This is to certify that the

dissertation entitled

Characterization and cDNA Cloning of Proteases Involved in Mosquito Vitello--genesis

presented by

CHO, WEN-LONG

has been accepted towards fulfillment
of the requirements for

Ph. D. degree in Genetics

Alexander S. Raikhel

Dr. Alexander S. Raikhel
Major professor

Date May 20, 1992

# CHARACTERIZATION AND cDNA CLONING OF PROTEASES INVOLVED IN THE MOSQUITO VITELLOGENESIS

By

Wen-Long Cho

A DISSERTATION

Submitted to

Michigan State University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

Genetics Graduate Program

1992

## ABSTRACT

# CHARACTERIZATION AND cDNA CLONING OF PROTEASES INVOLVED IN THE MOSQUITO VITELLOGENESIS

By

Wen-Long Cho

In the mosquito, *Aedes aegypti*, a lysosomal aspartic protease (mLAP), which is involved in the termination of vitellogenesis in the fat body, was purified and characterized. The native molecular weight of the purified mLAP determined by PAGE under non denaturing conditions was 80,000. The enzyme was resolved into a single 40 kDa peptide with SDS-PAGE under either reducing or non-reducing conditions. This mosquito LAP has a pI of 5.4 and its optimal condition for enzymatic reaction is pH 3.0 at 45$^{o}$C using hemoglobin as a substrate. The cDNA corresponding to mLAP was isolated and sequenced. The deduced amino acid sequence exhibits 92% and 81% similarity to human cathepsin D (HCD) and E (HCE) respectively. Kinetic analysi of mLAP in the mosquito fat body at the mRNA and protein levels revealed a 12 hr lag in its translation. Translational regulation of mLAP mRNA may occur. The 5'-untranslated region of mLAP mRNA is similar to elements conferring negative translational control by steroids.

A cDNA encoding a mosquito vitellogenic 53 kDa protein was cloned and sequenced. This cDNA hybridizes to a 1.5 kb mRNA present only in the fat body of vitellogenic females. The analysis of the deduced amino acid sequence from this cDNA indicates that it is a serine carboxypeptidase.

The enzymatic property of this mosquito vitellogenic carboxypeptidase (VCP) was identified using an assay with a serine protease inhibitor, [$^{3}$H] diisopropyl fluorophosphate. The finding that an enzyme required for embryonic development is synthesized outside of oocytes represents a biological phenomenon previously unknown.

## ACKNOWLEDGMENTS

I sincerely appreciate Dr. Alex Raikhel's guidance, support, and encouragements . I would like to thank the members of my committee, Dr. Tom Friedman, Dr. Will Kopachik, Dr. Natasha Raikhel and Dr. Charles Sweeley, for their comments, suggestions and careful reading of my manuscripts. I would like to extend my special thanks to Dr. Tochi Dhadialla and Mr. Alan Hays for their expert technical instructions in protein biochemistry. I appreciate the help in reviewing manuscripts from my fellow students, Kirk Deitsch and Neal Dittumer. I would also like to thank all the people from Alex's laboratory for their friendship.

Finally, I would like to thank my mother for her understanding and my wife, Chuen-Hui Su whom I owe my greatest appreciation for her love and making my life more comfortable during this period.

## CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

**CHAPTER 4**

## ABBREVIATIONS

**CAT**..... Chloramphenicol acetyltransferase

**CNP**..... 2', 3'-cyclic nucleotide 3'-phosphodiesterase

**DFP**..... Diisopropyl fluorophosphate

**DTT**..... 1,4-Dithiothreitol

**ECL**..... Enhanced chemiluminescence

**EDTA**..... Ethylenedinitrilo tetraacetic acid

**GCG**..... Computsoftware from University of Wisconsin Genetics Computer Group

**HCD**..... Human cathepsin D

**HCE**..... Human cathepsin E

**20-HE**..... 20-hydroxyecdysone

**IEF**..... Isoelectric focusing

**kDa**..... kilo-Dalton

**MBP**..... Myelin basic protein

**mLAP**..... Mosquito lysosomal aspartic protease

**PBM**..... Post blood meal

**PCR**..... Polymerase chain reaction

**pI**..... Isoelectric point

**SDS**..... Sodium dodecyl sulfate

**SDS-PAGE**..... SDS polyacrylamide gel electrophoresis

**VCP**..... Vitellogenic carboxypeptidase

**Vg**..... Vitellogenin

**WCP**..... Wheat carboxypeptidase

# CHAPTER 1

# INTRODUCTION

## 1. *Mosquito-borne diseases:*

Mosquito-borne human diseases remain a major international health problem. It is estimated that there are 200 million new cases of malaria annually (Kemp et al., 1987). Another disease, filariasis, carried by mosquitoes, affects more than 100 million people annually. In addition, many arboviral diseases, such as yellow fever, dengue fever and encephalitis, mediated by mosquito transmission are still jeopardizing human life (Tabachnick, 1991).

Malaria is a mosquito-borne infection caused by protozoa of the genus *Plasmodium*. Four species of malaria infect human: *P. falciparum, P. vivax, P. ovale* and *P. malariae*. Control measures directed at the parasite itself (drug therapy) or the mosquito (insecticide) have eradicated or controlled malaria in some countries, but have failed in much of the tropical world. Most forms of parasites are intracellular and therefore not directly accessible to the human immune system. Because of antigenic variation and /or strain difference, an effective vaccine is still being developed (Kemp et al., 1987).

Many arboviruses, such as dengue and yellow fever viruses, appear to be capable of indefinite survival in insects by transovarial transmission from one generation to the next or to be maintained in infection cycles (Fig. 1) without the intervention of the natural vertebrate or invertebrate hosts. (White and Fenner, 1986).

In addition to dengue fever and yellow fever, many other arbovirus infections causing human diseases are mediated by mosquitoes as shown in table 1 (White and Fenner, 1986). Control of arboviral disease rests upon (1) vector control and (2) vaccination. Temporally, these two strategies effectively control most of lethal encephalitis and hemorrhagic fevers. Nevertheless, the insecticide resistance of mosquitoes and side effects of vaccination are still potential problems for control of these diseases.

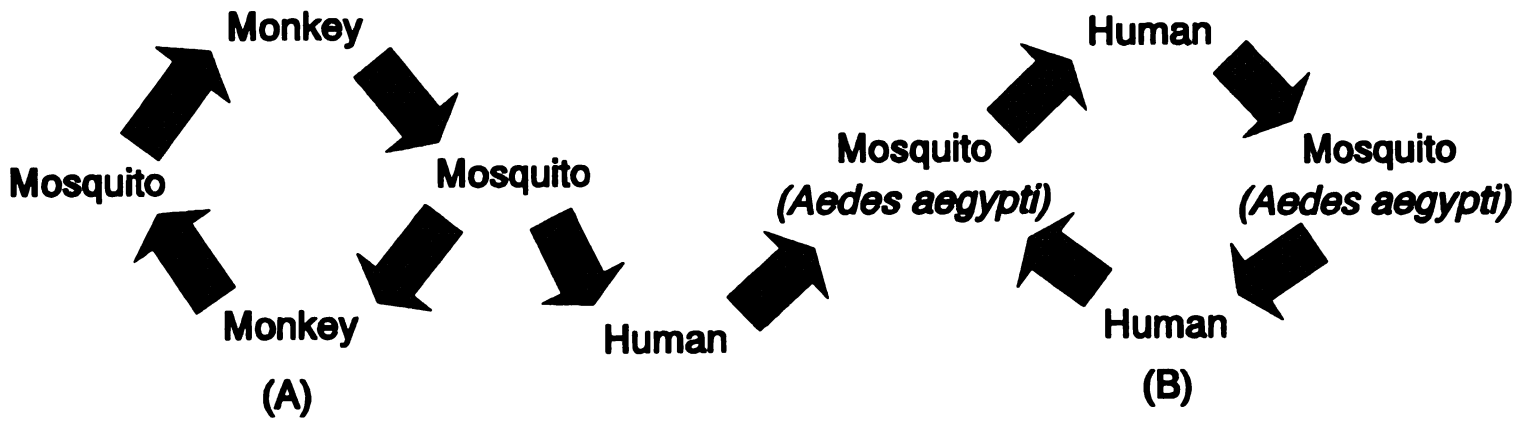



Fig. 1 Jungle (A) and urban (B) cycles of yellow fever and dengue fever.

**Table 1. Mosquito-borne diseases mediated by arboviruses.**

| Human Alphaviruses | | | | |
|---|---|---|---|---|
| VIRUS | DISTRIBUTION | INVERTEBRATE Vector | VERTEBRATE RESERVIOR | DISEASE IN MAN |
| Chikungunya | Asia,Africa | Mosquito | Man | Fever,arthritis, hemorrhagic fever |
| O'nyong-nyong | Africa | Mosquito | Man | Fever,arthritis |
| Ross River | Australia, Pacific | Mosquito | Marsupials | Arthritis |
| Sindbis | Africa,Asia, Australia | Mosquito | Birds | Fever |
| Easter equine encephalitis | Americas | Mosquito | Birds | Encephalitis |
| Western equine encephalitis | Americas | Mosquito | Birds | Encephalitis |
| **Human Flaviviruses** | | | | |
| VIRUS | DISTRIBUTION | INVERTEBRATE VECTOR | VERTEBRATE RESERVOIR | DISEASE IN MAN |
| Yellow fever | Africa, South America | Mosquito | Monkeys, man | Hepatities, hemorrhagic fever |
| Dengue 1-4 | Asia, Pacific, Caribbean, Africa | Mosquito | Monkeys, man | Fever, hemorrhagic fever |
| Japanese encephalitis | East and SE Asia | Mosquito | Birds, pigs | Encephalitis |
| Murray valley encephalitis | Australia, New Guinea | Mosquito | Birds | Encephalitis |
| West Nile fever | Africa, Europe, Middle East | Mosquito | Birds | Fever, encephalitis |
| St. Louis encephalitis | Americas | Mosquito | Birds | Encephalitis |
| Rocio | Brazil | Mosquito | Birds | Encephalitis |
| Venezuelan equine encephalitis | Americas | Mosquito | Rodents, horses | Encephalitis |

An understanding of the molecular basis of major physiological processes in mosquitoes could lead to develop novel strategies in vector management.

## 2. *Vitellogenic cycle of mosquito, Aedes aegypti:*

The maintenance and dispersal of mosquito-borne diseases depend upon the successful reproduction of mosquitoes. The foundation of the reproductive cycle is vitellogenesis involving massive production of yolk protein precursors and their accumulation by developing oocytes. The vitellogenesis of mosquitoes is triggered by a blood meal and, as a consequence, is linked to transmission of pathogens. Therefore, elucidation of aspects related to vitellogenesis is critical for the successful development of vector control.

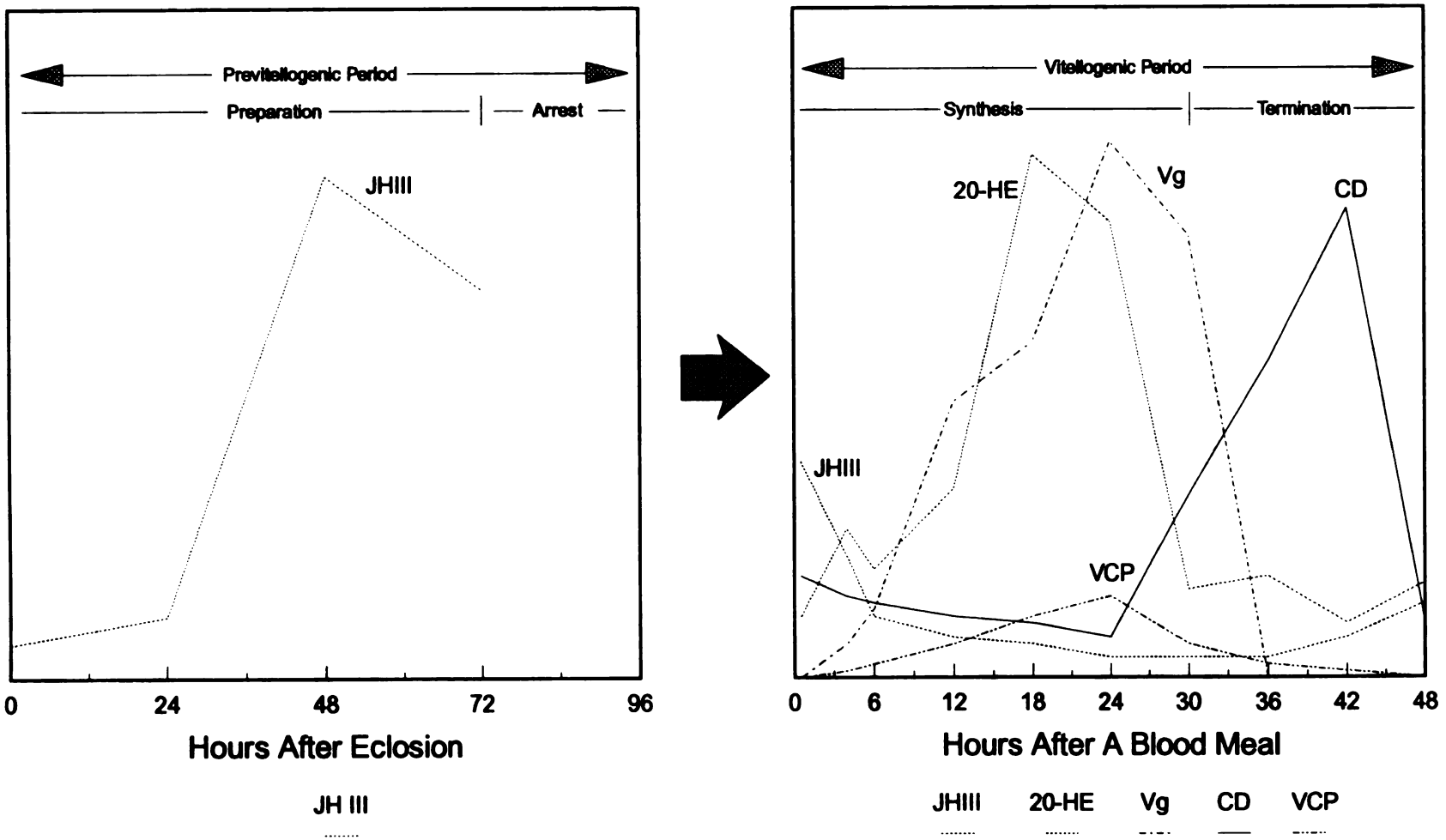
The mosquito fat body is a major organ involved in the synthesis of macromolecules and the metabolism of nutrients and other compounds. These functions are hormonally controlled and change successively according to the demands of the insect at different stages. In adult mosquitoes, the fat body is located in the head, thorax and predominantly in the abdomen. The primary important function of the fat body of an oogenic female is the synthesis of the yolk protein precursor, vitellogenin (Vg) and other proteins crucial for oocyte maturation.

In the female of the yellow fever mosquito, *Aedes aegypti,* used for this study, egg maturation proceeds through two developmental phases (Fig 2). The previtellogenic period begins at the emergence of the adult female and it is completed within 60-72 hr. During this phase, both fat bodies and ovaries become competent for subsequent vitellogenesis. The female then enters a state-of-arrest. The ingestion of vertebrate blood initiates the vitellogenesis.

**Fig. 2. Summary of events during the first cycle of egg maturation in the anautogeneous mosquito, *Aedes aegypti*.**

The first cycle proceeds through two developmental periods. The previtellogenic period begins at the eclosion of the adult female and it includes preparation stage and arrest stage. The vitellogenic period is initiated by a blood meal. It is divided into a synthetic stage (synthesis) and a termination stage (termination).

The relative concentrations of Juvenile hormone III (JHIII) and 20-hydroxyecdysone (20-HE) measured in mosquito whole bodies are indicated in both panels. In fat bodies, the cathepsin D-like activity (CD), the rate of vitellogenic carboxypeptidase (VCP) synthesis and the rate of vitellogenin (Vg) synthesis, are relatively presented in the right panel. The big arrow indicates that the developmental cycle continues from previtellogenic stage to vitellogenic stage when the mosquito receives a blood meal. (With modifications from Hagedorn *et al*; Raikhel, 1992 and Shapiro *et al.*, 1986)


Previtellogenic Period
Preparation
Arrest
JHIII
0
24
48
72
96
Hours After Eclosion
JH III
Vitellogenic Period
Synthesis
Termination
20-HE
Vg
CD
JHIII
VCP
0
6
12
18
24
30
36
42
48
Hours After A Blood Meal
JHIII
20-HE
Vg
CD
VCP

The vitellogenic period proceeds for 48 hr and it is divided into a synthesis stage and a termination stage based on the Vg synthesis in the fat body (Raikhel, 1992).

### 3. *Hormonal regulation of fat body activities during vitellogenesis:*

In the adult mosquito of *A. aegypti*, the levels of juvenile hormone III (JHIII) produced by the corpora allata (CA) were measured in whole bodies (Shapiro *et al.*, 1986). The amount of JHIII rises from 0.7 ng/g body weight (gbw) to 7.5 ng/gbw in 2 days after eclosion, then declines slowly over 5 days in non-blood fed mosquitoes. After a blood meal, the levels of JHIII decrease rapidly to 2.3 ng/gbw during the first 3 hr, followed by a slow decrease to the lowest point (0.4 ng/gbw) at 24 hr PBM. From 48 hr PBM on, JHIII starts to rise again until 96 hr. (Fig 2). Moreover, JHIII is shown to stimulate the increase of ploidy and the proliferation of ribosomes of fat bodies (trophocytes) during previtellogenic development (Dittmann *et al.*, 1989; Raikhel and Lea, 1990).

The ovaries of adult vitellogenic mosquitoes synthesize and secrete ecdysone which is converted to an active hormone, 20-hydroxyecdysone (20-HE), by fat bodies (Hagedorn, 1989). The levels of ecdysteroid in *A. aegypti* were measured in whole body preparations by Hagedorn et al. (1975). In mosquitoes, the levels of ecdysteroid are low during the first 8-10 hr after a blood meal with only a small peak at 4 hr PBM. Thereafter, it rises dramatically to the highest level at 16-20 hr PBM, then declines rapidly to previtellogenic levels (Fig. 2).

In mosquitoes, *A. aegypti*, 20-HE was proposed as the primary factor that initiates the Vg synthesis in fat bodies (Fallon, 1986; Hagedorn and Fallon, 1973; Hagedorn *et al.*, 1973). Dhadialla and Raikhel (1990) used an *in vitro* system to verify that Vg production can be initiated by physiological doses of 20-HE. This

hormone affects the transcription of Vg mRNA and resulted in greater accumulation of pro-Vg than its mature subunits (Dhadialla and Raikhel, 1990).

The recent finding of elements similar to *Drosophila* ecdysteroid response elements in regulatory sequences of Vg genes of *Anopheles* and *Aedes* (Romans and Miller, unpublished; Romans and Hagedorn, unpublished) strengthens the proposed role of 20-HE in Vg gene expression.

4. ***Biosynthesis of mosquito vitellogenin in the fat body:***

The major protein synthesized by fat bodies for egg development is Vg. Hagedorn *et al.*(1973, 1985) demonstrated that Vg synthesis in the mosquito fat bodies starts at 3-4 hr post blood meal (PBM), reaches the highest level at 24-28 hr then declines to background levels by 36-40 hr PBM (Fig 2).

The stages in the biosynthesis of Vg, starting from mRNA to the formation of the mature secreted protein, have been elucidated (Bose and Raikhel, 1988; Dhadialla and Raikhel, 1990; Raikhel *el al.*, 1990). The proposed pathway for Vg biosynthesis in the mosquito fat body is shown in Fig. 3. The precursor of primary translated Vg has a $M_r$ = 224,000. It is processed through a series of complicated modifications, such as glycosylation, phosphorylation, proteolytic cleavage and sulfation then it is transformed into mature secreted Vg subunits of 200- and 66- kDa.

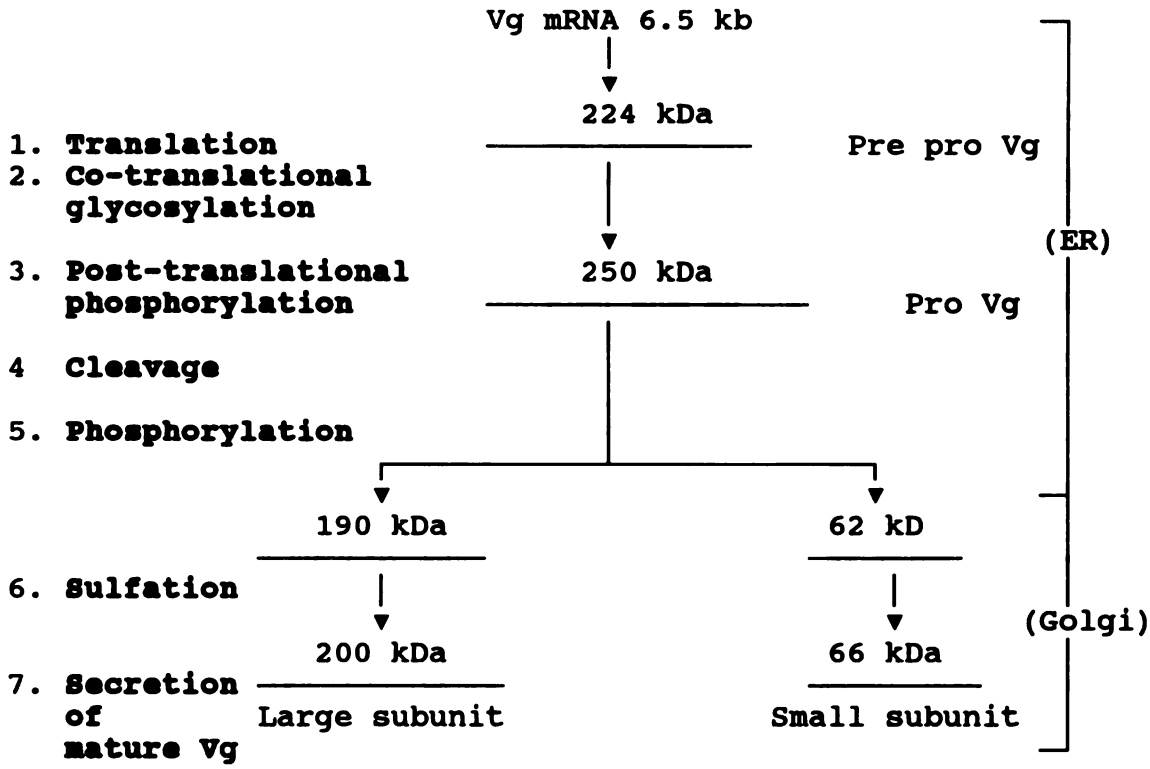



**Fig. 3. A schematic represntation of biosynthetic pathway for vitellogenin in mosquito fat bodies** (From Dhadialla and Raikhel, 1990).

**5. *Identification of an additional novel vitellogenic protein:***

Analysis of proteins secreted by mosquito fat bodies has demonstrated that several proteins, in addition to Vg, are synthesized and secreted into the hemolymph during vitellogenesis. One of these proteins is accumulated by ovaries similar to Vg (Hays and Raikhel, 1990).

This protein has a molecular weight of 53,000 as resolved by SDS-PAGE under reducing conditions. It is glycosylated, but not detectably phosphorylated or sulfated. Radioimmunoassay, using antibodies against this 53-kDa protein, has demonstrated that it is synthesized only by the fat body of vitellogenic female mosquitoes. Both the synthesis and secretion of this protein could be stimulated by a physiological dose of 20-HE in the previtellogenic female fat body cultured *in vitro*. The properties of this sex-, stage-, and tissue- specific protein, including its size, kinetics, and regulation of its synthesis drew our attention to investigate its expressional control during vitellogenesis.

**6. *Lysosomal activity in the fat body:***

The lysosomal system plays an important role in the turnover and catabolism of various subcellular components (Locke and Collins, 1980; Brainton, 1981; Farquhar and Palade, 1981; Glaumann *et al.*, 1981). During metamorphosis and egg maturation of insects, when the fat body switches from one function to another, lysosomal activity of fat bodies rises sharply and cellular organelles undergo massive autophagical degradation (Locke and Collins, 1980; Raikhel and Lea, 1983; Dean *et al.*, 1985; Raikhel, 1986a, 1986b).

Activities of several lysosomal enzymes were found to increase in the mosquito fat body during the cessation of vitellogenesis (Fig. 2: Raikhel, 1986a; 1986b). Moreover, analysis using video-enhanced fluorescent microscopy and electron microscopy revealed that lysosomes execute two important functions

during the termination of Vg synthesis: (1) interruption of Vg secretion by degrading the Vg-containing secretory granules, (2) destruction of the biosynthetic machinery, rough endoplasmic reticulum (RER) and Golgi complex, for trophocyte remodelling (Raikhel, 1986a; 1986b). Regulation of lysosomal activity is stimulated by the high titer of Vg concentration in hemolymph at the termination stage of vitellogenesis (Raikhel, 1986c). In contrast, 20-HE was shown to initiate lysosomal activity in the fat body of immature insects (Sass and Kovacs 1980; Locke and Collins, 1980; Tojo *et al.*, 1981). Therefore, 20-HE may also be involved in the stimulation of lysosomal activity in fat bodies of adult mosquitoes.

To understand the regulation of lysosomal activity in the female fat body at molecular level, we chose a lysosomal enzyme, cathepsin D, which is well characterized in the vertebrate system, for this project. Therefore, we decided to purify a cathepsin D-like enzyme from mosquitoes then characterized it and analyzed its expression in the fat body. Moreover, we planed to investigate the expression and regulation of this enzyme at the molecular level through the cDNA cloning, RNA analysis and studies of control mechanisms.

## REFERENCES


Bose, S. G. and Raikhel, A. S. (1988) *Biochem. Biophys. Res. Comm.* **155**, 436-42.

Brainton, D. F. (1981) *J. Cell Biol.*, **91**, 66s-76s.

Dean, R. L., Locke, M. and Collins, D. V. (1985) In: *Comprehensive Insect Physiology Biochemistry and Pharmacology* (Eds.) Kerkut, G. A. and Gilbert, L. I., Pergamon Press, Oxford, New York, Toronto, Sydney, Paris, Frankfurt. pp. 155-210.

Dhadialla, T. S. and Raikhel, A. S. (1990) *J. Biol. Chem.* **265**, 9924-9933.

Dittmann, F., Kogan, P. H., and Hagedorn, H. H. (1989) *Arch. Insect Biochem. Physiol.*, **12**, 133-143.

Fallon, A. M. (1986) *Exp. Cell Res.* **166**, 535-542.

Farquhar, M. G. and Palade, G. E. (1981) *J. Cell Biol.*, **91**, 77s-103s.

Glaumann, H., Ericsson, J. L. T., and Marzella, L. (1981) *Int. Rev. Cytol.* **73**, 149-182.

Hagedorn, H. H., O'Conner, J. D., Fuchs, M. S., Sage, B., Schlaeger, D. A., and Bohm, M. K. (1975) *Proc. Natl. Acad. Sci.* USA **72**, 3255-3259.

Hagedorn, H. H. (1985) In: *Comprehensive Inseect physiology, Biochemistry, and Pharmacology*, (eds) Kerkut, G. A., Gilbert, L. I., Vol. 8, pp.205-261, Pergamon Press, Oxford.

Hagedorn, H. H., Fallon, A. M., and Laufer, H. (1973) *Develop. Biol.*, **31**, 285-294.

Hagedorn, H. H. (1989) In: *Ecdysone from Chemistry to Mode of Action* (ed.) Koolman, J. Thieme Medical Publishers, Inc., New York, pp. 279-289.

Hagedorn, H. H. and Fallon, A. M. (1973) *Nature*, 244, 103-105.

Hays, A. R. and Raikhel, A. S. (1990) *Roux's Arch. Dev. Biol.*, **199**, 114-121.

Kemp, D. J., Coppel, R. L. and Ander, R. F. (1987) *Ann. Rev. Microbiol.* **41**, 181-208.

Locke, M. and Collins, J. V. (1980) In: *Pathobiology of Cell Membrane*, Academic Press, New York, Vol. 2, pp. 223-248.

Raikhel, A. S. (1986a) *J. Insect Physiol.* **32**, 297-604.

Raikhel, A. S. (1986b) *Tissue & Cell,* **18**, 125-142.

Raikhel, A. S. (1986c) In: *Host Regulated Developmental Mechanisms in Vector Arthropods.* (Eds.) Borovsky and A. Spielman, University of Florida, pp25-31.

Raikhel, A. S. and Lea, A. O. (1983) *Tissue & Cell*, **15**, 281-300.

Raikhel, A. S. and Lea, A. O. (1990) *Gen. Comp. Endocrinol.* **77**, 423-434.

Raikhel, A. S., Dhadialla, T. S., Cho, W. -L., Hays, A. R. and Koller, C. N. (1990) In: *Molecular Insect Science* (eds.) Hagedorn, H. H. Hildebrand, J. G., Kidwell, M. G., and Law, J. H. Plenum Press, New York, pp. 147-154.

