cDNA Cloning, Heterogeneous Expression and Biochemical Characterization of a Novel Trypsin-Like Protease from *Nilaparvata lugens*

Zhifan Yang[a,*], Xunli Xia[b], Xiaoxue Wang[a], and Guangyuan He[b,*]

[a] College of Life Sciences, Hubei University, Wuhan 430062, China.
E-mail: sailyangzhf@yahoo.com.cn
[b] School of Life Science and Technology, Huazhong University of Science and Technology, Wuhan 430074, China. E-mail: hegy@hust.edu.cn

* Authors for correspondence and reprint requests

Z. Naturforsch. **65c**, 109–118 (2010); received August 17/September 24, 2009

A reverse transcription-polymerase chain reaction (RT-PCR) strategy was used to clone diverse trypsin-like protease gene transcripts from midguts of the brown planthopper *Nilaparvata lugens* Stål (Homoptera: Delphacidae). Six individual trypsin-like protease transcripts were identified. On the basis of one nucleotide sequence of the six clones, a full-length cDNA sequence (1902 bp) was obtained by rapid amplification of cDNA ends (RACE). The cDNA contained an 1128-bp open reading frame encoding a putative protein of 375 amino acids with typical features of the trypsin-like protease. Heterogeneous expression of the coding sequence for the mature peptide in *Escherichia coli* cells showed that the expressed protease with a molecular weight of 27.0 is active, for its BApNAse activity assayed by using BApNA (*N*-benzoyl-D,L-arginine-*p*-nitroanilide) as substrate. The protease had its maximum activity at pH 8.0 and 35 °C. A much better stability was observed at pH values above 4.0 and temperatures below 40 °C. The enzyme was strongly inhibited by serine protease inhibitor. The trypsin-like protease is therefore likely one of the major digestive proteases responsible for protein hydrolysis in *N. lugens* gut, and multiple gene families encoding digestive proteases may help in adaptation of this sap-sucker to different rice varieties.

Key words: Trypsin-Like Protease, Heterogeneous Expression, BApNA (*N*-Benzoyl-D,L-arginine-*p*-nitroanilide)

Introduction

The brown planthopper *Nilaparvata lugens* (Homoptera: Delphacidae) is a serious insect pest of rice (Rubia-Sanchez *et al.*, 1999). To replace application of chemical protectants during the management of this pest, more and more environment-friendly technologies such as planting-resistant rice varieties were emphasized (Cuong *et al.*, 1997). However, this specialist sap-sucker has evolved mechanisms to adapt to various rice varieties by regulating its detoxification and toxin-tolerance enzyme genes (Yang *et al.*, 2005, 2006, 2007), including a trypsin-like protease gene always reported (Foissac *et al.*, 2002). Therefore, the proteolysis in the digestive system may contribute significantly to the adaptation of this pest to various rice varieties, especially resistant rice.

Trypsin is an important digestive proteinase, and it is highly specific towards the positively charged side chains of lysine and arginine (Brown and Wold, 1973). Foods ingested by chewing pests are mostly macromolecules broken down by digestive enzymes, such as trypsin-like proteases (Terra and Ferreira, 1994; Chapman, 1998). Sap-suckers such as *N. lugens* of the order Homoptera have been widely considered to lack proteolysis in the gut (Srivastava and Auclair, 1963; Terra *et al.*, 1996). However, Foissac *et al.* (2002) revealed that the trypsin-like protease activity is a major factor in gut protease activity in *N. lugens*, and a trypsin-like protease cDNA was isolated by single sequencing runs from the 5′ end of 100 randomly selected cDNA clones from the gut cDNA library. The cDNA-predicted protein is more similar to other vertebrate trypsins than to invertebrate serine proteases. The protein has the normal catalytic triad of residues (Asp, His, Ser) wich determine specificity of cleavage. But an abnormal feature of this protein is a C-terminal region which shows no similarity to other serine proteases. A further unusual feature of this *N. lugens* serine protease

is that the C-terminal residue of the propeptide is an asparagine residue; thus this trypsin-like protease is likely incapable of self-activation and is not cleaved by trypsin-like enzymes. The sequence possibly represents a small gene family of serine proteases (Foissac *et al*., 2002). Additionally, limited to the method applied in the present study, more cDNAs encoding trypsin-like proteases might be missed. Therefore, more information is necessary to reveal the digestive proteases in *N. lugens* and to detail the characteristics of the protease genes within the insect gut.

In this study, we analyzed the diversity of trypsin-like proteases in *N. lugens* by obtaining various cDNA sequences and cloning of a novel trypsin-like protease cDNA, and biochemical characterization of the heterogeneously expressed trypsin-like protease. The implications of our results for controlling this pest are discussed.

Material and Methods

Experimental insects

A colony of brown planthoppers (*N. lugens*) was mass-reared on plants of susceptible rice, Taichung Native 1 (TN1, an Indica rice variety maintained in our laboratory), at (25 ± 2) °C, 80% relative humidity, under a photoperiod of 16 h light and 8 h dark. The fourth instar larvae of *N. lugens* were collected for dissecting the gut tissues in 0.9% (w/v) NaCl and RNA extraction.

Cloning of trypsin-like proteinase cDNA fragments

Total RNA was extracted from the gut tissues according to instructions of the manufacturer of Trizol reagent (Invitrogen, Carlsbad, CA, USA). Trypsin-specific primers (DmTF/SerPR, Table I) were used as described elsewhere (Mazumdar-Leighton and Broadway, 2001). Reverse transcription polymerase chain reaction (RT-PCR) was carried out on gut total RNA using an Access RT-PCR kit (Promega, Madison, USA) in accordance with the manufacturer's instructions. PCR products were checked by electrophoresis on 0.8% (w/v) agarose gel in TAE buffer (40 mmol/L Tris-acetate, 2 mmol/L $Na_2EDTA \cdot 2H_2O$), and the resulting DNA band of about 470 bp was recovered and cloned into the pGEM-T vector (Promega). Seventy-four randomly selected inserts were sequenced at both ends. The sequences were compared with all databases in GenBank using the BLAST server command "blastx".

