MOLECULAR MECHANISMS GOVERNING THE DIFFERENTIAL REGULATION OF CYSTEINE PROTEASES IN INSECT ADAPTATION TO A SOYBEAN PROTEASE INHIBITOR

A Dissertation

by

JI EUN AHN

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

August 2008

Major Subject: Entomology

MOLECULAR MECHANISMS GOVERNING THE DIFFERENTIAL REGULATION OF CYSTEINE PROTEASES IN INSECT ADAPTATION TO A SOYBEAN PROTEASE INHIBITOR

A Dissertation

by

JI EUN AHN

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Co-Chairs of Committee,	Keyan Zhu-Salzman
	Linda A. Guarino
Committee Members,	Craig J. Coates
	Hisashi Koiwa
Head of Department,	Kevin Heinz

August 2008

Major Subject: Entomology

ABSTRACT

Molecular Mechanisms Governing the Differential Regulation of Cysteine Proteases in Insect Adaptation to a Soybean Protease Inhibitor.

(August 2008)

Ji Eun Ahn, B.S.;M.S., Jeonbuk National University, Korea

Co-Chairs of Advisory Committee: Dr. Keyan Zhu-Salzman
Dr. Linda A. Guarino

Under challenge by a dietary soybean cysteine protease inhibitor (scN), cowpea bruchids overcome the inhibitory effects by reconfiguring the expression profiles of their major digestive enzymes, the cathepsin L-like cysteine proteases (CmCPs). In addition, cowpea bruchids activate transcription of the counter-defensive cathepsin B-like cysteine protease (*CmCatB*). I undertook an interest in understanding the molecular mechanisms utilized by bruchids to differentially regulate cysteine proteases in response to plant inhibitors. First, to investigate the functional significance of the differential regulation of CmCPs, I expressed CmCP proprotein isoforms (proCmCPs) in *E. coli*, and characterized their activities. Among proCmCPs, proCmCPB1 exhibited the most efficient autocatalytic processing, the highest proteolytic activity, and was able to degrade scN in the presence of excessive CmCPB1. Second, to dissect the molecular mechanisms behind the differential function of CmCPs, I swapped domains between two representative subfamily members B1 and A16. Swapping the propeptides did not

qualitatively alter autoprocessing in either protease isoform. Incorporation of either the N- or C-terminal mature B1 segment into A16, however, was sufficient to prime autoprocessing of A16. Bacterially expressed isolated propeptides (pA16 and pB1) showed that pB1 inhibited B1 enzyme less than pA16 due to its protein instability. Taken together, these results suggest that cowpea bruchids selectively induce specific cysteine proteases for their superior autoprocessing, proteolytic efficacy, and scN-degrading activities, and modulate proteolysis of their digestive enzymes by controlling cleavage and stability of propeptides to cope with plant inhibitors. Third, to understand the transcriptional regulatory mechanisms of *CmCatB* hyperexpression that underlies bruchid adaptation, I cloned a portion of its promoter and demonstrated its activity in *Drosophila* S2 cells using a CAT reporter system. Gel shift assays identified cowpea bruchid Seven-up (CmSvp, chicken ovalbumin upstream promoter transcription factor homolog) in scN-unadapted insect midgut, and cowpea bruchid HNF-4 (CmHNF-4, hepatocyte nuclear factor 4) in scN-adapted insect midgut. When transiently expressed in S2 cells, CmSvp repressed, while CmHNF-4 activated CmCatB expression. CmSvp antagonized CmHNF-4-mediated transactivation when they were present simultaneously in the cell. Thus, the data suggest that transcriptional regulation of *CmCatB* in response to plant inhibitor depends, at least partly, on the cellular balance between positive and negative regulators.

ACKNOWLEDGEMENTS

I would like to thank my committee co-chairs, Dr. Zhu-Salzman and Dr. Guarino, and my committee members, Dr. Coates and Dr. Koiwa for their guidance and support throughout the course of this research.

Thanks also go to my friends and colleagues and the department faculty and staff for making my time at Texas A&M University a great experience. I also want to extend my gratitude to the National Research Initiative of the USDA Cooperative State Research, Education and Service (USDA-NRI) and the DuPont Young Professor Grant Program, which supported the research described in this dissertation.

Finally, thanks to my mother and father for their endless encouragement and support, and to my husband for his patience and love.

NOMENCLATURE

Anti-AaSvp	Mosquito *Aedes aegypti* Seven-up antibody
CAT	Chloramphenicol acetyl transferase
CmCatB	Cowpea bruchid cathepsin B-like cysteine protease
CmCP	Cowpea bruchid cathepsin L-like cysteine protease
CmCPA	Cowpea bruchid cathepsin L-like cysteine protease from subfamily A
CmCPB	Cowpea bruchid cathepsin L-like cysteine protease from subfamily B
CmHNF-4	Cowpea bruchid hepatocyte nuclear factor 4
CmSvp	Cowpea bruchid Seven-up, chicken ovalbumin upstream promoter-transcription factor (COUP-TF) homolog
COUP	Chicken ovalbumin upstream promoter *cis*-element
EMSA	Electrophoretic mobility shift assay
HNF-4	Hepatocyte nuclear factor 4 *cis*-element
proCmCP	Cowpea bruchids cathepsin L-like cysteine protease proprotein isoform
pA16	propeptide cathepsin L-like cysteine protease 16 from subfamily A
pB1	propeptide cathepsin L-like cysteine protease 1 from subfamily B
S2 cells	*Drosophila* Schneider 2 cells

scN	Soyacystatin N, Soybean cysteine protease inhibitor

TABLE OF CONTENTS

Page

ABSTRACT iii

ACKNOWLEDGEMENTS v

NOMENCLATURE vi

TABLE OF CONTENTS viii

LIST OF FIGURES x

LIST OF TABLES xii

CHAPTER

I INTRODUCTION 1

1.1 Control of pest cowpea bruchid using plant defense compounds 1
1.2 Insect adaptation to plant defense system 17
1.3 Objectives of this work 22

II FUNCTIONAL ROLES OF SPECIFIC BRUCHID PROTEASE ISOFORMS IN ADAPTATION TO A SOYBEAN PROTEASE INHIBITOR 29

2.1 Introduction 29
2.2 Results 32
2.3 Discussion 42
2.4 Experimental procedures 45

III COWPEA BRUCHID *CALLOSOBRUCHUS MACULATUS* CONTERACTS DIETARY PROTEASE INHIBITORS BY MODULATING PROPEPTIDES OF MAJOR DIGESTIVE ENZYMES 51

3.1 Introduction 51
3.2 Results 54
3.3 Discussion 65

CHAPTER Page

3.4 Experimental procedures……... 72

IV SEVEN-UP FACILITATES INSECT COUNTER-DEFENSE BY SUPRESSING CATHEPSIN B EXPRESSION……..... 79

4.1 Introduction……... 79
4.2 Results……... 81
4.3 Discussion. ……... 94
4.4 Experimental procedures..... 100

V ANTAGONISM BETWEEN HEPATOCYTE NUCLEAR FACTOR 4 AND SEVEN-UP MODULATES INSECT COUNTER -DEFENSE CATHEPSIN B EXPRESSION……..... 112

5.1 Introduction……... 112
5.2 Results……... 114
5.3 Discussion..... 125
5.4 Experimental procedures……... 133

VI SUMMARY……. 142

REFERENCES……... 145

VITA..... 176

LIST OF FIGURES

FIGURE		Page
2.1	Bacterial expression of the major digestive cysteine proteases of cowpea bruchids	33
2.2	proCmCPB1 can autoprocess within a broader pH range than proCmCPA9	35
2.3	CmCPs differ in their enzymatic activities as well as in autoprocessing capability	37
2.4	Two-step autoprocessing of CmCPA9 precursor	38
2.5	CmCPA16 precursor can be processed by insect gut cysteine or aspartic proteases	39
2.6	CmCPB1 is scN-sensitive but can degrade scN when in excess	41
3.1	Design and production of recombinant CmCP chimeras	55
3.2	Differential autoprocessing and enzymatic activities of chimeras	58
3.3	Activating A16 by incorporating single or double B1 residues	60
3.4	The propeptide of A16 has higher inhibitory activity and is more resistant to proteolysis than that of B1	62
3.5	B1 propeptide is required for scN degrading activity	64
4.1	Architecture of genomic DNA upstream of *CmCatB* coding region	82
4.2	Illustration of the promoter activity of the 493 bp fragment in *Drosophila* cells	83
4.3	Probe dissection to locate *cis*-elements using EMSA	85
4.4	Nuclear protein factors specifically interact with *COUP* element	87

FIGURE Page

4.5 Anti-AaSvp serum abolished the *COUP*-nuclear protein association….. 89

4.6 CmSvp is more abundant in scN-unadapted cowpea bruchid midgut than scN-adapted midgut..... 89

4.7 CmSvp shares high sequence similarity with COUP-TF/Svp members from red flour beetle TcSvp, mosquito AaSvp, *Drosophila* DmSvp, as well as human COUP-TF..... 91

4.8 CmSvp represses *CmCatB* expression 92

4.9 CmSvp repression of *CmCatB* requires binding at the *COUP* element..... 94

5.1 Defining regulatory *cis*-element(s) in *CmCatB* promoter that specifically interact with scN-adapted insect midgut nuclear protein(s)..... 115

5.2 A nuclear protein other than CmSvp interacts with *HNF-4/COUP* element in nuclear extract of scN-adapted insect midgut..... 117

5.3 Conservation of cowpea bruchid HNF-4 protein sequence……. 119

5.4 CmHNF-4 is a functional DNA binding protein that activates *CmCatB* expression……. 121

5.5 CmHNF-4 activation of *CmCatB* requires binding at the *HNF-4/COUP* element……..... 123

5.6 CmSvp antagonizes trans-activation activity of CmHNF-4……. 125

LIST OF TABLES

TABLE		Page
3.1	Proteolytic activities on Z-Phe-Arg-pNA of pBmAB, pBmBA and B1 following 1, 2 and 6 hr autoprocessing	64
3.2	Specific activities of selected CmCPs on Z-Phe-Arg-pNA at pH 5.0.....	64
3.3	Summary of autoprocessing and activity of chimeric and original proteins……...	67
4.1	Oligonucleotide primers synthesized for EMSAs……..	106
5.1	Oligonucleotide primers synthesized for EMSAs in scN-adapted bruchids …….....	135

CHAPTER I

INTRODUCTION

1.1 Control of pest cowpea bruchid using plant defense compounds

The grain legume cowpea, *Vigna unguiculata* L. (Walp.), is widely grown in the tropics and subtropics where this crop represents an important source of proteins in human diet. It is also known as southern pea or black-eyed pea and is an important grain legume in developing countries, particularly in Africa. Cowpea is tolerant to poor, dry soil conditions and provides a protein and energy rich food source for human and livestock. The demand for cowpeas is high but yields remain critically low, mainly due to insect pests, which attack virtually every developmental stage of the crop. Infestations are commonly so heavy that the seeds are unsuitable for use as food, feed, or planting.

The most damaging storage pest in cowpea is bruchid beetle, *Callosobruchus maculatus* (F.) (Coleoptera: Chrysomelidae: Bruchidae). Cowpea bruchid adults are reddish-brown slightly elongate and measure 2.0-3.5 mm in length. Females can lay up to 100 eggs and glue oval eggs on seed surface. The eggs are translucent grey in color and have a flat base. Within 5-6 days the eggs become white and the larvae hatch, and the first instar larvae bore through the testa and cotyledons. Cowpea bruchid is an internal-feeding insect, which spends its entire larval stage (four instars) feeding within the seed. Developing larvae feed only in a single seed, and chew near the surface and

---

This dissertation follows the style and format of *Insect Molecular Biology*.

leave a thin covering uneaten which appears as a "window". Pupation occurs inside the seeds for 3-4 days, and adults emerge from the "window". Adults do not feed on stored products and usually live not longer than 12 days. Development of cowpea bruchids takes about 20- 30 days depending on the temperature. Thus, on average, one life cycle is completed every month under tropical condition. Bruchid thus can cause severe loss of stored cowpea grain due to its short generation time, multiple generations, and high reproductive ability.

Bruchid controls are performed by treating stored seeds with fungicides, methyl bromide, carbon disulfide, and several other chemicals. These are considered environmentally undesirable and are too expensive for subsistence farmers. Therefore, transgenic cowpeas may represent the next promising strategy for control of cowpea bruchids. Genetic transformation of cowpea has not been successful (Garcia et al. 1986; Ikea et al., 2003). Recently, however, there has been progress in production of transgenic cowpea using *Agrobacterium*-mediated genetic transformation, and this demonstrates for the first time that stable transmission of the transgenes to progeny is possible (Popelka et al., 2006). This technology paves the way to develop cowpea germplasm with increased insect resistance.

Plants have evolved a certain degree of resistance to herbivores and pathogens through production of defense compounds, which may be toxic secondary compounds (e.g. antibiotics, alkaloids, terpenes, cyanogenic glucosides) and protein antimetabolites (e.g. chitinases, β-1,3-glucanases, lectins, arcelins, vicilins, systemins, enzyme inhibitors

such as protease inhibitors and α-amylase inhibitors). Among these compounds, lectins, protease inhibitors, α-amylase inhibitors and vicilins have been studied as candidate genes for transgenic plants resistant to cowpea bruchids.

1.1.1 Lectins

Lectins are a class of proteins of non-immune origin that possess at least one non-catalytic domain that specifically and reversibly binds to mono- or oligosaccharide (Peumans & Van Damme, 1995). Lectins are extensively distributed in nature and several hundred of these molecules have been isolated so far from plants, viruses, bacteria, invertebrates and vertebrates, including mammals. Seeds, particularly those of the Leguminoseae, are rich sources of lectins. Because of the specific interaction of lectins with glycoconjugates at a cell surface, they have capability to serve as recognition molecules within a cell, between cells, or between organisms.

There are still controversies regarding the biological roles of plant lectins in the parent organisms. Many studies, however, have indicated that plant lectins have a role in plant defense against herbivores and pathogens (Chrispeels & Raikhel, 1991; Peumans & Van Damme, 1995). Various plant lectins have shown entomotoxic effects when fed to insects from Coleoptera, Homoptera, and Lepidoptera orders. The common bean *Phaseolus vulgaris* contains phytohemagglutinin (PHA), an abundant hemagglutinin and mitogen. This tetrameric lectin is composed of five isoforms of the polypeptides PHA-E and PHA-L in different combinations. PHA-E and PHA-L both

recognize terminal galactose residues on complex glycans of mammalian glycoproteins. PHA was the first lectin to which anti-insect activities were ascribed on the basis of its deleterious effect on the cowpea bruchid larvae (Janzen et al., 1976; Gatehouse et al., 1984). Ironically, this first indication of a plant defense role for lectins against insects was based on false-positive result, since the effects were due to contamination with α-amylase inhibitor during protein purification (Huesing et al., 1991c). Murdock et al. (1990) observed that plant lectins from peanut (*Arachis hypogaea*), osage orange tree (*Maclura pomifera*), wheat germ (*Triticum aestivum)*, jimson weed (*Datura stramonium*) and potato (*Solanum tuberosum*), had an inhibitory effect on the larval development of cowpea bruchid. These lectins have specificities for *N*-acetylgalactosamine/galactose residues (GalNAc/Gal) and *N*-acetylglucosamine residues (GlcNAc). The most potent lectin was from hexaploid wheat, *T. aestivum*, known as wheat germ agglutinin (WGA). WGA is one of the chitin-binding lectins that contain one or more hevein domains referring to a 43 amino acids chain-binding polypeptide present in latex of the rubber tree *Hevea brasiliensis* (Chrispeels & Raikhel, 1991). WGA specifically binds the sugar GlcNAc, and consists of three isolectins (A, B, and D). All three isolectins of WGA equally contributed to inhibit the growth and development of cowpea bruchid (Huesing et al., 1991a). The chitin-binding lectins from rice (*Oryza sativa*) and stinging nettle (*Urtica dioica*) also inhibited the larval growth of cowpea bruchid (Huesing et al., 1991b). GSII, a GlcNAc-specific lectin from the Africa legume *Griffonia simplicifolia* has insecticidal effect on cowpea bruchid larvae (Zhu et al., 1996). Galactose-specific lectins from African yam beans (*Sphenostylis stenocarpa*)

delayed larval development of cowpea bruchid and larval mortality also was detected (Machuka et al., 2000). Lectin from *Talisia esculenta* produced approximately 90% mortality to cowpea bruchid at dietary level of 2% (w/w), and was resistant to midgut proteolysis of cowpea bruchid (Macedo et al., 2002).

The epithelial cells along the digestive tract of herbivores are directly exposed to the contents of the diet and, therefore, are possible target sites for plant defense proteins. Since glycoproteins are major constituents of these membranes, the luminal side of the midgut is literally covered with potential binding sites for dietary lectins. However, the precise mechanism(s) of action of lectins in insects are not well understood. There are possible four types of interactions, which cause disruption in feeding and digestion, and thereby growth and development. The first is the binding of lectins to the chitin in the peritrophic membrane (PM) of the insect digestive system (only for chitin-binding lectins). The PM is a selectively permeable structure that is secreted by the mesenteron cells of the midgut. Most insect PMs consists of chitin (a polymer of N-acethyl-D-glucosamine) and proteins, some of which may be glycosylated. The PM aids in digestion and absorption of nutrients, and protects the delicate microvillar brush border of the midgut epithelium from contact with rough food particles and microbes. Ingestion of WGA by European corn borer *Ostrinia nubilalis* (ECB) larvae caused inhibition of larval growth and appearance of abnormalities in the PM structure (Harper et al., 1998). The PM was thin and compact in the control larvae, whereas the PM of the WGA-fed larvae was multilayered and discontinuous, which allowed plant cell-wall fragments to penetrate into the microvilli of the epithelium. Most of ingested WGA bound to the

nascent chitin microfibrillar meshwork of the PM, which resulted in a breakdown in the chitinous meshwork and hypersecretion of abnormal PM. Inhibition of ECB growth by WGA appears to be due to disruption of PM formation that allows physical contact of ingested material with the microvillar brush border and eventual cessation of feeding by the larvae.

The second type of inhibitory interaction is through binding of lectins to the brush border membrane of the digestive epithelial cells in the insect midgut. These cells secrete digestive enzymes and absorb the chemical products of digestion. Numerous studies have demonstrated that dietary lectins can bind to the surface of the midgut epithelial cells. GNA (*Galanthus nivalis* agglutinin; snowdrop lectin) is a mannose-binding lectin, and has been shown to be toxic to rice brown planthoppers *Nilaparvata lugens* (Powell et al., 1998). GNA binding was concentrated on the luminal surface of the midgut epithelial cells, suggesting that GNA binds cell surface carbohydrate moieties in the gut. GNA binding also caused disruption of the microvilli brush border region and abnormalities in the midgut epithelial cells. These morphological changes would explain the delay nymphal development caused by GNA, and suggest a mechanism for mortality, based on breakdown of gut function. Zhu-Salzman et al. (1998) showed that the insecticidal activity of GSII against cowpea bruchids requires carbohydrate binding activity and biochemical stability of the GSII to insect digestive proteolysis. The insecticidal effect of TEL on cowpea bruchid involves a specific carbohydrate-lectin interaction with glycoconjugates on the surface of digestive tract epithelial cells, as well as resistance to enzymatic digestion by insect cysteine proteases (Macedo et al., 2004).

The third mechanism is lectin action within the insect body. If dietary lectin can survive proteolysis by insect digestive enzymes in the alimentary tract and be absorbed into the circulation, then it might pass to any site within the body through the circulating the hemolymph. In rice brown planthopper fed on GNA-diet, GNA binding was observed in the fat bodies, the ovaries, and through out the hemolymph (Powell et al., 1998). This suggests that GNA is able to cross the midgut epithelial barrier, and pass into the circulatory system, resulting in a systemic toxic effect. The fourth possibility is binding of lectins to glycosylated digestive enzymes, thus inhibiting their activity.

Chronic ingestion of lectins can cause hypertrophy of the insect gut. For example, Concanavalin A (Con A), the lectin present in the jackbean *Canavalia ensiformis*, acts as a growth factor causing stimulation of larval gut growth in tomato moth *Lacanobia oleracea* fed on Con A-diet for 16 days (Fitches & Gatehouse, 1998).

1.1.2 Protease inhibitors

Proteolytic enzymes, also called proteases or peptidases, catalyze the hydrolytic cleavage of specific peptide bonds in their target proteins. Proteases are thus responsible for the complete hydrolysis of proteins down to amino acids. They include the proteinases (endopeptidases) and the exopeptidases. Proteinases are the enzymes that cleave protein chains at specific sites within proteins and exopeptidases remove amino acids sequentially from their carboxyl or amino-terminus. Proteases are classified according to their mechanism of catalysis and amino acid present in the active center: (1)

serine proteases, with a serine and histidine; (2) cysteine proteases, with a cysteine; (3) aspartic proteases, with an aspartate group and (4) metalloproteaes, with a metallic ion (Zn^{2+}, Ca^{2+}, Ni^{2+} or Mn^{2+}) (Neurath, 1984). All of these types have been found in insect guts (Terra & Ferreira, 1994).

Proteases play key roles in many biological processes, such as blood coagulation and fibrinolysis, the release of protein hormones from precursor molecules, the transport of secretory proteins across membrane, the assembly of macromolecular structures (collgen fibers or certain viruses) and the control of proteolytic digestion itself. Proteases are not only a physiological necessity but also a potential hazard, since, if uncontrolled, they can destroy the protein components of cells and tissues. Two principal regulatory mechanisms of proteases are the synthesis of proteases as inactive proenzymes (zymogens) and activation of proenzymes by limited proteolysis. But theses do not fulfill the desired level of regulation, and cells and organisms require additional means of control. Protease inhibitors (PIs) are, thus, considered an important control method for protease activity. PIs are classified according to the type(s) of enzyme they inhibit (Ryan, 1990). A large number of naturally occurring PIs have been described in animals, plants and microorganisms. Plant PIs are generally small proteins that mainly occur in storage tissues, such as tubers and seeds. They are also induced as plant defensive compounds in response to insect and pathogen attacks. The defensive capabilities of plant PIs depend on types of proteases present in insect midguts or secreted by microorganisms, causing a reduction in the availability of amino acids necessary for their growth and development (Lawrence & Koundal, 2002).

Lepidopteran, hymenopteran, orthopteran and dipteran species mainly use serine proteases to digest plant foods, beause the pH of insect guts is usually in the alkaline range (8- 11) where serine proteases and metalloproteases are most active (Ryan, 1990). Plant serine PIs fall into a number of structurally distinct subfamilies based on their amino acid sequences (Bode & Huber, 1992). Bowman-Birk type inhibitors are small polypeptides (8 kD), typically found in legume seeds. Kunitz-type inhibitors are 20 kD proteins.

Cysteine proteases are the major digestive enzymes in the Coleoptera, which characteristically have mildly acidic pHs in their midguts near the pH optima of cysteine proteases (pH~5) (Murdock et al., 1987; Terra & Ferreira, 1994; Koiwa et al., 2000). Cysteine proteases are also active in the reducing environments. Thiol reducing compounds including DTT, cysteine, glutathione and β-mercaptoethanol are their activating agents (Kitch & Murdock et al., 1986). Based on structural features and subcellular localization, members of the cystatin or cysteine PIs superfamily in mammals are divided into three different subfamilies: family I or Stefins (11 kDa), family II (11-13 kDa) and family III or kininogens (60-120 kDa) (Barrett, 1987). In plants, they are known as phytocystatins (5-87 kDa) and show characteristics found in cystatins subfamilies I and II (Turk et al., 1997).

The protein inhibitors belonging to the stefin family are all single-chain proteins and lack carbohydrate and disulfide bonds. They consist of approximately 100 amino acid residues and are the smallest among the members of the cystatin superfamily. They are mainly acidic proteins and are found in various cells and tissues in animals. In

addition to the animal stefins, some plant inhibitors have been identified. Members of family II are slightly larger than the stefins and contain approximately 115 amino acid residues. They are usually nongylcosylated, single chain proteins and have two intramolecular disulfide bridges. The family consists mainly of species variants of cystatin C, cystatin S and its variants, and cystatin D. Kininogens are large multifunctional glycoproteins in mammalian blood plasma, and were first known as parent molecules for vasoactive peptides, the kinins. Three different types of kininogens have been identified, high molecular weight kininogen (HK), low molecular weight kininogen (LK) and T-kininogen (TK).

Aspartic proteases (cathepsin D-like proteases) were found along with cysteine proteases in species of six families of Hemiptera (Houseman & Downe, 1983). Aspartic protease activity has been observed in midgut homogenates of cowpea bruchids (Silva & Xavier-Filho, 1991). Zhu-Salzman et al. (2003) also showed that aspartic protease activity was significantly increased in cowpea bruchid midgut when major digestive enzymes were inhibited by dietary scN. The low pH of midguts of many members of coleoptera and hemiptera provides more favorable environments for aspartic proteases (pH optima ~ 3-5) than the high pH of most insect guts (pH optima ~ 8-11) (Houseman et al., 1987). The cathepin D inhibitor is unusual in that it inhibits trypsin and chymotrypsin as well as cathepsin D, but does not inhibit aspartic proteases such as pepsin, rennin or cathepsin E.

Both proteinaceous and nonproteinaceous cysteine PIs frequently have deleterious effects on cowpea bruchids, which may include impaired digestion, reduced

fecundity, delayed growth and development and increased mortality (Murdock et al., 1988; Koiwa et al., 1998; Zhu-Salzman et al., 2003). Soyacystatin N (scN), a soybean cysteine protease inhibitor, significantly has deleterious effects on cowpea bruchid (Koiwa et al., 1998). Biochemical analyses of proteolytic activities indicate that cysteine PIs inhibit digestive cysteine proteases in cowpea bruchid gut (Kitch & Murdock et al., 1986; Zhao et al., 1996; Zhu-Salzman et al., 2003). When cowpea bruchids were reared on artificial seeds containing a recombinant scN, a pepstatin A (aspartic PI), and soybean Kunitz trypsin inhibitor, pyramiding different classes of PIs caused synergistic insecticidal effects (Amirhusin et al., 2007).

1.1.3 α-Amylase inhibitor

α-Amylases (α-1,4-glucan-4-glucanohydrolases) are widespread hydrolytic enzymes found in microorganisms, animals and plants. They catalyze the initial hydrolyses of α-1,4-linked sugar polymers, such as starch and glycogen, into shorter oligosaccharides, an important step towards transforming sugar polymers into single units that can be assimilated by the organism. The enzyme plays a key role in carbohydrate metabolism. Furthermore, several insects such as seed weevils that feed on starchy seeds during larval and/or adult stages rely on their α-amylases for survival.

Multiple forms of α-amylases have been detected in the midgut lumen of cowpea bruchid and Mexican bean weevil *Zabrotes subfasciatus* (Campos et al., 1989; Silva et al., 1999). Starch digestion in cowpea bruchid larvae is accomplished by the joint action

of at least four types of α-amylases, and all enzymes have pH optima in the 5.2- 6.0 range, which is close to the acidic pH of cowpea bruchid midgut (Campos et al., 1989). An α-amylase inhibitor from wheat inhibited all enzymes except the α-amylase with the highest molecular weight among them. The carbohydrate digestion of cowpea bruchids and *Z. subfasciatus* occurs mainly in the lumen of the midgut whereas protein digestion should take place partly in the lumen and partly at the cell surface (Silva et al., 1999).

The class of nonproteinaceous α-amylase inhibitors contains diverse types of organic compounds such as acarbose, isoacarbose, acarviosine-glucose, hibiscus acid and the cyclodextrins (Franco et al., 2002).

Proteinaceous α-amylase inhibitors are found in microoragnisms, plants and animals (Ryan, 1990; Franco et al., 2000). Different plant α-amylase inhibitors exhibit different specificities against α-amylases from diverse sources. α-Amylase inhibitors can be classified according to their tertiary structure into six classes: lectin-like, knottin-like, cereal-type, Kunitz-like, γ-purothionin-like and thaumatin-like (Franco et al., 2002). The lectin-like α-amylase inhibitors (α-AIs) have been purified and characterized from different accessions and varieties of the *P. vulgaris*. The best-characterized isoform, known as α-AI1, was cloned and identified as an α-amylase inhibitor homologous to PHA (Moreno & Chrispeels, 1989). The knottin-type α-amylase inhibitor from *Amaranthus hypocondriacus* seeds (AAI) is the smallest proteinaceous inhibitor of α-amylase yet described, with just 32 residues and three disulfide bonds (Chagolla-Lopez et al., 1994). It contains a knottin fold; three antiparallel β strands and a characteristic disulfide topology. AAI specifically inhibits insect α-amylases and is inactive against

mammalian enzymes. Thus, this provides a highly specific potential weapon in plant defense. The cereal-type α-amylase inhibitors are composed of 120- 160 amino acid residues forming five disulfide bonds (Franco et al., 2000). The Kunitz-like α-amylase inhibitors contain around 180 residues and four cysteines. The barely α-amylase/subtilisin inhibitor (BASI) is involved in regulating the degradation of seed carbohydrate, preventing the endogenous α-amylase 2 from hydrolysing starch during premature sprouting (Kadziola, et al., 1998). The thaumatin-like α-amylase inhibitors are proteins with molecular masses of about 22 kDa with significant sequence similarity to pathogenesis-related group 5 (PR-5) proteins and to thaumatin, an intensively sweet protein from *Thaumatococcus danielli* fruit. The γ-purothionins-like α-amylase inhibitors have 47 or 48 residues, are sulfur-rich and form part of γ-thionin superfamily. Members of this superfamily are involved in plant defense through a variety of mechanisms: modification of membrane permeability (Thevissen et al., 1996), inhibition of protein synthesis (Mendez et al., 1996) and protease inhibition (Wijaya et al., 2000).

The lectin-like α-amylase inhibitors from *P. vulgaris* seeds have been shown to be detrimental to the development of cowpea bruchid and Azuki bean weevil *Callosobruchus chinensis* (Ishimoto & Kitamura, 1989; Huesing et al., 1991c; Shade et al., 1994). In transgenic pea *Pisum sativum* seeds, complete resistance to cowpea bruchids was detected at dietary levels of 0.8-1.0% (w/w) αAI1 with complete larval mortality of the first or second instars (Shade et al., 1994). γ-Thionins purified from cowpea seeds (*V. unguiculata*) inhibit α-amylase from cowpea bruchid larvae (Melo et al., 1999). Four different classes of α-amylase inhibitors from baru nut seeds (*Dipteryx*

alata) were isolated. They inhibited α-amylase activity of cowpea bruchid, and bioassays showed an enhanced mortality rate and reduced insect longevity in cowpea bruchid fed on diet containing baru crude extract (Bonavides et al., 2007).

1.1.4 Vicilin

Seed storage proteins are mostly the classically known globulins, which are insoluble in water and are typically present in leguminous seeds. Legume globulins are a multigene family, and usually occur as two size classes, 7S and 11S, according to their sedimentation coefficients. The two classes also differ in other physical and chemical characteristics. The 11S is generically deginated as legumin. The legumins are generally encoded by a large number of genes and translation products are proteolytically processed to the final two-chain proteins, which are linked by disulfide bonds. They aggregate to form hexamers and are also known as 11S storage globulins (Sales et al., 2000).

Vicilins are also known as 7S storage globulins and are composed of single-chain proteins without disulfide linkage. They aggregate to form trimers of subunits with varying molecular masses (45-53 kD). They are also encoded by a large number of genes (Sales et al., 2000).

Three cowpea cultivars were found in Nigeria that showed a high level of resistance to cowpea bruchid and they were incorporated as a trait in commercial cowpea lines (Singh et al., 1985; Singh & Singh, 1992). IT81D-1045, a resistant seed

descendant from the Nigerian resistant line, was studied as a source of resistance of domesticated lines to cowpea bruchid through its defensive proteins, such as enzyme inhibitors and vicilins. Macedo et al. (1993) first showed that vicilin is involved in the resistance of cowpea seeds to bruchid. A globulin fraction from IT81D-1045 seeds was detrimental to cowpea bruchid when incorporated in artificial seeds, due to the presence of vicilins in its globulin fractions. Cowpea bruchid emerged from resistant cowpea seeds excreted 7 times higher vicilin and 0.4 times less trypsin inhibitor than that emerged from susceptible seeds. This finding indicates that vicilins from resistant seeds are involved to significantly delay larval development and reduce insect emergence (Sales et al., 2005). Vicilins of other legume species that are non-host seeds to cowpea bruchid, such as jack bean (*Canavalia ensiformis*), soybean (*Glycine max*), common bean (*P. vulgaris*) and lima bean (*Phaseolus lunatus*), when added to artificial seeds, had strong deleterious effects on cowpea bruchid larvae. However, another vicilin from adzuki bean (*Vigna angularis*), which is a species closely related to cowpea (*V. unguiculata*) had no effect on bruchid larvae (Yunes et al., 1998). This finding probably explains why some vicilins are detrimental and others have no effect, which is a consequence of the association of the bruchids with their legume seed hosts. Consistently, analysis of the testa (seed coat) of *P. vulgaris* suggests that phaseolin (vicilin-like 7S storage globulin) detected in the testa is contributes to the resistance of *P. vulgaris* to bruchid attack (Silva et al., 2004).

The mechanism of action of the vicilins is not yet understood, however, several possible mechanisms have been proposed. First, the antimetabolic effects of legume

vicilins may be due to resistance to proteolysis by insect proteases. The vicilins isolated from resistant seeds had no effect on the development of the Mexican bean weevil *Z. subfasciatus*, which can also attack cowpea seeds (Macedo et al., 1993). A possible explanation is that vicilins from resistant seeds are less susceptible to digestion by the midgut proteases of cowpea bruchid than by those of *Z. subfasciatus* (Sales et al., 1992).

Second, the toxicity of vicilins seems to be correlated to their chitin-binding property and binding affinity to the epithelial cell surface of insect midgut. Cowpea seed vicilins bound to the chitin of microvilli in the larval midgut of bruchid. Bruchid-susceptible vicilins bound less to insect midgut than bruchid-resistant vicilins. In contrast to chitin-binding lectins, the hemolymph surrounding the gut was completely free of vicilin binding, suggesting that no alteration of the midgut structure with disruption of the microvilli or abnormalities in epithelial cells had occurred (Sales et al., 2001). Bruchid-resistant vicilins bound more strongly to chitin structures of the bruchid midgut than those of *Z. subfasciatus* (Firmino et al., 1996).

Moura et al. (2007) showed that toxicity of vicilin from *Enterolobium contortisiliquum* seeds on cowpea bruchid could be explained by the binding to chitin of insect midgut and low vicilin digestibility by insect proteases.

Third, there is the possibility of systemic effects of toxic vicilins on insect. Vicilins were detected not only in the midgut but also in hemolymph, fat body and Malpighian tubules of bruchid larvae fed on diet containing vicilins, suggesting that vicilins are able to cross the larval midgut epithelium. Vicilins bound to brush border membrane vesicles (BBMVs), suggesting the existence of specific receptors (Uchoa et

al., 2006). These findings suggest that systemic effects of toxic vicilins can be obtained through transport of vicilins from alimentary tract across the gut epithelial cells into the hemolymph and internal organs. There is little direct evidence to indicate the role of membrane proteins such as BBMVs that function as receptors in the transport of intact proteins into the hemolymph.

1.2 Insect adaptation to plant defense system

Insects and plants have been coevolving over 400 million years. And as a result, insects have very diverse tactics to neutralize plant defense mechanisms, which allow insects to minimize the impact of plant defense molecules and increase feeding efficiency. Counter-defense strategies can be grouped into physical and biochemical adaptation, morphological adaptation, symbiont acquisition, and manipulation of host plants such as induction of plant galls and plant susceptibility, trenching, and gregarious feeding (Karban & Agrawal, 2002).

Physical and biochemical adaptation comprise multiple mechanisms that include contact and ingestion avoidance, excretion, sequestration, metabolic resistance, and target-site mutation (Després et al., 2007). Avoidance of plant-defense chemicals can be genetically determined or gained by a learning process after experience with the toxic plant foods. Contact avoidance involves unique behaviors such as vein cutting or early deactivation of plant-defense chemicals before feeding. The caterpillar of the monarch butterfly *Danaus plexippus* feeds exclusively on milkweeds, which contains a variety of

toxins stored in pressurized latex canals. To avoid these toxins, late-instar larvae deactivate the latex defense by cutting veins before feeding (Helmus & Dussourd, 2005). Herbivores can also avoid contact with toxin by suppressing plant defense. Glucose oxidase is the principal salivary enzyme in tobacco earworm *Helicoverpa zea*, and it reduces the amount of toxic nicotine released by the tobacco plant *Nicotiana tabacum* through suppression of induced plant resistance (Musser et al., 2002).

In many insects, a large proportion of the accumulated toxic plant compounds can be excreted, or lost with exuvia during the molt (Zagrobelny et al., 2004). Insects can also sequester plant toxins, and subsequently use it as a defensive substance against predators or pathogens. Monarch butterfly larvae sequester milkweed cardenolides that render adults unpalatable to predators (Nishida, 2002). Cardenolides are also cardio-active steroid causing toxic effects on various vertebrates. Monarch adult thus can use sequestered cardenolides as a defense against predatory birds.

Herbivorous insects utilize the detoxification of plant toxins as one of major weapons to evolve in their coevolutionary arms race with plants. Metabolic resistance mainly involves overproduction of detoxification enzymes that can metabolize xenobiotics. Detoxification enzymes usually belong to one of three families: the cytochrome P450 monooxydases (P450s or *CYPs* for genes), the glutathione S-transferases (GSTs), and the carboxylesterases (CARs). P450s appear to play a key role in plant-insect interactions, and thus have been intensively investigated. P450s comprise a large superfamily of heme-thiolate enzymes that metabolize a wide range of both endogenous and exogenous hydrophobic compounds by incorporating oxygen into a

functionalized product. The adaptation of Papilionidae lepidopterans to overcome plant toxin is a very well documented example of the roles of P450 in plant-insect interactions (Petersen et al., 2001). The black swallowtail butterfly *Papilio polyxenes* induces furanocoumarin metabolism in detoxifying organs such as midguts and fat bodies allowing them to feed exclusively on plants containing furanocoumarins. *Aedes aegypti* larvae are able to ingest toxic arborescent leaf litter via overproduction of P450s, and this adaptive response enabled mosquitoes to colonize new habitats (David et al., 2006). The glutathione transferases (GSTs) are a large family of multifunctional enzymes involved in the detoxification of a variety of xenobiotics including insecticides (Enayati et al., 2005). GSTs primarily catalyze the conjugation of electrophilic toxins with the thiol group of reduced glutathione (GSH), generally increasing solubility of toxins and facilitating their elimination by insects. The adaptation of numerous crop-feeding lepidopteran species to plant chemicals through induction of GSTs has been studied (Yu, 1996). Carboxylesterases (CARs) can hydrolyze ester bonds from various substrates with a carboxylic ester. When brown planthopper *Nilaparvata lugens* (BPH) fed on BPH-resistant rice plants, Northern hybridization analysis showed that the expression of carboxylesterase was up-regulated in 4th instar nymphs (Yang et al., 2005).

Other metabolic resistance mechanisms have been identified in various plant-insect interactions. Specialized *Heliconius* caterpillars use a novel enzymatic system to convert cyanogenic glycosides into thiols (Engler *et al*., 2000). This process blocks cyanide release from plants, thus allowing insects to utilize the normally toxic compounds as a nitrogen source. Polyphenol oxidase conjugates phenolics into dietary

proteins, and decreases digestibility (Felton *et al*., 1992). In addition, phenolic acids induce oxidative stress in herbivorous lepidopteran larvae (Summers & Felton, 1994). Recent studies have shown that lepidopteran larvae counteract this tactic by maintaining reducing conditions in the gut, thus defeating the effects of dietary oxidized phenolics (Barbehen *et al*., 2001).