Raikhel, A. S. (1992) *Adv. Disease Vector Res.*, **9**, (in press).

Sass, M. and Kovacs, J. (1980) *J. Insect Physiol.* **26**, 569-577.

Shapiro, A. B., Wheelock, G. D., Hagedorn, H. H., Baker, F. C., Tsai, L. W., and Schooley, D. A. (1986) *J. Insect Physiol.*, **32**, 867-877.

Tabachnick, W. J. (1991) *American Entomologist*, **37**, 14-24.

Tojo, S., Kiguchi, K. and Kimura, S. (1981) *J. Insect Physiol.* **27**, 491-

Wite, D. O. and Fenner, F. J. (1986) In: *Medical Virology*, Academic Press, Edition 3, pp. 479-506.

# CHAPTER 2*

# PURIFICATION AND CHARACTERIZATION OF A LYSOSOMAL ASPARTIC PROTEASE WITH CATHEPSIN D ACTIVITY FROM THE MOSQUITO

*Originally published in Insect Biochem.

Reference: Cho, W. -L., Dhadialla, T. S. and Raikhel, A. S. (1991) Insect Biochem. **21**, 165-176.

## ABSTRACT

A lysosomal aspartic protease with cathepsin D activity, from the mosquito, *Aedes aegypti*, was purified and characterized. Its isolation involved ammonium sulfate (30%-50%) and acid (pH 2.5) precipitations of protein extracts from whole previtellogenic mosquitoes followed by cation exchange chromatography. Purity of the enzyme was monitored by SDS-PAGE and silver staining of the gels. The native molecular weight of the purified enzyme as determined by polyacrylamide gel electrophoresis under non-denaturing conditions was 80,000. SDS-PAGE resolved the enzyme into a single polypeptide with $M_r$=40,000 suggesting that it exists as a homodimer in its non-denatured state. The pI of the purified enzyme was 5.4 as determined by isoelectric focusing gel electrophoresis. The purified enzyme exhibits properties characteristic of cathepsin D. It utilizes hemoglobin as a substrate and its activity is completely inhibited by pepstatin-A and 6M urea but not by 10 mM KCN. Optimal activity of the purified mosquito aspartic protease was obtained at pH 3.0 and 45°C. With hemoglobin as a substrate the enzyme had an apparent Km of 4.2 μM. Polyclonal antibodies to the purified enzyme were raised in rabbits. The specificity of the antibodies to the enzyme was verified by immunoblot analysis of crude mosquito extracts and the enzyme separated by both non-denaturing and SDS-PAGE. Density gradient centrifugation of organelles followed by enzymatic and immunoblot analyses demonstrated the lysosomal nature of the purified enzyme. The N-terminal amino acid sequence of the purified mosquito lysosomal protease (19 amino acids) has 74% identity with N-terminal amino acid sequence of porcine and human cathepsins D.

# INTRODUCTION

The lysosomal system plays a significant role in the regulation of important physiological processes in both immature and adult insects. The fat body, which is functionally analogous to the vertebrate liver, undergoes dramatic remodelling associated with metamorphosis and the termination of egg maturation cycles. It is during these periods that the lysosomal activity rises sharply and cellular organelles undergo massive autophagical degradation (Locke, 1980; Raikhel and Lea, 1983; Dean *et al.*, 1985; Raikhel, 1986a; 1986b). In adult mosquitoes, there is a dramatic increase in lysosomal activity in the fat body during cessation of synthesis of the yolk protein precursor, vitellogenin (Raikhel, 1986a). This lysosomal activity is directed towards specific degradation of organelles involved in biosynthesis and secretion of vitellogenin (Raikhel, 1986b).

It is important to elucidate the molecular mechanisms regulating this specific lysosomal activity. Cathepsin D, which is one of the most abundant lysosomal proteases (Barrett, 1970), is a good candidate for developing a molecular probe for these studies. Furthermore, the kinetics of the specific activity of this enzyme in the mosquito fat body were demonstrated and the factor triggering the rise in cathepsin D activity identified (Raikhel, 1986a; 1986c). Although cathepsins D from various tissues of many vertebrates have been extensively characterized (Kirschke and Barrett, 1987), little is known about characteristics of similar enzymes from insect tissues. The only report on the purification and partial characterization of a cathepsin D-like enzyme was from pupae of the blowfly, *Aldrichina grahami* (Kawamura *et al.*, 1987). We report here the isolation and characterization of a lysosomal aspartic protease from the mosquito, *Aedes aegypti*. This lysosomal protease exhibits unique features by

being similar in its molecular composition to a vertebrate cathepsin E, but having enzymatic properties of cathepsin D. Furthermore, its N-terminal amino acid sequence has 74% identity with N-terminal amino acid sequence of porcine and human cathepsins D but only limited similarity with cathepsin E.

# EXPERIMENTAL PROCEDURES

## *Materials*

Bovine cathepsin D, bovine hemoglobin, pepstatin A, pepstatin A-agarose, S-sepharose, concanavalin A-sepharose, Percoll and chicken egg albumin were purchased from Sigma Chemical Co.. SDS-PAGE molecular weight markers, ampholytes and protein assay dye reagent were from Bio-Rad Laboratories. Non-denaturing PAGE high molecular weight markers, pI calibration kit (pI 3-10), PhastGel IEF 3-9 were obtained from Pharmacia. Immobilon transfer membrane (PVDF), 0.45-µm pore size, were from Millipore. All other reagents employed were of analytical grade from Sigma and Baker Co..

## *Insects*

Headless, previtellogenic mosquitoes (*Aedes aegypti*), frozen in liquid nitrogen and shipped on dry ice, were kindly provided by Dr A.O. Lea (Department of Entomology, University of Georgia). The frozen mosquitoes were stored at -70°C until needed. Vitellogenic mosquitoes were obtained from our laboratory reared colony. To initiate vitellogenesis, mosquitoes were fed on rats.

## *Purification of the Mosquito Aspartic Protease*

All the purification steps were carried out at 4°C. About 20 g of frozen mosquitoes were homogenized in 100 ml homogenization buffer ( 0.2 M sodium acetate, pH 3.5 ,0.15 M NaCl and 0.02% sodium azide). The homogenate was centrifuged at 27,000 xg for 30 min. The supernatant was collected and recentrifuged at 60,000 xg for 1 hr. Proteins in the supernatant thus obtained were precipitated with 30 to 50 percent ammonium sulfate. The precipitated

proteins were dissolved in 10 ml of the above homogenization buffer. The protein solution was dialyzed against 3 x 1 liter of sodium phosphate buffer (0.1 M Na phosphate, pH 2.5, 0.15 M NaCl). Peptides precipitated during dialysis were removed by centrifugation at 20,000xg. The supernatant was concentrated by ultrafiltration using Amicon YM-10 membrane and the buffer in the protein solution exchanged with 0.05 M Na-citrate buffer, pH 4.0, containing 0.02% $NaN_3$ in the same ultrafiltration cell. About 5-10 ml of the concentrated solution was centrifuged at 20,000 xg to remove any aggregated proteins before applying to a S-Sepharose fast flow column (1.0 x 10 cm) equilibrated with 0.05 M Na-citrate buffer, pH 4.0, 0.02% $NaN_3$. Unbound proteins, which did not contain detectable cathepsin D activity, were washed from the column with the equilibration buffer. About 50% of bound cathepsin D activity was eluted with the equilibration buffer containing 0.1 M NaCl. Fractions with cathepsin D activity were pooled and concentrated by ultrafiltration using Amicon YM-10 membranes and Centricon-30 concentrators (Amicon). SDS-PAGE followed by silver staining showed that at this stage of purification, the concentrated fraction predominantly consisted of a single polypeptide with only a few minor contaminating peptides. Further purification of the enzyme was achieved by its electroelution from polyacrylamide gels after electrophoresis.

Although, finally the above scheme was adopted for the purification of a mosquito aspartic protease with cathepsin D activity, gel filtration and affinity chromatographic techniques (below) were also attempted to purify the enzyme. These steps were extremely useful in establishing the identity of the aspartic protease on SDS gels as well as revealing some of its characteristics.

***Gel permeation chromatography***

The supernatant obtained after acid precipitation of the initial mosquito

homogenate was dialysed against 0.15 M sodium phosphate buffer, pH 7.0, 0.15 M NaCl. 0.02% $NaN_3$ and concentrated by ultrafiltration. Four milligram protein was applied onto a Sephadex G-100-120 column (2.5 cm x 50 cm) equilibrated and eluted with the dialysis buffer.

***Pepstatin-affinity chromatography***

Fractions containing cathepsin D activity, collected after gel permeation, were pooled and concentrated by ultrafiltration. During concentration, the buffer was exchanged by washing three times with 50 ml pepstatin A-agarose binding buffer (0.05 M Na-acetate, pH 3.5, 0.2 M NaCl, 0.02% $NaN_3$) and then applied as such onto a pepstatin A-agarose column (1 cm x 10 cm) equilibrated with binding buffer. Unbound proteins were washed out completely with the same buffer. The bound protein, an aspartic protease with cathepsin D activity, was eluted with a high pH and high salt buffer (0.05 M Tris-HCl, pH 8.5, 0.6 M NaCl). A complete elution of bound proteins from the column was achieved by using a chaotropic buffer (0.12 M Tris-HCl, pH 6.8, 4% SDS, 0.15 M DTT).

***Concanavalin A-Sepharose chromatography***

Fractions containing cathepsin D activity, collected after cation-exchange chromatography (above) were loaded on a Con A-Sepharose column (1.0 cm x 10 cm) equilibrated with Con A buffer (0.2 M Tris-HCl, pH 7.4, 0.5 M NaCl, 10 mM $CaCl_2$, 10 mM $MgCl_2$ and 0.02% $NaN_3$). The mosquito aspartic protease was eluted with 0.1 M methyl $\alpha$-D-mannopyranoside in Con A buffer.

***Assay of cathepsin D activity***

Whole mosquitoes or their dissected body parts were homogenized in the homogenization buffer. The homogenate was centrifuged at 13,600 xg for 20

min and the supernatant was assayed. The proteolytic activity of cathepsin D-like enzymes was measured in an incubation mixture (200 µl) containing 10 µl of extract or fraction obtained after column chromatography, 140 µl of 0.2 M citrate-phosphate buffer, pH 2.9 or 0.6 M Na-acetate buffer, pH 2.4, and 50 µl of 2% (w/v) aqueous hemoglobin solution as substrate resulting in a final pH 3.0. After 1 hour of incubation at 45°C, the reaction was stopped by addition of 40 µl of 18% (w/v) ice-cold trichloroacetic acid (TCA) to precipitate peptides. After 5 min on ice the precipitated peptides were centrifuged at 12,000 xg for 5 min. The digested small peptides released in 100 µl of supernatant were mixed with 700 µl of 3% TCA and the protein content determined by the Bio-Rad protein assay. To determine the activity specific to cathepsin D, 1 µl of 1.5 mM pepstatin A, an inhibitor of cathepsin D activity, was added to the control samples prior to the addition of hemoglobin. One unit of cathepsin D-like enzyme is defined as the amount required to release 1 µg of small peptides from the substrate in 1 h.

***Assay for acid phosphatase activity***

Acid phosphatase was assayed as described by Raikhel (1986a). The reaction mixture, containing 50 µl sodium citrate buffer (pH 4.8), 50 µl 0.45 p-nitrophenol phosphate, and 20 µl sample, were incubated at 37°C for 30 min. The reaction was terminated by addition of 60 µl of 0.5 N NaOH and then the absorbance was measured at 420 mn.

***Protein determination***

The concentration of soluble peptides was measured by a protein-dye assay (Bradford, 1976) using Bio-Rad reagents and bovine serum albumin (BSA) as a standard.

***Polyacrylamide gel electrophoresis***

The purity and the apparent molecular weight of the enzyme during its fractionation was monitored by SDS-PAGE on 10-15% gradient gels unless specified otherwise under reducing conditions (Laemmli, 1970) and the peptides visualized by silver staining according to the supplier's instructions (Sigma). The native molecular weight of the enzyme was estimated by electrophoresis under non-denaturing conditions on 5-15% gradient polyacrylamide gels.

***Isoelectric focusing (IEF)***

The isoelectric point (pI) of native form of the aspartic protease was determined by electrophoresis on IEF PhastGel (pH 3-9) using Phast-System (Pharmacia). Standard pI marker proteins were also separated on each gel.

***Two-dimensional gel electrophoresis***

Tube gels for the first dimension (IEF) were prepared by modification of O'Farrell's (1975) protocol. The gel mix was made by dissolving 5.5 g urea in 1.0 ml 40% acrylamide/5% bisacrylamide, 2.0 ml 10% NP-40, 0.4 ml Ampholyte 5-7, 0.1 ml Ampholyte 3.5-10 and 2.5 ml double distilled water. The anode solution was 85% phosphoric acid and the cathode solution was 20 mM NaOH. The gels were pre-run at 200 V for 15 min, at 300 V for 30 min and then at 400 V for 30 min. Protein samples were then loaded and focusing was carried out at 400 V for 16 hours followed by 1 hour at 800 V. Polypeptides separated in the first dimension were then separated by SDS-PAGE in the second dimension using 10-15% gradient slab gels.

***Electroelution of the enzyme from polyacrylamide gels***

The 40 kDa enzyme band was cut from SDS gels and electroeluted with

ISCO Model 1750 Electrophoretic Concentrator. According to the manufacturer's instructions, 40 mM Tris-Acetate buffer containing 2 mM EDTA and 4 mM Tris-Acetate buffer containing 0.2 mM EDTA (pH 8.6) were used for outer and inner compartments of the concentrator, respectively. The addition of 0.1% lithium dodecyl sulfate to both sample cup and inner cathode compartment buffers improved the efficiency of elution.

***Production of polyclonal antibodies***

The enzyme purified by electroelution from polyacrylamide gels was used to produce polyclonal antibodies in rabbits. Female New Zealand white rabbits were injected with 40 µg enzyme mixed with Freund's Complete Adjuvant (1:1 v/v). The rabbits were boosted at two weekly intervals with the enzyme preparation mixed with Freund's Incomplete Adjuvant (1:1 v/v). The rabbits were bled two weeks after the second booster injection and the sera were collected. The immunoglobulin (IgG) fraction was precipitated from the sera with 35% saturated ammonium sulfate solution. The precipitated IgG was solubilized in and dialysed against 0.2 M Na-phosphate, pH 7.8, 1 mM EDTA, 0.02% $NaN_3$, and stored at -20 $^oC$ in 0.5 mg/ml aliquots. The specificity of the antibodies was determined by immunoblot analysis.

***Immunoblot***

The proteins resolved by SDS-PAGE or non-denaturing PAGE were transferred to nylon-reinforced nitrocellulose (NitroScreen West, Dupont) according to Burnette (1981). After the transfer, proteins were visualized by Ponceau S stain and the molecular weight standards were marked. The blots were blocked with 5% non-fat dry milk in sodium phosphate buffer (10 mM

sodium phosphate, 0.9 % NaCl, pH 7.2) for 1.5 h. Primary antibodies were applied at a 1:500 dilution in blocking solution containing 0.3% Tween-20. The primary antibodies were detected using goat anti-rabbit antibodies conjugated with alkaline phosphatase (1:1,000 dilution). The antigen-antibody complex was then visualized by a substrate developing cocktail containing 0.005% 5-bromo-4-chloro-3-indolye phosphate, P-toluidine salt, and 0.01% Nitroblue tetrazolium-sodium.

***Subcellular fractionation***

The separation of lysosomes was performed according to Alquier *et al.*(1985). The mosquitoes were homogenized in 10 mM Tris buffer, pH 7.4 containing 0.25 M sucrose (TS buffer) using a glass/Teflon Potter-Elvehjem homogenizer. The resulting tissue homogenate was centrifuged at 800 xg for 20 min. The supernatant was collected and centrifuged at 26,000 xg for 20 min. The pellet was resuspended in 0.5 ml TS buffer and then mixed with 10 ml 30% Percoll in TS buffer. The suspension was centrifuged in a fixed-angled rotor (Beckman Type 65) at 60,000 xg for 45 min. A test tube prepared as for the samples but containing density marker beads (Pharmacia) was also centrifuged simultaneously to obtain a density profile. The gradients were collected into 21 fractions. The fractions corresponding to cathepsin and acid phosphatase activity were pooled. Although, Percoll in the gradient fractions did not effect the enzyme activity assays, it was removed by filteration through glass microfibre filters (GF/D, Whatman) before samples were used for SDS-PAGE.

***N-terminal amino acid sequencing***

After S-sepharose chromatography, the partially purified enzyme (50 pM) was resolved by 10-15% gradient SDS-PAGE and then electroblotted onto

PVDF membrane (Matsudaira, 1987). The membrane was stained with 0.1% Coomassie Blue R-250 in 50% methanol for 5 min, and then destained in 50% methanol/10% acetic acid for 10 min at room temperature. After rinsing with deionized $H_2O$ for 5-10 min, the membrane was air dried and the area with the enzyme band was cut out with a clean razor. The band on the cut membrane was sequenced on an Applied Biosystem model 477A protein sequencer directly and resulting PTH-amino acids were analyzed on an on-line 120A analyzer (Macromolecular Facilities, Department of Biochemistry, Michigan State University). The computer-aided comparison of the mosquito aspartic protease sequence with known cathepsin D and cathepsin E sequences from other species was done using a GCG program from the University of Wisconsin Genetics Computer Group and FASTP program (PIR, NBRF, Georgetown University, Washington D.C.). Two libraries were used for this comparison: GenEMBL protein library containing 28,968 sequences and Amino Acid Bank Protein library containing 5,415 sequences.

# RESULTS

***Purification of the mosquito aspartic protease with cathepsin D activity***

In our initial attempts to purify mosquito cathepsin D, we utilized ammonium sulfate precipitation of proteins from crude extracts of headless previtellogenic adult mosquitoes followed by gel permeation chromatography. Fractions from gel permeation column containing cathepsin activity were pooled together and applied on pepstatin A-agarose affinity column. The protein eluted from the affinity column by the high-salt elution buffer, when analyzed by SDS-PAGE, revealed a single band with an apparent molecular weight of 40,000 (Fig. 2, lane 6). The identity of the 40 kDa peptide as cathepsin was additionally verified by inhibition of its activity with pepstatin A in the enzymatic assay. However, the yield by the affinity chromatographic step was extremely low. Even after pH of the elution buffer was elevated to 9.0, less that 10% of cathepsin activity eluted from the column and the enzyme rapidly lost its activity. When the bound protein was denatured and completely eluted from the column by the chaotropic buffer, the 40 kDa peptide was revealed by SDS-PAGE. Therefore, due to low yields and increased instability of the enzyme, this procedure was not used. Similarly, peptstatin A-affinity chromatography could not be used for purification of a cathepsin D-like enzyme from another insect, blowfly, *Aldricina grahami* (Kawamura *et al.*, 1987). In contrast, this method was effectively used to purify mammalian cathepsins (Afting and Becker, 1981; Takahashi and Tang, 1981).

Finally, cation-exchange chromatography was used for routine purification of the mosquito enzyme after two precipitation steps of the crude extracts.

**Table 1. Purification of mosquito lysosomal aspartic protease.**

| Procedures | Protein (mg) | Total Activity (Units) | Specific Activity (Units/$\mu$g) | Yield (%) | Purification (Folds) |
|---|---|---|---|---|---|
| 1. Crude extract | 793.0 | 87,100 | 0.111 | 100 | 1 |
| 2. $(NH_4)_2SO_4$ ppt | 146.0 | 69,680 | 0.333 | 80 | 3 |
| 3. Acid precipitation | 26.0 | 65,325 | 2.56 | 75 | 23 |
| 4. Ultrafiltration | 3.9 | 34,545 | 8.7 | 40 | 78 |
| 5. S-sepharose | 0.013 | 6,000 | 461.0 | 6.9 | 4,153 |

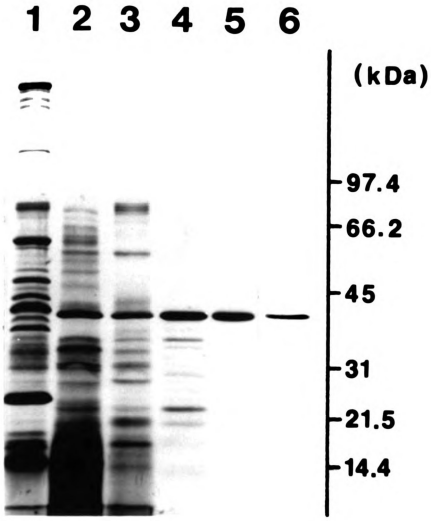
**Fig. 1: Elution profile of mosquito proteins on S-sepharose fast flow column (1.0 cm x 10 cm).**

Sample, proteins obtained after acid precipitation step (Fig. 2, lane 3). Arrows indicate elution with, (1) start buffer, 0.05 M Na-citrate, pH 4.0, 0.02% $NaN_3$ and (2) elution buffer, start buffer containing 0.1 M NaCl. Flow rate, 25 ml/h. Fraction size, 1 ml. A bar denotes the fractions pooled for further analyses.

Absorbance (595 nm)
0.5
0.4
0.3
0.2
0.1
0.0
CD-Activity
0
10
20
0
10
20
30
40
50
60
70
Fraction Number
Relative Absorbance (280 nm) ---

**Fig. 2: SDS-PAGE (10-15%) of mosquito lysosomal aspartic protease with cathepsin D activity after various purification steps as revealed with silver staining.**

Lanes: (1) soluble proteins of headless previtellogenic mosquitoes; (2) 30-50% ammonium sulfate precipitated proteins from step 1; (3) supernatant after acid precipitation at pH 2.5 of proteins from step 2; (4) pooled fractions numbers 10-45 containing cathepsin D activity as shown by dotted profile in Fig. 1; (5) pooled fractions indicated by a bar in Fig. 1 after S-sepharose chromatography; and (6) purified enzyme after affinity chromatography on pepstatin A-sepharose. The molecular weights on the right, in order of decreasing $M_r$, are of phosphorylase b, BSA, ovalbumin, carbonic anhydrase, soybean trypsin inhibitor and lysozyme.


1 2 3 4 5 6
(kDa)
97.4
66.2
45
31
21.5
14.4

Results obtained after such a purification are summarized in Table 1 and Fig. 2. Precipitation of the crude extract with 30-50% ammonium sulfate followed by lowering the pH of the resuspended proteins to 2.5 resulted in a 23-fold increase in the purity of the enzyme. A further 4000-fold purification of the enzyme with a 7% yield of the original extract was achieved by cation-exchange chromatography of the acid precipitated protein extract on a S-sepharose column. A typical elution profile of proteins from the S-sepharose column is shown in Fig. 1. The fractions containing peak cathepsin activity (indicated by a bar in Fig. 1) were highly enriched for the enzyme as judged by electrophoresis (Fig. 2).

***Electrophoretic analysis of the purified mosquito enzyme***

The purified enzyme was analyzed by polyacrylamide gel electrophoresis under both denaturing and non-denaturing conditions. In either case, the enzyme resolved into a single protein band. By SDS-PAGE the mosquito enzyme had an apparent molecular weight of 40,000 (Fig. 2). Its native molecular weight was estimated to be 80,000 (Fig. 3). To confirm that the 80 kDa protein band had cathepsin D activity, it was cut from the gel and the gel piece homogenized in buffer was used for cathepsin D enzymatic assay. The supernatant separated from the homogenized gel showed cathepsin D activity (results not shown).

Pooled fractions obtained after cation-exchange chromatography, containing partially purified enzyme, were also analyzed by 2-dimensional gel electrophoresis. After electrophoresis in the second dimension under reducing conditions a predominant spot with $M_r$=40,000 was revealed (Fig. 4).

**Fig. 3: Non-denaturing polyacrylamide gel electrophoresis of purified enzyme stained with silver.**

Lanes: (1) mosquito aspartic protease; (MW) molecular weight markers (from top to bottom), thyroglobulin, ferritin, catalase, aldolase, bovine serum albumin and chicken egg albumin.

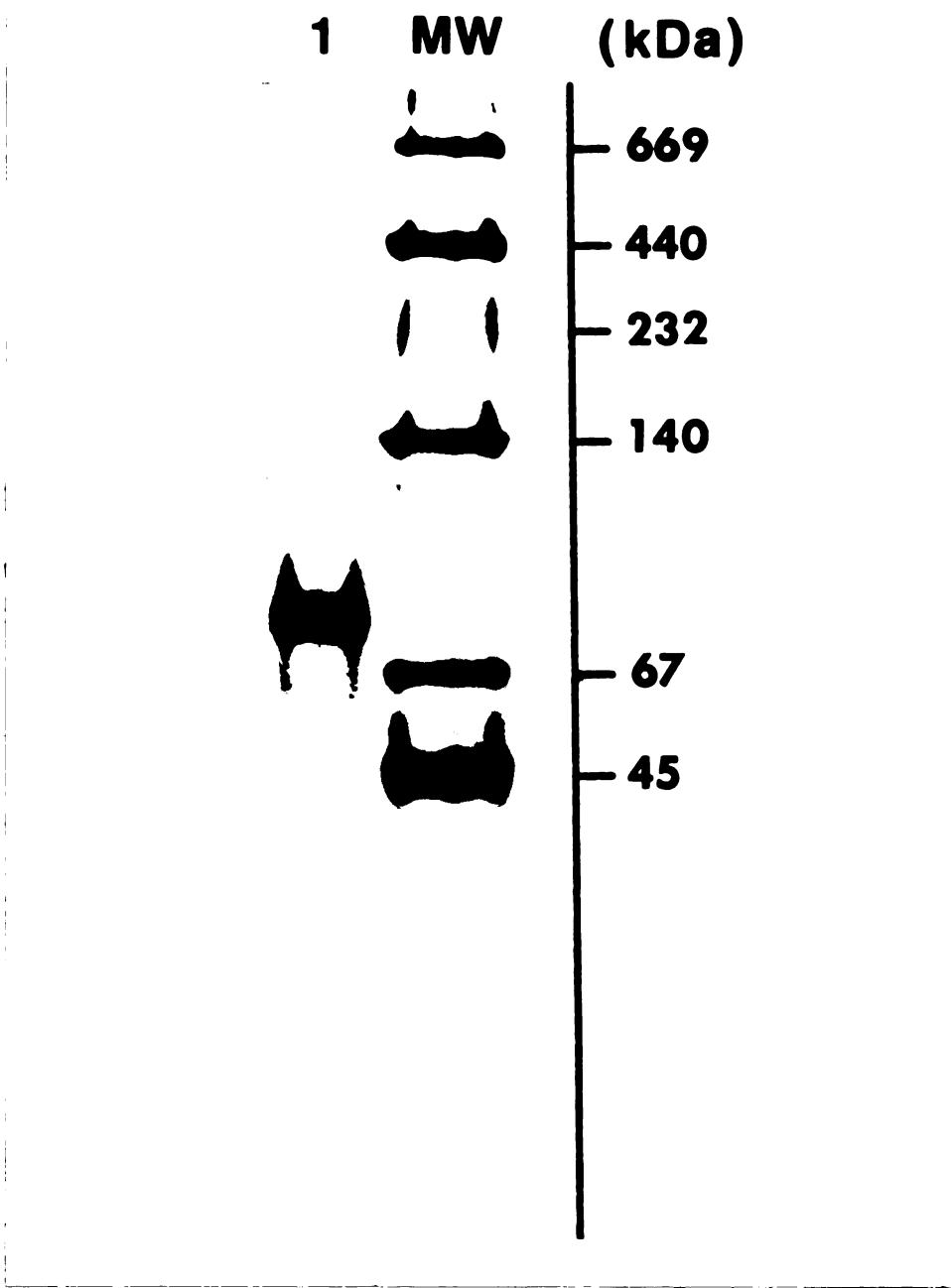

1
MW
(kDa)
669
440
232
140
67
45

Fig. 4: **Two dimensional gel electrophoresis of proteins in fractions pooled after S-sepharose chromatography lane 4 in Fig. 2).**

First dimension (1-D): IEF separation with ampholyte, pH 5-8; Second dimension (2-D): 10-15% gradient SDS-PAGE. The arrow shows the mosquito enzyme. The molecular weights on the right are the same as in Fig. 2.