The 5' and 3' rapid amplification of cDNA ends (RACE)

A cDNA fragment of 483 bp in length was chosen as template to design sense and anti-sense gene-specific primers (GSP-F/GSP-R, Table I) for 3' and 5' rapid amplification of cDNA ends (RACE). This clone represents a secondary abundant mRNA species in the six clones. The 3' and 5' RACE reactions were performed according to the instruction manual of the SMART RACE cDNA amplification kit (BD Bioscience Clontech, Palo Alto, CA, USA). Amplified fragments were cloned and sequenced. More than 3 independent clones of 3' and 5' end of cDNAs were sequenced to eliminate possible PCR mutations.

Computer-assisted analyses of the cDNA-predicted protein sequence

Full length nucleotide sequence analysis was carried out by BLAST similarity searches (http://www.ncbi.nlm.nih.gov/BLAST/). The molecular mass of the putative protein was predicted by Compute pI/Mw tool (http://us.expasy.org/tools/pi_tool.html). The signal peptide was predicted by SignalP 3.0 Server (http://www.cbs.dtu.dk/

Table I. Primers used in this study.

Primer	Sequence (5'-3')
DmTF	TCGAATTCATTGTGACCGCCGCTCAYT
DmTR	GGTAGATCTCACGGCTGGACAYT
GSP-F	CACCTGCGCTTCGTACAGTTGTGACGGG
GSP-R	CCAATCTGTTTGCCAGCATCTACCG
mTRY-F	GGTGAATCATGGGTGCGTCTGTGTGGTACTCTGT
mTRY-R	GCTAAGCTTGTATTTATTAGGTATATACTGCTCAGCAA

Y = C/T, R = A/G.

services/SignalP). Clustal X v1.8 (Thompson *et al.*, 1997) was used for analyzing the alignment and edition.

Protein expression in E. coli

A PET system (Novagen, Madison, WI, USA) was used to express the trypsin gene. The coding region of the mature enzyme was amplified by PCR with the primers mTRY-F and mTRY-R that contain the *Hin*dIII and *Eco*RI site (Table I). The PCR product (735 bp) was digested and inserted into the expression vector pET-23a, which inserts 6 histidine residues to the amino terminus of a target protein, to construct pET-735. The constructed vector was used to transform *Escherichia coil* BL21 (DE3) pLysS for overexpression. The recombinant cells at the optical density (OD) 0.8 induced by adding isopropylthio-β-D-galactoside (IPTG) (0.1 mmol/L) were performed at 25 °C for 8 h, and the cells from 100 mL culture were collected by centrifugation at 5,500 × *g* at 4 °C for 5 min. The cells were resuspended and rinsed for three times in 50 mM sodium phosphate buffer (pH 8.0) containing 300 mM NaCl, 15 mM imidazole, 15 mM 2-mercaptoethanol, 0.1 mM EDTA, 0.1% (v/v) Triton X-100, and 10% (v/v) glycerol, then stored at –80 °C for 30 min and thawed at room temperature for 10 min. The operation was performed three times and the cells were lyzed on ice by ultrasonic wave treatment (30 W) for 3 min at 30-s intervals. The lysate was centrifuged at 11,000 × *g* at 4 °C for 15 min, and the supernatant was recovered for purification by using the high-affinity Ni-NTA resin (Qiagen, Hamburg, Germany) according to the manufacturer's instructions. Elution of protein was monitored by measuring the absorption at 280 nm in each fraction; the fractions containing protein were further analyzed by SDS-PAGE following a standard protocol (Sambrook *et al.*, 1989). Fractions displaying massive target protein bands on the SDS-polyacrylamide gel were pooled. The protein concentration was monitored by the Bio-Rad protein assay (Bradford, 1976), using bovine serum albumin as a standard.

The trypsin-like activity was assayed using 40 μL of the 1:10 diluted enzyme sample plus 160 μL BApNA (*N*-benzoyl-D,L-arginine-*p*-nitroanilide; Sigma-Aldrich, St. Louis, MO, USA) [1.0 mM in 50 mM Tris-HCl (2-amino-2-hydroxymethyl-1,3-propanediol hydrochloride; Sigma-Aldrich), pH 8.0, containing 20 mM $CaCl_2$] at 25 °C; the absorbance at 405 nm was continuously read using a MicroElisa auto-reader (Dynatech, Alexandria, VA, USA). An absorbance curve was made, the initial rate was calculated, and a common unit was adopted. One enzyme unit was defined as the quantity of enzyme necessary to degrade 1.0 μM BApNA per min at 25 °C and pH 8.0. The specific activity was defined as the number of units per mg protein. In all subsequent assays, 1.0 unit of the purified enzyme was used.

Activity and stability at different pH values

The pH activity was assayed in a MicroElisa plate by adding 0.5 μL H_2O, 4.5 μL of 10 × buffer solution (McIlvaine's buffer for pH 5.0, 6.0, 7.0, 8.0 and 0.2 M glycine/NaOH buffer for pH 9.0, 10.0, 11.0) to 40 μL of the 1:4 diluted enzyme sample. After 1 h incubation at 37 °C the samples were placed on ice. Then 160 μL of 1.0 mM BApNA in 50 mM Tris-HCl, pH 8.0 (pre-warmed to 25 °C), were added to each reaction, sample well and kept at 25 °C for 20 min. The reaction was stopped and the absorbance at 405 nm was measured. The stability at different pH values was determined in the same way except that buffer solutions with pH values ranging from 2.0 to 12.0 were used.

