Fragment	993	1004	(SECV. ID NR. 23; 12/13)
Fragment	1297	1312	(SECV. ID NR. 18)
Fragment	1390	1409	(SECV. ID NR. 39)
Fragment	1532	1554	(SECV. ID NR. 21; 19/23)

39 Leu Ile Gly Tyr Asn Asn Gln Phe Ser Gly Arg Ala Ser Gln Tyr Val
1 5 10 15

41 Ala Pro Gly Thr Val Ser Ser Met Phe Ser Pro Ala Ala Tyr Leu Thr
20 25 30

43 Glu Leu Tyr Arg Glu Ala Arg Asn Leu His Ala Ser Asp Ser Val Tyr
35 40 45

45 Tyr Leu Asp Thr Arg Arg Pro Asp Leu Lys Ser Met Ala Leu Ser Gln
50 55 60

47 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

RO 121280 B1

	85	90	95	1
Val Met Glu Met Leu Ser Thr Phe Arg Pro Ser Gly Ala Thr Pro Tyr	100	105	110	3
His Asp Ala Tyr Glu Asn Val Arg Glu Val Ile Gln Leu Gln Asp Pro	115	120	125	5
Gly Leu Glu Gln Leu Asn Ala Ser Pro Ala Ile Ala Gly Leu Met His	130	135	140	7
Gln Ala Ser Leu Leu Gly Ile Asn Ala Ser Ile Ser Pro Glu Leu Phe	145	150	155	9
Asn Ile Leu Thr Glu Glu Ile Thr Glu Gly Asn Ala Glu Glu Leu Tyr	165	170	175	11
Lys Lys Asn Phe Gly Asn Ile Glu Pro Ala Ser Leu Ala Met Pro Glu	180	185	190	13
Tyr Leu Lys Arg Tyr Tyr Asn Leu Ser Asp Glu Glu Leu Ser Gln Phe	195	200	205	15
Ile Gly Lys Ala Ser Asn Phe Gly Gln Gln Glu Tyr Ser Asn Asn Gln	210	215	220	17
Leu Ile Thr Pro Val Val Asn Ser Ser Asp Gly Thr Val Lys Val Tyr	225	230	235	19
Arg Ile Thr Arg Glu Tyr Thr Thr Asn Ala Tyr Gln Met Asp Val Glu	245	250	255	21
Leu Phe Pro Phe Gly Gly Glu Asn Tyr Arg Leu Asp Tyr Lys Phe Lys	260	265	270	23
Asn Phe Tyr Asn Ala Ser Tyr Leu Ser Ile Lys Leu Asn Asp Lys Arg	275	280	285	25
Glu Leu Val Arg Thr Glu Gly Ala Pro Gln Val Asn Ile Glu Tyr Ser	290	295	300	27
Ala Asn Ile Thr Leu Asn Thr Ala Asp Ile Ser Gln Pro Phe Glu Ile	305	310	315	29
Gly Leu Thr Arg Val Leu Pro Ser Gly Ser Trp Ala Tyr Ala Ala Ala	325	330	335	31
Lys Phe Thr Val Glu Glu Tyr Asn Gln Tyr Ser Phe Leu Leu Lys Leu	340	345	350	33
Asn Lys Ala Ile Arg Leu Ser Arg Ala Thr Glu Leu Ser Pro Thr Ile	355	360	365	35
Leu Glu Gly Ile Val Arg Ser Val Asn Leu Gln Leu Asp Ile Asn Thr	370	375	380	37
Asp Val Leu Gly Lys Val Phe Leu Thr Lys Tyr Tyr Met Gln Arg Tyr	385	390	395	39
Ala Ile His Ala Glu Thr Ala Leu Ile Leu Cys Asn Ala Pro Ile Ser	405	410	415	41
Gln Arg Ser Tyr Asp Asn Gln Pro Ser Gln Phe Asp Arg Leu Phe Asn	420	425	430	43
Thr Pro Leu Leu Asn Gly Gln Tyr Phe Ser Thr Gly Asp Glu Glu Ile	435	440	445	45
Asp Leu Asn Ser Gly Ser Thr Gly Asp Trp Arg Lys Thr Ile Leu Lys	450	455	460	47

RO 121280 B1

1 Arg Ala Phe Asn Ile Asp Asp Val Ser Leu Phe Arg Leu Leu Lys Ile
 465 470 475 480
 3 Thr Asp His Asp Asn Lys Asp Gly Lys Ile Lys Asn Asn Leu Lys Asn
 485 490 495
 5 Leu Ser Asn Leu Tyr Ile Gly Lys Leu Leu Ala Asp Ile His Gln Leu
 500 505 510
 7 Thr Ile Asp Glu Leu Asp Leu Leu Leu Ile Ala Val Gly Glu Gly Lys
 515 520 525
 9 Thr Asn Leu Ser Ala Ile Ser Asp Lys Gln Leu Ala Thr Leu Ile Arg
 530 535 540
 11 Lys Leu Asn Thr Ile Thr Ser Trp Leu His Thr Gln Lys Trp Ser Val
 545 550 555 560
 13 Phe Gln Leu Phe Ile Met Thr Ser Thr Ser Tyr Asn Lys Thr Leu Thr
 565 570 575
 15 Pro Glu Ile Lys Asn Leu Leu Asp Thr Val Tyr His Gly Leu Gln Gly
 580 585 590
 17 Phe Asp Lys Asp Lys Ala Asp Leu Leu His Val Met Ala Pro Tyr Ile
 595 600 605
 19 Ala Ala Thr Leu Gln Leu Ser Ser Glu Asn Val Ala His Ser Val Leu
 610 615 620
 21 Leu Trp Ala Asp Lys Leu Gln Pro Gly Asp Gly Ala Met Thr Ala Glu
 625 630 635 640
 23 Lys Phe Trp Asp Trp Leu Asn Thr Lys Tyr Thr Pro Gly Ser Ser Glu
 645 650 655
 25 Ala Val Glu Thr Gln Glu His Ile Val Gln Tyr Cys Gln Ala Leu Ala
 660 665 670
 27 Gln Leu Glu Met Val Tyr His Ser Thr Gly Ile Asn Glu Asn Ala Phe
 675 680 685
 29 Arg Leu Phe Val Thr Lys Pro Glu Met Phe Gly Ala Ala Thr Gly Ala
 690 695 700
 31 Ala Pro Ala His Asp Ala Leu Ser Leu Ile Met Leu Thr Arg Phe Ala
 705 710 715 720
 33 Asp Trp Val Asn Ala Leu Gly Glu Lys Ala Ser Ser Val Leu Ala Ala
 725 730 735
 35 Phe Glu Ala Asn Ser Leu Thr Ala Glu Gln Leu Ala Asp Ala Met Asn
 740 745 750
 37 Leu Asp Ala Asn Leu Leu Leu Gln Ala Ser Ile Gln Ala Gln Asn His
 755 760 765
 39 Gln His Leu Pro Pro Val Thr Pro Glu Asn Ala Phe Ser Cys Trp Thr
 770 775 780
 41 Ser Ile Asn Thr Ile Leu Gln Trp Val Asn Val Ala Gln Gln Leu Asn
 785 790 795 800
 43 Val Ala Pro Gln Gly Val Ser Ala Leu Val Gly Leu Asp Tyr Ile Gln
 805 810 815
 45 Ser Met Lys Glu Thr Pro Thr Tyr Ala Gln Trp Glu Asn Ala Ala Gly
 820 825 830
 47 Val Leu Thr Ala Gly Leu Asn Ser Gln Gln Ala Asn Thr Leu His Ala
 835 840 845

RO 121280 B1

Phe	Leu	Asp	Glu	Ser	Arg	Ser	Ala	Ala	Leu	Ser	Thr	Tyr	Tyr	Ile	Arg	1
850						355					860					
Gln	Val	Ala	Lys	Ala	Ala	Ala	Ala	Ile	Lys	Ser	Arg	Asp	Asp	Leu	Tyr	3
865					370				875					880		
Gln	Tyr	Leu	Leu	Ile	Asp	Asn	Gln	Val	Ser	Ala	Ala	Ile	Lys	Thr	Thr	5
				885					890					895		
Arg	Ile	Ala	Glu	Ala	Ile	Ala	Ser	Ile	Gln	Leu	Tyr	Val	Asn	Arg	Ala	7
			900					905					910			
Leu	Glu	Asn	Val	Glu	Glu	Asn	Ala	Asn	Ser	Gly	Val	Ile	Ser	Arg	Gln	9
		915					920					925				
Phe	Phe	Ile	Asp	Trp	Asp	Lys	Tyr	Asn	Lys	Arg	Tyr	Ser	Thr	Trp	Ala	11
930						935					940					
Gly	Val	Ser	Gln	Leu	Val	Tyr	Tyr	Pro	Glu	Asn	Tyr	Ile	Asp	Pro	Thr	13
945					950					955				960		
Met	Arg	Ile	Gly	Gln	Thr	Lys	Met	Met	Asp	Ala	Leu	Leu	Gln	Ser	Val	15
				965					970					975		
Ser	Gln	Ser	Gln	Leu	Asn	Ala	Asp	Thr	Val	Glu	Asp	Ala	Phe	Met	Ser	17
				980				985					990			
Tyr	Leu	Thr	Ser	Phe	Glu	Gln	Val	Ala	Asn	Leu	Lys	Val	Ile	Ser	Ala	19
		995					1000						1005			
Tyr	His	Asp	Asn	Ile	Asn	Asn	Asp	Gln	Gly	Leu	Thr	Tyr	Phe	Ile	Gly	21
	1010					1015					1020					
Leu	Ser	Glu	Thr	Asp	Ala	Gly	Glu	Tyr	Tyr	Trp	Arg	Ser	Val	Asp	His	23
1025					1030					1035				1040		
Ser	Lys	Phe	Asn	Asp	Gly	Lys	Phe	Ala	Ala	Asn	Ala	Trp	Ser	Glu	Trp	25
				1045				1050						1055		
His	Lys	Ile	Asp	Cys	Pro	Ile	Asn	Pro	Tyr	Lys	Ser	Thr	Ile	Arg	Pro	27
			1060					1065					1070			
Val	Ile	Tyr	Lys	Ser	Arg	Leu	Tyr	Leu	Leu	Trp	Leu	Glu	Gln	Lys	Glu	29
		1075					1080					1085				
Ile	Thr	Lys	Gln	Thr	Gly	Asn	Ser	Lys	Asp	Gly	Tyr	Gln	Thr	Glu	Thr	31
	1090					1095					1100					
Asp	Tyr	Arg	Tyr	Glu	Leu	Lys	Leu	Ala	His	Ile	Arg	Tyr	Asp	Gly	Thr	33
1105					1110					1115				1120		
Trp	Asn	Thr	Pro	Ile	Thr	Phe	Asp	Val	Asn	Lys	Lys	Ile	Ser	Glu	Leu	35
				1125					1130					1135		
Lys	Leu	Glu	Lys	Asn	Arg	Ala	Pro	Gly	Leu	Tyr	Cys	Ala	Gly	Tyr	Gln	37
			1140					1145					1150			
Gly	Glu	Asp	Thr	Leu	Leu	Val	Met	Phe	Tyr	Asn	Gln	Gln	Asp	Thr	Leu	39
		1155					1160					1165				
Asp	Ser	Tyr	Lys	Asn	Ala	Ser	Met	Gln	Gly	Leu	Tyr	Ile	Phe	Ala	Asp	41
	1170					1175					1180					
Met	Ala	Ser	Lys	Asp	Met	Thr	Pro	Glu	Gln	Ser	Asn	Val	Tyr	Arg	Asp	43
1185					1190					1195				1200		
Asn	Ser	Tyr	Gln	Gln	Phe	Asp	Thr	Asn	Asn	Val	Arg	Arg	Val	Asn	Asn	45
				1205					1210					1215		
Arg	Tyr	Ala	Glu	Asp	Tyr	Glu	Ile	Pro	Ser	Ser	Val	Ser	Ser	Arg	Lys	47
			1220					1225						1230		

RO 121280 B1

1 Asp Tyr Gly Trp Gly Asp Tyr Tyr Leu Ser Met Val Tyr Asn Gly Asp
 1235 1240 1245
 3 Ile Pro Thr Ile Asn Tyr Lys Ala Ala Ser Ser Asp Leu Lys Ile Tyr
 1250 1255 1260
 5 Ile Ser Pro Lys Leu Arg Ile Ile His Asn Gly Tyr Glu Gly Gln Lys
 1265 1270 1275 1280
 7 Arg Asn Gln Cys Asn Leu Met Asn Lys Tyr Gly Lys Leu Gly Asp Lys
 1285 1290 1295
 9 Phe Ile Val Tyr Thr Ser Leu Gly Val Asn Pro Asn Asn Ser Ser Asn
 1300 1305 1310
 11 Lys Leu Met Phe Tyr Pro Val Tyr Gln Tyr Ser Gly Asn Thr Ser Gly
 1315 1320 1325
 13 Leu Asn Gln Gly Arg Leu Leu Phe His Arg Asp Thr Thr Tyr Pro Ser
 1330 1335 1340
 15 Lys Val Glu Ala Trp Ile Pro Gly Ala Lys Arg Ser Leu Thr Asn Gln
 1345 1350 1355 1360
 17 Asn Ala Ala Ile Gly Asp Asp Tyr Ala Thr Asp Ser Leu Asn Lys Pro
 1365 1370 1375
 19 Asp Asp Leu Lys Gln Tyr Ile Phe Met Thr Asp Ser Lys Gly Thr Ala
 1380 1385 1390
 21 Thr Asp Val Ser Gly Pro Val Glu Ile Asn Thr Ala Ile Ser Pro Ala
 1395 1400 1405
 23 Lys Val Gln Ile Ile Val Lys Ala Gly Gly Lys Glu Gln Thr Phe Thr
 1410 1415 1420
 25 Ala Asp Lys Asp Val Ser Ile Gln Pro Ser Pro Ser Phe Asp Glu Met
 1425 1430 1435 1440
 27 Asn Tyr Gln Phe Asn Ala Leu Glu Ile Asp Gly Ser Gly Leu Asn Phe
 1445 1450 1455
 29 Ile Asn Asn Ser Ala Ser Ile Asp Val Thr Phe Thr Ala Phe Ala Glu
 1460 1465 1470
 31 Asp Gly Arg Lys Leu Gly Tyr Glu Ser Phe Ser Ile Pro Val Thr Leu
 1475 1480 1485
 33 Lys Val Ser Thr Asp Asn Ala Leu Thr Leu His His Asn Glu Asn Gly
 1490 1495 1500
 35 Ala Gln Tyr Met Gln Trp Gln Ser Tyr Arg Thr Arg Leu Asn Thr Leu
 1505 1510 1515 1520
 37 Phe Ala Arg Gln Leu Val Ala Arg Ala Thr Thr Gly Ile Asp Thr Ile
 1525 1530 1535
 39 Leu Ser Met Glu Thr Gln Asn Ile Gln Glu Pro Gln Leu Gly Lys Gly
 1540 1545 1550
 41 Phe Tyr Ala Thr Phe Val Ile Pro Pro Tyr Asn Leu Ser Thr His Gly
 1555 1560 1565
 43 Asp Glu Arg Trp Phe Lys Leu Tyr Ile Lys His Val Val Asp Asn Asn
 1570 1575 1580
 45 Ser His Ile Ile Tyr Ser Gly Gln Leu Thr Asp Thr Asn Ile Asn Ile
 1585 1590 1595 1600
 47 Thr Leu Phe Ile Pro Leu Asp Asp Val Pro Leu Asn Gln Asp Tyr His

RO 121280 B1

	1605	1610	1615	1
Ala Lys Val Tyr Met Thr Phe Lys Lys Ser Pro Ser Asp Gly Thr Trp	1620	1625	1630	3
Trp Gly Pro His Phe Val Arg Asp Asp Lys Gly Ile Val Thr Ile Asn	1635	1640	1645	5
Pro Lys Ser Ile Leu Thr His Phe Glu Ser Val Asn Val Leu Asn Asn	1650	1655	1660	7
Ile Ser Ser Glu Pro Met Asp Phe Ser Gly Ala Asn Ser Leu Tyr Phe	1665	1670	1675	9
Trp Glu Leu Phe Tyr Tyr Thr Pro Met Leu Val Ala Gln Arg Leu Leu	1685	1690	1695	11
His Glu Gln Asn Phe Asp Glu Ala Asn Arg Trp Leu Lys Tyr Val Trp	1700	1705	1710	13
Ser Pro Ser Gly Tyr Ile Val His Gly Gln Ile Gln Asn Tyr Gln Trp	1715	1720	1725	15
Asn Val Arg Pro Leu Leu Glu Asp Thr Ser Trp Asn Ser Asp Pro Leu	1730	1735	1740	17
Asp Ser Val Asp Pro Asp Ala Val Ala Gln His Asp Pro Met His Tyr	1745	1750	1755	19
Lys Val Ser Thr Phe Met Arg Thr Leu Asp Leu Leu Ile Ala Arg Gly	1765	1770	1775	21
Asp His Ala Tyr Arg Gln Leu Glu Arg Asp Thr Leu Asn Glu Ala Lys	1780	1785	1790	23
Met Trp Tyr Met Gln Ala Leu His Leu Leu Gly Asp Lys Pro Tyr Leu	1795	1800	1805	25
Pro Leu Ser Thr Thr Trp Ser Asp Pro Arg Leu Asp Arg Ala Ala Asp	1810	1815	1820	27
Ile Thr Thr Gln Asn Ala His Asp Ser Ala Ile Val Ala Leu Arg Gln	1825	1830	1835	29
Asn Ile Pro Thr Pro Ala Pro Leu Ser	1845	1849		31
(2) INFORMAȚIA PENTRU SECV. ID. NR. 50:				33
(i) CARACTERISTICI DE SECVENȚĂ:				
(A) LUNGIME: 1740 baze perechi				35
(B) TIP: acid nucleic				
(C) ÎMPLETIRE: dublă				37
(D) TOPOLOGIE: lineară				
(ii) TIPUL MOLECULEI: ADN (genomic)				39
(xi) DESCRIEREA SECVENȚEI SECV. ID. NR. 50				
(regiune de codificare tcdA _{III}):				41
TTG CGC AGC GCT AAT ACC CTG ACT GAT CTC TTC CTG CCG CAA ATC AAT 48				43
Leu Arg Ser Ala Asn Thr Leu Thr Asp Leu Phe Leu Pro Gln Ile Asn	1	5	10	15
GAA GTG ATC ATG AAT TAC TGG CAG ACA TTA GCT CAG ACA CTA TAC AAT 96				45
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				47

