CHICKPEA – *Helicoverpa armigera*: A SYSTEM TO ELUCIDATE PLANT- INSECT PEST INTERACTIONS

A thesis submitted to The University of Pune

for the Degree of

DOCTOR OF PHILOSOPHY

in Biotechnology

AJAY SRINIVASAN,

Plant Molecular Biology Group,

Division of Biochemical Sciences,

National Chemical Laboratory,

Pune – 411008, INDIA.

February 2006

CERTIFICATE

Certified that the work incorporated in the thesis entitled "**Chickpea – *Helicoverpa armigera*: a system to elucidate plant- insect interactions**" submitted by Mr. Ajay Srinivasan was carried out under my supervision. The material obtained from other sources has been duly acknowledged in the thesis.

Dr. Vidya S. Gupta
(Research Guide)

DECLARATION

I hereby declare that the thesis entitled "**Chickpea – *Helicoverpa armigera*: a system to elucidate plant- insect interactions**" submitted for the degree of Ph. D. at the University of Pune has not been submitted by me for a degree at any other university.

Date:

Ajay Srinivasan,
National Chemical Laboratory, Pune.

AN INTER-INSTITUTIONAL COLLABORATIVE EFFORT

The research embodied in this thesis was carried out at **NATIONAL CHEMICAL LABORATORY, PUNE, INDIA** and **UNIVERSITY OF DURHAM, DURHAM, U.K.**

This research project is part of collaborative projects, on **CHICKPEA IMPROVEMENT** supported by funds from **THE DEPARTMENT OF BIOTECHNOLOGY, GOVERNMENT OF INDIA, NEW DELHI, INDIA** and **THE MCKNIGHT FOUNDATION, U.S.A.**

The Principal Scientists, from each institute, who are involved directly with the research in this thesis, are:

Dr. V. S. Gupta,
Plant Molecular Biology Group,
Division of Biochemical Sciences,
National Chemical Laboratory,
INDIA.

Dr. J. A. Gatehouse
School of Biological Sciences,
University of Durham,
Durham,
U.K.

INDEX

ACKNOWLEDGEMENTS

During the years as a Ph.D. student, many people have contributed towards a meaningful and successful outcome of this research project. However, very few have stood apart by sheer virtue of excellence in their field and I have been fortunate to benefit from their guidance. Undoubtedly, working under the supervision of Dr. Vidya Gupta, as my Research Guide, has been the driving force behind all aspirations. Words would probably not suffice to express my gratitude towards her for accepting me as a student and guiding me through every step of learning.

It was a rare opportunity to work under the guidance of Dr. John Gatehouse, while at Durham. His strongly scientific temperament continues to be a benchmark, for others to follow.

Working with Abhay and Ashok during the course of this thesis has always encouraged me to strive further. For their willingness to come out of the way and offer help, I remain indebted to them.

The conception and progress of this research project has been positively gained due to valuable and critical inputs from Dr. Prabhakar Ranjekar, Dr. Mohini Sainani and Dr. Vasanti Deshpande. I am grateful to Dr. David Bown, Dr. Hillary Wilkinson and Dr. Elaine Fitches for all the help and guidance they have offered me. I reserve my special thanks for Dr. Meena Lagu, who continues to accurately assess and manage the scientific requirements of the group, in addition to research.

My friends have always stood by my side through good times and rough patches. Suhas and Bhushan keep reminding me that life very much exists outside the confines of science, while Janneth has never failed to raise my spirits with her morale-boosters.

In NCL, I enjoyed the company of seniors, Aparna, Arundhati, Bimba, Rajesh, Vrinda, Dipak, Sami, Mukund, Shashi, KN Rao, Raju, Gauri A., Vijay, Armaity, Swati, Aditi G., CV, Sudeep, as also Sanjay, Rajasekhar, Manasi, Ritika, Rahul, Mona, Ashwini H., Vaiju, Gauri B., Naren, Radhika, Manje, Sagar, Varsha, Hemangi, Gayatri, Prashant P., Prashant H., Pavan, Pankaj, Krishna, Elan, Neeta, Rashmi,

Ramya, Richa, Aditi C., Ashwini R., Tejas, Ram, Savita, Sofia, Atul, Santosh, and, of course, Ajit and Jagtap.

I would like to thank the Director, National Chemical Laboratory for making available all the facilities for my research. I also acknowledge financial grants from the Council of Scientific and Industrial Research, New Delhi in the form of Research Fellowships during the tenure of my work. I am grateful to Dr. Jamadagni, Mahatma Phule Agricultural University, Rahuri for extending all co-operation in making available the field facilities.

Further I also acknowledge all friends and well-wishers, whom I might have left out due to oversight, but whose best wishes have always encouraged me.

Last, and definitely by no means the least, I would like to make a mention of my parents, who have been eagerly looking forward to this personal milestone with much patience. No one else would ever be more pleased than them.

Ajay Srinivasan.

LIST OF ABBREVIATIONS

AI	amylase inhibitor
BA*p*NA	benzoyl-DL-arginyl-*p*-nitroanilide
BBI	Bowman-Birk type (proteinase) inhibitor
Bt	*Bacillus thuringiensis*
cakpi	*Cicer arietinum* Kunitz type proteinase inhibitor (gene)
CaKPI	*Cicer arietinum* Kunitz type proteinase inhibitor (protein)
CBB-R250	Coomasssie Brilliant Blue-R250
cDNA	complementary deoxyribonucleic acid
CI	chymotrypsin inhibitor
DAF	days after flowering
DEPC	diethyl pyrocarbonate
DNA	deoxyribonucleic acid
EDTA	ethylenediamine tetra acetate
g, mg, µg, ng	gram, milligram, microgram, nanogram
HCl	hydrochloric acid
HGPs	*Helicoverpa armigera* gut proteinases
HGPI	*Helicoverpa armigera* gut proteinase inhibitor
HRP	horseradish peroxidase
IC_{50}	inhibitor concentration at 50% inhibition
IPTG	isopropyl-thio-galactoside
JA	jasmonic acid, jasmonate
L, mL, µL	liter, milliliter, microliter
λ_{260}, λ_{280}	absorption at 260nm, 280nm
M, mM, µM	molar, millimolar, micromolar
mmole, µmole	millimole, micromole
MALDI-TOF	matrix assisted laser desorption ionization-time of flight
MeJ	methyl jasmonate
NaCl	sodium chloride
NaOH	sodium hydroxide
nm	nanometer
PAGE	polyacrylamide gel electrophoresis

PCR	polymerase chain reaction
PI	proteinase inhibitor
PIN-I/II	potato (*Solanum tuberosum*) proteinase inhibitor I/II
PMSF	phenyl methyl sulphonyl fluoride
RNA	ribonucleic acid
rpm	revolutions per minute
RT-PCR	reverse transcriptase-polymerase chain reaction
SA	salicylic acid, salicylate
SAAA*p*NA	succinyl-alanyl-alanyl-alanyl-*p*-nitroanilide
SAAAPL*p*NA	succinyl-alanyl-alanyl-alanyl-prolyl-leucyl-*p*-nitroanilide
SDS	sodium dodecyl sulphate
TI	trypsin inhibitor
TLCK	N-*p*-tosyl-l-lysine chloromethyl ketone
TPCK	N-*p*-tosyl-l-phenylalanine chloromethyl ketone
TRIS	tris-hydroxymethyl aminomethane
WBPI	winged bean (*Psophocarpus tetragonolobus*) proteinase inhibitor
X-Gal	2-bromo-3-chloro-4-indolyl-galacto-pyranoside

***Note:** The terms **HGPI**, *Helicoverpa armigera* Gut Proteinases Inhibitor, and **CaKPI**, *Cicer arietinum* Kunitz-type Proteinase Inhibitor, are to be considered as synonymous within the context of this manuscript.

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

SECTION 1: Introduction

Chickpeas are a very versatile legume and form an important part of the diet in the Indian sub-continent as well as many Middle-Eastern countries. It is widely accepted that chickpeas originated in the Middle East, though cultivation may have begun in the Mediterranean basin. During the 16th century, chickpeas were brought to other subtropical regions of the world by Spanish explorers as well as Indians who emigrated to other countries. Today, India ranks the first among the main commercial producers of chickpeas.

Chickpeas (*Cicer arietinum* L.), are mainly considered as an important part of a vegetarian diet due to the high quality protein content, and is also known to many as "the poor man's meat". But, as represented in **Table 1.1.1**, there is much more to the humble chickpeas than what meets the eye.

Table 1.1.1 Nutritional significance of chickpeas (*Cicer arietinum*)

Nutrient	Amount	% RDV	Nutrient	Amount	% RDV
The values indicate nutritional content of 1 cup (164 g) of cooked chickpeas.			VITAMINS		
			Beta carotene	28.37 mcg	
			Thiamin	0.19 mg	12.67
			Riboflavin	0.10 mg	5.88
			Niacin	0.86 mg	4.3
GENERAL			Vitamin B6	0.23 mg	11.5
protein	14.53 g		Vitamin C	2.13 mg	3.55
carbohydrates	44.95 g		Vitamin E	1.92	
Dietary fiber	12.46 g	49.84	Folate	282 mcg	70.52
Sugar	1.97 g		Pantothenic acid	0.47 mg	4.7
total fat	4.25 g				
Ω-3 fatty acids	0.07 g	2.8	AMINO ACIDS		
Cholesterol	0.00 mg		Alanine	0.62 g	
Water	98.74 g		Arginine	1.37 g	
Ash	1.71 g		Aspartate	1.71 g	
calories	268.96		Cysteine	0.20 g	48.78
			Glutamate	2.54 g	
MINERALS			Glycine	0.61 g	
Calcium	80.36 mg	8.04	Histidine	0.40 g	31.01
Copper	0.58 mg	29.00	Isoleucine	0.62 g	53.91
Iron	4.74	26.33	Leucine	1.03 g	40.71
Magnesium	78.72 mg	19.68	Lysine	0.97 g	41.28
Manganese	1.69 mg	84.5	Methionine	0.19 g	25.68
Molybdenum	123.00 mcg	164.00	Phenylalanine	0.78 g	65.55
Phosphorus	275.52 mg	27.55	Proline	0.60 g	
Potassium	477.24 mg		Serine	0.73 g	
Selenium	6.07 mcg	8.67	Threonine	0.54 g	43.55
Sodium	11.48 mg		Tryptophan	0.14 g	43.75
Zinc	2.51 mg	16.73	Tyrosine	0.36 g	37.11
			Valine	0.61 g	41.5

mcg, millicentigram; %RDV, percent daily recommended (intake) value

As evident from **Table 1.1.1**, apart from the nutritious protein content, chickpeas are also a good source of soluble and insoluble fiber, as well as carbohydrates, vitamins and minerals. The high fiber content prevents blood sugar levels from rising too rapidly after a meal, making chickpeas an especially good choice for individuals with diabetes, insulin resistance or hypoglycemia. Soluble fiber is associated with control of cholesterol, and thus reduction in risk of coronary ailments. On the other hand, insoluble fiber not only helps prevent digestive disorders like irritable bowel syndrome and diverticulosis, but has also been associated with reducing the risk of colon cancer. While the folate content could help to lower the levels of homocysteine, and reduce the risk factor for heart attack and stroke, the magnesium could aid relaxation of blood vessels and improve the flow of blood, oxygen and nutrients throughout the body. The iron content of chickpeas can be considered as a safer alternative to red-meat, due to the low calorie-count and almost negligible fat content in the former. Chickpeas are an excellent source of the trace mineral manganese, which is an essential cofactor in a number of enzymes important in energy production and antioxidant defenses, e.g., superoxide dismutase. Apart from these health benefits, chickpea crop is also beneficial to the environment in many ways. Being a legume, the plant produces nitrogen-fixing nodules, which can enrich the soil with at least 50 kg of N/ha every season. Further, chickpea root exudates, which are rich in citric acid, help to dissolve calcium phosphates and help in mobilization of phosphorous to readily usable forms.

In addition to the dry seeds that are commonly used for cooking, green seeds of chickpeas are also consumed raw, as are the young shoot tips, as 'green' vegetable. It is no surprise then, that chickpeas are the third most important legume crop with a worldwide production of about 9.2 million Mt (Million tones) (**Fig. 1a**), in addition to being the most important legume crop in India. In fact, India alone accounts for ~60% of the world's total production of chickpea (**Fig. 1b**) (http://faostat.fao.org, 13thFebruary 2005). Though, chickpeas are grown and locally consumed, India is also the world's largest importer of chickpeas accounting for about 20% of global imports (Source: http://faostat.fao.org, 13th February 2005). These figures reflect on the growing demand for chickpea as well as the immense strain on crop production and

Fig. 1.1.1. World chickpea scenario.
Chickpeas (*Cicer arietinum*) ranks third among the most important legumes in the world. It is, however, the most important legume crop in India; India is the world's largest producer of chickpeas, contributing to ~60% of the total global produce.
(A) Global production of chickpea vis-à-vis other legume crops.
(B) Contribution of Indian agriculture to global production of chickpea.
ROW, Rest of the world.
Source: FAO Data (http://faostat.fao.org)

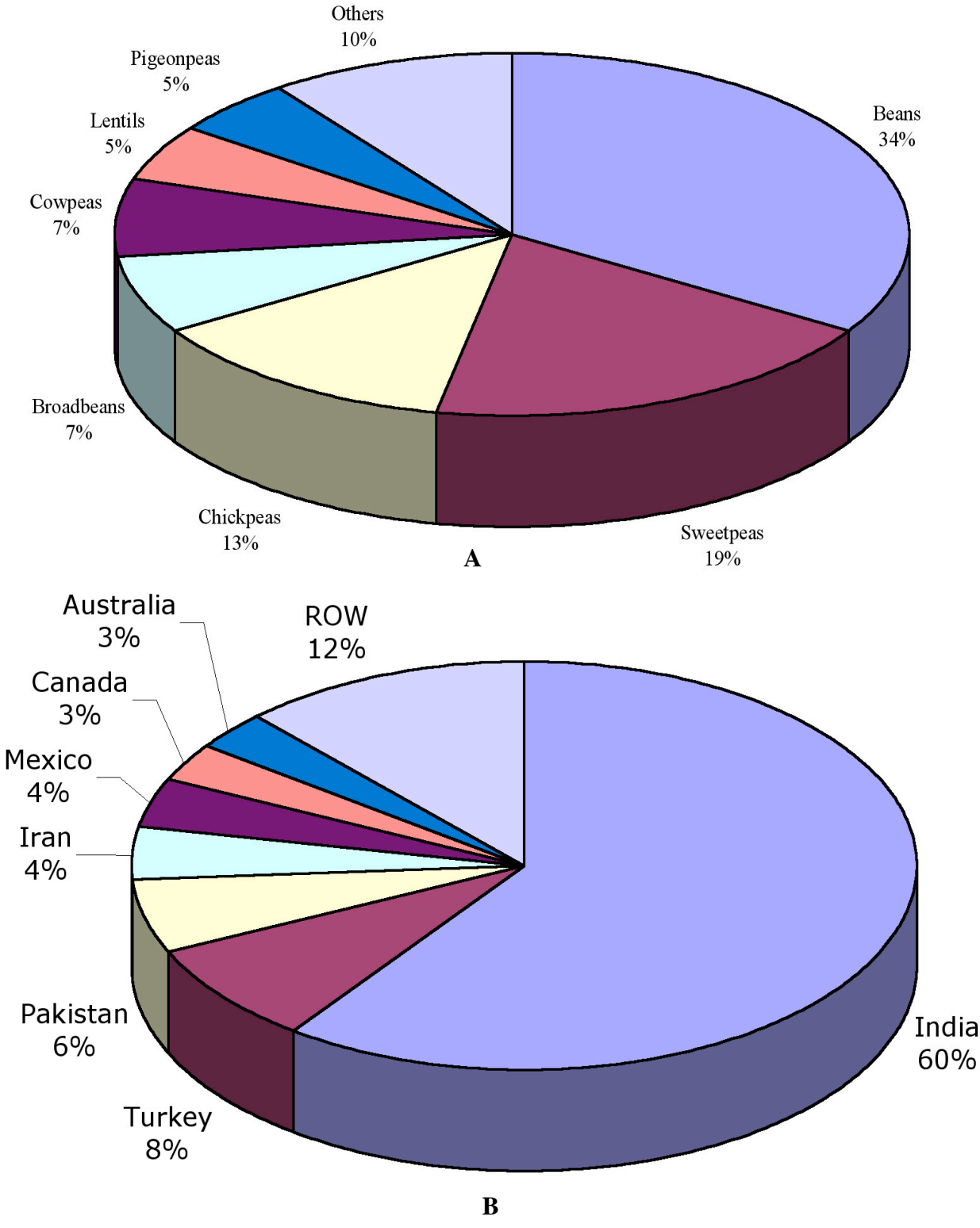

Others
10%
Pigeonpeas
5%
Lentils
5%
Cowpeas
7%
Broadbeans
7%
Chickpeas
13%
Sweetpeas
19%
Beans
34%
A
Australia
3%
ROW
12%
Canada
3%
Mexico
4%
Iran
4%
Pakistan
6%
Turkey
8%
India
60%
B

Fig. 1.1.1

yield. Since chickpeas are mainly cultivated as a rain-fed crop, annual production has been observed to suffer due to disturbances in weather patterns. Chickpea crops face losses from abiotic stresses such as drought and salinity as well as serious threats from fungal pathogens and insect pests. The prevalent fungal pathogens include *Fusarium oxysporum* sub-sp. Ciceri, which causes Fusarium Wilt disease, and *Aschochyta rabiei*, which causes Aschochyta Blight. Among the insect pests, Podborer (*Helicoverpa armigera* Hübn.) (**Fig. 2**) causes most devastating damages to standing crops than any other insect and amounts to almost 30% losses in yield (Thomas, 1999). The *H. armigera* (female) moth lays over 500 eggs and, upon hatching, the larvae feed voraciously on leaves, flowers as well as developing chickpea seeds. Due to continuous availability of alternate host plants like pigeonpea (*Cajanus cajan*), cotton (*Gossypium hirsutum*), sweetpea (*Pisum sativum*), tomato (*Lycopersicon esculentum*) and okra (*Abelmoschus esculentus*), *H. armigera* is a continuous menace throughout the cropping season. This polyphagous nature of the insect pest allows for survival in a diverse cropping environment, with no dearth of readily available food material in the form of standing crops. It has been known for more than a decade that *H. armigera* can infest and feed on as many as 180 different plant species (Manjunath *et al*, 1989), thus earning it the dubious distinction of being the "number one" insect pest. It is predicted that this already broad host range will continue to expand in the future to cover other agronomically important crops as well.

While the "green revolution" ushered in a new era of modern agricultural practices coupled with traditional knowledge to increase agricultural output, such advances have not born any fruit in the control of insect pests. The continuous abuse of chemical pesticides has led to development of pesticide-resistant insect pests, which in turn has necessitated a shift in focus towards other sustainable means such as those enunciated by the "Integrated Pest Management" (IPM) practices. These include strategic use of traditional insect control means (birds, parasitoid insect species) along with semi-modern practices such as pheromone traps and *Nuclear Polyhedrosis Virus* (NPV) applications. Advances in genetic engineering and related tools have opened up new vistas by facilitating transfer of coding DNA across different species, thus allowing us to "engineer" insect resistance in target crop plants by the introduction of genes, whose products could be toxic or antagonistic towards

Fig. 1.1.2. Damage to chickpea crop by podborer (*Helicoverpa armigera*).
Podborer (*H. armigera*) is a polyphagous Lepidopteran pest of more than 180 different plant species. Reports describe development of tolerance towards most of the synthetic pesticides as well as some biological agents such as the *Bt* toxin. Podborer is the most significant insect pest on chickpeas and is present throughout the cropping season. Young larvae of this insect initially feed on leaves and flowers before migrating to the developing seeds at the third instar.
(A) Healthy developing chickpea pods.
(B) Third instar larva feeding voraciously on pod.
(C) Pods damaged due to larval feeding.
(D) Pheromone traps and predatory birds in chickpea field.
(E) Healthy mature seeds

A

B

E

C

D

Fig. 1.1.2

insect growth and development; one of the foremost being the *Bacillus thuringiensis* δ-endotoxin (*Bt* toxin) gene. In addition, other potential agents belonging to varied mechanistic classes such as proteinase inhibitors (PIs), amylase inhibitors (AIs) and lectins have also been identified for this purpose. Simultaneously, adverse effects of transgenic approaches on friendly or useful insects as well as any negative impacts on the ecosystem are also being probed. The introduction of foreign gene(s) into plants is scrutinized for any unfavorable effects on crop yield, owing to channeling of metabolic pathways for synthesis of the transgene product. These investigations would go a long way in addressing many of the social, legal and ethical issues that surround the development of transgenic crop varieties. Nevertheless, it has been realized that in the absence of viable options to counter the challenges of insect pests, we would face greater dilemmas in future.

Organization of thesis:

The objective of the current research project is to identify and characterize defensive proteinase inhibitor (PI) from chickpea, *viz*, CaKPI, as well as to probe it's efficacy as a potential candidate to develop insect resistance in chickpea as well as other host plants. This thesis has been divided as further:

Chapter I: Introduction and Review of Literature

- **Section 1:** Introduction (the current section)
- **Section 2: "**Molecular basis of pest offenses, host defenses and engineering of resistance - a review of literature" is a concise essay on our current understanding of Lepidopteran insect pests and various approaches to control herbivore infestation, with emphasis on plant proteinase inhibitors (PIs), which are natural defensive biomolecules.

Chapter II: Atypical Features of a Proteinase Inhibitor from Chickpeas

This chapter describes the novel features encountered during the processes of identification, purification and genetic characterization of CaKPI. This chapter is divided into three sections (the titles are self-explanatory):

- **Section 1:** Purification and identification of CaKPI
- **Section 2:** Isolation and properties of *cakpi*

- **Section 3:** Genetic diversification and differential expression of *cakpi*

Chapter III: Defensive Aspects of CaKPI in Chickpea-Podborer Association

This chapter describes the functional characterization of CaKPI, *viz*., inhibitory activity against commercial and insect digestive proteinases, as well as antagonistic nature towards the physiology of the insect pest, *Helicoverpa armigera*. This chapter consists of two sections (the titles are self-explanatory):

- **Section 1:** *in vivo* and *in vitro* activities of CaKPI
- **Section 2:** *Helicoverpa armigera* does not adapt to dietary CaKPI

Chapter IV: "Significance of Structural and Functional Diversities on Proteinase-Proteinase Inhibitor and Plant-Pest Interactions: a theoretical study"

This discussion deals with the activity and inhibition of serine proteinases with respect to structural and functional variations arising due to mutations/natural variations. This is correlated to the multitude of serine proteinase genes observed in Lepidopterans like *H. armigera* and the implications on adaptation to host-plants and responses to/fate of dietary PIs.

Chapter V: Thesis Summary and Future Directions

The penultimate chapter summarizes the salient features of the chickpea-podborer association that have been uncovered during the course of this study. Approaches based on current findings, that could delve deeper into this association as well as contribute to the control of herbivorous insect pests are also explored.

Chapter VI: Bibliography

This chapter enumerates the previously published research articles that have been invaluable during conception and progress of the current study. Manuscripts arising out of the research work embodied in this thesis have also been enlisted.

SECTION 2: Molecular basis of pest offences, host defenses and engineering of resistance – a review of literature

More often than not, plant-insect associations have disastrous consequences from an agroeconomic perspective. The damages caused by herbivorous insects, in terms of crop losses, translates not only into lowered yield of the produce, but also into severe economic losses for farmers. These losses lead to a detrimental cascade of other socio-economic issues like inability to repay agricultural loans and other debts; these problems are especially significant when considering that the majority of Indian farmers are smaller stakeholders, who carry out farming in small areas with very little access to modern practices. This far-reaching threat of insect attack to agriculture has prompted studies to dissect and understand the physiological mechanisms of host plant-insect pest interactions, which determines the fate of the association. As a result of relentless efforts, crucial insights have been gained into the molecular processes underlying these interactions (Gatehouse, 2002).

1) <u>Lepidopteran herbivorous pests</u>

Lepidoptera (moths and butterflies) is the second largest order in the class Insecta. Most Lepidopteran larvae are herbivores; some species eat foliage, some burrow into stems or roots, and some are leaf-miners. Some adult Lepidopteran moths can lay up to 500 eggs at a time, and, upon hatching and emergence, the larvae immediately begin foraging and feed voraciously on available food sources (Waterhouse, 1957; Telang *et al*, 2000). The larvae go through various stages of development (instars), punctuated by molting events, before entering the dormant (pupal) stage. During the pupal stage, the insect undergoes dramatic changes in morphology and physiology and terminates in the adult (moth/butterfly) phase, thus completing a life cycle.

(a) <u>Economic importance</u>

Although many Lepidoptera, especially butterflies, are valued in the aesthetic sense for their beauty and from a horticultural point as pollinators, only a few such as the silkworm (*Bombyx mori*) are exploited commercially. Otherwise, this order represents one of the most destructive groups of insects; the ubiquitous nature of Lepidopteran insect pests is a threat to many agriculturally important crops. The host plant range for Lepidopteran pests may either be narrow (monophagous), e.g., *Manduca sexta*, which exhibits preference for solaneceous plants, or diverse (polyphagous), e.g., *Helicoverpa armigera*, which feeds on various legumes, vegetables and fruits. Lepidopteran pest mediated damage accounts for the highest

Table 1.2.1 Major Lepidopteran insect pests (http://www.aisglobal.net/lists.html)

Insect Pest		Target Crop(s)	Digestive Proteases*
Scientific Name(s)	**Common name(s)**		
Helicoverpa armigera, H. zea, eliothis virescens	Podborer, Tobacco budworm, Corn earworm, Tomato fruitworm, Sorghum headworm, Cotton bollworm	Among the 180 different reported hosts, the major ones are: Alfalfa (*Medicago sativa*), Broccoli (*Brassica oleracea* var. *italica*), Cabbage (*Brassica oleracea*), Chickpea (*Cicer arietinum*), Chrysanthemum (*Chrysanthemum coronarium*), Cotton (*Gossypium hirsutum*), Gardenpea (*Pisum sativum*), Kidney beans (*Phaseolus vulgaris*), Lentils (*Lens culinaris*), Lettuce (*Lactuca sativa*), Maize (*Zea mays*), Okra (*Abelmoschus esculentus*), Peanut (*Arachis hypogea*), Pepper (*Capsicum annum*), Potato (*Solanum tuberosum*), Pigeonpea (*Cajanus cajan*), Sorghum (*Sorghum bicolor*), Soybean (*Glycine max*), Strawberry (*Fragaria virginiana*), Sweet potato (*Ipomoea batatas*), Tobacco (*Nicotiana tobacum*), Tomato (*Lycopersicon esculentum*), Watermelon (*Citrullus lanatus*)	trypsin (90%), chymotrypsin (5%), elastase (1%), carboxypeptidase (1%), aminopeptidase (1%), cathepsin B-like (1%), metalloproteases (1%) (Gatehouse *et al*, 1997)
Spodoptera litura, S. exigua, S. frugiperda	Tobacco cutworm, Cotton bollworm, Beet armyworm, Fall armyworm	Beet (*Beta vulgaris*), Cabbage, Cotton, Cowpea (*Vigna unguiculata*), Eggplant (*Solanum melongena*), Gardenpea, Kidney bean, Onion (*Allium cepa*), Peanut, Pepper, Potato, Radish (*Raphanus sativus*), Safflower (*carthamus tinctorius*), Soybean, Sweet potato, Tobacco, Tomato	trypsin (7%), chymotrypsin (85%), elastase (1%), aminopeptidase (5%) carboxypeptidase (1%) (Broadway and Duffey,1986)
Manduca sexta	Tobacco hornworm	Eggplant, Tobacco, Tomato	trypsin (10%), chymotrypsin (80%), elastase (1%), aminopeptidase (Johnson *et al*, 1989)
Pectinophora gossypiella	Pink bollworm	Cotton	(none reported)
Pieris rapae	Imported cabbageworm	Alfalfa (*Medicago sativa*), Broccoli, Brussels sprout (*B. oleracea* var. *gemmifera*), Cabbage, Cauliflower (*B. oleracea* var. *botrytis*), Horseradish (*Armorecea rusticana*)	(none reported)
Plutella xylostella	Diamondback moth	Broccoli, Brussels sprout, Cabbage, Cauliflower, Horseradish, Mustard (*Brassica nigra*)	trypsin (major), chymotrypsin (major), elastase, aminopeptidase (DeLeo *et al*, 2001)
Agrotis ipsilon	Black cutworm	Broccoli, Cabbage, Carrot (*Daucus carota*), Eggplant, Green beans, Mustard, Potato, Spinach (*Spinacea oleracea*), Sugarcane (*Saccharum officinarum*)	trypsin (major), chymotrypsin (Mazumdar-Leighton and Broadway, 2001)
Anticarsia gemmatalis	Velvetbean caterpillar	Cowpea, Horsebean (*Parkinsonia aculeate*), Peanut, Soybean, Velvet bean (*Mucuna pruriens*),	trypsin (major), chymotrypsin, cathepsin B-like (Oliviera *et al*, 2005)

***Note:** Figures in parentheses represent approximate percent contribution to total gut protease activity.

in legumes like chickpea (*Cicer arietinum*), pigeonpea (*Cajanus cajan*), sweetpea (*Pisum sativum*) as well as in other important crops like cotton (*Gossypium hirsutum*), tobacco (*Nicotiana tobacum*), okra (*Abelmoschus esculentus*) and tomato (*Lycopersicon esculentum*) (Manjunath *et al*, 1989). A list of some of the major lepidopteran pests of crops is given in **Table 1.2.1**.

(b) Digestive mechanism and components

The insect digestive system can be roughly divided into the fore-, mid- and hindgut with respect to its position along the anterior-posterior axis of the body (Waterhouse, 1957). The major dietary components of Lepidopteran larvae are proteins from plant tissue, which is completely macerated by strong mandibles into a semi solid form prior to entering the foregut. As the food passes through the midgut, it is subject to enzymatic breakdown, the products of which are then assimilated. The hindgut shows very little digestive function, but may be involved in assimilation. The presence and/or contribution of gut microflora, to digestive processes have not yet been unambiguously proven. The undigested matter is excreted along with residual, but active enzymes - a peculiar feature of some Lepidopterans like *H. armigera* (Patankar *et al*, 2001). The high gut pH (Gringorten *et al*, 1993) and presence of free glycine in the digestive juices (Konno *et al*, 1997), which are hypothesized to protect the insect physiology from various anti-nutritional factors, are other peculiar characteristics. The inner side of the digestive gut is lined by a seamless component known as the peritrophic matrix (PM), which forms the barrier between the ingested material and the gut cell wall. The PM is known to contain digestive proteases (Bolognesi *et al*, 2001) as well as transporter proteins, which contribute to the passage of digested proteins (amino acids) to the gut epithelial cells for assimilation (Lehane, 1997; Tellam *et al*, 1999; Terra, 2001; Wang and Granados, 2001), in addition to protecting the insect from plant-derived toxic allelochemicals (Barbehenn, 2001).

(c) Digestive gut proteinases

Lepidoptera rely heavily on proteases for their digestive processes (Waterhouse, 1957; Telang *et al*, 2000). The digestive complement consists of endo-peptidases like serine, metallo- and cathepsin-B like proteinases, and, the exo-peptidases, *viz.*, amino- and carboxypeptidases. Many of these proteinases have been isolated, identified and the coding DNA/cDNA fragments have also been well

characterized (Zhu *et al*, 1997; 2000a; 2000b; Bown *et al*, 1998; Girard *et al*, 1999; Valaitis *et al*, 1999; Mazumdar-Leighton *et al*, 2000; Patankar *et al*, 2001; Bayes *et al*, 2003; Hegedus *et al*, 2003; Herrero *et al*, 2005); serine proteinases form the dominant mechanistic class (>95%) in the gut environment. The digestion in the larval gut might follow a logical trend, where by the complex proteinaceous material is broken down into much smaller oligo-peptides by the endo-peptidases like trypsins and chymotrypsins, and these oligo-peptides are further digested by the exo-peptidases liberating free amino acids, from the N- and C-termini. In such a situation, the endo-peptidases like trypsins and chymotrypsins contribute to the primary digestive system in larvae, whereas the exo-peptidases constitute the secondary digestive system. Although such a "division of labor" works well to the benefit of streamlining the digestive process, compartmentalization of the primary and secondary digestive proteinases has not yet been proven, and instead both processes appear to be operating in parallel, seamlessly throughout the midgut.

(d) <u>Differential proteinase genes expression and polyphagy</u>

Lepidopteran pests are armed with a fearful arsenal of digestive proteinases; while the gut of *H. armigera*, for example, consists of at least thirty active isoforms of various mechanistic classes. However, the relative levels of expressed gut proteinase genes vary between larvae feeding on various host plants (Harsulkar *et al*, 1999; Patankar *et al*, 2001; Chougule *et al*, 2005). This interesting feature was attributed to differential expression of digestive proteinases in response to a change in the chemical properties of dietary protein (Broadway *et al*, 1996). Such a dynamic system of gene expression allows the insect to effectively digest the available plant proteins with a minimal complement of proteinases and also avoids unnecessary expenditure of metabolic energy by curtailing expression of other proteinase genes with similar functions, and is immensely helpful in adapting to the varying protein composition seen in different hosts; this mechanism holds good for explaining the polyphagous nature of insects like *H. armigera*, which allows it to infest a wide variety of agriculturally important crops (Patankar *et al*, 2001). Although the exact nature of the regulatory mechanism governing differential expression of proteinase genes is not known, recent studies have identified neuro-peptide like proteins (Harshini *et al*, 2002a; 2002b; Davey *et al*,

2005), which have the ability to "flick the switch of molecular expression", resulting in polyphagy. Considering the significance of this regulatory mechanism, any inroads towards its understanding would prove beneficial in strategies to control herbivore mediated damage to agriculturally important crops.

2) Control of herbivorous insects

The agronomic importance of herbivorous insect pests cannot be exaggerated; the worldwide losses pose a massive challenge to global agricultural produce and threaten economies of developing countries. It is no surprise, then, that control of such pests have received constant attention and have been integral parts of programs such as the green revolution, which aimed at improving the quality and quantity of agriculture. Some of the means to control insect pests are summarized as follows:

(a) Traditional and semi-modern practices

Traditional practices largely represent approaches based on minimal intervention by humans and make use of natural resources and basic knowledge and are best exemplified by use of natural enemies to insect pests such as birds, predatory insects such as wasps and ladybirds, as well as natural parasitoid insects. Another effective means is to physically dislodge larvae from plants by vigorous shaking, as is commonly practiced with and has been observed to be effective in crops like pigeonpea and tobacco. These basic methods are usually employed in tandem with semi-modern means, *e.g.*, use of bird perches along with pheromone traps to attract moths, and, use of 'trap crops' that attract the insect pests and divert them from the main crop (Grundy *et al*, 2004; Jallow *et al*, 2004). These means rely on existing populations of natural enemies of insect pests and can hence be unpredictable. Crude extract of plant parts, such as neem leaves, are also commonly employed to discourage larvae from crop plants. In general, these practices represent an economical means to afford basal level of protection to crops.

(b) Chemical pesticides

The first synthetic compound to be widely employed as pesticide was dichloro-diphenyl-trichloroethane (DDT), which was introduced in 1946 (Elzinga, 1978). Following the perceived success of DDT, other generic chemicals such as benzene hexachloride, organophosphates, synthetic pyrethroids were introduced as

alternatives. However, lack of thorough understanding of these chemicals led to uncontrolled usage, resulting in side-effects, which were more damaging than helpful in the long run. One of the first drawbacks resulting from overuse of these recalcitrant chemicals was their persistence in vegetation as well as contamination of soil and water bodies. These residues found way into human consumption via foodgrains, vegetables, milk, poultry products, fish, etc leading to symptoms of chronic as well as acute poisoning. The disastrous consequences of these chemicals included ailments to eyes, skin and lungs as well as causing a variety of cardiovascular, immunological and neurological disorders (Soon, 1997). From an agricultural perspective, synthetic insecticides have a non-specific mode of action and hence have also been observed to affect beneficial insects (Armes *et al*, 1996). Secondly, emerging incidences of pesticide resistance (Jadhav and Armes, 1996) in insect have forced a re-assessment of such chemicals for agricultural use as well as prompted research towards identifying other environment friendly means.

(c) <u>Biological agents</u>

The initial failure of chemical pesticides to provide a safe, long-term and effective means for insect control has led to focus towards biological agents. These include saprophytic fungi, insect viruses, bacteria as well as toxic proteins derived from these organisms. Some saprophytic fungi that have been identified for controlling insect pests include *Metarhizhium* (Nahar *et al*, 2004), *Nomuraea* (Devi *et al*, 2003), *Beauveria* (Hazzard *et al*, 2003), etc., which secrete enzymes like chitinases that degrade the protective cuticular layer of the insect larvae and lead to innervations of larval mycelia followed, ultimately, by death of the insects. Fungal chitinases and related enzymes are also being evaluated for their stand-alone effectiveness in controlling insect populations. The most well studied and characterized among biological agents is the *Bacillus thuringiensis* δ-endotoxin (*Bt* toxin), which is accumulated as a crystalline inactive pre-protein (inclusion bodies) in the bacterium. The bacterium *B. thuringiensis* was discovered 1901 in Japan and 1911 in Germany by Ernst Berliner, who reported a disease called "Schlaffsucht" in caterpillars of the flour moth, *Anagasta kuehniella* (http://en.wikipedia.org/wiki/Bacillus_thuringiensis). When the *Bt* pro-toxin is ingested by the insect, the digestive proteinases breakdown the pre-protein into the active form, which binds to membrane receptors on the gut epithelial cells and

causes formation of pores leading to cytolysis and death of the insect (Dean *et al*, 1996; Brousseau *et al*, 1999; Frutos *et al*, 1999; Aronson and Shai, 2001). The purified form of the *Bt* toxin protein has also been shown to be stable and effective for 'stand-alone' application. The nuclear polyhedrosis virus (NPV) is a commonly used viral agent for insect control – replication of the virus within the insect causes a cellular breakdown and death of the insect. NPV represents an economical as well as an effective means to control insect pests with minimal technological setup, though its specificity towards pests is debatable. However, biological agents are a viable and healthier alternative to the synthetic chemical pesticides, because they cause little or no damage to the environment and are generally harmless to other animals, including humans.

(d) <u>Transgenic approaches</u>

Among many applications of transgenic technology, one of the foremost areas of research aims to develop resistance in plant species towards insect pests. This approach makes use of known bio-molecules, which posses antagonistic activity against insects and has been proven effective in insect control (Boulter, 1993; Duffey and Stout, 1996; Jouanin *et al*, 1998; Hilder and Boulter, 1999; Carlini and Grossi-de-Sa, 2002; Groot and Dicke, 2002; Lajolo and Genovese, 2002; Murdock and Shade, 2002; Babu *et al*, 2003; Simmonds, 2003; Ferry *et al*, 2004; Hartmann, 2004; Romeis *et al*, 2004). Following the initial success of *Bt* toxin against insect pests, it was employed for genetic modification of plants. Development of *Bt*-transgenics in crops like tobacco (Barton *et al*, 1987), cotton (Gupta *et al*, 2000; Carriere *et al*, 2003) was followed by thorough trials to determine the afforded insect resistance. As of today, commercial transgenic crops carry variants including *cry1Aa* (cotton, potato), *cry1Ab* (cotton, maize, potato, tobacco, tomato, rapeseed), *Cry1Ac* (cotton, maize, potato, rapeseed, tobacco, tomato) and *Cry3A* (alfalfa, canola, cotton, eggplant, maize, potato, rice, tomato) (http://www.icrisat.org/gt-bt/ResearchBreifs/TABLE2.HTM). However, in order to develop a library of naturally occurring antagonistic agents, attention was also given towards other biomolecules – one of the foremost examples would be the plant derived proteinase inhibitor proteins (PIs) (Garcia-Olmedo *et al*, 1987;

Fig. 1.2.1. Detrimental effects of various antagonistic agents on digestion, growth and development in Lepidopteran larvae

Herbivorous Lepidopteran larvae feed voraciously on plant parts to derive nutrients for optimum growth and development. The primary constituent of their diet is protein, which is digested into amino acids by proteases. Similarly, complex polysaccharides are broken down into simple sugars by amylases. The monomeric forms, i.e., amino acids and sugars are absorbed and assimilated for growth processes, that lead to normal development into healthy adult moth. Any impairment in digestion by antagonistic agents like proteinase inhibitors (PIs) and amylase inhibitors (AIs) leads to developmental malformations. Other agents like lectins, that affect nutrient absorption across the midgut epithelium also lead to similar effects.

