



# Early Trypsin, a Female-specific Midgut Protease in *Aedes aegypti*: Isolation, Amino-terminal Sequence Determination, and Cloning and Sequencing of the Gene

FERNANDO G. NORIEGA,\* XIAO-YU WANG,\* JAMES E. PENNINGTON,\*  
CAROLINA V. BARILLAS-MURY,† MICHAEL A. WELLS\*

Received 14 March 1995; revised and accepted 2 August 1995

Early trypsin is a female-specific protease present in the *Aedes aegypti* midgut during the first hours after ingestion of a blood meal. It plays an essential role in the transcriptional activation of the late trypsin form, the major midgut endoprotease involved in the blood meal digestion. Early trypsin is the most abundant midgut polypeptide isolated by benzamidine-sepharose affinity chromatography 3 h after feeding. The amino-terminal sequence of the early trypsin protein matches that of the 3a1 cDNA for a putative trypsinogen described by Kalhok *et al.* (*Insect. Molec. Biol.*, 2, 71–79, 1993). The early trypsin cDNA was over expressed in *Escherichia coli*. Polyclonal antibodies generated against this recombinant protein were used to show that the enzyme was present in the midgut during the first 4 h after feeding. A 2.5 kb genomic clone of the early trypsin was isolated, mapped and subcloned. A 1.56 kb subclone, corresponding to 1303 bp of the upstream regulatory region and 265 bp of the coding region, was sequenced. The gene contains a 64 nucleotide intron which interrupts the codon for Val at position 18 of the protein. This Val is located toward the end of the putative signal sequence of the protein.

Mosquito Trypsin Gene

## INTRODUCTION

Ingestion of a blood meal by the female mosquito, *Aedes aegypti*, results in an increase in trypsin activity (Fisk, 1950). Post-feeding induction of trypsin activity in the midgut of *A. aegypti* is separable into two phases (Felix *et al.*, 1991), and each phase is characterized by the presence of a specific group of trypsin forms (Graf and Briegel, 1989). The early phase begins immediately after feeding (Rudin and Hecker, 1979); is characterized by the presence of the early trypsin forms (Graf and Briegel, 1989); and involves translation of mRNA already present in the midgut before feeding (Felix *et al.*, 1991).

Late trypsins are present during the late phase of digestion, and are responsible for the majority of endoproteolytic activity present in the midgut during blood-meal

digestion. The late trypsin described by Graf and Briegel (1988, 1989) is the major trypsin form present in the midgut during the peak of trypsin activity, which occurs 24 h after blood feeding. The late trypsin cDNA (Barillas-Mury *et al.*, 1991) and gene (Barillas-Mury and Wells, 1993) sequences have been reported, as well as the transcriptional regulation of this gene (Barillas-Mury *et al.*, 1991; Noriega *et al.*, 1994).

Although the activity of early trypsin plays an indispensable role in inducing late trypsin gene transcription, and is, therefore, essential for blood digestion (Barillas-Mury *et al.*, 1995), little is known about this midgut protease. Recently, Kalhok *et al.* (1993) reported the cDNA sequence for a midgut trypsin present in unfed *A. aegypti* (GenBank X64362), which is different from the late trypsin sequence reported by Barillas-Mury *et al.* (1991). In the present paper, we describe the isolation of the most abundant early trypsin form and show that its amino-terminal sequence is identical to that of the X64362 cDNA. We also analyzed the expression of this protein using polyclonal antibodies. In addition, we report the isolation

\*Department of Biochemistry and Center for Insect Science, Biological Sciences West, University of Arizona, Tucson, AZ 85721, U.S.A.

†Present address: European Molecular Biology Laboratory, Heidelberg 6900, Germany.

and sequencing of the early trypsin gene, which contains a 64 nucleotide intron.

## MATERIALS AND METHODS

### *Insects*

*A. aegypti* of the Rockefeller strain were reared at 27°C and 80% relative humidity under a 12 h light: 12 h dark photoperiod regime. Adults were supplied with a cotton wool pad soaked in a 10% sucrose solution until 12–16 h before the experimental feeding. In this paper, we will refer to the sucrose-fed females as “unfed”.

### *Materials*

Pig  $\gamma$ -globulin (Cohn fraction II, III) and pig hemoglobin were from Sigma (St Louis, MO). RNA-binding glass powder (RNAid™ Kit) was purchased from BIO 101 (La Jolla, CA).

### *Mosquito meals*

Mosquitoes were fed either pig blood (supplemented with 2 mg/ml of isoleucine); 150 mg/ml solutions of pig  $\gamma$ -globulin or pig hemoglobin in 100 mM NaHCO<sub>3</sub> and 150 mM NaCl, pH 7.4; or the artificial blood meal described by Kogan (1990). All meals were equilibrated to 37°C and ATP was added to a final concentration of 1 mM immediately before use.