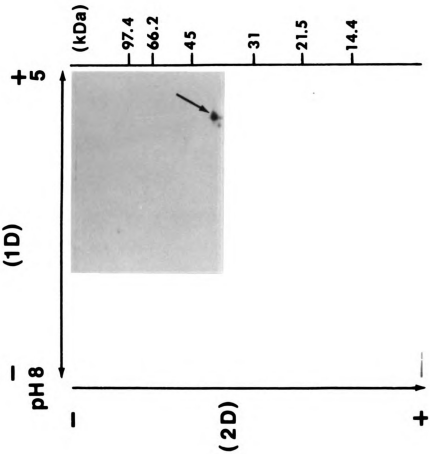
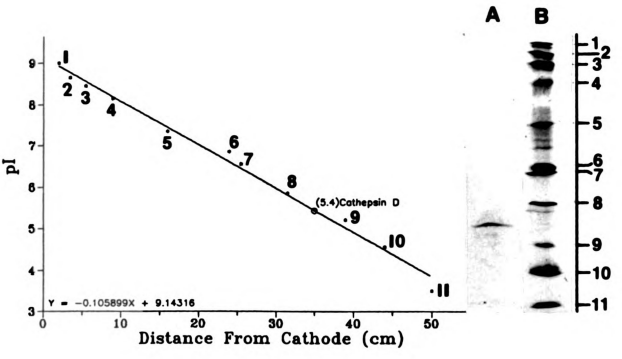

(kDa)
97.4
66.2
45
31
21.5
14.4
+
5
(1D)
–
pH 8
–
(2D)
+

**Fig. 5: Determination of isoelectric point (pI) of mosquito aspartic protease.** Lanes: (A) purified enzyme; (B) pI markers on the right and on the graph: (1) trypsinogen, 9.3; (2) lentil lectin-basic band, 8.65; (3) lentil lectin-middle band, 8.45; (4) lentil lectin-acidic band, 8.15; (5) myoglobin-basic band, 7.35; (6) myoglobin-acidic band, 6.85; (7) human carbonic anhydrase B, 6.55; (8) bovine carbonic anhydrase B, 5.85; (9) beta-lactoglobulin A, 5.2; (10) soybean trypsin inhibitor, 4.55; (11) amyloglucosidase, 3.5. The pI of the mosquito enzyme, 5.4, was determined from the plot on the left.


pI
Distance From Cathode (cm)
Y = −0.105899X + 9.14316
(5.4)Cathepsin D
A
B

***Isoelectric point of the purified enzyme***

Analytical IEF of the enzyme on pH 3-9 Phast IEF gels showed that the native enzyme has a pI of 5.4 (Figure 5). A similar pI value for the partially purified enzyme was obtained by two dimensional gel electrophoresis when it was separated in the first dimension (IEF) in the presence of 6 M urea (Fig. 4).

***Evidence for glycosylation of the purified enzyme***

Fractions obtained after cation-exchange chromatography (lane 4 in Fig. 2) and enriched for cathepsin D activity were applied on Con A-Sepharose column. The absence of cathepsin activity in the unbound fraction and its presence in fractions eluted with Con-A buffer containing 0.1 M methyl α-D-mannopyranoside indicated that the enzyme is a glycoprotein with mannose on the glycosyl part.

***Enzymatic properties of the purified enzyme***

The purified enzyme displayed a pH optimum close to 3.0 when assayed with hemoglobin as a substrate (Fig. 6). At 4°C it maintained its activity between pH 2.5 - 4.0, but lost its activity and degraded rapidly at pH higher than 7.

The optimum temperature for the hydrolysis of hemoglobin by the purified enzyme was 45°C as shown in Fig. 7. The enzyme was inactivated at temperatures above 60°C.

The purified mosquito enzyme was further characterized by its activity towards hemoglobin as a substrate. The activity of the purified enzyme (0.5 µg) was assayed with concentrations of hemoglobin increasing from 0 µM to 77.5 µM (based on the molecular weight of hemoglobin to be 64,500). The Michaelis Constant (Km), determined by the Michaelis-Menten plot as well as Lineweaver-Burk plot, was calculated to be 4.2 µM (Fig. 8).

**Fig. 6: Effect of pH on the activity of mosquito aspartic protease.**

The pH dependence of cathepsin D activity was determined by using stock solutions of 0.6 M phosphoric acid, 0.6 M acetic acid, 0.6 M $KH_2PO_4$ and 0.6M Tris-base to prepare buffers with different pH (1.5-9.0). The assay mixture was 140 ul of desired pH buffer, 10 ul (0.5 ug) of purified enzyme and 50 ul of 2% hemoglobin. Each point is a mean $\pm$ S.E.M. of 3 assays.

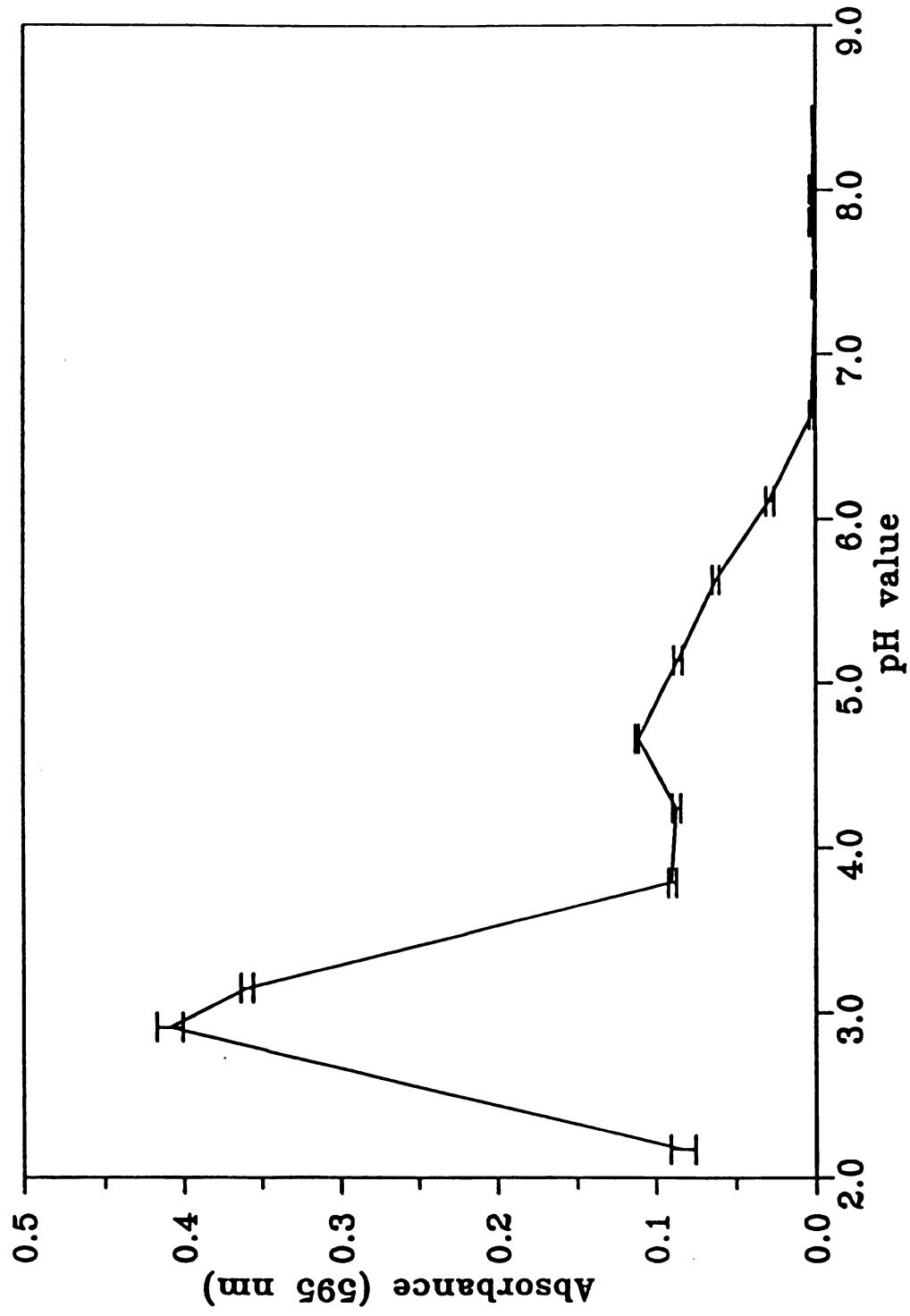

Absorbance (595 nm)
0.0
0.1
0.2
0.3
0.4
0.5
2.0
3.0
4.0
5.0
6.0
7.0
8.0
9.0
pH value

**Fig. 7. Effect of temperature on the activity of mosquito aspartic protease.** The reaction mixtures were the same as described in figure 3, except that only 0.6 M Na-acetate buffer, pH 2.4, was used in all reactions. Incubation of the reaction mixtures was for 1 hr at the indicated temperatures. Each point is a mean $\pm$ S.E.M of 3 determinations.

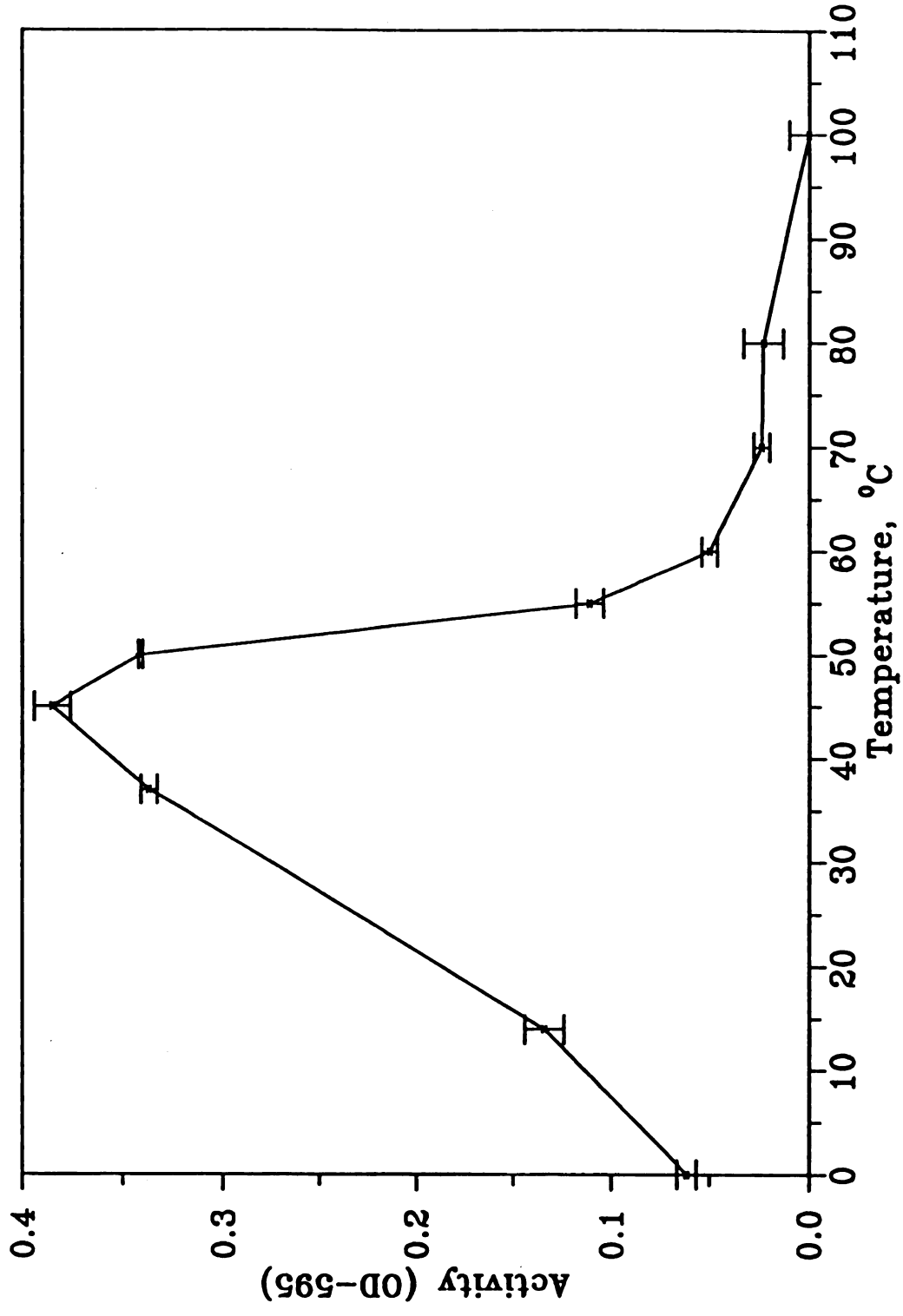
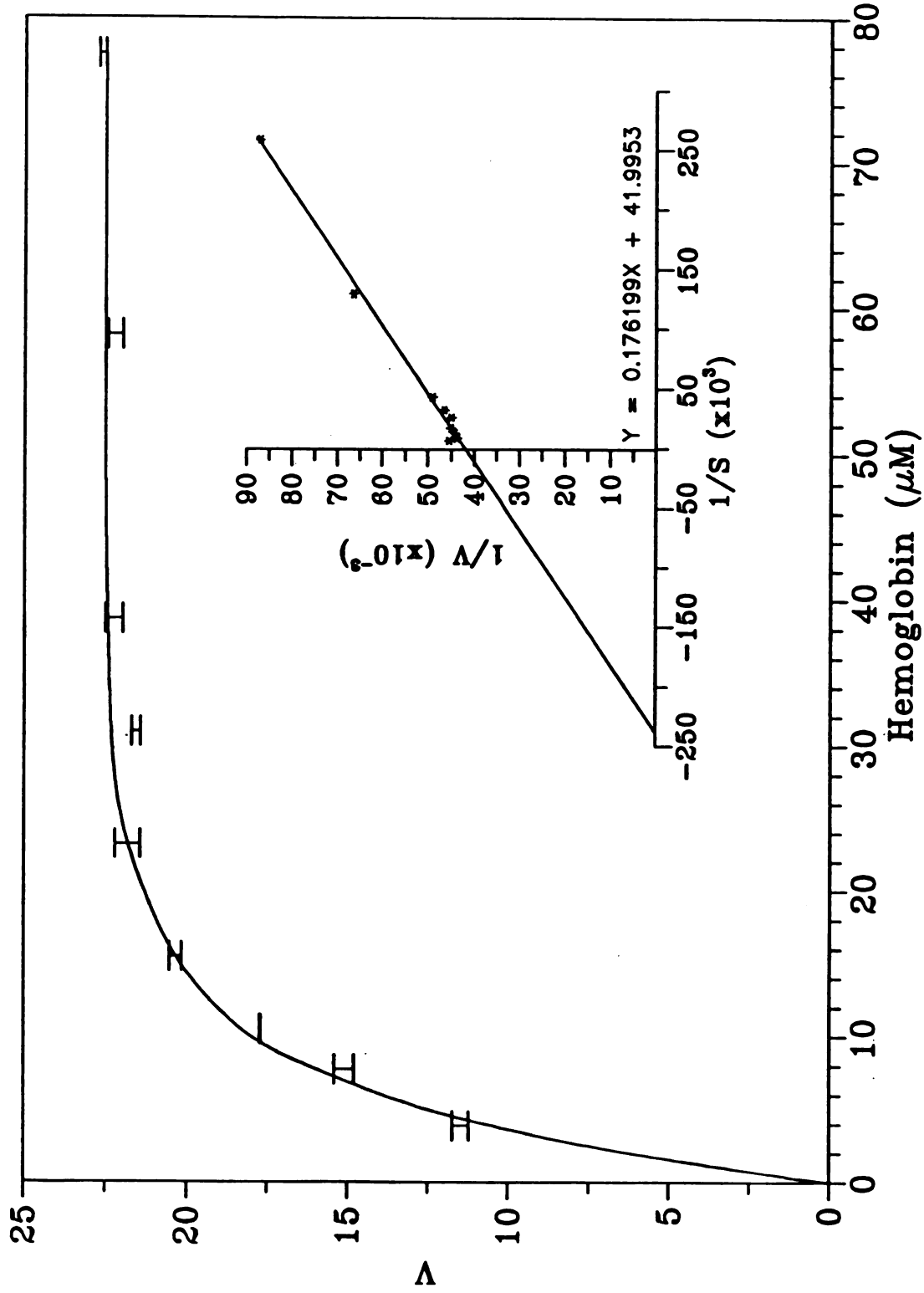

Activity (OD-595)
0.0
0.1
0.2
0.3
0.4
0
10
20
30
40
50
60
70
80
90
100
110
Temperature, °C

**Fig. 8. Michaelis-Menten and Lineweaver-Burk plots for mosquito aspartic protease with increasing concentrations of hemoglobin.**

The reaction mixture contained 10 ul (0.5 ug) of the purified mosquito enzyme, 140 ul of 0.6 M acetate buffer (pH 2.4) and 50 ul of hemoglobin (0-77.5 uM). The incubation time was 10 min. Inset: Double-reciprocal (Lineweaver-Burk) plot of the kinetics of mosquito aspartic protease using the Michaelis-Menten data.


V
Hemoglobin (μM)
1/V (x10⁻³)
1/S (x10³)
Y = 0.176199X + 41.9953

The activity of the mosquito enzyme to hydrolyze hemoglobin was also tested in the presence of various inhibitors. In contrast to pepstatin A (7.5 μM) and urea (6 M) which completely inhibited the enzyme activity, KCN (10 mM) did not have any inhibitory effect.

***Production and characterization of polyclonal antibodies against the mosquito aspartic protease***

Polyclonal antibodies against the gel-purified mosquito aspartic protease were produced in rabbits. The specificity of the produced antibodies was established by immunoblot analysis. Both crude mosquito extracts and purified aspartic protease were resolved by SDS-PAGE and non-denaturing PAGE and transferred to nitrocellulose membranes. These protein blots were probed with antibodies to the enzyme. The antibodies reacted with 80 kDa and 40 kda bands in both samples separated on non-denaturing and SDS-polyacrylamide gels, respectively (Fig. 9). The molecular weights of these protein bands correspond to the native and the reduced state of the aspartic protease (Figs. 3 and 4).

***Tissue and subcellular distribution of the aspartic protease***

Female mosquitoes 40 h post blood feeding were used to analyze the distribution of the enzyme in different tissues. Although cathepsin D activity was present in all assayed tissues (whole body, head and thorax, isolated abdomens free of mid-gut and ovaries, gut and ovaries) it was the highest in isolated abdomens, consisting largely of adhering fat body. When the same tissue extracts were analyzed on immunoblots using antibodies against the purified aspartic protease, the enzyme was detected in all of them. However, it was clearly the most abundant in the isolated abdomen extracts (not shown).

**Fig. 9. Immunoblot analysis of polyclonal antibodies against the mosquito aspartic protease.**

A. Non-denaturing PAGE; B. SDS-PAGE under reducing conditions. Lane 1, crude tissue extract after acid precipitation; 2, purified enzyme. Molecular weight standards on the right are as in Figs. 1 and 2. Eighty and 40 refer to the native and sub-unit molecular weight (kDa) of the enzyme, respectively.

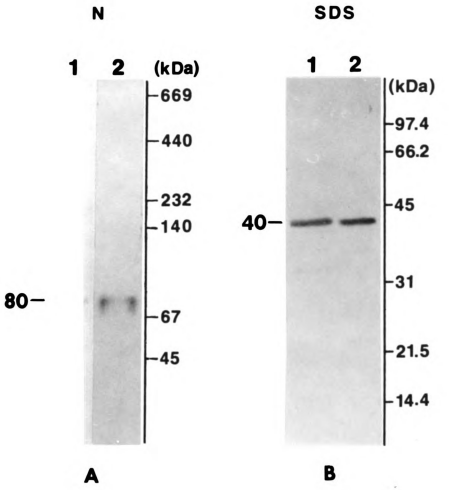
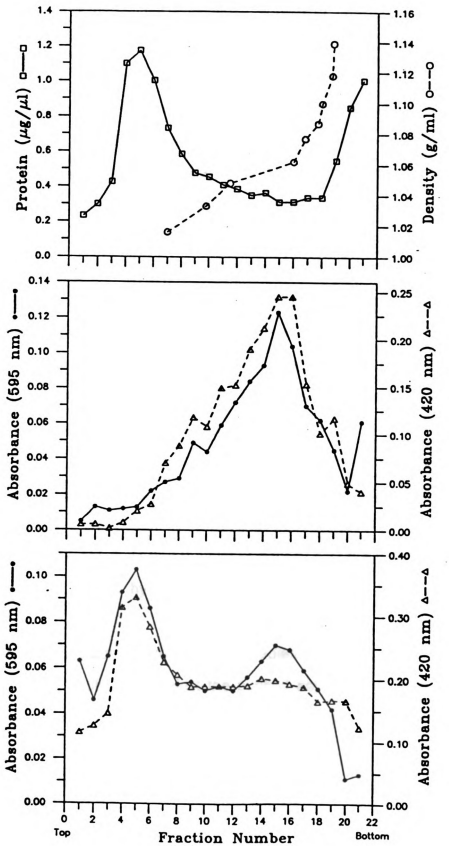

N
SDS
1
2
(kDa)
669
440
232
140
80—
67
45
1
2
(kDa)
97.4
66.2
45
40—
31
21.5
14.4
A
B

**Fig. 10. Protein and enzyme distributions after isopycnic centrifugation on 30% Percoll gradients of the mosquito organelles.**

A, Density and protein profiles; B and C, acid phosphatase (-Δ-) and cathepsin D (-•-) activity distribution before and after sonication of organelle preparation, respectively.


Protein (μg/μl)
Density (g/ml)
Absorbance (595 nm)
Absorbance (420 nm)
Absorbance (595 nm)
Absorbance (420 nm)
Top
Fraction Number
Bottom

In order to locate the subcellular source of cathepsin D activity in our purified preparations, we centrifuged TS buffer extracted mosquito homogenates (Materials and Methods) on isopycnic Percoll gradients. The distribution of protein, acid phosphatase (a lysosomal marker) and cathepsin D activities are reported in Fig. 10. The protein profile exhibited a large peak at low density. Both acid phosphatase and cathepsin D activities banded at a density higher than where the proteins floated on the gradients. The density (1.058 g/ml) at which the two enzymes peaked was reproducible in several experiments. However, in some experiments, in addition to this stable peak of enzyme activity another peak with acid phosphatase and cathepsin D activities was obtained which banded at the same position where peak protein concentration was (not shown). We argued that this second peak of enzymatic activity at lower density resulted from enzymes released from damaged lysosomes. In order to verify this, the organelle preparation was sonicated prior to Percoll gradient centrifugation. Analysis of such Percoll gradients showed a considerable reduction of the high density acid phosphatase/cathepsin D activity peak and an increase in the lower density enzymatic peak (Fig. 10). Based on these experiments, we conclude that the higher density enzyme activity peak identified by its content of acid hydrolases and its sensitivity to sonication represents lysosomes.

Next, the fractions from Percoll gradient were subjected to immunoblot analysis using antibodies against the purified aspartic protease. The enzyme was only detected in fractions which exhibited cathepsin D activity (Fig. 11).

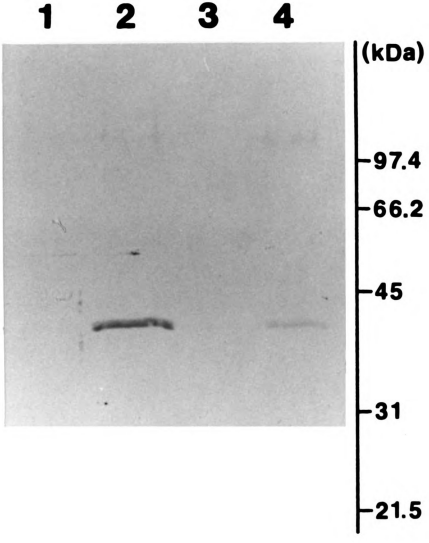
***Analysis of N-terminal amino acid sequence***

The N-terminal amino acid sequence of the electroblotted mosquito enzyme was determined. The 19 N-terminal amino acid residues of the mosquito

enzyme were compared with sequences from two computer protein libraries with GCG and FASTP programs. The results show that the mosquito sequence has the highest functional similarity with porcine and human cathepsin D. Comparison of the aligned amino acid sequences revealed a 74% identity between N-terminal amino acid sequence of the mosquito aspartic protease and porcine and human cathepsins D (Fig. 12A), and a 53% identity between N-terminal amino acid sequence of the mosquito enzyme and an internal sequence of human cathepsin E (Fig. 12B).

**Fig. 11. Immunoblot analysis of fractions from Percoll gradients of the mosquito organelles.**

Samples of fractions from Percoll gradient (Fig. 10) were separated on 12% polyacrylamide-SDS gels, transferred onto nitrocellulose membrane and probed with antibodies to the purified mosquito aspartic protease. Lanes 1, 2 and 3 correspond to fractions 5, 15 and 21, respectively, in Fig. 10 A, B. Lane 4, sample used for isopycnic centrifugation. Molecular weight standards are as in Fig. 2.


1
2
3
4
(kDa)
97.4
66.2
45
31
21.5

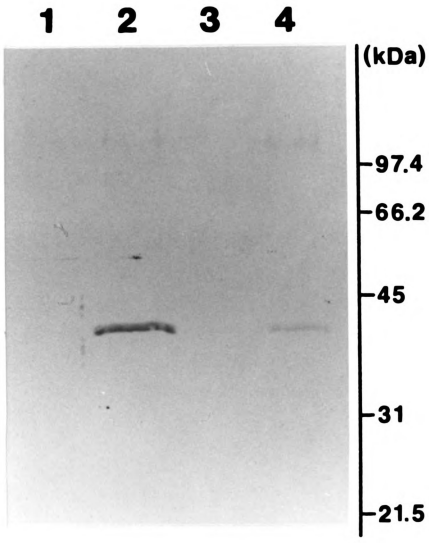
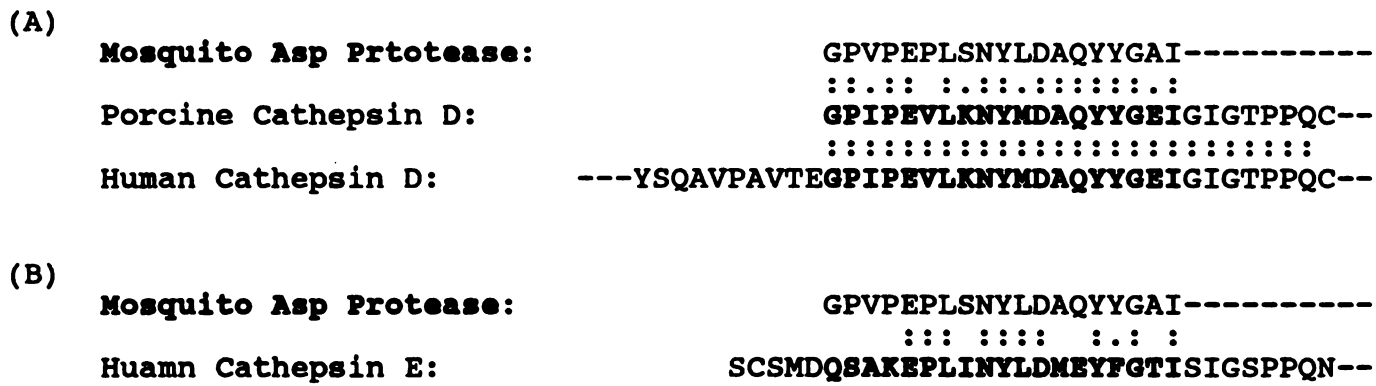

1
2
3
4
(kDa)
97.4
66.2
45
31
21.5

**Fig. 12. Comparison of N-terminal amino acid sequence of mosquito aspartic protease with sequences of vertebrate cathepsin D(A) and cathepsin E (B).** The marks, (:) and (.), between aligned sequences indicate identical and functionally related residues, respectively. Amino acids are represented by standard one-letter abbreviations.


(A)
Mosquito Asp Prtotease: GPVPEPLSNYLDAQYYGAI----------
Porcine Cathepsin D: GPIPEVLKNYMDAQYYGEIGIGTPPQC--
Human Cathepsin D: ---YSQAVPAVTEGPIPEVLKNYMDAQYYGEIGIGTPPQC--
(B)
Mosquito Asp Protease: GPVPEPLSNYLDAQYYGAI----------
Huamn Cathepsin E: SCSMDQSAKEPLINYLDMEYFGTISIGSPPQN--

# DISCUSSION

In this report, we present data on the purification and characterization of an aspartic lysosomal protease from the mosquito, *Aedes aegypti*. In general, its enzymatic characteristics are similar to those reported for cathepsins D from mammalian tissues (Takahashi and Tang, 1981; Kirschke and Barrett, 1987) and a cathepsin D-like enzyme from the blowfly, *A. grahami* (Kawamura *et al.*, 1987). The mosquito enzyme hydrolyzed hemoglobin as a substrate with pH and temperature optima of 3.0 and 45°C, respectively. Yamamoto *et al.* (1978) established differences between the activities of rat spleen cathepsin D and E based on inhibition produced by pepstatin A, urea or KCN. We found that in the case of the mosquito enzyme, complete inhibition with pepstatin-A and urea and not with KCN indicated that our purified enzyme was cathepsin D. Furthermore, and more importantly, a comparison of 19 residues of the N-terminal sequence of mosquito aspartic protease revealed a high degree of sequence and functional similarity with porcine and human cathepsins D. Similar comparison with the recently published amino acid sequences of cathepsins E (Azuma *et al.*, 1989; Yonezawa *et al.*, 1990) showed lesser degree of similarity. To our knowledge this is the first time that the N-terminal sequence of an insect aspartic protease has been obtained.