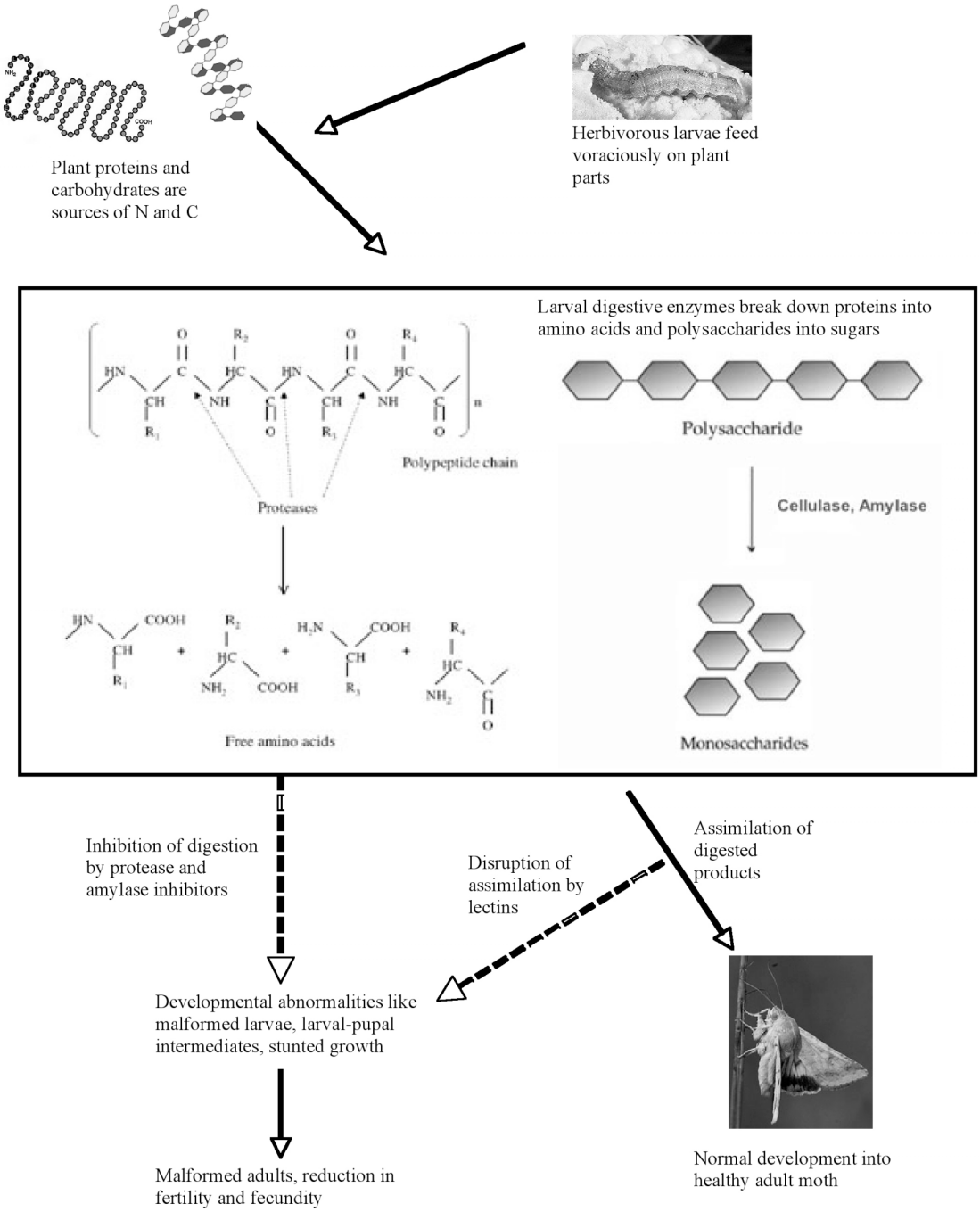

Plant proteins and carbohydrates are sources of N and C
Herbivorous larvae feed voraciously on plant parts
Larval digestive enzymes break down proteins into amino acids and polysaccharides into sugars
Polypeptide chain
Proteases
Free amino acids
Polysaccharide
Cellulase, Amylase
Monosaccharides
Inhibition of digestion by protease and amylase inhibitors
Disruption of assimilation by lectins
Assimilation of digested products
Developmental abnormalities like malformed larvae, larval-pupal intermediates, stunted growth
Malformed adults, reduction in fertility and fecundity
Normal development into healthy adult moth

Fig. 1.2.1

Hilder *et al*, 1987; Ryan, 1990; Ussuf *et al*, 2001; Haq *et al*, 2004; Christeller and Laing, 2005; Giri *et al*, 2005). PIs are naturally encountered in many plant species and varieties; their expression varies between different plant tissues, including leaves, flowers and fruits/seeds. Since PIs are derived from various plants, they were and still are considered as a safer alternative to agents derived from bacteria (*e.g.*, *Bt* toxin). The possible role of PIs in plant protection was investigated following early observations (Mickel and Sytandish, 1947) on abnormal development of insect larvae on soybean-based products. Subsequently the trypsin inhibitors present in soybean were shown to be toxic to the larvae of flour beetle, *Tribolium confusum* (Lipke *et al*. 1954). Following these early studies, there have been many examples of the antagonistic activity of PIs against insect pests, by *in vitro* assays against insect gut proteases as well as by *in vivo* (insect-feeding) assays. Most of the reported plant PIs originate from three main families namely Leguminosae, Solanaceae and Gramineae (Richardson, 1991); some PIs from Cucurbitaceae have also been described (Telang *et al*, 2003). One of the first PIs to be used for genetic transformation was the cowpea (*Vigna unguiculata*) trypsin inhibitor (CpTI), (Hilder *et al*, 1987) which afforded protection to tobacco against insect pests. Further research and development led to establishment of many PI-transgenic varieties as elaborated in **Table 1.2.2**. Amylase inhibitors (AIs), such as the barley (*Hordeum vulgare*) bi-functional α- amylase/trypsin inhibitor are effective against pests of stored grains, especially weevils, which rely on digestive enzymes like α-amylases to meet their nutritional requirements (Garcia-Olmedo *et al*, 1987). Lectins are another class of potential biomolecules due to their proven antagonistic activity against insects, *e.g.*, the Snowdrop lectin (GNA, *Galanthus nivalis* Agglutinin) (Down *et al*, 1996; 2000; 2001; Fitches *et al*, 1997, 1998, 2001a; 2001b; Bell *et al*, 1999; Birch *et al*, 1999; Foissac *et al*, 2000). Lectins deal a blow to insect physiology by binding to gut epithelial receptors and interfering with nutrient assimilation processes. Thus, identification of multiple classes of biomolecules has led to a virtual repository of antagonistic domains, which can be employed in various combinations (Fitches *et al*, 2002; 2004; Brunelle *et al*, 2005) by means of gene-fusion techniques. The resulting "mega-molecules" posses multiple domains with different activities and could prove omnipotent in control of more than one species of insect pests.

Table 1.2.2 PIs used to develop transgenic plants

PI source	Transformed Plant	Target insect class(es)
Cowpea (*Vigna unguiculata*) TI	Apple (*Malus domestica*), Lettuce (*Laciuca sativa*), Oilseed rape (*Brassica napus*), Potato, Rice Strawberry (*Fragaria virginiana*), Sunflower (*Helianthus annus*), Sweet potato (*Ipomea batatas*), Tobacco, Tomato	Coleoptera, Lepidoptera
Soybean (*Glycine max*) TI	Oilseed rape, Poplar (*Populus tremuloides*)[1], Potao, Rice[2], Tobacco, Tomato	Coleoptera, Lepidoptera
Potato (*Solanum tuberosum*) PI	Birch (*Betula spps.*), Lettuce, Petunia (*Picea spps.*), Rice, Tobacco	Lepidoptera, Orthoptera
Tomato (*Lycopersicon esculentum*) PI	Alfalfa (*Medicago sativa*), Tobacco, Tomato	Lepidoptera
Mustard (*Brassica nigra*) TI	Arabidopsis, Tobacco[3]	Lepidoptera
Barley (*Hordeum vulgare*) TI	Tobacco	Lepidoptera
Giant taro (*Alocasia macrorrhiza*) TI	Tobacco[4]	Lepidoptera
Tobacco (*Nicotiana alata*) PI	Peas (*Pisum sativum*)[5]	Lepidoptera
Squash (*Cucurbita maxima*) TI	Tobacco	Lepidoptera
Rice (*Oryza sativa*) PI	Arabidopsis[6], Oilseed rape, Poplar, Potato[7], Tobacco	Coleoptera, Homoptera

Adapted from Schuler *et al* (1998). Specific references marked in table: [1]Confalonieri *et al*, 1998; [2]Lee *et al*, 1999; [3]DeLeo *et al*, 1998; [4]Wu *et al*, 1997; [5]Charity *et al*, 1997; [6]Walker *et al*, 1999; [7]Cloutier *et al*, 1999

(e) Insect resistance and Integrated Pest Management

In addition to disastrous effects on environment and human health, the reckless and irresponsible use of synthetic pesticides have led to development of insect pests resistant to most of these chemicals, as in case of *H. armigera*, which has become recalcitrant towards the commonly used synthetic pyrethroids among other types of pesticides (Armes *et al*, 1996; Jadhav and Armes, 1996). The recent decades have also seen emerging events of insect pests becoming resistant to some variants of the *Bt* toxin. Though *Bt* offers insect resistance, the 'wipe-out' effect of *Bt* due to the direct killing exerts a strong selection pressure on the insects due to which *Bt* resistance is also commonly observed (Milne *et al*, 1995; Forcada *et al*, 1996; Keller *et al*, 1996; Oppert *et al*, 1997; Brousseau *et al*, 1999; Frutos *et al*, 1999; Zhu *et al*, 2000c; Akhurst *et al*, 2003; Bird *et al*, 2005; Gahan *et al*, 2005; Ma *et al*, 2005). Likewise negative impacts to the environment (Saxena *et al*, 1999) and on populations of beneficial insects (Losey *et al*, 1999) have also forced a reassessment of *Bt*. In this evolutionary race between insects

and their antagonists, insects seem to have the leading edge. However, prudent planning and careful use of available resources seems to be the solution to keep this menace, in check. Although, it is impossible to predict a clear winner in the long term, a winning strategy would employ a combination of means to achieve long term protection, while providing immediate relief. Such are the principles of Integrated Pest Management (IPM), which exhorts development of sustainable strategies that employ a combination of traditional, semi-modern as well as transgenic approaches (Soon *et al*, 1997; Thomas, 1999), *e.g.*, rotation of crops along with use of pheromone traps and encouragement of natural parasitoids and enemies along with limited use of pesticides, including those derived from natural (plant) sources, and finally pyramiding of different mechanistic classes of genes, such as PIs and *Bt*. The advantage of such approaches lies in their flexibility to be adapted to various insect pests depending on the varying levels of threat. Such multi-pronged strategies would ensure crop survival under extreme threat as well as present limited opportunities for the insect to gain an immediate upper hand.

3) <u>Legume proteinase inhibitors</u>

The group Leguminoseae (Fabaceae) represent a major source of dietary protein-rich crops such as lentils (*Lens culinaris*), soybeans (*Glycine max*), sweetpeas, chickpeas and pigeonpeas. The bulk of the dietary proteins in legumes are derived from the seeds where they constitute up to 25% of total seed content and are mainly present as the inactive storage proteins (Altschul *et al*, 1965; Sales *et al*, 2000). Legume PIs are exclusively found in seed tissue, where they constitute upto 5% of the total soluble protein content. Most legume seed PIs are inhibitors of serine proteinases like trypsin or chymotrypsin, although some of them may also exhibit inhibitory activity against elastases, subtilisin-like proteinases, etc. Initial hypotheses suggested that legume seed PIs function to protect the storage proteins from endogenous plant proteinases. However, two observations contrast this belief, (i) most endogenous plant proteinases which govern protein mobilization in storage tissues, are cysteines proteinases, and, (ii) serine proteinases are not involved in large scale protein digestion in plants (Reeck *et al*, 1997) – thus the presence of significant quantities of serine PIs in plants do not appear to function as inhibitors of endogenous proteinases. It is widely accepted

now that serine PIs have a defensive role herbivorous (especially, Lepidopteran) insect pests, which rely on serine proteinases for protein digestion. PIs act as substrate mimics and are hence able to bind stably with the proteinases; once ingested by the insects, these PIs bind to and inhibit the digestive serine proteinases in the insect (larval) gut, due to which protein digestion is blocked. Inhibition of digestive proteinases in the larval gut, causes a depletion in the pool of amino acids, otherwise available for assimilation and anabolic activities (Broadway and Duffey, 1986; Hilder *et al*, 1987; Broadway, 1996). This depletion leads to impaired intake of amino acids, which exerts a profoundly detrimental effect on larval physiology and retards insect growth and development (Harsulkar *et al*, 1999; Telang *et al*, 2003), and decreases the fertility and fecundity of the adult moths (DeLeo *et al*, 2001; Telang *et al*, 2003). **Table 1.2.3** lists effects of few representative plant PIs (leguminous as well as non-leguminous) on insect physiology. Further, the decreased fertility and fecundity of the adult moths causes a gradual decimation of the insect populations, and unlike the *Bt* toxin, do not cause direct killing.

(a) <u>Types and properties of legume seed PIs</u>

Legume seeds PIs are usually inhibitors of serine proteinases and exhibit strong inhibitory activity against trypsin, chymotrypsin or both. Inhibitory activity against proteinases of other mechanistic classes such as cysteine- or metallo-proteinases is rarely observed. Biochemical characterization of these inhibitor proteins resulted in identification of two subtypes, *viz.*, Bowman-Birk-type (BBI) and Kunitz-type; these subtypes have similar specificities, but they differ in their biochemical and physical properties. The Bowman-Birk type inhibitors were initially described in soybeans (*Glycine max*) as inhibitors of trypsin (Bowman, 1946; Birk, 1963) and subsequently reported with dual specificity towards trypsin and chymotrypsin in soybeans (Birk *et al*, 1961; Birk, 1985), as well as other legumes like chickpeas (*C. arietinum*) (Belew *et al*, 1975; Belew and Eaker, 1976; Smirnoff *et al*, 1976; Jibson *et al*, 1981; Birk, 1985) are single chain polypeptides of 6-12 kDa and possess seven intra-chain disulphide bridges. These PIs are unique in possessing two mutually exclusive domains, each with a single active site. Due to presence of two active sites, these PIs can bind simultaneously to two separate proteinase molecules. Commonly observed specificities in Bowman-Birk

Table 1.2.3 Some PIs and their effects on insect physiology

PI	Insect	Effect	Reference
Soybean Kunitz TI	*Helicoverpa armigera*	50-99% weight reduction	Johnston *et al*, 1993
	Heliothis zea	30% weight reduction	Broadway and Duffey, 1986
	Spodoptera exigua	15% weight reduction	
Soybean BBI	*Callosobruchus maculatus*	55-85% mortality	Gatehouse and Boulter, 1983
	H. armigera	10-50% weight reduction	Johnston *et al*, 1993
Bittergourd PI	*H. armigera*	43 % weight reduction, 67% decrease in fertility	Telang *et al*, 2003
	Spodoptera litura	70% weight reduction, 67% decrease in fertility	
Groundnut PI	*H. armigera*	75-80% weight reduction, delay in onset of pupation	Harsulkar *et al*, 1999
Mustard PI	*Spodoptera litura*	40-60% reduction in fertility	DeLeo and Gallerani, 2002
Oryzacystatin I	*Leptinotarsa decemlineata*	30% weight reduction	Michaud *et al*, 1995
Potato PinII	*Spodoptera exigua*	15% weight reduction	Broadway and Duffey, 1986
Winged Bean PI	*H. armigera*	Weight reduction	Giri *et al*, 2003
Capsicum PI		42-55% weight reduction, delayed pupation	Tamhane *et al*, 2005
Tomato PI		17-32% weight reduction	Damle *et al*, 2005

type inhibitors include trypsin/trypsin, chymotrypsin/chymotrypsin and trypsin/chymotrypsin, although some are known to posses a specificity for elastase on one of the domains. A related class of BBI is represented by the barley (*H. vulgare*) bi-functional α-amylase/trypsin inhibitor, which has inhibitory activity against two entirely different mechanistic classes of enzymes. Kunitz type inhibitors were also first purified and crystallized from seeds of soybean (*Glycine max*) and described as inhibitors of trypsin (Kunitz, 1945). Subsequent reports detailed purification of similar inhibitors from other legumes as well as having specificity against other serine proteinases like chymotrypsin. Kunitz-type PIs have only recently been reported in chickpeas (the research embodied in this thesis is the first report on identification and functional characterization of a

Kunitz-type PI from chickpeas). These PIs are single chain polypeptides of ~20 kDa, characterized by presence of two intra-chain disulphide bridges, and usually having a single activity site (inhibitory loop) although secondary activity has also been reported (Franco *et al*, 2002). Owing to the single active site, these inhibitors bind to proteinases in a simple 1:1 fashion. At the molecular level, Kunitz type PIs have a roughly spherical shape and the structure is characterized as a 'β-trefoil fold'. The amino acids responsible for serine proteinase binding and inhibition are located on an extended 'binding' loop, which is structurally similar across the Kunitz-type PIs. The specificity towards target proteinase is determined by the nature of the amino acid residue at the P1 position on the 'binding' loop – basic side-chain amino acids like arginine or lysine are usually associated with trypsin specificity whereas hydrophobic amino acids (phenylalanine>others) are linked to chymotrypsin specificity. Legume PIs follow the classical mechanism of proteinase inhibition, *i.e.*, the inhibitory loop of the PI mimics the substrate and binds to the protease, forming a stable complex, thus inactivating the protease. A detailed discussion on the structural and functional aspects of Kunitz-type PIs as well as the mechanism of their inhibitory action has been included in **Chapter IV**.

(b) <u>Genetics of expression of PIs</u>

Though plant PI are known to be encoded by multiple genes, studies indicate that these genes evolved by duplication events followed by rapid diversifications (Laskowski *et al*, 1998; Heibges *et al*, 2003a; Mello *et al*, 2003; Zupunski *et al*, 2003; Lopes *et al*, 2004; Wang *et al*, 2004). Differences in post-translational modifications are also responsible for the various iso-inhibitors that are commonly observed (Domoney *et al*, 1995; Patankar *et al*, 2001; Page *et al*, 2002; Deshimaru *et al*, 2004). In fact, presence of iso-inhibitors is somewhat akin to the multiple proteinase isoforms in insects. The diversification into PI genes into multi-gene families was attributed to the adaptive co-evolution of plants in response to attack herbivorous insects and other pathogens (Ehrlich and Raven, 1964). Plants respond to such attacks by activation of appropriate defense- or wound-responsive genes locally and systemically (Karban, 1989; Stotz *et al*, 1999; Chamberlain *et al*, 2000; Kessler and Baldwin, 2002; Arimura *et al*, 2005). Systemic responses to herbivory are mediated by various macromolecules of the well characterized known cell-to-cell 'octadecanoid' signaling pathway; this

pathway is activated in response to wound signaling resulting from tissue damage caused by insect feeding that culminates in jasmonate biosynthesis leading to production of PI among other molecules. In fact, herbivore-induced signaling has also been shown to attract other parasitoid insects, which prey on the herbivore (DeMoares *et al*, 1988); plant signaling responses can hence be considered as a double edged sword. Thus, a positive correlation exists between herbivore infestation, plant signaling and synthesis of PIs; this has been shown to be beneficial to plants by providing resistance against the herbivore (Agrawal, 1998; DeMoares *et al*, 1998).

(c) <u>Host and Non-host PIs in defense against insect pests</u>

Although many species of plants express PIs, not all are effective in defense against Lepidopteran pests. It has been observed that PIs of "host plants", *i.e.*, those prone to herbivore infestation, usually fail in offering protection (Giri *et al*, 1998; Harsulkar *et al*, 1999). On the other hand, the "non-host" plants, *i.e.*, which are not attacked by insects, are usually the better performers against insects (Harsulkar *et al* 1999). The corollary that follows from this statement is that adaptability to plant PIs determines the choice of host plants for the herbivorous insect pest. In case of chickpea, the seeds are abundant in Bowman-Birk type trypsin inhibitors (TIs). However, these TIs are ineffective against the insect pest *H. armigera*, whose digestive proteinases can easily degrade these TIs (Giri *et al*, 1998). Thus, *H. armigera* is able to successfully colonize chickpea plants and gain nutrients. It is clear that non-host plant PIs would be the agent of choice in programs to develop insect resistance in such plants. It is well accepted that the PIs, which exhibit strong inhibition of insect proteinases, *in vitro*, do not necessarily have similar effects *in vivo* (Edmonds *et al*, 1996). Similarly, PIs with moderate *in vivo* activity have been reported to exert significant antagonism to insects, *in vivo*. Usually a combination of PIs have worked as a better strategy (Johnson *et al*, 1989; Abdeen *et al*, 2005) because, they not only cause inhibition of a wider scope of proteinases, but also afford protection to each other, thus avoiding any loss in potency due to insect proteinase mediated degradation (Abdeen *et al*, 2005). Likewise, various combinations of PIs, lectins and the *Bt* toxin (Zhang *et al*, 2000) have shown promising results.

(d) Understanding insect resistance

The inability of host as well as non-host plant PIs to afford resistance to plants against insect pests have illustrated the adaptability of insects to these PIs (Bolter *et al*, 1995; Broadway, 1995; Bown *et al*, 1997; Jongsma and Bolter, 1997; Deleo *et al*, 1998; Lara *et al*, 2000; Paulillo *et al*, 2000; Brito *et al*, 2001; Volpicella *et al*, 2005). Qualitative and quantitative analyses of changes in insect gut proteinase activities and expression have proved invaluable in our understanding of insect responses and have helped identified the possible means by which an insect may adapt to dietary PIs. **Table 1.2.4** lists a few incidences of insects adapting to PIs.

Table 1.2.4 Adaptive responses in Lepidopteran insect pests towards PIs.

Insect	PI type	Adaptive response	Reference
Helicoverpa armigera	SKTI	Upregulation of Chymotrypsins	Bown *et al*, 1997
	Aprotinin	Upregulation of chymotrypsins and downregulation of trypsins	Gatehouse *et al*, 1997
	PinII		
	SKTI		
	PinI		
Heliothis virescens	Nicotiana leaf PIs	Synthesis of PI insensitive trypsins	Brito *et al*, 2001
Helicoverpa zea	SKTI	Upregulation of trypsins	Broadway and Duffey, 1986
	PinII		
	SKTI	Adapted trypsins are insensitive to SKTI, BBI, PinII and MTI-II	Volpicella *et al*, 2003
	SKTI	Increased gut proteolytic activity	Broadway, 1997
Spodoptera exigua	SKTI	Upregulation of trypsins	Broadway and Duffey, 1986
	PinII		
	PinII	Upregulation of trypsins	Jongsma *et al*, 1995
Agrotis ipsilon	SKTI	Increased gut proteolytic activity	Broadway, 1997
		Synthesis of PI insensitive trypsins, upregulation and downregulation of various chymotrypsins	Mazundar-Leighton and Broadway, 2001a
Trichoplusia ni	SKTI	Increased gut proteolytic activity	Broadway, 1997
		Synthesis of PI insensitive trypsins, upregulation and downregulation of various chymotrypsins	Mazundar-Leighton and Broadway, 2001a

To deal with these PIs, the insects could use any of the choices in a three-pronged strategy. The first choice available to insects is to increase the population of gut proteinases by a generalized over-expression, to compensate for the loss of activity due to inhibitor binding (Broadway and Duffey, 1986; Broadway, 1997; Gatehouse *et al*, 1997; Girard *et al*, 1998b). This simple game of numbers would be effective in dealing with a limited concentration of PIs. The second option for

the insects would be *de novo* synthesis of proteinases, which are insensitive to PIs (Jongsma *et al*, 1995; Broadway, 1996; Bown *et al*, 1997; Mazumdar-Leighton and Broadway, 2001a; 2001b; Volpicella *et al*, 2003). These proteinases usually have an altered substrate recognition site and hence possess no affinity for the PI in consideration. Finally, the insects have the option of over-expressing those proteinases, which could bind to and degrade the PI (Ishimoto and Chrispeels, 1996; Girard *et al*, 1998a; Giri *et al*, 1998; Moon *et al*, 2004; Telang *et al*, 2005). The insect also has the choice of using any or all of these approaches in different combinations to adapt to the PI (Zhu-Salzman *et al*, 2003). In fact, the process of adapting to PIs has often been compared to the same process that is responsible for polyphagy, as seen in many insects like *H. armigera*. These findings have laid more stress on understanding the total complement of digestive proteinases available to insect pest species, and their regulation in response to dietary components.

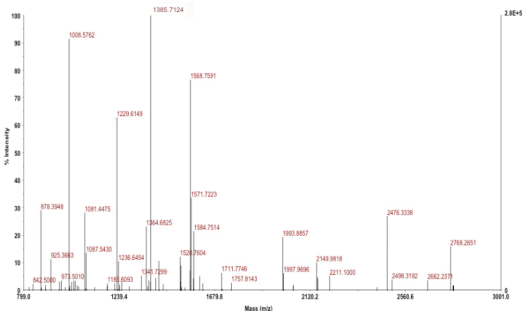
In this review, I have attempted to summarize our current understanding as well as to provoke scientific thought on the future of crop protection. Although our understanding of the processes underlying plant defenses as well as insect offenses has vastly improved over the last many years, a significant proportion is not yet known about the interactions at the molecular level, which decide the outcome of the battle between plants and pests. It is in the interest of humans to ensure protection of crops from insect pests, and this would be only possible through continuous efforts to delve deeper to decipher the molecular aspects of plant-pest association.

CHAPTER II:

Atypical Features of a Proteinase Inhibitor from Chickpeas

SECTION 1: Purification and identification of CaKPI

The research work described in this section is part of a full-length paper, which has been published in Plant Molecular Biology (Srinivasan et al, 2005a)



Introduction

Water-soluble seed proteins constitute about 15% of the total content of legume seeds, and proteinase inhibitor (PI) proteins generally account for about 5-10% of total soluble seed protein content (Sales *et al*, 2000). Since PIs form such a minor component of total seed material, they might not be visualized by conventional protein-staining methods; instead, they can be identified by virtue of their activity. The gel-X-ray film contact print technique (Pichare and Kachole, 1994) has proven invaluable in detection of PIs in legume seeds such as chickpea, pigeonpea and winged bean. In addition, *in vitro* assays have aided determination of specificity and activity of PIs. Obtaining the PI in a (semi) pure form is of paramount importance for determination of various physiochemical properties such as amino acid content and sequence, molecular weight, isoelectric constant etc. Purification of these PIs is routinely carried out by conventional means such as ammonium sulphate fractionation, dialysis and chromatography on various media, as necessitated by the physical properties of the protein. Among chromatographic techniques, gel-filtration (based on molecular weight) and ion-exchange (based on charge) are most commonly employed. Sequencing of N-terminal amino acids aids identification of the protein based on homology to other reported proteins. Mass spectrometry techniques such as matrix assisted LASER desorption ionization-time of flight (MALDI-TOF) also help to accurately identify the protein of interest. Through a combination of these various approaches in a sequential manner, it is possible to select and characterize single species of proteins in the vast population of those found in tissues such as seeds. This section describes the experimental approach for purification and identification of one such low expressing Kunitz type inhibitor from the mature seeds of chickpea.

***Note:** The acronyms/terms **HGPI**, *Helicoverpa armigera* Gut Proteinases Inhibitor, and **CaKPI**, *Cicer arietinum* Kunitz-type Proteinase Inhibitor, are to be considered as synonymous within the context of this manuscript.

Materials and methods

Materials

Seeds of chickpea (*C. arietinum* cv. Vijay) were obtained from the Pulses Research Station at Mahatma Phule Agricultural University, Rahuri, India. SephacrylS-200HR, DEAE-SepharoseCL6B, Superdex-200 and Phenyl-Sepharose were obtained from GE Healthcare (previously Amersham Biosciences). Protein molecular weight standard ("SDS-7") was procured from Sigma Chemicals, St. Louis, USA. AG-501X8 mixed bed resin was obtained from Bio-Rad, UK. X-ray films and developer were purchased from Kodak, Chennai, India. All other chemicals were procured locally and were of analytical grade.

Extraction of chickpea PIs

Soluble proteins were extracted from milled, defatted and depigmented seeds of chickpea in 10 volumes of deionized water, with constant stirring at 4°C for 4 h (Giri *et al*, 1998). Inhibitor proteins were enriched with Ammonium Sulphate fractionation (0-50% saturation), recovered by centrifugation at 10,000 rpm for 10 min at 4°C, dissolved in minimal amount of deionized water and desalted by dialysis against 3 changes of deionized water at 4°C for 30 min, each. Total protein was assayed by use of Bradford reagent.

Preparation of fresh Helicoverpa armigera (larval) gut proteinase extract, HGPs

500 mg tissue was weighed out from freshly dissected *H. armigera* larval midguts. Gut tissue was frozen in liquid nitrogen, ground to a fine powder and extracted in 5 mL of 200 mM Glycine-NaOH buffer (pH=10.0) for 2 h at 4°C. The extract was then centrifuged at 10,000 rpm (4°C, 10 min) and the supernatant was used as source of HGPs for *in gel* visualizations (Patankar *et al*, 2001).

Visualization of PIs

Approximately 5 µg of extracted seed protein was electrophoresed on 12% native polyacrylamide gel (pH=8.8) following which, the entire gel was immersed in trypsin

or HGPs for 10 to 30 min, rinsed in buffer (100 mM Tris-HCl, pH=8.0 for trypsin and 200 mM Glycine-NaOH buffer, pH=10.0 for HGPs) to remove excess proteinase and the gel finally overlaid on X-ray film whilst taking care not to introduce any air bubbles between the gel and the X-ray film. After incubation at 25°C for 5, 15 and 30 min respectively, the X-ray film was washed with cold or lukewarm water until gelatin digested by trypsin or HGPs had cleared. Inhibitors of trypsin (TI) or HGPs (HGPI) were visualized as bands of undigested gelatin against the hydrolyzed background on the surface of the X-ray film (Pichare and Kachole, 1994). For rapid detection of HGPs inhibitors, 2 µL of each sample was pre-incubated with 2 µL of HGPs preparation for 15 min at 25°C and the entire amount was spotted onto an X-ray film. After incubation for 10 min, the X-ray film was washed in lukewarm water and HGPs inhibitors detected by inhibition of gelatin hydrolysis in the region of the spot.

Purification of the HGPs inhibitor (CaKPI) from chickpea seeds

50 mg extracted seed protein (~1 mL) was loaded on SephacrylS-200HR gel filtration column and eluted under a constant flow (9 mLh^{-1}) of 25 mM Tris-HCl (pH 7.5). PI containing fractions were collected across six runs of gel filtration, pooled and loaded on DEAE-Sepharose ion exchange column. Elution was under linear (0 - 500 mM) NaCl gradient in 25 mM Tris-HCl (pH 7.5) at constant flow rate (60 mLh^{-1}). PI fractions from ion exchange column were pooled, dialyzed against three changes of deionized water and concentrated by lyophilization. This preparation was subjected to gel filtration on Superdex-200 FPLC column, with elution under 50 mM Tris-HCl, 150 mM NaCl (pH=7.5), at constant flow rate (15 mLh^{-1}), to obtain enriched CaKPI protein. At all stages of purification, the PI was monitored by electrophoretic separation on 12% native and/or 15% sodium dodecyl sulphate (SDS) polyacrylamide gels (Laemmli, 1970) and visualized by staining with silver nitrate or Coomassie Brilliant Blue R-250 (CBB R-250).

N-terminal sequencing of the purified HGPs inhibitor (CaKPI)

Partially purified PI was denatured with 2-mercaptoethanol and electrophoresed on 16% low-pH (7.5) Tricine SDS-PAGE. Prior to casting of gel, acrylamide solution

was de-ionised for 30 min using a commercial mixed-bed resin (*e.g.*, AG-501X8, Bio-Rad). 100 mM sodium thioglycollate was included in the upper tank buffer for scavenging oxidizing agents that could cause N-terminal blockage (Moos *et al*, 1998). The separated proteins were electro-blotted onto polyvinylene difluoride (PVDF) membrane under constant current. The blot was stained with CBBR-250 to identify the fragment of interest, which was sequenced at the amino-terminal by the Edman degradation scheme in a gas phase sequenator. This protocol was designed as per instructions and suggestions given on the University of Cambridge proteomics and protein sequencing facility web page (http://www.bioc.cam.ac.uk/pnac).

MALDI-TOF analysis of the purified HGPs inhibitor (CaKPI)

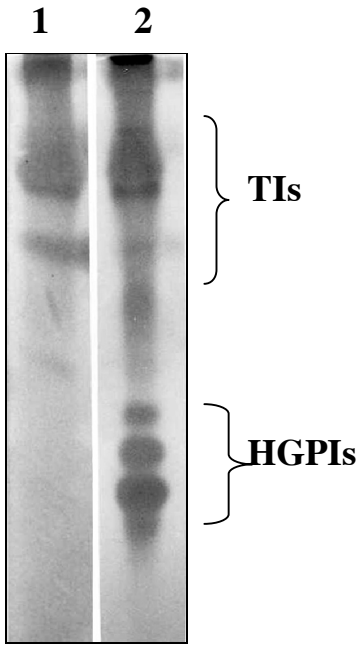
The purified PI was denatured with 2-mecrcaptoethanol, separated by 16% Tricine SDS PAGE, and visualized by staining with CBBR-250. Protein band was carefully excised with a sharp scalpel and completely destained with several washes of destaining solution (30% methanol, 10% acetic acid). The de-stained and denatured polypeptide fragment was digested by trypsin *in gel* and the resultant peptides were analyzed by MALDI-TOF. This protocol was designed as per instructions and suggestions given on the University of Cambridge proteomics and protein sequencing facility web page (http://www.bioc.cam.ac.uk/pnac). The spectrum obtained was interpreted online by the Mascot software (http://www.matrixscience.com) that retrieved matching polypeptide sequences from a global database. Amino acid sequences of the most likely identities were analyzed for probable internal trypsin cleavage sites using Peptide-Cutter (http://us.expasy.org/tools/peptidecutter/), so as to yield virtual tyrptic digests. The molecular masses of the 'virtual digests' of the matching polypeptide fragments were compared with the mass ions of putative amino acid sequences deduced by MALDI-TOF analysis of purified PI protein.

Results

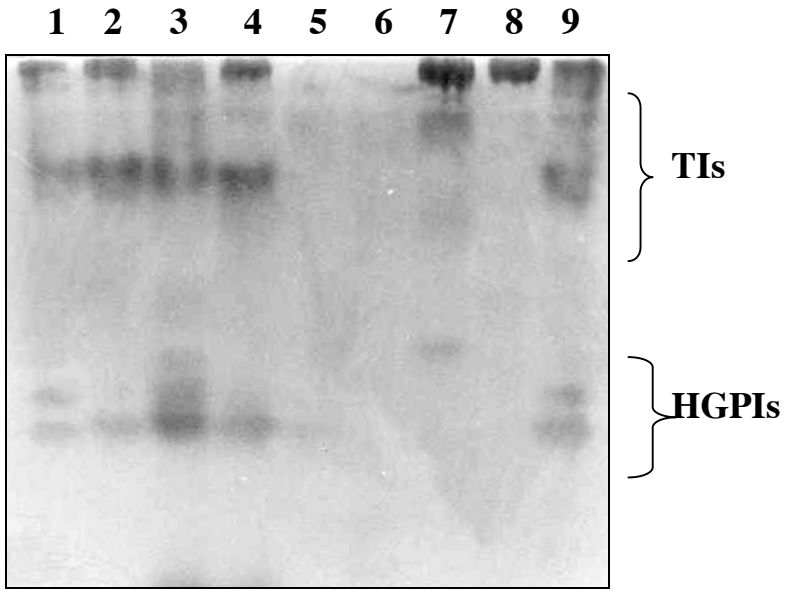
HGPs inhibitors are distinct from the Bowman Birk TIs reported in chickpeas

Chickpea seed extracts were separated by electrophoresis on 12% native

Fig. 2.1.1. TI and HGPI profiles in chickpeas
TIs and HGPI activity bands were detected by the Gel-X-ray film contact print technique.
(A) *Cicer arietinum*: Lane 1, visualization of TIs by treatment of gel with trypsin; Lane 2, visualization of HGPIs by treatment of gel with HGPs.
(B) Wild *Cicer* species. HGPI isoforms were detected in the following wild *Cicer* species by treatment of the gel with HGPs. Lane 1, *C. pinnatifidum*; Lane 2, *C. songaricum*; Lane 3, *C. bijugum*; Lane 4, *C. reticulatum*; Lane 5, *C. oxydon*; Lane 6, *C. microfilum*; Lane 7, *C. echinispermum*; Lane 8, *C. anatolicum*; Lane 9, *C. arietinum* (control).



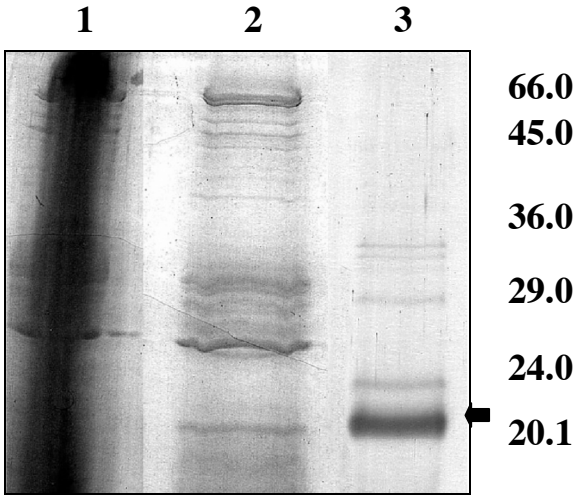
A



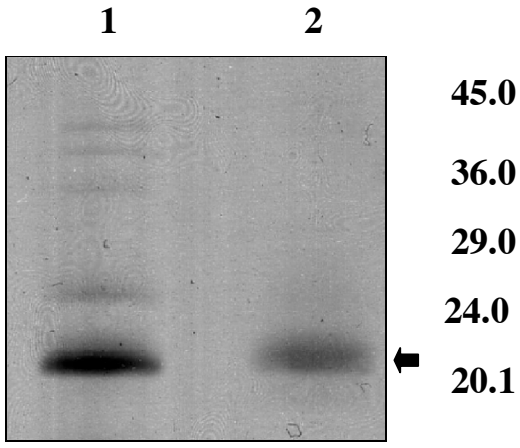
B

Fig. 2.1.1

Fig. 2.1.2. Purification of the HGPs inhibitor protein.
(A) Different stages of purification of the PI from chickpea seeds. Lane 1, crude seed extract; lane 2, enrichment followed by ion exchange on DEAE Sepharose; lane 3, final step of Gel-Filtration on Superdex200.
(B) 15% SDS-PAGE shows that the PI is made up of a single chain polypeptide. Lane 1, purified inhibitor treated with 2-mercaptoethanol; lane 2, untreated purified inhibitor.
Proteins were visualized by staining with coomassie brilliant blue R-250. (Band of interest is indicated by an arrow); molecular weight standards (kDa) are indicated to the right.



A



B

Fig. 2.1.2

polyacrylamide gels, following which inhibitors of trypsin and HGPs were visualized by the gel-X-ray film contact print technique (Pichare and Kachole, 1994; Harsulkar *et al*, 1999). Two distinct groups of PI activities were visualized by this technique. When seed-proteins were extracted from 100 mg meal in 1 mL water, the major activity-bands were of low mobility, with inhibitory activity towards trypsin (**Fig. 2.1.1A**, lane 1). These bands have previously been identified as the Bowman-Birk type (BBI) trypsin inhibitors (TIs). They had very little or no inhibitory effect on HGPs (Patankar *et al*, 1999). However, at a higher concentration of the seed-meal extract, (500 mg mL^{-1}), three distinct bands of inhibitory activity against HGPs, which showed faster migration, were observed (**Fig. 2.1.1A**, lane 2). These bands were perceived as having higher inhibitory activity against HGPs than trypsin and thus represented novel inhibitors in the seed extract. Since these PI specifically inhibited insect proteinases, it was initially designated as *H. armigera* gut proteinase inhibitor (HGPI). Different wild species of chickpea were also analyzed for isoforms of this inhibitor and it was observed that their profile varied significantly among the wild species (**Fig. 2.1.1B**).

Identification of the HGPs inhibitor

The novel PI proteins were purified to ~90% homogeneity, as judged by protein staining after SDS-PAGE (**Fig. 2.1.2A**). The three bands of HGPI activity observed on native PAGE co-purified under these conditions and visualized after native PAGE by staining the gel with silver nitrate or Coomassie Brilliant Blue R-250 (figure not shown). The purified PI protein had a molecular mass of approximately 20 kDa. Treatment of this inhibitor with 2-mercaptoethanol did not alter the migration on SDS-PAGE, indicating a single polypeptide chain (**Fig. 2.1.2B**). The N-terminal amino acid residues of PI polypeptide were determined by subjecting the purified protein to automated protein sequencing. The sequence determined resolved into two overlapping sequences, *viz.*, NEDVEQVLDINGNPIFPGGK and EDVEQVLDINGNPIFPGGKY, corresponding to predicted amino acids 24-43 and 25-44 of the reported chickpea α-fucosidase amino acid sequence (CAB76907, AJ276263; Dopico *et al*, unpublished data; Labrador, 2001, direct submission), but

Fig. 2.1.3 MALDI-TOF profile of the HGPs inhibitor.
The purified PI was trypsinized and spectrum of the resulting fragments was obtained by MALDI-TOF. (also refer **Table 2.1.1**)

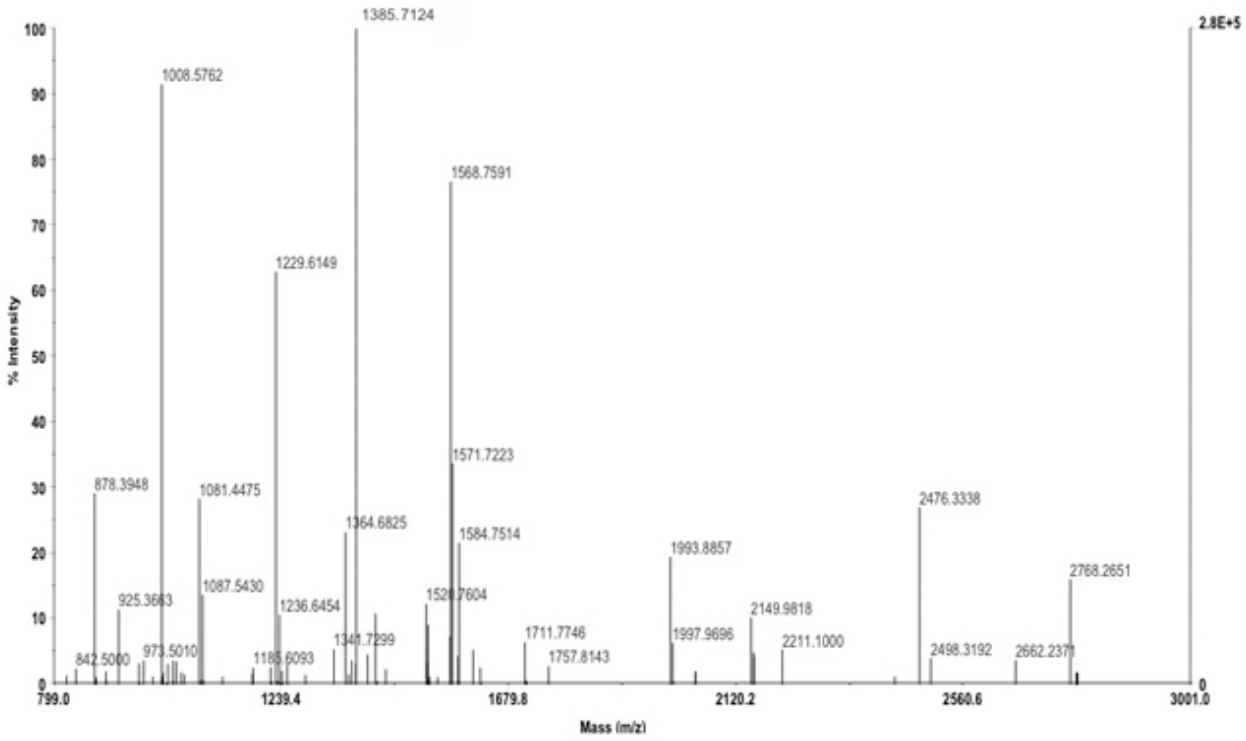

1385.7124
2.8E+5
1008.5762
1568.7591
1229.6149
1571.7223
878.3948
1081.4475
2476.3338
1364.6825
1584.7514
1993.8857
2768.2651
925.3663
1087.5430
1236.6454
2149.9818
1711.7746
1997.9696
2211.1000
842.5000
973.5010
1757.8143
2498.3192
2662.2378
% Intensity
Mass (m/z)
799.0
1239.4
1679.8
2120.2
2560.6
3001.0

Fig. 2.1.3

Table 2.1.1 Interpretation of MALDI-TOF data. Trypsin cleavage sites and resulting peptides in the HGPs inhibitor; theoretical masses predicted by "Peptide Cutter" as well as calculated by the "Mascot" software were compared with observed values in the MALDI-TOF spectra of the inhibitor purified from chickpea seeds. Carbamidomethyl substitutions in tryptic fragments are indicated by "(C)". Mass-ions observable in the spectrum and not assigned identities, but were close to "Peptide-Cutter" predicted values are marked by as asterisk "*". Note: Differences in calculated and observed masses arise due to limitations in calibration of the instrument and are hence recurring

Fragment residues	Resulting peptide	Predicted mass (Peptide cutter)	Calculated mass (Mascot)	Observed mass (MALDI-TOF)
1-20	NEDVEQVLDINGNPIFPGGK	2155.349	-	2155.06*
21-28	YYILPAIR	1008.228	1007.57	1008.58
29-36	GPPGGGVR	695.776	-	-
37-39	LDK	374.437	-	-
40-54	TGDSECPVTVLQDYK (C)	1654.809	1710.77	1711.77
55-63	EVINGLPVK	968.161	-	964.58*
64-86	FVIPGISPGIIFTGTPIEIEFTK	2476.938	2475.33	2476.33
87	K	146.189	-	-
88-95	PNCAESSK	834.899	-	-
87-95	KPNCAESSK (C)	981.088	1019.47	1020.48
96-106	WLIFVDDTIDK	1364.561	1363.68	1364.68
96-120	WLIFVDDTIDKACIGIGGPENYSGK(C)	2730.085	2767.26	2768.27
107-120	ACIGIGGPENYSGK (C)	1365.524	1421.65	1422.66
121-131	QTLSGTFNIQK	1236.390	1035.64	1236.65
132-139	YGSGFGYK	877.952	877.39	878.39
132-145	YGSGFGYKLGFCVK (C)	1543.801	1581.74	1582.75
140-145	LGFCVK	665.849	-	-
146-155	GSPICLDIGR (C)	1030.207	1086.54	1087.54
146-163	GSPICLDIGRYDNDEGGR (C)	1955.086	1992.88	1993.89
156-163	YDNDEGGR	924.879	924.36	925.37
156-164	YDNDEGGRR	1099.082	1080.44	1081.45
164	R	174.203	-	-
164-174	RLNLTEHEAFR	1403.561	1384.71	1385.71
165-174	LNLTEHEAFR	1229.358	1228.61	1229.61
175-189	VVFVDASSYEDGIVK	1627.812	-	1625.78*
190-192	(end of sequence) SVV	303.359	-	-

also similar to other Kunitz type PIs commonly observed in legumes. Analysis of the sequence (CAB76907) by "SignalP" (http://www.cbs.dtu.dk/services/SignalP/) (Bendtsen *et al*, 2004) identified a signal region, with equal probability of cleavage after residue 21 or 23, and thus residue 24 is a predicted N-terminal residue of the mature protein. The purified PI was excised from a 15% Tricine SDS polyacrylamide gel, subjected to in-gel trypsin digestion, and the mass ions of the resulting peptide fragments analyzed by MALDI-TOF. Interpretation of the MALDI-TOF spectrum was carried out using the "Mascot" software (Perkins *et al*, 1999), which compared the recorded spectrum of mass ions (**Fig. 2.1.3**) to existing entries in a global database

and, based on the degree of similarity, reported scores corresponding to the probability of identity. Based on such a search, the analysis indicated identity to the reported chickpea α-fucosidase, and once again, homology to reported legume Kunitz type serine PIs. *In silico* trypsin digestion of the predicted α-fucosidase polypeptide sequence by the "Peptide-Cutter" algorithm (http://us.expasy.org/tools/peptidecutter/), identified putative trypsin-cutting sites and generated fragments, which were comparable with the calculated and observed mass ions in the "Mascot" interpreted MALDI-TOF spectrum of the purified inhibitor protein (**Table 2.1.1**).

