Eur. J. Biochem. **262**, 627–636 (1999) © FEBS 1999

Sequencing and characterization of the citrus weevil, *Diaprepes abbreviatus*, trypsin cDNA

Effect of *Aedes* trypsin modulating oostatic factor on trypsin biosynthesis

Xin-Hua Yan[1]*, Hendrik L. De Bondt[2], Charles C. Powell[3], Robert C. Bullock[3] and Dov Borovsky[1]*

[1]*University of Florida-IFAS, Florida Medical Entomology Laboratory, Vero Beach FL, USA;* [2]*Laboratory of Analytical Chemistry and Medicinal Physicochemistry, Katholieke Universiteit, Leuven, Belgium;* [3]*University of Florida-IFAS, IRREC, Fort Pierce, FL, USA*

Trypsin mRNA from the citrus weevil, *Diaprepes abbreviatus*, was reverse transcribed and amplified by PCR. A cDNA species of 513 bp was cloned and sequenced. The 3′ and 5′ ends of the gene (262 bp and 237 bp, respectively) were amplified by rapid amplification of cDNA ends, cloned and sequenced. The deduced sequence of the trypsin cDNA (860 bp) encodes for 250 amino acids including 11 amino acids of activation and signal peptides and exhibited 16.8% identity to trypsin genes of selected Lepidoptera and Diptera. A three-dimensional model of *Diaprepes* trypsin contained two domains of β-barrel sheets as has been found in *Drosophila* and *Neobellieria*. The catalytic active site is composed of the canonical triad of His41, Asp92 and Ser185 and a specificity pocket occupied by Asp179 with maximal activity at pH 10.4. Southern blot analysis indicated that at least two copies of the gene are encoded by *Diaprepes* midgut. Northern blot analysis detected a single RNA band below 1.35 kb at different larval ages (28–100 days old). The message increased with age and was most abundant at 100 days. Trypsin activity, on the other hand, reached a peak at 50 days and fell rapidly afterwards indicating that the trypsin message is probably regulated translationally. Feeding of soybean trypsin inhibitor and *Aedes aegypti* trypsin modulating oostatic factor affected trypsin activity and trypsin biosynthesis, respectively. These results indicate that *Diaprepes* regulates trypsin biosynthesis with a trypsin modulating oostatic factor-like signal.

Keywords: cDNA sequence; trypsin modulating oostatic factor; peptide hormone; trypsin gene; weevil.

Insect growth and development depends on how efficiently food is digested by midgut proteinases [1]. Knowledge and understanding of how insects use proteinases facilitated the development of plant proteinase inhibitors [2], and the use of hormones that control the biosynthesis of trypsin-like enzymes such as, trypsin modulating oostatic factors (TMOF) [3–7]. The research progress of these two approaches can ultimately lead to birational insecticides that will control insect growth and reproduction.

Trypsin is an abundant and important enzyme that is found in insect midguts [8,9]. Lepidopteran gut has an alkaline pH, and thus favors a high level of serine protease activity, e.g. trypsin [10]. In most Coleoptera, however, cysteine proteases are the predominant enzymes that are found in the midgut [11]. However, in some Coleoptera, e.g. the rice weevil (*Sitophilus oryzaes*) and granary weevil (*Sitophilus granarius*) trypsin activity in the midgut is very high [12,13]. Similarly, in the boll weevil (*Anthonomus grandis*) serine proteases appear to be the major enzymes in the midgut [14]. Despite the intense effort that has been applied to the characterization of trypsin-like enzymes in insects, only a few trypsin genes have been fully sequenced and characterized; a few examples are: Lepidoptera, *Manduca sexta* [15], *Choristoneura fumiferana* [16]; Diptera, *Aedes aegypti* [17,18], *Anopheles gambiae* [19], *Neobellieria bullata* [20]. These studies showed that these enzymes are highly conserved and closely related [16,20]. Studies on trypsin expression and its message in Diptera indicated that an early trypsin regulated the transcription of a late trypsin [21] and that juvenile hormone controlled the transcription of the early trypsin [22]. Several models for the activation and the control of trypsin biosynthesis were proposed [23,24]. Although information is available on trypsin activity, gene expression and feeding conditions in haematophagous insects [20,21] little is known about phytophagous species.

TMOF, isolated from the mosquito *A. aegypti* was shown to be the physiological signal that terminates trypsin biosynthesis in mosquitoes and fleshflies [4–7,20]. These observations were viewed as strong evidence that TMOF-like hormones controlled the biosynthesis of trypsin in insects [25]. In *N. bullata* TMOF was shown to exert a translational control on the trypsin gene [20], whereas TMOF analogs reduced trypsin biosynthesis significantly in mosquitoes, biting midges, flies and fleas [4,5,7,20].

The citrus weevil, *Diaprepes abbreviatus*, is an important pest of citrus and other agricultural crops. Thus, we felt that it would be beneficial to explore new environmentally-friendly methods of controlling this pest insect [26]; one such possibility is to shut down its digestive system. To determine whether serine

Correspondence to D. Borovsky, University of Florida-IFAS, FMEL 200 9th St. S.E., Vero Beach, FL 32962, USA. Fax: +1 561 7787205.
Tel.: +1 561 778 7200. E-mail: Dobo@Icon.Vero.Ufl.Edu

Abbreviations: TMOF, trypsin modulating oostatic factor; SBTI, soybean trypsin inhibitor; BApNA, Nα-benzoyl-DL-arginine-4-nitroanilide; RT, reverse transcription; RACE, rapid amplification of cDNA ends.

**Note*: These two authors contributed equally to this paper.

Note: The novel nucleotide sequence published here has been submitted to the GenBank and is available under accession number AF072690.

(Received 30 September 1998, revised 12 February 1999, accepted 1 March 1999)

proteases play a role in food digestion in *D. abbreviatus* a trypsin-like enzyme gene from the larval midgut was isolated and characterized; a three-dismesional model was built and the effects of soybean trypsin inhibitor and TMOF on larval growth and development were studied.

MATERIALS AND METHODS

Insects and diet

D. abbreviatus larvae and pupae were obtained from the USDA Horticultural Research Laboratory (Orlando, FL, USA) and reared at 27 °C on an autoclaved artificial diet [27]. Soybean trypsin inhibitor (SBTI) or TMOF were dissolved in distilled water (10 $mg \cdot mL^{-1}$), adsorbed onto preweighed blocks (2 mg each) to a final concentration of 0.2% (w/w) and 0.04% (w/w), respectively. Diet blocks of controls were treated with water.

Reagents

Nα-benzoyl-DL-arginine-4-nitroanilide (BApNA), tosyl-L-lysine chloromethyl ketone, and SBTI were purchased from Sigma (St Louis, MO, USA). TMOF was synthesized as described previously [4].

Preparation of midgut homogenates

Seven guts each, from different age groups (28, 44, 50, 58, 71, 85 and 100 days old) of *D. abbreviatus* larvae were dissected in physiological saline, immediately transferred into 700 µL of 50 mM Tris/HCl, pH 7.9, containing 5 mM $CaCl_2$, homogenized in a glass homogenizer at 4 °C and centrifuged at 10 000 ***g*** for 20 min at 4 °C. The supernatants (600 µL) were stored at −20 °C until use. First-instar larvae (13 days old) were homogenized whole because the insects were too small for dissection.

Enzyme assay

Trypsin activity was assayed with BApNA as described earlier [28]. For the pH activity profile the following buffers were used: 0.2 M citrate/phosphate (pH 3.2–7.2), 0.2 M Tris/HCl (pH 7.2–8.9) and 0.2 M glycine/NaOH (pH 8.9–10.9).

Feeding SBTI to larvae

Two feeding trials with 24- and 53-day-old larvae were used to determine the effect of feeding SBTI on larval growth and trypsin activity. In trial 1, 20 larvae were fed a diet with SBTI and in trial 2, 10 larvae were fed a diet with SBTI. Each larva was individually reared in a small plastic cup (20 cm^3). Control groups were not fed SBTI. Each larva was weighed before feeding SBTI (Wt_0) and then at 3-day intervals for 12 days. The relative growth rate (G_r) was determined as follows:

$$G_r = [\log_{10}(Wt_e) - \log_{10}(Wt_0)]/D$$

where Wt_e is the weight at the end of each interval, Wt_0 is the weight at the beginning of each interval and D is the length of each interval in days [29]. Trypsin activity in larval guts of control and experimental groups was followed every 3 days (trial 1) and on day 12 (trial 2).

Statistical analysis

Data were analyzed using Student's *t*-test.

Feeding TMOF to larvae

TMOF solution (10 $mg \cdot mL^{-1}$) was adsorbed onto diet blocks at a final concentration of 0.04% (w/w). Each larva (23 or 33 days old) was weighed before feeding and at intervals of 3 days (trials 1–3) or weighed before feeding at days 9 and 11 and at the end of the feeding intervals (trial 4). Guts were dissected out and trypsin activity in each gut (trial 1 and 2) or in groups of guts (3–5 guts per group; trials 3 and 4) was also measured. Controls were fed diet without TMOF.

