Archives of Insect Biochemistry and Physiology 66:169–182 (2007)

Molecular and Structural Characterization of a Trypsin Highly Expressed in Larval Stage of *Zabrotes subfasciatus*

Cláudio P. Magalhães,[1,5] Rodrigo R. Fragoso,[1,2,7] Djair S.L. Souza,[1] Aulus E.A.D. Barbosa,[1,3] Carlos P. Silva,[4] Flávio Finardi-Filho,[5] Maria C.M. da Silva,[1] Thales L. Rocha,[1] Octavio L. Franco,[3,6] and Maria F. Grossi-de-Sa[1,3]*

The Mexican bean weevil, *Zabrotes subfasciatus,* feeds on several seeds such as *Vigna unguiculata, Phaseolus vulgaris,* and *Pisum sativum,* causing severe crop losses. This ability to obtain essential compounds from different diets could possibly be explained due to a wide variability of digestive proteinases present in the weevil's midgut. These may improve digestion of many different dietary proteins. Coleopteran serine-like proteinases have not been thoroughly characterized at the molecular level. In this report, a full-length cDNA encoding a trypsin-like protein, named ZsTRYP, was isolated from *Z. subfasciatus* larvae using RT-PCR, 5′ and 3′ RACE techniques. The quantitative real-time PCR analysis strongly correlated the *Zstryp* transcript accumulation to the major feeding developmental larval stage. *Zstryp* cDNA was subcloned into pET101 vector and expressed in a *Escherichia coli* BL21(DE3) strain. Nickel-nitrilotriacetic acid (Ni-NTA) affinity chromatography was used to purify a 29.0-kDa recombinant enzyme. The purified ZsTRYP was then assayed with several synthetic peptide substrates and also challenged with different inhibitors. The biochemical data allowed us to classify ZsTRYP as a trypsin. Moreover, homology modeling analysis indicated a typical trypsin structural core and a conserved catalytic triad (His_{41}, Asp_{86}, and Ser_{182}). Arch. Insect Biochem. Physiol. 66:169–182, 2007.

KEYWORDS: *Zabrotes subfasciatus*; trypsin; enzyme characterization; homology modelling

INTRODUCTION

Polyphagous insects, especially the post-harvest storage-pests, have the ability to synthesize in their digestive tracts several hydrolytic enzymes, which can process a wide range of dietary inputs (Terra and Ferreira, 1994; Franco et al., 2004; Pelegrini et al., 2006). Among them, the Mexican bean weevil, *Zabrotes subfasciatus,* is capable to attack cotyledons from common beans (*Phaseolus vulgaris*), cowpeas (*Vigna unguiculata*), and peas (*Pisum sativum*) (Silva et al., 2001a,b). This process directly involves the production of different classes of digestive proteolytical enzymes (Silva and Xavier-

[1]Embrapa Recursos Genéticos e Biotecnologia, Brasilia-DF, Brazil

[2]Universidade de Brasília, UnB, Brasilia-DF, Brazil

[3]Centro de Análises Proteômicas e Bioquímicas, Pós-Graduação em Ciências Genômicas e Biotecnologia, UCB, Brasilia-DF, Brazil

[4]Universidade Estadual do Norte Fluminense Darcy Ribeiro, Rio de Janeiro-RJ, Brazil

[5]Faculdade de Ciências Farmacêuticas, USP, São Paulo-SP, Brazil

[6]Departamento de Biologia, Universidade Federal de Juiz de Fora, UFJF, Juiz de Fora-MG, Brazil

[7]EMBRAPA Cerrados, Brasília-DF, Brazil

Contract grant sponsor: EMBRAPA; Contract grant sponsor: UCB; Contract grant sponsor: CNPq; Contract grant sponsor: CAPES.

Abbreviations used: AMV-RT = avian myeloblastosis virus reverse transcriptase; GSH = glutathione reduced form; GSSG = glutathione oxidized form; LB = Luria Bertani Broth; PMSF = phenylmethylsulphonylfluoride; SKTI = Soybean Kunitz Trypsin inhibitor; S1 = specificity pocket; TLCK = N-tosyl-L-lysine-chloromethyl ketone; TPCK = L-1-Chloro-3-[4-tosylamido]-4-phenyl-2-butanone; ZsTRYP = *Z. subfasciatus* trypsin.

*Correspondence to: Maria F. Grossi-de-Sa, Embrapa Recursos Genéticos e Biotecnologia, PqEB, Final W3 Norte, 70770-900, Brasília, DF, Brazil. E-mail: fatimasa@cenargen.embrapa.br

Received 29 January 2007; Accepted 3 May 2007.

DOI: 10.1002/arch.20208
Published online in Wiley InterScience (www.interscience.wiley.com)

Filho, 1991; Silva et al., 2001a,b). In Coleoptera, these digestive enzymes are secreted by a typical exocytosis from posterior midgut (Cristofoletti et al., 2001). These proteinases usually help in understanding insect–pests adaptation, possibly driven by the presence of numerous anti-nutritional compounds such as proteinase inhibitors, typically found in host plants (Ishimoto and Chrispeels, 1996; Silva et al., 1999, 2001a,b; Brito et al., 2001; Mazumdar-Leighton and Broadway, 2001; Patankar et al., 2001; Moon et al., 2004; Li et al., 2005). Furthermore, trypsins could also be involved in the mediation of immune/developmental signals (De Gregorio et al., 2001; Dimopoulos et al., 2001) and in extra- and intracellular protein breakdown, catalyzing specific cleavages on the edge of cationic residues (Arimatsu et al., 2007).

Despite a significant literature on trypsin sequences from other insect groups, coleopteran serine-proteinases have not been thoroughly characterized and rarely purified. This may be due to the very low quantities of these enzymes in their digestive tracts (Tsybina et al., 2005). Here, we report the cDNA cloning, expression pattern determination, heterologous expression, and refolding of a trypsin-like serine proteinase, ZsTRYP, from *Z. subfasciatus* larvae.

MATERIALS AND METHODS

Cloning *Z. subfasciatus* Trypsin cDNA

Larvae of *Z. subfasciatus* (third to fourth instars) were reared on *P. vulgaris* seeds in a Plant Pest Interaction Laboratory (Cenargen/ Embrapa) at 30°C and total darkness. Midguts were removed and ground in liquid nitrogen, and the total RNA was extracted using RNAeasy kit (Qiagen). Five micrograms of total RNA was reverse transcribed using Avian Myeloblastosis virus Reverse Transcriptase (AMV-RT, Promega) at 42°C for 1 h using oligo dT primers. cDNA fragments were amplified by polymerase chain reaction (PCR) using two oligonucleotide degenerate primers (5′-ACTGCTGCHCAYTG-3′ and 5′-GGRCCACCAGAGTCRCC-3′) designed according to conserved insect serine-proteinases motifs. The fragment amplified by PCR was cloned into pGEM-T easy vector (Promega) and the sequence was determined by using an automated sequencer. The cDNA sequence was analyzed in GenBank provided by the National Center for Biotechnology Information using the BLASTx program (Altschul et al., 1997; Gish and States, 1993). Based on the previous amplified sequences, specific oligonucleotides were synthesized for rapid cDNA amplification using RACE-PCR. The 5′ ends were amplified using a 5′RACE system (Gibco-BRL). Three specific primers were designed (ZSREVSERI: 5′-GCACAGATCATATTCC-3′; ZSREVSERII: 5′-CCA GTATGTCCATAC-3′; and ZSREVSERIII: 5′-GCA AGACCCTCCCTCCGG-3′). The first was used to synthesize the cDNA with AMV-RT. The cDNAs were purified and a homopolymeric tail was added to the 3′ end of the cDNA using Terminal deoxynucleotidyl transferase (TdT) and dATP. A first step of PCR was done using the specific reverse primer ZSREVSERII and poly dT-Ancor primer. Then 1.0 μL of the product of the initial step was used for a second step PCR with the nested primer (ZSREVSERIII) and the anchor primer at the same conditions. A 554-bp product was amplified, cloned, and sequenced. To obtain full trypsin-like protein cDNA, one specific forward primer (TRIPZABFOR1) located at the 5′-end of the trypsin-like precursor cDNA was designed and used in PCR amplification with an oligo-dT reverse primer. In order to improve the specificity, second primer (TRIPZABFOR2) and poly dT primer were used with 1.0 μL of the first step as a template. The cDNA clone was sequenced from both directions in an automated ABI DNA sequencer, using T7 and SP6 primers.

