Pergamon

PII: S0965-1748(97)00001-5

Insect Biochem. Molec. Biol. Vol. 27, No. 4, pp. 283–289, 1997
© 1997 Elsevier Science Ltd
All rights reserved. Printed in Great Britain
0965-1748/97 $17.00 + 0.00

cDNA Cloning and Pattern of Expression of an Adult, Female-Specific Chymotrypsin from *Aedes aegypti* Midgut

QIJIAO JIANG,† MARTIN HALL,‡ FERNANDO G. NORIEGA,† MICHAEL WELLS†*

Received 22 August 1996; revised and accepted 6 December 1996

A cDNA for a midgut chymotrypsin, induced by a blood meal, has been cloned and sequenced from the mosquito *Aedes aegypti*. The 938 base sequence codes for a 268 amino acid protein, which contains an 18-residue signal peptide and a seven-residue activation peptide. The deduced amino acid sequence contains several features typical of chymotrypsin proteases, including the catalytic triad of serine proteases and the residues that determine the chymotrypsin substrate specificity pocket. The chymotrypsin mRNA, absent in larvae, pupae, males and newly emerged females, reaches detectable levels within 24 h post-emergence and attains a maximum level 3–7 days after emergence. Translation of the chymotrypsin mRNA is induced by feeding a protein meal, and there is a dramatic increase in midgut chymotrypsin enzymatic activity after feeding. Chymotrypsin activity remained high during protein digestion, but chymotrypsin protein levels and enzymatic activity were almost undetectable once digestion was completed, 48 h after feeding. © 1997 Elsevier Science Ltd

protease Chymotrypsin Mosquito digestion *Aedes aegypti*

INTRODUCTION

Previous analyses of mosquito proteolytic enzymes, using crude extracts from the midgut of blood fed *Aedes aegypti*, demonstrated the presence of trypsin, chymotrypsin and aminopeptidase activities (Gooding, 1969; Graf and Briegel, 1982; Borovsky and Schlein, 1988). Unlike trypsin and aminopeptidase activities, which showed significant increases after blood feeding, chymotrypsin activity was described as present at a low and constant level throughout blood digestion (Gooding, 1969; Yang and Davies, 1971).

Molecular studies of digestive proteases in adult female *Ae. aegypti* midgut have identified three trypsin genes — one encoding an *early* form (Kalhok *et al.*, 1993; Noriega *et al.*, 1996a) and the other two encoding *late* forms (Barillas-Mury *et al.*, 1991; Kalhok *et al.*, 1993). The different expression patterns of the early and late trypsins correlate with the observed two phases of tryptic activity during blood meal digestion (Felix *et al.*, 1991; Barillas-Mury *et al.*, 1995). On the other hand, no molecular analyses have been carried out for chymotrypsin in *Ae. aegypti* midgut. Here we present the results of cDNA cloning and sequencing, and studies on the pattern of gene expression, protein synthesis and enzymatic activity for a midgut-specific chymotrypsin from adult female *Ae. aegypti* midgut.

MATERIALS AND METHODS

Insects

Ae. aegypti of the Rockefeller strain were reared as previously described (Noriega *et al.*, 1994). Unless indicated, all females used in these experiment were 5 days old. Female mosquitoes were fed an artificial meal containing 10% pig γ-globulin (Cohn fraction II, III; Sigma, St Louis, MO, U.S.A.) in 120 mm $NaHCO_3$ and 100 mm NaCl, pH 7.4. The meal was equilibrated to 37°C and ATP was added to a final concentration of 1 mm immediately before feeding.

CDNA cloning

A 670 bp polymerase chain reaction (PCR) product, encoding for the 3′ end of a cDNA of a chymotrypsin-like protease, was obtained while searching for *Ae.*

*Author for correspondence. Tel.: +1-520 621-3847; Fax: +1-520 621-9288; Email: wells@mozart.biosci.arizona.edu
†Department of Biochemistry and Center for Insect Science, University of Arizona, Tucson, AZ 85721, U.S.A.
‡Department of Physiological Botany, University of Uppsala, Villavagen 6, 752 36 Uppsala, Sweden.

aegypti proteins homologous to α-macroglobulins and complement proteins C_3 and C_4. Reverse transcriptase–polymerase chain reaction (RT–PCR) was carried out using reverse transcriptase, an oligo dT primer (first strand cDNA synthesis kit, Stratagene, La Jolla, CA, U.S.A.) and mRNA from whole fed mosquitoes. The cDNA generated was subjected to PCR amplification using an oligo dT primer and a 5′ degenerate primer, GGITGYGGIGARCARAAYATG (Hall *et al.*, 1995). The 670 bp PCR product was cloned using a TA cloning kit (Invitrogen, San Diego, CA, U.S.A.).

A midgut-specific cDNA library was constructed from females collected 4–6 h after feeding, using the Superscript cDNA Synthesis and Plasmid Cloning System (Gibco BRL, Gaithersburg, MD, U.S.A.). Total RNA from the midgut tissue was extracted using RNAzol B (Leedo Medical Laboratories, Houston, TX, U.S.A.). Poly$(A)^+$ RNA was purified using oligo dT cellulose (Stratagene). The 670 bp PCR fragment was labelled with [α-^{32}P]-dATP using a Random Primer DNA labelling kit (Gibco BRL). This probe was then used to screen the midgut cDNA library.

DNA sequencing and sequence analysis

Nucleotide sequences were determined from both strands by the dideoxy chain termination method (Sanger *et al.*, 1977), using Sequenase 2.0 (United States Biochemical Co., Cleveland, OH, U.S.A.). Nested deletions of the cDNA clone were generated by unidirectional digestion with exonuclease III (Henikoff, 1984).

