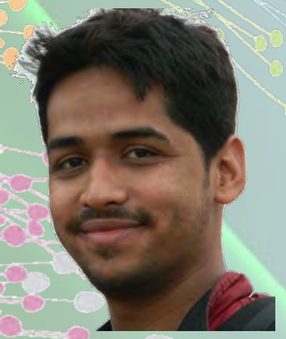


Rakesh Shamsunder Joshi
Senior Research Fellow
Division of Biochemical Sciences
CSIR-National Chemical Laboratory
Pune 411008, (MH) India
rakeshjoshi687@gmail.com

Dr. Ashok P. Giri
Senior Scientist
Division of Biochemical Sciences
CSIR-National Chemical Laboratory
Pune 411008, (MH) India
ap.giri@ncl.res.in

Council of Scientific and Industrial Research
National Chemical Laboratory

Ph.D. THESIS

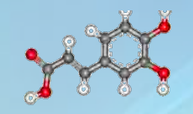
RAKESH SHAMSUNDER JOSHI

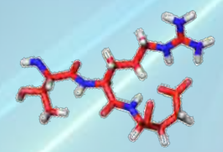
SEPTEMBER 2014

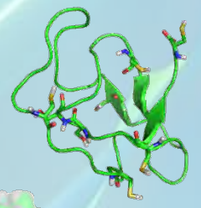
Tailoring Potent Plant Protease Inhibitor against *Helicoverpa armigera* Proteases

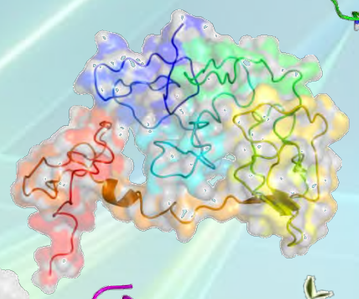
Rakesh Shamsunder Joshi

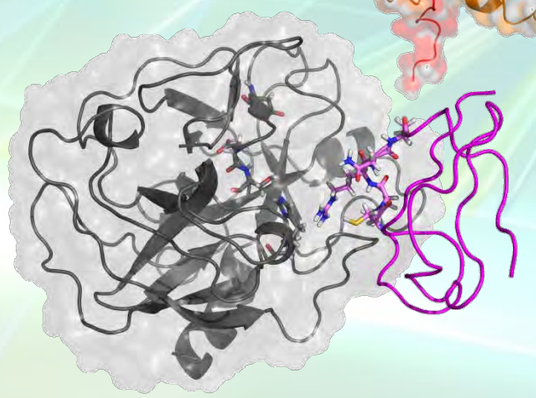
Research Supervisors
Dr. Ashok P. Giri
Dr. C. G. Suresh
Dr. Vidya S. Gupta











TAILORING POTENT PLANT PROTEASE INHIBITOR AGAINST *HELICOVERPA ARMIGERA* PROTEASES

A THESIS SUBMITTED TO THE

SAVITRIBAI PHULE PUNE UNIVERSITY

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

BIOTECHNOLOGY

BY

Mr. RAKESH SHAMSUNDER JOSHI

UNDER THE GUIDANCE OF

DR. ASHOK P. GIRI

DIVISION OF BIOCHEMICAL SCIENCES

CSIR-NATIONAL CHEMICAL LABORATORY

PUNE-411008, INDIA

SEPTEMBER 2014

राष्ट्रीय रासायनिक प्रयोगशाला
(वैज्ञानिक तथा औद्योगिक अनुसंधान परिषद)
डॉ. होमी भाभा रोड, पुणे - 411 008. भारत
NATIONAL CHEMICAL LABORATORY
(Council of Scientific & Industrial Research)
Dr. Homi Bhabha Road, Pune - 411008. India

Date: 15th September 2014

CERTIFICATE

Certified that the work in the Ph.D. thesis entitled '**Tailoring Potent Plant Protease Inhibitor against *Helicoverpa armigera* Proteases'** submitted by **Mr. Rakesh Shamsunder Joshi** was carried out by the candidate under my supervision. The material obtained from other sources has been duly acknowledged in the thesis.

Date:
Place: Pune

Dr. Ashok P. Giri
(Research Guide)

Dr. C. G. Suresh
(Co-guide)

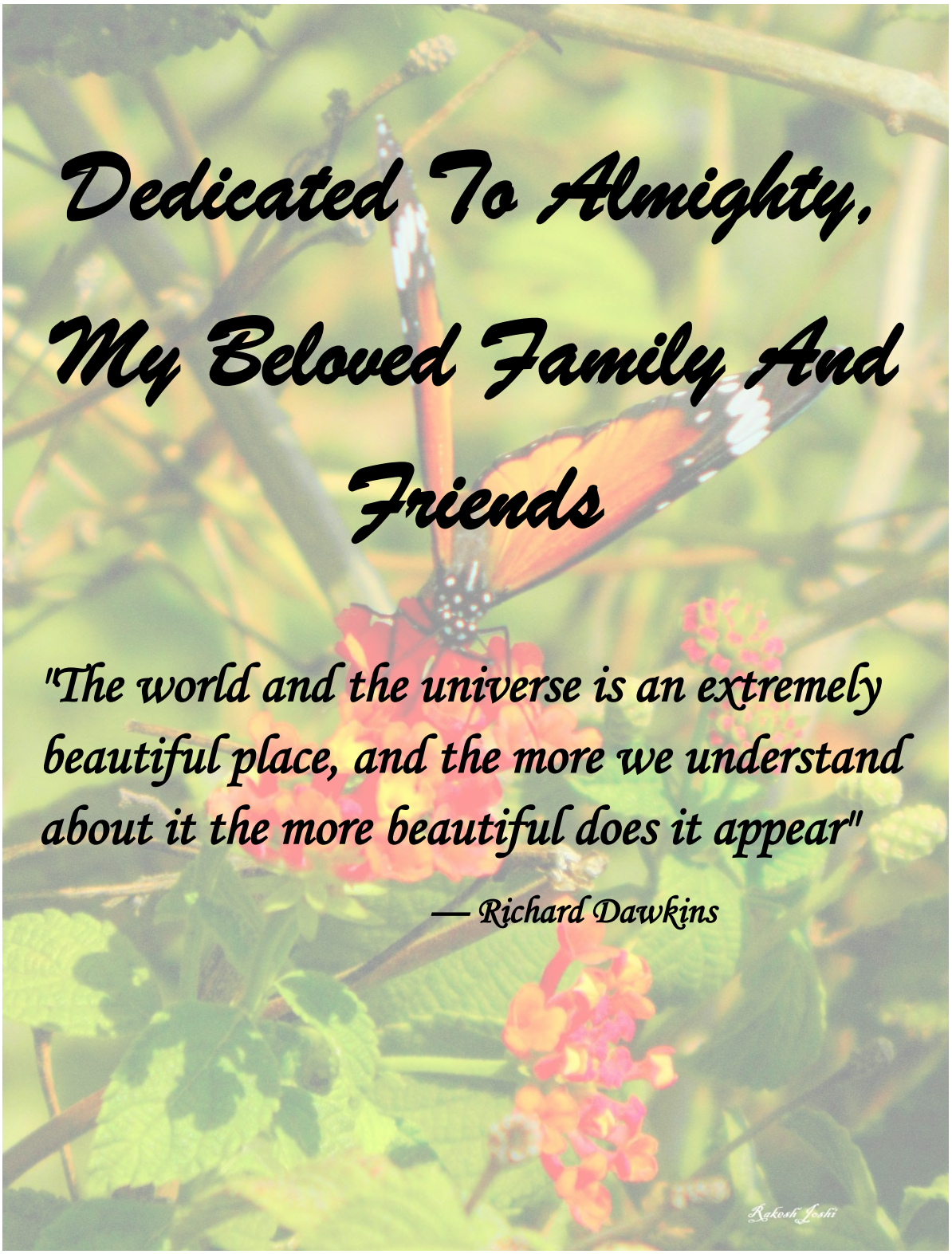
Communications Channels ☎ +91 20 25902000 +91 20 25893300 +91 20 25893400 Fax +91 20 25902601 (Director) +91 20 25902660 (Admin.) +91 20 25902639 (Business Development) URL : www.ncl-india.org

DECLARATION BY THE CANDIDATE

I hereby declare that the thesis entitled, **"Tailoring potent plant protease inhibitor against *Helicoverpa armigera* proteases"** submitted by me for the degree of Doctor of Philosophy to the Savitribai Phule Pune University, is the record of work carried by me at Biochemical Sciences Division, CSIR-National Chemical Laboratory, Pune - 411008, Maharashtra, India, under the supervision of Dr. Ashok P. Giri (research guide) and Dr. C. G. Suresh (research co-guide). The work is original and has not formed the basis for the award of any degree, diploma, associateship, and fellowship titles in this or any other university or other institute of higher learning. I further declare that the material obtained from other resources has been duly acknowledged in the thesis.

Rakesh Shamsunder Joshi
Senior research Fellow
Biochemical Sciences Division
CSIR-National Chemical Laboratory
Pune - 411008, India

Place: Pune
Date: 15th September 2014


Dedicated To Almighty,
My Beloved Family And
Friends
"The world and the universe is an extremely beautiful place, and the more we understand about it the more beautiful does it appear"
— Richard Dawkins
Rakesh Joshi

Contents

Acknowledgements

The joy of completion of my Ph.D thesis is to flick through this beautiful journey and remember all the friends and family who have supported and helped me along this long road. At the end of my thesis, it is a pleasant task to express my thanks to all those who have contributed by many ways for the success of this study and made it an unforgettable experience for me.

This thesis was made possible by the immense and much appreciated help and support of a number of people and I take this opportunity to thank them. First of all, I would like to express my sincere gratitude to my supervisor, Dr. Ashok Giri for his continuous support, encouragement, guidance and advice throughout the course of my doctoral candidature. He has been a true friend and mentor throughout my candidature. He always comes up with new ideas and will help you to solve that problem. But at the same time he will never force his ideas upon you and always welcome and appreciate your thoughts. His enthusiasm has always inspired me and I thank him for helping me in developing a good research plan and its efficient execution. I am especially indebted to my Co-supervisors, Dr. C. G. Suresh and Dr. Vidya Gupta, their punctuality, enthusiasm, and devotion to work sets an ideal example before me, which has always given me a boost for carrying out the Ph.D. work. The joy and enthusiasm they have for their research was contagious and motivational for me. I really thank my supervisors for their support during my good and bad times.

I am extremely grateful to, Dr. Kiran Kulkarni, for his valuable guidance, scholarly inputs and consistent encouragement I received throughout the research work. I highly appreciate the constructive suggestions and input offered by Dr. Rajendra Joshi, and Dr. Uddhavesh Sonavane, CDAC, Pune in carrying out simulation studies. I express my heartfelt gratitude to Prof. K. V. R. Chary, TIFR, Mumbai for providing his expertise and guidance in protein structure elucidation by Nuclear magnetic resonance. I would also like to acknowledge the exceptional teaching and motivation provided by Dr. Rajesh Kumar Singh (AGM, Gennova Biopharmaceuticals Ltd, Pune) to pursue research career. The valuable and critical inputs in the synthesis work by Dr. H. V. Thulasiram and Dr. Moneesha Fernandes are highly acknowledged. I would like to thank Dr. Mahesh J. Kulkarni, Dr. Narendra Kadoo and Dr. Sushma Gaikwad for their all timely support throughout my tenure.

I feel happy to express my thanks to Dr. Sourav Pal, Director, CSIR-NCL for providing me the research facilities. I must thank the office staff of Biochemical Sciences Division and Students Academic office, which was always ready to help whenever required. I thank UGC, India for supporting my research work through Research fellowships.

This work would never be complete without expressing my gratitude to Dr. Manasi Mishra, Dr. Purushottam Lomate, Dr. Vishal Dawkar and Dr. Vaijayanti Tamhane for their contribution

in developing my PhD objectives and for her suggestions and constant encouragement throughout the work. This work would never be complete without expressing my gratitude to Aarohi, Anirban, Fayaj, Trushna, Rajan, Nidhi, Neeru, Preeti, Mehul, Ranjit and Saikat for their valuable input and constant support. It would not have been possible for me to accomplish my thesis without them. I express my heartfelt thanks to Rahul who was my first friend in the lab and supported me from the first day in every aspect. I am thankful to Kedar, Sachin, Tejas and Yashwant, who made my stay in lab as well as in the hostel joyous. I would like to thank Amey, Atul, Anand, Ashwini B., Amit, Neha M., Neha K., Priyanka S., Sheon, Sharmila, Yojana and Zenia,for always being helpful. I thank Aditi, Amol K., Amol J., Ajit, Ashish,Ashwini R.,Deepika, Hemangi, Gouri Pathak, Gayatri, Hemalata, Medha, Priyanka J., Pranjali, Rakesh P., Ram, Ramesha, Rasika, Ramya, Saleem, Sonal, Smiriti, Sucheta, Uma, Varsha, Yamini and all other labmates for their timely help. I reserve special thanks to Indira mam, Sheetal and Yashshree who continue to manage the scientific requirements of the group. I would like to thanks all supporting staff including Jagtap kaka, Ramakant, Giri and Trehan.

I owe my thanks to Sneha, who have always stood by my side through good times and rough patches. I am thankful to Anupama and Vitthal for providing me homely environment. I take this opportunity to acknowledge my friends Aarti, Ajay, Ashish, Anuja, Bhavin, Dipesh, Harshal, Gajendra, Gouri, Pradeep, Pratik, Rubina, Rashmi, Mahesh, Mahendra, Rupesh, Rohan, Shweta, Vijay, Vikrant and Yogesh for their support. I am also thankful to Dr. Bhushan Dholakia, Dr. Arvinda Korwar, Dr. Yogesh Kolekar, Sandip, Deepak, Priyabrata, Manu S., Ranu, Tulika, Manas, Prabhakar and Devdatta for their help.

Last but not least, I would like to pay high regards to my family. There are no words to acknowledge my family for their constant support and encouragement. I owe my deepest gratitude to Maa, Pappa, Harish bhai, Rani tai, Aarti tai, Dadi and the list is endless, for their support and love which made me to reach up to this level.

Besides this, several people have knowingly and unknowingly helped me in the successful completion of this project. I thank all of them for their assistance.

Rakesh

We must, however, acknowledge, as it seems to me, that man with all his noble qualities, still bears in his bodily frame the indelible stamp of his lowly origin.....Charles Darwin

Abbreviations

AA	Amino Acids
ACN	Acetonitrile
BA*p*NA	Benzoyl-DL-Arginyl-P-Nitroanilide
BMMY	Buffered Methanol-Complex Medium
Bt	Bacillus Thuringiensis
CA	Caffeic Acid
CanPI	Capsicum Annuum Proteinase Inhibitor
CBB-R250	Coomasssie Brilliant Blue-R250
CD	Circular Dichroism
cDNA	Complementary Deoxyribonucleic Acid
CI	Chymotrypsin Inhibitor
DNA	Deoxyribonucleic Acid
DTT	Dithiothreitol
g, mg, µg, ng	Gram, Milligram, Microgram, Nanogram
GXCT	Gel-X Ray Film Contact-Print Technique
HCl	Hydrochloric Acdi
HGPs	*Helicoverpa armigera* Gut Proteinases
HIC	Hydrophobic Interaction Chromatorgraphy
IC_{50}	Inhibitor Concentration At 50% Inhibition
IMAC	Immobilized Metal Ion Affinity Chromatorgraphy
IPTG	Isopropyl B-D-1-Thiogalactopyranoside
IRD	Inhibitory Repeat Domain
kDa/kD	Kilo Dalton
K_i	Inhibition Constant
L, mL, µL	Liter, Milliliter, Microliter
M, mM, µM	Molar, Millimolar, Micromolar

MALDI-TOF-MS	Matrix Assisted Laser Desorption Ionization – Time of Flight Mass Spectrometry
MGY	Minimal Glycerol Medium
MM	Minimal Methanol Medium
MRE	Mean Residual Ellipticity
NaCl	Sodium Chloride
NaOH	Sodium Hydroxide
NaProPI	*Nicotiana alata* Pin-II Precursor Form
NaTI	*N. alata* Trypsin Inhibitor
NMR	Nuclear Magnetic Resonance
OD	Optical Density
PAGE	Poly-Acrylamide Gel Electrophoresis
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
PI	Proteinase Inhibitor
PIN-I/II	Potata Type Inihibitor I And II
rCanPI	Recombinant *C. annuum* Proteinase Inhibitor
RDB	Regeneration
RNA	Ribonucleic Acid
rpm	Revolutions Per Minute
RSL	Reactive Site Loop
RT-PCR	Reverse Transcriptase-Polymerase Chain Reaction
SDS	Sodium Dodecyl Sulfate
SAX	Small Angle X-ray Scattering
TFA	Trifluoroacetic Acid
TI	Trypsin Inhibitor
TIU	Trypsin Inhibitory Unit
TRIS	Tris-Hydroxymethyl Aminomethane
2-D	2-Dimensional Gel Electrophoresis

List of Figures

List of Tables

Thesis Summary

Chapter I: Introduction

Ecosystem provides an opportunity to various organisms to interact and adapt. Present scenario of an ecosystem is perhaps the manifestation of only success stories within organismal relations and their environmental interactions. Co-existence of plants and insects is most primitive and co-evolved interaction, constantly influenced by surrounding changes. Plants and insects have always maintained their co-existence through mutually benefiting from interacting partners. However, excess exploitation of services or resources without reasonable reward led to the development of morphological, physiological and molecular defense strategies in both the systems. An attack from a pest on crop plants leads to severe economic losses. Many lepidopteran larvae are herbivores and major pests in agriculture. Plant protects them from herbivore attack; plants use constitutive and induced defenses.

Amongst different lepidopteran pests, *Helicoverpa armigera* (Lepidopteran) is a serious polyphagous pest that affects multiple crops like cotton, pulses, vegetables, maize, tobacco, soybean, rapeseed, safflower, sunflower, sorghum, forest trees, and fruits (Tamhane *et al*., 2005a, Damle 2005). These insects comprises of many serine protease isoforms especially of trypsin and chymotrypsin's with varied reactivity and specificity (Srinivasan *et al*., 2006). This enormous gut protease diversity is key factor for insects to survive and adapt on a range of host plants (Giri *et al*. 1998; Patankar *et al*. 2001; Chougule *et al*. 2005). Insects have also adapted to host plant's protease inhibitor (PI) based defense strategy by: (i) the over-expression of proteases; (ii) the synthesis of proteases that are insensitive to inhibitors (Jongsma *et al*. 1995) or inhibitor-resistant proteases (Broadway 1996); and (iii) the production of proteases that possess the capacity to degrade PIs (Giri *et al*. 1998). Dynamic regulation of protease expression in response to ingested PIs or alteration in food content might be the key regulation factor for adaptation of this

insect (Patankar *et al*., 2001; Lomate and Hivrale 2011; Hivrale *et al*. 2013; Mahajan *et al*. 2013; Chikate *et al*. 2013).

Protease inhibitors (PI) are one of the most well studied plant defense proteins. PIs are abundantly present in leaves, flowers, seeds and tubers and are induced/up-regulated in response to herbivore attack (Green and Ryan, 1972; Ryan 1990). PIs function by blocking the digestive proteases in the larval gut thereby limiting the release of amino acids from dietary proteins (Hilder and Boulter, 1999; Tamhane *et al*., 2005b), thus reducing overall fitness of the insects. Moreover, PIs interfere with proteolytic activation of enzymes as well as molting and water-balance of the insects (Boulter, 1993) and hence lead to defect in the development and metamorphosis. Among the various classes of protease inhibitor, potato type II protease inhibitors (Pin-II) inhibitor is one of predominant class. In general, the Pin-IIs are serine protease inhibitors, which are wound-inducible and play a role in plant defense against insect pests (Green and Ryan, 1972). Pin-II genes are predominantly found in solanaceae family. The Pin-II proteins have few conserved characters *viz*. a signal peptide (≈ 25 aa) followed by inhibitory repeat domains 1 to 8 in precursors. Each repeat domain (≈55 aa) is connected to other by linkers (≈5aa) and also have a conserved structure having 8 cysteine residues forming 4 disulfide bridges which confer resistance to heat, extreme pH, proteolysis and stabilize the repeat structure. Each repeat domain has an active site either for trypsin or chymotrypsin inhibition depending upon the presence of arginine/lysine or leucine at the P1 position of the reactive site. Reactive sites and the linker regions of Pin-II inhibitors bare the signatures of having been under positive selection pressure indicating evolutionary significance of their role in plant protection (Kong and Rangnathan, 2008). Extensive screening of several non-host plants for potential insect gut PIs has resulted in the identification of pepper (*Capsicum annuum*) as good sources of diverse pool of plant PIs. The efficacy of *C. annuum* protease inhibitors (*Can*PIs*)* on *H. armigera* gut proteases has been reported (Chougule *et al*., 2005; Tamhane *et al*., 2005b). *C. annuum* predominantly express Potato Type II inhibitor (Pin-II). *C. annuum* have 67 novel *Can*PIs which vary in the number of repeats per gene (1 to 4 IRDs) and variations in the IRDs (55 unique IRDs) were

identified (Tamhane *et al*., 2009; Mishra *et al*., 2013). It has also been established that these genes show differential expression patterns spatially and temporally (Tamhane *et al*., 2009; Mishra *et al*., 2012).

There are few reports, which depicted the structural information about protease-PI interaction. Barrette *et al*., in 2003 showed the structure of multidomain tomato inhibitor and elucidated the molecular framework of interaction between reactive sites of inhibitors with active site of proteases. Mishra *et al*., (2013) showed the interaction of 4-domain CanPI-7 with proteases. It is observed that spatial orientation and distribution of CanPI-7 reactive loop is key factor in its reactivity against proteases. In 2010, Schirra *et al*., had elucidated NMR structures of disulfide bond mutants of *Nicotiana alata* Pin-II inhibitor and provided information about indispensability of disulfide bonds in the structure and function of these inhibitors. The selective gain or loss of disulfide bond(s) in SPIs can be associated with their functional differentiation, which is likely to be compensated by non-covalent interactions (hydrogen bonding or electrostatic interactions). Thus, these intramolecular interactions are collectively responsible for the functional activity of Pin-IIs, through the maintenance of scaffold framework, conformational rigidity and shape complementarities of reactive site loop (Li *et al*., 2011). Reactive loop variation among all the Pin-II PIs might provide us diverse pool of variations, to design synthesizes and assess their potential against insect proteases. Mystifying structural and functional diversity amongst these Pin-II PIs and to analyze their interactions with gut protease of *H. armigera* is challenging.

To overcome the limitation of existing pest management strategies and to ensure a continued supply of high–quality food, efforts are going on to utilize synthetic peptides and other small molecules to develop new insect–control agents. Different approaches of structure-based screening and designing are used to develop novel arena of the insect protease inhibitors. Production of secondary metabolites in response to insect herbivory is one of the defense strategies of plants. Insect growth inhibitory effect of plant secondary metabolites exerts weak selection pressure and thus avoids the chances of emergence of resistance against these compounds (Boerjan *et al*. 2003). These compounds exhibit dynamic expression and

distribution pattern in response to insect attack. (Maffei *et al*. 2007, Usha Rani and Jyothsna 2010; Ballhorn *et al*. 2011). Pivotal role of natural phenols in plant–environment interactions is evidenced from plants defense responses against biotic and abiotic stresses, including ultraviolet radiation, pathogen infection and herbivore damage. Plant phenolic compounds are categorized at polyphenols and simple phenolic compound. Polyphenols are subdivided as tannins and flavonoids, while Simple phenolic are further categorized in phenolic acid and coumarins. In phenolic acid, Hydroxybenzoic acid and Hydroxycinnamic acid (HCA) contributes to population of compounds (Magnani *et al*., 2014). HCA serve as precursor for synthesis of plant phenolic acid compounds. Recent reports showed that plant phenolic compound leads to reduction in palatability of the plant tissues to the herbivore and thus results in their growth retardation (Berenbaum 1995, Boerjan *et al*. 2003, Kubo 2006, Appel 1993, Mallikarjuna *et al*. 2004; War *et al*. 2012).

Chapter II: Characterization of inhibitory repeat domain (IRDs) of *Capsicum annuum* protease inhibitors

C. annuum expresses diverse Pin-II PIs comprising of IRD as basic functional unit. Most IRDs contain eight conserved cysteine forming four disulfide bonds, which are indispensible for their stability and activity. We investigated the functional significance of evolutionary variations in IRDs and their role in mediating interaction between the inhibitor and cognate protease. Among the 67 IRDs encoded by *C. annuum*, IRD-7, -9, and -12 were selected for further characterization on the basis of variation in their reactive site loop, number of conserved cysteine residues, and higher theoretical ΔG_{bind} for interaction with *H. armigera* trypsin. Moreover, inhibition kinetics showed that IRD-9, despite loss of two disulfide bonds, was a more potent among the three selected IRDs. Molecular dynamic simulations revealed that serine residues in the place of cysteines at seventh and eighth positions of IRD-9 resulted in an increase in the density of intramolecular hydrogen bonds and reactive site loop flexibility. Reduction in IRD-9 inhibitory activity after chemical modification of the serine residues also supported this observation and provided a possible explanation for the remarkable inhibitory potential of IRD-9.

IRD-9 showed special attributes like stability to proteolysis and synergistic inhibitory effect on other IRDs. Furthermore, *H. armigera* fed on artificial diet containing 5 TIU/g of recombinant IRD proteins exhibited negative impact on larval growth, survival rate and other nutritional parameters. Major digestive gut trypsin and chymotrypsin genes were down regulated in the IRD fed larvae, while few of them were up-regulated, this indicate alterations in insect digestive physiology. The results corroborated with protease activity assays and zymography. Altogether, these findings suggest that the sequence variations among IRDs reflect in their efficacy against proteases *in vitro* and *in vivo*, which also could be used for developing tailor-made multi-domain inhibitor gene(s).

Chapter III: Structural and functional insights of single and multi-domain CanPIs and their interactions with proteases

In order to study the structural basis of insect protease inhibition by a single and multi-domain CanPI inhibitor, we have expressed *C. annuum* single and multi-domain inhibitor along with *H. armigera* trypsin. Single domain (IRD-7, -9, -12) and multidomain (CanPI-7) inhibitor constructs were designed and expressed in bacterial system to obtain recombinant proteins. Three dimension structures of expressed single and multidomain inhibitors were predicted by homology modelling and validated by small angle X-ray scattering (SAX) analysis. IRDs were enriched with isotope labelling and used for structural characterization using high resolution NMR. Preliminary HSQC spectra showed well-dispersed peaks, and most of the individual peaks can be distinguished indicates the folded conformation of proteins. Thermodynamics parameters obtained from isothermal colorimetry analysis of IRD binding with porcine trypsin revealed the exothermic and sntaneous nature of interaction, indicated by negative change in enthalpy and positive change in entropy. Molecular dynamic simulation of insect protease and IRD complex showed that composition and flexibility of reactive site loop is determining factor in inhibitory activity. To obtain insights about interaction of multidomain inhibitor (CanPI-7) docking and MALDI TOF analysis were clubbed. This combined analysis exhibited that although CanPI-7 has 4 reactive sites, but due to steric hindrance it

can interact with only three proteases, thus stoichiometry ratio is 1:3 for CanPI-7 (Mishra *et al*., 2013).