Discussion

It has been proved that the more abundant Bowman-Birk type PIs from chickpea are unstable towards, and thus, not active against, digestive proteinases of *H. armigera* (Giri *et al*, 1998). However, there exists a distinct class of inhibitors that do have inhibitory activity against HGPs, although they are present at low levels in both developing as well as dry chickpea seeds. As a consequence, these PIs are not easily detected in seed extracts by staining or activity gels, and therefore, would probably have not been observed earlier. The HGPs inhibitor protein was purified to homogeneity by a combination of gel filtration and ion exchange chromatographic techniques. The purity of the protein was determined to be ~90%, by SDS-PAGE and staining with Coomassie Brilliant Blue R250. The initial gel-filtration chromatography (SephacrylS200-HR) was instrumental in enriching the PI containing fractions as both TIs (~10kDa) and HGPIs (~20kDa) were observed to co-purify (data not shown). However, ion-exchange chromatography on DEAE-Sepharose (anion exchanger) followed by a second, and higher resolution step of gel-filtration chromatography on a Superdex200 FPLC column, yielded the inhibitor protein at sufficient purity for further studies. An unambiguous identification of the purified inhibitor was possible through MALDI-TOF and analysis of the resultant spectrum, which gave a series of mass ions, corresponding to tryptic peptides from the purified inhibitor protein. On the basis of matching mass ion fragments in the protein database(s), the Mascot software assigned high probability of identity to the reported chickpea α-fucosidase (CAB76907, Dopico *et al*, unpublished data; Labrador, 2001, direct submission). Further, the "Peptide-Cutter" software was employed to predict

putative sites of trypsin cleavage in the reported α-fucosidase sequence (CAB76907); this *in silico* trypsin cleavage generated peptide fragments, which were then compared to the mass ions observed in the MALDI-TOF spectrum. Essentially, while the MALDI-TOF analysis and the "Mascot" software pertain to identifying the peptide sequences from the mass-spectrum obtained for the purified inhibitor protein, the "Peptide-Cutter" search generated a "virtual spectrum" of peptides from a known polypeptide sequence. The "Mascot" and "Peptide-Cutter" algorithms co-related with each other's results within an acceptable margin of error arising due to limitations of sensitivity and calibration of the instrument (**Table 2.1.1**). Whereas the "Mascot" software was advantageous in that it could account for the carbamidomethyl substitutions in the tryptic fragments, the "Peptide-Cutter" could predict observable mass-ions, which were not initially assigned matches by the "Mascot" software. Contaminant masses arising from trypsin degradation and probably from other minor impurities were expected, and also observed (not shown), but these did not interfere with or hinder the prediction. Further, the results of the MALDI-TOF analysis were confirmed by N-terminal amino acid sequencing of the purified inhibitor protein, which similarly reported similarity to legume Kunitz PIs. However, available database information (http:/www.ncbi.nlm.nih.gov/) suggests that there is just one other reported sequence for an α-fucosidase, *viz*., (Augur *et al*, 1995) from *Pisum sativum*, that bears strong homology to the purified polypeptide, in the current study. None of the other reported α-fucosidases, from plants and animals, show any homology to the *P. sativum* α-fucosidase, or the currently described polypeptide. Moreover, while the majority of α-fucosidases reported in the database are polypeptides larger than 50kDa, the purified protein was a single chain polypeptide of ~20kDa. Finally, Tarrago *et al* (2003) have already proven unambiguously, on the basis of sequence homology as well as absence of α-fucosidase activity, that the polypeptide reported as the *P. sativum* α-fucosidase is in fact a Kunitz-type proteinase inhibitor. The authors (Tarrago *et al*, 2003) further supported their finding by demonstrating that of α-fucosidase-specific antibodies did not bind to the alleged '*P. sativum* α-fucosidase'. In view of this indirect, but compelling, evidence, the purified PI was hence named *Cicer arietinum* (putative) Kunitz-type PI (CaKPI*).

The preliminary data obtained by analysis of the MALDI-TOF spectrum as well as by sequencing of N-terminal amino acid residues were used to perform database searches

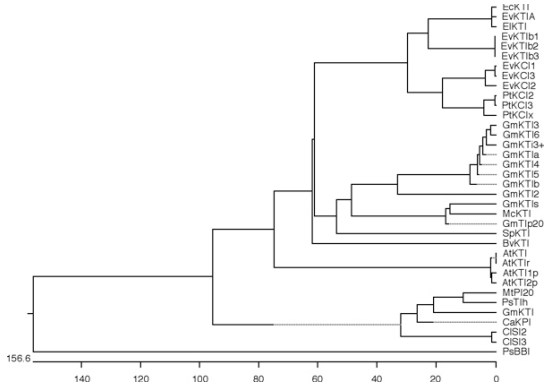
to identify homologous sequences in same or related species or groups. The investigations carried out to confirm the nature of this putative PI, as well as isolation cloning and nucleotide sequence analysis of the coding DNA are elaborated in the subsequent section.

Conclusion

The Kunitz type PI from chickpea seeds was purified to homogeneity by a combination of contemporary techniques. It was observed to be a single chain polypeptide consisting of about 200 amino acids (20kDa). Analysis of N-terminal amino acids revealed striking similarity to legume Kunitz type inhibitors, but was wrongly identified as α-fucosidase. Similar results were obtained by interpretation of MALDI-TOF spectrum. Hence the purified polypeptide was tentatively named as a putative Kunitz-type PI.

SECTION 2: Isolation and properties of *cakpi*

The research work described in this section is part of a full-length paper, which has been published in Plant Molecular Biology (Srinivasan et al, 2005a)



Introduction

As described in the previous section, purification of a stable and biochemically active protein from natural sources is the first step towards characterization of the various physico-chemical properties. Important clues to the identity of the protein can be obtained from sequencing of the N-terminal amino acids as well as by MALDI-TOF analysis; these aid to isolate the coding DNA by means of database searches, e.g., the Basic Local Alignment of Sequences Tool (BLAST) at the NCBI server (http://www.ncbi.nlm.nih.gov.in/BLAST). Several serine PIs have previously been studied in legume seeds and reported in these databases; most notable among these are soybean (*Glycine max*), winged bean (*Psophocarpus tetragonolobus*), pea (*Pisum sativum*), arabidopsis (*Arabidopsis thaliana*), bauhinia (*Bauhinia variegata*), canavalia (*Canavalia lineata*) and medicago (*Medicago truncatula*). Though comparison of protein sequences gives information on structural/functional variations, similar analysis of coding DNA sequences provides insights into the mutations occurring at the genetic level, from an evolutionary perspective. Vectors are also available that permit transcription and translation of the coding DNA ('expression') in heterologous systems such as bacteria and yeasts, under controlled conditions (bio-fermenter). Such 'expression vectors' allow us to obtain larger quantities of the (recombinant) protein, than the naturally occurring levels that are observed in seed tissue. Addition of a 'tag' of known amino acids (such as 6xHis) at the C-terminal of the recombinant protein, not only aid immunological detection but also function as a ligand for affinity based purification techniques. These modifications permit simpler, faster and accurate identification and purification of the recombinant protein. This section describes the isolation and characterization of the gene coding for CaKPI (*viz*., *cakpi*) from chickpea, as well as the expression of recombinant CaKPI in a yeast (*Pichia pastoris*) based bio-fermenter system.

Materials and Methods

Materials

Kanamycin, isopropyl-thio-galactoside (IPTG), 2-bromo-3-chloro-4-indolyl-galacto-pyranoside (X-Gal), plant DNA isolation kit and plasmid purification kit were procured from Sigma Chemicals, St. Louis, USA. DNA gel elution and purification kit was obtained from QIAgen, UK. pCR2.1 TOPO vector, pGAPZαB vector, Zeocin and *Pichia pastoris* X-133 were obtained from Invitrogen, UK. Anti C-term 6xHis as well as HRP conjugated Anti-Mouse (secondary) antibodies were procured locally. Enhanced Chemi-Luminiscence (ECL) blot development kit was obtained from GE Healthcare (previously, Amersham Biosciences). All other chemicals used were procured locally and were of analytical grade.

PCR amplification and cloning

Based on available sequence data, following gene specific primers; 5' CGC G**CT GCA G**GG AAC GAA GAT GTT GAA CAA G 3' (forward, with *Pst*I site) and 5' CGC G**GT CGA C**AA CAA CAG ATT TAA CAA TTC C 3' (reverse, with *Sal*I site), were synthesized and 10 µM each was used for PCR amplification on a chickpea genomic DNA template (25 ng), with a proofreading DNA polymerase. PCR cycling was as follows: initial denaturation at 94°C for 4 min, followed by 35 cycles of 94°C, 60 sec; 55°C, 60 sec; and 72°C, 60 sec. A final extension step at 72°C for 10 min was also included at the end of cycling. The products of the PCR were separated by electrophoresis on 1% agarose gel under 0.5X Tris-Acetate-ethylenediamine tetra-acetate (TAE) buffer (pH=8.0) at constant current. A standard DNA size marker was used to determine size of the amplified fragment, which was eluted and purified from agarose gel using a commercial kit. The purified DNA fragment was ligated into pCR2.1 TOPO vector by virtue of A/T overhangs and the construct used to transform competent *Escherichia coli* TOP10 cells. Transformants were selected under Kanamycin (20 µg mL^{-1}) by blue-white screening on Luria Bertonii (LB) medium. Plasmid was prepared from transformants by the alkaline lysis method (Sambrook *et al*, 1989) and the insert sequenced on an automated fluorescent sequencer by dideoxy chain termination method. The insert was then excised out of pCR2.1-construct by a double restriction digestion (at the *Pst*I and *Sal*I sites introduced during PCR amplification) and directionally cloned into another plasmid, *viz*., pGAPZαB,

between *Pst*I and *Xho*I sites. This step eliminated the C-terminal *myc* epitope coding region on the plasmid, but retained the subsequent 18-nucleotides coding for His_6, which was fused C-terminally to the expressed recombinant protein. The pGAPZαB-construct was used for transformation of competent *E. coli* TOP10 cells, followed by Zeocin (20 μg mL^{-1}) selection on Low-Salt Luria Bertonii (LSLB) medium. Plasmids were prepared from transformants and sequenced to confirm orientation and reading frame of the coding DNA. 10 μg of this construct was linearized with *Bln*I and used for transformation of the yeast *Pichia pastoris* X-133. Transformed yeast was selected on Yeast Extract Peptone Dextrose Sorbitol (YPDS) medium containing Zeocin (100 μg mL^{-1}) and were confirmed by colony PCR (Gussow and Clackson, 1989). Protocols followed for ligation, restriction digestion, transformation, etc. were as previously described by Sambrook *et al* (1989).

Identification of expressed recombinant protein by western blot hybridization

Yeast transformants were grown in individual 10 ml cultures for 36 h and the supernatants harvested by centrifugation (10,000 rpm, 10 min, 4°C). Soluble proteins in the supernatant were separated by 15% SDS-PAGE and electro-blotted onto nitrocellulose membrane at constant current. Standard protein size markers were identified by transient staining with Ponceau-S. After binding of mouse anti-6xHis (primary) antibodies to the C-terminal His_6 tag in the recombinant protein, horseradish peroxidase (HRP) conjugated goat anti-mouse (secondary) antibodies facilitated visualization of the recombinant protein by a hydrolysis of a commercial chemi-luminiscent substrate, as detected by exposure of the blot to X-ray film. Relative level(s) of recombinant protein expression were directly inferred from the intensity of the signal on the X-ray film. The yeast clone showing the highest expression was selected for over-expression of the recombinant protein in a laboratory-scale fermenter culture.

Large-scale heterologous expression of protein

Well-isolated colonies of *P. pastoris* transformants were used to inoculate 50 mL Yeast Extract Peptone Glycerol (YPG) medium containing Zeocin (25 μg mL^{-1}) and the culture was allowed to grow for 48 h at 30°C. These cultures were used to seed

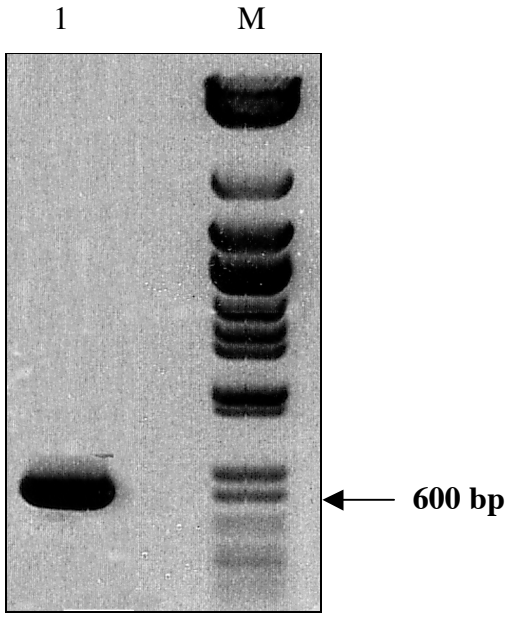
1500 mL of sterile Basal Salts Medium (pH adjusted after sterilization) in the fermenter vessel and allowed to grow overnight under constant agitation (250 rpm) at 30°C. On the second day, 50% glycerol feed was started at 10 mL h^{-1} and the pH was maintained by automated addition of 30% ammonia. Automatic foam control was maintained by addition of 10% silicone antifoam agent. Fermentation was terminated on fourth day and the culture centrifuged at 10,000 rpm for 50 min. Clear supernatant was harvested, to which, sodium chloride was added to a final concentration of 2 M. The entire volume was loaded onto a pre-equilibrated (2M sodium chloride) Phenyl-Sepharose column (Hydrophobic Interaction Chromatography) and the bound proteins were eluted under a reverse sodium chloride gradient (2.0 M – 0.0 M). Quantity and purity of the recombinant protein was analyzed in aliquots of the eluted fractions by 15% SDS-PAGE, followed by staining with CBB-R250 and/or western blot hybridization with Anti-C-term 6xHis, as described previously. Fermenter run was repeated to obtain sufficient quantity of expressed protein. The recombinant protein was freeze-dried and stored at 4°C for use in further assays.

Results

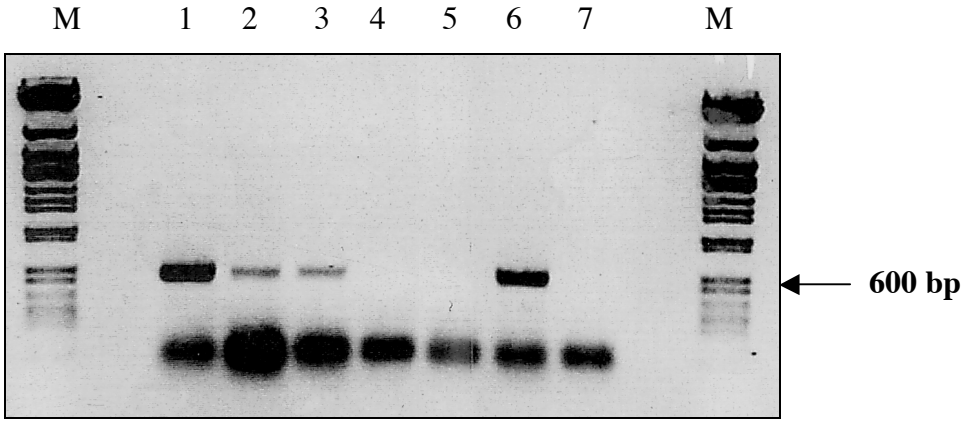
Cloning of the HGPs inhibitor gene and expression of the polypeptide product

Available database information, along with the experimentally deduced amino acid sequences (N-terminal sequencing and MALDI-TOF, ref. previous section) was used to design gene specific primers for PCR amplification, on chickpea genomic DNA as template, of the coding sequence of the mature protein, corresponding to residues 24-215 of the polypeptide CAB76907. A single product of ~600 base pairs was obtained (**Fig. 2.2.1A**), which was initially cloned into pCR2.1-TOPO vector and the construct used for bacterial (*Escherichis coli* TOP10) transformation. Plasmids were prepared from the positive transformants and the insert sequenced. The cloned DNA fragment was then directionally sub-cloned into binary (*E. coli-Pichia pastoris*) yeast expression vector pGAPZαB, which was used for transformation of *E. coli* for sequencing as well as large scale plasmid isolation. The pGAPZαB construct was

Fig. 2.2.1. Isolation and cloning of the coding DNA for CaKPI.
(A) PCR amplification using gene specific primers on a chickpea genomic DNA template yielded a fragment of ~600 bp. Lanes: 1, PCR product; M, DNA size marker. This fragment was subsequently inserted into a yeast expression vector, pGAPZαB and used to transform competent cells of *Pichia pastoris*.
(B) Putative *P. pastoris* genomic clones transformed with the pGAPZαB-construct were screened by colony PCR to confirm transformants.
Lanes 1-3 and 6, positives; lanes 4-5 and 7, negatives; M, λ-*Hin*DIII/*Eco*RI DNA size marker.



A



B

Fig. 2.2.1

Fig. 2.2.2. Recombinant protein expression by *Pichia pastoris*.
Supernatant from a 10 mL culture of *P. pastoris*, carrying the gene construct, was harvested and aliquots electrophoresed on a 15% SDS-polyacrylamide gel. Proteins were electroblotted onto nitrocellulose membrane. Blot was hybridized with Anti C-term 6xHis and then with horseradish peroxidase (HRP) conjugated secondary antibody to detect recombinant proteins.
Lane 1, Proteins in culture supernatant stained with Coomassie Brilliant Blue R250; lane 2, Positive expression; lane 3, negative control.
(Band of interest is indicated)

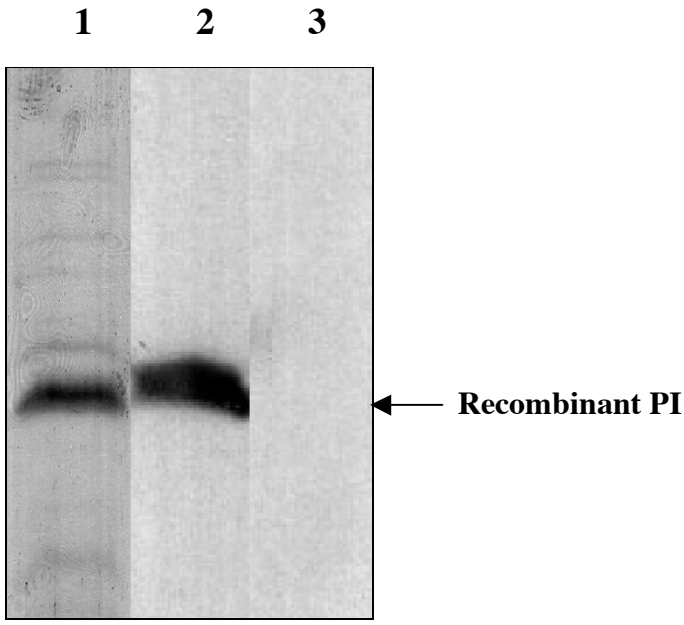


Fig. 2.2.2

finally linearized with *Bln*I and used for yeast (*P. pastoris*) transformation. Positive-transformant yeast clones carrying the cloned fragment were confirmed by colony PCR (**Fig. 2.2.1B**). Recombinant protein was expressed in *P. pastoris* for large-scale production in a bio-fermenter and the expression was detected by western blotting (**Fig. 2.2.2**). The bio-fermenter yielded ~25 mg inhibitor protein per fed-batch run, which were then pooled, quantified and used for all subsequent *in vitro* and *in vivo* assays (following section).

Analysis of the putative gene sequence

Sequencing of the PCR product revealed a 579 base pair open reading frame (ORF) terminated by a STOP codon (UAA), which coded for a putative polypeptide of 192 amino acids. When the deduced polypeptide sequence was queried in 'ScanProsite' (http://us.expasy.org/prosite/) (Gattiker *et al*, 2002), the program identified a signature pattern found in the Kunitz-type PI family (**Table 2.2.1**).

Table 2.2.1. Identification of Kunitz Inhibitor Family Signature Sequence in deduced HGPI protein sequence by Scan-Prosite.

Kunitz type PI Signature	[LIVM]-x-D-x-[EDNTY]-[DG]-[RKHDENQ]-x-[LIVM]-$(x)_5$-Y-x-[LIVM]
Observed Pattern	V-L-D-I-N-G-N-P-I-F-P-G-G-K-Y-Y-I
Location in *cakpi* sequence	NEDVEQVLDINGNPIFPGGKYYILPAIRGPPGGGVRLDKTGDSECPV TVLQDYKEVINGLPVKFVIPGISPGIIFTGTPIEIEFTKKPNCAESSKWL IFVDDTIDKACIGIGGPENYSGKQTLSGTFNIQKYGSGFGYKLGFCVK GSPICLDIGRYDNDEGGRRLNLTEHEAFRVVFVDASSYEDGIVKSVV

A homology search using the Basic Local Alignment of Sequences Tool for polypeptides (BLASTp) (http://www.ncbi.nlm.nih.gov/BLAST/) (Altschul *et al*, 1990) revealed that the predicted polypeptide is most similar to reported Kunitz type PIs in legumes such as *Medicago truncatula* (AAM88404), *Glycine max* (AAF87095), *Pisum sativum* (S49578), and *Canavalia lineata* (P81726, JX0311), which appeared to belong to a distinct clade, as revealed by Clustal analysis (**Fig. 2.2.3A**). This PI exhibited a critical sequence variation at the putative active site region, wherein the conventional arginine (Arg, R) or lysine (Lys, K) has been replaced by a glycine-isoleucine-serine (Gly-Ile-ser, G-I-S) motif. Similar variation

Fig. 2.2.3. Sequence analysis of CaKPI.
(A) Phylogenetic analysis of CaKPI polypeptide sequence and other Kunitz trypsin inhibitors (TIs) was carried out using the Clustal algorithm. ("CaKPI", indicated by the arrow)
Legend: AtKTI, *Arabidopsis thaliana* Kunitz Trypsin Inhibitor (AY054566, BT000366, NM105985, G96758); BvKTI, *Bauhinia variegata* Kunitz Trypsin Inhibitor (P83595); CaKPI, *Cicer arietinum* Kunitz Proteinase Inhibitor (CAB76907); ClSI, *Canavalia lineata* Subtilisin Inhibitor (P81726, JX0311); EcKTI, *Erythrina caffra* Kunitz Trypsin Inhibitor (P09943); ElKTI, *E. latissima* Kunitz Trypsin Inhibitor (P07475); EvKCI, *E. variegata* Kunitz Chymotrypsin Inhibitor (AAB25433, JC4990, P34952); EvKTI, E. variegata Kunitz Trypsin Inhibitor (JC5562, JH0781, P81365, P81366); GMKTI, *Glycine max* Kunitz Trypsin Inhibitor (AAF87095, S45035, AF233296, S45092, X64447, AB070269, AF314823, AF314824, X64448, X80039, AB029441); McKTI, *Matricaria chamomilla* Kunitz Trypsin Inhibitor (AB076806); MtKPI, *Medicago truncatula* Kunitz Proteinase Inhibitor (AAM88404); PsBBI, *Pisum sativum* Bowman Birk Inhibitor (2123385); PsTI, *P. sativum* Trypsin Inhibitor (S49578); PtKCI, *Psophocarpus tetragonolobus* Kunitz Chymotrypsin Inhibitor (S96733, S96732, S96735); SpKTI, *Swartzia pickelii* Kunitz Trypsin Inhibitor (doSoccoro *et al*., 2002).
(B) Alignment of various Kunitz Proteinase Inhibitors. Shaded regions indicate sequence similarity. (Putative active site region is indicated by "▼").
Legend: BvKTI, *Bauhinia variegata* Kunitz Trypsin Inhibitor (P83595); CaKPI, *Cicer arietinum* Kunitz Proteinase Inhibitor (CAB76907); ClSI, *Canavalia lineata* Subtilisin Inhibitor (P81726); EvKTI, E. variegata Kunitz Trypsin Inhibitor (P34952); GMKTI, *Glycine max* Kunitz Trypsin Inhibitor (AAF87095, S45035, AF233296); MtPI, *Medicago truncatula* Kunitz Proteinase Inhibitor (AAM88404); PsBBI, *Pisum sativum* Bowman Birk Inhibitor (2123385); PsTI, *P. sativum* Trypsin Inhibitor (S49578); SpKTI, *Swartzia pickelii* Kunitz Trypsin Inhibitor (doSoccoro *et al*., 2002).

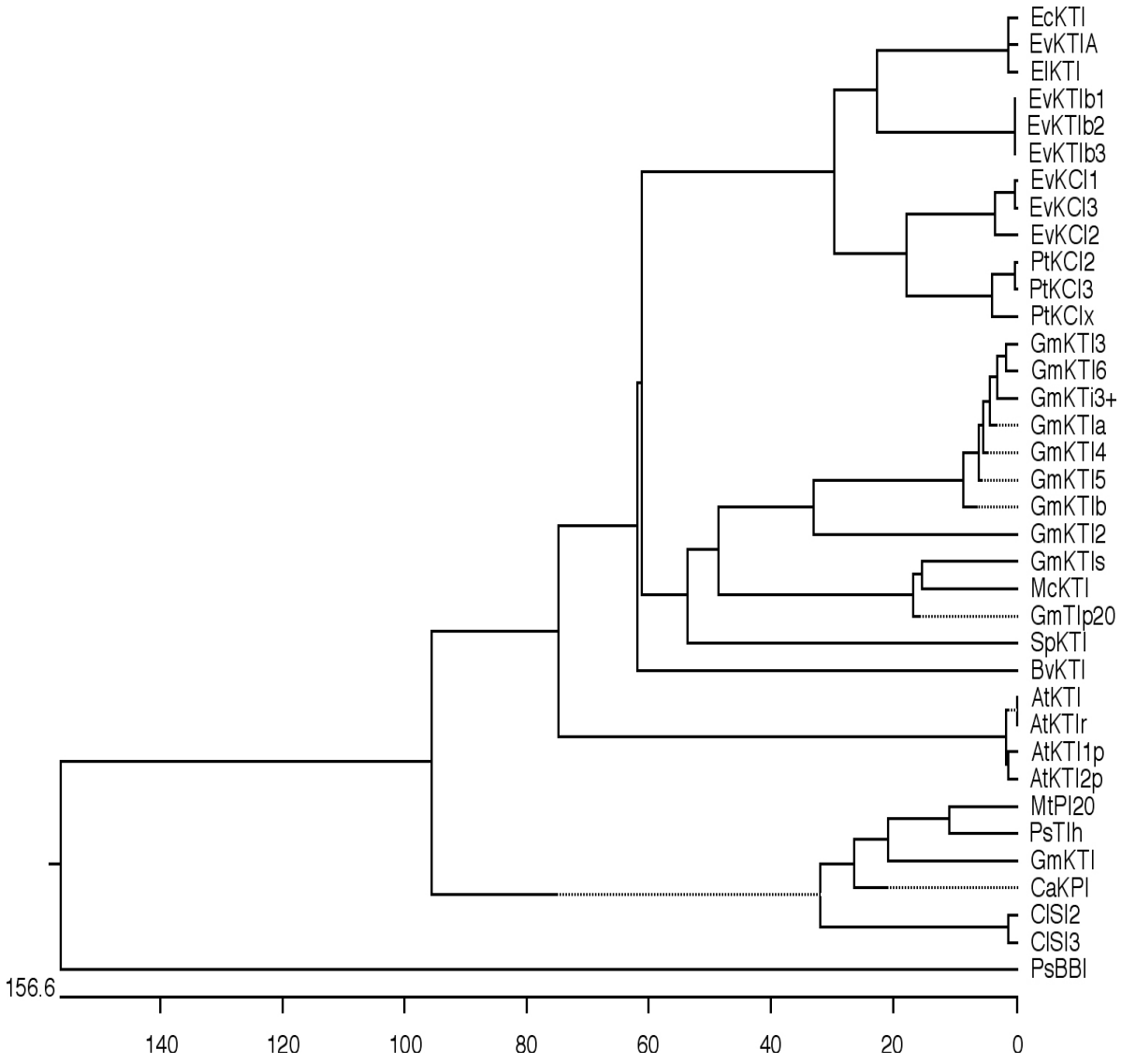

EcKTI
EvKTIA
ElKTI
EvKTIb1
EvKTIb2
EvKTIb3
EvKCI1
EvKCI3
EvKCI2
PtKCI2
PtKCI3
PtKCIx
GmKTI3
GmKTI6
GmKTi3+
GmKTIa
GmKTI4
GmKTI5
GmKTIb
GmKTI2
GmKTIs
McKTI
GmTIp20
SpKTI
BvKTI
AtKTI
AtKTIr
AtKTI1p
AtKTI2p
MtPI20
PsTIh
GmKTI
CaKPI
ClSI2
ClSI3
PsBBI
156.6
140
120
100
80
60
40
20
0

Fig. 2.2.3A

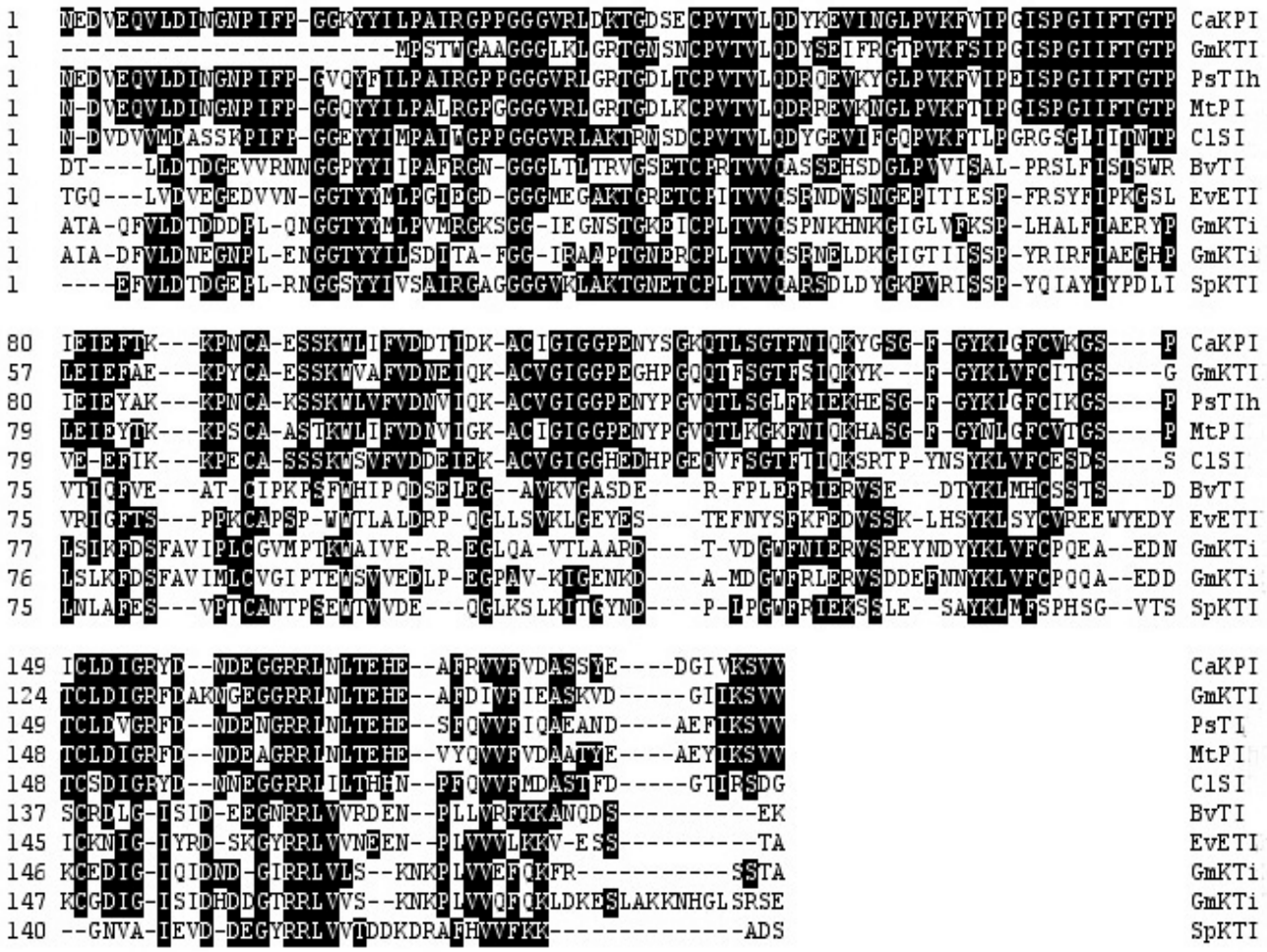


Fig. 2.2.3B

was also observed in the Kunitz-type PIs from *G. max*, *M. truncatula, C. lineata* and *P. sativum* (Fig. 2.2.3B) that belonged to the same clade, suggesting an evolutionary link to this sequence variation. Finally, as described in the previous section, since the amino acid as well as DNA sequences showed no similarity to any of the reported α-fucosidases, it was concluded that the reported sequence (CAB76907) is most likely to be a Kunitz-type PI, belonging to a distinct sub-group, the members of which share the putative active site variation. The isolated coding DNA (gene) was hence annotated as *cakpi,* for *Cicer arietinum* Kunitz proteinase inhibitor (CaKPI), which has been assigned accession number AY635930 in the NCBI database.

Discussion

The putative polypeptide encoded by PCR amplified DNA fragment is similar to other legume Kunitz inhibitors in being composed of single subunit polypeptide chain of ~200 amino acids, of which the first ~20 amino acids form the leader or signal peptide. Kunitz type type PIs have four cysteine residues that contribute to the two intra-chain disulphide bridges; four cysteines was also observed in the deduced amino acid sequence (CaKPI). The Kunitz type inhibitors usually have a single active domain region responsible for binding and inhibiting the target protease. However, some inhibitors are also known to have a secondary active site that functions independently (Dattagupta *et al*, 1999). Overlapping of both active sites is observed in case of the *Prosopis juliflora* Kunitz trypsin-chymotrypsin inhibitor leading to loss of independent activities (Franco *et al*, 2002). Other variations are known to exist as in case of the Kunitz TI from *Swartzia pickellii* that has only one intra chain disulphide bridge (doSocorro *et al*, 2002) and *Delonix regia* that has an amino acid insertion between the P1 and P2 residues (Krauchenko *et al*, 2003). Alignment of CaKPI amino acid sequence with legume Kunitz inhibitor sequences revealed a distinct alteration at the putative active site region, wherein the conventional arginine or lysine has been replaced by a glycine-isoleucine-serine (G-I-S) motif. This altered sequence is conserved in CaKPI coding region and also found conserved in the closely related Kunitz-type PIs from *P. sativum*, *G. max*, *C. lineata* and *M. truncatula*. Phylogenetic analysis of mature polypeptide sequences from representative Kunitz type PIs from various species was carried out to probe relationships between inhibitors and to

identify any evolutionary traits. It was observed that the CaKPI formed a distinct cluster with the PIs from *P. sativum*, *G. max*, *C. lineata* and *M. truncatula*, which also possess the G/E-I-S modification at the active site region. PIs are known to be encoded by multiple genes belonging to the same family (Page *et al*, 2002; Heibges *et al*, 2003a; Deshimaru *et al*, 2005). The databases suggest that the chickpea genome contains at least one other gene similar to *cakpi* (CAB76906), and although its product was not detected in this study, it may encode one of the inhibitor isoforms observed on activity staining of seed extracts (**Fig. 2.1.1**). A further complication in correlating genes and proteins for PIs is that the proteins undergo different post-translational modifications, such as proteolytic cleavage(s) at varying sites resulting in different isoinhibitor forms (Domoney *et al*, 1995; Heibges *et al*, 2003b). Evidence for this process was obtained in the present study, where purified CaKPI has a mixed N-terminal sequence caused by "trimming" of a residue from the initial product after signal peptide removal (**Section 2.1**). These N-terminal sequence differences are not expected to directly or significantly alter function, though their contribution towards stability/activity of CaKPI cannot be presently gauged. Differences in amino acid residues at the active site can be expected to have a direct effect on the specificity of the inhibitor towards a target proteinase, *e.g.*, replacement of an arginine by phenylalanine at the P1 position could alter specificity from trypsin to chymotrypsin. Studies on Kunitz inhibitors have revealed that the region in the vicinity of the active site (scissile peptide bond) shows very little variation whereas amino acids in other folds are more prone to accumulate mutations rendering them hyper-variable (Laskowski *et al*, 1998). This suggests that Kunitz inhibitors evolved by duplication of the putative ancestral gene following which the active site remained largely unchanged, but mutations in other regions caused rapid diversifications that gave rise to various iso-inhibitors (Mukhopadhyay, 2000; Zupunski *et al*, 2003). This diversification into a multi-gene family can be attributed to the adaptive co-evolution of plants in response to insects and other predators (Lopes *et al*, 2004), a concept which is supported by recent studies on functional comparison of Kunitz proteinase inhibitors as well as structural diversity and organization of their genes in potato (*Solanum tuberosum* L.) tubers (Heibges *et al*, 2003a; 2003b). Similar studies on possible evolutionary aspects of *cakpi*, with respect to diversification into multiple copies form the basis of the subsequent section. Further, the differential expression of

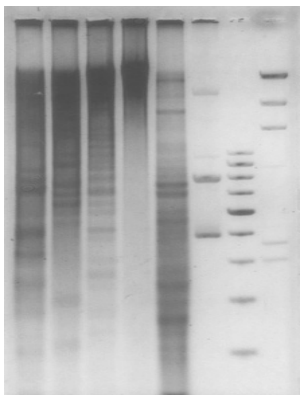
cakpi, as a possible marker of other endogenous functions, in addition to a putative defensive role has also been discussed.

Conclusion

The DNA coding for mature CaKPI protein was isolated by PCR amplification using sequence specific primers, based on available GenBank information. Sequence analysis of the obtained DNA fragment (coding region) as well as the putative gene product (translated polypeptide) confirmed similarity to legume Kunitz type PIs. The translated CaKPI polypeptide was observed to have a unique sequence variation that is shared with few Kunitz-type PIs from legumes - these PIs appeared to form a distinct group, which could imply an evolutionary aspect. The isolated DNA was cloned into an expression vector and mature CaKPI polypeptide was expressed *in vitro*, in a heterologous (yeast, *Pichia pastoris*) system, as a recombinant protein.

SECTION 3: Genetic diversification and differential expression of *cakpi*

The research work described in this section will be communicated as a full-length paper.



