

MOLECULAR CHARACTERIZATION OF DIGESTIVE PROTEASES OF THE  
YELLOW MEALWORM, *TENEBRIO MOLITOR* L.

by

SHEILA PRABHAKAR

B.Sc., Osmania University College for Women, 1999  
M.Sc., Osmania University, 2001

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department Of Entomology  
College Of Agriculture

KANSAS STATE UNIVERSITY  
Manhattan, Kansas

2006

## Abstract

Coleopteran insects compensate for dietary protease inhibitors by a number of mechanisms. To study this compensation response at the molecular level, the digestive proteases of *Tenebrio molitor* were studied. Biochemical studies of the pH optima and inhibitor sensitivity of proteases indicated the cysteine proteases were mostly in the anterior and serine proteases were in the posterior midgut of *T. molitor* larvae. Expressed Sequence Tags (ESTs) from *T. molitor* larval midgut cDNA libraries contained sequences encoding putative digestive proteases. Of a total of 1,528 cDNA sequences, 92 cDNAs encoded proteases, and 50 full-length cDNAs were grouped into serine, cysteine and metallo protease classes. Sequences *tmt1a*, *tmt1b* and *tmt1c* were identified as genes encoding isoforms of *T. molitor* trypsin, and *tmc1a* encoded *T. molitor* chymotrypsin. The general distribution cysteine protease transcripts in the anterior and serine protease transcripts in the posterior midgut, of *T. molitor* larvae, was in agreement with the biochemically-characterized compartmentalization of proteases. Expression analyses of selected transcripts demonstrated varied expression patterns across five developmental stages of *T. molitor*, with maximal expression of most protease transcripts in first instar larvae. Dietary serine and cysteine protease inhibitors fed in combination to early-instar *T. molitor* larvae caused a significant delay in larval growth in 21-day-old larvae. Real-time quantitative PCR analysis of RNA isolated from larvae fed different protease inhibitor treatments indicated that dietary inhibitors affected the expression of serine and cysteine proteases. Larvae fed soybean trypsin inhibitor, a serine protease inhibitor, compensated by the hyperproduction of proteases from the same class, as well as the upregulation of cysteine proteases. A cysteine protease inhibitor, E-64, caused a reduction in the hyperproduction of all proteases, and, in combination with the soybean trypsin inhibitor, lowered the compensation response of *T. molitor* larvae to negligible levels. These data suggest that *T. molitor* larvae are more sensitive to the effects of cysteine protease inhibitors, perhaps because these proteases are the first line of defense for larvae against plant protease inhibitor. The bioassay and molecular studies suggested

that combinations of inhibitors that target both serine and cysteine proteases are needed to effectively control larval infestations of *T. molitor*.

Key words:

Expressed Sequence Tag (EST), protease inhibitors (PIs), real-time quantitative PCR, *L-trans*-epoxysuccinylleucylamide [4-guanidino] butane (E-64), soybean trypsin inhibitor (STI)

MOLECULAR CHARACTERIZATION OF DIGESTIVE PROTEASES OF THE  
YELLOW MEALWORM, *TENEBRIO MOLITOR* L.

by

SHEILA PRABHAKAR

B.Sc., Osmania University College for Women, 1999  
M.Sc., Osmania University, 2001

A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Entomology  
College of Agriculture

KANSAS STATE UNIVERSITY  
Manhattan, Kansas

2006

Approved by:

Co-Major Professor  
Brenda Oppert

Approved by:

Co-Major Professor  
C. Michael Smith

## Abstract

Coleopteran insects compensate for dietary protease inhibitors by a number of mechanisms. To study this compensation response at the molecular level, the digestive proteases of *Tenebrio molitor* were studied. Biochemical studies of the pH optima and inhibitor sensitivity of proteases indicated the cysteine proteases were mostly in the anterior and serine proteases were in the posterior midgut of *T. molitor* larvae. Expressed Sequence Tags (ESTs) from *T. molitor* larval midgut cDNA libraries contained sequences encoding putative digestive proteases. Of a total of 1,528 cDNA sequences, 92 cDNAs encoded proteases, and 50 full-length cDNAs were grouped into serine, cysteine and metallo protease classes. Sequences *tmt1a*, *tmt1b* and *tmt1c* were identified as genes encoding isoforms of *T. molitor* trypsin, and *tmc1a* encoded *T. molitor* chymotrypsin. The general distribution cysteine protease transcripts in the anterior and serine protease transcripts in the posterior midgut, of *T. molitor* larvae, was in agreement with the biochemically-characterized compartmentalization of proteases. Expression analyses of selected transcripts demonstrated varied expression patterns across five developmental stages of *T. molitor*, with maximal expression of most protease transcripts in first instar larvae. Dietary serine and cysteine protease inhibitors fed in combination to early-instar *T. molitor* larvae caused a significant delay in larval growth in 21-day-old larvae. Real-time quantitative PCR analysis of RNA isolated from larvae fed different protease inhibitor treatments indicated that dietary inhibitors affected the expression of serine and cysteine proteases. Larvae fed soybean trypsin inhibitor, a serine protease inhibitor, compensated by the hyperproduction of proteases from the same class, as well as the upregulation of cysteine proteases. A cysteine protease inhibitor, E-64, caused a reduction in the hyperproduction of all proteases, and, in combination with the soybean trypsin inhibitor, lowered the compensation response of *T. molitor* larvae to negligible levels. These data suggest that *T. molitor* larvae are more sensitive to the effects of cysteine protease inhibitors, perhaps because these proteases are the first line of defense

for larvae against plant protease inhibitor. The bioassay and molecular studies suggested that combinations of inhibitors that target both serine and cysteine proteases are needed to effectively control larval infestations of *T. molitor*.

Key words:

Expressed Sequence Tag (EST), protease inhibitors (PIs), real-time quantitative PCR, *L-trans*-epoxysuccinylleucylamide [4-guanidino] butane (E-64), soybean trypsin inhibitor (STI)

# Table of Contents

List of Figures .....	xii
List of Tables .....	xv
Acknowledgements .....	xvi
Dedication .....	xviii
CHAPTER 1 - INTRODUCTION AND LITERATURE REVIEW .....	1
1. Pest Control in the Stored Product Industry .....	1
1.1. Need for Sustainable Stored Grain Insect Pest Management .....	1
1.2. Economics of Food Safety .....	1
1.3. Biological and Eco-friendly Methods of Pest Control.....	2
2. Insect Digestive System.....	2
2.1. The Insect Digestive System - A Model for Recent Study .....	2
2.2. The Insect Digestive Process .....	3
2.3. Role of Proteases in Digestion.....	3
3. Protease Inhibitors (PIs).....	4
3.1. Introduction to Protease Inhibitors as Natural Biopesticides.....	4
3.2. Factors Influencing the Success or Failure of PIs as Bioinsecticides .....	5
3.3. Adaptation to Protease Inhibitors.....	5
3.4. Conflicting Results in the <i>in vitro</i> and <i>in vivo</i> Evaluation of PIs.....	7
3.5. Differential Gene Regulation Influencing Adaptation.....	7
3.6. Transcriptomic Studies on Insect Digestive Proteases .....	9
3.7. Use of an Inhibitor Combination Strategy .....	10
4. <i>Tenebrio molitor</i> Linnaeus (Yellow Mealworm).....	10
4.1. Life cycle of <i>Tenebrio molitor</i> .....	10
4.2. <i>Tenebrio molitor</i> Digestive Proteases .....	11
4.2.1. Occurrence and Midgut Compartmentalization of Serine and Cysteine Proteases .....	11
4.2.2. Early Studies on <i>Tenebrio molitor</i> Proteolytic Enzymes.....	12
4.2.3. Recent Studies on <i>T. molitor</i> Midgut Proteases.....	13

4.2.4. Molecular Studies on <i>T. molitor</i> Midgut Proteases .....	13
5. Molecular Characterization of Digestive Proteases.....	14
5.1. Sequence Information on cDNAs Encoding Digestive Protease in Other Insects .....	14
5.2. Benefits of Characterizing cDNAs Encoding Digestive Enzymes .....	14
6. Research Hypotheses and Objectives .....	14
6.1. Rationale and Hypotheses.....	14
6.2. Research Objectives.....	15
6.3. Scope of the Project .....	15
7. References.....	16
 CHAPTER 2 - BIOCHEMICAL STUDY OF THE DIGESTIVE PHYSIOLOGY OF THE YELLOW MEALWORM, <i>TENEBRIO MOLITOR</i> L.....	
2.1. Abstract.....	39
2.2. Introduction.....	40
2.2.1. Proteases .....	40
2.2.2. Rationale .....	42
2.3. Materials and Methods.....	43
2.3.1. Isolation of Gut Enzymes.....	43
2.3.2. Microplate Assay-pH Curve .....	43
2.3.3. pH profile Comparison of <i>T. molitor</i> Larvae Reared on Wheat Flour and Wheat Germ .....	44
2.3.4. Effect of Variability of Gut Samples on Total Proteolytic Activity .....	45
2.3.5. Inhibitor Assays .....	45
2.3.6. Activity Blots .....	45
2.3.7. Zymograms .....	46
2.4. Results.....	46
2.4.1. Microplate Assay .....	46
2.4.2. pH Profile Comparison of <i>T. molitor</i> Larvae Reared on Wheat Flour and Wheat Germ .....	47
2.4.3. Effect of Storage of Gut Samples on Total Proteolytic Activity .....	47
2.4.4. Inhibitor Assays .....	47

2.4.5. Activity Blots .....	48
2.4.6. Zymograms .....	48
2.5. Discussion .....	48
2.6. References .....	51
CHAPTER 3 - TRANSCRIPTOME ANALYSIS OF DIGESTIVE PROTEASES FROM THE YELLOW MEALWORM, <i>TENEBRIO MOLITOR</i> L. (COLEOPTERA: TENEBRIONIDAE) .....	
3.1. Abstract .....	65
3.2. Introduction .....	66
3.3. Materials and Methods .....	67
3.3.1. Insect Rearing and Dissection .....	67
3.3.2. cDNA Library Construction and Sequencing .....	68
3.3.3. Computer-based Sequence Analysis .....	68
3.3.4. Isolation and Purification of <i>T. molitor</i> Digestive Trypsin and Chymotrypsin .....	69
3.3.5. In Gel Tryptic Hydrolysis of Purified Trypsin and Chymotrypsin .....	69
3.3.6. MALDI-TOF MS and MS/MS Analysis .....	70
3.3.7. RNA Isolation and Northern Blot Analysis .....	70
3.3.8. Phylogenetic Analysis .....	71
3.4. Results .....	71
3.4.1. Composition of <i>T. molitor</i> Transcripts .....	71
3.4.2. Distribution of Protease Transcripts in <i>T. molitor</i> .....	71
3.4.3. Classification of Protease Transcripts .....	72
3.4.3.1. Serine Proteases .....	72
3.4.3.2. Cysteine Proteases .....	74
3.4.3.3. Metalloproteases .....	75
3.4.4. Comparison of Sequences from Predicted Serine Proteases with Purified TmT1 and TmC1 .....	75
3.4.5. Phylogenetic Analysis of Protease cDNAs from <i>T. molitor</i> .....	76
3.4.6. Expression Analyses .....	77
3.5. Discussion .....	78

3.6. References.....	80
CHAPTER 4 - EVALUATION OF THE COMPENSATORY RESPONSE OF	
<i>TENEBRIO MOLITOR</i> LARVAE TO DIETARY PROTEASE INHIBITORS .....	113
4.1. Abstract.....	113
4.2. Introduction.....	114
4.2.1. Protease Inhibitors- Insect Adaptation.....	114
4.2.2. Multiple PI Strategy .....	114
4.2.3. <i>In vitro</i> and <i>in vivo</i> Response to Inhibitors .....	115
4.2.4. Real-time qPCR in Gene Expression Fold Change Determination .....	115
4.3. Materials and Methods.....	116
4.3.1. Inhibitors .....	116
4.3.2. Insect Bioassays .....	116
4.3.3. Preparation of Treatment Diets .....	116
4.3.4. Larval Weight Measurement.....	116
4.3.5. Statistical Analysis.....	117
4.3.6. Isolation of RNA.....	117
4.3.7. Primer design for Real-time qPCR .....	118
4.3.8. Real-time qPCR .....	118
4.3.9. Calculation of Fold Change Level .....	119
4.4. Results.....	119
4.4.1. Bioassays.....	119
4.4.2. Standard Curves for Real-time qPCR.....	120
4.4.3. Relative Fold Change Level.....	121
4.5. Discussion.....	121
4.6. References.....	123
CHAPTER 5 - SUMMARY .....	
5.1. Practical Implications and Application.....	134
5.1.1. Testing Effect of Other Inhibitors <i>in vitro</i> and <i>in vivo</i> .....	134
5.1.2. Application in Transgenics-Concerns and Considerations .....	135
5.2. Suggestions for the Future .....	136
5.3. Conclusions.....	137

5.4. References.....	138
Appendix A - List of Protease Inhibitors.....	142

## List of Figures

Figure 2.1. Hydrolysis of fluorescent-casein by enzymes in the anterior midgut (AM) extract of <i>T. molitor</i> larvae, in buffers containing or lacking L-cysteine. ....	56
Figure 2.2. Hydrolysis of fluorescent-casein by enzymes in the posterior (PM) midgut extract of <i>T. molitor</i> larvae, in buffers containing or lacking L-cysteine. ....	57
Figure 2.3. Hydrolysis of fluorescent casein by enzymes in the midgut extract of <i>T. molitor</i> larvae reared on (a) wheat flour (b) wheat germ in buffers containing or lacking L-cysteine. ....	58
Figure 2.4. Hydrolysis of fluorescent-casein by enzymes in the anterior midgut (AM) extracts of guts (a) stored at -80°C (b) freshly dissected guts of <i>T. molitor</i> larvae, in buffers containing L-cysteine. ....	59
Figure 2.5. Hydrolysis of fluorescent casein by enzymes in the posterior midgut (PM) extracts of guts (a) stored at -80°C (b) freshly dissected guts of <i>T. molitor</i> larvae, in buffers containing L-cysteine. ....	60
Figure 2.6. Hydrolysis of fluorescent-casein by enzymes in <i>T. molitor</i> gut extracts of freshly dissected guts or guts stored at -80°C that were in water or in saline buffer	60
Figure 2.7. A microplate inhibitor assay of <i>T. molitor</i> midgut incubated with inhibitors and enzyme activity measured with fluorescent-casein, in non-reducing buffers...	61
Figure 2.8. A microplate inhibitor assay of <i>T. molitor</i> midgut extracts incubated with inhibitors and enzyme activity measured with fluorescent-casein, in reducing buffers. ....	62
Figure 2.9. Activity blot incubated with SAAPFpNA showing six different chymotrypsin-like activities (numbered 1-6) in <i>T. molitor</i> midgut extract.....	63
Figure 2.10. Zymogram Gel picture showing <i>T. molitor</i> midgut protease (AM & PM) activities (1-8) when inhibited with five different inhibitors (lanes a-e). ....	64
Figure 3.1. Categories of sequences from randomly picked clones of cDNA libraries from <i>T. molitor</i> larvae (BLASTX matches with E<10).....	87

Figure 3.2. Distribution of sequences encoding digestive proteases from the anterior and posterior midgut cDNA libraries of *T. molitor* ..... 88

Figure 3.3. Alignment of predicted proteases from *T. molitor* larvae A) Alignment of serine proteases from *T. molitor* with trypsins from *Drosophila melanogaster* (Dmatryp, accession # CAA26732) and *Homo sapiens* (HsPRSS2b, accession # AAL14244); B) Alignment of predicted cathepsin B sequences with cathepsin B proteins from *Diabrotica virgifera virgifera* (DvvcathB accession #CAE47498) and *Araneus ventricosus* (AvcathB, accession # AAP59456); C) Alignment of predicted cathepsin L sequences with cathepsin L from *T. molitor* (TmcathL, accession # AAP94048); D) Alignment of predicted carboxypeptidase sequences with carboxypeptidase A from *Aedes aegypti* (AeCPA, accession #AAT36730) and *Bos taurus* (BtCPA- Accession# POO730). Shaded regions at the N-terminus are predicted signal peptides and within the sequences represent potential N-glycosylation sites; bolded sequences are conserved cysteines or critical residues; active site residues are boxed; arrow indicates activation site; stars indicate residues that determine specificity; loop regions are indicated in brackets..... 89

Figure 3.4. Alignment of TmT1 sequences with predicted trypsins (A) and TmC1 sequences with predicted chymotrypsins (B) from *T. molitor* cDNAs. Shaded sequences represent peptides that differ from their homologs in TmT1 and TmC1 ..... 100

Figure 3.5. Phylogenetic tree generated using MEGA 2.1 program. Bootstrap values were calculated based on the distance/neighbor-joining method. Protease cDNAs were A) serine; B) cathepsin B C) cathepsin L D) carboxypeptidase A. Abbreviations are the same as those described in Figure 3..... 102

Figure 3.6. Northern blot analysis of RNA extracted from first-instar larvae (L1), mid-instar larvae (L2) and late-instar larvae (L3), pupae and adults, as indicated in the figure. Probes specific to selected cDNAs from representative clusters were used for analysis. Tentative identification of the gene product is on the right. SP= serine protease homolog; SPH = serine protease homolog ..... 107

Figure 4.1. *T. molitor* eggs and newly emerged neonates (1<sup>st</sup> instar ) larvae used to infest treatment diets ..... 126

Figure 4.2. Contrast of *T. molitor* larvae fed a combination of 1% STI and 0.1% E-64 with those fed either 1% STI or 0.1% E-64 or control diet (no inhibitor)..... 127

Figure 4.3. A comparison of the weights of *T. molitor* larvae fed control, 1% STI, 0.1% E-64, or 1% STI + 0.1% E-64 at 7, 14 and 21 days post-hatch. Weights are the mean + S.E., n=15. .... 128

Figure 4.4. Real-time quantitative PCR analysis of the relative quantitative fold change expression of two *T. molitor* serine protease transcripts, *tmt1* and *tmc1* and two cysteine protease transcripts, *cathB* and *cathL* in guts of 21-day old larvae fed on diet containing a single inhibitor (1% STI or 0.1% E-64) or a combination of inhibitors (1% STI+0.1% E-64)..... 129

## List of Tables

Table 1.1. Table of proteases identified from <i>Tenebrio molitor</i> .....	27
Table 3.1. cDNAs encoding serine proteases from <i>T. molitor</i> larvae, with characteristics of the predicted peptides (contd.).....	109
Table 3.2. cDNAs encoding cysteine and metalloproteases from <i>T. molitor</i> larvae, with characteristics of the predicted peptides .....	111
Table 3.3. Relative abundance of mRNA transcripts in different developmental stages of <i>Tenebrio molitor</i> larvae.....	112
Table 4.1. Primer names, sequence IDs and nucleotide sequences of selected <i>T. molitor</i> digestive protease cDNAs. Primers were designed using a computer program, Beacon Designer (version 2.1, Premier Biosoft International, Palo Alto, CA).....	130

## Acknowledgements

I am deeply indebted to my co-major advisors, Dr. Brenda Oppert and Dr. C. Michael Smith for their invaluable input, help, support, guidance, patience and encouragement. I owe an extra-special, big debt of gratitude to Dr. Ming-shun Chen for opening his doors, taking me under his wings, and for giving direction to the project. He graciously allowed me to use the lab supplies and was a constant encouragement spurring me on to success and to achieve greater things. I especially would like to acknowledge the timely expertise and advice extended by my committee members, Dr. Yoonseong Park, Dr. Kun Yan Zhu and Dr. Mike Kanost. I also wanted to thank Dr. Eric A. Maatta for serving in the capacity as the Chairperson of the Examination Committee. My coming to Kansas State University would not have been possible had it not been for Dr. Sonny Ramaswamy. He stepped in at crucial phases during the project and made a way for me to see the completion of this project.

I take this opportunity to also thank Xiang Liu for teaching me all the molecular techniques and patiently helping me learn them. Thank you Xiang! I also thank my labmates (Dr. Chen's lab), Ashoka A. Maddur, Lieceng Zhu, Huixian Zhao, Xuming Liu, Raman, Chitvan and Hari Pillai for their help in the lab. I extend a very special thank you to Michele Zuercher for introducing me to laboratory life (USDA lab) and for her friendship and encouragement. I would like to express my gratitude to Tom Morgan for helping me with insect bioassays and to Roy Speirs for helping me master insect dissections. I humbly thank Yasuyuki Arakane (Dr. Kramer's lab), Dave Hogenkamp for their priceless help with RNA work. I thank Dr. Aqeel Ahmad for helping me with data analysis using SAS. I also thank Dr. Richard Beeman and Dr. Yoonseong Park for their collaboration and help with the sequencing of cDNAs and for letting me work on *T. castaneum* dissections as a summer project. A word of mention of Dr. James Higgins for helping me analyze my data using SAS program.

The entomology department provided a niche for me to grow in knowledge and develop professional skills and I offer my sincere thanks to each person- faculty, staff,

post-doctorates and fellow graduate students that is part of this family. I thank Dr. James R. Nechols and Dr. Gregory Zolnerowich for playing an important role in supervising my teaching as I served as a teaching assistant for Insect Biocontrol Lab (Spring 2005) and General Entomology Lab (Spring 2006).

I would like to acknowledge Dr. Youping Deng's help with analyzing 500 cDNA sequences. I thank Dr. Jianfa Bai (Gene Expression Facility) for introducing me to Microarrays and Real-time qPCR and for letting me use the Nanodrop spectrophotometer. Dr. Smith, Dr. Damon Crook, and Dr. Yoonseong Park deserve special mention as they provided me an opportunity to work collaboratively on the greenbug, red oak borer project and *T. castaneum* project respectively.

I thank Mr. Joe Horan and Mr. Martin Courtois for their assistance in formatting my thesis.

Last but not least, I consider it a privilege and honor to thank my friends, Bikat and Edna, Sharon, Kay and Charles, Chris, Zoma, Radhika, Sara, Barb, Romal, Martin, Punya, and family, Daddy, Mommy and Anita, my sister for being behind me in all my endeavors. Above all, I thank God for His grace, mercy, faithfulness and love in sustaining me, equipping me with strength to endure and persevere and for being with me every step of the way.

## **Dedication**

I dedicate this book to my father, Dr. Prabhakar Babu, mother, Dr. Mrs. Linda Babu and sister Anita Prabhakar.

# CHAPTER 1 - INTRODUCTION AND LITERATURE REVIEW

## 1. Pest Control in the Stored Product Industry

### *1.1. Need for Sustainable Stored Grain Insect Pest Management*

Flour and other grain-based processed foods represent value-added products, and the food industry does not tolerate even low levels of insect infestation. There is essentially zero tolerance for insect fragments or insect damage, irrespective of the established federal defect action levels (Kenkel *et al.*, 1994). The occurrence of a single insect-infested sample may result in the rejection of an entire truck, train or barge load of grain (Hagstrum *et al.*, 1999). Insect pest damage to grain, stored grain and grain by-products from grain causes significant losses to the food industry, estimated in millions of dollars annually. Insect pest control agencies are working toward sustainable stored grain insect pest management strategies that deliver high quality, safe, competitively priced food and grain products, within the regulatory framework of production for global markets and governments.

### *1.2. Economics of Food Safety*

In a study conducted by Storey *et al.* (1983) in the U.S., insects were detected in about one-fourth of over 4,000 grain samples from wheat stored for 1-4 years. A more recent study involving sampling in commercial grain elevators in Kansas revealed that more than 50% of infested samples of grain contained >10 insects/kg, and that grain residues likely were either densely infested with pest insects or uninfested (Arthur *et al.*, 2006).

Stored grain, free from insect fragments and contamination exceeding the normal tolerable limits, is not only a critical food safety issue but also important for trading of commodities in the international market. The annual worldwide pre-harvest losses due to insect pests, despite the use of insecticides, are approximately 15% of the total production costs, and are valued at over \$ 100 billion (Krattiger, 1997). Additionally, post-harvest losses caused by insect infestation, a majority of which occur in the developing world, is estimated to be 15% of the world's production (Herrera-Estrella, 2000).

### ***1.3. Biological and Eco-friendly Methods of Pest Control***

With increases in human population and rapid depletion of natural resources, there is a demand for increased sustainable agricultural productivity. However, world events also have mandated new procedures to ensure food biosafety. With stringent regulations monitoring synthetic chemical-based insect pest management, and with the continued development of pesticide resistance, the need for alternative pest control options has become urgent. Although transgenic cotton and maize crops are being produced widely in many developed countries, there is a continued need to discover new plant genes that provide resistance against pests because of regulatory and pest resistance issues. Plant genes encoding protease inhibitors (PIs) are prime candidates, with demonstrated activity against insect pests, with the added benefit of improving the nutritional quality of food in some cases (Lawrence and Koundal, 2002).

## **2. Insect Digestive System**

### ***2.1. The Insect Digestive System - A Model for Recent Study***

A dramatic growth of knowledge in the area of insect digestive physiology has occurred during the past decade. Interest in insect digestion waned after the development of synthetic insecticides in the 1940s. However, the environmental effects of some of these pesticides promoted an enthusiastic endeavor to pursue new avenues for insect control, one of them targeting the insect gut. The habitat and feeding diversity of insects, and the fact that the insect gut is a large and susceptible interface between the insect and its environment, have stimulated an increased understanding of gut function (Terra and Ferreira, 1994; 2005). The insect digestive system is the first line of defense against a broad spectrum of toxins and anti-nutritional dietary factors that break down complex food components into simple products, and also serving as a protective barrier to toxins for vulnerable insect tissues and cells (Moon *et al.*, 2004). Herbivorous insects feeding on plant tissues are attacked by a range of biochemical plant defenses, both constitutive and induced, in response to their attack on plants. The insect gut is therefore a viable target in the development of pest control methods, such as the use of transgenic plants expressing insect-specific toxins to control phytophagous insects. In order to be successful herbivores, insects need to be able to detoxify these plant compounds, or become insensitive to their toxic effect.

## ***2.2. The Insect Digestive Process***

The ingestion and digestion of plant material by insects triggers a series of physiological and biochemical events that result in the conversion of protein, carbohydrate and lipid in food to utilizable nutrients. These processes involve enzymes, and proteases are those enzymes that act on long polypeptide chains or on small peptides. The digestive process in insects is well organized and depends on the compartmentalization of digestive enzymes and on midgut fluid fluxes that translocate enzymes and products of digestion (Terra and Ferriera, 1994). Terra and Ferriera also described evolutionary trends in and spatial distribution of insect gut enzymes. Generally, initial digestion occurs inside the peritrophic membrane, intermediate digestion in the ectoperitrophic space and final digestion at the surface of the midgut cells by integral microvillar enzymes or by enzymes trapped in the glycocalyx (Terra and Ferriera, 2005). For example, digestion in the coleopteran pest, *Tenebrio molitor* L. is thought to occur in the endoperitrophic space, except for final digestion that takes place on the surface of midgut cells under the action of a microvillar aminopeptidase.

Most food molecules are complex polymers of protein and starch that are broken down sequentially in a three-step manner (Terra and Ferriera, 2005). Primary digestion is the dispersion and reduction in molecular size of food from polymers to oligomers. The reduction of oligomers to dimers is the intermediate phase in digestion, and the final step in digestion results in the production of monomers. All steps occur in the midgut under the influence and action of digestive enzymes.

## ***2.3. Role of Proteases in Digestion***

In the initial digestion of proteins, proteases (endopeptidases) are involved in breaking the internal bonds in proteins. Dipeptides are hydrolyzed by dipeptidases. Oligopeptides formed as a result of protease action are attacked from the N-terminal end by aminopeptidases and from the C-terminal end by carboxypeptidases releasing one amino acid at each step.

Proteases are classified according to their mechanistic action into serine, cysteine, aspartic and metallo-proteases (Barrett, 1986). In insects, proteases are secreted only in response to dietary protein entering the midgut, yet Bown *et al.* (1997) found 20% differentially regulated proteases from among the cDNAs expressed in the insect midgut. Terra and Ferriera (1994)

emphasized the need to study insect digestive enzymes in order to further understand their mechanisms, substrate interactions and inhibitor specificity.

### **3. Protease Inhibitors (PIs)**

#### ***3.1. Introduction to Protease Inhibitors as Natural Biopesticides***

Protease inhibitors (PIs) are synthesized by plants and are the most-studied class of plant-defense proteins. They have long been regarded of as natural, phytochemical defenses against arthropod attack, primarily because of their anti-nutritional effects that in turn negatively impact insect growth and development. The expression of PIs in plants is induced under stress-prone conditions, such as mechanical wounding, insect feeding, pathogen attack, drought and UV exposure (Schaller and Ryan, 1996; Conconi *et al.*, 1996; Giri *et al.*, 1998). These plant compounds are the end-products of the many defense cascades activated by plant elicitors, such as systemin, ethylene, methyl jasmonate, abscisic acid, salicylic acid, fungal cell wall oligomers, larval oral secretions, and electrical and hydraulic signals, leading to their accumulation in plant tissues (Patankar *et al.*, 1999; Ryan, 1990; Wildon *et al.*, 1992; Schaller and Ryan, 1996; Korth and Dixon, 1997). The toxicity of PIs to insects and their role in defense is attributed to their accumulation in quantities that are greater than those required for inhibiting endogenous proteases. The inhibition of proteases by inhibitors disrupts the normal insect digestive process, depletes the amino acid reserve, and affects other processes, such as proteolytic activation of enzymes, and molting (Hilder *et al.*, 1993). A PI found in soybean first reported by Read and Haas (1938) has been well studied. Among the inhibitors of insect digestive enzymes, the role of trypsin inhibitors has been more thoroughly investigated. The toxicity of the soybean trypsin inhibitor to larvae of the confused flour beetle, *Tribolium confusum* Jacquelin du Val, was demonstrated by Lipke *et al.* in 1954. The majority of plant protease inhibitors have been described in plants belonging to families Leguminosae, Solanaceae and Graminaceae (Richardson, 1977; 1991). Many experiments have been conducted to identify the function and distribution of PIs in higher plants. PLANT-PIs, a database of all known plant inhibitors, was developed to facilitate retrieval of such information (De Leo *et al.*, 2002).

### ***3.2. Factors Influencing the Success or Failure of PIs as Bioinsecticides***

The exploitation of plant defenses against insects is a potential insect pest control option, similar to cultivating and breeding insect-resistant plant varieties. However, insect counter-adaptation to plant defense mechanisms has been a major obstacle in insect pest management. Entomologists concerned with crop protection and plant resistance are faced with a perennial struggle of overcoming insect tolerance or resistance to integrated approaches. Jongsma and Bolter (1996) outlined factors to be considered before incorporating PIs artificially into an insect diet, or expressing a PI gene in a plant. These include the PI concentration ( $K_i$ ) for interaction with a protease, the stability of the PI in the insect gut, the complement of inhibitors needed to inhibit all gut proteases, and the insect's ability to adapt to protease inhibition. Factors affecting the potency of a protease inhibitor include the structural compatibility of the PIs with the protease(s) in the target organism, the physiological conditions within the midgut, and the dietary quality (e.g. polyphenoloxidase activity, protein quality, and quantity) (Laskowski, 1985; Broadway and Duffey, 1988; Felton *et al.*, 1989; Oppert *et al.*, 1993; Christeller *et al.*, 1994; Broadway, 1995). A review of insect proteases and protease inhibitors included approximately 11 biochemical and physiological factors that contribute to the effectiveness of PIs (Reeck *et al.*, 1997). Information about protease-protease inhibitor interactions between an insect and its target crop often is lacking, and much of the research conducted on PIs in insects has taken into account only one or a few of the aforementioned factors (Jongsma and Bolter, 1997). The co-evolution of plant PIs and insect proteases provides an interesting new paradigm for new ecological, physiological and biochemical research.

### ***3.3. Adaptation to Protease Inhibitors***

The effects of biopesticides based on proteinaceous PIs of insect digestive proteases have, in many cases, been ineffective because of an insect's ability to digest or tolerate these proteins (Konarev, 1996; Oppert, 2000). Responses to ingested PIs include the expression of proteases for which the plant has no inhibitors, the proteolytic degradation of PIs, and mutations in insect protease genes that render the proteases less sensitive to PIs without the loss of protease activity (Jongsma and Bolter, 1997). In some cases, insects have adapted to plant protease inhibitors by producing inhibitor-insensitive proteases (Bolter and Jongsma, 1995; Jongsma *et al.*, 1995; Bown *et al.*, 1997; Cloutier *et al.*, 2000; Mazumdar-Leighton and Broadway, 2001).

In others, insects combat the plant PI by producing inhibitor-degrading/fragmenting proteases in the midgut to compensate for the effects of the inhibitors (Michaud *et al.*, 1995; Giri *et al.*, 1998; Franco *et al.*, 2003; Girard *et al.*, 1998; Patankar *et al.*, 2001). De Leo *et al.*, (1998) found *Spodoptera littoralis* larvae that over-produce existing digestive enzymes to essentially 'titre' out the inhibitors. Some studies have indicated that cowpea bruchids, *Callosobruchus maculatus* (F.), employ all three of the aforementioned strategies as a response to being fed dietary soyacystatin (scN) (Moon *et al.*, 2004). Studies on the dietary factors that influence the number and level of activity of protease inhibitor resistant-enzymes in midguts of larval lepidopterans revealed a prolongation of the retention time of food in the digestive tract as well as the initiation of an increased level of activity of proteolytic enzymes not susceptible to inhibition by a particular protease inhibitor (Broadway, 1997). The two responses were proposed to be linked, because food was retained in the gut prior to the secretion of new proteolytic enzymes. Furthermore, the quantity and quality of protein had no significant effect on the level of activity of inhibitor-resistant enzymes. High levels of PIs targeting cysteine and aspartic proteases induced the production of papain inhibitor-insensitive enzymes in the Colorado potato beetle, *Leptinotarsa decemlineata* (Say) (Bolter and Jongsma, 1995). In addition, Broadway (1996b) noted that isozymes of gut proteases, the regulation of the level of activity of inhibitor-resistant enzymes, and preadaptation to PIs from non-host plants were a few factors critical in rendering insects resistant to PIs. Other studies have demonstrated that insect gut proteases degrade inhibitors and thus neutralize their toxic effects. Orr *et al.* (1994) have shown that *Diabrotica spp.* larval proteases are capable of degrading multicystatins of potato tubers. Girard *et al.* (1998) demonstrated that in interactions between larval proteases of a coleopteran, the mustard beetle, *Phaedon cochleariae* (F.), and PIs, serine proteases in association with leucine aminopeptidases rapidly cleaved the inhibitors. It was suggested that, in fact, insects derive dual benefit from digesting PIs, in that the digestion is not disrupted, and sulfur-rich amino acids provide additional nutrition for the insect.

This compensatory, adaptive response to PIs has been found even in tritrophic ecological interactions. In a study that monitored the inhibitory activity of recombinant oryzacystatin (OCI) along the potato - herbivore - predator continuum, predatory organisms adapted their digestive metabolism to the presence of plant anti-digestive proteins ingested by their herbivorous prey (Bouchard *et al.*, 2003). Therefore, studies that investigate protease responses to inhibitors

should monitor all the inhibitory effects of PI-expressing plants, not only on the herbivorous insects targeted, but also to the organisms likely to consume these pests.

### ***3.4. Conflicting Results in the in vitro and in vivo Evaluation of PIs***

Results of a study by Ortego *et al.*, 1998, to evaluate the effects of protease inhibitors on the larvae of the weevil *Aubeonymus mariaefrancisciae* Roudier, indicated that, although *in vitro* screening may be useful in selecting effective PIs, *in vivo* validation through feeding trials is necessary. It appears that insects have evolved well-developed compensatory responses to enzyme inhibitors under *in vivo* environments that cannot be tested effectively under *in vitro* conditions. The efficacy of a protein when expressed in a crop plant also may differ from that demonstrated in an artificial diet (Edmonds *et al.*, 1996). To test adaptive responses of insects to plants containing protease inhibitors, Broadway (1995) evaluated the effects of cabbage PIs on the cabbage looper, *Trichoplusia ni* (Hübner) and the corn earworm, *Helicoverpa zea* (Boddie). PIs reduced the growth and development of *T. ni* but had no influence on the level of digestion in *H. zea*. Apparently, insects not affected by protease inhibitors developed compensatory responses in the form of a hyper production of proteolytic enzymes. These data indicated that *in vitro* inhibition is not a good predictor of toxicity and were the first report of a shift in the relative proportion of proteolytic enzymes following insect ingestion of PIs (Broadway, 1995). Although the predominant trypsin-like enzymes were susceptible to trypsin inhibitors under 'normal' conditions, they were at least partially resistant to inhibition (i.e. weak binding capacity for inhibitors) following the ingestion of inhibitors. In another study, larvae of tobacco budworm, *Helicoverpa armigera* (Hübner) were stunted, but no mortality was observed when fed transgenic tobacco expressing a PI from the giant taro, *Alocasia macrorrhiza* Schott (Wu *et al.*, 1997). Larvae fed the transgenic plants had an 8% increase in inhibitor-insensitive proteases, establishing the existence of an adaptive mechanism.

### ***3.5. Differential Gene Regulation Influencing Adaptation***

Ingestion of soybean Kunitz trypsin inhibitor (SKTI) by *H. armigera* larvae induces the production of inhibitor-insensitive protease activity (Bown *et al.*, 1997). Gatehouse *et al.* (1997) also found a significant down-regulation of *H. armigera* larval trypsin genes and an up-regulation of chymotrypsin in response to dietary trypsin inhibitors. To phytophagous insects such as *H. armigera*, the presence of a wide array of potential proteolytic enzymes with differing

sensitivity to PIs is a distinct advantage. Recently, a study by Bown *et al.* (2004), revealed the complexity of *H. armigera*'s response to the ingested SKTI inhibitor. The response consisted of an initial up-regulation of all protease genes, followed by a longer-term down-regulation of inhibitor-specific protease genes, leading to a final up-regulation in inhibitor-insensitive proteases. The presence of non-target proteases in the gut extract of the Colorado potato beetle, *Leptinotarsa decemlineata* (Say), has presented questions about the stability of target protease inhibitors, such as cystatins, in the insect gut (Michaud *et al.*, 1995). Ahn *et al.* (2004) investigated the response of *C. maculatus* to a cysteine PI, scN, and found that bruchids reconfigure their major digestive proteases by inducing unique cysteine protease isoforms with superior auto-processing efficacy. Cowpea bruchids are able to overcome plant PIs by overproducing existing digestive enzymes, producing inhibitor-insensitive enzymes and by activating the expression of hydrolyzing enzymes that degrade enzymes when challenged by dietary scN (Zhu-Salzman *et al.*, 2003). Moon *et al.*, (2004) conducted a transcriptomic analysis of the response of *C. maculatus* guts challenged by scN, and found that numerous anti-microbial peptide genes were upregulated. This strengthening of insect defense against microbes was seen as an unintended negative side effect to using PIs. Other responses to plant PIs included an increased ferritin expression a decreased cytochrome C oxidase expression. Adaptation to protease inhibitors in larvae of the tobacco budworm, *Heliothis virescens* (Fabricius), was mediated by the synthesis of new proteases (Brito *et al.*, 2001). *H. virescens* larvae express new trypsin molecules that form oligomers and are apparently less affected by PIs because of tighter binding to the substrate (lower  $K_m$  values) and a putative decreased affinity for PIs.

Several molecular studies have characterized insect digestive carboxypeptidases (Ramos *et al.*, 1993; Bown *et al.*, 1998; Edwards *et al.*, 2000; Bown and Gatehouse, 2004; Bayes *et al.*, 2005). Bayes *et al.* (2005) identified a carboxypeptidase B from *H. zea* and elucidated the structural features that would render it insensitive to potato carboxypeptidase inhibitor (PCI) and other plant inhibitors. Four aminopeptidase N isoforms were identified from *T. ni* (Wang *et al.*, 2005). Because this enzyme also has been implicated in a receptor for some of the microbial toxins of *Bacillus thuringiensis* (*Bt*), this study has useful implications in *T. ni*-*Bt* toxin interactions, and in developing the role of the aminopeptidases in resistance management of *Bt* technology.

From all of the aforementioned examples of research on insect digestive protease-PI interaction, it is clear that the mechanism(s) controlling and regulating the insect's response to protease inhibitors is complex. The biochemical and molecular bases of these plastic responses to enzyme inhibitors are not well understood. Studying the adaptation mechanism to a particular inhibitor needs to be put into context as part of an overall response that may include more than one adaptation mechanism.

### ***3.6. Transcriptomic Studies on Insect Digestive Proteases***

In addition to studying specific enzyme/protein responses to protease inhibitors, larger-scale efforts to study the entire transcriptome of the gut of some insects have been made in an attempt to identify potential targets for pest control. Gene expression patterns in response to PIs have been studied in various lepidopteran insects (Bown *et al.*, 1997, 2004; Gatehouse *et al.*, 1997; Chougule *et al.*, 2005) and in some coleopterans (Michaud *et al.*, 1995; Zhu-Salzman *et al.*, 2003). Pedra *et al.* (2003) conducted a transcriptomic analysis of a cDNA library from the cowpea bruchid, *Callosobruchus maculatus* (Fabricius), which revealed digestive enzymes that could be targeted for future control strategies. In another study, the transcriptome of salivary glands from the adult female mosquito, *Anopheles* (Nyssorhynchus) *darlingi* was analyzed (Calvo *et al.*, 2004), specifically identifying secreted salivary products in order to further understand salivary gland functioning and salivary gland constitution. This information provided potential tools for the systematic analysis of molecules that may play an active role in blood feeding and the pathogenesis of malaria. Transcriptome information and analysis from midguts of the adult female biting midge, *Culicoides sonorensis* (Noriega *et al.*, 1994) revealed that of 1,719 serum-fed adult female midgut ESTs analysed, the most abundant functional group contained 600 ESTs that encoded putative proteins with proteolytic and peptidolytic functions (Campbell *et al.*, 2005). These experiments were conducted to help elucidate possible genetic determinants of arbovirus infection and to devise potential strategies to control vector-based infection. Xu *et al.* (2005) identified the molecular components of several *Anopheles* processes related to blood digestion, midgut expansion, and response to Plasmodium-infected blood, including digestive enzymes and other factors, using a microarray-based transcriptomic analysis. Hence, transcriptomic studies have provided a basis for the identification of new potential targets for pest/vector control.

### ***3.7. Use of an Inhibitor Combination Strategy***

The use of multiple PIs to inhibit the full spectrum of insect gut proteases may provide an effective control strategy for some pests. A combination of inhibitors targeting more than one protease class was successful in synergistically reducing the growth and causing mortality in the red flour beetle, *Tribolium castaneum* (Herbst) (Oppert *et al.*, 1993; 2003). A follow-up *in vitro* study of the *T. castaneum* gut proteases revealed a shift in the proteolytic enzyme profile when beetles were fed inhibitors of different protease classes, but an overall decrease in protease activities when beetles were fed a combination of inhibitors (Oppert *et al.*, 2005). Therefore, an appropriate combination of inhibitors can have a synergistic negative effect on growth and development. Similarly, Markwick *et al.* (1995) reported that the most significant reductions in the growth rate of larvae of the codling moth, *Cydia pomonella* (Linnaeus), occurred when combinations of protease inhibitors were fed in their diet. In general, proteolytic enzymes are transcriptionally regulated, and both the quantity and quality of dietary protein influence regulation (Liddle *et al.*, 1986; Tsuzuki *et al.*, 1991; Sharara *et al.*, 1993; Noreiga *et al.*, 1994). Insects often possess multiple protease genes encoding various protease isoforms with different sensitivities to a particular inhibitor. Undoubtedly, a complex of regulatory mechanisms at the level of transcriptional, translational, or enzymatic activation are involved in detoxification of the ingested inhibitors. Questions regarding the physiological pathway(s) responsible for the regulation of proteases in midguts of herbivorous insects remain unanswered (Broadway, 1997). Understanding these regulatory mechanisms will provide necessary information to improve the efficacy of protease inhibitors as insect control agents. Inhibitors effective against both ‘normal’ and induced proteases may need to be designed through multiple gene insertions to provide for effective protection.

## ***4. Tenebrio molitor Linnaeus (Yellow Mealworm)***

### ***4.1. Life cycle of Tenebrio molitor***

Insects can cause serious infestation problems in stored products that are of both plant and animal origin. Coleopteran pests belonging to the family Tenebrionidae occur in stored grain as well as in animal produce. The yellow mealworm, *Tenebrio molitor* L., is a freeze-intolerant omnivorous storage pest that damages grains, grain products, grain by-products, meat, and

feathers. Pests that infest animal products have been demonstrated to increase digestive protease and aminopeptidase enzyme activities (Baker, 1986).