### *Affinity chromatography*

Groups of 200–1000 mosquitoes were homogenized in 5 ml of 300 mM imidazole buffer pH 7.4, containing 20 mM CaCl<sub>2</sub> and 0.1 mM phenylthiourea, using a Brinkmann polytron. The homogenates were centrifuged at 10,000 g for 30 min. The trypsin in the soluble fraction was purified on a 2 ml benzamidine-Sepharose 6B affinity column (Pharmacia, Piscataway, NJ). The sample was applied to the column and allowed to incubate on the column for 2 h at 4°C. Then the column was washed with 6 column volumes of the homogenization buffer. The trypsin was eluted with 2 column volumes of 200 mM benzamidine in the same buffer and desalted into 100 mM Tris buffer pH 7.4, containing 150 mM NaCl, using a PD10 column (Pharmacia). In some cases, after desalting, diisopropylfluorophosphate (DFP) was added to the samples to a final concentration of 25 mM.

### *Protein sequencing*

After benzamidine-affinity purification, the proteins were separated by SDS-PAGE, blotted onto Polyvinylidene Difluoride membranes (PVDF) (Bio-Rad) and the major band, a 30 kDa protein, was subjected to amino acid sequence analysis performed at the Harvard Microchemistry Facility.

### *Preparation of an early trypsin cDNA probe*

Total RNA was isolated from females, 2 h after feeding a 15%  $\gamma$ -globulin meal, using RNA-binding glass

powder (Noriega and Wells, 1993). First strand cDNA was synthesized using reverse transcriptase and 100 pmol of a 34-mer oligonucleotide with the sequence 5'-(T)<sub>16</sub>CCAAGCTTGGATCCATCG-3'. This first strand-cDNA was used as a template for PCR, using as one primer a degenerate oligonucleotide designed to the highly conserved region flanking the active site serine present in trypsins (Peterson *et al.*, 1994), and, as the other primer, the sequence used for first strand cDNA synthesis (34-mer from which the oligo-dT sequence was removed). Conditions for PCR were as described by Peterson *et al.* (1994). PCR products were subcloned and sequenced using the dideoxy chain termination method (Sanger *et al.*, 1977).

### *Northern blots*

Total RNA was separated by electrophoresis on 1.5% agarose gels under denaturing conditions using the formaldehyde method (Fourney *et al.*, 1988). The RNA was transferred to a Nytran membrane (Schleicher & Schuell, Keene, NH) and hybridized with an early trypsin cDNA or a late trypsin cDNA (Barillas-Mury *et al.*, 1991), each of which were labeled using a Random Primer Kit (BRL) and ( $\alpha$ -<sup>32</sup>P)-dATP (ICN, Irvine, CA). Hybridizations were carried out under high stringency conditions, as previously described (Noriega and Wells, 1993).

### *Screening the genomic library*

Approximately  $5 \times 10^5$  plaques were plated from an amplified mosquito genomic library (Barillas-Mury and Wells, 1993), using as host cells *Escherichia coli* strain P2393 (Stratagene, La Jolla, CA). The plaques were screened with the radioactive early trypsin cDNA probe described above, and two positive clones were obtained.

### *Mapping and subcloning of the genomic DNA clone*

The two positive clones were purified and phage DNA was obtained using the method of Benson and Taylor (1984). The DNA was digested with different restriction enzymes and the digestion products subjected to Southern analysis (Sambrook *et al.*, 1989), using the labeled early trypsin cDNA as a probe. The results showed that both clones were identical. A 2.5 kb Sall to PSTI fragment was subcloned into the Bluescript SK<sup>+</sup> plasmid (Stratagene) for sequence analysis.

### *Sequencing the genomic DNA clone*

Exonuclease III deletion (Henikoff, 1984) was used to generate different sized clones. Double-stranded DNA was used as template and sequenced by the dideoxy chain termination method (Sanger *et al.*, 1977).

### *Data analysis*

The sequence of the upstream region of the gene was analyzed for the presence of consensus sequences for DNA binding proteins using the GCG sequence analysis software package (Genetics Computer Group Inc., Madison, WI).

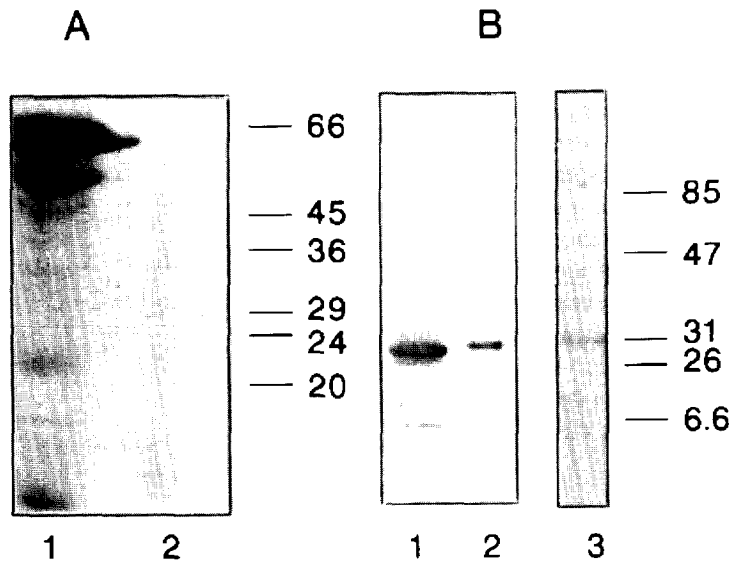


FIGURE 1. SDS-PAGE and Western blot analysis of mosquito benzamide purified homogenates. (A) Coomassie blue staining. 1: female homogenate (one mosquito equivalent) 3 h after feeding an artificial blood meal (ABM) (before benzamide purification); 2: benzamide-purified proteins from homogenates of females (10 female equivalents) 3 h after feeding an ABM. (B) Western blot probed with an early trypsin antiserum (1 : 8000 dilution). 1: benzamide-purified proteins from homogenates of females (10 female equivalents) 3 h after feeding an ABM; 2: recombinant early trypsin; 3: midgut extract (10 female equivalents) 3 h after feeding an ABM.