The mosquito aspartic protease had an apparent molecular weight of about 80,000 as determined by PAGE under non-reducing conditions and consisted of two subunits identical in molecular weight ($M_r$=40,000). For the blowfly cathepsin D-like enzyme, an apparent molecular weight of 41,000 was estimated by SDS-PAGE, but the native molecular weight was not determined (Kawamura *et al.*, 1987). Of the two mammalian aspartic proteases termed cathepsin D and E, the former has a molecular weight of about 40,000-50,000

and exists as a monomer in its native state (Berrett, 1977). Whereas, cathepsin E has a molecular weight of about 86,000-100,000 in its native form and consists of two identical subunits (Yonezawa *et al.*, 1987; Muto *et al.*, 1987).

The homogeneity of the purified mosquito aspartic protease after cation-chromatographic step was confirmed by isoelectric-focusing and electrophoresis under both denaturing and non-denaturing conditions. The existence of isoenzymes for the blowfly (Kawamura *et al.*, 1987) and mammalian (Berrett, 1970; Smith and Turk, 1974; Whitaker and Seyer, 1979; Yamamoto *et al.*, 1979) cathepsins D was demonstrated based on pI differences between isoenzymes. However, the mosquito aspartic protease in our purified preparations displayed a single pI of 5.4, indicating the absence of isoenzymes. This conclusion was also supported by kinetic data analysis (Bell and Bell, 1988).

The purified mosquito aspartic protease from the mosquito was shown to be a lysosomal enzyme. It has ubiquitous distribution in mosquito tissues. The enzyme was the most abundant in the fat body which is characterized by a high lysosomal activity (Raikhel, 1986a). Subcellular fractionation using Percoll gradients demonstrated the lysosomal nature of this enzyme. It was co-localized with the activity peak of lysosomal acid hydrolysis, acid phosphatase and cathepsin D.

The aspartic protease that we have purified from the mosquito is, therefore, characterized by unique features. Being a dimer, consisting of two identical subunits, this enzyme is similar to a mammalian cathepsin E which is an intracellular but not a lysosomal enzyme associated with lymphoid tissues (Yonezawa *et al.*, 1988). The mosquito enzyme is, however, a lysosomal enzyme and has enzyme characteristics similar to those of a lysosomal cathepsin D. Furthermore, the N-terminal of the mosquito protease has a much higher degree of similarity with cathepsin D than cathepsin E.

## ACKNOWLEDGEMENTS

We thank Dr. Arden O. Lea for providing us with large quantities of mosquitoes for enzyme purification, Mr. Alan R. Hays for the assistance and advice on the production of antibodies, and Dr. Charles Sweeley for his critical reading of the manuscript. This research was supported by grants AI-24716 from the National Institutes of Health and MSU Biotechnology Center to ASR

## REFERENCES


Afting E. G. and Becker M. L. (1981) *Biochem. J.* 197, 519-522.

Alquier C., Guenin P., Munari-Silem, Y., Audebert C. and Rousset B. (1985). *Biochem. J.* **232**, 529-537.

Azuma, T., Pals, G., Mohandas, T. K., Couvreur J. M., and Taggart, R. T. *J. Biol. Chem.* **264**, 16748-16753.

Barrett A. J. (1970) *Biochem. J.* **117**, 601-607.

Barrett, A. J. (1977) In: P*roteinases in Mammalian Cells and Tissue* (ed. A. J. Barrett), pp. 209-248, North-Holland Publishing Co., Amsterdam.

Bell J. E. and Bell E. T. (1988) In: *Proteins and Enzymes.* pp. 370-376. Englewood Cliffs, New Jersey.

Bradford M. M. (1976) *Analyt. Biochem.* **72**, 248-254.

Burnette, W. N. (1981). *Anal. Biochem.* **112**, 195-203.

Dean R. L., Locke M. and Collins J. V. (1985) In: *Comprehensive Insect Physiology Biochemistry and Pharmacology.* (Eds. Kerdut G. A. and Gilbert L. I.), Vol. 3, pp.155-210. Pergamon Press, Oxford, New York.

Kawamura M., Wadano A. and Miura K. (1987) *Insect Biochem.* **17**:77-83.

Kirschke H. and Barrett A. J. (1987) In: *Their Role in Protein Breakdown.* (Eds. Glaumann H. and Ballard F. J.), pp. 222-238, Academic Press, New York.

Laemmli U. K. (1970) *Nature* **227**, 680-685.

Locke M. (1980) The cell biology of fat body development. In: *Insect Biology in the Future* (Eds. Locke M. and Smith D. S.), pp. 227-252 Academic Press. New York.

Matsudaira P .(1987) *J. Biol. Chem.* **262**, 10035-10038.

Muto N., Yamamoto M., and Tani S. (1987) J. Biochem. **101**, 1069-1075.

O'Farrel P. H. (1975) *J. Biol. Chem.* **250**, 4007-4021.

Raikhel A. S. (1986a) *J. Insect Physiol.* **32**, 597-604.

Raikhel A. S. (1986b) *Tiss. Cell* **18**, 125-142.

Raikhel A. S. (1986c) *In: Host regulated Developmental Mechanism in Vector Arthropods.* Proceedings of the Vero Beach Symposium (Eds. Spielman, A. and Borovsky, D.), pp.25-31, University of Florida.

Raikhel A. S. and Lea A. O. (1983) *Tiss. Cell* **15**, 281-300.

Smith R. and Turk V. (1974) *Eur. J. Biochem.* **48**, 245-254.

Takahashi T. and Tang J. (1981) *Methods Enzymol.* Vol 80, 565-581.

Whitaker J. N. and Seyer J. M. (1979) *J. Neurochem.* **32**, 325-333.

Yamamoto K., Katsuda N. Himeno M. and Kato K. (1979) *Eur. J. Biochem.* **95**, 459-467.

Yamamoto K., Katsuda N., and Kato K. (1978) *Eur. J. Biochem.* **92**, 499-508.

Yonezawa S., Tanaka T., Muto N. and Tani S. (1987) *Biochem. Biophys. Res. Commun.* **144**, 1251-1256.

Yonezawa, S., Fujii K., Maejima Y., Tamoto K., Mori Y., and Muto N. (1988). *Arch. Biochem. Biophy.* **267**, 176-183.

Yonezawa, S., Takahashi T., Ichinose M., Miki K., Tanaka J., and Gasa S. (1990). *Biochem. Biophys. Res. Commun.* **166**, 1032-1038.

# CHAPTER 3*

# CLONING OF CDNA FOR MOSQUITO LYSOSOMAL ASPARTIC PROTEASE

## SEQUENCE ANALYSIS OF INSECT LYSOSOMAL ENZYME SIMILAR TO CATHEPSINS D AND E

* This manuscript is submitted to The Journal of Biological Chemistry.

## ABSTRACT

A CDNA coding for the lysosomal aspartic protease, from the mosquito (mLAP), was cloned and sequenced. The mLAP cDNA is 1,420 base pairs long with an open reading frame of 387 amino acids. The deduced amino acid sequence contains a signal pre-propeptide sequence of 18 amino acids followed by 369 amino acids with a 35 amino acid putative pro-enzyme domain in the N-terminal. The amino acid sequence of mLAP is 92% and 81% similar to human cathepsin D and cathepsin E, respectively. Typical cleavage sites for cathepsin D processing into light and heavy chains are lacking in mLAP. A single glycosylation site occurs in the mLAP sequence, at a position corresponding to the first glycosylation site of cathepsins D. The mLAP sequence shares putative phosphorylation determinants, which in cathepsins D are linked to the formation of mannose-6-phosphate.

In the mosquito fat body, lysosomal enzymes specifically degrade organelles involved in the biosynthesis and secretion of vitellogenin (Vg). The mLAP mRNA accumulates to its highest level 24 hr after initiation of Vg synthesis and 12 hr before the peak of mLAP protein accumulation and its enzymatic activity. Translational regulation of mLAP mRNA may occur. The 5'-untranslated region of mLAP mRNA is similar to elements conferring negative translational control by steroids.

# INTRODUCTION

Although lysosomes are ubiquitous cellular organelles, in several tissues they participate in various developmental and physiological processes (Smith and Farquhar, 1966; Glaumann *et al.*, 1981; Orci *et al.*, 1984). In the fat body of insects, the lysosomal system is involved in cellular remodelling, which is associated with metamorphosis and termination of egg maturation cycles (Locke, 1980; Dean *et al.*, 1985; Raikhel, 1986a; 1986b; 1992). The fat body of insects is a functional analogue to the vertebrate liver. It is responsible for metabolism and storage of carbohydrates, lipids and proteins, and synthesis and regulation of hemolymph proteins and sugars (Wyatt, 1980). In oogenic females an important function of the fat body is synthesis of yolk protein precursors, mainly vitellogenin (Vg) (Kunkel and Nordin, 1985; Raikhel and Dhadialla, 1992). In the mosquito fat body, the rise in specific activities of several lysosomal enzymes coincides with a dramatic decline in Vg synthesis (Raikhel 1986a). Fluorescent and electron microscopic analysis revealed two important roles of lysosomes during the termination of Vg production: (1) interruption of Vg secretion by degrading the Vg-containing secretory granules; (2) destruction of the biosynthetic machinery, RER and Golgi complexes, and subsequent remodelling of fat body trophocytes (Raikhel, 1986b).

Due to the medical importance of mosquitoes as vectors of numerous devastating human diseases, the elucidation of the biochemical and molecular basis of vitellogenesis and egg maturation is critical to the successful development of novel strategies in vector management. Mechanisms for altering the lysosomal activity in the mosquito fat body to cause premature interruption of vitellogenesis and egg maturation in mosquitoes may emerge.

In mammalian cells, targeting of newly synthesized enzymes to the

lysosome depends on phosphorylated mannose oligosaccharides (Kornfeld and Mellman, 1989). In yeast and plants, the signal targeting enzymes to the vacuole, the equivalent of lysosomes, is contained in polypeptide domains (Chrispeels and Raikhel, 1992; Bednarek and Raikhel, 1992). Nothing is known about the mechanism of lysosomal targeting in insects. The mechanisms may be different in insects because of inability to modify high-mannose sugar residues to complex sugars (Nordin *et al.*, 1984; Osir *et al.*, 1986; Nagao *et al.*, 1987).

A lysosomal aspartic protease from the mosquito *Aedes aegypti* was purified and characterized (Cho *et al.*, 1991a). The structure of this enzyme is similar to mammalian cathepsin E, in native molecular weight of 80,000 with two identical 40-kDa subunits. The purified enzyme, however, exhibits properties characteristic of cathepsin D. This mosquito aspartic protease does not have isozymes, and its pI is 5.4. Density gradient centrifugation of the organelles, followed by enzymatic and immunoblot analyses, localized the enzyme in lysosomes.

Here, we report the cloning and analysis of the cDNA coding for mLAP. This is the first report of the sequence for an insect lysosomal protease which provides insight into the evolution of aspartic proteases. The temporal expression of mLAP mRNA in mosquito fat bodies during the Vg synthetic cycle shows the peak of mLAP mRNA occurs 12 hr before the peak of mLAP protein and enzymatic activity. This suggests possible translational regulation of mLAP in the fat body. The 5'-untranslated region of mLAP mRNA has sequences which are similar to those implicated into negative translational control by steroids.

# EXPERIMENTAL PROCEDURES

## *Animals*

Mosquitoes, *Aedes aegypti*, were reared as described by Hays and Raikhel (1990). Larvae were fed on a standard diet (Lea, 1964). Vitellogenesis in adult females 3-5 days after eclosion was initiated by feeding them on rats.
In the anautogenous mosquito, *A. aegypti*, vitellogenesis proceeds through two distinct developmental stages. As a result of the previtellogenic stage, which begins at adult eclosion and is completed within 72 hr, the tissues participating in vitellogenesis, the fat body and the ovaries, become competent: the fat body for the synthesis of yolk protein precursors and the oocytes for their internalization. The ingestion of vertebrate blood triggers a cascade of hormonal signals that culminate in the massive synthesis of Vg and other yolk protein precursors in the fat body and their accumulation by the oocytes. The peak of these activities occurs at 24 hr after the initiation of vitellogenesis by a blood meal. Later, these activities dramatically decline. In the fat body this decline in the production of Vg and other yolk protein precursors coincides with the rising activity of lysosomal enzymes (Raikhel, 1992).

## *Materials*

Molecular weight markers for SDS-PAGE were purchased from Bio-Rad Laboratories. RNA ladder ( 0.24 - 9.5 kb) was from Bethesda Research Laboratories (BRL). Modified T7 DNA polymerase, Sequenase, was supplied by United State Biochemical Co. Horseradish peroxidase-conjugated affinity-purified rabbit anti-mouse IgG and goat anti-rabbit IgG were from Organon Teknika Corp. The enhanced chemiluminescence (ECL) Western blotting detection system was purchased from Amersham Corp. Perkin Elmer Cetus

was the source of the reagents for the PCR work. [$^{32}$P]dATP ( 3,000 Ci/mmol) for labeling and [$^{35}$S]dATP (1,000-1,500 Ci/mmol) for sequencing were from New England Nuclear (DuPont). All other reagents used were of analytical grade from Sigma Chemical Co. and Baker Co.

***Polymerase Chain Reaction***

The N-terminal amino acid sequence of mLAP (Cho *et al.*, 1991a) was used to design the 25 bp sense primer (primer-1). The 28 bp anti-sense primer (primer-2) was synthesized based on the amino acid sequence of the first catalytic center conserved among aspartic proteases (Azuma *el al.*, 1989). Both primers include a XbaI restriction site at their 5' ends (Fig. 1). The primers were synthesized at the Macromolecular Structure, Sequencing, and Synthesis Facility of Michigan State University.

DNA amplification was carried out on a Perkin Elmer thermal cycler with mosquito cDNA templates. The latter was obtained from 20 ug of total RNA isolated from the fat bodies of female mosquitoes 36 hr post-blood meal. All steps were performed according to the manufacturer's protocol (Cetus Co). The first five cycles of the primary amplification were carried out at 94$^{o}$C for 40 sec, at 48$^{o}$C for 2 min, and at 72$^{o}$C for 3 min; then the reaction was shifted to the second set of conditions for 20 additional cycles at 94$^{o}$C for 40 sec, at 53$^{o}$C for 2 min, and at 72$^{o}$C for 3 min. The final polymerase extension step was carried out at 72$^{o}$C for 10 min. A 96 base pair fragment was predominantly amplified during the first PCR cycle (Fig. 2A). This fragment was gel-purified from 10% acrylamide gel and was used for the secondary amplification. For the secondary amplification, twenty-five cycles were carried out at 94$^{o}$C for 40 sec, at 60$^{o}$C for 2 min, and at 72$^{o}$C for 3 min. The conditions for the final extension reaction

(A)

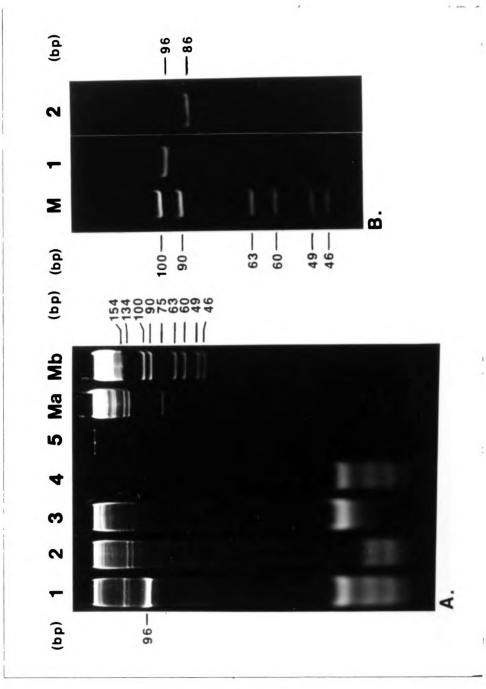
1 10
GPVPEPLSNYLDAQYYGAI----------FDTGSSNLWVP---------
Primer-1 Primer-2

(B)

| Primers | Primer Sequence | Degeneracy |
|---|---|---|
| Primer-1 | 5' CGTCTAGAGATGCNCAATATTATGGG 3'<br>XbaI C G C C | 64 |
| Primer-2 | 5' ATTCATGATTNCTNCTNCCNGTATCAAA 3'<br>XbaI GA GA G G | 16,384 |

**Fig. 1. Design of primers for amplification of mLAP cDNA fragment by the polymerase chain reaction.** (A) The primer-1 is based on the underlined stretch of amino acids from the N-terminal sequence of mLAP (Cho *et al.*, 1991a). The primer-2 is based on the underlined sequence from the first catalytic center of human cathepsin D which is conserved among aspartic proteases (Faust *et al.*, 1985). (B) Nucleotide sequences of primer-1 and primer-2. Sequences which are underlined and marked by XbaI indicate the sites for this restriction enzyme. Numbers indicate the level of degeneracy for each primer.

**Fig. 2. Results of the PCR amplification and isolation of the mLAP cDNA fragment.** (A) The first PCR amplification using: lane 1, cDNA from vitellogenic fat bodies and both primers; lane 2, the cDNA and the primer-1; lane 3, the cDNA and the primer-2; lane 4, both primers without the cDNA; lane 5, the cDNA without primers. A predominantly amplified band of 96 base pairs is marked by this number on the left. (B) The secondary PCR amplification utilizing the 96 bp cDNA fragment and both primers is shown in lane 1. Lane 2 shows a 86 bp cDNA fragment which was obtained after subcloning in pUC 119 and releasing by the restriction enzyme. The molecular sizes of both fragments are shown on the left. Lane Ma in A is the 1 kb DNA ladder (BRL), lane Mb in A and lane M in B are DNA markers made from pBR322 DNA digested with Alu I. The sizes of the markers are indicated by the numbers.


(bp)
1
2
3
4
5
Ma
Mb
(bp)
154
134
100
90
75
63
60
49
46
96
A.
(bp)
M
1
2
(bp)
100
90
63
60
49
46
96
86
B.

were the same as in the primary amplification (Fig. 2B).

***Construction of the cDNA library from fat bodies of vitellogenic female mosquitoes***

Total RNA from mosquito fat bodies was isolated by the guanidine thiocyanate method as described by Bose and Raikhel (1988). Mosquito fat bodies for cDNA library construction were collected from females at 0 to 48 hr after a blood meal with 6 hr intervals. Polyadenylated mRNA [poly(A)$^+$RNA] was obtained from total RNA by chromatography on an oligo(dT)-cellulose column (Ausubel *el al.*, 1990). The integrity of poly(A)$^+$ RNA was confirmed by Northern blot hybridization, using both Vg and mLAP probes. A cDNA library was constructed from 20 ug of mosquito poly (A)$^+$ RNA in a λZAP II vector (Stratagene Co). The oligo-dT- primed cDNA was fractionated to collect fragments bigger than 400 base pairs; these were then inserted into the EcoRI site of the λZAP II. The titer of the unamplified and amplified libraries are 1.7 X $10^6$ pfu/ml and 6 X $10^9$ pfu/ml, respectively.

***Cloning and sequencing of cDNA***

The 96-bp amplified cDNA fragment which was obtained by the PCR, was digested with XbaI. The released 86-bp fragment (Fig. 10) was subcloned into pUC 119 and sequenced. After positive confirmation of its sequence, this 86 bp mLAP cDNA fragment was used for screening of the λZAP II cDNA library. With this probe, 40 putative mLAP clones were isolated from the cDNA library. The cDNA clone with the longest insert (1420 base pairs) was sequenced, using a single-strand dideoxy chain termination method (Sanger *et al.*, 1977; Sambrook *et al.*, 1989). The results were confirmed by sequencing in both directions (Fig. 3).

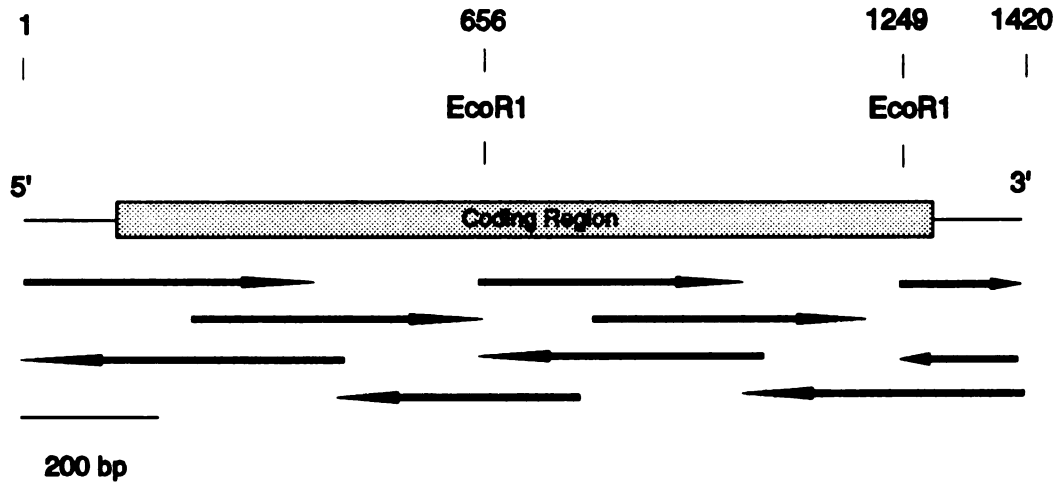



**Fig. 3. Sequencing strategy of mLAP cDNA clone.**

Numbers at the top correspond to the nucleotides in the cDNA. The coding region is indicated by a box. Positions of EcoR1 restriction sites are denoted. Horizontal arrows indicate the direction and extent of each sequencing determination.

***Analysis of nucleotide and amino acid sequences of mLAP***

The analysis of the deduced amino acid sequence of mLAP was performed by using the FASTA program (University of Wisconsin Genetics Computer Group Software), according to the algorithm of Lipman and Pearson (1985). The deduced amino acid sequence of mLAP was compared to deduced amino acid sequences translated from GenBank (release 68.0) and Swissprot (release 18.0) databases (Table 1). The initial score was calculated using the best sequence alignment of two sequences. The optimized score, calculated by considering insertions or deletions, gives a better indication of functional relations between proteins. The statistical significance of the scores was evaluated by Z value [(similarity score - mean of random scores)/(standard deviation of random scores); Z > 10 indicates statistical significance]. The similarities between mLAP and other sequences were calculated as the percentage of amino acid number having identical residues or functional substitutions, relative to the total number of amino acids in mLAP (387).

The dendrogram showing the relationship between members of a family of aspartic proteases was generated by the pairwise alignment of several aspartic proteases, with a computer GCG program (University of Wisconsin Genetics Computer Group Software). The similarity scores were used to create a clustering order that was represented as a dendrogram. Aspartic proteases revealing higher similarity scores to mLAP as indicated in Table 1, were chosen for this analysis. The following sequences were used in this analysis: mLAP, human cathepsin D (Faust *et al.*, 1985), pig cathepsin D (Shewale and Tang, 1984), mouse cathepsin D (Grusby *et al.*, 1990), Barley aspartic protease (Runeberg-Roos *et al.*, 1991), human cathepsin E (Azuma *et al.*, 1989), bovine chymosin A (Harris *et al.*, 1982), human pepsinogen A (Sogawa *et al.*, 1983), human renin (Imai *et al.*, 1983), yeast proteinase A (Ammerer *et al.*, 1986), and

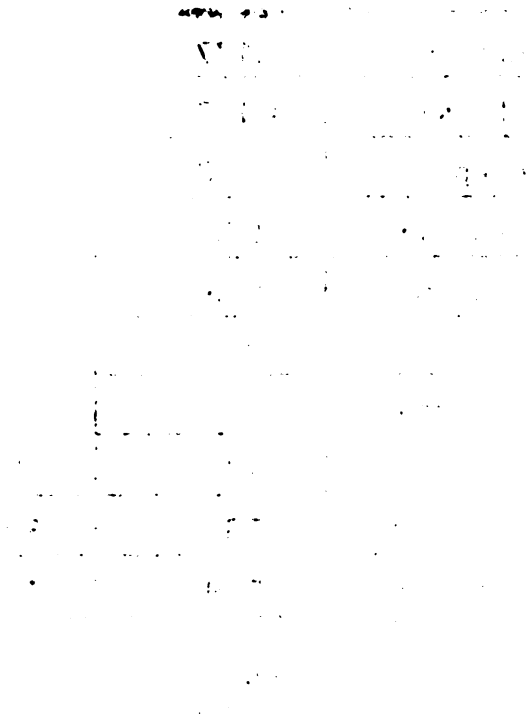


**Table 1. Homology of mLAP to aspartic proteases.**

| SOURCE | PROTEIN | SCORE[a] | Z[b] VALUE | IDENTITY % |
|---|---|---|---|---|
| Mouse | *Cathepsin D | 1,177 | 143.74 | 55.8 |
| Human | *Cathepsin D | 1,170 | 137.56 | 57.8 |
| Porcine | *Cathepsin D | 1,143 | 160.49 | 52.7 |
| Human | Cathepsin E | 988 | 100.46 | 43.4 |
| Yeast | *Proteinase A | 885 | 164.48 | 42.1 |
| Human | Renin | 894 | 116.29 | 40.5 |
| Human | Pepsinogen A | 833 | 78.90 | 42.1 |
| Barley | *Asp Protease | 775 | 40.00 | 35.1 |
| Bovine | Chymosin A | 763 | 39.00 | 47.8 |
| Rhizopus | Asp Protease | 549 | 40.70 | 31.0 |

[a]Only optimized scores are presented.
[b]Z value is a statistical significance of the optimized score (Z > 10 indicates statistical significance).
*Indicates the lysosomal or vacuolar enzyme.

Rhizopus aspartic protease (Horiuchi *et al.*, 1988). The dendrogram demonstrates the relationship among proteins based on their similarity of amino acid sequences and it also implies the evolutionary relationship among these proteins.

***Preparation of RNA and Northern blot analysis of mLAP and Vg mRNAs***

The isolation of total and polyadenylylated RNA from mosquito fat bodies was performed according to the method described by Bose and Raikhel (1988). Fat bodies were dissected from mosquito females collected from 0 to 48 hr after a blood meal with 3 hr or 6 hr intervals. Non-blood fed mosquitoes used for experiments were 3 to 4 day old. The abdominal wall with the attached fat body was removed and used for the isolation of the RNA and protein. For the Northern analysis, the total RNA was separated by 1.2% formaldehyde/agarose gel electrophoresis, transferred to nitrocellulose membrane, and hybridized with either a [$^{32}$P]-labeled 593 bp fragment of mLAP cDNA or a [$^{32}$P]-labeled 1.8 kb fragment of Vg gene A1 (gift from H.H. Hagedorn, University of Arizona) under high-stringency conditions (Fig. 4).

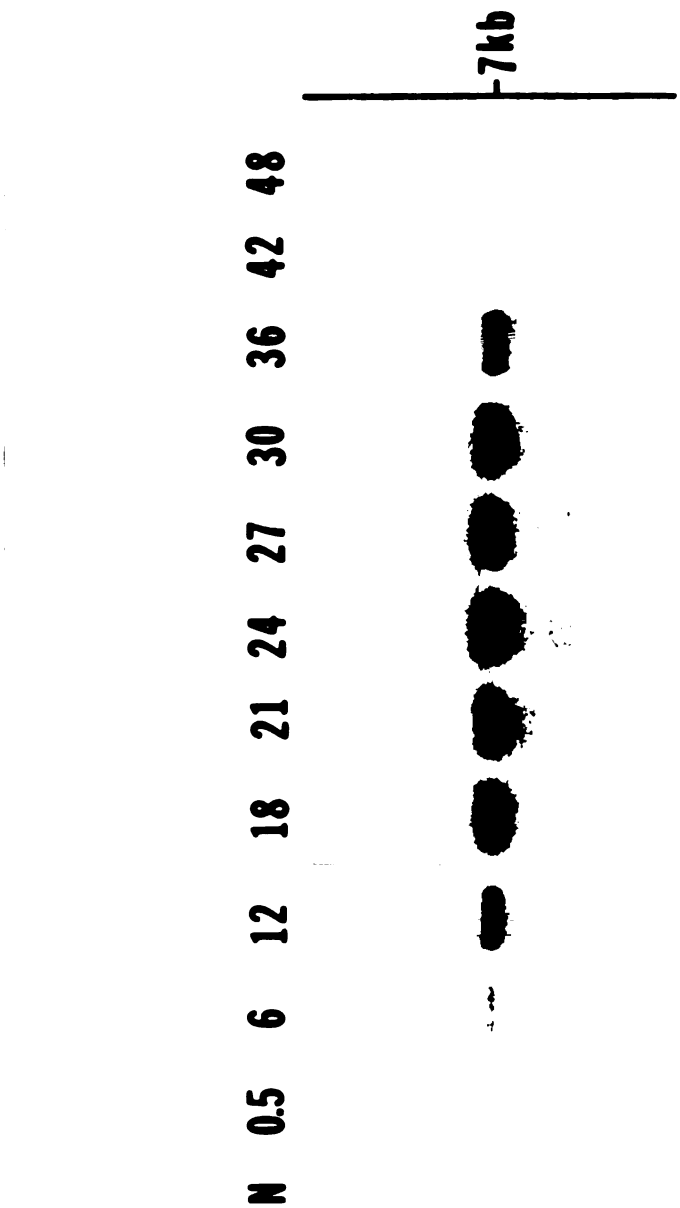
For quantitation, radioactive Northern blots of mLAP and Vg were exposed on "phosphor" screens (Molecular Dynamics Co) for several hours and then the screens were scanned with "Phosphor" Imager (Molecular Dynamics). The data were analyzed using the ImageQuant$^{TM}$ program.