*Tenebrio molitor* is one of the largest pests found in stored-grain products. The insect is indigenous to Europe, but is currently cosmopolitan in distribution. *T. molitor* is univoltine, with adults laying 200-300 eggs that are oval, opaque and shiny in appearance. The eggs hatch into yellowish, brown larvae that grow up to 30 mm in length and progress through several instars varying from 8 to more than 20 (Cotton and St George, 1929), over a period of approximately 3 months or more, depending on physical conditions and food source. In the spring, the larvae move to the surface of food, molt to the prepupal stage for a few days before becoming pupae. Pupae may also be observed on the surface of flour and are up to 16 mm in length with a row of characteristic lateral lamellae on each side. The pupal stage lasts approximately 3 weeks before pupae develop into black beetles in the summer. The length of the life cycle depends on food availability and the environmental conditions. The shortest period from the egg to adult stage is about 120 days. The longest period recorded has been about 629 days. Under optimal laboratory conditions (25°C, 60-70% RH, 16:8 L: D cycle) the time for completion of its life cycle is about 6 months.

## ***4.2. Tenebrio molitor Digestive Proteases***

### ***4.2.1. Occurrence and Midgut Compartmentalization of Serine and Cysteine Proteases***

*Tenebrio molitor* belongs to the coleopteran family Tenebrionidae, and its Cucujiformia ancestor was proposed to be a beetle adapted to ingest seeds rich in naturally occurring PIs by producing a cysteine protease for food digestion, instead of the common insect serine proteases (Terra and Cristofolletti, 1996). Aspartic proteases were found in the gut of insects in most families of Coleoptera including *T. castaneum* (Blanco-Labra *et al.*, 1996). However, further studies clarifying the origin, specificity and structure of insect cathepsin-D like enzymes, are necessary. *T. molitor* has been used as a model system to study insect proteases since 1964. Thie and Houseman (1990) were the first to report compartmentalization of both serine and cysteine proteases in the *T. molitor* midgut. Proteases from the cysteine and serine class were found in areas of the gut with optimal conditions for catalytic activity of the enzymes. Cysteine proteases were localized to the anterior midgut, and serine proteases were found in the posterior midgut. This was the first documented case of serine and cysteine proteases in an insect midgut. This

study provided the basis for further protein and molecular characterization of *T. molitor* digestive proteases. Terra and Cristofolletti (1996) later confirmed this compartmentalization, and proposed that the maximal activity of cysteine protease was in the acidic anterior midgut but that instability of the enzyme occurs in the posterior midgut, which has a higher pH. This compartmentalization was presumed to be advantageous to *T. molitor*, as the activity of cysteine proteases in the anterior midgut may inactivate ingested serine protease inhibitors prior to their contact with serine proteases in the posterior midgut.

#### **4.2.2. Early Studies on *Tenebrio molitor* Proteolytic Enzymes**

Applebaum *et al.* (1964) identified the proteolytic components trypsin, carboxypeptidase B and aminotripeptidase in the larval midgut of *T. molitor*. *T. molitor* larval and imaginal chymotrypsin, as well as imaginal trypsin have been isolated and described (Garty 1979, Ureili 1982, Golan 1981). These studies suggested that in *T. molitor* adult trypsin and chymotrypsin differ distinctly from comparable larval proteases, although both developmental stages inhabit the same milieu (Applebaum, 1985). Therefore, the genetic complement of *T. molitor* encodes information for trypsins and chymotrypsins that are differentially expressed in larvae and adults. A protocol developed by Zwillig (1968) for the purification of *T. molitor*  $\alpha$ - and  $\beta$ -proteases demonstrated their unusually high stability against inactivation. Zwillig *et al.* (1972) later identified a protease from adult *T. molitor* ( $\beta$ -protease) that lacked tryptic cleavage specificity and suggested that the enzyme was a “missing link” in protease evolution, based on differences in the structural properties of the catalytic site. However, the primary structural features homologous to the trypsin/chymotrypsin group were conserved. Levinsky *et al.* (1977) isolated and characterized a trypsin-like enzyme from the midgut of *T. molitor* larvae and concluded that *T. molitor* was the only stored-product pest with digestive trypsin, chymotrypsin and carboxypeptidase B. However, many insects with these enzyme spectra have since been described. Cristofolletti and Terra (2000) examined the thermodynamic and substrate-binding properties of a midgut microvillar aminopeptidase from *T. molitor*. Ferriera *et al.* (1990) detected the presence of digestive enzymes in secretory vesicles in the *T. molitor* larval midgut. Intracellular membrane-bound enzymes were secreted by exocytosis, including amylase from the anterior midgut, and carboxypeptidase and trypsin from the posterior midgut. Cristofolletti *et al.* (2001) elucidated an exocytic route for digestive enzyme secretion in *T. molitor* larvae, using immunocytolocalization studies, to demonstrate that protease-containing vesicles are discharged

by apocrine secretion from the anterior midgut, and that trypsin is released by exocytosis from the posterior midgut. A summary of the different *T. molitor* proteases and their properties is provided in Table 1.1.

#### ***4.2.3. Recent Studies on T. molitor Midgut Proteases***

A novel trypsin-like serine protease from the posterior midgut of *T. molitor* larvae, TmT1, was purified to homogeneity, and the N-terminus was identified as IVGGSSSISISSVPXQIXLQY (Tsybina *et al.*, 2005). The comparative N-terminal analysis also revealed that TmT1 contained 50-72% amino acid residue identity with other insect trypsin-like enzymes and 44-50% amino acid residue identity with mammalian trypsins. Similarly, the chymotrypsin-like protease, TmC1, also was isolated from the posterior midgut, purified to homogeneity, and the N-terminus was IISGSAASKGQFPWQ (Elpidina *et al.*, 2005).

An elaborate study of the spectrum of *T. molitor* larval digestive proteases has determined the distribution of total proteolytic activity in the midgut regions (Vinokurov *et al.*, 2006a & b, in press). The majority of trypsin-like protease activity, was due to three anionic and one cationic protease activities, and was mostly in the posterior midgut. Four anionic and one cationic chymotrypsin-like activities were found in the posterior midgut. Six different cysteine protease activities were detected, predominantly in the anterior midgut. These data further elucidate the complexity and compartmental regulation of digestive proteases in *T. molitor* larvae.

#### ***4.2.4. Molecular Studies on T. molitor Midgut Proteases***

Five cDNAs encoding procathepsin-L- (pCAL) like proteases were cloned and sequenced from the *T. molitor* larval midgut (Cristofolletti *et al.*, 2005). The characterization and immunolocalization of two of the major corresponding enzymes indicates that *T. molitor* has both lysosomal and digestive cathepsin-L like proteases (CALs) and that digestive CALs originate from lysosomal CALs by gene duplication and independent evolution. To date, there are only five digestive protease cDNAs from *T. molitor* available at GenBank (AY207373, AY332270, AY33271, AY337517 and AY332272).

## **5. Molecular Characterization of Digestive Proteases**

### ***5.1. Sequence Information on cDNAs Encoding Digestive Protease in Other Insects***

According to the review of Terra and Ferreira (2005), a total of 109 insect trypsin sequences (complete and incomplete), 63 complete sequences of insect chymotrypsins, and 12 cathepsin-L like sequences corresponding to eight species of coleopterans and hemipterans, have been registered in GenBank. The insect digestive carboxypeptidases A and B have been cloned and sequenced from Diptera (Ramos *et al.*, 1993; Bown *et al.*, 1998; Edwards *et al.*, 2000; Bown and Gatehouse, 2004), and over 20 midgut aminopeptidase cDNAs have been cloned from 10 lepidopteran species (Nakanishi *et al.*, 2002; Agrawal *et al.*, 2002; Rajagopal *et al.*, 2003).

### ***5.2. Benefits of Characterizing cDNAs Encoding Digestive Enzymes***

The molecular characterization of *T. molitor* digestive proteases is a prerequisite to understanding the nature and properties of individual digestive proteases. This information would be useful in the elucidation the compensatory response of *T. molitor* to PIs. Advances in molecular biological techniques have facilitated the validation of biochemical data and enzyme characterization. For example, cloning cDNA sequences encoding digestive enzymes enables the expression of recombinant enzymes that may used to synthesize antibodies, which can then be used in western blots or localization studies. RNAi studies or site-directed mutagenesis can also be used to evaluate and elucidate structure/function studies of digestive proteases. Microarrays and real-time quantitative PCR can be used to evaluate patterns of gene regulation.

## **6. Research Hypotheses and Objectives**

### ***6.1. Rationale and Hypotheses***

As was previously discussed, a single digestive enzyme protease inhibitor (PI) cannot be used against all herbivorous insects because an inhibitor may lose its potency due to changes in the relative proportion of digestive enzymes in the midgut. Thus, the success rates of PIs may be improved when combinations of inhibitors are expressed to cover the full spectrum of proteases in the gut. Therefore, this thesis is a study of the genetic complement encoding digestive proteases in the midgut of *T. molitor* larvae, as well as an evaluation of the response of these proteases to single and double PI treatments. After considering the previously reviewed

mechanisms of adaptation that insects have adopted in response to PIs, the following hypotheses are proposed:

1. Protease inhibitors have a negative impact on development of *T. molitor* larvae.
2. Cysteine proteases are upregulated in response to dietary serine protease inhibitors.
3. Serine proteases are upregulated as a response to dietary cysteine protease inhibitors.

Therefore, the general objective of this study was to understand the compensatory mechanism that *T. molitor* uses for adaptation to dietary protease inhibitors.

### **6.2. Research Objectives**

The specific objectives are four-fold:

1. To understand the basic digestive physiology of *T. molitor* larvae and to characterize biochemical properties of *T. molitor* digestive proteases.
2. To identify genes expressed in the gut of *T. molitor* larvae with an EST approach, and to characterize the cDNAs that encode putative digestive proteases.
3. To investigate the role of individual protease genes in food digestion by studying their temporal and spatial expression patterns.
4. To investigate and evaluate compensatory mechanisms of *T. molitor* larvae to dietary serine and cysteine protease inhibitors.

### **6.3. Scope of the Project**

The aforementioned objectives are proposed to explore the prospect of developing transgenic grain expressing protease inhibitors as stored grain biocontrol agents. The results and inferences from this study will broaden the basis of our understanding of the co-evolving interaction between PIs and insects. The information provided by this study may be applied to field pests as well. This research will help to understand the long-term complex compensatory regulatory mechanism to digestive protease inhibitors in coleopteran insects. Understanding the expressed protease profile in the midgut of *T. molitor*, which utilizes primarily two classes of proteases for digestion under laboratory dietary conditions, will significantly enhance further studies aimed developing new targets for coleopteran pest control based on PIs.

## 7. References

- Ahn, J-E., Salzman, R.A., Koiwa, H. and Zhu-Salzman, K. (2004) Functional roles of specific bruchid protease isoforms in adaptation to a soybean protease inhibitor. *Insect Mol. Biol.* 13: 649–657.
- Applebaum, S.W., Birk, Y., Harpaz, I. and Bondi, A. (1964) Comparative studies on proteolytic enzymes of *Tenebrio molitor* L. *Comp. Biochem. Physiol.* 11: 85-103.
- Applebaum, S.W. (1985) Biochemistry of digestion. In: Kerkot, G.A. and Gilbert, L.I., eds. *Comprehensive insect physiology; Biochemistry and Pharmacology*. New York, Pergamon Press, vol. 4, pp. 279-311.
- Agrawal, N., Malhotra, P. and Bhatnagar, R.K. (2002) Interaction of gene-cloned and insect cell-expressed aminopeptidase N of *Spodoptera litura* with insecticidal crystal protein Cry1C, *Appl. Environ. Microbiol.* 68: 4583-4592.
- Arthur, F.H., Hagstrum, D.W., Flinn, P.W., Reed, C.R., and Philipps, T.W. (2006) Insect populations in grain residues associated with commercial Kansas grain elevators. *J Stored Prod Res.* 42: 226-239.
- Barrett, A.J. (1986) The classes of proteolytic enzymes. In: Dalling, M.J ed. *Plant proteolytic enzymes*, Florida, CRC Press Inc., 1986, vol. 1, p. 1-16.
- Bayes, A., Comellas-Bigler, M., Rodriguez de la Vega, M., Maskos, K., Bode, W., Aviles, F.X., Jongmsa, M.A., Beekwilder, J. and Vendrell, J. (2005) Structural basis of the resistance of an insect carboxypeptidase to plant protease inhibitors. *Proc. Natl. Acad. Sci.* 102: 16602-16607.
- Bouchard, E., Cloutier, C. and Michaud, D. (2003) Oryzacystatin I expressed in transgenic potato induces digestive compensation in an insect natural predator via its herbivorous prey feeding on the plant. *Mol. Ecol.* 12:2439-2446.
- Baker, J. E. (1986) Amylase/proteinase ratios in larval midguts of ten stored-product insects. *Entomol. Exp. Appl.* 40: 41-46.
- Blanco-Labra, A., Martinez-Gallardo, N.A., Sandoval-Cardoso, L. and Delano-Frier., J. (1996) Purification and characterization of a digestive cathepsin D proteinase isolated from *Tribolium castaneum* larvae (Herbst). *Insect Biochem. Mol. Biol.* 26: 95-100.

Bolter, C. J. and Jongsma, M. A. (1995) Colorado potato beetles (*Leptinotarsa decemlineata*) adapt to proteinase inhibitors induced in potato leaves by methyl jasmonate. *J. Insect Physiol.* 41:1071-1078.

Bown, D.P., Wilkinson H.S. and Gatehouse J.A. (1997) Differentially regulated inhibitor-insensitive and insensitive protease genes from the phytophagous insect pest, *Helicoverpa armigera*, are members of complex multigene families. *Insect Biochem. Mol. Biol.* 27: 625-638.

Bown, D.P., Wilkinson H.S. and Gatehouse J.A. (1998) Midgut carboxypeptidase from *Helicoverpa armigera* (Lepidoptera: Noctuidae) larvae: enzyme characterisation, cDNA cloning and expression. *Insect Biochem. Mol. Biol.* 28: 739-749.

Bown, D.P., Wilkinson H.S. and Gatehouse J.A. (2004) Regulation of expression of genes encoding digestive proteases in the gut of a polyphagous lepidopteran larva in response to dietary protease inhibitors. *Physiol. Entomol.* 29: 278-290.

Bown, D.P. and Gatehouse, J.A. (2004) Characterization of a digestive carboxypeptidase from the insect pest corn earworm (*Helicoverpa armigera*) with novel specificity towards C-terminal glutamate residues. *Eur. J. Biochem.* 271: 2000-2011.

Bown, D.P., Wilkinson, H.S., Jongsma, M.A. and Gatehouse, J.A. (2004) Characterisation of cysteine proteinases responsible for digestive proteolysis in guts of larval western corn rootworm (*Diabrotica virgifera*) by expression in the yeast *Pichia pastoris*. *Insect Biochem. Mol. Biol.* 34:305-320.

Broadway, R. M. (1995) Are insects resistant to plant proteinase inhibitors? *J. Insect Physiol.* 41: 107-116.

Broadway, R. M. (1996a) Dietary proteinase inhibitors alter complement of midgut proteases. *Insect Biochem. Physiol.* 32: 39-53.

Broadway, R. M. (1996b) Resistance of plants to herbivorous insects: Can this resistance fail? *Can. J. Plant Pathol.* 18: 476-481.

Broadway, R. M. (1997) Dietary regulation of serine proteinases that are resistant to serine proteinase inhibitors. *J. Insect Physiol.* 43: 855-874.

Broadway, R.M. & Duffey, S.S. (1988) The effect of plant protein quality on insect digestive physiology and the toxicity of plant proteinase inhibitors. *J. Insect Physiol.* 34: 1111–1117.

- Broadway, R. M., Duffey, S. S., Pearce, G. & Ryan, C. A. (1986) Plant proteinase inhibitors: a defense against herbivorous insects? *Entomol. Exp. Appl.* 41: 33-38.
- Brito, L. O., Lopes, A.R., Parra, J. R. P., Terra, W. R. and Silva-Filho, M.C. (2001) Adaptation of tobacco budworm *Heliothis virescens* to proteinase inhibitors may be mediated by the synthesis of new proteinases. *Comp. Biochem. Physiol.* 128B: 365-375.
- Calvo, E., Andersen, J., De Francischetti, I.M.L.C.M., Debianchi, A.G., James, A.A., Ribeiro, J.M. and Marinotti, O. (2004) The transcriptome of adult female *Anopheles darlingi* salivary glands. *Insect Mol. Biol.* 13: 73–88.
- Campbell, C.L, Vandyke, K.A., Letchworth, G.J., Drolet, B.S., Hanekamp T. and Wilson, W.C. (2005) Midgut and salivary gland transcriptomes of the arbovirus vector *Culicoides sonorensis* (Diptera: Ceratopogonidae). *Insect Mol. Biol.* 14: 121-36.
- Conconi, A., Miquel, M., Browse, J.A. and Ryan C. A. (1996) Intracellular levels of free linolenic and linoleic acids increase in tomato leaves in response to wounding. *Plant Physiol.* 111: 797–803.
- Cotton, R.T. and St George, R.A. (1929) The meal worms. *Tech Bull. U.S. Dept. Agric.* 95: 1-37
- Chougule, N. P., Giri, A. P., Sainani, M. N. and Gupta, V.S. (2005) Gene expression patterns of *Helicoverpa armigera* gut proteases. *Insect Biochem. Mol. Biol.* 35: 355-367.
- Christeller, J.T., Farley, P. C., Ramsay, R. J., Sullivan, P.A., Laing, W.A. (1994) Midgut proteinase activities of three keratinolytic larvae, *Hofmannophila pseudospretalla*, *Tineola bisselliella*, and *Anthrenocerus australis*, and the effect of proteinase inhibitors on proteolysis. *Arch. Insect Biochem. Physiol.* 25: 159-173.
- Cloutier, C., C. Jean, M. Fournier, S. Yelle, and Michaud, D. (2000) Adult Colorado potato beetles, *Leptinotarsa decemlineata* compensate for nutritional stress on oryzacystatin I-transgenic potato plants by hypertrophic behavior and over-production of insensitive proteases. *Arch. Insect Biochem. Physiol.* 44: 69–81.
- Cloutier C, Fournier M, Jean C, Yelle S, Michaud, D. (1999) Growth compensation and faster development of Colorado potato beetle (Coleoptera: Chrysomelidae) feeding on potato foliage expressing oryzacystatin I. *Arch. Insect Biochem. Physiol.* 40: 69–79.

Cristofolletti, P.T. and Terra, W.R. (1999) Specificity, anchoring, and subsites in the active center of a microvillar aminopeptidase purified from *Tenebrio molitor* (Coleoptera) midgut cells. *Insect Biochem. Mol. Biol.* 29: 807-819.

Cristofolletti, P.T. and Terra, W.R. (2000) The role of amino acid residues in the active site of a midgut microvillar aminopeptidase from the beetle *Tenebrio molitor*. *Biochim. Biophys. Acta* 1479: 185-195.

Cristofolletti, P. T., Ribeiro, A. F. & Terra, W. R. (2001) Apocrine secretion of amylase and exocytosis of trypsin along the midgut of *Tenebrio molitor* larvae. *J. Insect Physiol.* 47: 143–155.

Cristofolletti, P.T., Ribeiro, A.F. and Terra, W.R. (2005) The cathepsin L-like proteinases from the midgut of *Tenebrio molitor* larvae: Sequence, properties, immunocytochemical localization and function. *Insect Biochem. Mol. Biol.* 35: 883-901.

De Leo F, Bonade-Bottino MA, Ceci LR, Gallerani R, Jouanin L (1998) Opposite effects on *Spodoptera littoralis* larvae of high expression level of a trypsin proteinase inhibitor in transgenic plants. *Plant Physiol.* 118: 997-1004.

De Leo F, Volpicella M, Licciulli F, Liuni S, Gallerani R, Ceci LR. (2002) PLANT-PIs: a database for plant protease inhibitors and their genes. *Nucleic Acids Res.* 30: 347-8.

Edmonds, H. S., Gatehouse, L. N., Hilder, V.A., and Gatehouse, J. A. (1996) The inhibitory effects of the cysteine protease inhibitor, oryzacystatin, on the digestive proteases and on the larval survival and development of the southern corn rootworm (*Diabrotica undecimpunctata howardi*). *Entomol. Exp. Appl.* 78: 83-94.

Edwards, M. J., Moskalyk, L. A., Donnelly-Doman, M., Vlaskova, M., Noriega, F.G., Walker, V.K. and Jacobs-Lorena M. (2000) Characterization of a carboxypeptidase A gene from the mosquito, *Aedes aegypti*. *Insect Mol. Biol.* 9: 33-38.

Elpidina, E. N., Tsybina, T. A., Dunaevsky, Y.E., Belozersky, M.A., Zhuzhikov, D.P. and Oppert, B. (2005) A chymotrypsin-like proteinase from the midgut of *Tenebrio molitor* larvae. *Biochimie* 87: 771-779.

Felton, G. W., Donato, K., del Vecchio, R. J. & Duffey, S. S. (1989) Activation of plant foliar oxidases by feeding reduces nutritive quality of foliage for noctuid herbivores. *J. Chem. Ecol.* 15: 2667-2694

Ferriera C., Bellinello, G.L., Ribeiro, A.F. and W. R. Terra. (1990) Digestive enzymes associated with glycocalyx, microvillar membranes and secretory vesicles from midgut cells of *Tenebrio molitor* larvae. *Insect Biochem.* 20: 839-847.

Franco, O.L., Dos Santos, R.C., Batista, J.A.N., Mendes, A.C.M., De Araujo, M.A.M., Monnerat, R.G., Grossa-se-Sa, M.F. and De Freitas, S.M. (2003) Effects of black-eyed trypsin/chymotrypsin inhibitor on the proteolytic activity and on the development on *Anthonomus grandis*. *Phytochem.* 63: 343-349.

Garty, N. (1979). Isolation and characterization of a chymotrypsin-like enzyme from *Tenebrio molitor* larvae. M.Sc. thesis, Faculty of Agriculture, The Hebrew University, Rehovot, Israel.

Gatehouse, L.N., Shannon, A.L., Burgess, E.P.J. and Christeller, J.T. (1997) Characterization of major midgut proteinase cDNAs from *Helicoverpa armigera* larvae and changes in gene expression in response to four proteinase inhibitors in the diet. *Insect Biochem. Mol. Biol.* 27: 929-944.

Girard, C., Le-Mateyar, M., Monade-Bottino, M., Pham-Delegue, M.H. and Jouanin, L. (1998) High level of resistance to proteinase inhibitors may be conferred by proteolytic cleavage in beetle larvae. *Insect Biochem. Mol. Biol.* 28: 229-237.

Giri A.P., Harsulkar A.M., Deshpande V.V., Sainani M.N., Gupta V.S., Ranjekar P.K. (1998) Chickpea defensive proteinase inhibitors can be inactivated by podborer gut proteinases. *Plant Physiol.* 116: 393-401

Golan, R. (1981) Isolation, characterization and comparative study of proteolytic enzymes from the midguts of *Tenebrio molitor* adults and larvae as a basis of possible biological pest control with naturally occurring protease inhibitors from plant sources. M.Sc. Thesis, Faculty of Agriculture, The Hebrew University, Rehovot, Israel.

Hagstrum, H.D., Reed, K.R. and Kenkel, P. (1999) Management of stored wheat insect pests in the U.S.A. *Integrated Pest Management Reviews* 4: 127-143.

Herrera-Estrella, L.R. (2000) Genetically Modified Crops and Developing Countries. *Plant Physiol.* 124: 923-926.

Hilder, V.A., Gatehouse A.M.R., Sheerman, S.E., Barker, R.F. and Boulter, D. (1987) A novel mechanism of insect resistance engineered from tobacco. *Nature* 330: 160-163.

Hilder, V.A. (1993) Plant genes for protecting transgenic crops from chewing and sap-sucking insect pests. *Workshop on Engineering Plants against Pests and Pathogens*, Instituto Juan March de Estudios e Investigaciones, Madrid, pp. 68-70

Johnston, K.A., Lee, M. J., Brough, C. L., Hilder V.A., Gatehouse A.M.R. and Gatehouse J. A. (1995) Protease activities in the larval midgut of *Heliothis virescens*: Evidence for trypsin and chymotrypsin-like enzymes. *Insect Biochem. Mol. Biol.* 25: 375–383.

Jongsma M.A., Bakker P.L., Peters J., Bosch D., and Stiekema W.J. (1995) Adaptation of *Spodoptera exigua* larvae to plant proteinase inhibitors by induction of gut proteinase activity insensitive to inhibition. *Proc. Natl. Acad. Sci. USA*, 92: 8041-8045

Jongsma, M.A., Stiekema, W.J., Bosch, D. (1996) Combatting inhibitor-insensitive proteases of insect pests. *Trends Biotechnol.* 14: 331-333

Jongsma, M.A. and Bolter, C. (1997) The adaptation of insect to plant protease inhibitors. *J Insect Physiol.* 43: 885-896

Kenkel, P., Criswell, J.T., Cuperus, G.W., Noyes, R.T., Anderson, K. and Fargo, W.S. (1994) Stored product integrated pest management. *Food Rev. Int.* 10: 177–93.

Konarev, A.V. (1996) Interaction of insect digestive enzymes with plant protein inhibitors from plants and host-parasite co-evolution. *Euphytica* 92: 89-94.

Korth, K. L., and Dixon, R.A. (1997) Evidence for chewing insect-specific molecular events distinct from a general wound response in leaves. *Plant Physiol.* 115: 1299-1305.

Krattiger, A. F. (1997) Insect resistance in crops: a case study of *Bacillus thuringiensis* (Bt) and its transfer to developing countries. *ISAAA Briefs*, 1997, no. 2, p. 42.

Laskowski, M. (1985) Protein inhibitors of serine proteinases: Mechanism and classification. In Friedman M (ed): *Nutritional and Toxicological Significance of Enzyme Inhibitors in Foods*. New York: Plenum Press, pp. 1-17.

Lawrence, P. K. and Koundal, K.R. (2002) Plant protease inhibitors in control of phytophagous insects. *Electron. J. Biotechnol.* 5: 93-109.

Levinsky, H., Birk, Y. and Applebaum, S.W. (1977) Isolation and characterization of a new trypsin-like enzyme from *Tenebrio molitor* L. larvae. *Int. J. Peptide Protein Res.* 10: 252-264.

Liddle, R. A., Green, G. M., Conrad C. K. and William J. A. (1986) Proteins but not aminoacids, carbohydrates, or fats stimulate cholecystokinin secretion in the rat. *Am. J. Physiol.* 251: G243-G248.

Lipke, H., Liener, I. E. and Fraenkel, G. (1954) Effects of soybean inhibitors on growth of *Tribolium confusum*. *J. Agri. Food Chem.* 2: 410-414.

Markwick, N. P., Reid, S. J., Laing, W. A. and Christeller, J. T. (1995) Effects of dietary protein and protease inhibitors on codling moth (Lepidoptera: Tortricidae). *J. Econ. Entomol.* 88: 33–39.

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

Michaud, D., Cantin, L. and Vrain, T.C. (1995) Carboxy-terminal truncation of oryzacysatin-II by oryzacytatin-insensitive insect digestive proteinases. *Arch. Biochem. Biophys.* 322: 469–474.

Moon, J., Salzman, R.A., Ahn, J-E., Koiwa, H. and Zhu-Salzman, K. (2004) Transcriptional regulation in cowpea bruchid guts during adaptation to a plant defence protease inhibitor. *Insect Mol. Biol.* 13: 283-291.

Nakanishi, K., Yaoi, K., Nagino, Y., Hara, H., Kitami, M., Atsumi, S., Muira, N. and Sato, R. (2002) Aminopeptidase N isoforms from the midgut of *Bombyx mori* and *Plutella xylostella*—their classification and the factors that determine their binding specificity of *Bacillus thuringiensis* Cry1Ac toxin, *FEBS Lett.* 519: 215–220.

Noriega, F.G., Barillas-Mury C. and Wells M. A. (1994) Dietary control of late trypsin gene transcription in *Aedes aegyptii*. *Insect Biochem. Mol. Biol.* 24: 627-631.

Oppert, B., Morgan, T.D., Culbertson, C. and Kramer, K. J. (1993) Dietary mixtures of cysteine proteinase and serine proteinase inhibitors exhibit increased toxicity toward the red flour beetle, *Tribolium castaneum*. *Comp. Biochem. Physiol.* 105C: 379-385.

Oppert, B. (2000) Transgenic plants expressing enzyme inhibitors and the prospects for biopesticide development. In *Phytochemical Biopesticides* (Koul, O. and Dhaliwal, G. S., eds.), Harwood Academic, The Netherlands, pp. 83-95.

Oppert, B., Morgan, T.D., Hartzer, K., Lenarcic, B., Galesa, K., Brzin, J., Turk, V., Yoza, K., Ohtsubo, K. and Kramer, K.J. (2003) Effects of proteinase inhibitors on growth and digestive

proteolysis of the red flour beetle, *Tribolium castaneum* (Herbst) (Coleoptera: Tenebrionidae). *Comp. Biochem. Physiol.* 134C: 481-490.

Oppert, B., Morgan, T.D., Hartzler, K. and Kramer, K. J. (2005) Compensatory proteolytic responses to dietary proteinase inhibitors in the red flour beetle, *Tribolium castaneum* (Herbst) (Coleoptera: Tenebrionidae). *Comp. Biochem. Physiol.*, 140C: 53-58.

Orr, G. L., Strickland J. A., Walsh T.A. (1994) Inhibition of *Diabrotica* larval growth by a multicystatin from potato tubers. *J. Insect Physiol.* 40: 893-900.

Ortego, F., Farinos, G. P., Ruiz, M., Marco, V. & Castanera, P (1998) Characterization of digestive proteases in the weevil *Aubeonymus mariafranciscae* and effects of proteinase inhibitors on larval development and survival. *Entomol. Exp. Appl.* 88: 265–274.

Patankar, A. G., Harshulkar, A. M., Giri, A. P., Gupta, V. S., Sainani, A. M., Ranjekar, P. K., and Deshpande, V. V. (1999). Diversity in inhibitors of trypsin and *Helicoverpa armigera* gut proteinases in chickpea (*Cicer arietinum*) and its wild relatives. *Theor. Appl. Genet.* 99: 719–726.

Patankar, A.G., Giri, A.P., Harsulkar, A.M., Sainani, M.N., Deshpande V.V., Ranjekar, P.K. and Gupta, V.S. (2001) Complexity in specificities and expression of *Helicoverpa armigera* gut proteinases explains polyphagous nature of the insect pest. *Insect Biochem. Mol. Biol.* 31: 453-464.

Pedra, J.H., Brandt, A., Westerman, R., Lobo, N., Li, H. -M., Romero-Severson, J., Murdock, L.L. and Pittendrigh, B.R. (2003) Transcriptome analysis of the cowpea weevil bruchid: Identification of putative proteinases and  $\alpha$ -amylases associated with food breakdown. *Insect Mol. Biol.* 12: 405–412.

Rajagopal, R., Agrawal, N., Selvapandiyan, A., Sivakumar, S., Ahmad, S. and Bhatnagar, R.K. (2003) Recombinantly expressed isoenzymic aminopeptidases from *Helicoverpa armigera* (American cotton bollworm) midgut display differential interaction with closely related *Bacillus thuringiensis* insecticidal proteins, *Biochem. J.* 370: 971–978.

Ramos, A., Mahowald, A. and Jacobs-Lorena, M. (1993) Gut-specific genes from the black fly *Simulium vittatum* encoding trypsin-like and carboxypeptidase-like proteins. *Insect Mol. Biol.* 1: 149-163.

Read, J.W. and Haas, L.W. (1938) The baking quality of flour as affected by certain enzyme actions. V. Further studies concerning potassium bromate and enzyme activity. *Cereal Chem.* 15: 59-68

Reeck, G.R., Kramer, K.J., Baker, J.E., Kanost, M.R., Fabrick, J.A and Behnke, C.A. (1997). Proteinase inhibitors and resistance of transgenic plants to insects. In: Carozzi, N and Koizel, M., eds. *Advances in insect control; the role of transgenic plants*. London, Taylor and Francis, pp. 157-183.

Richardson, M. (1977) The proteinase inhibitors of plants and micro-organisms *Phytochem.* 16: 159-169.

Richardson, M.J. (1991) Seed storage proteins: The enzyme inhibitors. In: Richardson, M.J, ed. *Methods in Plant Biochemistry*, eds. P.M. Dey and J. B. Harborne, p. 259-305. Academic Press, New York.

Ryan, C.A. (1990) Protease inhibitors in plants: Genes for improving defenses against insects and pathogens. *Annu. Rev. Phytopathol.* 28: 245-249.

Schaller A, and Ryan, CA. (1996) Systemin—a polypeptide defense signal in plants. *BioEssays* 18: 27–33.

Sharara, A. I., Bouras E. P., Misukonis M. A. and Liddle, R. A. (1993) Evidence for indirect dietary regulation of cholecystokinin release in rats. *Am. J. Physiol.* 265: G107-G112.

Storey, C. L., Sauer, D.B. and Walker, D. (1983) Insect populations in wheat, corn, and oats stored on the farm. *J. Econ. Entomol.* 76: 1323-1330.

Terra, W.R. and Ferreira, C. (1994) Insect digestive enzymes: properties compartmentalization and function. *Comp. Biochem. Physiol.* 109B: 1-62.

Terra, W.R. and Cristofolletti, P.T. (1996) Midgut proteinases in three divergent species of Coleoptera. *Comp. Biochem. Physiol.* 113B: 725-730.

Terra, W.R. and Ferreira, C. (2005) Biochemistry of digestion. In: L.I. Gilbert, K. Iatrov and S.S. Gill, Editors, *Comprehensive Molecular Insect Science, vol. 4, Biochemistry and Molecular Biology*, Elsevier, Oxford, pp. 171–224.

Thie, N.M.R. and Houseman, J.G. (1990) Cysteine and serine proteolytic activities in larval midgut of yellow mealworm, *Tenebrio molitor* L. (Coleoptera: Tenebrionidae). *Insect Biochem.* 20: 741-744.

Tsuzuki, S., Fushiki T., Kondo A., Muruyama H. and Sugimoto E. (1991) Effect of a high protein diet on the gene expression of a trypsin-sensitive, cholecystokinin-releasing peptide (monitor peptide) in the páncreas. *Eur. J. Biochem.* 199: 245-252.

Tsybina, T.A., Dunaevsky, Y.E., Belozersky, M.A., Zhuzhikov, D.P., Oppert, B. and Elpidina, E.N. (2005) Digestive proteinases of yellow mealworm (*Tenebrio molitor*) larvae: Purification and characterization of a trypsin-like proteinase. *Biochemistry (Mosc.)* 70: 300-305.

Urieli, N. (1982). Isolation and characterization of a chymotrypsin-like enzyme from the digestive tract of *Tenebrio molitor* and *Locusta migratoria*. In: Applebaum, S.W. (1985). *Biochemistry of digestion*. In: G.A. Kerkut and L.I. Gilbert, Editors, *Comprehensive Physiology, Biochemistry and Pharmacology of Insects*, Pergamon Press, Oxford, 4: pp. 279–311.

Vinokurov, K. S., Elpidina, E. N., Oppert, B., Prabhakar, S., Zhuzhikov, D. P., Dunaevsky, Y. E. and Belozersky, M. A. (2006a) Diversity of digestive proteinases in *Tenebrio molitor* (Coleoptera: Tenebrionidae) larvae. *Comp. Biochem. Physiol., Part B*, in press.

Vinokurov, K. S., Elpidina, E. N., Oppert, B., Prabhakar, S., Zhuzhikov, D. P., Dunaevsky, Y. E. and Belozersky, M. A. (2006b) Fractionation of digestive proteinases from *Tenebrio molitor* (Coleoptera: Tenebrionidae) larvae and role in protein digestion. *Comp. Biochem. Physiol., Part B*, in press.

Wang, P., Zhang, X. and Zhang, J. (2005) Molecular characterization of four midgut aminopeptidase N isozymes from the cabbage looper, *Trichoplusia ni*. *Insect Biochem. Mol. Biol.* 35: 611-620.

Wildon, D. C., Thain, J. F., Minchin, P. E. H., Gubb, I.R., Reilly, A. J., Skipper, Y. D., Doherty, H.M., O'Donnell, P.J., and Bowles, D.J. (1992) Electrical signaling and systemic proteinase inhibitor induction in the wounded plant. *Science* 360: 62-65.

Wu, Y., Llewellyn D., Matthews, A. and Dennis, E. S. (1997) Adaptation of *Helicoverpa armigera* (Lepidoptera: Noctuidae) to a proteinase inhibitor expressed in transgenic tobacco. *Mol. Breed.* 3: 371-380.

Xu, X., Dong, Y., Abraham, E.G., Kocan, A., Srinivasan, P., Ghosh, A.K., Sinden, R.E., Ribeiro, J.M., Jacobs-Lorena, M., Kafatos, F.C. and Dimopoulos, G. (2005) Transcriptome analysis of *Anopheles stephensi-Plasmodium berghei* interactions. *Mol. Biochem. Parasitol.* 142: 76-87.

Zhu-Salzman, K., Koiwa, H., Salzman, R.A., Shade, R.E. and Ahn, J.-E. (2003) Cowpea bruchid *Callosobruchus maculatus* uses a three-component strategy to overcome a plant defensive cysteine protease inhibitor. *Insect Mol. Biol.* 12: 135–145.

Zwilling, R. (1968) Zur Evolution der Endopeptidasen-IV.  $\alpha$ - and  $\beta$ -protease aus *Tenebrio molitor*. *Z. Physiol. Chem.* 349: 326-332.

Zwilling, R., Medugorac, I. and Mella, K. (1972) The evolution of endopeptidases-XIV. Non-tryptic cleavage specificity of a BAEE-hydrolyzing enzyme ( $\beta$ -protease) from *Tenebrio molitor*. *Comp. Biochem. Physiol.* 43: 419-424.

**Table 1.1. Table of proteases identified from *Tenebrio molitor***

	Protease Name	Source	Method of Isolation	Substrates Hydrolysed	Enzyme assay conditions	Inhibitors	Activators	Molecular Mass	pI	Optimal pH	Optimal temperature	Reference	Other information
1	Aminopeptidase	Microvilli of larval midgut membrane	Differential centrifugation, calcium precipitation, partial ultrasound disruption	LpNA, 1 mM	50 mM Tris-HCl, pH 7.8							Ferreira, C. et al., (1990) Digestive enzymes associated with the glycocalyx, microvillar membranes and secretory vesicles from midgut cells of <i>Tenebrio molitor</i> larvae. <i>Insect Biochem</i> 20: 839-847.	Membrane-bound, associated with microvilli. More active in the posterior midgut.
2	Aminopeptidase	Posterior third of larval midgut	*Papain-release from midgut cell membranes, chromatography on Alkyl-Superose, Mono Q, Phenyl-Superose (all FPLC); **CHAPS-solubilization from midgut cell membranes, preparative electrophoresis, Mono Q	Broad specificity towards N-terminal aminoacyl residue of peptides except of acidic residues	Routine: 1 mM LpNA, 100 mM Tris-HCl, pH 7.8	Amastatin - slowly tight-binding, bestatin and aminoacyl hydroxamates - rapidly reversible		*90000 **103000		8.0 stable at 5.0-10.8		Cristofolletti, P. T. and Terra, W. R. (1999) Specificity, anchoring, and subsites in the active center of a microvillar aminopeptidase purified from <i>Tenebrio molitor</i> (Coleoptera) midgut cells. <i>Insect Biochem Mol Biol</i> 29: 807-819.	A single membrane-bound AP in larval midgut. 55% of the microvillar proteins in posterior third of midgut. Glycoprotein rich in mannose. Is inserted into the microvillar membranes by a C-terminal anchor. Resembles mammalian aminopeptidase N (EC 3.4.11.2). Broad specificity with exception of acidic aminoacyl residues, prefers tri- and tetrapeptides relative to di- and pentapeptides. Has 4 subsites.

Table of proteases identified from *Tenebrio molitor* (contd.)

	Protease Name	Source	Method of Isolation	Substrates Hydrolysed	Enzyme assay conditions	Inhibitors	Activators	Molecular Mass	pI	Optimal pH	Optimal temperature	Reference	Other information
3	Aminopeptidase	Posterior third of larval midgut	Papain-release from midgut cell membranes, chromatography on Alkyl-Superose, Mono Q, Phenyl-Superose (all FPLC); sometimes CHAPS-solubilization from midgut cell membranes, preparative electrophoresis, Mono Q	LpNA	Routine: 1 mM LpNA, 100 mM Tris-HCl, pH 7.8	1,10-Phenantroline, K <sub>i</sub> 1.2 mM (2 mM causes 40% inhibition); EDTA (reversible) - is affected by pH (pK 5.8)				8.0		Cristofolletti, P.T. and Terra, W.R. (2000) The role of amino acid residues in the active site of a midgut microvillar aminopeptidase from the beetle <i>Tenebrio molitor</i> . <i>Biochim Biophys Acta</i> 1479: 185-195.	Is a metallopeptidase close to M1 family of peptidases with 1 Zn atom in the active site. Catalysis depends on a catalytic metal, a carboxylate and a protonated imidazole group, substrate binding relies in 1 phenol and 1 carboxylate groups. Resembles mammalian aminopeptidase N, but differs in details of substrate binding and in residues involved in catalysis. Isoamyl alcohol and sodium fluoride do not affect activity.
4	Aminotripeptidase (Exopeptidase)	Larval midgut	Column chromatography on ECTEOLA-cellulose and CM-cellulose	Leu-gly								Applebaum S.W. et al. (1964) Comparative studies on proteolytic enzymes of <i>Tenebrio molitor</i> L. <i>Comp Biochem Physiol</i> 11: 85-103.	SBTI-resistant.
5	Carboxypeptidase A	Larval posterior midgut	Differential centrifugation, calcium precipitation, partial ultrasound disruption	ZGlyPhe, 15 mM	50 mM Tris-HCl, pH 8.0, 50 mM NaCl							Ferreira, C. et al. (1990) Digestive enzymes associated with the glycocalyx, microvillar membranes and secretory vesicles from midgut cells of <i>Tenebrio molitor</i> larvae. <i>Insect Biochem</i> 20: 839-847.	Soluble. Glycocalyx-associated. More active in posterior midgut.

Table of proteases identified from *Tenebrio molitor* (contd.)

	Protease Name	Source	Method of Isolation	Substrates Hydrolysed	Enzyme assay conditions	Inhibitors	Activators	Molecular Mass	pI	Optimal pH	Optimal temperature	Reference	Other information
6	Carboxypeptidase B (Exopeptidase)	Larval midgut	Column chromatography on ECTEOLA-cellulose and CM-cellulose	Benz-gly-lys								Applebaum S.W. et al. (1964) Comparative studies on proteolytic enzymes of <i>Tenebrio molitor</i> L. <i>Comp Biochem Physiol</i> 11: 85-103.	SBTI-resistant.
7	Chymotrypsin	Larvae				DFP, SBTI, BBI, CI		23,400				Garty, N., 1979. Isolation and characterization of a chymotrypsin-like enzyme from <i>Tenebrio molitor</i> larvae. M.Sc. thesis, Faculty of Agriculture, the Hebrew University, Rehovot, Israel. Quoted in Applebaum, S.W., 1985. Biochemistry of digestion. In: Comprehensive Physiol., Biochem. and Pharmacol. of Insects (Edited by Kerkut, G.A. and Gilbert, L.I.), v. 4, pp. 279-311. Pergamon Press, Oxford.	10 half-cystines. Dissociates during SDS-PAGE into 3 subunits with M, 400, 11000 and 13000. Has higher affinity for ATEE than bovine chymotrypsin, but similar charge properties.

Table of proteases identified from *Tenebrio molitor* (contd.)

	Protease Name	Source	Method of Isolation	Substrates Hydrolysed	Enzyme assay conditions	Inhibitors	Activators	Molecular Mass	pI	Optimal pH	Optimal temperature	Reference	Other information
8	Chymotrypsin	Imago	Chromatography on CM-cellulose	ATEE, ATpNA		TPCK, PABA	Ca <sup>2+</sup>					Urieli, N., 1982. Isolation and characterization of a chymotrypsin-like enzyme from the digestive tract of <i>Tenebrio molitor</i> and <i>Locusta migratoria</i> . M.Sc. thesis, Faculty of Agriculture, the Hebrew University, Rehovot, Israel. Quoted in Applebaum, S.W., 1985. Biochemistry of digestion. In: Comprehensive Physiol., Biochem. and Pharmacol. of Insects (Edited by Kerkut, G.A. and Gilbert, L.I.), v. 4, pp. 279-311. Pergamon Press, Oxford.	Lacks half-cystines

Table of proteases identified from *Tenebrio molitor* (contd.)