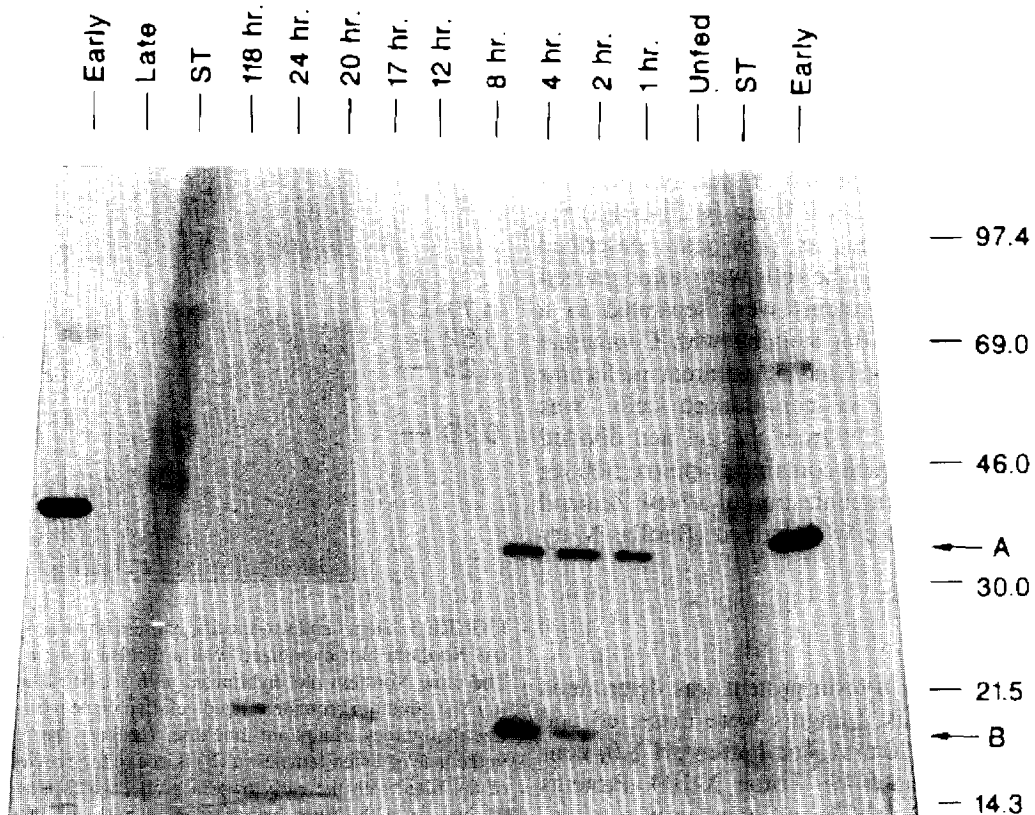


FIGURE 2. Early trypsin protein expression. Western blot of benzamide-purified extracts of females (20 females equivalents/lane) after feeding a 15%  $\gamma$ -globulin meal; probed with an early trypsin polyclonal antibody. Early: recombinant early trypsin. Late: recombinant late trypsin. ST: molecular weight standards. Unfed: unfed female. 1, 2, 4, ••, 118 h: different hours after feeding; (A) early trypsin; (B) 19 kDa fragment.

*Primer extension*

An oligonucleotide complementary to the gene sequence from bases 63–39, which was end-labeled with [ $\gamma$ - $^{32}$ P]-ATP, using T4 kinase (Maniatis *et al.*, 1982), was used as a primer with poly A<sup>+</sup> RNA isolated from unfed mosquitoes. The poly A<sup>+</sup> RNA and the primer were annealed at 65°C for 5 min in the presence of an RNAase inhibitor (40 U/ml, United States Biochemical Corp.), then the primer was extended with reverse transcriptase at 42°C for 1 h. The extension product was ethanol precipitated and analyzed on a 6% polyacrylamide gel simultaneously with a sequencing reaction done with the oligonucleotide primer and the genomic clone.

*Polyclonal antibodies against early trypsin*

Early trypsin was over-expressed in an *E. coli* expression system as previously described for late trypsin (Barillas-Mury *et al.*, 1995). *A. aegypti* genomic DNA was used as template for a PCR reaction to produce an early trypsin DNA containing the appropriate restriction sites for cloning into the expression vector pET 11-a (MRL2) (Novagen, Madison, WI). The primers were based on the sequence of X64362. One primer had an NdeI restriction site, an initiating methionine codon, and 19 bases complementary to the sequence of the lower strand between positions 91 and 111. The second primer introduced a BglIII site and was complementary to the upper strand between positions 777 and 796. This DNA encodes for the mature enzyme plus 4 amino acids of the activation peptide. It was necessary to produce this truncated DNA because the early trypsin gene contains an intron in the activation peptide.