Sequence analysis was performed with the GCG software package (University of Wisconsin Genetic Computer Group). Database searches were conducted using the Blastp program (NCBI, NIH). Additional deduced amino acid sequences were obtained from the GenEMBL database using Stringsearch. All invertebrate and vertebrate chymotrypsins were subjected to multiple sequence alignments using the PileUp program (Feng and Doolittle, 1987). The *Ae. aegypti* midgut chymotrypsin sequence has been deposited in GenBank under the accession number U56423.

In vitro *transcription and translation of the chymotrypsin-like cDNA clone*

The *Ae. aegypti* chymotrypsin-like cDNA clone (pMC-8) and the parental plasmid (as a negative control) were linearized using Not I (Gibco BRL) and transcribed using T_7 RNA polymerase (Promega). The transcripts were treated with RQ_1 RNase-free DNase (Promega), and further purified using phenol/chloroform extraction and ethanol precipitation. The mRNAs were then translated using rabbit reticulocyte lysate (Promega) in the presence and absence of microsomal membranes from Canine Pancreas (Promega), according to the manufacturer's protocol. A luciferase mRNA was used as a positive control in the translation step. The ^{35}S-labelled translational products ([^{35}S]-L-Cys, Amersham, Arlington Heights, IL, U.S.A.) were analysed by SDS–PAGE and autoradiography.

Expression of the recombinant chymotrypsin-like protease

A cDNA encoding the zymogen for the chymotrypsin-like protease was generated by PCR using as template the cDNA cloned into pMC-8. The 5′ primer, 5′CCC*GAATTC*TCCTCCAGAGCCACCCACAA3′, contains the sense strand sequence encoding part of the activation peptide and an EcoR I site (italics). The 3′ primer, 5′TCGCATGCACGCGTACGTAA3′, corresponds to a vector sequence flanking the cloning site. Twenty-five PCR cycles were carried out in steps of 1 min at 94°C, 2 min at 57°C and 2 min at 72°C. The amplified PCR product was digested with EcoR I and Hind III (Gibco BRL), and the digestion product was cloned into the expression vector pDS56/RBSII/6HIS (kindly provided by Dr Hans-Michael Muller, University of La Sapienza, Italy). The final plasmid construct carries the *Ae. aegypti* chymotrypsinogen cDNA and a sequence encoding six histidine residues at the amino terminus.

Five hundred millilitres of the *Escherichia coli* host strain M15 [pREP4] (Qiagen, Chatsworth, CA, U.S.A.), carrying the $(his)_6$-chymotrypsinogen expression vector were grown at 37°C in LB broth containing ampicillin (100 μg/ml) and kanamycin (25 μg/ml) until A_{590}=0.4. IPTG was then added to a final concentration of 0.5 mM. After 3 h of IPTG induction, the cells were harvested by centrifugation (5000 rpm, 10 min at 4°C), suspended in 50 ml of buffer A (6 M GuHCl, 0.1 M Na-phosphate, 0.01 M Tris-HCl, pH 8.0), and the cell suspension was stirred overnight at 4°C. The lysate was then centrifuged (15 000 rpm, 20 min at 4°C) and the supernatant was loaded onto a Ni-NTA sepharose column (1.5×4 cm), pre-equilibrated with buffer A. The column was washed with buffer A and then with buffer B (8 M urea, 0.1 M Na-phosphate, 0.01 M Tris–HCl, pH 8.0). The bound proteins were eluted with a step pH gradient from pH 6.5 to 4.0. Fractions of 1 ml were collected and analysed on both reducing and non-reducing SDS–PAGE with Coomassie blue staining. Purified recombinant chymotrypsinogen fractions were stored at 4°C.

Enzyme assay

Midguts were dissected and mixed with cold-buffer (15 mM acetic acid, 15 mM MES, 15 mM MOPS and 15 mM CHES, pH 9.0, containing 50 mM $CaCl_2$). Samples were homogenized using a hand-held plastic pestle, and sonicated for 1 min. After centrifugation at 15 000***g***, 4°C, for 15 min, the supernatant was used for enzymatic assays.

Chymotrypsin activity was measured using the synthetic substrate N-succinyl-ala-ala-pro-phe-nitroanilide (AAPF). Samples contained 440 μl of homogenization buffer and 10 μl of substrate (200 mM AAPF dissolved in DMSO). Reactions were initiated by the addition of 1 μl of midgut extract (equivalent to 1/15 of a midgut),

and carried out at room temperature for 5 min. Each reaction was repeated in triplicate. The rate of formation of the hydrolysis product, nitroaniline, was measured at 410 nm in a Beckman DU-40 spectrophotometer, and corrected for spontaneous hydrolysis of the substrate. Absorbancy readings were converted to amount of the product using an extinction coefficient of 10 000 M^{-1} cm^{-1}.

Western analysis

Purified $(his)_6$-chymotrypsinogen was used as an antigen to generate rabbit anti-*Ae. aegypti* chymotrypsin polyclonal antibodies (prepared by Cocalico Biological Inc., Reamstown, PA, U.S.A.). On immunoblots, the antibody, at a 1:5000 dilution, detected 50 ng of purified $(his)_6$-chymotrypsinogen, but did not react with recombinant early or late trypsins from *Ae. aegypti* (Barillas-Mury *et al.*, 1991; Noriega *et al.*, 1996a).