***Protein preparation and Western blot analysis of mLAP and Vg***

Mosquito fat bodies were homogenized in the homogenization buffer: 100 mM sodium phosphate buffer, pH 7.0, 150 mM NaCl, 0.02% sodium azide. The homogenization buffer was supplemented with a mixture of protease inhibitors: 1 mM phenylmethylsulfonyl fluoride, 5 mM α-amino-caprionic acid, 1 mM

**Fig. 4. Northern blot analysis of Vg mRNA in mosquito fat bodies during vitellogenesis.**

Fat bodies, dissected from non-blood fed mosquitoes (N) and from vitellogenic mosquitoes at indicated hours after a blood meal, were used for extractions of total RNA Total RNA was resolved by 1.2% formaldehyde agarose gel electrophoresis. The amount of RNA corresponding to the fat body of one mosquito was loaded into each lane. The RNA was transferred to a nitrocellulose membrane. A 1.8 kb fragment of the mosquito Vg gene A1, labeled with [$^{32}$P], was used for hybridization. The molecular weight of Vg RNA is marked on the right.

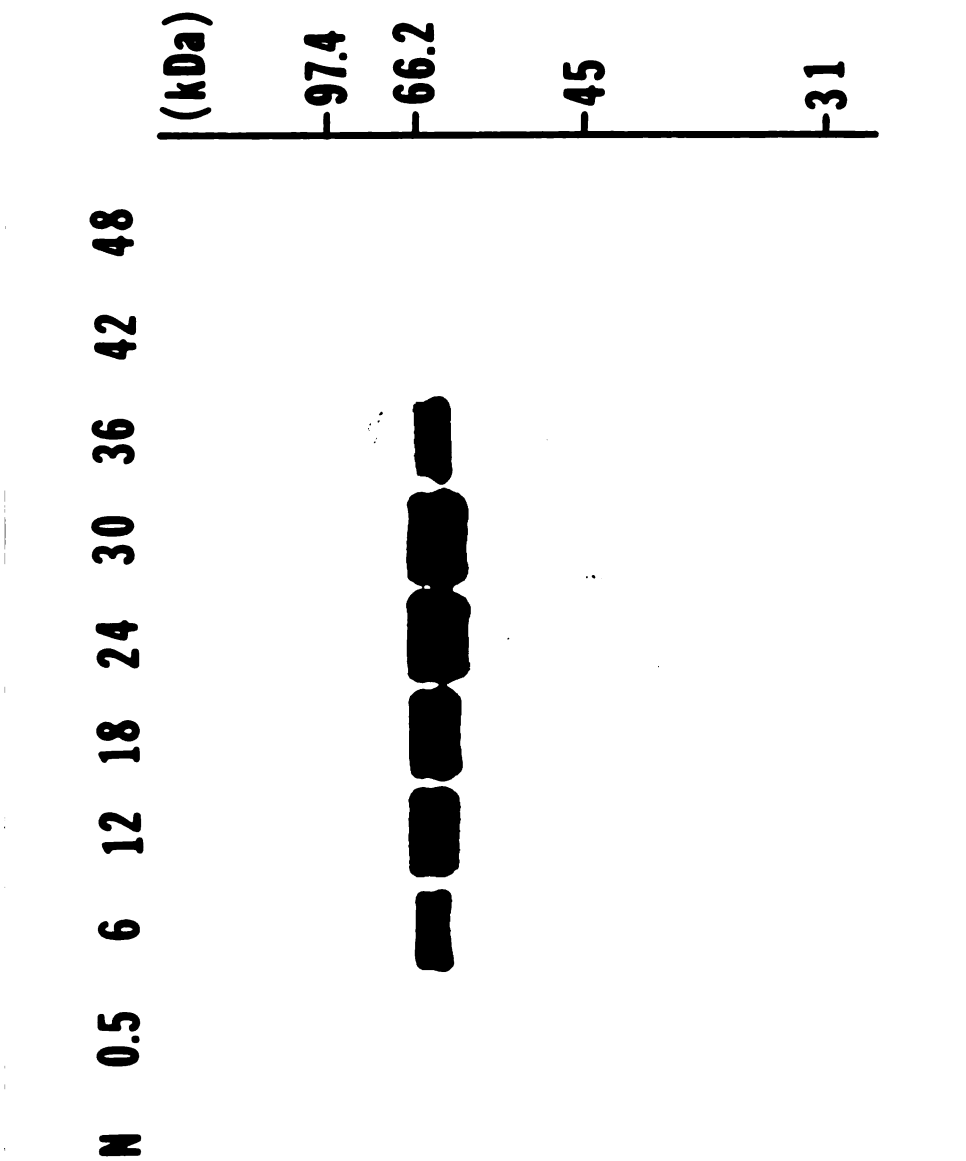

7kb
N 0.5 6 12 18 21 24 27 30 36 42 48

benzamidine, 10 ug/ml aprotinin, and 2 ug/ml each antipain, leupeptin and chymostatin. The homogenate was centrifuged at 27,216 x g for 30 min. The supernatant was used for 9% SDS-PAGE under reducing conditions. The amount of total protein corresponding to the fat body of one mosquito was loaded into each lane. Resolved proteins were transferred to nitrocellulose membrane using Bio-Rad Blotting Apparatus. Mosquito LAP was detected with polyclonal antibodies (Cho *et al.*, 1991a) and a ECL Western blotting detection system. Vitellogenin was detected with a mixture of two monoclonal antibodies (ratio 1:1), both directed against a small Vg subunit, 65 kDa (clone 2G1 and B11D12; Raikhel et al., 1986) and the same detection system (Fig. 5). The ECL Western blotting detection system was applied using the manufacturer's protocol (Amersham).

Chemiluminescent immunoblots were exposed to X-ray films. The films with bands below saturated exposure were subjected to Computing Densitometer Model 300A (Molecular Dynamics). The quantitation was performed with the program that was used for analyzing the Northern blot data.

**Fig. 5. Western blot analysis of Vg in the mosquito fat body during vitellogenesis.**

Fat bodies collected as described in Fig. 4 were subjected to protein extraction. Total proteins were resolved by 9% SDS-PAGE under reduction conditions. The amount of total proteins equal to the fat body of one mosquito was loaded into each lane. Vitellogenin is detected with monoclonal antibodies against to a small Vg subunit. The molecular weight markers are in order of decreasing $M_r$ phosphatase b, bovine serum albumin, ovalbumin, and carbonic anhydrase.


(kDa)
97.4
66.2
45
31
N
0.5
6
12
18
24
30
36
42
48

# RESULTS AND DISCUSSION

## I. ANALYSIS OF cDNA CODING FOR MOSQUITO LYSOSOMAL ASPARTIC PROTEASE

### *Cloning and analysis of mLAP cDNA*

The cDNA encoding mLAP was cloned by an approach using the polymerase chain reaction (PCR) and screening of a λZAP II cDNA library specific to the fat body of vitellogenic female mosquitoes.

The PCR primers were designed and synthesized based on the amino acid sequence of the N-terminal of mLAP (primer 1), and the first catalytic domain of aspartic proteases (primer 2). A 96 base pair fragment was obtained by PCR using these primers (for more details see "EXPERIMENTAL PROCEDURES"). The sequence of the 96 bp fragment revealed a 70% identity with human cathepsin D. Both ends of this fragment matched the PCR primer sequences. The PCR-amplified fragment was subcloned in pUC119 and a 86 bp fragment was released by enzyme digestion. On a Northern blot of total RNA from vitellogenic fat bodies 36 hr post-blood meal (PBM), it hybridized with a 1.5 kb mRNA. This probe was used to screen a λZAP II cDNA library specific to the vitellogenic fat bodies. Of forty putative mLAP clones obtained the clone with the longest insert (1,420 base pairs) was sequenced. The sequence of this cDNA, obtained with the strategy shown in Fig. 3, is presented in Fig. 6. There is an open reading frame of 1,164 nucleotides and a polyadenylation signal (AATAAA) at 19 nucleotides upstream of the poly(A) tail.

### *Analysis of the deduced amino acid sequence of mLAP*

The predicted amino acid sequence of mLAP is 387 residues (predicted

**Fig. 6. Nucleotide and deduced amino acid sequence of mosquito lysosomal aspartic protease (mLAP).**

The putative signal peptide is indicated in bold letters. The putative pro-enzyme sequence is underlined. The amino acid sequence matching the N-terminus of the purified mLAP, determined by microsequencing is boxed. A potential glycosylation residue (N) is marked by a solid square. Amino acids (D) of conserved catalytic sites are marked by diamond-shaped marks. The putative polyadenylylation signal (AATAAA) is underlined by a thick broken line. Positive numbers are counted from the first amino acid of the mature enzyme.

```
   1  GGCACGAGCGGCACGAGGGGTTTCACAGCAAGTAGAACGTCCTTTTTATTGTTCCTCTTT   60

  61  CCAAAGGAAGAATCCTTCTAATAGCAGACGTTGACGATAAGCCTTTGAAATCATCGCTCT  120

 121  CTAGCCGGCGAGATGCTAATTAAATCAATTATTGCCCTCGTTTGCTTGGCCGTTCTATCC  180
 -53                M  L  I  K  S  I  I  A  L  V  C  L  A  V  L  S  -38

 181  CAGGCGGACTTTGTTAGAGTTCAGCTGCATAAAACTGAAAGTGCGCGTCAGCATTTTCGA  240
 -37   Q  A  D  F  V  R  V  Q  L  H  K  T  E  S  A  R  Q  H  F  R  -18

 241  AATGTAGACACCGAGATCAAACAATTGCGTCTCAAGTATAATGCTGTATCTGGACCGGTG  300
 -17   N  V  D  T  E  I  K  Q  L  R  L  K  Y  N  A  V  S  G  P  V    3

 301  CCTGAGCCGCTTTCGAACTATTTGGATGCCCAATACTATGGAGCTATCACCATCGGAACA  360
   4   P  E  P  L  S  N  Y  L  D  A  Q  Y  Y  G  A  I  T  I  G  T   23

 361  CCACCGCAGAGCTTCAAAGTTGTGTTCGATACGGGATCATCTAACCTTTGGGTGCCCTCG  420
  24   P  P  Q  S  F  K  V  V  F  D  T  G  S  S  N  L  W  V  P  S   43

 421  AAGGAGTGCTCATTCACCAACATCGCTTGCTTGATGCACAACAAATACAATGCCAAGAAG  480
  44   K  E  C  S  F  T  N  I  A  C  L  M  H  N  K  Y  N  A  K  K   63

 481  TCATCGACGTTCGAAAAGAACGGAACAGCTTTCCATATTCAATATGGATCTGGTAGCTTA  540
  64   S  S  T  F  E  K  N  G  T  A  F  H  I  Q  Y  G  S  G  S  L   83

 541  TCTGGTTACTTGTCAACTGACACCGTTGGTTTGGGAGGGGTTTCCGTTACGAAACAAACC  600
  84   S  G  Y  L  S  T  D  T  V  G  L  G  G  V  S  V  T  K  Q  T  103

 601  TTCGCTGAAGCCATCAATGAACCAGGATTGGTATTCGTTGCGGCCAAGTTTGACGGAATT  660
 104   F  A  E  A  I  N  E  P  G  L  V  F  V  A  A  K  F  D  G  I  123

 661  CTCGGATTAGGCTACAGCTCGATTTCAGTAGATGGCGTCGTACCAGTATTCTACAATATG  720
 124   L  G  L  G  Y  S  S  I  S  V  D  G  V  V  P  V  F  Y  N  M  143

 721  TTCAACCAGGGTCTCATCGATGCTCCCGTTTTCTCTTTCTATTTGAATCGTGATCCAAGT  780
 144   F  N  Q  G  L  I  D  A  P  V  F  S  F  Y  L  N  R  D  P  S  163

 781  GCTGCTGAGGGTGGCGAAATTATTTTCGGTGGATCAGACTCGAATAAGTATACTGGGGAC  840
 164   A  A  E  G  G  E  I  I  F  G  G  S  D  S  N  K  Y  T  G  D  183

 841  TTTACTTATCTGTCGGTGGACCGTAAAGCCTACTGGCAATTCAAAATGGACTCCGTTAAG  900
 184   F  T  Y  L  S  V  D  R  K  A  Y  W  Q  F  K  M  D  S  V  K  203

 901  GTTGGCGATACTGAGTTCTGCAACAATGGATGCGAAGCAATTGCCGATACCGGCACCAGC  960
 204   V  G  D  T  E  F  C  N  N  G  C  E  A  I  A  D  T  G  T  S  223

 961  TTGATTGCCGGCCCAGTGTCGGAGGTCACCGCTATCAACAAGGCTATCGGTGGCACTCCT 1020
 224   L  I  A  G  P  V  S  E  V  T  A  I  N  K  A  I  G  G  T  P  243

1021  ATTATGAACGGAGAATACATGGTTGACTGCTCGTTGATTCCCAAACTGCCAAAGATCTCA 1080
 244   I  M  N  G  E  Y  M  V  D  C  S  L  I  P  K  L  P  K  I  S  263

1081  TTCGTTTTGGGAGGAAAATCATTCGATCTCGAAGGTGCTGATTACGTACTGCGTGTGGCT 1140
 264   F  V  L  G  G  K  S  F  D  L  E  G  A  D  Y  V  L  R  V  A  283

1141  CAAATGGGTAAAACCATCTGCCTGTCTGGGTTCATGGGAATCGATATTCCACCGCCTAAT 1200
 284   Q  M  G  K  T  I  C  L  S  G  F  M  G  I  D  I  P  P  P  N  301

1201  GGACCGTTGTGGATTTTGGGAGACGTTTTCATTGGTAAATATTACACCGAATTCGATATG 1260
 304   G  P  L  W  I  L  G  D  V  F  I  G  K  Y  Y  T  E  F  D  M  323

1261  GGCAATGATCGCGTTGGATTTGCCACTGCTGTCTAAAGAATTGATAGATTTGTATTGGTA 1320
 324   G  N  D  R  V  G  F  A  T  A  V                             334

1321  AAAATACCTGCACATTTCCAGTTCCAAAAATATTATTAGAAAGTGTGCATTACTGAAAAT 1380

1381  AAAATGTGAGTTAAACTCGTGCCAAAAAAAAAAAAAAAAA                     1420
```

$M_r$=41,890). Hydropathy analysis of the deduced amino acid sequence (Kyte and Doolittle 1982) indicated the N-terminal portion of mLAP (amino acid residues -53 to -36) is highly hydrophobic. This feature is typical of a signal pre-propeptide sequence (Fig. 7). A 18-residue N-terminal portion of mLAP is a putative signal peptide, according to the (-3, -1) rule for signal sequence cleavage sites (von Heijne, 1986).

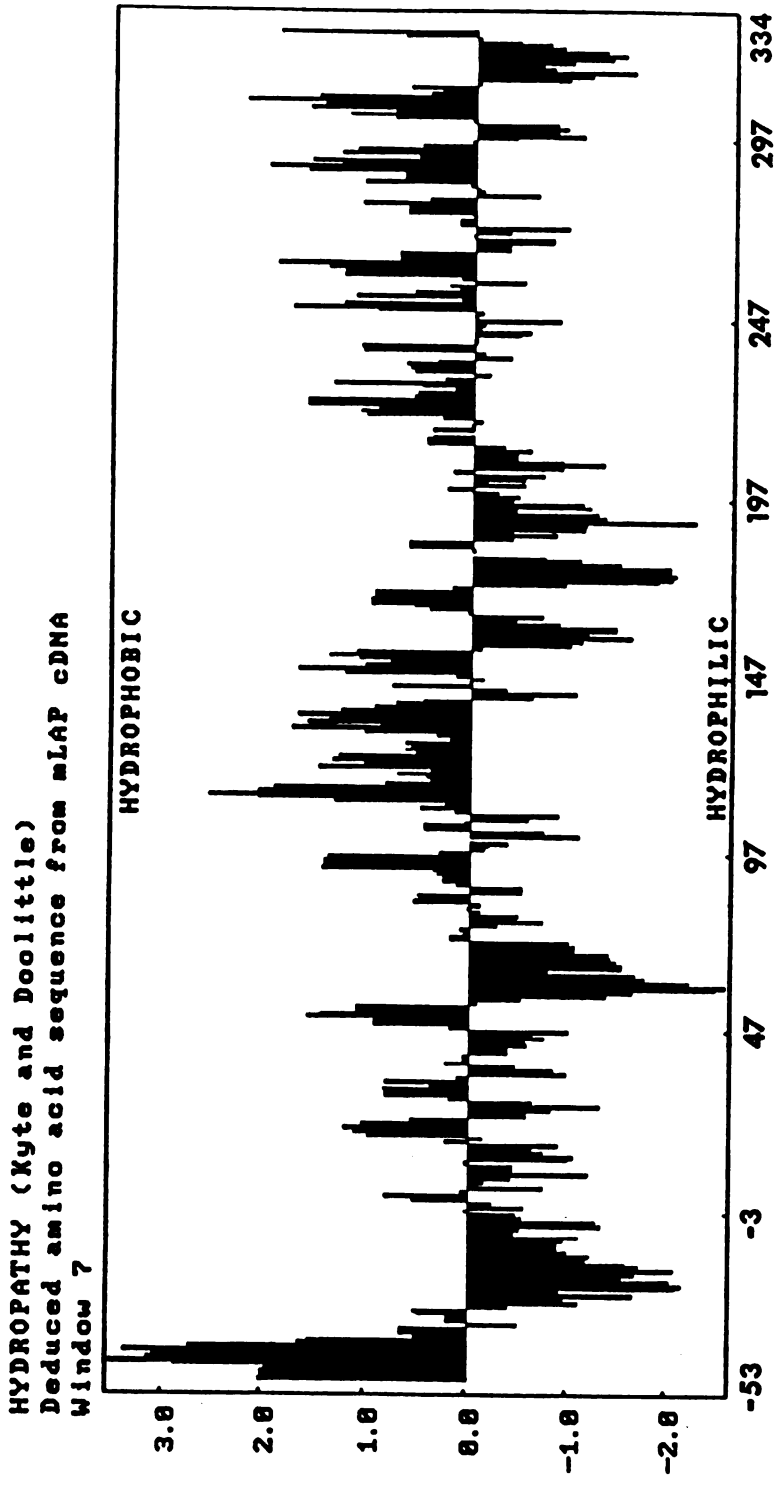
Lysosomal aspartic proteases, such as cathepsins D, are synthesized as pro-enzymes and activated in lysosomes by cleavage of pro-enzyme sequences (Yonezawa *et al.*, 1988). A comparison of the deduced amino acid sequence of mLAP with the N-terminal sequence of a mature mLAP (Cho *et al.*, 1991a), suggests that 35 amino acids of mLAP is a putative pro-enzyme sequence (Fig. 6). The mature subunit form of the mLAP, therefore, is 334 amino acids with the predicted $M_r$ of 35,812 Da. The mLAP subunit, estimated using reducing SDS-PAGE, is 40 kDa (Cho *et al.*, 1991a). There is a single potential N-linked glycosylation site at the position Asn-70 (Fig. 6). We expected a glycosylation site because mLAP is a glycoprotein (Cho *et al.*, 1991a). The increase in mature mLAP $M_r$ is likely from glycosylation.

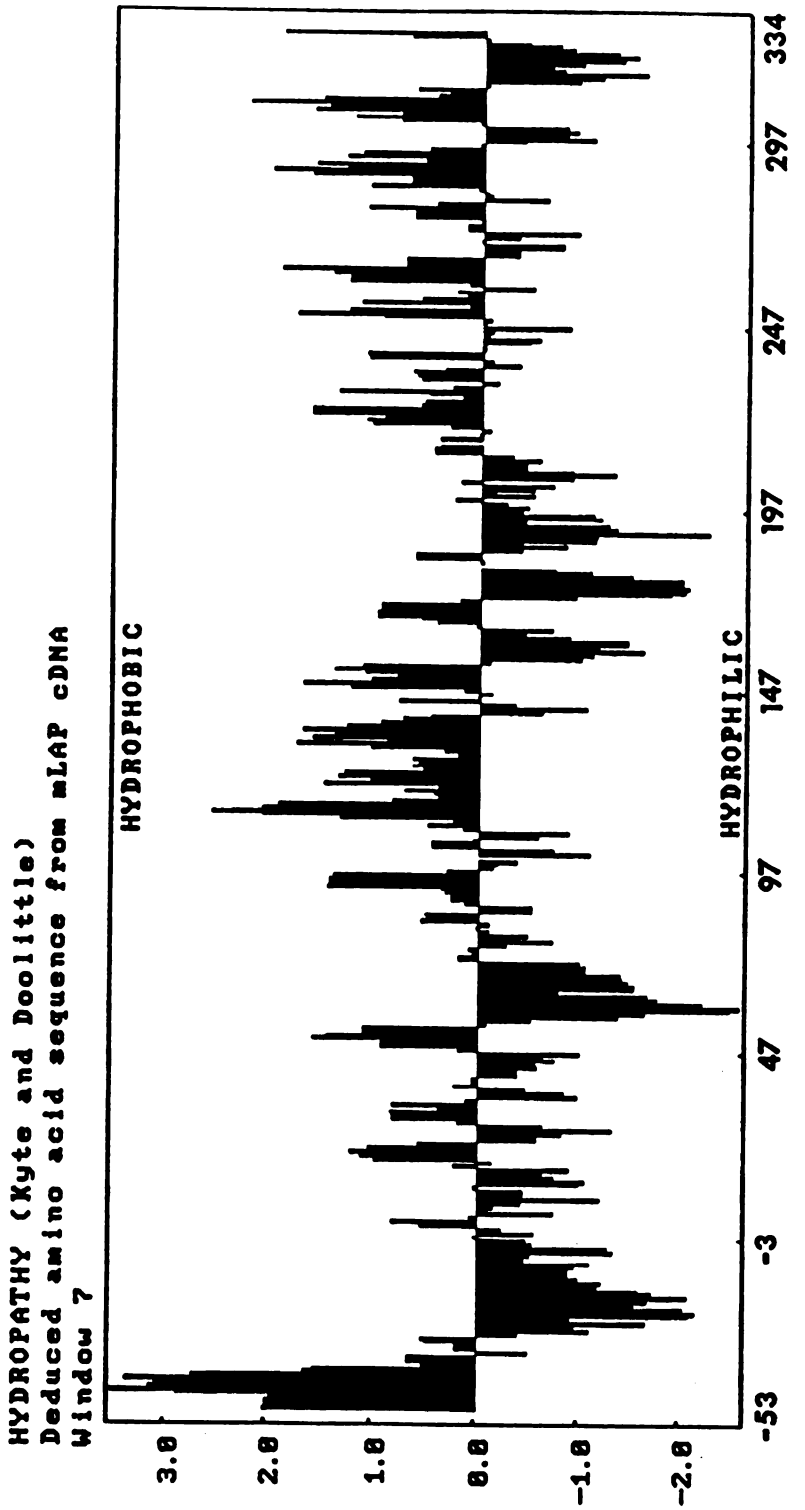
***Comparison of the mLAP amino acid sequence to aspartic protease sequences***

Comparison of the mLAP amino acid sequence to deduced amino acid sequences of lysosomal and non-lysosomal aspartic proteases revealed significant similarities (Table 1). High similarities exist between mLAP and cathepsins D from several mammalian species, human cathepsin E, and yeast vacuolar proteinase A (Table 1). The sequence identity is 58% between mLAP and human cathipsin D (HCD), and 44% between mLAP and human cathepsin E (HCE). However, considering both identical and conserved replacements, the

**Fig. 7. Hydropathy plot of the deduced amino acid sequence of mLAP.**

The hydropathy plot was obtained by using the algorithm of Kyte and Doolittle (1982). The ordinate gives the hydropathic index. The abscissa indicates the amino acid position corresponding to the deduced sequence in Fig. 6.


HYDROPATHY (Kyte and Doolittle)
Deduced amino acid sequence from mLAP cDNA
Window 7
HYDROPHOBIC
HYDROPHILIC
3.0
2.0
1.0
0.0
-1.0
-2.0
-53
-3
47
97
147
197
247
297
334


HYDROPATHY (Kyte and Doolittle)
Deduced amino acid sequence from mLAP cDNA
Window 7
HYDROPHOBIC
HYDROPHILIC
3.0
2.0
1.0
0.0
-1.0
-2.0
-53
-3
47
97
147
197
247
297
334

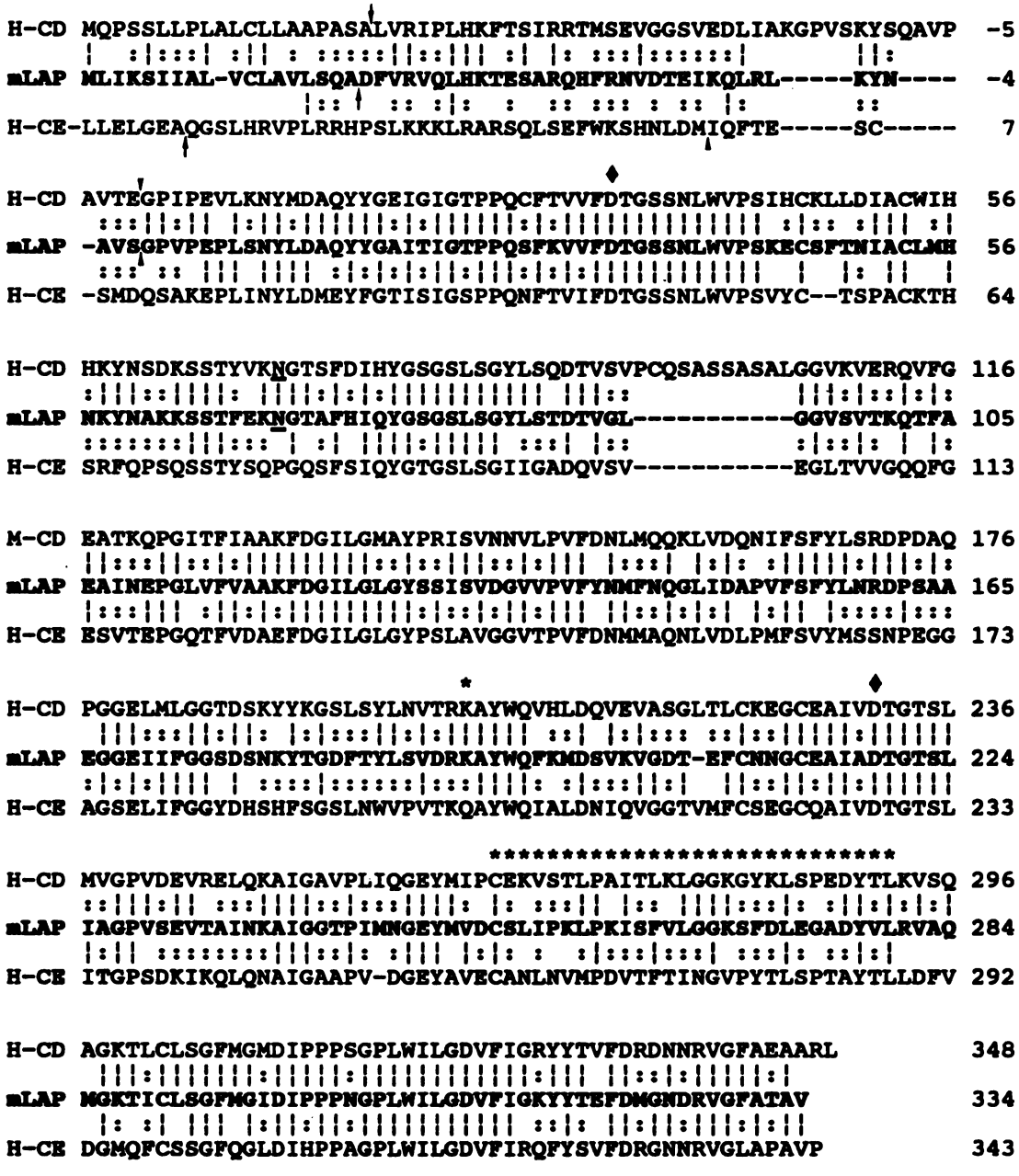
similarity is 92% between mLAP and HCD and 81% between mLAP and HCE (Fig. 8). Although the putative cleavage sites for the signal peptide and the pro-enzyme in mLAP are at the positions similar to those in HCD (Faust *et al.*, 1985), these sites are not conserved in the HCE sequence (Athauda *et al.*, 1990; Azuma *et al.*, 1989).