RO 121280 B1

1 Leu Arg His Asn Leu Ser Ile Asp Gly Gln Pro Leu Tyr Leu Pro Ile
 35 40 45
 3 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
 5 ACT TCT CAA GGT GGA GGC AAG CTA CCG GAA TCA TTT ATG TCC CTG TGG 240
 Thr Ser Gln Gly Gly Lys Leu Pro Glu Ser Phe Met Ser Leu Trp
 55 70 75 80
 7 CGT TTC CCG CAC ATG CTG GAA AAT GCG CCG 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
 9 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
 11 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
 13 ACT AAC CTG AGC ATT CAG CAC AAA ACC ATT GAA GAA TTC GAT CCC CAG 432
 Thr Asn Leu Ser Ile Gln Asp Lys Thr Ile Glu Glu Leu Asp Ala Glu
 130 135 140
 15 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
 17 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
 19 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
 21 GCA TCC CGT CTG GCC GGT CCG GCG GCT GAT CTG GTG CCT AAC ATC TTC 624
 Ala Ser Arg Leu Ala Gly Ala Ala Asp Leu Val Pro Asn Ile Phe
 195 200 205
 23 GGC TTT GCC GGT GCC GCG AGC CGT TCG GCG 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
 25 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
 27 AAA ATT AGC CAA TCT GAA ACC TAC CGT CGT CCG CGT CAG GAG TCG GAG 768
 Lys Ile Ser Gln Ser Glu Thr Tyr Arg Arg Arg Arg Gln Glu Trp Glu
 245 250 255
 29 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
 31 CTC AAA TCA CTC GCT GTA CCG 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
 33 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
 35 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

RO 121280 B1

CTG GCG GCG ATT TAC TTC CAG TTC TAC GAT TTG GCC GTC GCG CGT TGC 1308	1
Leu Ala Ala Ile Tyr Phe Gln Phe Tyr Asp Leu Ala Val Ala Arg Cys	3
325 330 335	
CTG ATG GCA GAA CAA GCT TAC CGT TCG GAA CTC AAT GAT GAC TCT GCC 1056	5
Leu Met Ala Glu Gln Ala Tyr Arg Trp Glu Leu Asn Asp Asp Ser Ala	
340 345 350	
CGC TTC ATT AAA CCG GCG GCC TCG CAG GGA ACC TAT GCC CGT CTG CTT 1104	7
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	9
Ala Gly Glu Thr Leu Met Leu Ser Leu Ala Gln Met Glu Asp Ala His	
370 375 380	
CTG AAA GCG GAT AAA GCG GCA TTA GAG GTT GAA CCG ACA GTA TCG CTG 1200	11
Leu Lys Arg Asp Lys Arg Ala Leu Glu Val Glu Arg Thr Val Ser Leu	13
385 390 395 400	
GCC GAA GTT TAT GCA GGA TTA CCA AAA GAT AAC GGT CCA TTT TCC CTG 1248	15
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 GCG AGT GCC GGC 1296	17
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 GCG GCC GCG ACC GAC ACT AAA ACC 1344	19
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	21
Ser Leu Gln Ala Ser Val Ser Phe Ala Asp Leu Lys Ile Arg Glu Asp	23
450 455 460	
TAC CCG GCA TCG CTT GCG AAA ATT CGA CGT ATC AAA CAG ATC AGC GTC 1440	25
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	27
Thr Leu Pro Ala Leu Leu Gly Pro Tyr Gln Asp Val Gln Ala Ile Leu	
485 490 495	
TCT TAC GCG GAT AAA GCC GGA TTA GCT AAC GCG TGT GAA GCG CTG GCA 1536	29
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	31
Val Ser His Gly Met Asn Asp Ser Gly Gln Phe Gln Leu Asp Phe Asn	33
515 520 525	
GAT GCG AAA TTC CTG CCA TTC GAA GCG ATC GCC ATT GAT CAA GCG ACG 1632	35
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	37
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 GCG TAC 1728	39
Ala Thr Met Leu Lys Thr Leu Asn Asp Ile Ile Leu His Ile Arg Tyr	
565 570 575	
ACC ATT AAA TAA 1740	41
Thr Ile Lys ...	
579	
(i) CARACTERISTICI DE SECVENȚĂ:	43
(A) LUNGIME: 579 aminoacizi	45
(B) TIP: aminoacid	
(C) ÎMPLETIRE: simplă	47
(D) TOPOLOGIE: lineară	

RO 121280 B1

(ii) TIPUL MOLECULEI: proteină

(xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 51 (TcdA_{III}):

1	Leu	Arg	Ser	Ala	Asn	Thr	Leu	Thr	Asp	Leu	Phe	Leu	Pro	Gln	Ile	Asn
5	1				5				10						15	
7	Glu	Val	Met	Met	Asn	Tyr	Trp	Gln	Thr	Leu	Ala	Gln	Arg	Val	Tyr	Asn
			20					25					30			
9	Leu	Arg	His	Asn	Leu	Ser	Ile	Asp	Gly	Gln	Pro	Leu	Tyr	Leu	Pro	Ile
			35					40					45			
11	Tyr	Ala	Thr	Pro	Ala	Asp	Pro	Lys	Ala	Leu	Leu	Ser	Ala	Ala	Val	Ala
	50					55						60				
13	Thr	Ser	Gln	Gly	Gly	Gly	Lys	Leu	Pro	Glu	Ser	Phe	Met	Ser	Leu	Trp
	65				70						75				80	
15	Arg	Phe	Pro	His	Met	Leu	Glu	Asn	Ala	Arg	Gly	Met	Val	Ser	Gln	Leu
				85						90					95	
17	Thr	Gln	Phe	Gly	Ser	Thr	Leu	Gln	Asn	Ile	Ile	Glu	Arg	Gln	Asp	Ala
			100						105					110		
19	Glu	Ala	Leu	Asn	Ala	Leu	Leu	Gln	Asn	Gln	Ala	Ala	Glu	Leu	Ile	Leu
		115						120					125			
21	Thr	Asn	Leu	Ser	Ile	Gln	Asp	Lys	Thr	Ile	Glu	Glu	Leu	Asp	Ala	Glu
	130						135					140				
23	Lys	Thr	Val	Leu	Glu	Lys	Ser	Lys	Ala	Gly	Ala	Gln	Ser	Arg	Phe	Asp
	145					150					155				160	
25	Ser	Tyr	Gly	Lys	Leu	Tyr	Asp	Glu	Asn	Ile	Asn	Ala	Gly	Glu	Asn	Gln
				165						170					175	
27	Ala	Met	Thr	Leu	Arg	Ala	Ser	Ala	Ala	Gly	Leu	Thr	Thr	Ala	Val	Gln
			180						185					190		
29	Ala	Ser	Arg	Leu	Ala	Gly	Ala	Ala	Ala	Asp	Leu	Val	Pro	Asn	Ile	Phe
			195					200					205			
31	Gly	Phe	Ala	Gly	Gly	Gly	Ser	Arg	Trp	Gly	Ala	Ile	Ala	Glu	Ala	Thr
	210						215					220				
33	Gly	Tyr	Val	Met	Glu	Phe	Ser	Ala	Asn	Val	Met	Asn	Thr	Glu	Ala	Asp
	225				230						235				240	
35	Lys	Ile	Ser	Gln	Ser	Glu	Thr	Tyr	Arg	Arg	Arg	Arg	Gln	Glu	Trp	Glu
				245					250					255		
37	Ile	Gln	Arg	Asn	Asn	Ala	Glu	Ala	Glu	Leu	Lys	Gln	Ile	Asp	Ala	Gln
			260					265					270			
39	Leu	Lys	Ser	Leu	Ala	Val	Arg	Arg	Glu	Ala	Ala	Val	Leu	Gln	Lys	Thr
		275						280					285			
41	Ser	Leu	Lys	Thr	Gln	Gln	Glu	Gln	Thr	Gln	Ser	Gln	Leu	Ala	Phe	Leu
	290					295						300				
43	Gln	Arg	Lys	Phe	Ser	Asn	Gln	Ala	Leu	Tyr	Asn	Trp	Leu	Arg	Gly	Arg
	305					310					315				320	
45																
47																

RO 121280 B1

Leu Ala Ala Ile Tyr Phe Gln Phe Tyr Asp Leu Ala Val Ala Arg Cys 325 330 335	1
Leu Met Ala Glu Gln Ala Tyr Arg Trp Glu Leu Asn Asp Asp Ser Ala 340 345 350	3
Arg Phe Ile Lys Pro Gly Ala Trp Gln Gly Thr Tyr Ala Gly Leu Leu 355 360 365	5
Ala Gly Glu Thr Leu Met Leu Ser Leu Ala Gln Met Glu Asp Ala His 370 375 380	7
Leu Lys Arg Asp Lys Arg Ala Leu Glu Val Glu Arg Thr Val Ser Leu 385 390 395 400	9
Ala Glu Val Tyr Ala Gly Leu Pro Lys Asp Asn Gly Pro Phe Ser Leu 405 410 415	11
Ala Gln Glu Ile Asp Lys Leu Val Ser Gln Gly Ser Gly Ser Ala Gly 420 425 430	13
Ser Gly Asn Asn Asn Leu Ala Phe Gly Ala Gly Thr Asp Thr Lys Thr 435 440 445	15
Ser Leu Gln Ala Ser Val Ser Phe Ala Asp Leu Lys Ile Arg Glu Asp 450 455 460	17
Tyr Pro Ala Ser Leu Gly Lys Ile Arg Arg Ile Lys Gln Ile Ser Val 465 470 475 480	19
Thr Leu Pro Ala Leu Leu Gly Pro Tyr Gln Asp Val Gln Ala Ile Leu 485 490 495	21
Ser Tyr Gly Asp Lys Ala Gly Leu Ala Asn Gly Cys Glu Ala Leu Ala 500 505 510	23
Val Ser His Gly Met Asn Asp Ser Gly Gln Phe Gln Leu Asp Phe Asn 515 520 525	25
Asp Gly Lys Phe Leu Pro Phe Glu Gly Ile Ala Ile Asp Gln Gly Thr 530 535 540	27
Leu Thr Leu Ser Phe Pro Asn Ala Ser Met Pro Glu Lys Gly Lys Gln 545 550 555 560	29
Ala Thr Met Leu Lys Thr Leu Asn Asp Ile Ile Leu His Ile Arg Tyr 565 570 575	31
Thr Ile Lys ... 579	33
	35
(2) INFORMAȚIA PENTRU SECV. ID. NR. :52:	
(i) CARACTERISTICI DE SECVENȚĂ:	37
(A) LUNGIME: 5532 baze perechi	
(B) TIP: acid nucleic	39
(C) ÎMPLETIRE: dublă	
(D) TOPOLOGIE: lineară	41
(ii) TIPUL MOLECULEI: ADN (genomic)	
(xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 52	43
(regiune de codificare tcdA _{III}):	
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	45
GCC GCG CCG GGC TCG GTT GCA TCG ATG TTC TCA CCG GCG GCT TAT TTG 96	47

RO 121280 B1

1	Ala	Ala	Pro	Gly	Ser	Val	Ala	Ser	Met	Phe	Ser	Pro	Ala	Ala	Tyr	Leu	
				20					25					30			
3	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	
5			35					40					45				
7	TAT	TAC	CTA	GAT	AAA	CGT	CGC	CCG	GAT	TTA	GCA	AGC	TTA	ATG	CTC	AGC	192
	Tyr		Leu	Asp	Lys	Arg	Arg	Pro	Asp	Leu	Ala	Ser	Leu	Met	Leu	Ser	
		50					55					60					
9	CAG	AAA	AAT	ATG	GAT	CAG	GAA	ATT	TCA	ACC	CTG	GCT	CTC	TCT	AAT	CAA	240
	Gln	Lys	Asn	Met	Asp	Glu	Glu	Ile	Ser	Thr	Leu	Ala	Leu	Ser	Asn	Glu	
	65					70					75					80	
11	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		
13	GTG	ATC	GAT	ATC	TTG	TCA	ACT	TAT	CGT	TTA	ACT	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			
15	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				
17	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					
19	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	
21	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		
23	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			
25	TCC	CCA	AGT	TAT	CTG	GCC	CGG	TAT	TAT	CGC	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				
27	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					
29	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	
31	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		
33	GAG	CTG	TAT	CCA	CAG	GGT	GCC	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			
35	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				
37	TCC	GCT	GAT	TGG	ACT	GAG	ATT	CCC	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					
41																	
43																	
45																	
47																	

RO 121280 B1

ATA	AAT	CAA	AAG	TAT	GAA	TCA	CAG	GCG	ACA	ATC	AAA	CGT	AGT	GAC	TCT	360	1
Ile	Asn	Gln	Lys	Tyr	Glu	Ser	Gln	Ala	Thr	Ile	Lys	Arg	Ser	Asp	Ser	320	3
395					310					315							
GAC	AAT	ATA	CTC	AGT	ATA	GGG	TTA	CAA	AGA	TGG	CAT	AGC	GGT	AGT	TAT	1008	5
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	7
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	GCC	1104	9
Phe	Leu	Leu	Lys	Met	Asn	Lys	Ala	Ile	Arg	Leu	Leu	Lys	Ala	Thr	Gly		
		355					360					365					11
CTC	TCT	TTT	GCT	ACG	TTG	GAG	CGT	ATT	GTT	GAT	AGT	GTT	AAT	AGC	ACC	1152	13
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	15
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	17
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	19
Asn	Ile	Asn	Ile	Ser	Gln	Gln	Ala	Val	Gly	Asn	Gln	Leu	Ser	Gln	Phe		
			420					425					430				21
GAG	CAA	CTA	TTT	AAT	CAC	CCG	CCG	CTC	AAT	GGT	ATT	CGC	TAT	GAA	ATC	1344	23
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	25
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	27
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	29
Phe	Gln	Val	Asn	Ala	Ser	Glu	Leu	Tyr	Gln	Met	Leu	Leu	Ile	Thr	Asp		
				485					490					495			31
CGT	AAA	GAA	GAC	GGT	GTT	ATC	AAA	AAT	AAC	TTA	GAG	AAT	TTG	TCT	GAT	1536	33
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	35
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	37
Glu	Leu	Asn	Ile	Leu	Leu	Val	Ile	Cys	Gly	Tyr	Gly	Asp	Thr	Asn	Ile		
	530					535					540						39
TAT	CAG	ATT	ACC	GAC	GAT	AAT	TTA	GCC	AAA	ATA	GTG	GAA	ACA	TTG	TTG	1680	41
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	43
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	45
Phe	Leu	Met	Thr	Thr	Ala	Thr	Tyr	Ser	Thr	Thr	Leu	Thr	Pro	Glu	Ile		
			580					585					590				47