	Protease Name	Source	Method of Isolation	Substrates Hydrolysed	Enzyme assay conditions	Inhibitors	Activators	Molecular Mass	pI	Optimal pH	Optimal temperature	Reference	Other information
9	Cysteine proteinase	Anterior half of larval midgut	Homogenization and centrifugation of peritrophic membrane contents	BANA	100 mM succinate, pH 5.0, 1 mM DTT, 1.5 mM EDTA	IAA, E-64 and leupeptin (which does not inhibit cathepsin H)	Thiol compounds DTT, cysteine, glutathione, mercaptoethanol			5		Thie, N. M. R. and Houseman, J. G. (1990) Cysteine and serine proteolytic activities in larval midgut of yellow mealworm, <i>Tenebrio molitor</i> L. (Coleoptera: Tenebrionidae). <i>Insect Biochem</i> 20: 741-744.	First case where cysteine and serine digestive proteinases are found together in an insect midgut. Anterior midgut contained 71.3% of cysteine proteinase activity. 93.6% was found in the gut contents compared to the gut wall. Cathepsin B-like enzyme. Unaffected by PMSF, aprotinin, pepstatin.
10	Cysteine proteinase	Posterior third or whole larval midgut	Chromatography on Superose 12 (FPLC), pH 7.0	BANA, 0.5 mM		E-64, 5 $\mu$ M	EDTA + cysteine	31,000		6.8		Terra W. R and Cristofoletti, P. T. (1996) Midgut proteinases in three divergent species of Coleoptera. <i>Comp. Biochem. Physiol.</i> 113B: 725-730.	Total cysteine proteinases activity predominates in anterior two thirds of the midgut, and trypsin and chymotrypsin - in posterior third. Hypothesis that cysteine proteinase undergoes autolysis at alkaline pH. <i>T.molitor</i> midgut lacks aspartic proteases.
11	Cysteine proteinase	Anterior two-thirds or whole larval midgut	Chromatography on Superose 12 (FPLC), pH 7.0	BANA, 0.5 mM	50 mM citrate-sodium phosphate, pH 7.0, 3 mM EDTA, 3 mM cysteine	E-64, 5 $\mu$ M	EDTA + cysteine	51,000		6.8		Terra W. R and Cristofoletti, P. T. (1996) Midgut proteinases in three divergent species of Coleoptera. <i>Comp. Biochem. Physiol.</i> 113B: 725-730.	Total cysteine proteinases activity predominates in anterior two thirds of the midgut, and trypsin and chymotrypsin - in posterior third. Hypothesis that cysteine proteinase undergoes autolysis at alkaline pH. <i>T.molitor</i> midgut lacks aspartic proteases.

Table of proteases identified from *Tenebrio molitor* (contd.)

	Protease Name	Source	Method of Isolation	Substrates Hydrolysed	Enzyme assay conditions	Inhibitors	Activators	Molecular Mass	pI	Optimal pH	Optimal temperature	Reference	Other information
12	Endopeptidase $\alpha$ -protease (EC 3.4.21.18)	Digestive tract of imago				PMSF, SBTI, LBTI		24,000				Zwilling, R. (1968) Zur Evolution der Endopeptidasen. IV. $\alpha$ - und $\beta$ -Protease aus <i>Tenebrio molitor</i> . <i>Hoppe-Seiler's Z. Physiol Chem</i> 349: 326-332.	Does not hydrolyze BAEE and ATEE. Is not inhibited by TLCK and TPCK.
13	Endopeptidase $\beta$ -protease	Digestive tract of imago	Gel filtration, paper electrophoresis, electrophoresis	BAEE, poly-lysine		SBTI, LBTI, BPTI, ovomucoid		60,000				Zwilling, R. et al. (1972) The evolution of endopeptidases. XIV. Non-tryptic cleavage specificity of a BAEE-hydrolyzing enzyme ( $\beta$ -protease) from <i>Tenebrio molitor</i> . <i>Comp Biochem Physiol</i> 43B: 419-424.	Certain similarities to pancreatic serine proteases is 'missing link', having changed by mutation certain properties of their catalytic site but conserved in primary structure. Features homologue to trypsin/chymotrypsin group of proteases. Lacks bovine trypsin cleavage specificity on the B-chain of oxidized insulin.
14	Serine proteinase (trypsin-like)	Posterior half of larval midgut	Homogenization and centrifugation of peritrophic membrane contents	BApNa	100 mM Tris-HCl, pH 8.0, 1 mM DTT, 3 mM EDTA	Aprotinin, SBTI, leupeptin, PMSF (5 mM); TLCK (0.5 mM) - slightly	Cysteine			8		Thie, N. M. R. and Houseman, J. G. (1990) Cysteine and serine proteolytic activities in larval midgut of yellow mealworm, <i>Tenebrio molitor</i> L. (Coleoptera: Tenebrionidae). <i>Insect Biochem</i> 20: 741-744.	First case where cysteine and serine (trypsin-like) digestive proteinases are found together in an insect midgut. Posterior midgut contained 66.9% of trypsin-like activity. 90.2% was found in the gut contents compared to the gut wall. Unaffected by pepstatin, IAA or E-64.

Table of proteases identified from *Tenebrio molitor* (contd.)

	Protease Name	Source	Method of Isolation	Substrates Hydrolysed	Enzyme assay conditions	Inhibitors	Activators	Molecular Mass	pI	Optimal pH	Optimal temperature	Reference	Other information
15	Trypsin	Larval posterior midgut	Differential centrifugation, calcium precipitation, partial ultrasound disruption	BApNA, 0.8 mM	50 mM Tris-HCl, pH 8.0							Ferriera, C. et al. (1990) Digestive enzymes associated with the glycocalyx, microvillar membranes and secretory vesicles from midgut cells of <i>Tenebrio molitor</i> larvae. <i>Insect Biochem</i> 20: 839-847.	Soluble. Glycocalyx-associated. More active in posterior midgut.
16	Trypsin	Posterior third of larval midgut contents	Chromatography on Econo Pac High Q (Bio-Rad), pH 7.0, Mono Q (FPLC), pH 9.0	BApNA, 0.83 mM	100 mM Tris-HCl, pH 7.5			25,000				Cristofolletti, P.T. et al. (2001) Apocrine secretion of amylase and exocytosis of trypsin along the midgut of <i>Tenebrio molitor</i> larvae. <i>J of Insect Physiol</i> 47: 143-155.	Pure preparation was isolated with a yield of 15% and specific activity 0.97 U/mg protein. Antibodies to purified trypsin showed only 1 band in posterior midgut contents after Western blotting. Secreted by exocytosis in posterior third of midgut and this is efficient secretory mechanism in a water-secreting epithelium. Occurs as a precursor in posterior midgut cells.

Table of proteases identified from *Tenebrio molitor* (contd.)

	Protease Name	Source	Method of Isolation	Substrates Hydrolysed	Enzyme assay conditions	Inhibitors	Activators	Molecular Mass	pI	Optimal pH	Optimal temperature	Reference	Other information
17	Trypsin	Imago midgut	Chromatography on DEAE-cellulose, affinity chromatography on PABA-Sepharose	BApNA, TAME		TLCK, PABA (slow rate), BBI, CI		16 500 (SDS-PAGE), 18400 (Sephadex G-50)				Golan, R., 1981. Isolation, characterization and comparative study of proteolytic enzymes from the midguts of <i>Tenebrio molitor</i> adults and larvae as a basis of possible biological pest control with naturally occurring protease inhibitors from plant sources. M.Sc. thesis, Faculty of Agriculture, the Hebrew University, Rehovot, Israel. Quoted in Applebaum, S.W., 1985. Biochemistry of digestion. In: Comprehensive Physiol., Biochem. and Pharmacol. of Insects (Edited by Kerkut, G.A. and Gilbert, L.I.), v. 4, pp. 279-311. Pergamon Press, Oxford	8 tryptophan residues. No half-cystine or methionone. pI is higher than of larval trypsin.

Table of proteases identified from *Tenebrio molitor* (contd.)

	Protease Name	Source	Method of Isolation	Substrates Hydrolysed	Enzyme assay conditions	Inhibitors	Activators	Molecular Mass	pI	Optimal pH	Optimal temperature	Reference	Other information
18	Trypsin (Endopeptidase)	Larval midgut	Column chromatography on ECTEOLA-cellulose and CM-cellulose	Polylysine		DFP, BBI, SBTI, LBTI				6.3-6.5	37°C	Applebaum S.W. et al. (1964) Comparative studies on proteolytic enzymes of <i>Tenebrio molitor</i> L. <i>Comp Biochem Physiol</i> 11: 85-103.	Results substantiate presence of a trypsin-like enzyme in the digestive system of <i>Tenebrio</i> . Proteolytic activity of midgut wall is distinct from that of midgut lumen in its resistance to inhibition by SBTI. Effect of inhibitors depends on their relative concentration per enzyme and combination of pH, temperature and length of incubation. Significance of sulphhydryl group for the integrity of proteolytic activity is due to effect of thiol compounds, such as cysteine and glutathione. SBTI-susceptible. Striking overall similarity between <i>Tenebrio</i> trypsin and bovine trypsin.
19	Trypsin-like enzyme (TLE)	Larval midgut	Chromatography on DEAE-cellulose, (pH 6.5, hold-up volume), rechromatography	BApNA, 50 mM; TAME, 15 mM; poly-L-lysine hydrobromide	For BApNa 140 mM Tris-HCl, pH 8.0, 25 mM CaCl <sub>2</sub>	DFP, TLCK, BBI (K <sub>i</sub> 5.87*10 <sup>-7</sup> ), BPTI (K <sub>i</sub> 7.92*10 <sup>-7</sup> ), CI, SBTI		24 300 (SDS-PAGE) 18 300 (ultra-centrifugation)	8	8	50°C	Levinsky, H. et al. (1977) Isolation and characterization of a new trypsin-like enzyme from <i>Tenebrio molitor</i> L. larvae. <i>Int J Peptide Protein Res</i> 10: 252-264.	A <sup>1%</sup> <sub>1cm</sub> = 24.1 at 280 nm. No carbohydrates in the molecule. No tryptophan, 4 half-cystine residues, N-terminal isoleucine, C-terminal asparagine or threonine. K <sub>m</sub> with BApNA 0.93 mM, with TAME 0.08 mM. Stable at neutral pH. Unstable in 4 M urea and at -20° in 1 mM HCl. Unaffected by Ca <sup>2+</sup> and by 10 mM DTT. Different in conformation from bovine trypsin, no common antigenic determinants.

Table of proteases identified from *Tenebrio molitor* (contd.)

	Protease Name	Source	Method of Isolation	Substrates Hydrolysed	Enzyme assay conditions	Inhibitors	Activators	Molecular Mass	pI	Optimal pH	Optimal temperature	Reference	Other information
20	Trypsin-like proteinase	Posterior midgut	ion-exchange chromatography on DEAE-Sephadex A-50 and gel filtration on Superdex-75	Bz-Arg-pNA	The Km values determined for Bz-Arg-pNA and Bz-Lys-pNA substrates were 0.04 and 0.09 mM, respectively	PMSF, and specific trypsin inhibitor, TLCK		25.5 kD	7.4	8.5	55°C	Tsybina, T.A., Dunaevsky, Y.E., Belozersky, M.A., Zhuzhikov, D.P., Oppert, B. and Elpidina, E.N. (2005) Digestive proteinases of yellow mealworm ( <i>Tenebrio molitor</i> ) larvae: purification and characterization of a trypsin-like proteinase. <i>Biochemistry</i> (Mosc.) 70: 300-305.	The enzyme hydrolyzes peptide bonds formed by Arg or Lys residues in the P1 position with a preference for relatively long peptide substrates. The N-terminal amino acid sequence, IVGGSSISISSVPXQIXLQY, shares 50-72% identity with other insect trypsin-like proteinases, and 44-50% identity to mammalian trypsins.

Table of proteases identified from *Tenebrio molitor* (contd.)

	Protease Name	Source	Method of Isolation	Substrates Hydrolysed	Enzyme assay conditions	Inhibitors	Activators	Molecular Mass	pI	Optimal pH	Optimal temperature	Reference	Other information
21	Chymotrypsin-like proteinase	Posterior midgut	ion-exchange and gel filtration chromatography	SucAAPFpNA, SucAAPLpNA and GlpAALpNA	Kinetic parameters of the enzymatic reaction demonstrated that the best substrate was SucAAPFpNA, with $k(\text{cat app})$ 36.5 s <sup>-1</sup> and $K(\text{m})$ 1.59 mM.	PMSF		23.0 kDa	8.4	9.5	51°C	Elpidina, E.N., Tsybina, T.A., Dunaevsky, Y.E., Belozersky, M.A., Zhuzhikov, D.P. and Oppert, B. (2005) A chymotrypsin-like proteinase from the midgut of <i>Tenebrio molitor</i> larvae. <i>Biochimie</i> 87: 771-779..	The activity of TmC1 was reduced with sulfhydryl reagents. Several plant and insect proteinaceous proteinase inhibitors were active against the purified enzyme, the most effective being Kunitz soybean trypsin inhibitor (STI). The N-terminal sequence of the enzyme was IISGSAASKGQFPWQ, which was up to 67% similar to other insect chymotrypsin-like proteinases and 47% similar to mammalian chymotrypsin A. The amino acid composition of TmC1 differed significantly from previously isolated <i>T. molitor</i> enzymes. The proteinase displayed high stability at temperatures below 43 degrees C and in the pH range 6.5-11.2, which is inclusive of the pH of the posterior and middle midgut.

\* Recently, Vinokurov *et al.* (2006a, in press) have described one cationic and three anionic proteinases in the posterior midgut of *T. molitor* that are responsible for trypsin-like activity, one cationic and four anionic proteinases with chymotrypsin-like activity.

**Abbreviations:**

**ATEE** - N-acetyl-L-tyrosine ethyl ester

**ATpNA** - N-acetyl-L-tyrosine-para-nitroanilide

**BAEE** - N-benzoyl-L-arginine ethyl ester

**BBI** - Bowman-Birk soybean inhibitor

**BANA** - N-alpha-benzoyl-L-arginine naphthylamide

**BApNA** - N-alpha-benzoyl-DL-arginine-para-nitroanilide

**BPTI** - bovine pancreatic trypsin inhibitor

**CHAPS** - 3-[(cholamidopropyl)-dimethylammonio]-1-propan sulfonate

**DFP** - diisopropylphosphofluoridate

**DTT** - dithiothreitol

**E-64** - *trans*-epoxysuccinyl-L-leucyl-amido(4guanidino)butane

**ECTEOLA** - anion exchange adsorbent epichlorhydrin triethanolamine

**IAA** - Iodoacetamide

**LBTI** - Lima bean trypsin inhibitor

**LpNA** - L-leucine -para-nitroanilide

**PABA** - para-aminobenz-amidine

**PMSF** - phenylmethyl sulfonyl fluoride

**SBTI** - soybean trypsin inhibitor

**TAME** - tosyl-L-arginine-methyl-ester

**TLCK** - N-tosyl-L-lysine chloromethyl ketone

**TPCK** - N-tosyl-L-phenylalanine chloromethyl ketone

**ZGlyPhe** - N-carbobenzoxy-glycyl-L-phenylalanine

## **CHAPTER 2 - BIOCHEMICAL STUDY OF THE DIGESTIVE PHYSIOLOGY OF THE YELLOW MEALWORM, *TENEBRIO MOLITOR* L.**

### **2.1. Abstract**

*Tenebrio molitor* is a univoltine stored product pest causing a significant economic loss in food production worldwide. Larvae of *T. molitor* have a unique digestive physiology and utilize a complex of digestive proteases, including those from the serine and cysteine classes. The pH environment in the larval midgut varies in different regions. The anterior midgut is acidic and optimal for cysteine protease activity; the posterior part is alkaline and optimal for serine protease activity. Using inhibitors of specific proteases, the activity of the various classes of proteases was measured in gut extracts from *T. molitor* larvae. Protease activity was susceptible to inhibition by SBTI, PMSF, TPCK, chymostatin, suggestive of serine proteases. Inhibition by E-64 in reducing buffers suggested the presence of cysteine proteases in the *T. molitor* midgut. These results define and establish reliable experimental conditions for future studies involving the regulatory and compensatory responses of *T. molitor* larvae to dietary protease inhibitors. The results have implications in improving the efficacy of biopesticides based on protease inhibitors for the control of coleopteran pests.

Key words:

Yellow mealworm, serine proteases, cysteine proteases, compensation, pH, enzyme localization

## 2.2. Introduction

### 2.2.1. *Proteases*

Food digestion is facilitated by a wide array of enzymes that break down the carbohydrate, fat and protein content of food into smaller units. Proteases are defined as those enzymes that catalyze the hydrolytic breakdown of proteins into amino acids or peptides. Proteases, enzymes that hydrolyze peptide bonds and disassemble proteins, have been isolated from plants, animals, bacteria and fungi. Proteases are extremely important, as more than 60% of the total commercial enzyme market is made up of proteases. The sub-subclasses of peptidases or proteases are the exopeptidases and the endopeptidases (Barrett, 1994). Proteases that hydrolyze amino acids sequentially from the ends of polypeptide chains are exopeptidases. Aminopeptidases are examples of proteases acting at a free N-terminus that liberate a single amino acid residue, a dipeptide or a tripeptide. They are classified on the basis of their dependence on metal ions (usually  $Zn^{2+}$  or  $Mn^{2+}$ ) and substrate specificity. Carboxypeptidases are an example of exopeptidases acting at a free C-terminus liberating a single residue or a dipeptide.

Endopeptidases (sometimes referred to as “proteinases”) act preferentially in the inner regions of peptide chains, away from the termini, as the presence of free  $\alpha$ - amino or  $\alpha$ - carboxyl groups has a negative effect on the activity of the enzyme. These include enzymes such as pepsin, trypsin or papain, and they catalyze the splitting of proteins into smaller peptide fractions and amino acids by a process known as proteolysis. There are four mechanistic classes of endopeptidases: aspartic, cysteine, serine, and amino (Barrett, 1986).

Proteases have optimal activity in a specific pH environment or with specific reagents. Serine proteases have conserved sequences and regions around the residues that form the so-called catalytic triad, those residues most directly involved in catalytic activity, and these include His57, Asp102 and Ser195 (chymotrypsin numbering) (Barrett and Rawlings, 1991). Serine proteases are inactivated irreversibly by diisopropylphosphofluoridate (DFP) or phenylmethyl sulfonyl fluoride (PMSF), or by ketones like L-1-chloro-3-(4 tosyl amido)-4 phenyl-2-butanone (TPCK) or L-1-chloro-3-[4-tosyl-amido]-7 amino-2 heptanone-hydrochloride (TLCK). Trypsins are serine proteases that exist as an inactive precursor and are processed to a soluble form before being secreted into the lumen midgut. They preferentially cleave protein chains on the carboxyl

side of basic L-amino acids such as arginine or lysine. Trypsin is specifically inhibited by TLCK interaction with a histidine residue in the active site (Shaw *et al.*, 1965). Chymotrypsins preferentially cleave protein chains on the carboxyl side of aromatic amino acids (Terra and Ferriera, 1994). The enzyme is specifically inhibited by TPCK, which also interacts with a specific histidine. Cysteine proteases (EC 3.4.22) include cathepsins B, H, L and S and have striking conservation of structure around the Cys25 and His159 (papain numbering) residues that are directly involved in catalysis. In insects, cysteine proteases are used in digestive processes (Rawlings and Barrett, 1994), but also are found in several other tissues and play other roles (Kanost and Clarke, 2005; Yamamoto *et al.*, 1994). Cysteine proteases are processed by removing the amino-terminal fragment to produce the active enzyme. The pro-region, or the amino-terminal region, plays important roles not only as inhibitors of enzyme activity but also in the correct folding of the newly synthesized protein, to protect it against denaturing effects in sudden changes in pH conditions (Oliviera *et al.*, 2003). Cathepsin L-like enzymes and cathepsin B-like enzymes are subfamilies of the papain family. The best-known animal cysteine protease is cathepsin B (EC 3.4.22.1.) (Barrett 1977a). Cysteine proteases are inhibited by mercurial compounds. Cystatins, cysteine protease inhibitors, include proteins that inhibit the activity of papain and related cysteine proteases (cathepsin B, H and L, ficin, Bromelain). The knowledge of the properties of a protease are important, especially since most organisms produce a number of different proteases, and it is essential that their activities be distinguished to understand biological activity (North, 1982). Thie and Houseman (1990) suggested that coleopteran species could be subdivided into at least three different groups, based on digestive proteolysis by either serine or cysteine proteases, or a combination of both. They reported *T. molitor* as the first case where both serine and cysteine proteases are found together in an insect midgut; however, we now know that other coleopterans have this type of digestion, although the relative activities vary. Protein digestion in *T. molitor* occurs partly in the lumen and partly at the cell surface of the posterior midgut (Ferriera *et al.*, 1990). Proteolytic enzymes are the most abundant in the lumen of the *T. molitor* larval midgut (Applebaum *et al.*, 1964). The names of the classes of proteases reflect the key catalytic moiety (an amino acid) participating in the cleavage of peptide bonds (Wolfson and Murdock, 1990). In practice, the mechanistic class to which a protease belongs is based on the following *in vitro* characteristics: 1. pH range over which it is maximally active; 2. sensitivity to various inhibitors; 3. ability to hydrolyze specific proteins or peptides;

and 4. similarity to well-characterized proteases (North, 1982; Barrett, 1986; Wagner, 1986). The susceptibility of midgut homogenates to inhibition can add information to the prediction of the type of protease as suggested by the pH optimum curves (Wolfson and Murdock, 1990). pH optima and inhibitor assays can be determined in the presence of a reducing agent that usually enhance the activity of cysteine proteases. The pH of gut contents is one of the important internal environmental properties that affects digestive enzymes (Terra and Ferreira, 1994).

### ***2.2.2. Rationale***

Insects are excellent models for studying gut function, as their feeding habits have allowed them to adapt to almost all kinds of habitats. This understanding has been used in the development of transgenic plants to control phytophagous insects. The digestive proteases of insects catalyze the release of free amino acids from dietary protein and thereby provide a supply of nutrients essential for normal growth and development (Wolfson and Murdock, 1990). As pointed out by Wolfson and Murdock, information on insect digestive proteolytic activity in a given insect species must exist, before the possible importance of protease inhibitors, either induced or constitutive, is considered a factor in aberrant growth patterns. Characterizing a protease by mechanistic class establishes experimental conditions to study properties of the enzyme that is unique to its biological function (Barrett, 1986). A study of gut proteases and the biochemical environment along the midgut is crucial to planning further molecular studies of spatial protease gene expression. The pH of the gut environment is one of the physiological conditions that determines the activity of digestive enzymes. Terra and Ferreira (1994) stated that the organization of the digestive process depends on compartmentalization of digestive enzymes and on midgut fluid fluxes that are responsible for the translocation of enzymes and products of digestion. Attempts were made to separate the complex of proteolytic enzymes into individual components as a prerequisite to the elucidation of mechanism of proteolytic digestion. Applebaum *et al.*, 1964 identified three distinct proteolytic components based on selective inhibition by specific inhibitors. The effectiveness of protease inhibitors as insecticides depends on the affinity or specificity of an inhibitor to the main gut proteases of the insect (Burgess *et al.*, 1991; Gatehouse *et al.*, 1993; McManus *et al.*, 1994; Jongsma *et al.*, 1995). Therefore, a comprehensive study of the insect digestive proteases and their mechanistic classes, and the effect of protease inhibitors is essential.

The purpose of this study is to understand the biochemical aspect of gut physiology of larval *T. molitor* by testing several parameters. Preliminary *in vitro* experiments have standardized conditions for biochemical assays, including microplate assays and pH curves, inhibitor assays, activity blots and zymograms. Results revealed the overall proteolytic activities in *T. molitor* larval midguts (both anterior and posterior) and the biochemical nature and properties of the gut environment.

## **2.3. Materials and Methods**

### ***2.3.1. Isolation of Gut Enzymes***

Stock cultures of the yellow mealworm, *Tenebrio molitor*, were collected from eastern Kansas in 1970 and have since been reared at a relative humidity of 60-70% at 25°C under photoregime conditions of a 16:8 (L:D) photoperiod. Insects were transferred from a wheat flour diet to a coarse oatmeal diet two days before biochemical studies. They were then allowed to feed overnight on a wet diet. Fully grown larvae of each sex weighing  $1.131 \pm 0.025$  g (n=20) were used. Actively feeding late instar larvae were immobilized on ice, after which they were dissected by excising the anterior and posterior ends and by pulling out the gut in cold 342 mM NaCl. Only samples from actively feeding larvae (i.e. those with full guts) were taken. The foregut was about 3%, midgut 67%, and the hindgut 30% of the entire gut in length. The whole midgut was removed and divided into anterior (AM) and posterior (PM) sections of identical length by cutting midway between the most anterior region and the insertion of the Malpighian tubules. Since the size of the midgut and the biochemical characteristic necessitated homogenization of subsamples, the supernatants of the subsamples were pooled, mixed and divided into aliquots prior to freezing. The tissue was then manually homogenized with a pestle, and samples were pooled, vortexed, centrifuged in a Hermle Z230 MA centrifuge at 15,000 x g for 5 min, and the supernatant was either used immediately or stored at -20°C for proteolytic assays. Protease activity in each midgut was diluted to be within the linear range of the assay.

### ***2.3.2. Microplate Assay-pH Curve***

A pH curve based on the microplate assay method of Oppert *et al.* (1996) was used to evaluate the hydrolysis of substrate and effects of different inhibitors on the crude gut extract of *T. molitor*. For substrate analysis, 2 µl of the *T. molitor* gut extract containing 0.18 gut

equivalents of AM and 0.30 gut equivalents of PM were diluted into buffers ranging from acidic to basic, using a universal buffering system (Frugoni, 1957). The universal buffers of increasing pH within a range of 1-10 were made using varying volumes of stock solution (54 ml H<sub>3</sub>PO<sub>4</sub>, 45.8 ml H<sub>3</sub>OAc, 49.6 g H<sub>3</sub>BO<sub>3</sub>), 2N NaOH, H<sub>2</sub>O and 0.01M KCl. In a black 96-well microplate, 88 µl of buffer with 10% distilled water (non-reducing medium) and 2 µl of the gut extract was added to first 4 rows (A-D) with pH increasing from left to right. In rows E-H 88 µl of buffer with 10% 5mM L-cysteine (reducing medium) and 2 µl of the gut extract was added in the same ascending order of pH. Ten µl of the substrate was added to rows A-H and the plate was covered. The process of adding the substrate was done in dim light, as the fluorescently labeled casein (BODIPY TR-X) substrate is photosensitive. In reducing buffers, the covered tray was pre-incubated at 37°C, for at least 15 min before adding the substrate. Fluorescence of dye-labeled peptides was measured at time intervals of 30, 60, 120, 180 and 240 min. Total proteolytic activity was measured using fluorescently labeled casein (BODIPY TR-X casein, Molecular Probes' Enz Chek™ Protease Assay Kits, Eugene, OR) to detect protease activity. Protease-catalyzed hydrolysis releases highly fluorescent BODIPY TR-X dye-labeled peptides. The accompanying increase in fluorescence was measured with a microplate reader and was proportional to protease activity. A 10 µg/ml working solution of the BODIPY-caesin was prepared. BODIPY TR-X has excitation and emission maxima of approximately 590±10 nm and 645±10 nm respectively. The instrument was a Fluoroskan Ascent FL microplate reader (Thermo Electron Corporation, Milford, MA) with a highly focused light beam for fluorometric measurements. The output data was copied onto and analyzed in a Microsoft Excel (Redmon, WA) spreadsheet. There were four replicates for each pH curve (with and without L-cysteine) and the microplate assay was duplicated at two different times using a new pool of guts.

### ***2.3.3. pH profile Comparison of T. molitor Larvae Reared on Wheat Flour and Wheat Germ***

To examine the effect of diet on differences in enzymatic activity, midgut homogenates of guts that were fed on wheat flour and wheat germ, were used in a microplate assay. The proteolytic activity of gut extracts of insects that fed on wheat flour and germ in the presence and absence of 5mM L-cysteine was measured in four replicates using a buffer with pH 8.5, and fluorescence was measured at 6 time points ranging from 0 time to 4 h after incubation at 37°C.

#### ***2.3.4. Effect of Variability of Gut Samples on Total Proteolytic Activity***

To examine the effect of temperature and storage conditions of midgut extracts on their enzymatic activity, microplate assays were done using freshly dissected guts and frozen guts. The experiment was repeated using one set of pooled AM extracts and one set of pooled PM samples, stored at -80°C. Results of these assays were compared to those using freshly dissected guts. In a similar experiment, guts that were freshly dissected or frozen stored in either distilled water or physiological buffered saline were used in a microplate assay.

#### ***2.3.5. Inhibitor Assays***

The effect of various protease inhibitors on gut proteolytic activity was determined. For inhibitor assays, midgut extracts were incubated at 37°C, with or without 5 mM L-cysteine, and with increasing dilutions of inhibitor (and a no inhibitor control) for 15 min in universal buffer (pH 8.5) in reducing or non-reducing media (Frugoni, 1957). Protease inhibitors were as follows (Appendix A): aprotinin, chymostatin, E-64 (trans-epoxysuccinyl-1-leucylamido [4-guanidino] butane), leupeptin, pepstatin, phenylmethyl sulfonyl fluoride (PMSF), TLCK, (L-1-Chloro-3-[4-tosyl-amido]-7 amino-2 heptanone-hydrochloride); TPCK, (L-1-Chloro-3-(4 tosyl amido)-4 phenyl-2-butanone), soybean Bowman-Birk trypsin-chymotrypsin inhibitor (SBTI) (Boehringer Mannheim, Indianapolis, IN and Sigma Chemical Co., St Louis, MO). Means were separated using LSD and were analyzed statistically using PROC GLM (SAS Institute, 1985).

#### ***2.3.6. Activity Blots***

An activity blot assay of the crude gut extract was performed using a method adapted from Ohlsson *et al.* (1985) as described by Oppert and Kramer, 1998. Briefly, soluble extracts from the entire gut were subjected to sodium dodecyl sulphate (SDS) polyacrylamide gel electrophoresis (Laemlii, 1970). Following electrophoresis and electroblotting, nitrocellulose blots were incubated with a substrate solution (a p-nitroanilide-conjugated substrate, N-succinyl-ala-ala-pro-phe-p-nitroanilide (SAAPFpNA) at a concentration of 0.5 mg/ml in 0.1 M Tris, pH 8.1, 0.02M CaCl<sub>2</sub>) by placing blots in an Econoblot (LabLogix, Belmont, CA) tray and layering 5 ml of substrate solution directly over the blot. Blots were then covered with an Econoblot plastic sheet and incubated at 37°C until a faint yellow color appeared (15-30 min) depending upon the amount of protease activity in the sample. Liberated nitroanilide was diazotized for visualization by subsequent incubations for 5 min each in 0.1% sodium nitrite in 1 M HCl, 0.5%

ammonium sulphamate in 1 M HCl, and 0.05% N- (1-naphthyl)-ethylenediamine in 47.5% ethanol. Membranes were then placed on a plastic Gel-bond sheet (FMC, Rockland, ME) in heat-sealed bags at -20°C. All substrates and chemicals were obtained from Sigma Chemical Co. (St. Louis, MO).

### **2.3.7. Zymograms**

In zymogram gels, proteases are characterized as clear bands in a stained protein background, where the protease has digested the substrate. NOVEX Zymogram gels (ZBC Caesin, Invitrogen, Carlsbad, CA) were used to detect proteases. *T. molitor* larval gut extracts were separated in a 4-16% Tris-Glycine gel containing blue-stained beta-casein in a 12-well format. There were four replicates of three lanes each with the first lane loaded with molecular marker, the second lane with 4 µl and the third well with 10 µl of the mixture of sample and sample buffer (2x). The gel was run with 1x Tris-Glycine SDS running buffer at 125 V for 90 min. After electrophoresis, the gel was transferred to Triton-X 100 (1:10 dilution) with gentle agitation at room temperature. The gel was then cut into four strips/sets with 3 lanes each, and strips were incubated at 37°C, each with a different treatment. Set 1 was fixed with 45% methanol, 10% acetic acid and stained with 0.5% Coomassie in 50% methanol and 10% acetic acid. Set 2 was incubated with universal buffer, pH 9.2, to evaluate serine protease activity. Set 3 was incubated in universal buffer, pH 6.2, with 5mM L-cysteine to evaluate cysteine protease activity. Set 4 was incubated in universal buffer, pH 9.2, with 25 µl of the yellow mealworm gut extract to evaluate serine protease inhibitory activity from the gut extract.

## **2.4. Results**

### **2.4.1. Microplate Assay**

Hydrolysis of the substrate, fluorescently labeled casein (BODIPY TR-X casein), by the crude gut extract of *T. molitor* was used to determine the optimal pH for activity of enzymes in the AM and PM. For AM proteases, there were two distinct activities, one with optimal activity in acidic buffers, with an optimum at pH 5.6, and another small increase in activity in basic buffers, around pH 7.9 (Figure 2.1). Activity also was examined in reducing buffers, because cysteine proteases have optimal activity under reducing conditions. However, there was no significant activation of activity for proteases from the AM in reducing buffers. This lack of

activation may have been due to the instability of cysteine proteases under the isolation conditions for gut extracts.

For proteases from the PM, a minor peak of activity was observed in acidic buffers, with an optimum at pH 5.6, and major peak of activity in alkaline buffers (Figure 2.2). The optimum pH for activity of PM proteases in the alkaline region was 7.9 in reducing buffers and 9.3 in nonreducing buffers. This shift in pH optimum may have been due to the increase of trypsin activity in reducing buffers, as has been previously reported for *T. molitor* trypsin (Elpidina *et al.*, 2005).

#### ***2.4.2. pH Profile Comparison of T. molitor Larvae Reared on Wheat Flour and Wheat Germ***

The total proteolytic activity of midgut extracts was measured and estimated in larvae that were fed two different diets, wheat flour (Figure 2.3A) or germ (Figure 2.3B), differing in oil content, fiber, vitamins, and minerals (Slavin *et al.*, 1999). Proteolytic activity in buffers with reducing reagent was not different that that in nonreducing buffer for either extract. There also was no difference in total proteolytic activity of gut extracts from larvae reared on wheat flour or wheat germ.

#### ***2.4.3. Effect of Storage of Gut Samples on Total Proteolytic Activity***

The proteolytic activity of freshly dissected gut extracts was compared with the proteolytic activity of guts extracts stored at -80°C over a variable period of time to determine if proteolytic activity decreases during storage. At acidic pH 4.0, the total proteolytic activity of freshly dissected larval gut extracts doubled in the AM and at alkaline pH 8.0, total proteolytic activity increased ~4-fold in the PM when compared to the proteolytic activity of guts stored at -80°C (Figures 2.4, 2.5). Results of a similar experiment indicated that when guts were dissected in water and immediately used for activity assays instead of first freezing at -80°C, maximal activity of enzymes was retained (Figure 2.6).

#### ***2.4.4. Inhibitor Assays***

A microplate assay was used to test the efficacy of different inhibitors on the crude gut extract of *T. molitor* and determine which of the inhibitors from various classes was most

effective in inhibiting protease activity of the *T. molitor* midgut. In nonreducing buffer, effective inhibition of general caesinolytic activity by *T. molitor* extracts was observed with aprotinin, chymostatin, PMSF, TLCK and SBTI. (Figure 2.7). However, in reducing buffer, E-64, aprotinin, chymostatin, PMSF, TPCK, and TLCK were effective inhibitors of general caesinolytic activity (Figure 2.8). The results from inhibitor assays showed effective inhibitors of general caesinolytic activity to be aprotinin, chymostatin, PMSF, TLCK and SBTI.

#### **2.4.5. Activity Blots**

An activity blot was used to visualize the development of protease activity in the *T. molitor* midgut extract when incubated with specific substrates. Hydrolysis of BApNA was nondetectable (data not shown). However, SAAPFpNA was hydrolyzed by at least six different chymotrypsin-like activities, with approximate molecular masses of 20, 25, 50, 90, 100 and 210 kDa (numbered 1-6, respectively, in Figure 2.9).

#### **2.4.6. Zymograms**

A zymogram, a strip or band of electrophoretic medium containing a stained protein, was used to observe the pattern of proteases after their separation by electrophoresis. A zymogram analysis of the gut extracts from the AM and PM demonstrated differences in the relative number and intensity of caesinolytic enzymes (Figure 2.10). At least eight different proteolytic activities were resolved in the AM and activities 2-8 were found in the PM. Activities 3, 7 and 8 were prominent in the AM and activities 3, 4, 7 and 8 were prominent in the PM. Inhibitors specific for selected cathepsins had no effect on *T. molitor* gut activity. All caseinolytic activities, except # 4 (seen as a white band in lane b of AM), were reduced when E-64 was added to the buffer.

### **2.5. Discussion**

This study identified several biochemical parameters useful in conducting further *in vitro* *T. molitor* gut-protease inhibitor studies. Zymogram assays revealed a diversity of digestive protease forms. Although some digestive proteases from *T. molitor* were described and identified by researchers in the past, a complete representation of all the digestive proteases in the midgut was lacking until recently. Results of a study by Vinokurov *et al.* (2006) have revealed a complex enzyme diversity through a combined use of general proteinaceous and a set of specific substrates, together with inhibitor analyses.

The susceptibility to inhibition by midgut homogenates was in agreement with the type of proteases suggested by the pH optimum curves. In the inhibitor assays, the whole midgut was used when, in fact, other studies have used AM and PM separately. However, the presence of especially trypsin and chymotrypsin serine proteases was indicated by inhibition with SBTI, PMSF, TLCK, chymostatin and aprotinin. Studies by Tsybina *et al.*, 2005 and Elpidina *et al.*, 2005 involved purification of a trypsin-like protease and a chymotrypsin-like protease from the *T. molitor* midgut. It should be noted that in reducing buffer, which enhances the effect of cysteine proteases, there was increased inhibition by E-64 and TPCK as compared to inhibition in a non-reducing buffer. Effective inhibition by PMSF, aprotinin, chymostatin, and TLCK also was observed in reducing medium. SBTI, on the other hand, inhibited less effectively in reducing medium as compared to inhibition in non-reducing buffers. This observation supported the hypothesis that the Cucujiformia ancestor was proposed to be a beetle adapted to ingest seeds rich in naturally occurring proteinase inhibitors by producing a cysteine protease for food digestion, instead of the common insect serine proteases (Terra and Cristofolletti, 1996). The pH of the buffer used for the microplate inhibitor assay was 8.5. However, the inhibitors may have not been stable in this pH. As cautioned by Barrett (1985), data presented for partial inhibition is difficult to evaluate, as it is incomplete and a time-dependant process with simple kinetics. Based on these data, the inhibitors for in *in vitro* and *in vivo* inhibitor assays should be selected with care.

The spatial distribution of the digestive proteases, such as cysteine proteases in the AM and serine proteases in the PM, was observed in our microplate assays with AM and PM extracts. The pH profile of the AM and PM was similar to an earlier study of the spatial distribution of two different pH environments on the midgut (Thie and Houseman, 1990). The evaluation of pH optimal activity can be a useful characteristic, although it cannot be used for the diagnostic characteristic of the different classes of proteases (Wolfson and Murdock, 1990). The increased hydrolysis by freshly dissected gut extracts demonstrated that the midguts of *T. molitor* larvae should be used immediately for *in vitro* enzymatic studies. Tissues kept frozen at  $-80^{\circ}\text{C}$  may preserve genetic material intact but lose proteolytic activity. Furthermore, freshly dissected guts stored in distilled water had maximal enzymatic activity as compared to freshly dissected gut stored in physiological buffered saline and frozen guts. These results suggest that dissecting guts

into water and using freshly dissected guts placed at 4°C are the best isolation conditions for protease assays.

In summary, these results suggested that the larvae of *T. molitor* rely on a diversity of serine proteases. Future studies are needed to identify the genes encoding these proteases *in vivo* and the regulatory expression of these proteases in response to ingested protease inhibitors. The enzymology of the insect digestive tract in relation to initial stages of digestion of large food polymers (starch or protein) reflects the biochemical adaptation of these post harvest insects to their preferred food (Baker, 1986). This basic biological data of insect digestive physiology can be made available and included in the computer models of population dynamics of stored-grain pest management.

## 2.6. References

- Applebaum, S. W., Y. Birk, I. Harpaz, and Bondi, A. (1964) Comparative studies on proteolytic enzymes of *Tenebrio molitor* L. *Comp. Biochem. Physiol.* 11: 85-103.
- Baker, J. E. (1986) Amylase/proteinase ratios in larval midguts of ten stored product insects. *Entomol. Exp. Appl.* 40: 41-46.
- Barrett, A.J. (1977b) Cathepsin D and other carboxyl proteinases. In: Proteinases in Mammalian Cells and Tissues. pp. 207-248. North Holland, Amsterdam.
- Barrett, A. J. (1986) The classes of proteolytic enzymes. In: M.J. Dalling (ed.). Plant Proteolytic Enzymes 1: 37-39. CRC Press, Boca Raton.
- Barrett, A.J. and Rawlings, N.D. (1991) Types and families of endopeptidases. In "Proteinases". *Biochem. Soc. Trans.* 19: 707-715.
- Boulter, D., Edwards, G. A., Gatehouse, A. M. R., Gatehouse, J. A. and Hilder, V. A. (1990) Additive protective effects of different plant-derived insect resistance genes in transgenic tobacco plants. *Crop Prot.* 9: 351-354.
- Burgess, E. P. J., Stevens, P. S., Keen, G. K. Laing, W. A., and Christeller, J. T. (1991) 61: *Entomol. Exp. Appl.* 123-130.
- Cristofaletti, P.T. and Terra, W.R. (1999) Specificity, anchoring, and subsites in the active center of a microvillar aminopeptidase purified from *Tenebrio molitor* (Coleoptera) midgut cells. *Insect Biochem. Mol. Biol.* 29: 807-819
- Cristofaletti, P.T. and Terra, W.R. (2000) The role of amino acid residues in the active site of a midgut microvillar aminopeptidase from the beetle *Tenebrio molitor*. *Biochim. Biophys. Acta* 1479: 185-195.
- Cristofaletti, P. T., Ribeiro, A. F. and Terra, W. R. (2001) Apocrine secretion of amylase and exocytosis of trypsin along the midgut of *Tenebrio molitor* larvae. *J. Insect Physiol.* 47: 143-155.
- Ferriera C., Bellinello, G. L. , Ribeiro, A. F. and Terra, W. R. (1990) Digestive enzymes associated with glycocalyx, microvillar membranes and secretory vesicles from midgut cells of *Tenebrio molitor* larvae. *Insect Biochem.* 20: 839-847.

Frugoni, J.A.C. (1957) Tampone universale di Britton a Robinson a forza ionica costante. *Gazz. Chim.* 87: 403-407.

Garty, N. (1979) Isolation and characterization of a chymotrypsin-like enzyme from *Tenebrio molitor* larvae. Biochemistry of digestion. In: Comprehensive Physiology, Biochemistry and Pharmacology of Insects 4: 279-311, Pergamon Press, Oxford

Gatehouse, A. M. R., Shi, Y., Powell, K.S., Brough, C., Hilder, V.A., Hamilton, W.D.O., Newell, C. A., Merryweather, A., Boulter, D. and Gatehouse, J. A. (1993) *Philos. Trans. R. Soc. London Ser. B* 342: 279-286.

Golan, R. (1981) Isolation, characterization and comparative study of proteolytic enzymes from the midguts of *Tenebrio molitor* adults and larvae as a basis of possible biological pest control with naturally occurring protease inhibitors from plant sources. Biochemistry of digestion. In: Comprehensive Physiology, Biochemistry and Pharmacology of Insects 4: 279-311, Pergamon Press, Oxford

Guiderdoni, E., Cordero, M.J., Vignols, F., Garcia-Garrido, J. M., Lescot, M., Tharreau, D., Meynard, D., Ferriere N., Notteghem, J. L., Delseny, M. (2002) Inducibility by pathogen attack and developmental regulation of the rice *Ltp1* gene. *Plant Mol. Biol.* 49:683-699.

Hilder, V. A., Gatehouse A. M. R., Sheerman, S. E., Barker R. F. and Boulter D. (1987) A novel mechanism of insect resistance engineered from tobacco. *Nature* 330: 160-163.

Johnson, R. J., Narvaez, J., An, G., Ryan, C. A. (1989) Expression of proteinase inhibitors I and II in transgenic tobacco plants: Effects on natural defense against *Manduca sexta* larvae. *Proc. Natl. Acad. Sci.* 86: 9871-9875.

Johnston, K. A., Lee, M. J., Gatehouse J.A. and Anstee, J. H. (1991) The partial purification and characterization of serine protease activity in midgut of larval *Helicoverpa armigera*. *Insect Biochem.* 21: 389-397.