Topical treatment of larvae with TMOF

Groups of larvae (12–13 per group; 32 and 43 days old) were reared individually in plastic cups containing artificial diet. TMOF (10 $mg \cdot mL^{-1}$) was dissolved in Me_2SO, and was applied onto the cuticle of each larva once a day for 4.0 days. Controls were treated similarly with Me_2SO without TMOF. Group 1 consisted of 32-day-old larvae, whereas group 2 and 3 consisted of 45-day-old larvae which were treated daily with 2.5 or 7.5 µg TMOF, respectively. Each larva was weighed at the beginning of the experiment and at the end of the experiment, and trypsin like activity was measured at the end of each experiment in individual or groups of 3–4 guts.

Genomic DNA extraction

D. abbreviatus genomic DNA was isolated using QIAamp tissue kit (Qiagen). Larvae (50 mg total weight) were homogenized in 180 µL TE buffer (10 mM Tris/HCl, 1 mM EDTA, pH 8.0) and processed according to the manufacturer's instructions. The genomic DNA was resuspended in TE buffer (100 µL) and an aliquot (10 µL) was run on 0.8% agarose gel electrophoresis. Genomic DNA samples of 5–35 kb were stored at −20 °C.

RNA extraction

Midguts (10–20 per group) were dissected from *D. abbreviatus* larvae (28–100 days old) in 150 mM NaCl, pH 7.2. Guts were opened, washed out with 150 mM NaCl, pH 7.2, and homogenized in 1.0 mL Trizol reagent (Gibco-BRL). Gut RNA was prepared as was described by Chomczynski and Sacchi [30], and stored in 100 µL RNase-free TE buffer, pH 8.0. Aliquots containing 400–500 ng RNA were amplified by PCR.

Southern blot analysis

D. abbreviatus genomic DNA was digested with 5 U of *Eco*RI, *Bam*HI or *Eco*RI/*Bam*HI (Gibco-BRL). DNA digests (5 µg per lane) were separated on 0.8% agarose gels and transferred to Hybond-N^+ nylon membrane using 0.4 M NaOH according to the manufacturer's instructions (Amersham) using a Turboblotter (Schleicher & Schuell). The membrane was blocked and hybridized with a purified 432 bp [^{32}P]cDNA-labeled probe (nucleotide 1–432, see Fig. 6) that was prepared using the *redi*-prime DNA-labeling kit [20]. Prehybridization and hybridization was performed at 55 °C with [^{32}P]-labeled cDNA [20]. After washing, the filter was air dried and exposed to X-ray film.

Northern blot analysis

Ambion Northern Max Kit was used for the Northern analysis (Ambion; Austin, TX, USA). Total RNA was extracted in Trizol from larval midgut tissues from different age groups as described above. RNA (15 μg per lane) was separated on denaturing 1.0% formaldehyde agarose gels at 100 V for 1.5 h [20], transferred to Hybond-N$^+$ nylon membrane, hybridized with a 432-bp [^{32}P]-labeled cDNA probe and washed according to the manufacturer's instructions and exposed to X-ray film for 48 h at −80 °C.

PCR primers

Forward and reverse primers were synthesized at the University of Florida DNA synthesis core (Gainesville, FL, USA). The primer sequences (5′ to 3′) and their melting temperatures (t_m) are: DB117 (reverse), CCIGA(AG)TCICC(CT)TG(AG)CA (t_m, 71.5 °C), derived from the consensus active site sequences of *A. aegypti*, *Anopheles gambiae*, *M. sexta* and *Drosophilla melanogaster* [15,18,19,31]. DB111 (forward), AA(AG)A-T(ACT)GTIGGIGG (t_m, 59 °C), DB112 (forward) (AC)GIA-T(ACT)GTIGGIGG (t_m, 56 °C) and DB46 (forward), AA(GA)GA(CT)TCITG(CT)CA(GA)GGIGA(TC) (t_m, 74 °C), derived from the N-terminal sequences and the conserved active site of *Aedes*, *Anopheles*, *Manduca* and *Drosophila*. (RIVGG, KIVGG and KDSCQGD, respectively) [15,18,19,31].

For rapid amplification of the 3′ and 5′ cDNA ends (RACE) the following primers were synthesized and used: dT_{17} adapter (reverse), GACTCGAGTCGACATCGATTTTTTTTTTTTT-TTTT (t_m, 74 °C); adapter (reverse), GACTCGAGTCGA-CATCG (t_m, 66 °C) [32]; DB177 (reverse) GCGAACATTTAATGTAGCAAGTGTGGG (t_m, 77.4 °C); DB178 (reverse), AAATAGACGTAGTGTAACACAGT-GAGCTGC (t_m, 78.8 °C) and DB179 (reverse), AGCTGCAGT-CAGAACTACTCTAGGAGC (t_m, 80.4 °C).

Trypsin cDNA synthesis by PCR

The method was adapted from Borovsky *et al*. [20]. A modified Gene Amp RNA PCR kit (Perkin-Elmer) was used. Briefly, a mixture of 4 μL 25 mM $MgCl_2$, 2 μL 500 mM KCl, 100 mM Tris/HCl, pH 8.3, 6 μL sterile distilled water, 4 μL dNTP (10 mM each of dATP, dTTP, dCTP, and dGTP), 1 μL RNase inhibitor (20 U), 1 μL reverse transcriptase (50 U) and 1 μL 15 μM primer DB117 was prepared. To each reaction tube, 19 μL of this mixture was added. Template RNA (400 ng in 1 μL) was added and the samples were overlaid with 50 μL light mineral oil. Reverse transcription (RT) was performed in a DNA thermal cycler 480 (Perkin-Elmer) at 42 °C for 15 min, 99 °C for 5 min and 5 °C for 5 min. After RT, 29 μL 50 mM KCl, 10 mM Tris/HCl, pH 8.3, Ampli*Taq* DNA polymerase (2.5 U), 1 μL primer DB111 or DB112 were added to each reaction. PCR was carried out as follows: denaturation for 3 min at 95 °C (1 cycle), annealing for 5 min at 48 °C and extension for 40 min at 60 °C (1 cycle), followed by denaturation for 1 min at 95 °C, annealing for 1 min at 48 °C and extension for 3 min at 60 °C (40 cycles); the final cycle extension was for 15 min at 60 °C. After PCR, tubes were incubated at 4 °C. Amplified cDNA was recovered and stored at −20 °C.

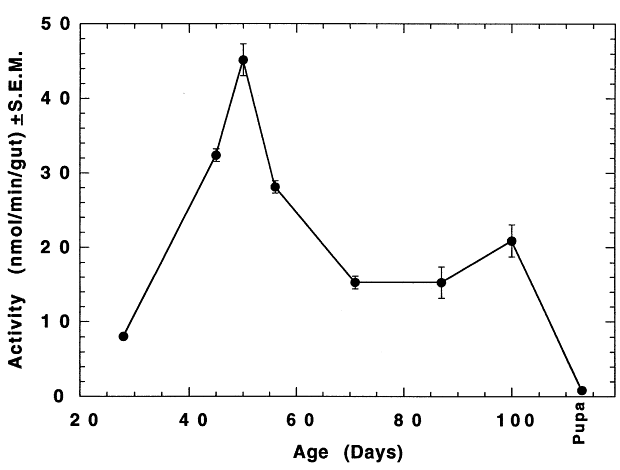

Fig. 1. Trypsin activity in the gut of *D. abbreviatus* during different developmental stages of the larva and pupa. *D. abbreviatus* (eight per group) were fed a diet and at different stages during larval development: I (28 days), II (45 days), III (50 days), IV (58 days), V (72 days), VI (87 days),VII (100 days) and pupal stage (over 100 days). Guts were removed and assayed for trypsin activity using BApNA. Results are an average of three determinations and are expressed as trypsin activity (nmol·min^{-1}) at 410 nm per gut ± SEM.

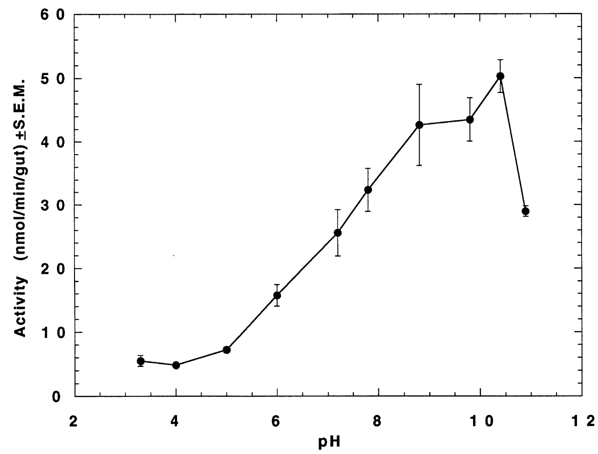

Fig. 2. The effect of pH on trypsin-like activity in the midgut of *D. abbreviatus* larva. Three groups (seven guts per group) of 43-day-old larvae were assayed for trypsin-like activity using BApNA at pH 3.2–11 in the presence of citrate/phosphate buffer (pH 3.2–7.2), Tris/HCl (pH 7.2–8.9), glycine/NaOH (pH 8.9–11). Results are expressed as trypsin activity (nmol·min^{-1} per gut) at 410 nm, and are an average of three determinations ± SEM.