Developmental Expression Pattern Determination by Quantitative Real Time PCR

Developmental stages of *Z. subfasciatus* were collected and immediately frozen in liquid nitrogen. Total RNA from eggs, larvae, pupae, and adult insects were extracted using RNeasy kit (Qiagen) and cDNA was synthesized with oligo dT primer and reverse transcriptase (MMLV-RT) as described

above. A control reaction using genomic DNA was performed to certify that analysis proceeds just by cDNA amplification and not by genomic DNA amplification. The chosen method for quantitative real time (RT) PCR was the common primers and the SYBR Green I (BIO-RAD), an intercalating base-pair DNA, whose fluorescent signal corresponds to FAM-490, read at the elongation phase. Real-time quantitation of *Zstryp* mRNA was performed with an iCycler iQ detection system (BIO-RAD, Hercules, CA) according to Yalcin et al. (2004). Initially, four combinations of primer pairs were evaluated, targeting the trypsin gene plus the actin primers. These five reactions (done in duplicate) were submitted to annealing temperature optimization with gradient of annealing temperature (50.0, 55.5, 59.3, and 65.0°C). The PCR program consisted of an initial denaturation at 95°C for 5 min, 35 cycles of amplification at 95°C for 30 sec, 55°C for 45 sec, and 72°C for 1 min, followed by a final elongation step at 72°C for 5 min. The best annealing temperature and trypsin primer combination were determined as the minor Ct and sharper sigmoid curve generated by RT PCR. The melt curve of PCR products, whose graph indicates the number of fragments amplified, was also determined. This PCR program consisted of a range of 50 to 90°C by adding 0.5°C at each 30 sec. This control in addition to the previous one using genomic DNA can ensure that the quantitative PCR data correspond to cDNA amplification, related to expression pattern. A standard curve of Ct as a function of template dilution factor (10e0, 10e3, 10e6, and 10e9), necessary to help the threshold-line determination, was done. Quantitative PCR was done in duplicate (reference gene) and triplicate (target gene) for the several analyzed life stages. The primers used were ZTrypForw2 5′-AATGTCCTCCACAACGC-3′ and ZtrypRev3 5′-GTAACATATCCCCAACC-3′ to the target gene and ActinForw 5′-GATCTGGCATCACACCTTCTAC-3′ with ActinRev 5′-AGGAAGCTCGTAGCTCTTCTC-3′ to the reference gene. The PCR program utilized was the same as the gradient experiment, except the annealing temperature, which was 55°C. Finally, the data were submitted to statistical, mathematical analysis following the Livak Method (Livak and Schmittgen, 2001), to infer a developmental expression pattern by comparing mRNA accumulation in different life stages of *Z. subfasciatus*.

Expression of ZsTRYP enzyme

The cDNA encoding *Z. subfasciatus* pro-mature serine proteinase was amplified by PCR using p*fu* polymerase enzyme with the oligonucleotide primers 5′-CACCATGAACCTACCTCGTCCAGACGGA-3′ (forward) and 5′-GATATGTGCTATCCAAGCCCTG-3′ (reverse) to obtain a blunt end sequence. The PCR products were isolated, then linked into the expression vector, pET101 (Invitrogen), using the Topo isomerase reaction and transformed into a competent *E. coli* BL21 (DE3) strain (Invitrogen) using heat shock treatment. Transformed *E. coli* cells were first grown on Luria–Bertani (LB) medium plates supplemented with 100 $\mu g.mL^{-1}$ ampicilin (Sigma). Single colonies were inoculated into 5.0 mL LB medium containing 100 $\mu g.mL^{-1}$ ampicilin. After overnight growth, this culture was transferred to a 2-liter flask containing 500 mL LB medium supplemented with 100 $\mu g.mL^{-1}$ ampicilin. Cells were grown with vigorous shaking (200 rpm) to a density of $O.D._{600}$ = 0.6–0.8, then the expression of recombinant protein was induced by 1 mM IPTG, and the incubation was continued for several hours. Samples were collected in different times and the level of expression was visualized by SDS-PAGE (Laemmli, 1970). The *Z. subfasciatus* serine proteinase was fused to His-tag and expressed under control of the expression plasmid T7 promoter.

Purification and Refolding of the Recombinant 6(His) Serine Proteinase

The pellet of induced bacteria was suspended in 20 mL of 20 mM Tris-HCl, 500 mM NaCl, 6M urea, pH 8.0 buffer. The supernatant was transferred into a tube containing Ni^{2+}-charged His-Bind Resin and the mixture was centrifuged on a Rotamix for 1 h. The pellet was washed twice with the same buffer to eliminate the contaminants. The bound proteins were eluted with the same buffer contain-

ing 200 mM imidazole and the ZsTRYP purity degree was SDS–PAGE checked (Laemmli, 1970). The expressed recombinant protein was refolded according to the following steps: the eluted protein was reduced adding DTT to a final concentration of 5.0 mM. The extract was added dropwise, under stirring, at 4°C into an oxidation buffer of 100 mM Tris–HCl, pH 8.0, 0.78 mM GSSG (glutathione disulfide), and 7.8 mM GSH (reduced glutathione). After 24 h, the refolded protein was dialyzed in water and lyophilized.

Proteinase Activity of the ZsTRYP Recombinant Protein

The lyophilized ZsTRYP was resuspended in 0.1M Tris-HCl, pH 8.0. The proteolytic activity was measured using fluorogenic substrates at a standard concentration of 0.25 mM (Z-Phe-Arg-MCA, Z-Gly-Gly-Arg-MCA, Z-Gly-Pro-Arg-MCA, Z-Arg-MCA, Z-Ala-Arg-Arg-MCA, and Z-Arg-Arg-MCA). The assays were buffered in 0.05 M Tris-HCl, $CaCl_2$ 100 μM buffer, pH 8.0, and then reactions were begun with the addition of the substrates. After 10 min at 37°C, the reactions were stopped by addition of 100% ethanol. Fluorescence was measured for the free release of methyl-coumarin on a HITACHI F2000 fluorimeter at 440 nm after an excitation at 380 nm. All assays were performed in triplicate. The optimal pH was determined in enzymatic assays as described before, buffered with 0.05 M Tris-HCl, buffer (pH 8.0 and 9.0), 0.05M sodium acetate buffer (pH 5.0 and 6.0), and 0.05 M phosphate buffer (pH 7.0). One enzymatic unit was defined as the amount of enzyme that catalyzes the hydrolysis of 1 μmol of substrate/min. Analysis of trypsin activity in the presence of proteinaceous and synthetic inhibitors was done using azocasein as substrate (Tomarelli et al., 1949). The synthetic inhibitors PMSF, TLCK, and TPCK (2 mM) and soybean Kunitz trypsin inhibitor-SKTI (100 μg) were incubated for 30 min with ZsTRYP. The reaction was carried out for 1 h and stopped by adding 20% trichloroacetic acid. The reaction medium was centrifuged at 10,000*g* for 10 min at 4°C and the supernatant was read at 440 nm.