Jongsma, M.A, Bakker, P.L., Peters, J., Bossch, D. and Stiekema, W. H. (1995) Adaptation of *Spodoptera exigua* larvae to plant proteinase inhibitors by induction of gut proteinase activity insensitive to inhibition. *Proc. Natl. Acad. Sci.* 92: 8041-8045.

Jongsma, M.A. And Bolter, C. (1997) The adaptation of insects to plant protease inhibitors. *J. Insect Physiol.* 43: 885-895.

Levinsky, H., Birk, Y. and Applebaum, S. W. 1977. Isolation and characterization of a new trypsin-like enzyme from *Tenebrio molitor* L. larvae. . *Int. J. Peptide Protein Res.* 10: 252-264.

Law, J.H., Dunn, P.E. and K.J. Kramer. (1977) Insect Proteases and Peptidases. In *Advances In Enzymology.* 389-425, John Wiley and Sons Inc., U.S.A.

McManus M.T., White, D. W.R. and Mc Gregor, P.G. (1994) Accumulation of a chymotrypsin inhibitor in transgenic tobacco can affect the growth of insect pests. *Transgenic Res.* 3:50-58

North, M. J. (1982) Comparative biochemistry of the proteinases of eukaryotic microorganisms. *Microbiology Rev.* 46:308-340.

Oliviera, A.S., Xavier-Filho, J. And Sales, M. P. (2003) Cysteine Proteinases and Cystatins. *Braz. Arch. Biol. Technol.* 46: 91-104.

Oppert, B., Kramer, K. J. and McGaughey, W. H. (1997) Rapid microplate assay for substrates and inhibitors of proteinase mixtures. *Biotechniques* 23: 70-72.

Oppert, B. (2001) Transgenic plants expressing enzyme inhibitors and the prospects for biopesticide development, In *Phytochemical Biopesticides*, O. Koul and G. S. Dhaliwal (eds.), Harwood Academic Publishers, Singapore, pp. 83-95.

Page, D., Quillien, L. and Duc, G. (2000) Trypsin inhibitory activity measurement: simplifications of the standard procedure used for pea seed. *Crop Science.* 40: 1482-1485.

Ramos-Eldorduy, J., Gonzalez, E. A., Hernandez, A. R., and Pino, J.M. (2002) Use of *Tenebrio molitor* (Coleoptera: Tenebrionidae) to recycle organic wastes and as feed for broiler chickens. *J. Econ. Entomol.* 95: 214-220.

Rawlings, N.D. and Barrett, A.J. (1994) Families of Serine Peptidases. In "Proteolytic Enzymes: Serine and Cysteine Peptidases." Ed. By A.J. Barrett. Academic Press, Inc., *Methods in Enzymology* 244:19-61.

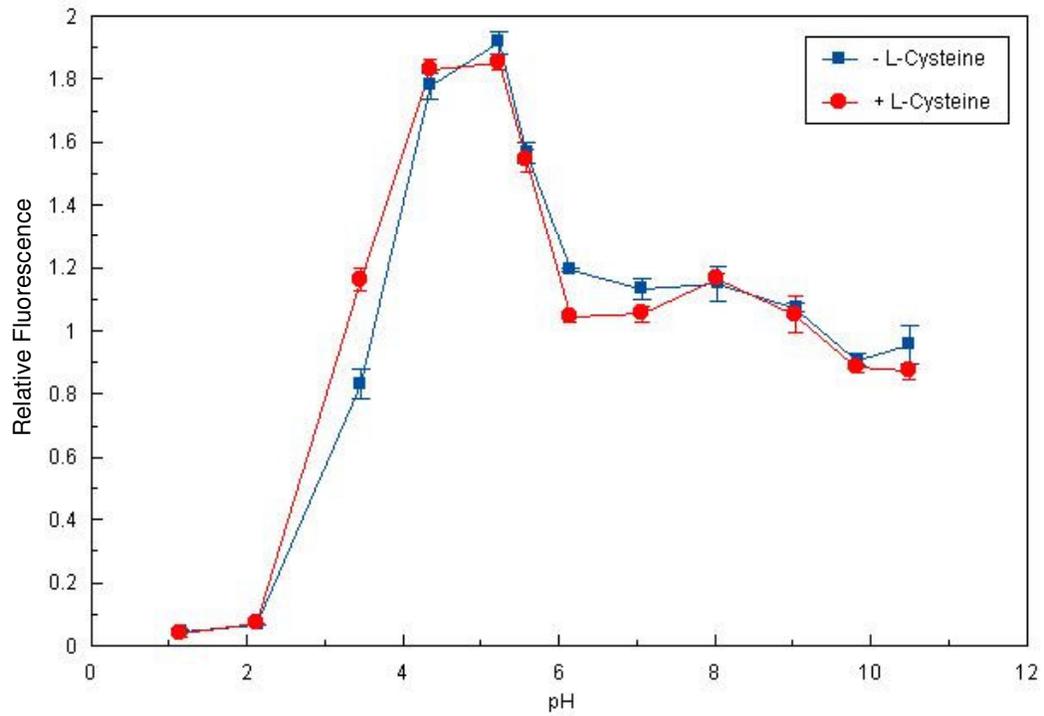
Richardson, M. (1977) The proteinase inhibitors of plants and micro-organisms. *Phytochem.* 16: 159-169.

Shaw, E., Mares-Guia, M. and Cohen, W. (1965) Evidence for an active center histidine in trypsin through use of a specific reagent 1-chloro-3-tosylamino-7 amino-2 heptanone, the chloromethyl ketone derived from N- $\alpha$ -tosyl-L-lysine. *Biochemistry* 4: 2219-2224.

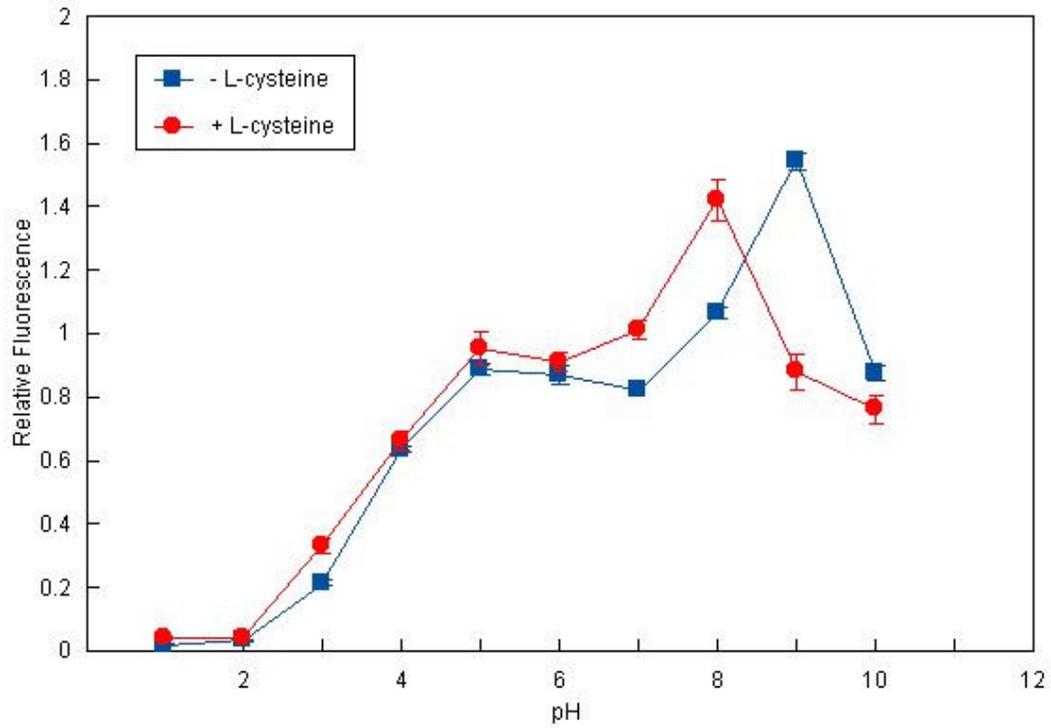
- Slavin, J. J., Martini, M. C., Jacobs, D. R., Jr & Marquart, L. (1999) Plausible mechanisms for the protectiveness of whole grains. *Am. J. Clin. Nutr.* 70: 459S-463S.
- Swegle, M., Kramer K.J, Muthukrishnan, S. 1992. Properties of barley seed chitinases and release of embryo-associated isoforms during early stages of imbibition. *Plant Physiol.* 99: 1009-1144.
- Terra, W. R. and Cristofaletti, P.T. (1996) Midgut proteinases in three divergent species of Coleoptera. *Comp. Biochem. Physiol.* 113B: 725-730.
- Terra, W. R. and Ferreira, C. (1994) Insect digestive enzymes: properties compartmentalization and function. *Comp. Biochem. Physiol.* 109B: 1-62.
- Thie, N. M. R. and Houseman, J. G. (1990) Cysteine and serine proteolytic activities in larval midgut of yellow mealworm, *Tenebrio molitor* L. (Coleoptera: Tenebrionidae). *Insect Biochem.* 20: 741-744.
- Throne, J. E., (1995) Computer modeling of the population dynamics of stored-product pests. In: *Stored-Grain Ecosystems*. Ed. By Jayas, D. S.; White, N. D. G.; Muir, W. E. New York, Basel, Hong Kong: Marcel Dekker, Inc., pp. 169 195.
- Urieli, N. (1982) Isolation and characterization of a chymotrypsin-like enzyme from *Tenebrio molitor* larvae. Biochemistry of digestion. In: *Comprehensive Physiology, Biochemistry and Pharmacology of Insects 4*: 279-311, Pergamon Press, Oxford.
- Wagner, F.W. (1986) Assessment of methodology for the purification, characterization, and measurement of proteases. In: M.J. Dalling (ed.). *Plant Proteolytic Enzymes 1*: 37-39. CRC Press, Boca Raton.
- Warchalewski, J.R., Paogon, Dzynska.A., Gralik, J., Nawrot, J. (2000b) The effect of gamma and microwave irradiation of wheat grain on development parameters of some stored grain pests. *Nahrung-Food* 44 : 411-414.
- Wolfson J. L. and Murdock, L. L. (1990) Diversity in digestive proteinase activity among insects. *J. Chem. Ecol.* 16: 1089-1102
- Yamamoto, Y., Takimoto, K., Izumi, S. Toriyama-Sakurai, M., Kageyama, T. and Takahashi, S. Y. (1994) Molecular cloning and sequencing of cDNA that encodes cysteine proteinase in the eggs of the silkworm, *Bombyx mori*. *J. Biochem.* 116: 1330-1335.
- Zwilling, R. (1968) Zur Evolution der Endopeptidasen. IV. A and  $\beta$ -protease aus *Tenebrio molitor*. *Hoppe-Seyler's Z. Physiol. Chem.* 349: 326-332.

Zwilling, R., Medugorac, I. and Mella, K. (1972) The evolution of endopeptidases-XIV. Non-tryptic cleavage specificity of a BAEE-hydrolyzing enzyme (B-protease) from *Tenebrio molitor*. *Comp. Biochem. Physiol.* 43: 419-424.

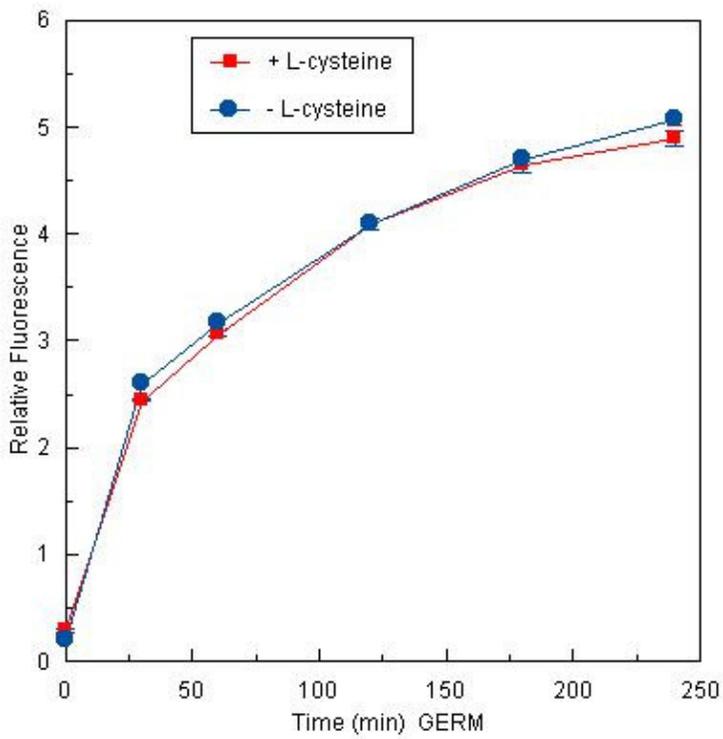
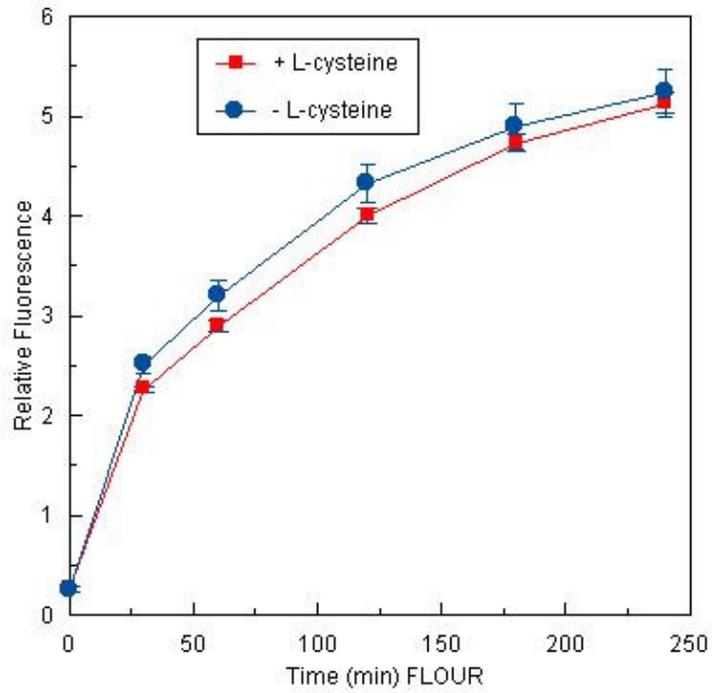
**Figure 2.1. Hydrolysis of fluorescent-casein by enzymes in the anterior midgut (AM) extract of *T. molitor* larvae, in buffers containing or lacking L-cysteine.**



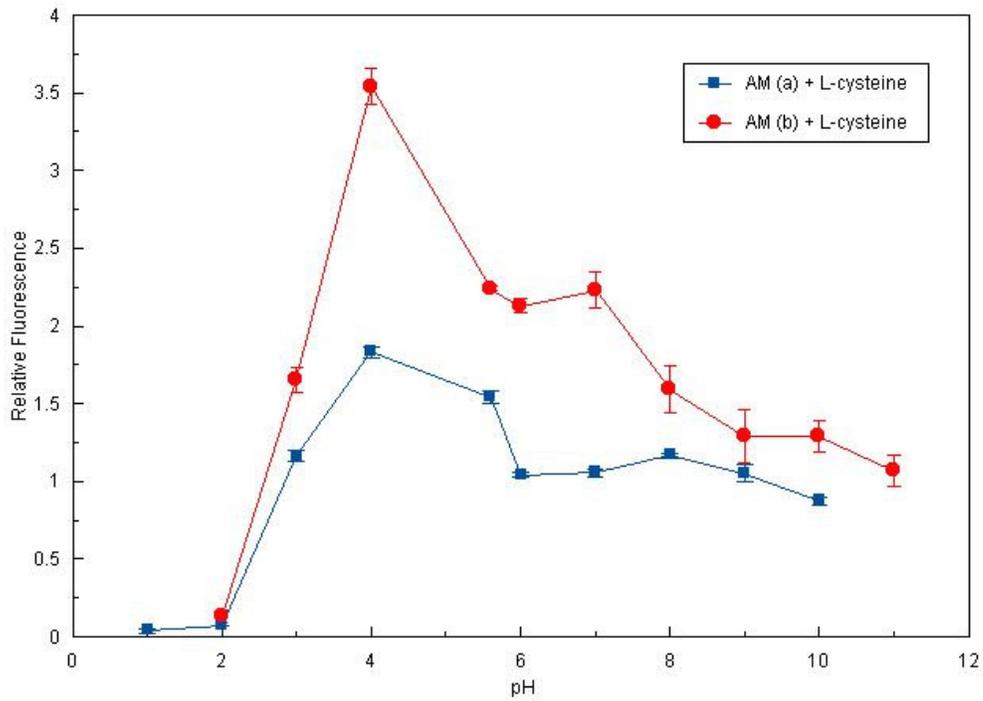
**Figure 2.2. Hydrolysis of fluorescent-casein by enzymes in the posterior (PM) midgut extract of *T. molitor* larvae, in buffers containing or lacking L-cysteine.**



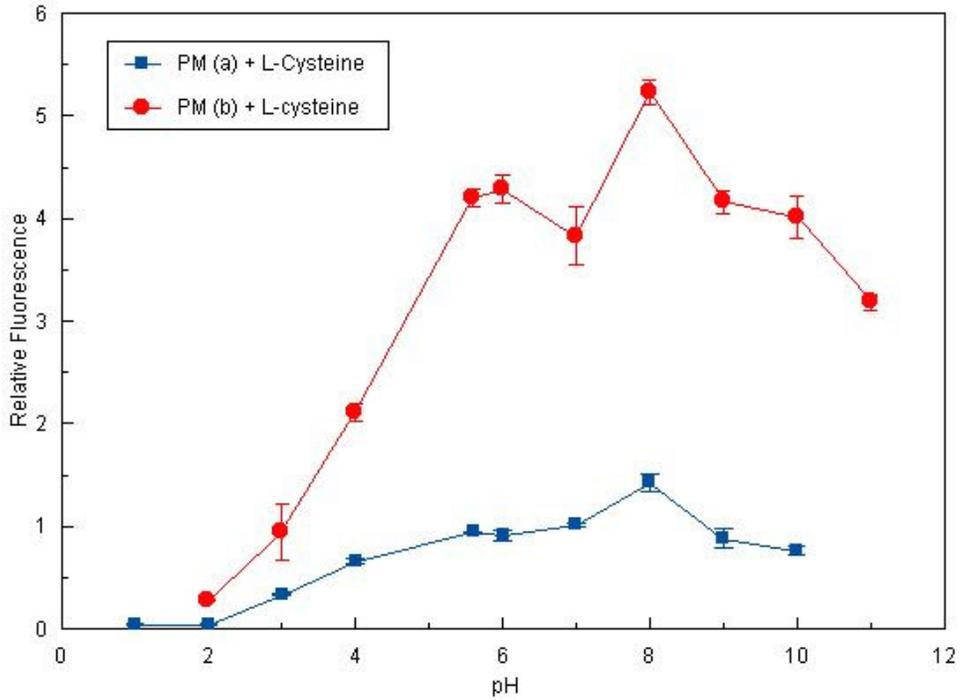
**Figure 2.3. Hydrolysis of fluorescent casein by enzymes in the midgut extract of *T. molitor* larvae reared on (a) wheat flour (b) wheat germ in buffers containing or lacking L-cysteine.**



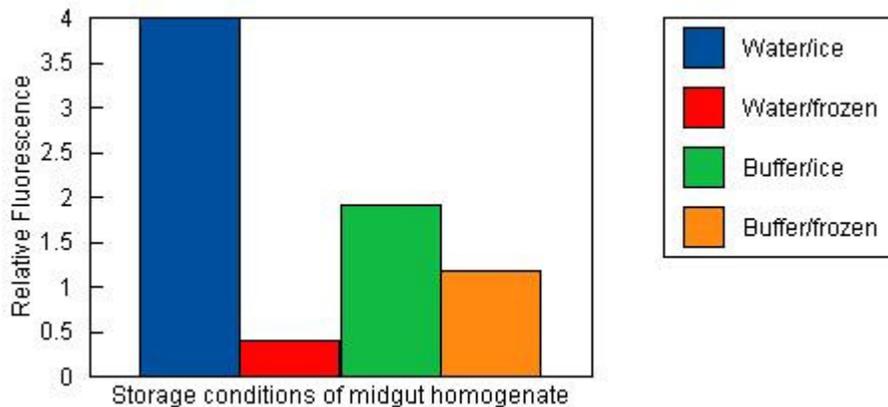
**Figure 2.4. Hydrolysis of fluorescent-casein by enzymes in the anterior midgut (AM) extracts of guts (a) stored at -80°C (b) freshly dissected guts of *T. molitor* larvae, in buffers containing L-cysteine.**



**Figure 2.5. Hydrolysis of fluorescent casein by enzymes in the posterior midgut (PM) extracts of guts (a) stored at -80°C (b) freshly dissected guts of *T. molitor* larvae, in buffers containing L-cysteine.**

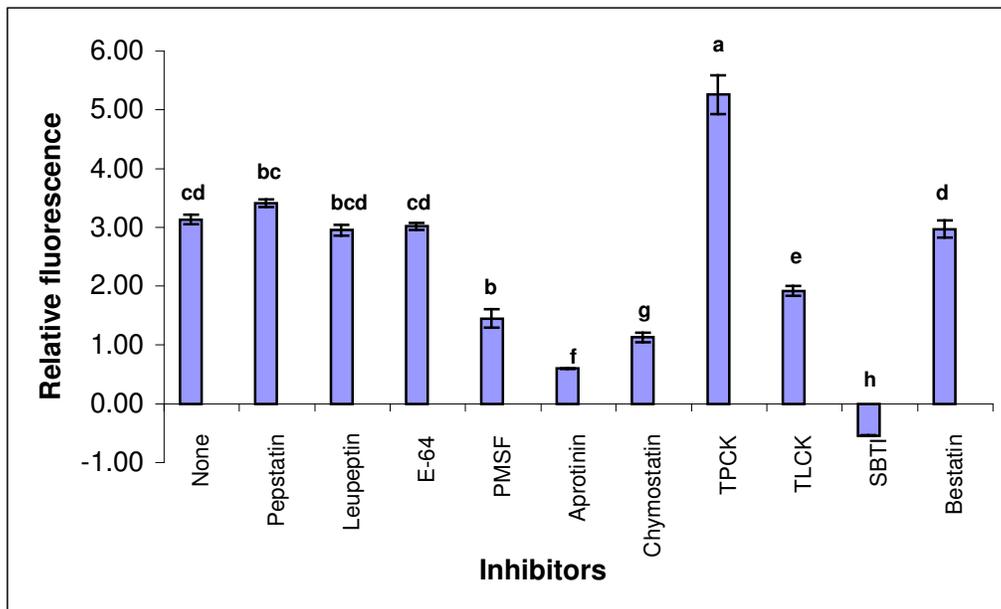


**Figure 2.6. Hydrolysis of fluorescent-casein by enzymes in *T. molitor* gut extracts of freshly dissected guts or guts stored at -80°C that were in water or in saline buffer**



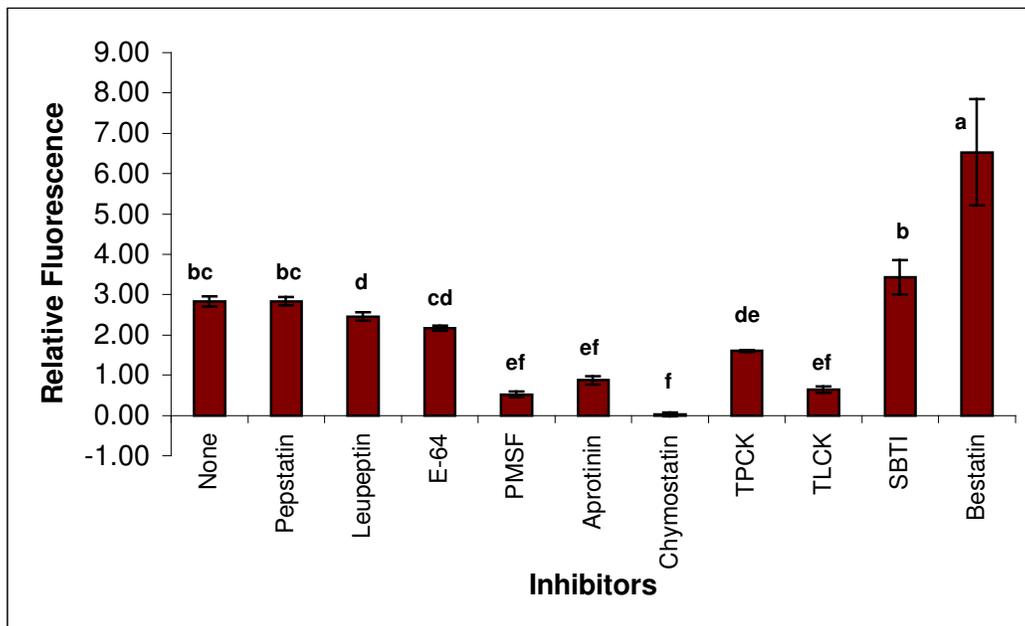
**Figure 2.7. A microplate inhibitor assay of *T. molitor* midgut incubated with inhibitors and enzyme activity measured with fluorescent-casein, in non-reducing buffers.**

\*Means followed by the same letter are not significantly different ( $P>0.05$ ; LSD). Grouping analysis was done statistically using PROC GLM (SAS Institute 1985) and means were separated using LSD.

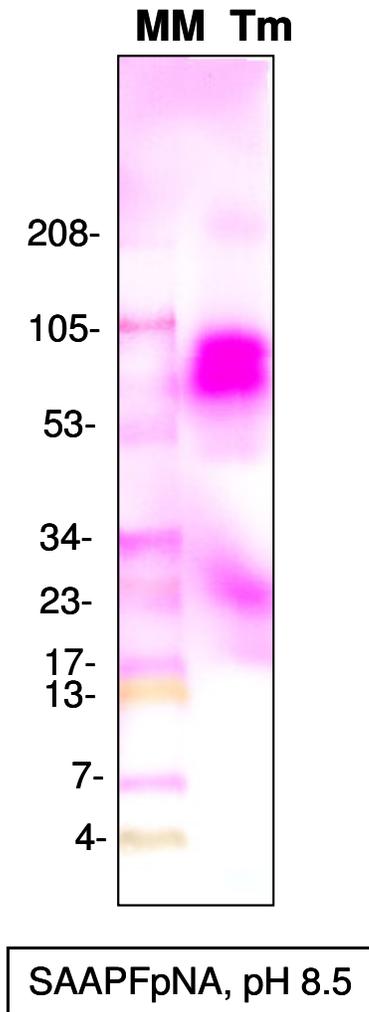


**Figure 2.8. A microplate inhibitor assay of *T. molitor* midgut extracts incubated with inhibitors and enzyme activity measured with fluorescent-casein, in reducing buffers.**

\* Means followed by the same letter are not significantly different ( $P>0.05$ ; LSD). Grouping analysis was done statistically using PROC GLM (SAS Institute, 1985) and means were separated using LSD.



**Figure 2.9.** Activity blot incubated with SAAPFpNA showing six different chymotrypsin-like activities (numbered 1-6) in *T. molitor* midgut extract.



**Figure 2.10. Zymogram Gel picture showing *T. molitor* midgut protease (AM & PM) activities (1-8) when inhibited with five different inhibitors (lanes a-e).**

**Lane    Inhibitors**

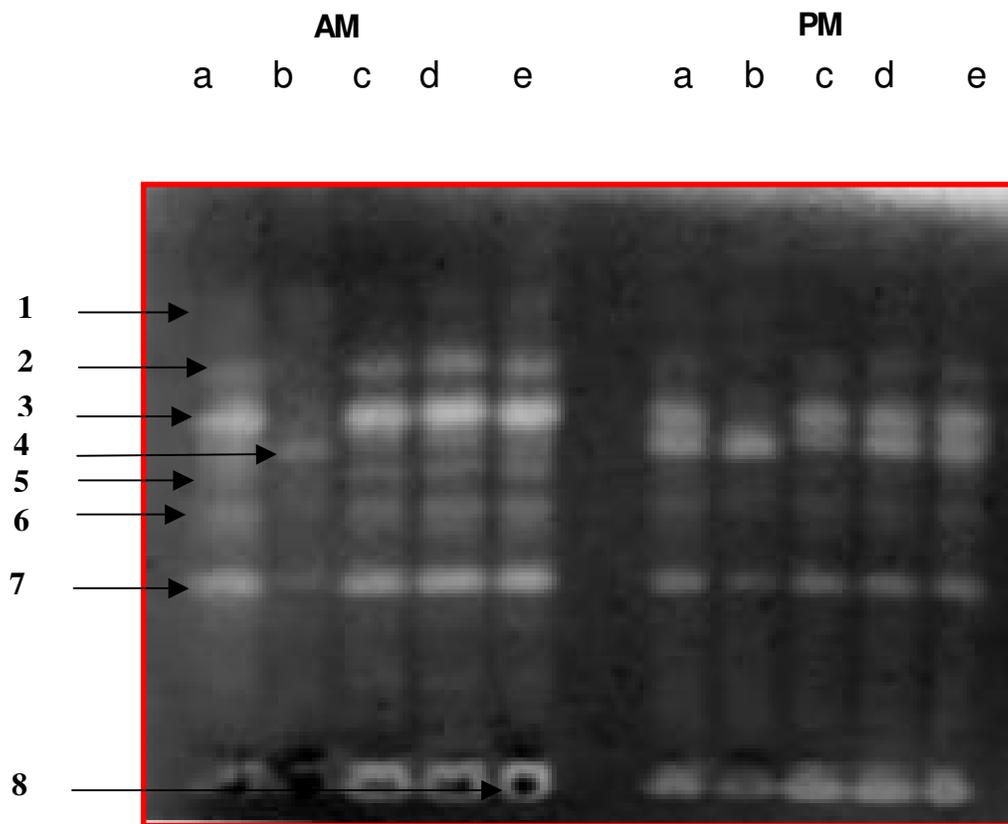
Lane a. Control

Lane b. E-64

Lane c. Cathepsin B II

Lane d. Cathepsin L I

Lane e. Cathepsin L II



# **CHAPTER 3 - TRANSCRIPTOME ANALYSIS OF DIGESTIVE PROTEASES FROM THE YELLOW MEALWORM, *TENEBRIO MOLITOR* L. (COLEOPTERA: TENEBRIONIDAE)**

## **3.1. Abstract**

As part of the study of the digestive transcriptome of the yellow mealworm, *Tenebrio molitor* Linnaeus, sequences from randomly picked clones from midgut cDNA libraries of various larval stages were analyzed. Of a total of 1,528 sequences, 92 sequences encoded potential proteases expressed in the midgut of *T. molitor*, and clones of these 92 were further sequenced to obtain 50 full-length cDNA sequences. Protease sequences were grouped into functional classes, and sequences were selected for further analysis. Northern blot analysis of representative sequences demonstrated the differential expression profile of selected transcripts across five developmental stages of *T. molitor*, suggesting an active regulation of digestive proteases. These sequences provide information on digestive proteases in coleopteran insects as a basis to study the response of coleopteran larvae to external stimuli and to evaluate regulatory features of the response.

### 3.2. Introduction

Proteases have important functions in all living systems. Insect digestive proteases catalyze the breakdown of protein during digestion and provide amino acids for growth and development. Therefore, proteases are attractive candidates for the development of new pest control proteins based on synthetic or natural inhibitors. Proteases include exopeptidases, such as amino- and carboxypeptidases, and endopeptidases, also referred to as proteinases, and are grouped as serine, cysteine, metallo, or aspartic, as defined by the active site and catalytic mechanism (Neurath, 1982). Protease gene families have been subdivided into clans based on sequence similarity and cataloged in the MEROPS database (Rawlings *et al.*, 2006).

Genome studies are providing information on the complexity and diversity of protease genes in insects. In a comparison of genome sequences, there were 305 serine protease (chymotrypsin or trypsin family) genes predicted in *Anopheles gambiae*, 199-210 in *Drosophila melanogaster*, and 118 in *Homo sapiens*, but only 13 in *Caenorhabditis elegans* and one in *Saacharomyces cerevisiae* (Rubin *et al.*, 2000; International Human Genome Sequencing Consortium, 2001; Holt *et al.*, 2002; Zdobnov *et al.*, 2002; Ross *et al.*, 2003). Multigene families of trypsins were found in *Lucilia cuprina* and *Haematobia irritans exigua* (Casu *et al.*, 1994; Elvin *et al.*, 1993). In *L. cuprina*, 125-220 different serine protease genes were expressed (Elvin *et al.*, 1994). A cluster of seven trypsin and two chymotrypsin genes in *Anopheles gambiae* was induced by feeding (Müller *et al.*, 1993). Transcriptome analysis of the insect midgut has provided further functional and physiological information related to digestion (Pedra *et al.*, 2003; Campbell *et al.*, 2005; Xu *et al.*, 2005).

Twenty-eight different serine proteinase genes in *Helicoverpa armigera* were differentially expressed in response to dietary inhibitors (Bown *et al.*, 1997). Inclusion of soybean trypsin inhibitor in the diet stimulated an initial up-regulation of protease genes and a longer down-regulation of inhibitor-sensitive protease genes (Bown *et al.*, 2004). The adaptation of insects to protease inhibitors suggests that specific transcriptional or translational regulation of protease genes enables feeding on plants that have protease inhibitor defenses.

The yellow mealworm, *Tenebrio molitor*, is a univoltine stored-product pest found in stored grain products, causing economic loss in food production worldwide. The earliest studies on insect larval digestive proteases were of *T. molitor* (Applebaum *et al.*, 1964). Larvae of *T.*

*molitor* have a digestive physiology that incorporates a complex of digestive proteases, including those from serine and cysteine protease classes, operating in a midgut with a sharp pH gradient (Terra *et al.*, 1985). Cysteine protease activity was compartmentalized to the anterior region of the larval midgut of *T. molitor*, whereas serine protease activity was found in the posterior region (Thie and Houseman, 1990; Terra and Ferreira, 1994). Multiple digestive protease activities have been reported in other coleopteran pests (Oppert *et al.*, 1993, 2003, 2005; Bouchard *et al.*, 2003; Zhu-Salzman *et al.*, 2003). A combination of biochemical studies on proteolytic activity in the midgut of larval *T. molitor* has identified at least 20 proteases, although some may not represent unique enzymes, as the methodology of study was not uniform. Of these proteases, four are aminopeptidases, two are carboxypeptidases, six are cysteine proteases, and the rest are serine proteases (Applebaum *et al.*, 1964; Zwillling, 1968; Zwillling *et al.*, 1972; Levinsky *et al.*, 1977; Garty, 1979; Golan, 1981; Urieli, 1982; Ferriera *et al.*, 1990; Thie and Houseman, 1990; Terra and Cristofolletti, 1996; Cristofolletti and Terra, 1999, 2000; Cristofolletti *et al.*, 2005; Elpidina *et al.*, 2005; Tsybina *et al.*, 2005). A recent comprehensive biochemical analysis suggested that at least 13 different protease activities are expressed simultaneously in *T. molitor* larvae under normal dietary conditions, including six cysteine proteases, four trypsin-like, and five chymotrypsin-like proteases (Vinokurov *et al.*, 2006a, 2006b). Recent studies have provided the N-terminal sequences of the major trypsin and chymotrypsin digestive proteases in this group (Tsybina *et al.*, 2005; Elpidina *et al.*, 2005). Five procathepsin L-like proteases from a cDNA library included the predicted sequence of a major luminal proenzyme (Cristofolletti *et al.*, 2005).

A comprehensive analysis of the cDNAs encoding proteases in the midgut of *T. molitor* larvae was conducted by a transcriptomic analysis of the regulation of insect proteases in the coleopteran insect. The structural motifs and expression patterns of representative genes encoding proteases from different classes were demonstrated and compared in five developmental stages of *T. molitor*.

### **3.3. Materials and Methods**

#### ***3.3.1. Insect Rearing and Dissection***

*T. molitor* larvae used in this research were derived from a laboratory colony. Insects were reared at a relative humidity of 60-70% at 25°C. Fully grown late instar larvae of both sexes, weighing  $1.13 \pm 0.03$  g (n=15), were used for AM and PM midgut cDNA library

construction and isolation of digestive trypsin and chymotrypsin. “Small” ( $0.89 \pm 0.24$  mg,  $n=27$ ) and “middle” ( $168 \text{ mg} \pm 0.25$ ,  $n=10$ ) instar larvae were used for the construction of small and middle cDNA libraries (S and L, respectively). Larvae from three different instars, pupae and adults were used for RNA isolation for Northern blots and qPCR.

Actively feeding larvae were immobilized on ice and were dissected by excising the anterior and posterior ends, removing the gut with forceps, and placing the gut into a solution of DEPC-water. The entire midgut was divided into AM and PM sections of identical length by cutting midway between the most anterior region and the insertion of the Malpighian tubules. The contents of the gut were gently forced out with a pair of forceps, and the tissue was rinsed twice in DEPC-water, dried on filter paper and immediately transferred into TRI Reagent™ (Molecular Research Center, Inc., Cincinnati, OH).

### ***3.3.2. cDNA Library Construction and Sequencing***

Total cellular RNA was extracted from dissected gut tissue using TRI reagent, according to the procedure provided by the manufacturer. Four cDNA libraries (AM, PM, S and L) were constructed from RNA samples using a SMART™ cDNA library construction kit, which synthesizes cDNA using oligo-dT primers (Clontech, Palo Alto, CA), according to the protocol provided by the manufacturer, with one modification. Instead of the original phage vector, PCR fragments were cloned directly into a plasmid using a TOPO TA cloning kit (Invitrogen, Carlsbad, CA). Approximately 500 clones from each of the four libraries were picked randomly, and unidirectional sequences were obtained commercially (MegaBACE 4000, Amersham Biosciences, Piscataway, NJ, Rexagen DNA Sequencing Service, Seattle, WA; ABI Prism™ 3730xl, SeqWright DNA Technology Services, Houston, TX). Universal primers M13 F (-20) and M13 R were used for further sequencing of predicted protease clones. Specific synthetic primers were used to confirm the sequences from both directions. Some sequence analysis also was performed using an ABI 3700 DNA sequencer at the DNA Sequencing and Genotyping Facility, Dept. of Plant Pathology, Kansas State University, Manhattan, KS.

### ***3.3.3. Computer-based Sequence Analysis***

Vector sequences were trimmed using Sequencher (Gene Codes Corporation, Ann Arbor, MI). Sequences were grouped into clusters and redundant cDNA sequences were identified and analyzed using a customized BLASTN program that produced outputs with sequence assembly

parameters similar to those of the CAP3 assembly program (Huang and Madan, 1999). Sequences also were analyzed using the BLASTX algorithm (<http://www.ncbi.nlm.nih.gov/BLAST>). Sequences were grouped by their similarity, based on predicted open reading frame (ORF), to sequences in the database at the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>; Bethesda, MD). Full-length sequences were obtained with synthetic primers for clones encoding nonredundant proteases.

Secretion signal peptides were predicted using PSORT or SignalP (<http://psort.nibb.ac.jp/form2.html>; <http://www.cbs.dtu.dk/services/SignalP>). The 'Compute pI/Mw tool' ([http://us.expasy.org/tools/pi\\_tool.html](http://us.expasy.org/tools/pi_tool.html)) was used for molecular mass calculations and pI prediction of mature proteins. The ExPASy Proteomics tools on the website of the Swiss Institute of Bioinformatics (<http://www.expasy.ch/tools>) were used to process data of deduced protein sequences. Alignments were made with MULTALIN (PBIL, Lyon, France) or CLUSTALW in pairwise comparisons (gap weight=8, gap length weight=2; Thompson *et al.*, 1994) was used in pairwise comparison.

#### ***3.3.4. Isolation and Purification of *T. molitor* Digestive Trypsin and Chymotrypsin***

Isolation and purification of digestive trypsin (TmT1) and chymotrypsin (TmC1) was performed as described earlier (Tsybina *et al.*, 2005 and Elpidina *et al.*, 2005, respectively). Two hundred PM of late instar larvae were homogenized in 0.75% NaCl and clarified with centrifugation. The extract was dialyzed against 20 mM potassium sodium phosphate buffer, pH 6.9, containing 0.02% sodium azide. Dialyzed extracts were subjected to batch chromatography on DEAE-Sephadex A-50 in the same buffer, followed by gel filtration FPLC on a Superdex-75 column in the same buffer. Aliquots of resolved and purified enzymes were analyzed by MALDI-TOF MS or subjected to SDS-PAGE according to Laemmli (1970), followed by in-gel tryptic hydrolysis.

#### ***3.3.5. In Gel Tryptic Hydrolysis of Purified Trypsin and Chymotrypsin***

Tryptic peptides of purified digestive trypsin and chymotrypsin were obtained after hydrolysis of proteins in a 0.15% Coomassie Brilliant Blue R-250 stained gel. A 1 mm 2 gel piece was excised and washed twice in 150  $\mu$ l of 40% (v/v) acetonitrile in 0.1 M  $\text{NH}_4\text{HCO}_3$  for 20 min at 56 °C. The gel was further dehydrated in 150  $\mu$ l of the same buffer, was dried, and was supplemented with 3  $\mu$ l of modified sequencing-grade trypsin (Promega, Madison, WI,

USA) dissolved in 0.05 M  $\text{NH}_4\text{HCO}_3$  to a final concentration of 10  $\mu\text{g/ml}$ . The hydrolysis was performed for 15 h at 37 °C and was stopped by the addition of 5  $\mu\text{l}$  of 0.1% of trifluoroacetic acid in a 10% solution (v/v) of acetonitrile and water, followed by thorough mixing. The gel solution was used for MALDI-TOF MS and MS/MS analysis of peptides.

### ***3.3.6. MALDI-TOF MS and MS/MS Analysis***

Aliquots (1 ml) of the sample were mixed on a steel target with an equal volume of 2,5-dihydroxybenzoic acid (Sigma-Aldrich, St. Louis, MO) solution (10 mg/ml in 30% acetonitrile/0.5% trifluoroacetic acid), and the droplet was left to dry at room temperature. Mass spectra were recorded on an Ultraflex MALDI-TOF-TOF mass spectrometer (Bruker Daltonik, Germany) equipped with a 337 nm laser with positive ion detection. Each mass spectrum was obtained as the sum of a minimum of 200 laser shots. Fragment ion spectra were generated by laser-induced dissociation slightly accelerated by low-energy collision-induced dissociation using helium as a collision gas. The  $\text{MH}^+$  molecular ions of the tryptic digest were measured in reflector mode, and the accuracy of mass peak measurement was 0.02%. Mass spectra of the intact proteins were obtained in linear mode. Mass peak accuracy of the measurement of intact protein and MS/MS fragmentation of peptides was 0.05 %. Correspondence of the masses and MS/MS peptide fragments to predicted protein peptides was manually interpreted with GPMW 4.04 (Lighthouse Data, Denmark).

### ***3.3.7. RNA Isolation and Northern Blot Analysis***

Total cellular RNA from five different developmental stages (first-instar, mid-instar, late-instar larvae, pupae and adults) of *T. molitor* was extracted as previously described. For Northern blots, equal amounts (5  $\mu\text{g}$ ) of total RNA were separated in a 1.2% agarose gel containing formaldehyde and were transferred to a nylon membrane. The membrane was baked at 80°C for 2 h to fix RNA onto the membrane. The membranes were then hybridized separately to 20 individual cDNA probes, each representing a different cluster of full-length proteases. The cDNA probes were labeled with  $^{32}\text{P}$ -dCTP using a random labeling kit from Stratagene (La Jolla, CA). Hybridization was carried out overnight at 42°C in a plastic bag containing 15 ml hybridization solution (10% dextran sulfate/1% SDS/1 M NaCl, pH 8.0). After hybridization, the membranes were washed twice with 2×SSC at room temperature for 30 min, twice with 2×SSC plus 1%SDS at 65 °C for 30 min, and twice with 0.1×SSC plus 1%SDS at room temperature for

30 min. The membranes were exposed to Kodak SR-5 X-ray film overnight. The expression of mRNA in developmental stages of *T. molitor* larvae was measured as band intensity per unit (NucleoTech, San Mateo, CA).

### **3.3.8. Phylogenetic Analysis**

To compare the relationships among proteases isolated from the *T. molitor* midgut and those of other species, phylogenetic trees were constructed. For serine proteases, trypsin from *D. melanogaster* and human trypsin HsPRSS1 were included as outgroups. Sequences were aligned using the Alignment Explorer/CLUSTAL tool in Molecular Evolutionary Genetics Analysis (MEGA 3.1). MEGA 2.1 was used to generate a phylogenetic tree calculated based on the distance/neighbor-joining method. Bootstrap values were obtained with 1000 replications.