Midgut tissues from female mosquitoes (10 mosquitoes per pool) 1, 2, 3, 4, 5 and 6 days after emergence, and 1, 3, 5, 7, 13, 15, 16, 18, 20, 22, 24, 48 and 72 h after feeding, were homogenized in ice-cold 10 mM Tris–HCl, pH 7.6, containing 0.1 M NaCl, 1 mM ethylenediamine tetraacetic acid (EDTA), 1 μg/ml aprotinin and 100 μg/ml PMSF. Samples were sonicated (2 min) and centrifuged (10 000 rpm, 10 min at 4°C). Proteins were separated on 12.5% SDS–PAGE and transferred to nitrocellulose (Schleicher and Schuell, Keene, NH, U.S.A.). Chymotrypsin was detected using rabbit anti-*Ae. aegypti* chymotrypsin polyclonal antisera (diluted 1:5000) and goat anti-rabbit IgG conjugated with horseradish peroxidase (BioRad, Richmond, CA, U.S.A.).

Northern analysis

Total RNA from larvae, pupae, adult males, unfed adult females and adult females at different times after feeding an artificial meal were extracted using RNA binding glass powder (BIO 101, La Jolla, CA, U.S.A.), as previously described (Noriega and Wells, 1993). Northern analysis was carried out using an amount of total RNA equivalent to one mosquito (Noriega *et al.*, 1996b). The chymotrypsin cDNA and the ribosomal probes were labelled using a Random primer kit (BRL) and [α-^{32}P]dCTP (ICN, Irvine, CA, U.S.A.). The amount of radioactivity bound to individual samples on the filter was quantified using a Betascope (Betagene, Waltham, MA, U.S.A.).

RESULTS

Isolation of the chymotrypsin-like serine protease cDNA

The 670 bp PCR product, described in Materials and methods, was used as a probe to screen an *Ae. aegypti* midgut cDNA library. A 938 bp clone with an open reading frame of 804 bp was isolated (Fig. 1). The sequence encodes a pre-proenzyme with 268 amino acids and contains an 18-residue signal peptide sequence, followed by a seven-residue activation peptide. The deduced amino acid sequence of the mature enzyme unambiguously identifies it as a member of the serine protease family, as it contains the catalytic residues (His/Asp/Ser) (Barrett and Rawlings, 1995) and approximately 32% sequence identity to mammalian chymotrypsins, trypsins and elastases (Wilcox, 1970; Del Mar *et al.*, 1980; Craik *et al.*, 1984).

Comparison and analysis of the deduced protein sequence of the putative serine protease

A multiple sequence alignment of the amino acid sequence of the putative *Ae. aegypti* serine protease with the sequences of other chymotrypsins is shown in Fig. 2. The *Ae. aegypti* sequence (AAC) has the catalytic triad typical of serine proteases at His 57, Asp 102 and Ser 195 (following the bovine chymotrypsin numbering system). Asp 194, which forms a salt bridge with the N-terminal Ile in the mature bovine enzyme, and stabilizes the catalytic site, is also conserved. The *Ae. aegypti* enzyme lacks a negatively charged carboxylate (e.g. Asp) at position 189, which is present at the bottom of the substrate binding pocket of all known trypsins. Instead, it has a hydrophobic substrate binding pocket which resembles that of chymotrypsins. The *Ae. aegypti* protein also contains the conserved Ser 189 at the base of the pocket, and the conserved Gly 216 residue lining the entry of the pocket. The presence of a Thr at position 226 in the *Ae. aegypti* sequence, replacing the smaller residues (Gly or Ala) found in the vertebrate chymotrypsins, may imply a somewhat different substrate preference for the *Ae. aegypti* enzyme. Ser 190 and Tyr 228, which are the side chains that extend into the substrate binding pocket in vertebrate chymotrypsins, are changed to Ala and Phe, respectively, in the *Ae. aegypti* enzyme. This might also imply a change in substrate specificity.

Three conserved cysteine bridges, common to all known invertebrate chymotrypsinogens, are also found in analogous positions in the *Ae. aegypti* sequence. In vertebrate enzymes, two more pairs of cysteine bridges are present, Cys 1–Cys 122 and Cys 136–Cys 201. These are known, in bovine chymotrypsin α, to hold the three polypeptide chains together. The loss of these extra cysteine pairs in invertebrate chymotrypsins may reflect a different conformational change during zymogen activation and processing.

Four other residues, that are conserved in all known chymotrypsins, and are substituted in the *Ae. aegypti* chymotrypsin sequence, are marked as "⇓" in Fig. 2; the significance of these residues is unknown.