In Silico Analysis

The full-length cDNA sequence was translated and pre- and pro-enzyme regions were identified using the SignalP 3.0 server software (http://www.cbs.dtu.dk/services/SignalP/) of the Center for Biological Sequence Analysis Prediction Server, BioCentrum-DTUTechnical, University of Denmark site (Bendtsen et al., 2004). Databank comparisons were done using the BLASTx software (Altschul et al., 1997) from the NCBI databank (http://www.ncbi.nlm.nih.gov). Ten trypsin-like sequences of the insects *Zabrotes subfasciatus, Rhyzopertha dominica* (AAD31268), *Phaedon cochleariae* (CAA 76929) (Coleoptera), *Anopheles gambiae* (XP_317171), *Drosophila melanogaster* (AAA17453), *Culex pipiens* (AAB37261), *Phlebotomus papatasi* (AAM 96942), *Aedes aegypti* (AAL93209) (Diptera); *Manduca sexta* (T10109) (Lepidoptera), and *Apis mellifera* (XP_397087) (Hymenoptera) were compared using CLUSTAL W software (Thompson et al., 1994) and were edited with the BOXSHADE software (http://www.ch.embnet.org/software/BOX_form.html). Protein Machine software, available at the Expasy site (http://us.expasy.org/tools/), was utilized to determine the molecular mass and p*I* of deduced ZsTRYP protein. Therefore, a molecular model of ZsTRYP was generated. BioInfo Meta Server (Ginalski et al., 2003) was used to find a specific template. Best scores of FFAS03, PDB-Blast, and 3DPSSM indicated that crayfish trypsin (Fodor et al., 2005) structure (PDB code: 1YR4) showed an enhanced structural similarity. The model was constructed using DeepView/Swiss PdbViewer Program, version 3.7, developed by the Swiss Institute of Bioinformatics (Guex and Peitsch, 1997). A raw sequence of Zstryp was loaded and fitted according to the structural alignment produced before. After superposition of atomic coordinates of 261 residues, an energy minimization was done using Gromos96, a force field that predicted the dependence of a molecular conformation on the type of environment (water, methanol, chloroform, DMSO, non-polar solvent, crystal, etc.). The program calculated the relative binding constants by evaluating free energy differences between various

molecular complexes using thermodynamic integration, perturbation, and extrapolation. The software predicts energetic and structural changes caused by modification of amino acids in enzymes. This method used six subsequent rounds, minimizing backbone and side chains (3,000 steps of steepest descent). Ramachandran plot and rmsd values were considered to validate the model.

RESULTS

Isolation and In Silico Characterization of ZsTRYP

The cDNA full length for a trypsin-like protein, designated ZsTRYP, obtained by PCR techniques, contains 867 nucleotides, including a 5′ untranslated region of 51 nucleotides and a 3′ UTR of 30 nucleotides (GenBank accession no. DQ458994). The complete open reading frame (786 nucleotides) encodes a proteinase with a typical zymogene-structure, which contains a pre-region (26 residues), a pro-region (8 residues), and a mature enzyme (227 residues) (Fig. 1). The pre-region was predicted by using SignalP program (http://www.cbs.dtu.dk/services/SignalP/ and flanking this region, a conserved zymogen motif (R-IVGG) was identified. A short conserved pro-region was also identified between the pre-region and the mature enzyme.

The mature enzyme has a theoretical molecular mass of 24.8 kDa and p*I* of 8.7 (http://www.expasy.ch/tools/pi_tool.html). Translated nucleotide sequence of *Zstryp* was also analyzed in the Data Bank through BLASTx, showing that the predicted protein shared similarities to insect trypsins. The higher scores were obtained for trypsins from dipterans as *Anopheles stephensi* (46%) and *Phlebotomus papatasi* (45%). Similarities to coleopterans such as the lesser grain borer *Rhyzopertha dominica* (42%), the cotton boll weevil *Anthonomus grandis* (42%), and the mustard beetle *Phaedon cochleariae* (37%) were also found (Fig. 2). The lack of similarity with more related insects from the Bruchidae could be justified by the absence of bruchid primary structure loaded in Data Bank. The similarity of Zstryp amino acid residues sequence to other organisms such as fungi *Aspergillus nidulans* (43%) and mammalian indicated highly enzyme-fold conservation during evolution. The alignment of ZsTRYP predicted amino acid sequence to other trypsin-like proteinases from insects showed an enhanced conserved C-terminal region in all sequences analyzed (Fig. 2). Six highly conserved cysteine residues, predicted to occur in disulfide bridges, were identified at positions Cys_{26}, Cys_{42}, Cys_{151}, Cys_{166}, Cys_{178}, and Cys_{202}. These positions are conserved in other trypsin-like enzymes as observed in *Ligius lineoralis* and *Rhyzopertha dominica* (Zeng et al., 2002; Zhu and Baker 1999). An extra pair of cysteine residues was identified at positions 70 and 93 in the ZsTRYP mature sequence. However, these cysteines are unable to form disulfide bridges according to the homology model, described below. Based on the presence of specific residues involved in the substrate recognition (Asp_{176}, Gly_{199}, and Gly_{209}), we suggest common trypsin substrate specificity for ZsTRYP.

Furthermore, the complete sequence of ZsTRYP allowed us to create a three-dimensional model, in order to understand functional and structural questions. Structural alignment against other trypsins indicated a high conservation of residues involved in the catalytic process (Fig. 3). According to the molecular model here proposed, ZsTRYP exhibits the conserved structural core of the chymotrypsin fold, consisting of two six-stranded β-barrel domains packed against each other, with the catalytic residues located at the junction of the two barrels (Fig. 3). Possible trypsin residues involved in catalysis (His_{41}, Asp_{86}, and Ser_{182}) are conserved, forming the catalytic triad of the active site. In a Ramachandran plot analysis 1.5% of residues were not found in allowed regions.

Expression of Zstryp mRNA in Developmental Stages of *Z. subfasciatus*

The quantitative RT-PCR revealed the expression pattern of mRNA *Zstryp* along the developmental stages of egg, larvae, pupae, and adult. Control reactions were performed with genomic DNA as template, which showed higher size fragments in agarose gels when compared to fragments obtained