RO 121280 B1

3	AGC Ser	AAT Asn	CTG Leu	ACG Thr	GCT Ala	ACG Thr	TTG Leu	TCT Ser	TCA Ser	ACT Thr	TTG Leu	CAT His	GGC Gly	AAA Lys	GAG Glu	AGT Ser	1824
5			595					600					605				
7	CTG Leu	ATT Ile	GGG Gly	GAA Glu	GAT Asp	CTG Leu	AAA Lys	AGA Arg	GCA Ala	ATG Met	GCG Ala	CCT Pro	TCC Cys	TTC Phe	ACT Thr	TCG Ser	1872
9			610				615					620					
11	GCT Ala	TTG Leu	CAT His	TTG Leu	ACT Thr	TCT Ser	CAA Gln	GAA Glu	GTT Val	GCG Ala	TAT Tyr	GAC Asp	CTG Leu	CTG Leu	TTG Leu	TGG Trp	1920
13						630					635						
15	ATA Ile	GAC Asp	CAC Gln	ATT Ile	CAA Gln	CCG Pro	GCA Ala	CAA Gln	ATA Ile	ACT Thr	GTT Val	GAT Asp	GGC Gly	TTT Phe	TGG Trp	GAA Glu	1968
17					645					650					655		
19	GAA Glu	GTG Val	CAA Gln	ACA Thr	ACA Thr	CCA Pro	ACC Thr	AGC Ser	TTG Leu	AAG Lys	GTG Val	ATT Ile	ACC Thr	TTT Phe	GCT Ala	CAG Gln	2016
21					660				665					670			
23	GTG Val	CTG Leu	GCA Ala	CAA Gln	TTG Leu	AGC Ser	CTG Leu	ATC Ile	TAT Tyr	CGT Arg	CGT Arg	ATT Ile	GGC Gly	TTA Leu	AGT Ser	GAA Glu	2064
25			675					680					685				
27	ACG Thr	GAA Glu	CTG Leu	TCA Ser	CTG Leu	ATC Ile	GTG Val	ACT Thr	CAA Gln	TCT Ser	TCT Ser	CTG Leu	CTA Leu	GTG Val	GCA Ala	GGC Gly	2112
29			690				695					700					
31	AAA Lys	AGC Ser	ATA Ile	CTG Leu	GAT Asp	CAC His	GGT Gly	CTG Leu	TTA Leu	ACC Thr	CTG Leu	ATG Met	GCC Ala	TTG Leu	GAA Glu	GGT Gly	2160
33			705			710					715					720	
35	TTT Phe	CAT His	ACC Thr	TGG Trp	GTT Val	AAT Asn	GGC Gly	TTG Leu	GGG Gly	CAA Gln	CAT His	GCC Ala	TCC Ser	TTG Leu	ATA Ile	TTG Leu	2208
37					725				730					735			
39	GCG Ala	GCG Ala	TTG Leu	AAA Lys	GAC Asp	GGA Gly	GCC Ala	TTG Leu	ACA Thr	GTT Val	ACC Thr	GAT Asp	GTA Val	GCA Ala	CAA Gln	GCT Ala	2256
41				740				745					750				
43	ATG Met	AAT Asn	AAG Lys	GAG Glu	GAA Glu	TCT Ser	CTC Leu	CTA Leu	CAA Gln	ATG Met	GCA Ala	GCT Ala	AAT Asn	CAG Gln	CTG Val	GAG Glu	2304
45			755				760					765					
47	AAG Lys	GAT Asp	CTA Leu	ACA Thr	AAA Lys	CTG Leu	ACC Thr	AGT Ser	TGG Trp	ACA Thr	CAG Gln	ATT Ile	GAC Asp	GCT Ala	ATT Ile	CTG Leu	2352
49			770			775						780					
51	CAA Gln	TGG Trp	TTA Leu	CAG Gln	ATG Met	TCT Ser	TGG Ser	GCC Ala	TTG Leu	GCG Ala	GTT Val	TCT Ser	CCA Pro	CTG Leu	GAT Asp	CTG Leu	2400
53			785			790					795					800	
55</																	

RO 121280 B1

865	870	875	880	1
ACT TCA CGT ATT GCA GAA GCT ATC GCC GGT ATT CAA CTG TAC GTT AAC 2638	Thr Ser Arg Ile Ala Glu Ala Ile Ala Gly Ile Gln Leu Tyr Val Asn			3
	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			5
	900	905	910	7
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			9
	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			11
	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			13
	945	950	955	15
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			17
	965	970	975	
AAA ACT TAT TTG ACC AGC TTT GAG CAG CTA GCA AAT CTG AAA GTA ATT 2976	Lys Thr Tyr Leu Thr Ser Phe Glu Gln Val Ala Asn Leu Lys Val Ile			19
	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			21
	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 Trp Arg Ser Val			23
	1010	1015	1020	25
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			27
	1025	1030	1035	
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			29
	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			31
	1060	1065	1070	33
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			35
	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			37
	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			39
	1105	1110	1115	41
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			43
	1125	1130	1135	
ACT CTA TTA GTT ATG TTC TAT TCG ATG CAG AGT AGT TAT ACC TCC TAT 3456	Thr Leu Leu Val Met Phe Tyr Ser Met Gln Ser Ser Tyr Ser Ser Tyr			45
	1140	1145	1150	
ACC GAT AAT AAT GCG CCG GTC ACT GCG CTA TAT ATT TTC GCT GAT ATG 3504				47

RO 121280 B1

1

3 Thr Asp Asn Asn Ala Pro Val Thr Gly Leu Tyr Ile Phe Ala Asp Met
1155 1160 1165

5 TCA TCA GAC AAT ATG ACG AAT GCA CAA GCA ACT AAC TAT TCG AAT AAC 3552
Ser Ser Asp Asn Met Thr Asn Ala Gln Ala Thr Asn Tyr Trp Asn Asn
1170 1175 1180

7 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

9 AAA AAA GTC ATA ACC AGA AGA GTT AAT AAC CGT TAT CCG GAG GAT TAT 3648
Lys Lys Val Ile Thr Arg Arg Val Asn Arg Tyr Ala Glu Asp Tyr
1205 1210 1215

11 GAA ATT CCT TCC TCT GTG ACA AGT AAC AGT AAT TAT TCT TCG GGT GAT 3696
Glu Ile Pro Ser Ser Val Thr Ser Asn Ser Asn Tyr Ser Trp Gly Asp
1220 1225 1230

13 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

15 GAA TCG GCG GCA GAA GAT TTA AGG CTA TCT ACC AAT ATG CCA TTG AGT 3792
Glu Ser Ala Ala Glu Asp Leu Arg Leu Ser Thr Asn Met Ala Leu Ser
1250 1255 1260

17 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 1275 1280

19 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

21 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

23 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

25 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

27 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

29 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

31 GTT ACT GGT GGG TAT TCG 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

33 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

35 GGT GGT GAC GAT CAA ATC TTT ACT GCT CAT 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

37 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

RO 121280 B1

CTC ACA ATA GAT TGT AAG AAT TTA AAT TTC ATC GAC AAT CAG GCA CAT 4303 Leu Thr Ile Asp Cys Lys Asn Leu Asn Phe Ile Asp Asn Gln Ala His 1445 1450 1455	1
ATT GAG ATT GAT TTC ACC GCT ACG GCA CAA GAT GGC CGA TTC TTG CGT 4416 Ile Glu Ile Asp Phe Thr Ala Thr Ala Gln Asp Gly Arg Phe Leu Gly 1460 1465 1470	3
GCA GAA ACT TTT ATT ATC CCG GTA ACT AAA AAA GTT CTC GGT ACT GAG 4464 Ala Glu Thr Phe Ile Ile Pro Val Thr Lys Lys Val Leu Gly Thr Glu 1475 1480 1485	5
AAC GTC ATT CCG TTA TAT AGC GAA AAT AAC GGT GTT CAA TAT ATC CAA 4512 Asn Val Ile Ala Leu Tyr Ser Glu Asn Asn Gly Val Gln Tyr Met Gln 1490 1495 1500	7
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	9
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	11
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	13
CTG TTC GAT AAA TAT GAT GAG TCT ATT CAT GCC 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	15
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	17
TAT CAA GGA GAA CTT AOC 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	19
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	21
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	23
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	25
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	27
CCG ATC GAT TTC TCT GGC GCC AAT GCT CTC TAT TTC TCG GAA CTC TTC 5040 Pro Met Asp Phe Ser Gly Ala Asn Ala Leu Tyr Phe Trp Glu Leu Phe 1665 1670 1675 1680	29
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	31
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	33
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	35
	37
	39
	41
	43
	45
	47

RO 121280 B1

1 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

3 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

5 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

7 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

9 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

11 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

13 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

15 AAA ACC CCG TTG 5532
 Lys Thr Pro Leu
 1844

(2) INFORMAȚIA PENTRU SECV. ID. NR. 53:

(i) CARACTERISTICI DE SECVENȚĂ:

(A) LUNGIME: 1844 aminoacizi

(B) TIP: aminoacid

(C) ÎMPLETIRE: simplă

(D) TOPOLOGIE: lineară

(ii) TIPUL MOLECULEI: proteină

(xi) DESCRIEREA SECVENȚEI: SECV. NR. 53: (TcdA_{ii})

Țesătura	De la	Până la	Descrierea
Peptidă	1	1844	Peptidă TcdA _{ii}
Fragment	1	11	(SECV. ID NR. 1)
Fragment	978	990	(SECV. ID NR. 23)
Fragment	1387	1401	(SECV. ID NR. 22)
Fragment	1448	1505	(SECV. ID NR. 24)
Fragment	1527	1552	(SECV. ID NR. 21)

39 Phe Ile Gln Gly Tyr Ser Asp Leu Phe Gly Asn Arg Ala Asp Asn Tyr
 1 5 10 15

41 Ala Ala Pro Gly Ser Val Ala Ser Met Phe Ser Pro Ala Ala Tyr Leu
 20 25 30

43 Thr Glu Leu Tyr Arg Glu Ala Lys Asn Leu His Asp Ser Ser Ser Ile
 35 40 45

45 Tyr Tyr Leu Asp Lys Arg Arg Pro Asp Leu Ala Ser Leu Met Leu Ser
 50 55 60

47 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

RO 121280 B1

85	90	95	1
Val Met Asp Met Leu Ser Thr Tyr Arg Leu Ser Gly Glu Thr Pro Tyr 100 105 110			3
His His Ala Tyr Glu Thr Val Arg Glu Ile Val His Glu Arg Asp Pro 115 120 125			5
Gly Phe Arg His Leu Ser Gln Ala Pro Ile Val Ala Ala Lys Leu Asp 130 135 140			7
Pro Val Thr Leu Leu Gly Ile Ser Ser His Ile Ser Pro Glu Leu Tyr 145 150 155 160			9
Asn Leu Leu Ile Glu Glu Ile Pro Glu Lys Asp Glu Ala Ala Leu Asp 165 170 175			11
Thr Leu Tyr Lys Thr Asn Phe Gly Asp Ile Thr Thr Ala Gln Leu Met 180 185 190			13
Ser Pro Ser Tyr Leu Ala Arg Tyr Tyr Gly Val Ser Pro Glu Asp Ile 195 200 205			15
Ala Tyr Val Thr Thr Ser Leu Ser His Val Gly Tyr Ser Ser Asp Ile 210 215 220			17
Leu Val Ile Pro Leu Val Asp Gly Val Gly Lys Met Glu Val Val Arg 225 230 235 240			19
Val Thr Arg Thr Pro Ser Asp Asn Tyr Thr Ser Gln Thr Asn Tyr Ile 245 250 255			21
Glu Leu Tyr Pro Gln Gly Gly Asp Asn Tyr Leu Ile Lys Tyr Asn Leu 260 265 270			23
Ser Asn Ser Phe Gly Leu Asp Asp Phe Tyr Leu Gln Tyr Lys Asp Gly 275 280 285			25
Ser Ala Asp Trp Thr Glu Ile Ala His Asn Pro Tyr Pro Asp Met Val 290 295 300			27
Ile Asn Gln Lys Tyr Glu Ser Gln Ala Thr Ile Lys Arg Ser Asp Ser 305 310 315 320			29
Asp Asn Ile Leu Ser Ile Gly Leu Gln Arg Trp His Ser Gly Ser Tyr 325 330 335			31
Asn Phe Ala Ala Ala Asn Phe Lys Ile Asp Gln Tyr Ser Pro Lys Ala 340 345 350			33
Phe Leu Leu Lys Met Asn Lys Ala Ile Arg Leu Leu Lys Ala Thr Gly 355 360 365			35
Leu Ser Phe Ala Thr Leu Glu Arg Ile Val Asp Ser Val Asn Ser Thr 370 375 380			37
Lys Ser Ile Thr Val Glu Val Leu Asn Lys Val Tyr Arg Val Lys Phe 385 390 395 400			39
Tyr Ile Asp Arg Tyr Gly Ile Ser Glu Glu Thr Ala Ala Ile Leu Ala 405 410 415			41
Asn Ile Asn Ile Ser Gln Gln Ala Val Gly Asn Gln Leu Ser Gln Phe 420 425 430			43
Glu Gln Leu Phe Asn His Pro Pro Leu Asn Gly Ile Arg Tyr Glu Ile 435 440 445			45
Ser Glu Asp Asn Ser Lys His Leu Pro Asn Pro Asp Leu Asn Leu Lys 450 455 460			47

RO 121280 B1

1 Pro Asp Ser Thr Gly Asp Asp Gln Arg Lys Ala Val Leu Lys Arg Ala
 3 465 470 475 480
 5 Phe Gln Val Asn Ala Ser Glu Leu Tyr Gln Met Leu Leu Ile Thr Asp
 485 490 495
 7 Arg Lys Glu Asp Gly Val Ile Lys Asn Asn Leu Glu Asn Leu Ser Asp
 500 505 510
 9 Leu Tyr Leu Val Ser Leu Leu Ala Gln Ile His Asn Leu Thr Ile Ala
 515 520 525
 11 Glu Leu Asn Ile Leu Leu Val Ile Cys Gly Tyr Gly Asp Thr Asn Ile
 530 535 540
 13 Tyr Gln Ile Thr Asp Asp Asn Leu Ala Lys Ile Val Glu Thr Leu Leu
 545 550 555 560
 15 Trp Ile Thr Gln Trp Leu Lys Thr Gln Lys Trp Thr Val Thr Asp Leu
 565 570 575
 17 Phe Leu Met Thr Thr Ala Thr Tyr Ser Thr Thr Leu Thr Pro Glu Ile
 580 585 590
 19 Ser Asn Leu Thr Ala Thr Leu Ser Ser Thr Leu His Gly Lys Glu Ser
 595 600 605
 21 Leu Ile Gly Glu Asp Leu Lys Arg Ala Met Ala Pro Cys Phe Thr Ser
 610 615 620
 23 Ala Leu His Leu Thr Ser Gln Glu Val Ala Tyr Asp Leu Leu Leu Trp
 625 630 635 640
 25 Ile Asp Gln Ile Gln Pro Ala Gln Ile Thr Val Asp Gly Phe Trp Glu
 645 650 655
 27 Glu Val Gln Thr Thr Pro Thr Ser Leu Lys Val Ile Thr Phe Ala Gln
 660 665 670
 29 Val Leu Ala Gln Leu Ser Leu Ile Tyr Arg Arg Ile Gly Leu Ser Glu
 675 680 685
 31 Thr Glu Leu Ser Leu Ile Val Thr Gln Ser Ser Leu Leu Val Ala Gly
 690 695 700
 33 Lys Ser Ile Leu Asp His Gly Leu Leu Thr Leu Met Ala Leu Glu Gly
 705 710 715 720
 35 Phe His Thr Trp Val Asn Gly Leu Gly Gln His Ala Ser Leu Ile Leu
 725 730 735
 37 Ala Ala Leu Lys Asp Gly Ala Leu Thr Val Thr Asp Val Ala Gln Ala
 740 745 750
 39 Met Asn Lys Glu Glu Ser Leu Leu Gln Met Ala Ala Asn Gln Val Glu
 755 760 765
 41 Lys Asp Leu Thr Lys Leu Thr Ser Trp Thr Gln Ile Asp Ala Ile Leu
 770 775 780
 43 Gln Trp Leu Gln Met Ser Ser Ala Leu Ala Val Ser Pro Leu Asp Leu
 785 790 795 800
 45 Ala Gly Met Met Ala Leu Lys Tyr Gly Ile Asp His Asn Tyr Ala Ala
 805 810 815
 47 Trp Gln 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

RO 121280 B1

Ile	Asn	Ala	Val	Val	Asp	Ser	Ala	Ala	Gly	Val	Arg	Asp	Arg	Asn	Gly	1
850						855					860					3
Leu	Tyr	Thr	Tyr	Leu	Leu	Ile	Asp	Asn	Gln	Val	Ser	Ala	Asp	Val	Ile	5
865					870					875					880	7
Thr	Ser	Arg	Ile	Ala	Glu	Ala	Ile	Ala	Gly	Ile	Gln	Leu	Tyr	Val	Asn	9
				885					890					895		11
Arg	Ala	Leu	Asn	Arg	Asp	Glu	Gly	Gln	Leu	Ala	Ser	Asp	Val	Ser	Thr	13
			900					905					910			15
Arg	Gln	Phe	Phe	Thr	Asp	Trp	Glu	Arg	Tyr	Asn	Lys	Arg	Tyr	Ser	Thr	17
		915					920					925				19
Trp	Ala	Gly	Val	Ser	Glu	Leu	Val	Tyr	Tyr	Pro	Glu	Asn	Tyr	Val	Asp	21
	930					935					940					23
Pro	Thr	Gln	Arg	Ile	Gly	Gln	Thr	Lys	Met	Met	Asp	Ala	Leu	Leu	Gln	25
945					950					955					960	27
Ser	Ile	Asn	Gln	Ser	Gln	Leu	Asn	Ala	Asp	Thr	Val	Glu	Asp	Ala	Phe	29
				965					970					975		31
Lys	Thr	Tyr	Leu	Thr	Ser	Phe	Glu	Gln	Val	Ala	Asn	Leu	Lys	Val	Ile	33
			980					985					990			35
Ser	Ala	Tyr	His	Asp	Asn	Val	Asn	Val	Asp	Gln	Gly	Leu	Thr	Tyr	Phe	37
		995					1000					1005				39
Ile	Gly	Ile	Asp	Gln	Ala	Ala	Pro	Gly	Thr	Tyr	Tyr	Trp	Arg	Ser	Val	41
	1010					1015						1020				43
Asp	His	Ser	Lys	Cys	Glu	Asn	Gly	Lys	Phe	Ala	Ala	Asn	Ala	Trp	Gly	45
1025					1030				1035						1040	47
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	