Activity and stability at different temperatures

In order to evaluate the optimum temperature for trypsin-like protease activity, nine portions of 40 μL each of 1:4 diluted ice-chilled enzyme samples were incubated at eight temperatures (5, 10, 15, 20, 25, 30, 35, and 40 °C) for 30 min. Then 160 μL of 1.0 mM BApNA in 50 mM Tris-HCl, pH 8.0 were added to each sample, mixed well and kept at 25 °C for 20 min. The reaction was stopped and the absorbance at 405 nm was measured. The stability at different temperatures was evaluated in the same way except that different temperatures (30, 35, 40, 45, 50, 60, 70, and 80 °C) were used. Controls were performed with samples maintained at 0 °C for 30 min before reaction.

Inhibitor studies

In order to confirm the nature of the expressed enzyme, 50 μL of 1 × McIlvaine's buffer, pH 8.0, containing 40 μL of 1:4 diluted enzyme samples were separately added to 10 μL of 0.001, 0.005,

0.01, 0.1, 1.0, 10.0, and 100.0 μM serine protease inhibitors (contained in 10 × Protease Inhibitor Cocktail Set I solution; Calbiochem, La Jolla, CA, USA). After a 30-min preincubation at 25 °C, 160 μL BApNA (1.0 mM in 50 mM Tris-HCl, pH 8.0) were added to the reaction solution and kept at 25 °C for 20 min. The reaction was stopped and the absorbance at 405 nm was measured. The inhibitory effect was measured as the percentage of activity remaining with respect to the control. Control consisted of the same amount of enzyme incubated in McIlvaine's buffer.

Results

Cloning and sequencing of trypsin-specific RT-PCR products from N. lugens

An about 470-bp DNA band was generated by RT-PCR. The DNA fragments were cloned and sequenced. Homologous analysis revealed that 68 out of the 74 inserts encoded a Part of trypsin-

```
          DmTF
TRY456    IVTAAHCA-VAGTANRFSIRAGLNT---LNAPESGFQESNVARVIVHPQYSDRTLQNDIA 56
TRY462    IVTAAHCA-VLGGASAYSIRAGVHN---IRFPEAESQEVDVALNIVPEFYSSGILLNDIA 56
TRY447    IVTAAHCVNDVLIASNIHVRAGEHY---LDGIDGHEQEHNAQKITVHPNYDASVFVNDIA 57
TRY453    VVTAAHCAIHITNYTDYYVALGSNK---LTN-SKALKKFAISKVTYHNGFSYSTLSNDIA 56
TRY474    IVTAAHCVQHMTKQDVMNLELRLGDHSLYQTGKVAHESRKASRVLYHKGFSMNHLRNDIA 60
TRY483    IVTAAHCIVREKLTIVRLGELDWN----TTDDNANHVDMPIEKAFPHPRYNPVKRATDVG 56
Drome     ILTAAHCVARMTSWDVAALTAHLGDYNIGTDFEVQHVSRRIKRLVRHKGFEFSTLHNDVA 338
          ::*****                .   .          :.     .*:.

TRY456    IMTLSTPLRLGADTRTEAARL-PSNRYDAKGD--ATIT-GWGTTSESG-SLS-QQLRTVT 110
TRY462    LMKLATPLKLGAAERTAAIKIAPSGTPIPSGN--ALVVSGWG-TVESG-GLSPAKLRAVD 112
TRY447    IIQIQGEFTL--NEWVNTIAL-PSDQEESNGT--CVVS-GWGSTFEGG-IIA-RNLHAVN 109
TRY453    IIKLKKPIRFNKNIKPKKIAT---RVPKQDTK--CIIS-GWGTWNYGD-HVIHDELKAAT 109
TRY474    LIHLSKPVEFKSNIQPVCLHTGSPGFEDGNSN---ADVTGWGALSEGG--SQPADLRTVR 115
TRY483    IIRLREPVRFSADIQPICLPA-STELRNKNLENISPYITGWGSFSYKSNLSYPSQLYEAQ 115
Drome     ILTLSEPVPFTREIQPICLPT-SPSQQSRSYSGQVATVAGWGSLRENG--PQPSILQKVD 395
          :: :   . :                .          ***     .       *  .

                                                  SerPR
TRY456    VPIITDAACRNAYG----QSDVADSMICAGLNAGGKDSCQGDSGGP 152
TRY462    VPAVTDDEGREAYG----ASAIYESMICAGRKEGGKDSCQGDSGGP 154
TRY447    VPVVDDVTCQGNYA----GNPIAPSMLCAG--VAGRDACQGDSGGP 149
TRY453    VLISNMTQCRANYS----DRVDPLTMICAGLVEGGVDSCQGDSGGP 151
TRY474    VRVITKSECLQKYA--SSTN-KPTDGMICAVSDNGEDSCQGDSGGP 158
TRY483    VNVKSNRDCAAAYARLGNKAGITIDDSVLCAGGKATDPCQGDSGGP 161
Drome     IPIWTNAECARKYG--RAAPGGIIESMICA-GQAAKDSCSGDSGGP 438
          :          *.                  . *.*.******
```

Fig. 1. Alignment of the amino acid residues encoded by the six trypsin-like protease cDNA fragments cloned from *N. lugens* larvae by degenerate RT-PCR. The conserved catalytic triads of histidine57, aspartate102, aspartate189, and serine195 are boxed. The amino acid residues used for designing the degenerate primers are flanked by arrows.

Table II. Sixty-eight trypsin cDNA fragments amplified by degenerate RT-PCR.