In vitro *transcription and translation of the chymotrypsin-like cDNA clone*

The full-length *Ae. aegypti* chymotrypsin cDNA was transcribed *in vitro*. This transcript was translated *in vitro* using a rabbit reticulocyte lysate. In the absence of microsomes, a pre-prochymotrypsin polypeptide product of ~34 kDa, was produced (Fig. 3, lane C). In the pres-

```
GCATCGAACAGTAATCG       17

ATG GCT TTC AAA CTC ACG GTA GCT TTC CTG CTC GTT GCC AGT TTG GCA CTC GCT TCC TCC   77
met ala phe lys leu thr val ala phe leu leu val ala ser leu ala leu ala Ser Ser   20

AGA GCC ACC CAC AAG ATC GTC GGT GGC GAT GAG GCC GAA GCG CAC GAA TTT CCC TAC CAA  137
Arg Ala Thr His Lys ILE VAL GLY GLY TYR GLU ALA GLU ALA HIS GLU PHE PRO TYR GLN   40

ATC TCG CTG CAG TGG AAC TTC AAC GAT GGA CAA ACG GAG ACC ATG CAC TTC TGC GGA GCT  197
ILE SER LEU GLN TRP ASN PHE ASN ASP GLY GLN THR GLU THR MET HIS PHE CYS GLY ALA   60

TCG GTG TTG AAC GAA AAC TTC GTC CTG ACG GCT GCT CAC TGC AAG ACC GCA TAC TCC AAT  257
SER VAL LEU ASN GLU ASN PHE VAL LEU THR ALA ALA HIS CYS LYS THR ALA TYR SER ASN   80

ACC GGG TTC ATC GAA GTG GTT GCC GCT GAA CAT GAT GTG GCC GTT GCG GAA GGA TCC GAA  317
THR GLY PHE ILE GLU VAL VAL ALA ALA GLU HIS ASP VAL ALA VAL ALA GLU GLY SER GLU  100

CAG CGT CGT TTG GTT GCG GAG TTC ATC GTC CAC GAG GAC TAT CAA GGA GGA GTC AGT CCC  377
GLN ARG ARG LEU VAL ALA GLU PHE ILE VAL HIS GLU ASP TYR GLN GLY GLY VAL SER PRO  120

GAT GAC ATT GCC GTC ATT CGT GTG GAC AAA CCC TTC GAA TTG AAC GAT AAG GTG AAG GCC  437
ASP ASP ILE ALA VAL ILE ARG VAL ASP LYS PRO PHE GLU LEU ASN ASP LYS VAL LYS ALA  140

GTT AAG CTG CCC AAG CAG TTG GAA CAA TTC GAT GGC GAT GTT ACT CTG AGT GGA TGG GGA  497
VAL LYS LEU PRO LYS GLN LEU GLU GLN PHE ASP GLY ASP VAL THR LEU SER GLY TRP GLY  160

TCC GTA TCG ACG ACG GTG TTC CCG GAC TAT CCT GAC AAA CTG AGG AAA GTC GTG CTT CCG  557
SER VAL SER THR THR VAL PHE PRO ASP TYR PRO ASP LYS LEU ARG LYS VAL VAL LEU PRO  180

CTG GTA GAC TAC GAA CAA TGT GAC ACC CTG TGG GGC AAC GAC AGT GCT CTA GCG AAG AGT  617
LEU VAL ASP TYR GLU GLN CYS ASP THR LEU TRP GLY ASN ASP SER ALA LEU ALA LYS SER  200

AAT GTC TGC GCT GGC CCG ATC GAT GGC TCC AAG TCG GCC TGC TCG GCT GAT TCA GGT GGC  677
ASN VAL CYS ALA GLY PRO ILE ASP GLY SER LYS SER ALA CYS SER ALA ASP SER GLY GLY  220

CCG TTG GTG AAA CAG TCC GGT GAA GAA GTG ATC CAG GTC GGT GTC GTG TCG TGG GGA GCC  737
PRO LEU VAL LYS GLN SER GLY GLU GLU VAL ILE GLN VAL GLY VAL VAL SER TRP GLY ALA  240

GTT CCA TGT GGA TCG CCA CGT CGT CCG ACC GTG TTT GCC GGA GTT TCC CAT TAC GTC GAT  797
VAL PRO CYS GLY SER PRO ARG ARG PRO THR VAL PHE ALA GLY VAL SER HIS TYR VAL ASP  260

TGG ATC GAG CAG CAG CTC CGT GCG TGAGAAAGAAACAGGACTCGGATGTGATGTTATTTGACGATGATGAC  868
TRP ILE GLU GLN GLN LEU ARG ALA                                                  268

CAATGCCTCTAAGCGGAATTAAAGTTATGGTGATGGCAACGAATTAAAACTTGAAAAAAAAAAAAAAAAA            938
```

FIGURE 1. Nucleotide sequence and deduced amino acid sequence of *Ae. aegypti* chymotrypsin cDNA. The 18 amino acids of the predicted signal sequence are in lower case and the following heptapeptide sequence is the putative activation sequence. The mature protein sequence (in capital letters) begins with Ile at residue 26. There are two possible polyadenylation sites, AATTAAA, underlined in the 3′ end of the sequence.

ence of microsomal membranes, post-translational processing yielded a smaller zymogen (M_r ~32 kDa) (Fig. 3, lane D).

Chymotrypsin gene and protein expression

Chymotrypsin gene expression was analysed by Northern blot analysis (Fig. 4, panel D). No chymotrypsin mRNA was detected in larvae, pupae and adult males (results not shown), however the transcript was readily detectable in unfed adult females. The chymotrypsin mRNA was abundant in the midgut, and was not detectable in carcasses (mosquitoes from which the midguts were removed, results not shown).

In adult females, the message could first be detected approximately 24 h after emergence, and increased dramatically within the next 48 h (Fig. 4, panel A). When a γ-globulin meal was ingested, the chymotrypsin mRNA remained at a nearly constant level throughout the following 22 h (Fig. 4, panel B).