An interesting example of insect metabolic resistance to a plant chemical defense is the glucosinolate-myrosinase system, also referred to as "the mustard oil bomb", which is induced in cruciferous plants. Upon plant tissue damage, the nontoxic glucosinolates are hydrolyzed by myrosinases into a variety of toxic products. The most common class of hydrolysis products, isothiocyanates (mustard oil), has been shown to be highly toxic to both generalist and specialist insect herbivores. However, the crucifer specialist insect, diamondback moth *Plutella xylostella* disarms the mustard oil bomb by utilizing glucosinolate sulfatase (Ratzka et al., 2002). Myrosinase cannot use the desulfated glucosinolates as substrates, allowing diamondback moth to avoid the formation of toxic glucosinolate breakdown products. Another lepidopteran specialist insect, larvae of cabbage white butterfly *Pieris rapae,* also circumvents the glucosinolate-myrosinase system (Wittstock et al., 2004). In contrast to *Plutella xylostella*, a larval gut protein, designated nitrile-specifier protein redirects the hydrolysis reaction of glucosinolate toward the formation of nitriles instead of isothiocyanates.

Protease inhibitors are an example of plant defensive chemicals that suppress the action of digestive proteases in herbivores and make dietary protein unavailable,

consequently, causing insect starvation. Plant protease inhibitors are developmentally regulated and are induced in response to insect attack (Ryan, 1990). Many transgenic plant experiments and insect-feeding bioassays have showed that plant protease inhibitors cause retardation of insect growth and development, insect starvation, and death (Hilder et al., 1987; Koiwa et al., 1998). Soyacystatin N suppresses insect digestive enzymes and subsequently causes delay of growth and development of cowpea bruchid, *Callosobruchus maculatus* (Fabricius) as well as other coleopterans such as western corn rootworm, *Diabrotica virgifera virgifera* LeConte and Colorado potato beetle, *Leptinotarsa decemlineata* Say (Koiwa et al., 1998; Zhao et al., 1996; Zhu-Salzman et al., 2003).

Studies on insect responses to the dietary incorporation of plant protease inhibitors have shown that insects adapt to protease inhibitors by utilizing several counter-defensive strategies: overproduction of existing digestive proteases (De Leo et al., 1998), differential expression of inhibitor-insensitive proteases (Cloutier et al., 2000; Zhu-Salzman et al., 2003), proteolytic degradation of protease inhibitors by insect protease (Michaud et al., 1995; Zhu-Salzman et al., 2003), and a biphasic expression of protease characterized by an initial up-regulation of all digestive protease specificities, followed by a simultaneous down-regulation of protease inhibitor-sensitive proteases and up-regulation of protease inhibitor-insensitive proteases (Bown et al., 2004).

Finally, mutation in the target of the plant-defense chemical can reduce or eliminate its inhibitory effects. Larvae and adults of Monarch butterfly *Danaus plexippus* are insensitive to dietary cardiac glycosides, whereas other Lepidoptera are

sensitive and intoxicated by ouabain (Holzinger & Wink, 1996). Ouabain inhibits Na^+, K^+-ATPase by binding to its α-subunit. The Monarch has a histidine in the putative ouabain binding site, instead of the asparagine in ouabain-sensitive insects. Therefore, this amino acid substitution in the ouabain binding site of Na^+, K^+-ATPase contributes to ouabain insensitivity in the Monarch.

1.3 Objectives of this work

The human population is expected to reach or even exceed 10 billion by the mid-21st century. Feeding this growing human population demands an increase in agricultural production and the improvement of crop protection. Current approaches to fulfill this demand are highly dependent on chemicals such as fertilizers, insecticides, fungicides, herbicides etc. The wide use of synthetic chemicals raises concerns regarding the rapid build-up of pest resistance and threatens the existing ecological balance. Furthermore, the annual production of synthetic chemicals now costs more than 10 billion dollars, even though estimated global pre- and post-harvest crop losses caused by pests (insects, nematodes, diseases, and weeds) are 30-40% in developed countries and 60-70% in developing countries annually. The global crop losses due to insect pests alone are 6-13% of total crop production (Thomas, 1999). The consumption of synthetic chemicals has started to decrease due to high concerns about their side effects. Therefore, it is necessary to develop new technologies in order to reduce the use of synthetic chemicals while increasing crop yields.

To address these issues, genetic engineers have produced transgenic plants for enhanced resistance to insect pests. In 2003, six transgenic crop plants (canola, corn, cotton, papaya, squash and soybean) produced an additional 2.4 million tons of food and fibers, and increased farm income by 1.9 billion dollars in the USA (Christou et al., 2006). Transgenic plants expressing insecticidal proteins, δ-endotoxins from the bacterium *Bacillus thuringiensis* (*Bt*) were first commercially released in 1996, and they have led to decreased insecticide usage. From 1996 to 2003, transgenic *Bt* crops were grown on more than 67.7 million ha worldwide (James, 2003). Evolution of insect resistance threatens the continued success of *Bt* crops, however. At least seven resistant laboratory strains of three pests (diamondback moth, *Plutella xylostella* [L.]; pink ballworm, *Pectinophora gossypiella* [Saunders]; cotton bollworm, *Helicoverpa armigera* [Hübner]) have developed resistance to *Bt* crops in laboratory and greenhouse tests (Tabashnik et al., 2003). Further, global monitoring data for six major lepidopteran pests (*Helicoverpa zea*, *Helicoverpa amigera*, *Heliothis virescens*, *Ostrinia nubilalis*, *Pectinophora gossypiella* or *Sesamia nonagrioides*) targeted by *Bt* crops showed field-evolved resistance in *Helicoverpa zea*, (Tabashnik et al., 2008).

It is clear that complementary durable and sustainable pest control strategies to insect resistance are necessary. Alternatives are suggested by the plant's own defense mechanisms, by manipulating the expression of their endogenous defense proteins, or introducing an insect control gene derived from other organisms (Christou et al., 2006; Ferry et al., 2006). The use of naturally occurring plant protease inhibitors to target insect digestive enzymes seemed promising, but efforts to achieve host plant resistance

by expressing protease inhibitors in transgenic plants have been largely unsuccessful due to rapid insect adaptation to the transgenic products (Jongsma & Bolter, 1997; De Leo et al., 1998; Cloutier et al., 2000). It is unlikely that the discovery of additional plant defense genes or attempts to improve the insecticidal activities of these proteins will solve the problem. Endogenous mechanisms of plant defense are likely to have only very limited success in insect pest control in agriculture contexts.

Reliance on the expression of a single gene product for pest control is a relatively short-term strategy that parallels the use of exogenously applied chemical pesticides. Therefore pyramiding (stacking) of genes encoding different anti-insect proteins has been developed (Zhao et al., 2003; Amirhusin et al., 2007; Tarver et al., 2007). Also such hybrid/fusion proteins to enhance toxin binding ability to target insects offer an alternative strategy to address potential limitations in conventional transgenic insect pest control (Mehlo et al., 2005). However, there is a report that insect can still develop resistance to transgenic plants expressing pyramided toxins (Zhao et al., 2005).

During the past 30 years, significant advances have been attained in our understanding of plant defenses and the mechanisms that provide protections from insect attacks. By comparison, much less attention has been paid to herbivore counter-defense, the mechanisms that herbivores employ to exploit host plants. For instance, the exact nature of the complex signaling events that are responsible for monitoring and coordinating nutrient uptake and gut proteolytic activities are not well understood, although it is clear that insects are able to express a variety of proteases in response to dietary protease inhibitor. Also many protease genes still have not yet been linked to

their function in terms of sensitivity to various inhibitors, substrate specificity or relative contribution to protein digestion, due to the lack of suitable expression system. Although current studies on the dynamic nature of the insect digestive proteases have offered knowledge on molecular strategies of insect counter-defense to plant defense, it is insufficient for the development of "fool-proof" strategies for insect pest control. Furthermore, the process of insect adaptation appears to induce diverse signaling pathways of counter-defense genes to evade plant defense molecules. Thus, to overcome this current problem in biotechnology-based insect control, we should acquire knowledge of the genetic and molecular bases of insect adaptation to plant defense molecules. Eventually, a deeper understanding of the mechanisms and signaling pathways governing the differential regulation of insect protease genes in response to plant defense will provide crucial knowledge of molecular mechanism of insect adaptation, and lead us to develop durable pest control strategies to insect resistance.

In this dissertation, the overall objective was to understand the molecular mechanisms used by cowpea bruchid to differentially regulate cysteine proteases during adaptation to soybean protease inhibitor.

In a world of worthy candidates, there are three compelling reasons to use cowpea bruchid as a model system to study insect adaptation to plant defense.

First, cowpea bruchid is a representative storage pest of cowpea and other grain legumes for human consumption. Cowpea is an important economic crop that provides a main source of dietary proteins for many poor people worldwide. The demand for cowpeas keeps growing, but cowpea bruchid is a major constraint for high yields.

Cowpea bruchid, therefore, has a long history of exposure to pesticides. Poor farmers in Africa cannot afford to buy pesticides, and chemical pesticides are environmentally unfavorable. Moreover, cowpea bruchid has proven to be readily adaptable to natural plant defensive compounds having developed resistance via target insensitivity and other mechanisms. This feature supports the use of cowpea bruchid as an ideal subject for the development of new strategies of pest control through knowledge of resistance mechanisms.

Second, cowpea bruchid is a species that represents the largest and most diverse of all eukaryotic orders, the Coleoptera. This order includes many beneficial and deleterious species, the latter associated with billions of dollars of agricultural losses annually. However, coleopteran insects have not been studied in many fields compared with different order insects, the Diptera. The genomes from 24 insect species have recently been sequenced, or are in the process of being sequenced (Grimmelikhuijzen et al., 2007). Among these insects are 12 fruit fly species belonging to the genus *Drosophila*, the best known being *Drosophila melanogaster* Meigen (Diptera), which was the first insect with a sequenced genome. Genomes from the malaria mosquito *Anopheles gambiae* Giles, the yellow fever mosquito *Aedes aegypti* L. and the mosquito *Culex pipiens* L. also have been sequenced. In contrast to the Diptera, the red flour beetle *Tribolium castaneum* Herbst is the only species whose genome has been sequenced in the Coleoptera. The biology of coleopteran insects is very little understood, even though they are very agriculturally important insects. This feature

recommends cowpea bruchid as a model system providing a much better understanding of practical matters such as insect resistance.

Third, the Zhu-Salzman's lab has developed important techniques for studying interactions between cowpea bruchid and potential plant defense compounds. An artificial seed system has been developed to perform carefully controlled and reproducible bioassays of the effects of plant allelochemicals on cowpea bruchids (Shade et al., 1986). With respect to growth and development rates and mortality, bruchid larvae develop in artificial cowpea seeds performed as well as those in intact parent seeds. Cowpea bruchids live hidden within seeds and their larval growth and development are thus impossible to be observed from outside seeds. Instead of visual observation of larval development, larval feeding behavior can be monitored using a biomonitor, a device that detects and records the ultrasonic feeding events generated by bruchid larvae hidden inside the artificial seeds and enables a precise determination of insect developmental stages (Shade et al., 1990).

My first objective was to analyze the functional significance of cathepsin L-like cysteine proteases (CmCPs) that are differentially regulated during cowpea bruchid adaptation to dietary scN. CmCPs are major digestive enzymes of cowpea bruchid, and their isoforms are further grouped into CmCPA and CmCPB subfamilies based on sequence similarity (Zhu-Salzman et al., 2003). The mRNAs of the major digestive CmCP enzymes are differentially expressed in response to dietary scN; CmCPB subfamilies are more abundant in scN-adapted bruchid gut than CmCPA subfamily members (Zhu-Salzman et al., 2003). Strong induction of the CmCPB genes suggests

that the CmCPB subfamily might play an important role in bruchids adaptation to dietary scN. Thus, I wanted to gain a comprehensive understanding of the differential transcriptional regulation of *CmCP*s.

My second objective was to define the mechanisms that regulate transcription of cathepsin B-like cysteine protease (CmCatB) during cowpea bruchid adaptation to dietary scN. DNA microarrays previously showed that cowpea bruchid was able to induce diverse counter-defense related genes (Moon et al., 2004). *CmCatB* was the most highly up-regulated gene identified in an scN-regulated EST collection from cowpea bruchid midguts. Northern blot analysis showed that *CmCatB* was highly induced in scN-adapted 4th instar larvae midgut, but undetectable in scN-unadapted larvae. The peak of *CmCatB* expression in 4th instar larvae was consistent with the bruchid adaptation to dietary scN during the 4th instar stage (Zhu-Salzman et al., 2003). Modeling of *CmCatB* on the human cathepsin B ortholog suggests an explanation for overexpression of this gene. The human enzyme contains a structural element termed the “occluding loop”, which blocks off one end of the substrate binding cleft. This domain is involved in the exopeptidase activity of cathepsin B, and blocks the access of substrates and inhibitors (Illy et al., 1997). Zhu-Salzman's lab recently confirmed that CmCatB is insensitive to scN due to the presence of this “occluding loop” through heterologous yeast expression system (Koo et al., 2008). Based on these findings suggesting *CmCatB* is involved in bruchids adaptation to plant inhibitor, I investigated the regulatory mechanisms that control transcription of *CmCatB* during bruchid adaptation to dietary scN.

CHAPTER II

FUNCTIONAL ROLES OF SPECIFIC BRUCHID PROTEASE ISOFORMS IN ADAPTATION TO A SOYBEAN PROTEASE INHIBITOR*

2.1 Introduction

An obvious function of insect digestive proteases is to break down dietary proteins into simple peptides and amino acids, which are then absorbed into the hemolymph. Proteases can be classified into serine, cysteine, aspartate and metallo proteases, and all of these types have been found in insect guts (Terra & Ferreira, 1994). Like many coleopteran insects, the cowpea bruchid beetle *Callosobruchus maculatus*, a pest of stored cowpea *Vigna unguiculata* and other grain legumes, predominantly uses cysteine proteases for food protein degradation (Murdock et al., 1987; Terra & Ferreira, 1994; Zhu-Salzman & Salzman, 2001). The insect completes its entire four-instar larval development inside the seeds, where it causes severe grain damage due to larval feeding. Molecular studies have led to the conclusion that the major digestive enzymes of the cowpea bruchid comprise multiple cathpsin L-like cysteine proteases (CmCPs), which can be further grouped into CmCPA and CmCPB subfamilies based on sequence similarity (Zhu-Salzman et al., 2003). Mammalian cathepsins have been shown to be synthesized as inactive, higher molecular weight proenzymes. The propeptide region

---

occludes the active site and prevents binding of the enzyme to its substrates (Groves et al., 1998). These enzyme precursors can be activated by removal of the propeptide regions either autocatalytically in a low pH environment or through the action of different proteases (Vernet et al., 1991; Rowan et al., 1992; Maubach et al., 1997; Rozman et al., 1999).

Less obvious, regarding the function of digestive enzymes of herbivore insects, are the roles they play in insect adaptation to plant defense proteins, although this situation has begun to change in recent years due to active research in this area. Using naturally occurring plant protease inhibitors to target insect digestive enzymes has received serious consideration as a means of insect pest management. However, efforts to achieve host plant resistance by expressing protease inhibitors in transgenic plants have been largely unsuccessful (Jongsma & Bolter, 1997; De Leo et al., 1998; Cloutier et al., 2000). Available molecular and biochemical studies indicate that modulation of the complement of midgut digestive enzymes contributes substantially to rapid insect adaptation to plant defensive protease inhibitors. Some insects adapt to inhibitors by overproducing existing digestive proteases (De Leo et al., 1998), while others activate different proteases that are not susceptible to a particular inhibitor (Bolter & Jongsma, 1995; Bown et al., 1997; Jongsma & Bolter, 1997; Cloutier et al., 2000; Mazumdar-Leighton & Broadway, 2001). Differential expression of protease isoforms varying in their sensitivity to specific protease inhibitors can also be developmentally programmed (Orr et al., 1994; Liu et al., 2004). Disarming plant protease inhibitors themselves through direct proteolytic fragmentation represents yet another strategy insects use to

counter the effect of the inhibitors (Michaud et al., 1995; Giri et al., 1998; Zhu-Salzman et al., 2003). These adaptive responses have hindered efforts in biotechnology-based insect control.

Soyacystatin N (scN), a soybean cysteine protease inhibitor, significantly delays the within-seed larval development of the cowpea bruchid and even causes insect death (Koiwa et al., 1998). However, the insect is able to initiate several counter-defensive strategies in response to dietary scN and recover normal feeding and growth at the 4th instar larval stage (Zhu-Salzman et al., 2003). Bruchids adapted to scN greatly increase their proteolytic capability through expression of both scN-sensitive and scN-insensitive enzymes. Adapted insects are also capable of hydrolyzing, and thus inactivating scN. mRNAs of the major digestive CmCP enzymes are differentially expressed in response to dietary scN; while transcripts of *CmCPB* subfamily were 116.3-fold more abundant in scN-adapted bruchid guts than in unadapted guts, *CmCPA* subfamily were only increased by 2.5-fold (Zhu-Salzman et al., 2003). Strong induction of *CmCPB1*, representing the *CmCPB* subfamily, in adapted bruchid guts suggests that it may play a role in the insect counter-defense response to evade the inhibitory effect of dietary scN. Yet no further evidence at protein level is available to support the involvement of CmCPB1 in the insect adaptive process. The current study aimed at obtaining a more comprehensive understanding of the differential transcriptional regulation of *CmCP*s. I bacterially expressed several CmCPs that exhibited varying characteristics in proenzyme processing and proteolytic activity. Although all CmCPs were susceptible to scN inhibition, CmCPB1 appeared to be more efficient in autocatalytic conversion from the

proenzyme to the mature form, and it showed higher protease activity compared to CmCPAs in the azocasein assay. Thus CmCPBs are most likely engaged in compensating for the scN inhibition of proteolytic activity in cowpea bruchid guts. The superiority of CmCPBs may not be required in the absence of dietary protease inhibitors, but it is apparently essential when insects are facing inhibitor challenge.

2.2 Results

2.2.1 Expression of recombinant proCmCPs

Cowpea bruchids utilize a multigene family to encode major digestive cathepsin L-like CmCPs, and these are differentially regulated by scN challenge (Zhu-Salzman et al., 2003). Further understanding of the distinct function of CmCPs in response to dietary scN requires characterization at the protein level. Cathepsin L proteins from human and rat as well as other papain-like cysteine proteases have been successfully expressed in *E. coli* (Smith & Gottesman, 1989; Dolinar et al., 1995; Barlic-Maganja et al., 1998; Ogino et al., 1999; Renard et al., 2000). It appears that removal of the signal peptide is critical for protein expression (Vernet et al., 1989), and that the presence of the full-length propeptide region is essential for formation of active secondary and tertiary structures during recombinant protein expression (Tao et al., 1994; Velasco et al., 1994; Santamaria et al., 1998; Ogino et al., 1999). From the total 11 full-length *CmCP* cDNA clones (10 *CmCPA*s and one *CmCPB*), three *CmCPA* cDNAs, i.e. *CmCPA9, A13* and

(A)

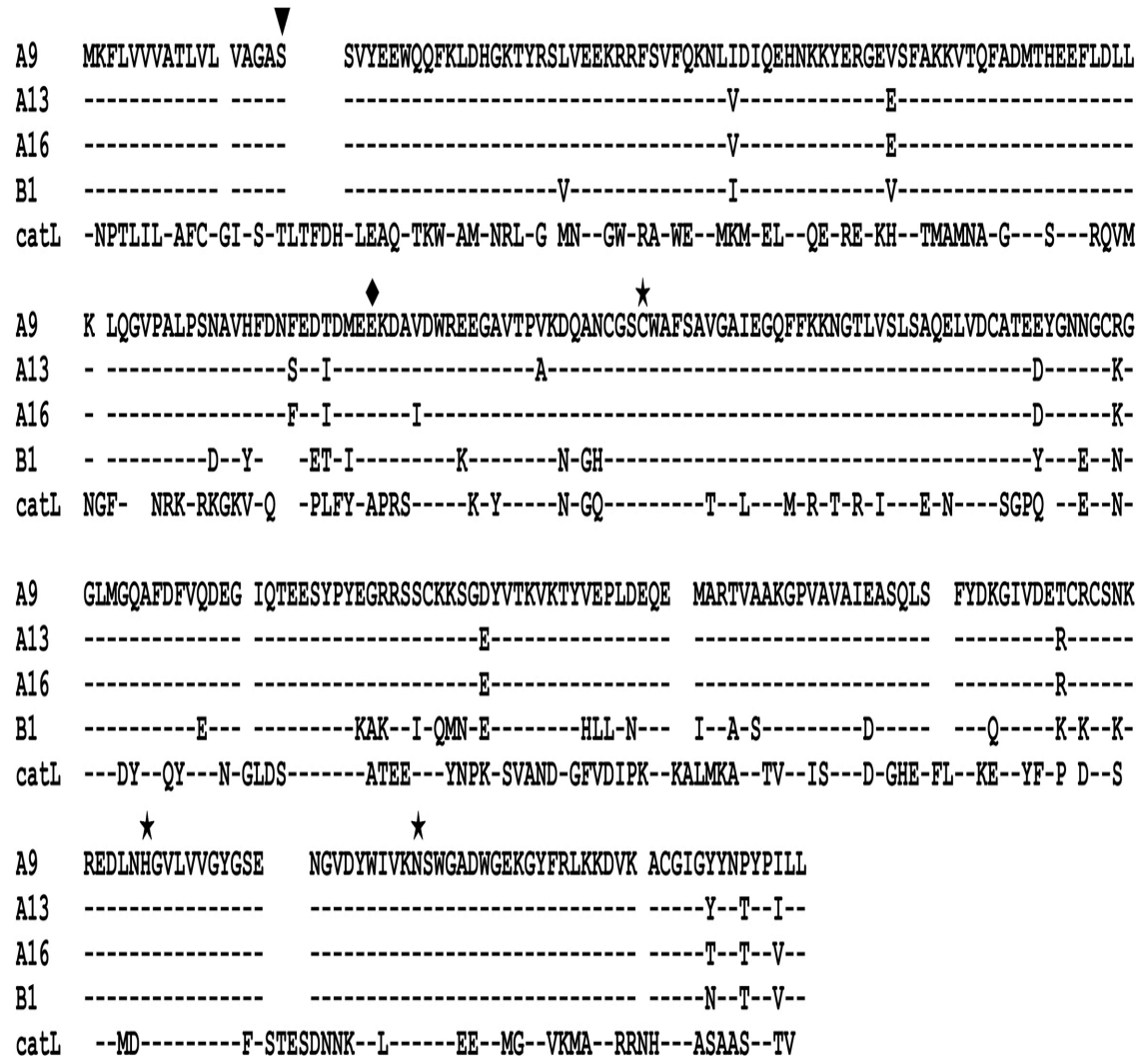

Figure 2.1. Bacterial expression of the major digestive cysteine proteases of cowpea bruchids.
(A) Sequence alignment of CmCPA9 (A9, AF544835), CmCPA13 (A13, AF544837), CmCPA16 (A16, AF544839), CmCPB1 (B1, AF544844) and human cathepsin L (catL, X12451). Dashes represent residues identical to CmCPA9. The stars mark the catalytic triad. The diamond indicates the first amino acid residue of mature CmCPs, and the triangle points to the beginning of the propeptide region.

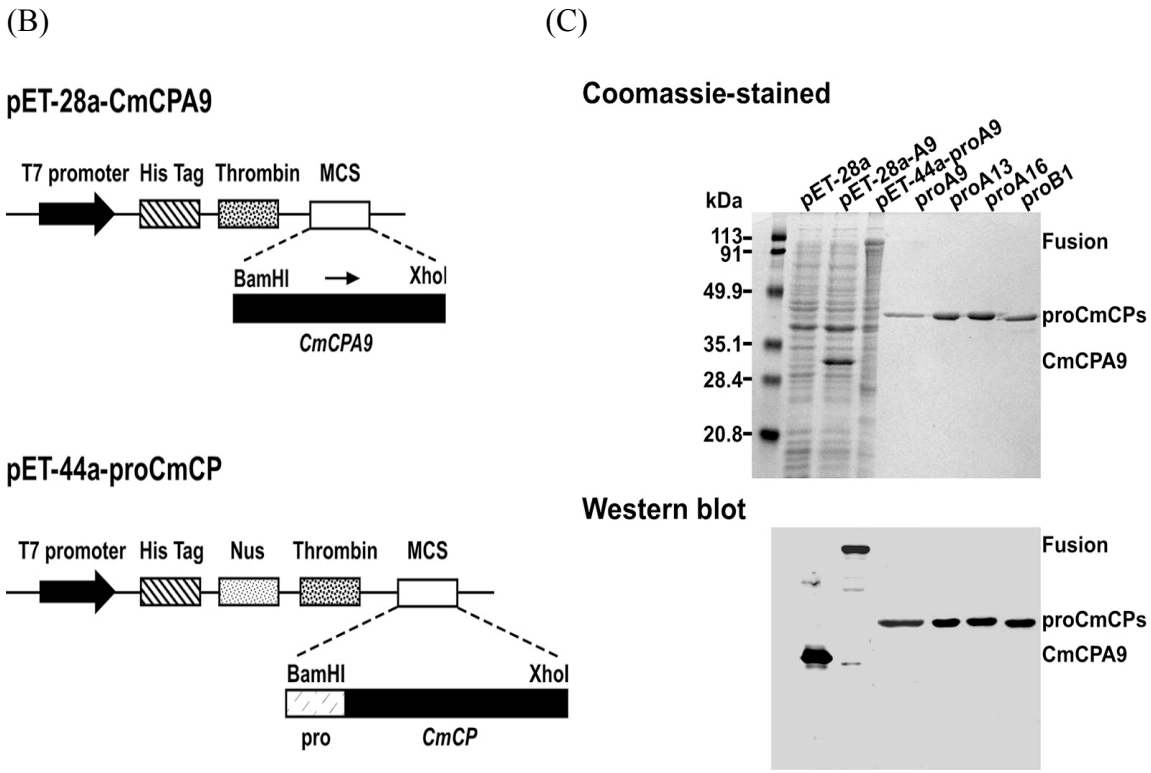

Figure 2.1. Continued. (B) Diagram of constructs for (i) expression of mature CmCPA9 in pET-28a vector for antibody production, and (ii) expression of proCmCP proteins fused with Nus A protein in pET-44a vector. His tag: 6x His for purification of recombinant protein via Ni^{2+} affinity chromatography; Nus: Nus A fused to recombinant proCmCPs; Thrombin: thrombin cutting site to release proCmCPs from Nus A; MCS: multiple cloning site; pro: propeptide region; *CmCP*: mature *CmCP* cDNAs. (C) SDS-PAGE and western blot analyses of recombinant proteins expressed in pET-28a and pET-44a, as well as purified proCmCPs following thrombin digestion. Whole cell extracts and recombinant proteins were resolved on 12.5% SDS-PAGE and stained with Commassie Brilliant Blue R-250, or transferred to nitrocellulose membrane for immunoblotting using anti-CmCPA9 antibody.

A16 and the *CmCPB1* cDNA (Fig. 2.1A) were expressed in *E. coli* as proproteins fused with the Nus A protein and the 6x His tag (Fig. 2.1B). All recombinant CmCP proproteins, released from fusion proteins by thrombin digestion, were soluble and reacted with the polyclonal anti-CmCPA9 antibody (Fig. 2.1C).

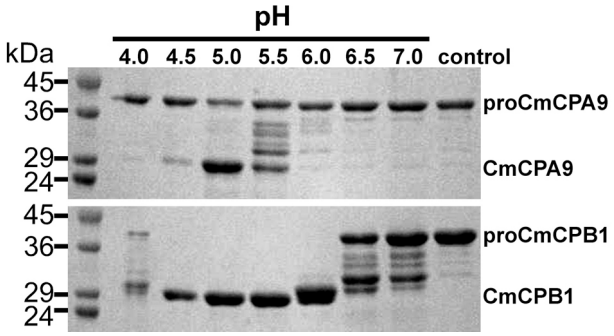

Figure 2.2. proCmCPB1 can autoprocess within a broader pH range than proCmCPA9. Each proCmCP was incubated at 37°C under various pHs for one hour, and were then resolved on 12.5% SDS-PAGE. proCmCPs without incubation were used as control.

2.2.2 Maturation of the CmCPs

To characterize the maturation process of various CmCP isoforms, I investigated the feasibility of conversion from inactive proproteins to enzymatically active mature proteins through autoprocessing as well as through external processing using cowpea bruchid gut extract. Complete autocatalytic conversion occurred between pH 4.5 and pH 6.0 for CmCPB1. CmCPA9, on the other hand, exhibited a much narrower autoprocessing pH range (Fig. 2.2). Thus I chose pH 5.0, the optimal pH of CmCPA9, for proenzyme processing and proteolytic reaction throughout my study. Two new bands were observed as a result of proCmCPA9 autoprocessing: a 28 kD band reacting with the anti-mature CmCPA9 antibody and corresponding to the size of the predicted mature enzyme, and a 15kD band corresponding to the cleaved propeptide but

not reacting with the antibody (Figs. 2.3A, 2.3B). The proCmCPB1 was readily converted to its mature form. Absence of the propeptide band during autoprocessing suggests that it may have been degraded by mature CmCPB1 (Fig. 2.3A). Two slightly different sizes of mature CmCPA9 bands were detectable in the immunoblot (Fig. 2.3B), indicating that more than one processing site may exist in proCmCPA9. Extended autoprocessing revealed a two-step mechanism to convert the proenzyme into a shorter but more stable mature CmCPA9 form (Fig. 2.4).

In contrast to CmCPA9 and CmCPB1, CmCPA13 and A16 were nearly incapable of autoprocessing (Fig. 2.3). However, when incubated with insect gut extract, both were completely converted to mature forms (Fig. 2.3). Conversion of Two new bands were observed as a result of proCmCPA9 autoprocessing: a 28 kD band reacting with the anti-mature CmCP9 antibody and corresponding to the size of the predicted mature enzyme, and a 15 kD band corresponding to the cleaved propeptide but proCmCPA16 to the mature form by gut extract could not be inhibited by pepstatin (an inhibitor of aspartic proteases) and was only partially inhibited by scN (Fig. 2.5A). Combination of the inhibitors, however, prevented the CmCPA16 maturation. To further dissect the mechanism of the intermolecular trans-processing, I examined the effect of mature CmCPA9 or/and CmCPB1 as well as pepsin (an aspartic protease) on proCmCPA16. Both cysteine and aspartic proteases readily converted the proprotein to mature CmCPA16 (Fig. 2.5B). These results indicate that either gut cysteine proteases or aspartic proteases can process CmCPA16.

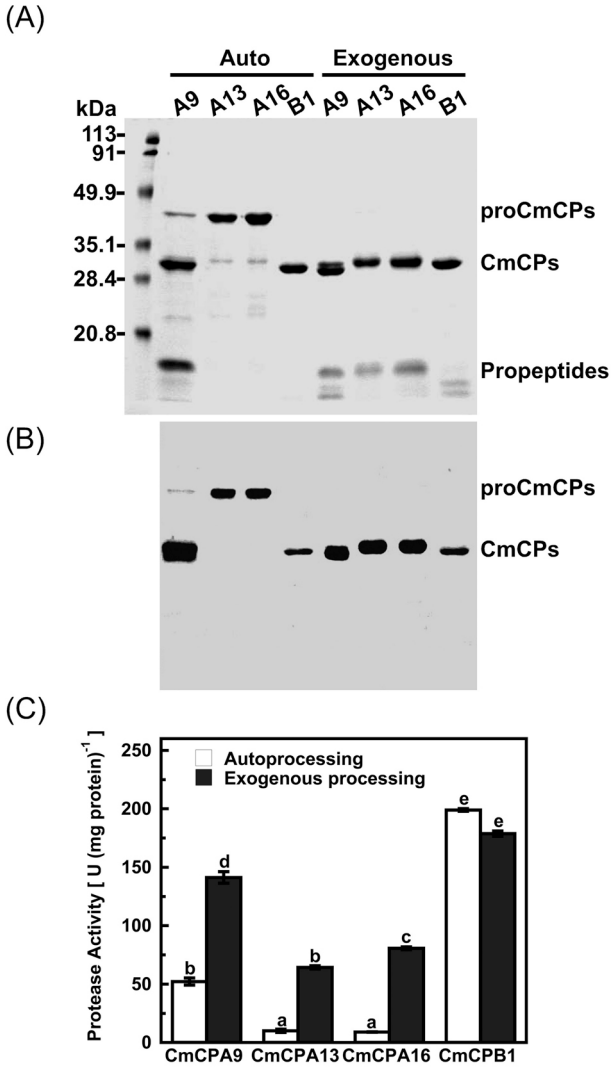

Figure 2.3. CmCPs differ in their enzymatic activities as well as in autoprocessing capability. Conversion from inactive proCmCPs into mature active CmCPs was examined through autocatalytic processing as well as action of exogenous insect gut enzymes. Purified recombinant proCmCPs were incubated under the acidic condition (pH 5.0) or with gut extract at 37°C. The samples were then subjected to (A) SDS-PAGE, (B) immunoblotting, and (C) enzymatic activity assays. Azocasein hydrolysis was used to measure protease activity. Absorbance at 440 nm caused by proteolysis of gut extract alone was subtracted from the total proteolysis of gut extract-processed CmCPs. Experiments were done in triplicate. Means were not significantly different when followed by the same letter (Fisher's protected LSD test, P=0.01). Error bars indicate standard error.

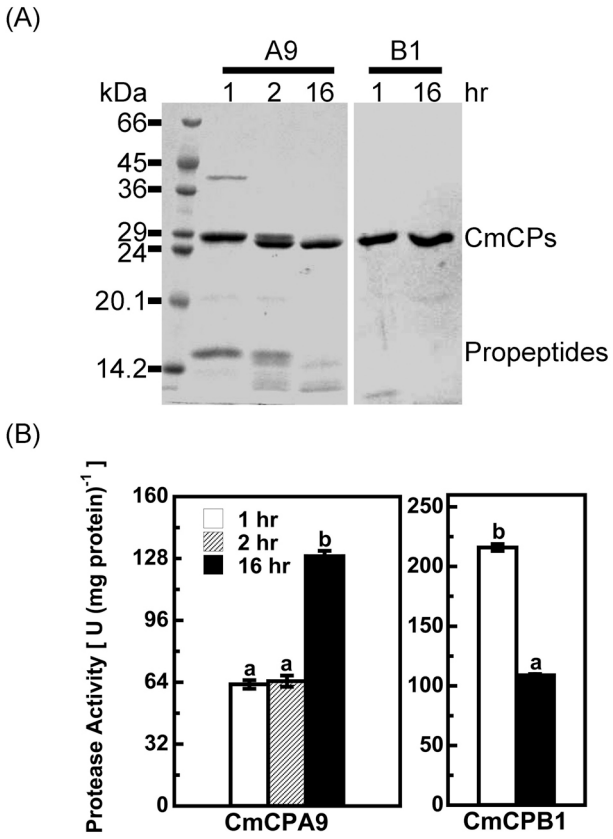

Figure 2.4. Two-step autoprocessing of CmCPA9 precursor.
(A) SDS-PAGE analysis of the purified recombinant proCmCPA9 and proCmCPB1 after incubation at pH 5.0 for different periods of time indicated. (B) Enzymatic activity of autoprocessed CmCPs. Enzymatic activity was measured in triplicate. Means are not significantly different when followed by the same letter (Fisher's protected LSD test, P=0.01). Error bars indicate standard error.

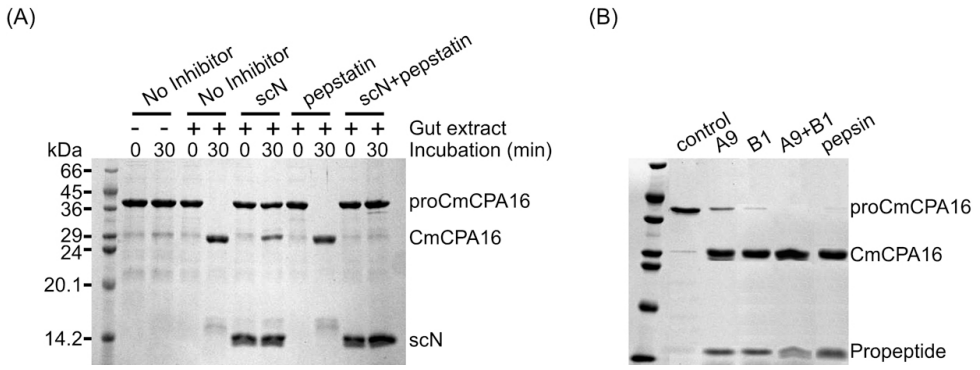

Figure 2.5. CmCPA16 precursor can be processed by insect gut cysteine or aspartic proteases. (A) Cowpea bruchid gut extracts were preincubated with the cysteine protease inhibitor scN, aspartic protease inhibitor pepstatin, or combination of inhibitors prior to addition of recombinant proCmCPA16. Samples were subjected to 15% SDS-PAGE. (B) Recombinant proCmCPA16 was incubated with autoprocessed CmCPA9 or/and CmCPB1 or pepsin, to examine possible trans-processing by these enzymes. Incubation of recombinant proCmCPA16 alone was used as control.

2.2.3. Proteolytic activities of CmCPs and scN inhibition

Azocasein assays revealed varying catalytic activities among mature CmCPs (Fig. 2.3C). Gut enzyme-processed CmCPAs had significantly higher proteolytic efficacy than autoprocessed CmCPAs. For CmCPA13 and A16, trans-processing by exogenous enzymes was essential for elicitation of CmCP enzymatic activity. For CmCPA9, enhanced activity possibly resulted from an expedited maturation process due to availability of diverse enzymes with different substrate specificities. Faster degradation of the CmCPA9 propeptide by gut extract vs. autoprocessing (Fig. 2.3) could be responsible for the elevated activity in CmCPA9, since during the initial period of

autoprocessing, cleaved but undegraded propeptides might still occupy the active site cleft of the newly produced cognate enzymes (Coulombe et al., 1996). It has been observed that human cathepsins B and L were inhibited by cleaved propeptides (Fox et al., 1992; Mach et al., 1994; Carmona et al., 1996; Menard et al., 1998). Physical association of the freed propeptide with the processed CmCPA9 was also detected by gel filtration chromatography after one-hour autoprocessing (data not shown). An alternative explanation for higher enzymatic activity of gut enzyme-processed CmCPA9 is that the slightly shorter mature CmCPA9 form may have higher proteolytic activity compared to the initial product of the two-step autoprocessing. Fig. 2.3 shows that 30 min processing through gut extract produces mainly the shorter CmCPA9 form. Processing at multiple sites of propeptide regions leading to products varying in their catalytic activities has also been observed in human procathepsin L, and lower activity was associated with incomplete N-terminal trimming of the mature cathepsin L (Ishidoh et al., 1994). Extended autoprocessing of proCmCPA9 that converts CmCPA9 to the shorter form indeed improved the enzymatic activity of the mature enzyme (Fig. 2.4), although this could be partly due to the fact that further degradation of the propeptide also occurred simultaneously.

CmCPB1, both autoprocessed and gut extract-processed, showed the greatest azocasein-digesting activity among all CmCPs tested (Fig. 2.3C), suggesting it has the highest intrinsic enzymatic activity, in addition to its more effective maturation process. However, its proteolytic activity decreased after extended autoprocessing (Fig. 2.4).