Catalytic activity of aspartic proteases depends on two aspartic acid residues at active centers (Shewale and Tang, 1984). These two catalytic centers with aspartic acid residues (Asp-33 and Asp-219) as well as surrounding residues, are conserved in mLAP (Fig. 8).

Human CD has two N-linked glycosylation sites, but mLAP and HCE each possesses only one glycosylation site. The position of the mLAP glycosylation site is identical to the first HCD site but not to the HCE one. In mammalian lysosomal enzymes, the mannose 6-phosphate residues serve as the recognition marker for the targeting of these enzymes to lysosomes (Kornfeld, 1987). Lysine 203 and amino acids 265-292 of HCD are required for phosphorylation of its high-mannose oligosaccharides (Baranski *et al.*, 1990; 1991). In contrast, the first phosphorylation determinant is missing and the second determinant is not well conserved in HCE. Cathepsin E is a non-lysosomal enzyme which is associated with either the plasma membrane or cytosol in gastric mucosa and neutrophils (Ueno *et al.*, 1989; Ichimaru *et al.*, 1990). The first determinant (Lys 203 of HCD) is found at Lys 192 in mLAP. The residues surrounding Lys 192 are conserved in mLAP (Fig. 8). The second phosphorylation determinant of mLAP (C-253 to L-283) is 43.3% identical to the corresponding sequence of HCD (C-265 to L-293). Considering functional substitutions, the similarity between these sequences of mLAP and HCD becomes 86.7% (Fig. 8).

**Fig. 8. Alignment of amino acid sequences of mLAP, human cathepsin D (HCD) and human cathepsin E (HCE).**

Amino acid sequences deduced from corresponding cDNAs were aligned by the FASTA computer program. Vertical dash lines indicate identical residues, and colons denote functional substitutions. Only a portion of HCE sequence is presented. The small arrow indicates the putative cleavage site for the signal peptide. The arrowhead refers to the cleavage sites for the pro-enzyme. The potential glycosylation residues are marked with thick lines. The conserved enzyme catalytic residues are denoted with diamond-shaped marks. Phosphorylation determinants, lysine 203 and the amino acid stretch 265-292 of HCD, are marked by asterisks. Dashed lines represent gaps.


H-CD MQPSSLLPLALCLLAAPASALVRIPLHKFTSIRRTMSEVGGSVEDLIAKGPVSKYSQAVP -5
mLAP MLIKSIIAL-VCLAVLSQADFVRVQLHKTESARQHFRNVDTEIKQLRL-----KYN---- -4
H-CE-LLELGEAQGSLHRVPLRRHPSLKKKLRARSQLSEFWKSHNLDMIQFTE-----SC----- 7
H-CD AVTEGPIPEVLKNYMDAQYYGEIGIGTPPQCFTVVFDTGSSNLWVPSIHCKLLDIACWIH 56
mLAP -AVSGPVPEPLSNYLDAQYYGAITIGTPPQSFKVVFDTGSSNLWVPSKECSFTMIACLMH 56
H-CE -SMDQSAKEPLINYLDMEYFGTISIGSPPQNFTVIFDTGSSNLWVPSVYC--TSPACKTH 64
H-CD HKYNSDKSSTYVKNGTSFDIHYGSGSLSGYLSQDTVSVPCQSASSASALGGVKVERQVFG 116
mLAP NKYNAKKSSTFEKNGTAFHIQYGSGSLSGYLSTDTVGL----------GGVSVTKQTFA 105
H-CE SRFQPSQSSTYSQPGQSFSIQYGTGSLSGIIGADQVSV----------EGLTVVGQQFG 113
M-CD EATKQPGITFIAAKFDGILGMAYPRISVNNVLPVFDNLMQQKLVDQNIFSFYLSRDPDAQ 176
mLAP EAINEPGLVFVAAKFDGILGLGYSSISVDGVVPVFYNMFNQGLIDAPVFSFYLNRDPSAA 165
H-CE ESVTEPGQTFVDAEFDGILGLGYPSLAVGGVTPVFDNMMAQNLVDLPMFSVYMSSNPEGG 173
H-CD PGGELMLGGTDSKYYKGSLSYLNVTRKAYWQVHLDQVEVASGLTLCKEGCEAIVDTGTSL 236
mLAP EGGEIIFGGSDSNKYTGDFTYLSVDRKAYWQFKMDSVKVGDT-EFCNNGCEAIADTGTSL 224
H-CE AGSELIFGGYDHSHFSGSLNWVPVTKQAYWQIALDNIQVGGTVMFCSEGCQAIVDTGTSL 233
H-CD MVGPVDEVRELQKAIGAVPLIQGEYMIPCEKVSTLPAITLKLGGKGYKLSPEDYTLKVSQ 296
mLAP IAGPVSEVTAINKAIGGTPIMNGEYMVDCSLIPKLPKISFVLGGKSFDLEGADYVLRVAQ 284
H-CE ITGPSDKIKQLQNAIGAAPV-DGEYAVECANLNVMPDVTFTINGVPYTLSPTAYTLLDFV 292
H-CD AGKTLCLSGFMGMDIPPPSGPLWILGDVFIGRYYTVFDRDNNRVGFAEAARL 348
mLAP MGKTICLSGFMGIDIPPPNGPLWILGDVFIGKYYTEFDMGNDRVGFATAV 334
H-CE DGMQFCSSGFQGLDIHPPAGPLWILGDVFIRQFYSVFDRGNNRVGLAPAVP 343

This high conservation of determinants of oligosaccharide phosphorylation suggests mLAP possesses a mannose-6-phosphate residue. The actual status of the mLAP oligosaccharide moiety, with respect to its phosphorylation, requires further investigation.

Multiple sequence alignment of both lysosomal (group I) and non-lysosomal (group II) aspartic proteases were done (Fig. 9). Only vertebrate cathepsins D have an additional stretch of amino acids the "ß-hairpin" which is a proteolytic processing region (Yonezawa *et al.*, 1988). In vertebrate lysosomes, the mature cathepsins D are generated after two steps: the first removes their pro-enzyme sequences, and the second cleaves at the ß-hairpin to form the heavy and light chain subunits (Faust *et al.*, 1985; Conner, 1989). Proteolytic conversion may be important for the stability of the cathepsin D tertiary structure (Yonezawa *et al.*, 1988). The maturation of the other aspartic proteases is different. The yeast vacuolar proteinase A is cleaved once at the pro-enzyme sequence (Faust and Kornfeld, 1989). The barley aspartic protease does not have a cleavage site similar to that of cathepsin D. Rather it is cleaved at new sites, resulting in a heterodimeric form of an active mature enzyme (Runeberg-Ross *et al.*, 1991). The mLAP forms a dimer from two identical 40-kDa subunits (Cho *et al.*, 1991a).

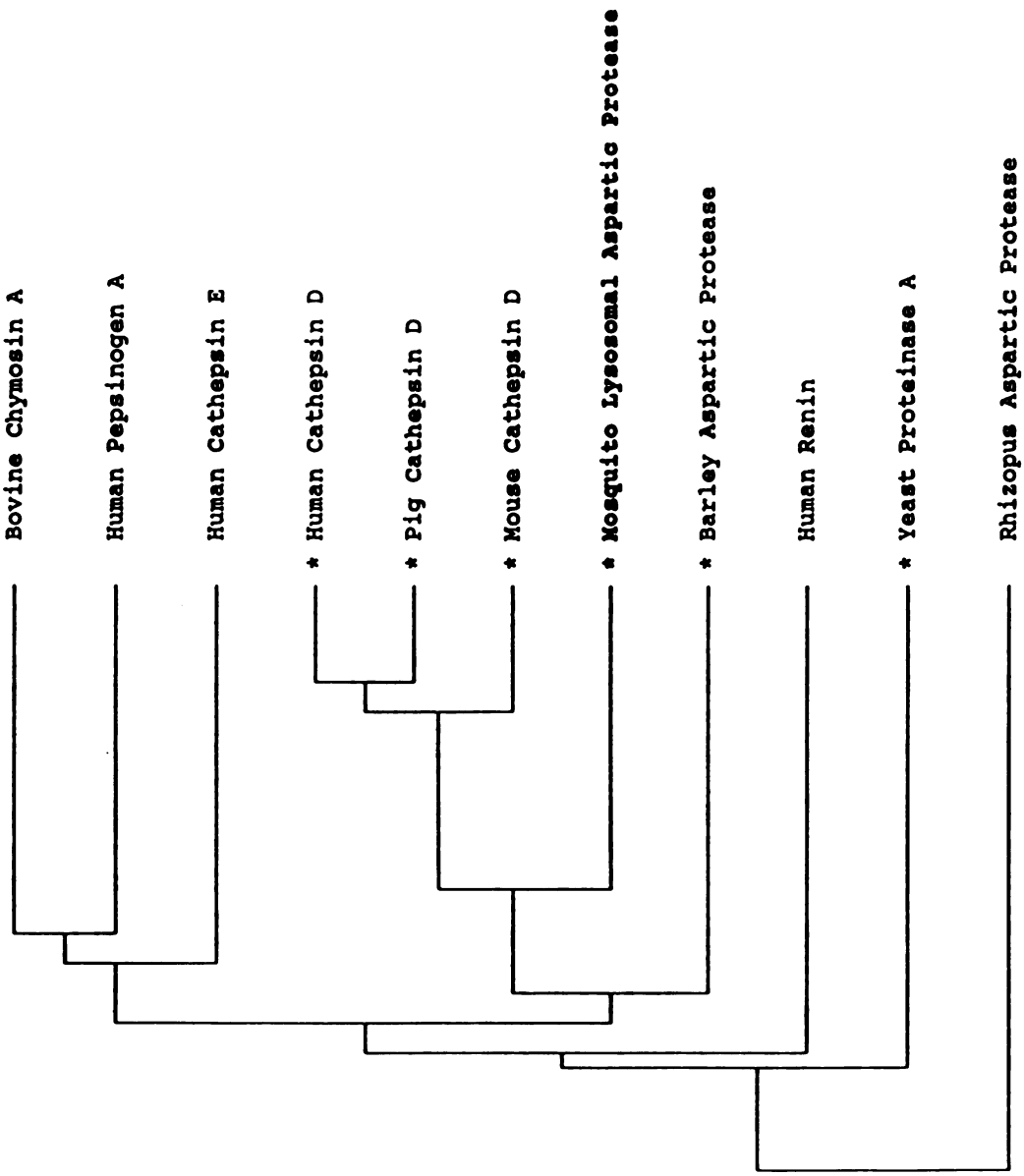
Tang and Wong (1987) suggested all aspartic proteases are derived from the same ancestral enzyme. It is clear that mLAP is most similar to cathepsin D. Mosquito LAP, however, shares with cathepsin E the dimeric structure, low optimal catalytic pH value, and lack of a ß-hairpin processing sequence. The pairwise computer comparison provided further evidence as for the evolutionary relationships between these members of a family of aspartic proteases. A dendrogram, generated by this analysis, is shown in Figure 10.

```
GROUP-I
Mosquito Aspartic Protease  --dtvgl ------------------------ ggvs--
Human Cathepsin D           --dtvsv P C Q'S A S S A S A'L ggvk--
Mouse Cathepsin D           --dtvsv P C K S D Q S K A ---- rgik--
Pig Cathepsin D             --dtvsv P C'N S A L S G'V ---- ggik--
Bovine Cathepsin D          --dtvsv P C N P S'S S'S P ---- ggvt--
Barley Aspartic Protease    --dsvtv ------------------------ gdlv--
Yeast Proteinase A          --dtlsi ------------------------ gdlt--
GROUP-II
Human Cathepsin E           --dqvsv ------------------------ eglt--
Human Renin                 --diitv ------------------------ ggit--
Human Pepsinogen A          --dtvqv ------------------------ ggis--
Bovine Chymosin A           --dtvtv ------------------------ sniv--
Rhizopus Aspartic Protease  --dnvnl ------------------------ ggll--
```

**Fig. 9. Sequence alignment of regions of aspartic proteases corresponding to the cleavage sites of cathepsins D.**

Several aspartic proteases were compared by a multiple alignment computer program. Corresponding partial sequences are presented. Lower case letters indicate homologous residues, capital letters are additional residues found in cathepsins D. Arrowheads denote the identified processing sites for the heavy and light chains of cathepsins D. Solid lines indicate gaps. The sequences of the following enzymes were used: mLAP, cathepsins D (Shewale and Tang, 1984; Faust *et al.*, 1985; Grusby *et al.*, 1990), yeast proteinase A (Ammere *et al.*, 1986), cathepsin E (Azuma *et al.*, 1989), renin (Imai *et al.*, 1983), pepsinogen (Sogawa *et al.*, 1983), barley-grain aspartic protease (Runeberg-Roos *et al.*, 1991), bovine chymosin (Harris *el al.*, 1982) and rhizopus aspartic protease (Horiuchi *et al.*, 1988). Group I includes aspartic proteases of lysosomal/vacuolar origin, group II non-lysosomal and secretory enzymes.

**Fig. 10. Relationship between members of a family of aspartic proteases.** The dendrogram of aspartic proteases was generated by pairwise comparison of deduced amino acid sequences of aspartic proteases. The distance along the vertical axis is proportional to the difference between sequences. Asterisks indicate lysosomal or vacuolar enzymes.


Bovine Chymosin A
Human Pepsinogen A
Human Cathepsin E
* Human Cathepsin D
* Pig Cathepsin D
* Mouse Cathepsin D
* Mosquito Lysosomal Aspartic Protease
* Barley Aspartic Protease
Human Renin
* Yeast Proteinase A
Rhizopus Aspartic Protease

## II. EXPRESSION OF MOSQUITO LYSOSOMAL ASPARTIC PROTEASE IN THE FAT BODY DURING VITELLOGENESIS

In fat bodies of vitellogenic female mosquitoes, the lysosomal activity increases dramatically by the end of Vg synthesis. The specific activities of several lysosomal enzymes, including cathepsin D-like activity, rise at 24-26 hr after the initiation of Vg synthesis by a blood meal and reach a maximum at 36-42 hr PBM (Raikhel, 1986a). In this study, we determined the expression kinetics of the mLAP mRNA and compared it to the kinetics of mLAP protein in mosquito fat bodies during the Vg synthetic cycle. As a control, we monitored the kinetics of Vg at both mRNA and protein levels in fat bodies for the same time periods.

Mosquito LAP mRNA levels increased between 6 and 12 hr PBM. The levels were highest at 24 hr PBM, and then gradually declined to the background level by 48 hr PBM (Figs. 11 and 12). In contrast, mLAP protein levels monitored by immunoblot increased 12 hr later than mLAP mRNA (Figs. 11 and 12). Mosquito LAP protein was elevated at 24 hr PBM, but the highest level was between 36 hr and 42 hr PBM. The increase in mLAP protein paralleled the activity of a cathepsin D-like enzyme (Raikhel, 1986a).

We examined the accumulation of Vg mRNA and Vg protein during Vg synthesis in the mosquito fat body. Unlike mLAP, both Vg mRNA and Vg protein accumulated at similar rates reaching maximal levels at 24 hr PBM, and then rapidly declining (Fig. 12). Racioppi *et al.* (1986) reported that the amount of Vg mRNA reached its peak levels at 36 hr PBM. Accoring to our data, however, both Vg mRNA and Vg protein have similar kinetics and peak at 24 hr PBM.

**Fig. 11. Expression of mLAP in the mosquito fat body during vitellogenesis.** Fat bodies, dissected from non-blood fed mosquitoes (N) and from vitellogenic mosquitoes at indicated hours after a blood meal, were used for extractions of total RNA and proteins. (A) Total RNA was resolved by 1.2% formaldehyde agarose gel electrophoresis. The amount of RNA corresponding to the fat body of one mosquito was loaded into each lane. The RNA was transferred to a nitrocellulose membrane, and hybridized with a [$^{32}$P]-labeled 593 bp mLAP probe. A RNA ladder (BRL) is shown to the right. (B) Total proteins were resolved by 9% SDS-PAGE under reducing conditions. The amount of total protein equal to the fat body of one mosquito was loaded into each lane. Polyclonal antibodies against mLAP were used for the immunoblot which was processed as described in "Experimental Procedures". The molecular weight markers are in order of decreasing $M_r$ phosphatase b, bovine serum albumin, ovalbumin, and carbonic anhydrase.

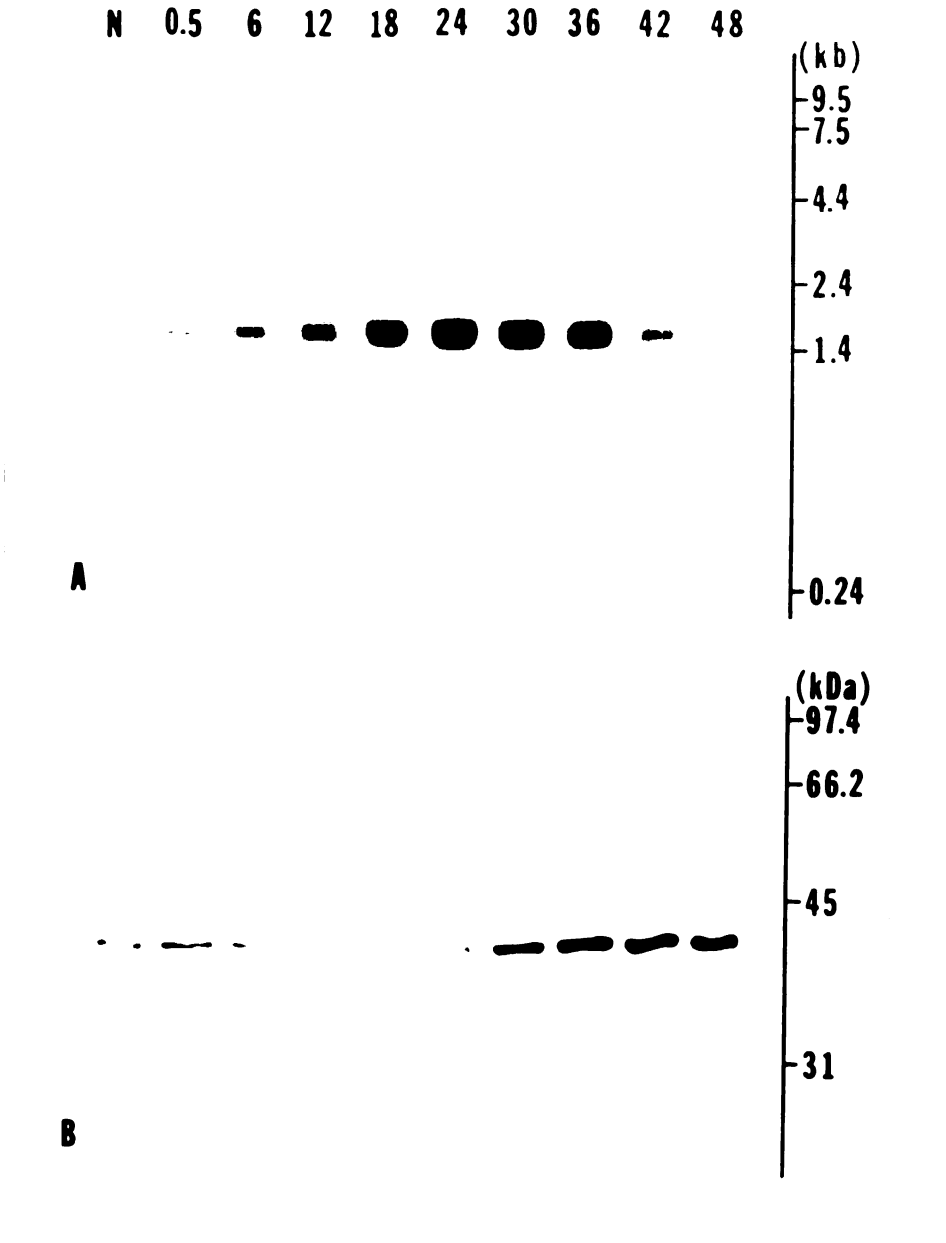
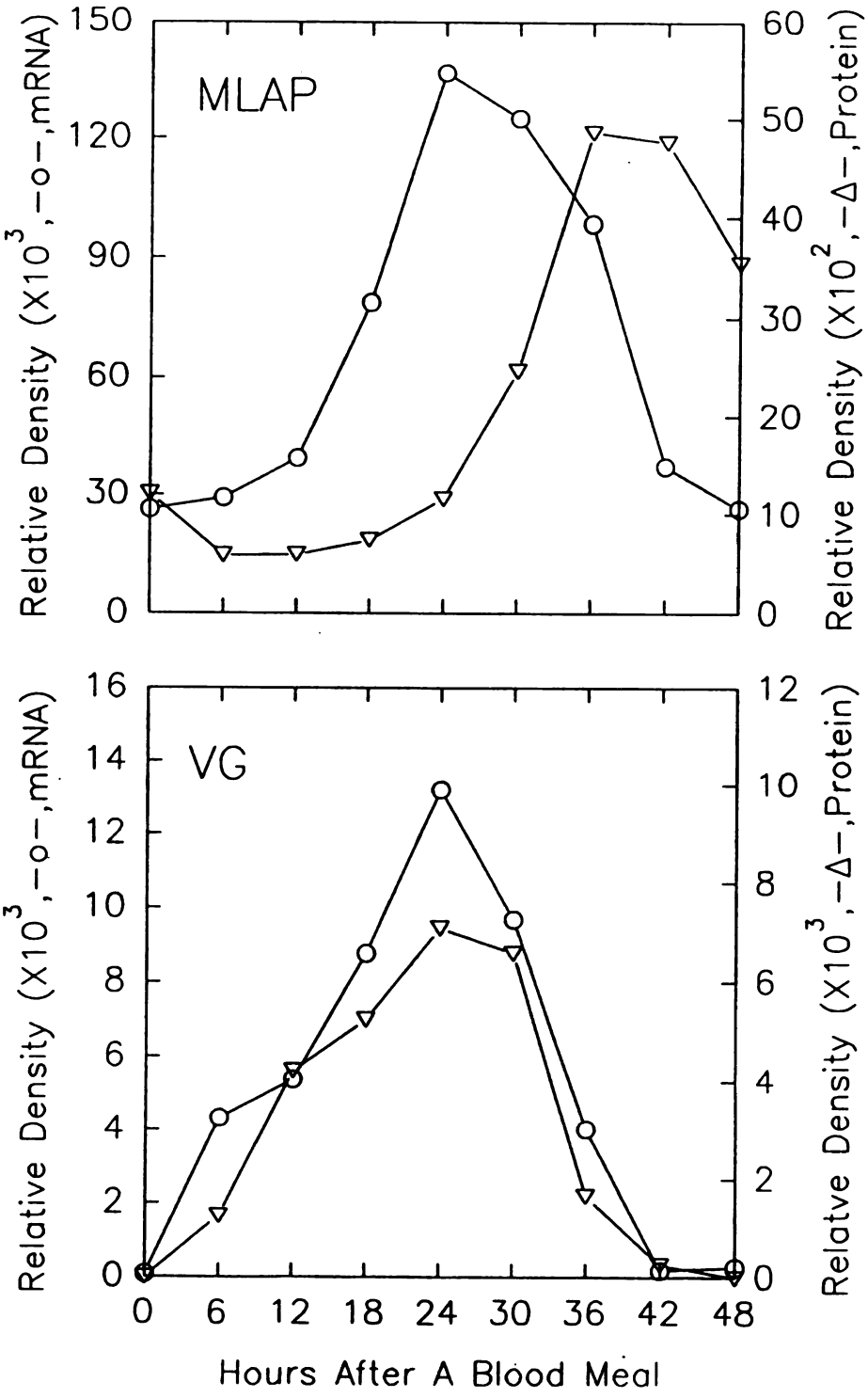

N
0.5
6
12
18
24
30
36
42
48
(kb)
9.5
7.5
4.4
2.4
1.4
0.24
A
(kDa)
97.4
66.2
45
31
B

**Fig. 12. Kinetics of mLAP and Vg at the mRNA and protein levels in the mosquito fat bodies during vitellogenesis.**

Quantitation of MRNA (-o-) and proteins (-∇-) were performed using methods as described in "Experimental procedures". Quantitation is based on Northern and Western blots of Figs. 4, 11 and 13. The upper panel is mLAP and the lower panel is Vg. Numbers on the bottom indicate the time after a blood meal when fat bodies were collected.


MLAP
VG
Relative Density (X10$^3$,-o-,mRNA)
Relative Density (X10$^2$,-Δ-,Protein)
Relative Density (X10$^3$,-o-,mRNA)
Relatve Density (X10$^3$,-Δ-,Protein)
Hours After A Blood Meal
0
6
12
18
24
30
36
42
48

Both Vg mRNA and mLAP mRNA reached their highest level 24 hr PBM. The hemolymph titer of an insect steroid hormone, 20-hydroxyecdysone (20-HE) is highest right before the peaks of mLAP and Vg mRNAs, at 18-20 hr PBM. After this time, 20-HE levels dramatically drop (Hagedorn, 1985). Since Vg synthesis is likely regulated at the transcriptional level by 20-HE (Hagedorn, 1985; 1989; Raikhel, 1992), 20-HE may also regulate mLAP gene expression. The 12 hr delay in the increase of mLAP protein, however, suggests mLAP synthesis is regulated at the translational level.

Steroid hormones have been implicated in regulation of gene expression at the transcriptional and post-transcriptional levels. Steroid hormones affect the stability of mRNAs (Brock and Shapiro, 1983; Gordon *et al.*, 1988; Nielsen and Shapiro, 1990) and also alter translation rates of certain mRNAs (Meyuhas *et al.*, 1987; Kumar *et al.*, 1989; Verdi *et al.*, 1989; Verdi and Campagnoni, 1990). Steroids effects can be either positive or negative.

The expression kinetics of mLAP at mRNA and protein levels suggests that 20-HE could be involved in the repressing translation of mLAP mRNA. It is unlikely that the stability of mLAP mRNA is affected since its level exhibits a steady decline after the peak, at 24 hr PBM (Fig. 12).

The translational inhibition of mLAP mRNA by 20-HE could be released when the titer of 20-HE declined 24 hr post-blood meal.

Verdi and Campagnoni (1990) identified a steroid modulatory element in the 5'-untranslated region of the myelin basic protein (MBP) mRNA. This element is involved into a positive regulation of translation by hydrocortisone. Putative modulatory sequences have been also identified in the 5'-untranslated region of 2',3'-cyclic nucleotide 3'-phosphodiesterase (CNP) mRNA whose translation is inhibited by hydrocortisone (Verdi and Campagnoni, 1990). We compared the 5'-untranslated region of mLAP mRNA to those of CNP and MBP. Three segments of the 5'-untranslated region of mLAP mRNA were found to be similar to sequences in CNP mRNA (Fig. 13).

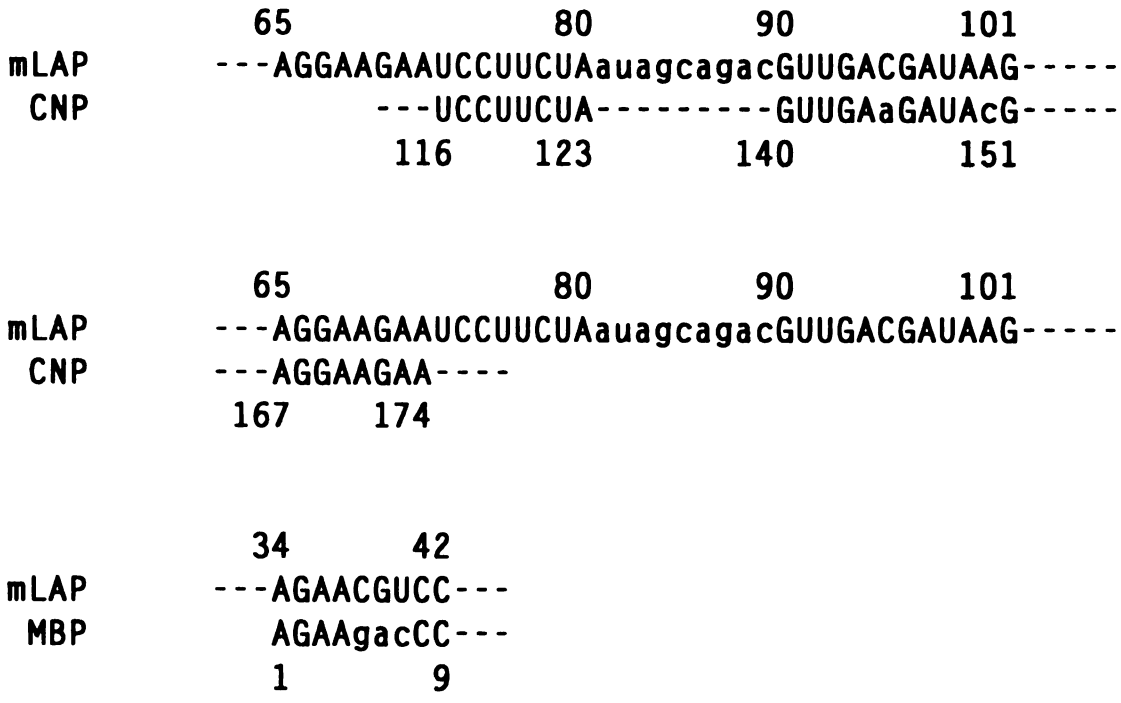



**Fig. 13. Putative steroid regulatory elements in the 5'-untranslated region of mLAP mRNA.**

The 5'-untranslated region of mLAP mRNA was compared to those of the mRNAs, which are translationally regulated by a steroid hormone: the rat 2', 3'-cyclic nucleotide 3'-phosphodiesterase (CNP) mRNA translation of which is negatively regulated by hydrocortisone (CNP cDNA sequence is taken from Bernier *et al.*, 1987); mouse myelin basic protein (MBP) mRNA is positively regulated by hydrocortisone (MBP cDNA sequence is taken from Newman *et al.*, 1987). Numbers indicate the position of the nucleotides from the beginning of each cDNA sequence. Capital letters denote identical nucleotides.