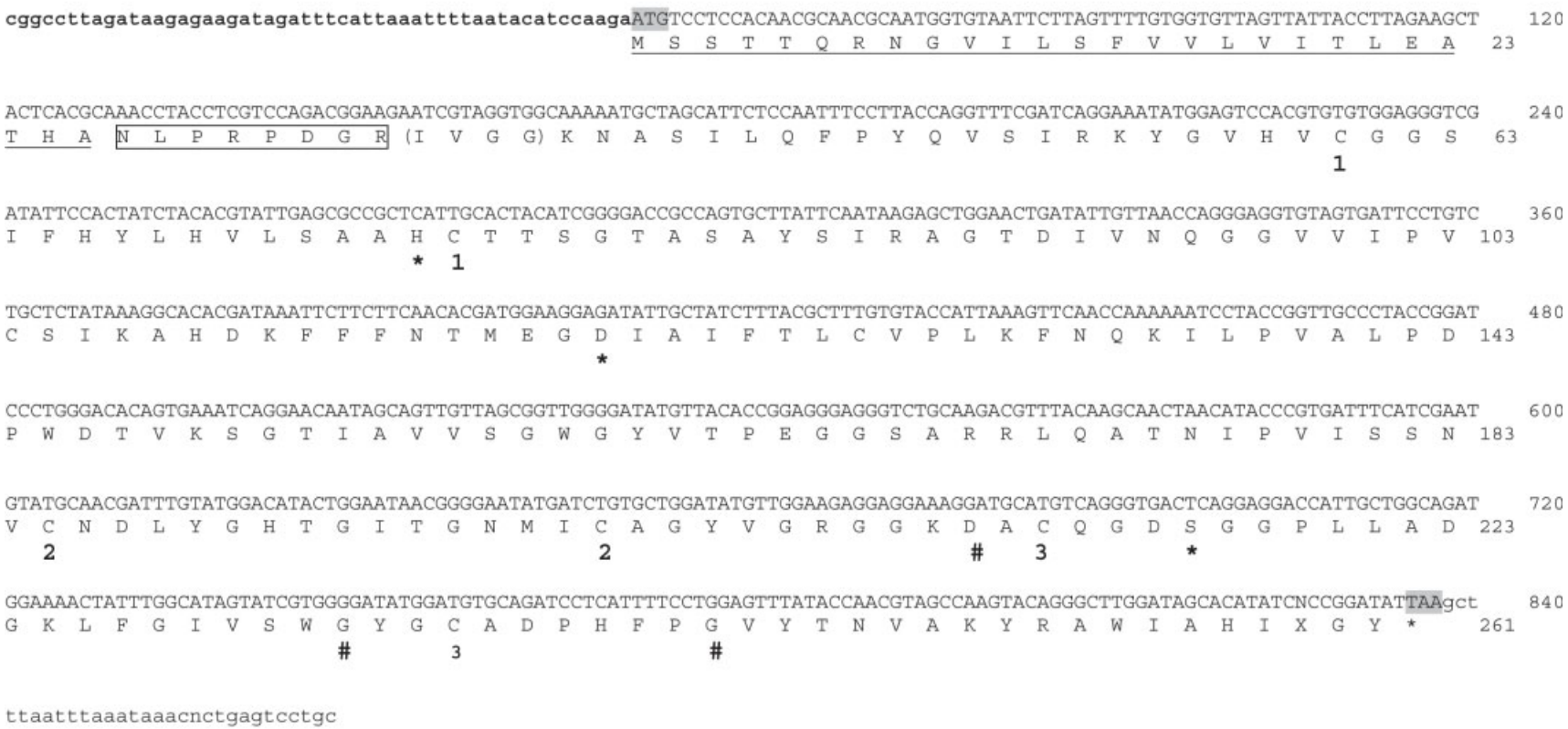

Fig. 1. The nucleotide sequence and derived amino acid sequence of a trypsin-like protein from *Z. subfasciatus* (ZsTRYP). 5′ non-translated regions are indicated in bold; start codon (ATG) and stop codon (TAA) are depicted in light gray; signal peptide is underlined; pro-region is boxed; activation motif is in parentheses; cysteine residues in conserved disulfide bridges are numbered; catalytic triad are indicated with an asterisk (*) (His_{41}, Asp_{86}, and Ser_{182}); the three residues that determine specificity in tripsin-like enzymes are marked with # (Asp_{176}, Gly^{200}, and Gly_{210}).

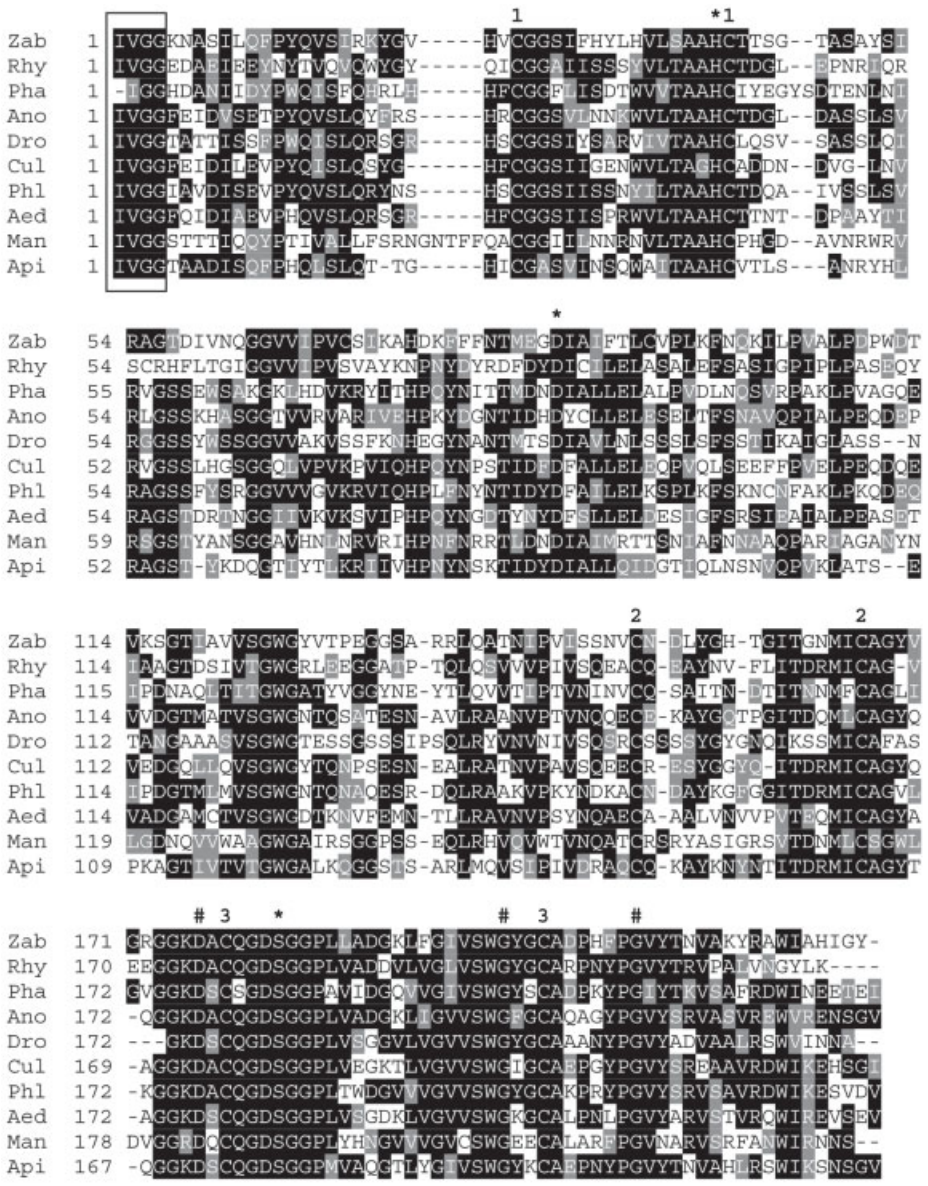

Fig. 2. Alignment of amino acid residues of ZsTRYP (against insect trypsin-like Rhy, *Rhyzopertha dominica* (AAD31268); Pha, *Phaedon cochleariae* (CAA76929); Ano, *Anopheles gambiae* (XP_317171); Dro, *Drosophila melanogaster* (AAA17453); Cul, *Culex pipiens* (AAB37261); Phl, *P. papatasi* (AAM96942); Aed, *Aedes aegypti* (AAL93209); Man, *Manduca sexta* (T10109); Apis, *Apis mellifera* (XP_397087). Conserved activation motifs are indicate in box; cysteine residues involved in disulphide bond formation are numbered; catalytic triad is indicate with asterisks (*); the three residues (Asp_{176}, Gly_{199}, Gly_{209}) that determine specificity in trypsin-like enzymes are indicated with #.