Clone name	cDNA length [bp]	Clone number	Accession number
tryp453	453	30	AJ629011
tryp483	483	28	AJ629016
tryp456	456	5	AJ629012
tryp447	447	3	AJ629014
tryp474	474	1	AJ629015
tryp462	462	1	AJ629013

like proteases and were classified into 6 groups (Table II). All the six encoded protein segments contained the corresponding conserved catalytic triad of histidine57, aspartate102, aspartate189, and serine195 (residue numbering after bovine chymotrypsin) (Fig. 1). The aspartate189, which is found in the specificity pocket of *Drosophila melanogaster* α-trypsin (accession number: X02989) (Davis *et al*., 1985), determined the six transcripts encoding trypsin-like proteases in *N. lugens*.

Cloning and characterization of trypsin-like protease cDNA from N. lugens

A 483-bp cDNA fragment was used to design gene-specific primers (GSPs) for RACE. The 5' and 3' cDNA ends were aligned to form a full-length cDNA sequence with 1902 bp in size. The nucleotide sequence was deposited in the GenBank with accession number AJ852425.

The cDNA-predicted prozymogen consists of a signal peptide of 18 amino acid residues at the N-terminal end predicted by SignalP 3.0 Server, an activation propeptide of 113 residues and a mature enzyme of 244 amino acids. Typical features of the prozymogen include the propeptide cleavage signal $IVGG^{62-65}$, a lysine residue, allowing trypsin to self-activate and separating the activation peptide from the mature protein (Zhu and Baker, 1999), the catalytic amino acid triad of serine protease active sites (His^{176}, Asp^{225}, Ser^{328}) (Greer, 1990), the three residues (Asp^{322}, Gly^{349}, Gly^{359}) that determine specificity in trypsin-like enzymes, and three pairs of conserved cysteine residues for disulfide bridges (Wang *et al*., 1995) (Fig. 2). The predicted molecular mass and isoelectric point of the mature enzyme are 27014.88 Da and 8.83. The deduced amino acid sequence has the highest homology with trypsin (AB073673-1) of *Bombyx mori* with an amino acid identity of 38%, the next highest level of identity (32%) to *Anopheles gambiae* (AF117749-1), and 30% identity to serine protease of *Ctenocephalides felis* (AF053921-1), when compared to databases in a BLAST search.

Expression and purification of the expressed enzyme in E. coli

The *E. coli* BL21 (DE3) pLysS cells were transformed with the recombinant expression vector pET-735. Expression of the protease was induced by adding 0.1 mM IPTG to the cultured cells. The His-tagged protein was purified using Ni^{2+} affinity columns. When electrophoresed, the purified protein migrated as a massive single band and showed a molecular weight (MW) of 27.0 kDa by measuring the distance migrated by the band and MW standards (Fig. 3).

Biochemical characterization of the expressed trypsin-like protease

The trypsin-proteolytic activity of the enzyme was detectable with BApNA as substrate, and the specific activity was 120 U/mg protein.

A rise of the proteolytic activity was observed at pH 5.0–8.0, then the activity decreased from pH 8.0 to 11.0. The maximum activity of the enzyme was measured at pH 8.0 (Fig. 4A). The stability of this enzyme was recorded at pH values from 4.0 to 12 and the best stability was observed at pH 6.0. However, inactivation of the enzyme occurred in acidic environments, *i.e.* at pH 2.0 (Fig. 4C).

A relatively high activity was observed at 25–35 °C, and the temperature for the highest activity was 30 °C. Furthermore, the enzyme still had 32% of its maximum activity at 5 °C (Fig. 4B). When the temperature increased to 35 °C, the residual activity of the trypsin-like protease was 95% of its maximum activity, and 80% at 40 °C (Fig. 4B). Maximum stability was recorded at 0 °C, *i.e.* the control temperature tested. At 30

```
Nlug    MSLTLIVILLSLATISYCDNIEFTVGNRFRRQAANTCETPSKQQGQCINIMGCKQLYDMLSNP-NRPPAQTSL
Bmor    -----MRILCALLCFILCTELVF------C-QFDEPCQTVENEVGSCLNLLQCKPYLKIVKEF-KTNPAAAVL
Agam    -MSPQAGRFLLLLVCVFCGVIGLS----DALNLQDACETPDGKVGTCVYLRSCLSIRNVLLKKENMTPEDRSL
Cfel    -------MFLLICALLFVKNIVLIN-AQISVIGDSCTRSYDGEAGECALITQCPSANRILQTT---------G
                :  :       : :          .  .: . : * *  :  *     ::

Nlug    LQGSFCGYENEKPRVCCPRQLISAPRPPSQ-------------PQPPSKPNPVNNQQQSQAN-----------
Bmor    LRKAHCGFEGSNPKVCCPLPGFPTVPPDTT-------------TTTTTTTTTTQAPAPTQSAKSAIEDLVPAF
Agam    VMKSKCGQEGRSVLVCCPLVRKLTGR-----------------FDAP-----VELPPPGECG-----------
Cfel    IRPDVCGYSTYEPIVCCVQQRYNSNWNSNREGNKRISEQKCDGYSTAVKQTLTVLPLVSDPNP----------
        :    ** .  .  ***      :                       ..     .      :

                                                          ▼          ○▼
Nlug    -----CGLSTVSINKIVGGRPAILRAWPWMALIGFNSMSR--PQWRCGGALVNTRHVITAAHCIVR----KKL
Bmor    PEPPYCGVSNASFSRVVGGVSAKLGDFPWMVNLGYKPRRGGGVRWLCGGSLITSRHVLTAAHCMHNH--DDDL
Agam    ---------KMQMDRIVGGEVAPIDGYPWLTRIQYYKGSNR-YGFHCGGVLIHNQYVLTAAHCIEGVPSSWIV
Cfel    ISFTVEKCDYNSVPLIVGGEVAKLGEFPHMAAIGWTETSG-AVNWWCGGTLISPEYVLTAAHCASVN--SEQP
                  ..  :***  * :  :* :. : :          : *** *:  .:*:*****

                                                          ○
Nlug    TIVRLGELDWNT-----TDDNANHV-DMPIEKAFPHPRYNPVK--RATDVGIIRLREPVRFSADIQPICLPAS
Bmor    YLVRVGELDLAR-----DDEGATPV-DVLIKTKIKHEQYDAAS--YTNDIGILVLEKDVPITDLIKPICIPKD
Agam    YQVRLGEFDTTTTIDCVEDDCADPVRDVPINAYVVHPDYYKQNGADYNDIALLQLSETVEFTDFIRPICLPTS
Cfel    DIVRLGEHNLKH-----SDDGADPI-DVPVDSVITHPSYHYPS--KYNDIALVKLRYPVSLSNSIRPSCLWAN
          **:** :         *: *  : *: :.  . *  *   .    .*:.:: *   * ::  *:* *:  .

                                                          ▼                   ▼
Nlug    TELRNKNLENISPYITGWGSFSYKSNLSYPSQLYEAQVNVKSNRDCAAAYARLGNKAGIT--IDDSVLCAGG-
Bmor    TELRSRSFEDYNPIIAGWGDTEFRG--PSATHLQVLQLPVVSNDFCAQAYSPYKNQK-----IDERVLCAGYK
Agam    EESRTVNLTGKYATVAGWGQTENST---SSTKKLHLRVPVVDNEVCADAFSSIRLEIIP------TQLCAGGE
Cfel    DEFDT----DS-SIATGWGKIDYAE--SRSDDLLKVVLKIIDNRQCAPLYVDQINRRRLRNGIVDTQMCAGEL
         *  .    .  .  :***. .       . .     : : .*  **  :     .           :***

           ●▼  ○                                ●   ▼     ●
Nlug    -EATDSCQGDSGGPLMIPIKQN------FYLFGVVSYG-HKCAEPGFPGVYTRVTEFVDWIQSNI----
Bmor    KGGKDACQGDSGGPLMQPIWNSQTYKTYFYQIGVVSFG-KKCAEAGFPGVYSRVTHFVPWIQEKVVGHA
Agam    -KGKDSCRGDSGGPLMRYGDGRSSTKS-WYLIGLVSFGLEQCGTDGVPGVYTRMSEYMDWVLDTME---
Cfel    DGGKDTCQGDSGGPLQITXQSN---KCIFYIVGITXFG-RGCGAPNSPGVYTRVSKYVDWIESVVWXN-
            ..*:*:*******              :* .*:.  :* .  *.   . ****:*::.:: *: .  :
```