## **3.4. Results**

### **3.4.1. Composition of *T. molitor* Transcripts**

cDNA libraries were constructed using mRNA from the anterior (AM) and posterior (PM) midgut of late-instar larvae, and entire midguts of early-(S) and mid-(L) instar *T. molitor* larvae. Sequences were obtained for a total of 301, 373, 272, and 330 clones selected randomly from the AM, PM, S, and L libraries, respectively. Each sequence was compared to the non-redundant GenBank database using the BLASTX algorithm. The threshold expected value (E-value) was set to 10, meaning that 10 such matches are expected to be found merely by chance alone (Karlin and Altschul, 1990). Of these, 26% (332) of the sequences had no significant match ( $E > 10$ ), and 67% encoded proteins that were not proteases (Figure 3.1). Seven percent of the cDNA sequences (92 total) encoded distinct sequences for putative midgut proteases. These sequences were grouped by similarity to proteases belonging to one of the four classes. Of these sequences, 73% encoded serine proteases, 13% encoded cysteine proteases, and 14% encoded metalloproteases. None of the sequences were similar to aspartic proteases.

### **3.4.2. Distribution of Protease Transcripts in *T. molitor***

Thirty-seven cDNAs predicted to encode serine proteases were identified in the PM cDNA library, whereas only 18 serine protease-encoding cDNAs were found in the AM cDNA library (Figure 3.2). In contrast, 10 sequences encoding cysteine proteases were from the AM

cDNA library and only two were from the PM cDNA library. A similar trend was observed for metalloproteases, with more sequences obtained from the AM than from the PM cDNA library.

### ***3.4.3. Classification of Protease Transcripts***

#### ***3.4.3.1. Serine Proteases***

A total of 50 full-length, nonredundant cDNAs encoding proteases from the midgut libraries were obtained. To characterize proteases further by their specificity, residues were identified that determine primary specificity (Rawlings and Barrett, 1993; Perona and Craik, 1995; Ross *et al.*, 2003; Jiang *et al.*, 2005).

All of the predicted serine protease (SP) sequences were aligned, and their similarity to the S1 serine peptidase family, clan SA, was compared with conserved sequence motifs TAAHC, DIAL and GDSGGP, containing the conserved catalytic residues His57, Asp102, and Ser195 (chymotrypsin numbering, Rawlings and Barrett, 1993). All sequences contained a typical secretion signal peptide for digestive serine proteases, and trypsin activation was predicted for all sequences, with an Arg and mostly Ile (and in a few cases, Val) bridging the activation site (Figure 3.3A). Therefore, the hydrophobic residue at the N-terminus of the mature enzymes, predicted to hydrogen bond to the Asp194 residue preceding the active site Ser (Kraut, 1971), was conserved in all of the sequences. In addition, the molecular masses of the predicted proenzymes and mature peptides were within the range of other digestive serine proteases, and each sequence contained a single catalytic domain (Table 3.1). Because these cDNAs were from midgut libraries, the enzymes they encode are predicted to participate in digestive processes.

Nine clones, including AM1-01, AM3-26, AM4-47, AM4-49, PM4-08, PM2-57, PM4-54, PM4-63, and PM5-90, encoded proteins with a substitution of Gln for the catalytic His57, and with Gly, Val, or Leu substitutions for the active site serine (Ser195) residue (Table 3.1). Therefore, these proteins were predicted to be catalytically inactive and were classified as serine proteinase homologs (SPHs; Ross *et al.*, 2003). PM1-93 also was a SPH, with the catalytic His57 substituted by Ser, and the active site Ser changed to Val. Clone PM4-36 had very different active site residues, VILEDP, and the sequence was truncated one amino acid residue past this sequence. Clone AM1-62 lacked two of the three critical residues of the specificity pocket due to a truncated sequence. Therefore, PM4-36 and AM1-62 were categorized as SPHs.

In most of the *T. molitor* serine protease sequences, six cysteine residues, which participate in three disulphide bonds, were found at positions conserved in S1 family proteases (Figure 3.3A; Kraut, 1971). The exceptions were AM1-11, PM4-06, and S3-80, each with one additional cysteine residue, and all were predicted to be functional serine proteases. PM4-06 is closely related to a purified chymotrypsin from *T. molitor* larval PM (Elpidina *et al.*, 2005). Sequences missing one conserved cysteine were AM1-62, PM4-36, PM4-60, and PM2-57, and all but PM4-60 were predicted to be SPH on the basis of deletions or mutations. It is not known whether the missing cysteine residue affects the activity of PM4-60, tentatively classified as a serine protease. Although 78% of the serine proteases in *D. melanogaster* contained the TAAHC region (Ross *et al.*, 2003), only 54% of the *T. molitor* sequences had this motif, with 25% containing TAGHC, and the rest containing TSGHC or TSAHC. Those sequences lacking conservation in the Asp194 position (AM1-21, PM5-80 and PM4-36) were all SPHs. One sequence, L3-34, lacked the conserved Asp194 but retained the active site serine. This sequence was classified as a serine protease, but the Asp194 to Gly substitution may disrupt the normal tertiary structure needed for activity.

Overall, there was less conservation in the DIAL motif than in the other conserved sequences. The DIAL motif was found in 37% of SPs from *D. melanogaster* (Ross *et al.*, 2003), but in only one of the 24 *T. molitor* serine protease sequences (PM4-06, chymotrypsin). The remainder of the serine proteases contained the motifs DISV (all of the trypsins), DVAL, DVGL, DVGM, DIGM, or DIGL, and these same motifs also have been found in *D. melanogaster* serine protease and SPHs. The exception was clone L4-25, with an Asn in the position of Asp102, a substitution that had been demonstrated to reduce enzyme activity 104-fold (Craik *et al.*, 1987).

Similar to those of *D. melanogaster*, most *T. molitor* S1 family proteases contained the active site Ser residue (85 and 79%, respectively; Ross *et al.*, 2003). However, five *T. molitor* S1 family proteases (L4-24, S3-72, PM2-01, PM4-31, and PM4-60) had a Thr substitution for the active site Ser. The Ser to Thr change in evolution would have occurred with only a single base change, and it has been hypothesized that an intermediate peptidase may have contained Thr as a nucleophile (Barrett *et al.*, 1998). In *D. melanogaster*, there are five potential sequences (AAF58664, 53414, 22154, 49207 and 44895) with either the GDTGGP or GDTGSP motif (<http://www.ento.okstate.edu/labs/jiang/table1.htm>). Serine proteases from *L. cuprina* contain the Ser to Thr substitution, but the activity of these proteases was not documented (Elvin *et al.*,

1994). We have tentatively classified *T. molitor* sequences as serine protease analogs, but their functional activity needs to be evaluated.

Residues in the S1 binding pocket and two loops that connect the walls of the binding pocket determine substrate specificity (Hedstrom *et al.*, 1992). Trypsins have a conserved Asp189, whereas chymotrypsins have a Ser at that position. The functional predictions for *T. molitor* sequences correlate to these conserved amino acids, inasmuch as only the *T. molitor* trypsin in cluster 1 (Table 3.1) have this conserved Asp (Figure 3.3A). Correspondingly, the *T. molitor* chymotrypsin (PM4-06) has a Ser at that position. AM4-49 was the only other sequence with a Ser in the same position, but it was predicted to be a SPH because of the lack of conservation in other critical residues. The amino acids in loop1 and loop2 of the *T. molitor* serine proteases have considerable variation within this group and also with corresponding sequences in mammalian enzymes. Thus, it is difficult to predict substrate specificity.

With the exception of AM1-21, the mature enzyme form of the serine proteases and SPHs ranged from 20,847 to 29,178 daltons (Table 3.1). There was a wide range in the pI of these enzymes, from 4.17 to 8.54. Serine proteases and SPH from the AM had acidic pI values, from 4.17 to 6.91; those in the PM grouped into either acidic (4.21-5.08) or basic (8.26-8.92) proteins.

#### **3.4.3.2. Cysteine Proteases**

An alignment of cathepsin B-like proteinases from *T. molitor* was made with cathepsin B from *Diabrotica virgifera virgifera* (DvvcathB; Bown *et al.*, 2004) and *Araneus ventricosus* (AvcathB), all belonging to the C1 cysteine peptidase family, clan CA (Figure 3.3B). *T. molitor* sequences (AM4-18 and AM3-87) had signal peptides and activation sites, an indication that they are secreted proteins. In addition, AM3-87 had the conserved dyad residues Cys25 and His159 (papain numbering), and Gln19 and Asn/Asp175 residues that stabilize the molecule (Rawlings and Barrett, 1993), but AM4-18 lacked the conserved His159 as well as the residues C-terminal to the conserved Asn175. However, AM4-18 contained the two His residues (His110/111) in the occluding loop region of cathepsin B that blocks the C-terminal end of the active site cleft and causes the enzyme to act as a dipeptidase (Musil *et al.*, 1991). AM3-87 lacked the conserved His residues and most of the occluding loop. Therefore, AM3-87 may be related to primitive forms of cathepsin B, such as those found in protozoa (Ward *et al.*, 1997). There were 12 Cys residues in AM3-87, similar to typical cathepsin B proteases, but there were

18 Cys residues in AM4-18. There was one potential glycosylation site in AM4-18 (at residue 297), but none in AM3-87.

Cathepsin L proteases belong to the same family and clan as cathepsin B and contain the same active site residues (Rawlings and Barrett, 1993). *T. molitor* cathepsin L sequences were aligned with another *T. molitor* cathepsin L (TmcathL, Cristofolletti *et al.*, 2005; Figure 3.3C). All *T. molitor* cathepsin L enzymes had the conserved residues Cys25, His159, Gln19 and Asn/Asp175, and all contained the conserved trio of cysteine residues found in mammalian homologs, Cys22/63, Cys56/95, and Cys153/Cys200. AM4-72 had 13 Cys residue, whereas AM3-32 and AM4-22 had only 8 residues. There were no N-glycosylation sites in any of the *T. molitor* cathepsin L sequences. AM3-32 and AM4-22 were most similar to TmcathL, only differing in the N-terminal regions. Although AM4-72 was similar to these cathepsin L sequences, the sequence was highly divergent in the N-terminus, and lacked the typical activation site as well as signal peptide.

#### **3.4.3.3. Metalloproteases**

Predicted carboxypeptidase sequences from *T. molitor* were aligned with *Aedes aegypti* carboxypeptidase A (AeCPA), members of the clan MC and family M14 metallopeptidase family, with an HXXE zinc-binding motif (Figure 3.3D). As with all other midgut protease cDNAs, a signal peptide and an activation site were identified. The numbering of residues important for enzyme specificity is from *Bos taurus* carboxypeptidase A (BtCPA- Accession #POO730; Titani *et al.*, 1975). The residues in the catalytic zinc-binding site included His69, Glu72 and His196. Residues involved in catalytic binding and cleavage include Arg71, Arg127, Arg145, X255 and Glu270.

#### **3.4.4. Comparison of Sequences from Predicted Serine Proteases with Purified TmT1 and TmC1**

An N-terminal sequence was obtained from the major trypsin, TmT1, from the *T. molitor* larval midgut (Tsybina *et al.*, 2005). This N-terminal sequence was identical to the predicted N-termini of the mature enzymes of PM1-83, PM1-75, PM 2-70, PM1-95 and PM2-03. The N-terminus of the major *T. molitor* larval midgut chymotrypsin, TmC1 (Elpidina *et al.*, 2005), corresponded to only one predicted mature protease, PM 4-06. For further comparison of the corresponding sequences, mass spectra of tryptic peptides of each purified protease were

obtained (data not shown). The most prominent peaks in the mass range of 700-3000 kDa from both spectra were further subjected to MS/MS fragmentation to obtain the partial sequence needed to verify their correspondence to predicted cDNA sequences of mature enzymes.

The masses of four peptides from the tryptic digest of TmT1 and masses of their MS/MS fragments were identical to those of predicted peptides 92-102, 103-112, 236-249, and 250-258 from sequences PM1-75, PM2-70, and PM1-95 (Figure 3.4A). The predicted C-terminal peptides from PM1-83 and PM2-03 differed from that of TmT1. However, PM1-75, PM2-70, and PM1-95 had similar predicted amino acid sequences, molecular masses, and pI values that also were similar to those calculated for TmT1 (Tsybina *et al.*, 2005). Therefore, TmT1 may be the protein product of PM1-75, PM2-70, or PM1-95. Alternatively, the purified protein may represent a mixture of trypsin isoforms. The differences in amino acid sequences of PM1-75, PM2-70, and PM1-95 are located in long tryptic peptides and were not resolved by the mass spectral analysis. Therefore, cDNAs PM1-75, PM2-70, and PM1-95 have been tentatively identified as “*tmt1a*, *tmt1b* and *tmt1c*”, respectively (Figure 3.3A).

The TmC1 tryptic digest consisted of four peptide masses (Figure 3.4B). The MS/MS of the tryptic peptides from TmC1 were identical to 112-118, 150-171, 253-260, and the C-terminal 266-274 predicted peptides of PM4-06 (Figure 3.4B). Therefore, PM4-06 was labeled “*tmc1a*” and was predicted to be the gene encoding the previously purified protein TmC1.

#### ***3.4.5. Phylogenetic Analysis of Protease cDNAs from T. molitor***

Phylogenetic analyses of *T. molitor* serine protease and homolog sequences demonstrated distinct groups (Figure 3.5A). All sequences in the branch PM 1-75, PM1-83, PM1-95, PM2-70, and PM2-03 were predicted to be trypsins, homologous to the biochemically characterized digestive trypsin TmT1 (Tsybina *et al.*, 2005). This hypothesis was supported by the presence of conserved residues critical for specificity, DGG (Table 3.1). PM4-06 was predicted to be chymotrypsin due to similarities to TmC1 (Elpidina *et al.*, 2005) and because of the presence of conserved residues in the specificity pocket, SGS (Table 3.1). The lineage of the chymotrypsin cDNA, *tmc1a*, grouped with SPHs from AM and PM (AM4-49, PM5-90, AM1-01, and AM3-26), all with His57  $\Rightarrow$  Glu and Ser195  $\Rightarrow$  Gly substitutions. The distinction between trypsin and chymotrypsin presumably resides between these two groups.

A prediction for the phylogeny of *T. molitor* cathepsin B is presented in Figure 3.4B. *T. molitor* sequence AM4-18 was most related to *D. virgifera* cathepsin B, while AM3-87 had similarity to cathepsin B from *D. virgifera* as well as that of *A. ventricosus*.

AM4-72 is related in sequence to the previously identified cathepsin L in *T. molitor* (Cristofaletti *et al.*, 2005; Figure 3.4C). AM3-32 and AM4-22, from the same linkage, also were in this group.

The bootstrap consensus phylogenetic tree prediction for the *T. molitor* carboxypeptidase A sequences revealed strong support for the presence of two closely related groups within the same cluster (91-100% identity). With the *A. aegypti* carboxypeptidase A as the outgroup, the first group was AM 1-72, and the second group consisted of two subgroups, with sequences AM 1-30 in the first subgroup, and AM1-60, AM 1-02, AM 2-60, AM 2-51, L4-60 and AM 3-75 in the second subgroup.

A preliminary comparison of all the sequences in the *T. molitor* gut database with *Tribolium castaneum* sequences in the EST databases (<http://www.bioinformatics.ksu.edu/BeetleBase>) revealed that 45.2% of the total *T. molitor* gut sequences have homologs in *T. castaneum* (data not shown). Further comparison of *T. molitor* sequences indicated that 29.2% of *T. molitor* sequences have homologs in the *D. melanogaster* genome (<http://flybase.bio.indiana.edu/>).

### **3.4.6. Expression Analyses**

To investigate the expression patterns of individual proteases and homologs representative of grouped sequences in five developmental stages (first-, mid-, and late-instar larva, pupa, and adult) of *T. molitor*, we examined mRNA levels by Northern analysis (Figure 3.6, Table 3.3). The seventeen individual protease transcripts chosen for analysis had variable expression patterns. The expression patterns of individual representative proteases across clusters/groups indicated no correlation with the phylogenetic relationships between and among clusters. Among the serine proteases, only transcript L3-34 was expressed in the pupal (non-feeding) stage. Transcripts encoding PM2-70 (*tmt1b*), PM4-06 (*tmc1a*), AM2-68, PM5-80, PM4-54, PM1-93, and PM3-37 shared a similar expression pattern: expression levels were highest in the first-instar larvae, decreased slightly in subsequent larval stages, not expressed in pupae, and a moderate level of expression was observed in adults. Transcripts encoding AM 1-

11, AM 1-62 and L3-34 were only expressed in certain developmental stages. Other transcripts had variable expression patterns: There was a high level of AM 1-11 expression in first and late-instar larvae as well as in the adults; AM 1-62 was expressed only in late-instar larvae; whereas L3-34 was only expressed in first instar larvae, pupae, and adults. Their transient expression in specific developmental stages suggested a role of developmental regulation for these proteases. Transcripts encoding PM4-86, PM5-90, and AM4-47 were primarily expressed in first instar larvae, whereas those encoding L3-34 and PM2-01 exhibited a relatively high level of expression in first instar larvae as well as a moderate level of expression in adults. Among the three cysteine proteases analyzed, AM 3-87, a predicted cathepsin B, was expressed in all developmental stages including the pupa. AM 4-18 (cathepsin B) and AM 3-32, a predicted cathepsin L, were expressed at increased levels in first-instar larvae and adults and at moderate levels in other larval instars.

### 3.5. Discussion

Our results demonstrate that the majority of cysteine protease cDNAs were isolated from the anterior midgut cDNA library, and most serine protease transcripts were from the posterior midgut library. These data provide molecular evidence in support of previous observations that digestion in *T. molitor* larvae is compartmentalized (Thie and Houseman, 1990; Terra and Cristofolletti, 1996; Vinokurov *et al.*, 2006a). Trypsin and chymotrypsin transcripts from the PM were related to previously purified and biochemically characterized proteases from the PM, TmT1 and TmC1 (Tsybina *et al.*, 2005, Elpidina *et al.*, 2005). Three closely related transcripts, *tmt1a*, *tmt1b* and *tmt1c*, may constitute individually or simultaneously the major digestive trypsin(s) TmT1, responsible for 84% of the total trypsin activity in the PM. One transcript, *tmc1a*, was predicted to encode a protein with an N-terminus and MS/MS identical to TmC1. TmC1 represents as much as 42% of the total chymotrypsin activity in the PM, and other chymotrypsins evidently were not represented in randomly selected clones from the cDNA library.

There were no cDNAs corresponding to the AM “heavy” trypsins, with apparent molecular masses of 59 kDa, observed in our previous biochemical studies (Vinokurov *et al.*, 2006b). The lack of cDNAs encoding heavy trypsins supports the hypothesis that these enzymes are complexes resulting from the association of trypsin monomers under certain gut conditions

(Wagner *et al.*, 2002; Brito *et al.*, 2001). However, it is still possible that transcripts for these enzymes may be of low abundance. Alternatively, these heavy trypsins may have sequences that are not similar to known proteases.

Our study complements others involving cDNAs encoding cathepsin L enzymes as well as the biochemical characterizations of these enzymes in *T. molitor* (Cristofaletti *et al.*, 2005; Vinokurov *et al.*, 2006b). All cathepsin transcripts were localized to the AM and correspond to biochemical characteristics of cysteine proteinases in the AM (Vinokurov *et al.*, 2006a).

This is the first report of carboxypeptidase A cDNA transcripts in *T. molitor* although several studies have developed a molecular characterization of digestive carboxypeptidase A in other insect midguts (Ramos *et al.*, 1993; Bown *et al.*, 1998; Edwards *et al.*, 2000; Bown and Gatehouse, 2004). All cDNAs for carboxypeptidases were from the AM, which suggests that these enzymes are localized to the AM in *T. molitor* larvae. Further work is necessary to fully characterize *T. molitor* carboxypeptidases, and to determine if these enzymes are involved in digestion, or if they function in immunity or detoxification.

The sequences of *T. molitor* midgut proteases, in conjunction with protein characterization of enzymes, provide the basis for studies of the nature of the compensatory response of *T. molitor* to dietary protease inhibitors and other antinutritional compounds. The data obtained from the Northern blot analysis revealed the overall maximal expression of protease transcripts in the first-instar larvae of *T. molitor* when compared with other developmental stages. The first instar represents the most active phase of growth for *T. molitor* larvae.

Our results emphasize the importance of serine and cysteine proteases in the growth, physiology and development of *T. molitor*. Gene expression patterns in response to protease inhibitors have been studied in various lepidopteran insects (Bown *et al.*, 1997, 2004; Gatehouse *et al.*, 1997; Chougule *et al.*, 2005) and in some coleopterans (Michaud *et al.*, 1995; Zhu-Salzman *et al.*, 2003). Therefore, information on the protease profile in the midgut of *T. molitor*, which utilizes primarily two classes of proteases for digestion under normal dietary conditions, will significantly enhance the development of new targets for coleopteran pest control based on protease inhibitors.

### 3.6. References

- Applebaum, S.W., Birk, Y., Harpaz, I. and Bondi, A. (1964) Comparative studies on proteolytic enzymes of *Tenebrio molitor* L. *Comp. Biochem. Physiol.* 11: 85-103.
- Barrett, A. J., Rawlings, N. D., and Woessner, J. F. (1998) Introduction: family S1 of trypsin (clan SA). In Handbook of Proteolytic Enzymes (Barrett, A. J., Rawlings, N. D, and Woessner, J. F., eds.), Academic Press, New York, pp. 3-12.
- Bouchard, E., Cloutier, C. and Michaud, D. (2003) Oryzacystatin I expressed in transgenic potato induces digestive compensation in an insect natural predator via its herbivorous prey feeding on the plant. *Mol. Ecol.* 12:2439-2446.
- Bown, D.P., Wilkinson H.S. and Gatehouse J.A. (1997) Differentially regulated inhibitor-insensitive and insensitive protease genes from the phytophagous insect pest, *Helicoverpa armigera*, are members of complex multigene families. *Insect Biochem. Mol. Biol.* 27:625-638.
- Bown, D.P., Wilkinson H.S. and Gatehouse J.A.(1998) Midgut carboxypeptidase from *Helicoverpa armigera* (Lepidoptera: Noctuidae) larvae: enzyme characterisation, cDNA cloning and expression. *Insect Biochem. Mol. Biol.* 28:739-749.
- Bown, D.P., Wilkinson H.S. and Gatehouse J.A. (2004) Regulation of expression of genes encoding digestive proteases in the gut of a polyphagous lepidopteran larva in response to dietary protease inhibitors. *Physiolog. Entomol.* 29:278-290.
- Bown, D.P. and Gatehouse, J.A. (2004) Characterization of a digestive carboxypeptidase from the insect pest corn earworm (*Helicoverpa armigera*) with novel specificity towards C-terminal glutamate residues. *Eur. J. Biochem.* 271:2000-2011.
- Bown, D.P., Wilkinson, H.S., Jongsma, M.A. and Gatehouse, J.A. (2004) Characterisation of cysteine proteinases responsible for digestive proteolysis in guts of larval western corn rootworm (*Diabrotica virgifera*) by expression in the yeast *Pichia pastoris*. *Insect Biochem. Mol. Biol.* 34:305-320.
- Brito, L.O., Lopes, A.R., Parra, J.R.P., Terra, W.R. and Silva-Filho, M.C. (2001) Adaptation of tobacco budworm *Heliothis virescens* to proteinase inhibitors may be mediated by the synthesis of new proteinases. *Comp. Biochem. Physiol.*, 128B:365-375.

Campbell, C.L, Vandyke, K.A., Letchworth, G.J., Drolet, B.S., Hanekamp T. and Wilson, W.C. (2005) Midgut and salivary gland transcriptomes of the arbovirus vector *Culicoides sonorensis* (Diptera: Ceratopogonidae). *Insect Mol. Biol.* 14:121-136.

Casu, R. E., Jarmey, J. M., Elvin, C. M. and Eisemann, C. H. (1994). Isolation of a trypsin-like serine protease gene family from the sheep blowfly *Lucilia cuprina*. *Insect Mol. Biol.* 3:159-170.

Chougule, N.P., Giri, A.P., Sainani, M.N. and Gupta, V.S. (2005) Gene expression patterns of *Helicoverpa armigera* gut proteases. *Insect Biochem. Mol. Biol.* 35:355-367.

Craik, C.S., Roczniak, S., Largman, C. and Rutter, W.J. (1987) The catalytic role of the active site aspartic acid in serine proteases. *Science* 228: 291-297.

Cristofolletti, P.T. and Terra, W.R. (1999) Specificity, anchoring, and subsites in the active center of a microvillar aminopeptidase purified from *Tenebrio molitor* (Coleoptera) midgut cells. *Insect Biochem. Mol. Biol.* 29:807-819.

Cristofolletti, P.T. and Terra, W.R. (2000) The role of amino acid residues in the active site of a midgut microvillar aminopeptidase from the beetle *Tenebrio molitor*. *Biochim. Biophys. Acta* 1479:185-195.

Cristofolletti, P.T., Ribeiro, A.F. and Terra, W.R. (2005) The cathepsin L-like proteinases from the midgut of *Tenebrio molitor* larvae: Sequence, properties, immunocytochemical localization and function. *Insect Biochem. Mol. Biol.* 35: 883-901.

Edwards, M.J., Moskalyk, L.A., Donnelly-Doman, M., Vlaskova, M., Noriega, F.G., Walker, V.K. and Jacobs-Lorena M. (2000) Characterization of a carboxypeptidase A gene from the mosquito, *Aedes aegypti*. *Insect Mol. Biol.* 9: 33-38.

Elpidina, E.N., Tsybina, T.A., Dunaevsky, Y.E., Belozersky M.A., Zhuzhikov, D.P. and Oppert, B. (2005) A chymotrypsin-like proteinase from the midgut of *Tenebrio molitor* larvae. *Biochimie* 87: 771-779.

Elvin, C. M., Whan, V. and Riddles, P.W. (1993) A family of serine protease genes expressed in adult buffalo fly *Haematobia irritans exigua*. *Mol. Gen. Genet.* 240:132-139.

Elvin, C. M., Vuocolo, T., Smith, W. J., Eisemann, C. H. and Riddles, P.W. (1994) An estimate of the number of serine protease genes expressed in sheep blowfly larvae (*Lucilia cuprina*). *Insect Mol. Biol.* 3:105-115.

Ferriera C., Bellinello, G.L., Ribeiro, A.F. and W. R. Terra. (1990) Digestive enzymes associated with glycocalyx, microvillar membranes and secretory vesicles from midgut cells of *Tenebrio molitor* larvae. *Insect Biochem.* 20: 839-847.

Franco, O.L., Dos Santos, R.C., Batista, J.A.N., Mendes, A.C.M., De Araujo, M.A.M., Monnerat, R.G., Grossa-se-Sa, M.F. and De Freitas, S.M. (2003) Effects of black-eyed trypsin/chymotrypsin inhibitor on the proteolytic activity and on the development on *Anthonomus grandis*. *Phytochemistry* 63: 343-349.

Garty, N. (1979) Isolation and characterization of a chymotrypsin-like enzyme from *Tenebrio molitor* larvae. M.Sc. Thesis, Faculty of Agriculture, The Hebrew University, Rehovot, Israel.

Gatehouse, L.N., Shannon, A.L., Burgess, E.P.J. and Christeller, J.T. (1997) Characterization of major midgut proteinase cDNAs from *Helicoverpa armigera* larvae and changes in gene expression in response to four proteinase inhibitors in the diet. *Insect Biochem. Mol. Biol.* 27:929-944.

Girard, C., Le-Mateyar, M., Monade-Bottino, M., Pham-Delegue, M.H. and Jouanin, L. (1998) High level of resistance to proteinase inhibitors may be conferred by proteolytic cleavage in beetle larvae. *Insect Biochem. Mol. Biol.* 28:229-237.

Golan, R. (1981). Isolation, characterization and comparative study of proteolytic enzymes from the midguts of *Tenebrio molitor* adults and larvae as a basis of possible biological pest control with naturally occurring protease inhibitors from plant sources. M.Sc. Thesis, Faculty of Agriculture, The Hebrew University, Rehovot, Israel.

Hedstrom L., Szilagyi, L. and Rutter, W. J. (1992) Converting trypsin to chymotrypsin: The role of surface loops. *Science* 255: 1249-1253.

Hilder, V.A., Gatehouse A.M.R., Sheerman, S.E., Barker R.F. and Boulter D. (1987) A novel mechanism of insect resistance engineered from tobacco. *Nature* 330:160-163.

Holt, R. A., Subramanian, G. M., Halpern, A., Sutton, G. G., Charlab, R., Nusskern, D. R., Wincker, P., Clark, A. G., Ribeiro, J. M. Wides, R., *et al.* (2002) The genome sequence of the malaria mosquito *Anopheles gambiae*. *Science* 298: 129-149.

Huang, X. and Madan, A. (1999) CAP3: A DNA Sequence Assembly Program. *Genome Res.* 9:868-877.

International Human Genome Sequencing Consortium. (2001) Initial sequencing and analysis of the human genome. *Nature* 409: 860-921.

Jiang, H., Wang, Y., Gu, Y., Guo, X., Zou, Z., Scholz, F., Trenczek, T.E. and Kanost, M.R. (2005) Molecular identification of a bevy of serine proteinases in *Manduca sexta* hemolymph. *Insect Biochem. Mol. Biol.* 35:931-943.

Johnston, K.A., Lee, M.J., Brough, C.L., Hilder V.A., Gatehouse A.M.R. and Gatehouse J.A. (1995) Protease activities in the larval midgut of *Heliothis virescens*: Evidence for trypsin and chymotrypsin-like enzymes. *Insect Biochem. Mol. Biol.* 25:375-383.

Konarev, A.V. (1996) Interaction of insect digestive enzymes with plant protein inhibitors from plants and host-parasite coevolution. *Euphytica* 92:89-94.

Karlin, S. and Altschul, S.F. (1990) Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes. *Proc. Natl. Acad. Sci.* 87: 2264-2268.

Kraut, J. (1971) Chymotrypsinogen: X-ray structure. In *The Enzymes*, vol. 3 (Boyer, P. D., ed.), Academic Press, New York, pp. 165-183.

Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227:680-685.

Levinsky, H., Birk, Y. and Applebaum, S.W. (1977) Isolation and characterization of a new trypsin-like enzyme from *Tenebrio molitor* L. larvae. *Int. J. Peptide Protein Res.* 10: 252-264.

Michaud, D., Cantin, L. and Vrain, T.C. (1995) Carboxy-terminal truncation of oryzacystatin-II by oryzacytatin-insensitive insect digestive proteinases. *Arch. Biochem. Biophys.* 322: 469-474.

Mittapalli, O., Wise, I.L. and Shukle, R.H. (2006) Characterization of a serine carboxypeptidase in the salivary glands and fat body of the orange wheat blossom midge, *Sitodiplosis mosellana* (Diptera: Cecidomyiidae). *Insect Biochem. Mol. Biol.* 36: 154-160.

Müller, H. -M., Crampton, J. M., Torre, A D., Sinden, R. and Crisnti, A. (1993). Members of a trypsin gene family in *Anopheles gambiae* are induced in the gut by blood meal. *EMBO J.* 12: 2891-2900.

Musil, D., Zucic, D., Eng, R. A., Mayr, I., Huber, R., Popovič, T., Turk, V., Towatari, T., Katunuma, N. and Bode, W. (1991) The refined 2.15 Å x-ray crystal structure of human liver cathepsin B: the structural basis for its specificity. *EMBO J.* 10: 2321-2330.

Neurath, H. (1982) The diversity of proteolytic enzymes. In *Proteolytic Enzymes, a Practical Approach* (Beynon, R.J. and Bond, J.S. eds.), IRL Press, Oxford, pp. 1-23.

Oppert, B. (2000) Transgenic plants expressing enzyme inhibitors and the prospects for biopesticide development. In *Phytochemical Biopesticides* (Koul, O. and Dhaliwal, G. S., eds.), Harwood Academic, The Netherlands, pp. 83-95.

Oppert, B., Morgan, T.D., Culbertson, C. and Kramer, K.J. (1993) Dietary mixtures of cysteine proteinase and serine proteinase inhibitors exhibit increased toxicity toward the red flour beetle, *Tribolium castaneum*. *Comp. Biochem. Physiol.* 105C: 379-385.

Oppert, B., Morgan, T.D., Hartzler, K., Lenarcic, B., Galesa, K., Brzin, J., Turk, V., Yoza, K., Ohtsubo, K. and Kramer, K.J. (2003) Effects of proteinase inhibitors on growth and digestive proteolysis of the red flour beetle, *Tribolium castaneum* (Herbst) (Coleoptera: Tenebrionidae). *Comp. Biochem. Physiol.*, 134C: 481-490.

Oppert, B., Morgan, T.D., Hartzler, K. and Kramer, K.J. (2005) Compensatory proteolytic responses to dietary proteinase inhibitors in the red flour beetle, *Tribolium castaneum* (Herbst) (Coleoptera: Tenebrionidae). *Comp. Biochem. Physiol.* 140C: 53-58.

Patankar, A.G., Giri, A.P., Harsulkar, A.M., Sainani, M.N., Deshpande V.V., Ranjekar, P.K. and Gupta, V.S. (2001) Complexity in specificities and expression of *Helicoverpa armigera* gut proteinases explains polyphagous nature of the insect pest. *Insect Biochem. Mol. Biol.* 31: 453-464.

Pedra, J.H., Brandt, A., Westerman, R., Lobo, N., Li, H.-M., Romero-Severson, J., Murdock, L.L. and Pittendrigh, B.R. (2003) Transcriptome analysis of the cowpea weevil bruchid: Identification of putative proteinases and  $\alpha$ -amylases associated with food breakdown. *Insect Mol. Biol.* 12: 405-412.

Perona, J. J. and Craik, C.S. (1995) Structural basis of substrate specificity in the serine proteases. *Protein Sci.* 4: 337-360.

Ramos, A., Mahowald, A. and Jacobs-Lorena, M. (1993) Gut-specific genes from the black fly *Simulium vittatum* encoding trypsin-like and carboxypeptidase-like proteins. *Insect Mol. Biol.* 1:149-163.

- Rawlings, N. D. and Barrett, A. J. (1993) Evolutionary families of peptidases. *Biochem. J.* 290: 205-218.
- Rawlings, N. D., Morton, F. R. and Barrett, A.J. (2006) MEROPS: the peptidase database. *Nucl. Acids Res.* 34: D270-D272.
- Ross, J., Jiang, H., Kanost, M. R. and Wang, Y. (2003) SPs and their homologs in the *Drosophila melanogaster* genome: an initial analysis of sequence conservation and phylogenetic relationship. *Gene* 304:117-131.
- Rubin, G. M., Yandell, M. D. Wortman, J. R., Miklos, G. L. G., Nelson, C. R. *et al.* (2000) Comparative genomics of the eukaryotes. *Science* 287:2204-2215.
- Terra, W.R. and Cristofolletti, P.T. (1996) Midgut proteinases in three divergent species of Coleoptera. *Comp. Biochem. Physiol.* 113B:725-730.
- Terra, W.R., Ferreira, C. and Bastos, F. (1985) Phylogenetic consideration of insect digestion. Disaccharidases and the spatial organization of digestion in the *Tenebrio molitor* larvae. *Insect Biochem.* 15:443-449.
- Terra, W.R. and Ferreira, C. (1994) Insect digestive enzymes: properties compartmentalization and function. *Comp. Biochem. Physiol.* 109B:1-62.
- Thie, N.M.R. and Houseman, J.G. (1990) Cysteine and serine proteolytic activities in larval midgut of yellow mealworm, *Tenebrio molitor* L. (Coleoptera: Tenebrionidae). *Insect Biochem.* 20:741-744.
- Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucl. Acid Res.* 22: 4673-4680.
- Titani, K., Ericsson, L.H., Walsh, K.A. and Neurath, H. (1975) Amino-acid sequence of bovine carboxypeptidase B. *Proc. Natl. Acad. Sci. USA* 72:1666-1670.
- Tsybina, T.A., Dunaevsky, Y.E., Belozersky, M.A., Zhuzhikov, D.P., Oppert, B. and Elpidina, E.N. (2005) Digestive proteinases of yellow mealworm (*Tenebrio molitor*) larvae: Purification and characterization of a trypsin-like proteinase. *Biochemistry (Mosc.)* 70: 300-305.
- Urieli, N. (1982) Isolation and characterization of a chymotrypsin-like enzyme from the digestive tract of *Tenebrio molitor* and *Locusta migratoria*. In: Applebaum, S.W. (1985). *Biochemistry of digestion*. In: G.A. Kerkut and L.I. Gilbert, Editors, *Comprehensive Physiology, Biochemistry and Pharmacology of Insects*, Pergamon Press, Oxford, 4: pp. 279-311.

Vinokurov, K. S., Elpidina, E. N., Oppert, B., Prabhakar, S., Zhuzhikov, D. P., Dunaevsky, Y. E. and Belozersky, M. A. (2006a) Diversity of digestive proteinases in *Tenebrio molitor* (Coleoptera: Tenebrionidae) larvae. *Comp. Biochem. Physiol., Part B*, in press.

Vinokurov, K. S., Elpidina, E. N., Oppert, B., Prabhakar, S., Zhuzhikov, D. P., Dunaevsky, Y. E. and Belozersky, M. A. (2006b) Fractionation of digestive proteinases from *Tenebrio molitor* (Coleoptera: Tenebrionidae) larvae and role in protein digestion. *Comp. Biochem. Physiol., Part B*, in press.

Wagner, W., Möhrlein, F., and Schnetter, W. (2002) Characterization of the proteolytic enzymes in the midgut of the European Cockchafer, *Melolontha melolontha* (Coleoptera: Scarabaeidae). *Insect Biochem. Mol. Biol.* 32:803-814.

Ward, W., Alvarado, L., Rawlings, N. D., Engel, J. C., Franklin, C., and McKerrow, J. H. (1997) A primitive enzyme for a primitive cell: the protease required for excystation of *Giardia*. *Cell* 89: 437-444.

Xu, X., Dong, Y., Abraham, E.G., Kocan, A., Srinivasan, P., Ghosh, A.K., Sinden, R.E., Ribeiro, J.M., Jacobs-Lorena, M., Kafatos, F.C. and Dimopoulos, G. (2005) Transcriptome analysis of *Anopheles stephensi-Plasmodium berghei* interactions. *Mol. Biochem. Parasitol.* 142: 76-87.

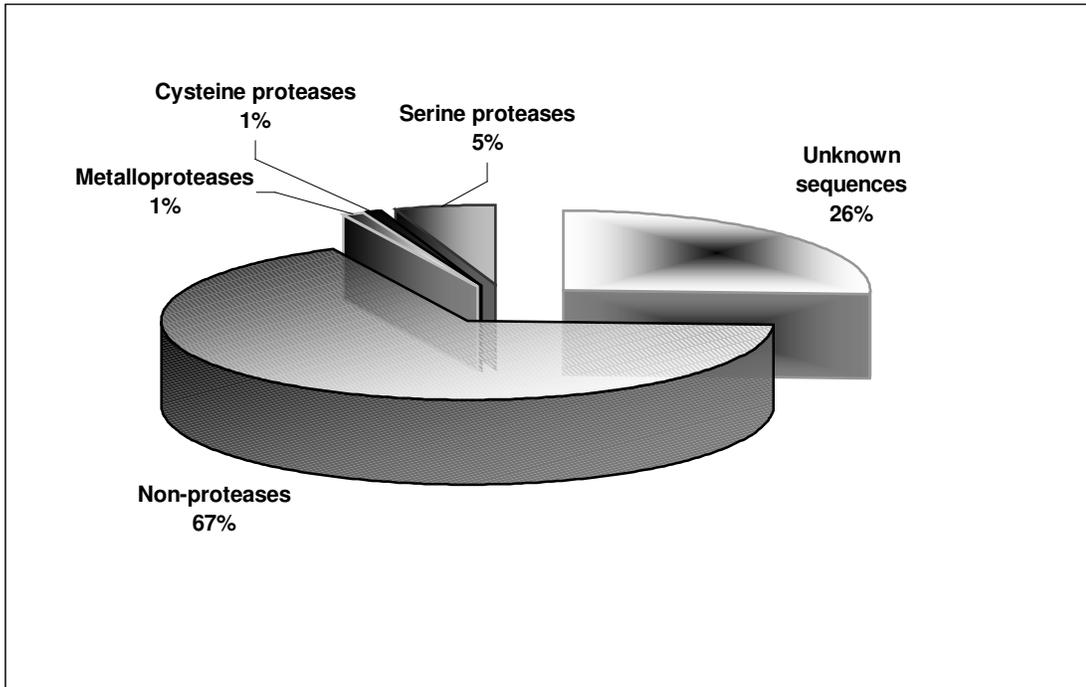
Zdobnov, E. M., Von Mering, C., Letunic, I., Torrents, D., Suyama, M., Copley, R. R., Christophides, G. K., Thomasova, D., Holt, R. A., Subramanian, G. M., *et al.* (2002) Comparative genome and proteome analysis of *Anopheles gambiae* and *Drosophila melanogaster*. *Science* 298: 149-159.

Zhu-Salzman, K., Koiwa, H., Salzman, R.A., Shade, R.E. and Ahn, J.-E. (2003) Cowpea bruchid *Callosobruchus maculatus* uses a three-component strategy to overcome a plant defensive cysteine protease inhibitor. *Insect Mol. Biol.* 12:135-145.

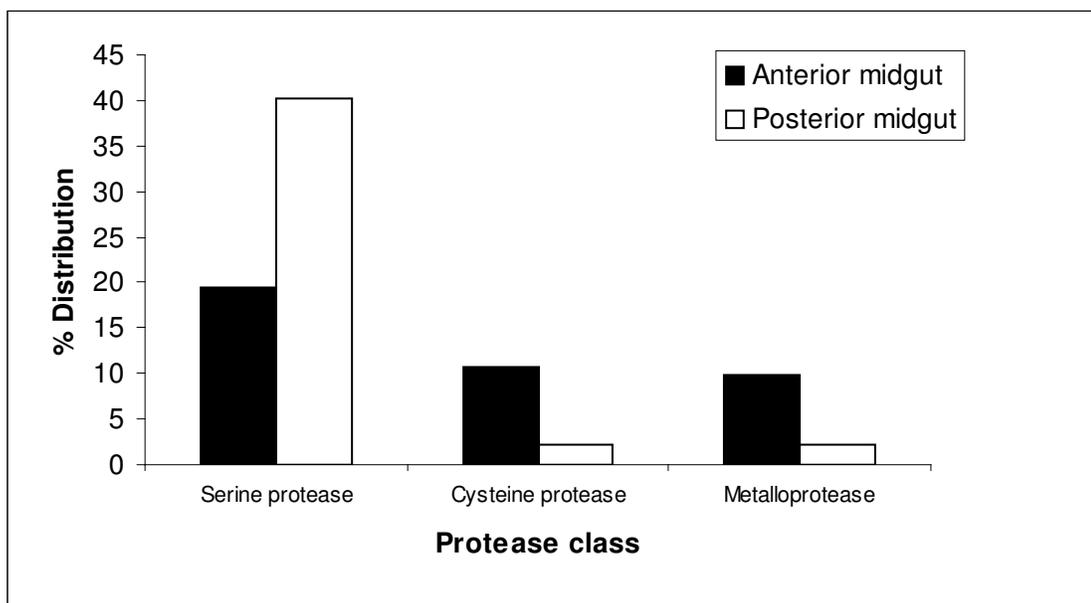
Zwilling, R. (1968) Zur Evolution der Endopeptidasen-IV.  $\alpha$ - and  $\beta$ -protease aus *Tenebrio molitor*. *Z. Physiol. Chem.* 349:326-332.

Zwilling, R., Medugorac, I. and Mella K. (1972) The evolution of endopeptidasen-XIV. Non-tryptic cleavage specificity of a BAEE-hydrolyzing enzyme ( $\beta$ -protease) from *Tenebrio molitor*. *Comp. Biochem. Physiol.* 43: 419-424.