RACE of trypsin cDNA 3′ and 5′ ends

The initial mixture (20 μL per sample) for the 3′ RACE was the same as described above for the RT-PCR except that the dT_{17} adapter (15 μM) was used. PCR conditions were as follows: 10 min at 24 °C, 1 h at 42 °C and 30 min at 52 °C, followed by

Table 1. Feeding of *D. abbreviatus* larvae with a diet containing SBTI. Groups of 24-day-old *D. abbreviatus* larvae (5–6 per group) were fed a diet containing SBTI (0.2%, w/w) for the number of days indicated and trypsin activity was measured at different intervals after feeding [20]. Control larvae were fed a diet without SBTI. Results are the means of two determinations.

STBI-containing diet for	Trypsin stimulation (%)	Trypsin inhibition (%)
3 days	14	0
6 days	21.6	0
9 days	11.5	0
12 days	43.7	0
12 days (53-day-old larvae)	0	47

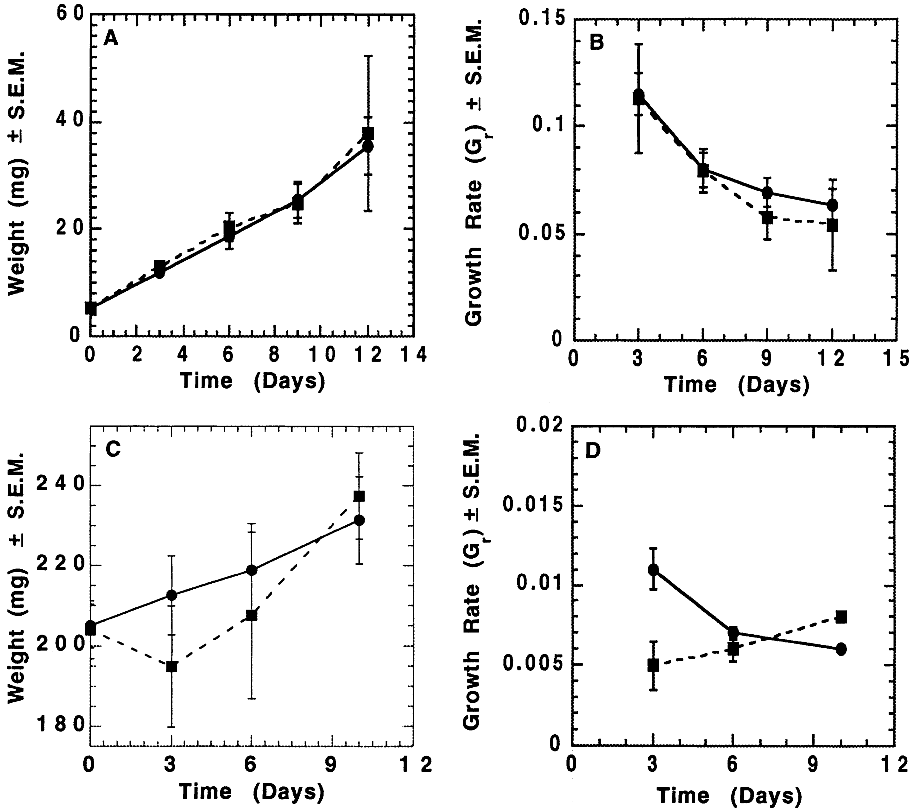

Fig. 3. The effect of feeding SBTI on growth rate (G_r) and weight. *D. abbreviatus* larvae 24 days old (A and B; $n = 20$) and 53 days old (C and D; $n = 10$) were fed 0.2% SBTI (■) and at different intervals after feeding the mean weight and growth rate (G_r) were assayed. Controls (●) were fed a diet without SBTI. Results are expressed as mean ± SEM of 20 determinations (A and B) or 10 determinations (C and D) of weight increase or growth rate (G_r).

denaturation at 99 °C for 5 min and incubation at 5 °C for 5 min. The reaction mixture was then diluted to 1.0 mL with Tris/EDTA (10 mM/1 mM), pH 8.0, and this solution was stored at −20 °C. PCR was carried out by means of the hot-start method using Ampli*Taq* Gold. Each reaction tube contained 4 µL 25 mM $MgCl_2$, 5 µL 500 mM KCL, 100 mM Tris/HCl, pH 8.3, 4 µL dNTP, 1 µL DB46 (15 µM), 1 µL adapter (15 µM), 29.5 µL sterile distilled water, 0.5 µL Ampli*Taq* Gold (2.5 U) and 5 µL of cDNA. To each tube a wax nugget (Perkin-Elmer) was added. PCR was carried out as follows: 3 min at 95 °C, 5 min at 55 °C and 40 min at 72 °C (1 cycle), denaturation at 95 °C for 1 min, annealing at 55 °C for 1 min and extension at 72 °C for 3 min (40 cycles). The final extension was at 72 °C for 15 min. After PCR, the reaction tubes were incubated at 4 °C. Samples were recovered and stored at −20 °C.

The initial RT-PCR mixture (20 µL/sample) for the 5′ RACE was the same as that described above for the 3′ RACE except that primer DB177 (15 µM) was used. The template RNA was denatured at 76 °C for 5 min before it was added to the reaction mixture. The RT-PCR conditions were as described above for the 3′ RACE. After incubation, the ssDNA was purified on QIAquick column (Qiagen) and concentrated by Speed Vac to 10 µL.

Polyadenylation of the ssDNA was carried out in a reaction mixture containing 10 µL ssDNA, 4.9 µL sterile distilled water, 4.0 µL tailing buffer (0.5 M potassium cacodylate, 10 mM $CoCl_2$, 1 mM dithiothreitol) (Gibco-BRL), 0.4 µL dATP and 0.7 µL terminal deoxynucleotide transferase (10.5 U) (Gibco-BRL) at 37 °C for 5 min, followed by 65 °C for 5 min. After incubation, the reaction mixture was diluted to 60 µL and stored at −20 °C. The polyadenylated ssDNA was amplified by PCR in a total volume of 50 µL: 4 µL 25 mM $MgCl_2$, 5 µL 500 mM KCl, 100 mM Tris/HCl, pH 8.3, 4 µL dNTP, 29.3 µL sterile distilled water, 1 µL 15 µM DB157 and 0.2 µL 15 µM DB158 forward primers, 0.5 µL Ampli*Taq* Gold (2.5 U), 1 µL 15 µM DB179 reverse nested primer and 5 µL polyadenylated ssDNA. PCR was carried out as follows: 95 °C for 5 min, 37 °C for 5 min and 72 °C for 40 min (1 cycle), denaturation at 95 °C for 1 min, annealing at 55 °C for 1 min and extension at 72 °C for 3 min (35 cycles); the final extension was for 15 min at 72 °C. After PCR, tubes were incubated at 4 °C. Samples were recovered and stored at −20 °C.

Table 2. Effect of topical treatment of TMOF on *D. abbreviatus* larval weight gain and trypsin-like activity. *D. abbreviatus* larvae were treated with Me_2SO/TMOF (2.5 $\mu g \cdot day^{-1}$ for 4.0 days). Controls were treated with Me_2SO. Each larva was weighed before and 4.5 days after treatments. Trypsin-like activity was checked in each gut 4.5 days after the first treatment. Results are expressed as mean of 12 determinations ± SEM.

	n	Weight (mg/larva)	Trypsin activity ($nmol \cdot min^{-1}$ per gut)
Before TMOF treatment	12	12.85 ± 2.25[a]	ND
4.5 days after TMOF treatment	12	15.58 ± 2.47[b]	3.7 ± 0.9[c]
Before Me_2SO treatment (control)	12	14.24 ± 1.98[a]	ND
4.5 days after Me_2SO treatment (control)	12	23.63 ± 3.95[b]	6.2 ± 1[c]

[a] Not significant $P > 0.05$. [b] Significant difference $0.005 < P < 0.01$. [c] Significant difference $0.025 < P < 0.05$.