Fig. 2. Alignment of *N. lugens* (Nlug), *B. mori* (Bmor), *A. gambiae* (Agam), and *C. felis* (Cfel) trypsin or serine protease sequences. Identical amino acids are indicated by asterisks and conservative substitutions by dots. The catalytic amino acid triads of serine proteinase active sites are marked with an open cycle (○), the three residues that determine specificity in trypsin-like enzymes are marked with a dot (●), and three pairs of conserved cysteine residues for disulfide bridges are marked with an inverted triangle (▼). The propeptide cleavage signal, IVGG, is underlined.

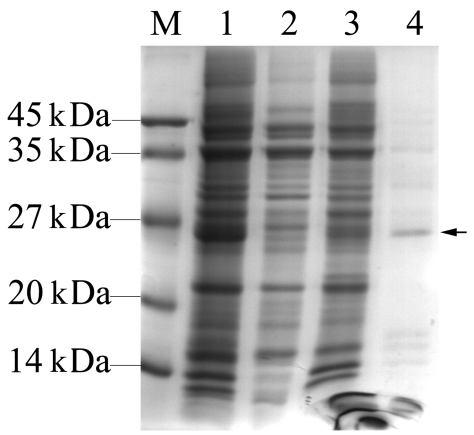

Fig. 3. SDS-PAGE analysis of the proteins produced in transformed *E. coli* and of the proteins purified by Ni^{2+} affinity spin columns. Lane M, low-molecular weight standards (Tiangen Laboratories, Beijing, China); lane 1, crude extract; lane 2, flowthrough; lane 3, wash; lane 4, eluate. The samples were loaded onto 10% polyacrylamide gel. The enzyme produced in *E. coli* is shown at 27.0 kDa (indicated by an arrow), which corresponds to the deduced molecular weight for the mature trypsin-like protease.