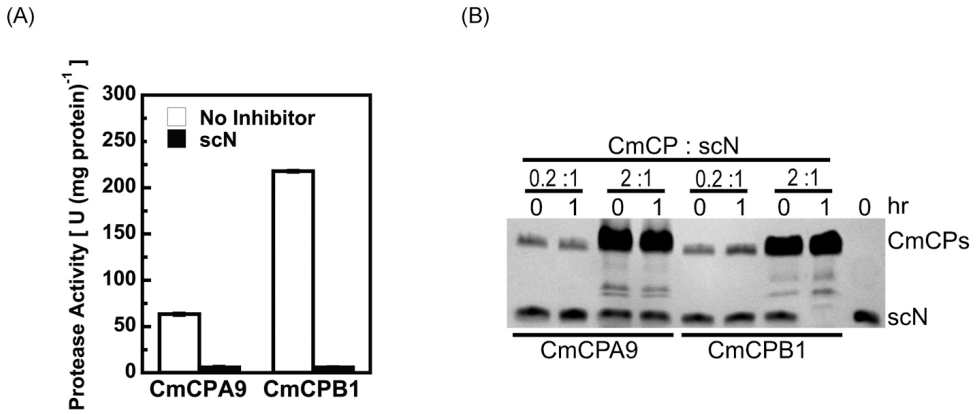

Figure 2.6. CmCPB1 is scN-sensitive but can degrade scN when in excess. (A) The autoprocessed CmCPA9 or CmCPB1 was preincubated with water or scN, and was then evaluated for enzyme activity as described in experimental procedures. Each treatment was done in triplicate. (B) Autoprocessed CmCPA9 (16 hr) or CmCPB1 (1 hr) was incubated with scN at molar ratios of 0.2:1 and 2:1 at 37°C for 0 and 1 hr, respectively. The samples were then subjected to 12% tricine SDS-PAGE. CmCPA9 autoprocessed for 16 hr was used to maximize its enzymatic activity.

Since no obvious protein degradation was detected (Fig. 2.4A), this decrease in CmCPB1 activity most likely was due to localized inactivation at the catalytic site rather than a collapse of the entire protein structure. Total gut activity assays indicated that a small portion of the cysteine protease activity was due to scN-insensitive proteolysis (Zhu-Salzman et al., 2003). To determine whether CmCPs are responsible for this scN-insensitive activity, I evaluated the enzymatic activity of autoprocessed CmCPA9 and CmCPB1 in the presence and absence of scN. Both enzymes were inhibited by scN (Fig. 2.6A), indicating that they are scN-sensitive.

Interestingly, an incubation containing a 2:1 ratio of mature CmCPB1 to scN led to the degradation of scN. However, this scN breakdown was not observed in the case of CmCPA9 (Fig. 2.6B).

2.3. Discussion

When challenged by dietary protease inhibitors, herbivore insects are able to modulate their digestive enzyme complements, qualitatively and quantitatively, to overcome the growth inhibition (Bolter & Jongsma, 1995; Michaud et al., 1995; Bown et al., 1997; Jongsma & Bolter, 1997; Giri et al., 1998; Cloutier et al., 2000; Mazumdar-Leighton & Broadway, 2001; Zhu-Salzman et al., 2003). Cowpea bruchids respond to scN by overproducing scN-sensitive proteases, meanwhile activating production of scN-insensitive and scN-degrading proteases. The major digestive cathepsin L-like CmCP enzymes are encoded by a multigene family in this insect. Differential regulation of *CmCP* transcripts suggests that these isoforms play distinct roles in the process of insect adaptation to this inhibitor (Zhu-Salzman et al., 2003). Selective induction of *CmCPB1* in scN-adapted insect guts underscores its unique position in the adaptive process. The current study presents the novel finding that superior functional aspects of CmCPB1 may underlie its increased expression under scN challenge. Although our initial assumption that CmCPBs might be scN-insensitive proved to be false, CmCPBs differ from the CmCPA subfamily in maturation mechanisms and efficacy, in intrinsic proteolytic activity, as well as in scN-degrading properties.

CmCPB1 appears to be the most self-sufficient enzyme among CmCPs tested. When bruchids are challenged by dietary scN and the majority of their gut enzymes are inhibited, rapid autocatalytic maturation and removal of its propeptide likely puts the newly formed mature CmCPB1 protease immediately to work. Dependence of CmCPA13 and CmCPA16 on exogenous enzymatic action, on the other hand, could impair their enzymatic activation when insects are facing scN in the diet and other proteases are inactive. Prolonged, two-step autoprocessing of CmCPA9 may delay its enzymatic activity. In addition, CmCPB1 has higher intrinsic proteolytic activity and possesses scN-degrading capability when in excess. Presumably, these characteristics render the CmCPB subfamily better isoforms to overexpress in the presence of scN to compensate for inhibited digestive capability. Although the CmCPB1 activity can be inhibited by scN, ramping up *CmCPB1* production in scN-adapted insects (Zhu-Salzman et al., 2003) could be the strategy cowpea bruchids use to circumvent this scN susceptibility. Continuous expression of scN-sensitive digestive enzymes could provide insects with more proteases to outnumber the inhibitors.

Proteolytic fragmentation of scN represents another strategy insects use to cope with scN inhibition, in addition to quantitative compensation. Degradation of scN can potentially free scN-inhibited proteases to resume their digestive capability. Cleavage of cystatin superfamily inhibitors by papain or other cathepsins has been previously reported (Lenarcic et al., 1991; Machleidt et al., 1995; Zhu-Salzman et al., 2003). Although no evidence indicates direct involvement of CmCPB1 in scN degradation *in vivo*, my observation could be physiologically relevant taking into the consideration that

over a 100-fold increase of *CmCPB1* transcript occurred in scN-adapted insects (Zhu-Salzman et al., 2003). Since compartmentalization of digestive enzymes is a common phenomenon in insect midgut (Terra & Ferreira, 1994; Elpidina et al., 2001), it is possible that high CmCPB1 to scN ratios may occur in localized gut regions, where CmCPB1 could effectively degrade the inhibitor to which it is sensitive. In fact, previous observation of scN truncation in scN-adapted insect gut extracts was thought to initiate at the anterior end of the cowpea bruchid gut (Zhu-Salzman et al., 2003).

Although the benefits of very high enzymatic activity of CmCPB1 under inhibitor stress are apparent, CmCPAs are most likely the predominant digestive enzymes cowpea bruchids employ in the absence of scN challenge, because *CmCPA* transcripts are 89-fold more abundant than those of *CmCPB*s in the unadapted insect group (data not shown). Various CmCP isoforms in the gut surroundings under normal physiological conditions should readily convert proCmCPA9 into its most efficient mature form, thus the protein's slow autocatalytic two-step processing may not reflect the scenario *in vivo*. With the same rationale, the lack of autocatalytic activity in CmCPA13 and CmCPA16 should not hinder their maturation when insects are not challenged. Differing activities among CmCPAs suggests that they have different substrate specificities despite high sequence similarity (Fig. 2.1A). Cooperation among CmCP isoforms in the insect guts presumably offers the insect much more efficient digestion capability than any individual isoform functioning alone.

The CmCPBs that have highest enzymatic activity may not be suitable enzymes in the absence of dietary scN. The capability of autoprocessing under a broad pH range,

in combination with its high proteolytic activity, could potentially lead to collateral damage of the insect digestive tract itself. Regulated pH values can promote as well as limit proteolytic activities at different regions of insect guts (Terra & Ferreira, 1994; Elpidina et al., 2001). It could be harder to control CmCPBs by a slight pH shift in insect guts due to its wide pH spectrum of activity. In comparison, a narrow activation pH range decreases the risk of self-destruction caused by CmCPA9. Under scN challenge, however, when the majority of normal digestive capability is paralyzed, expression of "emergency" enzymes such as CmCPBs is apparently beneficial, even critical, to the insect facing amino acid shortage.

The insect gut is not only the major digestive organ, but is a frontline of defense against a broad spectrum of dietary toxins and anti-nutritional factors. Functional characterization of recombinant CmCPs, in conjunction with analyses of transcriptional regulation, has further revealed the insect's capability of adjusting its gene expression profile and fine-tuning its digestive enzymatic activity to cope with dietary challenges.

2.4. Experimental procedures

2.4.1. Expression and purification of proCmCPs

The *CmCP* cDNA fragments encoding selected proCmCPs were PCR amplified (94°C for 30 sec, 54°C for 1 min, 68°C for 1 min for 35 cycles) using combinations of the following oligonucleotide primers: (1) 5′ -TTAATGGATCCTCTTCGGTCTACGA

AGAGTGGC- 3′; (2) 5′ -GTACGCTCGAGTTACAGAAGAATGGGGTAAGTG- 3′; and (3) 5′ -GTTGATCTCGAGTTACAGAAGAACGGGGATAGTGTT- 3′. Restriction sites BamHI and XhoI (underlined) were incorporated into the primers for directional subcloning. Primers 1 and 2 were used for amplification of *CmCPA9* (GENBANK accession number AF544835) and *CmCPA13* (AF544837), and primers 1 and 3 for amplification of *CmCPA16* (AF544839) and *CmCPB1* (AF544844). After restriction digestion, the fragments were ligated to a pET-44a(+) expression vector (Novagen) digested by the same enzymes. The constructs were then transferred to the *E. coli* strain Rosetta – gami (DE3) (Novagen), and correct DNA sequences were confirmed.

Bacterial strain Rosetta – gami (DE3) transformed with *proCmCP* constructs were grown at 37°C until OD_{600} reached between 0.4 and 1.0. Production of recombinant proCmCP proteins fused with Nus A was then induced by addition of isopropyl-β-D-thiogalactopyranoside (to 1.0 mM). Cells were grown overnight at room temperature, harvested and disrupted by sonication (Model 250 Sonifier, Branson). Recombinant proteins were purified using a Ni^{2+}-chelate affinity column according to the manufacturer's instructions (Amersham Pharmacia Biotech). The column was then equilibrated with thrombin cleavage buffer (20 mM Tris-HCl, 150 mM NaCl, 2.5 mM $CaCl_2$ and 0.1 mM DTT, pH 8.4). α-Thrombin (Sigma) was added to a final concentration of 0.5 units/mg fusion proteins. Following a 20 hr digestion at 4°C, proCmCPs were eluted from the column (The N-terminal Nus A protein remained

bound). The purified proteins were ammonium sulfate precipitated (Englard & Seifter, 1990) and analyzed on SDS-PAGE, with whole cell extracts for comparison.

2.4.2. Production of anti-recombinant CmCPA9 antibody and immnunoblot analysis

A cDNA encoding the mature protein region of CmCPA9 was amplified by PCR. The sense primer was 5′-AATATGGATCCAAGGACGCGGTCGACTGGAGAG-3′ and the antisense primer was 5′-CTGGCGCTCGAGTTACAGAAGAATGGGGTAAG TGTT-3′. The 645 bp PCR fragment was subcloned into the pET- 28a(+) expression vector (Novagen) at the BamHI and XhoI sites, and subjected to sequencing analysis to confirm the reading frame. *E. coli* strain BL21 (DE3) (Novagen) harboring the construct was grown and cells were disrupted by sonication as above. After centrifugation at 15,000 rpm at 4°C for 25 min, the pellet containing recombinant protein was resolved on 12.5% SDS-PAGE. Rabbit anti-recombinant CmCPA9 was produced by three intramuscular injections of the emulsified CmCPA9 (Corixa Ribi Adjuvant System) during the four-week interval, and used as primary polyclonal antibody. Bleeding was carried out four weeks after the final immunization. The antiserum was collected by centrifugation at 2,500 rpm, 4°C for 5 min and purified by adsorption with pET- 28a(+) cell lysate following the methods of Sambrook et al. (1989). The antiserum was then used as the primary polyclonal antibody.

For immunoblot analysis, protein samples were first resolved on 12.5% SDS-PAGE, then transferred to a nitrocellulose membrane. The above polyclonal rabbit anti-recombinant CmCPA9 was used as the primary antibody. Goat anti-rabbit IgG (H+L) coupled with horseradish peroxidase conjugate were used as secondary antibody (Kirkegaard Perry Laboratories). Immunosignals were visualized using Western Lightning Chemiluminescence Reagent (PerkinElmer Life Sciences, Inc.).

2.4.3. Obtaining cowpea bruchid gut extract

Guts from cowpea bruchid fourth instars were collected as described by Kitch & Murdock (1986). Five guts were homogenized in 30 μl buffer (100 mM sodium acetate, 1.0 mM EDTA, pH 5.0) and centrifuged at 10,000 g for 10 min. The supernatant was collected for various proteolytic activity tests.

2.4.4. *In vitro* processing of recombinant proCmCPs

The autoprocessing of proCmCPs (5 μg) was carried out according to the method of Nomura & Fujisawa (1997). The purified recombinant proCmCPs were incubated at 37°C in 100 mM sodium acetate, 2.0 mM EDTA and 2.5 mM DTT, pH 5.0 for 1, 2 or 16 hr, respectively. The buffers used for the pH profiling were 100 mM sodium acetate (ranging from pH 4.0 to 6.0) and 100 mM sodium phosphate (for pH 6.5 and 7.0). The exogenous processing of proCmCPs (5 μg) was performed by incubation with gut

extracts (1/60 gut equivalent per reaction) for 30 min, or with autoprocessed mature CmCPA9, CmCPB1 or pepsin (1 μg, respectively) for 1 hr at 37°C under the same buffer condition as autocatalytic processing. The samples were then analyzed on 12.5% SDS-PAGE and by immunoblotting as previously described.

2.4.5. Comparison of protease activity among mature CmCPs

Proteolytic activity of mature CmCPs was determined following Michaud et al. (1994). The processed CmCPs were incubated with 60 μl of 2% (w/v) azocasein diluted in assay buffer (100 mM sodium acetate, 5.0 mM L-cysteine, 0.1% Triton X-100, pH 5.0) for 4 hr at 37°C. After proteolysis, 300 μl of 10% trichloroacetic acid was added to the mixture and the residual azocasein removed by centrifugation at 13,000 g for 5 min. Supernatant (350 μl) was added to 200 μl of 50% ethanol, and the absorbance at 440 nm of this mixture was measured using a Beckman DU 64 spectrophotometer. Absorbance of the sample without proteolysis was used to zero the machine. Experiments were performed in triplicate. Proteolytic activity of CmCPs was plotted using KALEIDA-GRAPH (Abelbeck Software). One unit of protease activity was defined as the amount of mature CmCPs required to produce an absorbance change of 0.01 per hr in 1-cm cuvette at 37°C. A one-way ANOVA test was used to analyze the proteolytic activity data, and Fisher's protected LSD test (P=0.01) was used for mean separation (SPSS 11.1).

2.4.6. Inhibition of proCmCP processing and mature CmCP protease activities by scN

Gut extracts (1/60 gut per reaction) were preincubated with the cysteine protease inhibitor scN (5 μg), an aspartic protease inhibitor pepstatin (0.5 μg), or a combination of scN and pepstatin (5 μg + 0.5 μg), respectively at room temperature for 20 min, then mixed with proCmCPA16 (5 μg) in 100 mM sodium acetate, 2.0 mM EDTA, 2.5 mM DTT, pH 5.0 for 0 or 30 min at 37°C. The mixtures were then resolved on 15% SDS-PAGE.

The autoprocessed CmCPA9 or CmCPB1 (from 5 μg proCmCPs) were preincubated with scN (10 μg) at room temperature for 20 min. The protease activity of mature CmCPs, with or without preincubation with scN, was then measured and plotted as described. Two molar ratios of CmCP vs. scN (0.2:1 and 2:1) were examined for possible hydrolysis of scN by CmCPA9 or CmCPB1. Reactions were subjected to 12% tricine SDS-PAGE (Schagger & von Jagow, 1987).

CHAPTER III

COWPEA BRUCHID *CALLOSOBRUCHUS MACULATUS* CONTERACTS DIETARY PROTEASE INHIBITORS BY MODULATING PROPEPTIDES OF MAJOR DIGESTIVE ENZYMES*

3.1. Introduction

Insect pests cause US $45 billion in agricultural losses worldwide each year. One environmentally friendly response to counter this is to employ endogenous plant defense genes such as protease inhibitors aimed at inhibiting the insect digestive system. The soybean cysteine protease inhibitor N (scN) suppresses growth and development of cowpea bruchid (*Callosobruchus maculatus*) as well as other coleopterans such as western corn rootworm (*Diabrotica virgifera virgifera*) and Colorado potato beetle (*Leptinotarsa decemlineata*). The defensive function of protease inhibitors is attributed to their ability to inhibit insect digestive enzymes (Zhao et al., 1996; Koiwa et al., 1998). However, adaptive changes in insect alimentary tracts have been observed in many insects (Bown et al., 1997; Jongsma and Bolter, 1997; De Leo et al., 1998; Cloutier et al., 2000; Mazumdar-Leighton and Broadway, 2001; Zhu-Salzman et al., 2003; Brunelle et al., 2004), a phenomenon which becomes an obstacle in biotechnology-based pest management.

---

* Reprinted with permission from "Cowpea bruchid *Callosobruchus maculatus* counteracts dietary protease inhibitors by modulating propeptides of major digestive enzymes" by J.-E. Ahn, M. R. Lovingshimer, R.A. Salzman, J. K. Presnail, A. L. Lu, H. Koiwa and K. Zhu-Salzman 2007. *Insect Molecular Biology* 16(3): 295-304. Copyright [2007] by Blackwell Publishing.

Insect digestive enzymes can be classified into serine, cysteine, aspartate and metallo- proteases (Terra & Ferreira, 1994). The cowpea bruchid uses cathepsin L-like cysteine proteases of the papain family, named CmCPs, for food protein degradation (Murdock et al., 1987; Terra & Ferreira, 1994; Zhu-Salzman & Salzman, 2001). Cysteine proteases including CmCP are synthesized as inactive proproteins, e.g. proCmCP. These latent precursors have to undergo either inter- or intramolecular processing to be activated (Menard et al., 1998; Quraishi & Storer, 2001; Ahn et al; 2004). The N-terminal propeptides of cysteine proteases bind the substrate-binding cleft and block the catalytic Cys25 site rending it inaccessible. They also function in protein folding and intracellular sorting (Coulombe et al., 1996; Ogino et al., 1999). Studies of mammalian members of the papain family enzymes indicated that all proregions are potent inhibitors of their own mature enzymes, but cross-inhibition among cognate cathepsins varied. While cathepsin S propeptide inhibits cathepsin L with similar potency as cathepsin S, the reverse is not true (Carmona et al., 1996; Maubach et al., 1997). Also propeptides of cathepsins K, L and S don't significantly inhibit the activity of cathepsin B (Carmona et al., 1996; Billington et al., 2000). Thus inhibition by propeptides among cathepsins appears specific.

At least 30 different CmCP isoforms have been identified in the cowpea bruchid midgut and were grouped into CmCP subfamilies A and B (Zhu-Salzman et al., 2003). In response to dietary scN, cowpea bruchids dramatically remodeled their CmCP complement in the midgut. Induction of CmCPs upon adaptation to scN led to higher total gut protease activity of the adapted cohort than that of the unadapted group.

Interestingly, this insect selectively up-regulated CmCPs in subfamily B by over 100 fold, but only induced those in subfamily A by 2.5 fold (Zhu-Salzman et al., 2003). Comparison between B1 and several subfamily A members including A16 revealed that (i) CmCPB1 had higher proteolytic activity than CmCPAs; (ii) CmCPB1 was able to degrade scN while CmCPAs were not; and (iii) CmCPB1 proprotein exhibited more efficient autocatalytic conversion from the latent proenzyme to its active mature protease form than occurred in CmCPAs. A16, among other subfamily A members, is inert in automaturation (Ahn et al., 2004). Differential regulation of CmCPs is thought to enhance the efficacy of the insects in coping with inhibition by scN.

Identification of inhibitor-degrading isoforms of digestive enzymes has been reported. Human mesotrypsin PRSS3, one of the three digestive trypsin isoforms secreted from pancreas, is resistant to inhibition by Kunitz inhibitors. It is specialized for degradation of trypsin inhibitors, thus facilitating the digestion of foods rich in natural trypsin inhibitors. Molecular and biochemical studies indicated that a single amino acid alteration at Arg^{198} in place of highly conserved Gly is responsible for trypsin inhibitor-fragmentation in mesotrypsin (Nyaruhucha et al., 1997; Szmola et al, 2003). In cowpea bruchids, while it is apparently beneficial to produce CmCP subfamily B in the presence of scN, it is unclear whether such a single controlling amino acid exists. Also, X-ray crystallography has provided very detailed information on interactions of propeptides with their mature enzymes in several papain-like cysteine proteases (Musil et al., 1991; Coulombe et al., 1996; Turk et al., 1997; Groves et al., and

1998), yet little attention has been given to individual propeptides of multigene families, the different effects they may have on enzymatic activity.

In this study, I further dissected the molecular mechanisms behind the differential CmCP enzymatic activities. Apparently, autoprocessing activation and propeptide stability are among strategies cowpea bruchids use to regulate individual enzyme production and adjust overall proteolytic capacity to compensate for dietary challenge.

3.2. Results

3.2.1. Designing and expression of chimeric CmCPs and nomenclature

Despite rather high sequence similarity (Fig. 3.1A), a functional analysis of various CmCP isoforms showed differential autoprocessing, scN-degrading and proteolytic activities among them. A16 and B1 represent two extreme cases from two CmCP subfamilies; while B1 is most efficient in autoprocessing and has the highest enzymatic activity, A16 depends entirely on intermolecular processing to be activated and it has much lower enzymatic activity compared to B1. In addition, B1 is able to fragment scN when it outnumbers the inhibitor. This capacity is absent in all subfamily A members examined (Ahn et al., 2004). To further dissect the molecular basis for the differential activities, I designed six chimeric constructs by domain swapping between A16 and B1 (Fig. 3.1B). By investigating the autoprocessing, proteolysis and scN-degradation

(A)

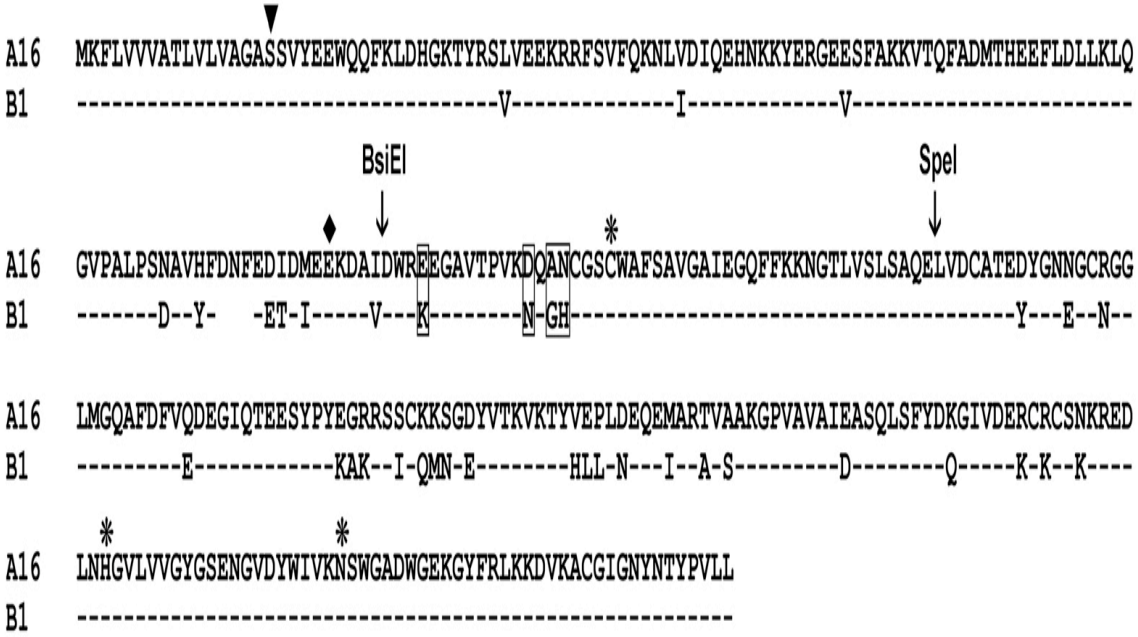

Figure 3.1. Design and production of recombinant CmCP chimeras. (A) Sequence alignment of A16 (AF544839) and B1 (AF544844). Dashes represent identical residues. The catalytic triad is marked by asterisks. The diamond indicates the first amino acid residue of mature CmCPs, and the triangle points to the beginning of the propeptide region. *Bsi*EI, *Spe*I: restriction sites.

activity of these chimeras, we anticipated determination of whether a specific segment(s) is responsible for the differential functionality of the digestive enzymes in cowpea bruchids.

I designated the chimeric proproteins by three capital letters indicating the source of the fragment regions. For example, in pAmBA the pA indicates the propeptide region of A16 and mBA denotes the mature N-terminal region from B1 and the C-terminal region from A16. Recombinant propeptides alone are shown as pA16 (propeptide of

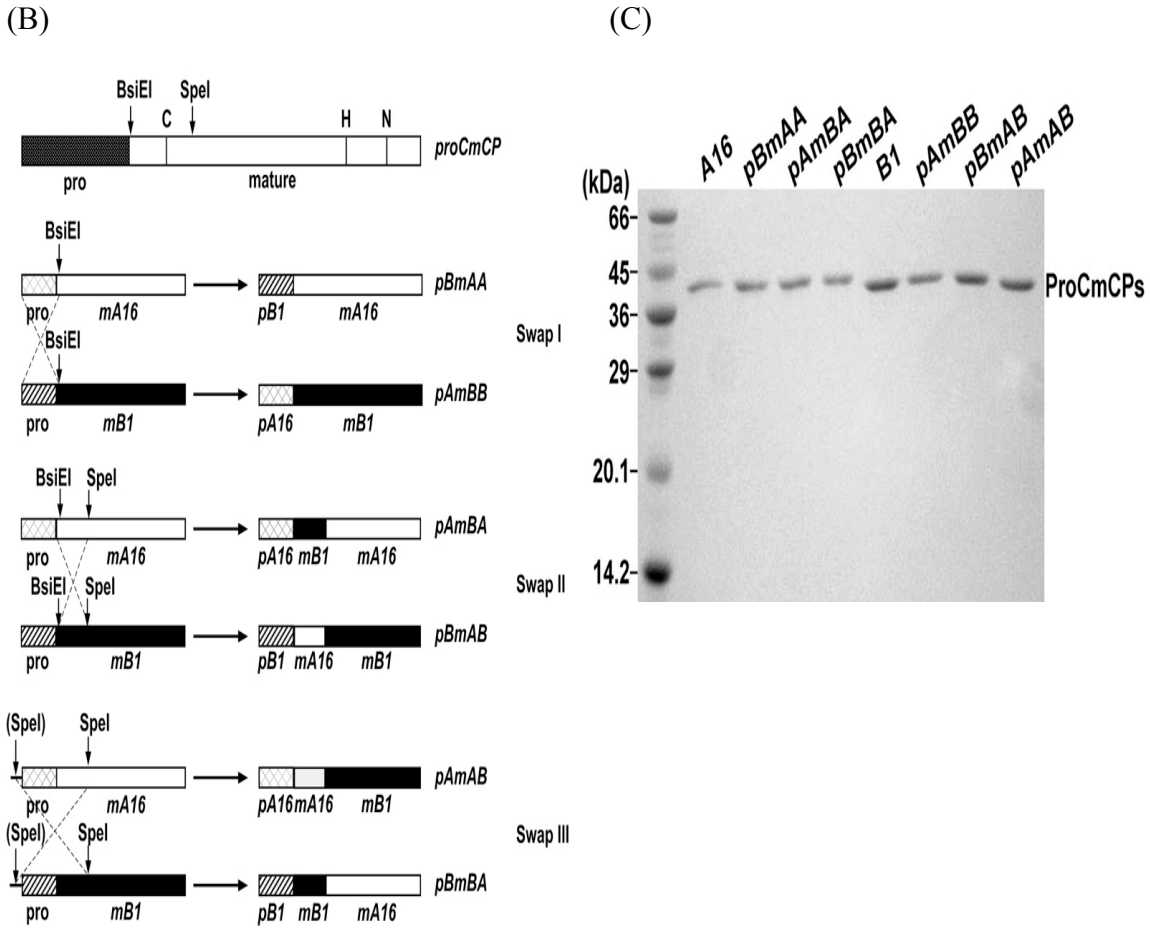

Figure 3.1. Continued. (B) Diagram of gene swapping between A16 and B1. C, H, N: the catalytic triad. A restriction digestion-based subcloning method was used to generate chimeric constructs. *Bsi*E1 allows exchange of the proregions between the two isoforms (swap I), and *Spe*I splits the CmCP between amino acid residues 52 and 53 of the mature proteins (swap III). *Spe*I in parentheses is located in the expression vector. Swap II permitted exchange of the N-terminal mature CmCP region via *Bsi*E1 and *Spe*I sites. p: propeptide; m: mature enzyme; A and B indicate the source of protein fragments in chimeras; A, from A16, B, from B1. (C) SDS-PAGE analysis of six chimeric and two parental proteins. The bacterial pET44 system was used for recombinant protein expression as previously described (Ahn et al., 2004). Proteins were resolved on 12.5% SDS-PAGE and stained with Commassie Brilliant Blue R-250.

A16) and pB1 (propeptide of B1). Swap I produced pBmAA and pAmBB, i.e. reciprocally switched propeptides. Swap II exchanged the N-terminal mature CmCP region (pAmBA and pBmAB), ultimately resulting in four amino acid changes between

the two CmCPs, upstream of and proximal to the key Cys25 residue. Swap III exchanged the C-terminus (pBmBA or pAmAB). The six chimeras and their parental proteins were all expressed as soluble proproteins after excision of the tag proteins from the expression vectors (Fig. 3.1C).

3.2.2. Propeptide swapping does not qualitatively alter the autoprocessing

Swap I, resulting in chimeras pBmAA and pAmBB, allowed us to determine whether the propeptide could impact the maturation process. Replacing the propeptide region of A16 with that of B1 did not initiate autoprocessing, as pBmAA remained in its proprotein form (Fig. 3.2A). Likewise, imposing the propeptide of A16 did not diminish maturation of the pAmBB, although its processed form showed substantially decreased protease activity in the azocasein assay compared to B1 (Fig. 3.2B). Decreased activity could be explained by the residual inhibitory activity of the pA16 propeptide (see below). Thus the propeptide region alone, while it could quantitatively affect enzymatic activity of the mature protein, was unable to reverse the maturation ability of the CmCP.

3.2.3. Single point mutations abolish autoprocessing inhibition in A16

Notably, chimera pAmBA (from swap II) differed by only four amino acid residues from A16, yet this was sufficient to prime removal of the propeptide from the autoprocessing-null parental protein (Fig. 3.2A). To determine whether this qualitative change was

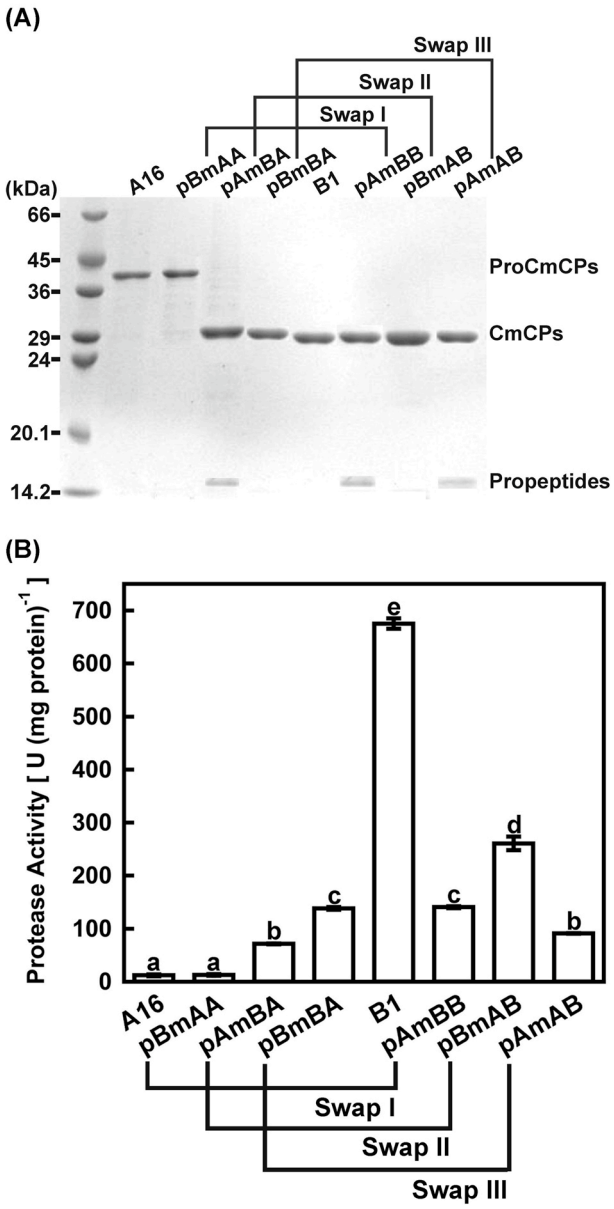

Figure 3.2. Differential autoprocessing and enzymatic activities of chimeras. (A) Proproteins (5 μg each) were incubated at 37°C in 100 mM sodium acetate, 2.0 mM EDTA and 2.5 mM DTT, pH 5.0 for 1 hr, and then resolved on 12.5% SDS-PAGE. (B) Azocasein assays. Absorbance at 440 nm resulted from azocasein hydrolysis by autoprocessed recombinant proteins was measured. Experiments were done in triplicate. Means are not significantly different when followed by the same letter as determined by the Fisher's protected LSD test (P=0.01). Error bars indicate standard errors.

triggered by a sole amino acid residue or as a result of a coordinated effect of all four residues, site-directed mutagenesis was performed to change A16 residues to those of B1, individually or in pairs. Single amino acid alterations proximal to the catalytic Cys25 residue (D18N, A20G and N21H) all initiated propeptide autoproteolysis, but were less effective than double (E9K&N21H, A20G&N21H) or tetra- (pAmBA) mutations (Fig. 3.3). E9K alone did not influence this process but it enhanced the effect of N21H when both residues were simultaneously changed. Thus it appears that conservation of all four residues in A16 is necessary to suppress autoprocessing. Structural changes introduced by single mutations, albeit small, could prompt autoproteolysis and result in enzymatic activity. This effect was increased as more residues were altered at the same time.

Swap III (replacing the C-terminal portion of A16 with the corresponding B1 section) also activated autoprocessing (Fig. 3.2A). Thus, beside those that were systematically studied as shown above, other residues further downstream could also trigger this response in CmCPB1. This indicates that multiple activation determinants are dispersed in the CmCPB1 mature region, and any changes that can potentially weaken the tight fit of the propeptide to the catalytic site result in increased autoprocessing. This offers more flexibility to the insects to regulate CmCP activity at the post-translational level. Consistently, diverse maturation patterns were observed in cowpea bruchids, e.g. the two-step event in CmCPA9 (Ahn et al., 2004).

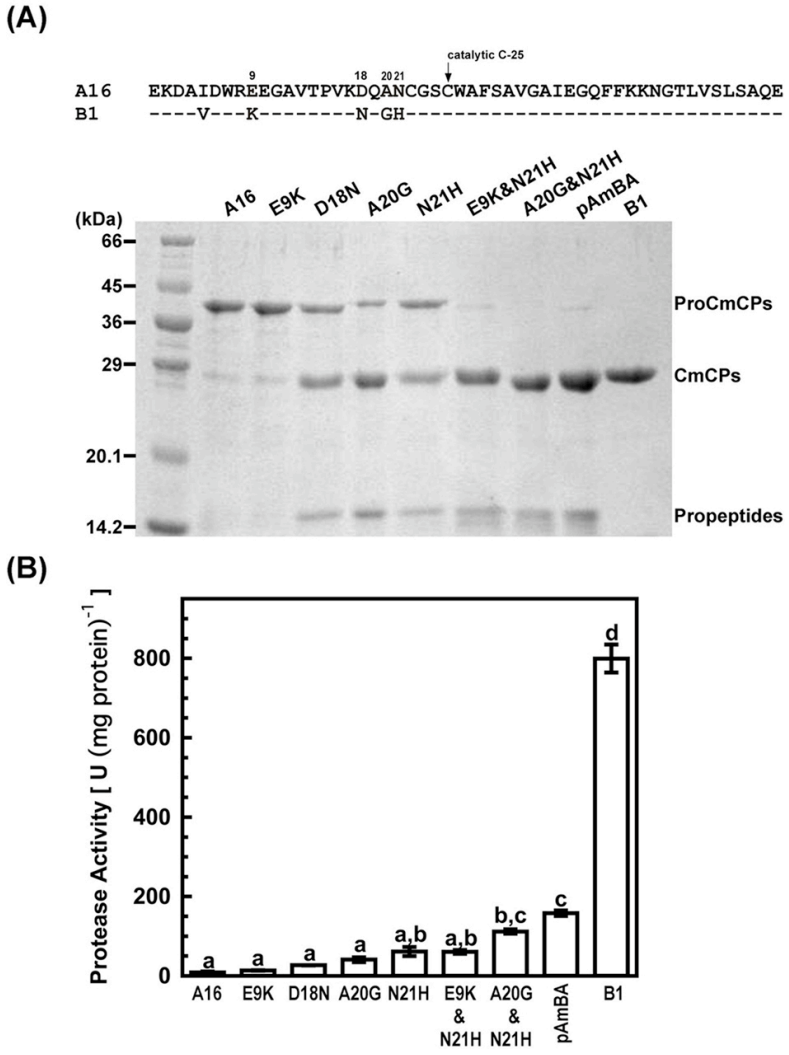

Figure 3.3. Activating A16 by incorporating single or double B1 residues. (A) Autoprocessing and (B) proteolytic activity of mutated proteins. Amino acid residues subjected to site-directed mutagenesis are indicated in the N-terminal mature protein fragments. The mutagenesis procedure was described in Experimental Procedures. Experiments were done in triplicate. Means are not significantly different when followed by the same letter as determined by the Fisher's protected LSD test (P=0.01). Error bars indicate standard errors.

3.2.4. Inhibitory activity and stability of cleaved propeptides determine CmCP functions

Enzymes preceded by pA16 propeptide exhibited lower proteolytic activity than when the same mature enzymes are preceded by pB1, e.g. B1 versus pAmBB (Fig. 3.2B). It is unlikely that the difference in activity resulted from differing autoprocessing because comparable conversion of proproteins to their mature forms was observed among chimeras capable of self-maturation (Fig. 3.2A). Excised propeptide pB1, however, was undetectable after 1 hr incubation during the autoprocessing reaction, while pA16 propeptide was readily visible, indicating that its degradation was at best partial. Thus, persistent pA16 propeptide, which was no longer covalently attached to the mature enzyme but incompletely degraded, could affect activity of cognate enzymes. In order to establish the role of the propeptides in proteolytic efficacy, I bacterially expressed isolated propeptides of A16 (pA16) and B1 (pB1), respectively. I then compared the recombinant propeptides for their inhibitory activity as well as protein stability when incubated with autoprocessed B1. Recent success in expressing propeptides of human cathepsins suggested the feasibility of this approach (Carmona et al., 1996; Billington et al., 2000).