Chymotrypsin protein expression in the midgut of unfed and fed females was determined by Western blot analysis using rabbit anti-*Ae. aegypti* chymotrypsin polyclonal antibody. The levels of immunoreactivity were quantified by densitometery (Fig. 4, panel C). Before feeding, chymotrypsin was detected at very low levels (Fig. 4, panel A). However, after feeding, the chymotrypsin protein was readily detectable within 3 h, reached a maximum value by 24 h and was below levels of detection 48 and 72 h after feeding (Fig. 4, panel B).

```
                                                 1          ∇1                                                        ⇓∇
RAC   IVNGEDAIPG SWPWQVSLQ. .....DKTGF HFCGGSLISE DWVVTAAHC. GVKTSD.... VVVAGEFDQG SDEENIQVLK IAQVFKNPKF NMFTVRNDIT
BOC   IVNGEDAVPG SWPWQVSLQ. .....DSTGF HFCGGSLISE DWVVTAAHC. GVTTSD.... VVVAGEFDQG LETEDTQVLK IGKVFKNPKF SILTVRNDIT
DOC   IVNGEDAVPG SWPWQVSLQ. .....DSTGF HFCGGSLISE DWVVTAAHC. GVRTTHQ... .VVAGEFDQG SDAESIQVLK IAKVFKNPKF NMFTINNDIT
GMC   IVNGEEAVPH SWSWQVSLQ. .....DQTGF HFCGGSLINE NWVVTAAHC. NVKNYH.... RVVLGEHDRS SNSEGVQVMT VGQVFKHPRY NGFTINNDIL
AAC   IVGGDEAEAH EFPYQISLQW NFNDGQTETM HFCGASVLNE NFVLTAAHC. KTAYSNTGFI EVVAAEHDVA VAEGSEQRRL VAEFIVHEDY QGGVSPDDIA
AGC1  VVGGEVAKNG SAPYQVSLQV P...GWG... HNCGGSLLND RWVLTAAHCL ..VGHAPGDL MVLVGTNSLK EGGEL...LK VDKLLYHSRY NLPRFHNDIG
AGC2  VVGGVVAKNC SAPYQVSLQV P...GWG... HNCGGSLLND RWVLTAAHCL ..VGYEPSDL MVLVGTNSLK EGGEL...LK VDKLLYHSRY NRPQFHNDIG
EHC   IVGGTDAPRG KYPYQVSLRA P.....K... HFCGGS.ISK RYVLTAAHCL ..VGKSKHQV TVHAGSVLLN KEEAV...YN AEELIVNKNY NSIRLINDIG
PVC1  IVGGVEATPH SWPHQAALFI D....DM... YFCGGSLISS EWVLTAAHC. ...MDGAGFV EVVLGAHNIR QNEASQVSIT STDFFTHENW NSWLLTNDIA
PVC2  IVGGVEATPH SWPHQAALFI D....DM... YFCGGSLISS EWVLTAAHC. ...MDGAGFV EVVLGAHNIR QNEASQVSIT STDFFTHENW NSWLLTNDIA
LCC   ITNGQDAVMG QFPYQVGLSL NLGNFKS... AWCGGSLIGN EWVLTAAHC. ...TDGVKSV TVFLGA.TYR TEAEVKYTVK PNDILIHPGW NNKTLKNDIS
MSC   IVGGSSSSVG QFPYQAGLVI TLPRGTA... A.CGGSLLSN RRVLTAAHCW WDGQNQASRF VVVLGSNRLF SGG...VRLE TRDIVMHGSW NPNLVRNDIA
      •  *                 *                   **   •       *•*****            *                         **

                                                 4                                2                    2    ⇓   3 ⇓ ∇
RAC   LLKLATPAQF SETVSAVCLP NVDDDFP..P GTVCATTGWG KTKYNAL.KT PEKLQQAALP IVSEADCKKS WG..SKITDV MTCAGA..SG VSSCMGDSGG
BOC   LLKLATPAQF SETVSAVCLP SADEDFP..A GMLCATTGWG KTKYNAL.KT PDKLQQATLP IVSNTDCRKY WG..SRVTDV MICAGA..SG VSSCMGDSGG
DOC   LLKLATPARF SKTVSAVCLP QATDDFP..A GTLCVTTGWG LTKHTNA.NT PDKLQQAALP LLSNAECKKF WG..SKITDL MVCAGA..SG VSSCMGDSGG
GMC   LVKLATPATL NMRVSPVCLA ETDDVFE..G GMKCVTSGWG LTRYNAA.DT PALLQQAALP LLTNEQCKKF WG..NKISDL MICAGA..AG ASSCMGDSGG
AAC   VIRVDKPFEL NDKVKAVKLP KQLEQFD..G DV..TLSGWG SVSTTVFPDY PDKLRKVVLP LVDYEQCDTL WGNDSALAKS NVCAGPIDGS KSACSADSGG
AGC1  LVRLEQPVQF SELVQSVEY. SEKAVP...A NATVRLTGWG RTS.AN.GPS PTLLQSLNVV TLSNEDCNKK GGDPGYTDVG HLCTLT.KTG EGACNGDSGG
AGC2  LMRLEQPVQF SELVQSVEY. LEKAVP...V NATVRLTGWG RTS.TN.GNV RTLLQSLNVV TLSNEDCKAK MGNPENVDFP DVCTLT.KAG EGACNGDSGG
EHC   LIRVSKDISY TQLVQPVKLP VSNTIK...A GDPVVLTGWG RIY.VN.GPI PNNLQQITLS IVNQQTCKFK ..HWGLTD.S QICTFT.KLG EGACDGDSGG
PVC1  LIRLPSPVSL NSNIKTVKLP SSDVS....V GTTVTPTGWG RPSDSA.SGI SDVLRQVNVP VMTNADCDSV YG...IVGDG VVCIDG.TGG KSTCNGDSGG
PVC2  LIKLPSPVSL NSNIKTVKLP SSDVA....V GTTVTPTGWG RPLDSA.GGI SDVLRQVDVP IMTNDDCDAV YG...IVGNG VVCIDS.EGG KGTCNGDSGG
LCC   LVKIPET.AY TALIQPVELP ALASSYPSFA GDEVIASGWG RISDSA.SGV TNYLQWARLE VISNAVCART YG..STITSS NLCVKT.PGG VSTCKGDSGG
MSC   MIRLPSNVGF NNNINVIALP SGSQLNLNFA GERAIASGFG RTRDGA..NI DGSLNHVTLD VIANNVCSRT FP..LLIQSS NICTSG.ANG RSTCHGDSGG
      ••••          •   •                  •*•*                  *•         •        *                    •*       •    *  ****

         4                      3          ⇓
RAC   PLVCQKDGVW TLAGIVSWGS GVC.STSTPA VYSRVTALMP WVQQILEAN. .
BOC   PLVCQKNGAW TLAGIVSWGS STC.STSTPA VYARVTALMP WVQETLAAN. .
DOC   PLVCQKDGAW TLVGIVSWGS GTC.STSTPG VYARVTKLIP WVQQILQAN. .
GMC   PLVCQKAGSW TLVGIVSWGS GTC.TPTMPG VYARVTELRA WVDQTIAAN. .
AAC   PLVKQSGEEV IQVGVVSWGA VPCGSPRRPT VFAGVSHYVD WIEQQLRA.. .
AGC1  PLVYE....G KLVGVVNFG. .VPCALGYPD GFARVSYYHD WVRTTMANNS K
AGC2  PLVYE....G KLVGVVNFG. .VPCGRGFPD GFARVSYYHE WVRTTMANNS .
EHC   PLVAN....G VQIGIVSYG. .HPCAVGSPN VFTRVYSFLD WIQKNQL... .
PVC1  PLNL..NG.. MTYGITSFGS SAGCEKGYPA AFTRVYYYLD WIQQKTGVTP .
PVC2  PLNL..NG.. MTYGITSFGS SAGCEVGYPD AFTRVYYYLD WIEQKTGVTP .
LCC   PLVLASSG.. VQVGLTSFGS ILGCEKGFPA AFTRVTSYLE WINEHTGISY .
MSC   PLAATRNNRP LLIGVTSFGH RDGCQRGHPA AFARVTSYDA WIRRNL.... .
      **               *    *              •  •*    *•
```