Two 8-nucleotide segments of mLAP mRNA, located in tandem at positions 65 to 80, were identical to the segments of CNP mRNA. The third 12-nucleotide segment of mLAP mRNA, located in positions 90-102, was only different in two nucleotides from the CNP element. There is also a 9-nucleotide segment of the 5'-region of mLAP mRNA which has a similarity with the steroid modulatory element of MBP (Fig. 13). Interestingly, the mLAP mRNA segment is different in three nucleotides, two of which (GA) were shown to be essential for the MBP element to function as a positive modulatory element (Verdi and Campagnon, 1990). Although final understanding of the functional significance of the putative steroid regulatory elements in the 5'-untranslated region of mLAP mRNA requires direct experimental proof, our finding provides strong evidence for involvement of 20-HE in the negative translational regulation of this mosquito lysosomal enzyme.

In conclusion, our analysis of mLAP cDNA has built a foundation for future studies of the regulation of the expression of mLAP at both transcriptional and translational levels. With the availability of mLAP cDNA, we can also address an important question about the mechanism of targeting of lysosomal enzymes in insects.

## Acknowledgments

We thank Dr. Henry Hagedorn for his kind gift of the gene probe of mosquito vitellogenin; and Drs. Will Kopachik and Suzanne Thiem for their critical reading of the manuscript.

# REFERENCES


Ammerer, G., Hunter, C. P., Rothman, J. H., Saari, G. C., Valls, L. A., and Steven, T. H. (1986) *Mol. Cell. Biol.*, **6**, 2490-2499.

Athauda, S. B. P., Matsuzaki, O., Kageyama, T., and Takahashi, K. (1990) *Biochem. Biophys. Res. Comm.*, **168**, 878-885.

Ausubel, F. M., Brent, R., Kingstone, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (1990) In: *Current Protocols in Molecular Biology* , Green Publishing Associates and Wiley-Interscience, Vol. 1, pp. 4.51-4.53, New Yolk

Azuma, T., Pals, G., Mohandas, T. K., Couvreur, J. M., and Taggar, R. T. (1989) *J. Biol. Chem.*, **264**, 16748-16753.

Baranski, T. J., Koelsch, G., Hartsuck, J. A., and Kornfeld, S. (1991) *J. Biol. Chem.*, **266**, 23365-23372.

Baranski, T. J., Faust, P. L., and Konfeld, S. (1990) *Cell*, **63**, 281-291.

Bednarek, S. Y., and Raikhel, N. V. (1992) *Plant Mol. Biol.*, (in press).

Bernier, L., Alvarez, F., Norgord, E. M., Raible, D. W., Mentaberry, A., Schembri,. J. G., Sabatini, D. D., and Colman, D. R. (1987) *J. Neurosci.*, **7**, 2703-2710.

Bose, S. G., and Raikhel, S. A. (1988) *Biochem. Biophys. Res. Commun.*, **155**, 436-442.

Brock, M. L., and Shapiro, D. J. (1983) *Cell*, **34**, 207-214.

Cardelli, J. A., and Dimond, R. L. (1988) In: *Protein Transfer and Organelle Biosynthesis* (Das, R. C., and Robbins, P. W., eds) pp. 363-399, Academic Press.

Cho. W. L., Dhadialla, T. S., and Raikhel, A. S. (1991a) *Insect Biochem.*, **21**, 165-176.

Cho, W. L., Deitsch, K. W., and Raikhel, A. S. (1991b) *Proc. Natl. Acad. Sci.*, **88**, 10821-10824.

Chrispeels, M. J., and Raikhel, N. V. (1992) *Cell*, **68**, 613-616.

Conner, G. E. (1989) *Biochem. J,*. **263**, 601-604.

Dean, R. L., Collins, J. V., and Locke, M. (1985) In: *Comprehensive Insect Physiology, Biochemistry and Pharmacology* (Kerkut, G. A., and Gilbert, L. I., eds) Vol. 3, pp. 155-210. Pergamon Press, Oxford.

Faust, P. L., and Kornfeld, S. (1989) *J. Biol. Chem.*, **264**, 479-488.

Faust, P. L., Kornfeld, S., and Chirgwin, J. M. (1985) *Pro. Natl. Acad. Sci.*, **82**, 4910-4914.

Glaumann, H., Ericsson, J. L. E., and Marzella, L. (1981) *Intern. Rev. Cytol.*, **73**, 149-182.

Gordon, D. A., Shelness, G. S., Nicosia, M., and Williams, D. L. (1988) *J. Biol. Chem.*, **263**, 2625-2631.

Grusby, M. J., Mitchell, S. C., and Glimcher, L. H. (1990) *Nucl. Acids Res.*, **13**, 4008.

Hagedorn, H. H. (1985) In: *Comprehensive Insect Physiology, Biochemistry, and Pharmacology* (Kerkut, G. A., and Gilbert, L. I., eds) Vol. 8, pp. 205-261, Pergamon Press, Oxford .

Hagedorn, H.H. (1989) In: *Ecdysone from Chemistry to Mode of Action* (Koolman J., ed.) Thieme Medical Publishers, Inc. New York, pp. 279-289.

Harris, T. J. R., Lowe, P. A., Lyons, A., Thomas, P. G., Eaton, M. A. W., Millican, T. A., Patel., T. P., Bose, C. C., Carey, N. H., and Doel, M. T. (1982) *Nucl. Acids Res.*, **10**, 2177-2187.

Hays, A. R., and Raikhel, A. S. (1990) *Roux's Arch. Dev. Biol.*, **199**, 114-121.

Horiuchi, H., Yanai, K., Okazaki, T., Takagi, M., and Yano, K. (1988) *J. Bacteriol.*, **170**, 272-278.

Ichimaru, E., Sakai, H., Saku, T., Kunimatsu, K., Kato, Y., Kato, I., and Yamamoto, K. (1990) *J. Biochem.*, **108**, 1009-1015.

Imai, T., Miyazaki, H., Hirose, S., Hori, H., Hayashi, T., Kageyama, R., Ohkubo, H., Nakanishi, S., and Murakami, K., (1983) *Proc. Natl. Acad. Sci.*, **80**, 7405-7409

Kornfeld, S. (1987) *FASEB J.*, **1**, 462-468.

Kornfeld, S., and Mellman, I. (1989) *Annu. Rev. Cell Biol.*, **5**, 483-525.

Kumar, S., Cole, R., Chiappelli, F., and De Vellis, J. (1989) *Proc. Natl. Acad. Sci.*, **86**, 6807-6811..

Kunkel, J. G., and Nordin, J. H. (1985) In: *Comprehensive Insect Physiology, Biochemistry, and Pharmacology* (Kerkut, G. A., and Gilbert, L. I., eds) Vol. 1, pp. 83-111, Pergamon Press, Oxford.

Kyte, J., and Doolittle, R. F. (1982) *J. Mol. Biol.*, **157**, 105-132.

Lea, A. O. (1964) *J. Med. Ent.*, **1**, 40-44.

Lipman, D. J.., and Pearson, W. R. (1985) *Science*, **227**, 1435-1441.

Locke, M. (1980) In: *Insect Biology in Future* (Locke, M., and Smith, D. S., eds) pp. 227-252, Academic Press.

Meyuhas, O., Thompson, JR, E. A., and Perry, R. (1987) *Mol. Cell. Biol.*, **7**, 2691-2699.

Nagao, E., Takahashi, N., and Chino, H. (1987) *Insect Biochem.*, **4**, 531-538.

Newman, S., Kitamura, K., and Campagnoni, A. T. (1987) *Proc. Natl. Acad. Sci.*, **84**, 886-890.

Nielsen, D. A., and Shapiro, D. J. (1990) *Mol. Cell. Biol.*, **10**, 371-376.

Nordin, J. H., Gochoco, C. H., Wojchowski, D. M., and Kunkel, J. G. (1984) *Comp. Biochem. Physiol.*, **79B**, 379-390.

Orci, L., Ravazzola, M., Amherdt, M., Yanaihara, C., Yanaihara, N., Halban, P., Renold, A. E., and Perrelet, A. (1984) *J. Cell. Biol.*, **98**, 222-228.

Osir, E. O., Anderson, D. R., Grimes, W. J., and Law, J. H. (1986) *Insect Biochem.*, **3**, 471-478.

Racioppi, J. V., Gemmill, R. M., Kogan, P. H., Calvo, J. M., and Hagedorn, H. H. (1986) *Insect Biochem.*, **16**, 255-262.

Raikhel, A. S., and Dhadialla, T. S. (1992) *Annu. Rev. Entomol.*, **37**, 217-251.

Raikhel, A. S. (1992) *Adv. Disease Vector Res.*, **9**, in press.

Raikhel, A. S. (1986a) *J. Insect Physiol.*, **32**, 597-604.

Raikhel, A. S. (1986b) *Tissue & Cell,* **18**, 125-142.

Raikhel, A. S., Pratt, L. H., Lea, A. O. (1986) *J. Insect Physiol.*, **32**, 879-890.

Runeberg-Roos, P., Tormakangas, K., and Ostman, A. (1991) *Eur. J. Biochem.*, **202**, 1021-1027.

Sambrook, J., Fritch, E. F., and Maniatis, T. (1989) In: *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor La., Cold Spring Harbor, NY) 2nd edition.

Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci.*, **74**, 5463-5467.

Shewale, J. G., and Tang, J. (1984) *Proc. Natl. Acad. Sci.*, **81**, 3703-3707.

Smith, R. E., and Farquhar, M. G. (1966) *J. Cell Biol.*, **31**, 319-345.

Sogawa, K., Fujii-Kuriyama, Y., MizuKami, Y., Ichihara, Y., and Takahashi, K. (1983) *J. Biol. Chem.*, **258**, 5306-5311.

Tang, J., and Wong, R. N. S. (1987) *J. Cell. Biochem.*, **33**, 53-63.

Ueno, E., Sakai, H., Kato, Y., and Yamamoto, K. (1989) *J. Biochem.*, **105**, 878-882.

Verdi, J. M., Kampf, K., and Campagnoni, A. T. (1989) *J. Neurochem.*, **52**, 321-324.

Verdi, J. M., and Campagnoni, A. T. (1990) *J. Biol. Chem.*, **265**, 20314-20320.

von Heijne, G. (1986) *Nucleic Acids Res.*, **14**, 4683-4690.

Wyatt, G. R. (1980) In: *Insect Biology in Future* (Locke, M., and Smith, D. S., eds) pp. 201-225, Academic Press.

Yonezawa, S., Takahashi, T., Wang, X. -J., Wong, R. N. S., Hartsuck, J. A., and Tang, J. (1988) *J. Biol. Chem.*, **263**, 16504-16511.

# CHAPTER 4*

# AN EXTRAOVARIAN PROTEIN ACCUMULATED IN MOSQUITO OOCYTES IS A CARBOXYPEPTIDASE ACTIVATED IN EMBRYOS

* Originally published in Proc. Natl. Acad. Sci.

Reference: Cho, W. -L., Deitsch, K. W., and Raikhel, A. S. (1991) Proc. Natl. Acad. Sci.,**88**,10821-10824.

## ABSTRACT

We report a novel phenomenon previously unknown for oviparous animals; in mosquitoes, a serine carboxypeptidase is synthesized extraovarially and then internalized by oocytes. The cDNA encoding mosquito vitellogenic carboxypeptidase (VCP) was cloned and sequenced. The VCP cDNA hybridizes to a 1.5 kb mRNA present only in the fat body of vitellogenic females. The deduced amino acid sequence of VCP shares significant homology with members of a serine carboxypeptidase family. Binding assays using a serine protease inhibitor, [$^{3}$H]-diisopropyl fluorophosphate (DFP), showed that VCP is activated in eggs at the onset of embryonic development. Activation of VCP is associated with the reduction in its size from 53-kDa (inactive pro-enzyme) to 48-kDa (active enzyme). The active 48-kDa form of VCP is maximally present at the middle and disappears by the end of embryonic development.

## INTRODUCTION

In oviparous animals, development of an embryo depends upon utilization of yolk proteins accumulated by oocytes during vitellogenesis. In most oviparous animals, nematodes, arthropods and vertebrates, the task of producing massive amounts of yolk protein precursors is performed by extraovarian tissues (1-5). The developing oocytes of these animals are highly specialized for the specific accumulation of extraovarian yolk protein precursors. The process of receptor-mediated endocytosis by which yolk protein precursors are accumulated in oocyte yolk bodies is well characterized for insects and vertebrates (6, 7). In contrast, the mechanisms of degradation of yolk proteins have received considerably less attention. Several enzymes that participate in hydrolysis of yolk proteins in insect embryos have been recently described (8-11). However, the origin of these enzymes and the time of their delivery into yolk bodies are unknown.

In this paper, we report a novel biological phenomenon previously unknown for insects or other oviparous animals. In vitellogenic female mosquitoes, an extraovarian tissue, the fat body, synthesizes not only the yolk protein precursor, vitellogenin, but also a pro-enzyme of a serine carboxypeptidase. We have named it "vitellogenic carboxypeptidase" (VCP). This enzyme is internalized by developing oocytes, accumulated in yolk bodies and activated during embryonic development.

## EXPERIMENTAL PROCEDURES

### *Animals*

Mosquitoes *Aedes aegypti* were reared as described by Hays and Raikhel (12). The embryonic stages of *A. aegypti* were determined according to Raminani and Cupp (13).

### *Cloning and sequencing of cDNA*

A cDNA clone of VCP was isolated by immunoscreening of a lambda gt11 library generated from fat bodies of vitellogenic female mosquitoes. This cDNA clone, pVCP-1, was used to isolate another cDNA clone, pVCP-2, encoding the 5'end 1020 bp of VCP from the same library. The former cDNA clone was sequenced with the dideoxy chain termination method in both directions (14).

The missing part of the 3' end sequence was obtained by the polymerase chain reaction using 20 nucleotides at the 3'end of pVCP-1 and poly-d(T)17 as primers including the XbaI digestion site at the 5' ends of both primers (15). Amplification was achieved by using cDNA reverse transcribed from 20 µg of total RNA prepared from whole bodies of mosquitoes 24 hr post blood meal as templates. The specific amplified band, identified by Northern blot hybridization to 1.5 kb mRNA, was subcloned into pUC119 for sequencing. Similar to the pVCP-2 cDNA clone, the 3'-end VCP cloned was sequenced in both directions.

### *Protein purification for partial sequencing*

Mosquito ovaries, 48 hr post blood meal, were homogenized in 20 mM Tris, 150 mM NaCl, pH 7.5, 0.02% sodium azide. The vitellogenic carboxypeptidase was isolated from the protein extract by anion-exchange,

affinity and hydrophobic chromatographic steps using DEAE Sepharose CL-4B, concanavalin A and Phenyl Sepharose columns, respectively. A fraction with nearly pure protein was resolved by SDS-PAGE and subjected to peptide sequencing using the method and facilities described by Cho *et al.* (16).

***Purification and analysis of RNA***

Isolation of total and polyadenylated RNA [poly(A)$^+$RNA] from mosquito tissues was performed as previously described (17).

*In vitro* translation of mosquito poly(A)$^+$RNA was performed as described by Bose and Raikhel (17) using rabbit reticulocyte lysate (Promega Biotech). The VCP precursor was identified in *in vitro* translation products of poly(A)$^+$RNA by immunoprecipitation with anti-VCP polyclonal antibodies (12).

The 840 bp insert from pVCP-1, a putative VCP cDNA clone, was used for hybrid selection of VCP mRNA from the fat body poly(A)$^+$RNA according to Sambrook *et al.* (15). The hybrid selected mRNA was translated *in vitro* using rabbit reticulocyte lysate and the protein product resolved by SDS-PAGE. The pUC119 plasmid without an insert was used for hybrid selection of mRNA as control.

For the Northern analysis, total RNA was separated by 1.2 % formaldehyde agarose gel electrophoresis, transferred to nitrocellulose paper and hybridized first to a [$^{32}$P]-labeled 840 bp VCP cDNA insert from pVCP-1 under high stringency conditions. RNA hybridized to the VCP cDNA probe was visualized by autoradiography of the blot. The probe was then stripped from the blot (14) and the blot rehybridized with a [$^{32}$P]-labeled 1.8 kb fragment of vitellogenin gene A1 (gift of Dr. H.H. Hagedorn) under high stringency conditions.

***Immunoblot Analysis***

Preparation of proteins, SDS-PAGE and immunoblotting were performed as described by Hays and Raikhel (12).

***Binding to Protein Inhibitor***

Protein extracts, prepared as for immunoblotting but without any protease inhibitors, were incubated with 0.1 μ Ci/μ l [$^3$H]-DFP (Dupont) for 1 hour at room temperature. These mixtures were then separated by SDS-PAGE under reducing conditions and processed for fluorography.

***Analysis of deduced amino acid sequence***

The analysis of deduced amino acid sequence was performed using FASTA program (GCG software) according to the algorithm by Lipman and Pearson (18). The deduced amino acid sequence of VCP was compared with amino acid sequences from the NBRF Protein Data Base (release 21.0) and with deduced amino acid sequences translated from GeneBank (release 66.0) and EMBL (release 19.0) Data Bases. The initial score was calculated from the best sub-sequence alignment of two sequences. The optimized score, calculated by considering insertions or deletions, gives a better indication about functional relation between proteins. The statistical significance of the scores was evaluated by Z value [(similarity score - mean of random scores)/(standard deviation of random scores); $Z > 10$ indicates statistical significance]. The similarities between VCP and other sequences were calculated as the percentage of amino acid number having identical residues or functional substitutions to the total number of amino acids in VCP (441).

## RESULTS AND DISCUSSION

Recently, a female-specific protein ($M_r$=53,000), initially called 53KP, has been found in the mosquito, *Aedes aegypti*. Like vitellogenin, this protein is synthesized by the fat body of vitellogenic females under the control of 20-hydroxyecdysone. The kinetics of the 53KP production by the vitellogenic fat body is also similar to that of vitellogenin: it is produced as early as 4 hr and reaches its peak near 24 hr after the initiation of vitellogenesis. Synthesis then drops to low levels by 36 hr and declines to background levels by 48 hr. This protein is secreted to the hemolymph and is selectively accumulated in yolk bodies of developing oocytes (12). Data presented here show that this protein (VCP) is a serine carboxypeptidase.

The cDNA encoding VCP was cloned by a combination of immunoscreening of a lambda gt11 cDNA library and the polymerase chain reaction. The identity of the cDNA was confirmed by direct sequencing of the N-terminus of the purified VCP and by in vitro translation of hybrid selected mRNA (data not shown).

Northern blot analysis demonstrated that transcription of the VCP mRNA is limited to only female fat bodies and is initiated after a blood meal. Similar to vitellogenin mRNA, the amount of VCP mRNA in the fat body is maximal at the peak of the protein production, 24 hr after a blood meal, then it declines to background levels by 48 hr after a blood meal (Figs. 2 and 3).

The sequence of full-length cDNA encoding VCP, confirmed by the sequencing of cDNA clones in both directions, is presented in Fig. 1. The size of VCP mRNA of 1.5 kb, estimated by Northern blot analysis, is in agreement with the 1,511 bp mRNA estimated from the cDNA sequence. The VCP cDNA has a single open reading frame that encodes a protein of 441 amino acids with a

**Fig. 1. Northern blot analysis of sex- and stage-specific expression of the VCP mRNA transcript.**

Total RNA was extracted from whole bodies of male mosquitoes (lane 1) and fat bodies of female mosquitoes before a blood meal (lane 2) or 24 and 48 hours post blood meal (lanes 3 and 4, respectively). RNA ladder (BRL) was resolved in lane M. A - agarose gel stained with ethidium bromide; B - hybridization with 0.84 kb fragment of VCP cDNA; C -hybridization with 1.8 kb fragment of mosquito vitellogenin A1 gene.

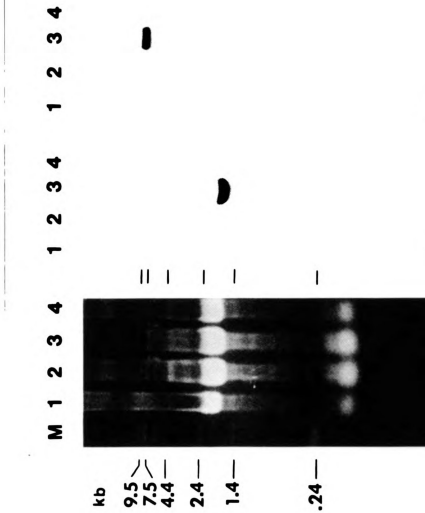

M 1 2 3 4
1 2 3 4
1 2 3 4
kb
9.5
7.5
4.4
2.4
1.4
.24
A
B
C

**Fig. 2. Northern blot analysis of tissue-specific expression of the VCP mRNA transcript.**

Total RNA was extracted from ovaries (lane 1), fat bodies (lane 2), and midgut (lane 3) of female mosquitoes 24 hours post blood meal. RNA ladder (BRL) was resolved in lane M. A -agarose gel stained with ethidium bromide; B - hybridization with 0.84 kb fragment of VCP cDNA.

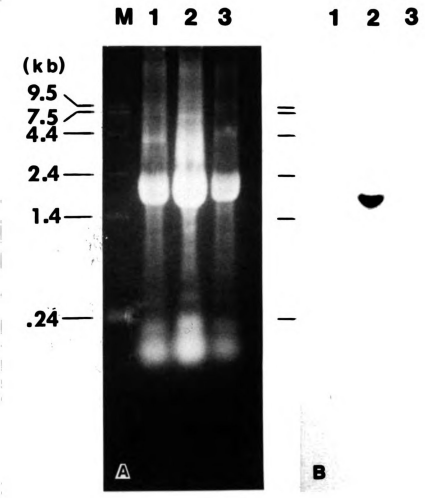

M 1 2 3
1 2 3
(kb)
9.5
7.5
4.4
2.4
1.4
.24
A
B

**Fig. 3. Nucleotide and deduced amino acid sequences of mosquito VCP.**

The amino acid sequence matching the N-terminus determined from purified VCP is underlined by a solid line. The signal peptide is boxed. The circled amino acid (N) is a potential glycosylation site. A putative polyadenylation signal, AATAAA, is shaded.

```
   1  TTCCCACTCGTTGTAAAGTGTAATCGAACAAGCATGGTGAAATTCCATTTACTAGTGCTG   60
                                        M  V  K  F  H  L  L  V  L     9

  61  ATTGCGTTCACCTGCTATACATGCAGCGACGCAACCTTATGGAATCCGTACAAGAAGCTA  120
  10   I  A  F  T  C  Y  T  C  S  D  A  T  L  W  N  P  Y  K  K  L    29

 121  ATGCGAGGATCGGCGTCTCCTCGTCGTCCAGGTGAAAGTGGTGAACCTTTGTTCCTGACT  180
  30   M  R  G  S  A  S  P  R  R  P  G  E  S  G  E  P  L  F  L  T    49

 181  CCACTGTTGCAGGATGGCAAAATTGAAGAGGCTCGCAACAAAGCCCGCGTCAACCATCCC  240
  50   P  L  L  Q  D  G  K  I  E  E  A  R  N  K  A  R  V  N  H  P    69

 241  ATGTTGAGCTCAGTGGAAAGCTACTCCGGTTTTATGACCGTTGATGCCAAGCACAACTCC  300
  70   M  L  S  S  V  E  S  Y  S  G  F  M  T  V  D  A  K  H  N  S    89

 301  AATTTGTTCTTCTGGTATGTTCCAGCGAAGAACAACCGCGAACAAGCGCCCATTCTTGTT  360
  90   N  L  F  F  W  Y  V  P  A  K  N  N  R  E  Q  A  P  I  L  V   109

 361  TGGCTGCAAGGAGGTCCAGGTGCGTCATCGCTGTTTGGAATGTTCGAAGAGAATGGACCG  420
 110   W  L  Q  G  G  P  G  A  S  S  L  F  G  M  F  E  E  N  G  P   129

 421  TTCCATATTCACAGGAACAACTCAGTGAAGCAACGTGAATATTCCTGGCATCAGAACCAT  480
 130   F  H  I  H  R  N  N  S  V  K  Q  R  E  Y  S  W  H  Q  N  H   149

 481  CACATGATCTACATCGATAATCCAGTTGGAACGGGATTCAGTTTCACCGATAGCGATGAA  540
 150   H  M  I  Y  I  D  N  P  V  G  T  G  F  S  F  T  D  S  D  E   169

 541  GGTTACTCCACCAACGAAGAACACGTGGGTGAAAATCTGATGAAATTCATTCAACAATTC  600
 170   G  Y  S  T  N  E  E  H  V  G  E  N  L  M  K  F  I  Q  Q  F   189

 601  TTCGTGCTGTTCCCCAATCTGTTGAAGCATCCATTCTACATCTCCGGTGAATCTTATGGT  660
 190   F  V  L  F  P  N  L  L  K  H  P  F  Y  I  S  G  E  S  Y  G   209

 661  GGAAAGTTTGTTCCTGCCTTCGGTTATGCCATCCACAACTCCCAAAGCCAACCTAAAATC  720
 210   G  K  F  V  P  A  F  G  Y  A  I  H  N  S  Q  S  Q  P  K  I   229

 721  AATCTGCAAGGATTGGCCATTGGTGATGGCTACACTGATCCGCTGAACCAACTTAACTAC  780
 230   N  L  Q  G  L  A  I  G  D  G  Y  T  D  P  L  N  Q  L  N  Y   249

 781  GGAGAATACCTGTATGAGCTGGGCCTGATCGATTTGAACGGAAGAAAGAAGTTCGACGAG  840
 250   G  E  Y  L  Y  E  L  G  L  I  D  L  N  G  R  K  K  F  D  E   269

 841  GATACGGCTGCTGCCATCGCCTGTGCCGAACGTAAGGACATGAAGTGCGCCAACCGCCTT  900
 270   D  T  A  A  A  I  A  C  A  E  R  K  D  M  K  C  A  N  R  L   289

 901  ATCCAAGGTCTGTTCGATGGACTCGATGGACAGGAATCGTACTTCAAGAAGGTCACCGGA  960
 290   I  Q  G  L  F  D  G  L  D  G  Q  E  S  Y  F  K  K  V  T  G   309

 961  TTCTCGTCCTACTACAACTTCATCAAGGGAGACGAGGAGAGCAAACAGGATAGTGTCCTG 1020
 310   F  S  S  Y  Y  N  F  I  K  G  D  E  E  S  K  Q  D  S  V  L   329

1021  ATGGAGTTCCTCAGCAACCCGGAGGTACGTAAGGGCATCCACGTTGGTGAACTGCCGTTC 1080
 330   M  E  F  L  S  N  P  E  V  R  K  G  I  H  V  G  E  L  P  F   349

1081  CACGACTCTGACGGTCACAACAAGGTCGCGGAAATGCTGTCCGAAGACACTCTGGACACC 1140
 350   H  D  S  D  G  H  N  K  V  A  E  M  L  S  E  D  T  L  D  T   369

1141  GTGGCTCCATGGGTCAGCAAGCTGCTCTCGCACTACCGCGTGCTGTTCTACAACGGTCAG 1200
 370   V  A  P  W  V  S  K  L  L  S  H  Y  R  V  L  F  Y  N  G  Q   389

1201  TTGGACATCATCTGCGCCTACCCGATGACGGTCGACTTCCTGATGAAGATGCCTTTCGAT 1260
 390   L  D  I  I  C  A  Y  P  M  T  V  D  F  L  M  K  M  P  F  D   409

1261  GGCGATAGCGAGTACAAGCGGGCCAATCGTGAGATCTACCGCGTGGATCGGAAATCGCCG 1320
 410   G  D  S  E  Y  K  R  A  N  R  E  I  Y  R  V  D  R  K  S  P   429

1321  GGTACAAGAAGCGGGCTGGTCGTCTGCAAGAGGTGCTGATCAGAAACGCCGGACACATGG 1380
 430   G  T  R  S  G  L  V  V  C  K  R  C

1381  TTCCGCGGGACCAGCCGAAATGGGCTTTCGACATGATCACCTCGTTCACTCACAAAAACT 1440

1441  ACTTGTGAGGGAAATGGTAATGATTTGGATGAATAAAGCTTTAAGCTGTAATTAAAAAAA 1500

1501  AAAAAAAAAAA
```

deduced molecular weight of 50,153. Cell-free translation of VCP mRNA in vitro revealed that the VCP precursor is a 50-kDa polypeptide (data not shown). Hydropathy analysis (19, 20) of the deduced amino acid sequence has shown that VCP exhibits properties typical of a secretory protein. The amino acid sequence of VCP has only one potential glycosylation site at position $Asn_{135}$. Glycosylation of VCP via high mannose, which accounts for about 2 kDa of the protein molecular mass, was demonstrated previously (12).