by PCR using cDNA as a template. These data could be explained by the presence of introns. The validation of SYBR Green I reactions throughout the melt curve graphic demonstrated one peak per reaction, indicating a high specificity of primer sets resulting in one amplicon (data not shown). Based on β-actin normalized data, the highest expression signal of mRNA *Zstryp* was recorded in larvae (Table 1). Larvae showed an expression level 176 times higher, when compared to the pupae stage. Egg expressed mRNA *Zstryp* 46 times higher than pupae. As expected, *Zstryp* mRNA expression detection in

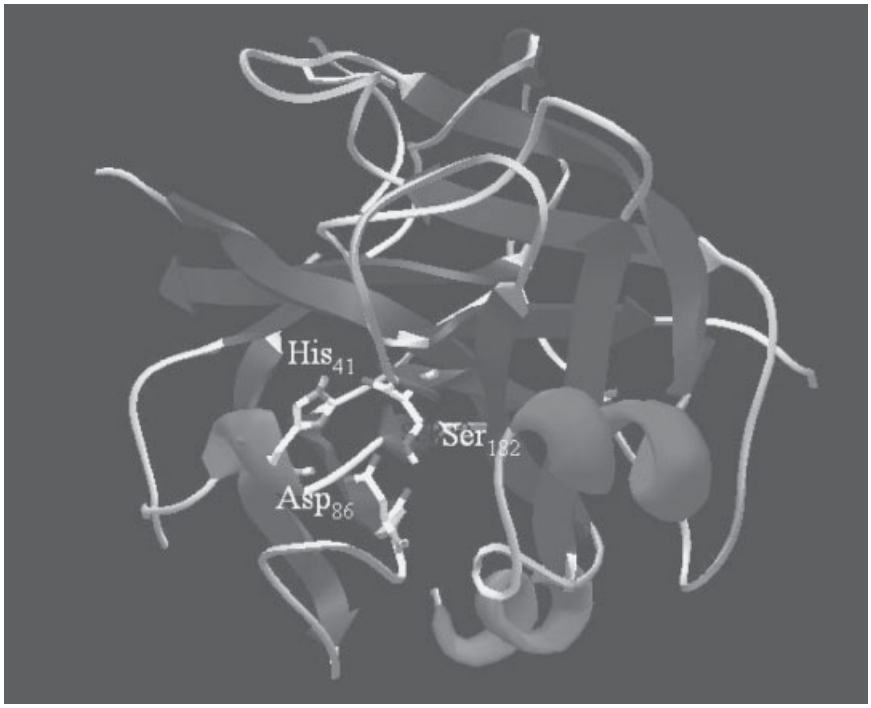

Fig. 3. Ribbon diagram of the predicted molecular model of ZsTRYP (gray). The residues involved in catalyses (His_{41}, Asp_{86}, and Ser_{182}) are shown in ball-and-stick representation. Model was visualized by using SPDB-viewer version 3.7 (Guex and Peltish, 1997).

TABLE 1. Fold Change in Expression of *Zstryp* cDNA Relative to the Reference Gene β-*actin* in Different Developmental Stages*

Developmental stages	β-*actin* C_T (reference gene)	*Zstryp* C_T (target gene)	ΔC_T (targ-refe)	$\Delta\Delta C_T$ (test-calib)	$2^{-\Delta\Delta C_T}$
Egg	25.2	26.6			
(test)	24.5	26.7			
		26.3			(31.78–67.18)
M. ± S.D.	24.85 ± 0.49	26.53 ± 0.21	1.68 ± 0.54	–5.53 ± 0.54	46.31
Larvae	26.3	25.7			
(test)	25.8	25.6			
		26.1			(130.69–240.52)
M. ± S.D.	26.05 ± 0.35	25.80 ± 0.26	–0.25 ± 0.44	–7.47 ± 0.44	176.88
Pupae	23.2	29.9			
(calibrator)	22.5	30.9			
		29.4			(0.53–1.88)
M. ± S.D.	22.85 ± 0.49	30.07 ± 0.76	7.22 ± 0.91	0.00 ± 0.91	1.00
Adult	19.8	25.7			
(test)	20.0	26.2			
		25.8			(1.89–2.87)
M. ± S.D.	19.90 ± 0.14	25.90 ± 0.26	6.00 ± 0.30	–1.22 ± 0.30	2.32

*Analysis following the Livak method for treatment of replicates data in which target and reference are amplified in separate wells. The ΔC_T normalize *Zstryp* amplification (target gene) in relation to β-*actin* amplification (reference gene) for each developmental stage. The $\Delta\Delta C_T$ normalize between different developmental stages in relation to pupae (calibrator), because pupae resulted in the highest ΔC_T. The $2^{-\Delta\Delta C_T}$ formula results in normalized *Zstryp* fold amount. M., mean; S.D., standard deviation.

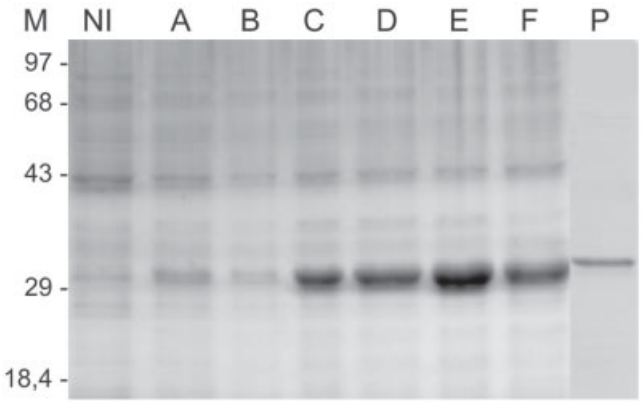

Fig. 4. SDS-PAGE of a time curse expression of the ZsTRYP protein. M, molecular weight marker; NI, non-induced *E. coli*; A–F, 1, 2, 3, 4, 12, and 24 h after IPTG 1-mM induction; P, purified protein in Ni-NTA matrix.

pupae and adult insect was lower than the expression observed in other developmental phases in *Z. subfasciatus* (Table 1).

Heterologous Expression and Purification

The *Zstryp* cDNA pre-mature sequence was inserted into pET 101, and expressed in *E. coli* BL21 (DE3). Following an induction time curve of ZsTRYP expression, the bacteria synthesized ZsTRYP His-tag marked after 3-h incubation. Proteins resulted in a molecular mass of 29.0 kDa, visualized by SDS-PAGE technique (Fig. 4). The highest protein content was observed after 12 h of incubation and the enzyme was further purified using a Ni-NTA column (Fig. 4, line P). The recombinant protein was expressed in large amounts in the soluble fraction; however, no proteolytic activity was detected (data not shown). The recombinant enzyme was extracted with denaturing agents and after purification it was treated with refolding buffer, recovering after this process almost all the expected enzyme activity.