Compared to pB1, pA16 exhibited higher inhibitory activity (Fig. 3.4A), which apparently was due to a higher proteolytic resistance than its counterpart. While efficient degradation of pB1 occurred within 1 hr of incubation with mature B1, the same enzyme could only partially fragment pA16 (Fig. 3.4B). Although it is unclear

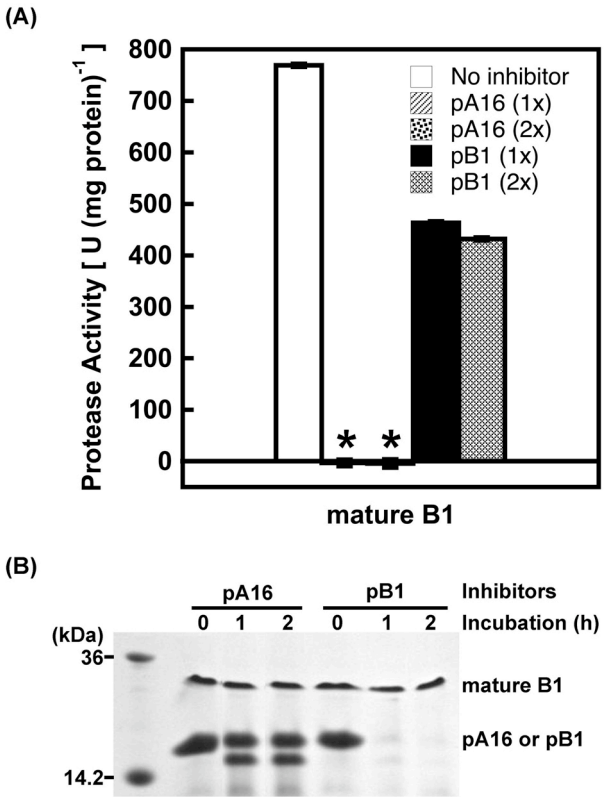

Figure 3.4. The propeptide of A16 has higher inhibitory activity and is more resistant to proteolysis than that of B1. (A) Proteolytic activity of mature B1 was determined using the azocasein assay at 1:1 (1x) and 2:1 (2x) molar ratios of propeptide: mature B1 (respectively). Activity without propeptide is the no-inhibitor control. * Indicates measurements when pA16 was added. (B) Recombinant propeptides were incubated with processed B1 for the indicated periods. The reaction mixture was resolved on 12% tricine SDS-PAGE.

whether the shortened pA16 retained its inhibitory function, the remaining intact pA16 could certainly block access of substrates and hinder the catalytic efficacy. Different residual inhibition likely explains the discrepancy in mature enzyme activity when preceded by different propeptides (Fig. 3.2B). This finding reveals yet another potential

mechanism that insects use to regulate their digestive enzymes; using propeptides of different stability to regulate activity of individual protease isoforms.

3.2.5. Subfamily B has higher intrinsic substrate affinity as well as enzymatic activity than subfamily A

To determine factors other than propeptide regions that contribute to proteolysis efficacy, the chromogenic substrate Z-Phe-Arg-pNA was selected to obtain kinetic parameters for pBmAB and pBmBA, as well as their parental B1. The chimeras contained the highly degradable pB1 propeptide (Fig. 3.2A), thus complications introduced by pA16 residual inhibitory activity are eliminated. Further, although the autoprocessed pB1 band was undetectable on SDS-PAGE, the possibility remained that a small amount of non-degraded pB1 could still persist and affect the kinetics of pBmAB, pBmBA and B1. Thus I performed enzymatic assays after 1, 2 and 6 hr (respectively) of autoprocessing (Table 3.1). Since no increase in activity was observed during this time course, the possibility of inhibition due to remaining pB1 in selected members can be excluded.

CmCPB1 showed the highest substrate affinity and Vmax. Incorporation of the A16 sequence, at either the N- or C-terminal mature peptide region, appeared to weaken the interaction of the enzyme with the substrate and decrease velocity (Table 3.2). This is in agreement with the observation of decreased proteolytic activity of the chimeric proteins in the azocasein assay (Fig. 3.2B). Our results indicated that differential

Table 3.1 Proteolytic activities on Z-Phe-Arg-pNA of pBmAB, pBmBA and B1 following 1, 2 and 6 hr autoprocessing

Enzyme	OD_{410} after autoprocessing (hr)		
	1	2	6
B1	0.54 ± 0.01	0.40 ± 0.00	0.39 ± 0.01
pBmAB	0.26 ± 0.00	0.26 ± 0.01	0.24 ± 0.00
pBmBA	0.13 ± 0.00	0.12 ± 0.00	0.09 ± 0.00

Table 3.2 Specific activities of selected CmCPs on Z-Phe-Arg-pNA at pH 5.0

Enzyme	Km (mM)	Vmax (nmol/min/mg)	Vmax/Km
mB1	0.017	2.6	157.5
mAB	0.023	2.5	108.4
mBA	0.043	0.3	7.7

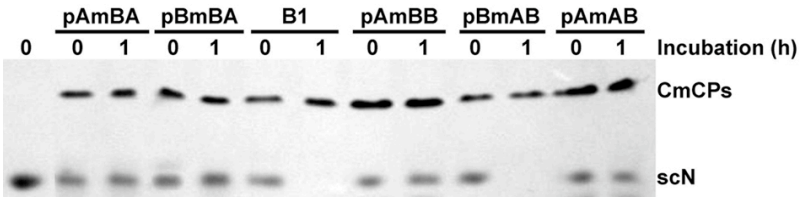

Figure 3.5. B1 propeptide is required for scN degrading activity. Autoprocessed chimeric and parental recombinant proteins were incubated with scN at 2:1 molar ratio. Reactions were subjected to 12% tricine SDS-PAGE.

proteolytic activity between subfamilies A and B in cowpea bruchid guts could also be controlled by sequences of the mature proteins, in addition to the propeptide effects.

One of the characteristics of CmCPB1 indicative of its role in insect adaptation to the protease inhibitor scN was that CmCPB1 was able to degrade scN when in excess (Ahn et al., 2004). The effect of the three regions defined by swapping on scN degradation was examined by checking scN integrity after incubation with each autoprocessing-capable CmCP chimeric protein (Fig. 3.5). None of the recombinant proteins preceded by proteolysis-resistant pA16 could hydrolyze scN, indicating highly degradable pB1 propeptide from CmCPB facilitates the disarming of plant defense proteins. However, having a degradable propeptide alone is insufficient for degrading scN. pBmBA did not hydrolyze scN, indicating that additional element(s) for scN degradation appear to reside in the C-terminus of B1 (Fig. 3.5, pBmAB). In contrast, the N-terminus of mature B1 enzyme was neither necessary nor sufficient for scN cleavage (Fig. 3.5, pBmBA versus pBmAB).

3.3. Discussion

Herbivorous insects utilize a suite of endo- and exopeptidases not only for breakdown of a variety of dietary proteins, but also as a counter-defense reservoir to help them cope with dietary toxins and anti-nutritional compounds they may encounter (Broadway, 2000). Numerous isoforms of major digestive CmCP proteases, varying in performance, have been identified from midguts of cowpea bruchids. It is intriguing that these

bruchids, under challenge by the dietary protease inhibitor scN, are able to selectively over-express a subset of the CmCP isoforms most beneficial to themselves. Here, I illustrated the underlying events that condition the superiority. I have focused on three testable biochemical properties of the chimeras as summarized in Table 3.3; a) autoprocessing, b) proteolytic activity, and c) scN-degradation. I have demonstrated that simply swapping propeptide regions between A16 and B1 did not qualitatively change the autoprocessing. The control of autoprocessing appears to reside in the mature region. Autoprocessing of CmCPA16 can be activated by such small changes as introducing a single mature B1 amino acid substitution, and substantially enhanced by two amino acid changes (Fig. 3.3). The four amino acid residues in proximity of Cys25 that underwent site-specific mutations are conserved in all CmCP subfamily A members which are unable or less efficient in autoprocessing (Ahn et al., 2004). In human cathepsin L, helix α3p of the propeptide is lodged between the loop formed by residue 20 to 24 and other mature peptide regions (Coulombe et al., 1996). Thus, alteration of the analogous residues in A16 located at or near the loop, i.e. D18, A20 and N21, could have interfered with the propeptide-mature A16 interaction, resulting in autoproteolysis. Since E9 is outside the interacting region, mutation at this site could have less impact on autoprocessing. Thus the unique mature B1 sequence confers not only higher intrinsic enzyme activity against the tested substrates but also higher autoprocessing efficiency than CmCPAs. However, ability to autoprocess did not always directly correlate with higher proteolytic activities of mature enzymes. We have shown that propeptides play critical functions in differentially adjusting proteolytic

Table 3.3 Summary of autoprocessing and activity of chimeric and original proteins

Chimera	Swap	Autoprocess	Propeptide remaining	Activity comparison	Degrading scN
A16 (pAmAA)	Original	No	_		_
pBmAA	I	No	_	–	_
pAmBA	II	Yes	Yes		No
pBmBA	III	Yes	No	pBmBA > pAmBA	No
pAmAB	III	Yes	Yes		No
pBmAB	II	Yes	No	pBmAB > pAmAB	Yes
pAmBB	I	Yes	Yes		No
B1 (pBmBB)	Original	Yes	No	B1 > pAmBB	Yes

activity after enzyme maturation. Propeptide pB1 was degraded quickly upon cleavage from the B1 proprotein, and was thus unable to interfere with the activity of mature proteases in both general proteolysis and scN degradation. In contrast, propeptide pA16 was highly stable and could inhibit mature CmCP activity even after cleaved from the mature proteins (Fig. 3.4). Therefore, autoprocessing efficiency and propeptide stability coordinately determined the protease activity.

3.3.1. Propeptides function as regulators of proteolysis to counter scN inhibition

Propeptides maintain the enzyme latency of proproteins; therefore, efficacy of maturation can differentially influence the digestive activity in members of a multigene family where variation in propeptide sequence exists. Little is known of the distinct function of propeptides within the same cathepsin family. Previously we have demonstrated differing transcriptional expression of CmCPs (Zhu-Salzman et al., 2003)*. Here I illustrate that insects also fine-tune CmCP function at the posttranslational level. Regulation of enzymatic activity by propeptides is apparently accomplished by utilizing propeptide isoforms with different stability (Fig. 3.4). Studies with mammalian cathepsins B, K, L and S showed that propeptides are all potent inhibitors of their own mature enzymes, but specificity toward cognate cathepsins can differ (Carmona et al., 1996; Maubach et al., 1997; Billington et al., 2000; Quraishi and Storer, 2001). However, among the cowpea bruchid cathepsin L-like CmCPs, the propeptide is not necessarily the most potent inhibitor of its own mature region, as I have shown that pA16 has higher inhibitory activity than pB1 toward mature B1. Although propeptides are detached from the mature CmCPs after autoprocessing, it is reasonable to assume that the binding mode is the same as when they are covalently linked. I have previously shown that A16 proprotein was readily processed by other cysteine and

---

* I performed screening of gut cDNA CmCP libraries from both scN-adapted and unadapted insects and DNA sequencing of 49 cDNA clones (Fig. 5), Southern blot analysis of CmCP (Fig. 6) and Northern blot analysis and Real-time PCR (Fig. 7) in Zhu-Salzman, K., Koiwa, H., Salzman, R.A., Shade, R.E. and Ahn, J-E. (2003). *Insect Mol Biol* 12: 135-14.

aspartic proteases presented in cowpea bruchid gut extract (Ahn et al., 2004). Thus pA16 should not inhibit protein digestion under normal conditions. Once insect digestive function is threatened, production of cathepsins with less stable propeptides could become very crucial for insect survival. The ERFNIN motif that forms the secondα-helix is conserved in CmCPs, and the altered residues are mainly located at the C-terminal segment of the propeptides (Fig. 3.1). These changes may loosened their interaction with the catalytic cleft, and promote its degradation. Presumably, decreased stability and/or inhibition capacity of pB1 freed more mature B1 enzyme, which also impacted scN-degradation.

3.3.2. Cooperative interactions of mature enzymes with propeptides

The trigger of autoprocessing appears to reside in the sequence of the protease proper. A higher proportion of B1 component in recombinant proteins always resulted in higher activity and more effective autoprocessing. Suppression of autoprocessing of A16 likely requires coexistence of multiple amino acids scattered through the entire mature protein region. Producing such autoprocessing-null proteases with relatively lower enzymatic activity must result from the necessity of strict control of proteolysis that could otherwise cause unwanted damage to cells or tissues in insect alimentary tracts.

One can envision that when insects are facing inhibitor challenge, the need for restricting activity becomes secondary. Expressing subfamily B enzymes that have scN-hydrolyzing ability and high intrinsic proteolytic activity as well as easily degradable

propeptides once matured is advantageous in digestion of a diet rich in protease inhibitor such as scN. It is worth noting that mature B1 was susceptible to scN inhibition, and only when it outnumbered the inhibitor could it degrade scN (Ahn et al., 2004). Presumably, the preferred enzyme-inhibitor interaction is via scN's blocking of mature B1's active site. Strong substrate affinity, however, may have forced excess CmCPB1 to bind to a less favorable location of scN already interacting with another CmCPB1 molecule. Degradation of scN (at the normally non-interacting site with cysteine proteases) likely is further facilitated by its high intrinsic velocity.

Characteristics of B1 are reminiscent of Kunitz inhibitor-degrading mesotrypsin PRSS3 (Nyaruhucha et al., 1997; Szmola et al, 2003). But in contrast to CmCPBs, in which multiple residues appear to be involved in scN-degrading activity, a single amino acid alteration through natural selection is responsible for trypsin inhibitor-fragmentation in mesotrypsin. Conceivably, enzymes like CmCPB1 and PRSS3 may also pose a hazardous rather than beneficial effect on the organisms in the absence of inhibitors. This may explain the relatively low levels of expression of B1 and PRSS3 compared to the major digestive enzymes under normal growth conditions (Szmola et al., 2003; Ahn et al., 2004). Premature activation of mesotrypsinogen in the pancreas is thought to cause the development of human pancreatitis by reducing the protective levels of pancreatic secretory trypsin inhibitor (Szmola et al., 2003). In parallel, CmCP subfamily B represents specialized isoforms that are ordinarily minor components of the enzymatic suite, to minimize their unwanted effect.

3.3.3. Implications in biotechnology

Insect counter-defense genes apparently are regulated at multiple levels, which enables them to effectively evade natural plant defense if needed. Propeptide regions have been considered for pest management due to the high potency and selectivity of inhibition (Michaud and Nguyen-Quoc, 2000). For instance, the proregion of the papaya proteinase IV was active against Colorado potato beetle digestive cysteine proteases (Visal et al., 1998). My study indicated that not all proregions from the same gene family have the same effects, nor are they necessarily the most potent inhibitors of their parental mature forms. CmCPs highly induced by scN likely possess less stable proregions. The most abundant isoforms under unchallenged conditions (such as CmCPAs) may be the source of inhibitors that are more resistant to proteolysis, due to the need for more strict control to prevent unwanted proteolytic activity. Individual inhibitors should be tested against the target pests, and their inhibitory activities are potentially improvable. For instance, inhibition of pA16 can be strengthened by removal of B1-processing site, rending it less susceptible to fragmentation (Fig.4B). Further, this effect can be potentiated by simultaneous treatment of other inhibitors such as scN.

Previous attempts to control insects through direct inhibition of insect digestive proteases have met with only limited success. Blocking upstream gene activation machinery could be more effective, because this could inhibit expression of a set of counter-defense-related genes, and thus could be potentially useful in biotechnology-based pest control. In mammals, a peptide hormone cholecystokinin (CCK) of the

gastrointestinal system causes the release of a variety of digestive enzymes from the pancreas and gallbladder (Iwai et al., 1987). CCK release is mediated by another peptide that is trypsin-sensitive. Despite its sensitivity to trypsin, dietary proteins prevent inactivation of the trypsin-sensitive CCK-releasing peptide, and its presence stimulates pancreatic enzyme secretion through its CCK releasing effect (Iwai et al., 1987; Tsuzuki et al., 1992). It is not yet known how insects regulate the differential secretion of digestive enzymes. It is possible that similar feedback regulation of major digestive enzymes exists, which explains at least the induction of CmCPs in subfamily A in the presence of dietary scN.

3.4. Experimental procedures

3.4.1. Gene swapping using restriction digestion-based subcloning

CmCP isoforms *A16* (GenBank accession number AF544839) and *B1* (AF544844), previously subcloned into pET44a(+) (Novagen Madison, WI, USA), were excised with *Bam*HI and *Xho*I restriction endonucleases, gel purified and restricted with *Bsi*EI that cuts once in both fragments. Two fragments, corresponding to the propeptide and mature regions (respectively) in each of the restriction reactions were purified. Cross ligations, (i.e. pA16 rejoined with mature B1 at *Bsi*EI site and vice versa), resulted in two chimeric fragments *pBmAA* and *pAmBB* (swap I). Swap II was performed similarly to swap I except that the two unique restriction enzymes *Bsi*EI and *Spe*I were used. The

N-terminal mature *CmCP* regions (142 bp) from *A16* and *B1* were switched again in the subsequent ligation reactions that produced *pAmBA* and *pBmAB*. For swap III, aimed at swapping the C-terminal sequences between A16 and B1, *pET44a-A16* and *pET44a-B1* were digested with *Spe*I. *Spe*I restricted between amino acids 52 and 53 of the mature CmCPs as well as the vector sequences upstream of the cloning sites. The *Spe*I fragments were exchanged, followed by religation reactions to generate *pBmBA* or *pAmAB*. All chimeric fragments were cloned back into pET-44a(+) for protein expression. The constructs were then transferred to *E. coli* DH5α cells, and subjected to sequencing reactions.

3.4.2. Expression of recombinant proteins

After correct DNA sequences were confirmed, constructs were transferred to *E. coli* expression host strain Rosetta-gami (DE3) (Novagen). Cells were grown at 37°C until OD_{600} reached between 0.4 and 1.0. Production of recombinant proteins fused with Nus A was induced by addition of isopropyl-β-D-thiogalactopyranoside (IPTG, 1.0 mM) overnight at room temperature. Cells were then harvested and disrupted by sonication (Model 250 Sonifier, Branson). Recombinant fusion proteins were purified using a Ni^{2+}-chelate affinity column according to the manufacturer's instructions (Amersham Pharmacia Biotech, Piscataway, NJ). The Nus A-bound column was reacted with 0.5 units/mg of α-thrombin (Sigma, St. Louis, MO) in thrombin cleavage buffer (20 mM Tris-HCl, 150 mM NaCl, 2.5 mM $CaCl_2$ and 0.1 mM DTT, pH 8.4) for 20 hr at 4°C.

Proteins dissociated from the Nus tag were concentrated by adding ammonium sulfate to 80% satuation (Englard and Seifter, 1990) and analyzed on 12.5% SDS-PAGE or 12% tricine SDS-PAGE (Schagger and von Jagow, 1987).

3.4.3. Autoprocessing of recombinant proproteins

Autoprocessing reactions were performed according to Ahn et al. (2004). The proproteins (5μg) were incubated at 37°C in 100 mM sodium acetate, 2.0 mM EDTA and 2.5 mM DTT, pH 5.0 for 1hr. For the time-course autoprocessing experiment, 2μg of proproteins were subjected to 1, 2 or 6 hr incubation. The processed CmCPs were then analyzed on 12.5% SDS-PAGE, or used for proteolytic reactions.

3.4.4. Azocasein assays and scN-degradation

The processed CmCPs (derived from 10 μg propropteins) were evaluated for proteolytic activity using azocasein as the substrate as previously described (Ahn et al., 2004). Experiments were performed in triplicate for statistical analysis. Enzymatic activity was plotted using KALEIDA-GRAPH (Synergy Software). One unit of protease activity was defined as the amount of mature CmCP required to produce an absorbance change of 0.01 per hr in 1-cm cuvette at 37°C. A one-way ANOVA test was used to analyze the proteolytic activity data, and Fisher's protected LSD test ($P = 0.01$) was used for mean separation (SPSS for Windows ver.11.1).

To determine whether the chimeras were able to degrade scN, processed recombinant proteins were incubated with scN for 1 hr at 37°C and 2:1 enzyme: inhibitor ratio. Reactions were subjected to 12% tricine SDS-PAGE.

3.4.5. Recombinant propeptides: production, inhibition and stability

cDNA fragments encoding propeptide regions of *A16* and *B1* were PCR-amplified (95°C for 30 s, 68°C for 1min for 35 cycles) using combinations of the following oligonucleotide primers: (1) 5′-TTAATGGATCCTCTTCGGTCTACGAAG AGTGGC-3′; (2) 5′-TTGATCTCGAGTTATTCCATATCAATGTCCTCAAAGTT-3′; (3) 5′-TTGATCTCGAGTTATTCGATATCAGTCTCCTCAAAGTA-3′. Restriction sites *Bam*HI and *Xho*I (underlined) were introduced into primers for directional subcloning. Primers 1 and 2 were used to obtain *pA16* and primers 1 and 3 for *pB1*. After restriction digestion, the fragments were subcloned into pET-44a(+). Procedures described above were followed for DNA sequence confirmation and recombinant propeptide production. The purified proteins (4 μg) were visualized on 12% tricine SDS-PAGE.

Propeptides, pA16 and pB1 respectively, were mixed with mature B1 at 1:1 and 2:1 of (inhibitor : enzyme) molar ratios respectively and preincubated at room temperature for 20 min. The mixtures were then subjected to azocasein assays and activity was plotted as previously described. Mature B1 without propeptides was used as no-inhibitor control.

Propeptide stability was evaluated by incubation of 4 μg recombinant pA16 or pB1 with processed B1 (from 4 μg proCmCPB1) for 0, 1 or 2 h respectively, at 37°C. The reaction mixtures were then resolved on 12 % tricine SDS-PAGE.

3.4.6. Site directed mutagenesis

Site-directed mutagenesis was carried out according to the PCR-based mutagenesis method (Fisher and Pei, 1997) with modifications. Briefly, the adjacent and phosphorylated oligonucleotide primers were designed on opposite strands with the desired mutation (small letter) located in the middle of nucleotides of one primer: (1) Antisense 5′- AGGTGTCACGGCTCCCTC TTt (E9K) TCTCCAGTC -3′; (2) Sense 5′- GTCAAGGACCAGGCAAATTGCGGATCA -3′; (3) Antisense 5′- AGGTGTCACGG CTCCCTCTTCTCTCCAGTC -3′; (4) Sense 5′- GTCAAG aAC (D18N) CAGGCAAA TTGCGGATCA -3′; (5) Sense 5′- GTCAAGGACCAG GgA (A20G) AATTGCGGATC A -3′; (6) Sense 5′- GTCAAGGACCAGGCA cAT (N21H) TGCGGATCA-3′; (7) Sense 5′- GTCAAGGACCAG GgA (A20G) cAT (N21H) TGCGGATCA -3′. The mutants were PCR amplified (95 °C for 30 s, 68 °C for 7 min for 35 cycles) using *pET44a-A16* as the template. Selected oligonucleotide primers were designed to introduce various point mutations: primers 1 and 2 for E9K mutation, primers 3 and 4 for D18N mutation, primers 3 and 5 for A20G mutation, primers 3 and 6 for N21H mutation, primers 1 and 6 for E9K&N21H mutations, and primers 3 and 7 for A20G and N21H mutations. Following PCR amplification, the product was self-ligated and treated with *Dpn*I

(specific for methylated and hemimethylated DNA in the parental DNA template) to select for mutation-containing synthesized DNA. The parental DNA isolated from *E. coli* DH5α cell (which is dam^+) is methylated and therefore susceptible to *Dpn*I digestion. The ligated products were transformed into *E. coli* DH5α cells and the site-directed mutations were confirmed by DNA sequencing. Recombinant site-directed mutants were expressed and purified in the *E. coli* strain Rosetta - gami (DE3) as described. The purified proteins (4 μg) were analyzed on 12.5 % SDS-PAGE.

3.4.7. Kinetics

Protease activity was evaluated with a synthetic chromogenic substrate, benzoyloxycarbonyl-L-phenylalanyl-L-arginine-*p*-nitroanilide (*Z*-Phe-Arg-*p*NA) (Bachem, King of Prussia, PA) as described (Tchoupe et al, 1991; Turk et al, 1993; Leippe et al, 1995; Paramá et al, 2004). Recombinant proproteins (B1, pBmBA and pBmAB) were incubated under acidic conditions (pH 5.0) for 1 hr at 37°C. The substrate, Z-Phe-Arg-pNA, at 1, 3, 10, 25, 30, 50, 75, 100, 250, 500 μM, respectively was prewarmed to 37°C in 100 mM sodium acetate buffer, pH 5.0, 1 mM EDTA and 2 mM DTT prior to addition of processed enzymes. Processed CmCP (10 μL) was added to 590 μl of the prewarmed solution. The rate of cleavage of the chromogenic *p*-nitroanilide group from the substrate was measured at 410 nm for 10 min with 7.1 sec intervals in a Beckman DU 64 spectrophotometer thermostated at 37°C. A total of 10 substrate concentrations (1 to 500 μM) were tested, and initial velocity for each substrate

concentration was calculated. To determine Vmax and Km, data at each concentration were fit to the Michaelis–Menten equation $v_0=V_{max}[S]/(K_m+[S])$ using the non-linear least-squares fitting analysis of KALEIDA-GRAPH (Synergy Software).

CHAPTER IV

SEVEN-UP FACILITATES INSECT COUNTER-DEFENSE BY SUPRESSING CATHEPSIN B EXPRESSION*

4.1. Introduction

Herbivorous insects are constantly challenged by a broad spectrum of toxins and anti-nutritional factors produced by their host plants. The insect alimentary tract thus becomes the front line of insect counter-defense. It actively responds to dietary challenges by readjusting expression of its transcriptome and changing the repertoire of proteins in cells that line the digestive tract. Insect digestive enzymes, broadly classified into serine, cysteine, aspartate and metallo- proteases (Terra & Ferreira, 1994), play an important role in protecting the vulnerable cells and tissues of the insect body, in addition to functioning in food breakdown. The cowpea bruchid *Callosobruchus maculatus* dramatically remodels its profile of midgut digestive enzymes in response to the soybean cysteine protease inhibitor scN. This insect not only reconfigures expression of its major digestive enzymes, the cathepsin L-like cysteine proteases, but also drastically induces a cathepsin B-like cysteine protease, namely *CmCatB* (Ahn et al., 2004; Moon et al., 2004; Ahn et al., 2007b). These changes apparently help the insect cope with nutrient deficiencies and resume normal feeding and development

---

(Zhu -Salzman et al., 2003).

Although undetectable in unchallenged insect guts, *CmCatB* was the most highly induced gene in microarrays designated to identify scN-regulated genes (Moon et al., 2004). This finding is intriguing, because its human ortholog, cathepsin B possesses an "occluding loop" that has been shown to block the access of substrates and inhibitors (Musil et al., 1991; Illy et al., 1997). It is likely that CmCatB enzymes play a role in cowpea bruchid adaptation by rendering cowpea bruchids less susceptible to scN inhibition. This hypothesis is supported by the presence of inhibitor-induced and -insensitive cysteine protease activity in challenged cowpea bruchids (Zhu-Salzman et al., 2003). Further, messenger RNA profiling through larval development under scN challenge revealed that accumulation of *CmCatB* transcript peaked in the 4^{th} instar, concordant with the time of adaptation (Zhu-Salzman et al., 2003; Moon et al., 2004). Together, these data suggest that *CmCatB* has a unique function in insect adaptation to dietary scN.

Genetic engineering for insect resistance using naturally-occurring plant defense genes represents an environmentally friendly approach for pest management. However, this biotechnology-based pest control strategy is threatened by insect adaptability. The Zhu-Salzman's lab, as well as others, has shown that insect adaptive response to dietary inhibitors is mediated through transcriptional activation of a number of genes, including proteases that are insensitive to the plant inhibitors and proteases that degrade the inhibitors. However, very little is known concerning how insects sense the challenge and direct the activation of counter-defense genes. Elucidation of the underlying

regulatory mechanisms will help identify new vulnerabilities in an insect, and may eventually be exploited for better insect management.

To deepen our understanding of insect counter-defense machinery, I investigated the transcriptional activation of *CmCatB*, a gene that is highly responsive to dietary scN treatment. I identified a *COUP* element in the *CmCatB* promoter that specifically interacted with a nuclear protein factor from unadapted insect guts. Consistently, a higher abundance of CmSvp, a COUP-transcription factor (COUP-TF) homolog was detected in unadapted insect guts, where *CmCatB* is not expressed, than in adapted insect guts, where *CmCatB* is highly expressed. Transient expression of CmSvp in *Drosophila* S2 cells efficiently repressed *CmCatB* expression. Thus I have demonstrated that CmSvp is involved in the negative regulation of insect counter-defense genes that help insects to cope with plant defense compounds.

4.2. Results

4.2.1. Isolation of *CmCatB* promoter

To understand how scN induces expression of *CmCatB*, I cloned an upstream region from the cowpea bruchid genomic DNA. A 1450 bp fragment containing 181 bp

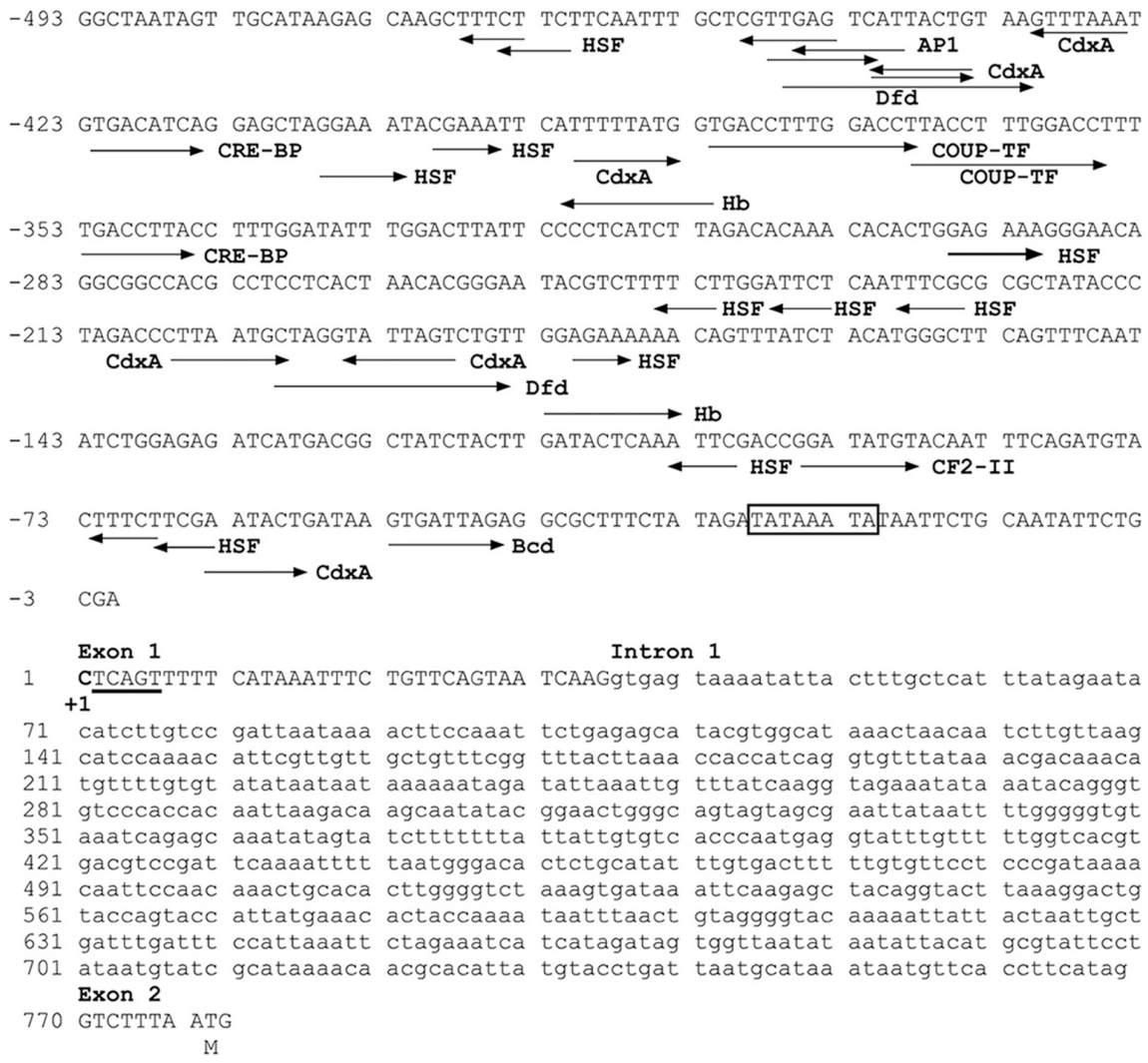

Figure 4.1. Architecture of genomic DNA upstream of CmCatB coding region. Transcription initiation site is marked as +1, and the upstream sequence is denoted with negative numbers. The intron sequence in the 5′ UTR is shown in lower case. Potential cis-regulatory elements in this putative CmCatB promoter are illustrated by arrows under the DNA sequence. A putative TATA motif is boxed, and a pentamer arthropod initiator sequence is underlined.

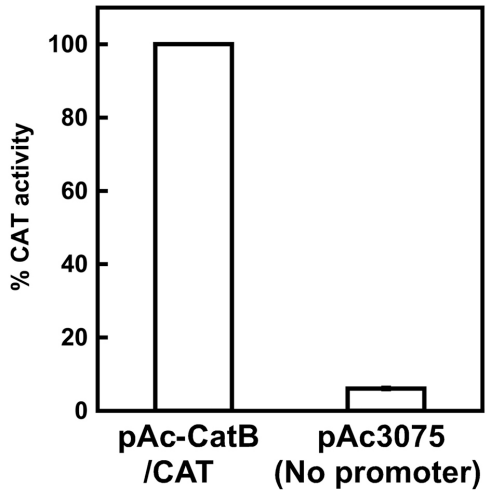

Figure 4.2. Illustration of the promoter activity of the 493 bp fragment in *Drosophila* cells. Construct pAc-CatB/CAT and reporter vector pAc3075 control was transfected into the S2 cells, respectively. CAT activity was measured and normalized as described in Materials and Methods.

of the coding region and 1269 bp of 5′ sequence was obtained by a PCR-based genome walking method (Fig. 4.1). The transcription initiation site was determined by 5′ RACE PCR. Comparison of genomic and cDNA sequences revealed a 35 bp untranslated exon as well as 734 bp intron. The 493 bp sequence flanking the 5′ end of exon 1 was thus assumed to function as the promoter for CmCatB. A potential TATA box is located between -29 and -22 position. A TCAGT pentamer was identified. This conserved sequence is known as the arthropod initiator sequence, and is important for promoter functions (Cherbas & Cherbas, 1993; Xiong & Jcacobs-Lorena, 1995). Numerous binding sites for putative trans-acting factors were identified in this promoter region.

To confirm the promoter activity of the 493 bp fragment, it was cloned into the vector pAc3075, which harbors the bacterial chloramphenicol acetyltransferase (*CAT*) reporter gene and a downstream cleavage/polyadenylation signal (Guarino & Summer, 1986). The resulting reporter construct was transiently transfected into *Drosophila* S2 cells and assayed for CAT activity. As expected, the activity of the reporter construct was significantly higher than the parental vector that contains *CAT* but no promoter (Fig. 4.2).

4.2.2. Nuclear protein factors interact specifically with CmCatB promoter region

Eukaryotic gene expression is typically regulated via interaction of *cis*-acting elements and trans-acting factors. Binding or release of the transcription factors to target promoter elements may induce or repress gene expression. To understand the interaction of nuclear proteins with the promoter elements of the scN-regulated *CmCatB*, I performed electrophoretic mobility shift assays (EMSAs). Two overlapping DNA fragments corresponding to the 493 bp promoter region were used for the binding assays (Fig. 4.3A). Nuclear extracts were prepared from guts of unadapted and adapted insects, 3 μg of which was determined to be optimal for the formation of DNA-protein complexes (data not shown). To avoid non-specific binding, 0.05 μg of poly(dI-dC) was added to all reactions. Shifted bands in adapted and unadapted extracts were detected with the upstream probe P1 but not with the promoter-proximal P2 (Fig. 4.3B).

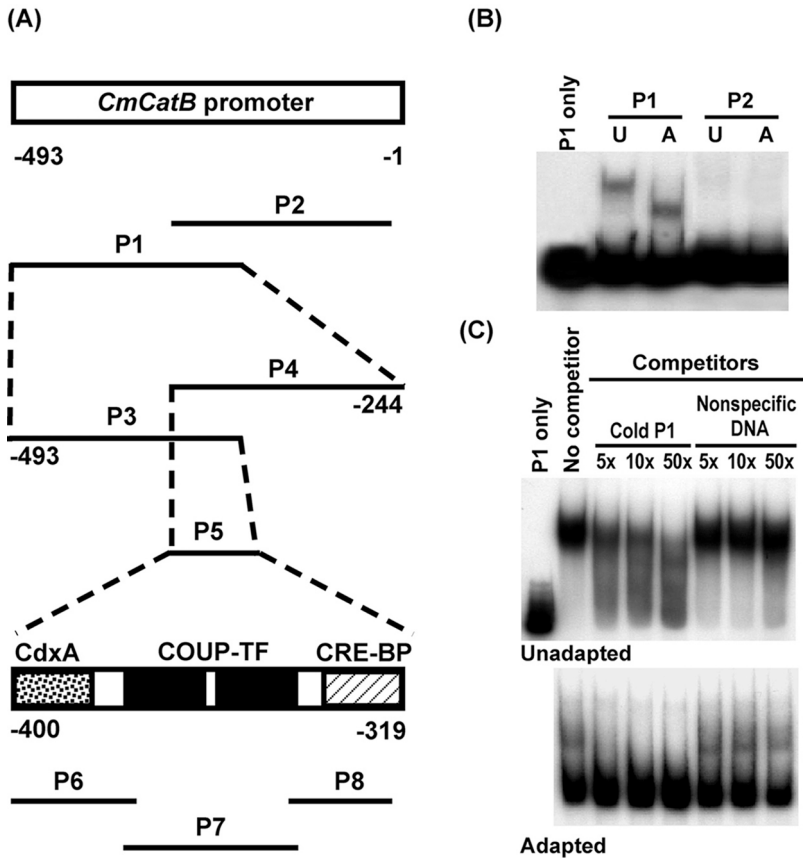

Figure 4.3. Probe dissection to locate *cis*-elements using EMSA. Nuclear extracts were obtained from freshly dissected adapted (A) and unadapted (U) guts. In competition assays, 5x, 10x or 50x molar excess of unlabeled probes, specific and nonspecific competitors, were preincubated with gut extract prior to the binding reaction. P: probe. CdxA, COUP-TF, and CRE-BP: putative *cis*-elements.

Competition assays using unlabeled probe or non-specific DNA verified that both of the P1-shifted bands were specific (Fig. 4.3C).

The observed difference in gel shift mobility suggested that different nuclear protein factors interact with the CmCatB promoter in these two extracts.

One scenario is that a negative regulator represses CmCatB expression in the unadapted gut nuclear extract through interactions with a negative element, while a factor in adapted insects binds to a positive *cis*-element that is responsible for activation of *CmCatB*. This is consistent with northern analysis showing that *CmCatB* expression is undetectable in unadapted 4th instar insect guts but highly induced in adapted insect guts (Moon et al., 2004). As an initial step in gaining a comprehensive understanding of insect adaptive mechanisms, in this study I focused on the potential negative regulation.

4.2.3. Nuclear factors of unadapted insect guts interact specifically with *COUP* element

To define the *cis*-elements, probe P1 was further dissected into two overlapping halves (P3 and P4 in Fig. 4.3A) and tested with the unadapted gut nuclear extracts (Fig. 4.4A). Both fragments bound specifically, indicating that the factor recognized the overlap between P3 and P4. Probe P5, roughly corresponding to the region common to both probes (Fig. 4.3A) indeed formed a DNA-protein complex (Fig. 4.4B). In this region, there were potential *cis*-elements corresponding to the known DNA-binding proteins CdxA, COUP-TF/Svp and CRE-BP (Fig. 4.3A). To determine which sequence within the P3-P4 overlap was responsible for the specific interaction, three probes, each encompassing one of the putative *cis*-elements, were synthesized and used in competition analysis. Only probe P7, which contains the two tandem *COUP* elements,

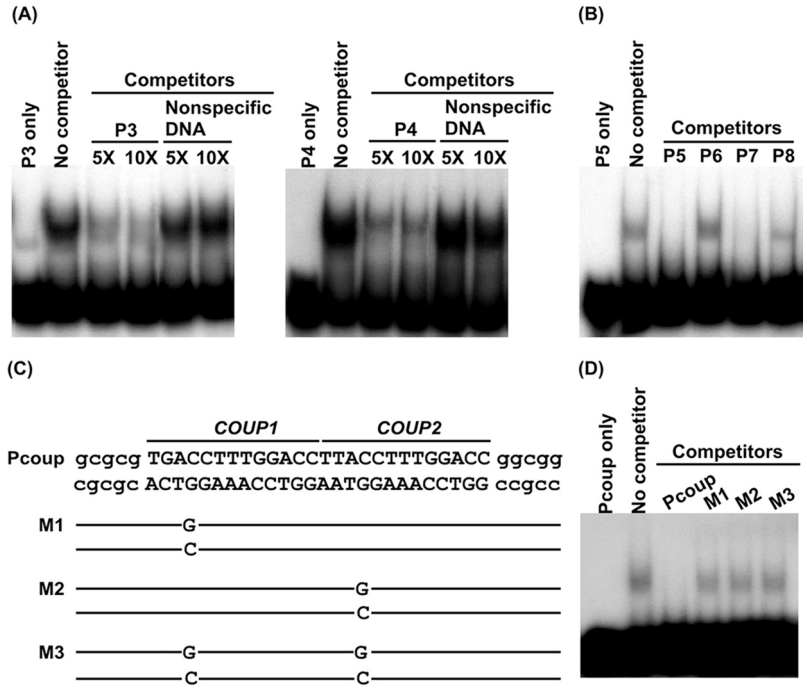

Figure 4.4. Nuclear protein factors specifically interact with *COUP* element. (A) EMSA with probes 3 and 4 to locate nuclear protein-binding site. (B) Only P7 (Pcoup) was able to compete for DNA binding of probe P5. (C) Alterations at *COUP* half-sites. (D) Mutations at *COUP* half-sites decreased the affinity of the nuclear protein factors.

could compete with P5 for protein binding (Fig. 4.4B). For the remainder of this capter, this probe is referred to as Pcoup.