H. armigera protease was expressed purified and checked for its secondary structure and kinetic parameters. Recombinant protease shows activity with Km~88 μM and Vmax ~ 43 μM. Estimation of different physical parameter shows that recombinant protease is active at alkaline pH and have wide range of thermal optima. CD spectroscopy showed that *H. armigera* trypsin shares secondary structure similarity with porcine trypsin. In order to generate structural insights about protease-PI interaction, we have purified complex of IRDs with porcine trypsin by size exclusion chromatography and set up the crystallization trials. To obtain diffraction quality crystals we are optimizing different conditions. In parallel, for producing realistic view of *H. armigera* protease-IRD interaction *H. armigera* trypsin with IRDs were co-expressed in single cassette. Furthermore it is expressed, purified and set for crystallization trails. Structure of IRD, *H. armigera* trypsin in unbound and IRD bound form will provide detailed information about their enzymatic and inhibition mechanism, which can be further, explored structure-based designing of the insect protease inhibitors.

Chapter IV: Evaluating natural and synthetic protease inhibitors against *H. armigera* proteases

In order to fulfill the critical need to develop new insect–control agents and to overcome the limitation of existing pest management strategies, we have explored two new approach of structure-based screening and designing of the insect protease inhibitors.

i. Synthetic peptide designing and application in pest control: Structural information of PIs-protease interaction was further extrapolated to design synthetic inhibitory peptides. Here we have proposed application of reactive site loop (RSL) peptide as insect protease inhibitors. For RSL inhibitory peptides designing, we retrieved all Pin-II inhibitors sequences form database and analysed. Among 397 Pin-II sequences, 287 IRDs and 23 RSL variants were obtained. Motif analysis showed that out of 23 RSL variant, 6 were predominant with 81% occurrence

frequency. RSL peptides were screened *in silico* against bovine trypsin, proteases from *Bombyx mori* (monophagous), *Meduca sexta* (oligophagous) and *H. armigera* (polyphagous). This screening revealed that two RSL peptide namely TRE and PRY had high affinity against most of the proteases. *In vitro* inhibition kinetic analysis exhibited TRE as potent inhibitor of bovine trypsin with Ki~25μM. Feeding bioassays and other experimental evidences depicted that TRE impedes the digestive physiology of insect and leads to differential molecular response at translational and transcriptional levels. The inhibition potential of this synthetic peptides can be further explored to control insect pests.

ii. Mechanistic insight into insecticidal action of natural phenols: Previous studies have shown that different natural compounds especially natural phenols, plays an important role in the plant defense against insect infestation. We investigated the mechanistic insights about insecticidal activity of a natural phenol, caffeic acid on *H. armigera*. *In silico* screening of several natural phenols indicated caffeic acid as an effective inhibitor of *H. armigera* gut proteases, which later validated by *in vitro* studies. *In vivo* inhibition of the gut protease activity by caffeic acid led to the reduced growth, survival and development of insects. Gut metabolite analysis of larvae fed showed that caffeic acid remained stable in the *H. armigera* gut environment. Dynamic alterations were observed in the insect in response to caffeic acid at transcriptional and translational level. In addition, structure-activity relationship of caffeic acid confirmed the significance of its functional groups in potency against proteases. Different biophysical studies such as fluorescence and circular dichorism spectroscopy showed that binding of caffeic acid causes the conformational change in protease structure. Molecular dynamic simulations revealed that binding of caffeic acid molecule at protease active site (Ser200) provides a nucleation center for sequential binding, which further blocks the active site access. This sequential binding at active site and heterogeneous mode of protease activity inhibition might be a key reason for insecticidal activity of natural phenols. Along with the inhibition of digestive activity, caffeic acid also causes a significant reduction in detoxifying enzyme activities, which may intensify the

detrimental effect on insect physiology. Our findings suggest that natural phenols can be utilized as potential insect controlling agent.

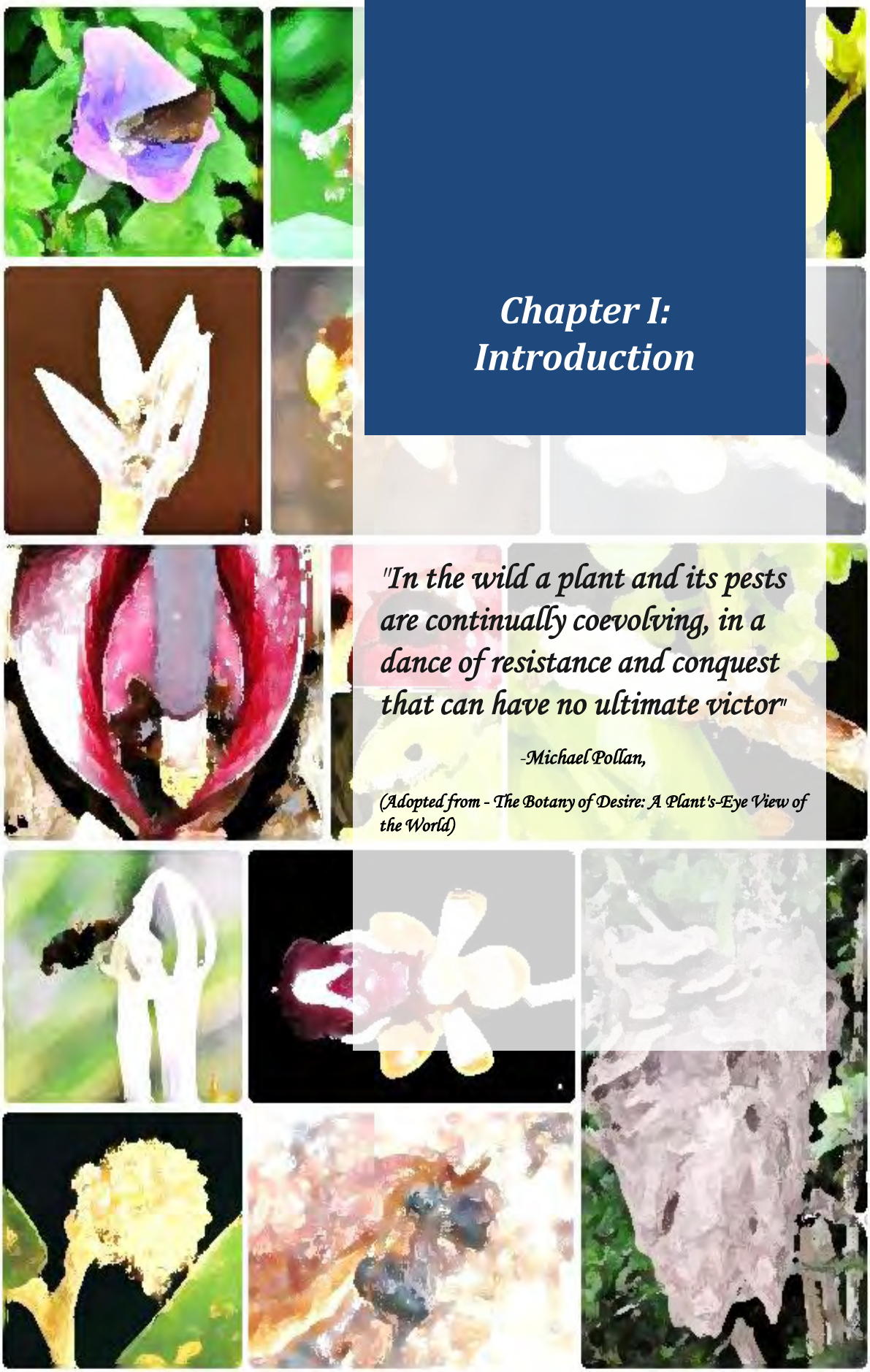
Chapter V: General discussion and future direction

General discussion

1) Sequence variation in IRD reflects differential inhibitory activity *in vitro* and *in vivo*. These variations give differential molecular response upon feeding to *H. armigera* at transcriptional and translation level.
2) Sequence variations in IRDs causes variable interaction with target proteases and this information could be used for developing tailor-made inhibitors
3) The remarkable efficiency of a Pin-II proteinase inhibitor sans two conserved disulfide bonds is due to enhanced flexibility and hydrogen bond density in the reactive site loop
4) In case of multidomain inhibitor spatial arrangement of reactive site is essential for its maximum activity
5) IRDs, *H. armigera* proteases and IRD-HaTry complex were successfully expressed and purified
6) Scaffolds of natural phenols and reactive loop might serve as potential candidates for development of new arena of insect control molecules

Future directions

1) Structural information of interaction of different range of protease inhibitor i.e. small molecules/peptides (Caffeic acid, TRE) and natural peptides (IRDs, CanPIs) with target insect proteases (HaTry4) using X-ray crystallography and NMR spectroscopy. It is hoped that this information would be helpful to design specific and effective pest controlling molecules
2) Lead optimization by means by QSAR or protein engineering to increase its potential against the target proteases.
3) Efficacy of newly identified/designed inhibitor molecules in green house/fields for assessment of its insect controlling potential



Chapter I: Introduction

"In the wild a plant and its pests are continually coevolving, in a dance of resistance and conquest that can have no ultimate victor"

-Michael Pollan,

(Adopted from - The Botany of Desire: A Plant's-Eye View of the World)

I.1 HISTORICAL PERSPECTIVE

Plant-insect interactions are considered to be one of the most primitive and thoroughly studied co-evolved systems (Ehrlich & Raven, 1964; Bronstein, 1994; Bronstein *et al.*, 2006). Various types of relationships have played their part in the evolutionary diversification of plants and insects. There is still considerable debate on the timeline and reasons behind the genesis or establishment of plant-insect interactions, primarily with respect to open questions on when and why these important interactions first evolved (Kasting & Catling, 2003).

Historical eras with drastic climatic conditions have been significant drivers for the evolution of plant-insect interactions. Climatic changes have historically influenced the taxonomic composition and geographic distributions of the plant taxa (McElwain & Punyasena, 2007). In the latter half of the Mesozoic era i.e. around 200 million years ago flowering plants (Angiosperm) have been evolved (**Fig. 1.1**).

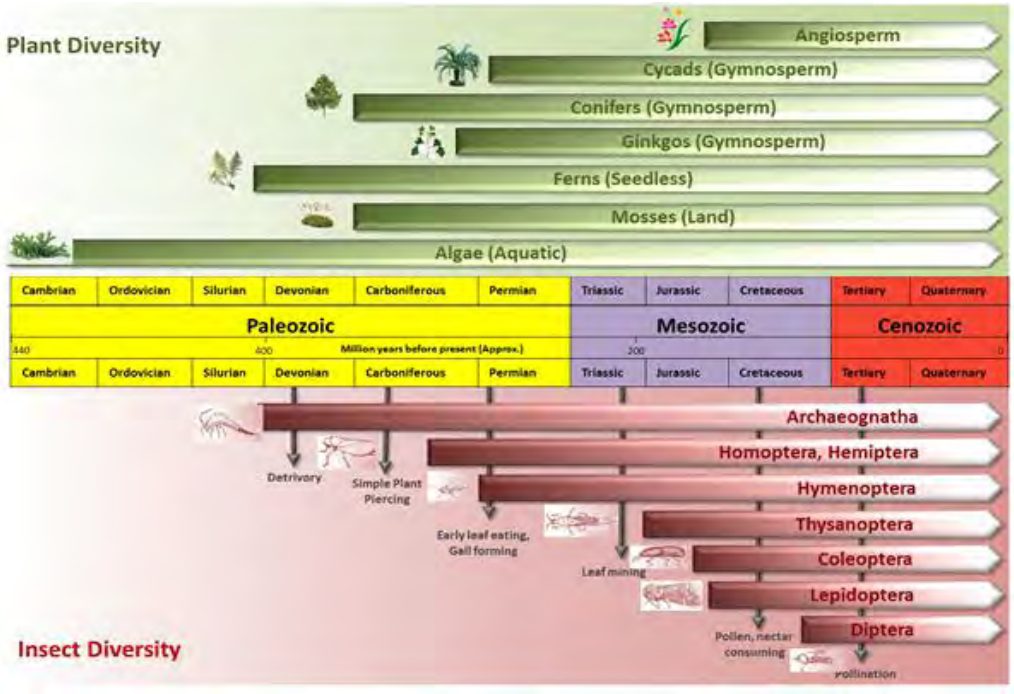


Fig.1.1: Timeline of plant and insect evolution and various feeding habits developed by insects throughout the evolutionary time. [Niklas *et al.* (1983), Tahvaninen and Niemela, (1987)]

High temperature conditions and increased carbon: nitrogen ratio in plants during the Paleocene-Eocene era resulted in a burst of insect herbivore diversity (Wilf & Labandeira, 1999; Currano *et al.*, 2008). Angiosperms underwent extreme diversification and spread widely owing to their ability to adapt to environmental instability (Heimhofer *et al.*, 2005; 2007). Relationships with corresponding pollinator insect taxa explained the congruent rise and diversification of flowering plants. These drastic changes in the host plants diversity and availability has directly influenced the balance between speciation and extinction of associated insect herbivores.

I.2 DYNAMICS OF PLANT-INSECT INTERACTION

Extensive fossil records and other evidences have highlighted the strong mutualistic associations of plants and insects sharing service-resource relationship. Various types of mutual interactions existed in different organisms depending on the nature of service and resources (**Fig. 1.2**).

Fig. 1.2: Plethora of plant-insect interactions. **A**. Skipper butterfly feeding on Ipomoea; **B**. Peacock pansy butterfly feeding on Leea; **C**. Honey bee feeding on Smithia; **D**. Fly (Milichidae) pollination in Brachystelma; **E**. Phorid fly with *Ceropegia pollinarium*; **F**. Banded blister beetle feeding on Alysicarpus; **G**. Beetle (Scarabids) pollination in Amorphophallus; **H**. Beetle feeding on stinky appendage; **I**. Milkweed caterpillar feeding Ceropegia leaves; **J**. Red tree ants harvesting honey dew from mealy-bugs; **K**. Crab spider, an ambush predator with fly on Ceropegia **L**. Gall induced by Bugs in *Hopeaponga*; **M**. Bee orchid (*Cottonia peduncularis*) an excellent mimic of bee; **N**. Hymentopterans harvesting resin from *Canarium strictum*; **O**. Pagoda ant nest built by Crematogaster ants using plant material (***Courtesy: Dr. Sachin Punekar's collection***).

Plants receive pollination services through these interactions and in turn most pollinators receive food from plants in the form of pollen and nectar or both, whereas others e. g. bees obtain gum, resin and wax from plants to build their hives (Michener, 2007). Seed dispersal is an important process and plants use several strategies for the spread of seeds. Ant-mediated seed dispersal known as 'myrmecochory' has been recorded in over 3000 plant species and more than 80 plant families (Giladi, 2006). Certain plants have seeds with arils which act as elaiosomes to attract ants. Ant workers take such seeds to their nests. After the consumption of elaiosomes, the seed is discarded and thus, is dispersed away from the parental plant (Takahashi & Itino, 2012). Altogether to confine herbivory, and survive and prosperous on available host plants for the insects eventually resulted in the appearance of 'biological warfare' or 'arms race' in native plants and insects.

I.3 PLANT DEFENSE VERSUS INSECT'S ADAPTATION

In order to reduce the impact of insects and improve their own reproduction and survival, plants have acquired a range of adaptations during the course of co-evolution.

I.3.1 Plant defense

As defense mechanism, plants exhibit several mechanical barriers, which restrict insects by deterring and/or injuring (Fernandes, 1994). Deterrents include defensive compounds released on the plant's surface; example, resins, lignins, silica and wax (Fernandes, 1994), making it difficult for the insects to consume (**Fig. 1.3**). For instance, holly plants bearing slippery leaves make insect difficult to feed. Touch mediated thigmonastic movements in some plants are also useful as defensive strategies. By providing housing and food for insect predators, plants have evolved "biotic" defense mechanisms to restrict insects (Heil *et al.*, 1997). Many plant species develop or produce extrafloral nectaries (EFN) in order to attract natural enemies of herbivore, particularly ants. Ants usually protect flowers from nectar robbing and prevent insects from laying eggs on the foliage. It has been widely observed that plants with EFNs are guarded by ants and have high success, while in

the absence of ants they are very heavily foraged by other insects. The examples of plant mechanical and morphological defenses that restrict the growth and feeding of insects are enlisted in **Fig. 1.3**.

In addition to mechanical defenses, plants use a dynamic range of chemical defense strategies against herbivores (Walling, 2000; Kessler & Baldwin, 2001; Mithofer & Boland, 2012) and foremost among these are the constitutive and/or induced production of defensive compounds. Plant secondary metabolites directly reduce the feeding, survival and reproduction of insects by either repelling or inhibiting physiological processes of insects (Karban & Baldwin, 1997; Mithofer *et al.*, 2009; Mithofer & Boland, 2012; Dawkar *et al.*, 2013). The diversity and complexity of secondary compounds have amplified over the evolutionary time scale resulting in the production of compounds that could be more challenging for herbivores to adapt (Becerra *et al.*, 2009). After damage plants emit numerous volatile compounds that act as signals to alert neighboring tissues and plants. These volatiles also attract predators and consequently involve in plant's indirect defenses (Baldwin *et al.*, 2006).

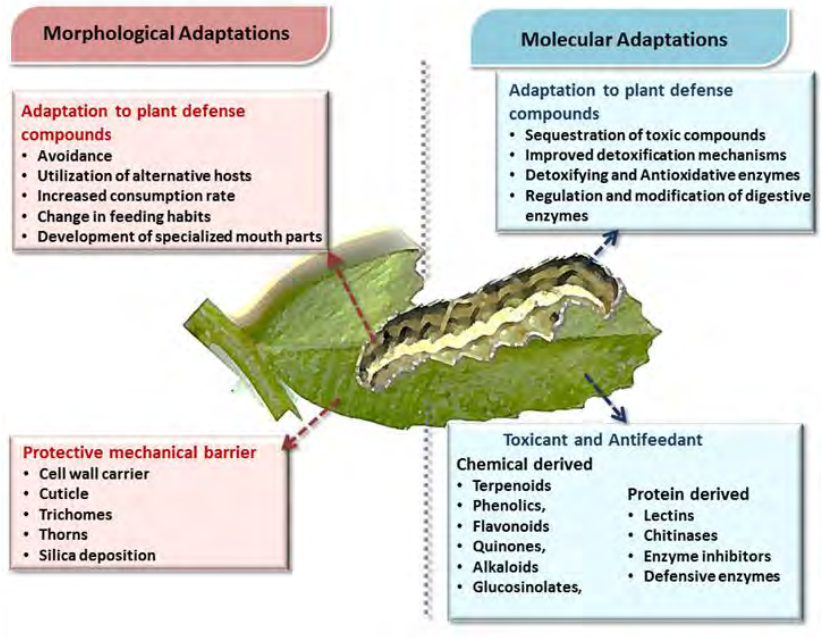


Fig. 1.3: Account of morphological and molecular adaptations of plants and insects during evolution of mutual interaction. Plant and insects use various strategies to get benefit and overcome to each other's defense.

Similarly, upon stress plants divert resources including higher defensive compounds and nutrients towards other tissues such as from local damaged leaves to systemic leaves and from root to shoot (Hummel *et al.*, 2007). Glandular trichomes of plants possess full of secondary metabolites with defensive functions, however these structures surprisingly act as the primary feast for Lepidopteran herbivores (Weinhold & Baldwin, 2011).

I.3.2 Insect adaptation to plant defense

In the phenomenon of co-evolution along with plants, insects also adapted to certain plant chemicals by developing mechanisms to metabolize, sequester, excrete, or selectively bind to plant defense compounds (**Fig. 1.3**). Herbivores have various means of manipulating their host plants, such as the modification of microhabitats in order to counter plant defenses and gain better use of the resources. For example, some caterpillars roll mature leaves around growing buds of the flowering shrub *Psychotria horizontalis* (Sagers, 1992). Gall formation is also one of the insect adaptation strategies to acquire food from plants. Galls not only provide insects with edible nutritious tissue embedded inside but also provide the protection from predators (Weis & Kapelinski, 1994). Sequestration is an important strategy to detoxify harmful metabolites, which insects often use for their own benefit against predators. Insects have evolved a way to suppress the plant defense by laying eggs leading to egg-induced salicylic acid (SA) accumulation, which negatively interferes with the jasmonic acid (JA) pathway, and thus eventually benefiting herbivores (Bruessow *et al.*, 2010). The adaptive responses in herbivorous insects often depend on the type of plant toxin and its mode of action. Dobler *et al.* (2012) demonstrated an example of convergent molecular evolution in cardenolide resistant herbivorous insects. Cardenolides are found in almost 12 families of flowering plants, which are inhibitors of insect (Na^+/K^+) ATPase. Cardenolide-resistant insect species from 15 genera within 4 orders (Coleoptera, Lepidoptera, Diptera and Hemiptera) were found to have same amino acid substitution (position 122; N122H) in the extracellular loop of (Na^+/K^+) ATPase (Dobler *et al.*, 2012). The absence of this particular substitution in any of the species which do not feed on

cardenolide producing plants suggests that the resistance conferring substitution in a specific subset of herbivorous insect species is a result of adaptive response across 300 million years of insect evolution and divergence. Such molecular adaptation may also have synecological implications when resistant insects sequester these compounds and use them for protection against predators and display warning coloration signifying their toxicity.

Similarly, several biochemical adaptation strategies have been developed by the insects to cope up with the plant defense chemicals (**Fig. 1.3**). Insects produce enzymes, which counteract and decrease the efficiency of numerous plant secondary metabolites. Insects use several enzymes such as cytochrome P450 oxidases, glutathione s-transferases and esterases for the detoxification of plant toxic compounds (Snyder & Glendinning, 1996; Feyereisen, 1999; Mithofer & Boland, 2012). Insects regulate expression of protease isoforms with varied properties and specificities in their gut to minimize plant defensive protease inhibitors (PIs) or depending upon dietary content (Broadway and Duffy 1986; Johnston *et al.,* 1991; Purcell *et al.*, 1992; Broadway 1995; 1996; Jongsma *et al.*, 1996; Bown *et al.*, 1997; DeLeo *et al.,* 1998; Lomate & Hivrale, 2010; 2011). Various overcoming strategies used by plants and insects against each other are shown in **Fig. 1.3**.

I.3.3 Insect selection of host leads to its specialization

Host selection by insects is governed by nutritional, geographical or ecological attributes. However, it has been suggested by Fraenkel (1959) that the host specificity of insects depends totally on the presence or absence of unusual secondary compounds in plants. Thus, it is likely that the qualitative and quantitative diversity of primary and secondary plant metabolites may influence the specific nutritional relationship between insects and plant hosts (Fraenkel, 1959; Ali and Agrawal, 2012). Based on the preference exhibited by insects for feeding on host plant species, they may be grouped as monophagous (feeding on a single or few closely related plant species), oligophagous (feeding on a certain group of plant families), or polyphagous (feeding on a wider group of plants) (Ali & Agrawal 2012).

Monophagous species or 'specialists' not only equipped with specialized mechanisms to deal with deterrents but also obtain sufficient nutrition from particular plant(s). Polyphagous species or 'generalists', on the other hand, rely more on general detoxifying mechanisms (such as enzymes such as, mixed functional oxidase) to neutralize adverse effects of range of phytochemicals as well as blend diverse digestive enzymes to acquire required nutrition (Krieger, 1971; Patankar *et al.*, 2001). Monophagy develops when herbivore species strongly compete for food and as a result some species develop specialized strategies and acquire strong competitive ability to obtain the resources from plants (Jaenike, 1990). Polyphagy or generalization frequently develops when herbivores have scarcity of preferred host plants and eventually they try to feed on alternative/available plants and obtain required nutrition and energy. In classical and pioneering work, Ehrlich & Raven (1964) stated that exploiting a particular plant as a food source involves metabolic adjustments/expenditures on the part of the insect, resulting in increased cost of detoxification systems needed by the generalists. Thus, generalists would show lower efficiencies or growth rates as compared to the specialists on any given host. Hence specialist herbivores are physiologically better adapted to their host plants defenses than the generalists (Ali & Agrawal 2012).

Whatever be the state/nature of interactions, plants and insects survived and maintained their niche in harmony until the onset of extreme interference on various ecosystems by humans. Human interference has prominent impact on plant and insect adaptability, which leads to reduction in the natural immunity of plant and emergence of pest status of the insect.

I.4 JOURNEY FROM INSECT TO PEST: CASE OF *HELICOVERPA ARMIGERA*

Biotic stresses such as pests, pathogens, weeds and herbivores have adversely affected the growth, development and productivity of plants. Insect pests, diseases and weeds are the major constraints limiting the potential agricultural production. It is estimated that herbivorous insects alone lead to about 26% loss of the food

production all over the world while India lose about 18% of the crop yield valued at Rs. 8,63, 884 crore INR due to pest attacks each year as (http://www.livemint.com/). Details of crop loss caused by insect pests to major agricultural crops in India are given in **Table 1.1** (Dahiya *et al.*, 2010). Among the various insects, the larvae of many lepidopteran species are extremely herbivores and major pests in agriculture. Lepidoptera is the second largest order in the class Insecta which includes moths and butterflies. Some of the major pests belong to the families *Tortricidae, Noctuidae*, and *Pyralidae*. The larvae of the *Noctuidae* genus *Spodoptera* (armyworms) and *Helicoverpa* (corn earworm) can cause extensive damage to certain crops. The host plant range for lepidopteran insects may either be narrow (monophagous or specialist), e.g., *Manduca sexta*, which exhibits preference for Solanaceous plants, or diverse (polyphagous or generalist), e.g., *Helicoverpa armigera*, which feeds on various legumes, vegetables and fruits (Ehrlich and Raven, 1964; Tamhane *et al.*, 2005a; Wu and Baldwin, 2010). Due to highly polyphagous nature of *H. armigera, H. zea, H. virescens* and *H. punctigera* they cause damage to a large number of plant species including food, fiber, oil, fodder and also horticultural and ornamental crops (Fitt, 1989).