PCR amplification, ligation into the plasmid pET 11-a and transformation into *E. coli* strain BL21-(DE3) for expression were carried out as previously described (Barillas-Mury *et al.*, 1995). The cells were homogenized after induction and the proteins were separated in a 12.5% SDS-PAGE gel. After staining with Coomassie Brilliant Blue, a prominent band of apparent molecular mass of 30 kDa, absent in non-induced cells, was observed. This band was cut from the gel and divided into three equal sections, each containing approx. 500  $\mu$ g of early trypsin protein. Immunization of a New Zealand rabbit was done as previously described (Barillas-Mury *et al.*, 1995).

*Immunoblotting*

The presence of early trypsin protein was determined by Western blot analysis. Midguts were dissected and sonicated for 20 s in 200 mM Tris-buffer pH 8.0, containing 0.1 M Ca<sub>2</sub>Cl and 0.4% Triton X-100. Proteins were separated by SDS-PAGE and transferred to nitrocellulose. Early trypsin was detected with the early trypsin polyclonal antibody and goat anti-rabbit IgG conjugated with alkaline phosphatase (Bio-Rad, Richmond, CA).

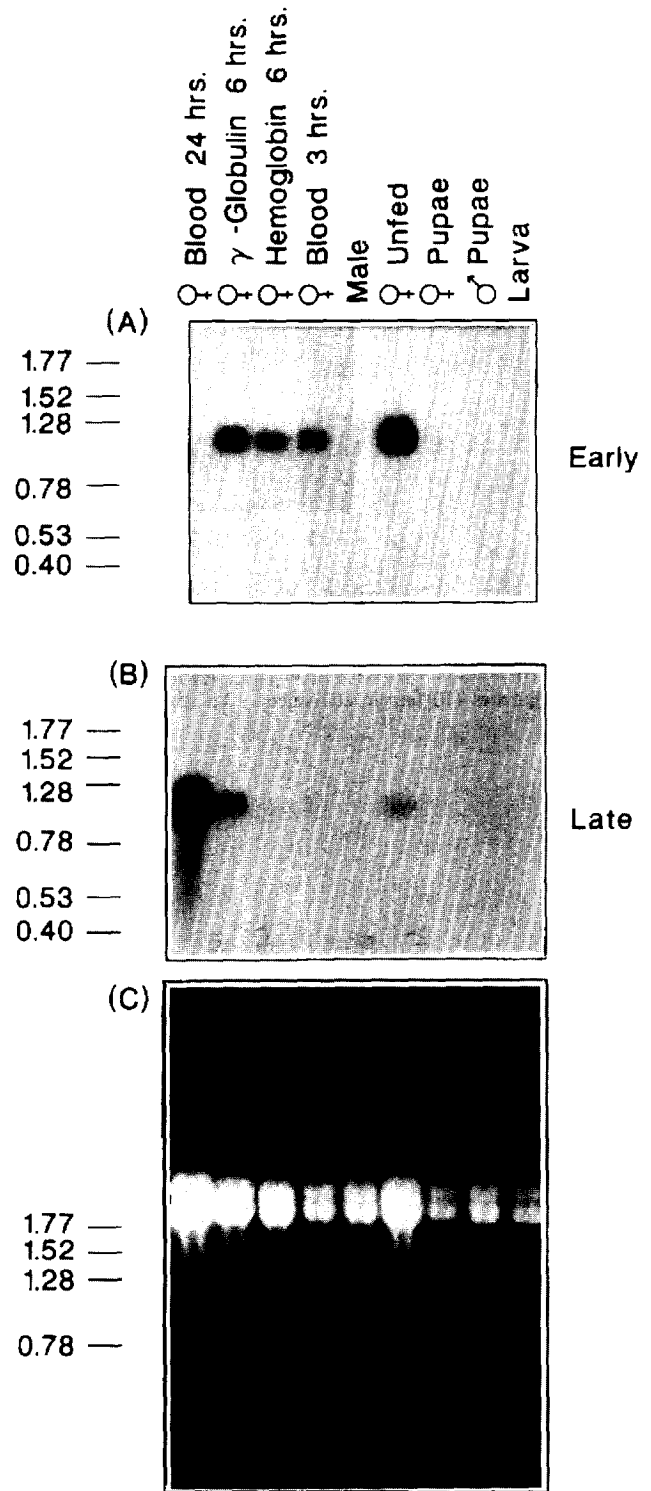


FIGURE 3. Stage- and sex-specific expression of early trypsin mRNA. (A) Northern blot hybridized with a labeled early trypsin probe. (B) The same Northern blot hybridized with a labeled late trypsin probe. (C) The ethidium bromide stained gel. Each lane represents RNA from five mosquitoes. Hours are time after feeding. The band that appears in the lane of unfed females in (B) is due to the previous hybridization of the filter with the early trypsin probe, which was not completely washed from the filter before probing with the late trypsin probe. Unfed females were 4 days old.