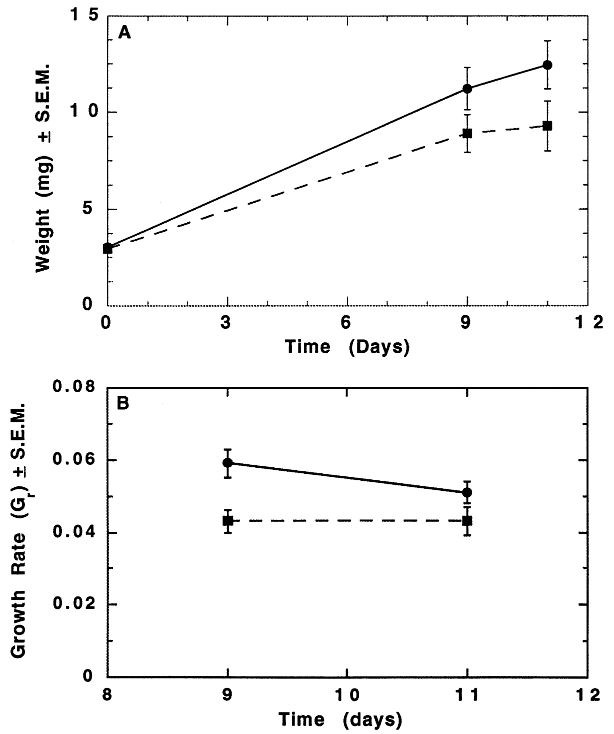

Fig. 4. The effect of feeding TMOF on the weight gain and growth rate (G_r). *D. abbreviatus* larvae (n = 29; 28 days old) were fed TMOF (0.04%) with their diet for 9 and 11 days. Larval mean weight (A) and growth rate (B) were monitored 9 and 11 days later. Results represent the mean of 29 determinations ± SEM The mean weight increase after 11 days (A) and the growth rate after 9 days (B) were significantly different ($P < 0.05$) between controls (●) and TMOF-fed insects (■).

Cloning, sequencing and analysis of PCR products

The cDNA species after PCR amplification of gut RNA were subcloned into pCR2.1 vector by means of TA cloning kit (Invitrogen, Inc.). Plasmid DNA was purified after lysis of the bacterial cells using QIAprep Spin Miniprep Kit. Plasmid DNA was digested with *Eco*RI (5 U) and analyzed by electrophoresis. Plasmids that contained inserts were sequenced by the dideoxynucleotide chain termination method [33] with [α^{35}S]dATP[S] and the enzyme sequenase (version 2.0; U.S. Biochemicals) [34]. Sequences were analyzed with MacDNASIS v 3.7 (Hitachi Software Engineering).

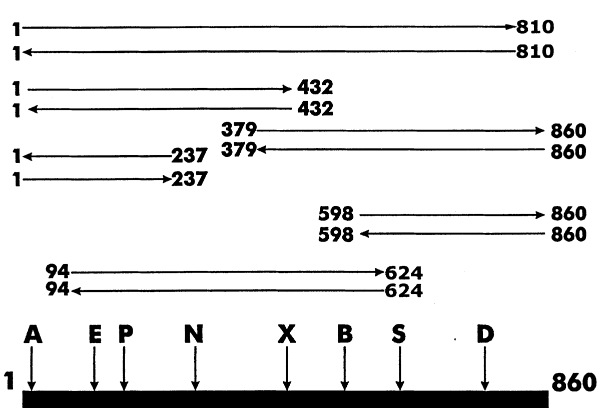

Fig. 5. Partial restriction map (bottom bar) and sequencing strategy (horizontal arrows) of the 860 bp trypsin of *D. abbreviatus*. Several overlapping PCR fragments (horizontal arrows) were amplified, cloned and sequenced as indicated. Restriction endonucleases: A, *Asu*II; B, *Bsty*I; D, *Dde*I; E, *Eco*RII; N, *Nsp*II; P, *Pst*I; S, *Sna*BI; X, *Xmn*I.

Three-dimensional model

A three-dimensional ribbon model of *D. abbreviatus* (L) trypsin was built using SYBYL molecular-modeling software (version 6.3) and composer as was carried out previously for *Neobellieria* and *Drosophila* trypsins [20]. Briefly, the model was aligned with 18 homologous chains in the Brookhaven protein data bank using a gap penalty of eight and homology matrix [20]. The topologically equivalent amino acids of the structurally conserved regions were iteratively refined by weight-least-square fit. The loops of the structurally variable regions were identified and the coordinates adjusted to the model. The model was refined by scanning the side-chain torsions to relieve weak van der Waals' contacts, and the energy of the backbone and the entire protein was minimized [20]. The three-dimensional structure was converted into a ribbon representation by the program MOLSCRIPT [35].

RESULTS

Trypsin-like activity in the midgut of *D. abbreviatus*

Trypsin-like activity was measured at 30 °C with BApNA during the larval pupal stages of *D. abbreviatus*. The highest level of trypsin-like activity was observed in the larval stage, and traces of trypsin-like activity were detected at the pupal stage (Fig. 1). During larval development, trypsin-like activity increased fivefold from the early larval instar (28 days old) to the third larval instar (50 days old). The activity declined 2.5-fold and then reached a low plateau between the fifth and seventh instar (71–100 days old) and fell to a minimum at the pupal stage (Fig. 1). The increase and decline in trypsin-like activity indicates that trypsin biosynthesis is regulated.

pH activity profile

Maximum activity of *D. abbreviatus* trypsin-like enzyme at different pH values occurred at pH 10.4 (Fig. 2). These results indicate that trypsin catalysis in *D. abbreviatus* favors the gut alkaline environment and is similar to trypsins from *Helicoverpa armigera* [36], *S. littoralis* [37] and *Heliothis virescens* [38].

The effect of feeding SBTI on trypsin-like activity and larval growth

Feeding 24-day-old larvae for 3, 6, 9 and 12 days with 0.2% (w/w) SBTI in the diet caused 14%, 21.6%, 11.5% and 43.7% stimulation of trypsin-like activity, respectively, as compared with controls that were fed a diet without SBTI (Table 1). Feeding SBTI to 53-day-old larvae for 12 days decreased trypsin-like activity by 47%, indicating that in 53-day-old larvae trypsin biosynthesis is probably down-regulated which prevented overstimulation of the trypsin gene. The mean weight gain of 24-day-old larvae was not affected (Fig. 3A). The relative rate of larval growth decreased slightly when larvae were fed SBTI for 9–12 days (Fig. 3B). When 53-day-old larvae were fed 0.2% SBTI, a decrease in larval mean weight was observed on days 3 and 6 but the decrease was

```
ATG CCA AGC AGA GAC TTC GAA AAT ACT AAT GGT ACC GTC TTT TAT CAT CCA ACT    54
 M   P   S   R   D   F   E   N   T   N   G   T   V   F   Y   H   P   T    18

ATT GGA GGT CGA ATT GTT GGT GGA GTT GCA ACT ACT ATT CAG GAC TTA CCC TGG   108
 I   G   G   R   I   V   G   G   V   A   T   T   I   Q   D   L   P   W    36
                +1
CAA GTA GCT ATA CTT CGT AAT GGA GCT CAA ATC TGT GGA GGA ATA CTC GTT GCT   162
 Q   V   A   I   L   R   N   G   A   Q   I   C   G   G   I   L   V   A    54

CCT AGA GTA GTT CTG ACT GCA GCT CAC TGT GTT ACA CTA CGT CTA TTT CCC ACA   216
 P   R   V   V   L   T   A   A   H   C   V   T   L   R   L   F   P   T    72

CTT GCT ACA TTA AAT GTT CGC ACT GGA TCA ACC ACT CAT AAT GCA GGA GGT ACC   270
 L   A   T   L   N   V   R   T   G   S   T   T   H   N   A   G   G   T    90

CGT GTA GCT GTC AGC AGC AGA ATC CTA CAT GCT CAA TAC CAA GAC TGC GAA ACC   324
 R   V   A   V   S   S   R   I   L   H   A   Q   Y   Q   D   C   E   T   108

TGT TCT CCC GAT TAC GAC ATT GCT GTT CTT CAT CTT GCT GCT AAT GCC AAT ATA   378
 C   S   P   D   Y   D   I   A   V   L   H   L   A   A   N   A   N   I   126

TCT CCT GCT GCG ACC ATT GCC CTC TGG GAC GAT AAT ACT GCC TTT GCT GCT GGT   432
 S   P   A   A   T   I   A   L   W   D   D   N   T   A   F   A   A   G   144

GTA GTT GGA ACT GTT TCA GGA TGG GGT GCT ACA AGC GAA GGT GGT GCT GGT TCC   486
 V   V   G   T   V   S   G   W   G   A   T   S   E   G   G   A   G   S   162

GTA ACA TTA AGA CGT GTA GAT GTA CCA GTT ATT GGT AAC GTT CAG TGT CGT AAT   540
 V   T   L   R   R   V   D   V   P   V   I   G   N   V   Q   C   R   N   180

GTA TAT GGA TCT ATA ATT ACC ACG AGA ACA ATT TGT GCT GGA TTA GCT CAA GGA   594
 V   Y   G   S   I   I   T   T   R   T   I   C   A   G   L   A   Q   G   196

GGT AGA GAC TCA TGC CAA GGA GAC TCT GGT GGT CCA TAC GTA ATT CAA AAC AGA   648
 G   R   D   S   C   Q   G   D   S   G   G   P   Y   V   I   Q   N   R   214

TTG GCT GGT ATT GTA TCA TTT GGA GCT GGT TGC GCT AGA GCT GGC TTG CCC GGT   702
 L   A   G   I   V   S   F   G   A   G   C   A   R   A   G   L   P   G   232

GTC TAT GCA AGT ATT CCC GGA TAC AGA GCT TGG ATC AGA CAA AAC GCT GGA CTT   756
 V   Y   A   S   I   P   G   Y   R   A   W   I   R   Q   N   A   G   L   250

TAA GTT AAC AAG AAT CCA AAG GAT GTT TTA CTG AGA ACA AAT TTC CTG TTG TCT   810

AAT AAA AAA TAT TAA AAA TAA ATG CAC TAA TGA AAA AAA AAA AAA AAA AA        860
```