**Figure 3.1. Categories of sequences from randomly picked clones of cDNA libraries from *T. molitor* larvae (BLASTX matches with  $E < 10$ )**

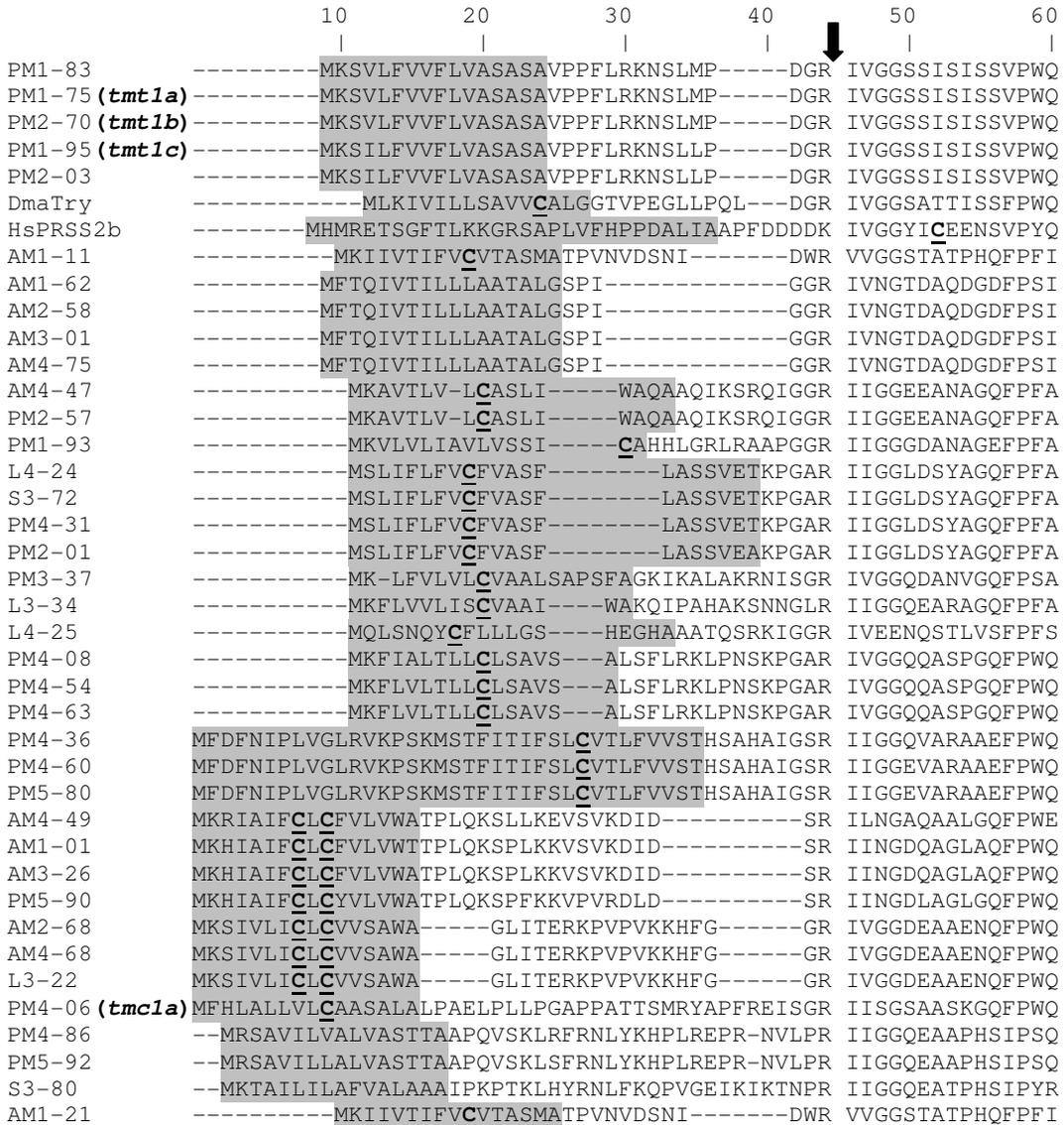


**Figure 3.2. Distribution of sequences encoding digestive proteases from the anterior and posterior midgut cDNA libraries of *T. molitor***



**Figure 3.3. Alignment of predicted proteases from *T. molitor* larvae** A) Alignment of serine proteases from *T. molitor* with trypsins from *Drosophila melanogaster* (Dmatryp, accession # CAA26732) and *Homo sapiens* (HsPRSS2b, accession # AAL14244); B) Alignment of predicted cathepsin B sequences with cathepsin B proteins from *Diabrotica virgifera virgifera* (DvvcathB accession #CAE47498) and *Araneus ventricosus* (AvcathB, accession # AAP59456); C) Alignment of predicted cathepsin L sequences with cathepsin L from *T. molitor* (TmcathL, accession # AAP94048); D) Alignment of predicted carboxypeptidase sequences with carboxypeptidase A from *Aedes aegypti* (AeCPA, accession #AAT36730) and *Bos taurus* (BtCPA- Accession# POO730). Shaded regions at the N-terminus are predicted signal peptides and within the sequences represent potential N-glycosylation sites; bolded sequences are conserved cysteines or critical residues; active site residues are boxed; arrow indicates activation site; stars indicate residues that determine specificity; loop regions are indicated in brackets.

**Figure 3.3A**



	70	80	90	100	110	120
PM1-83	ISLQY-Y--GSHICGGSIISANYVV	TAAHCTDGLTAGSLTVRAGTSTRGSGGQV	---	VNV		
PM1-75 ( <i>tmt1a</i> )	ISLQY-Y--GSHICGGSIISANYVV	TAAHCTDGLTAGSLTVRAGTSTRGSGGQV	---	VNV		
PM2-70 ( <i>tmt1b</i> )	ISLQY-Y--GSHICGGSIISANYVV	TAAHCTDGLTAGSLTVRAGTSTRGSGGQV	---	VNV		
PM1-95 ( <i>tmt1c</i> )	ISLQY-Y--GSHICGGSIISANYIV	TAAHCTDGLTAGSLTVRAGTSTRGSGGQV	---	VNV		
PM2-03	ISLQY-Y--GSHICGGSIISANYIV	TAAHCTDGLTAGSLTVRAGTSTRGSGGQV	---	VNV		
DmaTry	ISLQR-S--GSHSCGGSIYSANIIV	TAAHCLQSVSASVLQVRAGSTYWSGGV	---	AKV		
HsPRSS2b	VSLN--S--GYHFCGGSLISEQWV	SAGHCYKSRIQVRLGEHNIEVLE	---	INA		
AM1-11	VSLRTPY--DSHNCGGSIIAKNYVI	TAAHCVSGYAPSYTIVVAGTNQLN	---	ATNPLR		
AM1-62	VSVRF-L--NSHNCGGSILNERYIL	TAAHCVVSYFASFLSVQYDVTTI	---	SSGNAPNVLKV		
AM2-58	VSVRF-L--NSHNCGGSILNERYIL	TAAHCVVSYFASFLSVQYDVTTI	---	SSGNAPNVLKV		
AM3-01	VSVRF-L--NSHNCGGSILNERYIL	TAAHCVVSYFASFLSVQYDVTTI	---	SSGNAPNVLKV		
AM4-75	VSVRF-L--NSHNCGGSILNERYIL	TAAHCVVSYFASFLSVQYDVTTI	---	SSGNAPNVLKV		
AM4-47	AAIYKSTADGTYFCTGALMNTQWII	TAGQCV	---	GTLFTIRLGSNSLNSNDPNALRLST		
PM2-57	AAIYNSTADGTYFCTGALMNTQWII	TAGQCV	---	GTLFTIRLGSNSLNSNDPNALRLSA		
PM1-93	AAIQSQTINAGQYFCGGALVNTLFI	VTSGHCV	---	ATLFSIRLGITSLAETGQ		
L4-24	AAINVQTADSRFFCGGALLNHNWVI	TSGHCV	---	ATIFTIQLGSNTLTSADPDREIFST		
S3-72	AAINVQTADSRFFCGGALLNHNWVI	TSGHCV	---	ATIFTIQLGSNTLTSADPDREIFST		
PM4-31	AAINVQTADSRFFCGGALLNHNWVI	TSGHCV	---	ATIFTIQLGSNTLTSADPDREIFST		
PM2-01	AAINVQTADSRFFCGGALLNHNWVI	TSGHCV	---	ATIFTIQLGSNTLTSADPDREIFST		
PM3-37	AAIYVSTSTGTGYFCGGSLISQWV	TAAHC	---	IAG--GVAFQVILGSNTLKGTDPNRKT		
L3-34	AAITVQTETSQFFCGGALINNQWV	ITSAHCV	---	AVTIVIRLGSNNLQSGDPNRRITVAS		
L4-25	AAIYVQAASSTFFCGGALINNQWV	ITSAHCV	---	AVTIVIRLGSNSLVDSPNRRITVAS		
PM4-08	AAIYKYTADGRYFCGGTLYNEQWIL	TAGQCV	---	VID--ATEFTIQLGSNQLDSTDNNR		
PM4-54	AAIYKYTADGRYFCGGTLYNEQWIL	TAGQCV	---	VID--ATEFTIQLGSNQLDSTDNNR		
PM4-63	AAIYKYTADGRYFCGGSLFNEQWIL	TAGQCV	---	VID--ATEFTIQLGSNQLDSTDNNR		
PM4-36	VAIYVDTVDGKFFCGGSLNREWIL	TAAHCLYN--	---	GRLYTIQLGSTTLQSGDANR		
PM4-60	VAIYVDTVDGKFFCGGSLNREWIL	TAAHCLYN--	---	GRLYTIQLGSTTLQSGDANR		
PM5-80	VAIYVDTVDGKFFCGGSLNREWIL	TAAHCLYN--	---	GRLYTIQLGSTTLQSGDANR		
AM4-49	AALYVNIGTTTYFCSGNIISEEWIL	TAAQC	---	IIG--ADSIDVLAGLIDL---NGSGTVARG		
AM1-01	AALFI---GSYFVCSGSIISEEWIL	TAAQC	---	IEG--TSTVTVLGIVDL---NGTGAVVQS		
AM3-26	AALYI---GSYFVCSGSIISEEWIL	TAAQC	---	IDG--VGTVTVLGIVDL---NGSGAVVQS		
PM5-90	AAVYVPLGSSYFVCGGSIISEQWIL	TAAQC	---	IYG--VDSVTILAGVVDL---NGSGASAQS		
AM2-68	VAVYFDTS	SDGTYFCGGALVAENWVLTAGHC	---	VYH--AKVFTLHLGSNSLVDDDDNRVTLGA		
AM4-68	VAVYFDTS	SDGTYFCGGALVAENWVLTAGHC	---	VYH--AKVFTLHLGSNSLVDDDDNRVTLGA		
L3-22	VAVYFDTS	SDGTYFCGGALVAENWVLTAGHC	---	VYH--AKVFTLHLGSNSLVDDDDNRVTLGA		
PM4-06 ( <i>tmt1a</i> )	AALYLTVSGGTSFCGGALISSNWIL	TAAHCTQ	---	G--VSGITAYLGVVSL--SDSSRVTAQA		
PM4-86	AFLEMYTENEGWYCGGSLISENYVL	TAGHC	---	GED--VVKAMVSLGAHALSESVEGEITVDS		
PM5-92	AFLEMYTENEGWYCGGSLISENYVL	TAGHC	---	GED--VVKAVVALGAHALSESVEGEITVDS		
S3-80	TFLEVYSDSEGWYCGGSLISENYVL	TAGHC	---	GED--AVEAHVTLGAHKPLQTEDTQVQSVS		
AM1-21	VSLRTPY--DSHNCGGSIIAKNYVI	TAAHCV	---	SGYAPSYTIVVAGTNQLNATTPFRPFKSW		

	130	140	150	160	170	180
PM1-83	ARINQNPSYN-DRVIDY	<b>DISV</b> LQLSSSLSLGSSVAAVGL	--PSSST	SWSAGT	SVLVT	<b>GWG</b>
PM1-75 ( <i>tmt1a</i> )	ARINQNPSYN-DRVIDY	<b>DISV</b> LQLSSSLSLGSSVAAVGL	--PSSST	SWSAGT	SVLVT	<b>GWG</b>
PM2-70 ( <i>tmt1b</i> )	ARINQNPSYN-DRVIDY	<b>DISV</b> LQLSSSLSLGSSVAAVGL	--PSSST	SWSAGT	SVLVT	<b>GWG</b>
PM1-95 ( <i>tmt1c</i> )	ARINQNPSYN-DRLIDY	<b>DISV</b> LQLSSSLSLGSSVAAVGL	--PSSST	SWSAGT	SVLVT	<b>GWG</b>
PM2-03	ARINQNPSYN-DRLIDY	<b>DISV</b> LQLSSSLSLGSSVAAVGL	--PSSST	SWSAGT	SVLVT	<b>GWG</b>
DmaTry	SSFKNHEGYN-ANTMVN	<b>DIAV</b> IRLSSSLSFSSSIKAI	SL--ATYNPA	--NGASA	AAV	<b>SWG</b>
HsPRSS2b	AKIIRHPKYN-SRTLDN	<b>DILL</b> IKLSSPAVINSRVSAISL	--PTAPPA	--AGTESL	IS	<b>IGWG</b>
AM1-11	AQIIVHPEYS-SSLILN	<b>DVAL</b> LRLLETPIEESEEVQIVGL	--ETEYVD	--TVRD	<b>CVL</b>	<b>IGWG</b>
AM1-62	SSVIYNKDYTPGNGYIN	<b>DVAV</b> LKLQSPIIFGTNARPIKL	--PVAFN	STPENS	PAEL	<b>GGWG</b>
AM2-58	SSVIYNKDYTPGNGYIN	<b>DVAV</b> LKLQSPIIFGTNARPIKL	--PVAFN	STPENS	PAEL	<b>GGWG</b>
AM3-01	SSVIYNKDYTPGNGYIN	<b>DVAV</b> LKLQSPIIFGTNARPIKL	--PVAFN	STPENS	PAEL	<b>GGWG</b>
AM4-75	SSVIYNKDYTPGNGYIN	<b>DVAV</b> LKLQSPIIFGTNARPIKL	--PVAFN	STPENS	PAEL	<b>GGWG</b>
AM4-47	DTYFVHPEYDPL-TLIN	<b>DIGL</b> IKLRIAITLTDYISPIS	---LLAGST	LPDSS	SVLT	<b>IGWG</b>
PM2-57	DTYFVHPEYDPL-TLIN	<b>DIGL</b> IKLRIAITLTDYISPIS	---LLAGST	LPDSS	SVLT	<b>IGWG</b>
PM1-93	DKYVLHPLYNID-TLEN	<b>DIGV</b> LRLPVSFTDYIQPIG	---MPTRD	VQTN	AI	<b>AIGWG</b>
L4-24	NDYVIHPDFVPD-TIEN	<b>DIGL</b> IKLRLPVSFTSYIQPIN	---LPTVG	-LLNET	QVTAL	<b>GWG</b>
S3-72	NDYVIHPDFVPD-TIEN	<b>DIGL</b> IKLRLPVSFTSYIQPIN	---LPTVS	-LLNET	QVTAL	<b>GWG</b>
PM4-31	NDYVIHPDFVPD-TIEN	<b>DIGL</b> IKLRLPVSFTSYIQPIN	---LPTVS	-LLNET	QVTAL	<b>GWG</b>
PM2-01	NDYVIHPDFVPD-TIEN	<b>DIGL</b> IKLRLPVSFTSYIQPIN	---LPTVS	-LLNET	QVTAL	<b>GWG</b>
PM3-37	SIYVNHPDFNPD-TLEN	<b>DIGL</b> VKFLPIEYNDYIQPVY	---LPTVD	-LIDNL	GNTAI	<b>IGWG</b>
L3-34	SHVVPHPDFDPD-TSVN	<b>DIGI</b> VKL RMPVEFTDYIQPIN	---LASTP	-LPNSA	AAPT	<b>AI</b>
L4-25	SHYVAHPDYDPL-TLEH	<b>NI</b> GLIALRLPIQFTGYIQPIQ	---LTDKE	-ITTYN	HLTA	<b>IGWG</b>
PM4-08	TTYVVEPRFDPVSLRH	<b>DVGM</b> IKLPSVTVNDYIQPVRM	--LESMS	PIYKGV	AVETA	<b>GWG</b>
PM4-54	TTYVYVHPSFDPVSLRH	<b>DIGM</b> IKLSSPVTLTDYIQPVRM	--LESMS	PIYKGV	SVETA	<b>GWG</b>
PM4-63	TTYVYVHPSFDPVSLRH	<b>DIGM</b> IKLSSPVTLTDYIQPVRM	--LESMS	PIYKGV	SVETA	<b>GWG</b>
PM4-36	STAVIFPNFDPE-TLEH	<b>DIGL</b> IKLHMEITLTDYIQPIT	---LAEVGD	TVEG	MPAIA	<b>AVGWG</b>
PM4-60	STAVIFPNFDPE-TLEH	<b>DIGL</b> IKLHMEITLTDYIQPIS	---LAEVGD	TVEG	MPAIA	<b>AVGWG</b>
PM5-80	STAVIFPNFDPE-TLEH	<b>DIGL</b> IKLHMEITLTDYIQPIS	---LAEVGD	TVEG	MPAIA	<b>AVGWG</b>
AM4-49	TEIVLHGDDYDPD-AFNN	<b>DIGL</b> IKLSTPITFNVNVAPIA	----LAETL	LEDG	IDVR	<b>SVSWG</b>
AM1-01	SELILHKDYDPY-NFIN	<b>DIGL</b> VQLSTPLTFTRYLAPIA	----LADNL	LLEDG	LDVTV	<b>SWG</b>
AM3-26	SDLILHKDYDPD-YFLN	<b>DIGL</b> VQLRTPLTFTRYLAPIA	----LADNL	LLEDG	LDVTV	<b>ISWG</b>
PM5-90	SKLIVHNDYKLD-IPDN	<b>DIGL</b> VQLSTPLTFNQYVAAIT	----LAENL	LLEDG	VNVTV	<b>SWG</b>
AM2-68	SYSVPHPDYDPS-DLEN	<b>DIGL</b> IRIDTAYKTNDHIKVIP	----LASSEL	GADV	VDVIV	<b>SWG</b>
AM4-68	SYSVPHPDYDPS-DLEN	<b>DIGL</b> IRIDTAYKTNDHIKVIP	----LASSEL	GADV	VDVIV	<b>SWG</b>
L3-22	SYSVPHPDYDPS-DLEN	<b>DIGL</b> IRIDTAYKTNDHIKVIP	----LASSEL	GADV	VDVIV	<b>SWG</b>
PM4-06 ( <i>tmt1a</i> )	SRVVAHPSYSSS-TLAN	<b>DIAL</b> IQLSTSVATSTNIRTIS	----LSSST	LGTG	ASVTV	<b>SWG</b>
PM4-86	QDVTVHADYDGN-VIIN	<b>DI</b> AVIKLPEPVTLSDTIQPVALPTTADVDNTFTGEEARV	<b>SWG</b>			
PM5-92	QDVTVHADYDGN-VIIN	<b>DI</b> AVIKLPEPVTLSDTIQPVALPTTADVDNTFTGEEARV	<b>SWG</b>			
S3-80	KDIKIHEDYDGD-QVIN	<b>DVGL</b> IKPPESVTLNDAIKPVTLPSKADADNDFAGETARV	<b>SWG</b>			
AM1-21	AQNL-----					

	190	200	210	220	230	240
PM1-83	TTTEGSSSLPSALQGVNVQIVSQST	C-SSAYGSGS-ITDRMLCAGVTGGGKDA	C	C	C	CGDSGG
PM1-75 ( <i>tmt1a</i> )	TTTEGSSSLPSALQGVNVQIVSQST	C-SSAYGSGS-ITDRMLCAGVTGGGKDA	C	C	C	CGDSGG
PM2-70 ( <i>tmt1b</i> )	TTTEGSSSLPSALQGVNVQIVSQST	C-SSAYGSGS-NTDRMLCAGVTGGGKDA	C	C	C	CGDSGG
PM1-95 ( <i>tmt1c</i> )	TTTEGSSSLPSALQGVNVQIVSQST	C-SSAYGSGS-ITDRMLCAGVTGGGKDA	C	C	C	CGDSGG
PM2-03	TTTEGSSSLPSALQGVNVQIVSQST	C-SSAYGSGS-ITDRMLCAGVTGGGKDA	C	C	C	CGDSGG
DmaTry	TQSSGSSSIPSQLQYVNVNIVSQSQ	CASSTYGYGSIQIRNTMI	CAAAASG--KDA	C	C	CGDSGG
HsPRSS2b	NTLSSGADYPDELQCLDAPVLSQAE	CEASYPGK---ITNNMF	CVGFLEGGKDS	C	C	CGDSGG
AM1-11	RTSYPGS-IPNDLQFLNERTYPNDEC	-VSRWASAHAVYSSQICT	-LXKVGEGACH	C	C	CGDSGG
AM1-62	LPYSSGT-VMTHLQIVNITVFSDD	ECERIHAQTGPTSRKYHV	CAGVPQGGKQC	C	C	CGDSGG
AM2-58	LPYSSGT-VMTHLQIVNITVFSDD	ECERIHAQTGPTSRKYHV	CAGVPQGGKQC	C	C	CGDSGG
AM3-01	LPYSSGT-VMTHLQIVNITVFSDD	ECERIHAQTGPTSRKYHV	CAGVPQGGKQC	C	C	CGDSGG
AM4-75	LPYSSGT-VMTHLQIVNITVFSDD	ECERIHAQTGPTSRKYHV	CAGVPQGGKQC	C	C	CGDSGG
AM4-47	QIDDETAGLVDALNYYVLTLSNEE	C-RLAFG-DQVNDNMV	VDG-NYN-QGT	CF	CF	GDLS
PM2-57	QIDDETAGLVDALNYYVLTLSNEE	C-RLAFG-DQVNDNMV	VDG-NYN-QGT	CF	CF	GDLS
PM1-93	QIGDADAGLTNQLQTVLVALSNEE	C-RLTFG-NQIADTMV	VDG-NYN-EGAC	CI	CI	GDVGS
L4-24	QTSDDSDSALSETLQYVSATILSNA	AC-RLVYG-NQITDNMA	VEG-NYN-EGT	CI	CI	GDTGS
S3-72	QTSDDSDSALSETLQYVSATILSNA	AC-RLVYG-NQITDNMA	VEG-NYN-EGT	CI	CI	GDTGS
PM4-31	QTSDDSDSALSETLQYVSATILSNA	AC-RLVYG-NQITDNMA	VEG-NYN-EGT	CI	CI	GDTGS
PM2-01	QTSDDSDSALSETLQYVSATILSNA	AC-RLVYG-NQITDNMA	VEG-NYN-EGT	CI	CI	GDTGI
PM3-37	QTSDENAGIVNELNYVTVTISNAE	C-QLSYG-NTIFDTMV	CVAG-NYN-ERT	CF	CF	GDSGS
L3-34	QTSDDDPEMSNGLNYYVGLAVLSNEE	C-RMVG-NQLTDDMV	VEG-NFN-ERAC	CI	CI	GDSGS
L4-25	QTSADPELSDHLQYVSLITITNEE	C-KNVYG-FQVSDDMI	CATG-NYI-EGT	CI	CI	GDTGS
PM4-08	QTAD-SGDIVNDLNYVQLKIIANTE	C-QSYYG-DQFFGSMTC	TEGANYN-EGF	CF	CF	GDVGG
PM4-54	QTSN-NGDLVNDLNYVQLKIIANAE	C-KTYYG-NQFWGTMT	CTEGSNYN-EGF	CF	CF	GDVGG
PM4-63	QTSN-NGDLVNDLNYVQLKIIANAE	C-KTYYG-NQFWGTMT	CTEGSNYN-EGF	CF	CF	GDVGG
PM4-36	QISDSL SGLANDLHYVTVVVISNAE	C-RLTYG-DQVKSTMF	CTVG-NYN-EEFAL	CI	CI	VILED
PM4-60	QISDSL SGLANDLHYVTMVVISNAE	C-RLTYG-DQVKSTMF	CTVG-NYN-EGT	CI	CI	GDTGG
PM5-80	QISDSL SGLAHLHYVTMVVISNAE	C-RLTYG-DQVKSTMF	CTVG-NYN-EGT	CI	CI	GDTGG
AM4-49	ATSD-VGGVSEFLSYVDLVTIRNSE	C-IAVYG-NTIVDSIV	CAQSATALLKSV	CF	CF	GDGGS
AM1-01	ATVS-GGNEQLLLYADLVTIRNSE	C-TAIYG-NTILDSIV	CAESGTAVLKNI	CI	CI	GDGGA
AM3-26	ATDS-DGDESQFLNYADLVTIRNRE	C-TAIYG-N-IQDSSV	CAKSETVTVQNA	CI	CI	GDGGA
PM5-90	ATSD-DDDENQLLYVDLVTIRNSE	C-TAIYG-S-IQDSSV	CAESGTATVKN	CI	CI	GDGDD
AM2-68	ASGDWDG-VENHLRFVGLKTLNSDD	C-KAIYGEAVITDGMV	CAVGPNSS--EGT	CF	CF	GDSGG
AM4-68	ASGDWDG-VENHLRFVGLKTLNSDD	C-KAIYGEAVITDGMV	CAVGPNSS--EGT	CF	CF	GDSGG
L3-22	ASGDWDG-VENHLRFVGLKTLNSDD	C-KAIYGEAVITDGMV	CAVGPNSS--EGT	CF	CF	GDSGG
PM4-06 ( <i>tmt1a</i> )	RTSDSSSSISQTLNYYVGLSTISNTV	C-ANTYG-SIIQSGIV	CCTGSTI--QST	CF	CF	GDSGG
PM4-86	LTEGFDEIFSDVFNYYVDVKVISNEE	C-FRDYD--NVIDSIL	CTSGDART--GS	CF	CF	GDSGG
PM5-92	LTDGFDEILSDVFNYYVDVKVISNEE	C-LRDYD--NVIDSIL	CTSGDART--GS	CF	CF	GDSGG
S3-80	LTDGFDTDLSEVLNYYVDEVISNEK	C-EDTFG--SLVPSIL	CTSGDAYT--GS	CF	CF	GDSGG
AM1-21	-----	-----	-----	-----	-----	-----

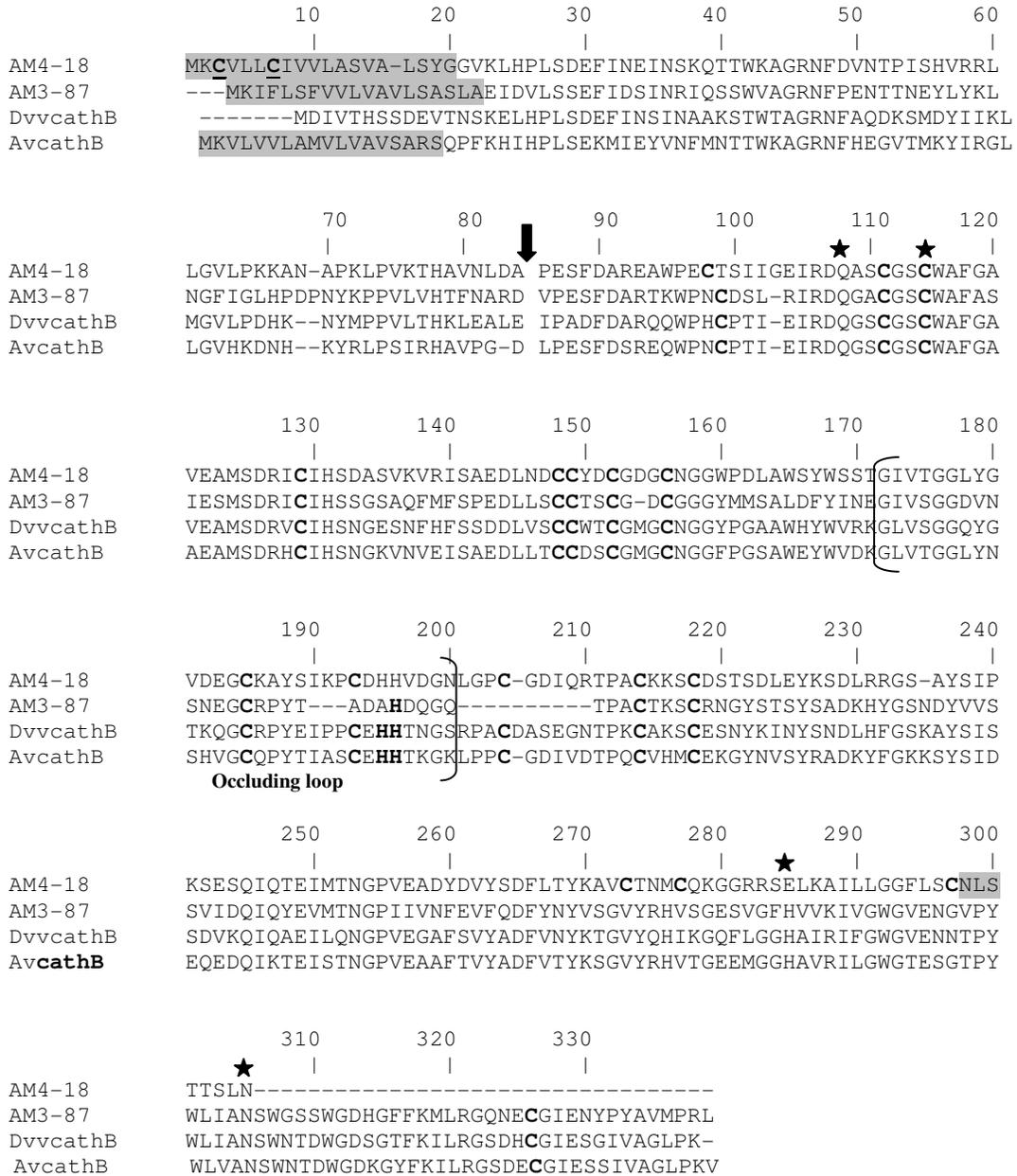
*Loop 1*

	250	260	270	280	290	300
PM1-83	PLVVGN	---VL--AGIVSWG	Y-GCARNG---	YPGVYSNVPALRSYFQQTAGI	-----	
PM1-75 ( <i>tmt1a</i> )	PLVVGN	---VL--AGIVSWG	Y-GCARNG---	YPGVYSNVPALRSYIQQTAGI	-----	
PM2-70 ( <i>tmt1b</i> )	PLVVGN	---VL--AGIVSWG	Y-GCARNG---	YPGVYSNVPALRSYIQQTAGI	-----	
PM1-95 ( <i>tmt1c</i> )	PLVVGN	---VL--AGIVSWG	Y-GCARNG---	YPGVYSNVPALRSYIQQTAGI	-----	
PM2-03	PLVVGN	---VL--AGIVSWG	Y-GCARNG---	YPGVYPNVPALRSYIQQPPEYKLFKYL	-----	
DmaTry	PLVSGG	---VL--VGIVSWG	Y-GCAYSN---	YPGVYADVAVLRSWVSTANSI	-----	
HsPRSS2b	PVVSNG	---EL--QGIVSWG	Y-GCAQKN---	RPGVYTKVYNYVDWIKDTIAANS	-----	
AM1-11	PLVVVVKDDKFSL	--IALVSWG	S-PGAR-G---	MPDVYTRVASFHEFITDNIKN	-----	
AM1-62	PLVVNG	---VQ--VGIVS			-----	
AM2-58	PLVVNG	---VQ--VGIVSWSVKP	ITVKG---	YPGVFTKVSSQVPWILEQIEDKN	-----	
AM3-01	PLVVNG	---VQ--VGIVSWSVKP	ITVKG---	YPGVFTKVSSQVPWILEQIGGV	-----	
AM4-75	PLVVNG	---VQ--VGIVSWSVKP	ITVKG---	YPGVFTKVSSQVPWILEQIGGV	-----	
AM4-47	PLIQYGGSSLIY	--HVGVS	FISSNGCEST---	DPSGFTRTAPYIEWLNNVTNN	-----	
PM2-57	PLIQYGGSSLIY	--HVGVS	FISSNGCEST---	DPSGFTRTAPYIEWLNNVTNN	-----	
PM1-93	PLIQYVNGHRAT	--IVGISS	FISSNGCEST---	DPSGYTRAFPVEWIQNTQVPS	-----	
L4-24	PLVEYLSRLYWI	--VGVSS	FLSNGCEST---	DPSGYTRIFPYTDWIRTIINP	-----	
S3-72	PLVEYLSRLYWI	--VGVSS	FLSNGCEST---	DPSGYTRIFPYTDWIRTIINP	-----	
PM4-31	PLVEYLSRLYWI	--VGVSS	FLSNGCEST---	DPSGYTRIFPYTDWIKTIINP	-----	
PM2-01	PLVEYLSRLYWI	--VGVSS	FLSNGMGA	KVLILQDIQEYSHIQIGSD-PSSILNL	CL-----	
PM3-37	PLMTTLNHHHWT	--VGVAS	FISTNGCEST---	DPSGYTRTFPYVDWIKTTAEIV	-----	
L3-34	PLVVRLIGGLFLQHVGV	FSFYSG	NGCETT---	DPSGNTRTYAYIDWIRETANITMF	-----	
L4-25	PLIQHIYNPQGV	RHAGIAS	FISGDG	CDQP---HPSGYTRTYLYLDWIANVTSGTY	-----	
PM4-08	ALLGDVPVGDYKI	QVGISS	FISQNGCESL---	DPTGYTRTDAYFQWMHNISKYG	-----	
PM4-54	ALLADVPVGDYKI	QVGISS	FISQNGCESL---	DPTGYTRVDGPYSWIVDTRNNSIV	-----	
PM4-63	ALLADVPVGDYKI	QVGISS	FISQNGCESL---	DPTGYTRVVATIHGLLIL	-----	
PM4-36	PW				-----	
PM4-60	PLVIAKGINSYV	--QIGVAG	FFSSQGCESM---	HPSGYIRTDVYNDWIWNTTQSL	-----	
PM5-80	PLVIAKGINSYV	--QIGVAG	FFSSQGCESM---	HPSGYIRTDVYNDWIWNTTQSL	-----	
AM4-49	PLVIDAGI	--SPV--LVGLV	SFISTDGCESG---	HPTGFTRTAAYRDWIRTNSGV	-----	
AM1-01	ALVLDVET	--DPV--HVGLV	SFFGGGDCESG---	YPPFGSTRTASFRIWIRDETGV	-----	
AM3-26	PLVLDVET	--DPV--HVGF	LSFISGDTDCESG---	YPSGFTRTASFRIWIRDETGV	-----	
PM5-90	PLVLDVET	--NPV--HVGL	LFLSFGGDRDCESR---	YPSGFTRTASFKNWIRDITDV	-----	
AM2-68	PLVIDDGSGNSV	--HVGVS	SWASASGCETN---	RPSGYTRTAAYRDWVESVIG	-----	
AM4-68	PLVTDGSGNSV	--HVGVS	SWASASGCETN---	HPSGYTRTAAYRDWVESVIG	-----	
L3-22	PLVIDDGSGNSV	--HVGVS	SWASASGCETN---	HPSGYTRTAAYRDWVESVIG	-----	
PM4-06 ( <i>tmt1a</i> )	PLVTGSGT	--SAV--HVGIV	SFGSSAGCAK	G---YPSAYTRTAAYRSWISSNAGV	-----	
PM4-86	PFILNGT	-----QIGIV	SYGI-TYCLPG---	YPSGFTRVTSFLEWIGENTDVQIE	-----	
PM5-92	PLILNGT	-----QIGIV	SYGI-TYCLPG---	YPSGFTRVTSFLDWIGENTDVQIE	-----	
S3-80	PLIKDDV	-----QIGV	SFGFWNHLLA---	WISLWLLQSHQLLGLDCHQFRCSNPINF	-----	
AM1-21					-----	

Loop 2

PM1-83 -----  
 PM1-75 (**tmt1a**) -----  
 PM2-70 (**tmt1b**) -----  
 PM1-95 (**tmt1c**) -----  
 PM2-03 PLIT -----  
 Dmtryp -----  
 HsPRSS1 -----  
 AM1-11 -----  
 AM1-62 -----  
 AM2-58 -----  
 AM3-01 -----  
 AM4-75 -----  
 AM4-47 -----  
 PM2-57 -----  
 PM1-93 -----  
 L4-24 -----  
 S3-72 -----  
 PM4-31 -----  
 PM2-01 -----  
 PM3-37 -----  
 L3-34 -----  
 L4-25 -----  
 PM4-08 -----  
 PM5-73 -----  
 PM4-54 -----  
 PM4-63 -----  
 PM4-36 -----  
 PM4-60 -----  
 PM5-80 -----  
 AM4-49 -----  
 AM1-01 -----  
 AM3-26 -----  
 PM5-90 -----  
 AM2-68 -----  
 AM4-68 -----  
 L3-22 -----  
 PM4-06 (**tmc1a**) -----  
 PM4-86 -----  
 PM5-92 -----  
 S3-80 IL-- -----  
 AM1-21 -----

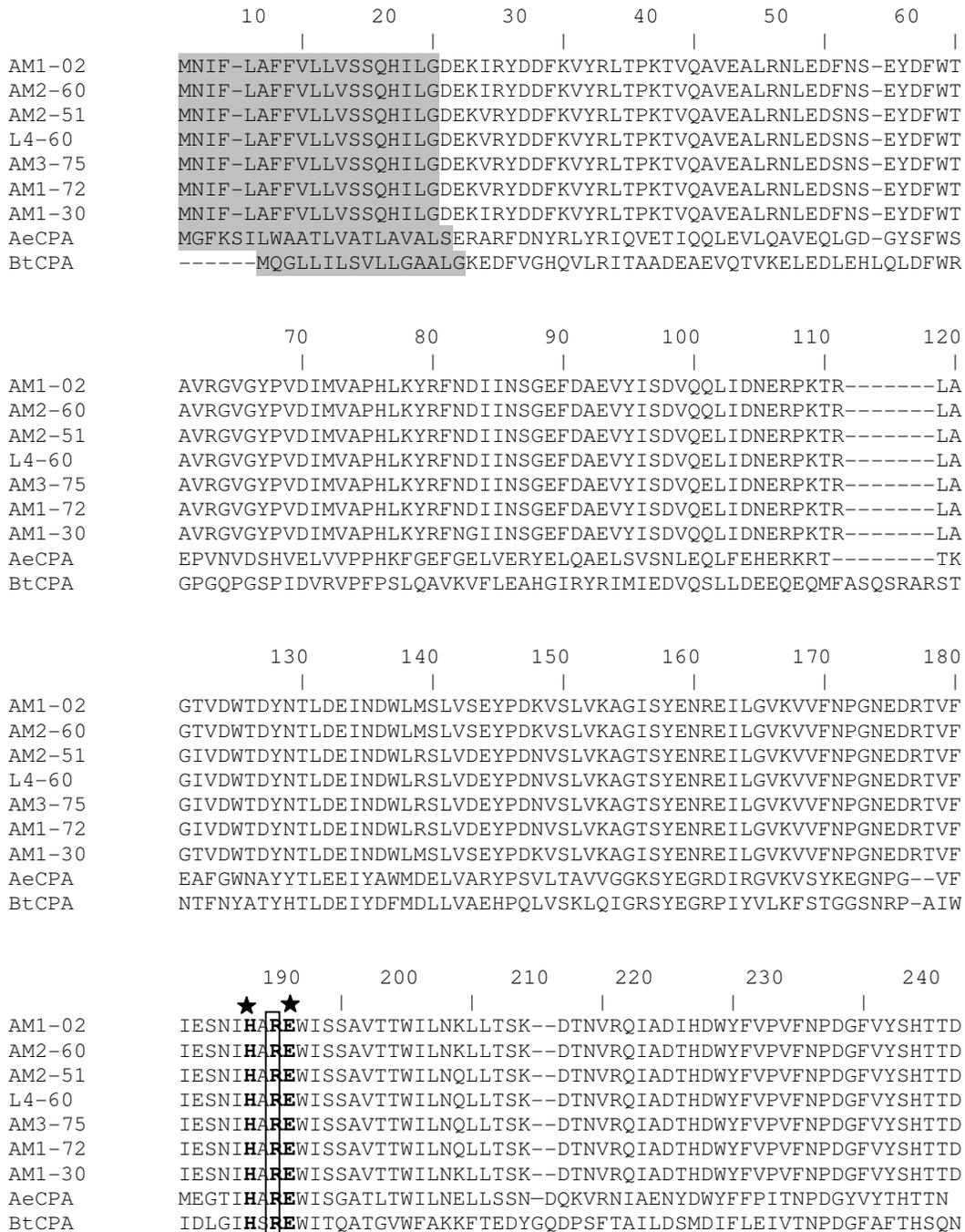
**Figure 3.3B**



**Figure 3.3C**

	10	20	30	40	50	60
AM4-72	MNQFGDMSKEEF	LAYVNRGKAQKPKHPENLRMPYVSSKKPLAASVDWRSNAVSEVKDQGG				
AM3-32	-----	-----	-----	-----	-----	-----
AM4-22	-----	-----	-----	-----	-----	-----
TmcatL	-----	-----	-----	-----	-----	-----
	70	80	90	100	110	120
AM4-72	<b>CGSC</b> CWSFSTTGAIEGQLALQRGRLTSLSEQNLID <b>CS</b> SSYGNAG <b>CD</b> GGWMDSAFSYIHDIYD					
AM3-32	-----	-----	-----	-----	-----	-----
AM4-22	-----	-----	-----	-----	-----	-----
TmcatL	-----	-----	-----	-----	-----	-----
	130	140	150	160	170	180
AM4-72	IMSESAYPYEAQGDY <b>CR</b> FDSSQSVTTLSGYYDLPSGGENSLADAVGQAGPVAVAI					
AM3-32	LTHKKSYSPIEEIRRQLIFKDNVAKIAEHNAKFEKGEVTYSKAMNQFGDMS					
AM4-22	LTHKKSYSPIEEIRRQLIFKDNVAKIAEHNAKFEKGEVTYSKAMNQFGDMS					
TmcatL	LTHKKSYSPIEEIRRQLIFKDNVAKIAEHNAKFEKGEVTYSKAMNQFGDMS					
	190	200	210	220	230	240
AM4-72	LQFYSGGLFYDQ <b>TC</b> NQSDLNHGVLVVGYSN <b>GD</b> GHWGSNAVSEVKDQGG <b>CG</b> SCWSFSTTGA					
AM3-32	LAYVNRGKAQKPKHPENLRMPYVFSKKPLAAS					
AM4-22	LAYVNRGKAQKPKHPENLRMPYVFSKKPLAAS					
TmcatL	LAYVNRGKAQKPKHPENLRMPYVFSKKPLAAS					
	250	260	270	280	290	300
AM4-72	IEGQLALQRGRLTSLSEQNLID <b>CS</b> SSYGNAG <b>CD</b> GGWMDSAFSYIPDYGIMSEFAYPYEAQ					
AM3-32	VEGQLALQRGRLTSLSEQNLID <b>CS</b> SSYGNAG <b>CD</b> GGWMDSAFSYIHDIYGIMSESAYPYEAQ					
AM4-22	VEGQLALQRGRLTSLSEQNLID <b>CS</b> SSYGNAG <b>CD</b> GGWMDSAFSYIHDIYGIMSESAYPYEAQ					
TmcatL	VEGQLALQRGRLTSLSEQNLID <b>CS</b> SSYGNAG <b>CD</b> GGWMDSAFSYIHDIYGIMSESAYPYEAQ					
	310	320	330	340	350	360
AM4-72	GDY <b>CR</b> FDSSQFVTTLSGYYDLPSGGENSLADAVGQAGPVAVAI					
AM3-32	DDY <b>CR</b> FDSSQSVTTLSGYYDLPSGDENSLADAVGQAGPVAVAI					
AM4-22	GDY <b>CR</b> FDSSQSVTTLSGYYDLPSGDENSLADAVGQAGPVAVAI					
TmcatL	GDY <b>CR</b> FDSSQSVTTLSGYYDLPSGDENSLADAVGQAGPVAVAI					
	370	380	390	400	410	420
AM4-72	<b>TC</b> NQSDLNHGVFVVGYSN <b>GD</b> NGQDYWILKNSWG <b>SG</b> WGESGYWRQVRNYGNN <b>CG</b> IATAASYP					
AM3-32	<b>TC</b> NQSDLNHGVFVVGYSN <b>GD</b> NGQDYWILKNSWG <b>SG</b> WGGENGYWTQVRNYGNN <b>CG</b> IATAASYP					
AM4-22	<b>TC</b> NQSDLNHGVLVVGYSN <b>GD</b> NGQDYWILKNSWG <b>SG</b> WGESGYWRQVRNYGNN <b>CG</b> IATAASYP					
TmcatL	<b>TC</b> NQSDLNHGVLVVGYSN <b>GD</b> NGQDYWILKNSWG <b>SG</b> WGESGYWRQVRNYGNN <b>CG</b> IATAASYP					
AM4-72	AL					
AM3-32	AL					
AM4-22	AL					
TmcatL	AL					