Table 1.1: Estimation of crop losses caused by insect pests in India

Crop	Actual production * (million tonnes)	Approximate estimated loss in yield		Hypothetical production in absence of loss	Monetary value of estimated loss (million INR)
		Percentage	Total (million tonnes)		
Cotton	44.03	30	18.9	62.9	339660
Rice	96.7	25	32.2	128.9	240138
Maize	19	20	4.8	23.8	29450
Sugarcane	348.2	20	87.1	435.3	70667
Rapeseed	5.8	20	1.5	7.3	26100
Groundnut	9.2	15	1.6	10.8	25165
Other oilseeds	14.7	15	2.6	17.3	35851
Pulses	14.8	15	2.6	17.4	43551
Coarse cereals	17.9	10	2	19.9	11933
Wheat	78.6	5	4.1	82.7	41368
Total		17.5			863884

*Production and minimum support price (MSP) fixed by Government of India for 2007-08, are adapted from Anonymous (2010)

H. armigera infests about 300 plant species and the most important host crops include cotton, legumes, tomato, tobacco, okra, potato, sunflower, safflower, maize, groundnut, etc. (Fitt, 1989; Rajapakse and Walter, 2007) (**Fig. 1.4A**). Preferentially larvae of *H. armigera* feed on the reproductive structures and economically important plant parts like seeds, fruits, pods which lead to prominent reduction in the yield of the crop. The losses due to *H. armigera* infestation are estimated to be 20 to 80% in cotton, 14 to 100% in pigeon pea, 18 to 26% in sorghum, 30 to 60% in sunflower, 15 to 46% in tomato and over 40% in okra (Reddy & Zehr, 2004). It inhabits many countries of Asia, Europe, Australia and Africa and is well known pest because of its wide geographic presence. The characteristic features which make it a successful pest are polyphagy, high reproductive rate, mobility and facultative diapause (Fitt, 1989). *H. armigera* has a life span ranging between 25 to 35 days. Each female can lay several hundred eggs on lower surface of leaves, flowers, shoot tips and young pods, where larva feeds preferentially. The eggs are hatched after 3-4 days, while the larval stage lasts for 12-16 days **(Fig 1.4B)**. The larval period is divided into six instars in which the fourth and the fifth are highly voracious and the most damaging ones (Tamhane *et al*, 2005a). The caterpillars are aggressive, occasionally carnivorous and, when the opportunity arises, cannibalistic. Pupa is a non-feeding stage which lasts normally for 6-10 days and the moth remains alive for 4-5 days (Gowda *et al.*, 2005; Tamhane *et al.*, 2005a). In case of unfavourable conditions, the larvae undergo facultative diapause i.e. a state of suppressed metabolism as pupae in order to survive. There is a great variation in the colour of larval instars and the moth depending on the environment and the food on which it feeds upon.

H. armigera feeds simultaneously on a number of host plants leading to high population build-ups. The insect's ability to survive on diverse host plants is an adaptive mechanism for their survival and propagation in the ecosystem (Tamhane *et al.*, 2005a). Polyphagy requires metabolic adjustments in insects in order to maximize the benefits from protein-rich plant reproductive structures, carbohydrate-rich leaves, and even diverse unbalanced diets (Sarate *et al.*, 2012). The total amount of protein, carbohydrate, and lipid in the diet directly influence the

insect growth and development. Digestive enzymes of *H. armigera* (amylases, proteases and lipases) display variable expression levels, regulated on the basis of macromolecular composition of the diet.

(A) **(B)**

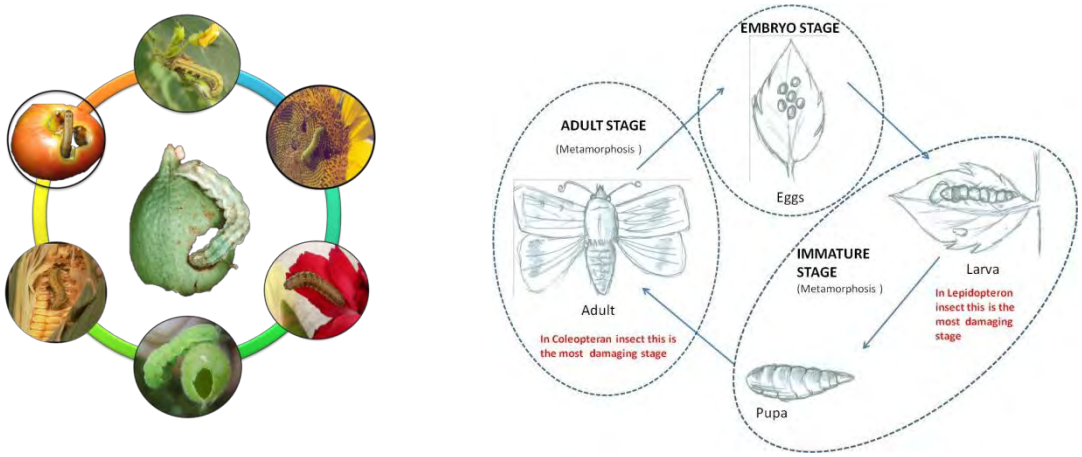


Fig. 1.4: (A) Polyphagous nature and (B) Life cycle of *Helicoverpa armigera*

The digestive complement of *H. armigera* mainly consists of endo-peptidases like serine, metallo-, cathepsin B like proteinases and exopeptidases. Serine proteinases, such as trypsin and chymotrypsin, form the dominant mechanistic class (>95%) in the gut environment (Johnston *et al.*, 1991; Purcell *et al.*, 1992; Harsulkar *et al.*, 1999; Patankar *et al.*, 2001). Many of these proteinases have been isolated, identified and the coding DNA/cDNA have also been well characterized (Gatehouse *et al.*, 1997; Bown *et al.*, 1997, 1998; Mazumdar-Leighton *et al.*, 2000; Patankar *et al.*, 2001; Bayes *et al.*, 2003; Chougule *et al.*, 2005; Telang *et al.*, 2005). As account of various reports on *H. armigera* digestive physiology, 32 trypsins, 22 chymotrypsin and few elastases gene have been identified, and addition of some new proteases isoform is still under process (Chikate *et al.* 2013). The complex proteinaceous food material is broken down in to smaller oligopeptides by endo-peptidases like trypsins and chymotrypsins. These oligo-peptides are further digested by exo-peptidases liberating the free amino acids. Thus, the digestion in the larval gut follows a rational trend as a consequence streamlining the digestive process. The

composition and relative levels of expressed gut proteinase vary according to the larval developmental stage and the dietary components (Patankar *et al.*, 2001; Chougule *et al.*, 2005; Sarate *et al.*, 2012) **(Fig. 1.5A)**. Variation in digestive complement pertaining to larval stage and diet explains well the polyphagous nature of *H. armigera* which allows it to infest a wide variety of agriculturally important crops simultaneously (Patankar *et al.*, 2001) (**Fig. 1.5B**).

(A) **(B)**

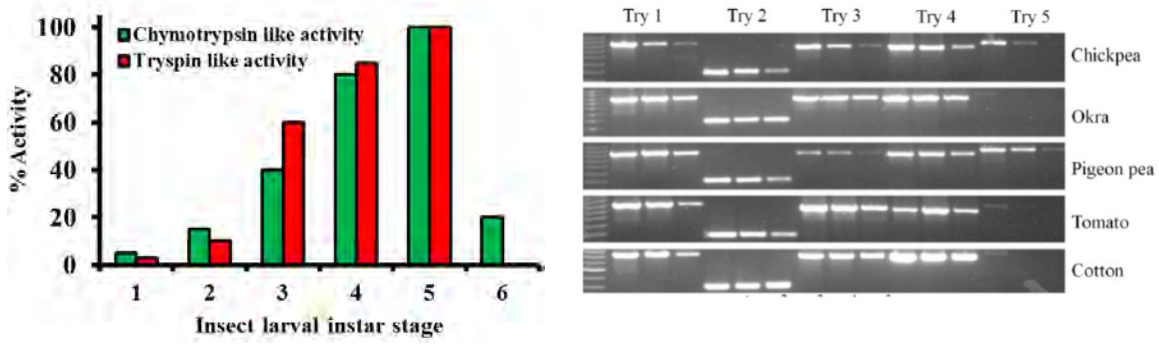


Fig. 1.5: Gut protease flexibility of H. armigera (A) Relative gut proteinase activity of *H. armigera* during the stages of larval development [Reproduced and modified from Patankar *et al.*, 2001]. **(B)** Expression of trypsin-like proteinases of *H. armigera* larvae feeding on various host plants [Adopted from Chougule *et al.*, 2005].

Plant has developed various defense strategies against *H. armigera*, which includes a range of adaptations evolved by plants in order to reduce the negative impact of herbivores and improve their survival and reproduction. Many plants produce secondary metabolites and proteins, for example: plant phenols, proteinase inhibitors, lectins and amylase inhibitors, which can influence digestive metabolism, growth and survival of *H. armigera*. These protease inhibitor metabolites and proteins cause indigestion and thereby disrupting the nutrient acquisition system in the infesting insects.

I.5 CONTROL OF *H. ARMIGERA*

Owing to the extensive loss in crop yield due to *H. armigera*, several strategies used for its management include: cultural practices, chemical insecticides, biological control, host resistance, biotechnological approaches and integrated pest management.

I.5.1 Cultural practices

This includes deep ploughing of soil, hand picking of large sized larvae, shaking of plants, weeding, intercropping, use of trap crops, following a good time of sowing (Dahiya *et al.*, 1999) and fertilizer application. The cultural practices are eco-friendly and economical, however, they are mostly very laborious. Use of chemical pesticides has been the main strategy for controlling these pests from the last few decades. The main advantage of chemical pesticides is that they are effective even when used in an advanced stage in advanced larval stages, however, of infestation but their continuous usage has led to the development of resistance against many chemical insecticides including organochlorides, organophosphates, carbamates, pyrethroids etc. in insect population (Mccaffery, 1998; Gunning *et al.*, 1998). Moreover importantly, chemical pesticides cause a severe threat to the environment by contaminating soil, water, vegetation and ecosystems and exert toxic effects on the biome including human beings and non-target organisms.

I.5.2 Biological control

This class of methods involves deploying the natural enemies of the pest viz. predators, parasitoids, biopesticides with live nematodes, fungi, bacteria, viruses or products derived from them and other plant products (Gurjar *et al.*, 2011). Being environment friendly, they are recently becoming very popular. For example, entomopathogenic fungus such as *Beauveria bassiana* and *Metarhizium anisopliae*, are treated as natural enemies of *H. armigera*. Among all the bio-pesticides, Nuclear Polyhedrosis Virus (NPV) and *Bacillus thuringiensis* (Bt) have been very popular. Bt produces insecticidal crystal proteins which can lead to death of a wide range of insect orders proteins. This causes death of a wide range of insect orders, and

Importantly these toxins could be specific to insect class (*Lepidoptera, Diptera, Coleoptera, Hymenoptera, Homoptera, Orthoptera* and *Mallophaga*) (Sanchis and Bourguet, 2008). Neem products, vegetable oils, crude plant oils are other popular plant products used for pest control with limited success. Host plant resistance to insect pests relies on the availability of resistance genes in the germplasm and their transfer through conventional breeding procedures or modern biotechnological approaches to high yielding cultivars. Few wild relatives of chickpea (*C. icer bijugum, C. judaicum* and *C. reticulatum*) have been reported to be potential sources of resistance/tolerance to *H. armigera*, which can be used to increase the levels of insect resistance tolerance in chickpea (Sharma *et al.*, 2007).

I.5.3 Biotechnological approach

This is one of the advanced and important approach, which utilizes recombinant DNA technology to enhance insect resistance of plants by transfer of heterologous genes from various other sources (Ferry *et al.*, 2006). Bt toxin has been the most commonly used gene for generating transgenic plants resistant to insect pests (Barton *et al.*, 1987; Gupta *et al.*, 2000). Commercial introduction of genetically modified maize, potato and cotton plants expressing Bt genes in the mid-1990s has been the most important landmark in crop improvement, which revolutionized agriculture by increasing its productivity. However, modification of a Bt toxin receptor site in insects has led to the development of resistance against transgenic Bt plants at least in laboratory conditions (McGaughey *et al.*, 1998; Frutos *et al.*, 1999). In anticipation of insect adaptation to Bt, toxin attention was shifted to other bio-molecules also. Inhibitors of amylases and of proteinases of insects and lectins are few of them are most studied plant defense molecules which exhibit potential to use them in insect and (Koundal, 2002; Carlini and Grossi-de-Sa, 2002; Babu *et al.*, 2003; Haq *et al.*, 2004; Giri *et al.*, 2005) and are under consideration for their in transgenic (Dunse *et al.*, 2010a). Lectins disrupt the gut cell wall of insects by binding to glycoproteins of brush border cells and thus, the nutrient uptake is hampered. α-amylase inhibitors (α-AIs) occur in many plants as a part of the natural defense mechanisms. They complex with the insect gut amylases and need further

research to identify potent AIs. Another attractive molecule is plant protease inhibitor (PI), for producing transgenic plants.

Proteinase inhibitors (PIs) are one of the most vital and well-studied plant defense molecules. PIs are ubiquitously found in various plant tissues and also typically produced in response to herbivore attack or damage to the plant tissues. PIs are naturally encountered in many plant species; their expression varies between various plant tissues including leaves, flowers and fruits/seeds. PIs are abundantly present in the storage tissues of plants and can represent up to 10% of total protein (Ussuf *et al.*, 2001). Many PIs are produced in response to various biotic and abiotic stress conditions, e.g. pathogen invasion, insect attack, wounding, and environmental stress. Plant PIs have been classified based on their specificity against four mechanistic classes of proteolytic enzymes, i.e. serine, cysteine, aspartic and metallo-protease (Ryan, 1990) (**Fig. 1.6**). Further, each type includes different inhibitor families which have been classified based on their molecular mass, sequence homology, structural characteristics and expression patterns (Garcia-Olmedo *et al.*, 1987; Laskowski *et al.*, 1980; Ryan, 1990).

In plants, ten protease-inhibitor families have been recognized. Most inhibitor proteins in these families exhibit molecular mass from 5 to 25 kDa. Many of them are products of multigene families and several iso-inhibitors have been found in a single species exhibiting different specificities towards proteases (Wu *et al.*, 2006; Tamhane *et al.*, 2009). Serine proteinase inhibitors are further classified in to various families such as Kunitz, Bowman-Birk, Wound-inducible (potato proteinase inhibitor [PIN] type I and II, Squash, Cereal trypsin /α-amylase inhibitor and Mustard seed trypsin inhibitor. Members of serine and cysteine proteinase inhibitor families have been more relevant to the area of plant defense because of their wide presence across many plants while metallo- and aspartic have limited occurrence (Ryan, 1990). Out of all these families Pin II represents one of the important class of PIs, because of their wound inducible expression they perform vital role in plant immune response against pests and pathogens.

The accumulation of trypsin and chymotrypsin-like proteinase inhibitors (PIs) throughout the aerial tissues of tomato and potato plants was demonstrated to

be a direct consequence of insect-mediated damage or mechanical wounding (Green and Ryan, 1972). The earliest report on possible role of PIs in plant protection dates back to 1947 when, abnormal development of insect larvae was observed on soybean-based products (Mickel and Sytandish, 1947). Subsequently, the toxic effects of trypsin inhibitors from soybean were shown on larvae of flour beetle, *Tribolium confusum* (Lipke *et al.*, 1954).

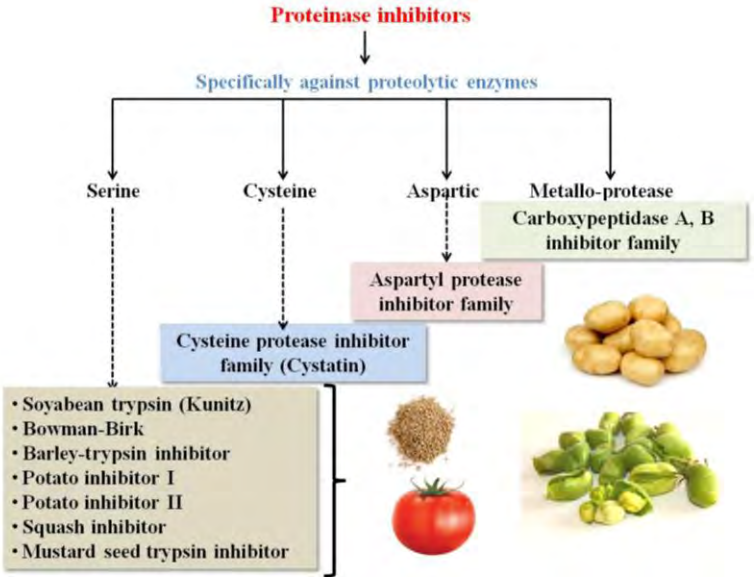


Fig 1.6: Plant proteinase inhibitor families

The direct evidence for inhibitory effect of PIs in plant leaves against insects was first demonstrated by Hilder *et al*. (1987) by expressing a cowpea trypsin inhibitor gene in transgenic tobacco plants. The leaves produced 1% of the leaf protein as inhibitors and showed more resistance to feeding by larvae of *Heliothis virescens* than the untransformed control tobacco plants. Cowpea inhibitor was established as an anti-nutrient agent against *Heliothis, Spodoptera, Diabrotica* and *Tribolium*, all agronomically important insect pests. Following these studies, there has been many examples of antagonistic activity of PIs against insect pests, by *in vitro* assays against insect gut proteases as well as by *in vivo* (insect feeding) assay (Broadway and Duffy, 1986a and 1986b; Johnston *et al.*, 1993; Harsulkar *et al.*,

1999; Bown *et al.*, 2004; Giri *et al.*, 2005; Tamhane *et al.*, 2007).PIs may adversely affect the proteinases of phytopathogenic microorganisms also. Plant derived trypsin and chymotrypsin inhibitors were found to suppress the activity of proteinases excreted by a fungus *Fusarium solani* and also the growth of hyphae and conidium germination in *F. culmorum*, and *Botrytis cinerea* (Mosolov *et al.*, 1976; Valueva *et al.*, 2004). PIs do not have a wipe-out effect like synthetic pesticides and thus, do not impose strong selection pressure. Many transgenic plants have been developed using the PI genes (Recent examples, Hilder *et al.*, 1987 and 1993; Dunse *et al.*, 2010a). Making use of combinations of a variety of defense molecules has risen as a futuristic approach to improve insect resistance in crop plants.

Transgenic expression of potato proteinase inhibitor II (PinII) PIs in plants has helped them in improving resistance against insect pests and showed how addition of a defensive trait can provide enhanced protection against the important predators. Pin-II PI genes of tomato (inhibitor I and II) and potato (inhibitor II) were expressed in transgenic tobacco plants (Johnson *et al.*, 1989). *Manduca sexta* larvae (tobacco hornworms) feeding on the leaves of transgenic plants containing inhibitor II, a powerful inhibitor of both trypsin and chymotrypsin, showed significant retardation in growth, compared to that of larvae fed on untransformed leaves. Interestingly, greater insecticidal effect was observed in tobacco plants transformed with the genomic sequence of tomato PI-II than in those transformed with the cDNA sequence indicating appropriate splicing of exogenous sequences in the transgenic plants to obtain the active protein (Zhang *et al.*, 2004). Tobacco plants transformed with Pin II gene, encoding a chymotrypsin inhibitor showed reduced larval growth for *Chrysodeixis eriosoma* (green looper) while no adverse effect on *S. litura* was observed (McManus *et al.*, 1994). Thus, a single proteinase inhibitor gene may not be universally effective against a range of insect pests. A transgenic expression of multi-domain proteinase inhibitor precursor from *N. alata* (Na-PI) in tobacco resulted in accumulation of 6 kD polypeptides in leaves to a level of 0.3%. *H. armigera* larvae feeding on such tobacco leaves, exhibited higher mortality or were delayed in growth and development relative to the control larvae (Charity *et al.*, 1999). Similar adverse effects of Na-PI were observed on *Epiphyas postvittiana*

(light-brown apple moth) when expressed in transgenic apple plants (Maheswaran *et al.*, 2007). Field trials with transgenic cotton plants, expressing the individual inhibitors and combination of inhibitors (NaPI-StPin1A), subjected to *H. armigera* infestation along with the natural prevalence of *H. punctigera* at the field site were conducted by Dunse *et al.* (2010a). They recorded an increase in number of cotton bolls in the transgenic line expressing both *StPin1A* and *NaPI* than the parental untransformed line and also a boost in lint weight per plant for these transgenic lines, when evaluated against the control line (**Fig. 1.7**).

Pin-II PIs of potato had also been used in transgenic rice and wheat plants to control biotic infestations by *Sesmia inferens* in rice (Duan *et al.*, 1996) and *Heterodera avenae*, a nematode in wheat (Vishnudasan *et al.*, 2005). Combined leaf specific over expression of potato PI-II and carboxypeptidase inhibitors in transgenic tomato resulted in increased resistance to *Heliothis obsolete* and *Liriomyza trifolii* larvae (Abdeen *et al.*, 2005).

Fig. 1.7: Growth of cotton balls on field-grown transgenic line expressing both *StPin1A* and *NaPI* cotton plants as compared to parental untransformed line. [From Dunse *et al.,* 2010].

Transgenic tobacco plants over expressing *SaPIN2*a (from *S. americanum*) under the control of CaMV 35S promoter were more resistant to *H. armigera* and *Spodoptera litura* larvae than the control plants (Luo *et al.*, 2009). They also reported an increase in the glandular trichome density along with the promotion of trichome branching in transgenic tobacco plants. An approach other than over expression of PI genes has also been considered to analyze the potential of PI in

insect tolerance. Enzymes lipoxygenase and linolenic acid are involved in JA biosynthesis. Anti-sense lines for lipoxygenase gene in potato plants abolished the accumulation of PIs on wounding to much extent (Royo *et al.*, 1999). As a consequence, the weight gain of Colorado potato beetles and beet armyworm fed on anti-sense plants was significantly larger than those fed on wild type plants.

I.6 PIN-II FAMILY OF PROTEINASE INHIBITORS

PIs of the potato inhibitor II family (PIN2/Pin-II) are the best characterized plant serine PIs with respect to their molecular properties (Bryant *et al.*, 1976; Xu *et al.*, 2001; Xu *et al.*, 2004; Sin and Chye, 2004) and have been interestingly explored at gene, protein and functional level. Pin-IIs have been found in various tissues, e.g. tubers (Bryant *et al.*, 1976), fruits (Richardson, 1979; Damle *et al.*, 2005; Tamhane *et al.*, 2007), wounded leaves (Pearce *et al.*, 1993), and flower (Sin and Chye, 2004). These Pin-II PIs have a unique single or multi domain repeat structure with variations, also a wound/insect infestation induced up-regulation and expression. All these features have made them a fascinating subject area to study plant-insect interactions as well as to understand the plant physiology. The focus is on potato inhibitor-II (Pin-II) family of serine PIs which displays a striking genetic and molecular diversity and a significant plant defense related role.

Pin-II PIs are predominantly found in Solanaceous plants and their wound induced up regulation has established their defense related role (Green & Ryan, 1972). Identification of many Pin-II homologs dispersed throughout the whole range of mono- and dicotyledonous plants has indicated more widespread occurrence of this family (Barta *et al.*, 2002). The striking feature of Pin-II PIs is the presence of variable number of inhibitory repeat domains (IRDs), structurally forming multi-domain proteins. The first two well characterized Pin-II PIs from tomato (Graham *et al.*, 1985) and potato (Sanchez-Serrano *et al.*, 1986) had two domains with trypsin and chymotrypsin reactive sites. Since then, PIs containing even up to 8 domains have been reported. Pin-II PIs are encoded by one, two or multi-gene families in *Nicotiana, Solanum* and *Capsicum* species, respectively. In *Nicotiana* most of the members show presence of single functional genomic copy of

Pin-II PIs, but varying number of IRDs (2 to 8). The numbers of IRDs in various *Nicotiana* species have shown no correlation to the phylogenetic associations, since plants with close phylogenetic relationships had different repeat numbers (Wu *et al.*, 2006). The evolutionary advantages of repetitive IRDs have not been established, but it has been predicted that the repetitive domains provide plants with a more efficient pool of inhibitory units which might play varied functions. Phylogenetic relationships of Pin-II PIs from various genera of Solanaceae are represented in **Fig. 1.8**. Around 67 Pin-II PIs having 1- to 4-IRDs (CanPIs) have been isolated from developing fruit and stem tissues of *C. annuum* (Tamhane *et al.*, 2009, Mishra *et al.*, 2013). Significantly higher level of PI activity as compared to the leaf, stem and fruit tissues have been found in flowers (Damle *et al.*, 2005; Tamhane *et al.*, 2009). Induced up regulation of Pin-II PIs has been noted in various Solanaceae plants like *Nicotiana sp., C. annuum, Solanum sp*. in response to insect attack, wounding, systemin, methyl jasmonate but also with respect to the gene as well as protein structure.

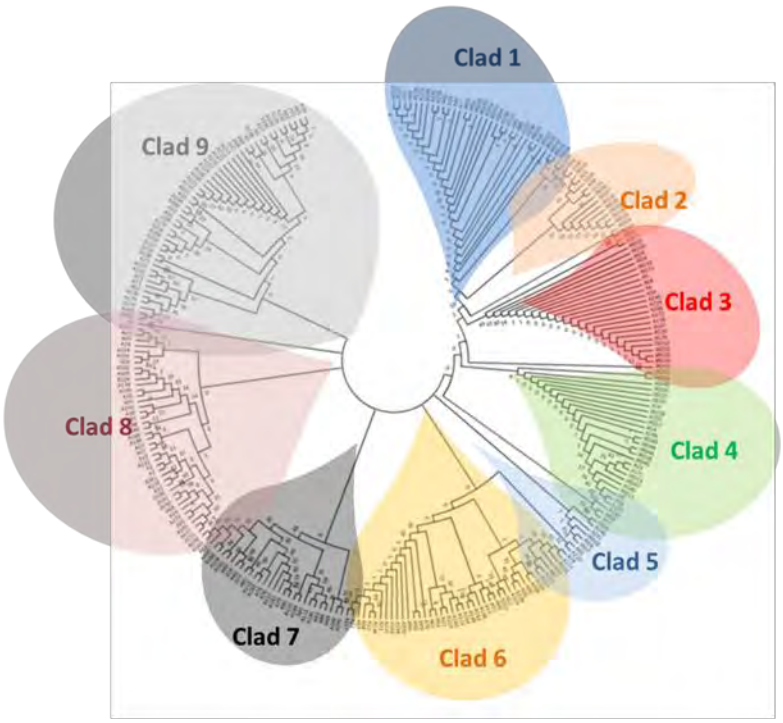


Fig. 1.8: Phylogenetic tree of Pin-II PIs family from *Solonaceae*

I.6.1 Basic structure of Pin II PIs gene and protein

The architecture of the Pin-II genes is conserved. The first exon encoding the N-terminus of the signal peptide and the second main exon encoding the C-terminus of the signal peptide followed by variable number of IRD's, are always separated by an intron of 100-200 bp (Barta *et al.*, 2002). Presence of variable number of repeat regions has suggested that duplication of Pin-II genes has occurred several times, especially in Solanaceae members. Single repeat genes are also found outside Solanaceae whereas 2 to 3 and 4 to 8 repeat unit PIs are exclusively present in Solanaceae. Single repeat Pin-II PIs are thought to be the ancestral members that have given rise to the other forms by a series of gene duplication and domain duplication events (Barta *et al.*, 2002). Kong & Ranganathan (2008) also support this gene duplication hypothesis for Pin-II family gene evolution. Unequal crossing over is presumed to be responsible for the expansion of the repeated domains. Inhibitory domain multiplication with the domains remaining fused plays a critical role in generating such diversity (Christeller, 2005).