Amydolitic Activity of ZsTRYP

The recombinant ZsTRYP showed high amidolytic activity against the fluorogenic substrates Gly-Pro-Arg-MCA and Phe-Arg-MCA. On the other hand, low proteolytic activity against the substrates Arg-MCA, Ala-Arg-Arg-MCA, Arg-Arg-MCA and Gly-Gly-Arg-MCA was obtained (Fig. 5). These data

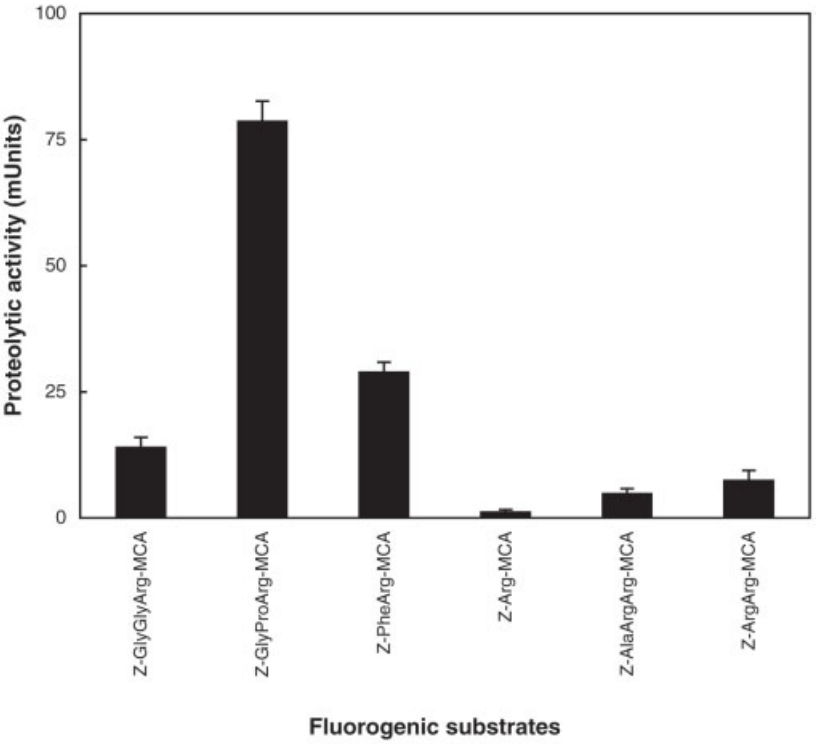

Fig. 5. Proteolytic assay ZsTRYP using different fluorimetric substrates. Each assay was carried out in triplicate and they did not differ more than 10%. Vertical bars correspond to standard deviation. One activity unit was defined as the amount of enzyme capable of catalyzing 1 µmol substrate.min^{-1}

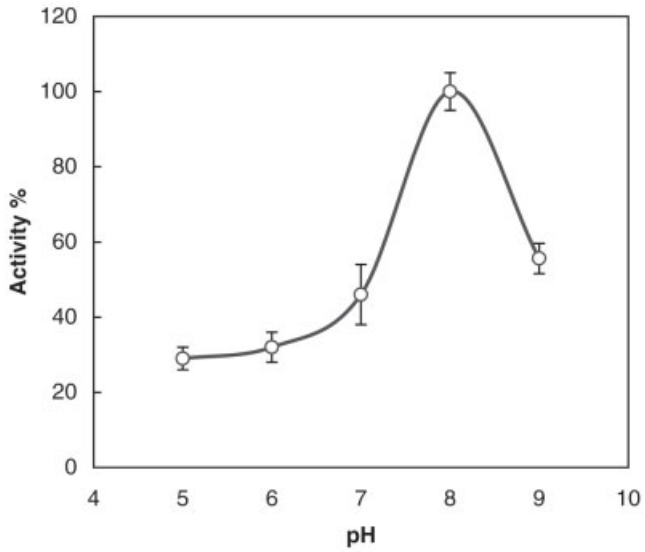

Fig. 6. Effect of pH on the *Zs*TRYP activity. Each assay was carried out in triplicate and did not differ more than 10%. Bars correspond to standard deviation.

demonstrate that this enzyme preferentially cleaved sequences that have residues with cyclic structures in the lateral chain before arginine. The enzyme is mainly active at pH 8.0 (Fig. 6). However, residual activity at pH 7.0 and 9.0 was also identified (Fig. 6). In relation to common synthetic serine-proteinase inhibitors, which included PMSF and TPCK, they were able to inhibit *Zs*TRYP. Otherwise, the proteinaceous inhibitor SKTI caused insignificant inhibition rates (Fig. 7).

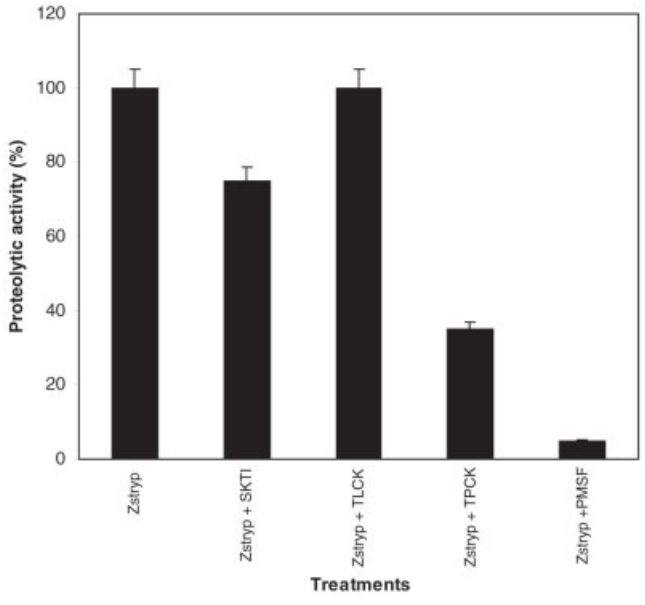

Fig. 7. *Zs*TRYP activity in the presence of synthetic and proteinaceous inhibitors. Each assay was carried out in triplicate and did not differ more than 10%. Vertical bars correspond to standard deviation.

DISCUSSION

Serine proteinases are involved in several biological events including digestion processes, blood clotting, pro-enzyme activation, signal transduction pathways, and the complement cascade of the immune system (Coughlin, 1999; Lopes et al., 2006). In insect digestion, several trypsins are synthesized in midgut (Terra and Ferreira, 1994). These are thought to be strictly involved in resistance to plant proteinase inhibitors (Jongsma et al., 1995; Zhu and Baker, 1999; Brito et al., 2001; Mazumdar-Leighton and Broadway, 2001). Among coleopterans, the trypsins were identified in *Rhyzopertha dominica*, *Tenebrio molitor*, *Phaedon cochleariae*, *Diaprepes abbreviata*, *Sitophilus oryzaes*, and *Anthonomus grandis*. The data obtained here indicate that the open reading frame of 786 nucleotides encodes an amino acid sequence with clear similarity to trypsin-like enzymes. Several experiments help to understand the function and structure of *Zs*TRYP. Initially, the quantitative RT-PCR was used to correlate mRNA accumulation with the developmental stages of *Zabrotes subfasciatus*. The larval expression levels were 3.8-fold higher, in comparison to expression levels in eggs. Lower mRNA *Zstryp* synthesis was obtained in pupae and adult. We infer that *Zs*TRYP acts in digestion. Otherwise, mRNA was observed in eggs, suggesting a probable *Zs*TRYP activity in protein recovery and amino acid distribution during *Z. subfasciatus* development. Previous reports had described variation of digestive trypsin-like enzymes in insects, including *Ostrinia nubilalis*. This insect is strongly resistant to *Bt* toxins, being able to increase the expression of endogenous digestive trypsin-like enzymes (Li et al., 2005).