The chicken ovalbumin upstream promoter-transcription factors (COUP-TF) are members of the nuclear steroid/thyroid hormone receptor superfamily (Wang et al., 1989). They bind to imperfect AGGTCA repeats, and play dual regulatory roles as activators or repressors depending on the promoter context, and are important for many biological functions (Tsai & Tsai, 1997). Therefore, we decided to test the hypothesis

that a COUP-TF interacts with the *cis*-element as a negative regulator in unadapted insect guts to repress *CmCatB* expression.

Both *COUP* elements contain direct imperfect repeats separated by two nucleotides, and there were a total of four AGGTCA half-sites in the -382/-357 region. To evaluate the effects of each individual *COUP* site on association with nuclear factors, we altered a G residue in the downstream half-site of each *COUP* site (Fig. 4.4C). It has previously been demonstrated that these residues are critical for binding of COUP-TFs (Hwung et al., 1988). In the M3 probe, both G residues were changed. None of the three mutagenized probes could compete with Pcoup probe for the protein binding (Fig. 4.4D), thus confirming that the trans factor was binding to the COUP element.

4.2.4. A COUP-TF interacts with *CmCatB* promoter

To identify the *COUP*-binding nuclear protein, I performed a supershift assay with a polyclonal anti-AaSvp antibody raised against a highly conserved region of the mosquito COUP-TF, AaSvp. Preincubation with anti-AaSvp abolished the DNA-protein association in unadapted insects, providing evidence that the binding protein is indeed a bruchid member of the COUP-TF/Svp family (Fig. 4.5). It should be noted that the shifted band in adapted insects was unaffected by anti-AaSvp antibody (data not shown). Thus, only the DNA-protein complex in unadapted insect gut cells was due to binding of a COUP-TF/Svp, and not the one formed in adapted insects. Because *COUP* binding

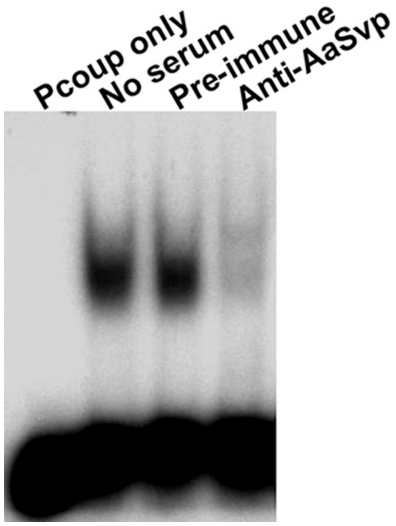

Figure 4.5. Anti-AaSvp serum abolished the *COUP*-nuclear protein association. AaSvp: COUP-TF homolog from mosquito *Aedes aegypti*. Anti-AaSvp antibody: polyclonal antibody raised against a highly conserved region of AaSvp. Antibody was preincubated with gut extract prior to the binding reaction with Pcoup.

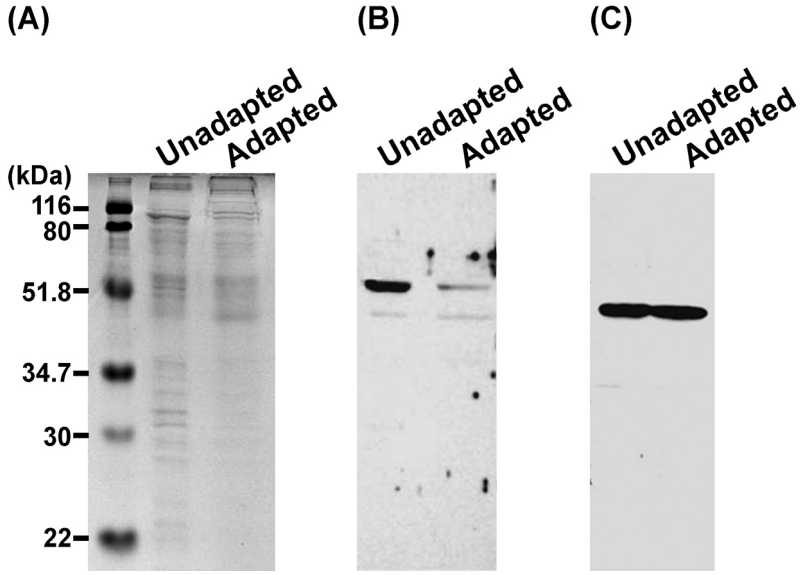

Figure 4.6. CmSvp is more abundant in scN-unadapted cowpea bruchid midgut than scN-adapted midgut. SDS-PAGE (A) and western blotting (B) of insect gut nuclear extract protein from adapted and unadapted guts. Polyclonal anti-AaSvp was used as primary antibody. (C) The protein blot was reprobed with anti-actin antibody to serve as loading control.

was not observed in adapted cowpea bruchids where CmCatB was dramatically induced, it suggests that the cowpea bruchid COUP-TF/Svp homolog may function as a repressor of *CmCatB* expression when insects are not challenged by dietary scN. Relief of repression in the adapted insect guts could then be due either to a decreased level of the transcription factor or to a post-translational modification of its activity. To test whether the cowpea bruchid COUP-TF/Svp was more abundant in unadapted insect guts than in scN-adapted guts, western blots were performed. Results revealed a significant decrease in accumulated levels in adapted insects, thus supporting the first possibility (Fig. 4.6).

4.2.5. CmSvp represses *CmCatB* expression

To provide definitive evidence that a COUP-TF/Svp negatively regulates *CmCatB* expression in cowpea bruchids, a cDNA clone encoding a putative COUP-TF was isolated by PCR using degenerate primers, followed by 5′ and 3′ RACE PCR. The resultant 1622 bp full-length cDNA clone contains an open reading frame of 1260 bp that encodes a protein of 419 amino acid residues (Fig. 4.7). Sequence alignment revealed a high degree of amino acid similarity to COUP-TFs, particularly with several insect Svp proteins, such as those from red flour beetle *Tribolium castaneum* (96%, GenBank accession number XM_962444), mosquito *Aedes aegypti* (78%) (Miura et al., 2002) and *Drosophila* (75%) (Moldzik et al., 1990). It also shares 71% amino acid sequence identity with human COUP-TF (Wang et al., 1989). We designated our clone

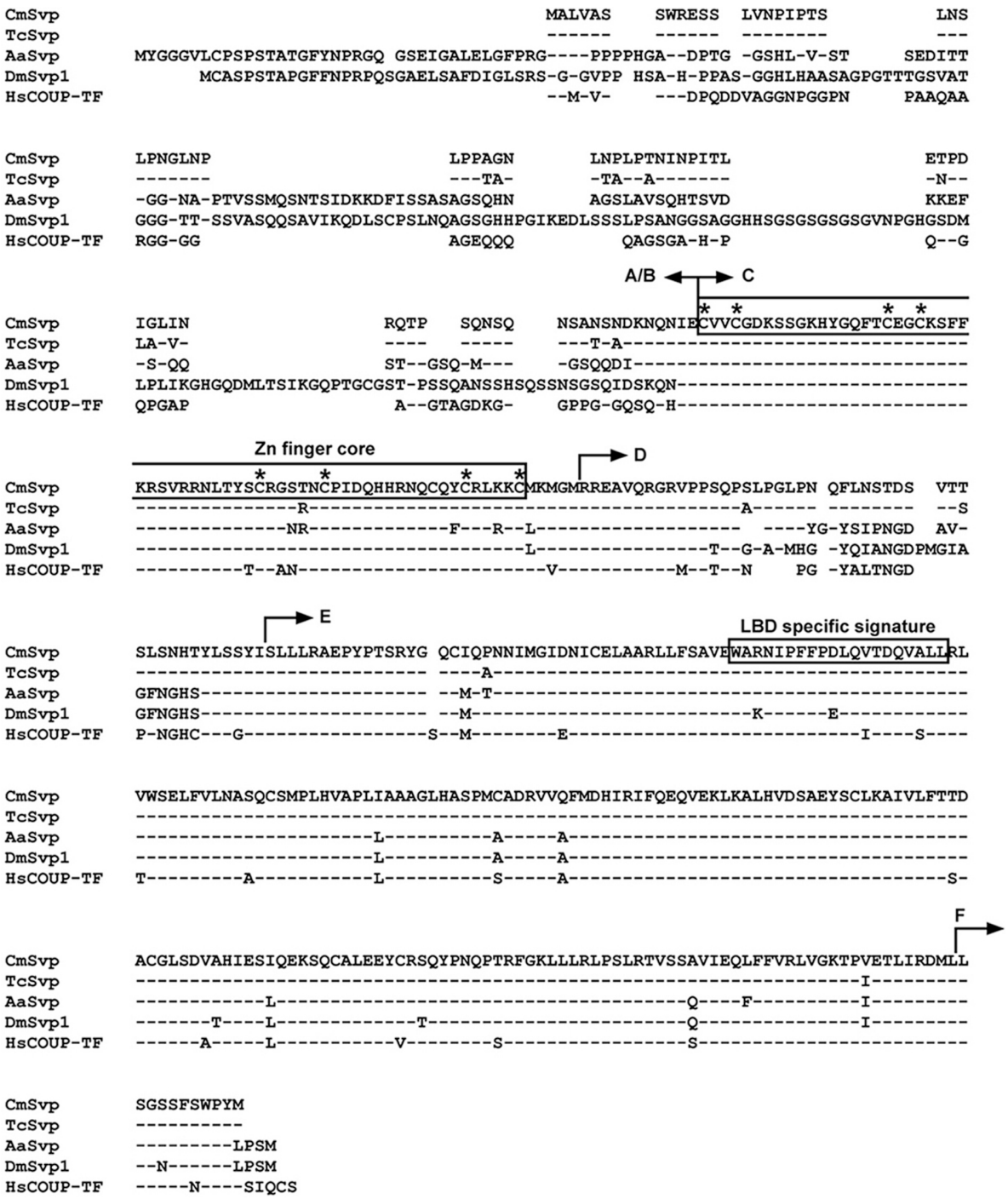

Figure 4.7. CmSvp shares high sequence similarity with COUP-TF/Svp members from red flour beetle TcSvp, mosquito AaSvp, *Drosophila* DmSvp, as well as human COUP-TF. The GENBANK accession number for *CmSvp* is EF372598. Dashes indicate identical residues. The boundaries of various regions are marked by bent arrows. Region C (the core of the DBD) and region E (the core of LBD) are the most conserved regions of COUP-TF/Svp proteins. The zinc-finger motif sequence of DBD is boxed. Eight highly conserved cysteine residues which form two zinc finger structures are indicated with asterisks. The LBD specific signature for the steroid/thyroid receptor superfamily is also boxed.

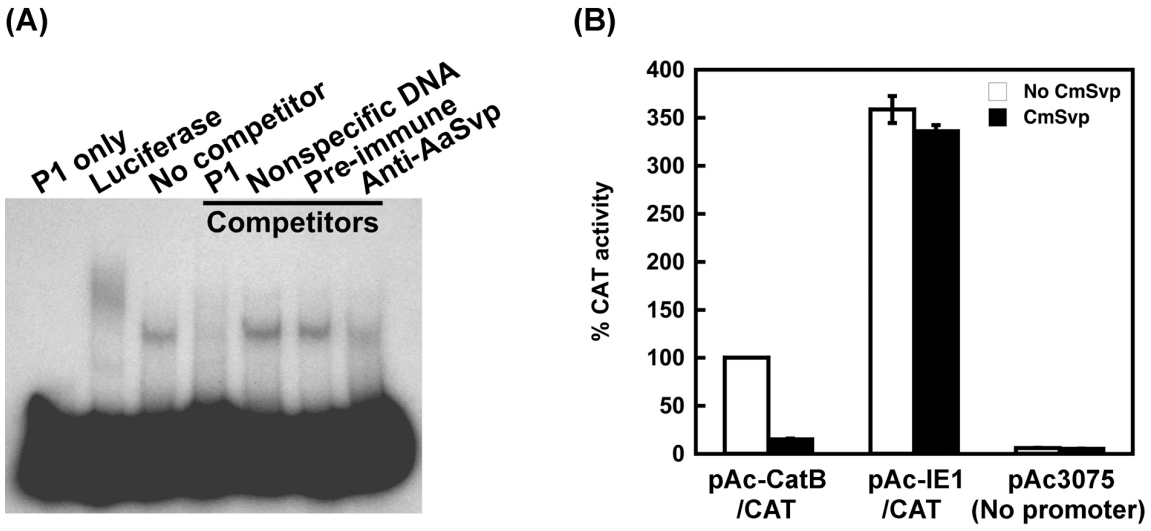

Figure 4.8. CmSvp represses *CmCatB* expression. (A) *In vitro* translated CmSvp was able to bind specifically at the *COUP* responsive element in P1 probe. Luciferase was used as a control for *in vitro* translation as well as for the EMSAs. (B) Transient expression of CmSvp abolished CAT activity (black bars). Cotransfection of empty expression vector with the reporter constructs (white bars) ensures comparable total DNA amounts in CmSvp-expressing and non-expressing S2 cells. The reporter plasmid pAc-IE1/CAT was used to determine specificity of the CmSvp and *CmCatB* promoter interaction. Transfection efficiency was standardized by β-galactosidase activity conferred by the control construct pAc5.1/V5-His/lacZ.

as CmSvp. Both the DNA-binding domain (DBD) and the ligand-binding domain (LBD) of CmSvp are highly conserved. The DBD has a typical zinc-finger motif sequence, $CX_2CX_{13}CX_2CX_{15}CX_5CX_{12}CX_4C$ (Mader et al., 1993). The 20 amino-acid residues, (F,W,Y)(A,S,I) (K,R,E,G)xxxx (F,L)xx(L,V,I) xxx(D,S)(Q,K)xx(L,V)(L,I,F), constitute an LBD specific signature for the steroid/thyroid hormone receptor superfamily (Wurtz et al., 1996). The most diverse regions among COUP-TF/Svp sequences are at the N-termini.

To demonstrate that CmSvp bound to *COUP* element, *in vitro* translated protein was used in EMSA assays. A shifted band, similar to that seen in unadapted gut extracts was observed (Fig. 4.8A). Competition assays confirmed binding specificity. CmSvp showed specific binding to the *COUP* responsive element.

To illustrate transcriptional repression of CmSvp, an expression construct with CmSvp under the control of the *Drosophila* actin 5 (Ac5) promoter was constructed. Co-transfection of pAc5-CmSvp with the reporter plasmid pAc-CatB/CAT into *Drosophila* cells showed that CmSvp efficiently abolished *CmCatB* expression (Fig. 4.8B). As a control for specificity, the IE1-CAT construct was also co-transfected with pAc5-CmSvp. CmSvp has no effect on the IE1 promoter, which does not contain *COUP* binding sites, indicating specific interaction between CmSvp and *CmCatB* promoter.

COUP-TF/Svp is able to regulate gene expression via *COUP* binding, as well as protein-protein interactions (Zelhof et al., 1995). To determine whether *COUP* binding is essential for CmSvp regulatory function, the cotransfections were also performed with construct pAc-CatBΔCOUP/ CAT, where the *cis*-element was removed. Although the promoter activity is drastically weakened in the absence of *COUP* element, it is clear that over-expression of CmSvp showed no repression on promoter activity. This result indicated that binding to the *COUP* site was required for CmSvp function (Fig. 4.9), in accordance with the EMSA results.

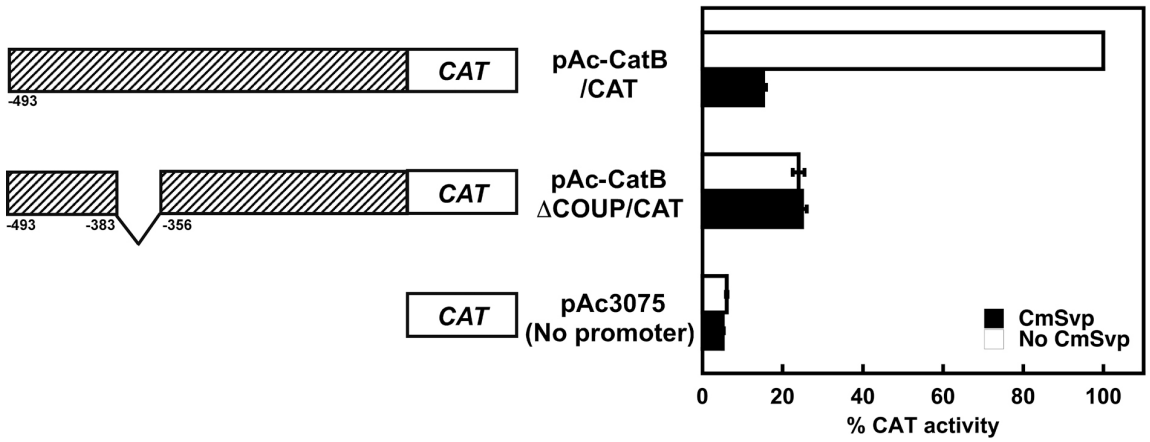

Figure 4.9. CmSvp repression of *CmCatB* requires binding at the *COUP* element. pAc-CatB/CAT and pAc-CatBΔCOUP/CAT were cotransfected with CmSvp-expressing pAc5-CmSvp (black bar) or non-expressing empty vector (white bar), respectively. The latter was to ensure comparable total DNA amounts in all transfected cells. Transfection efficiency was normalized as described for Figure 4.8.

4.3. Disscussion

Insects are capable of circumventing the negative effects of a wide range of plant toxins or other anti-nutritional factors. The Zhu-Salzman's lab has previously shown that the adaptive response in cowpea bruchids to dietary plant protease inhibitor challenge is mediated by transcriptional activation of a number of genes, including proteases that are insensitive to the inhibitors. Microarray studies revealed a cathepsin B-like *CmCatB* gene that is highly induced by a soybean cysteine protease inhibitor scN (Moon et al., 2004). The unique tertiary structure and developmental expression pattern of *CmCatB* renders it a suitable target for in-depth study on how insects regulate counter-defense

related genes. In searching for regulatory *cis*-elements in the *CmCatB* promoter and nuclear-localized trans-acting factors, I identified a COUP-TF binding site, and cloned CmSvp, the COUP-TF homolog from the cowpea bruchid midgut. I showed that CmSvp represses *CmCatB* expression, presumably via binding to the *COUP* responsive element. The inverse relationship, in adapted and unadapted insects, between *CmCatB* transcript and CmSvp protein levels suggested that CmSvp helps insects cope with dietary protease inhibitors by releasing *CmCatB* repression.

COUP-TF/Svp family belongs to the steroid/thyroid hormone receptor superfamily (Wang et al., 1989). This superfamily contains many ligand-activated transcription factors as well as a number of orphan nuclear receptors, the ligands of which have not been identified (Tsai & Tsai, 1997). COUP-TFs are among the best-studied orphan receptors. The *Drosophila Seven-up* (*Svp*) gene, encoding the COUP-TF ortholog, determines photoreceptor cell fate (Hiromi et al., 1993), controls cell proliferation in Malpighian tubules (Kerbers et al., 1998), and inhibits ecdysone-dependent transcription (Zelhof et al., 1995). Important roles of COUP-TF/Svp in neurogenesis, organogenesis and embryogenesis have been illustrated in mammals, chicken, zebrafish, frog and insects (Zelhof et al., 1995; Tsai & Tsai, 1997; Mouillet et al., 1999; Miura et al., 2002; Park et al., 2003; Raccurt et al., 2005). More recently, its involvement in regulating mobilization and utilization of glycogen and lipid in skeletal muscle cells has been reported (Stroup et al., 1997; Ferrer-Martinez et al., 2004; Myers et al., 2006).

COUP-TFs can act as activators as well as repressors. They were initially found to bind to imperfect direct repeats of AGGTCA in the chicken ovalbumin promoter, and this interaction is essential for *in vitro* transcription of chicken ovalbumin (Pastorcic et al., 1986). They also stimulate transcription of the rat cholesterol 7α-hydroxylase gene (Stroup et al., 1997), the phosphoenolpyruvate carboxykinase (Sugiyama et al., 2000), trout estrogen receptor gene (Lazennec et al., 1997), and HIV-1 long terminal repeat-directed genes in human microglial cells (Sawaya et al., 1996). Although COUP-TF was originally characterized as an activator of chicken ovalbumin gene expression, accumulated evidence indicates that COUP-TFs routinely function as negative regulators (Cooney et al., 1991; Zelhof et al., 1995; Miura et al., 2002; Ferrer-Martinez et al., 2004). In insects, COUP-TF/Svp function has been associated mainly with development. *Drosophila* Svp negatively regulates 20-hydroxyecdysone (20-E) signaling (Zelhof et al., 1995). Ecdysone-dependent signaling also plays a crucial role in the regulation of mosquito vitellogenesis. Mosquito AaSvp represses yolk protein production during mosquito vitellogenesis (Miura et al., 2002). *Tenebrio TmSvp* transcripts diminished when 20-E peaked, implying that TmSvp may negatively impact the ecdysone pathway (Mouillet et al., 1999).

In this study I demonstrated that COUP-TFs function beyond insect development. In cowpea bruchids, CmSvp normally blocks the expression of *CmCatB*, an scN inhibitor-induced gene. But when the major digestive enzymes (cathepsin L-like cysteine proteases) are inhibited, CmSvp becomes less abundant, possibly insufficient to regulate the *CmCatB* promoter, leading to *CmCatB* expression. Enlightened by the

structure of human cathepsin B, with which CmCatB shares high sequence similarity, I predict that CmCatB is insensitive to scN. Induction of such proteases would have an apparent advantage to insects in the presence of scN inhibitor.

Four modes of action of COUP-TF/Svp as repressors of gene expression have been proposed (Park et al., 2003). First, this nuclear protein can directly compete for binding sites with other nuclear hormone receptors, such as thyroid, retinoic acid and vitamin D3 receptors, that mediate hormone-induction of target gene expression (Cooney et al., 1991). Second, COUP-TFs can compete for the universal heterodimeric partner of nuclear receptors. Third, COUP-TFs can recruit corepressors and silencing mediators of the nuclear receptors through the C-terminus of the assumed ligand-binding domain (Shibata et al., 1997). Finally, COUP-TFs can repress transcription by binding directly to the ligand-binding domain of nuclear hormone receptors (Kimura et al., 1993; Achatz et al., 1997). Cotransfection of CmSvp expression vectors repressed *CmCatB* promoter activity. Direct binding of CmSvp to the *COUP* element appears to be essential for this function because deletion of the *COUP* element resulted in loss of CmSvp repression (Fig. 4.9). Whether CmSvp exerted this function through direct binding and/or through protein-protein interactions with corepressors of hormone receptors and/or receptors themselves, needs further investigation. Multiple modes of interaction have been observed in *Drosophila* Svp; this protein factor could compete with ecdysone receptor complex for the same DNA binding site, as well as forming heterodimers with the receptor (Zelhof et al., 1995).

When the *COUP* site was removed from the promoter, promoter activity decreased even in the absence of CmSvp coexpression, suggesting that a positive regulator also interacts with this responsive element. It is likely that under my experimental conditions, the activator interacts with *COUP* element more strongly than the repressor. But when CmSvp was transiently over-expressed, repression dominates. This explanation agrees with the inverse correlation between CmSvp protein and *CmCatB* expression levels (Fig. 4.6), i.e. the more CmSvp the stronger of the repression. Hepatocyte nuclear factor-4 (HNF-4) has been reported to antagonize the COUP-TF function via the same responsive element and enhance the ornithine transcarbamylase promoter (Kimura et al., 1993). It is possible that an activator of equivalent function plays a role in *CmCatB* regulation. Identifying the P1 probe-binding protein in adapted insect gut nuclear extract (Fig. 4.3) will shed some light on the activation of *CmCatB*.

It is well known that COUP-TFs are able to accommodate not only degeneracy in the consensus sequences but varied distances and orientations of the two AGGTCA half-sites as well (Tsai & Tsai, 1997; Hwung et al., 1988). In the -382/-357 region of the *CmCatB* promoter, there are a total of four AGGTCA imperfect direct repeats. Any two half-sites could, in theory, form a COUP site. The most distant two repeats are separated by 15 nucleotides, within the functional COUP-TF binding range (Cooney et al., 1992). Such an arrangement possibly offers more flexibility for regulation of *CmCatB* expression. Alternatively it may furnish a mechanism ensuring minimum expression of the *CmCatB*. This could be more efficient in nutrient uptake under normal feeding conditions because major digestive cathepsin L-like cysteine proteases are more

effective enzymes than CmCatB (Kirschke et al., 1995). Results obtained from mutagenesis at *COUP* sites supported this hypothesis (Figs. 4.4C, 4.4D).

The promoter of the human lysosomal cathepsin B has been studied for transcriptional regulation due to its association with tumor progress (Yan et al., 2000). Transcription factors Sp1 and Ets trans-activate cathepsin B in glioblastoma and in *Drsophila* cells. It is thought that this TATA-less promoter is activated and regulated via the Sp1 cluster near the transcription start site. I did not find an Sp1-binding site in *CmCatB* promoter, thus Sp1 is not likely to be involved in *CmCatB* regulation. As with *CmCatB*, expression of human cathepsin B is also impacted by a repressor element(s). Although it has not yet been determined, the *cis*-element was located in the intron 1 region rather than the upstream promoter (Yan et al., 2000). Apparent differences in expression mechanisms of human cathepsin B and cowpea bruchid CmCatB may reflect species- and/or tissue-specificity. It may also reflect their unique functions in each respective organism. Despite high amino acid sequence similarity, human cathepsin B, located in lysosomes, degrades proteins taken up by the cell, and recycles the amino acids and dipeptides for new protein synthesis, whereas CmCatB is believed to be secreted into the insect gut lumen for food protein digestion when major digestive enzymes are blocked by inhibitors.

It would be interesting to determine whether common *cis*-elements are shared by genes coordinately regulated by scN. Advances in bioinformatics and functional genomics have made it technically feasible to identify interlinked gene sets that are responsible for certain biological functions. Transcription factors that interact with

common *cis*-elements would make very attractive targets for further efforts in biotechnology-based insect control. Direct inhibition of insect digestive proteases has met with very limited success previously. Inhibition of these upstream regulators may be more effective, as they could potentially block expression of a subset of counter-defense-related genes. Inactivation of negative regulators like CmSvp may result in increased fitness cost in insects. Understanding regulation of the transcription factors thus becomes critical and requires more attention.

4.4. Experimental procedures

4.4.1. scN production and cowpea bruchid midgut and gut wall dissection

Bacterially expressed recombinant scN was purified as previously described (Zhu-Salzman et al., 2003). scN-Adapted cowpea bruchid larvae were obtained by having them feed on cowpea seeds with 0.2% scN incorporated, and scN-unadapted larvae were reared on regular diet. Adaptive feeding behavior occurred during the 4th instar (Zhu-Salzman et al., 2003), where midguts were dissected following the procedure of Kitch and Murdock (1986). To obtain gut wall tissue free of gut contents, midguts were gently cut open, and gut contents were removed by several rinses in the dissection buffer. Gut walls were then transferred to the hypotonic buffer (Active Motif, Carlsbad, CA) for nuclear extract preparation.

4.4.2. Identification of a transcription initiation site of *CmCatB*

Messenger RNA was extracted from adapted 4th instar larvae using a QuickPrep Micro mRNA Purification kit (Amersham Pharmacia Biotech). To locate the transcription start site of the *CmCatB* gene (GENBANK accession number AY429465), 1 μg of mRNA was reverse transcribed for amplification of its 5′ cDNA end with a SMART RACE cDNA Amplification kit (BD Biosciences Clontech, Palo Alto, CA). First strand cDNA synthesis was primed with a modified oligo(dT) primer. After template switch, 5′ RACE-PCR (94°C for 30 sec, 68°C for 30 sec, 72°C for 2 min for 35 cycles) was performed using the BD SMART II A oligonucleotide and an antisense gene-specific primer (5′ -TCTGAGAGGAAATCCAGCTCTGGTTGT- 3′). The PCR fragment was subcloned into the pCRII vector (Invitrogen, San Diego, CA) and subjected to sequencing analysis.

4.4.3. Cloning of the 5′ flanking region of *CmCatB*

To obtain genomic DNA, 50 cowpea bruchid midguts were homogenized in 1 ml of freshly made extraction buffer (50 mM EDTA, 0.5 % SDS, 0.2 % diethylpyrocarbonate, pH 8.0). The homogenate was incubated at 72°C for 30 min with occasional vortex mixing, followed by centrifugation at 13,000 rpm for 10 min. The supernatant was mixed with 100 μl of 5 M KOAc, incubated on ice for 15 min and centrifuged as above. After further extractions with phenol/chloroform/isoamyl alcohol (25:24:1) and

chloroform/isoamyl alcohol (24:1), the upper phase was mixed with an equal volume of isoprophyl alcohol, and centrifugated. The DNA pellet was washed with 70% ethanol, air-dried and finally resuspended in 100 μl of TE buffer.

A PCR-based genome walking method was performed to obtain DNA sequence upstream of the *CmCatB* coding region (Universal GenomeWalker kit; BD Biosciences Clontech). The primary PCR reaction (7 cycles of 94°C for 25 sec/70°C for 6 min, followed by 37 cycles of 94°C for 25 sec/65°C for 6 min) was performed with the adapter primer 1 (AP1) and a gene-specific, antisense primer (5′ -TTGATCCCTGATC TCCTTAATGCTTTC- 3′). AP2 primer and the nested antisense, gene-specific primer (5′ -CGCTAAGCAGTCGCTGGATATTATACA- 3′) were used in the subsequent PCR reaction. The PCR product was then ligated to pCRII vector and subjected to DNA sequencing analysis.

Potential *cis*-regulatory elements in the putative *CmCatB* promoter region were determined using the TFSEARCH ver.1.3 program (http://www.cbrc.jp/htbin/nphtfse-arch).

4.4.4. Construction of CAT reporter plasmids

The DNA sequence flanking the 5′ end of the *CmCatB* transcription initiation site was PCR amplified (95°C for 30 sec, 68°C for 1 min for 35 cycles) using the following oligonucleotide primers: (1) sense 5′ -CGTACCTGCAGGGCTAATAGTTGCATAAG AGCAAG- 3′; (2) antisense 5′ -GGCCTGTCGACTCGCAGAATATTGCAGAATTAT

AT- 3′. The PCR product, restricted with *Pst*I and *Sal*I (underlined) was subcloned into pAc3075, a vector that harbors the chloramphenicol acetyltransferase (CAT) reporter gene [10]. The resulting construct pAc-CatB/CAT was sequenced.

To construct pAc-CatBΔCOUP/CAT that lacks the 26 bp *COUP* site, the following oligonucleotide primers were designed: (3) antisense 5′ - AAGGTAAGGTC AAAACCATAAAAATGAATTTCGTATTT- 3′; (4) sense 5′ -AATTCATTTTTATG GTTTTCACCTTACCTTTG GATAT- 3′. The underlined nucleotides indicated the *COUP* deletion site. Two primary PCR reactions (95°C for 30 sec, 68°C for 1 min for 35 cycles) were performed with primers (1) and (3), as well as primers (2) and (4). Equal amounts of purified PCR products (QIAquick PCR Purification kit, QIAGEN, Valencia, CA) were mixed and subjected to the secondary PCR (95°C for 30 sec, 68°C for 1 min for 35 cycles) with primers 1 and 2. The PCR fragment was then subcloned into *Pst*I and *Sal*I sites of pAc3075, and the deletion of 26 bp COUP site was confirmed by DNA sequencing.

4.4.5. Transient transfection assays

Drosophila Schneider 2 (S2) cells were routinely maintained in Shields and Sang M3 insect medium (Sigma, St. Louis, MO), supplemented with 0.1 % (w/v) yeast extract, 0.25% (w/v) bactopeptone, 12.5% heat-inactivated fetal calf serum (FBS), penicillin (50 U/ml), streptomycin (50 μg/ml) and fungizone (0.25 μg/ml) at 27°C. For transfection experiments, the cells were seeded at a density of 1 x 10^6 cells per well on a six-well titer

plate, and allowed to attach for 1 h. The medium was then replaced twice with 2 mL fresh incomplete medium free of supplements, each for 15 min. Transfection was performed, in complete medium, by the calcium phosphate precipitation method (Graham & van der Eb, 1973; Wigler et al., 1977). Briefly, 8 μg of reporter plasmids were diluted in 254 μl of HEPES (*N*-2-hydroxyethylpiperazine-*N′*-2-ethanesulfonic acid)-buffered saline (26 mM HEPES, 0.78 mM Na_2HPO_4, 146.6 mM NaCl, pH 7.1) containing 135 mM $CaCl_2$, and incubated at room temperature for 30 min. Calcium phosphate precipitates formed due to the $CaCl_2$ added in the DNA tube and the phosphate in the HEPES-buffered saline. The mixture was then added dropwise to the attached cells. After 18 h incubation at 27°C, the transfection mixture was removed and replaced with complete medium. The control construct pAc5.1/V5-His/lacZ (1 μg, Invitrogen) was always cotransfected with all CAT constructs. Cells were then harvested 24 h post-transfection for CAT and β-galactosidase assays.

To measure CAT activity, the harvested cells were resuspended in 200 μl of PBS, broken by three cycles of freezing and thawing. After centrifugation at 13,000 rpm for 1 min, 30 μl of extracts were mixed with 20 μl of 100 mM Tris-HCl (pH 7.9) and heated at 65°C for 15 min to inactivate endogenous deacetylase activity. The extracts were then incubated with 200 μl of solution containing 100 mM Tris-HCl, 1 mM chloramphenicol, 0.1 μCi of ^{3}H-acetyl coenzyme A (pH 7.9) as well as 5 ml of a Betacount LSC cocktail (J.T. Baker, Phillipsburg, NJ) at 37°C for 0, 30, 60 or 90 min respectively. Enzymatic activity was measured by production of ^{3}H-acytylated

chloramphenicol using a Beckman LS 5000TD scintillation counter. Transfection assays were carried out in triplicate.

To normalize transfection efficiency, β-galactosidase activity was evaluated by measuring hydrolysis of the chromogenic substrate, o-nitrophenyl-β-D-galactopyranoside (ONPG) (Sigma). Cell extracts (10 μl) were incubated with 200 μl of 4 mg/mL of ONPG and 1 ml of Z buffer (100 mM sodium phosphate, 10 mM KCl, 1 mM $MgSO_4$, pH 7.0) containing 38.61 mM β-mercaptoethanol for 10 min at 37°C. Enzymatic reactions were terminated by addition of 0.5 ml of 1 M Na_2CO_3. Absorbance at 420 nm of this mixture was measured using a Beckman DU 64 spectrophotometer. Absorbance of the reaction mix without added cell extract was used calibrate the machine. Specific activity of β-galactosidase was defined as the amount of cell extract that hydrolyzed 1 nmol of ONPG to o-nitrophenol and D-galactose per min.

4.4.6. Electrophoretic mobility shift assays (EMSAs)

Insect gut nuclear extracts from gut walls of both scN-adapted and –unadapted 4th instar bruchid larvae were obtained using a Nuclear Extract kit (Active Motif, Carlsbad, CA). Gut nuclear extracts were aliquoted and stored at – 70°C until used for EMSA. Primers used for EMSA are listed in Table 4.1.

DNA probes larger than 60 bp were radiolabeled by PCR amplification (94°C for 30 sec, 55°C for 30 sec, 72°C for 30 sec for 25 cycles) containing 1.5 μM of [α-^{32}P] dCTP. Primers 1 and 2 were used for amplification of Probe 1 (P1, spanning -493 to -

Table 4.1. Oligonucleotide primers synthesized for EMSAs

Primer	Sequence
1	$G_{(-493)}$GCTAATAGTTGCATAAGAGCAAG$_{(-470)}$
2	$A_{(-244)}$AAAGACGTATTCCCGTGTTAGT$_{(-266)}$
3	$A_{(-302)}$CACTGGAGAAAGGGAACAGG$_{(-282)}$
4	$C_{(-41)}$GCCTCTAATCACTTATCAGTATTCG$_{(-66)}$
5	$C_{(-339)}$CAAAGGTAAGGTCAAAAGGTC$_{(-360)}$
6	$C_{(-400)}$GAAATTCATTTTTATGGTGACC$_{(-378)}$
7	$A_{(-319)}$GGGGAATAAGTCCAAATATCCAA$_{(-342)}$
8	$G_{(-378)}$GTCACCATAAAAATGAATTTCG$_{(-400)}$
9	$T_{(-382)}$GACCTTTGGACCTTACCTTTGGACC$_{(-357)}$
10	$G_{(-357)}$GTCCAAAGGTAAGGTCCAAAGGTCA$_{(-382)}$
11	$G_{(-360)}$ACCTTTTGACCTTACCTTTGG$_{(-339)}$
12	GCGCG $T_{(-382)}$GACCTTTGGACCTTACCTTTGGACC$_{(-357)}$GGCGG
13	CCGCC$G_{(-357)}$GTCCAAAGGTAAGGTCCAAAGGTCA$_{(-382)}$CGCGC
14	GCGCG$T_{(-382)}$GAgCTTTGGACCTTACCTTTGGACC$_{(-357)}$GGCGG
15	CCGCC$G_{(-357)}$GTCCAAAGGTAAGGTCCAAAGcTCA$_{(-382)}$CGCGC
16	GCGCG$T_{(-382)}$GACCTTTGGACCTTAgCTTTGGACC$_{(-357)}$GGCGG
17	CCGCC$G_{(-357)}$GTCCAAAGcTAAGGTCCAAAGGTCA$_{(-382)}$CGCGC
18	GCGCG$T_{(-382)}$GAgCTTTGGACCTTAgCTTTGGACC$_{(-357)}$GGCGG
19	CCGCC$G_{(-357)}$GTCCAAAGcTAAGGTCCAAAGcTCA$_{(-382)}$CGCGC

244), primers 3 and 4 for amplification of Probe 2 (P2, spanning -302 to -41), primers 1 and 5 for amplification of Probe 3 (P3, spanning -493 to -339), primers 2 and 6 for amplification of Probe 4 (P4, spanning -400 to -244), and primers 6 and 7 for amplification of Probe 5 (P5, spanning -400 to -319). The PCR fragments were purified

with a QIAquick PCR Purification kit (QIAGEN). See Results section for probe designs.