RO 121280 B1

	1220	1225	1230
1			
3	His Ser Leu Thr Met Leu Tyr Gly Gly Ser Val Pro Asn Ile Thr Phe 1235 1240 1245		
5	Glu Ser Ala Ala Glu Asp Leu Arg Leu Ser Thr Asn Met Ala Leu Ser 1250 1255 1260		
7	Ile Ile His Asn Gly Tyr Ala Gly Thr Arg Arg Ile Gln Cys Asn Leu 1265 1270 1275 1280		
9	Met Lys Gln Tyr Ala Ser Leu Gly Asp Lys Phe Ile Ile Tyr Asp Ser 1285 1290 1295		
11	Ser Phe Asp Asp Ala Asn Arg Phe Asn Leu Val Pro Leu Phe Lys Phe 1300 1305 1310		
13	Gly Lys Asp Glu Asn Ser Asp Asp Ser Ile Cys Ile Tyr Asn Glu Asn 1315 1320 1325		
15	Pro Ser Ser Glu Asp Lys Lys Trp Tyr Phe Ser Ser Lys Asp Asp Asn 1330 1335 1340		
17	Lys Thr Ala Asp Tyr Asn Gly Gly Thr Gln Cys Ile Asp Ala Gly Thr 1345 1350 1355 1360		
19	Ser Asn Lys Asp Phe Tyr Tyr Asn Leu Gln Glu Ile Glu Val Ile Ser 1365 1370 1375		
21	Val Thr Gly Gly Tyr Trp Ser Ser Tyr Lys Ile Ser Asn Pro Ile Asn 1380 1385 1390		
23	Ile Asn Thr Gly Ile Asp Ser Ala Lys Val Lys Val Thr Val Lys Ala 1395 1400 1405		
25	Gly Gly Asp Asp Gln Ile Phe Thr Ala Asp Asn Ser Thr Tyr Val Pro 1410 1415 1420		
27	Gln Gln Pro Ala Pro Ser Phe Glu Glu Met Ile Tyr Gln Phe Asn Asn 1425 1430 1435 1440		
29	Leu Thr Ile Asp Cys Lys Asn Leu Asn Phe Ile Asp Asn Gln Ala His 1445 1450 1455		
31	Ile Glu Ile Asp Phe Thr Ala Thr Ala Gln Asp Gly Arg Phe Leu Gly 1460 1465 1470		
33	Ala Glu Thr Phe Ile Ile Pro Val Thr Lys Lys Val Leu Gly Thr Glu 1475 1480 1485		
35	Asn Val Ile Ala Leu Tyr Ser Glu Asn Asn Gly Val Gln Tyr Met Gln 1490 1495 1500		
37	Ile Gly Ala Tyr Arg Thr Arg Leu Asn Thr Leu Phe Ala Gln Gln Leu 1505 1510 1515 1520		
39	Val Ser Arg Ala Asn Arg Gly Ile Asp Ala Val Leu Ser Met Glu Thr 1525 1530 1535		
41	Gln Asn Ile Gln Glu Pro Gln Leu Gly Ala Gly Thr Tyr Val Gln Leu 1540 1545 1550		
43	Val Leu Asp Lys Tyr Asp Glu Ser Ile His Gly Thr Asn Lys Ser Phe 1555 1560 1565		
45	Ala Ile Glu Tyr Val Asp Ile Phe Lys Glu Asn Asp Ser Phe Val Ile 1570 1575 1580		
47	Tyr Gln Gly Glu Leu Ser Glu Thr Ser Gln Thr Val Val Lys Val Phe 1585 1590 1595 1600		

RO 121280 B1

Leu Ser Tyr Phe Ile Glu Ala Thr Gly Asn Lys Asn His Leu Trp Val 1
 1605 1610 1615
 Arg Ala Lys Tyr Gln Lys Glu Thr Thr Asp Lys Ile Leu Phe Asp Arg 3
 1620 1625 1630
 Thr Asp Glu Lys Asp Pro His Gly Trp Phe Leu Ser Asp Asp His Lys 5
 1635 1640 1645
 Thr Phe Ser Gly Leu Ser Ser Ala Gln Ala Leu Lys Asn Asp Ser Glu 7
 1650 1655 1660
 Pro Met Asp Phe Ser Gly Ala Asn Ala Leu Tyr Phe Trp Glu Leu Phe 9
 1665 1670 1675 1680
 Tyr Tyr Thr Pro Met Met Met Ala His Arg Leu Leu Gln Glu Gln Asn 11
 1685 1690 1695
 Phe Asp Ala Ala Asn His Trp Phe Arg Tyr Val Trp Ser Pro Ser Gly 13
 1700 1705 1710
 Tyr Ile Val Asp Gly Lys Ile Ala Ile Tyr His Trp Asn Val Arg Pro 15
 1715 1720 1725
 Leu Glu Glu Asp Thr Ser Trp Asn Ala Gln Gln Leu Asp Ser Thr Asp 17
 1730 1735 1740
 Pro Asp Ala Val Ala Gln Asp Asp Pro Met His Tyr Lys Val Ala Thr 19
 1745 1750 1755 1760
 Phe Met Ala Thr Leu Asp Leu Leu Met Ala Arg Gly Asp Ala Ala Tyr 21
 1765 1770 1775
 Arg Gln Leu Glu Arg Asp Thr Leu Ala Glu Ala Lys Met Trp Tyr Thr 23
 1780 1785 1790
 Gln Ala Leu Asn Leu Leu Gly Asp Glu Pro Gln Val Met Leu Ser Thr 25
 1795 1800 1805
 Thr Trp Ala Asn Pro Thr Leu Gly Asn Ala Ala Ser Lys Thr Thr Gln 27
 1810 1815 1820
 Gln Val Arg Gln Gln Val Leu Thr Gln Leu Arg Leu Asn Ser Arg Val 29
 1825 1830 1835 1840
 Lys Thr Pro Leu 31
 1844

(2) INFORMAȚIA PENTRU SECV. ID. NR. 54:

(i) CARACTERISTICI DE SECVENȚĂ:

(A) LUNGIME: 1722 baze perechi

(B) TIP: acid nucleic

(C) ÎMPLETIRE: dublă

(D) TOPOLOGIE: lineară

(ii) TIPUL MOLECULEI: ADN (genomic)

(xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 54

(regiune de codificare tcbA_{III}):

CTA GGA ACA GCC AAT TCC CTG ACC GCT TTA TTC CTG CCG CAG GAA AAT 43
 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 CCG 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

RO 121280 B1

1
 3 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

5 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

7 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

9 CGC TTC CCT CAA ATG CTA GAA GGG GCA CCG 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

11 ATA CAG TTC GGT ACT TCA CTA TTG GGC TAC AGT GAG CGT CAG GAT CCG 336
 Ile Gln Phe Gly Ser Ser Leu Leu Gly Tyr Ser Glu Arg Gln Asp Ala
 100 105 110

13 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

15 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

17 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

19 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

21 GCG CTG GCG TTA CCG 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

23 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

25 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

27 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

29 AAA GTT GCT CAG TCG GAA ATA TAT CCG CGT CCG CGT CAA GAA TCG AAA 768
 Lys Val Ala Gln Ser Glu Ile Tyr Arg Arg Arg Arg Gln Glu Trp Lys
 245 250 255

31 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

33 CTG GAA TCA CTG TCT ATT CCG 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

35 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

37 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

RO 121280 B1

TTG TCA GGT ATT TAT TTC CAG TTC TAT GAC TTG GCC GTA TCA CGT TGC 1008	1
Leu Ser Gly Ile Tyr Phe Gln Phe Tyr Asp Leu Ala Val Ser Arg Cys	
325 330 335	3
CTG ATG GCA GAG CAA TCC TAT CAA TGG GAA GCT AAT GAT AAT TCC ATT 1056	5
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	7
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	9
Cys Gly Glu Ala Leu Ile Gln Asn Leu Ala Gln Met Glu Glu Ala Tyr	
370 375 380	11
CTG AAA TGG GAA TCT CGC GCT TTG GAA GTA GAA CGC ACC GTT TCA TTG 1200	13
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	15
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	17
Glu Gln Ile Pro Ala Leu Leu Asp Lys Gly Glu Gly Thr Ala Gly Thr	
420 425 430	19
AAA GAA AAT GGG TTA TCA TTG GCT AAT GCT ATC CTG TCA GCT TCG GTC 1344	21
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 ACC GAT TAT CCA GAC AGT ATC GTT 1392	23
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	25
Gly Ser Asn Lys Val Arg Arg Ile Lys Gln Ile Ser Val Ser Leu Pro	
465 470 475 480	27
GCA TTG GTT GGG CCT TAT CAG GAT GTT CAG GCT ATG CTC AGC TAT GGT 1488	29
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	31
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	33
Gly Thr Asn Asp Ser Gly Gln Phe Gln Leu Asp Phe Asn Asp Gly Lys	
515 520 525	35
TAC CTG CCA TTT GAA GGT ATT GCT CTT GAT GAT CAG GGT ACA CTG AAT 1632	37
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	39
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	41
Met Ser Asp Ile Ile Leu His Ile Arg Tyr Thr Ile Arg ...	
565 570 573	43
(2) INFORMAȚIA PENTRU SECV. ID. NR. 55:	45
(i) CARACTERISTICI DE SECVENȚĂ:	
(A) LUNGIME: 573 baze perechi	47
(B) TIP: acid nucleic	

RO 121280 B1

1 (C) ÎMPLETIRE: simplă
 (D) TOPOLOGIE: lineară
 3 (ii) TIPUL MOLECULEI: proteină
 (xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 55 (TcbA_{III})
 5 Leu Gly Thr Ala Asn Ser Leu Thr Ala Leu Phe Leu Pro Gln Glu Asn
 1 5 10 15
 7 Ser Lys Leu Lys Gly Tyr Trp Arg Thr Leu Ala Gln Arg Met Phe Asn
 20 25 30
 9 Leu Arg His Asn Leu Ser Ile Asp Gly Gln Pro Leu Ser Leu Pro Leu
 35 40 45
 11 Tyr Ala Lys Pro Ala Asp Pro Lys Ala Leu Leu Ser Ala Ala Val Ser
 50 55 60
 13 Ala Ser Gln Gly Gly Ala Asp Leu Pro Lys Ala Pro Leu Thr Ile His
 65 70 75 80
 15 Arg Phe Pro Gln Met Leu Glu Gly Ala Arg Gly Leu Val Asn Gln Leu
 85 90 95
 17 Ile Gln Phe Gly Ser Ser Leu Leu Gly Tyr Ser Glu Arg Gln Asp Ala
 100 105 110
 19 Glu Ala Met Ser Gln Leu Leu Gln Thr Gln Ala Ser Glu Leu Ile Leu
 115 120 125
 21 Thr Ser Ile Arg Met Gln Asp Asn Gln Leu Ala Glu Leu Asp Ser Glu
 130 135 140
 23 Lys Thr Ala Leu Gln Val Ser Leu Ala Gly Val Gln Gln Arg Phe Asp
 145 150 155 160
 25 Ser Tyr Ser Gln Leu Tyr Glu Glu Asn Ile Asn Ala Gly Glu Gln Arg
 165 170 175
 27 Ala Leu Ala Leu Arg Ser Glu Ser Ala Ile Glu Ser Gln Gly Ala Gln
 180 185 190
 29 Ile Ser Arg Met Ala Gly Ala Gly Val Asp Met Ala Pro Asn Ile Phe
 195 200 205
 31 Gly Leu Ala Asp Gly Gly Met His Tyr Gly Ala Ile Ala Tyr Ala Ile
 210 215 220
 33 Ala Asp Gly Ile Glu Leu Ser Ala Ser Ala Lys Met Val Asp Ala Glu
 225 230 235 240
 35 Lys Val Ala Gln Ser Glu Ile Tyr Arg Arg Arg Arg Gln Glu Trp Lys
 245 250 255
 37 Ile Gln Arg Asp Asn Ala Gln Ala Glu Ile Asn Gln Leu Asn Ala Gln
 260 265 270
 39 Leu Glu Ser Leu Ser Ile Arg Arg Glu Ala Ala Glu Met Gln Lys Glu
 275 280 285
 41 Tyr Leu Lys Thr Gln Gln Ala Gln Ala Gln Ala Gln Leu Thr Phe Leu
 290 295 300
 43 Arg Ser Lys Phe Ser Asn Gln Ala Leu Tyr Ser Trp Leu Arg Gly Arg
 305 310 315 320
 45 Leu Ser Gly Ile Tyr Phe Gln Phe Tyr Asp Leu Ala Val Ser Arg Cys
 325 330 335

RO 121280 B1

Leu Met Ala Glu Gln Ser Tyr Gln Trp Glu Ala Asn Asp Asn Ser Ile	1
340 345 350	
Ser Phe Val Lys Pro Gly Ala Trp Gln Gly Thr Tyr Ala Gly Leu Leu	3
355 360 365	
Cys Gly Glu Ala Leu Ile Gln Asn Leu Ala Gln Met Glu Glu Ala Tyr	5
370 375 380	
Leu Lys Trp Glu Ser Arg Ala Leu Glu Val Glu Arg Thr Val Ser Leu	7
385 390 395 400	
Ala Val Val Tyr Asp Ser Leu Glu Gly Asn Asp Arg Phe Asn Leu Ala	9
405 410 415	
Glu Gln Ile Pro Ala Leu Leu Asp Lys Gly Glu Gly Thr Ala Gly Thr	11
420 425 430	
Lys Glu Asn Gly Leu Ser Leu Ala Asn Ala Ile Leu Ser Ala Ser Val	13
435 440 445	
Lys Leu Ser Asp Leu Lys Leu Gly Thr Asp Tyr Pro Asp Ser Ile Val	15
450 455 460	
Gly Ser Asn Lys Val Arg Arg Ile Lys Gln Ile Ser Val Ser Leu Pro	17
465 470 475 480	
Ala Leu Val Gly Pro Tyr Gln Asp Val Gln Ala Met Leu Ser Tyr Gly	19
485 490 495	
Gly Ser Thr Gln Leu Pro Lys Gly Cys Ser Ala Leu Ala Val Ser His	21
500 505 510	
Gly Thr Asn Asp Ser Gly Gln Phe Gln Leu Asp Phe Asn Asp Gly Lys	23
515 520 525	
Tyr Leu Pro Phe Glu Gly Ile Ala Leu Asp Asp Gln Gly Thr Leu Asn	25
530 535 540	
Leu Gln Phe Pro Asn Ala Thr Asp Lys Gln Lys Ala Ile Leu Gln Thr	27
545 550 555 560	
Met Ser Asp Ile Ile Leu His Ile Arg Tyr Thr Ile Arg ...	29
565 570 573	
(2) INFORMAȚIA PENTRU SECV. ID. NR. 56:	31
(i) CARACTERISTICI DE SECVENȚĂ:	33
(A) LUNGIME: 2898 baze perechi	35
(B) TIP: acid nucleic	37
(C) ÎMPLETIRE: dublă	39
(D) TOPOLOGIE: lineară	41
(ii) TIPUL MOLECULEI: ADN (genomic)	43
(xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 56 (tccA):	45
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	

RO 121280 B1

1	33	Val Val Arg Met Pro Arg Glu Arg Phe Ile Arg Glu His Arg Ala Asp	43
3	145	CTC GGG CGC AGT GCT GAA AAA ATG TAT CAC CTG GCA GTG GGC TAT GCT	152
	49	Leu Gly Arg Ser Ala Glu Lys Met Tyr Asp Leu Ala Val Gly Tyr Ala	64
5	193	CAT CAG GTG TTA CAC CAT TTT CCG CGT AAT TCT CTT AGT CAA GCT GTT	240
7	65	His Gln Val Leu His His Phe Arg Arg Asn Ser Leu Ser Glu Ala Val	80
9	241	CAG TTT GCC TTG AGA AGT CCG TTC TCC GTA TCA GCG CCG GAT TAC GCC	288
	81	Gln Phe Gly Leu Arg Ser Pro Phe Ser Val Ser Gly Pro Asp Tyr Ala	96
11	289	AAT CAG TTT CTT GAT GCA AAC ACG CGT TCG 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
13	337	GGA TCA CCG GAA GCC AAT GAT CCG CCG GTA GCC TAT CTG ACT CAT ATT	384
15	113	Gly Ser Pro Glu Ala Asn Asp Ala Pro Val Ala Tyr Leu Thr His Ile	128
17	385	TAT CAA TTG CCC CTT GAA CAG GAA AAG AAT GCG GCC ACT ACC ATT ATG	432
	129	Tyr Gln Leu Ala Leu Glu Gln Glu Lys Asn Gly Ala Thr Thr Ile Met	144
19	433	AAT ACG CTG CCG GAG CGT CCG 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
21	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
23	529	ATT CTG TCC AAA GCT ATT CAG AAG AAA CTG AGT TTG ACT GAT CTG GAA	576
25	177	Ile Leu Ser Lys Ala Ile Gln Lys Lys Leu Ser Leu Thr Asp Leu Glu	192
27	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
29	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
31	673	ACT ACG TTG CAA GAT ATC ACT TTG CCA CAG ACG CTG GAT CTG CCG CAA	720
33	225	Thr Thr Leu Gln Asp Ile Thr Leu Pro Gln Thr Leu Asp Leu Pro Gln	240
35	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
37	769	GCT TTG ACC CGA CTG CAA ATC ATG GCG AGT CAG TTT TCG CCA CAG CAG	816
	257	Ala Leu Thr Arg Leu Gln Ile Met Ala Ser Gln Phe Ser Pro Glu Gln	272
39	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
41	865	TAT GGT GAC AGT TCG CTT ACT GTG AAT AGT TTC AGC GAC ATG ACC ATA	912
43	289	Tyr Gly Asp Ser Ser Leu Thr Val Asn Ser Phe Ser Asp Met Thr Ile	304