Introduction

Legume PIs, which seem to be exclusively expressed in seed tissue, have been implicated in defenses against herbivorous insect pests. Similar roles have been described for other plant PIs that are constitutively expressed in flowers (Damle *et al*, 2005), leaves (Tamhane *et al*, 2005) and fruits/seeds (Telang *et al*, 2003). Although many such reports have associated plant PIs with defensive role against herbivorous insect pests, not all PIs afford resistance against these herbivorous insect pests. The reasons for such a phenomenon may be innate, *viz*., lack of specificity towards the insect digestive proteinases or low levels of expression. The level of expression of PIs is not only governed by the strength of cis-acting regulatory elements (*e.g.,* promoter), but also by the actual number of copies of the PI gene. The commonly observed multiplicity of genes coding for legume PI isoforms probably arose due to duplication and rapid diversifications from a single ancestral gene (Mukhopadhyay, 2000; Zupunski *et al*, 2003). This process could also be linked to the adaptive co-evolution of plants and herbivorous insect pests (Lopes *et al*, 2004). Not surprisingly, plants express these PIs as numerous active isoforms, which are encountered during routine processes of purification and characterization. Further post-translational modifications of the primary gene-product(s) also add to the frequency of iso-inhibitors (Domoney *et al*, 1995; Heibges *et al*, 2003b). Finally, the phenotypic consequences of PI expression are also governed by the primary-structural features (amino acid sequence) of the polypeptide. The current section probes these very features of CaKPI and strives to decipher the genetic factors that govern the innately low levels of expression of the PI. The essential methodologies involved determination of the genomic copy number of *cakpi* as well and the profile of expression in various plant parts.

Materials and Methods

Materials

Dry, mature seeds of chickpea (*C. arietinum* cv. Vijay) were obtained from the Pulses Research Station at Mahatma Phule Agricultural University (MPKV), Rahuri, India. Perfecthyb-Plus and custom designed gene specific primers were procured from Sigma Chemicals, USA. Hybond N+ membrane, autoradiographic cassettes and intensifying screens were procured from Amersham Biosciences, USA. Cloning

vector pGEM-T Easy, restriction endonucleases and thermo-stable (*Taq*) DNA polymerase were obtained from Promega Inc., USA.. α^{32}P-dATP was procured from the Board of Radiation and Isotope Technology (BRIT), Hyderabad, India. X-ray film was procured from Kodak, Chennai, India. All other chemicals used were procured locally, where available and were of analytical grade.

Cultivation of chickpea plants and tissue collection

Chickpea plants were grown in pots under controlled lighting and temperature in a walk-in plant growth chamber. Plants were maintained at 25°C during day and at 16°C during night. A cycle of 16 h light period followed by 8 h dark period was automatically maintained. Only healthy plants were chosen for all studies. For DNA isolation, two weeks old plants were used – entire plants were carefully uprooted from the pots, the soil was washed from the roots using sterile deionized water, the plants transferred into 50 mL screw capped vials and immediately frozen under liquid nitrogen.

Similar precautions were taken for tissues selected for RNA isolation. Roots, leaves and flowers were collected from undamaged, healthy plants and frozen immediately under liquid nitrogen. Chickpea flowers were tagged on the day of flowering and healthy developing seeds were harvested after ~12, ~25 and ~40 days after flowering (DAF), corresponding to early- (E), mid- (M) and late- (Lt) mature stages. The harvested seed tissues were also immediately frozen under liquid nitrogen.

For longer storage, the tissue was transferred to and stored in deep freezers (–80°C) until required.

Isolation of genomic DNA

Genomic DNA was isolated from leaf tissue of chickpea grown under controlled conditions using the method previously described by Doyle and Doyle (1987). The isolated DNA was also treated with RNase-A (ab U, 2h, 37°C) to degrade contaminating RNA. The DNA was quantified spectrophotometrically (λ_{260}) and it's quality assessed by 0.8% agarose gel electrophoresis under native conditions.

Restriction enzyme digestion of DNA

Approximately 15 μg of genomic DNA was used for each of the restriction digests with 50U of *Eco*RI, *Eco*RV, *Bam*HI, *Dra*I and *Sal*I. Digestion was carried out at 37°C, with intermittent mixing, for a minimum of 48h or until completely digested, as determined by electrophoresis of an aliquot on 0.8% agarose gel and visualized by staining with ethidium bromide (Sambrook *et al*, 1989). The entire reaction mix was extracted with phenol:chloroform:iso-amyl alcohol (25:24:1) and chloroform:iso-amyl alcohol (24:1). The digested DNA was then precipitated under 0.1 vol. 5M sodium chloride, chilled 100% ethanol and the precipitated pellet was washed with chilled 70% ethanol. The pellet was air-dried (37°C) and dissolved in 20 μl sterile de-ionized water.

Southern Blotting

The genomic DNA digest obtained from the previous step was separated by slow electrophoresis on 0.8% agarose gel in tris-borate-EDTA buffer (pH=8.0) under constant current (30 mA). The entire gel was stained with ethidium bromide to visualize the digested DNA and size-markers under ultraviolet illumination. Following destaining, the gel was treated with 0.25M hydrochloric acid to depurinate the high molecular weight DNA fragments, this was followed by denaturation with 50mM sodium hydroxide, 1.5M sodium chloride. Finally the gel was equilibrated in 500mM Tris-HCl, 1.5M sodium chloride (pH=7.4) prior to electro-blotting of the separated DNA fragments onto a charged nitrocellulose membrane (Hybond N+). Semi-dry electro-blotting was carried out under 1.5M sodium chloride, 0.15 M sodium citrate (10X SSC), at constant current (1 mA cm^{-2}) for 2h. Following transfer, the DNA fragments were cross-linked to the membrane under ultraviolet irradiation (70,000 μJ cm^{-2}). The membrane was then baked at 80°C in a vacuum-oven for 2h and stored in a re-sealable polythene bag at 4°C until required (Sambrook *et al*, 1989).

Preparation of total RNA

Plant tissue frozen under liquid nitrogen was ground to a fine powder and total RNA prepared from all tissues using the method described by Wang *et al* (2005). Quantity of RNA was estimated by $\lambda_{260/280}$ and the quality was assessed by 1% agarose gel electrophoresis under native as well as denaturing conditions.

Northern Blotting

RNA samples were diluted with RNase free water and ~20μg aliquots were denatured with 70% deionized formamide, 8% formaldehyde, 1.5x MOPS and heating at 65°C for 5 min. The denatured sample was immediately chilled on ice and electrophoresed in 1% formaldehyde agarose gel under denaturing conditions. After the run, the entire gel was soaked sequentially in 200 mL RNase-free water for 15 min, 50 mM NaOH for 15 min and finally neutralized in10x SSPE for 30 min. Hybond N+ membrane was cut to the size of the gel and pre-equilibrated in 1x MOPS for 15 min. RNA from the gel was transferred to the membrane under constant current (1 mA cm^{-2}) in an electro-blotting apparatus for 2h. After the transfer, the position of the wells and the rRNA subunits were marked on the blot using a pencil. The blot was rinsed in 4x SSPE and the transferred RNA was cross-linked to the membrane under ultraviolet irradiation (70,000 μJ cm^{-2}). After cross-linking, the membrane was baked at 80°C in a vacuum-oven for 2h and stored in re-sealable polythene bags at 4°C until required (Sambrook *et al*, 1989).

Preparation of probe and hybridizations

The probe was prepared from the previously described (**Section 1**) *Escherichia coli* TOP10 transformant carrying the pCR2.1-*cakpi* construct. The construct was amplified by a polymerase chain reaction using previously described gene-specific oligonucleotide primers. α^{32}P-dATP was incorporated in the PCR mix so as to yield a radio-labelled double stranded DNA fragment for use as a probe. Hybridization of all Southern and Northern Blots were carried out with the generated probe using a commercial solution (PerfectHyb, Sigma-Aldrich) as per the manufacturers' instructions; initial prehybridization, 68°C, 4h; probe hybridization, 68°C, 4h; low stringency wash (2X SSC, 0.1% SDS), 30°C 10 min; medium stringency wash (0.5X SSC, 0.1% SDS), 68°C, 30 min; high stringency wash (0.1X SSC, 0.1% SDS), 68°C, 15 min. Finally the probed membranes were exposed to X-ray film for 120h for detection of probe-binding.

Results

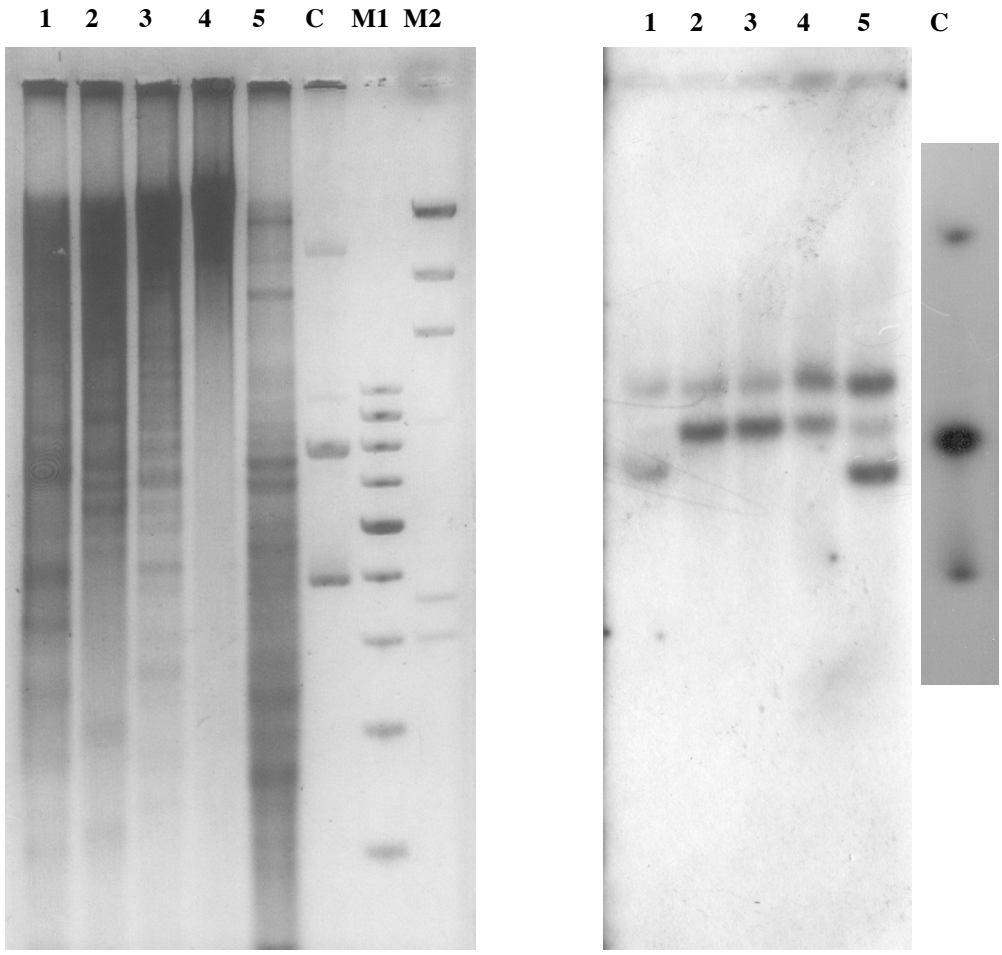
cakpi exhibits low degree of diversification in the chickpea genome

Fifteen micrograms of chickpea genomic DNA was digested, each with *Dra* I, *Bam*HI, *Eco*RI, *Eco*RV and *Sal*I, and the digestion products separated on agarose gel (**Fig. 2.3.1A**) and (Southern) blotted onto a nitrocellulose membrane. This was followed by hybridization of the membrane with a radiolabeled probe generated by PCR amplification of the *cakpi* coding region (~600 bp) using the previously described pCR2.1/*cakpi* construct. Exposure of the hybridized blot to X-ray film and analysis of the radiolabeled-probe binding pattern revealed a profile of low gene copy number (**Fig. 2.3.1B**). In addition to a strongly hybridizing fragment, another weak signal was also observed in genomic digests with *Dra* I, *Bam*HI, *Eco*RI and *Eco*RV (**Fig. 2.3.1B**, lanes 1-4). As expected, *Sal*I, which cuts approximately midway within the *cakpi* sequence yielded two fragments in the genomic digest (**Fig. 2.3.1B**, lane 5), which bound to the probe with almost equal affinity. These fragments strongly bound the probe even after high stringency washes with 0.5X SSC, 0.1% SDS, revealing a high degree of sequence homology between the probe and the target sequence(s).

cakpi is differentially expressed in developing seed tissue

Total RNA was isolated from various plant tissues including flowers, leaves and roots as well as seeds during various stages of development. For simplification, three arbitrarily defined stages were considered for developing seeds, *viz*., ~12, ~25 and ~40 days after flowering (DAF), corresponding to early- (E), mid- (M) and late- (Lt) matured seeds. Approximately 10 micrograms of total RNA (as estimated spectrophotometrically at λ260 and on agarose gel, by staining with ethidium bromide) was separated on a denaturing agarose gel and (Northern) blotted onto a nitrocellulose membrane. The RNA blot was probed as described in the previous section and the presence of a target sequence was determined by exposure of the hybridized blot to X-ray film. The profile of the northern blot is represented in **Fig. 2.3.2**. Binding of target to the probe was not observed in flower, leaf and root tissue RNA (lanes F, L and R). Likewise RNA from early (~12 DAF, lane E) and late (~40 DAF, lane Lt) stages of seed development also did not exhibit any signal. However, a weak signal did seem to emanate from mid stage (~25 DAF, lane M) seed RNA sample indicating presence of a homologous transcript in minimal levels.

Fig. 2.3.1 Southern Blot Hybridization profile of *cakpi*.
(A) Genomic digests. Restriction endonuclease digests of chickpea genomic DNA were separated by 0.8% agarose gel electrophoresis.
(B) The separated genomic digests were electroblotted onto nitrocellulose membrane; entire *cakpi* coding DNA (~600 bp) was radiolabeled by PCR amplification and used as probe.
In **(A)** and **(B)** Lanes: 1, *Dra*I; 2, *Bam*HI; 3, *Eco*RI; 4, *Eco*RV; 5, *Sal*I (cuts within *cakpi*).; C, Control (pCR2.1/*cakpi* construct); M1, 500 bp ladder; M2, λ DNA *Hin*DIII digest.



A **B**

Fig. 2.3.1

Fig. 2.3.2 Differential expression of *cakpi* in developing seeds.
Total RNA was isolated from chickpea flowers, leaves and roots, as well as from seeds at various stages of maturation – early (~12 DAF), mid (~25 DAF) and late (~40 DAF). The RNA was separated on a denaturing agarose gel and blotted onto nitrocellulose membrane. Radiolabeled coding DNA for *cakpi* was generated as a PCR product from the pCR2.1/*cakpi* construct and used as a probe.
(**A**) Hybridized blot showing putative hybridization signal (labeled with arrow),
(**B**) 18S rRNA from normalized (1μg each) total RNA samples visualized by staining with ethidium bromide after non-denaturing 0.8% agarose gel electrophoresis.

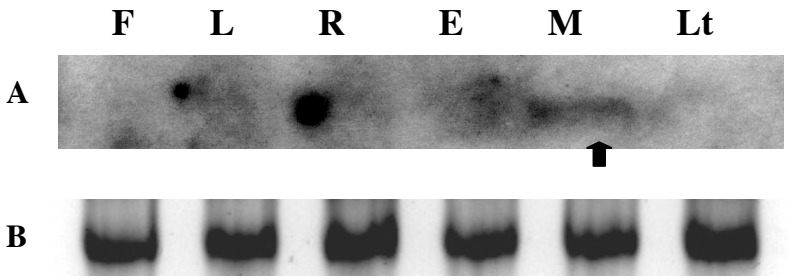


Fig. 2.3.2

Discussion

The phenotypic effects of gene expression are governed by genetic factors operating at various levels, which can be arbitrarily described for our convenience of understanding. At the first level, the sequence of the particular gene determines the primary structure of the polypeptide product, which contributes to protein stability and activity. Mutations at the activity-associated region(s) of a polypeptide may enhance, decrease or even completely abolish the biochemical function. The second level of genetic influence could be perceived as the number of individual copies of the gene in the genome. It can be expected that multiple copies of a gene, all of which are expressed similarly, would have more pronounced effects as compared to a single copy. The third level can be defined as the nature of upstream regulatory elements (*e.g.*, promoter region) which determine the transcription and, hence, rate of polypeptide synthesis; this transcriptional control is known to vary with innate factors such as age (maturation) and tissue-type as well as environmental factors such as nutrition and stress. The current section has probed these very factors that seem to govern the expression of the *cakpi* gene, which codes for a Kunitz-type proteinase inhibitor (CaKPI) that has been implicated with a defensive function against the larvae of the herbivorous insect pest *Helicoverpa armigera*.

CaKPI has been distinguished from conventional Kunitz type serine proteinase inhibitors by presence of a distinctly altered putative active site region; although arginine or lysine at P1 is considered a prerequisite for trypsin-binding and inhibitory activity, as in the soybean Kunitz type trypsin inhibitor, CaKPI exhibited comparable inhibitory activity against trypsin despite the active site variation (ref. **Chapter 3, Section 1**). It does appear conclusive that the minimal HGPs inhibitory activity in extracts of chickpea seeds are primarily the result of lower level of expression (CaKPI constitutes about 1% of total soluble protein content of seed), and that the amino acid alteration at the putative active site (inhibitory loop) does not have any noticeable effect on biochemical activity of CaKPI.

Available information in the GenBank (http://www.ncbi.nlm.nih.gov/) currently suggests the presence of two Kunitz type PIs in chickpea seeds – one of which (AY635930) is the currently described *cakpi* and the other a putative trypsin inhibitor (CAR276262). Both the nucleotide sequences exhibit ~70% sequence homology with each other, with the putative active site region exhibiting similar replacements of P1

amino acid residue. To confirm the actual number of genes similar to *cakpi* or presence of multiple alleles, blots of chickpea genomic DNA digests were hybridized with the previously cloned *cakpi*-coding region as a probe. The hybridization profile (**Fig. 2.3.1B**) clearly indicates two target sequences, one that strongly bound the probe and the other, weakly. The former is expected to be the *cakpi*-coding region (AY635930), whereas the latter is presumably the reported *cakpi* homolog (CAR276262). Contrary to existing thoughts on evolution of PI genes by multiple events of duplication and diversification (Mukhopadhyay, 2000; Zupunski *et al*, 2003; Lopes *et al*, 2004), the observed phenomenon in chickpea suggests that the level of diversification is very low for Kunitz type PIs, and perhaps even insignificant in comparison to other legumes like winged bean (Habu *et al*, 1992) and soybean, where up to five isoforms of Kunitz type PIs have been reported, so far. Similarly, about four different isoforms of Squash family PIs have been reported in bitter gourd (Telang *et al*, 2003), and at least eight different forms of wound-inducible PIs have been reported in capsicum (Tamhane *et al*, 2005). Though there are at least four different active isoforms of Bowman Birk type PIs (BBIs) reported in chickpea (Patankar *et al*, 1999), they are degraded by HGPs and hence their contribution towards resistance to herbivorous insect pests seems insignificant or none (Giri *et al*, 1998). It does seem ironical that the most diverse and abundant species of chickpea seed PIs (*i.e.*, the BBIs) do not contribute to insect resistance, whereas a low expressing Kunitz type PI seems promising in this approach. However, the low expression levels of CaKPI seem to be a boon in disguise because *H. armigera* could not adapt to higher levels of CaKPI (as described in **Chapter 3, Section 2**). The minimal diversification of the Kunitz type PIs in chickpea is reflected in the low copy number of observed homologs and may thus be the primary factor for low levels of CaKPI expression. It is entirely possible that the *cakpi*-like sequences represent a clade of 'young' genes, in the early stages of evolution – this could be partly inferred from the low copy number and partly from the presence of related sequences (that exhibit the putative active site variation) in very few of the studied legumes (previous section). In such a case, *cakpi* and the related sequences appear to be benchmarks that define a novel branch in the co-evolution of plant defenses in response to challenges by herbivorous insect pests.

Although expression of legume Kunitz type PIs have been previously reported from seed tissue (Jofuku and Goldberg, 1989), the possibility of expression may be considered in other tissues. Expression of higher amounts of PIs have been reported in tomato flowers, with respect to leaf tissue and this has been proven to possess defensive role against lower instar insects that feed on the delicate flower tissue (Damle *et al*, 2005). Likewise, expression of wound-inducible PIs has also been demonstrated in capsicum leaves, and it is known to contribute towards insect resistance (Tamhane *et al*, 2005). It is a well observed fact that the lower instar *H. armigera* larvae feed voraciously on the leaves and flowers of chickpea, before moving on to the developing seeds at the third instar. If CaKPI were expressed in these tissues (leaves, flowers), then the larvae would be exposed to the PI right from the neonate stage. Such continuous exposure to low levels of CaKPI could render the larvae insensitive to the PI. However, two observations are in conflict with this possibility, *viz*., (i) *H. armigera* larvae feeding voraciously on chickpea leaves exhibit vigorous growth, and, (ii) *H. armigera* larvae did not exhibit normal growth when exposed to dietary CaKPI from the neonate stage (**Chapter 3, Section 1**). Thus the possibility of early exposure of *H. armigera* larvae to CaKPI due to expression in leaves and flowers did not seem likely, although it could not be entirely ruled out. For this purpose expression analysis was carried out by Northern blotting of total RNA from flower and leaf tissue, followed by hybridization with *cakpi*-specific probe, as described previously. Expression of *cakpi* in roots was also probed to determine any secondary function. Transcription of *cakpi* was determined to be very low or absent in these tissues as evident from the lack of binding of probe to target transcript (**Fig. 2.3.2**, lanes F, L, R**)**. It does appear that *cakpi* expression in chickpea is confined to the seed tissue. Among the various stages of seed development studied (early, ~12 DAF; mid, ~25 DAF; late ~40 DAF), *cakpi* transcripts were detected in very low levels only in mid-mature seed RNA (**Fig. 2.3.2**, lanes E, M, Lt). Detection of CaKPI activity specifically in seed tissue, coupled with absence of transcripts in other plant parts indicates the presence of a seed-specific cis-acting regulatory element (promoter) probably governing *cakpi* expression. Since the mid and late mature seeds are characterized by active synthesis and accumulation of storage proteins, synthesis of PIs for protection of these storage proteins from endogenous plant proteinases appears likely. Since CaKPI was initially described by the HGPs inhibitory activity in

chickpea seeds (Patankar *et al*, 2001), and was not found to be associated with inhibition of endogenous proteinases, the hypothesis for a defensive role is strengthened. It has been previously reported that the HGPs inhibiting activity in chickpea seeds increases in response to insect chewing (Patankar er al, 2001). Since this property has been ascribed to CaKPI, it does appear possible that the regulatory elements governing expression of *cakpi* in chickpea seeds are also associated with defense responsiveness. Transgenic analysis of a wound-responsive promoter from poplar (*Populus tremuloides*) has shown that genes expressed under this promoter are transcribed similar to seed storage protein genes (Hollick and Gordon, 1995). Further, it has also been demonstrated that downregulation of PIs in *Solanum americanum* is innately related to defective seed development (Sin *et al*, 2006). Hence it does appear that seed-specificity and wound responsiveness may be linked to synthesis of storage proteins; although Reech *et al* (1997) have shown that legume seed PIs do not inhibit endogenous plant proteinases, the significance of PI expression along with storage proteins would lie in protecting the storage proteins against herbivorous insect pests.

The upregulation of plant PIs has been described as a defensive response, which aims to counter the arsenal of insect digestive proteinases. The elevated levels of PIs have been shown to afford protection to the plant against the infesting herbivorous insect pests. Subsequent studies on CaKPI involved *in vivo* assays to determine whether elevated levels of CaKPI do have any significant inputs on host plant (chickpea) defenses against *H. armigera* – these approaches have been detailed in the subsequent chapter.

Conclusion

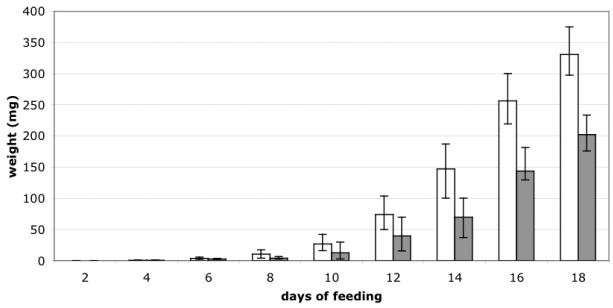
cakpi was observed to be a single copy gene with at least one detectable homolog (~70% sequence similarity) that has been previously reported in the GenBank. Thus, *cakpi* was found to exhibit low degree of diversification, indicative of possibly slow evolutionary rate or recent evolution. In such a case, *cakpi*, could be representative of recently evolved/evolving genes that may form part of the host-plant defenses against herbivorous insect pests. Further *cakpi* transcripts were not detected in any tissue apart from seeds, indicating possible presence of a tissue specific promoter.

CHAPTER III:

Defensive Aspects of CaKPI in Chickpea-Podborer Association

SECTION 1: *in vitro* and *in vivo* activities of CaKPI

The research work described in this section is part of two full-length papers, which has been published in Plant Molecular Biology (Srinivasan et all, 2005a) and Journal of Insect Physiology (Srinivasan et al, 2005b)



Introduction

The defensive role of plant PIs against herbivorous insects has been well studied; PIs bind to and inhibit the gut proteinases of insect larvae leading to a decrease in protein-digestive capacity. Reduced protein breakdown, in turn causes a depletion of the available levels of free amino acids that would have otherwise been assimilated (Broadway and Duffey, 1986; Hilder *et al*, 1987; Broadway, 1996). Under such conditions, the larval physiology faces a nutritional setback, which directly affects larval growth and development (Harsulkar *et al*, 1999; Giri *et al*, 2003; Telang *et al*, 2003; Srinivasan *et al*, 2005a). The outcomes of such antagonistic agents to larval growth and development are usually non-lethal and multiple; one of the most significant aspects is the lowered fertility/fecundity of the adult insect (DeLeo and Gallerani, 2002; Telang *et al*, 2003; Tamhane *et al*, 2005). These effects translate into a reduced capacity of the insects to thrive and reproduce resulting in a gradual decimation of the insect population over the next few generations. However, all PIs may not possess the capacity to cause such disastrous effects due to a variety of reasons. Though PIs may exhibit substantial activity against commercial proteinases, they might not be effective against the isoforms present in the larval gut (Edmonds *et al*, 1996). Secondly, *in vitro* inhibitory activity of PIs against larval gut proteinases may not translate into a strong *in vivo* action on the insect physiology, since the insects also have a remarkable ability to adapt to various PIs (Broadway *et al*, 1995; Jongsma and Bolter, 1997). Hence, screening of PIs for use in insect control programs is crucial and should be based on the ability of PIs to not only effectively inhibit larval gut proteinases, but also to exert significant antagonistic activity on growth and development of the larvae. In the current section, we explore the inhibitory activities of CaKPI on various commercial and *Helicoverpa armigera* gut proteinases (HGPs) as well as investigate the effects of dietary inclusion of CaKPI on *H. armigera*.

Materials and methods

Materials

Trypsin, chymotrypsin, elastase, Benzoyl-DL-arginyl-p-nitroanilide (BA*p*NA), Succinyl-alanyl-alanyl-alanyl-prolyl-leucyl-p-nitroanilide (SAAAPL*p*NA) and Succinyl-alanyl-alanyl-alanyl-p-nitroanilide (SAAA*p*NA) were procured from Sigma

Chemicals, USA. All other chemicals used were of analytical grade and procured locally.

Preparation of fresh HGPs

500 mg tissue was weighed out from freshly dissected *H. armigera* larval midguts. Gut tissue was frozen in liquid nitrogen, ground to a fine powder and extracted in 500 µL of 200 mM Glycine-NaOH buffer (pH 10.0) for 2 h at 4°C. The extract was then centrifuged at 10,000 rpm (4°C, 10 min) and the supernatant was used for all solution assays. 5 g *H. armigera* fecal matter was extracted separately in 20 mL of 200 mM Glycine-NaOH buffer (pH=10.0) for 2 h at 4°C. To this, 500 µL of gut extract, prepared as described above, was added, the volume made up to 25 mL with 200 mM Glycine-NaOH buffer (pH=10.0) and used for *in gel* inhibitor activity visualizations.

In vitro assays

The synthetic substrates sulphanilamide-azocasein, BA*p*NA, SAAAPL*p*NA and SAAA*p*NA were used for assaying total, trypsin-, chymotrypsin- and elastase- like activities of HGPs, respectively. For all assays, HGPs activity was first calibrated by end point titration of varying amounts (1, 2, 5 and 10 µL) of HGPs with excess substrate across a fixed time interval (10 min for BA*p*NA, SAAAPL*p*NA and SAAA*p*NA respectively and 30 min for sulphanilamide-azocasein) at 37°C. One unit of BA*p*NAase, SAAAPL*p*NAase and SAAA*p*NAase activities were defined as the amount (µL) of HGPs required for liberation of 125 mmoles of p-nitroaniline (extinction coefficient, 8800 $M^{-1}cm^{-1}$), leading to an increase in absorbance by 1 OD at 410 nm in 1 min, in 1mL reaction volume. One unit of azocaseinolytic activity was defined as the amount (µL) of HGPs required for liberation of 1 µmole of sulphanilamide (extinction coefficient, 900 $M^{-1}cm-^{1}$), leading to an increase in absorbance by 1 OD at 450 nm in 1 min, in 1mL reaction volume. Total HGPs activity was expressed as azocaseinase units (AzU $mL^{-1}min^{-1}$) by the equation AzU =1000 / *x*, where *x* represents volume of HGPs (µL) required for unit activity. Similarly, BA*p*NAase, SAAAPL*p*NAase and SAAA*p*NAase activities were also expressed as the corresponding trypsin (TU$mL^{-1}min^{-1}$), chymotrypsin (CU$mL^{-1}min^{-1}$) and elastase (EU$mL^{-1}min^{-1}$) activities. Proteinase activities were calculated for both, control diet fed (henceforth referred to as 'control') as well as for CaKPI-incorporated

diet fed (henceforth referred to as 'sensitized') HGP. For inhibitory assays, HGPs corresponding to 0.4 U of each proteinase activity was then incubated with increasing amounts of CaKPI (2, 4 and 6 μg) for 10 min and residual proteolytic activity was assayed as above. All *in vitro* assays were carried out independently in triplicates and average of all values were calculated. Standard error was determined and is indicated in all applicable figures and table.

Rearing of insects

Healthy, actively feeding *H. armigera* larvae collected from fields were transferred onto artificial diet and maintained. The composition of the diet was as described by Nagarkatti and Prakash (1974), the components of 650 mL diet being, chickpea seed meal, 77 g; wheat germ, 5.6 g; dried yeast powder, 19.2 g; casein, 12.8 g; ascorbic acid, 4.6 g; methyl para-hydroxy benzoate, 1.5 g; sorbic acid, 0.8 g; streptomycin sulphate, 0.2 g; cholesterol, 0.2 g; formaldehyde, 1 mL; multivitamin drops, 0.8 mL; vitamin E, 12 g, and agar, 10 g. The pupae were then transferred into individual 50 mL screw-capped vials containing about 5 g soil bed. Once adult moths emerged, they were transferred into polyethylene jars covered at the mouth with a black cotton cloth. Adult moths were provided with sterile cotton swabs imbibed with a liquid diet of 10% (w/v) honey or sucrose and 1% (w/v) vitamin E. To ensure greater genetic homogeneity among test populations, the insects were maintained on control diet for a minimum of three generations, after which they were used for *in vivo* assays.

H. armigera feeding assay

Anti-metabolic effects of CaKPI on growth of *H. armigera* larvae were investigated by insect feeding bioassays. For these assays, artificial diet as suggested by Nagarkatti and Prakash (1974) for *H. armigera* (as described in the previous paragraph) was used. The amount of CaKPI required per gram diet (1x concentration) for maximum inhibition of total HGP from a single larva at third instar was calculated by the formula $x = (T*I)/E$ where, x = amount of inhibitor required per gram diet, T = total volume of one insect gut, I = least amount (μL) of inhibitor (at a concentration of 1 μg μL^{-1}) required for maximum or complete inhibition of 0.4 U of enzyme, E = volume (μL) of enzyme corresponding to 0.4 U. Single larval midgut weighs ~40 mg

and corresponds to ~48 μL. Substituting this in the above equation as also 1.2 μL as volume of HGP preparation (0.4 U) and 2.0 μL CaKPI (2.0 μg), the equation solves to give 80.4 μg as the amount of CaKPI (1x concentration) required per gram diet for maximum inhibition of total gut proteolytic activity of one insect larva. For the actual feeding assay, CaKPI was used at a concentration of 0.5x (40.2 μg) per gram artificial diet. Two sets of 25 insects each was used for this bio-assay wherein, one set was allowed to feed on the artificial diet without inhibitor and the other fed on inhibitor containing diet. To minimize variations within and between sets, a homogenous third generation laboratory reared insect culture was employed. Soon after hatching, neonate larvae were transferred into individual vials containing either control or inhibitor containing diet. Diets were changed as and when they had been consumed or every alternate day. Larval instar was defined by visually monitoring the larvae for molting. Larval weights were recorded every alternate day beginning from after 48 h of feeding. Percent weight gain was calculated for every 48 h period, as also percent weight difference between control and inhibitor fed larvae. As a statistical method, standard error was calculated from average weights of 25 larvae.

Results

In vitro inhibition of trypsin and HGP activities

Purified recombinant protein was assayed for inhibitory activity against trypsin, chymotrypsin, subtilisin and HGP. Stoichiometric inhibition of (bovine) trypsin activity towards a synthetic substrate (BApNA) was observed, which was comparable to the inhibition of trypsin observed with SKTI; inhibition was approximately linear with increasing concentration of recombinant protein, and at an equimolar ratio of trypsin and inhibitor (1:1), trypsin activity was almost completely inhibited (**Fig. 3.1.1A**). Inhibition of chymotrypsin (bovine) or subtilisin (bacterial) was not recorded even with five-fold molar excess of inhibitor over that of the protease (data not shown). CaKPI also exhibited 60% maximum inhibition of trypsin-like activity in HGP (**Fig. 3.1.1B**), which was less than the 71% maximum inhibition exhibited by SKTI. The total proteolytic activity of HGP, as assayed with the substrate azocasein, was inhibited to a maximum of 69% by CaKPI (**Fig. 3.1.1B**), significantly higher than the 41% maximum inhibition observed with SKTI. The amount of inhibitor for 50%

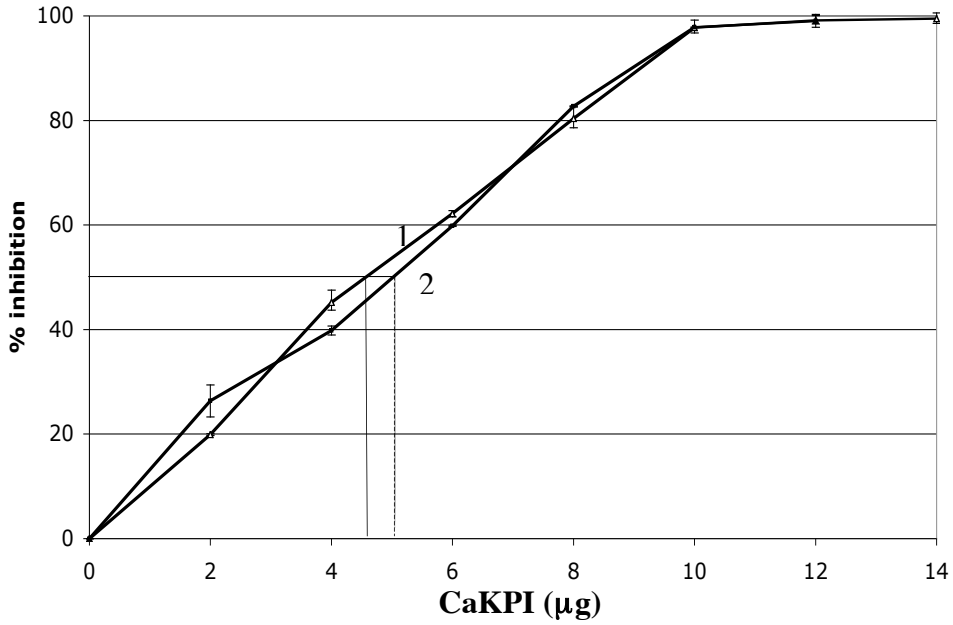
inhibition (IC_{50}), for standard trypsin activity (0.4U mL^{-1}) was estimated to be 5 μg mL^{-1} (2.5 x 10^{-7} M) for CaKPI, which is comparable to that of SKTI for trypsin (**Fig. 3.1.1A**). The IC_{50} value of CaKPI for total HGP activity was <2 μg mL^{-1} (<10^{-7} M) and that of SKTI was 2 μg mL^{-1} (10^{-7} M), respectively. Finally the IC_{50} values against trypsin like activity of HGP were 2 μg mL^{-1} (10^{-7} M) and <2 μg mL^{-1} (<10^{-7} M) for CaKPI and SKTI respectively (**Fig. 3.1.1B**). These results suggest that CaKPI is potentially an effective and better inhibitor of digestive proteolysis in *H. armigera* larvae, besides possessing inhibitory activity against bovine trypsin.

Inhibition of host and non-host gut proteinases by CaKPI

Inhibitory potential of CaKPI was assayed against total proteolytic activity of HGPs from fourth instar larvae fed on sweetpea (*Pisum sativum*), pigeonpea (*Cajanus cajan*), okra (*Abelmoschus esculentus*), as well as artificial diets incorporated with PIs from winged bean (*Psophocarpus tetragonolobus*, WBI), potato (*Solanum tuberosum*, PinII), a combination of WBI and PinII and finally, CaKPI incorporated diet. The sensitivity of the various HGPs activities to CaKPI is represented in **Fig. 3.1.2A**. Among the three host plants fed larvae, CaKPI inhibited sweetpea fed larval HGPs by upto 74%, pigeonpea fed larval HGPs by 72% and okra fed larval HGPs by 69%. CaKPI inhibited 52% of total gut activity of WBI fed larvae and 54% of total gut activity of PinII fed larvae. Comparatively, but not significantly, higher inhibition (57%) was observed with HGPs from larvae fed on a combination of WBI and PinII (**Fig. 3.1.2A**). HGPs activities from second, third, fourth, fifth and sixth instar larvae were also tested for their inhibition by CaKPI. No significant variations in sensitivities towards CaKPI were observed between the instars; the inhibition ranged from a maximum of 65%, at the fourth instar, to a minimum of 53%, at the fifth instar (**Fig. 3.1.2B**).

Fig. 3.1.1. Inhibitory activities of CaKPI and SKTI against various proteinases.
Recombinant CaKPI protein was tested for potential inhibitory activity against trypsin as well as HGP using different substrates. Increasing amounts of inhibitor led to complete inhibition of trypsin at equimolar ratio of inhibitor:enzyme. Trypsin like activity of HGP was inhibited by 60% and total HGP proteolytic activity was inhibited by 68%. IC_{50} values are marked on the X-axis. Standard error bars are indicated.
(A) Titration of against trypsin: 1, CaKPI; 2, SKTI;
(B) Titration of CaKPI and SKTI against HGP: 3, CaKPI inhibition of total gut proteinases; 4, CaKPI inhibition of gut trypsin; 5, SKTI inhibition of total gut proteinases; 6, SKTI inhibition of gut trypsin.

A


100
80
60
40
20
0
% inhibition
1
2
0
2
4
6
8
10
12
14
CaKPI (µg)

B

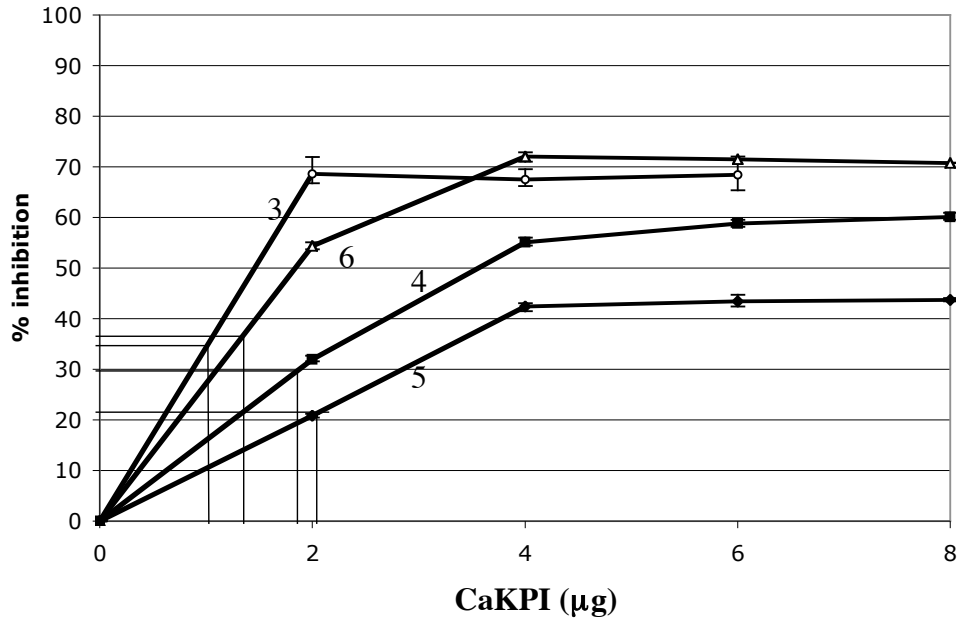
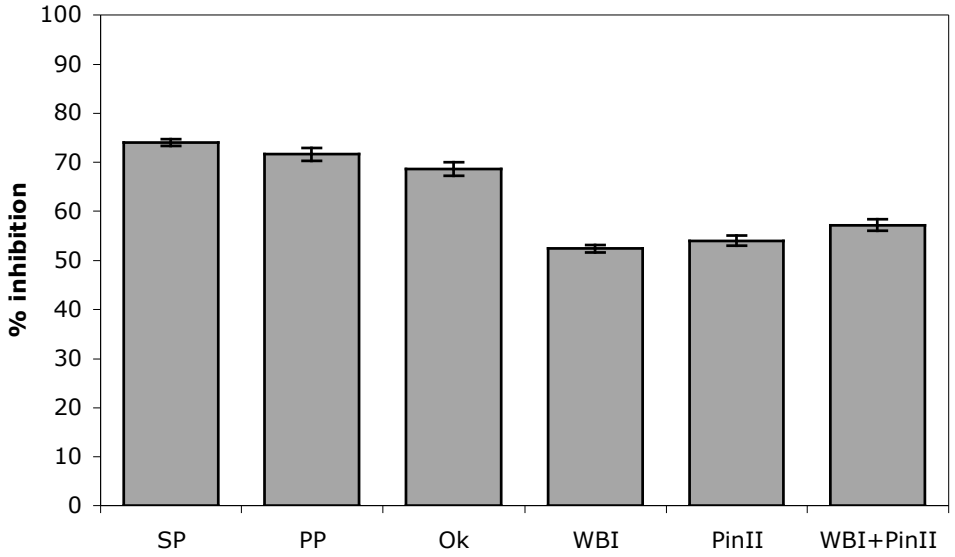

100
90
80
70
60
50
40
30
20
10
0
% inhibition
3
6
4
5
0
2
4
6
8
CaKPI (µg)

Fig. 3.1.1

Fig. 3.1.2. Inhibitory activity of CaKPI against various HGPs preparations.
(A) Inhibition of HGPs derived from 4th instar larvae fed on the host plants: sweetpea (SP); pigeonpea (PP); okra (Ok); and on artificial diets incorporated with the non host PIs: winged bean chymotrypsin inhibitor (WBI); potato proteinase inhibitor II (PinII); a combination of the two PIs (WBI+PinII).
(B) Inhibition of HGPs derived from larvae at second (2), third (3), fourth (4), fifth (5) and sixth (6) instars.
Standard error bars are indicated.