Fig. 6. Nucleotide sequence and predicted amino acid sequence of *D. abbreviatus* trypsin. Predicted site of cleavage of the signal peptide and activation peptide are denoted by arrows. The conserved sequence around Ser185, and the two polyadenylation signals are underlined by single and double lines, respectively. The position of the catalytic triad residues is denoted by closed triangles and the positions of the eight cysteines are marked by closed circles. Numbering of the mature enzyme starts at the isoleucine residue (+1) after the cleavage site of the activation peptide.

not significant. The inhibitory effect decreased with time and larval weight at day 10 was similar to that of control larvae (Fig. 3C). The same pattern was also observed with larval relative growth rate: this was twofold lower during the first 3 days of treatment in 53-day-old larvae and disappeared on days 6 and 10 (Fig. 3D). These results indicate that SBTI is more effective in inhibiting trypsin activity in older larvae.

The effect of feeding TMOF on larval growth and trypsin-like activity

Larval mean weight decreased by feeding larvae with a diet containing 0.04% (w/w) TMOF. Four feeding trials were run: in trials 1–3, larval mean weight and relative growth rate were lower when larvae were fed 0.04% (w/w) TMOF as compared with controls that were fed a regular diet; however, the differences were not statistically significant ($P > 0.05$). In trial 4, significant decrease in mean weight was observed on day 11 ($P < 0.05$, $n = 29$, d.f. = 28) (Fig. 4A) and significant decreases in growth rate were observed on days 9 and 11 ($P < 0.003$, $n = 29$, d.f. = 28) (Fig. 4B). The results from these feeding experiments show that TMOF affects trypsin biosynthesis in the gut of *D. abbreviatus* larva.

The effect of topical application of TMOF on larval growth and trypsin-like activity

When 32-day-old larvae were treated topically with 10 μg TMOF (2.5 μg per day for 4.0 consecutive days), trypsin-like activity and larval weight gain were 40% and 70%, respectively, lower than in controls (Table 2). The lower weight gain and lower trypsin-like activity were significant as was shown by Student's *t*-test (Table 2). When 45-day-old larvae were treated with 10 μg TMOF (2.5 μg per treatment for 4 consecutive days), larval mean weight was not affected; however, trypsin activity decreased by 19.5% (results not shown). Topical application of 45-day-old larvae with 30 μg TMOF (7.5 μg per treatment for 4 consecutive days) caused a decrease in larval mean weight gain and decreased trypsin-like activity in the gut by 65% (data not shown). These results indicate that TMOF affects trypsin biosynthesis in the gut when it is adsorbed through the cuticle into the hemolymph as was shown in mosquitoes and other dipterans.

D. abbreviatus-cDNA synthesis and characterization

Three fragments of cDNA (262 bp, 273 bp and 513 bp) of *D. abbreviatus* trypsin were synthesized from mRNA extracted from midguts of 25-day-old larvae. A 513-bp trypsin cDNA was

Fig. 7. Ribbon drawing, made by MOLSCRIPT, of *D. abbreviatus* trypsin. Two antiparallel β-barrel domains, β-strands in domain 1 (red) and β-strands in domain 2 (green) are shown. The side chains of the catalytic-triad residues and of Asp179 in the specificity pocket is shown in blue. The C-terminal (C) helix is in domain 1 and is shown in purple and the N-terminus (N) is in domain 2 and is shown in green. Coils that connect the two domains are shown in yellow.

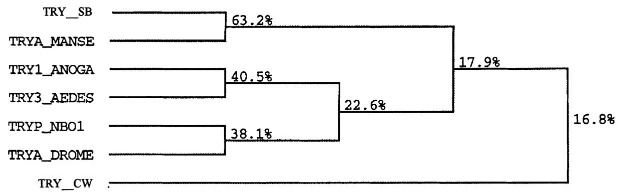

Fig. 8. Phylogentic tree based on multiple-sequence alignment according to the Higgins–Sharp algorithm (CLUSTRAL 4) of MacDNASIS 3.7. The branching-order similarities (%) are based on the sequences shown in [20].

reverse transcribed using downstream primer DB117 and amplified by PCR using upstream primer DB111 or DB112. The 513 bp amplified cDNA was separated by electrophoresis on an agarose gel (2%) and stained with ethidium bromide and purified. PCR with primer DB117 alone indicated that upstream primers DB111 and DB112 did not hybridize to the N-terminal of the trypsin gene and did not recognize *D. abbreviatus* trypsin mRNA. A 5′ RACE was used to amplify a 273-bp cDNA at the 5′ end using primer DB117 and primer DB179 or DB178. A 262-bp trypsin cDNA fragment at the 3′ end was amplified by PCR with primer pair DB46/adapter using 3′ RACE [32]. The 273 bp and 262 bp cDNA fragments at the 5′ and 3′ ends were separated by electrophoresis on agarose gel (2%), stained with ethidium bromide, eluted from the agarose by centrifugation using Amicon preparatory spin columns and further purified on QIAquick columns (Qiagen).

Cloning and sequencing of *D. abbreviatus* trypsin cDNA

The three cDNA fragments (262 bp, 273 bp and 513 bp) were subcloned and sequenced. The sequencing strategy and partial restriction map are shown in Fig. 5. The nucleotide sequence and deduced amino acid sequence are shown in Fig. 6. The 860 bp cDNA nucleotide sequence encodes for a 250-amino acids open reading frame that has a methionine codon at position 1. The N-terminal sequence is hydrophobic and represents a signal peptide with a cleavage site after Gly11

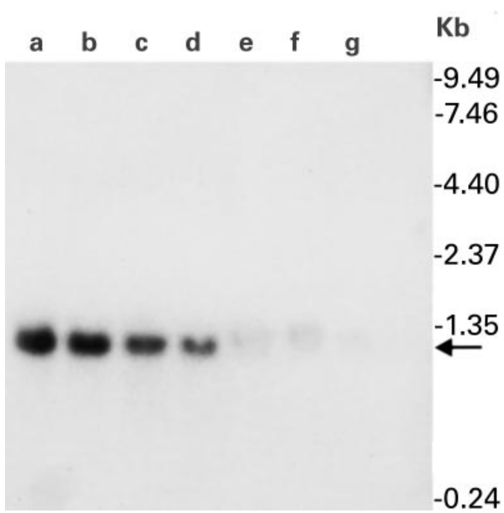

Fig. 9. Northern blot analysis of gut trypsin mRNA from *D. abbreviatus*. Guts were removed from larvae of different ages and RNA (15 µg per lane) was analyzed by Northern blot. Age of larval gut in lane: a, 100 days; b, 87 days; c, 70 days; d, 56 days; e, 50 days; f, 45 days; g, 28 days. The sizes of RNA markers are given on the right and *D. abbreviatus* trypsin mRNA is indicated by an arrow.

[39], which upon cleavage would release a trypsinogen with an 11-amino acid activation peptide. Cleavage after Arg22 would generate active trypsin (Fig. 6). The mature enzyme (residues 23–250) has 228 amino acids and eight cysteines that could form four cysteine bridges. Two consensus polyadenylation signals (AATAAA) are found at positions 811 and 826, and a poly A tail begins at position 844 (Fig. 6).