RO 121280 B1

313	ATG	ACT	GAT	CGA	ACA	AGT	TTG	ACT	GTA	CCC	CAG	GTA	GAA	CTG	ATG	TTG	350	1
305	Met	Thr	Asp	Arg	Thr	Ser	Leu	Thr	Val	Pro	Gln	Val	Glu	Leu	Met	Leu	320	
951	TGT	TCA	ACT	GTC	GGA	GGT	TCT	ACG	GTT	GTT	AAG	TCT	GAT	AAT	GTG	ACT	1903	3
321	Cys	Ser	Thr	Val	Gly	Gly	Ser	Thr	Val	Val	Lys	Ser	Asp	Asn	Val	Ser	335	5
1009	TCT	GGT	GAC	ACG	ACA	GGG	ACG	CCA	TTT	GGG	TAT	GGC	GCC	CGC	TTT	ATT	1056	7
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	GGG	ATT	ACC	CTG	AGT	CGC	AGT	GGT	GGG	GAG	1104	9
353	His	Ala	Gly	Lys	Pro	Glu	Ala	Ile	Thr	Leu	Ser	Arg	Ser	Gly	Ala	Glu	368	
1105	GGG	CAT	TTT	GCT	CTG	ACG	GTT	AAC	AAT	CTG	ACA	GAT	GAC	AAG	TTG	GAC	1152	11
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	13
385	Arg	Ile	Asn	Arg	Thr	Val	Arg	Leu	Gln	Lys	Trp	Leu	Asn	Leu	Pro	Tyr	400	15
1201	GAG	GAT	ATT	CAC	CTG	TTA	GTG	ACT	TCT	GCT	ATG	GAT	CGC	GAA	ACA	GGA	1248	17
401	Glu	Asp	Ile	Asp	Leu	Leu	Val	Thr	Ser	Ala	Met	Asp	Ala	Glu	Thr	Gly	416	
1249	AAT	ACC	GGG	CTG	TCG	ATG	AAC	GAC	AAT	ACG	CTG	CGT	ATG	TTG	GGA	GTG	1296	19
417	Asn	Thr	Ala	Leu	Ser	Met	Asn	Asp	Asn	Thr	Leu	Arg	Met	Leu	Gly	Val	432	
1297	TTC	AAA	CAT	TAT	CAG	GGG	AAG	TAT	GGT	GTT	AGC	GCT	AAA	CAA	TTT	GCT	1344	21
433	Phe	Lys	His	Tyr	Gln	Ala	Lys	Tyr	Gly	Val	Ser	Ala	Lys	Gln	Phe	Ala	448	23
1345	GGC	TGG	CTG	CGC	GTA	GTG	GCC	CCG	TTT	GCC	ATT	ACA	CCG	GCA	ACG	CCG	1392	25
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	27
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	29
481	Val	Ile	Asp	Asn	Gln	Asp	Phe	Val	Tyr	Thr	Leu	Thr	Thr	Gly	Gly	Asp	496	
1489	GGG	CCG	CGT	GTT	AAG	CAT	ATC	AGC	ACG	GCA	CTG	GGC	CTC	AAT	CAT	CGT	1536	31
497	Gly	Ala	Arg	Val	Lys	His	Ile	Ser	Thr	Ala	Leu	Gly	Leu	Asn	His	Arg	512	
1537	CAG	TTC	CTG	TTA	TTG	GGG	GAT	AAT	ATT	GCC	CGT	CAA	CAG	GGG	AAT	GTG	1584	33
513	Gln	Phe	Leu	Leu	Leu	Ala	Asp	Asn	Ile	Ala	Arg	Gln	Gln	Gly	Asn	Val	528	35
1585	ACG	CAA	AGC	ACA	CTC	AAC	TGT	AAT	CTG	TTT	GTG	GTG	TCA	GCT	TTT	TAC	1632	37
529	Thr	Gln	Ser	Thr	Leu	Asn	Cys	Asn	Leu	Phe	Val	Val	Ser	Ala	Phe	Tyr	544	
1633	CGT	CTG	GCT	AAT	TTG	GGG	CGC	ACA	TTG	GGG	ATA	AAT	CCA	GAG	TCT	TTG	1680	39
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	41

RO 121280 B1

1	561	Cys Ala Leu Val Asp Arg Leu Asp Ala Gly Thr Gly Ile Val Trp Gln	574
3	1729	CAA TTG GCA GGG AAA CCC ACA ATC ACG GTA CCA CAA AAA GAT TCC CCG	1776
5	577	Gln Leu Ala Gly Lys Pro Thr Ile Thr Val Pro Gln Lys Asp Ser Pro	592
7	1777	CTG GCG GCG GAT ATT CTG AGT TTG CTG CAA GCG CTA AGT GCG ATT GCT	1824
9	593	Leu Ala Ala Asp Ile Leu Ser Leu Leu Gln Ala Leu Ser Ala Ile Ala	608
11	1825	CAA TGG CAA CAA CAG CAC GAT TTA GAA TTT TCA GCA CTG CTT TTG CTG	1872
13	609	Gln Trp Gln Gln Gln His Asp Leu Glu Phe Ser Ala Leu Leu Leu Leu	624
15	1873	TTG AGT GAC AAC CCT ATT TCT ACC TCG CAG GGC ACT GAC GAT CAA TTG	1920
17	625	Leu Ser Asp Asn Pro Ile Ser Thr Ser Gln Gly Thr Asp Asp Gln Leu	640
19	1921	AAC TTT ATC CGT CAA CTG TCG CAG AAC CTA GCC ACT ACG TTT CTC GCT	1968
21	641	Asn Phe Ile Arg Gln Val Trp Gln Asn Leu Gly Ser Thr Phe Val Gly	656
23	1969	GCA ACA TTG TTG TCC CGC AGT GGG GCA CCA TTA GTC GAT ACC AAC GGC	2016
25	657	Ala Thr Leu Leu Ser Arg Ser Gly Ala Pro Leu Val Asp Thr Asn Gly	672
27	2017	CAC GCT ATT GAC TGG TTT GCT CTG CTC TCA GCA GGT AAT AGT CCG CTT	2064
29	673	His Ala Ile Asp Trp Phe Ala Leu Leu Ser Ala Gly Asn Ser Pro Leu	688
31	2065	ATC GAT AAG GTT GGT CTG GTG ACT GAT GCT GGC ATA CAA AGT GTT ATA	2112
33	689	Ile Asp Lys Val Gly Leu Val Thr Asp Ala Gly Ile Gln Ser Val Ile	704
35	2113	CCA ACG CTG CTC AAT ACA CAA ACC TTA TCT GAT GAA GAT AAG AAG CTC	2160
37	705	Ala Thr Val Val Asn Thr Gln Ser Leu Ser Asp Glu Asp Lys Lys Leu	720
39	2161	CCA ATC ACT ACT CTG ACT AAT ACG TTG AAT CAG GTA CAG AAA ACT CAA	2208
41	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

RO 121280 B1

2449	TTG CTG GAC TAT CCA GCC TAT TTT GGC GCT TCC GCA GAA ACA GTG ACC	2496	1
317	Leu Leu Asp Tyr Pro Ala Tyr Phe Gly Ala Ser Ala Glu Thr Val Thr	322	3
2497	GAT ATC AGT TTG TGG ATG CTT TAT ACC CTG AGC TGT TAT AGC GAT TTA	2544	5
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	7
849	Leu Leu Gln Met Gly Glu Ala Gly Gly Thr Glu Asp Asp Val Leu Ala	864	9
2593	TAC TTA CGC ACA GCT AAT GCT ACC ACA CCG TTG AGC CAA TCT GAT GCT	2640	11
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	13
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	15
897	Ala Ala Trp Ser Val Leu Gly Gly Ile Ala Lys Thr Thr Pro Gln Leu	912	17
2737	GAT GCG CTT CTG CGT TTG CAA CAG GCA CAG AAC CAA ACT GGT CTT GGC	2784	19
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	21
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	23
945	Thr Leu Trp Gln Ser Thr Gly Gln Ala Leu Val Ala Gly Val Ser His	960	25
2881	GTC AAG GGC AGT AAC TGA	2898	27
961	Val Lys Gly Ser Asn End	966	
(2) INFORMAȚIA PENTRU SECV. ID. NR. 57:			
(i) CARACTERISTICI DE SECVENȚĂ:			
(A) LUNGIME: 965 aminoacizi			
(B) TIP: aminoacid			
(D) TOPOLOGIE: lineară			
(ii) TIPUL MOLECULEI: proteină			
(xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 57 (peptidă TccA):			
Trăsătura	De la	Până la	Descrierea
	1	10	SECV. ID. NR. 8
1	Met Asn Gln Leu Ala Ser Pro Leu Ile Ser Arg Thr Glu Glu Ile His	16	39
17	Asn Leu Pro Gly Lys Leu Thr Asp Leu Gly Tyr Thr Ser Val Phe Asp	32	41
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	43
65	His Gln Val Leu His His Phe Arg Arg Asn Ser Leu Ser Glu Ala Val	80	45
81	Gln Phe Gly Leu Arg Ser Pro Phe Ser Val Ser Gly Pro Asp Tyr Ala	96	47

RO 121280 B1

1	97	Asn Gln Phe Leu Asp Ala Asn Thr Gly Trp Lys Asp Lys Ala Pro Ser	111
3	113	Gly Ser Pro Glu Ala Asn Asp Ala Pro Val Ala Tyr Leu Thr His Ile	128
5	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
7	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
9	193	Ala Val Asn Ala Arg Leu Ser Thr Thr Arg Tyr Pro Asn Asn Leu Pro	208
11	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
13	241	Asn Phe Trp Ala Thr Ala Lys Gly Lys Leu Ser Asp Thr Thr Ala Ser	256
15	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
17	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
19	321	Cys Ser Thr Val Gly Gly Ser Thr Val Val Lys Ser Asp Asn Val Ser	336
21	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
23	369	Ala His Phe Ala Leu Thr Val Asn Asn Leu Thr Asp Asp Lys Leu Asp	384
25	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
27	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
29	449	Gly Trp Leu Arg Val Val Ala Pro Phe Ala Ile Thr Pro Ala Thr Pro	464
31	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
33	497	Gly Ala Arg Val Lys His Ile Ser Thr Ala Leu Gly Leu Asn His Arg	512
35	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
37	545	Arg Leu Ala Asn Leu Ala Arg Thr Leu Gly Ile Asn Pro Glu Ser Phe	560
39	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
41	593	Leu Ala Ala Asp Ile Leu Ser Leu Leu Gln Ala Leu Ser Ala Ile Ala	608
43	609	Gln Trp Gln Gln Gln His Asp Leu Glu Phe Ser Ala Leu Leu Leu Leu	624

RO 121280 B1

625	Leu Ser Asp Asn Pro Ile Ser Thr Ser Gln Gly Thr Asp Asp Gln Leu	640	1
641	Asn Phe Ile Arg Gln Val Trp Gln Asn Leu Gly Ser Thr Phe Val Gly	655	3
657	Ala Thr Leu Leu Ser Arg Ser Gly Ala Pro Leu Val Asp Thr Asn Gly	672	5
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	7
705	Ala Thr Val Val Asn Thr Gln Ser Leu Ser Asp Glu Asp Lys Lys Leu	720	9
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	11
753	Ser Leu Pro Ala Leu Leu Leu Arg Trp Ser Gly Gln Thr Thr Tyr Gln	768	13
769	Trp Leu Ser Ala Thr Trp Ala Leu Lys Asp Ala Val Lys Thr Ala Ala	784	15
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	17
817	Leu Leu Asp Tyr Pro Ala Tyr Phe Gly Ala Ser Ala Glu Thr Val Thr	832	19
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	21
865	Tyr Leu Arg Thr Ala Asn Ala Thr Thr Pro Leu Ser Gln Ser Asp Ala	880	23
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	25
913	Asp Ala Leu Leu Arg Leu Gln Gln Ala Gln Asn Gln Thr Gly Leu Gly	928	27
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	29
961	Val Lys Gly Ser Asn	965	31

(2) INFORMAȚIA PENTRU SECV. ID. NR. 58:

(i) CARACTERISTICI DE SECVENȚĂ:

(A) LUNGIME: 4698 baze perechi

(B) TIP: acid nucleic

(C) ÎMPLETIRE: dublă

(D) TOPOLOGIE: lineară

(ii) TIPUL MOLECULEI: ADN (genomic)

(xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 58 (tccB):

1	ATG TTA TCG ACA ATG GAA AAA CAA CTG AAT GAA TCC CAG CGT GAT GCC	43	43
1	Met Leu Ser Thr Met Glu Lys Gln Leu Asn Glu Ser Gln Arg Asp Ala	16	
49	TTG GTG ACT GCC TAT ATG AAT TTT GTG GCG CCG ACG TTG AAA GGC GTC	96	45
17	Leu Val Thr Gly Tyr Met Asn Phe Val Ala Pro Thr Leu Lys Gly Val	32	47

RO 121280 B1

1	97	AGT	GCT	CAG	CCG	GTG	ACG	GTG	GAA	GAT	TTA	TAC	GAA	TAT	TTG	CTG	ATT	144
3	33	Ser	Gly	Gln	Pro	Val	Thr	Val	Glu	Asp	Leu	Tyr	Glu	Tyr	Leu	Leu	Ile	48
5	145	GAC	CCG	GAA	GTC	GCT	GAT	GAG	GTT	GAG	ACG	AGT	CGG	GTA	GCA	CAA	CCG	192
	49	Asp	Pro	Glu	Val	Ala	Asp	Glu	Val	Glu	Thr	Ser	Arg	Val	Ala	Gln	Ala	64
7	193	ATT	CCC	AGC	ATA	CAG	CAA	TAT	ATC	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
9	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
11	289	AAT	GAT	AAC	CAA	TAT	GCT	ATC	TGG	GCT	CCG	GGG	GCT	GAG	GTT	CGA	AAT	336
13	97	Asn	Asp	Asn	Gln	Tyr	Ala	Ile	Trp	Ala	Ala	Gly	Ala	Glu	Val	Arg	Asn	112
15	337	TAC	GCT	GAA	AAC	TAT	ATT	TCA	CCC	ATC	ACC	CCG	CAG	GAA	AAA	AGC	CAT	384
	113	Tyr	Ala	Glu	Asn	Tyr	Ile	Ser	Pro	Ile	Thr	Arg	Gln	Glu	Lys	Ser	His	128
17	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
19	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
21	481	GTG	AGT	AAT	CTA	TAT	GTG	CTC	AGT	GGT	TAT	ATT	AAT	CAG	GAT	AAA	TTT	528
23	161	Val	Ser	Asn	Leu	Tyr	Val	Leu	Ser	Gly	Tyr	Ile	Asn	Gln	Asp	Lys	Phe	176
25	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
27	577	CCC	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
29	625	GCA	GGG	AAT	CCG	GTG	ACG	CCA	AAT	TGC	TGG	AAT	GAT	TGC	CAG	GAA	ATC	672
	209	Ala	Gly	Asn	Pro	Val	Thr	Pro	Asn	Cys	Trp	Asn	Asp	Trp	Gln	Glu	Ile	224
31	673	ACT	TTG	CCG	CTG	TCT	GGT	GAT	ACG	GTG	CTG	GAG	CAT	ACA	GTT	CGC	CCG	720
33	225	Thr	Leu	Pro	Leu	Ser	Gly	Asp	Thr	Val	Leu	Glu	His	Thr	Val	Arg	Pro	240
35	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
37	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
39	817	TAC	AAC	ATA	AAG	TTT	GGT	TAT	AAA	CGT	TAT	GAT	GAT	ACT	TGG	ACA	CCG	864
	273	Tyr	Asn	Ile	Lys	Phe	Gly	Tyr	Lys	Arg	Tyr	Asp	Asp	Thr				

RO 121280 B1

913	ACA CAG CGA TCC AGC CTG CTG ATT GAT GAA TCT ACC ACC ACA TTG GGC	960	1
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	3
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	5
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	7
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	9
369	Tyr Gly Tyr Leu Tyr Cys Asp Ser Ala Phe Asn Arg His Val Leu Arg	384	
1153	CCC TTA AGT AAG AAC TTT TTG TTC AGT ACT TAC CGT GAT GAA ACG GAT	1200	11
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	13
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	15
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	17
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	19
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	21
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	23
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	25
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	27
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	29
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	31
545	Leu Leu Asp Thr Leu His Thr Val Thr Val Lys Gly Ser Tyr Ile Ala	560	
			33
			35
			37
			39
			41