For probes less than 60 bp, two complementary oligonucleotides were synthesized (Table 4.1). Probes 6, 7 and 8 covering CdxA, COUP and CRE-BP putative cis-elements were formed by annealing primers 6 and 8, 9 and 10, 5 and 11, respectively. The oligonucleotides were end-labeled separately with 0.73 μM of [γ-^{32}P]ATP using T4 DNA polynucleotide kinase, and then mixed in complementary pairs (0.35 μM of each). The oligonucleotides were annealed by incubation in TE buffer plus 100 mM NaCl at 65°C for 15 min, followed by gradual cooling to room temperature. After annealing, the double-stranded oligonucleotide probes were purified with QIAquick Nucleotide Removal kit (QIAGEN). Extra sequences were added to the end of the shortest probe P7 (underlined in Table 1), designated as Pcoup (primers 12 and 13), to ensure that the probes were double-stranded. Mutated Pcoup probes are named M1 (primers 14 and 15), M2 (primers 16 and 17) and M3 (primers 18 and 19).

EMSAs were performed by incubating 3 μg of gut nuclear extract, or 5 μl *in vitro* translated proteins, CmSvp or luciferase control (see below), for 20 min with labeled probes (20,000 cpm per reaction) in binding buffer [4% glycerol, 1 mM $MgCl_2$, 0.5 mM EDTA, 0.5 mM DTT, 10 mM Tris-HCl, pH7.5, 0.05 μg of poly(dI-dC)•poly(dI-dC)] at room temperature. Samples were resolved on 4% native polyacrylamide gel, followed by X-ray film exposure.

For competition assays, 5, 10 or 50-fold molar excess of specific or nonspecific competitors were incubated with nuclear extract for 20 min at room temperature prior to

the addition of probe. Nonspecific DNAs were prepared by PCR amplification of the coding regions of *CmCPA9* or *CmCPB1* genes. The sizes of the nonspecific DNAs were comparable to their corresponding competing probes. Pre-incubation of nuclear extract with polyclonal anti-AaSvp antibody raised against a highly conserved region of mosquito COUP-TF (kindly provided by Dr. Alexander Raikhel at the University of California) was also performed. One μl of pre-immune or anti-AaSvp serum, respectively, was used.

4.4.7. Cloning of *CmSvp* from cowpea bruchid midguts

Guts from scN-unadapted cowpea bruchid 4th instar larvae were used for total RNA extraction with the TRIzol Reagent (phenol and guanidine-isothiocyanate, Invitrogen). Two pairs of degenerate primers were designed based on highly conserved DNA-binding (DBD) and ligand-binding domains (LBD) of other COUP-TF family members: (1) sense 5′- AARCACTA YGGHCARTTYAC -3′; (2) antisense 5′- CADAT GTTSTCRATVCCCAT -3′; (3) sense 5′- TTYACBTGCGARGGNTGCAA -3′ and (4) antisense 5′- CCCATVATGTTGTTVGGYTGC -3′ (R = A,G; Y = C,T; H = A,C,T; D = A,G,T; S = C,G; V = A,C,G; B = C,G,T; N = A,C,G,T). The primary PCR (95°C for 30 sec, 55°C for 30 sec, 72°C for 1 min for 35 cycles) was conducted with primers 1 and 2 using a gene-specifically primed first strand gut cDNA mixture. The nested PCR (95°C for 30 sec, 42°C for 1 min, 72°C for 1 min for 35 cycles) then was performed with primers 3 and 4. The PCR fragment was subcloned into pCRII vector and sequenced.

The 5′ and 3′ cDNA ends were PCR amplified (94°C for 30 sec, 68°C or 60°C for 30 sec, 72°C for 1 or 2 min for 35 cycles) using the BD SMART RACE cDNA Amplification kit with the following gene-specific primers: (5) sense 5′- GTAACCACA CCTACCTCAGCAGCT -3′; (6) sense 5′- AGCTACATATCCTTGCTGTTGAG -3′; (7) antisense 5′- GGCTCTGCCCTCAA CAGCAAGGATATG -3′ and (8) antisense 5′- GA TATGTAGCTGCTGAGGTAGGTGTGG -3′. Primers 5 and 6 were used for PCR and nested PCR for 3′ RACE, and primers 7 and 8 for 5′ RACE. The PCR fragments were subcloned into pCRII vector and sequenced. The full-length coding region was obtained by RT-PCR (95°C for 30 s, 68°C for 2 min for 35 cycles) using the following primers: sense 5′- GAGTAGAGGCAAGAGTGTGTCCCTGG -3′, and antisense 5′- GTGAGGT AGTGAGTTGAGTCAGTTAGTTTCAGA -3′. It was then subcloned into pCRII vector, and the sequence was confirmed.

4.4.8. Western blot analyses

For western blotting, polyclonal anti-AaSvp antibody (kind gift from Dr. Alexander Raikhel) was used to detect the differential protein levels in the scN-adapted and –unadapted cowpea bruchid fourth instar guts. Eight μg of midgut nuclear extract protein from adapted or unadapted insects was first resolved on 12.5% SDS-PAGE, then transferred to a nitrocellulose membrane. The polyclonal chicken anti-AaSvp antibody was used as primary antibody at a 1:50 dilution. The secondary antibody was rabbit anti-chicken IgG conjugated with horseradish peroxidase (Sigma), and used 1: 160,000

dilution. Antigen-antibody complexes were detected using ECL Western Blotting Detection Reagents (Amersham Biosciences). To ensure equivalent protein loading of midgut nuclear extracts, the blot was reprobed with rabbit anti-actin primary antibody (1: 500 dilution, Sigma) and the secondary goat anti-rabbit IgG (H+L) antibody (1: 10,000 dilution) conjugated with horseradish peroxidase (Kirkegaard Perry Laboratories).

4.4.9. *In vitro* translation of CmSvp

CmSvp protein was produced by coupled *in vitro* transcription and translation. First, the plasmid pCRII-CmSvp was digested with *EcoR*V and *Sac*I restriction enzymes, and subcloned into pBluescript II-KS vector under the control of the T7 RNA polymerase promoter (Stratagene, La Jolla, CA). The TNT T7 Coupled Reticulocyte Lysate System (Promega, Madison, WI) was then used to produce CmSvp protein. The TNT reaction was incubated at 30°C for 90 min using 1 μg of DNA. From this reaction, 5 μl of protein was used for EMSA and competition assays as described above to evaluate DNA-binding specificity of CmSvp. Luciferase was used as a control for *in vitro* translation as well as for the EMSAs.

4.4.10. Cotransfection

To evaluate the effect of CmSvp on *CmCatB* expression, the expression construct pAc5-CmSvp was cotransfected with the reporter plasmid pAc-CatB/CAT into S2 cells. To construct pAc5-CmSvp, the entire coding region of *CmSvp* was amplified by PCR (94°C for 30 sec, 68°C for 2 min for 35 cycles) using the following oligonucleotide primers: sense 5′ -AAGCTGATATCGGTACCATGGCACTTGTGG- 3′, and antisense 5′ -GCTCATCTAGACATATACGGCCACGAGAATGAACT- 3′. *EcoR*V and *Xba*I (underlined) restriction sites were incorporated into primers for directional cloning. After restriction digestion, the PCR fragment was ligated to pAc5.1/V5-HisA vector (Invitrogen) and correct DNA sequence was verified. One μg of pAc5-CmSvp or pAc5.1/V5-HisA vector alone was cotransfected with the reporter plasmid, the latter ensuring comparable total DNA amounts in CmSvp-expressing and non-expressing cells. Cells were collected at 24 h post-transfection, and used for CAT activity assay. The reporter plasmid pAc-IE1/CAT (*CAT* gene placed under control of the promoter of a baculovirus immediate-early gene) (Guarino & Dong, 1991), was used for evaluation of specific interaction of CmSvp and *CmCatB* promoter. To test whether *COUP* binding is necessary for CmSvp regulation, construct pAc-CatBΔCOUP/ CAT, with the 26 bp *COUP* site removed, was also cotransfected with pAc5-CmSvp.

CHAPTER V

ANTAGONISM BETWEEN HEPATOCYTE NULEAR FACTOR 4 AND SEVEN-UP MODULATES INSECT COUNTER-DEFENSE CATHEPSIN B EXPRESSION

5.1. Introduction

Cowpea bruchid beetle, *Callosobruchus maculatus*, continuously encounters protease inhibitors that are produced by its host plants as defense against insect attack. To overcome the inhibitory effects of soybean cysteine protease inhibitor (scN), cowpea bruchids differentially regulate expression of their major digestive enzymes, cysteine proteases, by utilizing three mechanisms (Zhu-Salzman et al., 2003). First, Cowpea bruchids enhance proteolytic capability through quantitative overexpression of scN-sensitive cysteine poteases to outnumber scN. Second, cowpea bruchids induce scN-degrading cysteine proteases, which result in inactivation of scN. Third, cowpea bruchids induce scN-insensitive cysteine poteases leading to increase proteolytic capability.

Cowpea bruchids activate transcription of cathepsin L-like cysteine proteases (*CmCPs*) to encounter scN inhibition (Zhu-Salzman et al., 2003). *CmCPs* were grouped into subfamilies A and B based on protein sequence similarity. Cowpea bruchids quantitatively overexpressed scN-sensitive CmCPs or induced scN-degrading CmCP from subfamily B under scN challenge (Zhu-Salzman et al., 2003; Ahn et al., 2004).

When cowpea bruchids faced nutrient deficiency due to inhibition of their major digestive proteases such as CmCPs by scN, they induced very highly *CmCatB* expression, which was not detectable in cowpea bruchid reared on scN-free diet (Moon et al., 2004). We postulated that CmCatB is less susceptible to scN inhibition than CmCPs due to an occluding loop, which is a unique protein structure of cathepsin B. The occluding loop of human cathepsin B ortholog has been shown to prevent substrates and inhibitors from binding to the catalytic pocket (Musil et al., 1991; Illy et al., 1997). The Zhu-Salzman's lab recently confirmed that CmCatB is insensitive to scN inhibition through heterologous CmCatB protein expression in yeast (Koo et al., 2008).

Above findings suggest that both CmCPs and CmCatB are involved in bruchids adaptation to scN. Our knowledge, however, is lacking in understanding how insects are able to sense plant inhibitors in their food and the regulatory mechanisms of differential expression of CmCPs or transcription to coordinate counter-defense genes such as *CmCatB* to avoid the negative effect of plant inhibitors. To decipher the regulatory mechanisms behind insect adaptation at the protein-level, I investigated the functional significance of the differential regulation of CmCPs through expression of proprotein CmCPs (proCmCPs) in *E. coli* (Ahn et al., 2004; Ahn et al., 2007b). Cowpea bruchids selectively induced CmCPs from subfamily B for their superior protease activities that were regulated by controlling cleavage and stability of propeptides in bruchids adaptation to scN. To deeply understand the regulatory mechanisms behind insect adaptaion at the gene level, I studied the transcriptional regulation of *CmCatB,* which encodes an scN-insensitive protease and is highly responsive to dietary scN. I

previously identified that a regulatory chicken ovalbumin upstream promoter (*COUP*) element in the *CmCatB* promoter, negatively regulated *CmCatB* expression in unchallenged insect guts by CmSvp (Ahn et al., 2007a; see Chapter IV).

In this study, I investigated the transcriptional activation of *CmCatB* under challenge with the dietary scN to understand how upregulation of *CmCatB* helps cowpea bruchid to overcome the negative effect of plant inhibitors. I cloned a cowpea bruchid homolog to hepatocyte nuclear factor 4 (CmHNF-4), a nuclear receptor superfamily member that significantly activated *CmCatB* expression in transefected *Drosophil*a S2 cells. Furthermore, when CmHNF-4 transfected with CmSvp together into S2 cells, CmSvp significantly repressed CmHNF-4-mediated transactivation of *CmCatB* expression. These findings suggest that the expression of counter-defensive *CmCatB* is partly dependent on balance of the antagonistic effect between CmHNF-4 and CmSvp regulatory factors.

5.2. Results

5.2.1. Defining the *CmCatB* promoter *cis*-element

My previous EMSAs revealed differential mobility shifts in scN-adapted and –unadapted midgut nuclear extract, suggesting the existence of two different *CmCatB* promoter-interacting nuclear proteins. While the factor specific to scN-unadapted insects, i.e. CmSvp, has been identified (Ahn et al., 2007a), the one unique to scN-

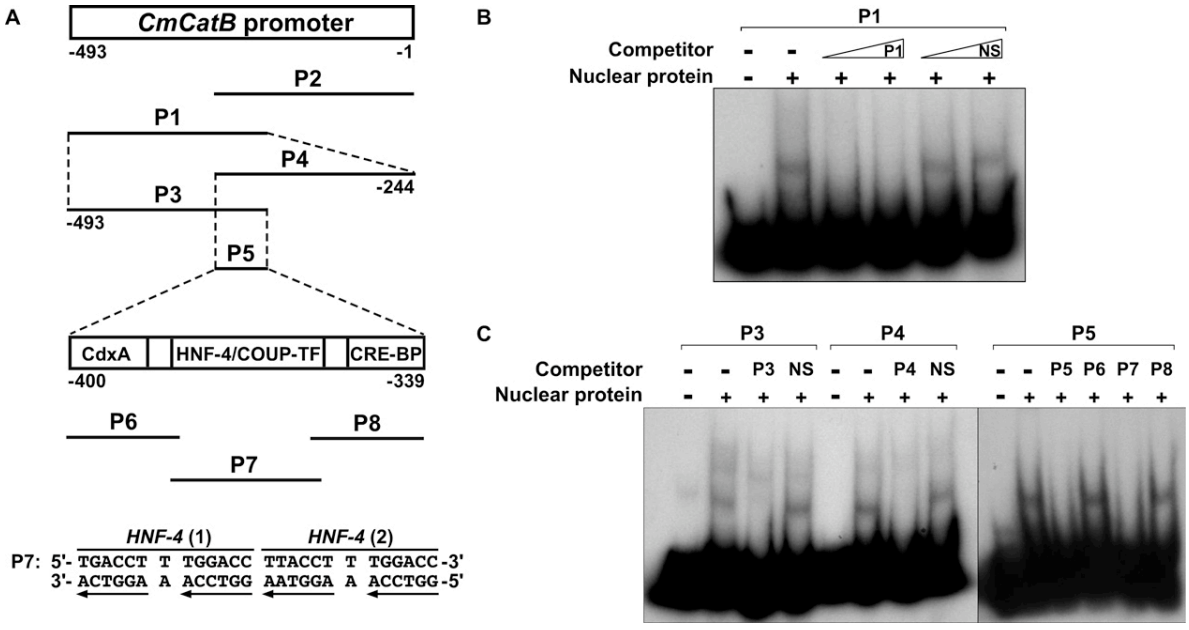

Figure 5.1. Defining regulatory *cis*-element(s) in *CmCatB* promoter that specifically interact with scN-adapted insect midgut nuclear protein(s). (A) Probe dissection for EMSA. P7 has two tandem HNF-4 binding sites. The arrows mark the half-sites and indicate their orientations. (B), (C) Tracking down a putative *cis*-element in P7 via EMSA and competition assays. Nuclear extracts were obtained from freshly dissected scN-adapted 4th instar larvae midguts. In competition assays, 5- or 10-fold molar excess of unlabeled specific or nonspecific competitors were preincubated with gut nuclear extract prior to the binding reaction. P: probe. NS: nonspecific DNA. CdxA, HNF-4/COUP-TF, and CRE-BP: putative *cis*-elements.

adapted insects remained to be elucidated. Like CmSvp, the nuclear protein from adapted insect guts specifically interacted with the probe P1, which spans –493 to –244 relative to transcription start (Figs. 5.1A and 5.1B), but did not bind the more proximal probe P2 (data not shown). Since the DNA-protein interaction occurred in adapted insects where *CmCatB* was highly induced, we hypothesized that this nuclear protein could serve as an activator that enhanced *CmCatB* expression through binding to target promoter elements.

To locate the putative *cis*-regulatory element, I performed a series of gel shift and competition assays (Figs. 5.1B and 5.1C). The overlapping probes P3 and P4, covering P1 region, both exhibited specific associations with nuclear extracts prepared from guts of adapted insects. These results suggested that the potential protein-binding region resides in the overlap between P3 and P4. Subsequent detection of protein-binding complex using P5 probe, roughly corresponding to the region common to P3 and P4, confirmed this idea. The specificity of the DNA-protein interaction was further verified by competition assays using unlabeled P5 or nonspecific DNA competitors.

I then used TFSEARCH (http://www.cbrc.jp/htbin/nphtfsearch), a *cis*-element predicting program to search for putative *cis*-elements and identified motifs corresponding to CdxA, HNF-4/COUP-TF and CRE-BP in P5 (Fig. 5.1A). P5 was then further refined into three probes P6, P7 or P8, each corresponding to one *cis*-element.

When applied to competition EMSAs, only probe P7 could compete with P5 for protein binding (Fig. 5.1C). This same *cis*-element was previously shown to mediate interactions between *CmCatB* promoter and CmSvp, a repressor from unadapted insect midguts. To confirm that the adapted nuclear extracts contained a factor that was distinct from CmSvp previously identified, I used antiserum raised against the *Aedes aegypti* homolog of CmSvp, which I previously showed interfered with DNA-protein interactions between P7 and CmSvp (Ahn et al., 2007a). The same serum was unable to interfere with complex formation between P7 and the adapted nuclear protein, nor did it cause a supershift (Fig. 5.2), strongly suggesting the presence of a novel.

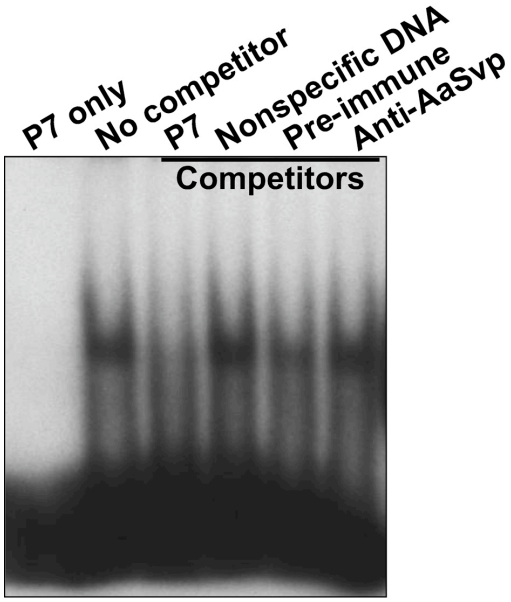

Figure 5.2. A nuclear protein other than CmSvp interacts with *HNF-4/COUP* element in nuclear extract of scN-adapted insect midgut. AaSvp: mosquito *Aedes aegypti* COUP-TF homolog. Anti-AaSvp: polyclonal antibody raised against a portion of the DNA and ligand-binding domains of mosquito AaSvp which cross-reacts with CmSvp. Antibody was preincubated with nuclear extract prior to the binding reactions with P7. DNA competition was performed by adding 10-fold molar excess of unlabeled DNA competitors to labeled P7.

5.2.2. HNF-4 is the potential DNA-binding protein in scN-adapted insect guts

Above results implied that the *CmCatB* transcript is regulated by different transcription factors binding to the identical *cis*-regulatory sequence. Indeed, the P7 probe contains two sets of imperfect AGGTCA direct repeats, a response element that can be targeted by receptors other than COUP-TF/Svp (Fig. 5.1A). Several features of the sequence suggested that it was an ideal binding site for hepatocyte nuclear factor 4 (HNF-4):

(1) the half-sites are separated by one nucleotide, (2) three consecutive adenosines are found in the center of the site, (3) the second positions of both half-sites maintain highly conserved a guanosine (Sladek and Seidel, 2001).

HNF-4 is a member of the steroid receptor superfamily initially identified in rat liver (Costa et al., 1989; Sladek et al., 1990), and was later cloned from a number of insects (Zhong et al., 1993; Kapitskaya et al., 1998; Swevers & Iatrou, 1998). It plays a critical role in nutrition and metabolism throughout animal development (Kapitskaya et al., 1998; Sladek & Seidel, 2001). Targeted disruption of the HNF-4 gene in mouse resulted in cell death in the embryonic ectoderm and impaired gastrulation. Thus HNF-4 also plays an important role in early embryogenesis (Chen et al., 1994; Duncan et al., 1997). Therefore we postulated that HNF-4 in cowpea bruchids activated *CmCatB* in insect midguts when insects encountered dietary protease inhibitors.

In order to test this hypothesis and illustrate the potential function of HNF-4 in trans-activating *CmCatB*, I cloned the putative transcription factor from scN-adapted cowpea bruchid midgut by a combination of PCR and RACE. PCR using degenerate primers, derived from ligand binding domain of HNF-4, which is conserved among vertebrate and invertebrate species, produced a 486 bp fragment. After DNA sequencing to confirm that the fragment corresponded to an HNF-4 family member, gene-specific primers for 5′ and 3′ RACE was synthesized. Finally, a 1,786 bp full-length cDNA sequence, named *CmHNF-4*, was amplified by PCR.

The translation of the cDNA sequence showed that *CmHNF-4* encodes a protein of 507 amino acid residues (Fig. 5.3A). The nuclear receptor superfamily is defined by

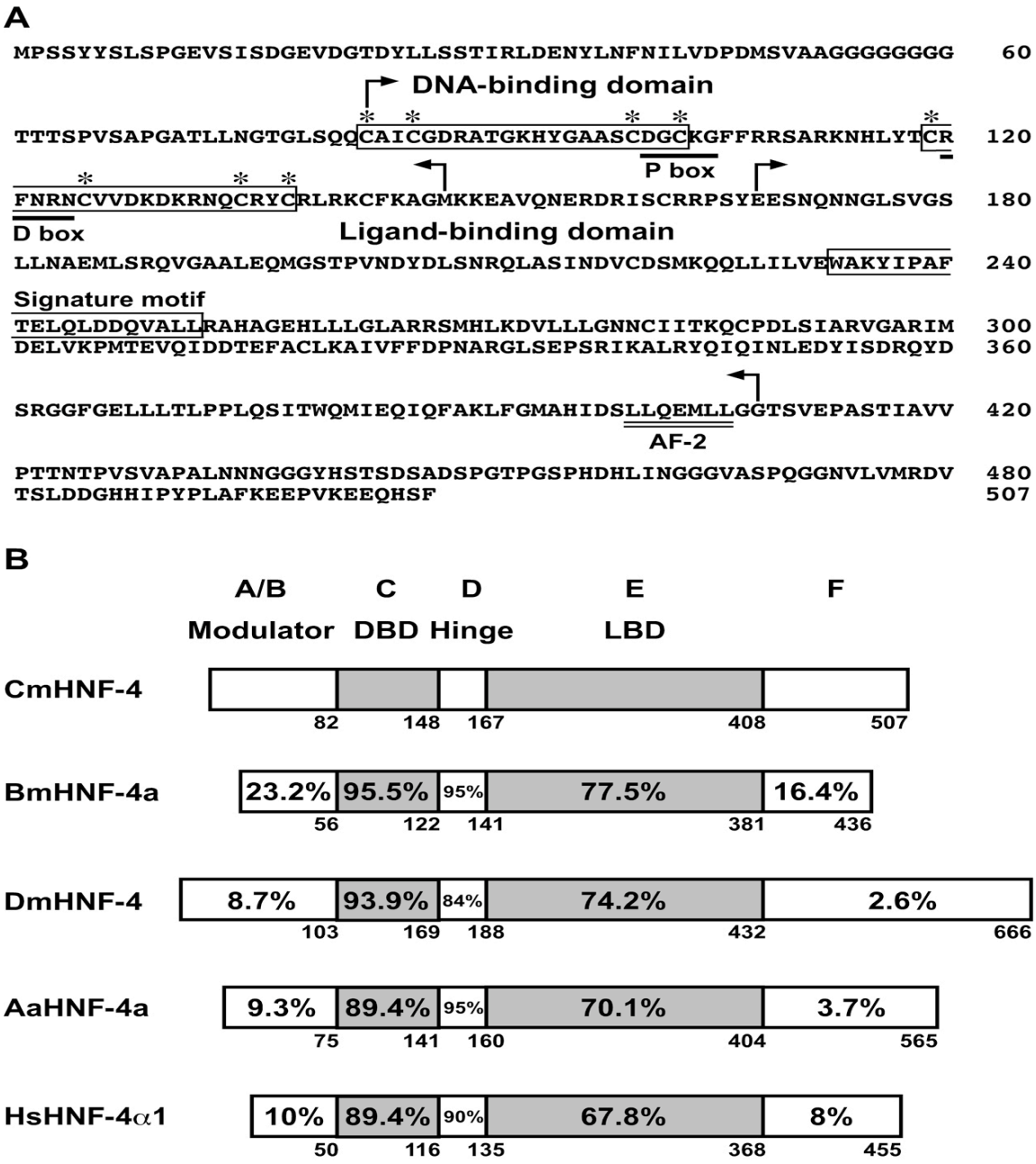

Figure 5.3. Conservation of cowpea bruchid HNF-4 protein sequence. (A) Amino acid sequence deduced from a full-length *CmHNF-4* cDNA (GENBANK accession number EU545256). The boundaries of the putative DNA-binding (DBD) and ligand-binding domains (LBD) are indicated by bent arrows. The two zinc finger modules in the DBD are boxed and its eight highly conserved cysteine residues are marked by asterisks. The P and D boxes in the DBD are underlined. The signature motif of the LBD of nuclear receptors is also boxed. The activation function-2 (AF-2) localized at the carboxyl-terminal end of the LBD is double-underlined. (B) CmHNF4 shares high protein sequence and domain structural similarity with HNF-4 members of silk moth *Bombyx mori* BmHNF-4a, *Drosophila melanogaster* DmHNF-4, mosquito *Aedes aegypti* AaHNF-4a and human *Homo sapiens* HsHNF-4α1. Letters above the CmHNF4 protein indicate conventional functional domains. Percentages indicate levels of identical amino acid residues shared between CmHNF4 and respective HNF4 proteins in functional domains. The start and end amino acids of the DBD and the LBD are numbered.

the presence of common structural elements. A highly conserved DNA-binding domain (DBD) and a less conserved C-terminal ligand-binding domain (LBD) are joined by a flexible hinge region. The DBD is rich in cysteines and basic amino acid residues. The positions of the cysteines are highly conserved in nearly all members of the nuclear receptor superfamily, and form two zinc-finger modules of the sequence, $CX_2CX_{13}C$ $X_2CX_{15}CX_5CX_9CX_2C$ (Mader et al., 1993). The first zinc finger provides DNA-binding specificity through the P (proximal)-box that recognizes the half-site DNA binding motif (Umesono & Evans, 1989). The D (distal)-box in the second zinc finger has a role in discrimination of the spacing between the half-sites of the DNA binding sequence through protein-protein interaction. The LBD can be defined by a signature motif that consists of the 20 amino-acid motif (F,W,Y)(A,S,I) (K,R,E,G)xxxx (F,L)xx(L,V,I)xxx (D,S)(Q,K)xx(L,V)(L,I,F) for nuclear hormone receptors (Wurtz et al., 1996). The LBD is a multifunctional domain that mediates ligand binding, dimerization, transactivation, and nuclear localization functions. The activation function-2 (AF-2), essential for ligand-dependant transactivation, has also identified at the C-terminus of the LBD. Its structural domains are also similarly organized as its homologues from silk moth *Bombyx mori* (Swevers & Iatrou, 1998), *Drosophila* (Zhong et al., 1993), mosquito (Kapitskaya et al., 1998), as well as human (Chartier et al., 1994) (Fig. 5.3B). The DBD shows the highest identity among different species (89.4-96%). The LBD is also highly conserved among homologues (68-78%), and very high identities are found in the hinge region (84.2-95%). Like most nuclear receptors, the potential transacting A/B (N-terminus) and F (C-terminus) domains are highly variable. Among HNF-4 homologs,

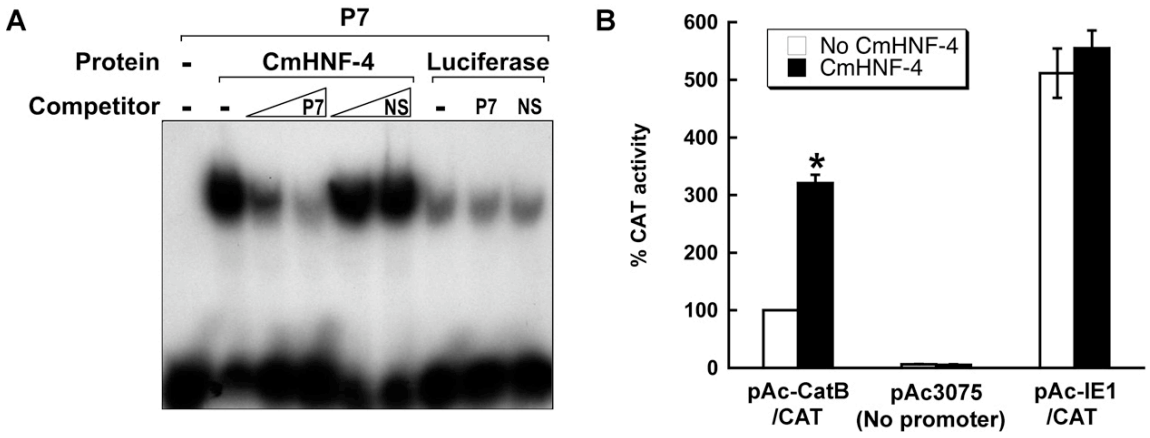

Figure 5.4. CmHNF-4 is a functional DNA binding protein that activated *CmCatB* expression. (A) *In vitro* translated CmHNF-4 specifically interacted with the *HNF-4* responsive element. Competition analysis was performed by adding 10- or 50-fold molar excess of unlabeled P7 or nonspecific DNA (NS), respectively. Luciferase, which was synthesized in the same in vitro translation system, was used as a negative control. (B) Activation of *CmCatB* expression by CmHNF-4. S2 cells were cotransfected with 8 μg of reporter plasmids and 0.5 μg of CmHNF-4 expression plasmid (black bars) or equivalent empty expression vector (white bars). The reporter plasmid pAc-IE1/CAT was used as a control for specificity of CmHNF-4 effect on *CmCatB* promoter. Transfection efficiency was normalized by β-galactosidase internal standard. Each bar represents the mean ± S.E. of three independent cotransfections. CAT activity driven by *CmCatB* promoter was significantly different (*t*-test, $P < 0.01$) in the presence vs. absence of CmHNF-4 expression plasmid, while no statistical difference was shown in non-promoter control or in the reporter plasmid pAc-IE1/CAT that lacks HNF-4 binding site.

the A/B domain shares only 9-23.2% identity, while the F domain is the most variable (3-16.4% identity).

To determine if CmHNF-4 was capable of DNA binding, the DNA was subcloned into a transcription vector, transcribed and translated in a rabbit reticulocyte lysate system. EMSA was performed using the P7 probe that contains the putative HNF-4-binding sites. The TNT expressed protein caused a band shift in EMSA. Competition

assays confirmed binding specificity of the *in vitro* translated protein to the P7 probe (Fig. 5.4A).

5.2.3. CmHNF-4 is a potential positive transcription regulator in scN-adapted insect guts

HNF-4 generally acts as a positive transcriptional regulator of its target genes (Ktistaki & Talianidis, 1997; Mcnair et al., 2000; Bartoov-Shifman et al., 2002; Watt et al., 2003). To test the role of CmHNF-4 in *CmCatB* activation, I utilized a transient transfection system, where the reporter gene chloramphenicol acetyltransferase (*CAT*) was cloned under the control of *CmCatB* promoter (Ahn et al., 2007a). *Drosophila* S2 cells were co-transfected with the reporter plasmid (pAc-CatB/CAT) and the CmHNF-4-expression plasmid (pAc5-CmHNF-4) and assayed for CAT activity. Transient expression of CmHNF-4 repeatedly caused 3 to 4-fold increase in reporter gene activity (Fig. 5.4B). This supports our hypothesis that CmHNF-4 promotes transcriptional activation of *CmCatB*.

Previous studies on human HNF-4 demonstrated that the transcription activation potential of three different HNF-4 isoforms was strictly dependent on the presence of HNF-4 binding site (Drewes et al., 1996). To investigate whether CmHNF-4 also required binding to the *HNF-4/COUP* element in the *CmCatB* promoter, cotransfections were performed with the *cis*-element was removed from the reporter plasmid pAc-

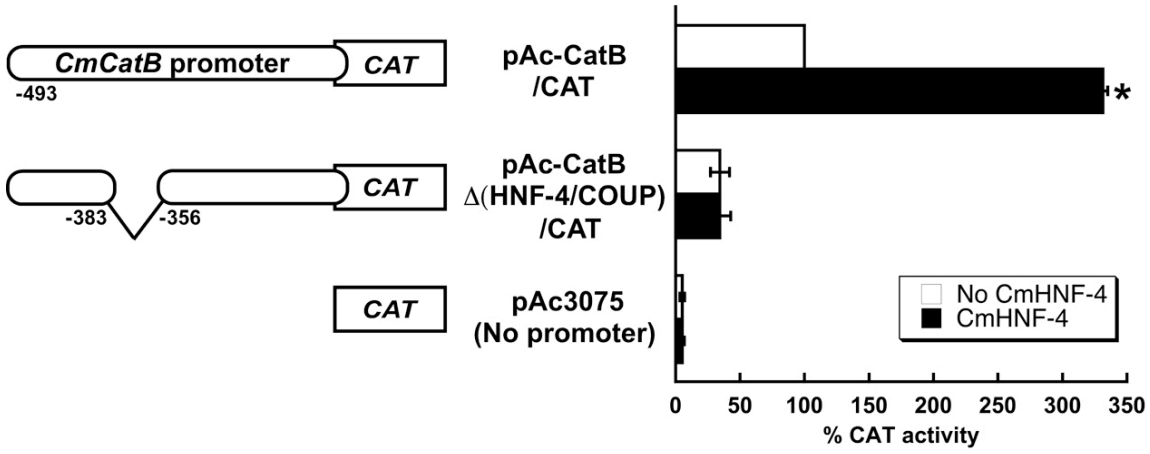

Figure 5.5. CmHNF-4 activation of *CmCatB* requires binding at the *HNF-4/COUP* element. Shown are schematic promoter structures in reporter plasmids, pAc-CatB/CAT, pAc-CatBΔ(HNF4/COUP)/CAT and pAc3075 that has no promoter. Numbering is relative to the transcription start point. Reporter plasmids (8 μg) were cotransfected with 0.5 μg of CmHNF-4-expressing plasmid (black bar) or 0.5 μg of non-expressing empty vector (white bar), respectively. The latter was to ensure comparable total DNA amounts in transfected S2 cells. Transfection and normalization of CAT activity was performed as above. Each bar represents the mean ± S.E. of four independent transfections. Only pAc-CatB/CAT report plasmid showed statistical significance (*t*-test, $P < 0.01$) in the presence vs. absence of CmHNF-4.

CatBΔ(HNF-4/COUP)/CAT. Quantitation revealed that the enhanced CAT activity was diminished, and also the basal activity was significantly decreased (Fig. 5.5). Thus binding to the *HNF-4*/*COUP* site appears to be necessary for CmHNF-4 positive regulatory function.

5.2.4. CmHNF4 activity is antagonized by CmSvp

Several gene promoters have been studied in which HNF-4 and COUP-TF bind to the same response element. For example, it was found that the ApoAI regulatory protein I (ARP-1) and Ear3/COUP-TF repressed HNF-4-mediated transactivation of the apolipoprotein CIII promoter through binding to their shared response element (Mietus-Snyder et al., 1992). To gain some insight into the regulatory relationship between the two transcription factors, which bind to the identical *cis*-regulatory sequence but elicit opposite effects, I transiently expressed both CmSvp and CmHNF-4 in S2 cells and evaluated the combined effect on expression from the *CmCatB* promoter. When CmHNF-4 and CmSvp were co-expressed in S2 cells, the enhancing effect of CmHNF-4 was diminished to a level nearly identical to transfected cells with no effector plasmid (Fig. 5.6). In contrast, when the IE1-CAT plasmid, which does not contain this specific *cis*-element, was co-transfected with CmSvp and CmHNF-4 expression plasmids, no significant change in CAT activity was observed. This result suggests that the two transcription factors, potentially both involved in *CmCatB* regulation, antagonizes each other's effect when co-expressed.

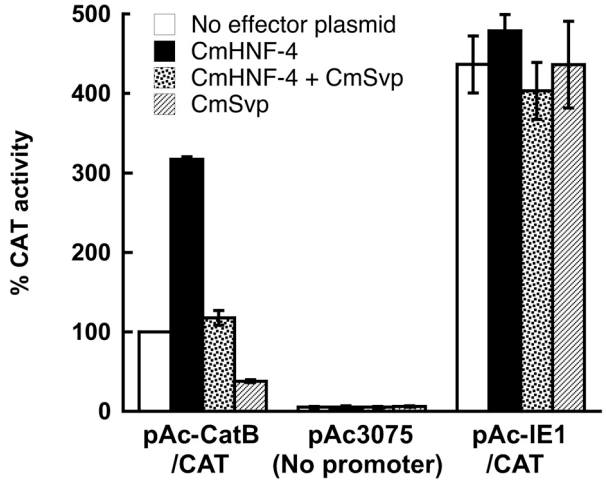

Figure 5.6. CmSvp antagonizes trans-activation activity of CmHNF-4. CmHNF-4 (0.5 μg) and/or 1 μg of CmSvp expression plasmids were cotransfected into S2 cells with 8 μg of reporter plasmids. Appropriate amount of empty expression vector was added to ensure equivalent total DNA amount transfected into S2 cells. Transfection and normalization was performed as above. Each bar represents the mean ± S.E. of three independent cotransfections. pAc-IE1/CAT was used as a control plasmid for specificity of CmHNF4 and/or CmSvp effects on *CmCatB* promoter.

5.3. Discussion

Insect adaptability to a variety of antinutritional factors or plant toxins has been a huge obstacle for insect pest management. Indeed, the biggest threat to agriculture may be limitations in our understanding of the insect pest, leading to improper strategies of pest control. Therefore, it is very important to understand the molecular bases of insect adaptive responses to plant defense. The Zhu-Salzman's lab undertook a study of the adaptive responses in cowpea bruchids to dietary plant protease inhibitor challenge and found that cowpea bruchids were able to overcome the negative effect of plant inhibitors

through transcriptional activation of counter-defense related genes including inhibitor-insensitive proteases. To understand how insects regulate counter-defense genes, I have investigated transcriptional regulation of the cathepsin B-like *CmCatB* gene that is highly responsive to a dietary soybean cysteine protease inhibitor scN, and expressed as an scN-insensitive protease (Moon et al., 2004; Koo et al., 2008). I previously showed that the cowpea bruchid COUP-TF homolog, CmSvp, repressed *CmCatB* expression through binding to the *COUP* responsive elements in unadapted insect midgut (Ahn et al., 2007a). In this study, I investigated a positive regulatory factor of *CmCatB* expression in scN-adapted insect midgut where the *CmCatB* gene was highly induced to cope with the effect of the dietary scN. I identified the *HNF-4* responsive elements in the *CmCatB* promoter, and cloned the regulatory transcription factor, CmHNF-4 from scN-adapted insect midgut. The CmHNF-4 activated *CmCatB* expression, and this transactivation required binding to *HNF-4* elements in the scN-adapted insect midgut.

Evolutionary conservation among HNF-4 tissue distribution suggests that the CmHNF-4 may be an imporatnt transcriptional regulator of digestive protease CmCatB in the midgut, as it in the vertebrate intestine. HNF-4 is predominantly expressed in liver, kidney and small intestine in vertebrates (Sladek et al., 1990; Drewes et al., 1996). In insects, it is found in fat body, Malpighian tubules and midgut, tissues with equivalent functions as their vertebrate counterparts (Zhong et al., 1994; Kapitskaya et al., 1998; Swevers & Iatrou, 1998). In vertebrates, HNF-4 expression is mainly restricted to liver, kidney, and intestine. Typically, HNF-4 expression was lacking in heart, brain, ovary, testis, and skeletal muscle (Zhong et al., 1994; Drewes et al., 1996). This postulation

was supported by CmHNF-4 expression in both scN-adapted and unadapted cowpea bruchid midguts through RT-PCR analysis (data not shown).