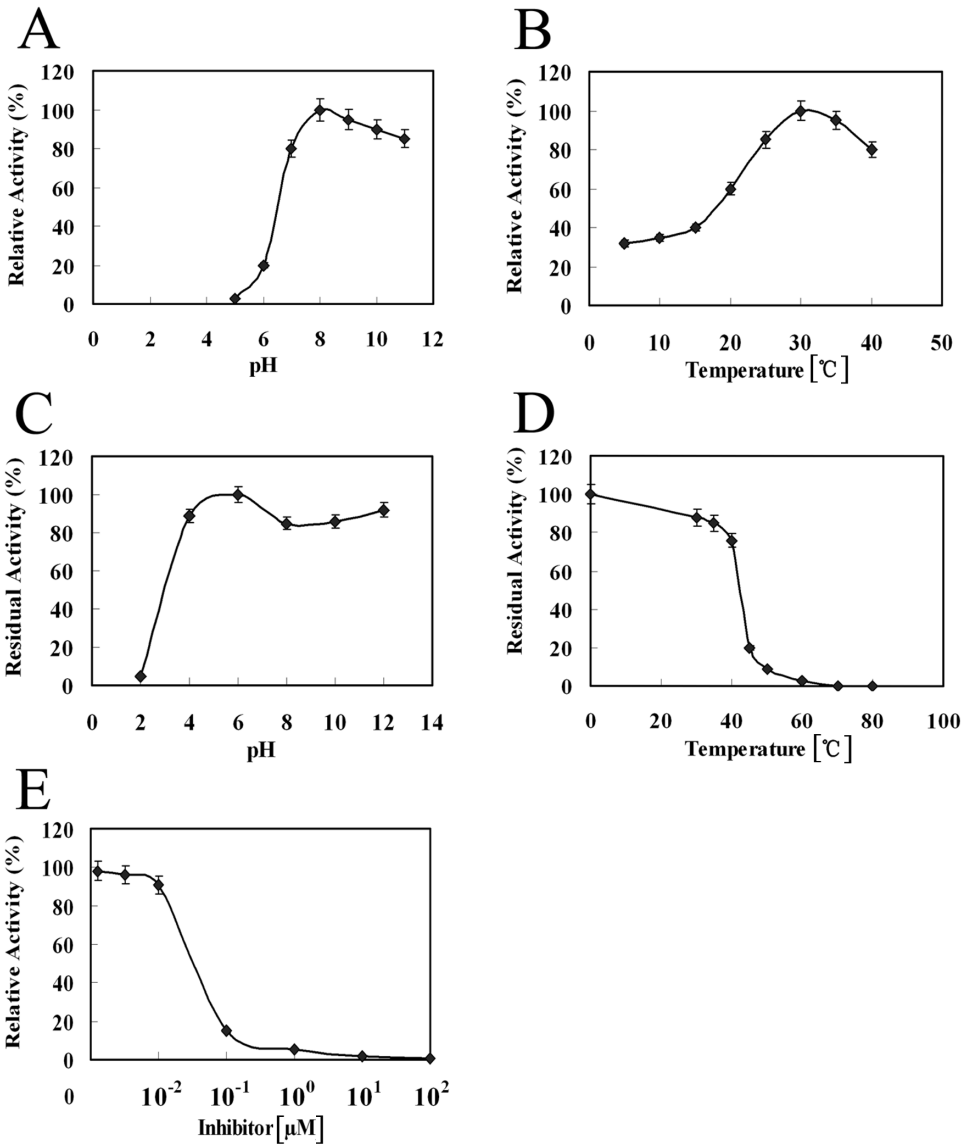

Fig. 4. Biochemical characterization of the expressed trypsin-like protease. (A) Activity of the trypsin-like protease at different pH values. Enzyme solutions were adjusted to the desired pH, incubated for 1 h at 37 °C, cooled on ice, and subsequently the optimal pH for the enzyme activity was determined at 25 °C for 20 min in several buffer systems from pH 5.0 to 11.0. (B) Activity of the trypsin-like protease at different temperatures. Enzyme solutions were incubated for 30 min at temperatures ranging from 5 to 40 °C, cooled and assayed with BApNA. Controls were prepared with samples maintained at 0 °C for 30 min before reaction. (C) Effect of pH value on the stability of the trypsin-like protease. pH stability was determined with a discontinuous buffer system (pH 2.0 – 12.0). (D) Stability of the trypsin-like protease at different temperatures. The temperature ranged from 30 to 80 °C. (E) Effect of the serine protease inhibitor on the activity of the trypsin-like protease. Reactions were performed in the presence of 1.0 unit of enzyme sample, 160 μL of 1.0 mM BApNA and 0.001 – 100 μM inhibitor. All values are means of three determinations ± SD.

or 35 °C, the enzyme was substantially stable, whereas the stability decreased sharply when the temperature increased above 40 °C. The enzyme was completely inactivated when the temperature was above 60 °C (Fig. 4D).

The slope of the inhibitory curve slightly declined when the inhibitor concentration increased from 0.001 to 0.01 μM. But the BApNAse activity significantly decreased when the inhibitor concentration increased from 0.01 to 0.1 μM. The enzyme activity was completely repressed in the presence of 10 μM inhibitor (Fig. 4E). The result, combined with the data concerning the molecular mass and substrate specificity (strong BApNAse activity), indicates that the enzyme can be grouped in the same class like bovine trypsin (EC 3.4.21.4).

Discussion

Previous studies revealed that a single digestive trypsin-like protease is present in some insects like *Musca domestica* (Lemos and Terra, 1992), *Choristoneura fumiferana* (Milne and Kaplan, 1993), *Ostrinia nubilalis* (Bernardi *et al*., 1996), *Periplaneta americana* (Lopes and Terra, 2003), *Pediculus humanus* (Kollien *et al*. 2004) and *Zabrotes subfasciatus* (Magalhães *et al*., 2007). But more studies showed that trypsin-like proteases frequently are present as multi-gene families (Gatehouse *et al*., 1997; Zhu *et al*., 2000, 2003, 2005; Mazumdar-Leighton and Broadway, 2001; Patankar *et al*., 2001; Oliveira *et al*., 2005; Vinokurov *et al*., 2006; Chougule *et al*., 2008; George *et al*., 2008). The gut of *Helicoverpa armigera* alone is known to contain about twenty different types of active serine protease isoforms at any developmental stage (Gatehouse *et al*., 1997; Patankar *et al*., 2001). The trypsin-like protease diversity in insects can be correlated with the adaptation to various proteins in plant diets as well as to exposure to plant-derived protease inhibitors (PIs) and lectins (Srinivasan *et al*., 2006). In the midguts of well-studied insects, such as the Lepidopteran larvae, serine proteases are known to dominate the insect larval gut environment and contribute to about 95% of the total digestive activity, and only a few of them in the trypsin-like protease family contribute significantly to digestion (Srinivasan *et al*., 2006). However, a broader array of protease isoforms could be advantageous for the insect in dealing with the diverse plant proteins, which may be harmful, given a narrow choice of proteases (Srinivasan *et al*., 2006). In the present study, we successfully isolated six individual trypsin-like protease cDNA fragments from *N. lugens* (Fig. 1 and Table II). The data demonstrated that multiple gene families encoding digestive proteases exist in *N. lugens* gut.