The conserved Pin-II PI protein consists of an endoplasmic reticulum signal peptide of 25 aa followed by variable number of IRDs, each of ~55 aa. The aa sequence of IRDs is highly variable, however, presence of eight cysteines, a single proline residue and an active site either for trypsin or chymotrypsin inhibition is a conserved feature of each IRD. The IRDs are separated by 5 aa linker regions. A vacuolar sorting signal at the C-terminal region is present in some Pin-II PIs. In most of the Pin-II PI precursors, there are partial IRDs at the N- and C- terminal which form covalent bonding to generate a functional IRD. Exceptionally, *C. annuum* PIs do not possess the N- and C- terminal partial IRDs (Tamhane *et al.*, 2009).For a multi IRD Pin-II protein, there are two possible domain organizations: (1) circularly permuted domain organization which is formed by the covalent joining of the two terminal half-repeats to form a repeat domain; or (2) tandem repeat domain organization where domains are arranged in beads-on-a-string way (Schirra and Craik, 2005) (**Fig. 1.9**). The 43 kDa precursor PI of *N. alata* (NaProPI) forms a circular "clasped bracelet" like structure as a result of formation of disulfide bridges between the partial repeat regions at the N- and C- terminals of the precursor

(Scanlon *et al.*, 1999; Schirra and Craik, 2005). Sequence of the linker regions between the IRDs is almost conserved in *Nicotiana sp.* (EEKKN), whereas show sequence variations in other genera. For example, in *C. annuum* sequences are QRNAK, EENAE, EASAE, EGNAE and EETQK. The linker region is very sensitive to proteases and is cleaved by endogenous proteinases in planta (Heath *et al.*, 1995) to release individual functional units, IRDs, which are able to inhibit a single proteinase molecule.

I.6.2 Structural characterization of PinII PIs

The three-dimensional structure of several Pin-II PIs has been elucidated, either by X-ray crystallography or NMR. The structure of chymotrypsin-binding domain (PCI-1) from potato PI-II, in complex with Streptomyces griseus proteinase B was solved to 2.1Å by X-ray crystallography (Greenblatt *et al.*, 1989).

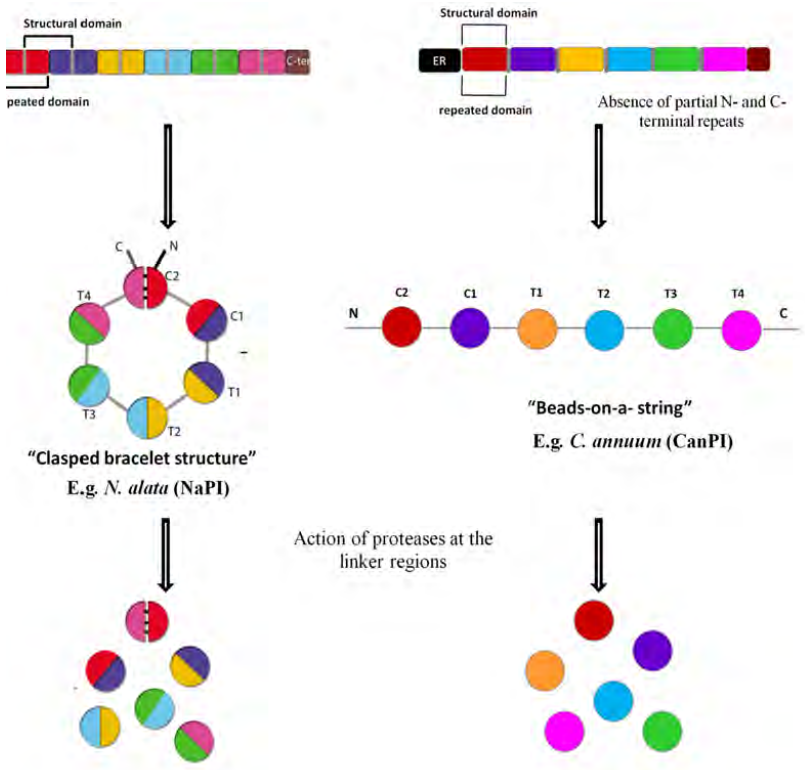


Fig. 1.9: Model of Pin-II PI precursors. Structure and processing of the 6 domain Na-pro PI. The 40.3 kDa PI precursor consists of 6 repeated domains preceded by an ER signal sequence and terminated by a vacuolar targeting signal. C2 is formed by joining of N- and C- terminal partial ends. In case of lack of these partial ends, beads on a string like structure are expected to be attained by the precursors. [Modified from Dunse and Anderson; ISB news report, June 2011].

The structures of single IRD from Pin-II PIs of *N. alata* (Nielsen *et al.*, 1994), two domain precursor PIs from tomato individually and in ternary complex with two molecules of *Subtilisin carlsberg* (Barrette-Ng *et al.*, 2003a) and 6-IRD PIs from *N. alata* (Schirra & Craik, 2005) have been determined. Single domain PIs having either trypsin inhibitory (TI) or chymotrypsin inhibitory (CI) reactive sites are formed from the multi-domain precursor protein by proteolytic cleavage at the linker regions (Heath *et al.*, 1995; Lee *et al.*, 1999). The eight conserved cysteines in each IRD are involved in formation of four disulfide bonds, which stabilize the repeat structure (**Fig. 1.10**). The eight conserved cysteines in each IRD are involved in formation of four disulfide bonds, which stabilize the repeat structure (**Fig. 1.10**). The single IRD CI protein contains a triple stranded β-sheet as the dominant secondary structural element, with several turns and a short region of helix.

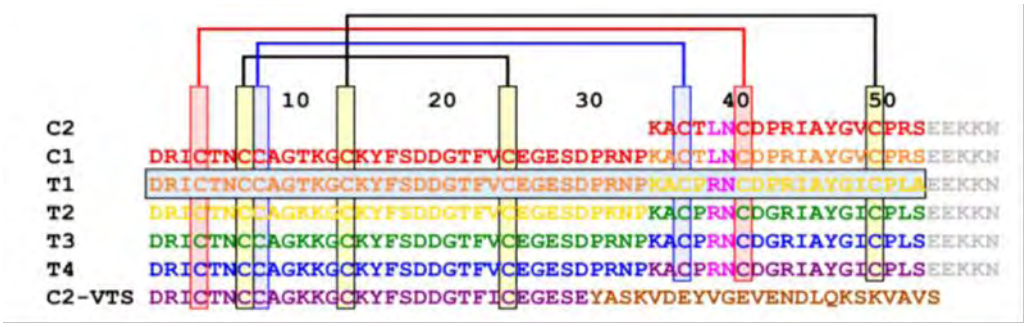


Fig. 1.10: **Amino acid sequence of NaProPI**. The reactive site of each inhibitor is colored magenta and the disulfide bridging patterns are indicated on top of the sequence. The two disulfide bonds (C8-C37 and C4-C41) anchoring the reactive-site loop to the core of the inhibitor are marked in blue and red, respectively. [From Schirra *et al.*, 2008].

The single IRD CI protein contains a triple stranded β-sheet as the dominant secondary structural element, with several turns and a short region of helix. The reactive site (CI) lies on an exposed loop, which is less defined than the rest of the protein. The overall shape of CI is disk like and the N- and C- terminals are exposed, indicating that this protein results from posttranslational processing of the precursor protein (Greenblatt *et al.*, 1989; Nielsen *et al.*, 1994, 1995). Due to the

high sequence identity between TI and CI IRDs, it has been anticipated that the TI domain also adopts 3D structures similar to CI (**Fig. 1.11**). Single domain PIs having either trypsin inhibitory (TI) or chymotrypsin inhibitory (CI) reactive sites are formed from the multi-domain precursor protein by proteolytic cleavage at the linker regions (Heath *et al.*, 1995; Lee *et al.*, 1999).

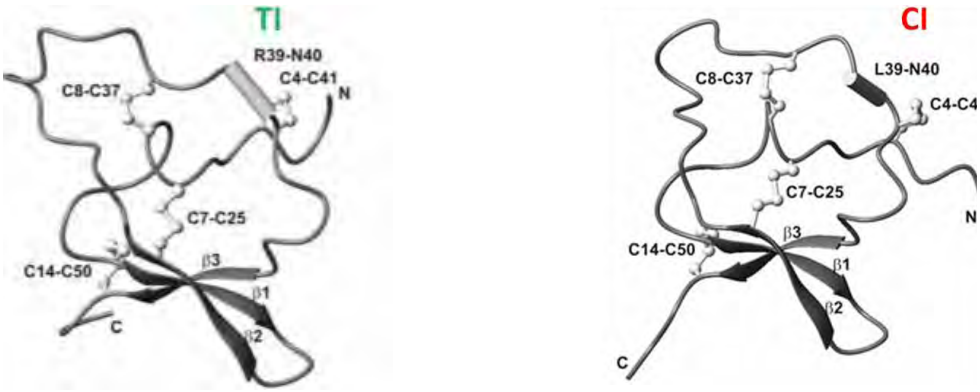


Fig. 1.11: 3-D structure of Pin-II PI proteins. The ribbon drawing of the representative models of solution structure of TI and CI, showing the regular secondary structure [From Schirra *et al.*, 2010].

Apart from these, other highly conserved residues which are structurally important are Pro-18, Gly-38 and Gly-46 as they belong to the three β-turns, respectively (Kong & Ranganathan, 2008). Normally in a TI domain, the RSL is constrained by both C8-C37 and C4-C41 disulfide bonds thereby limiting its flexibility in order to provide it specificity for binding. Recent NMR studies on the disulfide bond variant of TI (T1) from *N. alata*, (C4A/C41A-T1) showed similar conformation of the reactive loop as the normal TI, with moderately decreased inhibitory potential (**Fig. 1.12A**). In contrast, the C8A/C37A-T1 variant showed a disordered and highly flexible RSL making it's binding to trypsin more difficult and therefore, suppressing its inhibitory potency (**Fig.1.12B)** (Schirra *et al.*, 2010). NMR relaxation experiments confirmed the much increased flexibility of binding loop for C8A/C37A-T1 variant and confirmed the indispensability of C8-C37 disulfide bond in maintaining the stability and function of the protein whereas the C4-C41 was not much critical as being supplemented by other stabilizing interactions (**Fig. 1.12C**).

The effect of mutation/variation in the conserved residues is reflected not only in the structure but also in activity of Pin-II proteins. The selective substitution of conserved cysteine residues responsible for disulfide bond formation flanking the reactive site, resulted in poor inhibitory activity (K_i~ 1.8 µM) by C8A/C37A-T1 variant (Cys-C at positions 8 and 37 were replaced by Ala-A). On the other hand, substantial retention of TI activity by C4A/C41A-T1 variant (K_i~ 350 nM) as compared to the wild-type TI (K_i< 5 nM) affirmed the indispensability of C8-C37 bond in Pin-II proteins (Schirra *et al.*, 2010) (**Fig. 1.12)**. Flexibility of the RSL is remarkable in order to allow the binding of Pin-II PIs to wide range of proteinases but also to be specific for efficient binding. It was believed that there were no natural variants for cysteine residues likely because of the essential role of disulfide bonds in Pin-II PIs. Recent studies by Li *et al*. (2011) have identified six natural variants of Pin-II PIs from potato, missing cysteine residues involved in one or two disulfide bonds. The selective loss of cysteine residues at the first reaction center and in pairs has suggested that the loss of disulfide bonds is specific, and not random. Loss of disulfide bonds is often coupled with functional differentiation of the protein (Li *et al.*, 2011). Analysis of more naturally occurring variants can throw light on the evolution of disulfide bonds and the Pin-II PI protein evolutionary process.

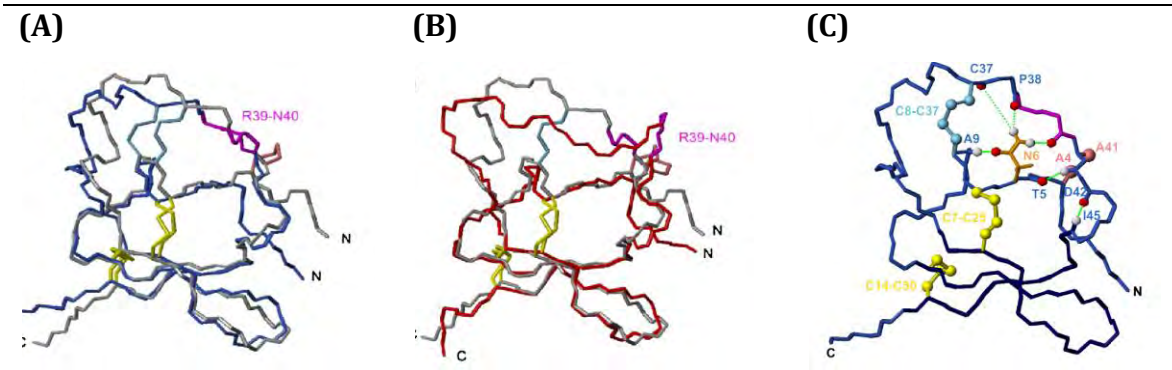


Fig. 1.12: Backbone superposition's of mutant and wild NaTI (A) C4A/C41A-T1 (blue) with the wild-type protein T1 (gray). **(B)** 8A/C37A-T1 (red) with the wild-type protein T1 (gray). C7–C25 and C14–C50 disulfide bonds colored yellow. The C4–C41 disulfide bonds are colored light red, and the C8–C37 disulfide bonds are colored light blue. (**C**) Detailed view of C4A/C41A-T1 showing the stabilization of the reactive-site loop. Hydrogen bonds stabilizing the reactive-site loop are indicated by continuous green lines. [From Schirra *et al.*, 2010]

I.6.3 Effect of dietary Pin II PI on insects

Most of the PIs are inhibitors of serine proteases like trypsin or chymotrypsin, although some of them may exhibit inhibitory activity against elastases, subtilisin-like proteases etc. They do not appear to function as inhibitors of endogenous plant proteinases as most of the latter ones are cysteine proteinases (Reeck *et al.*, 1997). Thus, it is widely acknowledged that serine PIs have a defensive role against herbivores (especially, Lepidopteran) which mainly rely on serine proteinases for digesting the dietary protein components. PIs act as substrate mimics and block the digestive proteinases in the larval gut thereby limiting the release of amino acids from food proteins (Broadway and Duffey, 1986a and 1986b; Hilder and Boulter, 1999). As a consequence, the depletion of amino acids exerts a profoundly detrimental effect on larval physiology and thereby retards the growth and development of the larvae (Broadway and Duffy, 1986; De Leo *et al.*, 2001; Telang *et al.*, 2003; Damle *et al.*, 2005; Tamhane *et al.*, 2007; Hartl *et al.*, 2010). Further, the decreased fertility and fecundity of the adult moths reduce the overall fitness of the insect populations and thus reduce their impact on crop. Additionally, antagonistic effects of PIs act synergistically with other components of the plant defense mechanism *viz.* retarded insects become easier targets for their parasites (Lewis *et al.*, 1997).

Wound induction of Pin-II PIs and their role in plant defense against insect pests has been established by several studies which suggests that induction of such PIs strengthen the defense of plants which are already partially protected by the constitutively expressed PI defense. Tomato PIs inhibited about 50–80% HGP activity of *H. armigera larvae* feeding on various host plants and of various larval instars (Damle *et al.*, 2005). Tomato PIs were found to be highly stable to insect gut proteases and bioassay using *H. armigera* larvae fed on artificial diet containing tomato PIs revealed adverse effect on larval growth and development, pupation, adult formation and fecundity in a dose-dependent manner. This demonstrated that host plant PIs also aid in their defense. The efficacy of *C. annuum* PIs against *H. armigera* gut proteases was demonstrated by Tamhane *et al.* (2005b). *In vitro*

assays showed 68-91%inhibition of trypsin activity and 39-85% inhibition of chymotrypsin-like activity of *H. armigera* gut proteases from larvae fed on various host and non-host plants by *C. annuum* PIs extracted from leaf. *H. armigera* fed on *C. annuum* PI incorporated diet, showed growth retardation as well as reduction in fertility and fecundity for two consecutive generations. Moreover, recombinant expressed diverse CanPIs, with 1- to4-IRDs were characterized for their insect inhibitory potential. *H. armigera* fed on rCanPI diet showed 30% mortality and 40% lower mass among the survivors, in the early instars (Tamhane *et al.*, 2007). Pupal mass reduction of 12-25% was recorded, leading to decreased fecundity. CanPI-7 with two CI sites and two TI sites showed the strongest anti-metabolic effect on *H. armigera.*

Incorporation of *N. alata* PIs in the artificial diet of the native budworm (*H. punctigera*) and the black field cricket (*Teteogryllus commodus*) revealed significant reduction in growth and failed to complete molting than insects on the control diet. *H. punctigera* fed with a cotton leaf-based artificial diet, composed of 0.26% (wt/vol) of *N. alata* proteinase inhibitor (NaPI) showed higher mortality (80%) as well as lower larval mass as compared to those fed with the control diet (Dunse *et al.*, 2010a). Interestingly, the consumption of NaPI by the larvae, led to induction of an insensitive chymotrypsin which was found to be resistant to inhibition by NaPI. The activity of this chymotrypsin was found to be inhibited by a Pin-I inhibitor (StPin1A) isolated from wound-induced leaves of *S. tuberosum*. *H. armigera larvae* fed with diets containing NaPI and StPin1A showed 50% and 40% weight loss, respectively than the larvae fed with control diet. Higher efficacy (90% loss in larval mass) was obtained, when fed with an artificial diet composed of both the types of inhibitor proteins, StPin1A and NaPI (**Fig. 1.13)**.

I.6.4 Insect response to PI

Although many plant species express PIs, including host plants, many lepidopteran pests feed and thrive well on them. An answer to this has been attributed to the adaptation of insect pests to host plant PIs by synthesizing proteinases that are either insensitive to inhibitors (Broadway, 1995, 1996, 1997; Jongsma *et al.*, 1995) or over-expression of proteinases (Broadway, 1997; Gatehouse *et al.*, 1997) or have

the capacity to degrade them (Michaud, 1997; Giri *et al.*, 1998; Harsulkar *et al.*, 1999). Host plants might possess effective PIs,

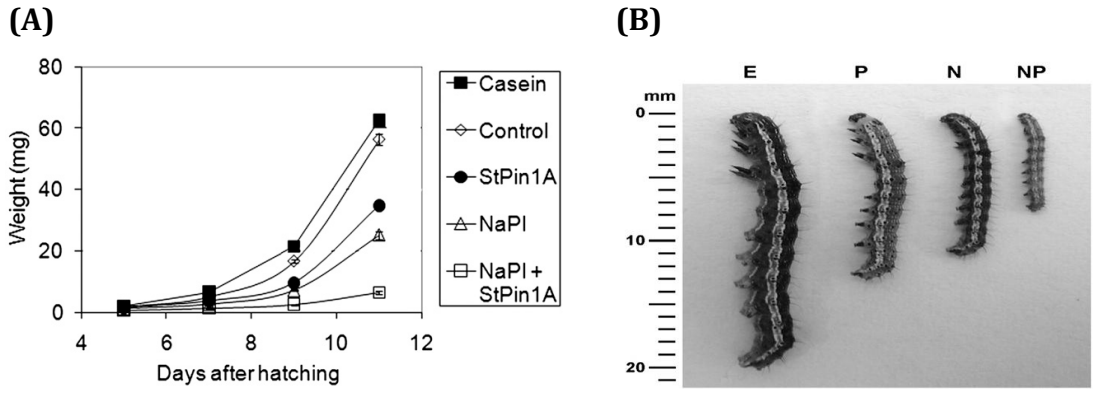


Fig. 1.13: **(A)** Growth and **(B)** phenotype of *H. armigera* larvae on cotton leaf-based diets containing PIs. [From Dunse *et al.*, 2010].

but their effects could be undermined due to low level of expression (Chougule *et al.*, 2003; Damle *et al.*, 2005; Srinivasan *et al.*, 2005). It can be overcome by expressing the host-PI genes under a stronger promoter to increase the inhibitor protein content. On the other hand, "non-host" plants PIs have been found to be effective against insects and responsible for rendering them unsuitable as a food for the insects. *H. armigera* larvae reared on a diet containing non-host PIs (groundnut, winged bean and potato) showed a reduction in total and trypsin-like proteinase activity, and the production of inhibitor-insensitive proteinases (Harsulkar *et al.*, 1999; Giri *et al.*, 2003). Non-host PIs added to larval di*et al*so inhibited growth of *H. armigera* and *S. litura* larvae (Harsulkar *et al.*, 1999; Telang *et al.*, 2003; Tamhane *et al.*, 2005b; 2007). Thus, stability of PIs to proteolytic degradation and the synergistic interaction of different PIs can drastically influence the efficacy of PIs.

The success or failure of the PI strategy can largely depend on the nature of the insect response to the ingested PIs. When incorporated into insect diets or expressed in transgenic plants and fed to insects, the efficacy of protease inhibitors has proved to be variable. High levels of growth retardation and mortality have been observed in some studies, whereas very minor or no deleterious effects seen even

when the inhibitor has been shown to be effective in preventing proteolysis, *in vitro* (Jongsma and Bolter, 1997).

I.7 EVALUATING SYNTHETIC AND NATURAL PROTEASE INHIBITORS AGAINST *H. ARMIGERA* PROTEASES

To overcome the limitation of existing pest management strategies and to ensure a continued supply of high–quality food, efforts are going on to utilize synthetic peptides and other small molecules to develop new insect–control agents.

I.7.1 Small peptide as insecticidal molecules

A] Natural peptide

Various natural peptides from plants and animals were examined for their insecticidal potential. Some of the important candidates are described below (**Fig. 1.14**).

i) Ureases (EC 3.5.1.5) associated peptides

These are nickel-dependent enzymes that catalyze urea hydrolysis resulting in ammonia and carbon dioxide. It has been postulated that these enzymes are involved in both, nitrogen bioavailability and defense processes in plants. Pires-Alves *et al.* (2003) identified a family of urease genes induced by abscisic acid in the legume *Canavalia ensiformis* (jackbean). Canatoxin (CNTX), is one of important member of the family, a toxic protein. CNTX-like proteins and ureases accumulate in the mature seed, suggesting a defense role associated with insecticidal properties. CNTX is lethal when ingested by bruchids (e.g., *Callosobruchus maculatus*) and bugs (Hemiptera, e.g., *Nezara viridula, Dysdercus peruvianus, Rhodnius prolixus*). Jaburetox-2, a recombinant peptide obtained from a truncated cDNA of a *C. ensiformis* urease gene, showed potent insecticidal activity when added at 0.01% w/w to the diets of sensitive insects, and it also killed insects that are not sensitive to native (full-length) ureases (Mulinari *et al*., 2008).

ii) Lectins: These are proteins of non-immune origin, displaying one or more non catalytic domains able to recognize and reversibly bind to free- and/or conjugate carbohydrates, without modification of their covalent structures. Lectins are able to interact with cell membranes of insects, specially glycoconjugates, glycolipids, glycoproteins, and polysaccharides, thus, working as defense proteins. Physiological functions proposed for plant lectins include participating in the colonization and nodulation of legume roots by rhizobial bacteria and participating in defense mechanisms against phytopathogens, nematodes, and phytophagous insects. The deleterious effect of lectins on insects from the orders Coleoptera (beetles), Homoptera (plant lice and aphids), and Lepidoptera (butterflies and moths) has been documented for several lectins, amongst them are the lectins of the Wheat germ agglutinin (WGA) and the snowdrop *Galanthus nivalis* (GNA). Down *et al.* observed that GNA inhibited the growth of the aphid *Aulacorthum solani* and reduced its fecundity by 65%. Among other insecticidal lectins are those from pitomba (*Talisia esculenta*) and *Annona coriacea*. Transgenic plants expressing different entomotoxic lectins are being tested in field (http://www.epa.gov/pesticides/biopesticides/ pips/index.htm). McCafferty *et al.* (2008) reported that a papaya tree transformed with the GNA (*Galanthus nivalis* L. agglutinin) gene expressed a biologically active lectin and presented acceptable control levels for the aphid *Tetranychus cinnabarinus*. Saha *et al.,* (2006) depicted that *Allium sativum* leaf agglutinin expressed in transgenic rice enhanced the resistance against to the major sap-sucking insects: brown plant hopper (BPH), green leafhopper (GLH), and white backed plant hopper (WBPH).

iii) Cyclotides: They are small circular polypeptides composed of 28 to 37 amino acid residues with a head-to-tail cyclic backbone in which the N and C termini are linked via a peptide bond. First reports on these proteins appeared in the early 1970s after studies on the medicinal properties of kalata-kalata, an African plant. Cyclotides contain six conserved cysteine residues connected in a knotted topology, which combined with the cyclic backbone form a structural motif known as the cyclic cystine knot, conferring to these molecules an exceptional stability. Cyclotides

can be classified into two subfamilies, Mobius and bracelet, according to the presence or absence of a twist formation in the backbone of the peptide and the presence of a cis-Pro motif, respectively.

In plants, cyclotides derive from precursor proteins which undergo cleavage and cyclization to produce mature molecules containing one, two or three cyclotide domains. This class of proteins, a group of 200 different cyclotides isolated from plants, include kalata B1, kalata B2, the circulins, cyclopsychotride, and several peptides from Viola species and recently also found in Fabaceae species. Cyclotide-like sequences, at the nucleic acid level, encoding six Cys residues in a similar spacing to that in known cyclotides, have been reported for variety of important monocotyledonous plants, including wheat (*Triticum spp*.), rice and maize (*Zea maize*) and from the leguminous species Butterfly pea (*Clitoria ternatea*). Initial assays with insect pests of the Lepidoptera family fed on artificial diets containing kalata B1 and B2, at concentrations similar to those naturally occurring in plants (0.15% w/v), demonstrated their insecticidal properties. *Helicoverpa punctigera* caterpillars fed on kalata B1-containing diets had 50% mortality rate and those surviving failed to progress past the first instar stage. Electron micrographs showed that this insecticidal activity involved disruption of the midgut membranes of the insects (Craik *et al.*, 2010).