HNF-4 plays an important role in nutrient transports, such as cholesterol and lipids by apolipoproteins, iron by transferrin, and thyroid hormone T_4 by transthyretin (Sladek & Seidel, 2001). HNF-4 is also a crucial regulator of many essential genes involved in nutrient metabolisms including lipid and steroid by cytochrome P450 family members; glucose by phosphoenolpyruvate carboxykinase, liver-type pyruvate kinase, and adolase B; amino acids by ornithine transcarbamylase, aldehyde dehydrogenase 2, and tyrosine aminotransferase. Furthermore, HNF-4 plays a central role in xenobiotic responses by inducing transcription of cytochrome P450 gene encoding a detoxification enzyme (Li & Chiang, 2006; Lee & Lee, 2007). Among these biological functions of HNF-4 homolog, a regulatory function in nutrient metabolism implys that CmHNF-4 is involved in transcriptional regulation of *CmCatB* encoding a digestive cysteine protease.

HNF-4 was initially considered to be an orphan nuclear receptor for which a ligand has not been determined. Recently, the crystal structures of HNF4α and HNF4γ revealed that an endogenous fatty acid constitutively locked in the LBD (Dhe-Paganon et al., 2002; Wisely et al., 2002). The fatty acid seems to function more like a structural cofactor than an exchangeable ligand, as it cannot be removed under physiological conditions. This may explain why HNF-4 is able to act as a constitutively active receptor even in the absence of exogenously added ligand. However, the effect of this molecule remains to be determined for other traditional effects of ligands, such as their ability to modulate HNF-4 activity *in vivo*.

To determine how, where and when nuclear receptors are regulated by small chemical ligands and/or protein partners, the regulatory activities of the *Drosophila* nuclear receptors were investigated in live developing transgenic *Drosophila* using a 'ligand sensor' system, which was able to visulalize spatial activity patterns for nuclear receptors (Palanker et al., 2006). Activation of HNF-4 ligand sensor was restricted to the yolk during embryogenesis, and junction of the midgut, proventriculus and gastric caeca at the onset of metamorphosis. The HNF-4 ligand sensor was also active in the larval fat bodies of feeding 3rd instar larvae, but showed reduced activation at the early prepupa. These findings indicate that HNF-4 is regulated by ligands and/or protein partners in the organs, which play a role as a nutrient source such as the yolk or nutrient absorption/transport such as the gastric caeca. Also HNF-4 regulation is correlated with insect feeding, as activation of HNF-4 ligand sensor was reduced after puparation when insects stop feeding. Therefore, these findings suggest that HNF-4 acts as a metabolic sensor.

These known biological functions of HNF-4 homologs suggest a model by which CmHNF-4 regulates *CmCatB* expression in scN-adapted insect midguts. When the main digestive enzymes, such as cathepsin L-like cysteine proteases (CmCPs), are not available due to dietary scN challenge, CmHNF-4 functions as a metabolic sensor that respond to nutrient deficiency in the cowpea bruchid midgut, and activates *CmCatB* expression, which encodes an scN-insensitive protease to compensate for CmCPs inactivation. Thus, HNF-4 eventually allows cowpea bruchids to resume normal larval feeding and development.

COUP-TFs have been known to antagonize HNF-4-dependent transactivation of many liver-enriched genes, and most of the *cis*-regulatory elements initially found to bind HNF-4 are also recognized by COUP-TF. COUP-TF inhibited the HNF-4-dependent activation of rat ornithine transcarbamylase (OTC) gene encoding an ornithine cycle enzyme, and two transcription factors recognized the same sites in the OTC promoter (Kimura et al., 1993). HNF-4 is a positive regulator in the tissue-specific and hypoxia-inducible expression of human erythropoietin, which is a hormone critical for the proliferation, and differentiation of red blood cells, and COUP-TF was shown to suppress HNF-4 transactivation (Galson et al., 1995). The human cytochrome P4502D6 promoter has a regulatory element that interacts with both HNF-4 and COUP-TFI, and COUP-TFI was shown to inhibit HNF-4 stimulation of the promoter (Cairns et al., 1996). HNF-4 and COUP-TF interacted with the same promoter regions including HNF-4 responsive elements of human angiotensinogen (ANG), and HNF-4 strongly activated ANG transcription, but COUP-TF dramatically repressed ANG transcription (Yanai et al., 1999). Cotransfection assays showed that COUP-TFI suppressed the ability of HNF-4 to activate the human aldehyde dehydrogenase 2 promoter (You et al., 2002). Intriguingly, CmHNF-4 and CmSvp have opposite regulatory functions bound to identical *cis*-regulatory sequences, *HNF-4/COUP* elements in the *CmCatB* promoter. When CmHNF-4 and CmSvp were expressed together in the cells, CmSvp antagonized CmHNF-4 activation of *CmCatB* expression (Fig. 5.6). Therefore, transcriptional regulation of *CmCatB* gene through *HNF-4/COUP* elements in response to the dietary scN effect depends on the intracellular balance between positive and negative regulators.

COUP-TFs repress HNF-4 transactivation via competitive binding to same regulatory element of many promoters (Mietus-Snyder et al., 1992; Cairns et al., 1996; Yanai et al., 1999), but are also able to inhibit transactivation without binding to a response element once tethered to a promoter through heterodimerization with retinoid acid receptor (RAR), thyroid hormone receptor (TR), and retinoid X receptor (RXR) (Leng et al., 1996; Tsai &Tsai, 1997). To date, however, heterodimer formation between HNF-4 and COUP-TF has not been detected. Indeed, HNF-4 is known to bind DNA exclusively as a homodimer (Jiang et al., 1995) even though heterodimerization between different HNF-4 has been observed (Zhong et al., 1993; Kapitskaya et al., 1998). Since both the CmHNF-4 and CmSvp required DNA binding for their regulatory activities (Fig. 5.5; Ahn et al., 2007a), it may possible that the CmSvp inhibits CmHNF-4 transactivation solely through competition for DNA binding site. We can't exclude, however, the possibility that CmSvp represses CmHNF-4 transactivation via formation of heterodimer based only on above findings. Whether CmSvp inhibits CmHNF-4 transactivation through competitive binding to same regulatory elements or heterodimerization with CmHNF-4 needs further investigation.

Nuclear receptors are transcription factors whose activities are dependent on direct binding of small, lipophilic ligands including steroids, thyroid hormone, retinoids, and vitamin D_3 (Mangelsdorf et al., 1995). Nuclear receptors control differentiation, development, cellular homeostasis, metabolism, and detoxification by directly regulating the expression of select target genes (Chawla et al., 2001). Mutations in nuclear receptors are associated with many human diseases such as cancer, diabetes,

inflammation, and heart disease due to their crucial regulatory roles. As an example, mutations in human HNF4α cause maturity onset diabetes in the young (MODY1), a rare form of non-insulin-dependent (type 2) diabetes mellitus (Yamagata et al., 1996). As a consequence, the pharmaceutical industry has made extensive efforts to produce medicines that modulate nuclear receptor functions. Recently, the studies of nuclear receptors of insects have drawn attention due to the important regulatory roles of vertebrate nuclear receptor counterparts. The 18 nuclear receptor genes were identified in *D. melanogaster*, and they represented all 6 subfamilies of vertebrate receptors (King-Jones & Thummel, 2005; Palanker et al., 2006). Since vertebrate nuclear receptors have central roles in growth regulation, metabolism, and detoxification, the extensive studies of insect nuclear receptors will provide new insight into insect development and physiology, and for insect pest control. Biotechnology-based pest control strategies such as transgenic plant protease inhibitor-expressing plants have not been successful due to insect adaptability. Therefore, the insect nuclear receptors that regulate insect counter-defense genes may contribute to the development of better strategies of insect pest control.

To investigate how CmSvp represses CmHNF-4-dependent activation of *CmCatB* expression, I performed EMSAs detecting *in vitro* interaction between CmHNF-4 and CmSvp on *COUP/HNF-4* elements. But I could not detect heterodimerization or protein-protein interaction between CmHNF-4 and CmSvp (data not shown). This finding, however, could have resulted from potentially weak interactions that do not survive in the electrophoretic conditions used. Therefore this

suggests investigating inhibitory mechanisms of CmSvp on CmHNF-4 transactivation *in vivo* system. I will perform cotransfections on CmHNF-4 and a chimeric protein construct (pAc5-Gal4-CmSvp) containing the LBD of CmSvp fused to the yeast Gal4 DNA-binding domain into S2 cells. Then the effect of Gal4-CmSvp on CmHNF-4 transactivation will be tested by CAT assay. A construct, pAc5-Gal4, expressing Gal4 DNA-binding domain will be contransfected as a control for specificity of Gal4-CmSvp effect on CmHNF-4 activity. If CmSvp repression requires a DNA binding on *CmCatB* promoter, CmHNF-4 transactivation should not be inhibited by Gal4-CmSvp due to loss of its DNA binding ability. CmHNF-4 transactivation, however, will be repressed by Gal4-CmSvp, if CmSvp repression depends on protein-protein interaction with CmHNF-4.

To facilitate further analysis of CmHNF-4 on transcriptional regulation of *CmCatB*, I will produce anti-recombinant CmHNF-4 antibody. Briefly, three cDNA fragments, the first encoding A/B domain (amino acids 1-82), the second encoding D, E and F domains (amino acids 149-507) and the third encoding entire domains of *CmHNF-4*, will be PCR amplified, and subcloned in pET-28a (+) expression vector. Constructs will be transferred to *E. coli* expression host strain BL21 (DE3). The recombinant proteins will be purified using a Ni^{2+} chelate affinity chromatography and the purified proteins will be used to produce rabbit polyclonal anti-recombinant CmHNF-4 antibody.

To demonstrate the functional significance of transcriptional regulators, CmSvp and CmHNF-4 in insect adaptation, RNA interference will be performed. Each cDNA fragment of above genes will be cloned into pBlueScript for *in vitro* transcription with

opposing T7 and T3 promoters. The resulting sense and antisense RNAs will be annealed, incorporated into artificial seeds with scN, and fed to cowpea bruchids. When larvae reach the 4th instar, insects will be removed from the seeds. Quantitative RT-PCR will be used to confirm silencing. Two negative controls will be included insects on control diet, and insects fed on dsRNA from green fluorescence protein gene, which does not share sequence homology with any known insect gene. If CmSvp regulates negatively *CmCatB* expression in scN-unadapted insects, CmCatB expression will be induced when CmSvp is silenced. If CmHNF-4 is a transcriptional activator of *CmCatB* in scN-adapted insects, CmCatB expression will be reduced when CmHNF-4 is silenced.

5.4. Experimental procedures

5.4.1. Insect treatment and midgut nuclear extract

Recombinant scN was expressed in *E. coli* BL21 (DE3) and purified via a Ni^{2+} chelate affinity chromatography as previously described (Zhu-Salzman et al., 2003). Cowpea bruchids were reared on artificial diet containing 0.2% scN until they reached the 4th instar larval stage when adaptive feeding behavior was observed. Insect midguts were dissected and gut contents were removed following the procedure of Ahn et al. (2007a). Nuclear extracts were prepared using a Nuclear Extract kit (Active Motif, Carlsbad, CA). Briefly, freshly dissected midguts free from gut contents were placed in a pre-chilled, sterile Dounce homogenizer, and homogenized in cold hypotonic buffer with 1

mM DTT and 0.1% (v/v) detergent. For every five guts, 100 μl buffer was added. The homogenate was incubated on ice for 15 min prior to a 10 min centrifugation at 850x g. The cell pellet was resuspended in 50 μl hypotonic buffer, incubated on ice for 15 min, and then lysed by adding detergent to 1% (v/v) followed by vigorous vortexing for 10 sec and centrifugation for 30 sec at 14,000x g. The nuclear pellet was resuspended in 30 μl complete lysis buffer (1 mM DTT, 1% (v/v) protease inhibitor cocktail, lysis buffer AM1). Nuclear proteins were then extracted by gently shaking on ice for 30 min. After centrifugation for 10 min at 14,000x g, the supernatant, containing nuclear proteins, was transferred to a new tube in aliquots and stored at – 70°C till use. All centrifugations were performed at 4°C.

5.4.2. Electrophoretic mobility shift assays (EMSAs)

DNA probes were radiolabeled either by PCR amplification or using T4 DNA polynucleotide kinase. Primers for EMSA are listed in Table 5.1. DNA probes larger than 60 bp were randomly labeled with 1.5 μM [α-^{32}P]dCTP by PCR amplification (94°C for 30 sec, 55°C for 30 sec, 72°C for 30 sec for 25 cycles). P1 (spanning -493 to -244) was amplified with primers 1 and 2, P2 (spanning -302 to -41) amplified with primers 3 and 4, P3 (spanning -493 to -339) amplified with primers 1 and 5, and P4 (spanning -400 to -244) amplified with primers 2 and 6. The overlapping region P5 (spanning 400 to -339) was obtained using primers 5 and 6. The PCR fragments were

Table 5.1. Oligonucleotide primers synthesized for EMSAs in scN-adapted bruchids

Primer	Sequence
1	$G_{(-493)}$GCTAATAGTTGCATAAGAGCAAG$_{(-470)}$
2	$A_{(-244)}$AAAGACGTATTCCCGTGTTAGT$_{(-266)}$
3	$A_{(-302)}$CACTGGAGAAAGGGAACAGG$_{(-282)}$
4	$C_{(-41)}$GCCTCTAATCACTTATCAGTATTCG$_{(-66)}$
5	$C_{(-339)}$CAAAGGTAAGGTCAAAAGGTC$_{(-360)}$
6	$C_{(-400)}$GAAATTCATTTTTATGGTGACC$_{(-378)}$
7	$G_{(-378)}$GTCACCATAAAAATGAATTTCG$_{(-400)}$
8	gcgcgT$_{(-382)}$GACCTTTGGACCTTACCTTTGGACC$_{(-357)}$ggcgg
9	ccgccg $_{(-357)}$GTCCAAAGGTAAGGTCCAAAGGTCA$_{(-382)}$cgcgc
10	$G_{(-360)}$ACCTTTTGACCTTACCTTTGG$_{(-339)}$

purified with a QIAquick PCR Purification kit (QIAGEN).

For probes less than 60 bp, two complementary oligonucleotides were synthesized (Table 5.1). Probes 6, 7 and 8 covering CdxA, HNF4/COUP-TF and CRE-BP putative *cis*-elements were formed by annealing primers 6 and 7, 8 and 9, 5 and 10, respectively. The oligonucleotides were end-labeled separately with 0.73 μM of [γ-^{32}P]ATP using T4 DNA polynucleotide kinase, and then mixed in complementary pairs as we previously described (Ahn et al., 2007a). After annealing, the double-stranded oligonucleotide probes were purified with QIAquick Nucleotide Removal kit (QIAGEN). Extra sequences (lower-case letters in Table 5.1) were added to the end of the shortest probe P7 to ensure that the probes were double-stranded.

EMSAs were performed by incubating 3 μg of gut nuclear extract, or 5 μl *in vitro* translated CmHNF-4 or luciferase, for 20 min with labeled probes (20,000 cpm per reaction) in binding buffer [4% glycerol, 1 mM $MgCl_2$, 0.5 mM EDTA, 0.5 mM DTT, 10 mM Tris-HCl, pH7.5, 0.05 μg of poly(dI-dC)•poly(dI-dC)] at room temperature. Samples resolved on 4% native polyacrylamide gel were exposed to X-ray film. Luciferase was used for experimental control. For competition assays, 5-, 10- or 50-fold molar excess of specific or nonspecific competitors were incubated with nuclear extract for 20 min at room temperature prior to the addition of probe. Nonspecific DNAs were prepared by PCR amplification of the coding regions of the bruchid cathepsins, *CmCPA9 or CmCPB1*, at the sizes of equivalent to their corresponding competing probes. One μl of anti-mosquito AaSvp serum raised against a portion of the DNA and ligand-binding domains of mosquito COUP-TF homolog or pre-immune (generously provided by A. Raikhel in University of California, Riverside, CA) was also pre-incubated with nuclear extract for 20 min.

5.4.3. Cloning of *CmHNF-4* from cowpea bruchid midguts

Total RNA was extracted from the midgut tissue of the scN-adapted cowpea bruchid 4th instar larvae using the TRIzol Reagent (Invitrogen, Carlsbad, CA), and reverse transcribed with Superscript^{TM} II reverse transcriptase (Invitrogen). Two pairs of degenerate primers were designed based on highly conserved ligand-binding domain (LBD) of other HNF-4 family members, but showed limited identity with CmSvp: (1)

sense 5′- GATGTBTGYGAKTCBATGAAG-3′; (2) antisense 5′- ATYTGYTSRATCA TYTGCCAKGT-3′; (3) sense 5′- GAKTCBATGAAGSARCAG-3′ and (4) antisense 5′- CATYTGCCAKGTRATDSWCTG -3′ (B = C,G,T; Y = C,T; K = G,T; S = C,G; R = A,G; D = A,G,T; W = A,T). The primary PCR (95°C for 30 sec, 42°C for 30 sec, 72°C for 40 sec for 35 cycles) was conducted with primers 1 and 2. The nested PCR (95°C for 30 sec, 42°C for 1 min, 72°C for 1 min for 35 cycles) was then performed with primers 3 and 4. The PCR fragment was subcloned into a pCRII vector and sequenced.

The 5′ and 3′ cDNA ends were PCR amplified (94°C for 30 sec, 68°C for 30 sec, 72°C for 1 or 3 min for 35 cycles) using the BD SMART RACE cDNA Amplification kit with the following gene-specific primers: (5) sense 5′- GCTTTTACGGAACTGCA ACTGGATGAC-3′; (6) sense 5′- GACCAGGTAGCTTTGTTACGAGCACAT-3′; (7) antisense 5′- GAGATGCATTGACCTTCTAGCTAATCC-3′ and (8) antisense 5′- AG CTAATCCTAAAAGTAGATGTTCTC C-3′. Primers 5 and 6 were sense primers used for primary PCR and nested PCR for 3′ RACE. Likewise, primers 7 and 8 were antisense primers for 5′ RACE nested PCR reactions. The PCR fragments were subcloned into pCRII vector and sequenced. The full-length coding region was obtained by RT-PCR (95°C for 30 s, 50°C for 30 sec, 72C for 2 min for 35 cycles) using the following primers: sense 5′- AAGCAGTGGTATCAACGCAGAGTACGC-3′, and antisense 5′- GGTAATGCATCAAAAGGGTTTTTCCTG-3′. It was then subcloned into pCRII vector, and the sequence of plasmid pCRII-CmHNF-4 was confirmed.

5.4.4. *In vitro* translation of CmHNF-4

CmHNF-4 protein was produced by coupled *in vitro* transcription and translation using the TNT T7 Quick Coupled Transcription/Translation System (Promega, Madison, WI). The TNT reaction was incubated at 30°C for 90 min using 1 μg of plasmid pCRII-CmHNF-4 containing T7 RNA polymerase promoter. From this reaction, 5 μl of protein was used for EMSA and competition assays as described above to evaluate DNA-binding specificity of CmHNF4. Luciferase was used as a control for *in vitro* protein synthesis as well as the CmHNF-4 binding specificity.

5.4.5. Transient transfection

Transfection experiments were performed as previously described (Ahn et al., 2007a). *Drosophila* Schneider 2 (S2) cells, maintained in Shields and Sang M3 insect medium (Sigma, St. Louis, MO) with supplements, were seeded at a density of 1 x 10^6 cells per well on a six-well titer plate, and allowed to attach for 1 h. The medium was then replaced twice with 2 mL fresh incomplete medium free of supplements, each for 15 min. The calcium phosphate precipitation method was used for transient transfection, where 8 μg reporter plasmids were diluted in 254 μL of HEPES (*N*-2-hydroxyethyl piperazine-*N*′-2-ethanesulfonic acid)-buffered saline (26 mM HEPES, 0.78 mM Na_2HPO_4, 146.6 mM NaCl, pH 7.1) containing 135 mM $CaCl_2$. During the 30 min

incubation at room temperature, calcium phosphate precipitates formed due to the $CaCl_2$ added in the DNA tube and the phosphate in the HEPES-buffered saline. The mixture was then added dropwise to the attached cells. After 18 h incubation at 27°C, the transfection mixture was removed and replaced with complete medium. As an internal standard, 1 μg of pAc5.1/V5-His/lacZ (Invitrogen) was cotransfected with all CAT reporter plasmids, pAc-CatB/CAT, pAc-CatBΔ(HNF4/COUP)/CAT, pAc3075, and pAc-IE1/CAT (Ahn et al., 2007). Cells were then harvested 24 h post-transfection for CAT and β-galactosidase assays.

To evaluate the potential combinatory effect of CmHNF-4 and CmSvp on *CmCatB* expression, the expression plasmid pAc5-CmHNF-4 was cotransfected with the reporter plasmid and/or pAc5-CmSvp (Ahn et al., 2007a) into S2 cells. To construct pAc5-CmHNF-4, the entire coding region of *CmHNF-4* was amplified by PCR (95°C for 30 sec, 68°C for 2 min for 35 cycles) using the following oligonucleotide primers: sense 5′ -AAGCTGAATTCCAAAAGATGCCTTCCTCGT - 3′, and antisense 5′ -GCT CACTCGAGAAAACTATGCTGCTCCTCCTT - 3′. *EcoR*I and *Xho*I (underlined) restriction sites were incorporated into primers for directional cloning. After restriction digestion, the PCR fragment was ligated to pAc5.1/V5-HisA expression vector (Invitrogen) and correct DNA sequence was verified. Five hundred ng of pAc5-CmHNF-4 and/or 1 μg of pAc5-CmSvp were cotransfected with the reporter plasmid. Appropriate amounts pAc5.1/V5-HisA vector alone were added to the transfection reactions to ensure comparable total DNA amount in transfected S2 cells. Cells were collected at 24 h post-transfection, and used for CAT and β-galactosidase assays (see

below). The reporter plasmid pAc-IE1/CAT (*CAT* gene placed under control of the promoter of a baculovirus immediate-early gene) (Guarino & Dong, 1991) was used to evaluate specificity of CmHNF-4 and CmSvp effects on *CmCatB* promoter.

5.4.6. CAT and β-galactosidase activity assays

CAT assays were performed following Ahn et al. (2007a). To measure CAT activity, the harvested cells were broken by three cycles of freezing and thawing in 200 μl of PBS. Endogenous deacetylase activity in the cell extract was heat inactivated. The extracts (30 μl) were then incubated with 200 μl of solution containing 100 mM Tris-HCl, 1 mM chloramphenicol, 0.1 μCi of ^{3}H-acetyl coenzyme A (pH 7.9) as well as 5 ml of Insta-Fluor Plus (PerkinElmer, Waltham, MA) at 37°C for 0, 30, 60 or 90 min respectively. Enzymatic activity was measured by production of ^{3}H-acytylated chloramphenicol using a Beckman LS6500 scintillation counter. Transfection assays were carried out independently for three times or more.

To normalize transfection efficiency, β-galactosidase activity was measured using the chromogenic substrate, o-nitrophenyl-β-D-galactopyranoside (ONPG) (Sigma). Cell extracts (10 μl) were incubated with 200 μl of 4 mg/mL of ONPG and 1 ml of Z buffer (100 mM sodium phosphate, 10 mM KCl, 1 mM $MgSO_4$, pH 7.0) containing 38.61 mM β-mercaptoethanol for 10 min at 37°C. Reactions were terminated by addition of 0.5 ml of 1 M Na_2CO_3. Absorbance at 420 nm of this mixture was measured using a Thermo GENESYS 10 UV spectrophotometer. Absorbance of the

sample without cell extract was used to zero the machine. Specific activity of β-galactosidase was defined as the amount of cell extract that hydrolyzed 1 nmol of ONPG to o-nitrophenol and D-galactose per min.

Paired-samples *t* test was used to evaluate the CAT activity data using SPSS for Mac. ver.16.0.

CHAPTER VI

SUMMARY

Due to insect adaptability to plant defense molecules, the use transgenic plants expressing natural plant protease inhibitors has not been a successful insect pest management strategy. Herbivorous insects are able to differentially express proteases such as inhibitor-insensitive or inhibitor-degrading proteases in response to plant inhibitors. Also herbivores are capable of inducing counter-defense related genes to cope with plant inhibitors. Therefore, to develop durable strategies of insect pest control, it is critical to understand thoroughly how insects differentially regulate proteases to overcome plant inhibitor, and the transcriptional regulatory mechanisms of counter-defensive genes behind insect adaptation. In this thesis, I used several approaches to investigate the molecular mechanisms of cowpea bruchid adaptation to dietary soybean cysteine protease inhibitor (scN).

The Zhu-Salzman's lab has demonstrated that cowpea bruchids circumvent the negative effect by dietary scN at 4th instar larval stage through differential expression of scN-insensitive protease or scN-degrading protease. Cathepsin L-like cysteine protease (CmCP) is a major digestive enzyme, and a multigene family grouped into CmCPA and CmCPB subfamilies based on protein sequence similarity. The Zhu-Salzman's lab previously found that transcripts of CmCPB subfamily were more highly abundant in scN-adapted bruchid midgut than CmCPA subfamily in response to dietary scN. Here I investigated the functional significance of CmCPs involved in bruchid resistance to scN

at the potein level.. In contrast to the CmCPA family, the CmCPB family showed more efficient autocatalytic processing from the latent proenzyme to its active mature protease form, and higher intrinsic proteolytic activity. Although all CmCPs were scN-sensitive, scN was degraded only by the presence of excessive CmCPB, but not by CmCPA. To overcome the negative effect of scN, cowpea bruchids prefer to express unique cysteine protease isoforms with superior autoprocessing, proteolytic activity, and scN-degrading activity.

I then further dissected the molecular mechanisms underlying the differential CmCP enzymatic activities. I performed domain swapping between the two respective subfamily members B1 and A16, the latter unable to autoprocess or degrade scN even after exogenous processing. Swapping the propeptide did not qualitatively alter autoprocessing activity, but incorporation of either the N- or C-terminal mature B1 segment into A16 was sufficient to prime autoprocessing of A16. Thus, the autoprocessing activity of B1 is controlled by the mature enzyme region. Bacterially expressed isolated propeptides (pA16 and pB1) differed in their ability to inhibit B1 proteolytic activity. Lower inhibitory activity in pB1 is likely due to its lack of protein stability. These findings suggest that cowpea bruchids modulate proteolysis of their digestive enzymes by controlling cleavage and stability of propeptides in response to plant inhibitors.

scN-insensitive CmCatB was the most highly up-regulated gene identified in an scN-regulated EST collection from cowpea bruchid midguts. I cloned a portion of its promoter and demonstrated its activity in Drosophila S2 cells using a CAT reporter

system. Gel shift assays identified two regulatory *cis*-elements, chicken ovalbumin upstream promoter (COUP) and hepatocyte nuclear factor 4 (HNF-4) in the CmCatB promoter. Interestingly, COUP and HNF-4 elements share the same regulatory DNA sequence in the CmCatB promoter. *COUP* elements interacted with cowpea bruchid Seven-up (CmSvp) in scN-unadapted insect midgut where CmCatB was not expressed. *HNF-4* elements bound to cowpea bruchid HNF-4 (CmHNF-4) in scN-adapted insect midgut where CmCatB was highly expressed. Transient expression in S2 cells showed that CmCatB expression was repressed by CmSvp and activated by CmHNF-4 through binding at HNF-4/COUP elements. Further, CmSvp repressed CmHNF-4-dependent transactivation when they were present simultaneously in the S2 cell. Thus, it suggests that transcriptional regulation of *CmCatB* is mediated with the interplay between positive and negative regulators in response to plant inhibitors.

REFERENCES

Achatz, G., Holzl, B., Speckmayer, R., Hauser, C., Sandhofer, F. and Paulweber, B. (1997) Functional domains of the human orphan receptor ARP-1/COUP-TFII involved in active repression and transrepression. *Mol Cell Biol* **17**: 4914-4932.

Ahn, J.-E., Guarino, L.A. and Zhu-Salzman, K. (2007a) Seven-up facilitates insect counter-defense by suppressing cathepsin B expression. *FEBS J* **274**: 2800-2814.

Ahn, J.-E., Lovingshimer, M. R., Salzman, R. A., Presnail, J. K., Lu, A. L., Koiwa, H. and Zhu-Salzman, K. (2007b) Cowpea bruchid *Callosobruchus maculatus* counteracts dietary protease inhibitors through modulating properties of major digestive enzymes. *Insect Mol Biol* **16**: 295-304.

Ahn, J.-E., Salzman, R.A., Braunagel, S.C., 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.

Amirhusin, B., Shade, R.E., Koiwa, H., Hasegawa, P.M., Bressan, R.A., Murdock, L.L. and Zhu-Salzman, K. (2007) Protease inhibitors from several classes work synergistically against *Callosobruchus maculatus*. *J Insect Physiol* **53**: 734-740.

Barbehenn, R.V., Bumgarner, S.L., Roosen, E.F. and Martin, M.M. (2001) Antioxidant defenses in caterpillars: role of the ascorbate-recycling system in the midgut lumen, *J Insect Physiol*. **47**: 349-357.

Barlic-Maganja, D., Dolinar, M. and Turk, V. (1998) The influence of Ala205 on the specificity of cathepsin L produced by dextran sulfate assisted activation of the recombinant proenzyme. *Biol Chem* **379**: 1449-1452.

Barrett, A.J. (1987) The cystatins: a new class of peptidase inhibitors. *TIBS* **12**: 193-196.

Billington, C.J., Mason, P., Magny, M.C. and Mort, J.S. (2000) The slow-binding inhibition of cathepsin K by its propeptide. *Biochem Biophys Res Commun* **276**: 924-929.

Bode, W. and Huber, R. (1992) Natural protein proteinase inhibitors and their interaction with proteinases. *Eur J Biochem* **204**: 433-451.

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.

Bonavides, K.B., Pelegrini, P.B., Laumann, R.A., Grossi-de-Sa, M.F., Bloch, C., Jr., Melo, J.A., Quirino, B.F., Noronha, E.F. and Franco, O.L. (2007) Molecular identification of four different α-amylase inhibitors from baru (*Dipteryx alata*) seeds with activity toward insect enzymes. *J Biochem Mol Biol* **40**: 494-500.

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

Bown, D.P., Wilkinson, H.S. and Gatehouse, J.A. (1997) Differentially regulated inhibitor-sensitive and insensitive protease genes from the phytophagous insect

pest, *Helicoverpa armigara*, are members of complex multigene families. *Insect Biochem. Mol. Biol.* **27**: 625-638.

Broadway, R.M. (2000) The response of insects to dietary protease inhibitors. In: *Recombinant Protease Inhibitors in Plants by* (Michaud, D., ed.), pp.80-88. Eurekah.com.

Brunelle, F., Cloutier, C. and Michaud, D. (2004) Colorado potato beetles compensate for tomato cathepsin D inhibitor expressed in transgenic potato. *Arch Insect Biochem Physiol.* **55**: 103-113.

Cairns, W., Smith, C.A., McLaren, A.W. and Wolf, C.R. (1996) Characterization of the human cytochrome P4502D6 promoter. A potential role for antagonistic interactions between members of the nuclear receptor family. *J Biol Chem* **271**: 25269-25276.

Carmona, E., Dufour, E., Plouffe, C., Takebe, S., Mason, P., Mort, J. S. and Menard, R. (1996) Potency and selectivity of the cathepsin L propeptide as an inhibitor of cysteine proteases. *Biochemistry* **35**: 8149-8157.

Campos, F.A.P., Xavier-Filho, J., Silva, C.P. and Ary, M.B. (1989) Resolution and partial characterization of proteinases and α-amylases from midguts of larvae of the bruchid beetle *Callosobruchus maculatus* (F.). *Comp Biochem Physiol* **92B**: 51-57.

Chagolla-Lopez, A., Blanco-Labra, A., Patthy, A., Sanchez, R. and Pongor, S. (1994) A novel α-amylase inhibitor from amaranth (*Amaranthus hypocondriacus*) seeds. *J Biol Chem* **269**: 23675-23680.

Chartier, F.L., Bossu, J.P., Laudet, V., Fruchart, J.C. and Laine, B. (1994) Cloning and sequencing of cDNAs encoding the human hepatocyte nuclear factor 4 indicate the presence of two isoforms in human liver. *Gene* **147**: 269-272.

Chawla, A., Repa, J.J., Evans, R.M. and Mangelsdorf, D.J. (2001) Nuclear receptors and lipid physiology: opening the X-files. *Science* **294**: 1866-1870.

Chen, W.S., Manova, K., Weinstein, D.C., Duncan, S.A., Plump, A.S., Prezioso, V.R., Bachvarova, R.F. and Darnell, J.E., Jr. (1994) Disruption of the HNF-4 gene, expressed in visceral endoderm, leads to cell death in embryonic ectoderm and impaired gastrulation of mouse embryos. *Genes Dev* **8**; 2466-2477.

Cherbas, L. and Cherbas, P. (1993) The arthropod initiator: the capsite consensus plays an important role in transcription. *Insect Biochem Mol Biol* **23**: 81-90.

Chrispeels, M.J. and Raikhel, N.V. (1991) Lectins, lectin genes, and their role in plant defense. *Plant Cell* **3**: 1-9.

Christou, P., Capell, T., Kohli, A., Gatehouse, J.A. & Gatehouse, A.M. (2006) Recent developments and future prospects in insect pest control in transgenic crops. *Trends Plant Sci* **11**: 302-308.

Cloutier, C., Jean, C., Fournier, M., Yelle, S. 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.

Cooney, A.J., Tsai, S.Y., O'Malley, B.W. and Tsai, M.J. (1992) Chicken ovalbumin upstream promoter transcription factor (COUP-TF) dimers bind to different

GGTCA response elements, allowing COUP-TF to repress hormonal induction of the vitamin D3, thyroid hormone, and retinoic acid receptors. *Mol Cell Biol* **12**: 4153-4163.

Cooney, A.J., Tsai, S.Y., O'Malley, B.W. and Tsai, M.J. (1991) Chicken ovalbumin upstream promoter transcription factor binds to a negative regulatory region in the human immunodeficiency virus type 1 long terminal repeat. *J Virol* **65**: 2853-2860.

Costa, R.H., Grayson, D.R. and Darnell, J.E., Jr. (1989) Multiple hepatocyte-enriched nuclear factors function in the regulation of transthyretin and alpha 1-antitrypsin genes. *Mol Cell Biol* **9**: 1415-1425.

Coulombe, R., Grochulski, P., Sivaraman, J., Menard, R., Mort, J.S. and Cygler, M. (1996) Structure of human procathepsin L reveals the molecular basis of inhibition by the prosegment. *EMBO J* **15**: 5492-5503.

David, J.P., Boyer, S., Mesneau, A., Ball, A., Ranson, H. and Dauphin-Villemant, C. (2006) Involvement of cytochrome P450 monooxygenases in the response of mosquito larvae to dietary plant xenobiotics. *Insect Biochem Mol Biol* **36**: 410-420.

De Leo, F., Bonade-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 proteinase inhibitor in transgenic plants. *Plant Physiol* **118**: 997-1004.

Després, L., David, J.P. and Gallet, C. (2007) The evolutionary ecology of insect resistance to plant chemicals. *Trends Ecol Evol* **22**: 298-307.

Dhe-Paganon, S., Duda, K., Iwamoto, M., Chi, Y.I. and Shoelson, S.E. (2002) Crystal structure of the HNF4 alpha ligand binding domain in complex with endogenous fatty acid ligand. *J Biol Chem* **277**: 37973-37976.

Dolinar, M., Maganja, D.B. and Turk, V. (1995) Expression of full-length human procathepsin L cDNA in *Escherichia coli* and refolding of the expression product. *Biol Chem Hoppe-Seyler* **376**: 385-388.

Drewes, T., Senkel, S., Holewa, B. and Ryffel, G.U. (1996) Human hepatocyte nuclear factor 4 isoforms are encoded by distinct and differentially expressed genes. *Mol Cell Biol* **16**: 925-931.

Duncan, S. A., Nagy, A. and Chan, W. (1997) Murine gastrulation requires HNF-4 regulated gene expression in the visceral endoderm: tetraploid rescue of Hnf-4 (-/-) embryos. *Development* **124**; 279-287.

Elpidina, E.N., Vinokurov, K.S., Gromenko, V.A., Rudenskaya, Y.A., Dunaevsky, Y.E. and Zhuzhikov, D.P. (2001) Compartmentalization of proteinases and amylases in *Nauphoeta cinerea* midgut. *Arch Insect Biochem Physiol* **48**: 206-216.

Enayati, A. A., Ranson, H. and Hemingway, J. (2005) Insect glutathione transferases and insecticide resistance. *Insect Mol Biol* **14**: 3-8.

Englard, S. and Seifter, S. (1990) Precipitation techniques. *Method Enzymol* **182**: 285-300.

Felton, G.W., Donato, K.K., Broadway, R.M. and Duffey, S.S. (1992) Impact of oxidized olant phenolics on the nutritional quality of dietary protein to a nocyuid herbivore, *Spodoptera exigua*. *J Insect Physiol.* **38**: 277-285.

Ferrer-Martinez, A., Marotta, M., Baldan, A., Haro, D. and Gomez-Foix, A.M. (2004) Chicken ovalbumin upstream promoter-transcription factor I represses the transcriptional activity of the human muscle glycogen phosphorylase promoter in C2C12 cells. *Biochim Biophys Acta* **1678**: 157-162.

Firmino F., Sales M.P., Gomes V.M., Miranda M.R.A., Domingues S.J.S. and Xavier-Filho J. (1996) Cowpea (*Vigna unguiculata*) vicilins associate with putative chitinous structures in the midgut and feces of the bruchid beetles *Callosobruchus maculatus* and *Zabrotes subfasciatus*. *Braz J Med Biol Res* **29**: 749-756.

Fisher, C.L. and Pei, G.K. (1997) Modification of a PCR-based site-directed mutagenesis method. *BioTechniques* **23:** 570-574.

Fitches, E. and Gatehouse, J.A. (1998) A comparison of the short and long term effects of insecticidal lectins on the activities of soluble and brush border enzymes of tomato moth larvae (*Lacanobia oleracea*). *J Insect Physiol* **44**: 1213-1224.

Fox, T., de Miguel, E., Mort, J. S. and Storer, A.C. (1992) Potent slow-binding inhibition of cathepsin B by its propeptide. *Biochemistry* **31**: 12571-12576.

Franco, O.L., Rigden, D.J., Melo, F.R. and Grossi-De-Sa, M.F. (2002) Plant α-amylase inhibitors and their interaction with insect α-amylases. *Eur J Biochem* **269**: 397-412.

Franco, O.L., Rigden, D.J., Melo, F.R., Bloch Jr, C., Silva, C.P. and Grossi de Sa, M.F. (2000) Activity of wheat α-amylase inhibitors towards bruchid α-amylases and structural explanation of observed specificities. *Eur J Biochem* **267**: 2166-2173.

Galson, D.L., Tsuchiya, T., Tendler, D.S., Huang, L.E., Ren, Y., Ogura, T. and Bunn, H.F. (1995) The orphan receptor hepatic nuclear factor 4 functions as a transcriptional activator for tissue-specific and hypoxia-specific erythropoietin gene expression and is antagonized by EAR3/COUP-TF1. *Mol Cell Biol* **15**: 2135-2144.

Garcia, J.A., Jacques, H. and Rob, G. (1986) Transformation of cowpea *Vigna unguiculata* cells with an antibiotic resistance gene using a Ti-plasmid-derived vector. *Plant Sci* **44**: 37-46.

Gatehouse, A.M., Gatehouse, J.A., Bharathi, M., Spence, J. and Powell, K.S. (1998) Immunohistochemical and developmental studies to elucidate the mechanism of action of the snowdrop lectin on the rice brown planthopper, *Nilaparvata lugens* (Stal). *J Insect Physiol* **44**: 529-539.

Gatehouse, A.M.R., Dewey, F.M., Dove, J., Fenton, K.A. and Pusztai, A. (1984) Effects of seed lectins from *Phaseolus vulgaris* on the development of larvae of *Callosobruchus maculatus*; mechanism of toxicity. *J Sci Food Agric* **35**: 373-380.