## RESULTS

*Early trypsin isolation and characterization*

Groups of 1000 females were homogenized 3 h after feeding an artificial blood meal (Kogan, 1990) and benzamidine affinity chromatography was used to isolate trypsins. When these affinity-purified samples were analyzed by SDS-PAGE, one major protein with an apparent molecular weight of 30 kDa was detected after Coomassie Blue staining (Fig. 1). We transferred this protein onto a PVDF membrane and the sequence of the first 20 amino terminal amino acids was IVGGFQIDIA-EVPHQVSLQR. When compared with the amino sequences present in the GenBank, it was identical to the derived amino terminal sequence of the 3a1 cDNA (Accession number X64362), that codes for a putative trypsinogen (Kalhok *et al.*, 1993).

*Early trypsin protein expression*

Rabbit polyclonal antibodies against the recombinant X64362 protein were used to study the induction of early trypsin by feeding. Early trypsin antiserum specifically recognized the recombinant early trypsin, the most abundant protein present in the 3 h benzamidine-purified fraction, and a 30 kDa protein present in a midgut extract (Fig. 1).

Females were fed 15%  $\gamma$ -globulin and sacrificed at different times. Early trypsin was concentrated using benzamidine-affinity columns and detected by Western blot

analysis. Early trypsin was present in relative large amounts in the midgut at 1, 2 and 4 h after feeding, but by 8 h, early trypsin levels were almost undetectable (Fig. 2). A second protein, of about 19 kDa, was detected by the antibody 2–24 h after feeding (Fig. 2); this protein may represent an early trypsin proteolytic degradation product.

*Preparation of the early trypsin cDNA probe*

A 293 bp DNA fragment obtained after PCR, was cloned and sequenced. A search of GenBank revealed a perfect match to a midgut *A. aegypti* cDNA for a putative trypsinogen (Accession number X64362). This PCR product was  $^{32}$ P-labeled and used as a probe for Northern blot analysis. The PCR product hybridized with an RNA of  $\approx$  900 bases, whose pattern of expression was completely different from that of the late trypsin gene (Fig. 3); no early trypsin mRNA was detected in larvae, pupae, or in male adults. However, the message was abundant in unfed females and the levels declined after ingestion of a blood or protein meal (Fig. 3). The early-trypsin probe did not cross-react with any of the two other *A. aegypti* midgut trypsin cDNAs described; the late trypsin cDNA (Barillas-Mury *et al.*, 1991) or the 5g1 trypsin cDNA (Kalhok *et al.*, 1993) (results not shown).

*Cloning and sequencing of the early trypsin gene*

Figure 4 shows a partial restriction map of the genomic clone from which a 2.5 kb subclone, containing the early