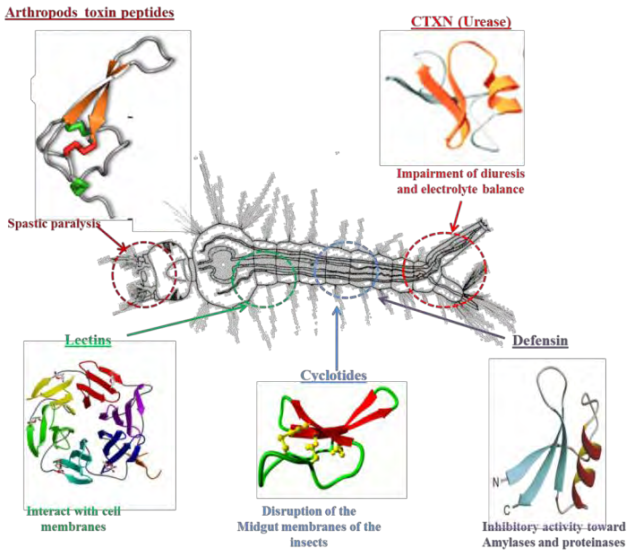


Fig 1.14: Effect of various peptides on insect metabolism and physiology

iv) Defensin: Plant defensins are present in leaves, floral organs, pods, and seeds. Plant defensins, or c-thionins, are basic peptides, with 45–54 amino acids residues, and eight conserved cysteine residues, forming disulfide bridges. They are resistant to adverse conditions, such as extreme pH (pH ~ 2), high temperature (~85°C), oxidation, and proteolytic action. Defensin share some structural features, like a cysteine-stabilized ab motif, with three antiparallel a-strands and one a-helix, except for the PhD1 protein from *Petunia hibrida*, which is a plant defensin with five disulfide bonds. Other residues are variable, leading to a wide biological diversity of defensins so that there are antifungal and bactericidal defensins. Others have no antimicrobial activity, but inhibit a-amylases or protein synthesis, and can play a role in plant defense against insects (Lay & Anderson, 2005).

VrD1 (VrCRP) from a bruchid-resistant mungbean was the first defensin from plant exhibiting *in vitro* and *in vivo* insecticidal activities. Some defensins are inhibitors of α-amylases and proteinases and can play a role in plant defense against insects. Defensins isolated from cowpea seeds display inhibitory activity toward amylases and proteinases. Structural analyses of cowpea defensins have shown a dependence on the high density of surface cationic residues to facilitate interaction with the catalytic site of target enzymes as well as their toxicity against bacteria pathogenic to plants and to humans. The essentiality of arginine and lysine residues for biological activities of defensins underlies attempts to improve the insecticidal or bactericidal properties of these molecules by site-directed mutagenesis (Stotz *et al.*, 2010)

v) Toxins from arthropods: Insect-selective toxins have been identified in venoms from a number of arthropods like bees, wasp, spider and scorpion etc. (Zlotkin, 1985). Insecticidal potential of spider venom peptide, AIP-1 shown by Hardy *et al.* 2013, is likely to be synergized by the gut-lytic activity of the *Bacillus thuringiensis* Cry toxin (Bt) expressed in insect-resistant transgenic crops, and consequently it might be a good candidate for trait stacking with Bt (King *et al.*, 2013). The scorpion venoms have three main groups of polypeptide neurotoxins which modify axonal sodium conductance. Out of these groups, third group of neurotoxins are the

excitatory insect selective toxins which cause an immediate (knock down) spastic paralysis of insects by the induction of repetitive firing in their motor nerves due to an increase of the sodium peak current and the voltage dependent slowing of its inactivation (Walther *et al.*, 1976; Pelhate & Zlotkin, 1981). Similarly peptides from polyhedrosis viruses have been identified as effective insecticidal molecule against lepidopteran insects. Thus, all these reports indicated that there is significant interest in the structure based peptide design approach to develop range of peptide insecticidal peptides with unique modes of action and wider phyletic selectivity.

B] Engineering and synthesis of insecticidal peptide

i) Tuning up the stability and activity of SPIs

In order to tune up the inhibitory potential, efforts have been made to increase protein stability by inserting novel disulfide cross-links. The effect of adding or removing disulfide cross-links on protein flexibility and stability was studied in *Cucurbita maxima* trypsin inhibitor-V (CMTI-V, MI: I07.005) (Qi *et al.*, 2005; Hogg *et al.*, 2003 and Costa *et al.*, 2006). By comparing two models of CMTI-V, it was concluded that cross-links enhance stability. Removing a natural disulfide bond (C3S/C48S) destabilized both the native proteins, while adding a disulfide bond (E38C/W54C) increased the stabilization of the native state. Thus, it is obvious that losing disulfide bond, despite replacing with a hydrophilic group, need not always result in positive effects. The loss of disulfide bond has to be coupled with evolution and selection for it to finally appear in the system. The stabilizing effect of a disulfide bond is due to both enthalpy and entropy contribution. This effect is also modulated by hydration (or hydrophobicity) and flexibility in the residual structure of SPIs (Philipp *et al.*, 1998; Qi *et al.*, 2005 and Hogg *et al.*, 2003). There is negative correlation between protein hydrophobicity and the stability of SPIs. In case of SPIs, the thermodynamic parameters of denaturation reveal varied effects of disulfide bond on protein stability. The processes of enthalpy and entropy changed accompanying alteration in disulfide bond, while protein stabilization was profoundly modulated by the altered hydrophobicity of both, native and denatured states (Hogg *et al.*, 2003; Costa *et al.*, 2006; Swedberg *et al.*, 2011).

ii) Maintenance of the canonical loop in sunflower trypsin inhibitor (SFTI) Three-dimensional structure of SFTI (MI: I12.002, PDB: 1JBL) showed two anti-parallel β-strands stabilized by a single disulfide bond and an extensive internal network of hydrogen bonds. The disulfide bonds or internal hydrogen bond network was modulated to study the importance of the RSL rigidity in inhibition mechanism (Swedberg *et al.*, 2011; Lesner *et al.*, 2011). Substitutions favoring the formation of internal hydrogen bonds are directly correlated with increased inhibition potency (Swedberg *et al.*, 2011, Lesner *et al.*, 2011). Such an understanding can be useful in developing a new range of SFTI inhibitors that are more effective against proteases. For example, against kallikrein-related peptidase (or trypsin), newly engineered SFTI showed 125-fold enhanced inhibitory potency and selectivity. In many cases, the position of the substitution influenced the frequency and distribution of internal hydrogen bonds across the entire scaffold. In case of SFTI, substitutions at position 14 (D14N) led to increase in the intramolecular hydrogen bond density, and this modified inhibitor exhibited excellent attributes, such as stability in a cancer cell milieu and bio-available by intra-peritoneal perfusion in mice, making it an attractive candidate for further therapeutic development (Swedberg *et al.*, 2011; Lesner *et al.*, 2011).

iii) Variation in reactive site loop as template for designing insecticidal tripeptides: The diversity in the number of repeats and diversity within IRD sequences is predominantly observed in Pin-II PIs. Different specificities within a multi repeat protein contribute to a PI cocktail to fight against varied pest/ pathogenic attacks. In various members of Pin-II precursors studied there is a combination of TI/ CI domains. The specificity of a Pin II PIs is governed mostly by a few amino acids in what is called the reactive site loop. Each IRD of the Pin-II precursor contains a single reactive site either for trypsin or chymotrypsin inhibition. In complex with a protease, an array of five amino acids involve in interaction with active site of protease. The two outer residues of these five are cysteines, which are invariable in all natural Pin-II variants. The central three amino

acids vary considerably in identity and found to be extremely flexible in solution. The P1 residue of the reactive site determines its specificity; presence of lysine "K" or arginine "R" in the P1 position confers TI potential and Leucine "L" bestows the inhibitor CI potential (**Fig. 1.11)**. The profound docking of the side chain of P1 residue in the S1 binding pocket of the protease plays an extensive role in the energetics of the specificity of PI protease interaction. Beekweilder *et al.* (2000) showed the role of aa variation of two reactive sites in the inhibition of trypsin and chymotrypsin by mutating each of the two reactive sites in various ways. Mutational studies on PI-II from potato have highlighted the importance of secondary contacts also not involving RSL as well, in determining the specificity of protease inhibition. In PI-II from tomato, the inhibition capacity of a TI domain could not be transferred to the other domain by mutating the P1 residue or the residues within the RSL (Beekwilder *et al.*, 2000; Schirra & Craik, 2005).

I.7.2 Plant secondary metabolites as insecticidal molecules

Plants have evolved a whole arsenal of defensive chemical compounds against herbivores. Chemical defense products may range from low molecular weight compounds, called secondary metabolites, to peptides and proteins that are active against insects (**Fig. 1.15**). Secondary metabolites are organic compounds that are not directly involved in the normal growth, development, or reproduction of plants (Pichersky *et al.*, 2000). The insecticidal properties of these compounds (hereafter called "plant toxins") are diverse: they may act as repellents or feeding deterrents, or induce direct toxicity leading to symptoms ranging from the inhibition of larvae or insect growth to death. Production of secondary metabolites in response to insect herbivory is one of the defense strategies of plants. Insect growth inhibitory effect of plant secondary metabolites exerts weak selection pressure and thus avoids the chances of emergence of resistance against these compounds (Boerjan *et al.* 2003). In some cases, recent advances in molecular biology have made it possible to accurately identify the cellular or molecular targets of these toxins. Another active field of investigation is the elucidation of the plastic nature of such defense responses, modulated both by phytophagous (and pathogens) attacks and abiotic

factors including nutrient availability, light and drought (Bennett *et al.*, 1994). Some of these compounds are always synthesized in the plant (constitutive resistance) while others are produced only after initial damage (induced resistance). Both constitutive and induced resistance have been shown to generate costs, described as allocation costs, resource-based tradeoffs between resistance and fitness, or as ecological costs, decreases in fitness resulting from interactions with other species. Where there is limited nutrient availability plants may accumulate secondary metabolites (Bennett *et al.*, 1994).

Because plants can produce a nearly inexhaustible number of metabolites, they possess an enormous reservoir of potentially defensive compounds, many of which have been described in the context of plant interactions with other organisms. These compounds belong to various chemical classes such as isoprene derived terpenoids including mono-, sesqui-, di-, and triterpenoids as well as steroids; N-containing alkaloids; phenolic compounds including flavonoids; and others. These compounds also differ in their structures, indicating the presence of different target structures (**Fig.1.15**). In addition, some compounds occur ubiquitously, whereas others are restricted to certain taxa, for example, cocaine is specific to the genus *Erythroxylum*, suggesting either broad bioactivity or functions in particular interactions. To minimize the risk of self-intoxication, many defense compounds are usually stored in compartments of limited metabolic activity, such as the vacuole or the apoplasm. This is obvious for alkaloids as well as phenolic substances.

In contrast to the large number of specialized compounds whose involvement in plant defenses against herbivorous insects and other arthropods is known, the exact mode of action on a molecular level as well as the related target structures of these compounds are still unknown. As a result, not all the different compounds or classes of compounds mentioned in the introduction can be discussed in detail, but some case studies are addressed in the following.

In general, the mode of action frequently includes membrane disruption, inhibition of nutrient and ion transport, inhibition of signal transduction processes, inhibition of metabolism, or the disruption of hormonal control of physiological

processes. Saponins such as avenacosides have an amphiphilic character and can disrupt cellular membranes (Rattan *et al.*, 2011). Cardenolides (cardiac glycosides) are specific inhibitors of the Na^+/K^+ATPase that maintains the electric potential in animal cells, from human to Drosophila (Dobler *et al.*, 2012). Cicutoxin, a polyacetylene, prolongs the repolarization phase of neuronal action potentials, very likely by blocking voltage-dependent potassium channels. Phytoecdysteroids represent a group of plant compounds that mimic insect hormones, ecdysteroids (including ecdyson), and interfere with the regulation of the periodical molting process (Rattan *et al.*, 2010).

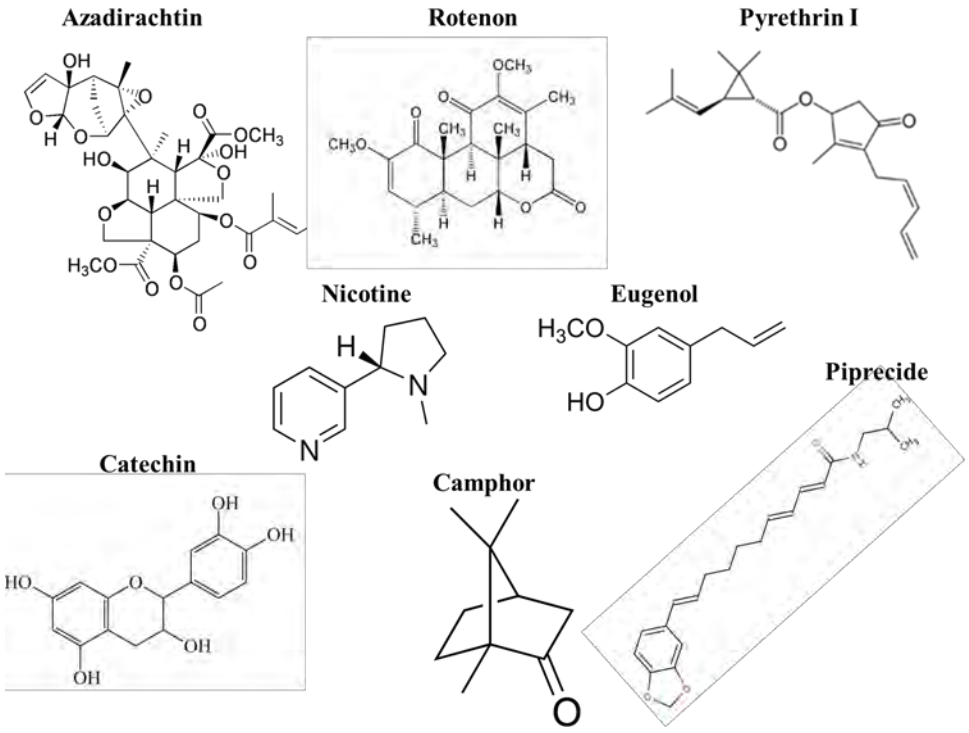


Fig. 1.15: Structures of insecticidal plant secondary metabolites

Without doubt, a plant's need to invest in defenses is costly regardless of whether the defense is constitutive or inducible. The costs are different with respect to the compounds synthesized, e.g., phenolics are suggested to be cheaper than

alkaloids and other nitrogen containing secondary metabolites because of the additional effort required for inorganic nitrogen to be made bioavailable. Ubiquitous and metabolic cost effective nature of phenolic compounds represents them as vital molecules in direct and constitutive defense of plant against herbivory.

A] Plant phenols: Plant phenols exhibit dynamic expression and distribution pattern in response to insect attack. (Sharma *et al.* 2009; Usha Rani and Jyothsna 2010; Ballhorn *et al.* 2011, Maffei *et al.* 2007; Barakat *et al.* 2010; He *et al.* 2011; War *et al.* 2012). Toxic or deterrent activity of phenols against insect pests facilitates the direct defense, while in indirect defenses they attract natural enemies of insect pests (Karban and Baldwin 1997; Heil 2008; Sharma *et al.* 2009; Barakat *et al.* 2010; War *et al.* 2012). Pivotal role of natural phenols in plant–environment interactions is evidenced from plants defense responses against biotic and abiotic stresses, including ultraviolet radiation, pathogen infection and herbivore damage (Simmonds 2003; Treutter 2006). Plant phenolic compounds are categorized at polyphenols and simple phenolic compound. Polyphenols are subdivided as tannins and flavonoids, while Simple phenolic are further categorized in phenolic acid and coumarins. In phenolic acid, hydroxybenzoic acid and hydroxycinnamic acid (HCA) contributes to population of compounds. HCA serve as precursor for synthesis of plant phenolic acid compounds. Recent reports showed that plant phenolic compound leads to reduction in palatability of the plant tissues to the herbivore and thus results in their growth retardation (Boerjan *et al.* 2003, Berenbaum 1995; Appel 1993, Treutter 2006).

Modern agricultural practices are depended on the extensive use of chemical pesticides which leads to resistance and resurgence in insects, also reason for substantial environmental and human health problems (Ahmad *et al.*, 2008). This scenario highlights the need of development and application of natural product-based and ecofriendly pesticides. Insect growth retardant potential of the plant phenols along with their anticancerous and antioxidant properties might provide us excellent natural compound library to formulate eco-human friendly "Dietary pesticides" (Ames *et al.*, 1999). Application of these molecules in context of pesticide

development, it is essential to have mechanistic insights about their mode of action as insect growth retardant molecule.

I.8 GENESIS OF THE THESIS

Keeping this in view following objectives were formulated to design molecules that can inhibit individually or in combination the compliment of diverse gut proteases of *H. armigera.*

I.8.1 Objectives

1) Biochemical and biophysical characterization of inhibitory repeat domains (IRDs) of *C. annuum* PIs
2) Assessment of distinct IRDs for their activity against *H. armigera* proteases *in vitro* and *in vivo*
3) Deciphering structural attributes of CanPIs, IRDs and their interactions with *H. armigera* proteases
4) Evaluating small synthetic peptides and natural molecules for their potential against *H. armigera* proteases

I.8.2 Organization of the thesis

Chapter I: Review of literature

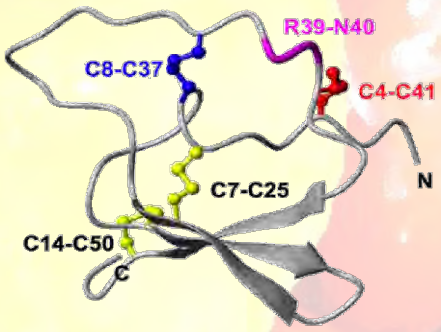
Chapter II: Characterization of inhibitory repeat domain (IRDs) of *Capsicum annuum* protease inhibitors

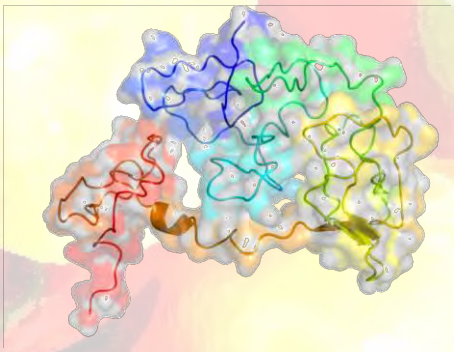
Chapter III: Structural and functional insights of single and multi-domain CanPIs and their interactions with proteases

Chapter IV: Evaluating natural and synthetic protease inhibitors against *H. armigera* proteases

Chapter V: General discussion and future direction

Chapter II: Characterization of Inhibitory Repeat Domain (IRDs) of Capsicum annuum Protease Inhibitors (CanPIs)





Chapter II: Characterization of inhibitory repeat domain (IRDs) of *Capsicum annuum* protease inhibitors (CanPIs)

Capsicum annuum expresses diverse Pin-II PIs comprising IRD as a basic functional unit. Most IRDs contain eight-conserved cysteine forming four disulfide bonds, which are indispensible for their stability and activity. The functional significance of evolutionary variations in IRDs and their role in mediating interaction between the inhibitor and cognate protease was investigated. Among the 55 IRDs encoded by *C. annuum*, IRD-7, -9, and -12 were selected for further characterization on the basis of variation in their reactive site loop, number of conserved cysteine residues, and higher theoretical ΔG_{bind} for interaction with *H. armigera* trypsin. Moreover, inhibition kinetics showed that IRD-9, despite loss of two disulfide bonds, was more potent among the three selected IRDs. Reduction in IRD-9 inhibitory activity after chemical modification of the serine residues suggests that C7S and C8S variation in IRD-9 has significant role in higher inhibitory potential. Furthermore, IRD-9showed special attributes like stability to proteolysis and synergistic inhibitory effect with other IRDs. Importantly, *H. armigera* fed on AD containing 5 TIU/g of recombinant IRD proteins exhibited negative impact on larval growth, survival rate and other nutritional parameters. Major digestive gut trypsin and chymotrypsin genes were down regulated in the IRD fed larvae, while few of them were up-regulated, this indicated alterations in insect digestive physiology. These results corroborated with protease activity assays and zymography. Altogether, these findings suggest that the sequence variations among IRDs reflect in their *in vitro* and *in vivo* efficacy against proteases, which could be used for developing tailor-made multi-domain inhibitor gene(s).

II.1 INTRODUCTION

Plants and insects have co-evolved in order to survive in their ever-changing niche. While insects evolved their proteinase enzymes to derive maximum nutritional benefit, plants evolved upregulating the expression of defense related molecules to counter it (Green & Ryan, 1972; Karban, 1989; Kessler & Baldwin, 2002). The

agronomically adverse insect pest *Helicoverpa armigera* (Lepidoptera: Noctuidae) has been widely studied for its polyphagy and adaptability on various host plants. In order to sustain on chemically varied dietary content, insects acquired molecular flexibilities of their gut enzymes (Zalucki *et al*., 1986; Patankar *et al*., 2001; Sarate *et al*., 2012).

Numerous investigations corroborate the fact that plants rely upon plant proteinase inhibitors (PIs) for protection against attack by phytophagous pests. PIs are ubiquitous in the plant kingdom and have been extensively studied as plant defense molecules, which inhibit hydrolytic enzymes of the insect gut (Green & Ryan, 1972; Ryan, 1990). In some systems, the role of PIs in plant defense is directly evident; whereas in others, the PI diversity and its differential expression may suggest their endogenous functions (Sin & Chye, 2004; Zavala *et al*., 2004a; Zavala *et al*., 2004b; Damle *et al*., 2005; Johnson *et al*., 2007). Among various PI families, Serine PI Pin-II/Pot-II family displays a remarkable structural and functional diversity at the gene and protein level (Kong & Ranganathan, 2008). Wound, herbivory and stress induced up-regulation of these PIs clearly link them to plant defense (Green & Ryan, 1972). Previous studies using transgenic systems or *in vivo* assays have positively correlated the advantage offered by Pin-II PI expression in plants against insect attack (Johnson *et al*., 1989; McManus *et al*., 1994; Duan *et al*., 1996). Recently, a combination of Pin-II PIs from *Nicotiana alata* and *Solanum tuberosum* expressed as transgenes in cotton and tested at the field level proved to enhance the productivity by 30% due to reduction in pest infestation (Dunse *et al*., 2010).

Precursor proteins of Pin-II PIs consist of 1- to 8- inhibitory repeat domains (IRDs) connected by proteolytic-sensitive linkers, which release IRD units upon cleavage. Each IRD is a peptide of around 50 amino acid (aa) length with a molecular mass of ~6 KDa. The aa sequence of IRDs shows variations, at the same time the 8 cysteine residues that form disulfide bridge are conserved (Nielsen *et al*., 1995; Scanlon *et al*., 1999; Lee *et al*., 1999; Schirra *et al*., 2001). One structural feature of

Pin-II IRD is a disordered loop with triple stranded β sheet scaffold. The disordered solvent exposed reactive loop is anchored by the four conserved disulfide bonds (C4-C41, C7-C25, C8-C37 and C14-C50) (Schirra *et al*., 2005; Schirra *et al*., 2008). Among the four disulfide bonds, C8-C37 has been found to be very crucial for maintaining active conformation, whereas C4-C41 has an important role in maintaining the flexibility of the reactive loop (Schirra *et al*., 2010). Thus, any selective loss of disulfide bond is expected to have evolutionary significance leading to functional differentiation of inhibitors (Li *et al*., 2011).

Each IRD possesses a single reactive site, which inhibits trypsin or chymotrypsin depending on whether a Lys/Arg or a Leu residue is present at the P1 position. In Pin-II precursor, the IRDs are capable of simultaneously inhibiting single or several protease molecules (Barrette *et al*., 2003a; Barrette *et al*., 2003b). In a well-known standard mechanism of protease inhibition by Pin-II PIs, the convex shaped reactive site of the inhibitor (P1 side chain) is recognized by the concave active site (S1 binding pocket) of the enzyme in a substrate like manner and plays a major role in the energetics of recognition (Czapinska *et al*., 1999). Protease-PI interaction is an entropy driven process and is further influenced by non-contact residues of the inhibitor by means of van der Waal's interaction and hydrogen (H) bonding (Otlewski *et al*., 2001). The structure of Pin-II IRDs or two domain PIs in complex with protease displays molecular framework of the protease-PI interaction (Barrette *et al*., 2003a; Dunse *et al*., 2010). Structural studies on unbound Pin-II inhibitor provide insight into the conformational flexibility of the reactive loop and its role in modulating protease binding efficiency (Barrette *et al*., 2003b).

Previous studies have shown that Pin-II PIs from *C. annuum* (CanPIs) and their recombinant forms exhibit anti-metabolic effects on *H. armigera* by inhibiting larval growth and development (Antcheva *et al*., 2001; Tamhane *et al*., 2005; Tamhane *et al*., 2007). CanPI precursor proteins interact with the gut proteases of *H. armigera* and get processed into their constituent IRDs (Mishra *et al*., 2010). Eighteen unique IRDs with aa variations in the reactive loop and/or in the number of cysteine

residues have been identified (Tamhane *et al*., 2009). Recent findings have led to the emergence of a hypothesis that functional specifications of the Pin-II PIs are closely associated with their sequence and structural variations (Schirra *et al*., 2010; Li *et al*., 2011). For investigating protease-PI interaction and identifying a potent PI, it is important to study inhibition potential of each PI. However, no previous study has specific details about the structure of *H. armigera* gut trypsin or chymotrypsin or their interaction with CanPIs. In this chapter the recombinant expression, inhibition mechanism and other biochemical characteristics of CanPI IRD(s) have been reported. Moreover the effect of IRD(s) sequence variation, with special reference to variants with different number of cysteine residues has been studied on the protease-PI interaction by using molecular dynamics simulation.

II.2 MATERIALS AND METHODS

II.2.1 Materials

All reagents, enzymes and substrates were obtained from Sigma-Aldrich, St. Louis, MO. Sterile plastics ware from Abdos, WB, India; expression vector pPIC9 and P. pastoris GS115 from Invitrogen (Invitrogen, Carlsbad, CA, USA); Bradford reagent and electrophoresis reagents were from Bio-Rad Laboratories, Hercules, CA; X-ray films and Kodak 163 DA developer were purchased from Kodak, Chennai, India; HIC matrix i.e. Phenyl Sepharose and disposable PD-10 Desalting Columns were from GE Healthcare Life Sciences, Uppsala, Sweden

H. armigera larvae were collected from Mahatma Phule Krishi Vidyapeeth, Rahuri, MS, India, and reared on Artificial Diet (AD). Laboratory conditions were maintained at humidity 60%, temperature 28 °C and photoperiod of 16 h light: 8 h dark for one generation as described earlier (Tamhane *et al.*, 2005).