A


100
90
80
70
60
50
40
30
20
10
0
% inhibition
SP
PP
Ok
WBI
PinII
WBI+PinII

B

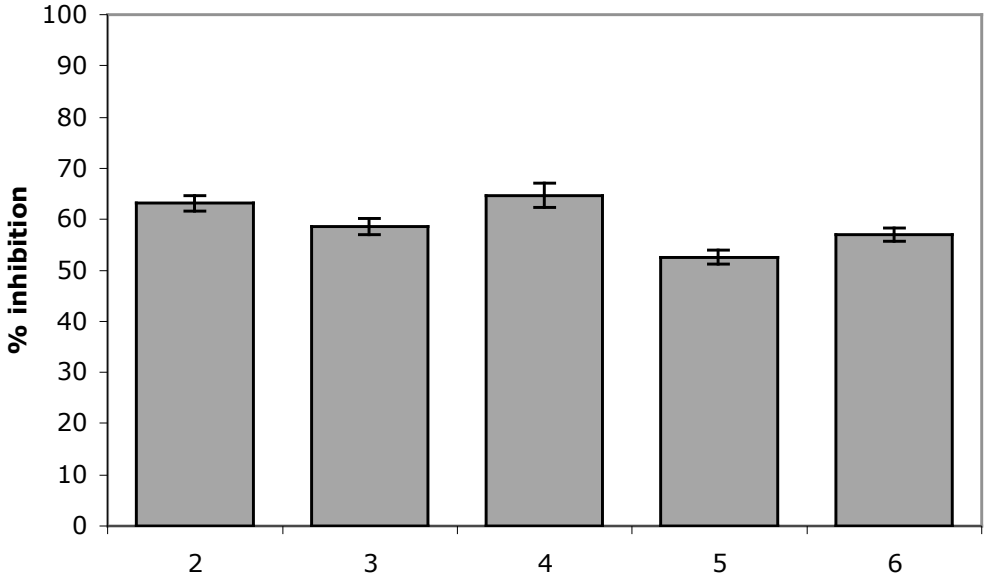

100
90
80
70
60
50
40
30
20
10
0
% inhibition
2
3
4
5
6

Fig. 3.1.2

Anti-metabolic effects of CaKPI on developing H. armigera larvae

Efficient *in vitro* inhibitors of insect gut proteinases may not be effective anti-metabolites *in vivo* (Edmonds *et al*, 1996). To test the potency of CaKPI as an insect-control agent, feeding assays were carried out with CaKPI-incorporated diet on a test population of *H. armigera* larvae. While the larvae feeding on control diet developed and gained weight similar to healthy field collected larvae, the inhibitor fed larvae showed a distinct lag. No significant weight differences were observed between control and inhibitor fed sets in the first and second instars (second day and sixth day respectively). However, at the third instar (eighth day) the average inhibitor fed larval weight was 3.8 mg, almost 64% lower than the average control weight of 10.5 mg. Similarly, at the fourth instar, average inhibitor fed larval weight was 39.7 mg, 47% lower than the average control weight of 74.1 mg. By the fifth instar (eighteenth day), the average inhibitor fed larva weighed 202 mg, 39% lower than the control at 330.8 mg (**Fig. 3.1.3**).

Discussion

During the course of inhibitory assays, though chymotrypsin- or subtilisin-inhibitory activities were not observed, CaKPI exhibited stoichiometric trypsin inhibition and, at equimolar concentrations, trypsin activity was completely inhibited by CaKPI. This suggested that CaKPI has a single inhibitory site for trypsin and, therefore, for trypsin-like enzymes. However, the amino acid sequence at the putative inhibitory site suggests no trypsin specificity, since there is no arginine (R) or lysine (K) residue present that would give specificity towards trypsin; the glycine-isoleucine-serine (G-I-S) motif at the active site of CaKPI does not contain a peptide bond cleavable by trypsin, leading to two possibilities: (i) the active site is in fact elsewhere on the molecule (there is a lysine residue 5 amino acids N-terminal to the putative active site motif); or (ii) inhibition is not based on the classical model for Kunitz inhibitors, where the scissile bond of the PI is positioned at the trypsin active site. A useful lead was provided by comparison of the trypsin inhibitory activities of CaKPI and SKTI; both exhibit a similar stoichiometric inhibition of trypsin activity, hinting that the actual mechanism of trypsin binding and inhibition by CaKPI was equally effective as the classical Kunitz model, *viz*., SKTI.

Fig. 3.1.3. Effect of CaKPI on growth of *H. armigera* larvae.
Recombinant CaKPI protein was employed in controlled feeding experiments. One set of larvae were allowed to feed on artificial diet without any inhibitor (Control, white bars) and another s*et al*lowed to feed on inhibitor incorporated diet (Test, grey bars). Weights were recorded every alternate day beginning from after 2 days of feeding. Graph shows average weights from each set of 25 larvae. Standard error bars are also indicated.

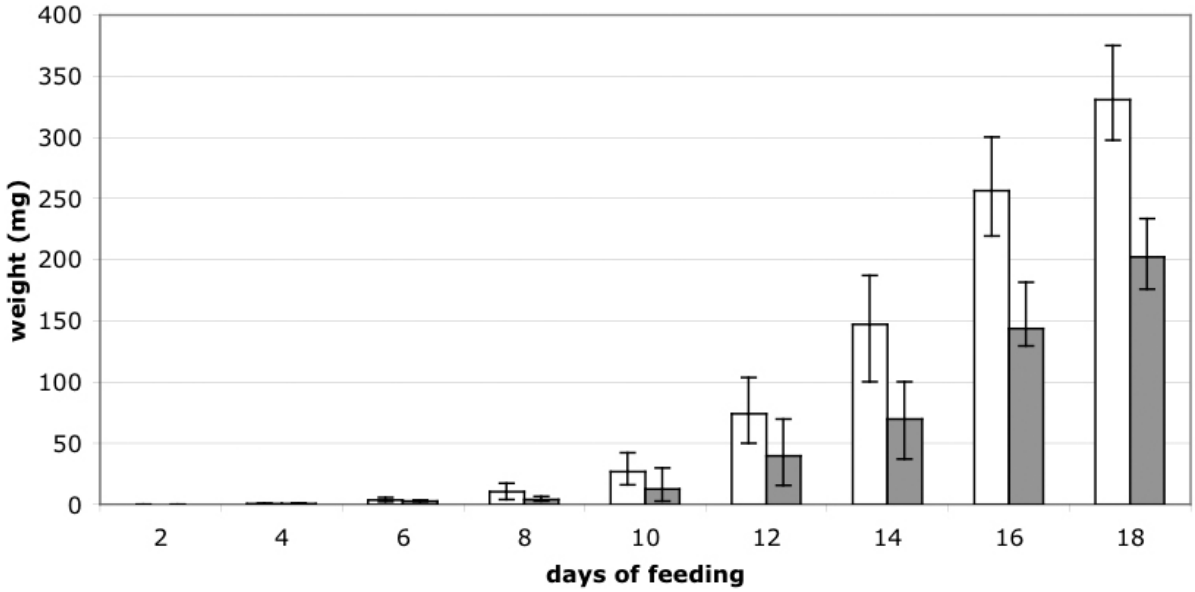

400
350
300
250
200
150
100
50
0
weight (mg)
2
4
6
8
10
12
14
16
18
days of feeding

Fig. 3.1.3

Inhibitory potential of CaKPI was investigated against an unfractionated gut proteinase extract from *H. armigera*, which aimed at providing an insight into the behavior and activity of the inhibitor in a complex proteolytic environment such as that exists in the insect gut. It was found that CaKPI is more active towards general proteolysis by insect gut extracts than towards trypsin-like activity, and more active towards insect gut trypsin-like activity than towards bovine trypsin, which suggests that its specificity may not be like that of a "classical" trypsin inhibitor. Inhibition of gut trypsin-like activity was higher with SKTI than CaKPI, but the total inhibition of gut proteolytic activity by CaKPI was significantly greater than with SKTI. The higher activity towards general proteolysis can be explained as the inhibitor being able to inhibit proteinases other than those, which possess trypsin-like activity. Even though majority of *H. armigera* gut proteinases are trypsins and -like enzymes (Patankar *et al*, 2001), higher inhibition of total proteolytic activity (including other minor specificities) would translate into more pronounced detrimental effects on larval digestion and physiology. In this context, CaKPI is evidently a better choice for inhibition of total gut proteolytic activity.

CaKPI not only exhibited high inhibition of host plant-fed HGPs activities, but was also equally effective in inhibiting non-host PI fed HGPs; even though sensitivity of non-host PI fed HGPs to CaKPI was expected to be less or absent, substantial inhibition of these HGPs was observed. It may be inferred that CaKPI was significantly, if not equally, effective against larvae, adapted to varying dietary protein composition of various host plants as well as varying PI content of non-host plants. Previously, Harsulkar *et al* (1999) have enunciated the importance of using non-host PIs in succession towards countering larval adaptation. However, CaKPI seems suitable for developing insect resistance in target crops, in spite of being host-plant derived. Further, CaKPI was consistently effective in inhibiting HGPs derived from various instar larvae, indicating that it could act as a continuous means of antagonism to the larvae during all stages of insect growth and can thus be expected to restrict larval growth and development.

As the next obvious step, the *in vivo* effects of CaKPI against developing larvae of the insect pest *H. armigera* was assayed by controlled feeding experiments. CaKPI was observed to inflict maximum adverse effect on third instar larvae and the effect progressively decreased until the fifth instar. Under chickpea field conditions, *H.*

armigera larvae feed on leaves of chickpea until the second instar and then, at the third instar and onwards, move on to the developing seeds. The gut proteinase profile is expected to change around the third instar when the larvae start feeding on seeds, which have a qualitatively and quantitatively different protein content. Any antagonistic activity exerted by the inhibitor on larval physiology at this stage would have the maximum impact in tilting the balance against the insect pest. Insect feeding bioassays have thus unambiguously shown that this inhibitor has the potential to be used as a source of developing *H. armigera* tolerance, in chickpea.

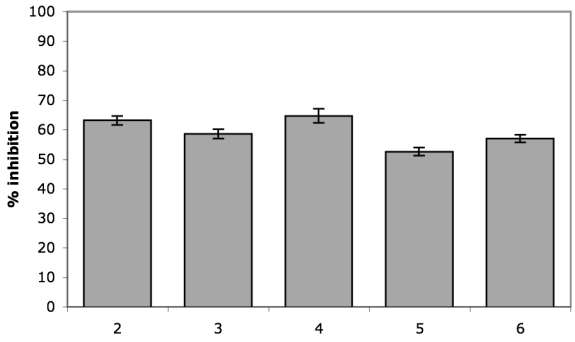
Although PIs exert strong antagonistic activity on larval physiology, the insects are known to respond by qualitatively or quantitatively changing the profile of digestive proteinases in response to the PI. In order to probe the possible adaptive responses of *H. armigera* and whether adaptation is actually achieved, studies were carried out to this effect and are described in the following section.

Conclusion

CaKPI exhibited stoichiometric inhibition of commercial trypsin, comparable to the activity of the soybean Kunitz trypsin inhibitor. CaKPI did not exhibit any activity against chymotrypsin or subtilisin. Similarly, CaKPI also showed inhibition of HGP trypsin activity as well as total HGP activity. Inhibitory activities of CaKPI on HGPs derived from larvae feeding on various host as well as non-host PI based diets were high. CaKPI also exhibited similar inhibitory profile on HGPs derived from larvae at various instars. Finally, the anti-metabolic activity of CaKPI on developing *H. armigera* larvae was demonstrated by feeding assays.

SECTION 2: *Helicoverpa armigera* does not adapt to dietary CaKPI

The research work described in this section is part of a full-length paper, which has been published in Journal of Insect Physiology (Srinivasan et al, 2005b)



Introduction

As discussed previously, ingestion of effective PIs is expected to derail the digestive process and the effect of this antagonistic pressure manifests in myriad forms of developmental as well as functional abnormalities. However, insects are known to respond to presence of dietary by upregulating the expression level of gut proteases or by synthesis of inhibitor-degrading or –insensitive proteinases (Jongsma and Bolter, 1997; Paulillo *et al*, 2000; Brito *et al*, 2001; Bown *et al*, 2004). Upregulation of digestive proteinases aims to fill the void in protein digestion caused due to binding of inhibitor to the target proteinase(s). While PI insensitive proteases remain free from inhibition due to lack of proper PI binding sites, PI degrading proteases which bind to PIs resume digestive activity after degrading and effectively neutralizing the PI. Once again the insect has a choice of adaptive measures and these can be used in all possible combinations to gain an upper hand over the PI. Though insects exhibit such responses, adaptation is achieved only if the outcome is a successful metabolism of the altered diet, resulting in normal growth and development. Thus, though response always precedes adaptation, the latter might not always be observed. Insect responses also remain unpredictable due to conflicting observations. Since the gut environment is a complex and dynamic mix of proteases it would be difficult to predict responses towards a specific PIs unless we have complete understanding of genetic and functional properties of insect digestive proteases as well as a well characterized dietary protein content.

Materials and methods

Materials

Bovine trypsin, chymotrypsin, elastase, Benzoyl-DL-arginyl-p-nitroanilide (BA*p*NA), Succinyl-alanyl-alanyl-alanyl-prolyl-leucyl-p-nitroanilide (SAAAPL*p*NA), and Succinyl-alanyl-alanyl-alanyl-p-nitroanilide (SAAA*p*NA) were procured from Sigma Chemicals, USA. mRNA purification kit, first strand cDNA synthesis kit and proof-reading DNA polymerase were from Clontech, USA. X-ray films and developer were purchased from Kodak, India. All other chemicals used were of analytical grade and procured locally.

Stability of inhibitor towards proteolytic degradation

Approximately 5 µg of CaKPI was mixed and incubated with 5 µg of each, trypsin and chymotrypsin, respectively, and incubated at 37°C for 0, 30 and 180 min. As a positive control, 5 µg of untreated CaKPI was incubated in sterile deionized water for 180 min. After incubation, the mixtures were immediately denatured with 2-mercaptoethanol and separated by 15% sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE), following which, the entire gel was stained with coomassie brilliant blue R-250 (CBB R250). Structural stability of CaKPI towards trypsin and chymotrypsin was determined by monitoring change in migration of the proteinase treated CaKPI sample vis-à-vis the untreated (control) sample. Similarly, 2 µg of CaKPI was mixed and incubated with 10 µL of a fresh HGPs preparation (corresponding to 0.05 units of azocaseinolytic activity, $AzUmin^{-1}$; described in the next section) and incubated at 37°C for 0, 30 and 180 min. As a positive control, 2 µg of untreated CaKPI was incubated for 180 min. After incubation, the mixtures were immediately separated by 12% native PAGE, following which the entire gel was equilibrated in 200 mM Glycine-NaOH (pH 10.0) for 5 min at 4°C, followed by a freshly prepared HGP solution (as described in previous section) for 10 min at 4°C and finally rinsed in 200 mM Glycine-NaOH (pH 10.0). The gel was overlaid on X-ray film whilst taking care not to introduce any air bubbles between the gel and the X-ray film (Pichare and Kachole, 1994). After incubation at 25°C for 5, 10 and 15 min, respectively, the X-ray film was washed with cold or lukewarm water until gelatin digested by the imbibed HGP had cleared. Activity bands were visualized as zones of undigested gelatin against the hydrolyzed background on the surface of the X-ray film.

In vitro assays

The synthetic substrates sulphanilamide-azocasein, BA*p*NA, SAAAPL*p*NA and SAAA*p*NA were used for assaying total, trypsin, chymotrypsin and elastase -like activities of HGPs, respectively. For all assays, HGPs activity was first calibrated by end point titration of varying amounts (1, 2, 5 and 10 µL) of HGPs with excess substrate across a fixed time interval (10 min for BA*p*NA, SAAAPL*p*NA and SAAA*p*NA respectively and 30 min for sulphanilamide-azocasein) at 37°C. 1 unit of BA*p*NAase, SAAAPL*p*NAase and SAAA*p*NAase activities were defined as the

amount (μL) of HGPs required for liberation of 125 mmoles of p-nitroaniline (extinction coefficient, 8800 $M^{-1}cm^{-1}$), leading to an increase in absorbance by 1 OD at 410 nm in 1 min, in 1mL reaction volume. One unit of azocaseinolytic activity was defined as the amount (μL) of HGPs required for liberation of 1 μmole of sulphanilamide (extinction coefficient, 900 $M^{-1}cm^{-1}$), leading to an increase in absorbance by 1 OD at 450 nm in 1 min, in 1mL reaction volume. Total HGPs activity was expressed as azocaseinase units (AzU $mL^{-1}min^{-1}$) by the equation AzU =1000 / *x*, where *x* represents volume of HGPs (μL) required for unit activity. Similarly, BA*p*NAase, SAAAPL*p*NAase and SAAA*p*NAase activities were also expressed as the corresponding trypsin ($TUmL^{-1}min^{-1}$), chymotrypsin ($CUmL^{-1}min^{-1}$) and elastase ($EUmL^{-1}min^{-1}$) activities. Proteinase activities were calculated for both, control as well as for CaKPI fed HGP. For inhibitory assays, HGPs corresponding to 0.4 U of each proteinase activity was then incubated with increasing amounts of CaKPI (2, 4 and 6 μg) for 10 min and residual proteolytic activity was assayed as above. All *in vitro* assays were carried out independently in triplicates and average of all values were calculated. Standard error was determined and is indicated in all applicable figures and table.

Preparation of RNA from insect gut tissue and synthesis of first strand cDNA

Midguts from fourth instar *H. armigera* were removed by dissecting the larvae mid-ventrally under sterile conditions. Food contents were removed from the dissected guts by gentle agitation of guts in diethyl pyrocarbonate (DEPC) treated water. The guts were then snap frozen in liquid nitrogen and stored at –80°C. Total RNA was extracted from dissected gut using the acid guanidinium thiocyanate phenol chloroform method (Chomczynski and Sacchi, 1986). Quality of the RNA was checked on agarose gel, by staining with ethidium bromide (EtBr) and quantified by λ_{260} on a spectrophotometer. After estimation, concentrations of RNA, prepared from control as well as CaKPI fed larval guts, were equalized by dilution with DEPC treated water. Contamination of genomic DNA in the RNA preparation was checked by PCR amplification using gene specific primers and a DNA dependent DNA polymerase (e.g., Taq polymerase) on the total RNA preparation as template. Once purity of the RNA preparation was established, first strand cDNA synthesis was carried out using1 μg of normalized total RNA and a commercially available reverse

transcriptase enzyme and oligo-deoxyThymidine (oligo-dT) primer, as per the manufacturers' instructions.

Quantitative PCR

Oligonucleotide primers were synthesized (Chougule *et al*, 2005) based on available sequence information of various *H. armigera* gut proteinases including trypsins (Y12271, Y12269, Y12277, Y12276, Y12270), chymotrypsins (Y12287, Y12281, Y12273), aminopeptidases (AF535165, AY038606, AF441377, AF535166, AY052651), carboxypeptidases (AJ005177, AJ005178, AJ005176), elastase (PhelasB2) and cathepsin-B like proteinase (AY222788). These primers were employed in a polymerase chain reaction using 10, 100 and 1000 fold dilutions of the cDNA derived from a midgut mRNA preparation, as described above. A typical cycle consisted of a denaturation step at 94°C for 1 min, followed by primer annealing at 45°C for 30 sec, and a final extension at 72°C for 1 min. The PCR reaction consisted of 30 such cycles and a final step of incubation at 72°C for 5 min. PCR reactions were carried out in three independent replicates to rule out inconsistencies. The number of cycles required for efficient PCR, have been previously standardized (Chougule *et al*, 2005) as to give 50% amplification of the target transcript.

Results

Stability of native CaKPI towards proteolytic digestion by various proteinases

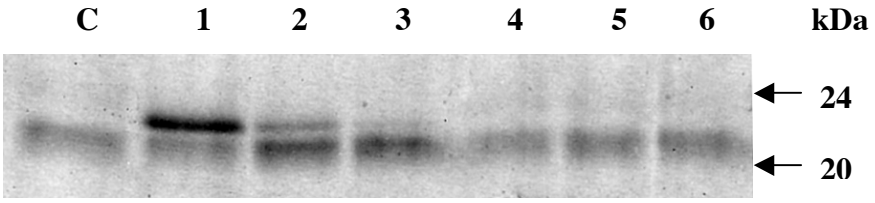
Stability of the native CaKPI polypeptide towards various proteolytic activities was assessed by co-incubation of CaKPI and the proteinase(s) for defined time intervals after which, the components of the mixture were separated by PAGE. CaKPI was either visualized, after SDS-PAGE, by staining with CBBR-250, or detected, after native-PAGE on the basis of its inhibitory activity by the gel X-ray film contact print method. Staining with CBBR250 showed that treatment of CaKPI with trypsin or chymotrypsin for up to three hours did not alter the electrophoretic mobility, implying that there was no degradation of native CaKPI by either trypsin or chymotrypsin (compare lanes of control and treated samples in **Fig. 3.2.1A**). CaKPI was observed to retain its activity as revealed by *in gel* HGP-inhibiting activity even after three hours

of incubation in presence of control diet fed (henceforth referred to as 'control') or CaKPI-incorporated diet fed (henceforth referred to as 'sensitized') HGPs; it is interesting to note that the average food retention time in *H. armigera* larval gut being three to four hours. This suggested that the inhibitory activity was not lost due to action of HGPs. Secondly, the activity band representing HGP inhibition by CaKPI was observed to co-migrate in case of both, the HGPs treated as well as the untreated samples (compare lanes of control and sensitized HGPs treated samples in **Fig. 3.2.1B**). This clearly indicated the stability of CaKPI structure and activity towards HGPs as against the chickpea Bowman-Birk type PIs, which were readily degraded (Giri *et al*, 1998).

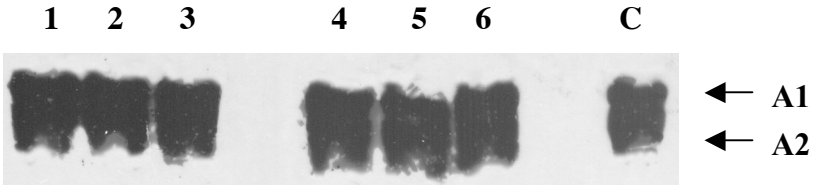
Changes in gut proteinase activities and sensitivities

CaKPI feeding caused marginal increase in the total- and trypsin-like activities of HGPs by about 7% each whereas the chymotrypsin-like activity remained virtually unchanged. Gut elastase activity was not measurable even with elevated levels (10x) of HGPs (**Table 3.2.1**). However, these HGPs were also more sensitive towards inhibition by CaKPI; 11% and 8% increase in CaKPI mediated inhibition of total- and trypsin-like activities of HGPs was recorded in CaKPI fed HGPs in comparison to control diet fed HGPs (**Fig. 3.2.2, Table 3.2.1**). Inhibition of chymotrypsin-like activity was not recorded with CaKPI in case of both, control as well as sensitized, HGPs.

Fig. 3.2.1. Stability of CaKPI towards proteolytic degradation.
(A) Stability towards trypsin for: Lane 1, 0min; Lane 2, 30 min; Lane 3, 180 min; and chymotrypsin for: Lane 4, 0 min; Lane 5, 30 min; Lane 6, 180 min. Lane C, control (untreated).
CaKPI bands were detected after separation on 15% reducing SDS-PAGE by staining with Coomassie Brilliant Blue R-250. Band at ~21kDa is CaKPI; other bands are derived from proteinases used in the assays.
(B) Stability towards control HGPs for: Lane 1, 0 min; Lane 2, 60 min; Lane 3, 180 min; and sensitized HGPs for: Lane 4, 0 min; Lane 5, 60 min; Lane 6, 180 min. Lane C, control (untreated).
CaKPI activity bands were detected by the gel-X-ray film contact print method. A1 and A2 represent CaKPI activity bands 1 and 2.



A



B

Fig. 3.2.1

Fig. 3.2.2. Inhibitory activity of CaKPI against control and sensitized HGPs. Inhibition of total gut proteolytic activity (HGP-total) and gut trypsin activity (HGP-trypsin) in HGPs derived from larvae fed on artificial diet (shaded bars) and CaKPI-incorporated diet (white bars). Standard error bars are indicated.

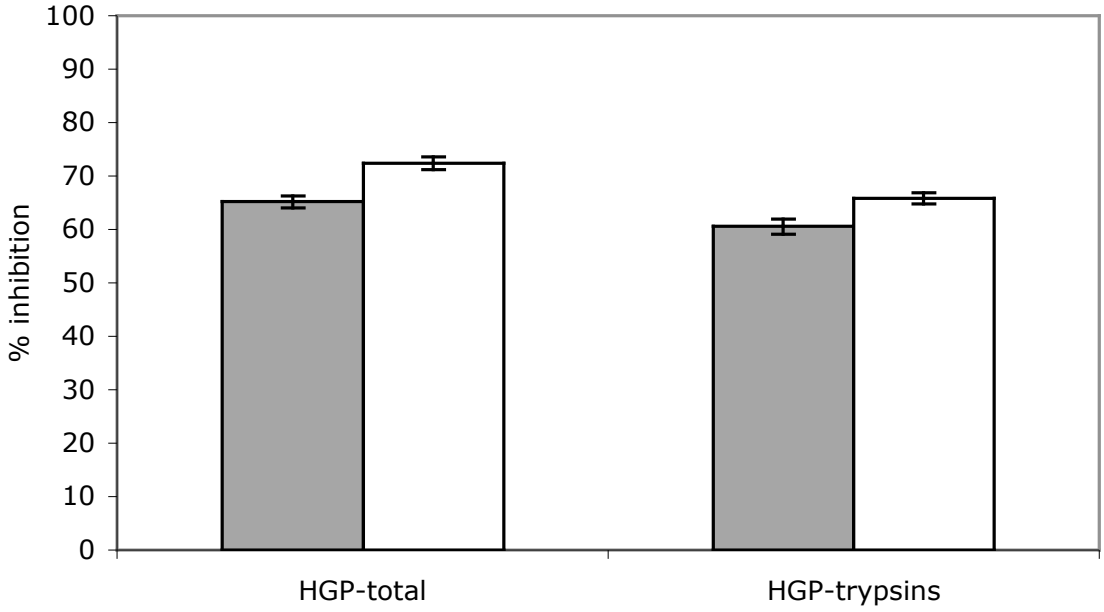

100
90
80
70
60
50
40
30
20
10
0
% inhibition
HGP-total
HGP-trypsins

Fig. 3.2.2

Table 3.2.1. Comparison of activities and CaKPI sensitivities of HGPs derived from control and sensitized larval guts. Total HGPs, HGPs-trypsin, HGPs-chymotrypsin and HGPs-elastase activities were assayed using chemical substrates. Sensitivity towards CaKPI is represented as the maximum observed inhibition of proteolytic activity by CaKPI. "N.D.", not detected.

Component	Proteolytic Activity ($UmL^{-1}min^{-1}$)			Sensitivity to CaKPI (% inhibition)		
	Control	Sensitized	Change %	Control	Sensitized	Change %
Total HGPs (AzU)	4.35 ± 0.01	4.67 ± 0.01	+7.35	65.91 ± 0.20	73.15 ± 0.15	+10.98
HGPs trypsin (TU)	10.67 ± 0.01	11.42 ± 0.01	+7.03	61.17 ± 0.09	65.79 ± 0.05	+7.55
HGPs-chymotrypsin (CU)	3.97 ± 0.01	4.0 ± 0.01	+0.76	N.D.	N.D.	-
HGPs-elastase (EU)	N.D.	N.D.	-	N.D.	N.D.	-

Differential expression of midgut protease genes

Eighteen different proteinase gene transcripts were studied, among which change in expression was observed in eight (**Table 3.2.2**). Only three (HaTry1, HaTry2 and HaTry3) of the five total trypsin transcripts were detectable, all of which showed marginal increase due to CaKPI feeding (**Fig. 3.2.3A**). Among the three isoforms of chymotrypsin, transcripts of two were detectable, of which, one (HaChy3) was over-expressed, and the other (HaChy1) was synthesized *de novo* in sensitized larvae (**Fig. 3.2.3B**). Among the five aminopeptidases, only two isoforms were detected (HaAmi2, HaAmi3), which were over-expressed in sensitized larvae (**Fig. 3.2.3C**). None of the carboxypeptidases, cathepsin or elastase transcripts were detectable in controlled or sensitized sets (Not shown).

Fig. 3.2.3. Differential expression of proteinase transcripts in response to CaKPI feeding.

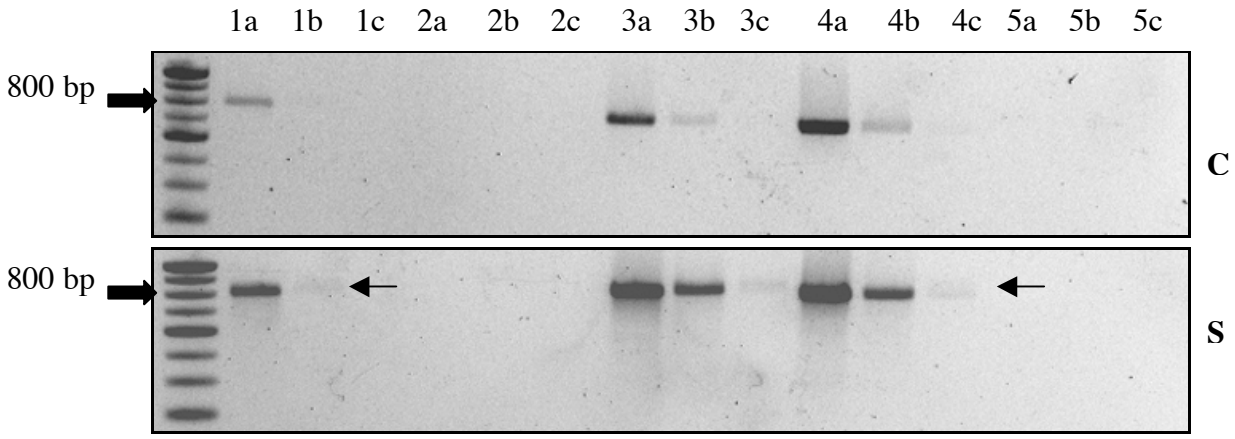
(A) trypsin isoforms HaTry1 (1), HaTry2 (2), HaTry3 (3), HaTry4 (4) and HaTry5 (5);

(B) chymotrypsin isoforms HaChy1 (6), HaChy2 (7) and HaChy3 (8);

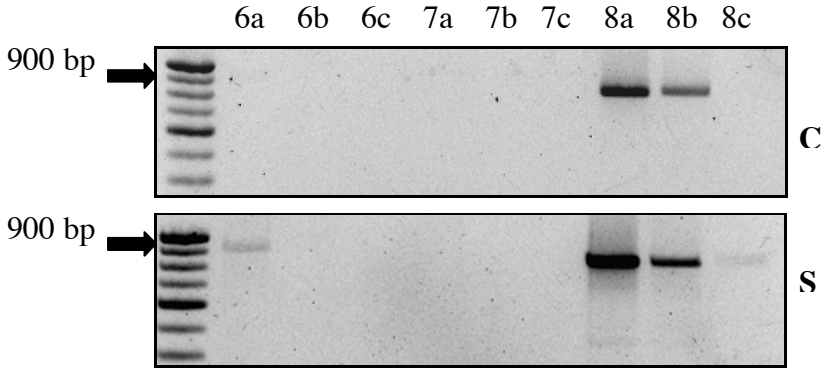
(C) aminopeptidase isoforms HaAmi1 (9), HaAmi2 (10), HaAmi3 (11), HaAmi4 (12) and HaAmi5 (13);

(D) optimized total RNA from control diet fed larvae;

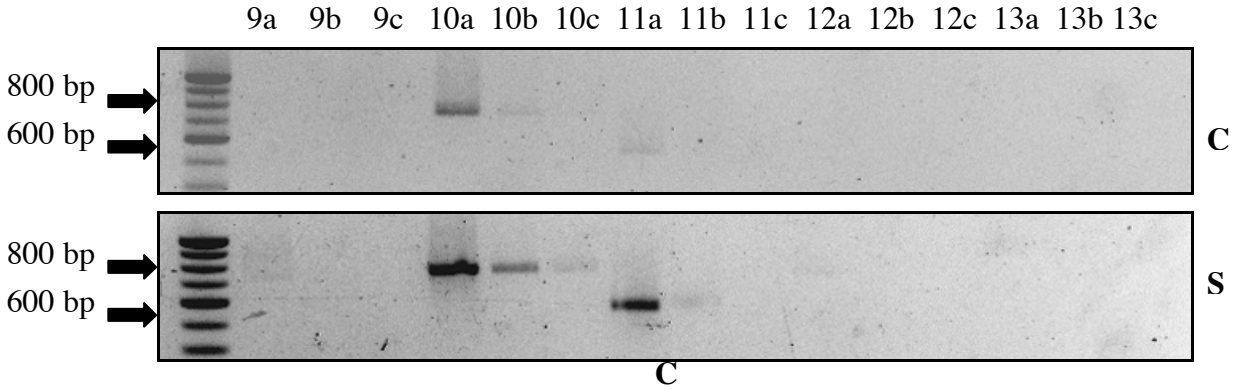
R_C, and CaKPI incorporated diet fed larval; Rs C, control larvae; S, sensitized larvae. a, 10^{-1} dilution; b, 10^{-2} dilution; c, 10^{-3} dilution respectively of template cDNA. Weakly amplified fragments are indicated by an arrow (←). Molecular weights of amplified fragments are as indicated to the left of 100bp DNA size marker lane (M).


1a 1b 1c 2a 2b 2c 3a 3b 3c 4a 4b 4c 5a 5b 5c
800 bp
800 bp
C
S

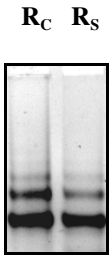
A


6a 6b 6c 7a 7b 7c 8a 8b 8c
900 bp
900 bp
C
S

B


9a 9b 9c 10a 10b 10c 11a 11b 11c 12a 12b 12c 13a 13b 13c
800 bp
600 bp
800 bp
600 bp
C
S

C


R_C R_S

D

Fig. 3.2.3

Table 3.2.2. Differential expression of midgut proteinase gene transcripts in *H. armigera*. Levels of proteinase transcripts in larvae fed on CaKPI as well as the host plants: CP, chickpea (*Cicer arietinum*); Ok, okra (*Abelmoschus esculentus*); PP, pigeonpea (*Cajanus cajan*); Tom, tomato (*Lycopersicon esculentum*); Cot, cotton (*Gossypium hirsutum*); an PIs from the non-host plants: GN, groundnut (*Arachis hypogea*); BG, bitter gourd (*Momordica charantia*); WB, winged bean (*Psophocarpus tetragonolobus*); Cap, capsicum (*Capsicum annum*). With reference to expression on control diet fed larvae: +, increase; -, decrease; NC, no change; ND, not detected (in control as well as test feeding sets).

Specificity	Isoform (Accession)	Expression of proteinase genes in *H. armigera* feeding on:									
		Host plants					Non-host PIs				
		CP	Ok	PP	Tom	Cot	GN	BG	WB	Cap	CaKPI
Trypsins	HaTry1 (Y12271)	NC	+	+	+	+	+	+	-	+	+
	HaTry2 (Y12269)	+	+	+	+	+	+	+	-	-	ND
	HaTry3 (Y12277)	NC	+	-	+	+	+	+	+	NC	+
	HaTry4 (Y12276)	NC	NC	NC	NC	NC	NC	NC	-	NC	+
	HaTry5 (Y12270)	-	-	NC	-	-	NC	-	-	-	ND
Chymotrypsins	HaChy1 (Y12287)	-	-	NC	-	-	NC	NC	-	-	+
	HaChy2 (Y12281)	-	NC	-	+	-	NC	+	-	NC	ND
	HaChy3 (Y12273)	NC	NC	NC	NC	NC	NC	NC	NC	NC	+
Aminopeptidases	HaAmi1 (AF535165)	NC	+	NC	+	+	+	+	NC	-	ND
	HaAmi2 (AY038608)	-	+	+	+	+	+	+	NC	-	+
	HaAmi3 (AF441377)	NC	+	+	+	+	+	+	NC	-	+
	HaAmi4 (AF535166)	-	+	NC	+	NC	+	+	-	-	ND
	HaAmi5 (AY052651)	-	+	-	+	+	+	+	-	-	ND
Carboxypeptidases	HaCar1 (AJ005177)	-	-	NC	NC	-	-	-	-	NC	ND
	HaCar2 (AJ005178)	ND	ND	ND	+	+	ND	ND	ND	ND	ND
	HaCar3 (AJ005176)	ND	ND	ND	+	ND	+	+	ND	ND	ND
Elastase	HaEla1 (AY222788)	-	-	-	NC	-	NC	NC	-	-	ND
Cathepsin B like	HaCat1 (PHelasB2)	-	+	NC	NC	+	+	-	-	-	ND

Discussion

Insect pests respond and very often adapt to ingestion of plant PIs by altering the complement of digestive gut proteinases. Faced with the prospect of reduced nutrient uptake and following consequences, quantitative and/or qualitative changes in gut proteinases are observed (Jongsma and Bolter, 1997; Paulillo *et al*, 2000; Brito *et al*, 2001; Bown *et al*, 2004b). Quantitative changes include an increase in synthesis of all or specific gut proteases to attain optimal levels of active proteases and rate of protein digestion (Broadway and Duffey, 1986; Broadway, 1997; Gatehouse *et al*, 1997; Girard *et al*, 1998b). On the other hand, qualitative responses include synthesis of "insensitive" protease isoforms to which the PI is unable to bind and inhibit (Jongsma *et al*, 1995; Broadway, 1996; Bown *et al*, 1997; Mazumdar-Leighton and Broadway, 2001a; b; Volpicella *et al*, 2003), or which have the ability to degrade the PI (Giri *et al*, 1998; Girard *et al*, 1998a; Zhu-Salzman *et al*, 2003; Moon *et al*, 2004; Telang *et al*, 2005). Thus, exposure to plant PIs could potentially result in insect adaptation by any or all of the available choices.

Since *H. armigera* is continuously exposed to basal levels of chickpea seed PIs in field or via artificial diet formulations, some PI degrading capability would be expected and such has been observed in case of the HGPs mediated degradation of chickpea seed BBIs (Giri *et al*, 1998). However, CaKPI was observed to be stable against degradation by artificial diet fed HGPs for up to three hours, which is the normal food retention time in the larval gut. Evidently, the insects do not respond to the lower levels of CaKPI by synthesis of inhibitor degrading proteinases. Though higher concentration of CaKPI caused detrimental effects on growth and development and thus warranted adaptive responses in larvae, newer isoforms, if any, in the sensitized HGPs also lacked any ability to degrade CaKPI. It does appear that continuous exposure to the low levels of CaKPI has rendered *H. armigera* passive towards not only the basal levels CaKPI, but also to significantly higher amounts. It is conclusive that *H. armigera* lacks constitutively and/or differentially expressed proteinases, which could degrade CaKPI.

Reports describe adaptation in Lepidopteran larvae by over-expression of trypsins (Broadway and Duffey, 1986) or chymotrypsins (Gatehouse *et al*, 1997). Sensitized HGPs exhibited only a marginal increase in the total proteolytic and trypsin-like activities over control HGPs, but even these changes would not amount to any

adaptive responses that would compensate for loss of HGPs activity due to CaKPI mediated inhibition. Secondly, the concurrent increase in sensitivity to CaKPI reinforces lack of a true physiological adaptive response. Increase in activity of chymotrypsins may be advantageous not only in being uninhibited by the TI, but may also cause proteolytic inactivation of the TI. But the virtual absence of any change in HGPs-chymotrypsin activity denies *H. armigera* larvae this chance to overcome the effect of CaKPI. The unchanged HGPs-chymotrypsin activity as well as their insensitivity to CaKPI indicate that these proteinases have very little or no relation with CaKPI feeding or insect adaptation. At this point, a corollary between the seemingly "atypical" properties of CaKPI, its inhibitory profile and insect adaptation can be drawn. The absence of a classical trypsin or chymotrypsin recognition sequence at the putative active site region seems to be one reason for the inability of trypsin, chymotrypsin or HGP to degrade native CaKPI, but it is still not clear as to how the modified active site sequence contributes to binding and inhibition of trypsin and trypsin like proteinases in HGP. This unusual sequence variation and atypical inhibitory activity could contribute to the inability of *H. armigera* larvae to adapt to CaKPI.

Gut proteinases can be classified as sensitive or insensitive to host/non-host plant PIs, on the basis of change in levels of their transcripts in response to PI feeding (Chougule *et al*, 2005). These sensitivities have also been linked to specific amino acid variation(s) in the active site and neighboring regions within the proteinase (Lopes *et al*, 2004). While quantitative increase of proteinase gene transcripts due to CaKPI feeding was not very high, CaKPI did cause up-regulation in a wider spectrum of proteinases (**Table 3.2.2**) in comparison to host plant (chickpea) as well as non-host plant (winged bean, *Psophocarpus tetragonolobus*) PI fed larvae. While differences were expected between winged bean PI (WBPI) and CaKPI fed larvae, it was interesting to note the differences between chickpea plant fed and CaKPI fed larvae. The simultaneous up-regulation of trypsin as well as chymotrypsin isoforms is indicative of a failed two-pronged strategy to overcome detrimental effect of CaKPI; the up-regulation of trypsins aimed to fill the void in gut trypsin activity due to CaKPI inhibition whereas, chymotrypsin up-regulation aimed to make up for loss of trypsin activity as well as try to degrade CaKPI. The significance of aminopeptidase up-regulation is not very clear at present, but presents an interesting option of exo-acting

proteinases to counter the depletion of endo-acting proteinases. It is conclusive that, in spite of being host-plant derived, CaKPI does fulfill the criteria for an effective candidate in crop protection programmes, *viz.*, (i) it causes severe antagonistic effect on developing larvae, and more importantly, (ii) the larvae do not adapt to this inhibitor.