Trypsins possess a catalytic triad that characterizes all serine proteases, consisting of His, Asp, and Ser amino acid residues, and typically shows two characteristics that distinguish them from other serine proteases: (1) specificity for the peptide bond formed by the carboxylic side of Arg or Lys residues and (2) ability to activate other pancreatic zymogens (de Albuquerque *et al.,* 2001). The full cDNA obtained in the present study encodes a novel protein with all characteristics of trypsin-like protease as described above. To confirm that the protease is active, the coding region of the mature trypsin was heterogeneously expressed in *E. coli* cells. The expressed enzyme was biochemically characterized (Fig. 4). The protein had the trypsin-like activity for its ability of hydrolysis of BApNA. It was strongly inhibited by the serine protease inhibitor. The results combining the characteristics of the predicted amino acid sequence strongly suggest that the trypsin-like protease in *N. lugens* gut is active and possibly responsible for digestion of at least part of the proteins in rice sap.

Transgenic rice plants constitutively expressing the soya bean Kunitz trypsin inhibitor (SKTI) were found to show resistance to *N. lugens* (Lee *et al*., 1999). The mortality of *N. lugens* is increased during insect development. And *N. lugens* infestation could induce the protease inhibitor gene in rice plants (Zhang *et al*., 2004). These data readily explain that *N. lugens* contains digestive protease activity. In this work, the multi-genes of trypsin-like proteases were confirmed in *N. lugens*. Their expression profiles in response to different families of PIs in cultivating rice and in transgenic rice plants will be investigated in our recent work. A comprehensive analysis of the trypsin-like protease/inhibitor interaction could help in the design of optimized insecticides and improving of the control of this rice pest.

Acknowledgement

This work was supported by grants from the National Natural Science Foundation of China (30500328).

Bernardi R., Tedeschi G., Ronchi S., and Palmieri S. (1996), Isolation and some molecular properties of a trypsin-like enzyme from larvae of European corn borer *Ostrinia nubilalis* Hübner (Lepidoptera: Pyralidae). Insect Biochem. Mol. Biol. **26**, 883–889.

Bradford M. M. (1976), A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72**, 248–254.

Brown W. E. and Wold F. (1973), Alkyl isocyanates as active-site-specific reagents for serine proteases. Reaction properties. Biochemistry **12**, 828–834.

Chapman R. F. (1998), The Insects: Structure and Function. Cambridge University Press, New York, p. 770.

Chougule N. P., Doyle E., Fitches E., and Gatehouse J. A. (2008), Biochemical characterization of midgut digestive proteases from *Mamestra brassicae* (cabbage moth; Lepidoptera: Noctuidae) and effect of soybean Kunitz inhibitor (SKTI) in feeding assays. J. Insect Physiol. **54**, 563–572.

Cuong N. L., Ben P. T., Phuong L. T., Chau L. M., and Cohen M. B. (1997), Effect of host plant resistance and insecticide on brown planthopper *Nilaparvata lugens* Stål and predator population development in the Mekong Delta, Vietnam. Crop Prot. **16**, 707–715.

Davis C. A., Riddell D. C., Higgins M. J., Holden J. J., and White B. N. (1985), A gene family in *Drosophila melanogaster* coding for trypsin-like enzymes. Nucleic Acids Res. **13**, 6605–6619.

de Albuquerque C., Muhlia-Almazán A., Hernández-Cortes P., and Garcia-Carreño F. L. (2001), Proteinases from marine organisms. In: Recent Advances in Marine Biotechnology (Fingerman M. and Nagabhushanam R., eds.). Science Publishers, Plymouth, pp. 209–238.

Foissac X., Edwards M. G., Du J. P., Gatehouse A. M. R., and Gatehouse J. A. (2002), Putative protein digestion in a sap-sucking homopteran plant pest (rice brown planthopper; *Nilaparvata lugens*: Delphacidae) – identification of trypsin-like and cathepsin B-like proteases. Insect Mol. Biol. **32**, 967–978.

Gatehouse L. N., Shannon A. L., Burgess E. P. J., and Christeller J. T. (1997), Characterization of major midgut proteinase cDNAs from *Helicoverpa armigera* larvae and changes in gene expression in response to four proteinase inhibitors in the diet. Insect Biochem. Mol. Biol. **27**, 929–944.

George D., Ferry N., Back E. J., and Gatehouse A. M. R. (2008), Characterisation of midgut digestive proteases from the maize stem borer *Busseola fusca*. Pest Manag. Sci. **64**, 1151–1158.

Greer J. (1990), Comparative modeling methods: application to the family of the mammalian serine proteases. Proteins **7**, 317–334.

Kollien A. H., Waniek P. J., Prols F., Habedank B., and Schaub G. A. (2004), Cloning and characterization of a trypsin-encoding cDNA of the human body louse *Pediculus humanus*. Insect Mol. Biol. **13**, 9–18.

Lee S. I., Lee S. H., Koo J. C., Chun H. J., Lim C. O., Mun J. H., Song Y. H., and Cho M. J. (1999), Soybean Kunitz trypsin inhibitor (SKTI) confers resistance to the brown planthopper (*Nilaparvata lugens* Stål) in transgenic rice. Mol. Breeding **5**, 1–9.

Lemos F. J. A. and Terra W. R. (1992), Soluble and membrane-bound forms of trypsin-like enzymes in *Musca domestica* larval midguts. Insect Biochem. Mol. Biol. **22**, 613–619.

Lopes A. R. and Terra W. R. (2003), Purification, properties and substrate specificity of a digestive trypsin from *Periplaneta americana* (Dictyoptera) adults. Insect Biochem. Mol. Biol. **33**, 407–415.

Magalhães C. P., Fragoso R. R., Souza D. S. L., Barbosa A. E. A. D., Silva C. P., Finardi-Filho F., da Silva M. C. M., Rocha T. L., Franco O. L., and Grossi-de-Sa M.F. (2007), Molecular and structural characterization of a trypsin highly expressed in larval stage of *Zabrotes subfasciatus*. Arch. Insect Biochem. Physiol. **66**, 169–182.