II.2.2 Selection and interaction analysis of IRDs

Eighteen sequentially unique IRD(s) were identified from 21 *CanPI* genes previously reported from our laboratory, (Tamhane *et al*., 2009). Phylogenetic analysis of these

IRDs was carried out using MEGA5 software (http://www.megasoftware.net/). Insect proteases were selected on the basis of their sequence divergence obtained from phylogenetic and sequence analysis. The sequences for trypsins, chymotrypsins and cathepsins were taken from Uniprot database (http://www.uniprot.org/). Three-dimensional models for IRDs and *H. armigera* proteases were generated using a protein structure prediction server (http://ps2.life.nctu.edu.tw/), which implements an approach to comparative modeling by satisfying spatial restraints (Chen *et al*., 2006). Sequence similarity search was performed for the selected model using PSI-BLAST against a database of known protein structures with default parameters for validation. Taking into consideration the maximum query coverage, the desired model was preferred. Predicted models were validated by MOLPROBITY (Chen *et al*., 2010).

A docking study was performed to determine the binding energy and interaction of IRDs with *H. armigera* proteases. Predicted structures of proteases were refined by energy minimization and restraint relaxation using Swiss PDB-Viewer (v4.1.0) (Guex *et al*., 1997). In order to perform molecular docking, models of *H. armigera* proteases and inhibitors were submitted to Patchdock online server (http://bioinfo3d.cs.tau.ac.il/PatchDock/) following the standard package protocols and further refined by FireDock online server (Schneidman-Duhovny *et al*., 2005). Binding energy obtained for each complex was normalized by mean values and represented in heat map format using MeV software packages (http://www.tm4.org/mev/). The gradient ruler from -5 to 5 is an indicator of interaction strength. Data were clustered using hierarchical clustering method (Saeed *et al*., 2006).

II.2.3 Cloning, expression and purification of IRD(s)

The mature peptide region of selected IRDs were cloned into expression vector pPIC9 for recombinant, extracellular expression in *P. pastoris* GS115 and purified by hydrophobic interaction chromatography as described previously (Tamhane *et al*.,

2005). The purified proteins were quantified by Bradford reagent and checked for purity on 15% Tricine-SDS-PAGE.

II.2.4 Inhibition assay and kinetics

H. armigera larvae were reared on AD and whole-gut tissue was dissected out from fourth instar larvae. *H. armigera* gut proteases (HGP) were extracted from 2 gm of gut tissue by homogenizing in 0.2 M Glycine-NaOH buffer, pH 10.0 in 1:1 ratios (w/v) and kept at 4 °C for 2 h (Tamhane *et al*., 2005). The suspension was centrifuged at 13,000 g, 4 °C for 20 min and the resulting supernatant was used as a source of gut proteases of *H. armigera* (HGP). Total proteolytic activity of 50 mM bovine trypsin/HGP and inhibition of their activity by IRDs (5 µg) was measured by Azocasein assays. Trypsin-like activity of the HGP and its inhibition by IRDs was also estimated using chromogenic substrate Benzoyl-L-arginyl*p*-nitroanilide (BA*p*NA). BA*p*NA assays were performed as described previously (Tamhane *et al*., 2005) and HGPI units of all the IRDs were determined. HGP inhibitory (HGPI) unit is defined as the amount of protein that will inhibit 1 unit of HGP activity using BA*p*NA as a substrate at 37°C, pH 7.8. Minimum three replicates of each experiment were performed.

Michaelis-Menten constant (K_m) for trypsin was calculated by using various concentrations of BA*p*NA substrate (1 to 5 mM), and then plotting double reciprocal curve with 1/(v) and 1/[S]. The kinetic properties of IRDs were analysed over a range of concentration of inhibitors (1 µM to 1 mM). IC_{50} values for each inhibitor were calculated from the sigmoid curve indicating the best fit for the percentage inhibition data obtained. The values of Ki values for each inhibitor were calculated directly from IC_{50} values using Cheng-Prusoff's classical equation (Copeland *et al*., 1995).

$$K_i = \frac{\left(IC50 - \frac{[Et]}{2}\right)}{\left(1 \frac{[S]}{Km}\right)}$$

Inhibition of bovine trypsin, chymotrypsin, elastase, cathepsin and HGP activities were determined using Azocasein as a substrate as described earlier (Tamhane *et al*., 2005) and inhibitory units of all the IRDs were determined. Inhibitory unit is defined as the amount of protein that will inhibit 1 unit of protease activity using azocasein as a substrate at 37°C, pH 7.8. Minimum three replicates of each experiment were performed.

II.2.5 Biochemical characterization of IRD(s)

II.2.5.1 Effect of reducing agents

For elucidating the role of disulfide bonds in the activity, IRD proteins were treated with β-mercaptoethanol followed by heating. These preparations were checked for inhibitory activity by gel X-ray film contact print technique (GXCT) (Pichare and Kachole, 1994). For this, 0.5 HGPI units from each sample were separated on the 15% native-PAGE gel. After electrophoresis, gel was equilibrated with 0.1 M Tris-HCl buffer (pH 7.8) for 10 min followed by incubation in 0.04% trypsin for 10 min and Tris-HCl wash for 2 min. The gel was exposed to X-ray film for the time intervals of 5, 10 and 15 min, respectively. The films were washed with warm water and inhibitory activity bands were visualized as unhydrolyzed gelatin on the X-ray film (Pichare and Kachole, 1994).

II.2.5.2 Estimation of free thiol content by Ellman's assay

Protein (2 µg) was mixed in 100 µl of Tris-HCl buffer (pH 7.8); to this Ellman's reagent (50 µl) and MQ water (840 µl) were added. The mixture was incubated at 37 °C for 10 min and absorbance was measured at 412 nm. The concentration of free thiol content [RSH] of sample was calculated using the following equation (Aitken *et al*., 1996).

$$\Delta A_{412} = E_{412}TNB^{2-}\,[RSH]$$

where, $\Delta A_{412} = A_{final} - (3.1/3.2)\,(A_{DTNB} - A_{buffer})$

and, $E_{412}TNB^{2-} = 1.415 \times 104\ cm^{-1}M^{-1}$

II.2.5.3 Effect of temperature

Each IRD protein (5 µg) was heated from 60 to 100 °C for 15 min. The treated samples were then used for inhibition assay using BA*p*NA and trypsin inhibition was estimated throughout the above mentioned range of temperature.

II.2.5.4 Proteolytic stability and HGP inhibition visualization

To study the interaction and stability of PIs with HGP *in vitro*, 0.5 HGPI units of individual IRDs (IRD-7, -9 and -12) were incubated with 0.5U HGP for two time points (5 min and 1 h) at 24 °C. These HGP- treated PIs were resolved on native-PAGE and processed for TI activity visualization as described above. This mixture of protease and PIs was also used for visualizing the remaining protease activity of HGP in the presence of inhibitor, on 8% native-PAGE using GXCT.

II.2.5.5 Combinatorial inhibition assay

In nature, PIs comprise different combination of IRDs. IC_{50} concentrations of each IRD(s) were used to formulate various combinations of IRDs to check their synergistic effect on HGP/bovine trypsin inhibition potential. The inhibition assay was carried out as already described. Four different formulations i.e. IRD-7+9, IRD-7+12, IRD-9+12 and IRD-7+9+12, were used for inhibition assay.

II.2.6 Modification of serine residues with phenylmethylsulfonyl fluoride (PMSF)

PMSF specifically binds irreversibly and covalently to serine and thus blocks its function. The reaction mixture containing inhibitor (IC_{50} concentration of each IRD in 1 ml) in 100 mM Tris-HCl buffer, pH 7.8 and 5, 10 and 15 mM each of PMSF was incubated at 30 °C for 1 hr. Aliquots were removed at various time intervals (15, 30, 45 and 60 min) and desalted thrice using disposable PD-10 desalting columns to remove excess PMSF from the samples. Residual activities of these modified inhibitors were determined under standard assay conditions. Inhibitor sample incubated in the absence of PMSF served as control. Pseudo first-order plot for

inactivation of IRDs by PMSF and second order plot of pseudo-first order rate constants (K app) (min^{-1}) as a function of log of PMSF concentration were plotted using the equations from Koller *et al*. (1982).

$$1/k_2(\text{Observed}) = 1/k_2 + [H+]/k_2K_{app}$$

Where, k2 is the second-order rate constant for modification of the residue. Inhibitory activity of untreated and PMSF treated (10 mM for 2 h at 30 °C) inhibitors were also assessed using GXCT method as described above.

II.2.7 Feeding assay

Bioassays were conducted by feeding *H. armigera* larvae on AD containing PIs (Tamhane *et al*., 2005). AD was supplemented with the recombinant IRD proteins in appropriate quantities to give equal TI units (5 TIU/g of feed). The neonates obtained from lab-reared moths were reared on control diet for the first 2 days and then transferred to IRD-containing diets and control diet (AD with BSA added in equal to PI amount). Feeding bioassay was performed in separate sets of 25 larvae each with three replicates of each set. Larval weights were meticulously recorded every alternate day. Weight reduction in the PI fed larvae was compared to that of the control group. The larval mortality and weights were recorded and compared with that of the control group. Different growth and survival parameters were used to estimate the effects of IRDs on the growth and development of *H. armigera.*

II.2.8 Nutritional parameters

Various nutritional parameters of fourth-instar larvae exposed to control and recombinant IRD-containing diet were compared. The feces and uneaten food were dried and weighed. Nutritional indices of consumption, digestion and utilization of food were calculated, as described previously (Farrar *et al*., 1989). The nutritional indices, namely efficiency of conversion of ingested food (ECI), efficiency of conversion of digested food (ECD) and approximate digestibility (Ad) were calculated as follows

$$ECI = (\Delta B/I) \times 100$$

$$ECD = [\Delta B/ (I - F)] \times 100; \text{ and}$$

$$Ad = [(I - F)/ I] \times 100$$

Where, I = weight of food consumed, Δ B = change in body weight, and F = weight of feces produced during the feeding period.

II.2.9 Estimation and visualization of proteinase activity

The gut tissue of recombinant IRD fed insects was homogenized and mixed with 0.2 M glycine-NaOH buffer (pH 10), and kept at 4°C for 2 h. The homogenate was centrifuged at 4°C for 20 min at 10,000 *g;* supernatant was used as HGP. Enzymatic assays using azocasein as substrates were performed in order to estimate proteinase activity of experimental tissue (Tamhane *et al*., 2005).

Equal activity units (0.5 TU) of HGP from insect fed on AD containing recombinant IRD-7, -9 and -12 were resolved on 10% native-PAGE; the gel was further processed for activity visualization using the gel X-ray film contact print technique (GXCT) (Pichare and Kachole 1994). To analyze the stability and residual inhibitory activity of IRDs in HGP, another set of gel was incubated in Tris-HCl buffer for 10 min, followed by 0.04% trypsin for 10 min and then washed with the buffer. Inhibitor activity visualization was carried out as described previously (Tamhane *et al*., 2005).

II.2.10 Quantitative real-time PCR

Total RNA was isolated from the inhibitor-fed insect gut tissues using Trizol reagent (Invitrogen, Carlsbad, CA, USA) and synthesis of the first strand cDNA was carried out with High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Foster, CA, USA) using random primers based on the manufacturer's protocol.

Relative transcript abundance of trypsins and chymotrypsins was determined by quantitative Real-Time PCR (qRT-PCR) using 7900HT Fast Real-Time PCR System (Applied Biosystems, Foster, CA, USA) and Fast start Universal SYBR Green Master (Rox) 2× concentrate (Roche Diagnostics, GmbH, Germany). The relative expression of trypsin and chymotrypsin genes (HaTry 1, 2, 3, 4, 5, 6, 7, 8; HaChy 1, 2, 3, and 4) was assessed (Chikate *et al*., 2013). For each gene, amplification efficiency was assessed by constructing a standard plot using 5 serial

dilutions of cDNApool which were prepared by combining aliquots from all the cDNA samples under study. β-actin (Accession no.: AF286059) was used as a reference gene for normalization. Reaction mix for qRT-PCR and thermal cycler conditions were followed as described by (Chikate *et al*., 2013). For each treatment/time point, two biological replicates (each constituting three technical replicates), the average transcript abundance and sub-sequent fold difference with respect to the control were calculated.

II.2.11 Statistical analysis

All data were statistically analyzed by independent sample t-test. Asterisks indicate significant differences (*, $p<0.05$; **, $p<0.01$).

II.3 RESULTS AND DISCUSSION

II.3.1 Sequence and structural variation in IRDs

Phylogenetic (**Fig. 2.1A**) and multiple sequence alignment (**Fig. 2.1B**) analysis of CanPI IRDs showed significant divergence due to sequence variations in the reactive loop regions and in the number of cysteine residues. Heat map provides an overview of the binding energetics of all the 18 IRDs with target proteases. The data indicated that IRD-7, -9 and -12 bind more strongly to HaTry compared to the other IRDs and thus selected for further analysis (**Fig. 2.1C**). The multiple sequence alignment of IRD-7, -9 and -12 with *N. alata* IRD (NaTI) showed over 90% sequence identity (**Fig. 2.1D**). In case of Pin-II PIs the major variation is found in the reactive loop (Kong and Ranganathan, 2008). The residues in the reactive loop of IRD-7 and -9 is "CPKNC", whereas in IRD-12 is "CPRNC". Another crucial variation is in the number of cysteine residues present. The number of conserved cysteine residues in IRD-7 and -12 is eight while the same in IRD-9 is only six, making the latter one unique among IRDs. Two cysteines at 7^{th} and 8^{th} position of IRD-9 are replaced by serine residues, disrupting two disulfide bonds.

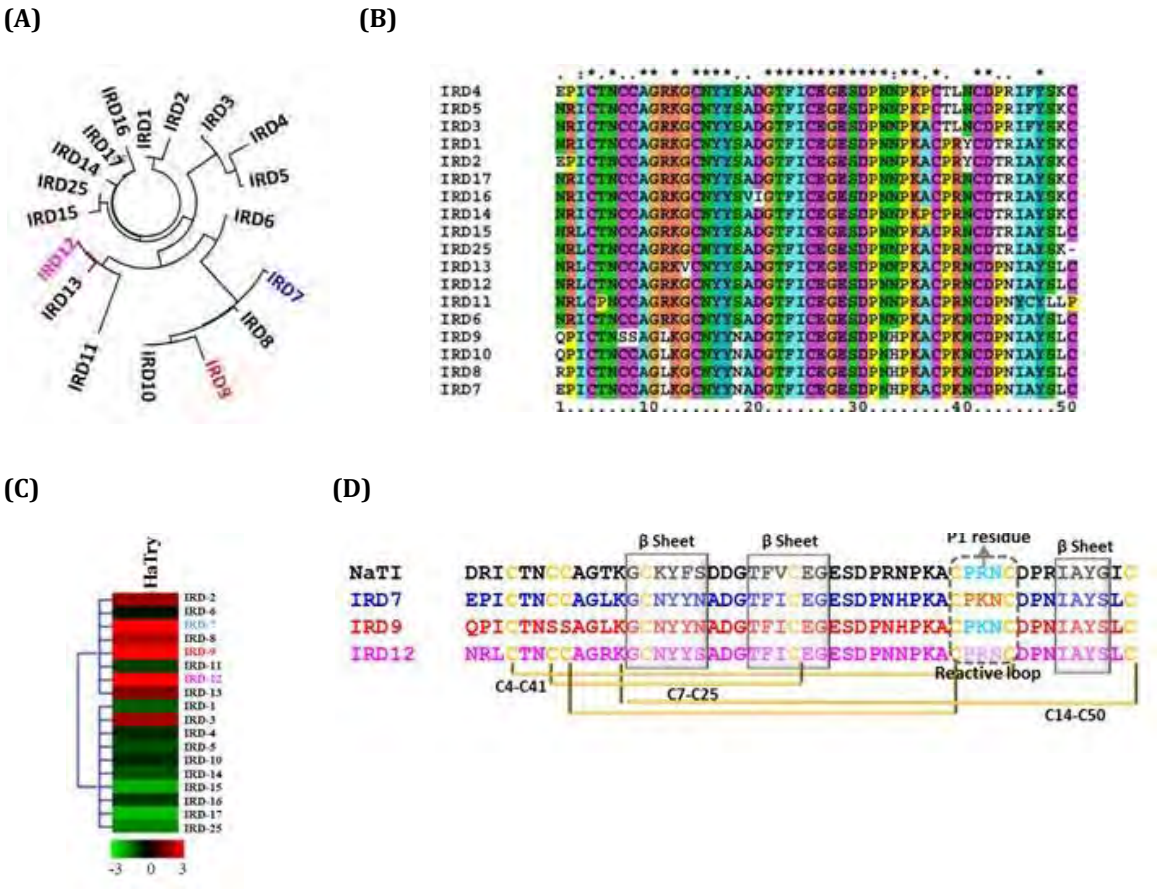


Fig. 2.1: Sequence and structural diversity of IRDs. (A) The circular dendrogram of deduced AA sequences of IRDs. The three IRDs studied here are marked in colors as IRD-7: blue, IRD-9: red and IRD-12: magenta. **(B)** Multiple sequence alignment of 18 IRDs encoded by *C. annumm* **(C)** Heat map analysis of the binding energetics of all the 18 IRDs with target proteases. **(D)** Multiple sequence alignment of IRD-7, -9, -12 and *Nicotiana alata* trypsin inhibitor (NaTI) using ClustalX2. Conserved cysteine residues are marked in yellow and the reactive loop region (residue 37-41) including the P1 residue is indicated with cyan. The number and position of cysteine residues are conserved in all except IRD-9 in which the cysteines at 7th and 8th positions are changed to serine. The sequence of the reactive loops of IRD-7 and -9 is "CPKNC" whereas that of IRD-12 and NaTI is "CPRNC".

II.3.2 Inhibition kinetics and biochemical characterization of IRD-7, -9 and -12 indicates variation in their properties

IRD-7, -9 and -12 were extracellularly expressed in *Pichia pastoris* and the soluble fraction in each case yielded the single protein band in each case corresponding to ~5.8 kDa on 15% Tricine-SDS-PAGE (**Fig. 2.2A**). Assays using BA*p*NA and azocasein as substrates showed that IRD-9 and -12 inhibited about 80 to 85% of HGP activity

while inhibition by IRD-7 was only 40 to 45% (**Fig.2.2B**). Both the substrates showed low inhibitory efficiency by IRD-7 and highest proficiency by IRD-9.

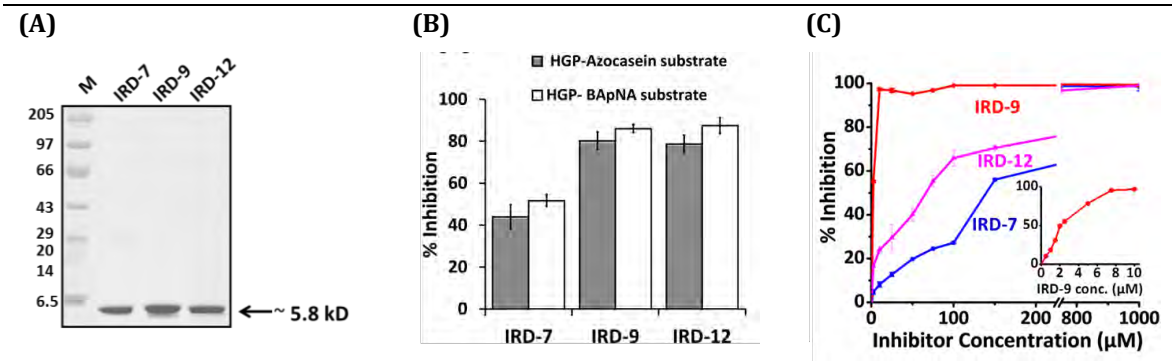


Fig. 2.2: Inhibition studies for IRD-7, -9 and -12 (A) Purified protein preparations of IRDs; IRD-7, -9 and -12 show single band of approximately ~5.8 kDa on 15% Tricine-SDS-PAGE. **(B)** Inhibition of HGP activity with 10 µg of BA*p*NA and azocasein substrate. **(C)** Estimation of IC_{50} and Ki values by using inhibition of bovine trypsin with various concentrations of inhibitors and with substrate BA*p*NA of concentration 1 mM.

Furthermore, the kinetic studies displayed a sigmoidal pattern with increasing concentrations of the inhibitors suggesting reversible and competitive inhibition with tight binding. IRD-9 turned out to be a stronger inhibitor of bovine trypsin (IC_{50} ~0.0022 mM) than IRD-7 (IC_{50}~0.135 mM) and IRD-12 (IC_{50}~0.065 mM) **(Fig. 2.2C)**. The inhibition constant Ki determined directly from IC_{50} by using the Cheng-Prusoff's equation also confirmed the same **(Table 2.1)**. Although the aa and structure variations of IRDs account for their differential binding efficiency, the exact molecular mechanism that contribute to binding efficiency is not understood.

It is known that the disulfide bonds are essential for the folding, function, and stability of IRDs (Schirra *et al.*, 2010). In the present study the activity of all three IRDs seen on 15% Native-PAGE was lost in the reduced state **(Fig. 2.3A).** Disulfide rich proteins are also known to show high thermal stability (Bronsoms *et al.*, 2011). Inhibition assays carried out at different temperatures showed that IRD-7 and -12 retained their inhibitory activity against trypsin even at 90 ^{0}C for 30 min whereas IRD-9 gradually lost activity starting from 70 ^{0}C (**Fig. 2.3B**). The reduced thermal instability of IRD-9 might be due to decrease in the number of disulfide bonds.

Table 2.1: IC_{50} and Ki value obtained from inhibition kinetics

Inhibitor	**Experimental**		
	IC_{50} (mM)	***Ki* (mM)**	**ΔG_{bind} (kcal/mol)**
IRD-7	0.135	0.126	-4.787
IRD-9	0.0022	0.002	-8.05
IRD-12	0.065	0.061	-5.917

Interestingly, IRD-9 exhibited proteolytic resistance for 60 min when incubated with HGP as compared to IRD-7 and -12, both of which submitted to instantaneous proteolysis (**Fig. 2.3C**). Gut extract of *H. armigera,* a complex mixture of various trypsin and chymotrypsins like proteases, displayed at least 7 isoforms (HGP-1 to -7) (**Fig. 2.3D**). These isoforms of HGP vary in terms of properties and specificity. Interestingly, HGP isoforms were differentially inhibited by various IRDs. The activities of HGP-3 and -4 were inhibited by all IRDs, whereas that of HGP-5, -6 and -7 were inhibited exclusively by only IRD-9. Protease activity band between HGP-6 and -7 was developed only in the case of IRD-9 treatment and was not present even in untreated HGP, indicating IRD-9 bound protease complex acquiring a different charge state. Thus, IRD-9 presented unique binding property and activity.

The synergistic effect of IRDs was analysed by performing inhibition assay with combination of different IRDs in IC_{50} concentration. The presence of IRD-9 in combination with IRD-7 and IRD-12 enhanced their corresponding HGPI activity from 49 to 65% and 51 to 63%, respectively (**Fig.2.3E**). Results obtained showed that IRD-9 might have a synergistic effect and can lead to higher potentiation of other IRDs. These biochemical evidences support the higher efficiency of varied combination of CanPIs/IRDs in inhibiting insect gut proteases, which signifies the biological relevance of sequence variation.

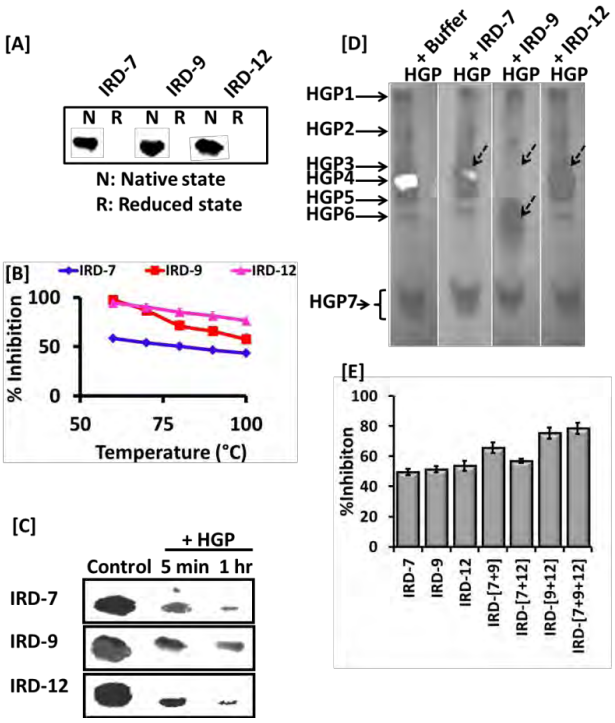


Fig. 2.3: Biochemical characteristics of IRDs. (A) Native-PAGE activity gel with equal units (i.e. 0.5 TIU) of a native and reduced sample of inhibitor IRD-7, IRD-9 and IRD-12 in consecutive lanes. Only native samples showed inhibitory activity, and not in the reduced state. **(B)** Inhibition activity of IRDs against HGP at different temperatures **(C)** In vitro stability of IRDs towards HGP. Equal HGPI units (0.5 Units) of IRD-7, -9 and -12 were incubated with 1 HGP unit at 24°C for 5min (lane 2), 1 hr (lane 3) and the reaction mixtures were resolved on 15% native-PAGE gel. Each IRD without HGP treatment (lane 1) was loaded as a control. The gels were processed for TI activity visualization by GXCT. IRD-9 shows higher intensities as compared with IRD-7 and -12 in the presence of the HGP. **(D)** Comparative inhibition of HGP isoforms by different IRDs. Equal HGPI units of IRD-7, -9, -12 were incubated with HGP for 30min at 24°C. The above reaction mixtures were then resolved on 8% native-PAGE. The gels were processed for protease activity visualization by GXCT. IRD-9 and -12 show inhibition of maximum HGP isoforms. **(E)** Synergistic inhibitory effect of different combinations of IRD-7, -9 and -12 using 0.5 mM of each protein.

In silico studies indicated that IRD-9 had two free cysteine residues which might be in the form of thiol. This observation is confirmed by Ellman's assay, which estimates free thiol groups in small peptides. In the present study, it showed that 3.8 µM of IRD-9 had ~ 7.9 µM of free thiol (~ 2 free cysteine residues) whereas a similar amount of IRD-7 and -12 had approximately ~0.155 and ~0.183 µM free thiol content (absence of any free thiol). These results provided additional support to the *in silico* predictions that IRD-7 and -12 had four disulfide bonds, whereas IRD-9 had only two leaving two remaining cysteines free.