Multiple sequence and three-dimensional structure analyses

The deduced amino acid sequence of *D. abbreviatus* trypsin was compared with the amino acid sequences from several other insects found in the National Biomedical Research Foundation Protein identification Resource database. The Higgin–Sharp alogarithm (CLUSTAL4) of MacDNASIS 3.7 was used in the multiple-sequence alignment against several other serine proteinases. The N-terminal sequence IVGG is conserved in all the sequences except that of *N. bullata* in which Asn is replaced by Gly [20]. The catalytic active center containing residues His41, Asp92 and Ser185 was highly conserved in other serine proteinases [15,16,18–20,31]. The specificity pocket sequence is RDSC for *D. abbreviatus* and RDQC for *M. sexta* and *C. fumiferana*. In Diptera, the sequence was either KDSC for *A. aegypti*, or KDAC for *A. gambiae*, *D. melanogaster* and *N. bullata* [18,19,31]. Asp179 lies at the bottom of the specificity pocket (Fig. 7) and is conserved in *C. fumiferna*, *M. sexta*, *A. gambiae*, *A. aegypti* and *D. melanogaster*. The 10 amino acids around Ser185 are conserved in most of the sequences examined and form a coil as part of the active site (Fig. 7). Asp179 sits at the bottom of the coil and Ser185 at the top and the oxygenation hole is located below Ser185 in published models of *Neobellieria* and *Drosophila* trypsins [20] and in this model (Fig. 7). A single helix is found at the C terminus. The active site is situated in a crevice between two antiparallel-β-barrel type domains. Domain 1 contains two residues of the catalytic triad, His41 and Asp92, whereas Ser185 is in the second domain (Fig. 7). These features are also similar in the *Neobellieria* and *Drosophila* models [20] and have been reported in models of other trypsins and chymotrypsins which are based on X-ray-diffraction data [40,41]. The model predicts that only three pairs of disulfide bonds form between Cys26 and Cys42, Cys156 and Cys170 and Cys181 and Cys205. The model predicts that Cys84 and Cys87 do not form a disulfide bond because they do not interact (they are further than 1.0 nm from each other). A phylogenetic tree, based on the homology of the amino acid sequences among seven insects (Fig. 8) indicates that *D. abbreviatus* trypsin has only 16.8% identity to that of Lepidoptera and Diptera species (Fig. 8).

D. abbreviatus trypsin mRNA levels during larval developmental stages

Northern analysis was carried out to determine trypsin mRNA expression level during larval development. Total RNA was extracted from midgut tissues of several different age groups. A single RNA band below 1.35 kb was detected in all seven age groups (Fig. 9). The level of the trypsin transcript increased with age from 28-day-old larvae to 100-day-old larvae (Fig. 9), whereas trypsin-like activity was highest in 50-day-old larvae and 2.5-fold lower in 100-day-old larvae (Fig. 1). These results indicate that in older larvae there is accumulation of trypsin transcript that is not translated.

Genomic organization of *D. abbreviatus* trypsin gene

Southern blots of *D. abbreviatus* genomic DNA with several restriction enzymes (*Eco*RI, *Xba*I, *Xbo*I, *Bam*HI) revealed two bands of 3 and 4 kb with *Eco*RI and one band with the other enzymes (4–7 kb) suggesting the existence of two trypsin genes in *D. abbreviatus* (data not shown). Several trypsin genes have been identified in mosquitoes [18,19]. PCR amplification of the genomic DNA with two primers that amplified the gene from the methionine start signal to the polyA adenylation signal (Fig. 6) amplified only one band of 810 bp (data not shown) which indicates that if there are two trypsin genes they probably have similar sizes lacking long introns – very short introns cannot be identified on a 0.8% agarose gel.

DISCUSSION

A trypsin-like enzyme of *D. abbreviatus* was characterized and its role in protein digestion was studied. Trypsin-like activity was found in all the developmental stages of *D. abbreviatus* larvae indicating that the enzyme is involved in food digestion (Fig. 1). The trypsin-like enzyme was most active at an alkaline pH of 10.4 (Fig. 2) indicating that the pH of *D. abbreviatus* gut is probably alkaline as shown for Lepidoptera and mosquitoes. Trypsin-like activity has been detected previously in several weevils, such as the rice weevil (*S. oryzae*), granarius granary weevil (*S. oryzea Summers*), and sweet potato weevil (*Cylas formicarius elegantus*) [13,41,42]. The major proteinase of the boll weevil (*Anthonomus grandis*) was found to be a trypsin-like enzyme, with maximal midgut activity at pH 10–11 [14]. McGhie *et al*. [43] detected trypsin-like activity in three soil-dwelling grubs which fed on plant roots, suggesting that food and living environment possibly played a critical role in the insect's digestive enzyme development. The results of this study show that the root feeder *D. abbreviatus* also uses trypsin-like enzymes in its digestive system. Trypsin-like activity in the midgut of *D. abbreviatus* larva was age related (Fig. 1). These results confirm previous reports that proteinase activity changed with larval development [44]. The trypsin-like activity profile of *D. abbreviatus* did not follow those of *Lutzomyia anthophora* [45] or *Culex pipiens* [46]. Trypsin-like activity in these insects was correlated positively with larval age and increased throughout the larval developmental stages, whereas trypsin-like activity in *D. abbreviatus* larva increased with age up to 50 days and then declined and reached a minimum in the pupal stage (Fig. 1). Minimal trypsin activity was also reported for mosquitoes and sand flies during the pupal stage [45,46].

The effect of SBTI on insect growth and trypsin activity has been reported for several phytophagous insects [47]. McManus and Burgess [29] proposed that age should be considered as a factor for the evaluation of proteinase inhibitors as potential insect control agents. Compared with older larvae, younger larvae were more sensitive to SBTI [29]. This may imply a different digestive physiology and adaptation to the inhibitor in older larvae. Our results indicated that feeding younger larvae (24 days old) with SBTI caused trypsin overproduction and a slight decrease in growth (Fig. 3 and Table 1). Similarly, trypsin activity increased twofold in 13-day-old larvae when they were fed SBTI (data not shown). These results support Broadway and Duffey's [44, 48] observation that proteinase inhibitors cause overproduction of trypsin. However, feeding SBTI to older *D. abbreviatus* larvae (53 days old) caused a 46% decrease in trypsin-like activity indicating that other proteinases in older larvae assume the role of trypsin-like enzyme after feeding them SBTI. This adaptive mechanism negated the effect of SBTI, and no significant growth inhibition was observed (Fig. 3 and Table 1). Similarly, Purcell *et al*. [14] and Wu *et al*. [49] reported that *in vivo*, SBTI had little effect on boll weevil and *Helicoverpa armigera* larval growth, respectively.

TMOF is the physiological signal that terminates trypsin biosynthesis in mosquitoes, flies, fleas and biting midges [4,5,50]. Trypsin-like activity, larval weight and growth significantly decreased after topical treating or feeding *D. abbreviatus* larvae with TMOF (Fig. 4 and Table 2). These results imply that TMOF-like hormones probably regulate trypsin-like enzyme biosynthesis in *D. abbreviatus*. Thus, the use of TMOF-like factors to control food digestion in *D. abbreviatus* larva is a better approach because these factors affect trypsin biosynthesis and not trypsin activity. Furthermore, proteinase inhibitors which inhibit trypsin activity may cause an over stimulation of the gene, as was shown in 24-day-old larvae (Table 1), or increase the level of activity of proteolytic enzymes that were not susceptible to inhibition by SBTI [51].

To study the properties of *D. abbreviatus* trypsin a full-length cDNA that encodes trypsin was sequenced. As many trypsins from mammals and insects exhibit sequence similarity around the serine residue in the active-site pocket and the N-terminus [17–19,31,52] (Fig. 7), degenerate primers from these two regions were designed and used to amplify a trypsin cDNA from *D. abbreviatus*. A cDNA band of 530 bp was identified by agarose gel electrophoresis. The 530 bp cDNA was initially considered to be too short to be trypsin cDNA, because in most insects and mammals there are approximately 600 nucleotides between the N-terminus and the serine at the active site. Subcloning and sequencing of the cDNA showed that the downstream primer (DB117) hybridized 94 nucleotides upstream of the methionine start signal (Figs 5 and 6). The sequence obtained from the clone was used to construct primers to be used for RACE [32] and to amplify the cDNA up to the poly(A) tail. Thus, a complete cDNA sequence was obtained (Figs 5 and 6). The deduced amino acid sequence encodes a preproenzyme of 250 amino acids ($M_r \approx 25\,000$) similar to trypsin from *N. bullata* [20] and from *M. sexta* [15]. The enzyme contains a string of hydrophobic amino acids characteristic of signal peptides [53] (Fig. 6). The signal peptide has a cleavage site between Gly11 and Thr12 [39] cleavage of which would liberate a trypsinogen with an 11 amino-acid activation peptide. In *M. sexta* a seven amino-acid activation peptide was reported and in *A. aegypti*, *D. melanogaster* and *N. bullata* 10–6 amino-acid activation peptide was suggested [15,17,20,31].