FIGURE 2. Multiple sequence alignment of chymotrypsins. Abbreviations are as follows: RAC, rat; BOC, bovine; DOC, dog; GMC, cod (*G. morhua*); AAC, mosquito (*Ae. aegypti*); AGC1 and AGC2, mosquito (*A. gambiae*); EHC, European hornet (*V. crabro*); PVC1 and PVC2, shrimp (*P. vannamei*); LCC, black fly (*L. cuprina*); MSC, tobacco hornworm (*M. sexta*). *Identical residue in all sequences; ●conserved residues identified by George *et al.* (1990); the residues of the catalytic triad are marked with ∇; residues unique to the *Ae. aegypti* chymotrypsin are marked with ⇓; three pairs of Cys conserved in both vertebrate and invertebrate chymotrypsins are labelled 1–3 — the extra Cys in vertebrate proteins are marked as 4.

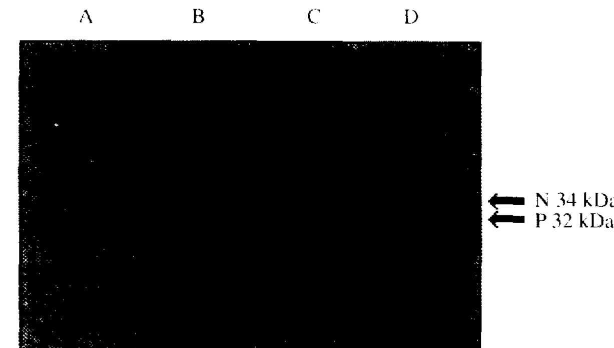

FIGURE 3. *In vitro* transcription and translation of the mosquito midgut chymotrypsin cDNA. The cDNA clone and the parental plasmid (as a negative control) were linearized and transcribed under control of the T_7 promoter. The transcripts were then translated using a rabbit reticulocyte lysate and [^{35}S]-L-Cys. The ^{35}S-labelled translational products were analysed by SDS–PAGE and autoradiography. Lane A is the negative control, lane B is a positive control using luciferase mRNA; lane C is the translation product produced from the chymotrypsin mRNA in the absence of microsomal membranes; lane D is the translation product produced from the chymotrypsin mRNA in the presence of microsomal membranes. N: non-processed form; P: processed form.

Chymotrypsin enzymatic activity

Enzymatic activity was measured in crude midgut extracts. (Fig. 4, panels A and B). In homogenates of midguts from unfed females there was no detectable chymotrypsin activity, whereas in homogenates of midguts from fed females there was enzymatic activity, which paralleled the changes in chymotrypsin protein concentration detected by Western blotting. Twenty-four hours after the meal, the chymotrypsin activity was 52.4 nmol/min/midgut, which, based on a specific activity of purified recombinant *Ae. aegypti* chymotrypsin against AAPF of 175 μmol/mg/min (Jiang, 1996), would amount to approximately 300 ng of protein per midgut. The hydrolysis of AAPF was not inhibited by soybean trypsin inhibitor (data not shown).

DISCUSSION

Chymotrypsin is the major midgut endoprotease present during larval development in *Ae. aegypti* (Yang and Davies, 1971; Kunz, 1978a), but the biological role of chymotrypsin during blood meal digestion in *Ae. aegypti* has been largely neglected. Yang and Davies