II.3.3 Serine residues modification influence inhibition potential of IRDs

The effect of PMSF on the activity of IRDs is shown in the **Fig.2.4**. Reaction of the inhibitor with PMSF led to modification of one serine residue (number of residues modified were deduced from graph of Log K_{app} against conc. of PMSF) in IRD-9 and resulted in 35 to 45% activity loss.

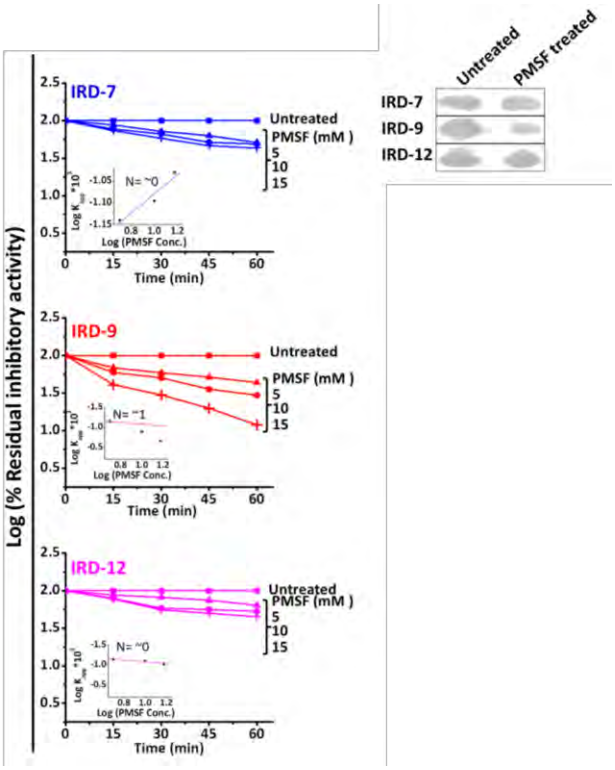


Fig. 2.4: Inactivation of IRDs by serine modification using PMSF. Pseudo first-order plot for inactivation of IRDs by PMSF of concentration 5, 10 and 15 mM for 15, 30, 45 and 60 min **(A)** IRD-7, IRD-9 and IRD-12. **(B)** Inset shows corresponding second order plot of pseudo-first order rate constants (K_{app}) (min^{-1}) as a function of log (PMSF) concentration.

Modification of IRD-7 and -12 did not show significant effect on the activity **(Fig. 2.4A)**. Furthermore, activity visualization assay showed that PMSF modified IRD-9 reduced inhibition potential as compared to modified IRD-7 and -12 **(Fig.2.4B)**. These results pointed out that, serine could be involved in holding the reactive loop in proper position through a network of hydrogen bonds which was blocked on treatment with PMSF and resulted in loss of inhibitory activity of IRD-9. Thus, result indicated that serine residues were not directly involved in the interaction, but they significantly affected the binding of inhibitor with a protease molecule.

II.3.4 *In vitro* assay indicates broad specificity and variable inhibition of activity of various proteases by selected IRD variants

Sequence alignment of cloned inhibitors represented conserved cysteine residue position, while reactive site loop was the most variable region in the sequence. Selected inhibitors demonstrated differential reactivity and specificity against various proteases. IRD-9 and -12 exhibited strong inhibition (70 to 90%) of trypsin, chymotrypsin and HGP, while IRD-7 (40 to 70%) showed comparatively less inhibition **(Fig. 2.5)**. IRD-9 and -12 inhibited 30 to 50% of cathepsin and elastase activity, whereas IRD-7 showed 10 to 20% inhibition. It was observed that IRD-9 and IRD-12 displayed strong inhibitory effect on most of the proteinases examined. Activity assay against various proteases indicated that sequence variation in the IRDs might affect its inhibitory efficiency and target specificity. Although the selected IRDs were primarily trypsin inhibitors, they exhibited activity against chymotrypsin, elastase and cathepsin.

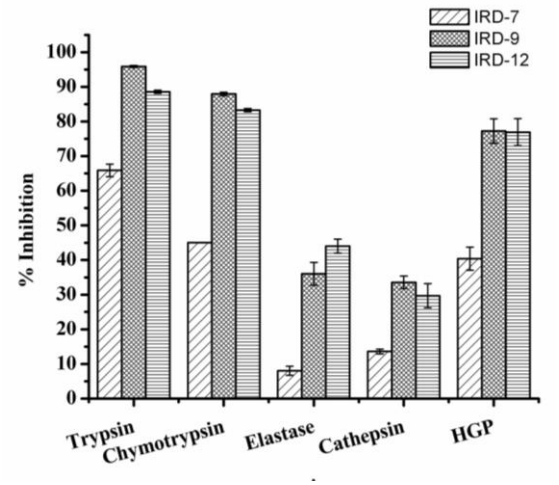


Fig. 2.5: Inhibition of bovine trypsin, bovine chymotrypsin, elastase, cathepsin and HGP by 20µg of IRDs in azo-caseinolytic assays. Each value is an average of six replicates IRD-9 and -12 showed significant inhibitions in all different proteases than that of IRD-7 against all proteases.

II.3.5 IRDs exhibit strong binding efficiency with *H. armigera* proteases

Docking and relative analysis displayed significant differences in binding energies suggesting that varied IRDs had variable interaction with *H. armigera* proteases.

Among the three IRDs, IRD-9 showed strong interaction with the lowest binding energy with various *H. armigera* proteases **(Fig. 2.6**). As evident from activity inhibition assays, docking studies revealed broad specificity of IRDs with chymotrypsins, cathepsins and other serine proteases. Strong binding of IRD-9 with trypsin and chymotrypsin among all the three IRDs motivated us to compare in *vivo* effect of the recombinant IRD proteins against *H. armigera* digestive physiology.

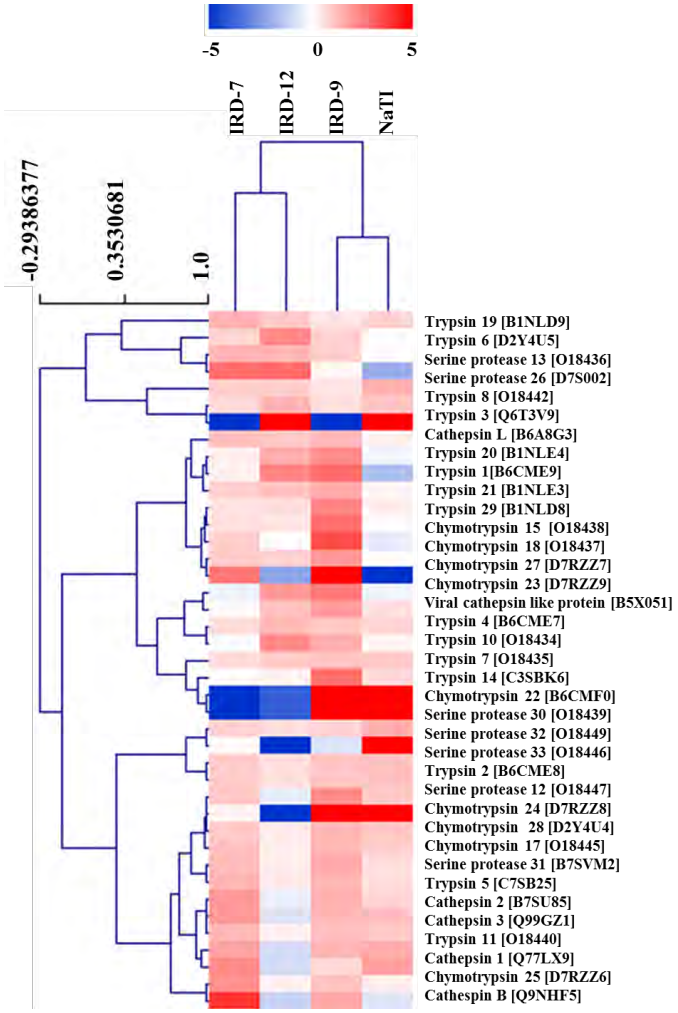


Fig. 2.6: Heat map with hierarchical clustering of relative free binding energy normalized by mean values of all *H. armigera* trypsins, chymotrypsin, other serine proteases and cathepsins with IRD-7, 9, 12 and NaTI. Accession number of each protease sequence (uniprot IDs) given in the bracket. The scale values (-0.29 to 1.0) on the cluster branches indicates the degree of correlation in different clusters.

Interaction pattern of IRD-7 and -12 with most of the proteases was similar, which led to their clustering together for all the analyzed proteases. Binding energy comparison and hierarchical clustering analysis provided wide overview of specific interaction of inhibitor with various proteases. Furthermore, it will also give

speculation about mode of action and effect of various inhibitors on *H. armigera* digestive proteases.

II.3.6 Incorporation of IRDs in diet retards the growth and development of *H. armigera* larvae

To understand the *in vivo* effect of IRDs on the development of *H. armigera* larvae, feeding experiments were conducted with appropriate controls. Active recombinant IRD proteins (5 TIU/g diet) were incorporated into diet to examine their *in vivo* potential against *H. armigera*. Development of larvae reared on a control and IRD protein-containing diets is presented in **Fig.2.7A**. Feeding of insects on IRD-containing diet caused reduction in larval mass gain and survival rate. On day 11, larvae fed on diets containing IRD-9 and -12 weighed ~40 and 35% less, respectively, than the control (fed on AD without IRD-protein) larvae. In comparison, larvae fed on AD containing IRDs were ~50 to 60% smaller than control larvae. Larvae fed on diet containing IRD-7 displayed ~ 20% and 15% reduction in larval mass and size, respectively **(Fig. 2.7B)**. Furthermore, larvae fed on inhibitor containing diet showed significant ($p \leq 0.05$) reduction in survival rate **(Fig. 2.7C)**. At day 11, there was ~ 20% reduction in controlled survival rate (survival rate normalized by control larvae survival rate) of larvae fed on IRD-9 and -12 containing AD as compared to control larvae. PIs fed larvae displayed early and sharp decrease in larval survival rates, in case of IRD-7 and -12, it was followed by partial recovery as the feeding period extends. This might be due to expression of PIs insensitive proteases and overexpression of proteases, which might help insect to overcome the lethal and detrimental effect of inhibitors (Dunse *et al*., 2010; de Oliveira *et al*., 2013).

Evaluation of nutritional parameter like Efficiency of Conversion of Ingested Food (ECI), Efficiency of Conversion of Digested Food (ECD) and Approximate Digestibility (Ad) revealed that the ingestion of IRD proteins had deleterious effect on growth and rudimentary metabolism of the insect **(Table 2.2)**.

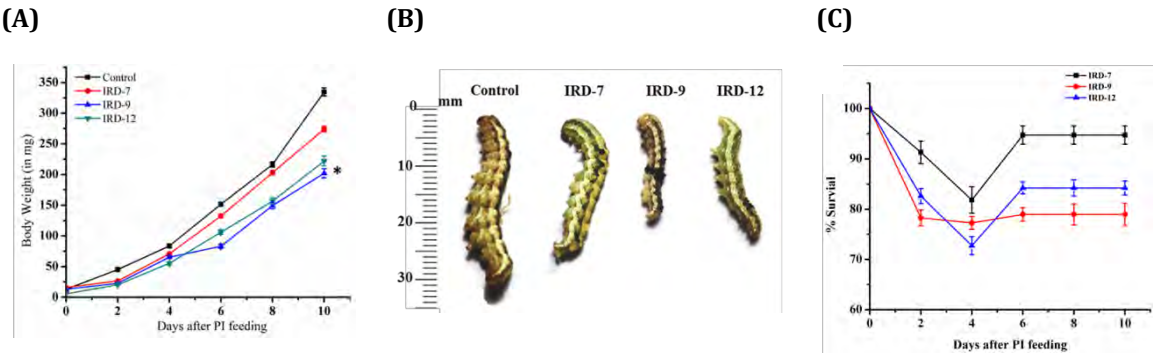


Fig. 2.7: ***In vivo* effect of IRDs ingestion of H. armigera,** Growth (A), average size of larvae recorded on day 10 (B) and Survival rate normalized with control (C) of *H. armigera* larvae raised on ADs containing 5 TIU of each IRD. Eggs were hatched, and neonates were transferred to ADs containing of PIs.

There was direct correlation in the inhibitory potential and reduction in ECI, ECD and Ad. Assessment of these parameters showed that IRD-9 and -12 negatively affect digestive physiology of insect and thus impede insect growth and survival **(Table 2.2)**. Inhibition of serine protease activities also obstructed normal developmental pathways leading to delay in pupation and molting, which was also evident from the data **(Fig. 2.7B)**. These results indicated that IRD-9 and -12 could serve to develop effective inhibitor molecules against gut proteases from *H. armigera*.

Table 2.2: Effect of IRDs ingestion on the *H. armigera* feeding behavior and dietary utilization.

	IRD-7	IRD-9	IRD-12
ECI	10.31 (±1.01)	8.35 (±1.12)	9.76(±1.32)
ECD	25.79 (±2.53)	18.89(±1.81)	22.42(±3.32)
Ad	56.15 (±1.95)	46.55(±0.36)	50.05(±1.15)

Abbreviations: ECI = Efficiency of Conversion of Ingested Food; ECD = Efficiency of Conversion of Digested Food; Ad = Approximate Digestibility.

II.3.7 *In vivo* inhibition of gut proteinases in *H. armigera* larvae reared on IRDs

In comparison with control HGP activity, HGP of larvae fed on IRD-9 and -12 showed ~70 and 80% reduction of activity, respectively. Larvae fed on IRD-7 displayed~40 to 50% of inhibition of HGP activity as compared to control larvae **(Fig.2.8A).** In case of trypsin-like proteinases activity of HGP from larvae fed on IRD-9 and -12 showed ~80 to 85% reduction as compared to control HGP activity, while it was moderately reduced in HGP of larvae fed on IRD-7.

The present study demonstrated the efficacy of three IRDs in inhibiting the proteinases and growth of *H. armigera*. Proteinase profiles revealed that IRDs inhibit various HGP isoforms with their differential activity **(Fig. 2.8B)**.

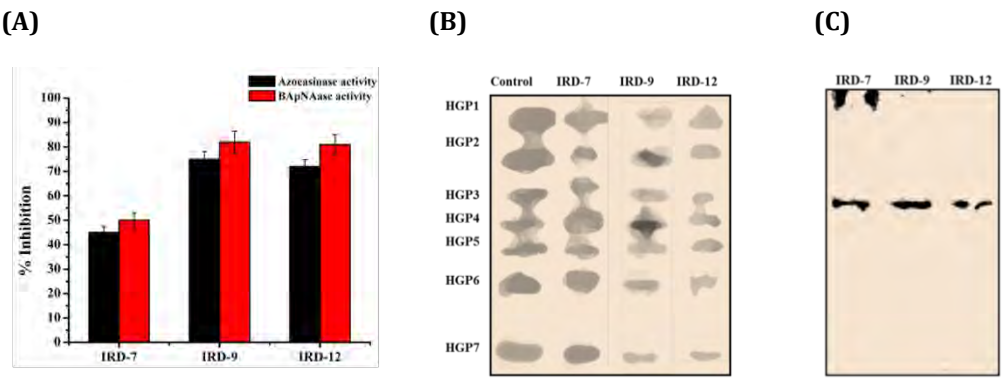


Fig. 2.8: Molecular response of H. armigera to ingested IRDs, *In vivo* proteolytic activity **(A)** was assessed by azocasein and BApNA assay. Differential protease activity **(B)** of HGP of larvae fed on AD containing IRD-7, -9 and -12, was visualized using gel X-ray contact print technique (GXCT), untreated X-ray film was used as the substrate. GXCT shows seven bands of control HGP, HGP1 to 7. In case of HGP of IRDs fed larvae, intensity fading of various HGP isoform indicates their inhibition. The residual protease inhibitor activity **(C)** in gut extract of larvae fed on three IRDs was also visualized by GXCT. Intense band of IRD-9 indicated that it was more resistant to gut proteolytic environment as compared to IRD-7 and -12.

Furthermore, the HGP profile of IRD-fed larvae revealed a number of isoforms differing in intensity and mobility. Zymography examination of HGP from control and IRD-fed larvae showed that most of the HGP isoforms are inhibited upon IRD ingestion **(Fig. 2.8B)**. Seven distinct isoforms, HGP1 to 7 were detected in the gut extract of larvae reared on control diet. HGP of larvae fed on IRDs showed

significant reduction in intensity of all HGP bands. In case of IRD-7 and -9, HGP-4 isoform showed overexpression, indicated insensitivity of HGP4 to IRD proteins. Proteinase inhibitor activity gel was performed to check the stability of IRDs in insects gut environment **(Fig. 2.8C)**.

Protease activity gel analysis provided qualitative account about expression of various active isoforms of proteases upon PI ingestion, and also insight about degradome dynamics of *H. armigera* gut to counter PIs. The difference in the electrophoretic mobility of protease isoforms was evident in HGP of IRD-9 and -12 fed larvae with their different interactions. This data indicated the differential sensitivity of proteases to inhibitors and also stability of protease-PI complexes. The expression of PI-insensitive proteases assisted in obtaining optimal nutritional requirements by utilizing diet protein sources as well as using proteinase inhibitors as source of amino acids under PI stress (Jongsma *et al*., 1995).

II.3.8 Ingestion of IRDs triggers differential gene expression of gut proteases in *H. armigera* larvae

Expression analysis of eight major gut trypsin genes further provided the quantitative evidence of altered digestive physiology of *H. armigera* **(Fig. 2.9)**. Real time PCR analysis showed that trypsins namely HaTry3, 6 and 8 were found to be down regulated in IRD-7, -9 and -12 fed larvae, at the same timeIRD-7 and -9 fed larvae additionally showed downregulation of HaTry1, 2 and 5. Downregulation of these trypsins might also correlate with drop in proteolytic activity of numerous HGP isoforms when countered with IRDs. HaTry-4 and especially HaTry7 were highly upregulated in all the IRD fed larvae. HaTry1, 2 and 5 were exclusively upregulated in IRD-12 fed larvae. Overexpression of protease transcripts could help the insect to overcome the inhibitory effect of PIs, which might in turn help in growth and development of insects.

Among four major chymotrypsin, three transcripts HaChy1, 2 and 3 showed down regulation in all (IRD-7, -9 and -12)-fed larvae. HaChy4 showed distinct

behavior with overexpression in IRD-7 and -9 fed larvae and down regulation upon IRD-12 exposure. Expression dynamics of gut protease genes in an attempt to overcome the effect of inhibition and to acquire optimum nutritional requirements for growth and development provides an excellent survival benefit to *H. armigera*.

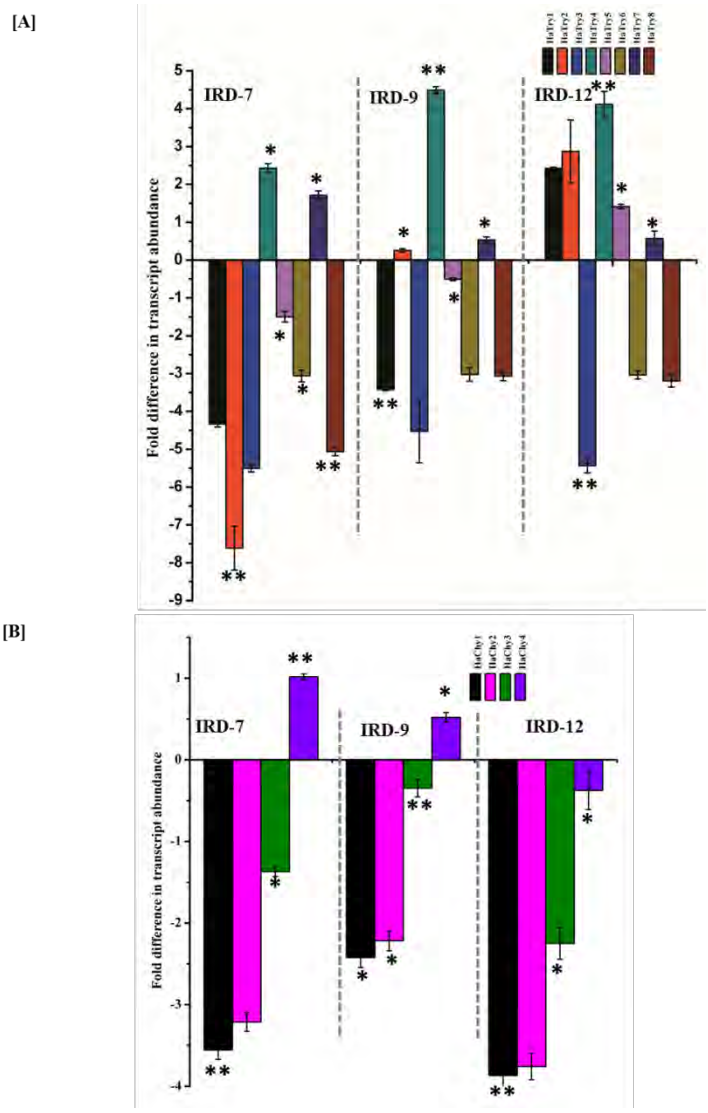
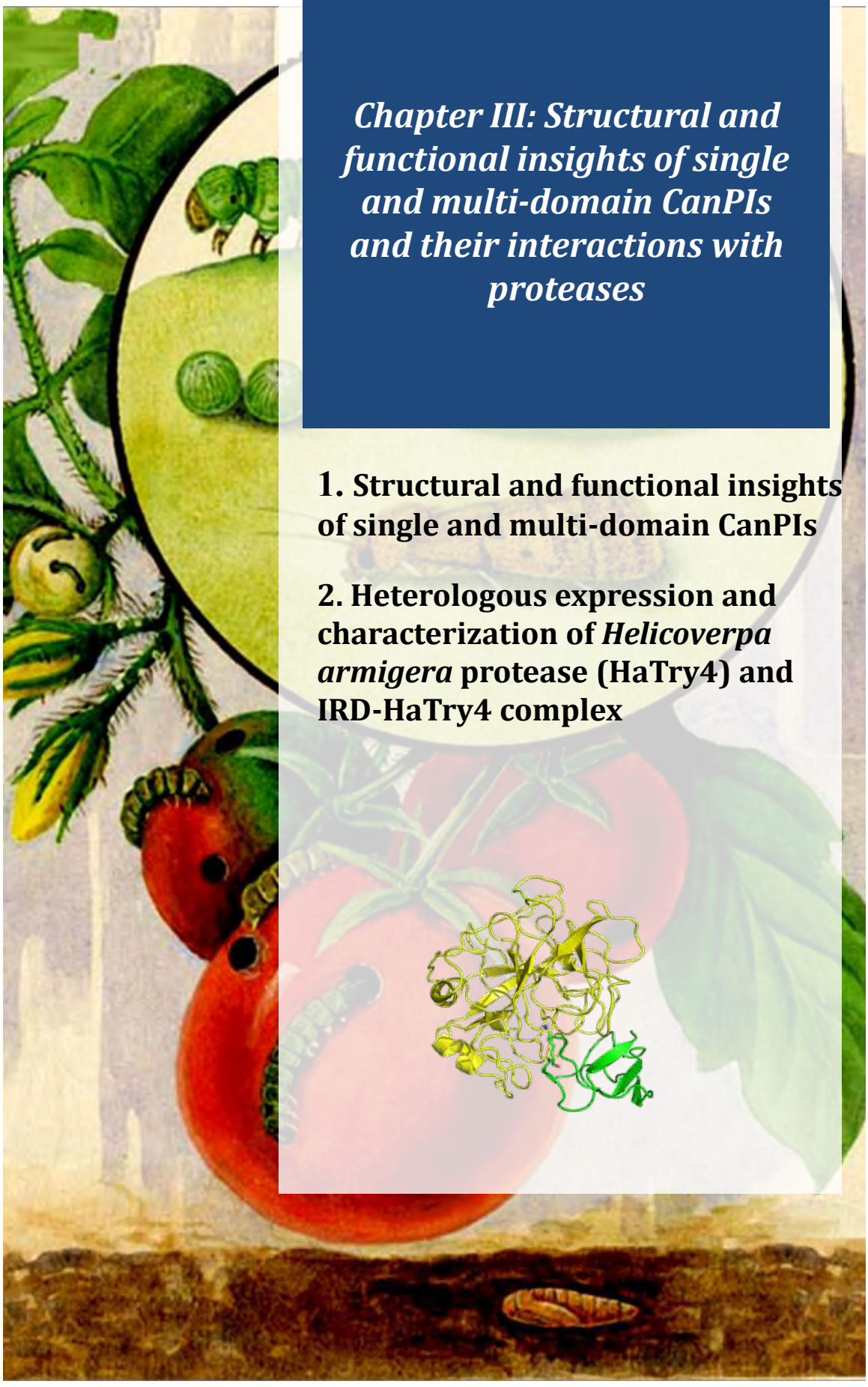


Fig. 2.9: Fold difference in transcript abundance of major digestive trypsins and chymotrypsins for larvae fed on AD containing IRD-7, -9 and -12 (5 TIU) on the 10th day after feeding. (A) Fold difference in transcript abundance of trypsin genes (B) Fold difference in transcript abundance of chymotrypsin genes. Most of trypsin and chymotrypsin gene showed down regulation; while few of them showed upregulation in all the three inhibitors fed insects tissues.

II.4 CONCLUSION

This chapter employed a combination of experimental and theoretical approaches to investigate the molecular details of HaTry-IRD interaction. Expression and biochemical characterization of IRD-7, -9 and -12 revealed IRD's sequence-dependant variation of inhibition. Furthermore, IRD-9 lacking two disulfide bonds shows phenomenal inhibition activity compared to other IRDs. This natural variant also exhibit special attributes like stability to proteolysis and inhibitory synergistic effect on other IRDs etc., which makes this molecule unique among the members of Pin-II inhibitor family. Chemical modification studies of serine residues combined with MD simulation confirms the role of serine residues that replaced cysteine in increasing the inhibition potential of IRD-9. Continuous exposure to IRDs since neonatal stage exhibited negative impact on *H. armigera* growth and development. This investigation provides insight into the differential efficacy and specificity of IRD proteins against target proteases, owing to the sequence variations within these proteins. Exposure of larvae to varied IRDs offered variable molecular response and resulted in alteration in digestive physiology, particularly with respect to protease expression. Current study establishes IRD-9 and -12 as effective growth inhibitors to *H. armigera* larvae. This phenomenon of performance variability depending on sequence variation can be explored for designing effective inhibitor candidate against *H. armigera*.