Host plant PIs have been largely ignored due to (i) absence of convincing action against herbivorous insect pests, and, (ii) their degradation by the insect gut proteinases. However, the current section has described how one such low expressing PI from chickpea could prove effective at higher levels than what the insect pest (*H. armigera*) is exposed to in chickpea plants and/or chickpea seed meal based diet (Nagarkatti and Prakash, 1974). This opens up a window of opportunity for study of other PIs from susceptible host plants and evaluate their potential against infestation by herbivorous insects.

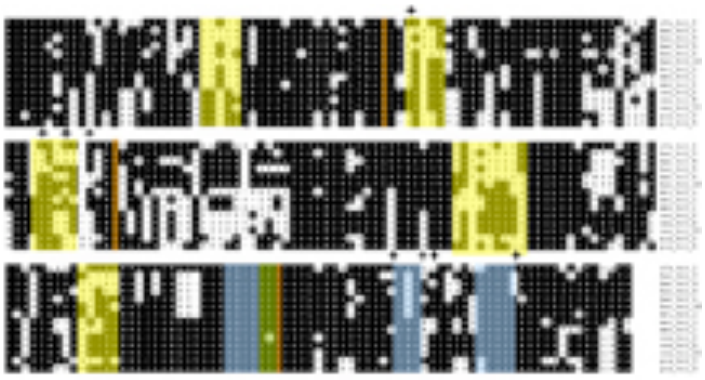
Conclusion

PI degrading/insensitive gut proteinases were found to be absent as well as uninducible on CaKPI feeding. *H. armigera* exhibited a generalized and non-specific upregulation of gut proteinase activity in response to CaKPI feeding, which were insufficient to mount a strong adaptive response. Thus, *H. armigera* larvae do not seem to possess a convincing mechanism that would help them overcome the inhibitory effects of CaKPI.

CHAPTER IV:

Significance of Structural and Functional Diversities in Proteinase-Proteinase Inhibitor and Plant-Pest Interactions: a theoretical study

This chapter has been communicated as a full-length review to Cellular and Molecular Biology Letters (Srinivasan et al, 2005c)



Structural characteristics of Kunitz type PIs

Kunitz type serine PIs are single-chain polypeptides of about 200 amino acid residues having a roughly spherical molecular structure ~40A° across (Onesti *et al*, 1991; Dattagupta *et al*, 1996). These PIs consist of 12 antiparallel β-strands, long loops connecting these β-strands and a 3_{10}-helix. Six of the strands form a short antiparallel β-barrel, with one side of the barrel being closed by a 'lid' consisting of the other six strands ('β-trefoil fold') (McLachlan *et al* 1979; Murzin *et al*, 1992). Kunitz type PIs display a three-fold internal symmetry, with the symmetry axis coinciding with the barrel axis. The repeating unit (or sub-domain) is a 4-stranded motif of ~60 amino acids, structurally organized as L-β_1-L-β_2-L-β_3-L-β_4, where L denotes the loop connecting consecutive β-strands. The 3 sub-domains usually exhibit similarity for the β-strands but not for the connecting loops (Song *et al*, 1998). The first ~20 amino acids form the leader or signal peptide region, which is usually cleaved to form the mature peptide; a putative signal peptide region of 21/23 amino acids was identified in the *Cicer arietinum* Kunitz type proteinase inhibitor (CaKPI) by SignalP (http://www.cbs.dtu.dk/services/SignalP/) (Bendsten *et al*, 2004). Kunitz type PIs are also identified by a seventeen amino acid motif, [LIVM]xDx[EDNTY][DG][RKHDENQ]x[LIVM](x)$_5$Yx[LIVM], where 'x' can be any residue; in CaKPI, ScanProsite (http://www.expasy.org/tools/scanprosite/) (Gattiker *et al*, 2003) identified the motif as 'VLDINGNPIFPGGKYYI'. Four well-conserved cysteines are usually involved in formation of the two intra-chain disulfide bridges, which contribute to structural stability and function. However, recent reports have also described Kunitz type PIs having one (DoSocorro *et al*, 2002) or no disulfide linkages (Araujo *et al*, 2005); in these polypeptides, other main- and side-chain interactions are presumed to stabilize the three-dimensional structure. The reactive site peptide bond (P1-P1') (Schechter and Berger, 1967), which behaves as a substrate for the cognate proteinase, is located on an exposed and convex loop that forms a simple recognition motif for the target proteinase (Qasim *et al*, 1995). The overall length of this loop is often defined as a hexa-peptide (P3-P3') (Apostoluk and Otlewski, 1998), although the determinant of principal proteinase-specificity is the P1 side-chain. PIs with trypsin specificity usually exhibit Arg or Lys at P1, whereas Phe is preferred over other hydrophobic residues for chymotrypsin specificity (Huber and

Fig. 4.1 Predicted structure of CaKPI.
The CaKPI polypeptide sequence was theoretically modeled using the Swiss-Model comparative protein modeling server (http://swissmodel.expasy.org) based on available structural information of homologous Kunitz type PIs from the Protein Data Bank (http://www.rcsb.org/pdb). Templates used for comparative modeling were the *Delonix regia* Kunitz type trypsin inhibitor (DrKTI, 1r8nA) and the soybean Kunitz type trypsin inhibitor (SKTI, 1avu). Expected inhibitory loop region is indicated.
Legend: CaKPI (red 3D backbone): spheres represent are Gly68, green; Ile69, red; Ser70, orange. Asn14 that contributes to stability of the canonical loop conformation is also indicated (blue, ball and stick). The backbone of the amino acid residues that constitute the Kunitz family signature sequence are highlighted in dark green. DrKTI, (blue discontinuous backbone): wireframe amino acids in blue are Arg63 and Ile64. SKTI (yellow tubular backbone): wireframe amino acids in light blue are Lys69 and Gln70. Structure was visualized with iMol (http://www.pirx.com/iMol/).

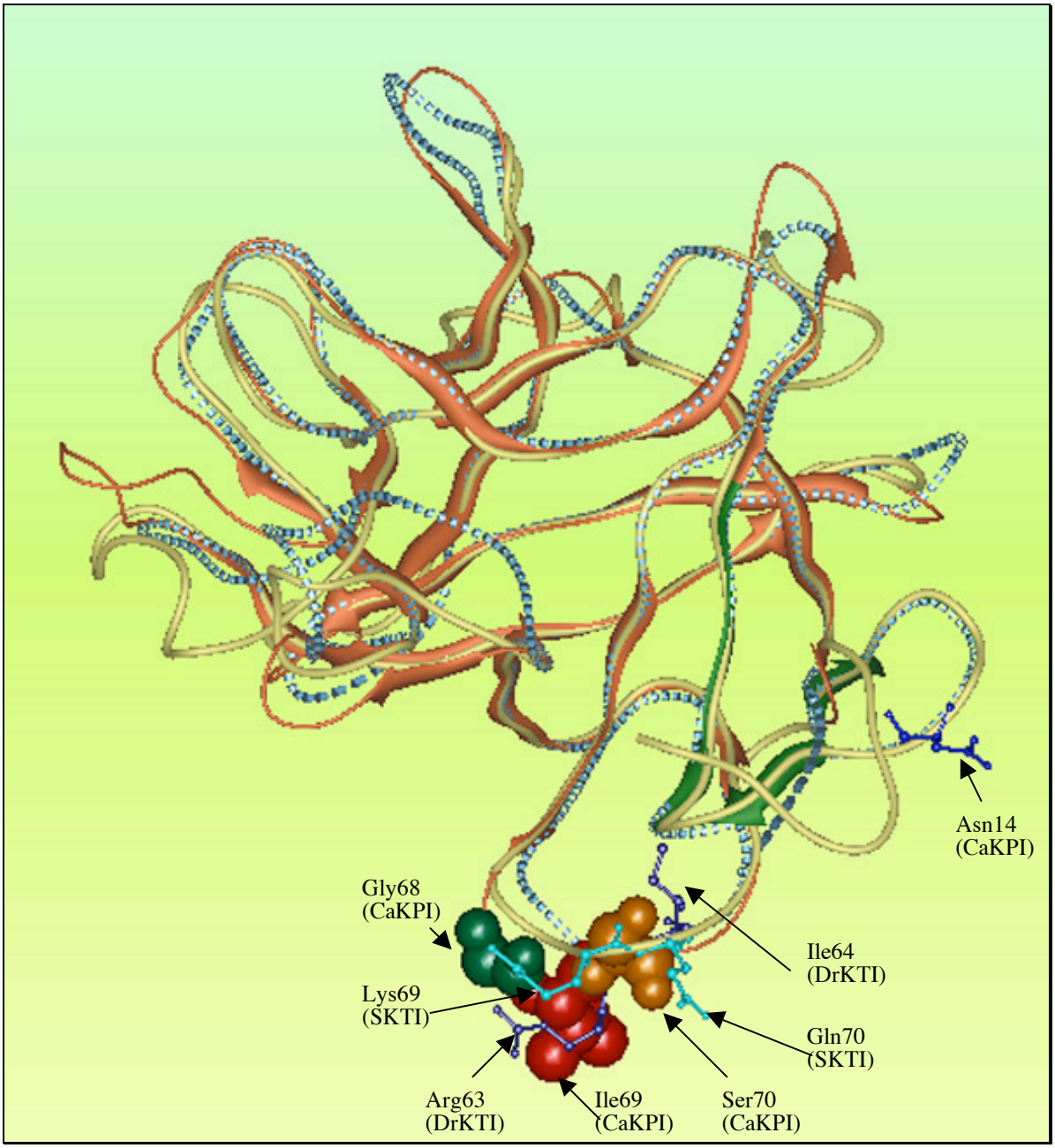

Asn14
(CaKPI)
Gly68
(CaKPI)
Ile64
(DrKTI)
Lys69
(SKTI)
Gln70
(SKTI)
Arg63
(DrKTI)
Ile69
(CaKPI)
Ser70
(CaKPI)

Fig. 4.1

Bode, 178). In case of CaKPI, a Gly-Ile-Ser motif was observed at the putative inhibitory site, due to which the exact mechanism of the observed trypsin inhibition was unclear (discussed subsequently); it was concluded that either the actual scissile bond lies elsewhere on the molecule or the Gly-Ile-Ser motif exhibits some novel mechanism of trypsin binding/inhibition. The amino acid side-chains in the inhibitory loop are similarly oriented in the STI family; the residues between P3-P3', which exhibit similar Ramachandran angles (~ 3_{10} helix-like for P1, & ~ β strand-like for others), are known to superpose well irrespective of the global structure (Laskowski *et al*, 2000). In the "lock and key" association mechanism of proteinase and PI, the Ramachandran angles of the combining loops of PI remain unchanged upon complex formation and hence these PIs are termed "canonical" (Bode and Huber, 1978). The canonical conformation is a conserved structural motif responsible for tight binding with serine proteinases but cannot be considered as a novel secondary structure element (Apostoluk and Otlewski, 1998). The conformational stability of the inhibitory loop depends on the local H-bonding interactions; the H-bonding interactions made by side chain of Asn (13/14) were found to be conserved and also more pronounced than those made by its main chain for stabilization and maintenance of the canonical loop conformation (Ravichandran *et al*, 2001; Dasgupta *et al*, 2003; Iwanaga *et al*, 2005). Although variations at P1 is well accepted to influence the primary-proteinase specificity, not much is known about the contribution(s) of P1' variants (Grzesiak *et al*, 2000). Although different amino acids may be conserved at P1' in Kunitz type serine PIs belonging to different families, it does not appear to affect binding to their cognate enzymes. P1' Ile is known to be conserved in some members of the STI family (Grzesiak *et al*, 2000) – in CaKPI, Ile is found in the vicinity of the putative P1' region, once again raising an interesting question about the probable contribution of the Gly and Ser residues towards trypsin binding and inhibition. Using known three-dimensional structures from previously resolved crystallographic data of the soybean Kunitz type TI (SKTI, 1avu) as well as the *Delonix regia* Kunitz type TI (DrKTI, 1r8nA) from the Protein Data Bank (http://www.rcsb.org/pdb) as templates, the CaKPI polypeptide sequence was theoretically modeled on the Swiss-Model comparative protein modeling server (http://swissmodel.expasy.org) and is represented in **Fig. 4.1**. As this structure is predicted based on sequence similarity and the assumption of conserved local

structural aspects (Ramachandran angles) across stretches of similar sequences, it would have to be validated by X-ray crystallographic studies on CaKPI and it's binding to target proteinase, *viz*., trypsin. However, it does provide us with an interesting template to analyze the possible orientation of the Gly-Ile-Ser motif at the putative active site (inhibitory loop) region, with the assumption that this region does assume the canonical conformation essential for trypsin binding and inhibition. At the same time, it would be necessary to investigate other regions of the CaKPI polypeptide that could permit binding to either the active site of trypsin or at other topological locations on trypsin that could eventually lead to inhibition of trypsin activity as observed with CaKPI.

Structure and activity of serine proteinases

The structures and activities of various serine proteinases have for long been the subject of intensive studies. As a result of these studies, our current understanding has helped unravel the finer intricacies of the properties of these proteinases. Serine proteinases like trypsin or chymotrypsin are polypeptides of about 200 amino acid residues, with a signal region that is cleaved off during conversion of the zymogen form into the active proteinase. In trypsins, post-translational cleavage by proteinases further results in three subunits held together by disulfide linkages. The structural features of a serine proteinase are represented in **Fig. 4.2** based on the reported X-ray crystallographic data for porcine trypsin (1S81) in the Protein Data Bank. Serine proteinases consist of 2 perpendicular β-barrel domains, each formed by six antiparallel β-strands, and the C-terminal α-helix (Czapinska and Otlewski, 1999). The catalytic and substrate-binding sites lie in the cleft between the β-barrels and the enzyme-substrate interactions bridge both the domains; functional residues are positioned mostly in the loops connecting the β-strands (Czapinska and Otlewski, 1999). The S1 pocket, built of three β-strands (usually 189-192, 214-216, 226-228), and the oxyanion-binding site (Gly193-Ser195) belong to the C-terminal β-barrel (Czapinska and Otlewski, 1999). Gly193, that is a part of the oxyanion-binding site has also been shown to be critical for structure as well as activity (Schmidt *et al*, 2004). Three disulfide bonds are usually conserved: 42-58, 168-182 and 191-220; the last one cross-links the S1 pocket in majority of serine proteinases and seems to

Fig. 4.2 Structural aspects of porcine trypsin.
Key features of porcine trypsin (1S81) are highlighted in this image. The catalytic triad is formed by local association of His57 (red), Asp102 (yellow) and Ser195 (blue). The amino acid residues Gly193 and Asp194 (green) that reside just before the catalytic Ser195 constitute the oxyanion binding site. The S1 pocket is a molecular cavity formed between three separate amino acid stretches, *viz.*, Ser190-Gln192 (white), Ser214-Gly216 (grey) and Gly226-Tyr228 (lilac). Asp189 (brown) lies at the distal end of the S1 pocket and contributes to association with the amino side chains of P1 Lys/Arg in substrate and trypsin inhibitor molecules. Chymotrypsins exhibit Asp189Ser substitution, which leads to altered specificity towards hydrophobic side chains (*e.g.*, Phe) at P1 in substrate. Structure was visualized with iMol (http://www.pirx.com/iMol/).

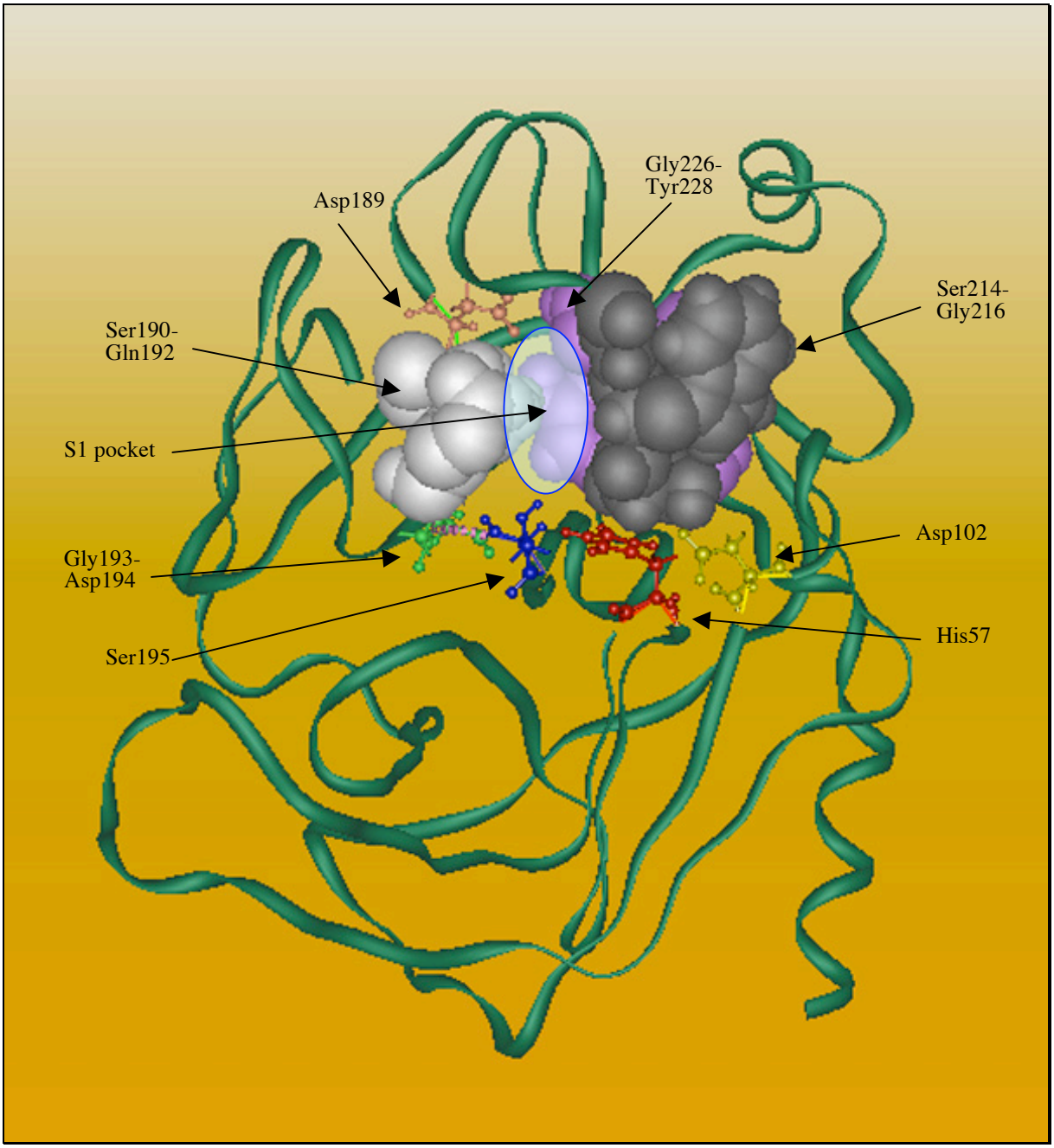

Asp189
Gly226-
Tyr228
Ser214-
Gly216
Ser190-
Gln192
S1 pocket
Asp102
Gly193-
Asp194
His57
Ser195

Fig. 4.2

influence catalytic efficiency (Varallyay *et al*, 1997; Czapinska and Otlewski, 1999). Though S1 sites of trypsins and chymotrypsins are similar, differences occur at positions 189, 192 and 138, which are Ser, Met and Thr in chymotrypsin, whereas trypsins exhibit Asp, Gln and Ile at these positions as well as a Ser218 deletion (Czapinska and Otlewski, 1999). Trypsins prefer basic side-chains at P1 (Arg>Lys) due to the conserved Asp189, at the base of the S1 pocket (Huber and Bode, 1978). Chymotrypsins, on the other hand, which exhibit Asp189Ser substitution with respect to trypsin, show a clear S1 preference for hydrophobic residues (Phe>others) at P1 (Huber and Bode, 1978) and a gradual increase in catalytic efficiency linked to the P1 side-chain volume (Schellenberger *et al*, 1991). The oxyanion hole at the entrance to the S1 site is important for binding of substrate and stabilization of the intermediate form. Gly216 located at the rim of the S1 pocket affects scissile bond positioning through an antiparallel β-sheet between the P1-P3 and S1-S3 sites; the absence of a side chain at Gly216 allows efficient access of the P1 side chain to the interior of the S1 pocket. In fact, replacement of Gly216 with other residues has been shown to affect S1 sub-site accessibility and, hence, proteinase activity (Hedstrom *et al*, 1994). In elastases, the isopropyl side chain of Val216 partially fills the S1 pocket, thereby limiting the specificity to substrates with small aliphatic residues at P1 (Sinha *et al*, 1987; Bode *et al*, 1988). The enzyme-substrate/inhibitor interfaces in trypsin and chymotrypsin complexes are similar, barring the S4'-P4' H bond in trypsin complexes (Grzesiak *et al*, 2000). The hydrolytic mechanism of serine proteinases is a function of the "catalytic triad" which results due to the spatial proximity of His57, Asp102 and Ser195 residues in the active conformation of the polypeptide (Matthews *et al*, 1967). This triad is a well-conserved feature of serine proteinases (Dodson and Wlodawer, 1998) and its mechanism has been well elucidated (Hunkapiller *et al*, 1973; Blow *et al*, 1974; Kraut, 1977; Baillargeon *et al*, 1980; Hedstrom, 2000; Ishida and Kato, 2004). The reaction, can be summed up as: (i) binding of substrate to the active site and formation of the P1 Arg-Asp189 salt bridge, (ii) nucleophilic attack of the Ser195 side chain oxygen upon the carbonyl carbon atom of the P1 residue leading to formation of the first oxyanion intermediate, (iii) formation of a quarternary amine at the P1 amide due to proton donation by the (acidic) His57, (iv) collapse of the tetrahedral oxyanion structure, cleavage of the P1-P1' bond and release

Fig. 4.3. Catalytic mechanism of trypsin.
Trypsin mediated substrate hydrolysis (proteolysis) can be defined in the following six steps.
(1) In the initial step, substrate binds to the active site. P1 arginine forms a salt bridge with Asp189 at the bottom of the S1 pocket.
(2) The oxygen of the serine side chain participates in a nucleophilic attack upon the carbonyl carbon atom of the P1 residue. This produces a covalent enzyme-substrate complex known as the 'oxyanion intermediate'.
(3) The protonated histidine at the active site acts as a general acid and donates its proton to the P1' amino group. A quaternary amine thus forms at the P1' amide.
(4) The quaternary amine and tetrahedral oxyanion structure collapses and results in cleavage of the P1-P1' bond and release of the P1' fragment of the substrate peptide. The P1 peptide fragment, however, remains as an acylenzyme intermediate.
(5) Water enters the active site and is deprotonated by histidine (thus acting as a base, in this step). The water concomitantly performs a nucleophilic attack on the carbon atom of the acylenzyme, thus forming another oxyanion intermediate.
(6) Finally, the tetrahedral oxyanion structure collapses, and the P1 terminus of the substrate peptide is released. Inhibitors of serine proteases act by forming a stable acyl-enzyme intermediate thus preventing dissociation of the cleaved fragments.

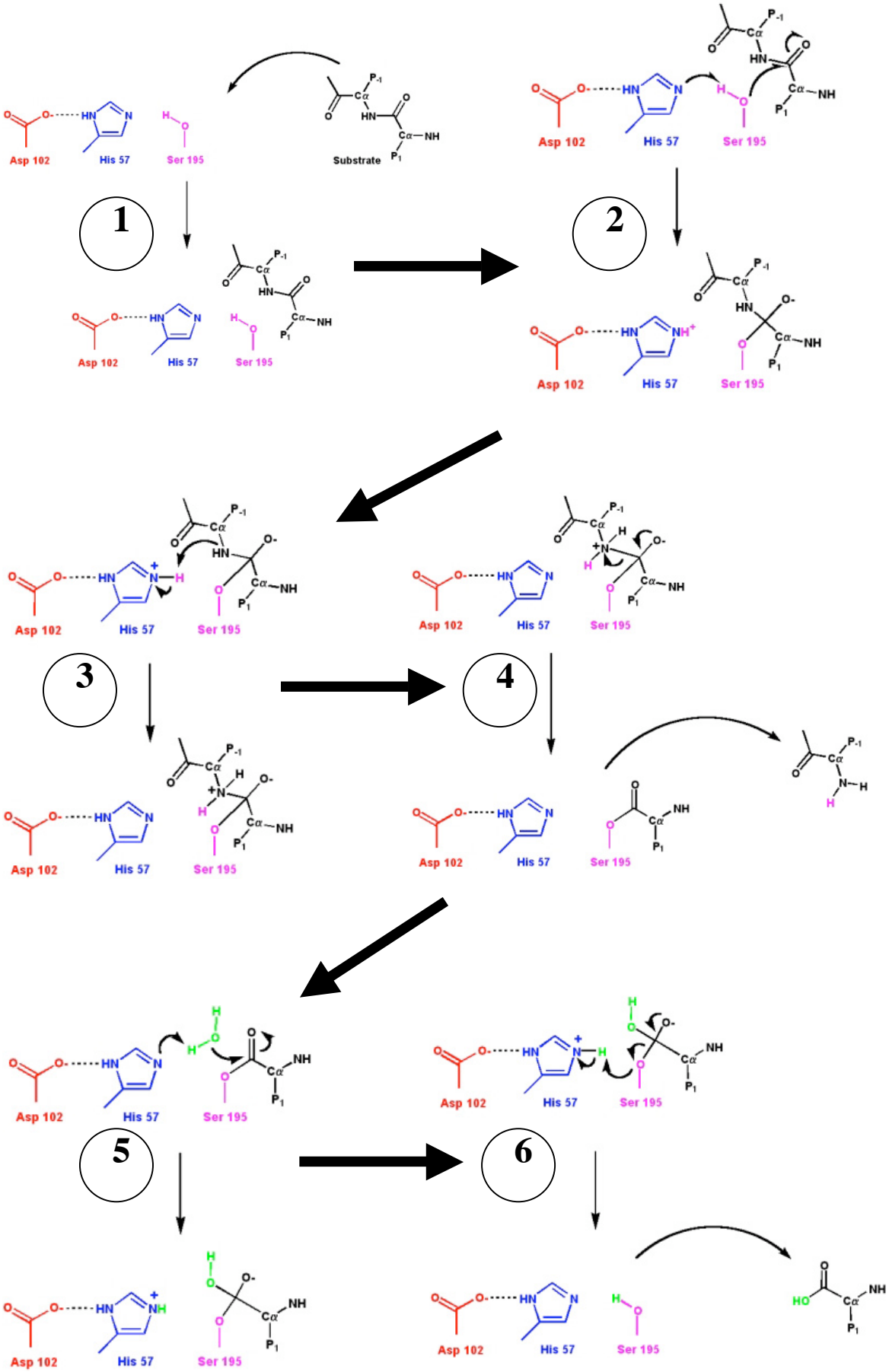


Fig. 4.3

of the P1' fragment, (v) His57 (basic) mediated deprotonation water, a concomitant nucleophilic attack on the carbon atom of the acylenzyme, formation of the second oxyanion intermediate, and, (vi) collapse of the tetrahedral oxyanion structure, followed by release of the P1 terminus of the substrate peptide. This mechanism of action of the catalytic triad of trypsin is represented in **Fig. 4.3**.

Binding and inhibition of serine proteinases by Kunitz type proteinase inhibitors

The mode of proteinase-PI association resembles that of an ideal substrate binding (Luthy *et al*, 1973; Laskowski, 1986; Otlewski *et al*, 2001): an anti-parallel β-sheet is formed between the P3-P1 residues and the 214–216 segment of the enzyme (Apostoluk and Otlewski, 1998). However, binding of PI to the target proteinase causes a destabilization or deformation (Plotnick *et al*, 1996, Huntington *et al*, 2000) of the structure leading to loss of activity. Majority of the proteinase-PI contacts (~50% of intermolecular van der Waals contacts and H-bonds) are made by the P1 residue, which penetrates deeply into the S1 pocket and becomes fully buried upon proteinase-PI complex formation (Krowarsch *et al*, 2005). The amino acids besides this segment and those from discontiguous (secondary contact) regions can also influence the association energy (Apostoluk and Otlewski, 1998). Many reports have supported the extended binding interactions between proteinase and PI, *i.e.*, in which secondary contact is involved and have demonstrated their importance to inhibitor binding (Imperiali *et al*, 1987; Capasso *et al*, 1997; Katz *et al*, 2003; Fodor *et al*, 2005). In trypsin-TI interaction, binding of P1 Arg at S1 differs from P1 Lys; the guanidinium group of P1 Arg forms a direct and tighter interaction with Asp189 resulting from a cyclic network of H bonds (Czapinska and Otlewski, 1999), unlike P1 Lys, which is water-bridged to Asp189 (Sichler *et al*, 2002). Ser190 positioned in the S1 pocket, allows an additional and stabilizing H bond to the already bound Arg/Lys (Evnin *et al*, 1990; Stubbs *et al*, 1995). Association energy for basic side chains (Arg/Lys) with trypsin thus results from (i) the ion-pair between Asp189 and charge of Arg/Lys, and, (ii) the burial of the hydrophobic part of the basic residue (Krowarsch *et al*, 2005). The influence of cognate (Lu *et al*, 1997) as well as non-cognate (Helland *et al*, 2003) amino acid residues and their effects on binding have also been well investigated. Polar side chains (Ser, Asn, His, Gln) at P1 also bind probably due to weak, water-mediated H bonding interactions with Asp189, whereas

β-branched side chains (Val, Leu, Ile) bind weakly due to steric hindrances at the narrow entrance to the S1 pocket (Krowarsch *et al*, 2005). It is hence assumed, in case of the putative inhibitory loop site in CaKPI, that Ser may contribute to trypsin binding by a virtue of the polar side chain. However the implications of the β-branched Ile, when adjacent to the polar and putative bond-forming Ser, is enigmatic. In chymotrypsins, hydrophobic effects drive the association reaction within the S1 pocket; optimal binding of hydrophobic side chains of cognate P1 residues (Phe, Leu, Trp) is reflected in the consequent shrinkage of the S1 pocket (Krowarsch *et al*, 2005). Since β-branched side chains (Val, Leu, Ile) at P1 are known to hinder binding, the Ile in the putative loop region of CaKPI could prohibit binding to chymotrypsin and explain the absence of chymotrypsin inhibitory activity in CaKPI. It has been proposed that basic side chains (Arg, Lys) at P1 may also bring about binding of PI to chymotrypsin due to interaction with carbonyl groups at the entrance of the S1 pocket (Krowarsch *et al*, 2005); since Arg/Lys is absent at the putative inhibitory loop of CaKPI, this option for association to chymotrypsin is also ruled out, concurring with the experimental observations. Though obvious differences exist, substrate/PI-binding interactions within the S1 pockets of trypsin and chymotrypsin are thought to follow a similar approach. While the interaction of P1 with the proteinase is energetically most important, the main- and side-chain interactions of secondary contact regions serve to (i) maintain the position of native P1-P1' bond for recognition by the cognate proteinase (Ardelt and Laskowski, 1985), and, (ii) ensure that substitutions at P1 do not shift the P1-P1' bond and the altered P1 side chain is still imbedded in the S1 cavity (Huang *et al*, 1995).

Due to the striking similarities in the association of proteinases with substrates as well as PIs, the latter can be considered as substrates that, at the optimal pH and temperature, undergo very slow hydrolysis due to the atypical partitioning of high K_{cat}/K_m values into very low K_{cat} and K_m values (Laskowski and Qasim, 2000). In proteinase-PI interaction a single peptide bond (P1-P1') is not only hydrolyzed in the PI (Laskowski and Kato, 1980) but also re-synthesized upon complex formation (Ardelt and Laskowski, 1985). Following proteinase-PI association, hydrolysis of the acyl-enzyme requires partial dissociation of the leaving group peptide to allow access by the hydrolytic water and hence, PIs with stronger binding affinity in the intact form may be more resistant to dissociation if, and when, cleavage occurs at the scissile

bond (Radisky *et al*, 2005). CaKPI was observed to be stable towards degradation by trypsin, chymotrypsin and the (largely trypsin-like) HGPs. Since binding of CaKPI to trypsin is then accepted as a prerequisite for inhibition, it could be interpreted that the Gly-Ile-Ser motif may be responsible for higher affinity for trypsin binding and stoichiometric inhibition, but the mechanism may not involve scissile bond hydrolysis (or the hydrolysis may proceed extremely slowly under experimental conditions). As mentioned previously, CaKPI may also exhibit binding at a different site on trypsin; this binding may not involve positioning of the scissile bond in the vicinity of the catalytic triad, and may be an alternate possibility for the absence of P1-P1' bond cleavage, or in other words, stability towards proteolytic degradation by trypsin or HGPs. Stability of CaKPI to chymotrypsin may result from the lack of recognition and binding interactions, at the active site and/or other topological regions of chymotrypsin. In this context, a secondary binding site has been recently reported in trypsin (Shamladevi *et al*, 2005). It would be interesting to identify and probe such alternate binding site(s) on the proteinase that could correlate with the atypical properties of CaKPI.

Structural diversities and functional specificities of Lepidopteran serine proteinases

Since the specificity of the serine proteinases is intricately linked to the S1 pocket residues, it may be extrapolated that the S1 pocket may be linked to functional diversity. In **Fig. 4.4 (A, B)**, key changes have been illustrated across representative Lepidopteran trypsins and chymotrypsins. It is clearly seen that the amino acids that form the catalytic triad (orange) are highly conserved in trypsins and chymotrypsins, as are those of the oxyanion hole (green); an exception is the trypsin from *Ostrinia nubilalis* (onu_try_6), which exhibits Asp194Gly substitution. The amino acids in the S1 pocket (blue) are well conserved in trypsins, whereas chymotrypsins show a much greater degree of diversity, and would be expected to exhibit greater flexibility in substrate recognition and/or activity as compared to trypsins. Perhaps this feature may also be linked to the relative populations of serine proteinase isoforms in Lepidopterans like *H. armigera*, where the sheer diversity of trypsins (Patankar *et al*, 2001) probably compensates for the limited flexibility at the S1 pocket, whereas the higher flexibility expected in chymotrypsins (Solivan *et al*, 2002) could compensate

Fig. 4.4. Molecular diversity in Lepidopteran serine proteases.
Sequences of representative Lepidopteran trypsins and chymotrypsins were aligned by the Clustal algorithm to illustrate the occurrence of amino acid variations in the catalytic triad (orange), oxyanion hole (green) and S1 pocket (blue), as well as those regions putatively involved in formation of adventitious contacts (yellow) and the 'hot-spots' which govern sensitivity to PIs (marked by '+').
(A) Trypsin sequences are from *Choristoneura fumiferana* (cfu_try_2, AAA81525); *Helicoverpa armigera* (har_try_3, CAA72956; har_try_4, CAA72955; har_try_5, CAA72954; har_try_6, CAA72949; har_try_10, CAA72962); *Heliothis virescens* (hvi_try_1, AAF43708), *Manduca sexta* (mse_try_1, P35047; mse_try_2, P35046); *Ostrinia nubilalis* (onu_try_6, AAX62035; onu_try_9, AAX62032; onu_try_22, AAR98918,); *Plodia interpunctella* (pin_try_3, AAF24226; pin_try_5, AAC36248).
(B) Chymotrypsin sequences are from *H. armigera* (har_chy_4, CAA72960; har_chy_5, CAA72959; har_chy_6, CAA72958; har_chy_7, CAA72952); *H. virescens* (hvi_chy_1, AAF43709); *M. sexta* (mse_chy_1, AAA58743); *O. nubilalis* (onu_chy_3, AAX62030; onu_chy_13, AAX62027; onu_chy_14, AAX62026); *P. interpunctella* (pin_chy_1, AAC36149); *Spodoptera frugiperda* (sfr_chy_1, AA075039; sfr_chy_3, AAC36150).