RO 121280 B1

1	1681	TGG	GAA	GGG	GAA	ACA	CCT	ACC	GGT	TAT	AAT	CTG	TAT	ATT	CCA	GAT	GGT	1728
3	561	Trp	Glu	Gly	Glu	Thr	Pro	Thr	Gly	Tyr	Asn	Leu	Tyr	Ile	Pro	Asp	Gly	576
5	1729	ACC	GTG	TTG	CTA	GAT	TGG	TTT	GAT	AAA	ATA	AAT	TTT	GCT	ATT	GCT	CTT	1776
	577	Thr	Val	Leu	Leu	Asp	Trp	Phe	Asp	Lys	Ile	Asn	Phe	Ala	Ile	Gly	Leu	592
7	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
9	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
11	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
13	1921	AGT	ACT	CAA	ACT	TTC	GGA	CTT	ACC	AGC	GGT	GCG	ACT	TAT	TCT	ACA	ACT	1968
15	641	Ser	Thr	Gln	Thr	Phe	Gly	Leu	Thr	Ser	Gly	Ala	Thr	Tyr	Ser	Thr	Thr	656
17	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
19	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
21	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
23	2113	GTT	CCG	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
25	2161	TCC	CGT	TAC	GAT	AGT	AAA	CGT	GGT	CTG	GTG	CAA	TAT	CTG	GAC	TTC	TGG	2208
27	721	Ser	Arg	Tyr	Asp	Ser	Lys	Arg	Gly	Leu	Val	Gln	Tyr	Leu	Asp	Phe	Trp	736
29	2209	ACC	TCA	TCA	TTA	CCC	GCG	AAA	ACC	COT	CTT	AAC	ACC	ACC	TTT	GTG	COT	2256
	737	Thr	Ser	Ser	Leu	Pro	Ala	Lys	Thr	Arg	Leu	Asn	Thr	Thr	Phe	Val	Arg	752
31	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
33	2305	ACC	TTG	CAG	GCA	GAT	CCT	TCT	CTG	GAA	OCA	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
35	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
37	2401	GAA	TTG	TTC	TTT	CAC	CTG	CCG	TTT	TTG	GTT	GCT	ACA	CGC	TTT	GCC	AAC	2448
39	801	Glu	Leu	Phe	Phe	His	Leu	Pro	Phe	Leu	Val	Ala	Thr	Arg	Phe	Ala	Asn	816
41	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	332

RO 121280 B1

2497 333	CCG GCG ATG AAA AAC AAG CCA CAC AAT GCC CCG GCT TAT TGG AAT GTA	2544 348	1
	Pro Ala Met Lys Asn Lys Pro His Asn Ala Pro Ala Tyr Trp Asn Val		3
2515 349	CGT CCG TTC GTT CAA CGA AAC AGC GAT TTG TCA CGT CAT TTG GAC GAT	2592 364	5
	Arg Pro Leu Val Glu Gly Asn Ser Asp Leu Ser Arg His Leu Asp Asp		
2593 365	TCT ATA GAC CCA GAT ACT CAA GCT TAT GCT CAT CCG GTG ATA TAC CAG	2640 380	7
	Ser Ile Asp Pro Asp Thr Gln Ala Tyr Ala His Pro Val Ile Tyr Gln		
2641 381	AAA GCG GTG TTT ATT GCC TAT GTC AGT AAC CTG ATT GCT CAG GGA GAT	2688 396	9
	Lys Ala Val Phe Ile Ala Tyr Val Ser Asn Leu Ile Ala Gln Gly Asp		
2689 397	ATG TGG TAT CGC CAA TTG ACT CGT GAC GGT CTG ACT CAG GCC CGT GTC	2736 412	11
	Met Trp Tyr Arg Gln Leu Thr Arg Asp Gly Leu Thr Gln Ala Arg Val		13
2737 413	TAT TAC AAT CTG GCC GCT GAA TTG CTA GGG CCT CGT CCG GAT GTA TCG	2784 420	15
	Tyr Tyr Asn Leu Ala Ala Glu Leu Leu Gly Pro Arg Pro Asp Val Ser		
2785 429	CTG AGT AGC ATT TCG ACG CCG CAA ACC CTG GAT ACC TTA GCA GCC GGG	2832 444	17
	Leu Ser Ser Ile Trp Thr Pro Gln Thr Leu Asp Thr Leu Ala Ala Gly		
2833 445	CAA AAA GCG GTT TTA CGT GAT TTT GAG CAC CAG TTG GCT AAT AGT GAT	2880 460	19
	Gln Lys Ala Val Leu Arg Asp Phe Glu His Gln Leu Ala Asn Ser Asp		
2881 461	ACC GCT TTA CCC GCA TTG CCG GGC CGC AAT GTC AGC TAC TTG AAA CTG	2928 476	21
	Thr Ala Leu Pro Ala Leu Pro Gly Arg Asn Val Ser Tyr Leu Lys Leu		23
2929 477	GCA GAT AAT GGC TAC TTT AAT GAA CCG CTC AAT GTT CTG ATG TTG TCT	2976 492	25
	Ala Asp Asn Gly Tyr Phe Asn Glu Pro Leu Asn Val Leu Met Leu Ser		
2977 493	CAC TGG GAT ACG TTG GAT GCA CCG TTA TAC AAT CTG CGT CAT AAC CTG	3024 1008	27
	His Trp Asp Thr Leu Asp Ala Arg Leu Tyr Asn Leu Arg His Asn Leu		
3025 1009	ACC GTT GAT GGC AAG CCG CTT TCG CTC CCG CTC TAT GCT GCG CCT GTT	3072 1024	29
	Thr Val Asp Gly Lys Pro Leu Ser Leu Pro Leu Tyr Ala Ala Pro Val		
3073 1025	GAT CCG GTA GCG TTG TTG GCT CAG CGT GCT CAG TCC GGC ACG TTG ACG	3120 1040	31
	Asp Pro Val Ala Leu Leu Ala Gln Arg Ala Gln Ser Gly Thr Leu Thr		33
3121 1041	AAT GGC CTC AGT GGC GCC ATG TTG ACC GTG CCG CCA TAC CGT TTC ACC	3168 1056	35
	Asn Gly Val Ser Gly Ala Met Leu Thr Val Pro Pro Tyr Arg Phe Ser		
3169 1057	GCT ATG TTG CCG CGA GCT TAC AGC GCC GTG GGT ACG TTG ACC AGT TTT	3216 1072	37
	Ala Met Leu Pro Arg Ala Tyr Ser Ala Val Gly Thr Leu Thr Ser Phe		
3217 1073	GGT CAG AAC CTG CTT AGT TTG TTG GAA CGT ACC GAA CGA GCC TGT CAA	3264 1088	39
	Gly Gln Asn Leu Leu Ser Leu Leu Glu Arg Ser Glu Arg Ala Cys Gln		41

RO 121280 B1

1	3265	GAA GAG TTG GCG CAA CAG CAA CTG TTG GAT ATG TCC AGC TAT GCC ATC	3311
3	1089	Glu Glu Leu Ala Gln Gln Gln Leu Leu Asp Met Ser Ser Tyr Ala Ile	1104
5	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
7	3361	CTG CTA GCT AGT CAG GCT ACG GCA CAA CAG COT CAT GAC CAT TAT TAC	3408
	1121	Leu Leu Ala Ser Gln Ala Thr Ala Gln Gln Arg His Asp His Tyr Tyr	1156
9	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
11	3457	ACC CAA ACG 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
13	3505	ACT GCC AGT GCG GCA CTG AAA GTG ATC CCG AAT ATC TTT GGT TTG GCT	3552
15	1169	Thr Ala Ser Gly Ala Leu Lys Val Ile Pro Asn Ile Phe Gly Leu Ala	1184
17	3553	GAT GGC GGC TCG CGC TAT GAA GGA GTA ACG GAA GCG ATT GCC ATC GCG	3600
	1185	Asp Gly Gly Ser Arg Tyr Glu Gly Val Thr Glu Ala Ile Ala Ile Gly	1200
19	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
21	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
23	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
25	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
27	3793	GCA CAG CAG GTA CAA ATT CGG ACC ATG CTG ACT TAC TTA ACT ACT CGT	3840
29	1265	Ala Gln Gln Val Gln Ile Arg Thr Met Leu Thr Tyr Leu Thr Thr Arg	1280
31	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
33	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
35	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
37	3985	CAG ACC GGT ACC TGG AAC GAC CAT TAC CGT GGT TTG CAA GTG GCG GAG	4032
	1329	Gln Thr Gly Thr Trp Asn Asp His Tyr Arg Gly Leu Gln Val Gly Glu	1344
39	4033	ACA CTG CAA CTC AAT TTG CAT CAG ATG GAA GCG GCC TAT TTA GTT CGT	4080
41	1345	Thr Leu Gln Leu Asn Leu His Gln Met Glu Ala Ala Tyr Leu Val Arg	1360

RO 121280 B1

4081	CAC GAA CGC CGT CTT AAT GTG ATC CGT ACT GTG TCG CTC AAA AGC CTA	4123	1
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	3
1377	Leu Gly Asp Asp Gly Phe Gly Lys Leu Lys Thr Glu Gly Lys Val Asp	1392	5
4177	TTT CCA TTA AGC GAA AAG CTG TTT GAC AAC GAC TAT CCG GGG CAC TAT	4224	7
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	9
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	11
1425	Pro Tyr Gln Asn Val Lys Ala Thr Leu Thr Gln Thr Ser Ser Ser Ile	1440	13
4321	TTG TTA GCA GCA GAT ATC AAT GGT GTT AAA CGT CTC AAT GAT CCG ACA	4368	15
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	17
1457	Gly Lys Glu Gly Asp Ala Thr His Ile Val Thr Asn Leu Arg Ala Ser	1472	19
4417	CAG CAG GTG GCG CTC TCT TCT GGC ATT AAT GAT GCC GGT AGC TTT GAG	4464	21
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	23
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	25
1505	Val Ser Lys Trp Thr Leu Asn Phe Pro Arg Ser Val Asp Glu His Ile	1520	27
4561	GAC GAT AAG ACA TTG AAA GCG GAT GAG ATG CAG GCC GCA CTG TTG GCG	4608	29
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	31
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	4698	33
1553	Gly Ala Ser Phe Ala Asn Gln Val Lys Lys Thr Leu Ser End	1566	35

(2) INFORMAȚIA PENTRU SECV. ID. NR. 59:

(i) CARACTERISTICI DE SECVENȚĂ:

(A) LUNGIME: 1665 aminoacizi

(B) TIP: aminoacid

(C) TOPOLOGIE: lineară

(ii) TIPUL MOLECULEI: proteină

(xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 59 (peptidă TccB):

Trăsătura	De la	Până la	Descrierea
	1	11	SECV. ID. NR. 7

RO 121280 B1

1
i 11 SEQ ID NO:7

3
1 Met Leu Ser Thr Met Glu Lys Gln Leu Asn Glu Ser Gln Arg Asp Ala 16
5 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
7 49 Asp Pro Glu Val Ala Asp Glu Val Glu Thr Ser Arg Val Ala Gln Ala 64
9 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
11 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
13 129 Tyr Phe Ser Glu Leu Glu Thr Thr Leu Asn Gln Asn Arg Leu Asp Pro 144
15 145 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
17 177 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
19 209 Ala Gly Asn Pro Val Thr Pro Asn Cys Trp Asn Asp Trp Gln Glu Ile 224
21 225 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
23 257 Ala Val Gln Lys Asp Ala Asp Gly Lys Asn Ile Gly Lys Thr His Ala 272
25 273 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
27 305 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
29 337 Glu Thr Asp Ser Asn Pro Tyr Gly Arg Leu Met Leu Gly Val Phe Val 352
31 353 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
33 385 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
35 417 Thr Lys Val Val Thr Gly Ala Thr Glu Asp Pro Glu Asn Thr Gly Trp 432
37 433 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
39 465 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
41 497 Tyr Leu Asp Tyr His Asp Gly Asn Tyr Tyr Thr Phe His Asn Ala Ile 512

RO 121280 B1

513	Ile Asn Tyr Tyr Pro Ser Gly Tyr Gly Gly Gly Ser Val Pro Asn Gly	523	1
529	Thr Trp Ala Leu Glu Gln Arg Ile Asn Glu Gly Trp Ala Ile Ala Pro	544	3
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	5
577	Thr Val Leu Leu Asp Trp Phe Asp Lys Ile Asn Phe Ala Ile Gly Leu	592	7
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	9
625	Glu Ile Asn Ala Glu Thr Ala Asp Gly Arg Asn Leu Phe Lys Arg Tyr	640	11
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	13
673	Leu Gln Val Cys Leu Asn Val Val Trp Asp His Tyr Asp Arg Pro Ser	688	15
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	17
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	19
753	Thr Leu Ile Glu Lys Ala Asn Leu Gly Leu Asp Ser Leu Leu Asp Tyr	768	21
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	23
801	Glu Leu Phe Phe His Leu Pro Phe Leu Val Ala Thr Arg Phe Ala Asn	816	25
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	27
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	29
881	Lys Ala Val Phe Ile Ala Tyr Val Ser Asn Leu Ile Ala Gln Gly Asp	896	31
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	33
929	Leu Ser Ser Ile Trp Thr Pro Gln Thr Leu Asp Thr Leu Ala Ala Gly	944	35
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	37
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	39
1009	Thr Val Asp Gly Lys Pro Leu Ser Leu Pro Leu Tyr Ala Ala Pro Val	1024	41
1025	Asp Pro Val Ala Leu Leu Ala Gln Arg Ala Gln Ser Gly Thr Leu Thr	1040	

RO 121280 B1

1 1041 Asn Gly Val Ser Gly Ala Met Leu Thr Val Pro Pro Tyr Arg Phe Ser 1056
3 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
5 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
7 1121 Leu Leu Ala Ser Gln Ala Thr Ala Gln Gln Arg His Asp His Tyr Tyr 1136
9 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
11 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
13 1201 Leu Met Ala Ala Gly Gln Ala Thr Ser Val Val Ala Glu Arg Leu Ala 1216
15 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
17 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
19 1281 Phe Thr Gln Ala Thr Leu Tyr Gln Trp Leu Ser Gly Gln Leu Ser Ala 1296
21 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
23 1329 Gln Thr Gly Thr Trp Asn Asp His Tyr Arg Gly Leu Gln Val Gly Glu 1344
25 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
27 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
29 1409 Leu Arg Gln Ile Lys Thr Val Ser Val Thr Leu Pro Thr Leu Val Gly 1424
31 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
33 1457 Gly Lys Glu Gly Asp Ala Thr His Ile Val Thr Asn Leu Arg Ala Ser 1472
35 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
37 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
39 1537 Asn Met Asp Asp Val Leu Val Gln Val His Tyr Thr Ala Cys Asp Gly 1552
41 1553 Gly Ala Ser Phe Ala Asn Gln Val Lys Lys Thr Leu Ser 1565

2) INFORMAȚIA PENTRU SECV. ID. NR. 60:

(i) CARACTERISTICI DE SECVENȚĂ:

(A) LUNGIME: 3132 baze perechi

(B) TIP: acid nucleic

(C) ÎMPLETIRE: dublă

(D) TOPOLOGIE: lineară

RO 121280 B1

(ii) TIPUL MOLECULEI: ADN (genomic)

(xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 60 (tccC):

1	ATG AGT CCG TCT GAG ACT ACT CTT TAT ACT CAA ACC CCA ACA GTC AGC	48	1
1	Met Ser Pro Ser Glu Thr Thr Leu Tyr Thr Gln Thr Pro Thr Val Ser	16	3
49	GTG TTA GAT AAT CGC GGT CTG TCC ATT CGT GAT ATT GGT TTT CAC CGT	96	5
17	Val Leu Asp Asn Arg Gly Leu Ser Ile Arg Asp Ile Gly Phe His Arg	32	7
97	ATT GTA ATC GGG GGG GAT ACT GAC ACC CGC GTC ACC CGT CAC CAG TAT	144	9
33	Ile Val Ile Gly Gly Asp Thr Asp Thr Arg Val Thr Arg His Gln Tyr	48	11
145	GAT GCC CGT GGA CAC CTG AAC TAC AGT ATT GAC CCA CGC TTG TAT GAT	192	13
49	Asp Ala Arg Gly His Leu Asn Tyr Ser Ile Asp Pro Arg Leu Tyr Asp	64	15
193	GCA AAG CAG GCT GAT AAC TCA GTA AAG CCT AAT TTT GTC TGG CAG CAT	240	17
65	Ala Lys Gln Ala Asp Asn Ser Val Lys Pro Asn Phe Val Trp Gln His	80	19
241	GAT CTG GCC GGT CAT GCC CTG CGG ACA GAG AGT GTC GAT GCT GGT CGT	288	21
81	Asp Leu Ala Gly His Ala Leu Arg Thr Glu Ser Val Asp Ala Gly Arg	96	23
289	ACT GTT GCA TTG AAT GAT ATT GAA GGT CGT TCG GTA ATG ACA ATG AAT	336	25
97	Thr Val Ala Leu Asn Asp Ile Glu Gly Arg Ser Val Met Thr Met Asn	112	27
337	GCG ACC GGT GTT CGT CAG ACC CGT CGC TAT GAA GGC AAC ACC TTG CCC	384	29
113	Ala Thr Gly Val Arg Gln Thr Arg Arg Tyr Glu Gly Asn Thr Leu Pro	128	31
385	GGT CGC TTG TTA TCT GTG AGC GAG CAA GTT TTC AAC CAA GAG AGT GCT	432	33
129	Gly Arg Leu Leu Ser Val Ser Glu Gln Val Phe Asn Gln Glu Ser Ala	144	35
433	AAA GTG ACA GAG CGC TTT ATC TGG GCT GGG AAT ACA ACC TCG CAG AAA	480	37
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 CGG 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	