Additionally, the biochemistry of heterologously expressed *Zs*TRYP was characterized. The use of heterologous expression systems has demonstrated remarkable advantages over the traditional biochemical approach (Dimopoulos et al., 1996; Mu-

lenga et al., 2001; Hanquier et al., 2003; Nakajima et al., 2003). Otherwise, despite the benefits of recombinant enzyme production, the trypsin expression in *E. coli* heterologous systems has also shown a common self-digestion enzyme and a multiple disulfide bound formation, due to the reducing environment in *E. coli* bacteria cytoplasm (Verheyden et al., 2000). Novel strategies have been applied to solving these difficulties, which included periplasmic expression, a known oxidant environment (Vasquez et al., 1989; Fukuoka et al., 2002), the use of a thioredoxin redutase deficient mutant strain (Verheyden et al., 2000; Prinz et al., 1997), and cytosolic expression followed by the use of oxidating buffers, as developed in this work (Shan et al., 2003; Hohenblum et al., 2004). The enzyme was mainly active at pH 8.0, but residual activity at pH 7.0 and 9.0 was recorded (Fig. 6). This enzyme shows a molecular mass of around 29 kDa. Among trypsin, the trypsin isolated from *T. molitor* showed molecular mass of 25.5 kDa, isoelectric point of 7.4, and an enhanced activity at pH 8.5 (Tsybina et al., 2005). Zymogram analysis of *Rhyzopertha dominica* indicated the presence of three trypsin-like isoenzymes with molecular masses ranging from 23.0 to 30.0 kDa with enhanced activity under basic reaction conditions (Zhu and Baker, 1999). Molecular analysis of digestive enzymes from *Phaedon cochleariae* also showed three trypsin-like cDNAs and one major activity band observed in zymogram with approximately 23.0 kDa (Girard and Jouanin, 1999). Furthermore, a family of serine proteinases was characterized in *A. grandis*, comprising 11 trypsin-like sequences and three chymotrypsin-like sequences (Oliveira-Neto et al., 2004).

Another important question concerns proteinases synthesized in response to the presence of antifeedant compounds. When *A. grandis* neonate larvae were reared on an artificial diet containing SKTI, a larval weight reduction of up to 64% was observed for the highest SKTI concentration of 500 μM (Franco et al., 2004). These data indicated that trypsin inhibitors probably are inefficient to control *Z. subfasciatus*. Nevertheless, novel proteinaceous inhibitors should be screened against *Z. subfasciatus* enzymes.

The proposed ZsTRYP model showed similarity to other well-characterized serine-proteinases such as bovine (Helland et al., 1999) and crayfish trypsin (Fodor et al., 2005). The Ca^{2+} binding loop is also conserved in all of them. The major binding feature of ZsTRYP to substrate appears to be identical to that in the crayfish trypsin, since the catalytic triad is identical. One important difference from vertebrate trypsins is that bovine trypsin has six disulphide bonds (Helland et al., 1999) while ZsTRYP has four. This is also different from the disulphide bond pattern observed for crayfish trypsin, which has three (Fodor et al., 2005). All evolutionarily conserved half-cysteines were close to the active site of the enzyme (Roach et al., 1997), suggesting their importance in trypsins structural stability (Kardos et al., 1999).

Since no three-dimensional structure was elucidated until now, theoretical homology modeling of ZSATryp was also utilized to support some hypothesis on the structural bases of differences on primary substrate specificity. The most important residues of specificity pocket (S1), Asp_{86} and Ser_{182}, are observed and numbered in Figure 3. Both residues made trypsin 4-fold more active on Arg than on Lys at S1, as observed for bovine trypsin (Lopes et al., 2006). Another important factor is that a Pro followed by an Arg in substrate is necessary for an increased hydrolysis. When hydrophobic residues as Phe are attached to Arg, the proteolytic activity is strongly reduced (Fig. 5). This was an expected result, since the preferences between polar and hydrophobic residues are markedly different among insect enzymes (Lopes et al., 2006). Otherwise, some reports considered that high polar sub-sites suggest a clear adaptation of insect-pest plant feeders, in order to reduce the deleterious effects caused by ingestion of plant material rich in trypsin inhibitors, which have their surfaces rich in hydrophilic residues (Lopes et al., 2006). We had found an opposite result, since *Z. subfasciatus* trypsin showed higher affinity to polar substrate. Due to the difficulty for localization of hydrophobic Zstryp subsites, our data were insufficient to carry out quantitative comparisons between polar and non-polar substrates. Crystallographic data

could shed some light over the results reported here. In conclusion, the theoretical structure proposed here, in addition to biochemical characterization, provides a basis for further studies of the inhibitor screening and structural aspects of proteinase inhibitor specificity.

LITERATURE CITED

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman, DJ 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402.

Arimatsu Y, Kotani E, Sugimura Y, Furusawa T. 2007. Molecular characterization of a cDNA encoding extracellular dsRNase and its expression in the silkworm, *Bombyx mori*. Insect Biochem Mol Biol 37:176–183.

Bendtsen JD, Nielsen H, von Heijne G , Brunak S. 2004. Improved prediction of signal peptides: SignalP 30. J Mol Biol 340:783–795.

Brito LO, Lopes AR, Parra JR, Terra WR, Silva-Filho MC. 2001. Adaptation of tobacco budworm *Heliothis virescens* to proteinase inhibitors may be mediated by the synthesis of new proteinases. Comp Biochem Physiol 128:365–375.

Coughlin, SR 1999. How the protease thrombin talks to cells. Proc Natl Acad Sci USA 96:11023–11027.

Cristofeletti PT, Ribeiro AF, Terra WR. 2001. Apocrine secretion of amylase and exocytosis of trypsin along the midgut of *Tenebrio molitor* larvae. J Insect Physiol 47:143–155.

De Gregorio E, Spellman PT, Rubin GM, Lemaitre B. 2001. Genome-wide analysis of the Drosophila immune response by using oligonucleotide microarrays. Proc Natl Acad Sci USA 98:12590–12595.

Dimopoulos G, Richman A, Torre AD, Kafatos FC. 1996. Identification and characterization of differentially expressed cDNAs of the vector mosquito, *Anopheles gambiae.* Proc Natl Acad Sci USA 93:13066–13071.

Dimopoulos G, Müller HM, Levashina EA, Kafatos FC. 2001. Innate immune defense against malaria infection in the mosquito. Curr Opin Immunol 13:79–88.

Fodor K, Harmat V, Hetenyi C, Kardos J, Antal J, Perczel A, Patthy A, Katona G, Graf L. 2005. Extended intermolecular interactions in a serine protease-canonical inhibitor complex account for strong and highly specific inhibition. J Mol Biol 350:156–169.

Franco OL, Dias SC, Magalhães CP, Monteiro AC, Bloch C Jr, Melo FR, Oliveira-Neto OB, Monnerat RG, Grossi-de-Sa MF. 2004. Effects of soybean Kunitz trypsin inhibitor on the cotton boll weevil *Anthonomus grandis*. Phytochemistry 65:81–89.

Fukuoka S, Nyaruhucha CM. 2002. Expression and functional analysis of rat P23, a gut hormone-inducible isoform of trypsin, reveal its resistance to proteinaceous trypsin inhibitors. Biochim Biophys Act 1588:106–112.

Ginalski K, Elofsson A, Fischer D, Rychlewski L. 2003. 3D-Jury: a simple approach to improve protein structure predictions. Bioinformatics 19:1015–1018.

Girard C, Jouanin L. 1999. Molecular cloning of cDNAs encoding a range of digestive enzymes from a phytophagous beetle, *Phaedon cochleariae.* Insect Biochem Mol Biol 29:1129–1142.

Gish W, States DJ. 1993. Identification of protein coding regions by database similarity search. Nat Genet 3:266–272.

Guex N, Peitsch MC. 1997. SWISS-MODEL, the Swiss-PdbViewer: An environment for comparative protein modeling. Electrophoresis 18:2714–2723.