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

Graham, F.L. and van der Eb, A.J. (1973) A new technique for the assay of infectivity of human adenovirus 5 DNA. *Virology* **52**: 456-467.

Grimmelikhuijzen, C.J., Cazzamali, G., Williamson, M. and Hauser, F. (2007) The promise of insect genomics. *Pest Manag Sci* **63**: 413-416.

Groves, M.R., Coulombe, R., Jenkins, J. and Cygler, M. (1998) Structural basis for specificity of papain-like cysteine protease proregions toward their cognate enzymes. *Proteins* **32**: 504-514.

Guarino, L.A. and Dong, W. (1991) Expression of an enhancer-binding protein in insect cells transfected with the *Autographa californica* nuclear polyhedrosis virus IE1 gene. *J Virol* **65**: 3676-3680.

Guarino, L.A. and Summers, M.D. (1986) Functional mapping of a *trans*-activating gene required for expression of a baculovirus delayed-early gene. *J Virol* **57**: 563-571.

Harper, M.S., Hopkins, T.L. and Czapla, T.H. (1998) Effect of wheat germ agglutinin on formation and structure of the peritrophic membrane in European corn borer (*Ostrinia nubilalis*) larvae. *Tissue & Cell* **30**: 166-176.

Helmus, M.R. and Dussourd, D.E. (2005) Glues or poisons: which triggers vein cutting by monarch caterpillars ?. *Chemoecology* **15**: 45-49.

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

Hiromi, Y., Mlodzik, M., West, S.R., Rubin, G.M. and Goodman, C.S. (1993) Ectopic expression of seven-up causes cell fate changes during ommatidial assembly. *Development* **118**: 1123-1135.

Holzinger, F. and Wink, M. (1996) Mediation of cardiac glycoside insensitivity in the monarch butterfly (*Danaus plexippus*): role of an amino acid substitution in the ouabain binding site of Na^+, K^+-ATPase. *J Chem Econol.* **22**: 1921-1937.

Houseman, J.G., Campbell, F.C. and Morrison, P.E. (1987) A preliminary characterization of digestive proteases in the posterior midgut of the stable fly *Stomoxys calcitrans* (L.) (Diptera: Muscidae). *Insect Biochem* **17**: 213-218.

Houseman, J.G. and Downe, A.E.R. (1983) Cathepsin D-like activity in the posterior midgut of hemipteran insects. *Comp Biochem Physiol* **75B**: 509-512.

Huesing, J.E., Murdock, L.L. and Shade, R.E. (1991a) Effect of wheat germ isolectins on development of cowpea weevil. *Phytochemistry* **30**: 785-789.

Huesing, J.E., Murdock, L.L. and Shade, R.E. (1991b) Rice and stinging nettle lectins: insecticidal activity similar to wheat germ agglutinin. *Phytochemistry* **30**: 3565-3568.

Huesing, J.E., Shade, R.E., Chrispeels, M.J. and Murdock, L.L. (1991c) α-Amylase inhibitor, not phytohemagglutinin, explains resistance of common bean seeds to cowpea weevil. *Plant Physiol* **96**: 993-996.

Hwung, Y.P., Wang, L.H., Tsai, S.Y. and Tsai, M.J. (1988) Differential binding of the chicken ovalbumin upstream promoter (COUP) transcription factor to two different promoters. *J Biol Chem* **263**: 13470-13474.

Ikea, J., Ingelbrecht, I., Uwaifo, A. and Thottappilly, G. (2003) Stable gene transfomation in cowpea (*Vigna unguiculata* L. walp.) using particle gun method. *Afr J Biotechnol* **2**: 211-218.

Illy, C., Quraishi, O., Wang, J., Purisima, E., Vernet, T. and Mort, J.S. (1997) Role of the occluding loop in cathepsin B activity. *J Biol Chem* **272**: 1197-1202.

Ishidoh, K. and Kominami, E. (1994) Multi-step processing of procathepsin L in vitro. *FEBS Lett* **352**: 281-284.

Ishimoto, M. and Kitamura, K. (1989) Growth inhibitory effects of an α-amylase inhibitor from the kidney bean, *Phaseolus vulgaris* (L.) on three species of bruchids (Coleoptera: Bruchidae). *Appl Ent Zool* **24**: 281-286.

Iwai, K., Fukuoka, S.I., Fushiki, T., Tsujikawa, M., Hirose, M., Tsunasawa, S. and Sakiyama, F. (1987) Purification and sequencing of a trypsin-sensitive cholecystokinin-releasing peptide from rat pancreatic juice - its homology with pancreatic secretory trypsin-inhibitor. *J Biol Chem* **262**: 8956-8959.

Janzen, D.H., Juster, H.B. and Liener, I.E. (1976) Insecticidal action of the phytohemagglutinin in black beans on bruchid beetle. *Science* **192**: 795-796.

James, C. (2003) Preview: Global ststus of commercialized transgenic crops. *ISAAA* briefs 30 *International Service for the Acqusition of Agri-biotech Applications,* Ithaca, NY.

Jiang, G., Nepomuceno, L., Hopkins, K. and Sladek, F.M. (1995) Exclusive homodimerization of the orphan receptor hepatocyte nuclear factor 4 defines a new subclass of nuclear receptors. *Mol Cell Biol* **15**: 5131-5143.

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

Kadziola, A., Sogaard, M., Svensson, B. and Haser, R. (1998) Molecular structure of a barley α-amylase-inhibitor complex: implications for starch binding and catalysis. *J Mol Biol* **278**: 205-217.

Kapitskaya, M.Z., Dittmer, N.T., Deitsch, K.W., Cho, W.L., Taylor, D.G., Leff, T. and Raikhel, A.S. (1998) Three isoforms of a hepatocyte nuclear factor-4 transcription factor with tissue- and stage-specific expression in the adult mosquito. *J Biol Chem* **273**:29801-29810.

Karban, R. and Agrawal, A.A. (2002) Herbivore offense. *Annu Rev Ecol.* **33**: 641-664.

Kerber, B., Fellert, S. and Hoch, M. (1998) Seven-up, the *Drosophila* homolog of the COUP-TF orphan receptors, controls cell proliferation in the insect kidney. *Genes Dev* **12**:1781-1786.

Kimura, A., Nishiyori, A., Murakami, T., Tsukamoto, T., Hata, S., Osumi, T., Okamura, R., Mori, M. and Takiguchi, M. (1993) Chicken ovalbumin upstream promoter-transcription factor (COUP-TF) represses transcription from the promoter of the gene for ornithine transcarbamylase in a manner antagonistic to hepatocyte nuclear factor-4 (HNF-4). *J Biol Chem* **268**: 11125-11133.

King-Jones, K. and Thummel, C.S. (2005) Nuclear receptors-a perspective from *Drosophila*. *Nat Rev Genet* **6**: 311-323.

Kirschke, H., Barrett, A.J. and Rawlings, N.D. (1995) Proteinases 1. Lysosomal cysteine proteinases. *Protein Profile* **2**:1587-1643.

Kitch, L.W. and Murdock, L.L. (1986) Partial characterization of a major gut thiol proteinase from larvae of *Callosobruchus maculatus* F. *Arch Insect Biochem*

Physiol **3**: 561-575.

Koiwa, H., Shade, R.E., Zhu-Salzman, K., D'Urzo, M.P., Murdock, L.L., Bressan R.A. and Hasegawa P.M. (2000) A plant defensive cystatin (soyacystatin) targets cathepsin L-like digestive cysteine proteinases (DvCALs) in the larval midgut of western corn rootworm (*Diabrotica virgifera virgifera*). *FEBS Lett* **471**: 67-70.

Koiwa, H., Shade, R.E., Zhu-Salzman, K., Subramanian, L., Murdock, L.L., Nielsen, S.S., Bressan, R.A. and Hasegawa, P.M. (1998) Phage display selection can differentiate insecticidal activity of soybean cystatins. *Plant J* **14:** 371-379.

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

Lazennec, G., Kern, L., Valotaire, Y. and Salbert, G. (1997) The nuclear orphan receptors COUP-TF and ARP-1 positively regulate the trout estrogen receptor gene through enhancing autoregulation. *Mol Cell Biol* **17**: 5053-5066.

Lee, S.B. and Lee, S.H. (2007) Identification of a putative nuclear receptor involved in cytochrome P450 induction by cypermethrin in IPLB-LdFB cells. *Pesticide Biochem Physiol* **88**:237-246.

Leippe, M., Sievertsen, H.J., Tannich, E. and Horstmann, R.D. (1995) Spontaneous release of cysteine proteinases but not of pore-forming peptides by viable *Entamoeba histolytica*. *Parasitology* **111**: 569-574.

Lenarcic, B., Krasovec, M., Ritonja, A., Olafsson, I. and Turk, V. (1991) Inactivation of human cystatin C and kininogen by human cathepsin D. *FEBS Lett* **280**: 211-215.

Leng, X., Cooney, A.J., Tsai, S.Y. and Tsai, M.J. (1996) Molecular mechanisms of COUP-TF-mediated transcriptional repression: evidence for transrepression and active repression. *Mol Cell Biol* **16**: 2332-2340.

Li, T. and Chiang, J.Y.L. (2006) Rifampicin induction of CYP3A4 requires pregnane X receptor cross talk with hepatocyte nuclear factor 4a and coactivators, and suppression of small heterodimer partner gene expression. *Drug Metabolism and Disposition* **34**: 756-764.

Liu, Y., Salzman, R.A., Pankiw, T. and Zhu-Salzman, K. (2004) Transcriptional regulation in southern corn rootworm larvae challenged by soyacystatin N. *Insect Biochem Mol Biol* **34**: 1069-1077.

Macedo, M.L.R., de Castro, M.M. and Freire, M.G.M. (2004) Mechanisms of the insecticidal action of TEL (*Talisia esculenta* lectin) against *Callosobruchu*s *maculatus* (Coleoptera: Bruchidae). *Arch Insect Biochem Physiol* **56**: 84-96.

Macedo, M.L.R., Freire, M.G.M., Novello, J.C. and Marangoni, S. (2002) *Talisia esculenta* lectin and larval development of *Callosobruchus maculatus* and *Zabrotes subfasciatus* (Coleoptera: Bruchidae). *Biochimica et Biophysica Acta* **1571**: 83-88.

Macedo, M.L.R., Andrade, L.B.S., Moraes, R.A. and Xavier-Filho, J. (1993) Vicilin variants and the resistance of cowpea (*Vigna unguiculata*) seeds to the cowpea weevil (*Callosobruchus maculatus*). *Comp Biochem Physiol* **105C**: 89-94.

Mach, L., Mort, J.S. and Glossl, J. (1994) Maturation of human procathepsin B-Proenzyme activation and proteolytic processing of the precursor to the mature

proteinase, in vitro, are primarily unimolecular processes. *J Biol Chem* **269**: 13030-13035.

Machleidt, W., Nagler, D.K., Assfalg-Machleidt, I., Stubbs, M.T., Fritz, H. and Auerswald, E.A. (1995) Temporary inhibition of papain by hairpin loop mutants of chicken cystatin. *FEBS Lett* **361**: 185-190.

Machuka J.S., Okeola O.G., Chrispeels M.J. and Jackai L.E. (2000) The African yam bean seed lectin affects the development of the cowpea weevil but does not affect the development of larvae of the legume pod borer. *Phytochemistr*y **53**: 667-674.

Mader, S., Chambon, P. and White, J.H. (1993) Defining a minimal estrogen receptor DNA binding domain. *Nucleic Acids Res* **21**: 1125–1132.

Mangelsdorf, D.J., Thummel, C., Beato, M., Herrlich, P., Schutz, G., Umesono, K., Blumberg, B., Kastner, P., Mark, M., Chambon, P. and Evans, R.M. (1995) The nuclear receptor superfamily: the second decade. *Cell* **83**: 835-839.

Maubach, G., Schilling, K., Rommerskirch, W., Wenz, I., Schultz, J. E., Weber, E. and Wiederanders, B. (1997) The inhibition of cathepsin S by its propeptide-specificity and mechanism of action. *Eur J Biochem* **250**: 745-750.

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.

Mehlo, L., Gahakwa, D., Nghia, P.T., Loc, N.T., Capell, T., Gatehouse, J.A., Gatehouse, A.M. and Christou, P. (2005) An alternative strategy for sustainable pest

resistance in genetically enhanced crops. *Proc Natl Acad Sci US*A **102**: 7812-7816.

Melo, F.R., Sales, M.P., Pereira, L.S., Bloch, C., Jr and Franco, O.L. (1999) α-Amylase inhibitors from cowpea seeds. *Prot Pept Lett* **6**: 387-392.

Menard, R., Carmona, E., Takebe, S., Dufour, E., Plouffe, C., Mason, P. and Mort, J.S. (1998) Autocatalytic processing or recombinant human procathepsin L - Contribution of both intermolecular and unimolecular events in the processing of procathepsin L *in vitro*. *J Biol Chem* **273:** 4478-4484.

Mendez, E., Rocher, A., Calero, M., Girbes, T., Citores, L. and Soriano, F. (1996) Primary structure of ω-hordothionin, a member of a novel family of thionins from barley endosperm, and its inhibition of protein synthesis in eukaryotic and prokaryotic cell-free systems. *Eur J Biochem* **239**: 67-73.

Michaud, D. and Nguyen-Quoc, B. (2000) Using natural and modified protease inhibitors. In: *Recombinant Protease Inhibitors in Plants by* (Michaud, D., ed.), pp. 114-127. Eurekah.com.

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.

Michaud, D., Nguyen-Quoc, B., Bernier-Vadnais, N., Faye, L. and Yelle, S. (1994) Cysteine proteinase forms in sprouting potato tuber. *Physiol Plant* **90**: 497-503.

Mietus-Snyder, M., Sladek, F.M., Ginsburg, G.S., Kuo, C.F., Ladias, J.A., Darnell, J.E., Jr. and Karathanasis, S.K. (1992) Antagonism between apolipoprotein AI

regulatory protein 1, Ear3/COUP-TF, and hepatocyte nuclear factor 4 modulates apolipoprotein CIII gene expression in liver and intestinal cells. *Mol Cell Biol* **12**: 1708-1718.

Miura, K., Zhu, J., Dittmer, N.T., Chen, L. and Raikhel, A.S. (2002) A COUP-TF/Svp homolog is highly expressed during vitellogenesis in the mosquito *Aedes aegypti*. *J Mol Endocrinol* **29**: 223-238.

Mlodzik, M., Hiromi, Y., Weber, U., Goodman, C.S. and Rubin, G.M. (1990) The *Drosophila* seven-up gene, a member of the steroid receptor gene superfamily, controls photoreceptor cell fates. *Cell* **60**: 211-224.

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.

Moreno, J. and Chrispeels, M.J. (1989) A lectin gene encodes the α-amylase inhibitor of the common bean. *Proc Natl Acad Sci USA* **86**: 7885-7889.

Mouillet, J. F., Bousquet, F., Sedano, N., Alabouvette, J., Nicolai, M., Zelus, D., Laudet, V. and Delachambre, J. (1999) Cloning and characterization of new orphan nuclear receptors and their developmental profiles during *Tenebrio* metamorphosis. *Eur J Biochem* **265**, 972-981.

Moura, F.T., Oliveira, A.S., Macedo, L.L., Vianna, A.L., Andrade, L.B., Martins-Miranda, A.S., Oliveira, J.T., Santos, E.A. and de Sales, M.P. (2007) Effects of a chitin-binding vicilin from *Enterolobium contortisiliquum* seeds on bean bruchid pests (*Callosobruchus maculatus* and *Zabrotes subfasciatus*) and

phytopathogenic fungi (*Fusarium solani* and *Colletrichum lindemuntianum*). *J Agric Food Chem* **55**: 260-266.

Murdock, L.L., Huesing, J.E., Suzanne, N.S., Pratt, R.C. and Shade, R.E. (1990) Biological effects of plant lectins on the cowpea weevil. *Phytochemistry* **29**: 85-89.

Murdock, L.L., Shade, R.E. and Pomeroy, M.A. (1988) Effects of E-64, a cysteine proteinase inhibitor, on cowpea weevil growth, development, and fecundity. *Environ Entomol* **17**: 467-469.

Murdock, L.L., Brookhart, G., Dunn, P.E., Foard, D.E., Kelley, S., Kitch, L., Shade, R.E., Shukle, R.H. and Wolfson, J.L. (1987) Cysteine digestive proteinases in Coleoptera. *Comp. Biochem. Physiol* **87B**: 783-787.

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

Musser, R.O., Hum-Musser, S.M., Eichenseer, H., Peiffer, M., Ervin, G., Murphy, J.B. and Felton, G.W. (2002) Caterpillar saliva beats plant defences. *Nature* **416**: 599-600.

Myers, S. A., Wang, S. C. and Muscat, G. E. (2006) The chicken ovalbumin upstream promoter-transcription factors modulate genes and pathways involved in skeletal muscle cell metabolism. *J Biol Chem* **281**: 24149-24160.

Neurath, H. (1984) Evolution of proteolytic enzymes. *Science* **224**: 350-357.

Nishida, R. (2002) Sequestration of defensive substances from plants by Lepidoptera. *Annu Rev Entomol* **47**: 57-92.

Nomura, T. and Fujisawa, Y. (1997) Processing properties of recombinant human procathepsin L. *Biochem Bioph Res Co* **230**: 143-146.

Nyaruhucha, C.N.M., Kito, M. and Fukuoka, S.I. (1997) Identification and expression of the cDNA-encoding human mesotrypsin(ogen), an isoform of trypsin with inhibitor resistance. *J Biol Chem* **272:** 10573-10578

Ogino, T., Kaji, T., Kawabata, M., Satoh, K., Tomoo, K., Ishida, T., Yamazaki, H., Ishidoh, K. and Kominami, E. (1999) Function of the propeptide region in recombinant expression of active procathepsin L in *Escherichia coli*. *J Biochem* **126**: 78-83.

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

Palanker, L., Necakov, A.S., Sampson, H.M., Ni, R., Hu, C., Thummel, C.S. and Krause, H.M. (2006) Dynamic regulation of *Drosophila* nuclear receptor activity in vivo. *Development* **133**: 3549-3562.

Parama, A., Iglesias, R., Alvarez, M.F., Leiro, J., Ubeira, F.M. and Sanmartin, M.L. (2004) Cysteine proteinase activities in the fish pathogen *Philasterides dicentrarchi* (Ciliophora : Scuticociliatida). *Parasitology* **128:** 541-548.

Park, J.-I., Tsai, S.Y. and Tsai, M.J. (2003) Molecular mechanism of chicken ovalumin upstream promoter-transcription factor (COUP-TF) actions. *Keio J Med* **52**: 174-181.

Pastorcic, M., Wang, H., Elbrecht, A., Tsai, S.Y., Tsai, M.J. and O'Malley, B.W. (1986) Control of transcription initiation in vitro requires binding of a transcription factor to the distal promoter of the ovalbumin gene. *Mol Cell Biol* **6**: 2784-2791.

Petersen, R.A., Zangerl, A.R., Berenbaum, M.R. and Schuler, M.A. (2001) Expression of CYP6B1 and CYP6B3 cytochrome P450 monooxygenases and furanocoumarin metabolism in different tissues of *Papilio polyxenes* (Lepidoptera: Papilionidae). *Insect Biochem Mol Biol* **31**: 679-690.

Peumans, W.J. and Van Damme, E.J.M. (1995) Lectins as plant defense proteins. *Plant Physiol* **109**: 347-352.

Popelka, J.C., Gollasch, S., Moore, A., Molvig, L. and Higgins, T.J. (2006) Genetic transformation of cowpea (*Vigna unguiculata* L.) and stable transmission of the transgenes to progeny. *Plant Cell Rep* **25**: 304-312.

Powell, K.S., Spence, J., Bharathi M., Gatehouse, J.A. and Gatehouse, A.M.R. (1998) Immunohistochemical and developmental studies to elucidate the mechanism of action of the snowdrop lectin on the rice brown planthopper, *Nilaparvata lugens* (Stal). *J Insect Physiol* **44**: 529-539.

Quraishi, O. and Storer, A.C. (2001) Identification of internal autoproteolytic cleavage sites within the prosegments of recombinant procathepsin B and procathepsin S - Contribution of a plausible unimolecular autoproteolytic event for the processing of zymogens belonging to the papain family. *J Biol Chem* **276:** 8118-8124.

Raccurt, M., Smallwood, S., Mertani, H.C., Devost, D., Abbaci, K., Boutin, J.M. and

Morel, G. (2005) Cloning, expression and regulation of chicken ovalbumin upstream promoter transcription factors (COUP-TFII and EAR-2) in the rat anterior pituitary gland. *Neuroendocrinology* **82**: 233-244.

Ratzka, A., Vogel, H., Kliebenstein, D.J., Mitchell-Olds, T. and Kroymann, J. (2002) Disarming the mustard oil bomb. *Proc Natl Acad Sci U S A* **99**: 11223-8.

Renard, G., Garcia, J.F., Cardoso, F.C., Richter, M.F., Sakanari, J.A., Ozaki, L.S., Termignoni, C. and Masuda, A. (2000) Cloning and functional expression of a *Boophilus microplus* cathepsin L-like enzyme. *Insect Biochem Mol Biol* **30**: 1017-1026.

Rowan, A.D., Mason, P., Mach, L. and Mort, J.S. (1992) Rat procathepsin B-proteolytic processing to the mature form in vitro. *J Biol Chem* **267**: 15993-15999.

Rozman, J., Stojan, J., Kuhelj, R., Turk, V. and Turk, B. (1999) Autocatalytic processing of recombinant human procathepsin B is a bimolecular process. *FEBS Lett* **459**: 358-362.

Ryan, CA. (1990) Protease inhibitors in plants: genes for improving defenses against insects and pathogens. *Annu. Rev. Phytopathol* **28**:425-449.

Sales, M.P., Andrade, L.B., Ary, M.B., Miranda, M.R., Teixeira, F.M., Oliveira, A.S., Fernandes, K.V. and Xavier-Filho, J. (2005) Performance of bean bruchids *Callosobruchus maculatus* and *Zabrotes subfasciatus* (Coleoptera: Bruchidae) reared on resistant (IT81D-1045) and susceptible (Epace 10) *Vigna unguiculata* seeds: relationship with trypsin inhibitor and vicilin excretion. *Comp Biochem Physiol A* **142**: 422-426.

Sales, M.P., Pimenta, P.P., Paes, N.S., Grossi-de-Sa, 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.

Sales, M.P., Gerhardt, I.R., Grossi-De-Sa, M.F. and Xavier-Filho, J. (2000) Do legume storage proteins play a role in defending seeds against bruchids? *Plant Physiol* **124**: 515-522.

Sales, M.P., Macedo, M.L.R. and Xavier-Filho, J. (1992) Digestibility of cowpea (*Vigna ungulculata*) vicilins by pepsin, papain and bruchid (Insect) midgut proteinases. *Comp Biochem Physiol* **103B**: 945-950.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Santamaria, I., Velasco, G., Pendas, A.M., Fueyo, A. and Lopez-Otin, C. (1998) Cathepsin Z, a novel human cysteine proteinase with a short propeptide domain and a unique chromosomal location. *J Biol Chem* **273**: 16816-16823.

Sawaya, B.E., Rohr O., Aunis D., and Schaeffer E. (1996) Chicken ovalbumin upstream promoter transcription factor, a transcriptional activator of HIV-1 gene expression in human brain cells. *J Biol Chem* **271**: 23572-23576.

Schagger, H. and von Jagow, G. (1987) Tricine sodium dodecyl-sulfate polyacrylamide-gel electrophoresis for the separation of proteins in the range from 1-kDa to 100-kDa. *Anal Biochem* **166:** 368-379.

Shade, R.E., Furgason, E.S. and Murdock, L.L. (1990) Detection of hidden insect infestations by feeding-generated ultrasonic signals. *Am Entomol* **Fall**: 231-234.

Shade, R.E., Murdock, L.L., Foard, D.E. and Pomeroy, M.A. (1986) Artificial seed system for bioassay of cowpea weevil (Coleoptera: Bruchidae) growth and development. *Environ Entomol* **15**: 1286-1291.

Shade, R.E., Schroeder, H.E., Pueyo, J.J., Tabe, L.M., Murdock, L.L., Higgins, T.J. and Chrispeels, M.J. (1994) Transgenic pea seeds expressing the α-amylase inhibitor of the common bean are resistant to bruchid beetles. *Bio Technology* **12**: 793-796.

Shibata, H., Nawaz, Z., Tsai, S. Y., O'Malley, B. W. and Tsai, M. J. (1997) Gene silencing by chicken ovalbumin upstream promoter-transcription factor I (COUP-TFI) is mediated by transcriptional corepressors, nuclear receptor-corepressor (N-CoR) and silencing mediator for retinoic acid receptor and thyroid hormone receptor (SMRT). *Mol Endocrinol* **11**: 714-724.

Silva, C.P., Terra, W.R., Xavier-Filho, J., Grossi de Sa, M.F., Lopes, A.R. and Pontes, E.G. (1999) Digestion in larvae of *Callosobruchus maculatus* and *Zabrotes subfasciatus* (Coleoptera: Bruchidae) with emphasis on α-amylases and oligosaccharidases. *Insect Biochem Mol Biol* **29**: 355-366.

Silva, C.P. and Xavier-Filho, J. (1991) Comparison between the levels of aspartic and cysteine proteinases of the larval midguts of *Callosobruchus maculatus* (F.) and *Zabrotes subfasciatus* (BOH.) (Coleoptera: Bruchidae). *Comp Biochem Physiol* **99B**: 529-533.

Silva, L.B., Sales, M.P., Oliveira, A.E., Machado, O.L., Fernandes, K.V. and Xavier-Filho, J. (2004) The seed coat of *Phaseolus vulgaris* interferes with the development of the cowpea weevil [*Callosobruchus maculatus* (F.) (Coleoptera: Bruchidae)]. *An Acad Bras Cienc* **76**: 57-65.

Singh, B.B. and Singh, S.R. (1992) Breeding for bruchid resistance in cowpea. *IITA Res* **5**: 1-5.

Singh, B.B., Singh, S.R. and Adjadi, O. (1985) Bruchid resistance in cowpea. *Crop Sci* **25**: 736-739.

Sladek, F.M. and Seidel, S.D. (2001) Hepatocyte nuclear factor 4α. In: *Nuclear Receptors and Genetic Diseases by* (Burris, T. and McCabe, E.R.B., ed.), pp. 309-361. Academic Press, London.

Sladek, F. M., Zhong, W. M., Lai, E. and Darnell, J. E., Jr. (1990) Liver-enriched transcription factor HNF-4 is a novel member of the steroid hormone receptor superfamily. *Genes Dev* **4**: 2353-2365.

Smith, S.M. and Gottesman, M.M. (1989) Activity and deletion analysis of recombinant human cathepsin L expressed in *Escherichia coli*. *J Biol Chem* **264**: 20487-20495.

Stroup, D., Crestani, M. and Chiang, J. Y. L. (1997) Orphan receptors chicken ovalbumin upstream promoter transcription factor II (COUP-TFII) and retinoid X receptor (RXR) activate and bind the rat cholestrol 7α-hydroxylase gene (*CYP7A*). *J Biol Chem* **272**: 9833-9839.

Sugiyama, T., Wang, J. C., Scott, D. K. and Granner, D. K. (2000) Transcription

activation by the orphan nuclear receptor, chicken ovalbumin upstream promoter-transcription factor I (COUP-TFI). Definition of the domain involved in the glucocorticoid response of the phosphoenolpyruvate carboxykinase gene. *J Biol Chem* **275**: 3446-3454.

Summers, C. B. and Felton, G. W. (1994) Prooxidant effects of phenolic acids on the generalist herbivore Helicoverpa zea (lepidoptera: noctuidae): potential mode of action for phenolic compunds in plant anti-herbivore chemistry. *Insect Biochem Mol Biol* **24**: 943-953.

Swevers, L. and Iatrou, K. (1998) The orphan receptor BmHNF-4 of the silkmoth *Bombyx mori*: ovarian and zygotic expression of two mRNA isoforms encoding polypeptides with different activating domains. *Mech Dev* **72**: 3-13.

Szmola, R., Kukor, Z. and Sahin-Toth, M. (2003) Human mesotrypsin is a unique digestive protease specialized for the degradation of trypsin inhibitors. *J Biol Chem* **278:** 48580-48589.

Tao, K., Stearns, N.A., Dong, J.M., Wu, Q.L. and Sahagian, G.G. (1994) The proregion of cathepsin L is required for proper folding, stability, and ER exit. *Arch Biochem Biophys* **311**: 19-27.

Tabashnik, B.E., Gassmann, A.J., Crowder, D.W. and Carriere, Y. (2008) Insect resistance to Bt crops: evidence versus theory. *Nat Biotechnol* **26**: 199-202.

Tabashnik, B.E., Carriere, Y., Dennehy, T.J., Morin, S., Sisterson, M.S., Roush, R.T., Shelton, A.M. and Zhao, J.Z. (2003) Insect resistance to transgenic Bt crops: lessons from the laboratory and field. *J Econ Entomol* **96**: 1031-1038.

Tarver, M.R., Shade, R.E., Shukle, R.H., Moar, W.J., Muir, W.M., Murdock, L.L. and Pittendrigh, B.R. (2007) Pyramiding of insecticidal compounds for control of the cowpea bruchid (*Callosobruchus maculatus* F.). *Pest Manage Sci* **63**: 440-446.

Tchoupe, J.R., Moreau, T., Gauthier, F. and Bieth, J.G. (1991) Photometric or fluorometric assay of cathepsin-B, cathepsin-L and cathepsin-H and papain using substrates with an aminotrifluoromethylcoumarin leaving group. *Biochim Biophys Acta* **1076:** 149-151.

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

Thevissen, K., Ghazi, A., De Samblanx, G.W., Brownlee, C., Osborn, R.W. and Broekaert, W.F. (1996) Fungal membrane responses induced by plant defensins and thionins. *J Biol Chem* **271**: 15018-15025.

Thomas, M. B. (1999) Ecological approaches and the development of "truly integrated" pest management. *Proc Natl Acad Sci U S A* **96**: 5944-5951.

Tsai, S. Y. and Tsai, M. J. (1997) Chick ovalbumin upstream promoter-transcription factors (COUP-TFs): Coming of age. *Endocrine Rev* **18**: 229-240.

Tsuzuki, S., Miura, Y., Fushiki, T., Oomori, T., Satoh, T., Natori, Y., and Sugimoto, E. (1992) Molecular cloning and characterization of genes encoding rat pancreatic cholecystokinin (CCK)-releasing peptide (monitor peptide) and pancreatic secretory trypsin-inhibitor (PSTI). *Biochim Biophys Acta* **1132**: 199-202.

Turk, B., Turk, V. and Turk, D. (1997) Structural and functional aspects of papain-like cysteine proteinases and their protein inhibitors. *Biol Chem* **378:** 141-150.

Turk, B., Dolenc, I., Turk, V. and Bieth, J.G. (1993) Kinetics of the pH-induced inactivation of human cathepsin L. *Biochemistry* **32**: 375-380.

Uchoa, A.F., DaMatta, R.A., Retamal, C.A., Albuquerque-Cunha, J.M., Souza, S.M., Samuels, R.I., Silva, C.P. and Xavier-Filho, J. (2006) Presence of the storage seed protein vicilin in internal organs of larval *Callosobruchus maculatus* (Coleoptera: Bruchidae). *J Insect Physiol* **52**: 169-178.

Umesono, K. and Evans, R. M. (1989) Determinants of target gene specificity for steroid/thyroid hormone receptors. *Cell* **57**: 1139-1146.

Velasco, G., Ferrando, A.A., Puente, X.S., Sanchez, L.M. and Lopez-Otin, C. (1994) Human cathepsin O-Molecular cloning from a breast carcinoma, production of the active enzyme in *Escherichia coli*, and expression analysis in human tissues. *J Biol Chem* **269**: 27136-27142.

Vernet, T., Khouri, H.E., Laflamme, P., Tessier, D.C., Musil, R., Gour-Salin, B.J., Storer, A.C. and Thomas, D.Y. (1991) Processing of the papain precursor. Purification of the zymogen and characterization of its mechanism of processing. *J Biol Chem* **266**: 21451-21457.

Vernet, T., Tessier, D.C., Laliberte, F., Dignard, D. and Thomas, D.Y. (1989) The expression in *Escherichia coli* of a synthetic gene coding for the precursor of papain is prevented by its own putative signal sequence. *Gene* **77**: 229-236.

Visal, S., Taylor, M.A. and Michaud, D. (1998) The proregion of papaya proteinase IV inhibits Colorado potato beetle digestive cysteine proteinases. *FEBS Lett.* **434:** 401-405.

Watt, A.J., Garrison, W.D. and Duncan, S.A. (2003) HNF4: a central regulator of hepatocyte differentiation and function. *Hepatology* **37**: 1249-1253.

Wigler, M., Silverstein, S., Lee, L.-S., Pellicer, A., Cheng, Y.-C. and Axel, R. (1977) Transfer of purified herpes virus thymidine kinase gene to cultured mouse cells. *Cell* **11**: 223-232.

Wijaya, R., Neumann, G.M., Condron, R., Hughes, A.B. and Polya, G.M. (2000) Defense proteins from seed of *Cassia fistula* include a lipid transfer protein homologue and a protease inhibitory plant defensin. *Plant Sci* **159**: 243-255.

Wisely, G. B., Miller, A. B., Davis, R. G., Thornquest, A. D., Jr., Johnson, R., Spitzer, T., Sefler, A., Shearer, B., Moore, J. T., Miller, A. B., Willson, T. M. and Williams, S. P. (2002) Hepatocyte nuclear factor 4 is a transcription factor that constitutively binds fatty acids. *Structure* **10**: 1225-1234.

Wittstock, U., Agerbirk, N., Stauber, E. J., Olsen, C. E., Hippler, M., Mitchell-Olds, T., Gershenzon, J. and Vogel, H. (2004) Successful herbivore attack due to metabolic diversion of a plant chemical defense. *Proc Natl Acad Sci U S A*. **101**: 4859-4864.

Wurtz, J. M., Bourguet, W., Renaud, J. P., Vivat, V., Chambon, P., Moras, D. and Gronemeyer, H. (1996) A canonical structure for the ligand-binding domain of nuclear receptors. *Nat Struct Biol* **3**: 87-94.

Xiong, B. & Jacobs-Lorena, M. (1995) Gut-specific transcriptional regulatory elements of the carboxypeptidase gene are conserved between black flies and *Drosophila*. *Proc Natl Acad Sci USA* **92**: 9313-9317.

Yamagata, K., Furuta, H., Oda, N., Kaisaki, P. J., Menzel, S., Cox, N. J., Fajans, S. S., Signorini, S., Stoffel, M. and Bell, G. I. (1996) Mutations in the hepatocyte nuclear factor-4alpha gene in maturity-onset diabetes of the young (MODY1). *Nature* **384**: 458-460.

Yan, S., Berquin, I. M., Troen, B. R. and Sloane, B. F. (2000) Transcription of human cathepsin B is mediated by Sp1 and Ets family factors in glioma. *DNA Cell Biol* **19**: 79-91.

Yanai, K., Hirota, K., Taniguchi-Yanai, K., Shigematsu, Y., Shimamoto, Y., Saito, T., Chowdhury, S., Takiguchi, M., Arakawa, M., Nibu, Y., Sugiyama, F., Yagami, K. and Fukamizu, A. (1999) Regulated expression of human angiotensinogen gene by hepatocyte nuclear factor 4 and chicken ovalbumin upstream promoter-transcription factor. *J Biol Chem* **274**: 34605-34612.

Yang, Z., Zhang, F., He, Q. & He, G. (2005) Molecular dynamics of detoxification and toxin tolerance genes in brown planthopper (*Nilaparvata lugens* Stal., Homoptera: Delphacidae) feeding on resistant rice plants. *Arch Insect Biochem Physiol* **59**: 59-66.

You, M., Fischer, M., Cho, W. K. and Crabb, D. (2002) Transcriptional control of the human aldehyde dehydrogenase 2 promoter by hepatocyte nuclear factor 4: inhibition by cyclic AMP and COUP transcription factors. *Arch Biochem Biophys* **398**: 79-86.

Yu, S. J. (1996) Insect glutathione S-transferases. *Zoological Studies* **31**: 9-19.

Yunes, A.N.A., Andrade, M.T., Sales, M.P., Morais, R.A., Fernandes, K.V., Gomes, V.M. and Xavier-Filho, J. (1998) Legume seed vicilins (7S storage proteins) interfere with development of the cowpea weevil [*Callosobruchus maculatus* (F)]. *J Sci Food Agric* **76**: 111-116.

Zagrobelny, M., Bak, S., Rasmussen, A.V., Jorgensen, B., Naumann, C.M. and Moller B.L. (2004) Cyanogenic glucosides and plant-insect interactions. *Phytochemistry* 6**5**: 293-306.

Zelhof, A. C., Yao, T. P., Chen, J. D., Evans, R. M. and McKeown, M. (1995) Seven-up inhibits ultraspiracle-based signaling pathways in vitro and in vivo. *Mol Cell Biol* **15**: 6736-6745.

Zhao, J.Z., Cao J., Collins, H.L., Bates, S.L., Roush, R.T., Earle, E.D. and Shelton, A.M. (2005) Concurrent use of transgenic plants expressing a single and two *Bacillus thuringiensis* genes speeds insect adaptation to pyramided plants. *Proc Natl Acad Sci U S A* **102**: 8426-8430.

Zhao, J.Z., Cao, J., Li, Y., Collins, H.L., Roush, R.T., Earle, E.D. and Shelton, A.M. (2003) Transgenic plants expressing two *Bacillus thuringiensis* toxins delay insect resistance evolution. *Nat Biotechnol* **21**: 1493-1497.

Zhao, Y., Botella, M.A., Subramanian, L., Niu, X.M., Nielsen, S.S., Bressan, R.A. and Hasegawa, P.M. (1996) Two wound-inducible soybean cysteine proteinase inhibitors have greater insect digestive proteinase inhibitory activities than a constitutive homolog. *Plant Physiol* **111:** 1299-1306.

Zhong, W., Mirkovitch, J. and Darnell, J. E., Jr. (1994) Tissue-specific regulation of mouse hepatocyte nuclear factor 4 expression. *Mol Cell Biol* **14**: 7276-84.

Zhong, W., Sladek, F. M. and Darnell, J. E., Jr. (1993) The expression pattern of a *Drosophila* homolog to the mouse transcription factor HNF-4 suggests a determinative role in gut formation. *EMBO J* **12**: 537-44.

Zhu, K., Huesing, J.E., Shade, R.E., Bressan, R.A., Hasegawa, P.M. and Murdock, L.L. (1996) An insecticidal N-acetylglucosamine-specific lectin gene from *Griffonia simplicifolia* (Leguminosae). *Plant Physiol* **110**: 195-202.

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.

Zhu-Salzman, K. and Salzman, R.A. (2001) Functional mechanics of the plant defensive *Griffonia simplicifolia* lectin II: Resistance to proteolysis is independent of glycoconjugate binding in the insect gut. *J. Econ Entomol* **94**: 1280-1284.

Zhu-Salzman, K., Shade, R.E., Koiwa, H., Salzman, R.A., Narasimhan, M., Bressan, R.A., Hasegawa, P.M. and Murdock, L.L. (1998) Carbohydrate binding and resistance to proteolysis control insecticidal activity of *Griffonia simplicifolia* lectin II. *Proc Natl Acad Sci U S A* **95**: 15123-15128.

VITA

Name: Ji Eun Ahn

Address: Department of Entomology
Texas A&M University
College Station, TX 77843-2475
Phone: 979-458-3359
Fax: 979-862-4790

Email Address: ji-eun-ahn@neo.tamu.edu

Education: B.S., Molecular Biology, Jeonbuk National University, Korea, 1992
M.S., Molecular Biology, Jeonbuk National University, Korea, 1994
Ph.D., Entomology, Texas A&M University, 2008