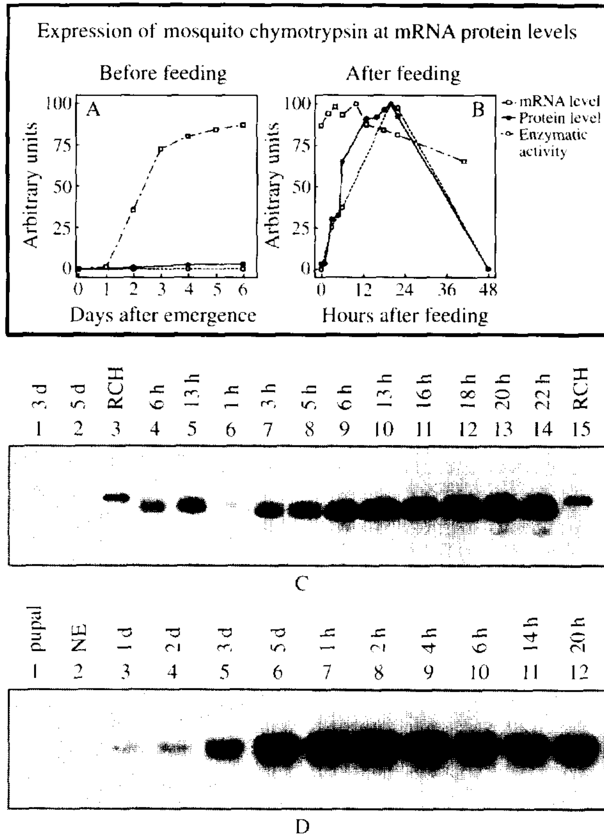

FIGURE 4. Pattern of midgut-specific chymotrypsin gene expression and protein synthesis. Chymotrypsin gene expression was determined by Northern blot analysis of total RNA from pupae, unfed adult females and adult females at different times after feeding an artificial blood meal. Each point represents the analysis of a group of 10 mosquitoes. Arbitrary units were calculated by first normalizing the cpm in chymotrypsin mRNA to the cpm in ribosomal RNA and then multiplying each value by 100/maximum value (10 h after feeding). Chymotrypsin protein expression in midguts from unfed adult females and from adult females at different times after feeding an artificial blood meal was determined by Western blot analysis (one midgut equivalent/lane), using rabbit anti-*Ae. aegypti* chymotrypsin polyclonal antibodies (1:5000 dilution). Protein levels were quantitated by densitometery. Arbitrary units were calculated by multiplying each value by 100/maximum value (22 h after feeding). Chymotrypsin enzymatic activity from midgut homogenates was measured using Suc-Ala-Ala-Pro-Phe-pNA. Arbitrary units were calculated by multiplying each value by 100/maximum value (22 h after feeding). Values are the mean of two individual assays. A: Levels before feeding; B: levels after feeding; C: Western blot; D: Northern blot. 1 d–5 d represent days after emergence; 1 h–22 h represent hours after feeding; NE: newly emerged female; RCH: recombinant chymotrypsin. Lanes 4 and 5 are from midguts in which lumen contents have been removed.

(1971), using N-glutaryl-L-phenylalanine *p*-nitroanilidine (GPNA) as a substrate, reported low chymotrypsin activity in unfed female adults, and noted that chymotrypsin activity did not rise after blood feeding. Briegel and Lea (1975), using benzoyl-tyrosine-ethyl ester (BTEE) as a substrate, found low chymotrypsin activity in crude midgut extracts, and also reported that this low activity could be further reduced using the chymotrypsin inhibitor TPCK. Kunz (1978b) noted that midgut extracts hydrolysed tosyl-arginine-methyl ester (TAME, a trypsin substrate) approximately 10 times more rapidly than BTEE. Borovsky and Schlein (1988), using an assay based on the binding of ^{3}H-DFP to crude midgut extracts and specific inhibition of the binding with TPCK and/or TLCK, concluded that 77% of the endoproteolytic activity was trypsin-like and less than 20% was chymotrypsin-like. All these reports supported the view that trypsin is the primary endoprotease present in the midgut of the female *Ae. aegypti* after a blood meal, and that chymotrypsin does not contribute significantly to the endoproteolytic cleavage of blood proteins (Clements, 1992).

Most of the previous studies measured chymotrypsin activity using BTEE. However, the specific activity of purified recombinant *Ae. aegypti* chymotrypsin against BTEE (0.66 μmol/mg/min) is about 3% of that of bovine chymotrypsin, whereas its specific activity against AAPF (175 μmol/mg/min) is six-fold higher than that for bovine chymotrypsin (Jiang, 1996). In practical terms, this means that using AAPF as the substrate increases the sensitivity of the assay at least 200-fold. This difference in sensitivity between assays using AAPF and BTEE explains why we were able to measure significant amounts of chymotrypsin activity compared with previous studies. The amount of chymotrypsin present in the midgut, approximately 300 ng at 24 h, is comparable to the amount of early trypsin present (Pennington *et al.*, 1995), but far less than the amount of late trypsin (Graf *et al.*, 1988). The induction of chymotrypsin synthesis and activity by blood feeding have recently been reported in *Anopheles gambiae* (Horler and Briegel, 1995; Muller *et al.*, 1995).

We have identified a full-length cDNA clone that encodes an adult female-specific midgut chymotrypsin. The deduced amino acid sequence of *Ae. aegypti* chymotrypsin showed several structural features typical of chymotrypsin proteases (Wilcox, 1970), including the catalytic triad of serine proteases and the amino acid residues considered to determine chymotrypsin substrate specificity. It also contains six cysteine residues located at conserved positions, common to all invertebrate chymotrypsins. From the results of the *in vitro* transcription and translation experiments, we infer that the *Ae. aegypti* chymotrypsin cDNA contains all the signals required for translation and the post-translational processing expected for a protein that is secreted.

The regulation of chymotrypsin synthesis in *Ae. aegypti* midgut is different from that of other adult female-specific endoproteases. Like early trypsin (Noriega *et al.*, 1996a, b), the chymotrypsin gene is first transcribed at the beginning of adult life; its mRNA reaches its maximal level before feeding, and the protein product is not translated until the ingestion of a meal. Unlike early trypsin, the chymotrypsin mRNA does not decrease after feeding and the chymotrypsin protein is produced continuously during meal digestion. The fact that protein expression and enzymatic activity in crude extracts increase after the ingestion of a protein meal, suggests an involvement in protein digestion.