D. abbreviatus trypsin has an even number of cysteines (eight) as was shown in *Bombyx mori* [54] in contrast with other reports that showed an odd number of cysteines (nine) in *N. bullata*, *C. fumiferana*, and two trypsins from *M. sexta* and seven cysteines in the early trypsin of *A. aegypti* [15,18,20]. A three-dimensional model (Fig. 7) predicts that only three pairs of disulfide bonds between Cys26 and Cys42, Cys156 and Cys168, and between Cys179 and Cys203. The distance between Cys84 and Cys87 is 12 Å which precludes a stable disulfide bond (2–3 Å; Fig. 7). The model exhibits the antiparallel β-barrel characteristic of the chymotrypsin superfamily that includes trypsin, chymotrypsin, elastase and thrombin [41] and was also shown for *Neobellieria* and *Drosophila* trypsins [20]. The theoretical isoelectric point (8.28) of *D. abbreviatus* trypsin compared with that of *N. bullata* trypsin (6.97) [20]; this may indicate that the pH in the gut of *D. abbreviatus* is basic and is probably close to the optimum pH of 10.4 that is required for maximal enzymatic activity (Fig. 2). *D. abbreviatus* trypsin has positively charged amino acids (15 arginines and four histidines) and negatively charged amino acids (nine asparagines, eight cysteines, two glutamic acids and six tyrosines).

Multiple-sequence alignment shows that the amino acid sequence of *D. abbreviatus* trypsin has a highly conserved sequence at the histidine and serine catalytic sites. The three amino acids GlnGlyAsp precede Ser185 at the active site and Asp92 and His41 complete the catalytic triad (Fig. 7). The aspartic acid residue that determines the enzyme specificity is found at position 179 of the mature enzyme (Fig. 7).

Northern blot analysis of guts that were removed from larvae of different ages showed that the trypsin mRNA in the midgut increased with larval age and reached a maximum at 100 days (Fig. 9). Trypsin-like enzyme activity in the midgut increased only during the early stages reaching a maximum at 50 days and then rapidly decreased (Fig. 1). These results indicate that trypsin mRNA is controlled translationally, probably by TMOF-like hormones as was shown in *N. bullata* [20]. In some haematophygous insects, the early trypsin mRNA was shown to be under a translational control [55]. Synthesis of other proteins has been shown to be regulated at the translational level including the ferritin/transferrin receptor [56], ribonucleotide reductase [57], heat-shock proteins [58], tubulin [59] and oocyte proteins [60]. Although late larvae might consume less food the abundant message in the late larval stages indicates that trypsin is probably under a translational control in *D. abbreviatus*.

PCR of *D. abbreviatus* genomic DNA using primers DB185 and DB182 (positions 1–810; Fig. 6) amplified one DNA band of 810 bp indicating that the *D. abbreviatus* trypsin gene probably does not have introns. Davis *et al*. [31] also reported that the trypsin gene in *Drosophila* does not have introns. Two introns have been reported in the *C. fumiferana* trypsin gene, and one in *A. aegypti* early trypsin [16,61]. Digestion of the genomic DNA by *Eco*RI produced two bands, possibly indicating that in *D. abbreviatus* there are at least two trypsin genes; however, point mutations or pseudogenes might also contribute to this observation [62].

ACKNOWLEDGEMENTS

This work was supported by grants from NATO CRG940057, Insect Biotechnology Inc., Bayer, NIH AI-41254, DACS and the University of Florida project enhancement award to D. Borovsky, and by the Florida Citrus Production Research Grant to C. A. Powell and D. Borovsky. This publication is Florida Agricultural Experimental Station Journal Series no. R-06494.

REFERENCES

1. Valaitis, A.P. (1995) Gypsy moth midgut proteinases: purification and characterization of luminal trypsin, elastase and the brush border membrane leucine aminopeptidase. *Insect Biochem. Molec. Biol.* **25**, 139–149.
2. Johnston, K.A., Gatehouse, J.A. & Anstee, J.H. (1993) Effect of soybean protease inhibitors on the growth and development of larval *Helicoverpa armigera. J. Insect Physiol.* **39**, 657–664.
3. Borovsky, D. (1985) Isolation and characterization of highly purified mosquito oostatic hormone. *Arch. Insect Biochem. Physiol.* **2**, 333–349.
4. Borovsky, D., Carlson, D.A., Griffin, P.R., Shabanowitz, J. & Hunt, D.F. (1990) Mosquito oostatic factor: a novel decapeptide modulating trypsin-like enzyme biosynthesis in the midgut. *FASEB J.* **4**, 3015–3020.
5. Borovsky, D., Carlson, D.A., Griffin, P.R., Shabanowitz, J. & Hunt, D.F. (1993) Mass spectrometry and characterization of *Aedes aegypti* trypsin modulating oostatic factor (TMOF) and its analogs. *Insect Biochem. Mol. Biol.* **23**, 703–712.
6. Borovsky, D., Song, Q., Ma, M. & Carlson, D.A. (1994) Biosynthesis, secretion and cytoimmunochemistry of trypsin modulating oostatic factor of *Aedes aegypti*. *Arch. Insect Biochem. Physiol.* **27**, 27–38.
7. Bylemans, D., Borovsky, D., Hunt, D.F., Shabanowitz, J., De Grauwels, L. & Loof, A. (1994) Sequencing and characterization of trypsin modulating oostatic factor (TMOF) from the ovaries of the grey fleshfly, *Neobellieria bullata. Regul. Pept.* **50**, 61–72.
8. House, H.L. (1974) Digestion. In *The Physiology of Insect*a (Rockstein, M., ed.) Vol. 5, pp. 63–117. Academic Press, New York.
9. McFarlane, J.E. (1985) Nutrition and digestive organs. In *Fundamentals of Insect Physiology* (Blum, M.S., ed.), pp. 59–89. Wiley, New York.
10. Applebaum, S.W. (1985) Biochemistry of digestion. In *Comparative Physiology, Biochemistry and Pharmacology of Insect*s (Kurkut, G.A. & Gilbert, L.I., eds) Vol. 4, pp. 279–311. Pergamon Press, London.
11. Murdock, L.L., Brookhart, G., Dunn, P.E., Foard, D.E., Kelley, S., Kitch, L., Shade, R.E., Schukle, R.H. & Wolfson, J.L. (1987) Cysteine digestive proteinase in Coleoptera. *Comp. Biochem. Physiol.* **87B**, 783–787.
12. Terra, W.R. & Cristofoletti, P.T. (1996) Midgut proteinases in three divergent species of Coleoptera. *Comp. Biochem. Physiol.* **113B**, 725–730.
13. Baker, J.E. (1982) Digestive proteinases of *Sitophilus* weevils (Coleoptera: Curculionidae) and their response to inhibitors from wheat and corn flour. *Can. J. Zool.* **60**, 3206–3214.
14. Purcell, J.P., Greenplate, J.T. & Sammons, R.D. (1992) Examination of midgut luminal proteinase activity in six economically important insects. *Insect Biochem. Molec. Biol.* **22**, 41–47.
15. Peterson, A.M., Barillas-Mury, C.V. & Wells, M.A. (1994) Sequence of three cDNAs encoding an alkaline midgut trypsin from. *Manduca sexta. Insect Biochem. Mol. Biol.* **24**, 463–471.
16. Wang, S., Young, F. & Hicky, D.A. (1995) Genomic organization and expression of a trypsin gene from the spruce budworm, *Choristoneura fumiferana. Insect Biochem. Molec. Biol.* **25**, 899–908.
17. Barillas-Mury, C., Graf, R., Hagedorn, H.H. & Wells, M.A. (1991) cDNA and deduced amino acid sequence of a blood meal-induced trypsin from the mosquito, *Aedes aegypti*. *Insect Biochem.* **21**, 825–831.
18. Kalhok, S.E., Tabak, L.M., Prosser, D.E., Brook, W., Downe, A.E.R. & White, B.N. (1993) Isolation, sequencing and characterization of two cDNA clones coding for trypsin-like enzymes from the midgut of *Aedes aegypti*. *Insect Mol. Biol.* **2**, 71–79.
19. Muller, H.M., Crampton, J.M., del la Torre, A., Sinden, R. & Crisanti, A. (1993) Members of a trypsin gene family in *Anopheles gambiae* are induced in the gut by the blood meal. *EMBO J.* **12**, 2891–2900.
20. Borovsky, D., Janssen, I., Vanden Broeck, J., Huybrechts, R., Verhaert, P., De Bondt, H.L., Bylemans, D. & De Loof, A. (1996) Molecular sequencing and modeling of *Neobellieria bullata* trypsin; evidence for translational control by *Neobellieria* trypsin-modulating oostatic factor. *Eur. J. Biochem.* **237**, 279–287.
21. Noriega, F.G., Pennington, J.E., Barillas-Mury, C., Wang, X.Y. & Wells. M.A. (1996) *Aedes aegypti* midgut early trypsin is post-transcriptionally regulated by blood feeding. *Insect Mol. Biol.* **5**, 25–29.
22. Noriega, F.G., Shah, D.K. & Wells, M.A. (1997) Juvenile hormone controls early trypsin gene transcription in the midgut of *Aedes Aegypti. Insect Mol. Biol.* **6**, 63–66.
23. Moffatt, M., Blakemore, D. & Lehane, M.J. (1995) Studies on the synthesis and secretion of trypsin in the midgut of *Stomoxys calcitrans*. *Comp. Biochem. Physiol.* **110B**, 291–300.
24. Barillas-Mury, C., Noriega, F.G. & Wells, M.A. (1995) Early trypsin activity is part of the signal transduction system that activates transcription of the late trypsin gene in the midgut of the mosquito, *Aedes aegypti*. *Insect Biochem. Molec. Biol.* **25**, 241–246.
25. Lehane, M.J., Blakemore, D., Williams, S. & Moffatt, M.R. (1995) Mini review: Regulation of digestive enzyme level in insects. *Comp. Biochem. Physiol.* **110B**, 285–289.
26. Ward, G.B., Mayer, R.T., Feldlaufer, M.F. & Svoboda, J. (1991) Gut chitin synthase and sterols from larvae of *Diaprepes abbreviatus*

(Coleoptera: Curculionidae). *Arch. Insect Biochem. Physiol.* **18**, 105–117.