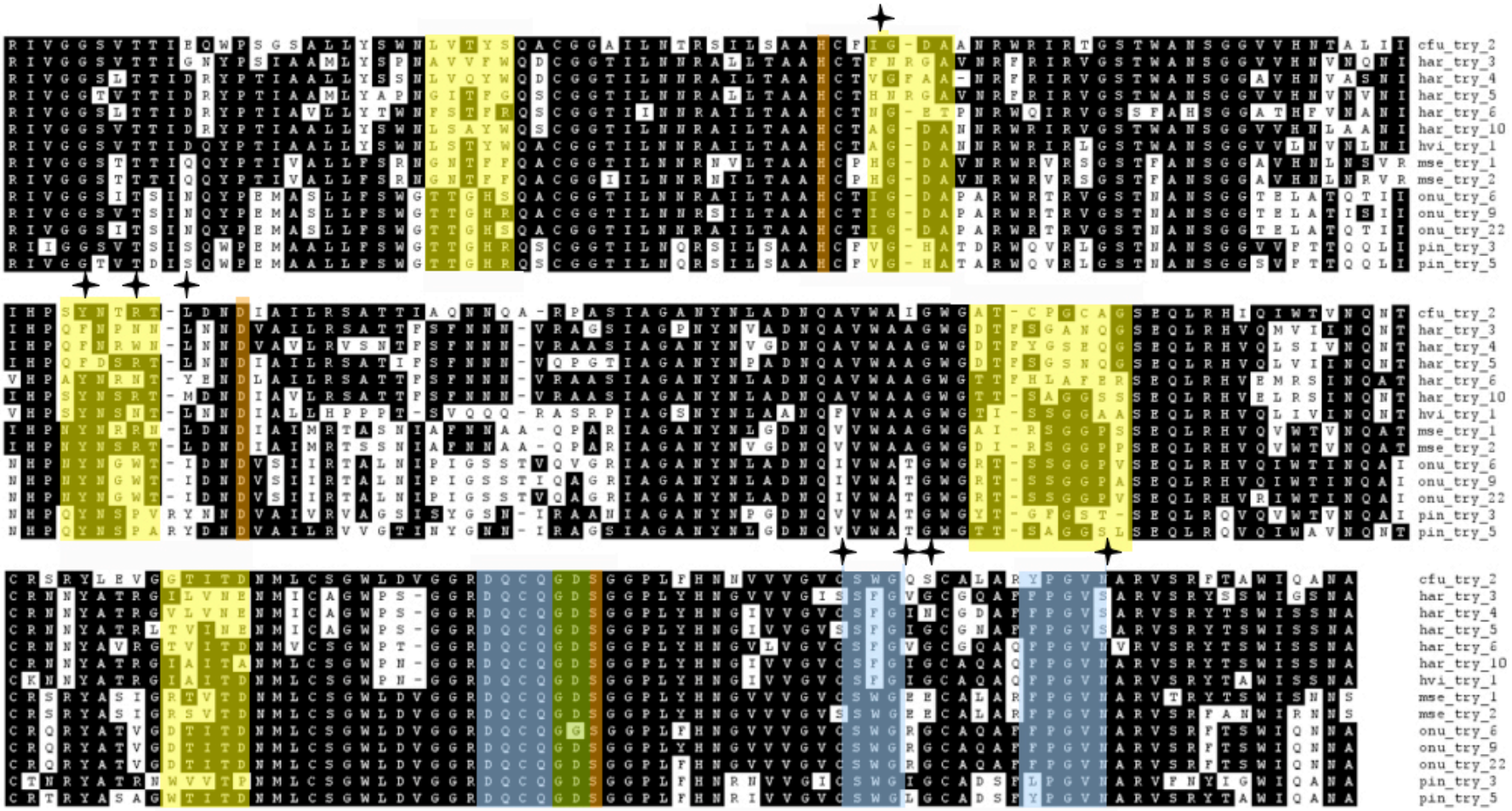

cfu_try_2
har_try_3
har_try_4
har_try_5
har_try_6
har_try_10
hvi_try_1
mse_try_1
mse_try_2
onu_try_6
onu_try_9
onu_try_22
pin_try_3
pin_try_5

Fig. 4.4A

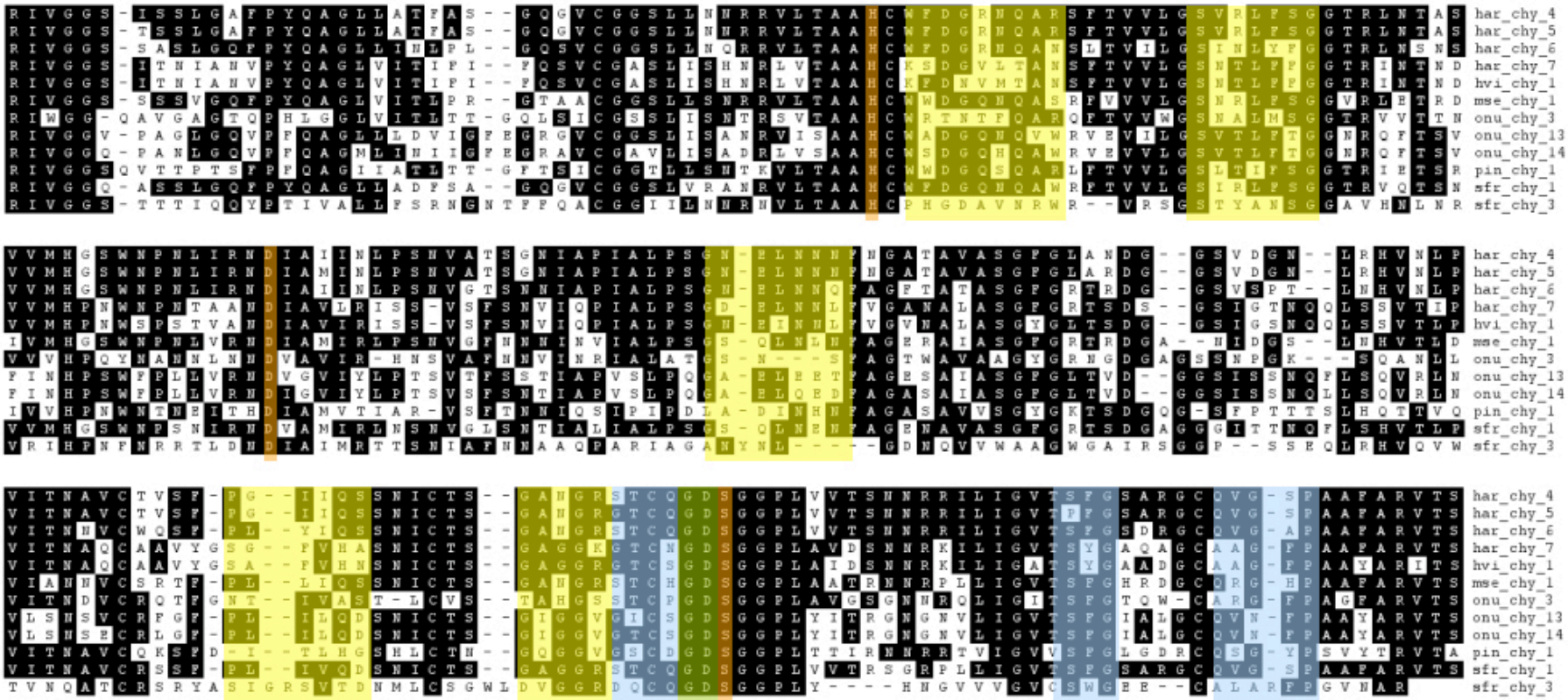

har_chy_4
har_chy_5
har_chy_6
har_chy_7
hvi_chy_1
mse_chy_1
onu_chy_3
onu_chy_13
onu_chy_14
pin_chy_1
sfr_chy_1
sfr_chy_3

Fig. 4.4B

for the relatively lower content. The non-contiguous amino acids (yellow) are thought to be involved in formation of adventitious/secondary contacts (Komiyama *et al*, 2003); not surprisingly, a high degree of variation is observed in these regions as well as in the 'hot spots' (indicated by a '+' in **Fig. 4.4A**), which determine (in)sensitivity to PIs. Thus a cursive study of the primary structure of proteinases reveals a readily apparent diversity in structure, and probably function (Lopes *et al*, 2006), attributable to the natural mutation events and selection of functionally active variants. The presence of multiple proteinase isoforms can be traced back to multi-copy proteinase genes that probably arose due to gene duplication and diversification events (Lopes *et al*, 2004). It has been proposed that the higher success rate in the incorporation of serine residues into catalytic centers coupled with an independent evolution of the various serine proteinase 'clans' (Barrett and Rawlings, 1995) are responsible for the higher diversity observed in serine proteinases (Krem and Cera, 2001). Attempts have also been made to associate structural motifs as 'markers' to trace the evolutionary history as well as inter-relationships between various specificities of serine proteinases (Krem and Cera, 2001). The impact of evolution on diversity in serine proteinases appears to be positive because this mechanistic class has successfully formed the dominant population in the Lepidopteran digestive environment. Though it is possible to visualize the amino acid alterations that influence the structural properties of the translated polypeptides, this does not necessarily provide any information on altered activities. In studies pertaining to diversity of gut proteinases in model insects, co-relation of structural diversity with the biologically more relevant functional diversity has always been a daunting task. Though newly identified putative proteinases are routinely annotated based on similarity of sequence to known proteinases, it may not always be accurate, and at worst may be totally misleading when it comes to predicting the function. Thus, anomalies such as 'functional variants', *viz.*, proteinases homologous to one particular type but having activity similar to another, are routinely observed as exemplified in **Fig. 4.5**, where few representative Lepidopteran serine proteinases have been depicted. It is seen in case of an elastase from *Manduca sexta*, (mse_ela_1, AAA67842), which is similar to a chymotrypsin from *Heliothis virescens* (hvi_chy_1, AAF43709). Another example

Fig. 4.5. Functional diversity in Lepidopteran serine proteases.
This phylogenetic tree was derived based on comparison of non-redundant Lepidopteran endopeptidase sequences and activities were co-related to sequence similarity.
Bombyx mandarina serine protease (bma_ser_1, AAX39408); *B. mori* serine proteases (bmo_ser_1, BAD93199; bmo_ser_3, AAB26023; bmo_ser_5, BAB91156); *Choristoneura furmiferana* trypsin (cfu_try_1, AAA81525); *Helicoverpa armigera* cathepsins (har_cat_1, AAQ75437; har_cat_2, AAF35867); *H. armigera* chymotrypsins (har_chy_1, CAA72960; har_chy_3, CAA72966; har_chy_6, CAA72958; har_chy_8, CAA72951); *H. armigera* serine proteases (har_div_1, CAA72953; har_div_2, CAA72965; har_ser_1, AAC02217; har_ser_2, AAD31713); *H. armigera* trypsins (har_try_1, AAR20817; har_try_3, CAA72956; har_try_4, CAA72955; har_try_5, CAA72954; har_try_6, CAA72949; har_try_11, CAA72957); *Heliothis virescens* chymotrypsin (hvi_chy_1, AAF43709); *H. virescens* trypsin (hvi_try_1, AAF43708); *Lonomia oblique* serine proteases (lob_ser_1, AAV91432; lob_ser_3, AAV91434; lob_ser_4, AAV91435; lob_ser_5, AAV91456; lob_ser_6, AAV91457; lob_ser_7, AAV91544); *Manduca sexta* chymotrypsin (mse_chy_1, AAA58743); *M. sexta* elastase (mse_ela_1, AAA67842); *M. sexta* trypsin (mse_try_1, P35047); *Ostrinia nubilalis* chymotrypsins (onu_chy_1, AAX62040; onu_chy_2, AAX62031); *O. nubilalis* trypsins (onu_try_1, AAX63384; onu_try_5, AAX62036; onu_try_19, AAR98921; onu_try_21, AAR98919); *Plodia interpunctella* chymotrypsin (pin_chy_1, AAC36149); *P. interpunctella* trypsins (pin_try_1, AAF24228; pin_try_3, AAF24226; pin_try_5, AAC36248); *Spodoptera frugiperda* chymotrypsin (sfr_chy_1, AAO75039); *Scirpophaga incertulas* serine protease (sin_ser_1, AAC02219); *S. incertulas* trypsins (sin_try_1, AAC02220; sin_try_2, AAC02218)

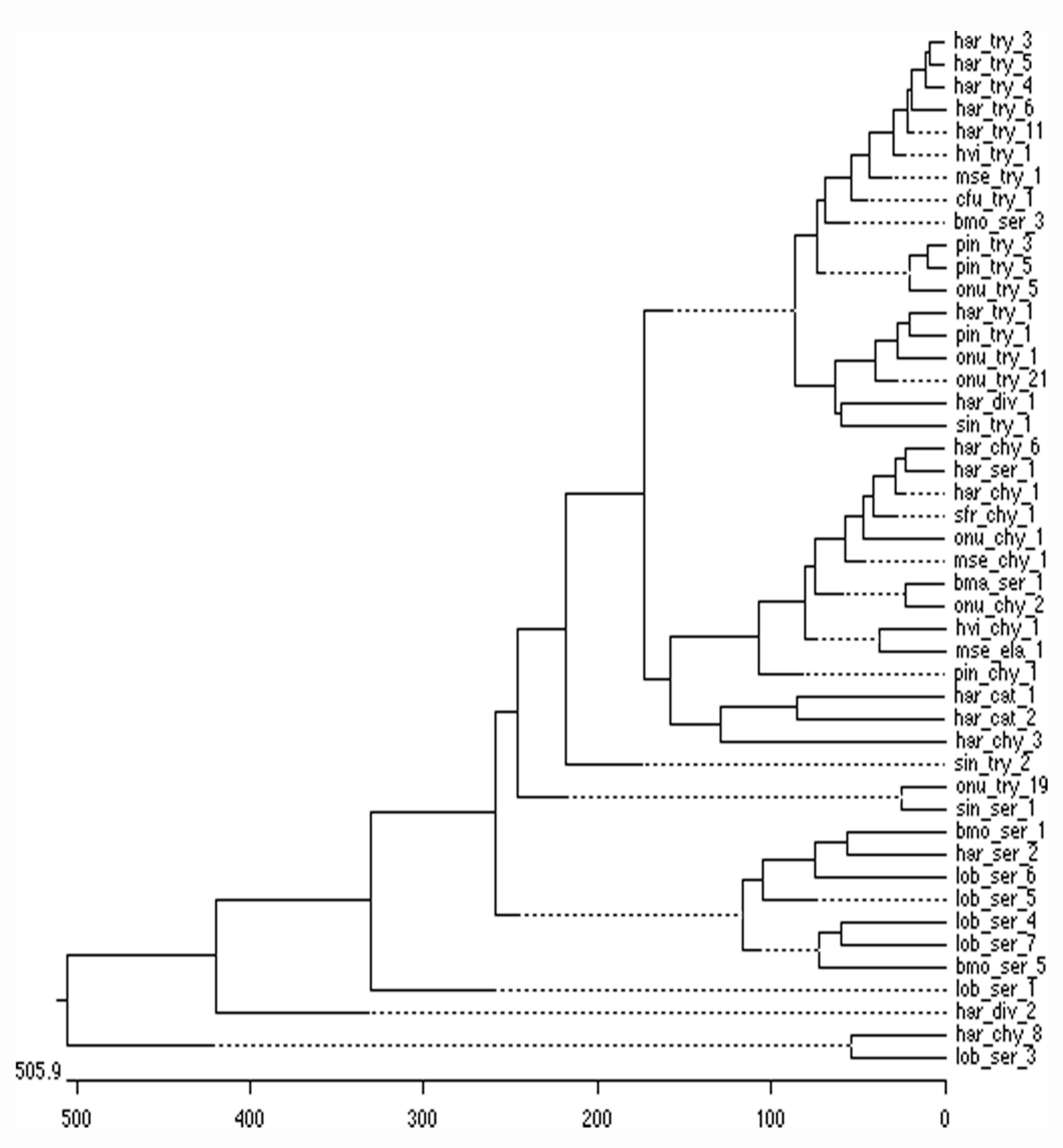

har_try_3
har_try_5
har_try_4
har_try_6
har_try_11
hvi_try_1
mse_try_1
cfu_try_1
bmo_ser_3
pin_try_3
pin_try_5
onu_try_5
har_try_1
pin_try_1
onu_try_1
onu_try_21
har_div_1
sin_try_1
har_chy_6
har_ser_1
har_chy_1
sfr_chy_1
onu_chy_1
mse_chy_1
bma_ser_1
onu_chy_2
hvi_chy_1
mse_ela_1
pin_chy_1
har_cat_1
har_cat_2
har_chy_3
sin_try_2
onu_try_19
sin_ser_1
bmo_ser_1
har_ser_2
lob_ser_6
lob_ser_5
lob_ser_4
lob_ser_7
bmo_ser_5
lob_ser_1
har_div_2
har_chy_8
lob_ser_3
505.9
500
400
300
200
100
0

Fig. 4.5

would be the chymotrypsin from *H. armigera* (har_chy_8, CAA72951), which seems to be unrelated to other chymotrypsins. Newly identified putative serine proteinase genes sequences carry the risk of being mis-annotated until their products are functionally characterized. Hence, activity-characterization of the other serine proteinases in our example, *viz*., those from *Bombyx mori* (bmo_ser_1, BAD93199; bmo_ser_5, BAB91156), *H. armigera* (har_div_2, CAA72965; har_ser_2, AAD31713) and *Lonomia oblique* (lob_ser_1, AAV911432; lob_ser_3, AAV91434; lob_ser_4 AAV91435; lob_ser_5, AV91456; lob_ser_6, AAV91457; lob_ser_7, AAV91544), which form a separate structural group, could further elucidate this feature. Although direct experimental evidence is always preferred for functional characterization, these approaches involve intricate procedures, often not entirely free of errors. In absence of a credible means to co-relate structural and functional aspects, theoretical studies have assumed a greater role; computer-aided modeling and a parallel dissection of structural features and activities in related proteinases have greatly helped these efforts (Atassi and Manshouri, 1993; Beveridge, 1998; Czapinska and Otlewski, 1999; Iengar and Ramakrishnan, 1999; Nishihira and Tachikawa, 1999; Aravind *et al*, 2002; Chou and Cai, 2004; Mekonnen at el, 2006). These studies are based on experimental evidence derived from known sequences and specificities of purified proteinases. A combination of theoretical and experimental studies is essential to derive algorithms that analyze the behavior of amino acids in local environments. Simplified empirical rules may also be helpful within the defined scope of the amino acid type (whether hydrophobic or acidic or aromatic, etc). Although empirical rules generally risk being invalidated by experimental evidence, such attempts would aid refinement of the complex algorithms. A holistic approach is thus necessitated for predicting the structure-function relationship of newly identified proteinase genes for a better understanding of the complexities involved in their evolution, expression and regulation.

Evolutionary aspects of proteinase and PI: implications on plant-herbivorous pest interactions

Due to the importance of digestive proteinases in larval physiology, plant derived PIs have received continuous attention in devising programs to control herbivorous insect pests. As has been discussed previously, binding of PIs to target proteinases causes

inhibition of proteinase activity, which retards digestion and leads to crippling effects not only on larval growth and development but also on the fertility and fecundity of the adult moths. However, in the insect gut, the PI encounters multiple isoforms of the target proteinase, each of which may have its own unique properties with respect to slight differences in association with PI/substrate. This is further complicated by adaptation events involving changes in digestive gut proteinase profile. Hence it would not be directly possible to predict the exact fate of the PI in such a dynamic environment. However, altered binding capabilities in proteinases may also lead to unexpected results, as exemplified by the inhibition of (predominantly trypsin-like) HGPs activity by the winged bean (*Psophocarpus tetragonolobus*) Kunitz type chymotrypsin inhibitor (Giri *et al*, 2003), and CaKPI, that possesses an active site variation and is not expected to inhibit HGP-trypsins (Srinivasan *et al*, 2005a). Evidently feeding habits of the insect, with respect to choice of host plant(s) as well as exposure and possible adaptation to PI(s) that lead to changes in gut proteolytic complement, can be expected to play an important role in determining adaptation/susceptibility to other PIs; apparently the extreme diversity in larval gut proteinases often works out in favor of the insect. Coupled with an effective, but not so well understood, signaling mechanism, the polyphagous larvae have an enviable ability to alter the digestive proteinase complement in response to change in the nutritional quality of the diet and/or towards nutritional challenge(s) arising due to ingestion of anti-metabolic agents like PIs; in many cases, insects are able to successfully escape the anti-metabolic effects of dietary PIs by altering the gut proteinase complement. Just as it is difficult to predict the fate of ingested PIs, insect responses may also be unforeseeable due to reasons of dynamism and diversity in digestive proteinases. The link between insect adaptability, herbivory and diversity of digestive proteinase genes is an interesting study from the evolutionary point of view. It is possible that random changes were followed by natural selection, determined by host plant availability and/or PIs. Whether the insects were pre-adapted to PIs or whether adaptation was gained during evolution is still not clear, although we can understand the role of PIs in contributing to differentiation between host and non-host plants. Adaptation to one particular type of PIs by one species of insects, could offer it a broader host range (Broadway, 1996), i.e. all plants that produce related PIs – in other words, the steady evolutionary responses appear to be the dominant factors in

insect evolution and adaptation. The faster life cycle of insects, coupled with their ability to procreate in large numbers, ensures a rich pool of genetic diversity resulting from mutations as well as DNA recombination events. Apparently, the effects of random mutations on population dynamics is amplified by the high reproductive rate. Evolutionary aspects of defense are also observed in plants, as evident from the structural and functional diversity in PIs. However, it is observed that plant defenses are not grossly overshadowed by herbivorous insects; evidently the plant also seems to have an option of producing a vast range of PIs, which may be activated in response to insect wounding. Thus, the co-adaptive evolutionary race between the plants and insects aims to compensate for each other's arsenal, and remains hard to predict owing to complexities.

Concluding remarks

Plant-insect pest interactions continue to remain a challenging topic for research. Two factors contribute mainly to the perceived challenges, *viz*., (i) the sheer types and number of genes that participate in this interaction, both, in the host as well as the pest, and, (ii) the complexities involved in the expression of these participatory genes. Among other interactions, the chickpea (*C. arietinum*)-podborer (*H. armigera*) association exemplifies the failure of the host defenses (chickpea PIs) in coping with herbivore (podborer) infestation. An approach towards understanding this phenomenon was chalked-out that resulted in conception of the current research-project and culminated in this thesis. In case of chickpea, the Kunitz-type protease inhibitor (CaKPI) was chosen for study as it represented a promising agent towards insect-defense. The innately low levels of CaKPI expression in chickpea coupled with the susceptibility of the more abundant Bowman-Birk type PIs towards degradation by the pest's (digestive) proteolytic machinery was hypothesized as a reason for lack of resistance towards podborer. During the course of *in vivo* and *in vitro* assays, higher levels of CaKPI was observed to cause antagonistic effects on the insect physiology; convincing adaptive responses that would help the insect overcome this defensive barrier were also found lacking. This could pave way for efforts to develop chickpea cultivars that constitutively express higher levels of CaKPI, and test the worthiness of these cultivars under cropping conditions. In fact, non-cultivable (wild) *Cicer* species could also be screened for content and activity of CaKPI-homologs

against *H. armigera*. Host-crops have been largely ignored due to the assumption, on the basis of few examples, that their defenses are bound to fail against insect pests. However, the perceived success of CaKPI against *H. armigera* could serve as a basis for screening host plant PIs for activity against herbivorous pests. In case of *H. armigera*, though the contribution of gut proteases to digestion as well as other key functions is well researched, some questions remain unanswered. Little is known about the actual basis for regulation and differential expression of gut proteases; when multiple mechanistic classes are involved, the lack of understanding is more evident. Finally, concrete relationships between structural and functional aspects of proteases have not been established, although there have been many attempts to develop algorithms on the basis of the contribution of amino acid residues and their contribution to proteinase activity.

Perhaps the biggest threat to agriculture may not be insect attack, but rather limitations in our understanding of the insect pest, which lead to improper strategies to control the insect. Current understanding of the dynamic nature of the Lepidopteran digestive proteases is insufficient for development of 'fool-proof' strategies for insect control; efforts towards deciphering the mechanisms and signaling pathways governing the digestive processes is necessitated. Although it would be a while before significant insights are achieved, currently available information suggests that it would be possible to control herbivore-mediated damage and stem crop-yield losses if not totally prevent infestation; this is especially vital in case of crop plants like chickpea, which are highly susceptible to herbivorous insect attack. Though the balance seems tilted in favor of insect pests at present, future studies could eventually help us to control and probably overcome the devastating effects of insect herbivory.

CHAPTER V:

Thesis Summary and Future Directions

Thesis summary

The objectives in undertaking the current research-project as well as the relevant findings associated with these objectives are listed below:

1) *To verify presence of possible defense related proteinase inhibitors (PIs) in seeds of chickpea*:

Kunitz type PIs in legume seeds have been implicated in a defensive role against herbivorous insect pests. Three isoforms of one such PI, having molecular weight close to that reported for legume Kunitz type PIs (~20 kDa), were identified by the gel-Xray film contact print technique to be present in low levels in chickpea seeds. Unlike the previously reported Bowman-Birk type trypsin inhibitors (TIs), the newly identified PI was stable towards degradation by *H. armigera* gut proteinases (HGPs) and were found to possess inhibitory activity against HGPs. This PI was initially named as *H. armigera* gut proteinase inhibitor (HGPI).

2) *To purify such PI(s), ascertain the identity and characterize further based on sequence information*:

The major isoform of HGPI was purified to ~90% purity by conventional techniques. Sequencing of the N-terminal 20 amino acids yielded a sequence similar to corresponding region in legume Kunitz type PIs. This was augmented by MALDI-TOF analysis of a tyrptic digest of denatured polypeptide, which also revealed similarity to legume Kunitz type PIs. Based on the experimentally deduced amino acid sequence as well as reported legume Kunitz type PI sequences in the GenBank, oligonucleotides primers were synthesized and used for PCR amplification on a chickpea genomic DNA template. This yielded a fragment of expected size (~600 bp), which was cloned into a commercial vector and sequenced. The sequence of the amplified DNA fragment was highly homologous to reported cDNA sequences of Kunitz type inhibitors and coded for a putative polypeptide product of ~200 amino acids. The 'translated' polypeptide had a sequence variation at the putative active site (inhibitory loop) region, wherein the expected arginine (R) or lysine (K) was replaced by a glycine-isoleucine-serine (GIS) motif. This fragment was hence assigned a new

name, *viz*., *cakpi* for (*C. arietinum* Kunitz type proteinase inhibitor) and its polypeptide product as CaKPI.

3) *To clone the coding DNA of the PI into a yeast expression vector for large scale production in a bio-fermenter*:

The coding DNA was cloned into a yeast expression vector and this construct was used to transform competent *Pichia pastoris* cells. After selection of positive transformants and confirmation by colony PCR, expression of recombinant CaKPI was assayed by western blot hybridization, using antibodies directed to the C-terminal His_6 tag that was introduced during cloning. The *P. pastoris* clone showing highest expression was used for large-scale expression of recombinant CaKPI in a laboratory scale (2000 mL) bio-fermenter. The expressed protein was purified from the culture supernatant by hydrophobic interaction chromatography (Phenyl-sepharose) and lyophilized.

4) *To assay in vitro inhibitory activities against commercial and insect gut proteinases as well as to probe in vivo activities, of the PI*:

CaKPI did not inhibit chymotrypsin or subtilisin activity but did exhibit stoichiometric inhibition of trypsin. Inhibitory action was also observed against trypsin-like and total HGP activities. Moreover, sequence data suggested that CaKPI should not have typical trypsin inhibitory activity, but observed data indicated otherwise, it was concluded that CaKPI probably exhibits an atypical mode of proteinase inhibition. *H. armigera* larvae fed on inhibitor-incorporated artificial diet showed significant reduction in weight gain. CaKPI was thus observed to demonstrate potent anti-metabolic activity by interfering with larval digestion and subsequent weight gain during growth and development. It was conclusive that the innately low levels of expression of CaKPI in chickpea seeds do not seem sufficient to contribute significantly to the defense mechanism in chickpea against *H. armigera*.

5) *To study responses of insect pest towards dietary PI so as to ascertain presence or absence of adaptive mechanisms*.

CaKPI degrading proteinases were not found to be constitutively or differentially expressed in *H. armigera* larvae. Insignificant increases in *H. armigera* gut

proteolytic activities following dietary incorporation of CaKPI ruled out adaptive responses based on over-expression of proteinases. The gut proteinase activities in CaKPI fed larvae were still sensitive towards inhibition by CaKPI, indicating that adaptive responses based on synthesis of PI insensitive proteinases were also not involved. Analysis of gut proteinase transcript levels by Quantitative Reverse Transcriptase Polymerase Chain Reaction (Q-RTPCR) confirmed that there was no major shift in the expression of at least 18 major gut proteinase genes in response to CaKPI feeding. Adaptation of *H. armigera* to CaKPI during larval development was ruled out by *in vitro* assays with CaKPI and HGP from larvae at various developmental instars. CaKPI inhibited gut proteinases from larvae fed on various host plants as well as artificial diets incorporated with non-host PIs; host/non-host plant PI induced changes in HGP did not decrease sensitivity towards CaKPI. Continuous exposure of *H. armigera* to basal levels of CaKPI in chickpea seeds, as well as the seemingly anomalous activity/behavior of CaKPI could explain the inability of *H. armigera* to overcome antagonistic effects of CaKPI by adaptation.

6) *To determine the genetic basis for low expression of cakpi*.

c*akpi* was determined to be a low copy number gene. There also appears to be one other gene that shares some sequence similarity with *cakpi*. In this respect *cakpi* shows less diversification into isoforms in comparison to other reported Kunitz type PIs as well as of other mechanical types such as Bowman Birk type, wound inducible type and Squash family type. The paradox in chickpea is that the Bowman Birk type PIs, which are known to exist in at least 4-5 different forms do not seem to contribute to defense against the Lepidopteran herbivore, *Helicoverpa armigera*, although the low copy number and low expressing Kunitz type CaKPI, seems to possess strong potential for defensive function. Further, *cakpi* transcripts were detected only in developing seeds, particularly during the mid-stage, wherein storage proteins are known to be synthesized and accumulated. *cakpi* transcripts were not detected in flowers, leaves or roots. This indicated the strong possibility of seed-specific regulatory elements that govern its expression. Since HGPs inhibitory activity was previously reported to be upregulated in chickpea seeds, when challenged with herbivore (*H. armigera*) attack, it appears that this may be a result of *cakpi* upregulation as a putative defensive response in chickpea. This would particularly

warrant the presence of defense-related regulatory elements, linked to the octadecanoid pathway, which is a well characterized signaling pathway in legume responses to insect herbivory.

7) *To study the differential expression of the defensive PI*

Transcripts of CaKPI were not detected in flower, leaf and root tissue by Northern Blot Hybridization; these transcripts were found to be present only in seed tissue at mid-maturation stage (~25 DAF). This indicates the probable presence of a seed specific promoter governing expression of *cakpi*. The role of CaKPI in developing seeds, may be viewed with respect to the concurrent expression of storage proteins in seeds. Whether CaKPI plays any role in protection of the storage proteins against endogenous proteinases is currently unknown, but seems unlikely. However, CaKPI does seem to have a role in protection of storage proteins against herbivorous insect pests – a function that appears to be belied by the very low level of transcription and translation. Expression of CaKPI also appears to be governed by wound-responsive (defense-related) regulatory elements, as inferable from previous reports.

Future directions

1) Deduce actual mechanism of trypsin inhibition exhibited by CaKPI; whether the novel glycine-isoleucine-serine motif contributes to trypsin binding and inhibition or if the actual inhibitory site is located elsewhere.

2) Probe the reason why *H. armigera* does not appear to respond to elevated levels of dietary CaKPI; to confirm whether this phenomenon is related to the continuous exposure *H. armigera* larvae to basal levels of CaKPI in chickpea plants/chickpea seed-meal based diet.

3) Screening germplasm of related *Cicer* species (cultivars or wilds) that exhibit higher expression of *cakpi* or it's homolog. Determine possibility and feasibility of *cakpi* introgression by inter-crosses.

4) Development of binary vector constructs for *Agrobacterium tumifasciens* mediated transfer of the *cakpi* gene into model systems (*e.g.*, *Arabidopsis thaliana*) or other host plants (*e.g.*, *Lycopersicon esculentum*) – evaluation of *H. armigera* tolerance/resistance in transformants.

5) Genetic enhancement of chickpea cultivars for higher expression of *cakpi* by use of promoter-mutants or by *A. tumifasceins* mediated transformation to increase gene copy number – evaluation for *H. armigera* tolerance/resistance.

6) Determine ratio of fitness cost to benefit (*H. armigera* resistance) in transgenics over-expressing CaKPI.

CHAPTER VI:

Bibliography

Publications arising from the current research work:

1) **Srinivasan A**, Giri AP, Harsulkar AM, Gatehouse JA, Gupta VS (2005) A Kunitz trypsin inhibitor from chickpea (*Cicer arietinum* L.) that exerts anti-metabolic effect on podborer (*Helicoverpa armigera*) larvae. Plant Molecular Biology 57: 359–374.

2) **Srinivasan A**, Chougule NP, Giri AP, Gatehouse JA, Gupta VS (2005) Podborer (*Helicoverpa armigera* Hübn.) responses to the *Cicer arietinum* Kunitz Proteinase Inhibitor. Journal of Insect Physiology 51: 1268-1276.

3) **Srinivasan A**, Giri AP, Gupta VS (2005) Structural and functional diversities in Lepidopteran gut proteinases – a review. (communicated to Cellular and Molecular Biology Letters)

4) **Srinivasan A**, Harsulkar AM, Giri AP, Gupta VS (2005) Genetic structure and differential expression of the *Cicer arietinum* Kunitz proteinase inhibitor gene (*cakpi*). (Manuscript under preparation)

Related publications:

1) Telang MA, **Srinivasan A**, Patankar AG, Harsulkar AM, Joshi VV, Damle A, Deshpande VV, Sainani MN, Ranjekar PK, Gupta GP, Birah A, Rani S, Kachole MS, Giri AP, Gupta VS (2003) Bitter gourd proteinase inhibitors: potential growth inhibitors of *Helicoverpa armigera* and *Spodoptera litura*. Phytochemistry 63: 643–652.

Cited references:

Abdeen A, Virgos A, Olivella E, Villanueva J, Aviles X, Gabarra R, Prat S (2005) Multiple insect resistance in transgenic tomato plants over-expressing two families of plant proteinase inhibitors. Plant Molecular Biology 57: 189-202.

Agrawal AA (1998) Induced responses to herbivory and increased plant performance. Science 279: 1201-1202.

Akhurst RJ, James W, Bird LJ, Beard C (2003) Resistance to the Cry1Ac δ-endotoxin of *Bacillus thuringiensis* in the cotton bollworm, *Helicoverpa armigera* (Lepidoptera: Noctuidae). Journal of Economical Entomology 96: 1290-1299.

Altschul AA, Yatsu LY, Ory RL, Engleman EM (1965) Seed proteins. Annual Review of Plant Physiology 17: 113-136.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. Journal of Molecular Biology 215: 403-410.

Apostoluk W and Otlewski J (1998) Variability of the canonical loop conformations in serine proteinases inhibitors and other proteins. Proteins: Structure, Function, and Genetics 32:459-474.

Araujo APU, Hansen D, Vieira DF, deOliviera C, Santana LA, Beltramini LM, Sampaio CAM, Sampaio MU, Oliva MLV (2005) Kunitz-type *Bauhinia bauhinioides* inhibitors devoid of disulfide bridges: isolation of the cDNAs, heterologous expression and structural studies. Biological Chemistry 386: 561-568.

Aravind L, Mazumder R, Vasudevan S, Koonin VE (2002) Trends in protein evolution inferred from sequence and structure analysis. Current Opinion in Structural Biology 12: 392-399.

Ardelt W and Laskowski M Jr. (1985) Turkey ovomucoid third domain inhibits eight different serine proteinases of varied specificity on the same ...Leu18-Glu19 ... reactive site. Biochemistry 24: 5313–5320.

Arimura G, Kost C, Boland W (2005) Herbivore-induced, indirect plant defences. Biochimica et Biophhysica Acta 1734: 91-111.

Armes NJ, Jadhav DR, DeSouza KR (1996) A survey of insecticide resistance in *Helicoverpa armigera* in the Indian subcontinent. Bulletin Entomological Research 86: 499-514.

Aronson AI and Shai Y (2001) Why *Bacillus thuringiensis* insecticidal toxins are so effective: unique features of their mode of action. FEMS Microbiology Letters 195: 1-8.

Atassi MZ and Manshouri T (1993) Design of peptide enzymes (pepzymes): surface-simulation synthetic peptides that mimic the chymotrypsin and trypsin active sites exhibit the activity and specificity of the respective enzyme. Proceedings of the National Academy of Sciences, USA 90: 8282-8286.

Augur C, Stiefel V, Darvill A, Albersheim P, Puigdomenech P (1995) Molecular cloning and pettern of expression of an α-L-fucosidase gene from pea seedlings. Journal of Biological Chemistry 270: 24839-24843.

Babu RM, Sajeena A, Seetharaman K, Reddy MS (2003) Advances in genetically engineered (transgenic) plants in pest management – an overview. Crop Protection 22; 1071-1086.

Baillargeon MW, Laskowski M, Neves DE, Porubcan MA, Santini RE, Markley JL (1980) Soybean trypsin inhibitor (Kunitz) and its complex with trypsin. Carbon-13 nuclear magnetic resonance studies of the reactive site arginine. Biochemistry 19: 5703-5710.

Barbehenn RV (2001) Roles of peritrophic membranes in protecting herbivorous insects from ingested plant allelochemicals. Archives of Insect Biochemistry and Physiology 47: 86-99.

Barrett AJ and Rawlings ND (1995) Families and clans of serine peptidases. Archives of Biochemistry and Biophysics 318: 247-250.

Barton KA, Whiteley HR, Yang NS (1987) *Bacillus thuringiensis* delta-endotoxin expressed in transgenic *Nicotiana tabacum* provides resistance to lepidopteran insects. Plant Physiology 85: 1103-1109.

Bayes A, Sonnenschein A, Daura X, Vendrell J, Aviles FX (2003) Procarboxypeptidase A from the insect pest *Helicoverpa armigera* and its derived enzyme. FEBS 270: 3026-3035.

Belew M and Eaker D (1976) The trypsin and chymotrypsin inhibitors in chickpeas (*Cicer arietinum* L). Identification of the trypsin-reactive site, partial-amino-acid sequence and further physico-chemical properties of the major inhibitor. European Journal of Biochemistry 62: 499-508.

Belew M, Porath J, Sundberg L (1975) The trypsin and chymotrypsin inhibitors in chickpeas (*Cicer arietinum* L). Purification and properties of the inhibitors. European Journal of Biochemistry 60: 247-258.

Bell HA, Fitches EC, Down RE, Marris GC, Edwards JP, Gatehouse JA, Gatehouse AMR (1999) The efect of snowdrop lectin (GNA) delivered via artificial diet and transgenic plants on *Eulophus pennicornis* (Hymenoptera: Eulophidae), a parasitoid of the tomato moth *Lacanobia oleracea* (Lepidoptera: Noctuidae). Journal of Insect Physiology 45: 983-991.

Bendtsen JD, Nielsen H, von Heijne G, Brunak S (2004). Improved prediction of signal peptides – SignalP 3.0. Journal of Molecular Biology. 340: 783-795.

Beveridge AJ (1998) A theoretical study of the initial stages of catalysis in the aspartic proteinases. Journal of Molecular Chemistry 453: 275-291.

Birch NA, Geoghegan IE, Majerus MEN, McNicol JW, Hackett, Gatehouse AMR, Gatehouse JA (1999) Tri-trophic interactions involving pest aphids, predatory2-spot ladybirds and transgenic potatoes expressing snowdrop lectin for aphid resistance. Molecular Breeding 5: 75–83.

Bird LJ and Akhurst RJ (2005) Fitness of Cry1A-resistant and –susceptible *Helicoverpa armigera* (Lepidoptera: Noctuidae) on transgenic cotton with reduced levels of Cry1Ac. Journal of Economical Entomology 98: 1311-1319.

Birk Y (1961) Purification and some properties of a highly active inhibitor of trypsin and alpha-chymotrypsin from soybeans. Biochimica et Biophysica Acta 54: 378-381.

Birk Y (1985) The Bowman-Birk inhibitor. Trypsin- and chymotrypsin-inhibitor from soybeans. International Journal of Peptide and Protein Research. 25: 113-131.

Birk Y, Gertler A, Khalef S (1963) A pure trypsin inhibitor from soya beans. The Biochemical Journal

87: 281-284.

Blow DM, Janin J, Sweet RM (1974) Mode of action of soybean trypsin inhibitor (Kunitz) as a model for specific protein-protein interactions. Nature 249: 54-57.

Bode W and Huber R (1978) Crystal structure analysis and refinement of two variants of trigonal trypsinogen: trigonal trypsin and PEG (polyethylene glycol) trypsinogen and their comparison with orthorhombic trypsin and trigonal trypsinogen. FEBS Letters 90: 265-269.

Bode W, Meyer E, Powers JC (1988) Human leukocyte and porcine pancreatic elastase: X-ray crystal structures, mechanism, substrate specificity, and mechanism-based inhibitors. Biochemistry 28: 1951-1963.

Bolognesi R, Ribeiro AF, Terra WR, Ferreira C (2001) The peritrophic membrane of *Spodoptera frugiperda*: Secretion of peritrophins and role in immobilization and recycling digestive enzymes. Archives of Insect Biochemistry and Physiology 47: 62-75.

Boulter D (1993) Insect pest control by copying nature using genetically engineered crops. Phytochemistry 34: 1453-1466.

Bowman DE (1946) Differentiation of soybean antitrypsin factors. Proceedings of the Society for Experimental Biology and Medicine 63: 547-550.

Bown DP, Wilkinson HS, Gatehouse JA (1997). Differentially regulated inhibitor-sensitive and insensitive protease genes from the phytophagous insect pest, *Helicoverpa armigera*, are members of complex multigene families. Insect Biochemistry and Molecular Biology 27: 625-638.

Bown DP, Wilkinson HS, Gatehouse JA (1998) Midgut carboxypeptidase from *Helicoverpa armigera* (Lepidoptera: Noctuidae) larvae: enzyme characterisation, cDNA cloning and expression. Insect Biochemistry and Molecular Biology 28: 739–749.

Bown DP, Wilkinson HS, Gatehouse JA (2004) regulation of expression of genes encoding digestive proteases in the gut of a polyphagous lepidopteran larva in response to dietary protease inhibitors. Physiological Entomology 29: 278-290.

Brito LO, Lopes AR, Parra JRP, Terra WR, Silva-Filho MC (2001) Adaptation of tobacco budworm *Heliothis virescens* to proteinase inhibitors may be mediated by the synthesis of new proteinases. Comparative Biochemistry and Physiology 128: 365–375.

Broadway RM (1995) Are insects resistant to plant proteinase inhibitors. Journal of Insect Physiology 41, 1107-116.

Broadway RM (1996) Plant dietary proteinase inhibitors alter complement of midgut proteases. Archives of Insect Biochemistry and Physiology 32: 39–53.

Broadway RM (1997) Dietary regulation of serine proteinases that are resistant to serine proteinase inhibitors. Journal of Insect Physiology 43: 855–874.

Broadway RM and Duffey SS (1986) Plant proteinase inhibitors: mechanism of action and effect on the growth and digestive physiology of larval *Heliothis zea* and *Spodoptera exigua*. Journal of Insect Physiology 32: 827–833.

Brunelle F, Girard C, Cloutier C, Michaud D (2005) A hybrid, broad-spectrum inhibitor of Colorado

potato beetle aspartate and cysteines digestive proteinases. Archives of Insect Biochemistry and Physiology 60: 20-31.

Capasso C, Rizzi M, Menegatti E, Ascenzi P, Bolognesi M (1997) Crystal structure of the bovine α-chymotrypsin:Kunitz inhibitor complex. An example of multiple protein:protein recognition sites. Journal of Molecular Recognition 10, 26-35.

Carlini CR, Grossi-de-Sa MF (2002) Plant toxic proteins with insecticidal properties. A review on their potentialities as bioinsecticides. Toxicon 40: 1515-1539.

Carriere Y, Ellers-Kirk C, Sisterson M, Antilla L, Whitlow M, Dennehy TJ, Tabashnik DE (2003) Long-term regional suppression of pink bollworm by *Bacillus thuringiensis* cotton. Proceedings of the National Academy of Sciences, USA 100: 1519-1523.

Chamberlain K, Pickett JA, Woodcock (2000) Plant signalling and induced defence in insect attack. Molecular Plant Pathology. 1: 67-72.

Chomczynski P, Sacchi N (1986) Single step method of RNA isolation by acid guanidium thiocynate-phenol-chloroform extraction. Analytical Biochemistry 162: 156–159.

Chou KC and Cai Y (2004) A novel approach to predict active sites of enzyme molecules. Proteins: Structure, function and bioinformatics 55: 77-82.

Chougule NP, Giri AP, Sainani MN, Gupta VS (2005) Gene expression patterns of *Helicoverpa armigera* gut proteases. Insect Biochemistry and Molecular Biology 35: 355-367.

Christeller J and Laing W (2005) Plant serine proteinase inhibitors. Protein and Peptide Letters 12: 1-9.

Cooksey KE. (1968) Purification of a protein from *Bacillus thuringiensis* toxic to larvae of lepidoptera. Biochemical Journal 106: 445-454.

Czapinska H and Otlewski J (1999) Structural and energetic determinants of the S1-site specificity in serine proteases. European Journal of Biochemistry 260: 571-595.

Damle MS, Giri AP, Sainani MN, Gupta VS (2005) Higher accumulation of proteinase inhibitors in flowers than leaves and fruits as a possible basis for differential feeding preference of *Helicoverpa armigera* on tomato (*Lycopersicon esculentum* Mill, Cv. Dhanashree). Phytochemistry 66: 2659-2667.

Dasgupta G, Sen U, Dattagupta JK (2003) *In silico* mutations and molecular dynamics studies on a winged bean chymotrypsin inhibitor protein. Protein Engineering 16: 489-496.

Dattagupta JK, Podder A, Chakrabarti C, Sen U, Dutta SK, Singh M (1996) Structure of a Kunitz-type chymotrypsin inhibitor from winged bean seeds at 2.95 A resolution. Acta Crystallographica D 52: 521-528.

Dattagupta JK, Podder A, Chakrabarti C, Sen U, Mukhopadhyay D, Dutta SK, Singh M (1999) Refined crystal structure (2.3 A) of a double-headed winged bean alpha-chymotrypsin inhibitor and location of its second reactive site. Proteins 35: 321-331.

Davey M, Duve H, Thorpe A, East P (2005) Helicostatins: brain-gut peptides of the moth, *Helicoverpa armigera* (Lepidoptera: Noctuidae). Archives of Insect Biochemistry and Physiology 58: 1-16.

David F, Bernard AM, Pierres M, Marguet D (1993) Identification of serine 624, aspartic acid 702 and histidine 734 as the catalytic triad residues of mouse dipeptidyl-peptidase IV (CD26): a member of the novel family of nonclassical serine hydrolases. Journal of Biological Chemistry 268: 17247-17252.

Dean DH, Rajamohan F, Lee MK, Wu SJ, Chen XJ, Alcantara E, Hussain SR (1996) Probing the mechanism of action of *Bacillus thuringiensis* insecticidal proteins by site directed mutagenesis – a minireview. Gene 179: 111-117.

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

DeLeo F, Bonade-Bottino M, Ceci LR, Gallerani R, Jouanin L (2001) Effects of a mustard trypsin inhibitor expressed in different plants on three lepidopteran pests. Insect Biochemistry and Molecular Biology 31: 593-602.

DeLeo F, Gallerani R (2002) The mustard trypsin inhibitor 2 affects the fertility of *Spodoptera littoralis* larvae fed on transgenic plants. Insect Biochemistry and Molecular Biology 32: 489-496.

DeMoraes CM, Lewis WJ, Pare PW, Alborn HT, Tumlinson JH (1998) Herbivore-infested plants selectively attract parasitoids. Nature. 393: 570-573.

Deshimaru M, Yoshimi S, Shioi S, Terada S (2004) Multigene family for Bowman-Birk type proteinase inhibitors of wild soja and soybean: the presence of two BBI-A genes and pseudogenes. Bioscience, Biotechnology and Biochemistry 68: 1279-1286.

Devi PS, Prasad YG, Chowdary DA, Rao LM, Balakrishnan K (2003) Identification of virulent isolates of the entomopathogenic fungus *Nomuraea rileyi* (F) Samson for the management of *Helicoverpa armigera* and *Spodoptera litura* (identification of virulent isolates of N. rileyi). Mycopathologia. 156: 365-373.

Dodson G and Wlodawer A (1998) Catalytic triads and their relatives. Trends in Biochemical Sciences 23: 347-352.