Hanquier J, Sorlet Y, Desplancq D, Baroche L, Ebtinger M, Lefevre JF, Pattus F, Hershberger CL, Vertes AA. 2003. A single mutation in the activation site of bovine trypsinogen enhances its accumulation in the fermentation broth of the yeast *Pichia pastoris.* Appl Environ Microbiol 69:1108–1113.

Helland R, Otlewski J, Sundheim O, Dadlez M , Smalas AO. 1999. The crystal structures of the complexes between bovine beta-trypsin, ten P1 variants of BPTI. J Mol Biol 287:923–942.

Hohenblum H, Vorauer-Uhl K, Katinger H, Mattanovich D. 2004. Bacterial expression and refolding of human trypsinogen. J Biotechnol 109:3–11.

Ishimoto M, Chrispeels MJ. 1996. Protective mechanism of the Mexican bean weevil against high levels of α-amylase inhibitor in the common bean. Plant Physiol 111:393–401.

Jongsma MA, Bakker PL, Peters J, Bosch D, Stiekema WJ.

1995. Adaptation of *Spodoptera exigua* larvae to plant proteinase inhibitors by induction of gut proteinase activity insensitive to inhibition. Proc Natl Acad Sci USA 92:8041–8045.

Kardos J, Bodi A, Zavodszky P, Venekei I, Graf L. 1999. Disulfide-linked propeptides stabilize the structure of zymogen and mature pancreatic serine proteases. Biochemistry 38:12248–12257.

Laemlli UK 1970. Clevage of structural proteins during the assembly of head of bacteriophage T4. Nature 227:680–685.

Li H, Oppert B, Higgins RA, Huang F, Buschman LL, Gao J, Zhu KY. 2005. Characterization of cDNAs encoding three trypsin-like proteinases and mRNA quantitative analysis in Bt-resistant and -susceptible strains of *Ostrinia nubilalis*. Insect Biochem Mol Biol 35:847–860.

Livak KJ , Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2-delta delta CT. Methods 254:402–408.

Lopes AR, Juliano MA, Marana SR, Juliano L, Terra WR. 2006. Substrate specificity of insect trypsins and the role of their subsites in catalysis. Insect Biochem Mol Biol 36:130–140.

Mazumdar-Leighton S, Broadway RM. 2001. Transcriptional induction of digestive midgut trypsin in larval *Agrotis ipsilon* and *Helicoverpa zea* feeding on the soybean trypsin inhibitor. Insect Biochem Mol Biol 31:645–657.

Moon J, Salzman RA, Ahn J-E, Koiwa H, Zhu-Salzman K. 2004. Transcriptional regulation in cowpea bruchid guts during adaptation to a plant defense inhibitor. Insect Mol Biol 13:283–291.

Mulenga A, Sugimoto S, Ingram G, Oshashi K , Misao O. 2001. Characterization of two cDNA encoding serine proteinase from the hard tick *Haemaphysalis longicornis.* Insect Biochem Mol Biol 31:817–825.

Nakajima N, Sugimoto M, Ishihara K. 2003. Earthworm-serine protease: characterization, molecular cloning and application of the catalytic functions. J Mol Catal B Enz 23:191–212.

Oliveira-Neto OB, Batista JA, Rigden DJ, Fragoso RR, Silva RO, Gomes EA, Franco OL, Dias SC, Cordeiro CM, Monnerat RG, Grossi-de-Sa MF. 2004. A diverse family of serine proteinase genes expressed in cotton boll weevil *Anthonomus grandis*: implications for the design of pest-resistant transgenic cotton plants. Insect Biochem Mol Biol 34:903–918.

Patankar AG, Giri AP, Harsulkar AM, Sainani MN, Deshpande VV, Ranjekar PK, Gupta VS. 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.

Pelegrini PB, Murad AM, Grossi-de-Sá MF, Mello LV, Romeiro LAS, Noronha EF, Caldas RA, Franco OL. 2006. Structure and enzyme properties of *Zabrotes subfasciatus* α-amylase. Arch Insect Biochem Physiol 61:77–86.

Prinz WA, Aslund F, Holmgren A, Beckwith J. 1997. The role of the thioredoxin, glutaredoxin pathways in reducing protein disulfide bonds in the *Escherichia coli* cytoplasm. J Biol Chem 272:15661–15667.

Roach JC, Wang K, Gan L, Hood L. 1997. The molecular evolution of the vertebrate trypsinogens. J Mol Evol 45:640–652.

Shan J, Baguinon M, Zheng L, Krishnamoorthi R. 2003. Expression, refolding, and activation of the catalytic domain of human blood coagulation factor XII. Protein Expres Purif 27:143–149.

Silva CP, Xavier-Filho J. 1991. Comparison between the levels of aspartic, cysteine proteinases of the larval midguts of *Callosobruchus maculatus* F., *Zabrotes subfasciatus* Boh. Coleoptera: Bruchidae. Comp Biochem Physiol 99:529–533.

Silva CP, Terra WR, Xavier-Filho J, Grossi-de-Sá MF, Lopes AR, Pontes EG. 1999. Digestion in larvae of *Callosobruchus maculatus* and *Zabrotes subfasciatus* Coleoptera: Bruchidae. with emphasis on α-amylases, oligosaccharidases. Insect Biochem Mol Biol 29:355–366.

Silva CP, Terra WR, Lima RM. 2001a. Differences in midgut serine proteinases from larvae of the bruchid beetles *Callosobruchus maculatus, Zabrotes subfasciatus.* Arch Insect Biochem Physiol 47, 18–28.

Silva CP, Terra WR, Xavier-Filho J, Grossi-de-Sá MF, Isejima EM, DaMatta RA, Miguens FC, Bifano TD. 2001b. Digestion of legume starch granules by larvae of *Zabrotes subfasciatus* (Coleoptera: bruchidae) and the induction of α-amylases in response to different diets. Insect Biochem Mol Biol 31:41–50.

Terra WR, Ferreira C. 1994. Insect digestive enzymes: properties, compartmentalization and function. Comp Biochem Physiol 109:1–62.

Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680.

Tomarelli RM, Charney J, Harding ML. 1949. The use of azoalbumin as a substrate in the colorimetric determination of peptic and tryptic activity. J Lab Clin Med 34:428–433.

Tsybina TA, Dunaevsky YE, Belozersky MA, Zhuzhikov DP, Oppert B, Elpidina EN. 2005. Digestive proteinases of yellow mealworm *Tenebrio molitor*. larvae: Purification and characterization of a trypsin-like proteinase. Biochemistry-Moscow 70:300–305.

Vasquez JR, Evnin LB, Higaki JN, Craik CS. 1989. An expression system for trypsin. J Cell Biochem 39:265–276.

Verheyden G, Volckaert G, Engelborghs Y. 2000 Expression of chymotrypsinogen in the thioredoxin reductase deficient mutant strain of *Escherichia coli* AD494DE3, purification via a fusion product with a hexahistidine-tail. J Chromatogr B Biomed Sci Appl 737:213–224.

Yalcin A, Kanit L, Sozmen EY. 2004. Altered gene expressions in rat hippocampus after kainate injection with or without melatonin pre-treatment. Neurosci Lett 359:65–68.

Zeng F, Zhu Y, Cohen A. 2002. Partial characterization of trypsin-like protease, molecular cloning of a trypsin-like precursor cDNA in salivary glands of *Lygus lineolaris.* Comp Biochem Physiol 131B:453–463.

Zhu YC, Baker JE. 1999. Characterization of midgut trypsin-like enzymes, three trypsinogen cDNAs from the lesser grain borer, *Rhyzopertha dominica* Coleoptera: Bostrichidae. Insect Biochem Mol Biol 29:1053–1063.