27. Beavers, J.B. (1982) Biology of *Diaprepes abbreviatus* (Coleoptera: Curculionidae) reared on an artificial diet. *Fla. Entomol.* **65**, 263–269.
28. Borovsky, D. & Schlein, Y. (1988) Quantitative determination of trypsin like and chymotrypsin like enzymes in insects. *Arch. Insect Biochem. Physiol.* **8**, 249–260.
29. McManus, M.T. & Burgess, E.P.J. (1995) Effects of the soybean (Kunitz) trypsin inhibitor on growth and digestive proteases of larvae of *Spodoptera litura. J. Insect Physiol.* **41**, 731–738.
30. Chomczynski, P. & Sacchi. N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate–phenol–chloroform extraction. *Anal. Biochem.* **162**, 156–159.
31. Davis, C.A., Riddler, D.C., Higgins, M.J., Holden, J.J.A. & White, B.N. (1985) A gene family *Drosophila melanogaster* coding for trypsin-like enzymes. *Nucleic Acids Res.* **13**, 6605–6619.
32. Frohman, M.A. (1993) Rapid amplification of complementary DNA ends for generation of full-length complementary DNAs: thermal RACE. *Methods Enzymol.* **8**, 340–356.
33. Sanger, F., Nicklen, S. & Coulson, A.R. (1977) DNA sequencing with chain terminating inhibitors. *Proc. Natl Acad. Sci. USA*. **74**, 5463–5467.
34. Tabor, S. & Richardson, C.C. (1987) DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. *Proc. Natl Acad. Sci. USA* **84**, 4767–4771.
35. Kraulis, P. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946–950.
36. Johnston, K.A., Lee, M.J., Gatehouse, J.A. & Anstee, J.H. (1991) The partial purification and characterization of serine protease activity in midgut of larval *Helicoverpa armigera. Insect Biochem.* **21**, 389–397.
37. Lee, M.J. & Anstee, J.H. (1995) Endoproteases from the midgut of larval *Spodoptera litoralis*, include a chymotrypsin-like enzyme with an extended binding site. *Insect Biochem. Molec. Biol.* **25**, 49–61.
38. Johnston, K.A., Lee, M.J., Brough, C., Hilder, V.A., Gatehouse, A.M.R. & Gatehouse, J.A. (1995) Protease activities in the larval midgut of *Heliothis virescens*: Evidence for trypsin and chymotrypsin like enzyme. *Insect Biochem. Molec. Biology* **25**, 375–383.
39. von Heijne, G. (1986) A new method for predicting signal sequence cleavage sites. *Nucl. Acids Res.* **14**, 4683–4690.
40. Blow, D.M. (1976) Structure and mechanism of chymotrypsin. *Acc. Chem. Res.* **9**, 145–152.
41. Branden, C. & Tooze, J. (1991) *Introduction to Protein Structure.* Garland Publishing, Inc., New York.
42. Baker, J.E., Woo, S.M. & Mullen, M.A. (1984) Distribution of proteinases and carbohydrases in the midgut of larvae of the sweetpotato weevil *Cylas formicarius elegantulus* and response of proteinases to inhibitors from sweet potato. *Entomol. Exp. Appl.* **36**, 97–105.
43. McGhie, T.K., Christeller, J.T., Ford, R. & Allsopp, P.G. (1995) Characterization of midgut proteinase activities of white grubs: *Lepidiota noxia*, *Lepidiota negatoria,* and *Antitrogus consanguineus* (Scarabaeidae, Melolonthin). *Arch. Insect Biochem. Physiol.* **28**, 351–363.
44. Broadway, R.M. & Duffy, S.A. (1986) The effect of dietary protein on the growth and digestive physiology of larval *Heliothis zea* and *Spodoptera exigua. J. Insect Physiol.* **32**, 678–680.
45. Mahmood, F. & Borovsky, D. (1992) Biosynthesis of trypsinlike and chymotrypsinlike enzyme in immature *Lutzomyia anthophora* (Diptera: Psychodidae). *J. Med. Entomol.* **29**, 489–495.
46. Spiro-Kern, A. (1974) Untersuchungen uber die proteasen bei *Culex pipiens. J. Comp. Physiol.* **90**, 53–70.
47. Larocque, A.M. & Houseman, J.G. (1990) Effect of ingested soybean, ovomucoid and corn protease inhibitors on digestive processes of the European corn borer, *Ostrinia nubilalis* (Lepidoptera: Pyralidae). *J. Insect Physiol.* **36**, 691–697.
48. Broadway, R.M. & Duffy, S.A. (1986) Plant proteinase inhibitors: mechanism of action and effect on the growth and digestive physiology of larval *Heliothis zea* and *Spodoptera exigua*. *J. Insect Physiol.* **32**, 827–833.
49. Wu, Y., Lewellyn, D.L., Mathews, A. & Elizabeth, S.D. (1997) Adaptation of *Helicoverpa armigera* (Lepidoptera: Noctuidae) to a proteinase inhibitor expressed in transgenic tobacco. *Mol Breed* **3**, 371–380.
50. Borovsky, D. (1988) Oostatic hormone inhibits biosynthesis of midgut proteolytic enzymes and egg development in mosquitoes. *Arch. Insect Biochem. Physiol.* **7**, 187–210.
51. Broadway, R.M. (1997) Dietary regulation of serine proteinases that are resistant to serine proteinase inhibitors. *J. Insect Physiol.* **43**, 855–874.
52. Hedstorm, L., Szilagyi, L. & Rutter, W.J. (1992) Converting trypsin to chymotrypsin: the role of the surface loops. *Science* **255**, 1249–1253.
53. Gierasch, L.M. (1989) Signal sequences. *Biochemistry* **28**, 923–930.
54. Sasaki, T., Hishida, T., Ichikawa, K. & Asari, S. (1993) Amino acid sequence of alkaliphilic serine protease from silkworm, *Bombyx mori*, larval digestion juice. *FEBS Lett.* **320**, 35–37.
55. Felix, C.R., Betschart, B., Billingsley, P.E. & Freyvogel, T.A. (1991) Post feeding induction of trypsin in the midgut of *Aedes aegypti* L. (Diptera: Culicidae) is separable into two cellular phases. *Insect Biochem.* **21**, 197–203.
56. Theil, E.C. (1990) Regulation of ferritin and transferrin receptor mRNAs. *J. Biol. Chem.* **265**, 4771–4774.
57. Berry, J.O., Carr, J.P. & Klessig, D.F. (1988) mRNAs encoding ribulose-1,5-bisphosphate carboxylase remain bound to polysomes but are not translated in amaranth seedlings transferred to darkness. *Proc. Natl Acad. Sci. USA* **85**, 4190–4194.
58. Lindquist, S. (1986) The heat shock response. *Annu. Rev. Biochem.* **55**, 1151–1191.
59. Yen, T.J., Machlin, P.S. & Cleveland, D.W. (1988) Autoregulated instability of β-tubulin mRNAs by recognition of the nascent amino terminus of β-tubulin. *Nature* **334**, 580–585.
60. Sachs, A. & Wahle, E. (1993) Poly (A) tail metabolism and function in eucaryotes. *J. Biol. Chem.* **268**, 22955–22958.
61. Noriega, F.G., Wang, X.Y., Pennington, J.E., Barillas-Mury, C.V. & Wells, M.A. (1996) Early trypsin, a female-specific midgut protease in *Aedes aegypti*: isolation, amino-terminal sequence determination, and cloning and sequencing of the gene. *Insect Biochem. Molec. Biol.* **26**, 119–126.
62. Mutimer, H., Deacon, N., Crowe, S. & Sonza, S. (1998) Pitfalls of processed pseudogenes in RT-PCR. *Biotechniques* **24**, 585–588.