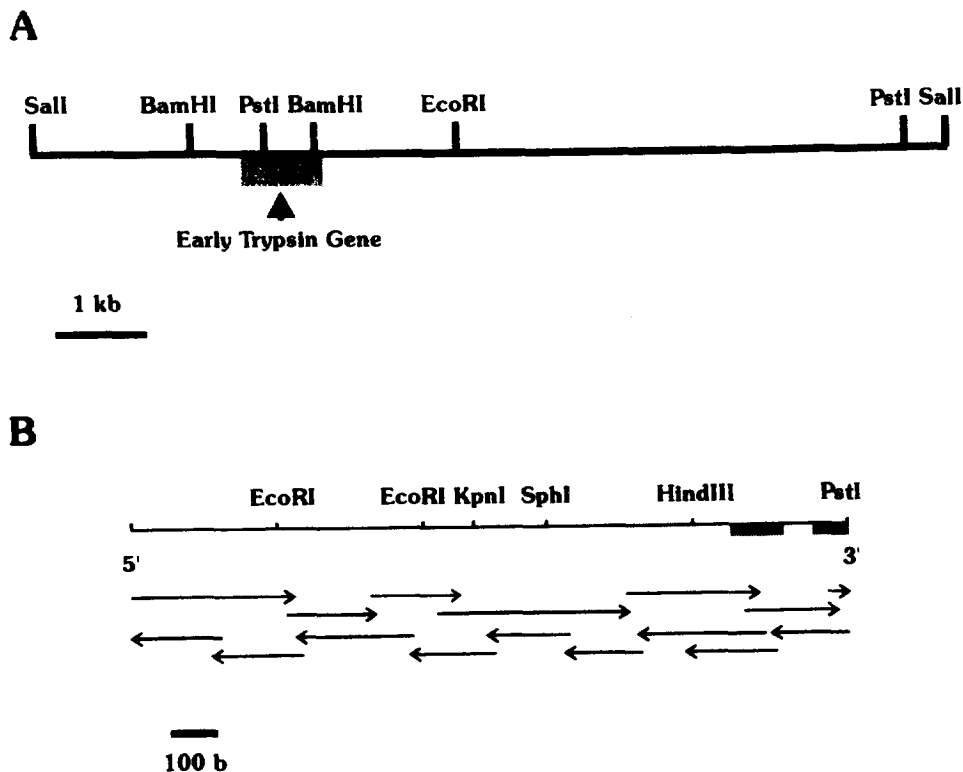


FIGURE 4. Characterization of the early trypsin genomic clone. (A) A partial restriction map of the genomic clone. The map was constructed using the indicated restriction enzymes and Southern blot analysis. The shade region shows the location of the early trypsin gene. (B) The sequenced region of the early trypsin gene and the sequencing strategy used. The shaded regions show the transcribed regions. The arrows indicate the direction and extent of sequence determined from each site.

```

-1303          TTTTCGTTTTGAAAATTGAAGCAATGGGCAAGGAAAGATTCCAC
-1260  TGATATCTAAGAGAATATAGAACAACGATATTTTTTTATGAAGGCTGCATTGATTCCAGGA
-1200  GTTGGAAATTC AATTTTGTCTTCCAAAATGAGAAGAACTTTACAGAGAAAAATAGTTAT
-1140  TGTAATCTATGAATTTGCATGAAAAGTTGAAATAATCAGTCTTGATAGGTTGCAATTTTT
-1080  GCTTGGTAGAACAGTCAATGGGTCAGATATGAGAAGAACTTCCAAAAGGGAAATCGATCA
-1020  CTGCTAAAATATATCATTATGATACTAAACGCTAGAATTCACGTGATGCCATCCGATGG
-960   ACTCTGCATCTGTCTTCAGAAGCCATATTTTCAATGCCCAACTACCACAAGTTCGTGGAC
-900   GAAGCATTTCTCGACTGTTGGAAATTACGTATTTTGAAGACCCATCTAAAAAATTTGACA
-840   TGGAAATATGGTCCAGGTGTTTCCGGAGTTATCTGGATTGCTTGGGGTATAAAAAATTG
-780   GCCACATCATCCCTTCAACGGATATCTTAGGATCCTTGGCGGTTGGTGTTCGGACAAAG
-720   TTCTAGCTCATCAGGATCTTGAGTTATGAAAAATTTAAGAAAAATGTTAGCGCCACGAA
-660   TTCTCATCAGATTTCCATTGCTAGATAGCCTTTGCCGAAAACACAAACATTCTAGCTCA
-600   TCAGGATCCTGAGATATCCGTTGAATGGATGATGTGGCCAATTTTGGTACCTCAAGACAA
-540   TCCGGAATAACTACGAAACCACCTAGACCATATCTCCGAAAGTTTCAGACCTTAAACTAG
-480   AAGAGTTTCTCGGGAATTTCTGAAAATTTAGCAAATTCATCAGGAAACAAAAATGCTA
-420   TGCAGGTTTTAGTGTGTTTCTCAGCATGCAAAGTATGTTGGCTAAATGAAGGTTAAGAT
-360   GGTAGTGTGCAAAATACAGAGCACATAGTGTAAATGTTTAAAATTATACATATAAAAGAA
-300   TTAAAAACAAGTAAAAGCATTTTATCCATATAATATTTGAAAGATTACTGGGTACTCCA
-240   AACATCAGATAATTGGTTTCCCTTCAATGTTTTGAAATTACCCATCCCACACGCGAAGA
-180   CGATAAAACCAC TTGGGAAAATCCATCTGCACGTGTGTACCGTAATCTCCACCAC TTCAA
-120   ATTAAGAATCCGCATCATAGGAATGACTTCTTGGCGTCTCTCGAAGCTTATCAGCCCA
-60   GAGCCGCCAAAGTGGGTTTCGATTAGCCACTATAAAAGTCCATTGCAACGTACACTCATC
1     CTTCCCAAAGCCAAACCACTTCTAACAGTTTACCGGCTACCGCATAACCTGAACCACA
61    GCCATGAACCAATTTCTCTTTGTGAGTTTTTTGTGCCCTTCTGGACTCAGCCAAAGGgtaag
121   ggacaaaatagtttcgagaggtgtacaacgctagatggttaaatttctaaaaatcgaatagT
181   TTCAGCTGCAACGCTGTCCAGCGGTGCATCGTTGGCGGATTCAGATCGACATCGCCGA
241   GGTCCCGCATCAGGTGTCCCTGCAG

```

FIGURE 5. Sequence of the early trypsin gene. The sequence encompasses 1303 bases upstream of the transcription start site (negative numbers) and 265 bases of the transcribed region (positive numbers), including a 64 base intron. The bases found in the mature mRNA are underlined and the TATA box is indicated by \*.

trypsin gene, was isolated. Figure 4 also shows the sequencing strategy used, and Fig. 5 shows the nucleotide sequence of 1568 bases of the subclone. The genomic sequence encompasses 1303 nucleotides upstream of the transcription initiation site, which was determined by primer extension analysis (Fig. 6), and 265 bases of the transcribed mRNA. The gene contains a 64 nucleotide intron which interrupts the codon for Val at position 18 of the protein.

Comparison of the upstream region of the early trypsin gene with known transcription factor recognition

sequences revealed 137 perfect matches, however, none of the putative transcription factors are of obvious relevance in mosquitoes. Sequence alignment of the upstream regions of the early and late trypsin genes revealed only two regions of high similarity (> 4 bases). One region of 59 bases, shown below, had 74% identity.

```

Early Gene -503  GARAGTTTCAGACCTTAAACTAGAGAGT...TTCTCGGGAAATTTCTGAAAATTTGAG -449
Late Gene -598  GAAAATTTGAG.CCTCAAATAGCACGTTTTGAGTCTTGGGAATTTTAGAAGTTTTAG -539

```

The second region consists of 9 bases which encompasses the TATA box with the following