RO 121280 B1

1	625	GAC GAA ACT GTC TGG CAG GGA ATG CTG GCA AGT GAG GTC TAT ACG ACA	672
3	269	Asp Glu Thr Val Trp Gln Gly Met Leu Ala Ser Glu Val Tyr Thr Thr	224
5	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
7	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
9	769	GGG AGT TGG TTG ACC GTG AAA GGC CAG ACT 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
11	817	TCC CTG AGC TGG TCA GCC GCA GGT CAT AAA TTG COT GAA GAG CAC GGT	864
	273	Ser Leu Ser Trp Ser Ala Ala Gly His Lys Leu Arg Glu Glu His Gly	288
13	865	AAC GGC GTG GTT ACG GAG TAC AGT TAT GAG CCG GAA ACT CAA CGT CTG	912
15	289	Asn Gly Val Val Thr Glu Tyr Ser Tyr Glu Pro Glu Thr Gln Arg Leu	304
17	913	ATA GGT ATC ACC ACC CCG CGT CCC GAA GCG 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
19	961	GTA TTG CAG GAT CTA CGC TAT AAG TAT GAT CCG GTG GCG AAT GTT ATC	1008
	321	Val Leu Gln Asp Leu Arg Tyr Lys Tyr Asp Pro Val Gly Asn Val Ile	336
21	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
23	1057	GTG GAG CCG GAG AAT CGC TAT GTT TAT GAT TCT CTG TAT CAG CTT ATC	1104
25	353	Val Glu Pro Glu Asn Arg Tyr Val Tyr Asp Ser Leu Tyr Gln Leu Met	368
27	1105	AGT GCG ACA GCG 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
29	1153	CTT CCC TCA CCC GTT ATA CCT GTT CGT 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
31	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
33	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
35	1297	ACC GTT TCA AGC CGC AGT AAC CCG 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
37	1345	GAT CCA ACC CGA GTG GAT GCG CTA TTT GAT TCC GGC GGT CAT CAG AAG	1392
39	449	Asp Pro Thr Arg Val Asp Ala Leu Phe Asp Ser Gly Gly His Gln Lys	464
41	1393	ATG TTA ATA CCG GGG CAA AAT CTG GAT TGG AAT ATT CGG GGT GAA TTG	1440

RO 121280 B1

455	Met	Leu	Ile	Pro	Gly	Gln	Asn	Leu	Asp	Trp	Asn	Ile	Arg	Gly	Glu	Leu	455	1
1431	CAA	CGA	GTC	ACA	CCG	GTG	AGC	CGT	GAA	AAT	AGC	AGT	GAC	AGT	GAA	TGG	1431	3
431	Gln	Arg	Val	Thr	Pro	Val	Ser	Arg	Glu	Asn	Ser	Ser	Asp	Ser	Glu	Trp	431	5
1489	TAT	COC	TAT	AGC	ACT	GAT	GOC	ATG	CGG	CTG	CTA	AAA	GTG	AGT	GAA	CAG	1489	7
497	Tyr	Arg	Tyr	Ser	Ser	Asp	Gly	Met	Arg	Leu	Leu	Lys	Val	Ser	Glu	Gln	497	9
1537	CAG	ACG	GCC	AAC	AGT	ACT	CAA	GTA	CAA	CGG	GTG	ACT	TAT	CTG	CCG	GGA	1537	11
513	Gln	Thr	Gly	Asn	Ser	Thr	Gln	Val	Gln	Arg	Val	Thr	Tyr	Leu	Pro	Gly	513	13
1585	TTA	GAG	CTA	CGG	ACA	ACT	GGG	GTT	GCA	GAT	AAA	ACA	ACC	GAA	GAT	TTG	1585	15
529	Leu	Glu	Leu	Arg	Thr	Thr	Gly	Val	Ala	Asp	Lys	Thr	Thr	Glu	Asp	Leu	529	17
1633	CAG	GTG	ATT	ACG	GTA	GGT	GAA	GCG	GGT	CGC	GCA	CAG	GTA	AGG	GTA	TTG	1633	19
545	Gln	Val	Ile	Thr	Val	Gly	Glu	Ala	Gly	Arg	Ala	Gln	Val	Arg	Val	Leu	545	21
1681	CAC	TGG	GAA	AGT	GGT	AAG	CCG	ACA	GAT	ATT	GAC	AAC	AAT	CAG	GTG	CGC	1681	23
561	His	Trp	Glu	Ser	Gly	Lys	Pro	Thr	Asp	Ile	Asp	Asn	Asn	Gln	Val	Arg	561	25
1729	TAC	AGC	TAC	GAT	AAT	CTG	CTT	GCC	TCC	AGC	CAG	CTT	GAA	CTG	GAT	AGC	1729	27
577	Tyr	Ser	Tyr	Asp	Asn	Leu	Leu	Gly	Ser	Ser	Gln	Leu	Glu	Leu	Asp	Ser	577	29
1777	GAA	GCG	CAG	ATT	CTC	AGT	CAG	GAA	GAG	TAT	TAT	CCG	TAT	GCC	GCT	ACC	1777	31
593	Glu	Gly	Gln	Ile	Leu	Ser	Gln	Glu	Glu	Tyr	Tyr	Pro	Tyr	Gly	Gly	Thr	593	33
1825	GCG	ATA	TGG	GCG	GCG	AGA	AAT	CAG	ACA	GAA	GCC	AGC	TAC	AAA	TTT	ATT	1825	35
609	Ala	Ile	Trp	Ala	Ala	Arg	Asn	Gln	Thr	Glu	Ala	Ser	Tyr	Lys	Phe	Ile	609	37
1873	CGT	TAC	TCC	GGT	AAA	GAG	CGG	GAT	GCC	ACT	GCA	TTG	TAT	TAT	TAC	GCC	1873	39
625	Arg	Tyr	Ser	Gly	Lys	Glu	Arg	Asp	Ala	Thr	Gly	Leu	Tyr	Tyr	Tyr	Gly	625	
1921	TAC	CGT	TAT	TAT	CAA	CCT	TGG	GTG	GGT	CGA	TGG	TTG	AGT	GCT	GAT	CCG	1921	
641	Tyr	Arg	Tyr	Tyr	Gln	Pro	Trp	Val	Gly	Arg	Trp	Leu	Ser	Ala	Asp	Pro	641	
1969	GCG	GGA	ACC	GTG	GAT	GCG	CTG	AAT	TTG	TAC	CGA	ATG	GTG	AGG	AAT	AAC	1969	
657	Ala	Gly	Thr	Val	Asp	Gly	Leu	Asn	Leu	Tyr	Arg	Met	Val	Arg	Asn	Asn	657	
2017	CCC	ATC	ACA	TTG	ACT	CAC	CAT	CAC	GGA	TTA	GCA	CCG	TCT	CCA	AAT	AGA	2017	
673	Pro	Ile	Thr	Leu	Thr	Asp	His	Asp	Gly	Leu	Ala	Pro	Ser	Pro	Asn	Arg	673	
2065	AAT	CGA	AAT	ACA	TTT	TGG	TTT	GCT	TCA	TTT	TTG	TTT	CGT	AAA	CCT	GAT	2065	
689	Asn	Arg	Asn	Thr	Phe	Trp	Phe	Ala	Ser	Phe	Leu	Phe	Arg	Lys	Pro	Asp	689	
2113	GAG	GGA	ATG	TCC	GCG	TCA	ATG	AGA	CGG	GGA	CAA	AAA	ATT	GCG	AGA	GCC	2113	
705	Glu	Gly	Met	Ser	Ala	Ser	Met	Arg	Arg	Gly	Gln	Lys	Ile	Gly	Arg	Ala	705	
2161	ATT	GCC	GGC	GCG	ATT	GCG	ATT	GCG	GGT	CTT	GCG	GCT	ACC	ATT	GCC	GCT	2161	
721	Ile	Ala	Gly	Gly	Ile	Ala	Ile	Gly	Gly	Leu	Ala	Ala	Thr	Ile	Ala	Ala	721	

RO 121280 B1

1	2209	ACG GCT GGC GCG GCT ATC CCC GTC ATT CTG GCG GTT GCG GCC GTA GGC	2256
	737	Thr Ala Gly Ala Ala Ile Pro Val Ile Leu Gly Val Ala Ala Val Gly	752
3			
	2257	GCG GCG ATT GCG GCG TTG ATG GGA TAT AAC GTC GGT AGC CTG CTG GAA	2304
5	753	Ala Gly Ile Gly Ala Leu Met Gly Tyr Asn Val Gly Ser Leu Leu Glu	768
	2305	AAA GGC GCG GCA TTA CTT GCT CGA CTC GTA CAG GCG AAA TCG ACG TTA	2352
7	769	Lys Gly Gly Ala Leu Leu Ala Arg Leu Val Gln Gly Lys Ser Thr Leu	784
	2353	GTA CAG TCG GCG GCT GCG GCG GCT GCC GGA GCG AGT TCA GCC GCG GCT	2400
9	785	Val Gln Ser Ala Ala Gly Ala Ala Ala Gly Ala Ser Ser Ala Ala Ala	800
11	2401	TAT GGC GCA CGG GCA CAA GGT GTC GGT GTT GCA TCA GCC GCC GCG GCG	2448
	801	Tyr Gly Ala Arg Ala Gln Gly Val Gly Val Ala Ser Ala Ala Gly Ala	816
13			
	2449	GTA ACA GCG GCT GTG GGA TCA TGG ATA AAT AAT GCT GAT CCG GCG ATT	2496
15	817	Val Thr Gly Ala Val Gly Ser Trp Ile Asn Asn Ala Asp Arg Gly Ile	832
	2497	GCG GCG GCT ATT GCG GCC GCG AGT GCG GTA GCG ACC ATT GAT ACT ATG	2544
17	833	Gly Gly Ala Ile Gly Ala Gly Ser Ala Val Gly Thr Ile Asp Thr Met	848
	2545	TTA GCG ACT GCC TCT ACC CTT ACC CAT GAA GTC GCG GCA GCG GCG GGT	2592
19	849	Leu Gly Thr Ala Ser Thr Leu Thr His Glu Val Gly Ala Ala Ala Gly	864
	2593	GCG GCG GCG GGT GCG ATG ATC ACC GGT ACG CAA GCG AGT ACT CCG GCA	2640
21	865	Gly Ala Ala Gly Gly Met Ile Thr Gly Thr Gln Gly Ser Thr Arg Ala	880
23	2641	GGT ATC CAT GCC GGT ATT GCG ACC TAT TAT GCG TCC TGG ATT GGT TTT	2688
	881	Gly Ile His Ala Gly Ile Gly Thr Tyr Tyr Gly Ser Trp Ile Gly Phe	896
25			
	2689	GGT TTA GAT GTC CCT AGT AAC CCC GCC GGA CAT TTA GCG AAT TAC GCA	2736
27	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
29	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 GCG CCG GTT GTC AGC CCA	2832
31	929	Met Gly Gly Gly Phe Leu Ser Arg Leu Leu Gly Arg Val Val Ser Pro	944
33	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
35			
	2881	CCT GTC TTT CAG CCG ATA TTT AGT GTT CTC GCG GCG CTT GTC GGT GGT	2928
37	961	Pro Val Phe Glu Pro Ile Phe Ser Val Leu Gly Gly Leu Val Gly Gly	976
	2929	ATT GGA ACT GCG CTG CAC AGA GTG ATG GGA AGA GAG AGT TGG ATT TCC	2976
39	977	Ile Gly Thr Gly Leu His Arg Val Met Gly Arg Glu Ser Trp Ile Ser	992
41	2977	AGA GCG TTA AGT GCT GCC GGT AGT GGT ATA GAT CAT GTC GCT GCG ATG	3024

RO 121280 B1

993	Arg Ala Leu Ser Ala Ala Gly Ser Gly Ile Asp His Val Ala Gly Met	1013	1
3025	ATT GGT AAT CAG ATC AGA GGC AGG GTC TTG ACC ACA ACC GGG ATC GCT	3071	3
1009	Ile Gly Asn Gln Ile Arg Gly Arg Val Leu Thr Thr Thr Gly Ile Ala	1024	5
3073	EAT GCG ATA GAC TAT GGC ACC AGT GCT GTG GGA GCC GCA CGA CGA GTT	3120	7
1025	Asn Ala Ile Asp Tyr Gly Thr Ser Ala Val Gly Ala Ala Arg Arg Val	1040	
3121	TTT TCT TTG TAA	3132	9
1041	Phe Ser Leu End	1043	

(2) INFORMAȚIA PENTRU SECV. ID. NR. 61:

(i) CARACTERISTICI DE SECVENȚĂ:

(A) LUNGIME: 1043 aminoacizi

(B) TIP: aminoacid

(C) TOPOLOGIE: lineară

(ii) TIPUL MOLECULEI: proteină

(xi) DESCRIEREA SECVENȚEI: SECV. ID. NR. 61 (peptidă TccC):

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	21
33	Ile Val Ile Gly Gly Asp Thr Asp Thr Arg Val Thr Arg His Gln Tyr	48	23
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	25
81	Asp Leu Ala Gly His Ala Leu Arg Thr Glu Ser Val Asp Ala Gly Arg	96	27
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	29
129	Gly Arg Leu Leu Ser Val Ser Glu Gln Val Phe Asn Gln Glu Ser Ala	144	31
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	33
177	Val Thr Arg Leu Met Ser Gln Ser Leu Ala Gly Ala Met Leu Ser Gln	192	35
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	37
225	Gln Ser Thr Thr Asn Ala Ile Gly Ala Leu Leu Thr Gln Thr Asp Ala	240	39
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	41
273	Ser Leu Ser Trp Ser Ala Ala Gly His Lys Leu Arg Glu Glu His Gly	288	43
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	45

RO 121280 B1

1	321	Val Leu Gln Asp Leu Arg Tyr Lys Tyr Asp Pro Val Gly Asn Val Ile	336
3	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
5	369	Ser Ala Thr Gly Arg Glu Met Ala Asn Ile Gly Gln Gln Ser Asn Gln	384
7	395	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
9	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
11	449	Asp Pro Thr Arg Val Asp Ala Leu Phe Asp Ser Gly Gly His Gln Lys	464
13	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
15	497	Tyr Arg Tyr Ser Ser Asp Gly Met Arg Leu Leu Lys Val Ser Glu Gln	512
17	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
19	545	Gln Val Ile Thr Val Gly Glu Ala Gly Arg Ala Gln Val Arg Val Leu	560
21	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
23	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
25	625	Arg Tyr Ser Gly Lys Glu Arg Asp Ala Thr Gly Leu Tyr Tyr Tyr Gly	640
27	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
29	673	Pro Ile Thr Leu Thr Asp His Asp Gly Leu Ala Pro Ser Pro Asn Arg	688
31	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
33	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
35	753	Ala Gly Ile Gly Ala Leu Met Gly Tyr Asn Val Gly Ser Leu Leu Glu	768
37	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
39	801	Tyr Gly Ala Arg Ala Gln Gly Val Gly Val Ala Ser Ala Ala Gly Ala	816
41	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

RO 121280 B1

849	Leu Gly Thr Ala Ser Thr Leu Thr His Glu Val Gly Ala Ala Ala Gly	864	1
865	Gly Ala Ala Gly Gly Met Ile Thr Gly Thr Gln Gly Ser Thr Arg Ala	880	3
881	Gly Ile His Ala Gly Ile Gly Thr Tyr Tyr Gly Ser Trp Ile Gly Phe	896	5
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	7
929	Met Gly Gly Gly Phe Leu Ser Arg Leu Leu Gly Arg Val Val Ser Pro	944	9
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	11
977	Ile Gly Thr Gly Leu His Arg Val Met Gly Arg Glu Ser Trp Ile Ser	992	13
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	15
1025	Asn Ala Ile Asp Tyr Gly Thr Ser Ala Val Gly Ala Ala Arg Arg Val	1040	17
1041	Phe Ser Leu 1043		

(51) Int.Cl.

C12N 5/14 (2006.01);

C12N 15/00 (2006.01);

A01G 13/00 (2006.01);

A01H 1/00 (2006.01)

```

- ATG CAG GAT TGT CCG GAA GTA TCG ATT ACA AGC CCG TCA CTT CCC AAA GGT GGC GGT
  TAC GTC CTA ACA GGC CTT CAT AGC TAA TGT TGC CAC AGT GAA GGG TTT CCA CCG CCG
  1 Met Gln Asp Cys Pro Glu Val Ser Ile Thr Thr Leu Ser Leu Pro Lys Gly Gly Gly
      P2Psh
58 GCT ATC AAT GGC ATG CCA GAA GCA CTG AAT GGT CCC GGC CTT GAT GGA ATG GGC TCC
  CGA TAG TTA CCG TAC CTT CTT CGT GAC TTA CCA CCG CCG GGA CTA CTT TAC CCG ACC
  20 Ala Ile Asn Gly Met Gly Glu Ala Leu Asn Ala Ala Gly Pro Asp Gly Met Ala Ser
  115 CTA TCT CTG CCA TTA CCC CTT TCG ACC GGC AGA CCG AGC CTT CCT GCA TTA TCG CTC
  GAT AGA GAC GGT AAT GGG GAA AGC TGG CCG TCT CTT TCC GGA GGA CTT AAT AGC GAC
  39 Leu Ser Leu Pro Leu Ser Thr Gly Asn Gly Thr Ala Pro Gly Leu Ser Leu
  172 ATT TAC ACC AAC AGT CCA GGT AAT GGS CTT TTC CCC ATC GGC TGG CAA TGC GGT GTT
  TAA ATG TCG TTG TCA CGT CCA TTA CCC GGA AAG CCG TAG CCG ACC GTT ACC CCA CAA
  56 Ile Tyr Ser Asn Ser Ala Gly Asn Gly Pro Phe Gly Ile Gly Trp Gln Cys Gly Val
  229 ATG TCC ATT AGC CCA CCC ACC CAA CAT GGC CTT CAA CAT TGA CCA GGT
  TAC AGG TAA TCG GCT CCG TGG GTT GTA CCG GAA GTT GTA ACT GCT CCA
  77 Met Ser Ile Ser Arg Thr Gln His Gly Leu Gln His ... Arg Arg
      P2.35R

```

Fig. 1

(51) Int.Cl.