Mazumdar-Leighton S. and Broadway R. M. (2001), Identification of six chymotrypsin cDNAs from larval midguts of *Helicoverpa zea* and *Agrotis ipsilon* feeding on the soybean (Kunitz) trypsin inhibitor. Insect Biochem. Mol. Biol. **31**, 633–644.

Milne R. and Kaplan H. (1993), Purification and characterization of a trypsin-like digestive enzyme from spruce budworm (*Choristoneura fumiferana*) responsible for the activation of delta-endotoxin from *Bacillus thuringiensis*. Insect Biochem. Mol. Biol. **6**, 665–673.

Oliveira M. G. A., De Simone S. G., Xavier L. P., and Guedes R. N. C. (2005), Partial purification and characterization of digestive trypsin-like proteases from the velvet bean caterpillar, *Anticarsia gemmatalis*. Comp. Biochem. Physiol. **140B**, 369–380.

Patankar A. G., Giri A. P., Harsulkar A. M., Sainani M. N., Deshpande V. V., Ranjekar P. K., and Gupta V. S. (2001), Complexity in specificities and expression of *Helicoverpa armigera* gut proteinases explains polyphagous nature of the insect pest. Insect Biochem. Mol. Biol. **31**, 453–464.

Rubia-Sanchez E., Suzuki Y., Miyamoto K., and Watanabe T. (1999), The potential for compensation of the effects of the brown planthopper *Nilaparvata lugens* Stål. (Homoptera: Delphacidae) feeding on rice. Crop Prot. **18**, 39–45.

Sambrook J., Fritsch E. F., and Maniatis T. (1989), Molecular Cloning: a Laboratory Manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

Srinivasan A., Giri A. P., and Gupta V. S. (2006), Structural and functional diversities in lepidopteran serine proteases. Cell Mol. Biol. Lett. **11**, 132–154.

Srivastava P. N. and Auclair J. L. (1963), Characteristics and nature of proteases from the alimentary canal of pea aphid, *Acyrthosiphon pisum* (Harr.) (Homoptera: Aphididae). J. Insect Physiol. **9**, 469–474.

Terra W. R. and Ferreira C. (1994), Insect digestive enzymes: properties, compartmentalization and function. Comp. Biochem. Physiol. **109B**, 1–62.

Terra W. R., Ferreira C., and Baker J. E. (1996), Compartmentalization of digestion. In: Biology of the Insect Midgut (Lehane M. J. and Billingsley P. F., eds.). Chapman and Hall, London, pp. 206–235.

Thompson J. D., Gibson T. J., Plewniak F., Jeanmougin F., and Higgins D. G. (1997), The ClustalX windows interface: flexible strategies for multiple sequence

alignment aided by quality analysis tools. Nucleic Acids Res. **24**, 4876–4882.

Vinokurov K. S., Elpidina E. N., Oppert B., Prabhakar S., Zhuzhikov D.P., Dunaevsky Y. E., and Belozersky M. A. (2006), Fractionation of digestive proteinases from *Tenebrio molitor* (Coleoptera: Tenebrionidae) larvae and role in protein digestion. Comp. Biochem. Physiol. **145B**, 138–146.

Wang S., Young F., and Hickey D. A. (1995), Genomic organization and expression of a trypsin gene from the spruce budworm, *Choristoneura fumiferana*. Insect Biochem. Mol. Biol. **25**, 899–908.

Yang Z. F., Zhang F. T., He Q., and He G. C. (2005), Molecular dynamics of detoxification and toxin-tolerance genes in brown planthopper (*Nilaparvata lugens* Stål., Homoptera: Delphacidae) feeding on resistant rice plants. Arch. Insect Biochem. Physiol. **59**, 59–66.

Yang Z. F., Zhang F. T., Zhu L. L., and He G. C. (2006), Identification of differentially expressed genes in brown planthopper *Nilaparvata lugens* (Homoptera: Delphacidae) responding to host plant resistance. Bull. Entomol. Res. **96**, 53–59.

Yang Z. F., Yang H. Y., and He G. C. (2007), Cloning and characterization of two cytochrome P450 CYP6AX1 and CYP6AY1 cDNAs from *Nilaparvata lugens* Stål (Homoptera: Delphacidae). Arch. Insect Biochem. Physiol. **64**, 88–99.

Zhang F. T., Zhu L. L., and He G. C. (2004), Differential gene expression in response to brown planthopper feeding in rice. J. Plant Physiol. **161**, 53–62.

Zhu Y. C. and Baker J. E. (1999), Characterization of midgut trypsin-like enzymes and three trypsinogen cDNAs from the lesser grain borer, *Rhizopertha dominica* (Coleoptera: Bostrichidae). Insect Biochem. Mol. Biol. **29**, 1053–1063.

Zhu Y. C., Kramer K. J., Dowdy A. K., and Baker J. E. (2000), Trypsinogen-like cDNAs and quantitative analysis of mRNA levels from the Indian meal moth, *Plodia interpunctella*. Insect Biochem. Mol. Biol. **30**, 1027–1035.

Zhu Y. C., Zeng F. R., and Oppert B. (2003), Molecular cloning of trypsin-like cDNAs and comparison of proteinase activities in the salivary glands and gut of the tarnished plant bug *Lygus lineolaris* (Heteroptera: Miridae). Insect Biochem. Mol. Biol. **33**, 889–899.

Zhu Y. C., Liu X., Maddur A. A., Oppert B., and Chen M. S. (2005), Cloning and characterization of chymotrypsin- and trypsin-like cDNAs from the gut of the Hessian fly [*Mayetiola destructor* (say)]. Insect Biochem. Mol. Biol. **35**, 23–32.