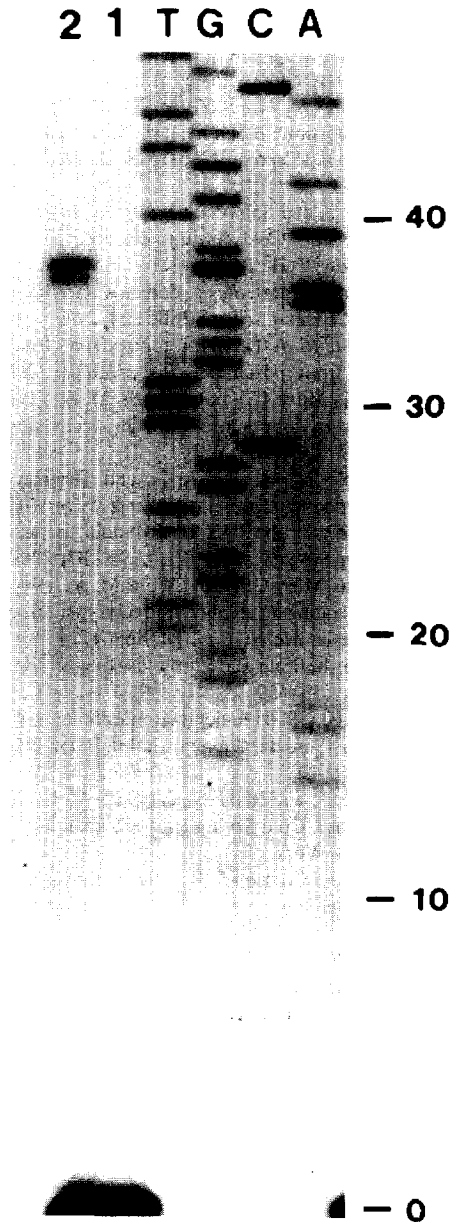


FIGURE 6. Primer extension analysis of the early trypsin transcript. The labeled oligonucleotide was hybridized with poly A<sup>+</sup> RNA from unfed mosquitoes (lane 2) or no RNA (lane 1) and extended as described in Materials and Methods. A sequencing reaction done with the oligonucleotide primer and the genomic clone was used as a size marker (right margin). Zero indicates the labeled oligonucleotide.

sequence: ACTATAAAA. This sequence matches the 9 bp degenerate consensus sequence, HCWATHAAA (where H is A, T or C, and W is A or T), for the binding of the engrailed homeodomain protein from *Drosophila melanogaster* (Ohkuma *et al.*, 1990).

#### DISCUSSION

According to the time of appearance in the mosquito midgut, trypsin forms can be divided in two groups: early and late trypsins. Early trypsins are defined as any trypsin form present in the midgut during the first hours following a blood meal. Early trypsin forms were first described by Graf and Briegel (1989), based on immunoprecipit-

ation and *in vitro* translation experiments; however, no further characterization of these forms has been published. Kalhok *et al.* (1993) speculated that one of the cDNAs they reported, the 3a1 clone, encoded an early trypsin form, but they did not describe the expression of the protein encoded by this cDNA.

We have purified the most abundant early trypsin form, and amino-terminal analysis showed that this protein is encoded by the 3a1 clone of Kalhok *et al.* (1993). In addition, the antibodies raised against the 3a1 recombinant protein recognized both our purified 30 kDa early trypsin and a 30 kDa protein present in the midgut only during the first hours of digestion.

The unavailability of Graf and Briegel's antiserum (supplies have been depleted), prevents us from unequivocally identifying the cDNA described in this paper as encoding for one of Graf and Briegel's early trypsins. Regardless of whether we can link our early trypsin to those described previously, it is clear that this early trypsin form is the major protein isolated by affinity chromatography from the mosquito midgut early during digestion of the blood meal. The fact that this has been the only early trypsin cDNA obtained from two different libraries, using two different cloning strategies, i.e. screening with a *Drosophila* trypsin cDNA probe (Kalhok *et al.*, 1993) and using PCR (this report), is a further indication that this is the major early trypsin form, which is responsible for the midgut early trypsin activity described by Felix *et al.* (1991) and Barillas-Mury *et al.* (1995).

Our analysis of the expression of the early trypsin gene confirmed the adult female-specificity described by Kalhok *et al.* (1993), and showed that the steady state levels of early trypsin mRNA were reduced after the ingestion of a protein meal.

We have isolated, subcloned and sequenced the early trypsin gene. The gene contains a 64 nucleotide intron. This is the first intron reported amongst insect trypsins. The intron interrupts the codon for Val at position 18 of the protein. This Val is located toward the end of the putative signal sequence of the protein; the mammalian serine proteases possess also an intron located in a similar position (out of 5–7 introns in total) (Swift *et al.*, 1984; Rogers, 1985; Neurath, 1986).

The physiological significance of any of the consensus sequences for DNA binding sites found in the upstream region is not evident and further experiments are necessary to define which, if any, of these sequences are important in early trypsin regulation. When the sequence of the 1303 bp upstream region was aligned with the upstream region of the late trypsin gene (Barillas-Mury and Wells, 1993), two sequences of high similarity were found. The first, consisting of a 59 bp stretch, might contain information for tissue, stage or sex specificity, but more data will be required to confirm such a speculation. The second region fits the consensus sequences for binding of the Engrailed protein. The Engrailed homeodomain protein has been shown, *in vitro*, to bind competitively with the TATA box-binding protein TFIID, and

repress transcription (Ohkuma *et al.*, 1990). It is possible that a engrailed-like protein may be involved in regulating early trypsin gene expression in *A. aegypti*.