**Figure 3.3D**



	250	260	270	280	290	300
AM1-02	RMWFKTRVPY--FLCAGADFN	NRNWGYFFNTGGSSSNPCSE	TYGGPSAFSEPSTKTLSEFI			
AM2-60	RMWFKTRVPY--FLCAGADFN	NRNWGYFFNTGGSSSNPCSE	TYGGPSAFSEPSTKTLSEFI			
AM2-51	RMWFKTRVPY--FLCAGADFN	NRNWGYFFNTGGSSSNPCSE	TYGGPSAFSEPSTKTLSEFI			
L4-60	RMWFKTRVPY--FLCAGADFN	NRNWGYFFNTGGSSSNPCSE	TYGGPSAFSEPSTKTLSEFI			
AM3-75	RMWFKTRVPY--FLCAGADFN	NRNWGYFFNTGGSSSNPCSE	TYGGPSAFSEPSTKTLSEFI			
AM1-72	RMWFKTRVPY--FLCAVPTF-	TELGLLFHTGGSSSNPCSE	TYGGPSAFSEPSTKTLSEFI			
AM1-30	RMWFKTRVPY--FLCAGADFN	NRNWGYFFNTGGSSSNPCSE	TYGGPSAFSEPSTKTLSEFI			
AeCPA	RQWFKTRTPHS-ILCVGADAN	NRNWAYNFMQGGASNVPCSD	TYAGPSAFSEEEETRLSEYF			
BtCPA	RLWFKTRSVTSSSLCVGVDA	NRNWDAGFGKAGASSSPCSE	TYHGKYANSEVEVKSIVDFV			

	310	320	330	340	350	360
AM1-02	TTIGPKLGAYIAFHSYSQ	LMLLPYGYSSSHLDNYQD	LYNVGVKAASSLSQKYG	TKFQVGN		
AM2-60	TTIGPKLGAYIAFHSYSQ	LMLLPYGYSSSHLDNYQD	LYNVGVKAASSLSQKYG	TKFQVGN		
AM2-51	TTIGPKLGAYIAFHSYSQ	LMLLPYGYSSSHLDNYQD	LYNVGVKAASSLSQKYG	TKFQVGN		
L4-60	TTIGPKLGAYIAFHSYSQ	LMLLPYGYSSSHLDNYQD	LYNVGVKAASSLSQKYG	TKFQVGN		
AM3-75	TTIGPKLGAYIAFHSYSQ	LMLLPYGYSSSHLDNYQD	LYNVGVKAASSLSQKYG	TKFQVGN		
AM1-72	TTIGPKLGAYIAFHSYSQ	LMLLPYGYSSSHLDNYQD	LYNVGVKAASSLSQKYG	TKFQVGN		
AM1-30	TTIGPKLGAYIAFHSYSQ	LMLLPYGYSSSHLDNYQD	LYNVGVKAASSLSQKYG	TKFQVGN		
AeCPA	TSVQPKISTYLSFHAYSQ	LMLLPYGHHTTEPLDNYDE	IMDIGRLAIAKLSERHGT	QYKIGN		
BtCPA	KDHG-NFKAFLSIHSYSQ	LLLLPYGYTTQSI	IPDKTELNQVAKSAVEAL	KSLYGTSYKYGS		

	370	380	390	400	410	420
AM1-02	IVELLYVASGG\$MDWVK	GTFKTPITYTYE	LRDTGRYGFILPADQI	IIPSAEETLDSLVTIL		
AM2-60	IVELLYVASGG\$MDWVK	GTFKTPITYTYE	LRDTGRYGFILPADQI	IIPSAEETLDSLVTIL		
AM2-51	IVELLYVASGG\$MDWVK	GTFKTPITYTYE	LRDTGRYGFILPADQI	IIPFAEETLDSFVTL		
L4-60	IVELLYVASGG\$MDWVK	GTFKTPITYTYE	LRDTGRYGFILPADQI	IIPFAEETLDSFVTL		
AM3-75	IVELLYVASGG\$MDWVK	GTFKTPITYTYE	LRDTGRYGFILPADQI	IIPSAEETLDSLVTIL		
AM1-72	IVELLYVASGG\$MDWVK	GTFKTPITYTYE	LRDTGRYGFILPADQI	IIPSAEETLDSLVTIL		
AM1-30	IVELLYVASGG\$MDWVK	GTFKTPITY--	YLTNLG-----	IL-VDMDLFCQLIKLF		
AeCPA	IAEAITYIASGG\$IDWIK	GVYKTPIVLCYE	LRDTGRYGFVLPDQI	IIPNSEETLDSIIIVIL		
BtCPA	IITTIYQASGG\$IDWSY	N-QGIKYSFTFE	LRDTGRYGFLLPASQI	IPTAQETWLGVLTIM		

	430
AM1-02	QEFDKIKKN--
AM2-60	QEFDKIKKN--
AM2-51	QEFDKIKKN--
L4-60	QEFDKIKKN--
AM3-75	QEFDKIKKN--
AM1-72	QEFDKIKKN--
AM1-30	RPLKKLSTLS-
AeCPA	EEGEKRLHVL
BtCPA	EHTLNNLY---

**Figure 3.4. Alignment of TmT1 sequences with predicted trypsins (A) and TmC1 sequences with predicted chymotrypsins (B) from *T. molitor* cDNAs. Shaded sequences represent peptides that differ from their homologs in TmT1 and TmC1**

**Figure 3.4A**

```

          10      20      30      40      50      60
          |      |      |      |      |      |
TmT1      IVGGSSISISSVPXQIXLQY-----
PM1-83    MKSVLFVVFVLVASASAVPPFLRKNLSLMPDGRIVGGSSISISSVPWQISLQYYGSHICGGS
PM2-70    MKSVLFVVFVLVASASAVPPFLRKNLSLMPDGRIVGGSSISISSVPWQISLQYYGSHICGGS
PM1-75    MKSVLFVVFVLVASASAVPPFLRKNLSLMPDGRIVGGSSISISSVPWQISLQYYGSHICGGS
PM1-95    MKSILFVVFVLVASASAVPPFLRKNLSLMPDGRIVGGSSISISSVPWQISLQYYGSHICGGS
PM2-03    MKSILFVVFVLVASASAVPPFLRKNLSLMPDGRIVGGSSISISSVPWQISLQYYGSHICGGS
          ***:*****:*****:*****:*****:*****

          70      80      90      100     110     120
          |      |      |      |      |      |
TmT1      -----GSGGQVVNVARINQNPSYNDR-----
PM1-83    IISANYVVTAAHCTDGLTAGSLTVRAGTSTRGSGGQVVNVARINQNPSYNDRVIDYDISV
PM2-70    IISANYVVTAAHCTDGLTAGSLTVRAGTSTRGSGGQVVNVARINQNPSYNDRVIDYDISV
PM1-75    IISANYVVTAAHCTDGLTAGSLTVRAGTSTRGSGGQVVNVARINQNPSYNDRVIDYDISV
PM1-95    IISANYIVTAAHCTDGLTAGSLTVRAGTSTRGSGGQVVNVARINQNPSYNDRLIDYDISV
PM2-03    IISANYIVTAAHCTDGLTAGSLTVRAGTSTRGSGGQVVNVARINQNPSYNDRLIDYDISV
          *****:*****:*****:*****:*****:*****

          130     140     150     160     170     180
          |      |      |      |      |      |
TmT1      -----
PM1-83    LQLSSSLSLGSSVAAVGLPSSSTSWAGTSVLVTGWGTTTEGSSSLPSALQGVNVQIVSQ
PM2-70    LQLSSSLSLGSSVAAVGLPSSSTSWAGTSVLVTGWGTTTEGSSSLPSALQGVNVQIVSQ
PM1-75    LQLSSSLSLGSSVAAVGLPSSSTSWAGTSVLVTGWGTTTEGSSSLPSALQGVNVQIVSQ
PM1-95    LQLSSSLSLGSSVAAVGLPSSSTSWAGTSVLVTGWGTTTEGSSSLPSALQGVNVQIVSQ
PM2-03    LQLSSSLSLGSSVAAVGLPSSSTSWAGTSVLVTGWGTTTEGSSSLPSALQGVNVQIVSQ
          *****:*****:*****:*****:*****:*****

          190     200     210     220     230     240
          |      |      |      |      |      |
TmT1      -----NGYYPG
PM1-83    STCSSAYGSGSITDRMLCAGVTGGGKDACQGDSGGPLVVGNVLAGIVSWGYGCARNGYPG
PM2-70    STCSSAYGSGSNTDRMLCAGVTGGGKDACQGDSGGPLVVGNVLAGIVSWGYGCARNGYPG
PM1-75    STCSSAYGSGSITDRMLCAGVTGGGKDACQGDSGGPLVVGNVLAGIVSWGYGCARNGYPG
PM1-95    STCSSAYGSGSITDRMLCAGVTGGGKDACQGDSGGPLVVGNVLAGIVSWGYGCARNGYPG
PM2-03    STCSSAYGSGSITDRMLCAGVTGGGKDACQGDSGGPLVVGNVLAGIVSWGYGCARNGYPG
          ***** *****:*****:*****:*****:*****

          250     260
          |      |
TmT1      VYSNVPALRSYIQQTAGI
PM1-83    VYSNVPALRSYFQQTAGI----- 258
PM2-70    VYSNVPALRSYIQQTAGI----- 258
PM1-75    VYSNVPALRSYIQQTAGI----- 258
PM1-95    VYSNVPALRSYIQQTAGI----- 258
PM2-03    VYPNVPALRSYIQQPPEYKLFKYLLPLIT 269
          ** . *****:***.

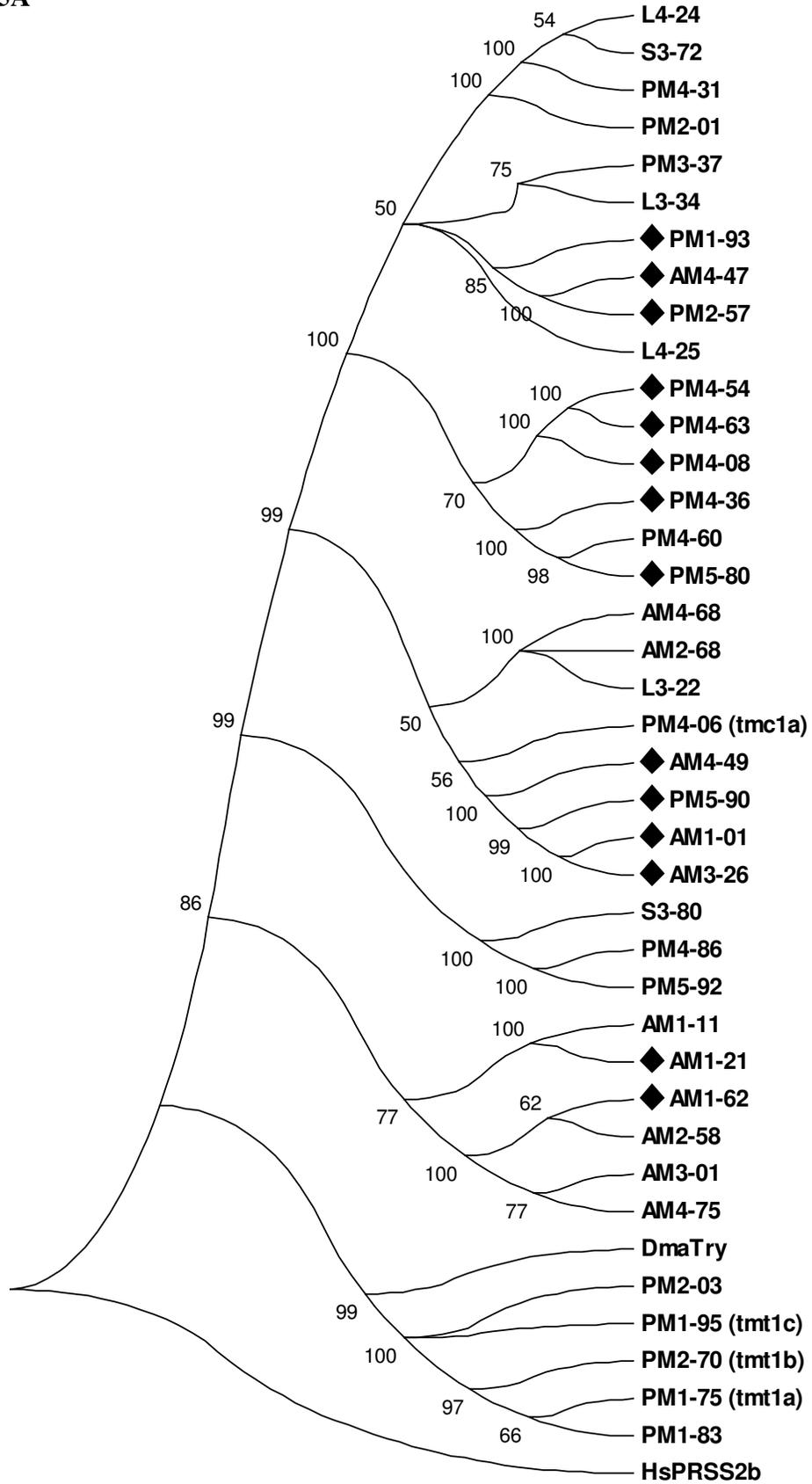
```

**Figure 3.4B**

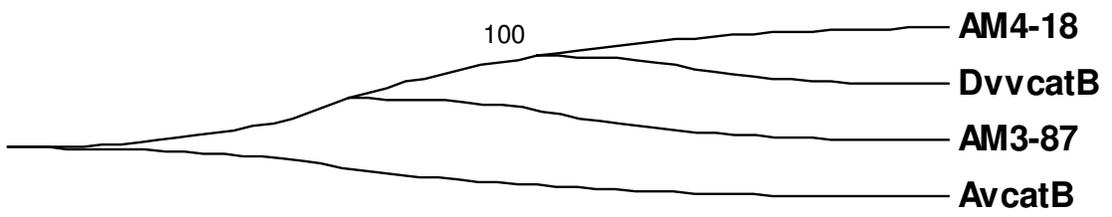
	10	20	30	40	50	60
TmC1					IISGSAASKGQFPWQ	
PM4-06	MFHLALLVLCAASALALPAELPLLPGAPPATTS	MRYAPFREISGRIIISGSAASKGQFPWQ				
	70	80	90	100	110	120
TmC1	-----	-----	-----	-----	-----	VTAQASR--
PM-06	AALYLTVSGGTSFCGGALISSNWILTAAHCTQGV	SGITAYLGVVSLSDSSRVTAQASRVV				
	130	140	150	160	170	180
TmC1	-----	-----	-----	-----	-----	-----
PM-06	APSYSSSTLANDIALIQLSTSVATSTNIRTISL	SSSTLGTGASVTVSGWGR-----				
	190	200	210	220	230	240
TmC1	-----	-----	-----	-----	-----	-----
PM-06	QTLNYVGLSTISNTVCANTYGSIIQSGIVCCTG	STIQSTCNGDGGPLVTGSGTSAVHVG				
	250	260	270			
TmT1	-----	GYPSAYTR-----	SWISSNAGV			
PM-06	IVSFGSSAGCAKGYPSAYTRTAAYRSWISSNAGV	-----	274			

**Figure 3.5. Phylogenetic tree generated using MEGA 2.1 program. Bootstrap values were calculated based on the distance/neighbor-joining method. Protease cDNAs were A) serine; B) cathepsin B C) cathepsin L D) carboxypeptidase A. Abbreviations are the same as those described in Figure 3.**

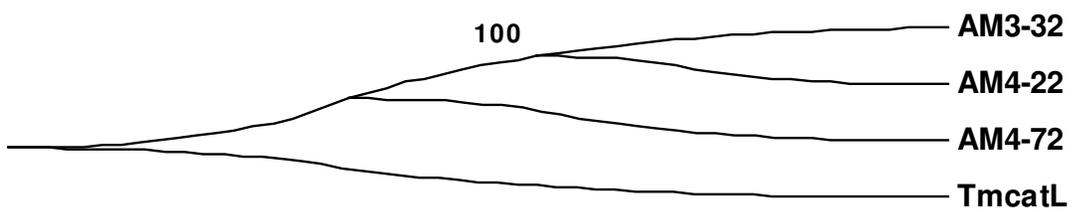
**Figure 3.5A**



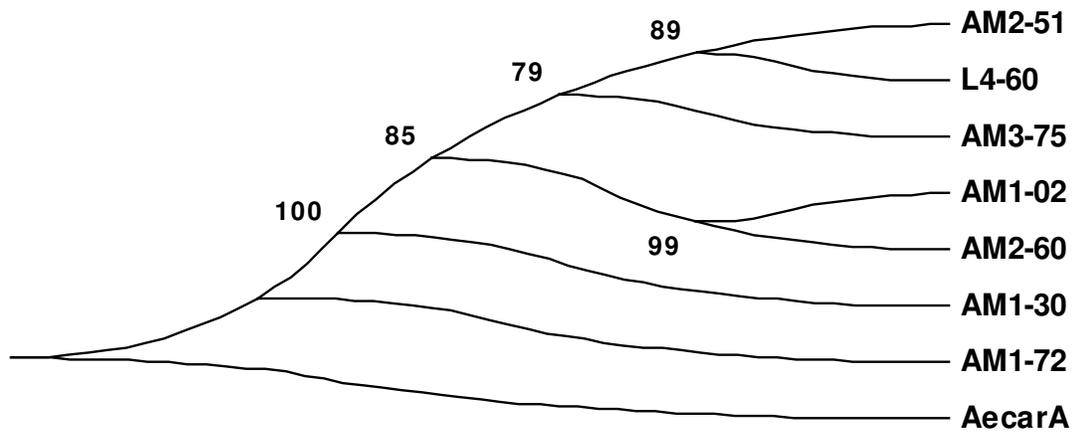
**Figure 3.5B**



**Figure 3.5C**



**Figure 3.5D**



**Figure 3.6. Northern blot analysis of RNA extracted from first-instar larvae (L1), mid-instar larvae (L2) and late-insatr larvae (L3), pupae and adults, as indicated in the figure. Probes specific to selected cDNAs from representative clusters were used for analysis. Tentative identification of the gene product is on the right. SP= serine protease homolog; SPH = serine protease homolog**

					<b>Cluster</b>	
<b>L1</b>	<b>L2</b>	<b>L3</b>	<b>P</b>	<b>A</b>		
					PM 2-70	<i>tmt1</i> (trypsin) 1
					AM1-11	SP 2
					AM1-62	SPH 3
					PM4-86	SP 4
					PM5-90	SPH 7
					PM4-06	<i>tmc1</i> (chymotrypsin) 9
					AM2-68	SP 10
					PM5-80	SPH 11
					PM4-54	SPH 13
					AM4-47	SPH 15
					PM1-93	SPH 16
					PM3-37	SP 17
					L3-34	SP 18
					PM2-01	SP 19
					AM4-18	cathepsin B 22
					AM3-87	cathepsin B 23
					AM3-32	cathepsin L 24
					PM2-27	ubiquitin

**Table 3.1. cDNAs encoding serine proteases from *T. molitor* larvae, with characteristics of the predicted peptides (contd.)**

Cluster	Clone ID	Accession Number	Mature Enzyme Mm Da	PI <sup>1</sup>	Signal Peptide (aa)	Conserved Residues <sup>2</sup>			Critical Residues <sup>2</sup>	Putative ID <sup>3</sup>
						TAAHC	DIAL	GDSGGP		
1	PM1-75 <sup>4</sup>	DQ356014	24,423.27	8.26	16		DISV		DGG	trypsin
	PM1-83	DQ356015	26,457.28	8.26	16		DISV		DGG	trypsin
	PM2-70	AY845177	26,424.21	8.26	16		DISV		DGG	trypsin
	PM2-03	DQ356016	25,918.18	8.54	16		DISV		DGG	trypsin
	PM1-95	DQ356017	24,407.21	8.26	16		DISV		DGG	trypsin
2	AM1-11	DQ356022	26,305.82	5.16	16		DVAL		GGD	SP
	AM1-21	DQ356023	9,660.82	8.71	16		missing	missing	---	SPH
3	AM1-62	DQ356018	21,794.49	6.16	17		DVAL		G-	SPH
	AM2-58	DQ356019	25,540.83	6.49	17		DVAL		GSG	SP
	AM3-01	DQ356020	25,227.56	6.85	17		DVAL		GSG	SP
	AM4-75	DQ356021	25,267.58	6.91	17		DVAL		GSG	SP
4	PM4-86	DQ356024	28,154.30	4.27	16	TAGHC	DI AV		GAS	SP
	PM5-92	DQ356025	27,671.83	4.21	16	TAGHC	DI AV		GGG	SP
5	S3-80	DQ356026	29,177.70	4.64	16	TAGHC	DVGL		GGG	SP
6	AM1-01	DQ356027	25,978.36	4.25	16	TAAQC	DIGL	GDGGAP	NFF	SPH
	AM3-26	DQ356028	26,290.43	4.17	16	TAAQC	DIGL	GDGGAP	NIS	SPH
7	PM5-90	DQ356029	26,541.68	4.22	16	TAAQC	DIGL	GDGGDP	NLS	SPH
8	AM4-49	DQ356030	26,073.51	4.26	16	TVAQC	DIGL	GDGGSP	SIT	SPH
9	PM4-06 <sup>5</sup>	DQ356031	25,991.02	8.92	16				SGS	chymotrypsin
10	AM2-68	DQ356033	26,294.98	4.52	16	TAGHC	DIGL		GAS	SP
	AM4-68	DQ356032	26,464.11	4.52	16	TAGHC	DIGL		GAS	SP
	L3-22	DQ356034	26,363.01	4.52	16	TAGHC	DIGL		GAS	SP

<sup>1</sup>Isoelectric point is for the mature enzyme.

<sup>2</sup>Only sequences differing from conserved regions are given; critical residues are found in the specificity substrate-binding pocket of the enzyme.

<sup>3</sup>Identification based on active site residues, as per Perona and Craik (1995) and Ross *et al.* (2003). SP = serine proteinase; SPH = serine proteinase homolog.

<sup>4</sup>All sequences in this cluster have a predicted N-terminus identical to TmT1 (Tsybina *et al.*, 2005).

<sup>5</sup>Sequence encodes TmC1 (Elpidina *et al.*, 2005).

Table 3.1. cDNAs encoding serine proteases from *T. molitor* larvae, with characteristics of predicted peptides.

Cluster	Clone ID	Accession Number	Mature Enzyme Mm (Da)	pI	Signal Peptide (aa)	Conserved Residues <sup>1</sup>			Critical Residues <sup>1</sup>	Putative ID <sup>2</sup>
						TAAHC	DIAL	GDSGGP		
11	PM4-60	DQ356038	25,848.23	4.54	40		DIGL	GDTGGP	GFS	SP
	PM5-80	DQ356039	25,871.27	4.62	40		DIGL	GTDGGP	GFS	SPH
	PM4-36	DQ356040	20,846.71	4.42	40		DIGL	VILEDP	E--	SPH
12	PM4-08	DQ356037	27,818.11	4.62	16	TAGQC	DVGM	GDVGGP	GIT	SPH
13	PM4-54	DQ356035	28,063.45	4.59	16	TAGQC	DIGM	GDVGGP	GIT	SPH
	PM4-63	DQ356036	27,253.75	4.74	16	TAGQC	DIGM	GDVGGP	GIT	SPH
14	L4-25	DQ356041	26,828.90	4.82	19		NIGL	GDTGSP	GIS	SP
15	AM4-47	DQ356042	26,350.40	4.28	17	TAGQC	DIGL	GDLGSP	GIS	SPH
	PM2-57	DQ356043	26,359.35	4.28	17	TAGQC	DIGL	GDLGSP	GIS	SPH
16	PM1-93	DQ356044	25,901.04	4.51	16	TSGSC	DIGV	GDVGSP	GIS	SPH
17	PM3-37	DQ356045	26,598.74	5.08	19		DIGL	DGSGSP	RIS	SP
18	L3-34	DQ356046	26,962.17	4.64	18	TSAHC	DIGL	GDSGSP	RYS	SP
19	PM2-01	DQ356047	25,933.28	4.42	21	TSGHC	DIGL	GDTGIP	GLQ	SP
20	PM4-31	DQ356048	25,654.70	4.37	21	TSGHC	DIGL	GDTGSP	GLS	SP
21	L4-24	DQ356049	25,656.68	4.37	21	TSGHC	DIGL	GDTGSP	GLS	SP
	S3-72	DQ356050	25,682.71	4.37	21	TSGHC	DIGL	GDTGSP	GLS	SP

**Table 3.2. cDNAs encoding cysteine and metalloproteases from *T. molitor* larvae, with characteristics of the predicted peptides**

Cluster	Clone ID	Accession Number	Mature Enzyme Mm (Da)	pI	Signal Peptide ID(aa)	Active Site Residues <sup>1</sup>	Critical Residues <sup>1</sup>	Putative ID <sup>2</sup>
22	AM4-18	DQ356051	30,680.48	5.41	19	QCEN	HH	cathepsin B
23	AM3-87	DQ356052	33,569.24	4.80	19	QCHN	AH	cathepsin B
24	AM3-32	DQ356053	34,468.89	4.87	16	QCHN		cathepsin L
	AM4-22	DQ356054	34,459.97	5.11	16	QCHN		cathepsin L
25	AM4-72	DQ356055	45,506.49	4.22	-	QCHN		cathepsin L
26	AM2-60	DQ356060	45,234.71	5.01	19	HEH	RRRME	carboxypeptidase A
	AM1-30	DQ356058	43,831.50	5.48	19	HEH	RRRME	carboxypeptidase A
	AM2-51	DQ356061	45,294.58	4.86	19	HEH	RRRME	carboxypeptidase A
	AM1-02	DQ356056	45,246.77	5.01	19	HEH	RRRME	carboxypeptidase A
	AM1-72	DQ356059	44,950.37	4.87	19	HEH	RRRME	carboxypeptidase A
	AM3-75	DQ356057	45,200.46	4.86	19	HEH	RRRME	carboxypeptidase A
	L4-60	DQ356062	45,294.58	4.86	19	HEH	RRRME	carboxypeptidase A

**Table 3.3. Relative abundance of mRNA transcripts in different developmental stages of *Tenebrio molitor* larvae.**

Clone ID	Tentative ID <sup>a</sup>	L1 <sup>b</sup>	L2	L3	Pupa	Adult
PM2-70	trypsin (tmt1b)	1	0.82	0.42	-	0.23
PM1-93	SPH	1	0.98	0.54	-	0.47
PM4-86	SP	1	0.25	-	-	-
PM5-80	SP	1	0.80	0.67	-	0.37
PM4-54	SPH	1	0.84	0.72	-	1.01
AM1-62 <sup>c</sup>	SPH	-	-	1	-	0.46
L3-34	SP	1	-	-	0.78	1.02
PM5-90	SPH	1	-	-	-	-
AM4-47	SPH	1	0.55	0.46	-	-
AM1-11	SP	1	0.21	0.76	-	0.95
PM2-01	SP	1	0.45	0.24	-	0.32
PM4-06	chymotrypsin (tmc1a)	1	0.99	0.95	-	0.74
AM2-68	SP	1	1.06	0.91	-	0.80
PM3-37	SP	1	1.04	0.90	-	1.11
AM3-87	cathepsin B	1	0.87	1.04	1.03	1.42
AM4-18	cathepsin B	1	1.10	0.94	-	1.03
AM3-32	cathepsin L	1	0.81	1.04	-	1.19
PM2-22	ubiquitin	1	1.06	1.06	1.10	1.20

<sup>a</sup>Tentative identification of the gene product. SP = serine protease; SPH = serine protease homolog.

<sup>b</sup>Relative intensity (pixels per unit area) of band in each developmental stage divided by L1. No expression of RNA is indicated by (-). L1, L2, and L3 refer to the larval stage, as defined in the experimental procedures.

<sup>c</sup>Expressed only in late instar larva and adult (value is relative to L3).

# CHAPTER 4 - EVALUATION OF THE COMPENSATORY RESPONSE OF *TENEBRIO MOLITOR* LARVAE TO DIETARY PROTEASE INHIBITORS

## 4.1. Abstract

The response of *T. molitor* larvae to dietary serine and cysteine protease inhibitors was investigated. Single inhibitor dietary treatments in larvae did not result in a significant growth reduction. However, a significant reduction in the mean weight of 21 day-old larvae was observed when larvae were fed a combination of inhibitors. mRNA extracted from gut tissues of *T. molitor* larvae (control and treated) was evaluated for the differential regulation of four selected protease cDNAs by real-time qPCR. These cDNAs included *tmt1* and *tmc1*, representing the serine proteases trypsin and chymotrypsin, respectively, and *cathB* and *cathL*, representing the cysteine proteases cathepsins B and L, respectively. Results of experiments demonstrated an upregulation of all proteases with either inhibitor. However, the upregulation response was increased with the serine protease inhibitors. When larvae were fed a combination of inhibitors, an the increase in the upregulation of all four cDNAs was reduced. Results from bioassays and gene quantitation studies support the hypothesis that *T. molitor* larvae compensate for dietary inhibitors through an upregulation of protease genes. Moreover, *T. molitor* larvae compensate for the effect of protease inhibitors via multiple mechanisms.

## 4.2. Introduction

### 4.2.1. *Protease Inhibitors- Insect Adaptation*

Plants synthesize protease inhibitors (PIs) as a defense against pathogen/pest attack. Early observations of the role of PIs in plant protection by Mickel and Standish (1947) indicated that larvae of certain insects were unable to develop normally on soybean products. Later research hypothesized that the insect can detoxify or develop insensitivity to the toxic products produced by the plant. Based on the feeding nature of the insect, the response can be specific or general. For example, the adaptive response of the insect can be finely tuned to the biochemical capacity of the host, such as insect species that are oligophagous and have a limited range of hosts (Rosenthal *et al.*, 1976). Polyphagous insects adapt by more general detoxification mechanisms. The mechanisms that insects use to respond to ingested PIs are: overproduction of endogenous enzymes to titre out the effect of plant inhibitors (De Leo *et al.*, 1998); expression of inhibitor-insensitive enzymes (Bolter and Jongsma, 1995; Jongsma *et al.*, 1995; Bown *et al.*, 1997; Cloutier *et al.*, 2000; Mazumdar-Leighton and Broadway, 2001); expression of inhibitor-hydrolyzing/degrading enzymes (Michaud *et al.*, 1995; Giri *et al.*, 1998), or all of the above (Moon *et al.*, 2004).

### 4.2.2. *Multiple PI Strategy*

The success of a multi-inhibitor approach to control coleopteran pests was first demonstrated in *Tribolium castaneum* (Oppert *et al.*, 1993). Since that study, the use of multiple PIs to inhibit a full spectrum of gut proteases has been proposed as a strategy to overcome protease compensation to PIs in pests (Jongsma and Bolter, 1997; Girard *et al.*, 1998). In support of this hypothesis, the use of a combination of serine and cysteine PIs resulted in significant growth retardation and mortality in the larvae of *T. castaneum* due to the disruption of compensation to each inhibitor (Oppert *et al.*, 2003; 2005). Results of the *in vitro* examination of gut proteolytic activity of these *T. castaneum* larvae suggested that larvae shifted their proteolytic enzyme profile when fed inhibitors of either protease class. When both classes of proteases were inhibited, larvae were unable to adapt by upregulating other classes of proteases.

To further understand the regulatory mechanism underlying the *in vitro* response, it is necessary to look at gene expression *in vivo*.

#### **4.2.3. *In vitro and in vivo Response to Inhibitors***

Candidate inhibitors for insect control can be identified by studying the effect of inhibitors on digestive enzymes, both *in vitro* and *in vivo* (Oppert, 2001). With the advent of techniques like real-time quantitative polymerase chain reaction (qPCR), microarray analysis, and RNA interference (RNAi), the *in vivo* studies of protease gene regulation and gene expression are possible. Results from high-throughput DNA microarray experiments (Moon *et al.*, 2004) indicated that insects were able to modulate the expression of numerous genes to overcome the effects of plant PIs. Bown *et al.* (2004) suggested that the insect is not only able to call upon a large number of protease-encoding genes, but also is capable of regulating the expression differentially when PIs are introduced into the diet. Bown *et al.* (2004) proposed a model that *H. armigera* larvae would respond to potent inhibitors by the production of insensitive proteases, but that inhibitors with lower affinity for insect proteases would not invoke this response. This type of compensation was proposed to result in a more successful adaptation of the insect to PIs with a lessened metabolic cost. The apparent complexity in the insect protease response to dietary PIs requires a careful understanding of the overall response *in vitro* and *in vivo*.

#### **4.2.4. *Real-time qPCR in Gene Expression Fold Change Determination***

A number of insect studies have incorporated real-time qPCR. Mahroof *et al.* (2005) determined a 1.7-4.3-fold increase in the expression of heat-shock transcripts in heat-shocked *T. castaneum* adults and larvae. Differences in mRNA levels of trypsin-like proteases between resistant and susceptible strains of *O. nubilalis* were demonstrated by Li *et al.* (2005) using real-time qPCR. Another study by Clermont *et al.*, 2004 demonstrated the induced expression of an insect metalloprotease inhibitor by a humoral immune response in *G. mellonella* larvae using real-time PCR.

The goal of this study was to evaluate the response of *T. molitor* larvae to dietary PIs, either singly or in combination, by bioassay and real-time qPCR.

## **4.3. Materials and Methods**

### ***4.3.1. Inhibitors***

The inhibitors used in this study were the cysteine PI, L-trans-epoxysuccinylleucylamide [4-guanidino] butane (E-64; Roche Applied Science, Indianapolis, IN, USA) and soybean trypsin inhibitor (STI, Kunitz; Sigma Chemical Co., St. Louis, MO, USA). The inhibitors were used singly or in combination in the preparation of diets for three treatments. The control diet was without inhibitor. The doses of the three treatments were as follows: 1% STI, 0.1% E-64 and 1% STI+0.1% E-64.

### ***4.3.2. Insect Bioassays***

Adult *T. molitor* derived from a laboratory culture reared continuously on 95% whole wheat flour mixed with 5% Brewer's yeast, were isolated and reared collectively in a jar kept at 28°C and 75% RH. Flour was sieved regularly to obtain eggs that were collected in a Petri dish and placed in a salt chamber at 28°C and 75% RH. Neonates (Figure 4.1) were transferred to diets, 15 larvae per treatment, contained in 4 oz cups.

### ***4.3.3. Preparation of Treatment Diets***

A stock diet was prepared with 1.42 g of whole-wheat flour and 0.071 g of Brewer's yeast. A total of 320 mg of stock diet was prepared and was aliquoted per treatment. Treatment 1 contained 320 mg only stock diet with no inhibitor and served as the control. Treatment 2 (1% STI) included 3.30 mg of STI per 337 mg of stock diet. Treatment 3 (0.1% E-64) included 0.364 mg of E-64 to 364 mg of stock diet. The combination inhibitor treatment 4 (0.1% E-64 + 1% STI) contained 0.372 mg of E-64, 3.70 mg of STI, and 370 mg of stock diet.

### ***4.3.4. Larval Weight Measurement***

Larvae were weighed individually on a Mettler Toledo microbalance (UMX) (Mettler Toledo Inc., Columbus, OH). Weight data were recorded on the 7th, 14th and 21st day post hatch. On day 21, larvae were dissected, and guts were collected for RNA isolation.

#### **4.3.5. Statistical Analysis**

The design of the experiment was a randomized complete block design with repeated measures. A mixed-model analysis was conducted using treatment, time (days post-hatch), and their interaction were as fixed effects, and replicated as random effects. Time was also a repeated measures factor. Models were fit assuming numerous standard structures for the serial correlation, and the best fitting model was selected using Akaike's Information Criterion (Littell *et al.*, 1996; Guerin and Stroup, 2000). F-tests for fixed effects and t-tests for pairwise comparisons were performed based on chosen correlation structure (Littell *et al.*, 1996). All tests used a 0.05 type I error rate. All computations were done using PROC MIXED in SAS (SAS Institute, 2001). Larval weight data was subjected to ANOVA using PROC GLM of SAS (SAS Institute, 2001).

#### **4.3.6. Isolation of RNA**

The anterior and posterior ends of guts from 8 larvae of each treatment were dissected, and the gut was pulled out with forceps and placed into a solution of DEPC-water. The guts were dried on filter paper and immediately transferred to 100  $\mu$ l of TRIReagent™ (Molecular Research Center, Inc., Cincinnati, OH). Total cellular RNA was extracted from dissected gut tissue according to the procedure provided by the manufacturer. Pooled guts were homogenized and incubated for 5 min. at room temperature (RT). To extract RNA, 20  $\mu$ l of chloroform was added and mixed vigorously. The sample was incubated for 2-15 min at RT, and then was centrifuged at 12,000 x g for 15 min. The supernatant was then transferred into a new tube and 24  $\mu$ l of isopropanol was added to each tube, mixed, and incubated for 10 min at RT. Samples again were centrifuged at 12,000 x g for 8 min. The precipitated RNA pellet was washed in 70% DEPC-ethanol and centrifuged at 7,500 x g for 5 min. The pellet was air-dried for about 30 min and was dissolved in 20  $\mu$ l of DEPC water. RNA samples were stored at -80°C. RNA samples were subjected to electrophoresis in an agarose gel to check for RNA quality. In order to ensure the absence of genomic DNA contamination, 2  $\mu$ l of each RNA sample was amplified by PCR using gene-specific primers for *tmt1*. The amplified products were analyzed on a 2% agarose gel containing ethidium bromide. RNA samples were treated with RNAase-free DNase I (Epicentre Technologies, Madison, WI) following the manufacturer's protocol. To 4  $\mu$ l of the RNA sample, 0.5  $\mu$ l of RNAase-free DNase I was added and incubated at 37°C for 20 min. Following

DNAase treatment, four DNAase-treated RNA samples were reverse transcribed using SuperScript™ III Reverse Transcriptase (RT) (Invitrogen, Carlsbad, California). 0.5 microliters of oligo-dT and dNTP were added to each tube and incubated at 65°C for 5 min. Samples were then placed on ice for 2 min. To each of the four samples, 1 µl of 10xRT buffer containing 2 µl of 25mM MgCl<sub>2</sub>, 1 µl of 0.1 mM DTT and 0.5 µl of RNase Out was added. The contents of the tube were mixed and incubated at 45°C for 2 min. After adding 0.5 µl of SuperScript™ III Reverse Transcriptase, samples were incubated at 45°C for 50 min for cDNA synthesis. Samples were then incubated for 5 min. at 85°C. The tubes were then placed on ice for 2 min, centrifuged, treated with RNase H and incubated at 37°C for 20 min. Single-stranded cDNA samples were used for real-time qPCR. The concentration of the first-strand cDNA was determined by a Nanodrop ND1000 spectrophotometer (NanoDrop Technologies, Roackland, DE, USA). A comparison of the absorbance at 260/280 nm was used to determine the purity of the samples.

#### ***4.3.7. Primer design for Real-time qPCR***

Three sets of primer pairs were designed by Beacon Designer software (PREMIER Biosoft International) for selected cDNA sequences representing *T. molitor* trypsin, *tmt1* (PM2-70), chymotrypsin, *tmc1* (PM4-06), cathepsin L, cathL (AM3-32), cathepsin B, cathB (AM3-87) (Table 4.1). Primer melting temperatures were T<sub>m</sub> 56-58°C and amplicons < 150 bp in length. A cDNA encoding an alpha-tubulin (AM4-80) normalizer gene was selected and primers were similarly designed. The specificity of each primer pair to a template was tested before gene amplification. Plasmids containing full-length transcripts of each cDNA were used to establish standard curves and determine their copy numbers in samples. A serial 10-fold dilution of cDNA samples was prepared, and six different dilutions in three technical replications were used to establish standard curves.

#### ***4.3.8. Real-time qPCR***

Real-time qPCR was performed to study the up and down regulation of selected protease cDNA sequences in guts of larvae that were treated with single and combination inhibitor treatments, with an Mx3000P QPCR system (Stratagene, La Jolla, CA) and a 25-µl reaction using a 2X Brilliant® SYBR® Green QPCR Master Mix kit (Stratagene, La Jolla, CA). After preliminary experiments to optimize the reaction, a range of DNA concentrations, from 0.4-1.7 ng to 10.0 fg (equivalent to 3.30-4.0×10<sup>9</sup> to 3.3-4.0×10<sup>4</sup> copies per reaction mixture) was

utilized to generate a set of standard curves for each PCR. PCR amplification included 1.0 µl of each forward and reverse primer (concentration of 10pmol/µl), 12.5 µl of 2X SYBR Green Brilliant qPCR Master Mix, and 2 µl cDNA template diluted to 10.5 µl. The PCR cycle was: denaturation at 95° C for 600 sec; 45 cycles of 95°C for 30 sec, annealing at 55°C for 60 sec, 72°C for 30 sec for final extension. Melting curve analysis determined the primer-dimer formation of gene-specific primers. Melting curves were obtained by increasing the temperature from 55 to 95°C (0.5°C/10 s) to denature double-stranded DNA. Experiments had three technical replicates. In gene expression experiments, the PCR cycle was run following the same program as mentioned above, and 12.5 µl of 2X SYBR Green Brilliant qPCR Master Mix, and 2 µl cDNA template diluted to 10.5 µl 50 x diluted cDNA template was used.

#### ***4.3.9. Calculation of Fold Change Level***

Mx3000P software was used to calculate fold change level, but relative quantitation also was calculated using a comparative Ct method developed by Liu and Saint (2002):  $(1 + E_{GOI})^{-\Delta\Delta Ct}$  where GOI is the gene of interest, E is the primer efficiency of the gene of interest, and  $\Delta\Delta Ct$  is the difference between the  $\Delta Ct$  of the gene of interest and  $\Delta Ct$  of the housekeeping gene [(Ct of GOI for treated-Ct of GOI for control) – (Ct of normalizer gene for treated- (Ct of normalizer gene for control))].

### **4.4. Results**

#### ***4.4.1. Bioassays***

Mortality was negligible and ranged from 6 to 13% in the control (no inhibitor) and 6% on the single inhibitor treatments (1% STI and 0.1% E-64). There was no mortality recorded in the combination inhibitor treatments.

The physical appearance of *T. molitor* larvae indicated the differential effects of inhibitors on larval growth (Figure 4.2). Larvae fed the combination of PIs (1% STI + 0.1% E-64) were smaller in size and were quivering when observed on day 21, when larvae were dissected. This combination of inhibitors resulted in a growth retardation similar to that observed in earlier studies of *T. castaneum* larvae (Oppert *et al.*, 1993, 2003). However, the larval quivering suggested that physiological effects of the inhibitor also involve the impairment of other physiological factors, such as the nervous system.

Statistical analysis revealed a significant effect of all four treatments on larval weights ( $F=29.71$ ;  $df=3,6$ ;  $P=0.0005$ ). There was a significant effect of time (days post hatch) on *T. molitor* larval weights ( $F=97.97$ ;  $df=2, 7.7$ ;  $P<0.0001$ ). There was also a significant interaction of treatment and time on larval weights ( $F=5.28$ ;  $df=6, 7.74$ ;  $P=0.01$ ). In a comparison of treatment means, at day 21, there was no significant difference between weights of larvae fed control and 1% STI ( $t=2.17$ ;  $df=5.24$ ;  $P=0.08$ ), control and 0.1% E-64 ( $t=1.47$ ;  $df=5.24$ ;  $P=0.20$ ). However, at day 21, there was a highly significant difference in weights of larvae fed control and the combination inhibitor treatment ( $t=7.36$ ;  $df=5.24$ ;  $P=0.0006$ ). A significant difference was also seen between 1% STI and the combination treatment ( $t=5.93$ ;  $df=22$ ;  $P<0.0001$ ) and 1% E-64 and the combination inhibitor treatment ( $t=6.82$ ;  $df=22$ ;  $P<0.0001$ ). There was a significant interaction of the components of the combination treatment (1% STI and 0.1% E-64) with each other ( $F=7.00$ ;  $df=1,6.1$ ;  $P=0.03$ ). This result indicated a synergistic action of STI and E-64 in reducing larval weights. At day 21, the difference in *T. molitor* larval weights between larvae fed control and single inhibitor diets and those fed the combination inhibitor treatment was most pronounced (Figure 4.3, Table 4.2). Overall, single inhibitor treatments resulted in a weight reduction of 30% or less, whereas the combination inhibitor treatment caused a 55-70% reduction in weight (data not shown).

#### **4.4.2. Standard Curves for Real-time qPCR**

The standard curve of the linear relationship between  $C_t$  and initial amounts of cDNA allows the determinations of concentrations of unknowns based on their  $C_t$  values (Wong and Medrano, 2005). Standard curves were obtained for all four protease transcripts as well as for the housekeeping/normalizer gene (graphs not shown). The log-linear phase is the period where PCR reaches optimal amplification with the PCR product doubling after every cycle, in ideal reaction conditions. A log-linear regression plot, where  $C_t$  values were plotted as a function of  $\log_{10}$  concentration of the template, was generated using  $C_t$  values. Relative quantitation involves a comparison between the sample cDNA of the gene of interest to that of the control/normalizer gene. Relative quantitation usually requires that PCR efficiencies of all genes be similar or preferably at or above 90%. The standard curve for alpha-tubulin demonstrated a PCR efficiency of 103.6% and a  $r^2$  of 0.994. A correlation coefficient greater than 0.99 is a measure of good primer efficiency and is one prediction of a successful real-time experiment. The PCR

efficiencies for the amplification of *tmt1*, *tmc1*, cathB and cathL were tested in five 10x serial dilutions of larval cDNA template, and resulted in 100.3% efficiency for *tmt1*, 96.2% for *tmc1*, 99.7% for cathB and 96.3% for cathL. Regression slopes were equal to -3.315 for *tmt1*, -3.238 for alpha-tubulin, -3.416 for *tmc1*, -3.329 for cathB, and -3.414 for cathL. A melting curve analysis indicated no evidence of the formation of primer-dimers in any of the primers tested.