In summary, we reported the cDNA and deduced amino acid sequences of a chymotrypsin from the midgut of *Ae. aegypti*. We described the patterns of gene expression, following adult eclosion and protein synthesis, after a blood meal. The pattern of gene expression following adult eclosion is similar to that of the early trypsin gene. However, the changes in protein synthesis and mRNA levels after a blood meal are different from those of either early or late trypsin. It is thus becoming evident that the regulation of protease synthesis in the female *Ae. aegypti* midgut after a blood meal is complex — three different patterns have been observed for the three different endoproteases studied.

REFERENCES

Barillas-Mury C., Graf R., Hagedorn H. H. and 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.

Barillas-Mury C., 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.

Barrett A. J. and Rawlings N. D. (1995) Perspectives in biochemistry and biophysics: families and clans of serine peptidases. *Arch. Biochem. Biophys.* **318,** 247–250.

Borovsky D. and Schlein Y. (1988) Quantitative determination of trypsin-like and chymotrypsin-like enzymes in insects. *Arch. Insect Biochem. Physiol.* **8,** 249–260.

Briegel H. and Lea A. O. (1975) Relationship between protein and proteolytic activity in the midgut of mosquitoes. *J. Insect Physiol.* **21,** 1597–1604.

Clements, A. N. (1992) *The Biology of Mosquitoes*, Vol. 1. Chapman and Hall, London.

Craik C. S., Choo Q. L., Swift G. H., Quinto C., Macdonna R. J. and Rutter W. J. (1984) Structure of two related rat pancreatic trypsin genes. *J. Biol. Chem.* **259,** 14255–14264.

Del Mar E. G., Largman G., Brodric J. W., Fasset M. and Geokas M. C. (1980) Substrate specificity of human pancreatic elastase 2. *Biochemistry* **19,** 468–472.

Felix C. R., Betschart B., Billingsley P. F. and 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.

Feng D. F. and Doolittle R. F. (1987) Progressive sequence alignment as a prerequisite to correct phylogenetic trees. *J. Molec. Evol.* **25,** 351–360.

George D. G., Barker W. C. and Hunt L. T. (1990) Mutation data matrix and its uses. *Methods Enzymol.* **183,** 333–351.

Gooding R. H. (1969) Studies on proteinases from some blood-sucking insects. *Proc. Entomol. Soc. Ontario* **100,** 139–145.

Graf R. and Briegel H. (1982) Comparison between aminopeptidase and trypsin activity in blood-fed females of *Aedes aegypti*. *Rev. Suisse Zool.* **89,** 845–850.

Graf R., Binz H. and Briegel H. (1988) Monoclonal antibodies as probes for *Aedes aegypti* trypsin. *Insect Biochem.* **18,** 463–470.

Hall M., Scott T., Sugumaran M., Söderhäll K. and Law J. H. (1995) Proenzyme of *Manduca sexta* phenol oxidase: purification, activation, substrate specificity of the active enzyme, and molecular cloning. *Proc. Natl. Acad. Sci. U.S.A.* **92,** 7764–7768.

Henikoff S. (1984) Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. *Gene* **28,** 351–359.

Horler E. and Briegel H. (1995) Proteolytic enzymes of female *Anopheles*: biphasic synthesis, regulation and multiple feeding. *Arch. Insect Biochem. Physiol.* **28,** 189–205.

Jiang, Q. (1996) Cloning and characterization of midgut-specific gene/gene products in the mosquito *Aedes aegypti*. PhD dissertation. University of Arizona.

Kalhok S. E., Tabak L. M., Prosser D. E., Brook W., Downe A. E. and White B. (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.

Kunz P. A. (1978a) Resolution and properties of the proteinases in the larva of the mosquito *Aedes aegypti*. *Insect Biochem.* **8,** 43–51.

Kunz P. A. (1978b) Resolution and properties of the proteinases in adult *Aedes aegypti*. *Insect Biochem.* **8,** 169–175.

Muller H.-M., Catteruccia F., Vizioli J., Della Torre A. and Crisanti A. (1995) Constitutive and blood-meal induced trypsin genes in *Anopheles gambiae*. *Exp. Parasitol.* **81,** 371–385.

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. and Wells M. A. (1994) Dietary control of late trypsin gene transcription in *Aedes aegypti*. *Insect Biochem. Molec. Biol.* **24,** 627–631.

Noriega F. G., Wang X., Pennington J. E., Barillas-Mury C. and Wells M. A. (1996a) 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.

Noriega F. G., Pennington J. E., Barillas-Mury C., Wang X.-Y. and Wells M. A. (1996b) *Aedes aegypti* midgut early trypsin is post-transcriptionally regulated by blood feeding. *Insect Molec. Biol.* **5,** 25–29.

Pennington J. E., Noriega F. G. and Wells M. A. (1995) The expression of early trypsin in *Aedes aegypti*. *J. Cell Biochem.* **21A (Suppl.),** 211.

Sanger F., Nicklen S. and Coulson A. R. (1977) DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* **74,** 5463–5467.

Wilcox P. E. (1970) The serine proteases. *Methods Enzymol.* **19,** 64–108.

Yang Y. J. and Davies D. M. (1971) Trypsin and chymotrypsin during metamorphosis in *Aedes aegypti* and properties of the chymotrypsin. *J. Insect. Physiol.* **17,** 117–131.

Acknowledgements—We thank John Norem for insect care. 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).