C12N 5/14 (2006.01);

C12N 15/00 (2006.01);

A01G 13/00 (2006.01);

A01H 1/00 (2006.01)

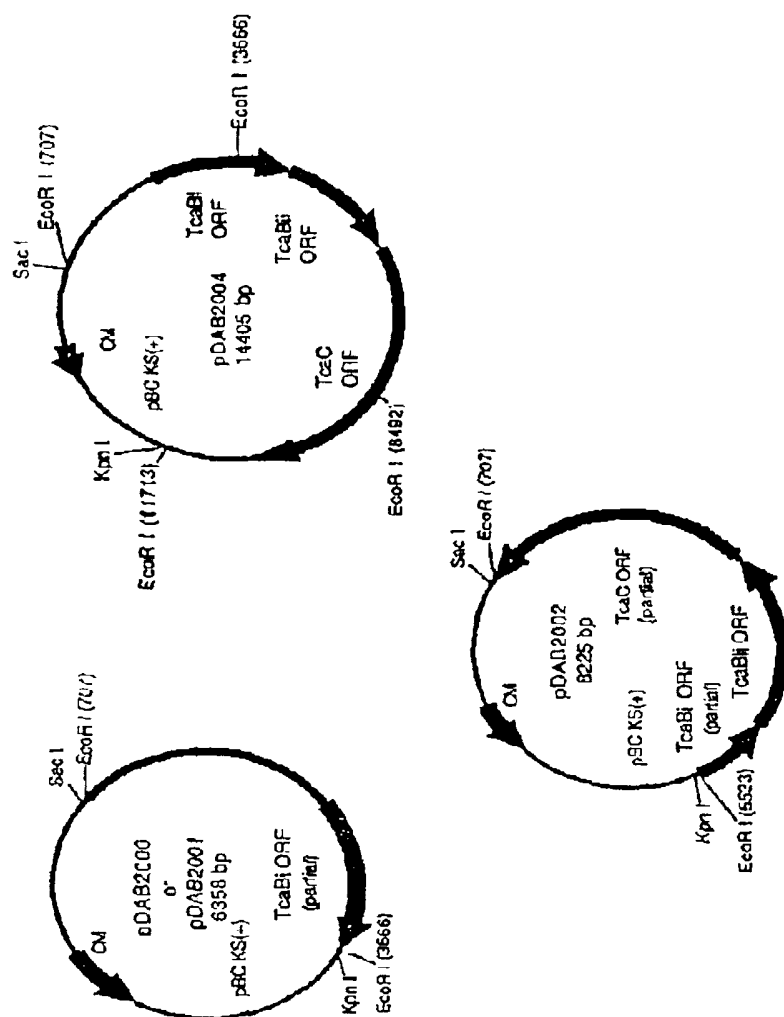


Fig. 2

(51) Int.Cl.

C12N 5/14 (2006.01);

C12N 15/00 (2006.01);

A01G 13/00 (2006.01);

A01H 1/00 (2006.01)

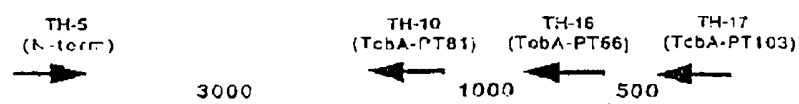


Fig. 3

(51) Int.Cl.

C12N 5/14 (2006.01);

C12N 15/00 (2006.01);

A01G 13/00 (2006.01);

A01H 1/00 (2006.01)

TcbA	1740	1750	1760	1770	1780
	SSAQALKNDS	ERMDFSCANA	LYFWELFYIT	PMMAHRLLO	EQNFDAANHW
TcaBi	450	460	470	480	
	SS	hPvDFSGpyy	LYWFIFFli	Pfllvtrng	EQryedadtW
TcbA	1790	1800	1810	1820	1830
	FRYVWSPSGY	IVDGKLAIVH	WVVRPLEEDT	SWNAQQLDST	DPDAVAQDDP
TcaBi	490	510	520	530	
	rdangql	ykyifrsagY	MTGskprry	WNVmPLqldT	awdttQpatT
TcbA	1840	1850	1860	1870	1880
	MHYKVAIPMA	TLDLLIARGD	AAYRQLERDT	LAEAKMYTQ	ALNLLGDEPO
TcaBi	540	550	560	570	580
	MHYKVAIPMA	TLDLLIARGD	sAYRQLERDT	LVEAKMYTQ	AqqlLQprPd
TcbA	1890	1900	1910	1920	1930
	VMLSTIWRNP	TLGNAASKIT	QQVRQQVLTQ	LRINSRVKTP	LLGTANSLTA
TcaBi	600				
	ihntnTWpNP	TLsk			
TcbA	1940	1950	1960	1970	1980
	LFLPQENSKL	KGYWRTLAQR	MFNLRRHNLST	DGQPLSLPLY	AKPADPKALL
TcaBi	20	30	40	50	60
	FLPpyndvL	LGYNckLelR	lynLRRHNLST	DGQPLSLPLY	AtPvDFKtLq
TcbA	1990	2000	2010	2020	2030
	SAAVSASQGG	ADLPKAPLTI	HREPCMLEGA	AGLWQLIQF	GSSLIGYSER
TcaBi	70	80	90	100	110
	rqggggdgtG	sapaggggsv	GRyPllvERa	RsaVslLtQF	GnSLqttilEn
TcbA	2040	2050	2060	2070	2080
	QDAEAMSQLL	QTQASELILT	SIRMQDNQLA	ELLESNTALC	VSLAGVQQR
TcaBi	120	130	140	150	160
	QDnEkMtLL	QTQqeaillK	ghdiQgWnLk	gLqnsITALQ	asrdGdtlRq

Fig. 4A

(51) Int.Cl.

C12N 5/14 (2006.01);

C12N 15/00 (2006.01);

A01G 13/00 (2006.01);

A01H 1/00 (2006.01)

TcBa	2090	2100	2110	2120	2130
	DSYSQLYEEN	INDAGEQRALA	LRSESAIESQ	GAQISRMAA	GVDMAFNIFG
TcABii	170	180	190	200	210
	khYSdLiagg	laAnEiaGLt	LRStamI-Tn	Gvarglli-ag	GinauPNVFG>
TcBa	2140	2150	2160	2170	2180
	LADGGMHYGA	IAYAIADGIE	LSASAKMVDA	EKVAQSEERY	RRRQENKICR
TcABii	220	230	240	250	260
	LAnGGsewGA	pligsqcatq	vgAgicdqsA	gisevtagYq	RRqeEWalQR>
TcBa	2190	2200	2210	2220	2230
	DNAQAEINQL	NAQLESLSIR	FEAAEMQKEY	LXTQQAQAQA	QLTFILSKFS
TcABii	270	280	290	300	310
	DIAdnEitQL	daQiqSLqeq	itwAqXqitl	seTeQAnAQA	iydlqttrFt>
TcBa	2240	2250	2260	2270	2280
	NQALYSWLRG	RLSGTYFQFY	DLAVSRCLMA	EQSYQWEAND	NSISFVKPGA
TcABii	320	330	340	350	360
	gQALYnWnaG	RLSalyYqMY	DstlpicLqD	kaalvqsgak	esdSifqvpu>
TcBa	2290	2300	2310	2320	2330
	WQGTYAGILC	GEALIONLAQ	MEEAYLKWES	RALEVERTVS	LAVVYDSLEG
TcABii	370	380	390	400	410
	WndlwqGtLa	GEoLsseLqk	ldaiwLargg	igLEaIRTYS	Ldtlfgc--G>
TcBa	2340	2350	2360	2370	2380
	NDRFNLAEQI	PALLDKQEGT	ACTKQNGLSL	ANAILSASVK	LSDLKLGTDY
TcABii	420	430	440	450	
	----tLsEnI	nkun-GEtV	spsgovtLaL	tpdIfqatld	LSqlqIdnsY>
TcBa	2390	2400	2410	2420	2430
	PDSTVGSNKV	RRIKQISVSL	FALVGFYQDV	QAMLSYGGST	QLPKGCSALA
TcABii	460	470	480	490	500
	-n--lGneKk	RRIKrIaVtL	PtLLGFYODI	eAtLvaGaee	aLshGvndqe>
TcBa	2440	2450	2460	2470	
	VSHGTNDSGQ	FQLDFNDGKY	LPFZGIALED	QGTNLNLFVN	
TcABii	510	520	530		
	rsvtdfinder	F-LpF-cOrd	accgtlaLn>		

Fig. 4B

(51) Int.Cl.

C12N 5/14 (2006.01);
C12N 15/00 (2006.01);
A01G 13/00 (2006.01);
A01H 1/00 (2006.01)

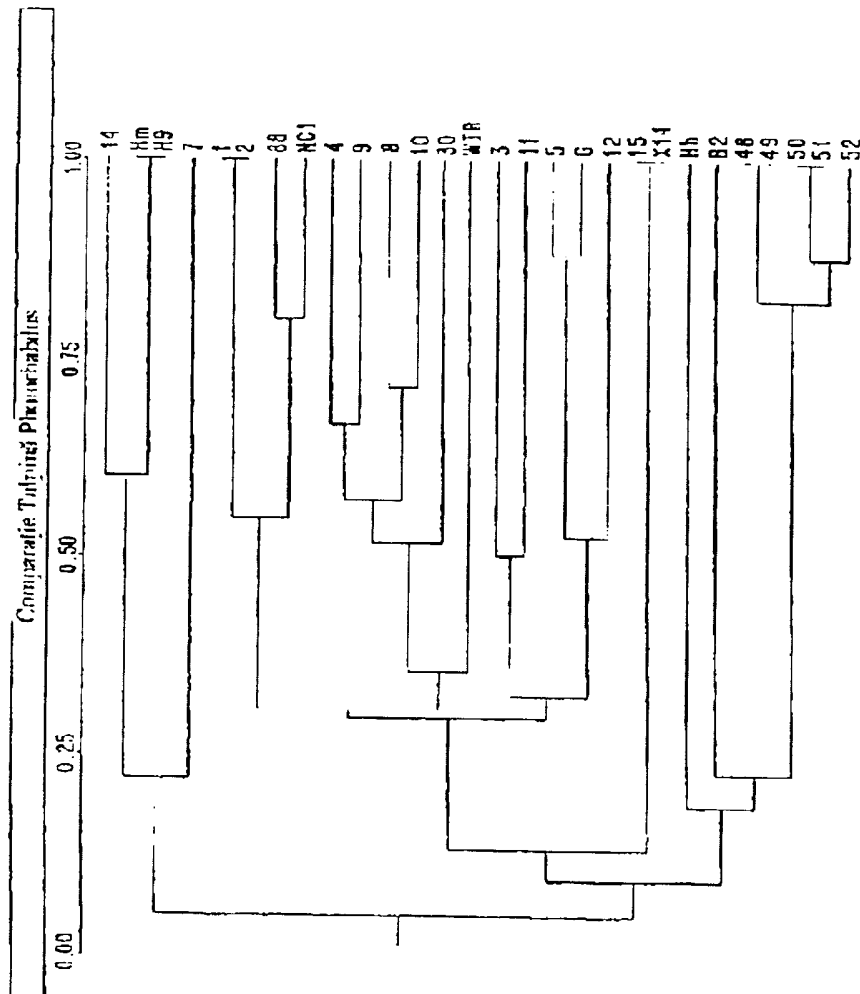


Fig. 5

(51) Int.Cl.

C12N 5/14 (2006.01);

C12N 15/00 (2006.01);

A01G 13/00 (2006.01);

A01H 1/00 (2006.01)

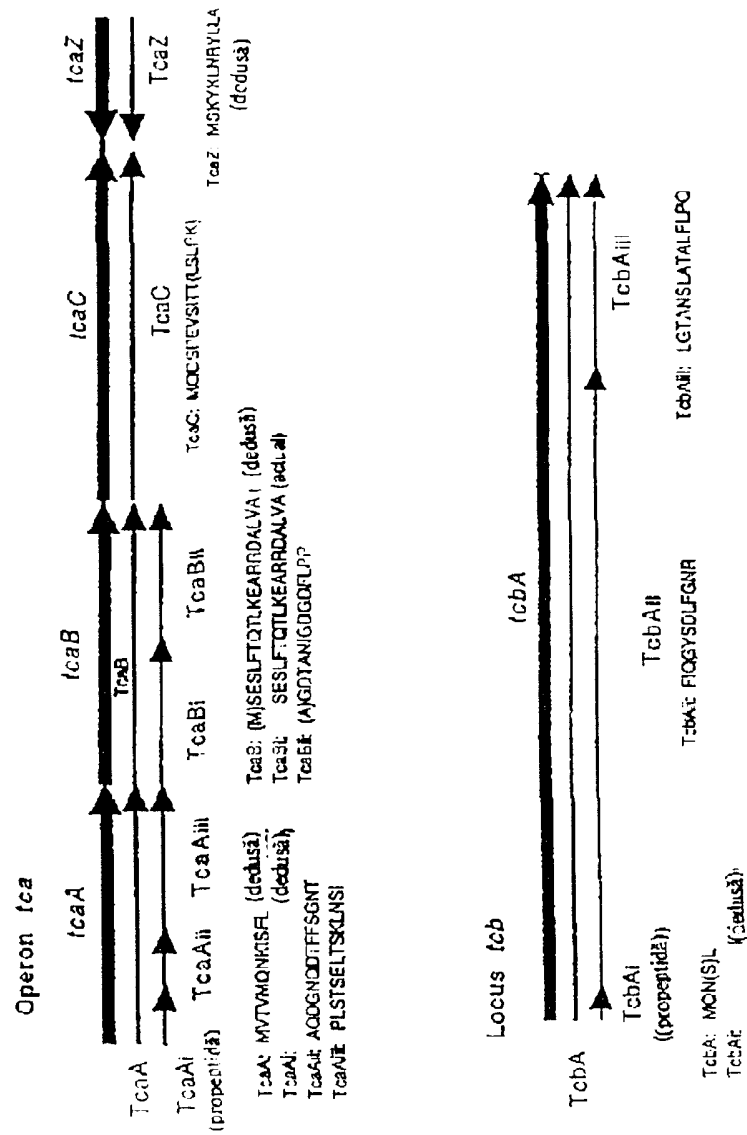


Fig. 6A

(51) Int.Cl.

C12N 5/14 (2006.01);
C12N 15/00 (2006.01);
A01G 13/00 (2006.01);
A01H 1/00 (2006.01)

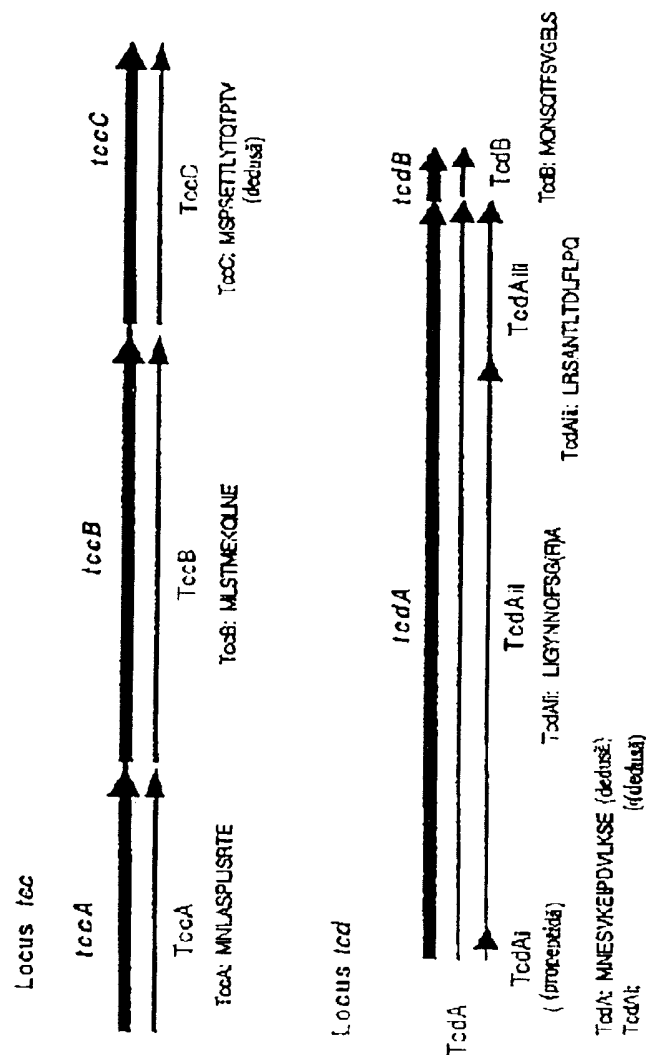


Fig. 6B