In summary, immunological techniques, in combination with amino-terminal sequencing, have allowed us to identify the major early trypsin form present in the female *A. aegypti* midgut during digestion. The identification of this trypsin form and the isolation of the early trypsin gene are important steps towards a better understanding of the blood-meal digestion process in *A. aegypti*.

## REFERENCES

- Barillas-Mury C. V. and Wells M. A. (1993) Cloning and sequencing of the blood meal-induced late trypsin gene from the mosquito *Aedes aegypti* and characterization of the upstream regulatory region. *Insect Molec. Biol.* **2**, 7–12.
- Barillas-Mury C. V., Graf R., Hagedorn H. H. and Wells M. (1991) cDNA and deduced amino acid sequence of a blood meal-induced trypsin from the mosquito, *Aedes aegypti*. *Insect Biochem.* **21**, 825–831.
- Barillas-Mury C. V., Noriega F. G. and 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.
- Benson S. A. and Taylor R. K. (1984) A rapid small-scale procedure for isolation of phage  $\lambda$  DNA. *BioTechniques* **2**, 126–127.
- Felix C. R., Betschart B., Billingsley P. F. and Freyvogel T. A. (1991) Post-feeding induction of trypsin in the midguts of *Aedes aegypti* L. (Diptera: Culicidae) is separable into two cellular phases. *Insect Biochem.* **21**, 197–203.
- Fisk F. W. (1950) Studies on proteolytic digestion in adult *Aedes aegypti* mosquitoes. *Ann. Entomol. Soc. Am.* **43**, 555–572.
- Fourney R. M., Miyakoshi J., Day III R. S. and Paterson M. (1988) Northern blotting: efficient RNA staining and transfer. *Focus* **10**, 5–7.
- Graf R. and Briegel H. (1988) Monoclonal antibodies as probes for *Aedes aegypti* trypsin. *Insect Biochem.* **18**, 463–470.
- Graf R. and Briegel H. (1989) The synthetic pathway of trypsin in the mosquito *Aedes aegypti* L. (Diptera: Culicidae) and *in vitro* stimulation in isolated midguts. *Insect. Biochem.* **19**, 129–137.
- Henikoff S. (1984) Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. *Gene* **28**, 351–359.
- Kalhok S. E., Tabak L. M., Prosser D. E., Brook W., Downe A. E. R. and White B. N. (1993) Isolation, sequencing and characterization of two cDNA clones coding for-trypsin-like enzymes from the midgut of *Aedes aegypti*. *Insect Molec. Biol.* **2**, 71–79.
- Kogan P. H. (1990) Substitute blood meal for investigating and maintaining *Aedes aegypti* (Diptera: Culicidae). *J. Med. Entomol.* **27**, 709–712.
- Maniatis T., Fritsch E. F. and Sambrook J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Springs Harbor Laboratory, Cold Springs Harbor, New York.
- Neurath H. (1986) Limited proteolysis, domains and the evolution of protein structure. *Chemica Scripta* **26B**, 221–227.
- Noriega F. G. and Wells M. A. (1993) A comparison of three methods for isolating RNA from mosquitoes. *Insect Molec. Biol.* **2**, 21–24.
- Noriega F. G., Barillas-Mury C. V. and Wells M. A. (1994) Dietary control of late-trypsin gene transcription in *Aedes aegypti*. *Insect Biochem. Molec. Biol.* **24**, 627–631.
- Ohkuma Y., Horikoshi M., Roeder R. G. and Desplan C. (1990) Engrailed, a homeodomain protein, can repress *in vitro* transcription by competition with the TATA-box-binding protein transcription factor IID. *Proc. Natn. Acad. Sci. U.S.A.* **87**, 2289–2293.
- Peterson A. M., Barillas-Mury C. V. and Wells M. A. (1994) Sequence of three cDNAs encoding an alkaline midgut trypsin from *Manduca sexta*. *Insect Biochem. Molec. Biol.* **24**, 463–471.
- Rogers J. (1985) Exon shuffling and intron insertion in serine protease genes. *Nature* **315**, 458–459.
- Rudin W. and Hecker H. (1979) Functional morphology of the midgut of *Aedes aegypti* L. (Insecta, Diptera) during blood digestion. *Cell Tissue Res.* **200**, 193–203.
- Sambrook J., Fritsch E. F. and Maniatis T. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Sanger F., Nicklen S. and Coulson S. R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natn. Acad. Sci. U.S.A.* **74**, 5463–5467.
- Swift G. H., Craik C. S., Stary J. S., Quinto C., Lahaie R. G., Rutter W. J. and MacDonald R. J. (1984) Structure of the two related elastase genes expressed in the rat pancreas. *J. Biol. Chem.* **259**, 14271–14278.

---

*Acknowledgements*—We thank John Norem for insect care, and Drs Ann Peterson, Alan Smith and Miranda Van Heusden for helpful comments. This work was supported by NIH grant AI31951 and a grant from the John D. and Catherine T. MacArthur Foundation to the Center for Insect Science (8900408).