#### **4.4.3. Relative Fold Change Level**

Following the method developed by Liu and Saint (2002), the amplification efficiency of an individual reaction was calculated from the kinetics of the reaction and was used for quantitation and normalization. This method is especially useful when the amplification efficiencies of reference and target gene are not equal to 1. In the gene expression experiment, the differential expression levels of *tmt1*, *tmc1*, cathB and cathL were normalized using alpha-tubulin as the housekeeping gene in 21-day old larvae that were fed single or combination inhibitor treatments. Results in Figure 4-4 indicate that when larvae were fed 1% STI (expected to inhibit *T. molitor* trypsin), *tmt1* was up regulated ~6 fold, *tmc1* ~5-fold, cathB 4.5 fold, and cathL ~15 fold. All four selected proteases assayed were up-regulated compared to the activity of the same proteases in larvae on control diets. *T. molitor* larvae that were fed treatment 2 (0.1% E-64) up-regulated *tmt1* 1.1 fold, *tmc1* 2 fold, cathB 1 fold and cathL 5 fold. Larvae fed both 1% STI and 0.1% E-64 had less than a 1 fold up-regulation of *tmt1*, *tmc1* and cathB, and a 1.4 fold upregulation of cathL.

### **4.5. Discussion**

Results of the larval feeding inhibitor bioassay as well as those of the real-time qPCR experiments support the hypothesis of a multiple-inhibitor approach to control tenebrionid pests. There was a synergistic growth retardation effect in *T. molitor* larvae that were fed both 0.1% E-64 and 1% STI, as measured by the mean weight of larvae. In contrast, there were minimal growth effects in larvae that were fed single inhibitors.

From the anecdotal observation of *T. molitor* larvae, in addition to negative effects on survival and growth, the combination of inhibitors fed to *T. molitor* larvae affected general larval fitness and metabolism. Other studies of the effects of inhibitors on insect physiology have reported adverse effects on reproduction (Ashouri *et al.*, 1998), learning (Pham-Delègue *et al.*, 2000), diapause, and digestion (Cloutier *et al.*, 2001). One hypothesis is that the quivering of

larvae resulted from a nervous system malfunction or diversion of energy for the physiological compensation response.

Real-time qPCR results indicated that insects were able to use more than one mechanism to compensate when fed either serine or cysteine PIs. *T. molitor* larvae upregulated serine proteases when fed STI, a serine PI. When fed E-64, a cysteine PI, larvae had a 5 fold upregulation of cathepsin L. These results indicated that 21 day-old larvae fed single serine or cysteine inhibitors use the hyperproduction of proteases from the same protease class as the inhibitor to ‘titer out’ the inhibitor molecules. However, larvae fed single inhibitors also appear to upregulate proteases from other classes.

There was a 15 fold upregulation of cathepsin L in *T. molitor* larvae in response 1% STI, and a 2 fold upregulation of chymotrypsin when larvae were fed 0.1% E-64. This result is similar to the results observed by Oppert *et al.* (2005) in *T. castaneum* larvae fed E-64, where larvae responded by increasing chymotrypsin activity. Additionally, the normal insect gut function was effectively disrupted by a combination of inhibitors targeting multiple protease classes. Three-week old larvae (3rd instar) were a vulnerable life stage of *T. molitor* that can be targeted for future pest control strategies. This was the first study to report the relative quantitation of changes in gene expression in insect larvae that were fed dietary inhibitors.

The results of this study have implications in the control of tenebrionid and other coleopteran stored-product pests. The multiple-inhibitor approach can be adopted as an improved method for pest management based on PIs. Suitable proteinaceous inhibitors are needed for transgenic technology, and cost effective inhibitors for sprays and powder treatments are needed. Further studies also are required before the effective doses of inhibitors can be recommended for use by entomologists involved in the development of plant resistance based on protease inhibitors.

## 4.6. References

- Ashouri A., Overney S., Michaud, D. & Cloutier, C. (1998) Fitness and feeding are affected in the two-spotted stinkbug, *Perillus bioculatus*, by the cysteine proteinase inhibitor, oryzacystatin I. *Arch. Insect Biochem. Physiol.* 38: 74–83.
- Bolter, C. J. and Jongsma, M. A. (1995) Colorado potato beetles (*Leptinotarsa decemlineata*) adapt to proteinase inhibitors induced in potato leaves by methyl jasmonate. *J. Insect Physiol.* 41: 1071-1078.
- Bown, D.P., Wilkinson H.S. and Gatehouse J.A. (1997) Differentially regulated inhibitor-insensitive and insensitive protease genes from the phytophagous insect pest, *Helicoverpa armigera*, are members of complex multigene families. *Insect Biochem. Mol. Biol.* 27: 625-638.
- Clermont, A., Wedde, M., Seitz, V., Podsiadlowski, L., Lenze, D., Hummel, M. and Vilcinskis, A. (2004) Cloning and expression of an inhibitor of microbial metalloproteinases from insects contributing to innate immunity. *Biochem J.* 382: 315–322.
- Cloutier, C., C. Jean, M. Fournier, S. Yelle, and Michaud, D. 2000. Adult Colorado potato beetles, *Leptinotarsa decemlineata* compensate for nutritional stress on oryzacystatin I-transgenic potato plants by hypertrophic behavior and over-production of insensitive proteases. *Arch. Insect Biochem. Physiol.* 44: 69–81.
- Girard, C., Le-Mateyar, M., Monade-Bottino, M., Pham-Delegue, M. H. and Jouanin, L. (1998) High level of resistance to proteinase inhibitors may be conferred by proteolytic cleavage in beetle larvae. *Insect Biochem. Mol. Biol.* 28: 229-237.
- Giri A. P., Harsulkar A. M., Deshpande V. V., Sainani M. N., Gupta V.S., Ranjekar P. K. (1998) Chickpea defensive proteinase inhibitors can be inactivated by podborer gut proteinases. *Plant Physiol.* 116: 393-401.
- Guerin, L. and Stroup, W. W. (2000) A simulation study to evaluate PROC MIXED analysis of repeated measures data. Proceedings of the 12<sup>th</sup> Annual Conference on Applied Statistics in Agriculture, Manhattan, KS: Kansas State University, 170-203.

Jongsma M.A., Bakker P.L., Peters J., Bosch D., and Stiekema W.J. (1995) Adaptation of *Spodoptera exigua* larvae to plant proteinase inhibitors by induction of gut proteinase activity insensitive to inhibition. *Proc. Natl. Acad. Sci.* 92: 8041-8045.

Jongsma, M.A., Bolter, C. (1997) The adaptation of insect to plant protease inhibitors. *J. Insect Physiol.* 43: 885-896.

Li, H., Oppert, B., Higgins, R.A., Huang, F., Buschman, L.L., Gao, J-R. and Zhu, K.Y. (2005) Characterization of cDNAs encoding three trypsin-like proteinases and mRNA quantitative analysis in *Bt*-resistant and -susceptible strains of *Ostrinia nubilalis*. *Insect Biochem. Mol. Biol.* 35: 847-860.

Littell, R. C. Milliken, G. A., Stroup, W. W. and Wolfinger, R. D. (1996) SAS System for Mixed Models, SAS Institute Inc. Cary, NC.

Liu, W. and Saint, D.A. (2002) A new quantitative method of real time reverse transcription polymerase chain reaction assay based on simulation of polymerase chain reaction kinetics. *Analytical Biochem.* 302: 52-59.

Mahroof, R., Zhu, K.Y., Neven, L., Subramanyam, B. and Bai, J. (2005) Expression patterns of three heat shock protein 70 genes among developmental stages of the red flour beetle, *Tribolium castaneum* (Coleoptera: Tenebrionidae). *Comp. Biochem. Physiol.* 141: 247-256.

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

Michaud, D., Cantin, L. and Vrain, T.C. (1995) Carboxy-terminal truncation of oryzacystatin-II by oryzacytatin-insensitive insect digestive proteinases. *Arch. Biochem. Biophys.* 322: 469–474.

Mickel, C.E. and Standish, J. (1947) Susceptibility of processed soy flour and soy grits in storage to attack by *Tribolium castaneum*. *Univ. of Minn. Agr. Exp. Sta. Tech Bull.* 178: 1-20.

Moon, J., Salzman, R.A., Ahn, J-E., Koiwa, H. and Zhu-Salzman, K. (2004). Transcriptional regulation in cowpea bruchid guts during adaptation to a plant defence protease inhibitor. *Insect Mol. Biol.* 13: 283-291.

Pham-Delègue M.H., Girard C., Le Métayer M., Picard-Nizou A.L., Hennequet C., Pons O., Jouanin L. (2000) Long-term effects of soybean protease inhibitors on digestive enzymes, survival and learning abilities of honeybees. *Entomol. Exp. Appl.* 95: 21-29.

Oppert, B., Morgan, T. D., Culbertson, C. and Kramer, K. J. (1993) Dietary mixtures of cysteine proteinase and serine proteinase inhibitors exhibit increased toxicity toward the red flour beetle, *Tribolium castaneum*. *Comp. Biochem. Physiol.* 105C: 379-385.

Oppert, B. (2000) Transgenic plants expressing enzyme inhibitors and the prospects for biopesticide development. In *Phytochemical Biopesticides* (Koul, O. and Dhaliwal, G. S., eds.), Harwood Academic, The Netherlands, pp. 83-95.

Oppert, B., Morgan, T.D., Hartzler, K., Lenarcic, B., Galesa, K., Brzin, J., Turk, V., Yoza, K., Ohtsubo, K. and Kramer, K.J. (2003) Effects of proteinase inhibitors on growth and digestive proteolysis of the red flour beetle, *Tribolium castaneum* (Herbst) (Coleoptera: Tenebrionidae). *Comp. Biochem. Physiol.* 134C: 481-490.

Oppert, B., Morgan, T.D., Hartzler, K. and Kramer, K. J. (2005) Compensatory proteolytic responses to dietary proteinase inhibitors in the red flour beetle, *Tribolium castaneum* (Herbst) (Coleoptera: Tenebrionidae). *Comp. Biochem. Physiol.* 140C: 53-58.

Rosenthal, G.A., Dahlman, D.L. and Janzen, D.H. (1976) A novel means for dealing with L-canavanine, a toxic metabolite. *Science* 192: 256–258.

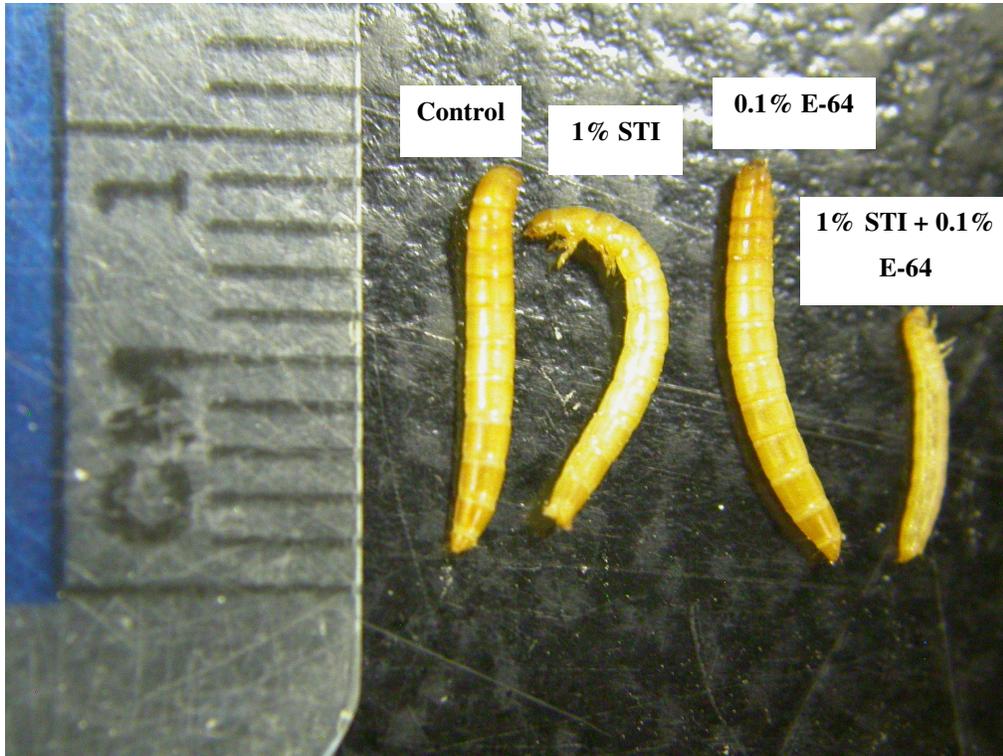
SAS Institute (2001) SAS User's guide: statistics. Version 8.2, SAS Institute, SAS Institute Inc., Cary, NC.

Wong, M. L. and Medrano, J. F. 2005. Real-time PCR for mRNA quantitation. *Biotech.* 39: 75-85

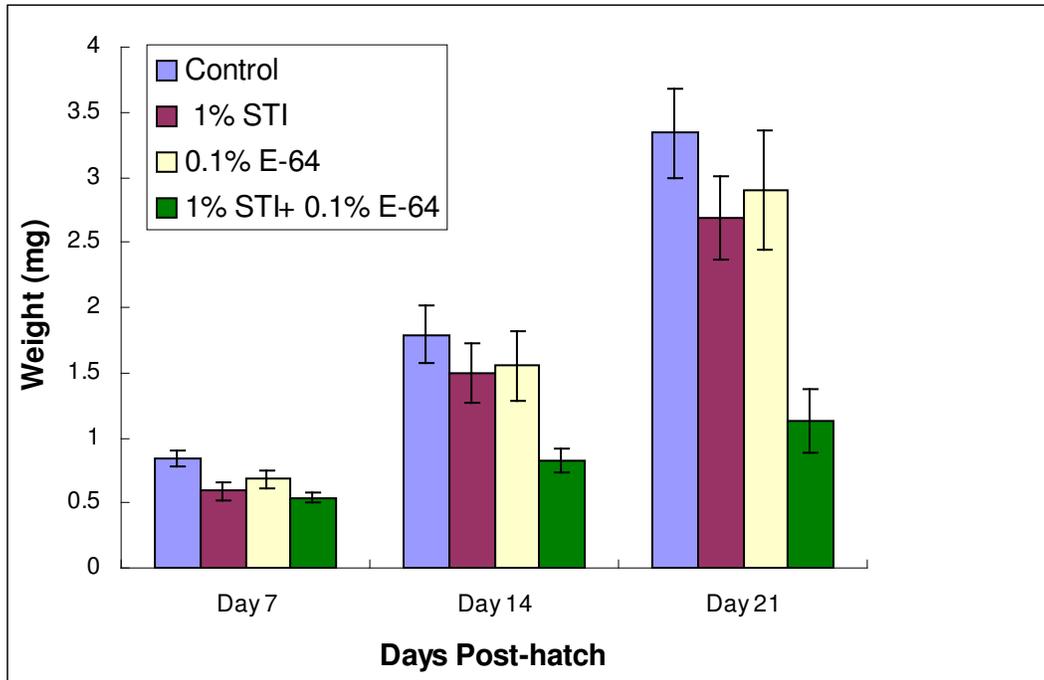
**Figure 4.1. *T. molitor* eggs and newly emerged neonates (1<sup>st</sup> instar ) larvae used to infest treatment diets.**



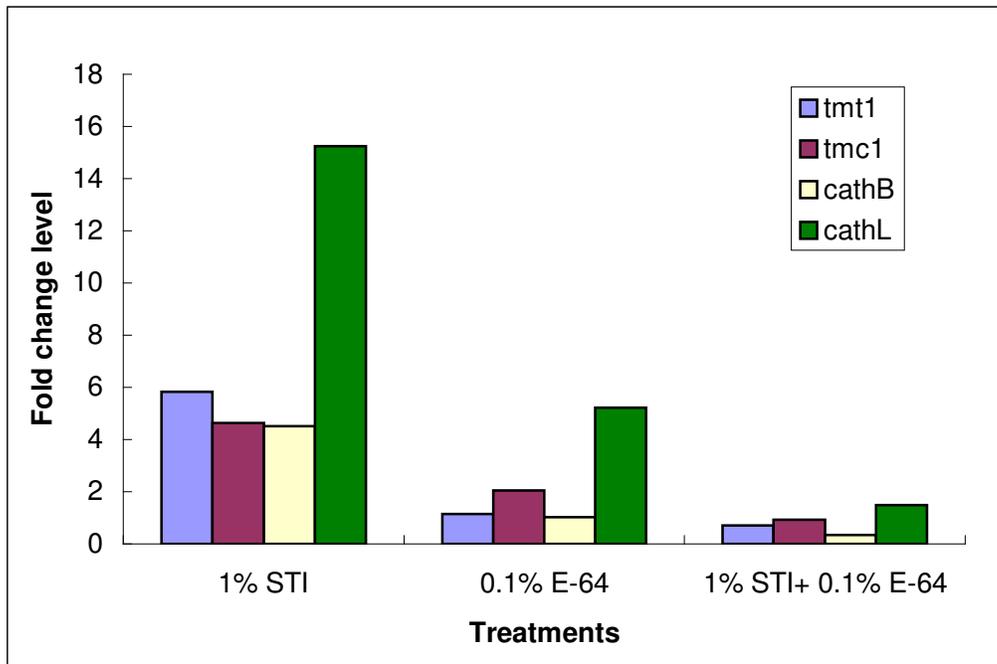
**Figure 4.2. Contrast of *T. molitor* larvae fed a combination of 1% STI and 0.1% E-64 with those fed either 1% STI or 0.1% E-64 or control diet (no inhibitor).**



**Figure 4.3.** A comparison of the weights of *T. molitor* larvae fed control, 1% STI, 0.1% E-64, or 1% STI + 0.1% E-64 at 7, 14 and 21 days post-hatch. Weights are the mean + S.E., n=15.



**Figure 4.4. Real-time quantitative PCR analysis of the relative quantitative fold change expression of two *T. molitor* serine protease transcripts, *tmt1* and *tmc1* and two cysteine protease transcripts, *cathB* and *cathL* in guts of 21-day old larvae fed on diet containing a single inhibitor (1% STI or 0.1% E-64) or a combination of inhibitors (1% STI+0.1% E-64).**



**Table 4.1. Primer names, sequence IDs and nucleotide sequences of selected *T. molitor* digestive protease cDNAs. Primers were designed using a computer program, Beacon Designer (version 2.1, Premier Biosoft International, Palo Alto, CA).**

<b>Primer ID</b>	<b>Plasmid cDNA ID</b>	<b>Forward primer sequence (5'-3' direction)</b>	<b>Reverse primer sequence (5'-3' direction)</b>	<b>Amplicon length</b>
<i>α</i> -tubulin	AM4-80	GCCGACTTCTTCGTAATCCTTC	TCCAACACCACCGCCATC	174
<i>tmt1</i>	PM2-70	GCGGCGGTTCAATAATCAG	ACGAGTAGAAGTTCCAGCAC	104
<i>tmc1</i>	PM4-06	CCTCATCCAACCTCTCCACTTC	ACTGTGCTGGTCCTTCC	124
cathB	AM3-87	ACGGCAGTAACGACTATGTG	TGAAATCCAACGCTCTCTCC	107
cathL	AM3-32	CTGGTGCCGTTGAAGGTC	CCGTAGTCGTGGATGTAGC	151

## CHAPTER 5 - SUMMARY

Protease inhibitors are an important component in the multi-mechanism insect defensive strategy employed by many plant species. Since the insect-plant interaction is a co-evolving process, combating insect damage to plants, processed food, and stored food is an ongoing challenge. The continual adaptation of insects to plant defenses has been a major research focus for many years. Insects employ many types of mechanisms to compensate for the detrimental effects of protease inhibitors. The rationale for this project was to evaluate insect compensation to protease inhibitors and to understand the compensatory mechanism in the response of *T. molitor* larvae to dietary protease inhibitors.

Enzyme assays described in chapter 2 indicated a diversity of proteases in *T. molitor* larvae, as has previously been reported. Proteases from the serine and cysteine classes were identified by inhibitor assays. Interestingly, the inhibition of protease activity in *T. molitor* gut extracts by a trypsin inhibitor, SBTI, was lessened when the experiments were conducted in reducing buffers, conditions that favor cysteine protease activity. Terra and Cristofolletti (1996) suggested that the Cucujiformia ancestor was a beetle that adapted to naturally occurring serine protease inhibitors by shifting from serine to cysteine proteases for digestion. *T. molitor* has retained serine protease activity and may represent an evolutionary intermediate. However, compartmentalization of the *T. molitor* midgut occurred because cysteine proteinases were unstable at higher pH found in the posterior midgut, and they are localized to the anterior midgut where they are most active. This compartmentalization also serves as a regulatory mechanism, in that any cysteine protease activity that migrates to the posterior midgut with the food bolus is inactivated. Compartmentalization of proteases in *T. molitor* is especially relevant when insects are fed serine protease inhibitors. If *T. molitor* is fed a single serine protease inhibitor, *T. molitor* cysteine proteases in the anterior midgut can hydrolyze serine protease inhibitors prior to their contact with serine proteases, thus preventing serine protease inhibition. The acidic pH of the anterior midgut also may contribute to the instability of serine protease inhibitors. However, if a cysteine protease inhibitor (E-64) in combination with a serine protease inhibitor is fed to *T. molitor* larvae, the cysteine protease inhibitor will target the anterior midgut cysteine proteases

and prevent hydrolysis by anterior cysteine proteases, and the effects of the serine protease inhibitor will be more substantial.

The information obtained from the cDNA libraries of *T. molitor* larval guts comprised an important aspect of this study. Of the 1,528 total gut cDNA sequences, 92 (7%) were identified that encode digestive proteases in *T. molitor* larvae. When the sequences were grouped into serine, cysteine or metallo proteases, 73% of the cDNAs that encoded proteases were homologous to serine proteases, 13% were homologous to cysteine proteases and 14% were homologous to metalloproteases. Of 50 full-length cDNA sequences that were submitted to GenBank, none encoded aspartic proteases.

Compartmentalization of serine and cysteine proteases in the posterior and anterior midgut, respectively, was examined by relative transcript abundance. Most serine protease transcripts were present in the posterior midgut cDNA library, while most cDNA encoding cysteine proteases were identified in the anterior midgut library. The distribution of these cDNAs indicated that serine proteases are the most diverse class of proteases in the *T. molitor* midgut, although the relative activity may be different. Sequence alignment of the cDNAs encoding serine proteases helped detect the presence/absence/variation in the sequence motifs that typify serine proteases- TAAGC, DIAL and GDSGDP. *T. molitor* cDNAs that had the serine residue in the active site motif substituted with threonine were tentatively classified as serine protease analogs, although their function as an active protease remains to be tested. Serine protease homology (SPHs) lacked one or more of the serine protease critical residues, and there were at least 14 *T. molitor* cDNAs that encoded putative SPHs. We hypothesize that, in addition to the previously proposed functions of SPHs in various defense responses, including hemolymph coagulation, melanotic encapsulation, induction of peptide antimicrobial synthesis and activation of cytokines (Kanost *et al.*, 2001; Jiang and Kanost, 2000), the SPHs expressed in the *T. molitor* gut participate in the compensational response to dietary fed protease inhibitors. If SPHs are expressed constitutively, or they are upregulated in response to serine protease inhibitors, they may bind to inhibitors with higher affinity and prevent the harmful targeting of active serine proteases. This hypothesis remains to be tested.

*T. molitor* cysteine proteases included cathepsin B and cathepsin L types. All *T. molitor* cDNAs encoding cysteine proteases, except AM4-72, had secretion signal peptides, an indication that these proteases originate in the membrane and are processed to the lumen by protease

activation. This study complements that of other researchers who have identified and localized cathepsin L cDNAs and proteins (Cristofolletti *et al.*, 2005; Vinokurov *et al.*, 2006b). This is the first study to report cDNAs encoding carboxypeptidase A from *T. molitor*. All of the *T. molitor* carboxypeptidase sequences had predicted signal peptide sequences and had homology to *Aedes aegypti* carboxypeptidase A and were members of the clan MC and family M14 metallopeptidase family.

Another significant finding of this study was the identification genes encoding major midgut proteolytic enzymes, trypsin (*tmt1a*, *tmt1b*, and *tmt1c*) and chymotrypsin (*tmc1*). The predicted N-termini of these sequences were identical to the N-termini of the purified proteins (Tsybina *et al.*, 2005; Elpidina *et al.*, 2005). Phylogenetic analysis of the *T. molitor* serine protease sequences revealed two distinct groups separated by sequences identified as SPHs. Based on the identification of *tmt1a*, *tmt1b*, *tmt1c*, and *tmc1a*, located in different areas of the tree, the structural distinction that contributes to *T. molitor* trypsin and chymotrypsin may be revealed by intervening sequences.

Expression analyses using northern blots demonstrated the temporal regulation of proteases in *T. molitor*. In general, proteases were most highly expressed in early-instar larvae, suggesting targeting early-instar larval stages with protease inhibitors may be more effective as a pest control strategy. This comprehensive information on the protease distribution and expression patterns in the midgut of *T. molitor* will impact improvement and development of pest control methods based on protease inhibitors.

In chapter 4, the individual and combined effects of STI, inhibiting serine proteases, and E-64, inhibiting cysteine proteases, was determined in *in vivo* assays. Weights of larvae fed either STI or E-64 were no different than weights of larvae fed control diet at 21 days post-infestation. However, the combined effect of these inhibitors significantly reduced *T. molitor* larval weights, suggesting that the larval digestive system is unable to compensate for the effects of inhibitors in two different classes of proteases.

These results support the hypotheses that larvae that are fed single inhibitors are able to regulate protein digestion by up-regulating proteases from another protease class, but are unable to compensate when fed multiple inhibitors. This hypothesis was further supported by results of real-time qPCR analyses, which demonstrated that insects use more than one mechanism to compensate when fed inhibitors of either serine or cysteine protease class. When fed E-64, larvae

had a 5 fold upregulation of the target protease, cathepsin L, indicating that larvae use the hyperproduction/overproduction of proteases from the same protease class as the inhibitor to compete with the inhibitor molecules. The results also indicate that *T. molitor* larvae upregulate proteases from another inhibitor class, which supported the hypothesis that *T. molitor* larvae upregulate cysteine proteases when fed serine protease inhibitors and serine proteases when fed cysteine protease inhibitors. However, this hypothesis was not entirely supported, because in our experiments, *T. molitor* larvae upregulated proteases from two different classes simultaneously in response to dietary inhibitors. Nevertheless, we conclude that the normal insect gut function is effectively disrupted when larvae were fed a combination of inhibitors targeting multiple protease classes. This is the first study to report relative quantitation of fold changes in gene expression in insect larvae fed dietary inhibitors.

## **5.1. Practical Implications and Application**

The results of this study have implications in the control of tenebrionid and other coleopteran stored-product pests. The multiple-inhibitor approach can be adopted as an improved method for pest management based on protease inhibitors. Further studies are required before the effective dose of inhibitors and the appropriate combination of inhibitors can be recommended for use by entomologists involved in crop improvement through plant resistance. Three-week old larvae (3<sup>rd</sup> instar) appear to be the lifestage of *T. molitor* that can be targeted for future pest control strategies.

It is a well-known fact that the high complexity of protease/inhibitor interactions in and the diversity of proteolytic enzymes used by pests need to be considered in order to make the right choice of the appropriate protease inhibitor. A thorough and detailed understanding of the biological system of the insect by assessment is critical to the selection of suitable for pest control. Understanding the molecular and genetic basis of plant resistance and insect counter-defense is imperative, not only for a basic science perspective but also for a biotechnology-based control practice (Zhu-Salzman *et al.*, 2003). Hence, this study was designed to study, in a comprehensive manner, the transcriptional protease profile in the midgut of *T. molitor*.

### ***5.1.1. Testing Effect of Other Inhibitors in vitro and in vivo***

An immediate application of the techniques developed in our research would be the *in vitro* and *in vivo* testing of several commercially available inhibitors for *T. molitor* digestive

response. These studies can be extended to other coleopteran insects (already studied in *T. castaneum*) as well as phytophagous insects from other insect orders. Northern blots and real-time PCR can be further used to provide expression profiles of different proteases in response to dietary protease inhibitors.

### ***5.1.2. Application in Transgenics-Concerns and Considerations***

One major method of applying inhibitor-based plant resistance is to pyramid multiple protease inhibitor genes into plants. Numerous laboratory studies have tested the development of pest-resistance programs based on inhibitor expression in transgenic plants and their impact in turn on non-target organisms (Alfonso-Rubi *et al.*, 2003; De Leo *et al.*, 1998; Irie *et al.*, 1996). Protease inhibitors have also been used in conjunction with other genes that target several physiological and biochemical processes.

Lawrence and Koundal (2004) compiled a list of the many PIs used to create transgenic crop plants and inhibitors isolated from plants with divergent modes of action against different species. Integrating transgenes of interest flanked by vector sequences usually requires a promoter to drive the expression of the protease inhibitor. An ideal promoter needs to be selected, based on its responsiveness to the invasion of the host plant by a pest, or based on regulation by inducers just prior to pest attack. Promoters should be sufficiently active to mediate a substantial defense, and localized to the site of pest invasion. The identification of suitable promoters can be achieved by using promoter-trapping techniques (Babiychuk *et al.*, 1997).

However, the general usefulness of recombinant protease inhibitors in plant protection has yet to be demonstrated. Using one or more genes in combination, whose products target several physiological and biochemical processes in the insect appears to be a possibility. In addition to PI genes, lectins, alpha-amylase inhibitors, cholesterol oxidases, peroxidases, spider venom, avidins, agglutinins, ribosome-inactivating proteins can be engineered into plants (Corbin *et al.*, 1994; Dowd *et al.*, 1998; Down *et al.*, 1996; Fitches *et al.*, 2002; Huang *et al.*, 2001; Ishimoto *et al.*, 1996; Kramer *et al.*, 2000; Yao *et al.*, 1996; Zhou *et al.*, 2001). The use of proteinase inhibitors not only would help in pest control, but also in protecting plants from fungal, viral and bacterial pathogens.

Another major concern in expressing protease inhibitors in transgenic plants would be the effect of protease inhibitors on non-target herbivores. Cowgill and Atkinson (2003) provided a

method to a sequential approach to risk management of transgenic plants expressing protease inhibitors. A tiered approach similar to that used to test the effect of pesticides on beneficial organisms was adapted to test the direct effects of plants expressing protease inhibitors on nontarget herbivorous insects. Such nontarget effects can be evaluated by designing experiments to identify hazard, determine levels of exposure and assess if the hazard constitutes a risk.

## 5.2. Suggestions for the Future

Future experiments can be directed toward testing the effect of several other inhibitors besides STI and E-64. Bioassays can be extended for a longer period of time to watch for potential behavioral, physiological changes manifested by the insect over time. It would be interesting to study the phenotypic effects in adult *T. molitor* that have been reared on inhibitor diets.

There is a need to verify if the inhibitors, in addition to blocking the enzymes directly by binding to them, bind to a chitin column of the peritrophic membrane in the *T. molitor* midgut. Studies by Sale *et al.*, (2001); Macedo *et al.*, (2002, 2004) have demonstrated that a protease inhibitor also binds to a chitin column and to chitinous structures in the midgut of the insect. Immunoblotting may be one method to use to demonstrate the presence or absence of binding of inhibitor to the protease inhibitors.

Microarray analyses, dot blot analysis, or northern blot analyses of *T. molitor* midgut gene expression in response to dietary inhibitors also may reveal new enzymes recruited by the insect to compensate for reduced digestion. In a study that examined the expression profiling in guts of *Callosobruchus maculatus* fed on a diet containing a cysteine protease inhibitor, a significant number of up- and down-regulated genes were described that failed to match any sequences that encoded known proteins (Moon *et al.*, 2004). Since the cDNA library in our study was isolated from guts of larvae fed control (no inhibitor) diet, transcripts that would be upregulated in response to inhibitors may not have been detected. Additional future experiments using microarray analysis and northern blots should be conducted to address this question. Furthermore, anterior midguts and posterior midguts can be used separately in future analyses to study differential up regulation and down-regulation in the compartmentalized midgut simultaneously.

Time-dependant and developmental stage-dependant changes in *T. molitor*'s response to dietary inhibitors may provide more specific clues as to the insects' use of multiple mechanisms or compensation by shifting to another mechanism (Zhu-Salzman *et al.*, 2004). Guts dissected at different time points after feeding on inhibitor-treated diets can be assayed to determine the proteolytic response of the insect over time. In addition, the responses of different larval stages to inhibitors may reveal the most susceptible life stage to target for the application of pest control methods based on PIs.

Protein characterization of the different proteins encoded by the transcripts identified in the course of this study is essential and would complement the molecular analysis of protease gene expression. The identification of protein isoforms and functional roles played by these proteases would be helpful in understanding the overall response of *T. molitor* to dietary inhibitors. Since the cDNA sequences of 50 proteases involved in *T. molitor* protein digestion are now known, dsRNA can be synthesized and used a genomic tool to ascertain the gene's function. Recently, RNAi has been widely used as a tool for probing gene function (Elbashir *et al.*, 2001).

### **5.3. Conclusions**

In conclusion, the digestive system of the yellow mealworm, *T. molitor*, was a model to explore the complexity of the compensational response of larvae to dietary inhibitors. The results from this study provided information on protease genes expressed in the midgut of *T. molitor* larvae, and illustrated the expression patterns of these genes in response to dietary protease inhibitors. This study evaluated the efficacy of a multiple inhibitor strategy to control *T. molitor* larval infestation, and the results indicated that a combination of inhibitors that targeted two different protease classes (serine and cysteine) contributed to significantly reduced larval weights. The molecular analysis suggested that the compensatory response of larvae to inhibitors was disrupted by the combination inhibitor strategy. This information can be valuable for the application of transgenic protease inhibitors for cereal crop improvement.

## 5.4. References

Alfonso-Rubi, J., Ortego, F., Castanera, P., Carbonero, P., Diaz, I. (2003) Transgenic expression of trypsin inhibitor CMe from barley in indica and japonica rice, confers resistance to the rice weevil *Sitophilus oryzae*. *Transgenic Res.* 12:23-31.

Babiychuk, E., Fuangthong, M., Van Montagu, M., Inzé, D., and Kushnir, S. (1997). Efficient gene tagging in *Arabidopsis thaliana* using a gene trap approach. *Proc. Natl. Acad. Sci. USA* 94: 12722–12727.

Bell HA, Down RE, Fitches EC, Edwards JP, Gatehouse AMR. 2003. Impact of genetically modified potato expressing plant-derived insect resistance genes on the predatory bug *Podisus maculiventris* (Heteroptera: Pentatomidae). *Biocontrol Sci. Technol.*13: 729–41.

Bown, D.P., Wilkinson H.S. and Gatehouse J.A. (2004) Regulation of expression of genes encoding digestive proteases in the gut of a polyphagous lepidopteran larva in response to dietary protease inhibitors. *Physiolog. Entomol.* 29: 278-290.

Corbin, D. R., Greenplate, J. T., Wong, E. Y. and Purcell, J. P. (1994) Cloning of an insecticidal cholesterol oxidase gene and its expression in bacteria and in plant protoplasts. *Appl. Environ. Microbiol.* 60: 4239-4244.

Cowgill, S.E. and Atkinson, H.J. (2003) A sequential approach to risk assessment of transgenic plants expressing protease inhibitors: effects on nontarget herbivorous insects. *Transgenic Res.* 12: 439-449.

Cristofolletti, P.T., Ribeiro, A.F. and Terra, W.R. (2005) The cathepsin L-like proteases from the midgut of *Tenebrio molitor* larvae: Sequence, properties, immunocytochemical localization and function. *Insect Biochem. Mol. Biol.* 35: 883-901.

De Leo, F, Bonadé-Bottino, M.A., Ceci, L.R., Gallerani, R. and Jouanin, L. (1998) Opposite effects on *Spodoptera littoralis* larvae of high expression level of a trypsin protease inhibitor in transgenic plants. *Plant Physiol.* 18: 997-1004.

Dowd, P. F., Lagrimini, L. M. and Herms, D. A. (1998) Differential leaf resistance to insects of transgenic sweetgum (*Liquidambar styraciflua*) expressing tobacco anionic peroxidase. *Nat. Toxins* 6: 241-249.

Dowd P. F., Mehta A. D. and Boston R. S. (1998 ) Relative toxicity of the maize endosperm ribosome-inactivating protein to insects. *J. Agric. Food Chem.* 46: 3775-3779.

Down, R. E., Gatehouse, A.M. R., Hamilton, W.D.O. and Gatehouse, J.A. (1996) Snowdrop lectin inhibits development and decreases fecundity of the glasshouse potato aphid (*Aulocorthum solani*) when administered *in vitro* and via transgenic plants both in laboratory and greenhouse trials. *J. Insect Physiol.* 42: 1035-1045.

Elbashir, S. M., Lendeckel, W. and Tuschl, T. (2001) RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev.* 15: 188-200.

Elpidina, E.N., Tsybina, T.A., Dunaevsky, Y.E., Belozersky M.A., Zhuzhikov, D.P. and Oppert, B. (2005) A chymotrypsin-like proteinase from the midgut of *Tenebrio molitor* larvae. *Biochimie* 87: 771-779.

Fitches, E., Audsley, N., Gatehouse, J.A. and Edwards, J. P. (2002) Fusion proteins containing neuropeptides as novel insect control agents: snowdrop lectin delivers fused allatostatin to insect haemolymph following oral ingestion. *Insect Biochem. Mol. Biol.* 32:1653-1651.

Huang, J.Q., Wel, Z.M., An, H.L. and Zhu, Y.X. (2001) *Agrobacterium tumefaciens*-mediated transformation of rice with the spider insecticidal gene conferring resistance to leafhopper and striped stem borer. *Cell Res.* 11: 149-155.

Irie K, Hosoyama H, Takeuchi T, Iwabuchi K, Watanabe H, Abe M, Abe K & Arai S (1996) Transgenic rice established to express corn cystatin exhibits strong inhibitory activity against insect gut proteases. *Plant Mol. Biol.* 30: 149–157.

Ishimoto, M, Sato, T., Chrispeels, M. J. and Kitamura, K. (1996) Bruchid resistance of transgenic azuki bean expressing seed alpha amylase inhibitor of common bean. *Entomol. Exp. Appl.* 79: 309-315.

Jiang, H and Kanost, M.R. (2000) The clip-domain family of serine proteinases in arthropods. *Insect Biochem. Mol. Biol.* 30: 95-105.

Kanost, M.R., Jiang, H., Wang, Y., Yu, X. Q., Ma, C., Zhu, Y. (2001) Hemolymph proteinases in the immune responses of *Manduca Sexta*. *Adv. Exp. Med. Biol.* 484: 319-328.

Kramer, K.J., Morgan, T.D., Throne, J.E., Dowell, F.E., Bailey, M and Howard, J.A. (2000) Transgenic avidin maize is resistant to storage insect pests. *Nature Biotechnology* 18: 670-674.

- Lawrence, P.K. and Koundal, K.R. (2002) Plant protease inhibitors in control of phytophagous insects. *Electron. J. Biotechnol.* 5: 93-109.
- Macedo, M. L.R. and Xavier-Filho, J. (1992) Purification and characterization of trypsin inhibitors from seeds of *Clitoria ternatea*. *J. Sci. Food Agri.* 58: 55-58.
- Macedo, M. L.R., Mara de Sá, C., Freire, M. G. M. and Parra, J.R. P. (2004) A Kunitz-Type Inhibitor of Coleopteran Proteases, Isolated from *Adenanthera pavonina* L. Seeds and Its Effect on *Callosobruchus maculatus*. *J. Agric. Food Chem.*, 52: 2533 –2540.
- Sales, M.P., Pimenta, P.P. , Paes, N.S. Grossi-de-Sá, M.F. and Xavier-Filho, J. (2001) Vicilins (7S storage globulins) of cowpea (*Vigna unguiculata*) seeds bind to chitinous structures of the midgut of *Callosobruchus maculatus* (Coleoptera: Bruchidae) larvae *Braz. J. Med. Biol. Res.* 34: 27-34.
- Terra, W. R. and Cristofolletti, P.T. (1996) Midgut proteinases in three divergent species of Coleoptera. *Comp. Biochem. Physiol.* 113B: 725-730.
- Tsybina, T.A., Dunaevsky, Y.E., Belozersky, M.A., Zhuzhikov, D.P., Oppert, B. and Elpidina, E.N. (2005) Digestive proteinases of yellow mealworm (*Tenebrio molitor*) larvae: Purification and characterization of a trypsin-like proteinase. *Biochemistry (Mosc.)* 70:300-305.
- Thie, N.M.R. and Houseman, J.G. (1990) Cysteine and serine proteolytic activities in larval midgut of yellow mealworm, *Tenebrio molitor* L. (Coleoptera: Tenebrionidae). *Insect Biochem.* 20: 741-744.
- Vinokurov, K. S., Elpidina, E. N., Oppert, B., Prabhakar, S., Zhuzhikov, D. P., Dunaevsky, Y. E. and Belozersky, M. A. (2006a) Diversity of digestive proteases in *Tenebrio molitor* (Coleoptera: Tenebrionidae) larvae. *Comp. Biochem. Physiol., Part B*, in press.
- Vinokurov, K. S., Elpidina, E. N., Oppert, B., Prabhakar, S., Zhuzhikov, D. P., Dunaevsky, Y. E. and Belozersky, M. A. (2006b) Fractionation of digestive proteases from *Tenebrio molitor* (Coleoptera: Tenebrionidae) larvae and role in protein digestion. *Comp. Biochem. Physiol., Part B*, in press.
- Yao, B., Fan, Y., Zeng, Q. and Zhao, R. (1996) Insect-resistant tobacco plants expressing insect-specific neurotoxin AaIT. *Chin. J. Biotechnol.* 12: 67-72.
- Zhou, Y.G., Tian, Y.C., Mang, K. Q. (2001) Cloning of AHA gene from *Amaranthus hypochondriacus* and it's aphid inhibitory effect in transgenic tobacco plants. *Sheng Wu Gong Cheng Xue Bao.* 17: 34-39.

Zhu-Salzman, K., Koiwa, H., Salzman, R.A., Shade, R.E. and Ahn, J. -E. (2003) Cowpea bruchid *Callosobruchus maculatus* uses a three-component strategy to overcome a plant defensive cysteine proteinase inhibitor. *Insect Mol. Biol* 12: 135–145.

## Appendix A - List of Protease Inhibitors

Inhibitor (abbreviated)	Inhibitor	Specificity	Concentratiion used
Pepstatin	-	Aspartic proteases	0.01 mM
Leupeptin	-	Serine and cysteine proteases such as plasmin, Trypsin, Papain, Cathepsin B	0.01 mM
E-64	trans-epoxysuccinyl-L-leucylamido-(4-guanidino) butane	Inhibits papain and other cysteine proteases like cathepsin B and L	0.1mM
PMSF	Phenyl methyl sulfonyl fluoride	Inhibits serine proteases(chymotrypsin, trypsin and thrombin).Also inhibits cysteine proteases such as papain (reversible by DTT treatment)	1 mM
Aprotinin	-	Trypsin, Plasmin, Chymotrypsin, kallikerin	0.01 m M
Chymostatin	-	Specific inhibitor of $\alpha$ -, $\beta$ -, $\gamma$ -, $\delta$ -chymotrypsin.	0.1 mM
TPCK	L-1-Chloro-3-(4-tosyl amido)-4-phenyl-2-butanone)	Irreversibly inhibits chymotrypsin. Also inhibits many other serine and cysteine preteases such as bromelain, ficin, and papain	0.2 mM
TLCK	(L-1-Chloro-3-(4-tosylamido)-7-amino-2-heptanone-HCl)	Irreversibly and specifically inhibits chymotrypsin. Also inhibits many other serine proteases such as bromeleain, ficin and papain	0.2 mM
SBTI	Soybean trypsin inhibitor	Inhibits trypsin. Soybean trypsin inhibitor also inhibits factor Xa, plasmin and plasma kallikerin. Neither inhibit matallo, cysteine, and aspartic proteases or tissue kallikerin	100 grams/ml
Bestatin	[(2S, 2R)-3-Amino-2-hydroxy-4-Phenylbutanoyl]-L-Leucine	Competitive inhibitor of aminopeptidases. Inhibits aminopeptidases e.g aminopeptidase B, leucine aminopeptidase, aminopeptidases on the surface of mammalian cells).	50 $\mu$ g/ml