In other insects, small extraovarian proteins are also involved in formation of yolk protein reserves. Microvitellogenin (31 kDa) of *Manduca sexta* and *Hyalophora cecropia*, and the 30-kDa protein of *Bombyx mori* are produced by the fat body and deposited in oocytes (21-23). The nucleotide sequence of *Manduca* microvitellogenin cDNA has 70% similarity to the cDNA sequence coding for the *Bombyx* 30-kDa protein, indicating a close evolutionary relationship between these proteins (24-26).

Comparison of amino acid sequences has not revealed any similarity between mosquito VCP and either *Manduca* microvitellogenin or *Bombyx* 30-kDa protein. Unexpectedly, the VCP sequence exhibited significant homology with members of a family of serine carboxypeptidases (Table 1). The homology between the amino acid sequences of mosquito VCP and these carboxypeptidases is the highest at the N-terminal portion which includes two conserved domains (27-32). The highest identity (28.1%) was found in a stretch of 392 amino acids of VCP and wheat carboxypeptidase Y homolog (WCP) (Fig. 4). However, considering both identical and conservative replacements, similarity between VCP and WCP is 62.3%. In mosquito VCP the conserved domains are at positions $Trp_{110}$ - $Ser_{119}$ and $Gly_{205}$ - $Gly_{210}$ (Fig. 4).

**Table 1. Homology of mosquito VCP to serine carboxypeptidase**

| Source | Protein | Score | Z Value | Similarity % |
|---|---|---|---|---|
| Wheat | Carboxypeptidase Y Homolog | 523 | 81.1 | 62.3 |
| Mouse | Protective Protein | 434 | 54.8 | 45.6 |
| Human | Protective Protein | 422 | 59.2 | 48.2 |
| Barley | Carboxypeptidase I | 394 | 42.9 | 56.7 |
| Yeast | Carboxypeptidase Y | 383 | 43.2 | 31.4 |
| Barley | Serine Carboxypeptidase II | 380 | 49.1 | 35.4 |
| Yeast | KEX1 Carboxypeptidase | 339 | 40.0 | 42.8 |

Only optimized score are presented in this table; Z value is a statistical significance of optimized scores (Z > 10 indicates statistical significance)

**Fig. 4. Alignment of amino acid sequences of mosquito VCP and wheat carboxypeptidase Y homolog (WCP).**

Amino acid sequences deduced from corresponding cDNAs were aligned by FASTA computer program. Vertical lines indicate identical residues and colons denote functional substitutions. Only portions of both amino acid sequences, flanking the overlapping regions, are presented. Two conserved domains found in sequences of serine carboxypeptidases are shaded. The serine protease catalytic center is marked by an asterisk.

```
VCP LWNPYKKLMRGSASPRRPGESGEPLFLTPLLQDGKIEEARNKARVNHPMLSSVESYSGFM 81
                                        ::  |::  |  |  :  :   :  :::::  ::|:
WCP FPGAQAERLIRALNLLPGRPRRGLGAGAEDVAPGQLLERRVTLPGLPEGVGDLGHHAGYY 91

VCP TVDAKHNSNLFFWYVPAKNNREQAPILVWLQGGPGASSLFGMFEENGPFHIHRNNSVKQR 141
     :  ::|::::|:::  :::::|:  |:::||  ||||  ||  :::|  |||||  |  :|  |:   :
WCP RLPNTHDARMFYFFFESRGKKED-PVVIWLTGGPGCSSELAVFYENGPFTIANNMSLVWN 150

VCP EYSWHQNHHMIYIDNPVGTGFSFTDSDEGYSTNEEHVGENLMKFIQQFFVLFPNLLKHPF 201
    :::|::  ::|::|  ::|||||::::|  :  :  :|:  |:::|  :|:|  ||    |:::|:  |
WCP KFGWDKISNIIFVDPATGTGFSYSSDDRDTRHDEAGVSNDLYDFLQVFFKKHPEFVKNDF 210

          *
VCP YISGESYGGKFVPAFGYAIHNSQSQ---PKINLQGLAIGDGYTDPLNQLN-YGEYLYELG 257
    :|:|||||:|:::|||:  :|:::::   ::|||:|:|||:|  |||   |  :  |::|   :::
WCP FITGESYAGHYIPAFASRVHQGNKKNEGTHINLKGFAIGNGLTDPAIQYKAYTDYALDMN 270

VCP LI---DLNGRKKFDEDTAAAIA-CAERKDMKC--ANRLIQGLFDGLDGQESYFKKVTGFS 311
    ||   | :  :||    : ||  |::  : :|  | :: :::|:::       :|::| :
WCP LIQKADYDRINKFIPPCEFAIKLCGTDGKASCMAAYMVCNSIFNSI-------MKLVGTK 323

VCP SYYNFIKGDEE--SKQDSVLMEFLSNPEVRKGIHVGELPFHDSDGHNKVAEMLSEDTLDT 369
    :||:  |: |:  : : | | :|::: :||::| ||:: | :::  ::| : : :| : :
WCP NYYDVRKECEGKLCYDFSNLEKFFGDKAVRQAIGVGDIEFVSCS--TSVYQAMLTDWMRN 381

VCP VAPWVSKLLSH-YRVLFYNGQLDIICAYPMTVDFLMKMPFDGDSEYKRANREIYRVDRKS 428
    ::  :: ||::  :||:|:|: |:||::  :  :: :| ::|:::: ::: : : ||
WCP LEVGIPALLEDGINVLIYAGEYDLICNWLGNSRWVHSMEWSGQKDFAKTAESSFLVDDAQ 441

VCP PGTRSGLVVCKRC                                                441
    :|
WCP AGVLKSHGALSFLKVHNAGHMVPMDQPKAALEMLRRFTQGKLKESVPEEEPATTFYAA   499
```

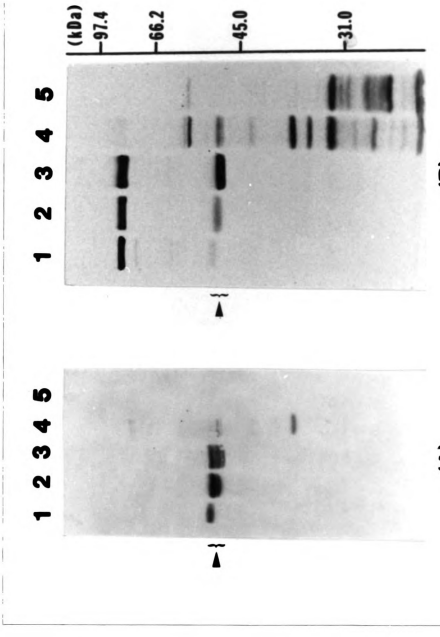
The $Ser_{207}$ of the VCP second domain corresponds to $Ser_{257}$ of yeast carboxypeptidase Y which was shown to be the catalytic center of serine carboxypeptidases (33). In contrast to mosquito VCP and WCP (Fig. 4), all other members of the serine carboxypeptidase family have three conserved domains (27-32). Functional implications of these differences are not clear. Unlike serine endopeptidases, the importance of three domains for carboxypeptidase activity has not been established (34). Similar to all other serine carboxypeptidases in the family (27-32), the amino acid sequence of mosquito VCP does not have any significant homology with any known sequences of trypsin-like serine endopeptidases or esterases.

Internalization of VCP in mosquito oocytes occurs without any changes in its size (12). The molecular mass of VCP, however, decreases by 0.5-1.0 kDa at the onset of embryonic development (Fig. 5). As embryogenesis progresses, VCP undergoes further reduction in its size and another, immunologically related 48-kDa band appears (Fig. 5). By the end of embryonic development, VCP degrades into smaller peptides which disappear once the first instar larva hatches (Fig. 5).

Radiolabeled serine protease inhibitor, [$^3$H]DFP, which binds to the active center of serine proteases (35), binds weakly to VCP in oocytes (Fig. 5; lanes 1). The intensity of binding to VCP, however, increases at the onset and reaches maximum in the middle of embryogenesis, when the inhibitor binds to a VCP band of 48 kDa (Fig. 5). The binding of DFP to VCP can be inhibited by prior treatment of protein extracts with other serine protease inhibitors, phenylmethylsulfonyl fluoride or leupeptin (date not shown). Based on above results, we conclude that VCP is synthesized by the fat body and internalized by oocytes as a pro-enzyme which is then activated in eggs at the onset of embryogenesis.

**Fig. 5. Immunoblot (A) and [$^3$H]-DFP binding (B) analyses of VCP during mosquito embryonic development.**

The lanes in both panels contain protein extracts from: 1, mosquito ovaries removed 24 hr post blood meal during peak yolk accumulation; 2, 0-3 hr post oviposition eggs at the onset of embryogenesis; 3, 44-47 hr post oviposition eggs during mid embryogenesis; 4, 94-97 hr post oviposition eggs at the end of embryogenesis; 5, first instar larva. In both figures, the mosquito VCP bands are shown by arrows. The high molecular weight polypeptide bound to DFP (Fig. 4B) is not immunologically related to VCP. In both panels, proteins were resolved by SDS-PAGE on 9% gels under reducing conditions. The molecular weight standards in order of decreasing Mr are phosphorylase B, bovine serum albumin, ovalbumin and carbonic anhydrase (Bio-Rad).


(kDa)
97.4
66.2
45.0
31.0
1 2 3 4 5
(B)
1 2 3 4 5
(A)

The activation of VCP is associated with the increase in its electrophoretic mobility. It is not clear, however, whether this increase in VCP mobility is due to its proteolytic cleavage or deglycosylation. The mouse carboxypeptidase (protective protein) is activated as a result of cleavage and DFP binds only to the activated carboxypeptidase subunit which contains the serine catalytic center (31, 32).

Some of the serine carboxypeptidases with which VCP shares significant homology, such as human and mouse protective proteins and yeast KEX1 carboxypeptidase, are implicated in proteolytic activation of a number of enzymes or other biologically active molecules (29-32). The mosquito VCP could play a similar role by activating hydrolytic enzymes which are involved in degradation of yolk proteins in developing embryos. Alternatively, it could function as an exopeptidase in sequential degradation of vitellogenin.

The mosquito VCP is different from proteases known to hydrolyze yolk proteins in insect embryos: an acidic cathepsin B-like protease of *Drosophila*, a thiol protease or cathepsin L-like protease of *Bombyx* (8-11). Similar to mosquito VCP, all these proteases are deposited as pro-enzymes in yolk bodies of eggs and activated in embryos. *Drosophila* cathepsin B-like protease is activated as a result of proteolysis by a serine protease ($M_r$=25,000) which is also present in egg yolk bodies (9). It is not known, however, whether any of these enzymes are of extraovarian origin or whether they are synthesized by developing oocytes themselves. The mosquito VCP, therefore, is the first example among oviparous animals of a proteolytic enzyme produced by an extraovarian tissue and accumulated by oocytes for use in embryonic development.

## ACKNOWLEDGEMENTS

We thank Dr. T. S. Dhadialla and Mr. A. R. Hays for their advice in protein purification, Dr. H. H. Hagedorn for his generous gift of vitellogenin gene probe and Drs. T. S. Dhadialla, T. B. Friedman, J. R. Miller and N. V. Raikhel for their critical reading of the manuscript. This work was supported by grants from the National Institutes of Health and from the Biotechnology Center of the Michigan State University to ASR.

## REFERENCES


1. Kunkel, J.G. & Nordin, J.H. (1985) In: *Comprehensive Insect Physiology, Biochemistry and Pharmacology,* eds. Kerkut, G.A. & Gilbert, L.I. (Plenum, New York), Vol. 1, pp. 83-111.

2. Wyatt, G.R. (1988) *Can. J. Zool.,* **66,** 2600-2610.

3. Dhadialla, T.S. & Raikhel, A.S. (1990) *J. Biol. Chem.,* **265,** 9924-9933.

4. Kanost, M.R., Kawooya, J.K. Law, J.H., Ryan, R.O., Van Heusden, M.C., & Ziegler, R. (1990) *Adv. Insect Physiol.,* **22,** 299-396.

5. Wahli, W. (1988) *TIG* **4,** 227-232.

6. Ferenz, H.-J. (1990) In: *Molecular Insect Science,* eds. Hagedorn, H.H., Hildebrand, J.G., Kidwell, M.G. & Law, J.H. (Plenum Press, New York and London), pp. 131-138.

7. Raikhel, A.S. & Dhadialla, T.S. (1992) *Annu. Rev. Entomol.,* **37,** 217-251.

8. Medina, M., Leon, P. & Vallejo, C.G. (1988) *Arch. Biochem. Biophys.,* **263,** 355-363.

9. Medina, M. & Vallejo, C.G. (1989) *Insect Biochem.,* **19,** 687-691.

10. Indrasith, L.S., Sasaki, T. & Yamashita, O. (1988) *J. Biol. Chem.,* **263,** 1045-1051.

11. Kageyama, T. and Takahashi, S. (1990) *Eur. J. Biochem.,* **193,** 203-210.

12. Hays, A.R. & Raikhel, A.S. (1990) *Roux's Arch. Dev. Biol.,* **199,** 114-121.

13. Raminani, L.N. & Cupp, E.W. (1978) *Int. J. Insect Morphol. Embryol.,* **7,** 273-296.

14. Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989) In: *Molecular Cloning: A Laboratory Manual,* Second Edition (Cold Spring Harbor Lab., Cold Spring Harbor, NY).

15. Lee, C.C., Wu, X., Gibbs, R.A., Cook, R.G., Muzny, D.M. & Caskey, C.T. (1988) *Science,* **239,** 1288-1291.

16. Cho, W.-L., Dhadialla, T.S. & Raikhel, A.S. (1991) *Insect Biochem.,* **21,** 165-

176.

17.Bose, S.G. & Raikhel, A.S. (1988) *Biochem. Biophys. Res. Commun.*, **155**, 436-442.

18.Lipman, D.J. & Pearson, W.R. (1985) *Science*, **227**, 1435-1441.

19.Kyte, J. & Doolittle, R.F. (1982) *J. Mol. Biol.*, **157**, 105-132.

20.von Heijne, G. (1983) *Eur. J. Biochem.*, **133**, 17-21.

21.Kawooya, J.K., Osir, E.O. & Law, J.H. (1986) *J. Biol. Chem.*, **261**, 10844-10849.

22.Zhu, J., Indrasith, L.S. & Yamashita, O. (1986) *Biochim. Biophys. Acta.*, **882**, 427-436.

23.Kulakosky, P.C. & Telfer, W.H. (1987) *Insect Biochem.*, **17**, 845-858.

24.Sakai, N., Mori, S., Izumi, S., Haino-Fukishima, K., Ogura, T., Maekawa, H. & Tomino, S. (1988) *Biochim. Biophys. Acta*, **949**, 224-232.

25.Wang, X.-Y., Cole, K.D. & Law, J.H. (1988) *J. Biol. Chem.*, **263**, 8851-8855.

26.Wang, X.-Y., Cole, K.D. & Law, J.H. (1989) *Gene*, **80**, 259-268.

27.Baulcombe, D.C., Backer, R.F. & Jarvis, M.G. (1987) *J. Biol. Chem.*, **262**, 13726-13735.

28.Sorensen, S.B., Svendsen, I.B. & Breddam, K. (1987) *Carlsberg Res. Commun.*, **52**, 285-295.

29.Dmochowska, A., Dignard, D., Henning, D., Thomas, D.Y. & Bussey, H. (1987) *Cell*, **50**, 573-584.

30.Valls, L.A., Hunter, C.P., Rothman, J.H. & Stevens, T.H. (1987) *Cell*, **48**, 887-887.

31.Galijart, N.J., Gillemans, N., Harris, A., van der Horst, G.T.J., Verheijen, F.W., Galjaard, H. & d'Azzo, A. (1988) *Cell*, **54**, 755-764.

32.Galijart, N.J., Gillemans, N., Meijer, D. & d'Azzo, A. (1990) *J. Biol. Chem.*, **265**, 4678-4684.

33.Hayashi, R., Moore, S. & Stein, W.H. (1973) *J. Biol. Chem.*, **248**, 8366-8369.

34.Wilson, K.P., Liao, D.-I., Bullock, T., Remington, S.J. & Breddam, K. (1990) *J. Mol. Biol.*, **211**, 301-303.

35.Scheiner, C.J. & Quigley, J.P. (1982) *Anal. Biochem.*, **122**, 58-69.

# CHAPTER 5

# SUMMARY AND FUTURE RESEARCH PROSPECTS

***1. Mosquito lysosomal aspartic protease:***

The termination stage of mosquito vitellogenesis involves the cessation of Vg gene expression and the dramatic increase of lysosomal activity. This lysosomal activity may interrupt Vg secretion by degrading the Vg-containing granules and cause trophocyte remodeling by autophagocytosis of biosynthetic machinery (Raikhel, 1986a, 1986b). In my study, one of the lysosomal enzymes, aspartic protease, was purified. The protein characterization and cDNA sequence analysis of mLAP reveal that mLAP shares similarity to both cathepsins D and E as listed in Table-1. The result implies that mLAP may be similar to the ancestral protein of vertebrate cathepsins D and E. During evolution, this protein lost a phosphorylation determinant resulting in a different subcellular distribution than lysosomal enzymes. In contrast, an insertion was introduced causing an additional processing step in vertebrate cathepsin D.

During the termination stage of vitellogenesis, mLAP levels rise 10 fold over the levels of the synthetic stage. Based on the profiles of mLAP mRNA (Fig.3 of Chapter 3) and 20-HE concentration (Fig.2 of Chapter 1), we postulated that 20-HE may regulate mLAP gene expression. This hypothesis will be investigated at the transcriptional level by using *in vivo* and *in vitro* systems. Microsurgical removal and the culture of mosquito fat body will be used to analyze the effect of 20-HE on mLAP transcription. Moreover, when the upstream region of mLAP gene is cloned, the prospective promoter and enhancer regions will be cloned in pCAT-Enhancer vectors and pCAT-Promoter vectors for functional analysis of regulatory elements in eukaryotic cell lines. (Rosenthal, 1987). The position and the size of the cis- regulatory elements of the mLAP gene will be localized and determined with combinative oligonucleotide-directed deletions (Wallrath and Friedman, 1992). To identify putative tissue specific regulatory elements, these sequence will be inserted into

**Table 1. Comparison among mLAP, HCD and HCE.**

| | | mLAP | HCD | HCE |
|---|---|---|---|---|
| 1 | Native $M_r$ | 80 kDa | 47 kDa | 76 kDa |
| 2 | $M_r$ of subunits | 40 kDa (Homodimer) | 14 kDa & 31 kDa | 38 kDa (Dimer) |
| 3 | Optimal reaction pH | 3.0 | 3.5 | 3.0-3.2 |
| 4 | Insertional cleavage sites | No | Yes | No |
| 5. | N-linked Glycosylation | 1 | 2 | 1 |
| 6 | Iso-enzymes | No | Yes (3) | Yes (2) |
| 7 | pI | 5.4 | 5.7; 6; 6.6 | 4.1-4.6 |
| 8 | Inhibition by 10 mM KCN | No | No | Yes |
| 9 | Inhibition by 6 M urea | Yes | Yes | No |
| 10 | Localization | Lysosome | Lysosome | Cytosol or membrane bound |
| 11 | Distribution | Ubiquitous | Ubiquitous | Tissue specific |
| 12 | Amino acid similarity to mLAP | | 92% | 81% |

P-elements containing a reporter gene, lac Z, for *D. melanogaster* germline transformation (Wallrath *et al.*, 1990). The expression can be analyzed by the color response caused by the β-galactosidase activity in dissected tissue from transformed flies. If the regulatory elements linked to the upstream region of the lac Z gene are functional in the specific tissue, the tissue will turn blue when it is stained with X-gal .

The increased translation of mLAP mRNA is postponed for 12 hr after mRNA levels increase. We have found putative steroid hormone regulation elements in the 5'-untranslated region of mLAP mRNA. This finding suggests that 20-HE may inhibit the translation of mLAP mRNA between 12 hr and 24 hr PBM. This regulation of mRNA translation by a steroid hormone is different from the receptor-mediated regulation of gene expression. It demonstrates that a steroid hormone can regulate the transcription in the nucleus but may also be involved with the translation in the cytosol (Verdi and Campagnoni, 1990). In our case, 20-HE may interact with the 5' untranslated region of mLAP mRNA to inhibit or decrease the translation. To study this mechanism, the approach used by Verdi and Campagnoni (1990) will be adopted. The mosquito *in vitro* translational system will be prepared from a mosquito cell line (Gillies and Stollar, 1981). The cDNA corresponding to the 5' untranslated region of mLAP mRNA will be ligated to the upstream coding region of CAT, and inserted into a pBluescript phagemid. With this construct, the fusion RNA can be generated in an *in vitro* system with T7 RNA polymerase. To increase the stability of mRNA and enhance the efficiency of translation, mCAP$^{TM}$ RNA Capping kit from Stratagene will be used. The translational regulation of 20-HE on the 5' untranslated region of mLAP mRNA will then be analyzed using a mosquito *in vitro* translation system by measuring CAT activity without the endogenous interference.

In vertebrates, the targeting of lysosomal enzymes from their site of synthesis in RER to lysosomes is mainly directed by mannose-6-phosphate receptors (cation dependent and cation independent receptors) which recognize the mannose-6-phosphate marker exposed on lysosomal enzymes (Kornfeld, 1987). Lysosomal enzymes contain common protein determinants that are recognized by lysosomal enzyme N-acetylglucosamine-1-phosphotransferase and UDP-GlcNAc (UDP-linked N-acetylglucosamine) in the formation of mannose-6-phosphate residues (Baranski *et al.*, 1990). Two determinants for mannose-6 phosphorylation similar to those of human cathepsin D are found in the mLAP amino acid sequence. The cDNA for an insect lysosomal enzyme which is well characterized at the protein level, can be used to investigate this sorting mechanism in insect cells. The first step is to analyze whether mLAP is phosphorylated at its glycan. Next we need to understand whether the N-linked phosphorylated glycan is essential for transport of mLAP to lysosomes. To do this, polyclonal antibodies which can recognize the native form of mLAP are required. With the purification protocol that I developed for mLAP, it is difficult to isolate enough mLAP for antibody production. The coding sequence of mLAP cDNA will be cloned into a baculovirus expression vector under the control of the polyhedrin promoter. The overexpressed protein therefore can be purified easily for polyclonal antibody production. Before a serious of analysis, a glycosylation inhibitor, tunicamycin, will be used to examine whether it can interrupt the transport of mLAP. Furthermore, to verify that mLAP is phosphorylated at the oligosaccharide chain, mLAP will be labeled with [$^{32}$P] orthophosphoric acid in fat bodies cultured *in vitro* and then mLAP will be immunoprecipitated. If mLAP is phosphorylated at the glycan only, the treatment of endoglycosidase H will remove the radioactivity from mLAP. The result can be simply examined with SDS-PAGE. If the peptide back bone of mLAP is phosphorylated too, then high

performance liquid chromatography (HPLC) has to be used for detecting the [32P] labeled oligosaccharide. If mLAP is proved to be phosphorylated at the oligosaccharide, site-directed mutagenesis will be used to change the potential N-linked glycosylation residue (Asn-70). The mutated and wild type cDNAs will be constructed into plasmids containing an inducible *Drosophila* hsp70 promoter respectively. The DNA prepared from these constructs will be used to transfect a mosquito cell line, *Ae. albopictus*, for transient expression (Durbin and Fallon, 1985). The difference of mLAP activity and mLAP protein detected in culture medium between wild type and mutant DNA transfections will elucidate the importance of phosphorylated glycan in insect lysosomal targeting. Alternatively, the mutated cDNA and wild type mLAP cDNAs can also be subcloned into a baculoviral vector under the control of early gene promoter, IE-1, or even the polyhedrin promoter for expression in an insect cell line (Steiner *et al, 1988*; Hammock *et al.*, 1990 ). The amount of secreted mLAP proteins and the mLAP activity measured in culture medium will allow us to understand the role of phosphorylated mannose(s) in targeting of insect lysosomal enzymes. More investigations of insect lysosomal enzyme trafficking can be achieved by methods as described in Baranski *et al* (1990 and 1991).

***2. Mosquito vitellogenic carboxypeptidase:***

Cloning and sequencing of an additional vitellogenic protein, VCP, have confirmed that it is synthesized extraovarially in a sex-, tissue- and stage-specific manner. Even more importantly , analysis of its cDNA sequence has lead to a discovery of a novel biological phenomenon previously unknown for any oviparous animals. It is the first observation that an enzyme, participating in embryonic development, is synthesized outside of the developing oocytes.

In the mosquito, VCP is synthesized as a proenzyme and processed to an

active enzyme during embryonic development. To analyze this activation mechanism, it is important to purify the active form of the enzyme from embryos and study its enzymatic properties. Furthermore, the subcellular localization of VCP in embryos and the substrate determination for VCP are also essential for the understanding of its physiological function in embryonic development.

The synthesis of VCP in fat bodies is controlled by 20-HE (Hays and Raikhel, 1990). Further analysis of the regulatory mechanism depends upon cloning of the corresponding gene(s). VCP is a tissue- sex- and stage- specific expressed gene(s). Therefore, the regulatory elements of VCP gene(s) will be constructed in plasmids containing CAT or luciferase genes. DNA prepared from these constructs will be subjected to *in vitro* transcription with nuclear extract prepared from mosquito fat bodies 24 hr PBM with a method modified from *Drosophila* extract (Heiermann and Pongs, 1985; Kamakaka *et al.*, 1991). The transcripts will be used for *in vitro* translation with rabbit reticulocyte lysate. The functional regulatory elements will be detected from enzymatic activities of CAT or luciferase. In addition, those elements inserted into P-elements containing a reporter gene, lac Z, for *D. melanogaster* embryo transformation as described in mLAP project will be also an alternative tool for us to examine the sex-, tissue- and stage- specific regulations of VCP gene(s) in *Drosophila* (Wallrath *et al.*, 1990).

The cornerstone of mosquito reproduction is vitellogenesis. Mosquito LAP and VCP are essential proteases involved in vitellogenesis. Therefore, to fully understand the regulation of these two proteins at the molecular level will be a potential tool for the development of strategies for the interruption of vitellogenesis or egg development in the mosquito population.

# REFERENCES


Baranski, T. J., Faust, P. L., and Kornfeld, S. (1990) *Cell,* **63**, 281-291.

Baranski, T. J., Koelsch, G., Hartsuck, J. A., and Kornfeld, S. (1991) *J. Biol. Chem.*, **266**, 23365-23372.

Durbin, J. E., and Fallon, A. M. (1985) *Gene* **36**, 173-178.

Gillies, S., and Stoller, V. (1981) *J. Biol. Chem.*, **256**, 13188-13192

Hays, A. R., and Raikhel, A. S. (1990) *Roux's Arch. Dev. Biol.,* **199**, 114-121.

Hammock, B. D., Bonning, B. C., Possee, R. D., Hanzlik, T. N., and Maeda, S. (1990) *Nature* **344**, 458-461

Heiermann, R., and Pongs, O. (1985) *Nucl. Acids Res.* **13**, 2709-2730.

Kamakaka, R. T., Tyree, C. M., and Kadonaga, J. T.(1991) *Proc. Natl. Acad. Sci.* **88**, 1024-1028.

Kornfeld, S. (1987) *FASEB J.* **1**, 461-468.

Raikhel, A. S. (1986a) *J. Insect Physiol.* **32**, 597-604.

Raikhel, A. S. (1986b) *Tissue & Cell*, **18**, 125-142.

Rosenthal, N. (1987) Meth. Enzymol. **152**, 714-720.

Steiner, H., Pohl, G., Gunne, H., Hellers, M., Elhammer, A., and Hansson, L. (1988) Gene **73**, 449-457.

Verdi, J. M. , and Campagnoni, A. T. (1990) *J. Biol. Chem.*, **265**, 20314-20320.

Wallrath, L.L., Burnett, J. B., and Friedman, T. B. (1990) Mol. Cell. Biol. **10**, 5114-5127.

Wallrath, L.L., and Friedman, T. B. (1992) Biotech. **12**, 214-216.