Domoney C, Welham T, Sidebottom C, Firmin JL (1995) Multiple isoforms of *Pisum* trypsin inhibitors result from modification of two primary gene products. FEBS Letters 360: 15-20.

doSocorro M, Cavalcanti M, Oliva ML, Fritz H, Jochum M, Mentele R, Sampaio M, Coelho LC, Batista IF, Sampaio CA (2002) Characterization of a Kunitz trypsin inhibitor with one disulfide bridge purified from *Swartzia pickellii*. Biochemical and Biophysical Research Communications 291: 635-639.

Down RE, Ford L, Bedford SJ, Gatehouse LN, Newell C, Gatehouse JA, Gatehouse AMR (2001) Influence of plant development and environment on transgene expression in potato and consequences for insect resistance. Transgenic Research 10: 223-236.

Down RE, Ford L, Woodhouse SD, Raemaekers RJM, Leitch B, Gatehouse JA, Gatehouse AMR (2000) Snowdrop lectin (GNA) has no acute toxic effects on a beneficial insect predator, the 2-spot ladybird (*Adalia bipunctata* L.). Journal of Insect Physiology 46: 379-391.

Down RE, Gatehouse AMR, Hamilton WDO, Gatehouse JA (1996) Snowdrop lectin inhibits development and decreases fecundity of the glasshouse potato aphic (*Aulacorthum solani*) when administered *in vitro* and via transgenic plants both in laboratory and glasshouse trials. Journal of Insect Physiology 42: 1035-1045.

Doyle JJ and Doyle JL (1987) A rapid DNA isolation procedure for small amount of fresh leaf tissue. Phytochemistry Bulletin 19: 11-15.

Duffey S, Stout M (1996) Antinutritive and toxic components of plant defense against insects. Archives of Insect Biochemistry and Physiology 32: 3–37.

Edmonds HS, Gatehouse LN, Hilder VA, Gatehouse JA (1996) The inhibitory effects of the cysteine protease inhibitor, oryzacystatin, on digestive proteases and on larval survival and development of the southern corn rootworm (*Diabrotica undecimpunctata howardi*). Entomologia et Experimentalis Applicata 78: 83-94.

Ehrlich PR and Raven PH (1964) Butterflies and plants: a study in coevolution. Evolution 18: 586-608.

Evnin LB, Vasquez JR, Craik CS (1990) Substrate specificity of trypsin investigated by using a genetic selection. Proceedings of the National Academy of Sciences, USA 87: 6653-6659.

Ferry N, Edwards MG, Gatehouse JA, Gatehouse AMR (2004) Plant–insect interactions: molecular approaches to insect resistance Current Opinion in Biotechnology 15: 155-161.

Fitches E and Gatehouse JA (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*). Journal of Insect Physiology 44: 1213-1224.

Fitches E, Audsley N, Gatehouse JA, Edwards JP (2002) Fusion proteins containing neuropeptides as novel insect control agents: snowdrop lectin delivers fused allatostatin to insect haemolymph following oral ingestion. Insect Biochemistry and Molecular Biology 32: 1653-1661.

Fitches E, Edwards MG, Mee C, Grishin E, Gatehouse AMR, Edwards JP, Gatehouse JA (2004) Fusion proteins containing insect-specific toxins as pest control agents: snowdrop lectin delivers fused insecticidal spider venom toxin to insect haemolymph following oral ingestion. Journal of Insect Physiology 50: 61-71.

Fitches E, Gatehouse AMR, Gatehouse JA (1997) Effects of snowdrop lectin (GNA) delivered via artificial diet and transgenic plants on the development of tomato moth (*Lacanobia oleracea*) larvae in laboratory and glasshouse trials. Journal of Insect Physiology 43: 727-739.

Fitches E, Ilett C, Gatehouse AMR, Gatehouse LN, Greene R, Edwards JP, Gatehouse JA (2001b) The effects pf *Phaseolus vulgaris* erythro- and leucoagglutinating isolectins (PHA-E and PHA-L) delivered via artificial diet and transgenic plants on the growth and development of tomato moth (*Lacanobia oleracea*) larvae; lectin binding to gut glycoproteins in vitro and in vivo. Journal of Insect Physiology 47: 1389-1398.

Fitches E, Woodhouse SD, Edwards JP, Gatehouse JA (2001a) In vitro and in vivo binding of snowdrop (*Galanthus nivalis* agglutinin; GNA) and jackbean (*Canavalia ensiformis*; Con A) lectins within tomato moth (*Lacanobia oleracea*) larvae; mechanisms of insecticidal action. Journal of Insect Physiology 47: 777-787.

Fodor K, Harmat V, Hetenyl C, Kardos J, Antal J, Perczel A, Patthy A, Katona G, Graf L (2005) Extended intramolecular interactions in a serine protease-canonical inhibitor complex account for strong and highly specific inhibition. Journal of Molecular Biology 350: 156-169.

Foissac X, Nguyen TL, Christoi P, Gatehouse AMR, Gatehouse JA (2000) Resistance to green leafhopper (*Nephotettix virescens*) and brown plant hopper (*Nilaparvata lugens*) in transgenic rice expressing snowdrop lectin (*Galanthus nivalis* agglutinin; GNA). Journal of Insect Physiology 46: 573–583.

Forcada C, Alcacer E, Garcera MD, Martinez R (1996) Differences in the midgut proteolytic activity of two *Heliothis virescens* strains, one susceptible and one resistant to *Bacillus thurigiensis* toxins. Archives of Insect Biochemistry and Physiology 31, 257–272.

Franco OL, Grossi de Sa MF, Sales MP, Mello LV, Oliveira AS, Rigden DJ (2002) Overlapping binding sites for trypsin and papain on a Kunitz-type proteinase inhibitor from *Prosopis juliflora*. Proteins 49: 335-341.

Frutos R, Rang C, Royer M (1999) Managing insect resistance to plants producing *Bacillus thuringiensis* toxins. Critical Reviews in Biotechnology 19: 227–276..

Gahan LJ, Ma YT, Coble MLM, Gould F, Moar WJ, Heckel DG (2005) Genetic basis of resistance to Cry1Ac and Cry2Aa in *Heliothis virescens* (Lepidoptera: Noctuidae) Journal of Econoical Entomology 98: 1357-1368.

Garcia-Olmedo F, Salcedo G, Sanchez-Monge R, Gomez L, Roys J, Carbonero P (1987) Plant proteinaceous inhibitors of proteases and amylases. Oxford Survey of Plant Molecular and Cellular Biology 4: 275–334.

Gatehouse JA (2002) Plant resistance towards insect herbivores: a dynamic interaction (Tansley review no. 140) New Phytologist 156: 145-169.

Gatehouse LN, Shannon AL, Burgess EPJ, Christeller JT (1997) Characterization of major midgut proteinase cDNAs from *Helicoverpa armigera* larvae and changes in gene expression in response to four proteinase inhibitors in the diet. Insect Biochemistry and Molecular Biology 27: 929-944.

Gattiker A, Gasteiger E, Bairoch A (2002) ScanProsite: a reference implementation of a PROSITE scanning tool. Applied Bioinformatics 1: 107-108.

Gennaro SD, Ficca AG, Panichi D, Poerio E (2005) cDNA cloning and heterologous expression of a wheat proteinase inhibitor of subtilisin and chymotrypsin (WSCI) that interferes with digestive enzymes of insect pests. Biological Chemistry 386: 383-389.

Girard C and Jouanin L (1999) Molecular cloning of cDNAs encoding a range of digestive enzymes from a phytophagous beetle, *Phaedon cochleariae*. Insect Biochemistry and Molecular Biology 29: 1129-1142.

Girard C, Le Metayer M, Bonade-Bottino M, Pham-Delegue MH, Jouanin L (1998a) High level of resistance to proteinase inhibitors may be conferred by proteolytic cleavage in beetle larvae. Insect Biochemistry and Molecular Biology 28: 229–237.

Girard C, Le Metayer M, Zaccomer B, Bartlet E, Williams I, Bonade-Bottino M, Pham-Delegue MH, Jouanin L (1998b) Growth stimulation of beetle larvae reared on transgenic oilseed rape

expressing a cysteine proteinase inhibitor. Journal of Insect Physiology 44: 263–270.

Giri AP, Chougule NP, Telang MA, Gupta VS (2005) Engineering insect tolerant plants using plant defensive proteinase inhibitors. in: Recent Research Developments in Phytochemistry, (Pandalai, S. G., Ed) Research Signpost, India, 8: 117-137.

Giri AP, Harsulkar AM, Deshpande VV, Sainani MN, Gupta VS, Ranjekar PK (1998) Chickpea defensive proteinase inhibitors can be inactivated by podborer gut proteinases. Plant Physiology 116: 393-401.

Giri AP, Harsulkar AM, Ku MSB, Gupta VS, Deshpande VV, Ranjekar PK, Franceschi VR (2003) Identification of potent inhibitors of *Helicoverpa armigera* gut proteinases from winged bean seeds. Phytochemistry 63: 523-532.

Gringorten JL, Crawford DN, Harvey WR (1993) High pH in the ectoperitrophic space of the larval lepidopteran midgut. Journal of Experimental Biology 183: 353-359.

Groot AT and Dicke M (2002) Insect-resistant transgenic plants in a multi-trophic context. The Plant Journal 31: 387-406.

Gruden K, Kuipers AGJ, Guncar G, Slapar N, Strukelj B, Jongsma MA (2004) Molecular basis of Colorado potato beetle adaptation to potato plant defence at the level of digestive cysteines proteinases. Insect Biochemistry and Molecular Biology 34: 365-375.

Grundy PR, Sequeira RV, Short KS (2004) Evaluating legume species as alternative trap crops to chickpea for management of *Helicoverpa* spp. (Lepidoptera: Noctuidae) in central Queensland cotton cropping systems. Bulletin of Entomological Research 94: 481-486.

Gupta GP, Mahapatro GK, Chandra A (2000). Bio-potency of insecticidal crystal proteins of *Bacillus thuringiensis* against cotton (*Gossypium hirsutum*) bollworms. Indian Journal of Agricultural Sciences 70: 194-196.

Gussow D, Clackson T (1989) Direct clone characterization from plaques and colonies by the polymerase chain reaction. Nucleic Acids Research 17: 4000-4008.

Habu Y, Peyachoknagul S, Umemoto K, Sakata Y, Ohno T (1992) Structure and regulated expression of Kunitz chymotrypsin inhibitor genes in winged bean (*Psophocarpus tetragonolobus* L. DC.). Journal of Biochemistry 111: 249-258.

Haq SK, Atif SM, Khan RH (2004) Protein proteinase inhibitor genes in combat against insects, pests and pathogens: natural and engineered phytoprotection. Archives of Biochemistry and Biophysics 431: 145-159.

Harshini S, Nachman RJ, Sreekumar S (2002a) Inhibition of digestive enzyme release by neuropeptides in larvae of *Opisina arenosella* (Lepidoptera: Cryptophasidae). Comparative Biochemistry and Physiology Part B 132: 353-358.

Harshini S, Nachman RJ, Sreekumar S (2002b) *In vitro* release of digestive enzymes by FMRF amide related neuropeptides and analogues in the lepidopteran insect *Opisina arenosella* (Walk.). Peptides 23: 1759-1763.

Harsulkar AM, Giri AP, Gupta VS, Sainani MN, Deshpande VV, Patankar AG, Ranjekar PK (1998) Characterization of *Helicoverpa armigera* gut proteinases and their interaction with proteinase

inhibitors using gel X-ray film contact print technique. Electrophoresis 19: 1397-1402.

Harsulkar AM, Giri AP, Patankar AG, Gupta VS, Sainani MN, Ranjekar PK, Deshpande VV (1999) Successive use of non-host plant proteinase inhibitors required for effective inhibition of *Helicoverpa armigera* gut proteinases and larval growth. Plant Physiology 121: 497-506.

Hartmann T (2004) Plant-derived secondary metabolites as defensive chemicals in herbivorous insects: a case study in chemical ecology. Planta 219: 1-4.

Hazzard RV, Schultz BB, Groden E, Ngollo ED, Seidlecki E (2003) Evaluation of oils and microbial pathogens for control of lepidopteran pests of sweet corn in New England. Journal of Econonomic Entomology 96: 1653-1661.

Hedstrom L (2000) Serine protease mechanism and specificity. Chemical Reviews 102: 4501 4523.

Hedstrom L, Farr-Jones S, Kettner CA, Rutter WJ (1994) Converting trypsin to chymotrypsin: ground-state binding does not determine substrate specificity. Biochemistry 33: 8764-8769.

Hegedus D, Baldwin D, O'Grady M, Braun L, Gleddie S, Sharpe A, Lydiate D, Erlandson M (2003) Midgut proteases from *Mamestra configurata* (Lepidoptera: Noctuidae) larvae: characterization, cDNA cloning and expressed sequence tag analysis. Archives of Insect Biochemistry and Physiology 53: 30–47.

Heibges A, Glaczinski H, Ballvora A, Salamini F, Gebhardt C (2003a) Structural diversity and organization of three gene families for Kunitz-type enzyme inhibitors from potato tubers (*Solanum tuberosum* L.) Molecular Genetics Genomics. 269: 526-534.

Heibges A, Salamini F, Gebhardt C (2003b) Functional comparison of homologous members of three groups of Kunitz-type enzyme inhibitors from potato tubers (*Solanum tuberosum* L.) Molecular Genetics and Genomics 269: 535-541.

Helland R, Czapinska H, Leiros I, Olufsen M, Otlewski J, Smalas AO (2003) Structural consequences of accommodation of four non-cognate amino acid residues in the S1 pocket of bovine trypsin and chymotrypsin. Journal of Molecular Biology 333: 845-861.

Herrero S, Combes E, Oers MMV, Vlak JM, deMaagd RA, Beekwilder J (2005) Identification and recombinant expression of a novel chymotrypsin from *Spodoptera exigua*. Insect Biochemistry and Molecular Biology 35: 1073-1082.

Hilder VA and Boulter D (1999) Genetic engineering of crop plants for insect resistance – a critical review. Crop Protection 18: 177-191.

Hilder VA, Gatehouse AMR, Sheerman SF, Barker RF, Boulter D (1987) A novel mechanism of insect resistance engineered into tobacco. Nature 330: 160–163.

Hollick JB, Gordon MP (1995) Transgenic analysis of a hybrid poplar wound-inducible promoter reveals developmental patterns of expression similar to that of storage protein genes. Plant Physiology. 109: 73-85.

Huang K, Lu W, Anderson S, Laskowski M Jr, James MNG (1995) Water molecules participate in proteinase-inhibitor interactions: crystal structures of Leu18, Ala18, and Gly18 variants of turkey ovomucoid inhibitor third domain complexed with Streptomyces griseus proteinase B. Protein Science 4: 1985–1997

Huang Y, Brown MR, Lee TD, Crim JW (1998) RF-amide peptides isolated from the midgut of the corn earworm, *Helicoverpa zea*, resemble pancreatic polypeptide. Insect Biochemistry and Molecular Biology 28: 345-356.

Huber R and Bode W (1978) Structural basis of the activation and action of trypsin. Accounts of Cheimcal Research 11: 114-122.

Hunkapiller MW, Smallcombe SH, Whitaker DR, Richards JH (1973) Ionization behaviour of the histidine residue in the catalytic triad of serine proteases. The Journal of Biological Chemistry 248: 8306-8308.

Huntington JA, Read RJ, Carrell RW (2000) Structure of a serpin–protease complex shows inhibition by deformation. Nature 407: 923-926.

Iengar P and Ramakrishnan C (1999) Knowledge-based modeling of the serine protease triad into non-proteases. Protein Engineering 12: 649-655.

Imperiali B and Abeles RH (1987) Extended binding inhibitors of chymotrypsin that interact with leaving group subsites S1'-S3'. Biochemistry 26: 4474-4477.

Ishida T and Kato S (2004) Role of Asp102 in the catalytic relay system of serine proteases: a theoretical study. Journal of the American Chemical Society 126: 7111-7118.

Ishimoto M and Chrispeels MJ (1996) Protective mechanism of the Mexican bean weevil against high levels of α-amylase inhibitor in the common bean. Plant Physiology 111: 393-401.

Iwanaga S, Yamasaki N, Kimura M, Kouzuma Y (2005) Contribution of conserved Asn residues to the inhibitory activities of Kunitz-type protease inhibitors from plants. Bioscience, Biotechnology and Biochemistry 69: 220-223.

Jadhav DR, Armes NJ, (1996) Comparative status of insecticide resistance in the *Helicoverpa* and *Heliothis* species (Lepidoptera: Noctuidae) of South India. Bulletin Entomological Research 86: 525-531.

Jallow MFA, Cunningham JP, Zalucki MP (2004) Intra-specific variation for host plant use in *Helicoverpa armigera* (Hubner) (Lepidoptera: Noctuidae): implications for management. Crop Protection 23: 955-964.

Jibson MD, Birk Y, Bewley TA (1981) Circular dichroism spectra of trypsin and chymotrypsin complexes with Bowman-Birk or chickpea trypsin inhibitor. International Journal of Peptide and Protein Research 18: 26-32.

Jofuku KD and Goldberg RB (1989) Kunitz trypsin inhibitor genes are differentially expressed during the soybean life cycle and in transformed tobacco plants. The Plant Cell 1: 1079-1093.

Johnson ET, Dowd PF (2004) Differentially enhanced insect resistance, at a cost, in *Arabidopsis thaliana* constitutively expressing a transcription factor of defensive metabolites. Journal of Agricultural Food Chemistry. 52: 5135-5138.

Johnson R, Narvaez J, An G, Ryan CA (1989) Expression of proteinase inhibitors I and II in transgenic tobacco plants: Effects on natural defense against *Manduca sexta* larvae. Proceedings of the National Academy of Sciences, USA 86: 9871-9875.

Jongsma MA, Bakker PL, Peters J, Bosch D, Stiekema WJ (1995) Adaptation of *Spodoptera exigua* larvae to plant proteinase inhibitors by induction of gut proteinase activity insensitive to inhibition. Proceedings of the National Academy of Sciences, USA 92: 8041-8045.

Jongsma MA, Bolter C (1997) The adaptation of insects to plant protease inhibitors. Journal of Insect Physiology. 43: 885-895.

Jouanin L, Bonade-Bottino M, Girard C, Morrot G, Giband M (1998) Transgenic plants for insect resistance. Plant Science. 131: 1–11.

Kaiser ET, Lawrence DS, Rokita SE (1985) The chemical modification of enzyme specificity. Annual Reviews in Biochemistry 54: 565-595.

Karban R (1989) Induced plant responses to herbivory. Annual Reviews in Ecological Systematics. 20: 331-348.

Katz BA, Elrod K, verner E, mackman RL, Luong C, Shrader WD, Sendzik M, Spencer JR, Sprengeler PA, Kolesnikov A, Tai VWF, Hui HC, Breitenbucher JG, Allen D, Janc JW (2003) Elaborate manifold of shirt hydrogen bind arrays mediating binding of active site-directed serine protease inhibitors. Journal of Molecular Biology 329: 93-120.

Keller M, Sneh B, Strizov N, Prudovsky E, Regev A, Koncz C, Schell J, Zilberstein A (1996) Digestion of d-endotoxin by gut proteases may explain reduced sensitivity of advanced instar larvae of *Spodoptera littoralis* to Cry1C. Insect Biochemistry and Molecular Biology 26: 365–373.

Kessler A, Baldwin IT (2002) Plant responses to insect herbivory: the emerging molecular analysis. Annual Reviews in Plant Biology. 53: 299-328.

Komiyama T, VanderLugt B, Fugere M, Day R, Kaufman RJ, Fuller RS (2003) Optimization of protease-inhibitor interactions by randomizing adventitious contacts. Proceedings of the National Academy of Sciences, USA. 100: 8205-8210.

Konno K, Hirayama C, Shinbo H (1997) Glycine in digestive juice; a strategy of herbivorous insects against chemical defense of host plants. Journal of Insect Physiology 43: 217-224.

Krauchenco S, Pando SC, Marangoni S, Polikarpov I (2003) Crystal structure of the Kunitz (STI)-type inhibitor from *Delonix regia* seeds. Biochemical and Biophysical Research Communications 312: 1303-1308.

Kraut J (1977) Serine proteases: structure and mechanism of catalysis. Annual Reviews in Biochemistry 46: 331-358.

Krem MW and Cera ED (2001) Molecular markers of serine protease evolution. The EMBO Journal 20: 3036-3045.

Krowarsch D, Zakrzewska M, Smalas OA, Otlewski J (2005) Structure-function relationships in serine protease-bovine pancreatic trypsin inhibitor interaction. Protein and Peptide Letters 12: 1-5.

Kunitz M (1945) Crystallization of a trypsin inhibitor from soyabean. Science 101: 668-669.

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature. 227: 680–685.

Lajolo FM and Genovese MI (2002) Nutritional significane of lectins and enzyme inhibitors from legumes. Journal of Agricultural and Food Chemistry 50: 6592-6598.

Lara P, Ortego F, Gonzalez-Hidalgo E, Castanera P, Carbonero P, Diaz I (2000) Adaptation of *Spodoptera exigua* (Lepidoptera: Noctuidae) to barley trypsin inhibitor BTI-CMe expressed in transgenic tobacco. Transgenic Research 9: 169–178.

Laskowski M Jr (1986) Protein inhibitors of serine proteinases - mechanism and function. Advances in Experimental and Medical Biology 199: 1-17.

Laskowski M Jr and Kato I (1980) Protein inhibitors of proteinases. Annual Reviews in Biochemitry 49: 593-626.

Laskowski M Jr and Qasim MA (2000) What can the structures of enzyme-inhibitor complexes tell us about the structures of enzyme substrate complexes. Biochimica et Biophysica Acta 1477: 324-337.

Laskowski M Jr., Port SJ, Tashiro M, Whatley HE (1998) Positive Darwinian selection in evolution of protein inhibitors of serine proteinases. Cold Spring Harbor Symposium on Plant Biology 52: 545-553.

Lipke, H, Fraenkel GS, Liener IE (1954) Effects of soybean inhibitors on growth of *Tribolium confusum*. Journal of the Science of Food and Agriculture 2: 410-415.

Lopes AR, Juliano MA, Juliano L, Terra WR (2004) Coevolution of insect trypsins and inhibitors. Archives of Insect Biochemistry and Physiology 55: 140-152.

Lopes AR, Juliano MA, Marana SR, Juliano L, Terra WR (2006) Substrate specificity of insect trypsins and the role of their subsites in catalysis. Insect Biochemistry and Molecular Biology 36: 130-140.

Losey JE, Raynor LS, Carter ME (1999) Transgenic pollen harms monarch larvae. Nature 399: 214-216.

Lu W, Apostol I, Qasim MA, Warne N, Wynn R, Zhang WL, Anderson S, Chiang YW, Ogin E, Rothberg I, Ryan K, Laskowski M Jr. (1997) Binding of amino acid side chains to S1 cavities of serine proteases. Journal of Molecular Biology 266: 441-461.

Luthy JA, Praissmax M, Finkenstadt WR, Laskowski M Jr. (1973) Detailed mechanism of interaction of bovine β-trypsin with soybean trypsin inhibitor (Kunitz) – stopped flow measurements. The Journal of Biological Chemistry 248: 1760-1771.

Ma G, Roberts H, Sarjan M, Featherstone N, Lahnstein J, Akhurst R, Schmidt O (2005) Is the mature endotoxin Cry1Ac from *Bacillus thuringiensis* inactivated by a coagulation reaction in the gut lumen of resistant *Helicoverpa armigera* larvae? Insect Biochemistry and Molecular Biology 35: 729-739.

Manjunath TM, Bhatnagar VS, Pawar CS, Sithanantham S (1989) Economic importance of *Heliothis* spp. in India and an assessment of their natural enemies and host plants. In: Proceedings of the workshop on biological control of *Heliothis*: Increasing the effectiveness of natural enemies. (Eds. King EG and Jackson RD) Far Eastern Regional Research Office, USDA, New Delhi pp. 197-227.

Matthews BW, Sigler PB, Henderson R, Blow DM (1967) Three dimensional structure of tosyl-alpha-chymotrypsin. Nature 214; 652–655.

Mazumdar-Leighton S and Broadway RM (2001a) identification of six chymotrypsin cDNAs from

larval midguts of *Helicoverpa zea* and *Agrotis ipsilon* feeding on the soybean (Kunitz) trypsin inhibitor. Insect Biochemistry and Molecular Biology 31: 633-644.

Mazumdar-Leighton S and Broadway RM (2001b) Transcriptional induction of diverse midgut trypsins in larval *Agrotis ipsilon* and *Helicoverpa zea* feeding on the soybean trypsin inhibitor. Insect Biochemistry and Molecular Biology 31: 645-657.

Mazumdar-Leighton S, Babu CR, Bennett J (2000) Identification of novel serine protease gene transcripts in the midguts of two tropical insect pests, *Scirpophaga incertulas* (Wk.) and *Helicoverpa armigera* (Hb.). Insect Biochemistry and Molecular Biology 30: 57-68.

McLachlan AD (1979). 3-fold structural pattern in the soybean trypsin inhibitor (Kunitz). Journal of Molecular Biology 133, 557-563.

Mekonnen SM, Olufsen M, Smalas AO, Brandsal BO (2006) Predicting proteinase specificities from free energy calculations. Journal of Molecular Graphics and Modelling 25: 176-185.

Mello MO, Tanaka AS, Silva-Filho MC (2003) Molecular evolution of Bowman-Birk type proteinase inhibitors in flowering plants. Molecular Phylogenetics and Evolution 27: 103-112.

Michaud D (1997). Avoiding protease-mediated resistance in herbivorous pests. Trends in Biotechnology 15: 4–6.

Mickel CE and Standish J (1947) Susceptibility of processed soy flour and soy grits in storage to attack by *Tribolium castaneum*. University of Minnesota Agricultural Experimental Station Technical Bulletin 178: 1-20.

Milne RE, Pang ASD, Kaplan H (1995) A protein complex from *Choristoneura fumiferana* gut juice involved in the precipitation of d-endotoxins from *Bacillus thuringiensis* subsp. *sotto*. Insect Biochemistry and Molecular Biology 25: 1101-1114.

Moon J, Salzman RA, Ahn JE, Koiwa H, Zhu-Salzman K (2004) Transcriptional regulation in cowpea bruchid guts during adaptation to a plant defence protease inhibitor. Insect Biochemistry and Molecular Biology 13: 283-291.

Moos M Jr., Nguyen NY, Liu TY (1998) Reproducible high yield sequencing of proteins electrophoretically separated and transferred to an inert support. Journal of Biological Chemistry 263: 6005-6008.

Mukhopadhyay D (2000) The molecular evolutionary history of a winged bean alpha-chymotrypsin inhibitor and modeling of its mutations through structural analyses. Journal of Molecular Evolution 50: 214-223.

Murdock LL, Shade RE (2002) Lectins and protease inhibitors as plant defenses against insects. Journal of Agricultural and Food Chemistry 50: 6605-6611.

Murzin AG, Lesk AM, Chothia C (1992) β-trefoil fold. Journal of Molecular Biology 223: 531-543.

Nagarkatti S and Prakash A (1974) Rearing *Heliothis armigera* (Hubn). on artificial diet. Technical Bulletin 17. Commonwealth Institute of Biological Control, Banglore.

Nahar P, Ghormade V, Deshpande MV (2004) The extracellular constitutive production of chitin deacetylase in *Metarhizium anisopliae*: possible edge to entomopathogenic fungi in the biological control of insect pests. Journal of Invertebrate Pathology 85: 80-88.

Nishihira J and Tachikawa H (1999) Theoretical evaluation of a model of the catalytic triads of serine and cysteine proteases by *an initio* molecular orbital calculation. Journal of Theoretical Biology 196: 513-519.

Oliveira MG, De Simone SG, Xavier LP, Guedes RN (2005) Partial purification and characterization of digestive trypsin-like proteases from the velvet bean caterpillar, Anticarsia gemmatalis. Comparative Biochemistry and Physiology B (Biochemistry and Molecular Biology) 140: 369-380

Onesti S, Brick P, Blow D M (1991) Crystal structure of a Kunitz-type trypsin inhibitor from *Erythrina caffra* seeds. Journal of Molecular Biology 217: 153-176.

Oppert B, Kramar KJ, Beeman RW, Johnson D, McGaughey WH (1997) Proteinase-mediated insect resistance to *Bacillus thuringiensis* toxins. Journal of Biological Chemistry 272: 23473-23476.

Otlewski J, Jaskolski M, Buczek O, Cierpicki T, Czapinska H, Krowarsch D, Smalas AO, Stachowiak D, Szpineta A, Dadlez M (2001) Structure-function relationship of serine protease-protein inhibitor interaction. Acta Biochimica Polonica 48: 419-428.

Page D, Auburt G, Duc G, Welham T, Domoney C (2002) Combinatorial variation in coding and promoter sequences of genes at the *Tri* locus in *Pisum sativum* accounts for variation in trypsin inhibitor activity in seeds. Molecular Genetics and Gneomics 267: 359-369.

Patankar AG, Giri AP, Harsulkar AM, Sainani MN, Deshpande VV, Ranjekar PK, Gupta VS (2001) Complexity in specificities and expression of *Helicoverpa armigera* gut proteinases explains polyphagous nature of the insect pest. Insect Biochemistry and Molecular Biology 31: 453-464.

Patankar AG, Harsulkar AM, Giri AP, Gupta, VS, Sainani M.N, Ranjekar PK, Deshpande VV (1999) Diversity in inhibitors of trypsin and *Helicoverpa armigera* gut proteinases in chickpea (*Cicer arietinum*) and it's wild relatives. Theoretical and Applied Genetics 99: 719-726.

Paulillo LCMS, Lopes AR, Cristofoletti PT, Parra JRP, Terra WR, Silva-Filho MC (2000) Changes in midgut endopeptidase activity of *Spodoptera frugiperda* (Lepidoptera: Noctuidae) are responsible for adaptation to soybean proteinase inhibitors. Journal of Econ. Entomology 93: 892–896.

Perkins, D.N., Pappin, D.J., Creasy, D.M. Cottrell, J.S. 1999. Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis 20: 3551-3567.

Pichare MM, Kachole MS (1994) Detection of electrophoretically separated proteinase inhibitors using X-ray film. Journal of Biochemical and Biophysical Methods 28: 215–224.

Plotnick MI, Mayne L, Schechter NM, Harvey R (1996) Distortion of the active site of chymotrypsin complexed with a serpin. Biochemistry 35: 7586-7590.

Polgar L (2005) The catalytic triad of serine peptidases. Cellular and Molecular Life Sciences 62: 1-12.

Presnell SR, Patil GS, Mura C, Jude KM, Conley JM, Bertrand JA, Kam CM, Powers JC, Williams LD (1998) Oxyanion-mediated inhibition of serine proteases. Biochemistry 37: 17068-17081.

Qasim MA, Ranjbar MR, Wynn R, Anderson S, Laskowski M Jr (1995) Ionizable P1 residues in serine proteinase inhibitors undergo large pK shifts on complex formation. The Journal of Biological Chemistry 270: 27419-27422.

Radisky ES, Lu CJK, Kwan G, Koshland DE Jr. (2005) Role of the intramolecular hydrogen bond network in the inhibitory power of chymotrypsin inhibitor 2. Biochemistry 44: 6823-6830.

Ravichandran S, Dasgupta J, Chakrabarti C, Ghosh S, Singh M, Dattagupta JK (2001) The role of Asn14 in the stability and conformation of the reactive-site loop of winged bean chymotrypsin inhibitor: crystal structures of two point mutations Asn14Lys and Asn14Asp. Protein Engineering 14: 349-357.

Reech GR, Kramer KJ, Baker JE, Kanost MR, Fabrick JA, Behnke CA (1997) Proteinase inhibitors and resistance of transgenic plants to insects. In: Carozzi N and Koziel M, eds. Advances in insect control: the role of transgenic plants. London, Taylor and Francis, p. 157-183.

Richardson MJ (1991) Seed storage proteins: The enzyme inhibitors. In: Richardson MJ, ed. Methods in Plant Biochemistry, New York, Academic Press, p. 259-305.

Romeis J, Sharma HC, Sharma KK, Das S, Sarmah BK (2004) The potential of transgenic chickpeas for pest control and possible effects on non-target arthropods. Crop Protection 23: 923-938.

Ryan CA (1990) Proteinase inhibitors in plants: genes for improving defenses against insects and pathogens. Annual Review of Phytopathology 28: 425–449.

Sales MP, Gerhardt IR, Grossi-de-Sa MF, Xavier-Filho J (2000) Do legume storage proteins play a role in defending seeds against bruchids? Plant Physiology 124: 515-522.

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning: A Laboratory Manual, second edition, Cold Spring Harbor Laboratory Press, Plainview, New York.

Saxena D, Flores S, Stotzky G (1999) Transgenic plants: Insecticidal toxin in root exudates from Bt corn. Nature 402: 480

Schechter I and Berger A (1967) Biochemical and Biophysical Research Communication 27: 157–162.

Schellenberger V, Braune K, Hofmann HJ, Jakubke HD (1991) The specificity of chymotrypsin. A statistical analysis of hydrolysis data. European Journal of Biochemistry 199: 623-636.

Schmidt AE, Ogawa T, Gailani D, Bajaj SP (2004) Structural role of Gly193 in serine proteases. The Journal of Biological Chemistry 279: 29485-29492.

Schuler TH, Poppy GM, Kerry BR, Denholm I (1998) Insect resistant transgenic plants. Trends in Biotechnology 16: 168–175.

Shamaladevi N and Pattabhi V (2005) Secondary binding site of trypsin: revealed by crystal structure of trypsin-peptide complex. Journal of Biomolecular Structure and Dynamics 22: 1-8.

Sichler S, Hopfner KP, Kopetzki E, Huber R, Bode W, Brandstetter H (2002) The influence of residue 190 in the S1 site of trypsin-like serine proteases on substrate selectivity is universally conserved. FEBS Letters 530: 220-224.

Simmonds MSJ (2003) Flavonoid-insect interactions: recent advances in our knowledge. Phytochemistry 64: 21-30.

Sin SF, Yeung EC, Chye ML (2006) Downregulation of *Solanum americanum* genes encoding proteinase inhibitor II causes defective seed development. The Plant Journal 45: 58-70.

Sinha S, Watorek W, Karr S, Giles J, Bode W, Travis J (1987) Primary structure of human neutrophil elastase. Proceedings of the National Academy of Sciences, USA 84: 2228-2232.

Smirnoff P, Khalef S, Birk Y, Applebaum SW (1976) A trypsin and chymotrypsin inhibitor from chickpeas (*Cicer arietinum*). Biochemical Journal 157: 745-751.

Solivan S, Selwood T, Wang ZM, Schechter NM (2002) Evidence for diversity of substrate specificity among members of the chymase family of serine proteases. FEBS Letters 512: 133-138.

Song J and Markley JL (2003) Protein inhibitors of serine proteinases: role of backbone structure and dynamics in controlling the hydrolysis constant. Biochemistry 42: 5186-5194.

Soon LG (1997) Integrated pest management in developing Asia. In Proceedings of the International Conference on Ecology and Agriculture: Towards Sustainable Development (Eds. Dhaliwal GS, Randhawa NS, Arora R, Dhawan AK) Chandigarh, India pp 3-16.

Srinivasan A, Chougule NP, Giri AP, Gupta VS (2005b) Podborer (*Helicoverpa armigera* Hubn.) does not show specific adaptations in gut proteinases to dietary *Cicer arietinum* Kunitz proteinase inhibitor. Journal of Insect Physiology 51: 1268-1276.

Srinivasan A, Giri AP, Gupta VS (2006) Structural and functional diversities in Lepidopteran serine proteases. (submitted to Cellular and Molecular Biology Letters)

Srinivasan A, Giri AP, Harsulkar AM, Gatehouse JA, Gupta VS (2005a) A Kunitz trypsin inhibitor from chickpea (*Cicer arietinum* L.) that exerts anti-metabolic effect on podborer (*Helicoverpa armigera*) larvae. Plant Molecular Biology 57: 359-374.

Srinivasan A, Harsulkar AM, Giri AP, Gupta VS (2006) Genetic structure and differential expression of the *Cicer arietinum* Kunitz proteinase inhibitor. (manuscript under preparation).

Stotz HU, Kroymann J, Mitchell-Olds T (1999) Plant-Insect interactions. Current Opinion in Plant Biology. 2: 268-272.

Stubbs MT, Huber R, Bode W (1995) Crystal structures of factor Xa specific inhibitors in complex with trypsin: structural grounds for inhibition of factor Xa and selectivity against thrombin. FEBS Letters. 375: 103-107.

Tamhane VA, Chougule NP, Giri AP, Dixit AR, Sainani MN, Gupta VS (2005) *In vivo* and *in vitro* effect of *Capsicum annum* proteinase inhibitors on *Helicoverpa armigera* gut proteinases. Biochimica et Biophysica Acta 1722; 156-167.

Tanaka T and Yada RY (2004) Redesign of catalytic center of an enzyme: aspartic to serine proteinase. Biochemical and Biophysical Research Communications 323: 947-953.

Tarrago T, Martinez I, Torrent M, Codina A, Giralt E, Puigdomenech P, Ludevid D (2003) The *fuc1* gene product (20 kDa FUC1) of *Pisum sativum* has no α-L-fucosidase activity. Plant Molecular Biology 51: 877-884.

Telang A, Booton V, Chapman RF, Wheeler DE (2000) How female caterpillars accumulate their nutrient reserves. Journal of Insect Physiology 47: 1055–1064.

Telang M, Srinivasan A, Patankar AG, Harsulkar AM, Joshi V, Damle A, Deshpande VV, Sainani MN, Ranjekar PK, Gupta GP, Birah A, Rani S, Kachole M, Giri AP, Gupta VS (2003) Bitter gourd proteinase inhibitors: potential growth inhibitors of *Helicoverpa armigera* and *Spodoptera litura*. Phytochemistry 63: 643-652.

Telang MA, Giri AP, Sainani MN, Gupta VS (2005) Elastase-like proteinase of *Helicoverpa armigera* is responsible for inactivation of a proteinase inhibitor from chickpea. Journal of Insect Physiology 51: 513-522.

Tellam RL, Wijfells G, Willadsen P (1999) Peritrophic matrix proteins. Insect Biochemistry and Molecular Biology 29: 87-101.

Terra WR (2001) The origins and functions of the insect peritrophic membrane and peritrophic gel. Archives of Insect Biochemistry and Physiology 47: 47-61.

Thomas MB (1999) Ecological approaches and the development of 'truly integrated' pest management. Proceedings of the National Academy of Sciences, USA 96: 5944-5951

Ussuf KK, Laxmi NH, Mitra R (2001) Proteinase inhibitors: Plant-derived genes of insecticidal protein for developing insect-resistant transgenic plants. Current Science 80: 847–853.

Vago C, Kurstak E. (1965) Mechanism of action of *Bacillus thuringiensis* introduced by a hymenopteran parasite into the hemolymph of a lepidopteron. Antonie Van Leeuwenhoek. 313: 282-294.

Valaitis AP, Augustin S, Clancy KM (1999) Purification and characterization of the western spruce budworm larval midgut proteinases and comparison of gut activities of laboratory-reared and field-collected insects. Insect Biochemistry and Molecular Biology 29: 405–415.

Varallyay E, Lengyel Z, Graf L, Szilagyi L (1997) the role of disulfide bond C191-C220 in trypsin and chymotrypsin. Biochemical and Biophysical Research Communications 230: 592-596.

Volpicella M, Ceci LR, Cordewener J, America T, Gallerani R, Bode W, Jongsma MA, Beekwilder J (2003) Properties of purified gut trypsin from *Helicoverpa zea* adapted to proteinase inhibitors. European Journal of Biochemistry 270: 10-19.

Volpicella M, Cordewener J, Jongsma MA, Gallerani R, Ceci LR, Beekwilder J (2005) Identification and characterization of digestive serine proteases from inhibitor-resistant *Helicoverpa zea* larval midgut. Journal of Chromatography B (in press)

Wang KJ, Yamashita T, Watanabe M, Takahata Y (2004) Genetic characterization of a novel *Tib*-derived variant of soybean Kunitz trypsin inhibitor detected in wild soybean (*Glycine soja*). Genome 47: 9-14.

Wang P and Granados RR (2001) Molecular structure of the peritrophic membrane (PM): identification of potential PM target sites for insect control. Archives of Insect Biochemistry and Physiology 47: 110-118.

Wang T, Zhang N, Du L (2005) Isolation of RNA of high quality and yield from *Ginkgo biloba* leaves. Biotechnology Letters 27: 629-633.

Waterhouse (1957) Digestion in insects Annual Review of Entomology 2: 1-18.

Wightman JA, Anders MM, Rao VR, Reddy LM (1995) Management of *Helicoverpa armigera* (Lepidoptera: Noctuidae) on chickpea in southern India: thresholds and the economics of host plant resistance and insecticide application. Crop Protection 14: 37-46.

Zhang JH, Wang CZ, Qin JD (2000) the interactions between soybean trypsin inhibitor and δ-endotoxin of *Bacillus thuringiensis* in *Helicoverpa armigera* larva. Journal of Invertebrate Pathology 75: 259-266.

Zhu Y, Kramer KJ, Dowdy AK, Baker JE (2000a) Trypsinogen-like cDNAs and quantitative analysis of mRNA levels from the indianmeal moth, *Plodia interpunctella*. Insect Biochemistry and Molecular Biology 30: 1027–1035.

Zhu Y, Kramer KJ, Oppert B, Dowdy AK (2000c) cDNAs of aminopeptidase-like protein genes from *Plodia interpunctella* strains with different susceptibilities to *Bacillus thuringiensis* toxins. Insect Biochemistry and Molecular Biology 30: 215–224.

Zhu Y, Oppert B, Kramer KJ, McGaughey WH, Dowdy AK (1997) cDNAs for a chymotrypsinogen-like protein from two strains of *Plodia interpunctella*. Insect Biochemistry and Moleular Biology 27: 1027-1037.

Zhu Y, Oppert B, Kramer KJ, McGaughey WH, Dowdy AK (2000b) cDNA sequence, mRNA expression and genomic DNA of trypsinogen from the indianmeal moth *Plodia interpunctella*. Insect Molecular Biology 9: 19-26.

Zhu-Salzman K, Koiwa H, Salzman RA, Shade RE, Ahn JE (2003) Cowpea bruchid *Callosobruchus maculatus* uses a three-component strategy to overcome a plant defensive cysteine protease inhibitor. Insect Biochemistry and Molecular Biology 12: 135-45.

Zupunski V, Kordis D, Gubensek F (2003) Adaptive evolution in the snake venom Kunitz/BPTI protein family. FEBS Letters 547: 131-136.