Chapter III: Structural and functional insights of single and multi-domain CanPIs and their interactions with proteases

1. Structural and functional insights of single and multi-domain CanPIs

2. Heterologous expression and characterization of *Helicoverpa armigera* protease (HaTry4) and IRD-HaTry4 complex

Chapter III: Structural and functional insights of single and multi-domain CanPIs and their interactions with proteases

III.1 Structural and functional insights of single and multi-domain CanPIs

III.1.1 INTRODUCTION

Each IRD of the Pin-II precursor contains a single reactive site either by trypsin or chymotrypsin inhibition. Single IRD PI of Pin-II family can bind to a single protease. Also it is well known that two domain PIs of tomato and potato can simultaneously inhibit two protease molecules (Greenblatt et al., 1989; Barette-Ng et al., 2003a). P1 residue of the reactive site determines its specificity; presence of lysine "K" or arginine "R" in the P1 position confers TI potential and Leucine "L" bestows the inhibitor CI potential. The profound docking of the side chain of P1 residue in the S1 binding pocket of the protease plays an extensive role in the energetics of the specificity of PI-protease interaction. In bovine pancreatic trypsin inhibitor (BPTI, Kunitz type), the substitution of P1 residue with amino acid (aa) other than K/R showed a weaker side chain interaction of the P1 residue in the S1 binding pocket of trypsin (Otlewski et al., 2001). A stretch of 5 aa close to the reactive site, P3 to P2' is important in determining enzyme specificity of the inhibitor. It is bound by two disulfide bonds which give the reactive site a considerable rigidity. While the aa in the adjoining region of this core segment show a very high sequence variability, conferring flexibility to provide broad inhibitory potential. The P4 to P3' of the inhibitor domain interacts with S6 to S2' of protease pocket to bring its inhibition by mimicking a substrate. Mutational studies on Pin-II from potato have highlighted the importance of secondary contacts also not involving RSL as well, in determining the specificity of the protease inhibition. In PI-II from tomato, the inhibition capacity of a TI domain could not be transferred to the other domain by mutating the P1 residue or residues within the RSL (Beekwilder et al., 2000; Schirra and Craik, 2005).

The effect of mutation/variation in the conserved residues is reflected not only in the structure but also in activity of Pin-II proteins. The selective substitution of conserved cysteine residues responsible for disulfide bond formation flanking the

reactive site, resulted in weak inhibitory activity (Ki~ 1.8 μM) by C8A/C37A-T1 variant (Cys-C at position 8 and 37 were substituted by Ala-A). On the other hand, substantial retention of TI activity by C4A/C41A-T1 variant (Ki~ 350 nM) as compared to the wild-type TI (Ki<5 nM) affirmed the indispensability of C8-C37 bond in Pin-II proteins (Schirra et al., 2010) (**Fig. 1.14**).

The structural basis of inhibition of a multi-domain Pin-II inhibitor has been shown by its ternary complex with two subtilisin Carlsberg molecules and revealed how it can be bound to and simultaneously inhibits two enzyme molecules within a single ternary complex (Barette-Ng et al., 2003a). The diversity in the number of repeats and diversity within IRD sequences is predominantly observed in Pin-II PIs. Different specificities within a multi repeat protein may contribute to a PI cocktail to tackle varied pest/ pathogenic attacks. Various members of Pin-II precursors studied, a combination of TI/ CI domains is predominantly noted. For example, the six domains *N. alata* PI (NaProPI) possesses four TI domains and two CI domains. The two, three and four domain PIs of *C. annuum* possesses varied combinations of inhibitory specificities. In potato PI which composes two IRDs, one is TI specific and the other is CI specific. In case of *N. attenuata* precursor, X and YIRDs inhibit trypsin and chymotrypsin, respectively. However, the NaProPI, which has 6-IRDs, cannot bind to six proteases simultaneously because of steric interference. This 6-IRD PI of *N. alata* could inhibit maximum of four chymotrypsin or 2.6 trypsin molecules (Heath et al., 1995; Lee et al., 1999). In order to attain absolute inhibition potential, individual IRDs must be released from the precursor by proteolytic processing. This also highlights the reason behind the absence of inter-domain interactions in Pin-II PIs.

Crystal or NMR structures of multidomain Potato inhibitor-II (Pin-II) are crucial for structure–function relationships concerning e.g., multivalent interaction with proteases, substrate specificity and spatial distribution of reactive sites. The three-dimensional structures of several Pin-II PIs, single- as well as two-domain, have been determined either by X-ray crystallography or NMR. However, there have been limited studies on multi-domain inhibitors giving insights into their domain orientations, binding to proteases and stoichiometry. Possibly because the

precursors PIs are processed at the linker region(s) by endogenous proteases to release IRD(s), and thus, it is difficult to characterize the multi-domain inhibitors from natural plant sources. The findings raise the vitality of conserved residues in Pin-II proteins on their structure and thus, their function. On the basis of this background information, it was attempted to elucidate the functional relevance of the natural variation present in CanPIs.

III.1.2 MATERIAL AND METHODS

III.1.2.1 Cloning of IRDs in bacterial expression system

Based on sequence analysis and previous study (Chapter II), IRD-7, -9 and -12 were chosen for the expression. The cDNA encoding the mature peptide region of IRDs was amplified by using specific flanking primers and cloned in p3E expression vector for recombinant expression in *Escherichia coli* (Kwan et al., 2007). Schematic of p3E_IRDs clone is given in **Fig. 3.1**. Initially vector was linearized using *PmeI* and *NbBsmI*, followed by ligation of the amplified product using T4 DNA ligase. The vector contains a GST solubilization tag at C-terminal and sequence encoding the cleavage site for precision protease. For ease of purification, we have incorporated double streptavidin (DS) tag was incorporated at the N-terminal of the cloned gene **(Fig. 3.1)**.

Fig. 3.1: Diagrammatic representation of the clone for p3E_IRDs clone

The cDNA encoding the mature peptide region of CanPI-7 was cloned in ligation independent cloning (LIC) in compatible expression vector pMCSG7 for recombinant expression in *E. coli*.

III.1.2.2 Expression and purification of recombinant IRDs

IRD-7, -9 and -12 were expressed and purified individually. All the recombinant proteins were expressed in *E.coli* B834 (DE3) pRare cells. The cells were initially grown overnight at 37 °C in 10 ml LB medium supplemented with antibiotics (Ampicillin-100 µg/ml, Chloramphenicol- 25 µg/ml). The pre-culture was used to inoculate 1 L LB broth medium with appropriate antibiotics and allowed to grow until the OD (600nm) reached 0.5-0.6. Cells were induced with IPTG (0.5 mM) overnight at 16^{o}C and harvested by centrifugation. The cell pellet was solubilized in ice cold cell lysis buffer A (50 mMTris-HCl, pH 8.0; 300 mM NaCl; 5% glycerol) and disrupted by sonication (10 s pulse with 10 s intervals for 10 min) using an Ultrasonic Disruptor UD-201 (Tomy, Tokyo, Japan). The supernatant and pellets were separately collected by centrifugation for 60 min at 14,000g, 4^{o}C. The supernatant was loaded on Glutathione Sepharose 4B (GE Life Sciences, City, State, Country) and purified using standard manufacturer's protocol. The fusion protein eluted with buffer B (50 mM Tris-HCl, pH 8.0; 300 mM NaCl; 5% glycerol; 250 mM Glutathione).

The GST tag was cleaved using precision protease at a protease to target protein ratio of 1:5 (w/w) at 4^{o}C for 12 h. Furthermore, additional purification was performed using Strep-Tactin Superflow Plus Cartridge (Quigen) to remove the cleaved tag and collect the protein in flow through. This was applied on pre-equilibrated (50 mM Tris-HCl, pH 8.0; 200 mM NaCl; 5% glycerol) Sephadex S-200 (size exclusion) for further purification.

CanPI-7 being disulfide rich proteins, *E. coli* Origami B (DE3) was chosen as a host to produce the recombinant proteins. Origami B (DE3) is a modified strain that has mutations in both, the thioredoxin reductase (trxB) and glutathione reductase (gor) genes thus maintaining a non-reducing condition in cytoplasm facilitating the S–S bond formation in the recombinant protein in the proper order.

III.1.2.3 Inhibition kinetics of CanPIs

Trypsin inhibitory (TI) activity assays for IRDs and CanPI-7 proteins were performed using BA*p*NA as a substrate (Detailed procedure is given in **section II 2.3**).

III.1.2.4 MALDI-TOF MS analysis

For the analysis of rCanPIs, 3 µg of protein sample was mixed with 20 µl of freshly prepared sinapinic acid (Sigma-Aldrich) (30% ACN, 0.1% TFA). About 2 µl aliquots of this mixture were spotted on the stainless steel MALDI plate by dried-droplet method and incubated at 37°C for 20 min. The MALDI target plate was further subjected to MALDI-TOF as specified above to get spectral profiles.

The mass spectral analysis was done on Q-TOF-MALDI-TOF-MS (ABSCIX, Framingham, MA, USA) equipped with 337-nm pulsed nitrogen laser. The mode of operation was in a positive linear mode with an accelerating voltage of 25 kV. All spectra were acquired by accumulating 50 single laser shots over each sample spot in the range of 1–30 kDa with the following settings: delayed ion extraction time of 1100 ns, grid voltage 93% and low-mass ion gate set to 1000 Da. They were processed for advanced baseline correction and noise removal using Data Explorer software (ABSCIX). The instrument was calibrated using apomyoglobin and BSA (both Sigma-Aldrich, St. Louis, MO, USA).

III.1.2.5 CD spectroscopy analysis

The CD spectra of the proteins were recorded in a J-815 spectropolarimeter (Jasco, Tokyo, Japan) at 28˚C in a quartz cuvette. Each CD spectrum was accumulated from three scans at 100nm/min with a 1 nm slit width and a time constant of 1s for a nominal resolution of 0.5nm. Far UV CD spectra of the IRDs and native (Extracted and purified from flower) PIs proteins (125µg/ml) were collected in the wavelength range of 190-250 nm using a cell path length 0.1cm for monitoring the secondary structure. All the spectra were corrected for buffer contributions and observed values were converted to mean residual ellipticity (MRE) in deg cm^2 $dmol^{-1}$ defined as

$$MRE = M\theta_\lambda / 10dcr$$

Where, M is the molecular weight of the protein, θ_λ is CD in millidegree, d is the pathlength in cm, c is the protein concentration in mg/ml and r is the number of residues in the protein. Secondary structure elements were measured by using CD pro software.

III.1.2.6 Isothermal titration colorimetry (ITC)

ITC was performed using a MicroCal Auto-iTC instrument (GE Healthcare, Buckinghamshire, UK). Nineteen injections of 2 µl bovine trypsin (Stock = 2.0 mM) were titrated against 0.1mM solution of CanPIs. Experiments were carried out at 37^{o}C in a Tris-HCl buffer, pH 7.8. Reference titration was carried out by injecting the same concentration of trypsin into the buffer. Reference titration was subtracted from the experimental titration. Origin 6.0 software was used to derive affinity constants (Kd) and the molar reaction enthalpy (CH).

III.1.2.7 NMR spectroscopy of IRDs

Protein samples were prepared in a mixed solvent of 90% H_2O and 10%2H_2O containing 50mM sodium phosphate (pH 6.4) and 100mM NaCl in the presence of 10mM MgCl2. The protein concentration was ~1mM. NMR experiments were recorded on a Bruker Avance 800 MHz NMR spectrometer equipped with a 5 mm triple-resonance cryogenic probe. Experiments were recorded at 298 K with uniformly $^{15}N/^{13}C$labeledIRDs and included sensitivity-enhanced 2D [$^{15}N–^{1}H$]–HSQC using water-flipback for minimizing water saturation, as well as a suite of 3D NMR experiments such as HNCO, HN(CA)CO, CBCANH, and CBCA(CO)NH (Bax and Ikura 1991; Bax and Grzesiek 1993), essentially for the assignment of backbone resonances.

III.1.2.8 Collection and processing of the SAXS data

Synchrotron X-ray scattering data were all collected at the EMBL X33 beamline (DESY, Hamburg, Germany). The scattering patterns of all samples were recorded in several solute concentrations ranging from 1 to 5 mg/ml. At sample-detector distance of 2.7 m, the range of momentum transfer $0.1 < s < 5$ nm-1 was collected.

Primary data processing and evaluation of the overall structural parameters were performed using standard procedure by the program package PRIMUS (Konarev et al., 2003). Structural data from experimental SAXS profiles can be compared to atomic structures, for which a theoretical SAXS profile must be calculated. In the present work, this process was carried out in the FoXS server (Schneidman-Duhovny et al., 2010).

III.1.2.9 Molecular docking and simulation study

The 3D models for IRD-7, -9, -12 and CanPI-7 were obtained using I TASSER server. All the models were energy minimized using 1000 steps of the conjugate gradient algorithm and short MD simulations. The stereo-chemical properties of the final selected models were validated using PROCHECK and ProSA (https://prosa.services.came.sbg.ac. at/prosa.php) analyses. MODELLER software was used for *in silico* point mutation and three variants of IRD-9 namely IRD-9A, -9B, and -9C were designed. These variants comprised mutations S7A, S8A in IRD-9A; C28S, C37S in IRD-9B and S7A, S8A, C28A, C37A in IRD-9C, respectively. Furthermore, each individual IRD was docked against HaTry using ZDOCK, while in CanPI-7 multiple HaTry molecules were used for docking. The binding site residues for HaTry and each of the IRDs were specified for docking, to allow the catalytic triad of HaTry (His69, Asp114 and Ser211) to interact with the reactive loop of the IRDs (CPxxC). After the initial docking, the best complex in each case was chosen based on the ZDOCK scores i.e. ZRANK, which is in the range of 15–31 for small proteins of 100 residues.

Explicit MD simulations were performed for exploring the molecular mechanism of the dynamic interactions, the importance of the interacting residues in binding, and the stability of the disulfide bridges. A set of six simulations was carried out, corresponding to the six protease-IRD complexes using the GROMACS 4.0.7 package with GROMACS ffG43a1 force field for 20 ns each. All the six systems were solvated with single-point charge water model and neutralized with proper counter-ions. All the six systems were then energy minimized using 10,000 steps of the steepest descent algorithm present in the GROMACS package followed by a 100

ps position restraining simulation – restraining the protein by a 1000 kJ/Mol harmonic constraint to relieve the close contacts with water molecules under canonical (NVT) ensemble conditions, where moles (N), volume (V) and temperature (T) are conserved. V-rescale (modified Berendsen) temperature coupler was utilized to couple the temperature. Another equilibration run under isothermal–isobaric ensemble (NPT) ensemble conditions was performed for 100 ps, before the final production run of 20 ns each for all the systems, where moles (N), pressure (P) and temperature (T) are conserved. V-rescale temperature and Parrinello-Rahman pressure couplers were used to maintain the temperature (293 K) and pressure (1 bar) values with the protein and non-protein (water and ions) molecules separately coupled with a coupling constant of Tt = 0.1 picoseconds (ps). The isotropic pressure coupling was set with Tp = 2 ps. A time step of 2 femtoseconds (fs) was used throughout with periodic boundary conditions and LINCS constraint algorithm was used to maintain the geometry of the molecules. The van der Waals interactions and Coulomb interactions were cutoff at 12Å with updates every five steps, while long-range electrostatic interactions were calculated using the particle-mesh Ewald method. All the simulations were performed on a PARAM Yuva cluster at the Centre for Development of Advanced Computing at Pune, India, using 64 Intel Xeon 2.93GHz Quad Core processors. The results were analyzed using the in-built analysis package of GROMACS, XMGRACE (http://plasmagate.weizmann .ac.il/Grace/) and in-house developed scripts (Ghosh et al., 2012). The trajectories were visualized using Visual Molecular Dynamics (VMD) software and all the images were rendered using PyMol. The overall stability of all the simulated systems was also checked with respect to temperature, pressure, and potential energy. All the six simulated systems were in thermodynamic equilibrium during the production simulation runs confirming the convergence of the individual trajectories (Figure S2). PDBePISA server (http://www.ebi.ac.uk/msd-srv/prot_int/cgi-bin/piserver) has been utilized for free-energy calculations of these complexes.

III.1.3 RESULTS AND DISCUSSION

III.1.3.1 Recombinant expression of disulfide rich IRD-7, -9, -12 and CanPI-7

The diagrammatic representation of the selected IRDs from *C. annuum* and sequence alignments of the constituent three unique IRDs is shown in **Fig. 3.2A**. All IRDs comprised a single trypsin inhibitory domain (TI). Each IRD was 50 aa polypeptide and consisted of 8 conserved cysteine (C) residues contributing to four disulfide bonds and a single reactive site (P1). The core reactive site of an IRD, 'PKN' or 'PRN' was confined by two conserved cysteine residues. In case of IRD-9, two cysteines at position seventh and eighth were replaced by serine (S) residues. The soluble fraction was purified and final preparations yielded single protein corresponding to 6 kDa for all three IRDs (**Fig. 3.2B**) and 25 kDa for CanPI-7 **(Fig. 3.2C)**.

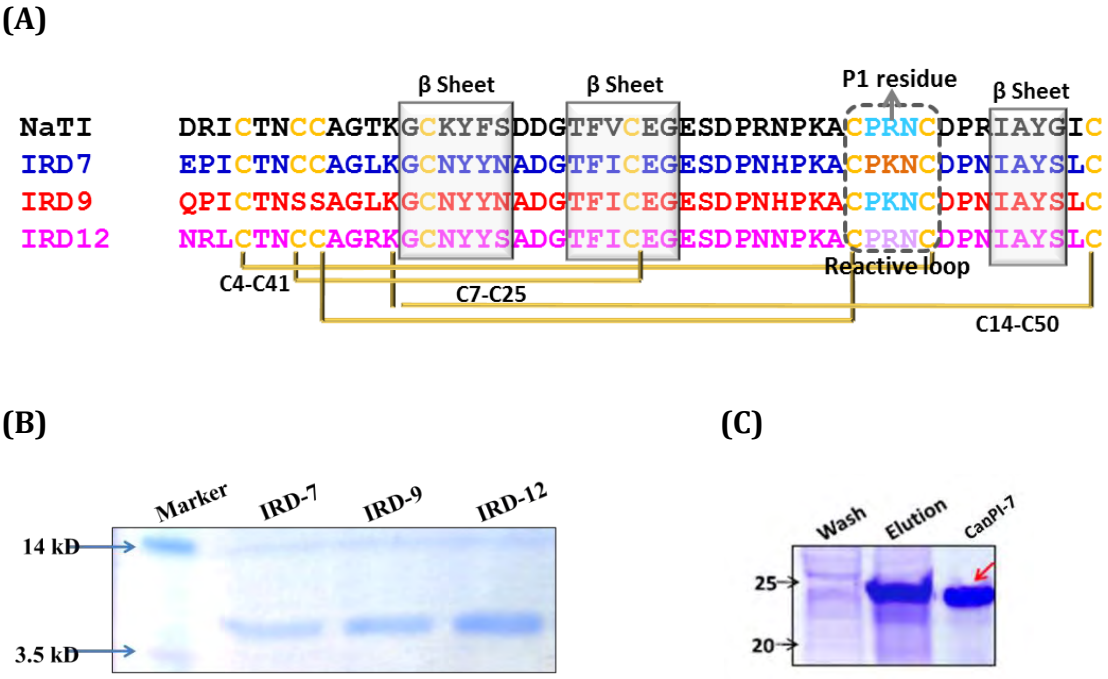


Fig. 3.2: (A) Multiple sequence alignment of IRD-7, -9, -12 and *Nicotiana alata* trypsin inhibitor (NaTI) using DNASTAR and ClustalX2 software. Conserved cysteine residues are marked in yellow colour and the reactive loop region (residue 37-41) including the P1 residue is indicated with cyan colour. The number and position of cysteine residues are conserved in all except IRD-9 in which the cysteines at seventh and eighth positions are changed to serine. The sequence of the reactive loops of IRD-7 and -9 is "CPKNC" whereas that of IRD-12 and NaTI is "CPRNC". **(B)** Purified recombinant IRD-7, -9 and -12 (6 kDa) and **(C)** CanPI-7 separated on 15% SDS–PAGE and stained with Coomassie Blue R250.

Expressed IRDs were characterized by MALDI-TOF mass spectrometry analysis. The molecular mass for expressed IRDs characterized by MALDI-ToF mass spectroscopy and theoretical mass calculated using ExPASy - ProtParam tool were given in **Table 3.1**.

Table 3.1: MALDI-TOF mass characterization data of the expressed IRDs

IRDs	**MALDI-ToFmass**	
	Calcd.	Obsd.
IRD-7	5359.00	5546.89
IRD-9	5325.80	5588.12
IRD-12	5424.00	5576.66

III.1.3.2 Inhibitory activities and kinetic analysis of IRDs

Assays using BA*p*NA and azocasein as substrates showed that IRD-9 and -12 inhibited about 80–85% of HGP activity while inhibition by IRD-7 was only 40–45% **(Fig. 3.3)**. Furthermore, kinetic studies displayed a sigmoid pattern of increasing concentrations of the inhibitors suggesting reversible and competitive inhibition with tight binding. IRD-9 turned out to be a stronger inhibitor of bovine trypsin (IC50 ~2 µM) than IRD-7 (IC50 ~47 µM) and IRD-12 (IC50 ~35 µM) **(Fig. 3.3)**.

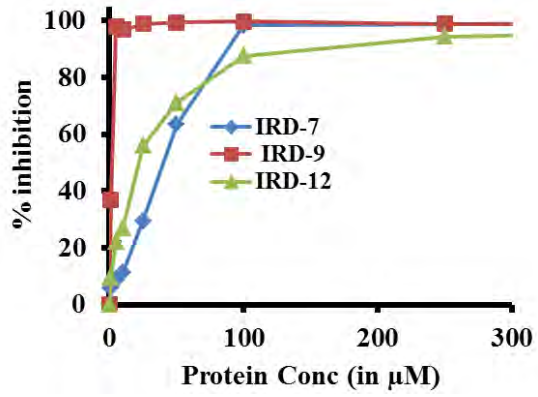


Fig. 3.3: Estimation of IC_{50} and K_i values by using inhibition of bovine trypsin with various concentrations of inhibitors and with substrate BA*p*NA of concentration 1 mM. The data represent the average of three independent experiments.

The inhibition constant K_i determined directly from IC_{50} by using the Cheng–Prusoff's equation also confirmed the same (**Table 3.2**). Although the aa and structure variations of IRDs account for their differential binding efficiency, the exact molecular mechanism that contributed to binding efficiency was not understood.

Table 3.2: IC50 and K_i value obtained from inhibition kinetics

Inhibitor	Experimental	
	IC_{50} (µM)	*Ki*(µM)
IRD-7	47	43
IRD-9	2	1.8
IRD-12	35	32

III.1.3.3 Secondary structure of rIRDs shares similarity with native IRDs

CD Pro analysis of the far UV CD spectrum **(Fig. 3.4)** yielded the values of secondary structure elements as: α-helix-3.8%, β-sheet-41.4%, turns-21.2% and unordered-33.4% for native protein and recombinant IRDs. The single negative band between 195 to 200nm in the CD spectrum and the CDPro analysis demonstrated that IRDs were β-sheet rich and unordered proteins. Similarity in CD spectra of native protein and recombinant expressed IRDs provided fact that protein expressed in *E. coli* had similar fold like native PIs proteins. Both native and recombinant IRDs showed minima at 196nm, which indicates that the overall secondary structure is similar.

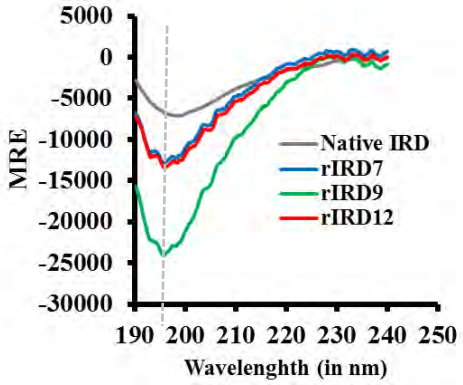


Fig.3.4: Far UV CD spectra of native IRD and recombinant IRD (rIRD). Dashed line in grey colour represents MRE minima. Similar minima for native and recombinant IRDs indicated the similar secondary structure. The data represent the average of three independent experiments

III.1.3.4 Binding of CanPIs to proteases is thermodynamically favorable

Thermodynamic studies of trypsin-IRD interaction was carried out using ITC as it is one of the most widely used quantitative technique for direct measurement of the enthalpy change when two species interact, allowing the determination of heat of association, stoichiometry, and binding affinity from a single experiment. The affinity constants (Kd), the molar reaction enthalpy (CH) and the stoichiometry of binding (N) for rIRDs are provided in **Table 3.3**.

Table 3.3: Thermodynamic evaluation of rIRDs binding with bovine trypsin

Sample	Affinity constants (Kd) in µM	Molar reaction enthalpy (CH) in Kcal/mol	Stoichiometry of binding (N)
rIRD-7	0.113	0.11	1
rIRD-9	0.002	0.23	1
rIRD-12	0.062	0.162	1

Binding was strongly exothermic and showed 1:1 stoichiometry for trypsin and IRD. The spontaneity of the process was evidenced by a negative change in the enthalpy, CH, and a positive change in the entropy, CS. Binding of IRD-9 found to be stronger as compared to IRD-7 and -12 as it is depicted from lowest ΔH for IRD-9.

III.1.3.5 Molecular mechanism of IRD(s)-HGP interaction

MD simulations were used to predict the binding affinities and hence, the inhibitory effects of the individual IRDs against HaTry. Molecular models of the IRD bound HaTry predicted several atomic interactions with a reactive loop of inhibitors that also explained the contribution of the solvent exposed reactive loop. In IRD-9-HaTry interaction, carbonyl oxygen atoms of Met-92 and Ser-207 of HaTry active site formed hydrogen bonds with inhibitor side chain of Asn40 and Lys39, while the side chain of Ser192 from HaTry formed a hydrogen bond with the side chain of Lys-39. Arg-39 from IRD-12 reactive site loop formed three hydrogen bonds with Ser207 and His50 of the HaTry active site (**Fig. 3.5**). In case of IRD-7, Lys39 amide and carbonyl of reactive site loop interact with the carboxyl carbonyl oxygen atom of His50. There are additional hydrogen bonds that exist between Cys37 from the reactive site loop of IRD-9 and -12, with the carbonyl oxygen atom of Ile210 or side

chain of Arg109 residue from HaTry. Although the interaction of the active site of enzymes in all the three inhibitors was similar in nature, significant differences were observed in making the weak interaction like hydrogen bonding and van der Waals interactions, which resulted in differential binding free energy of the complexes.

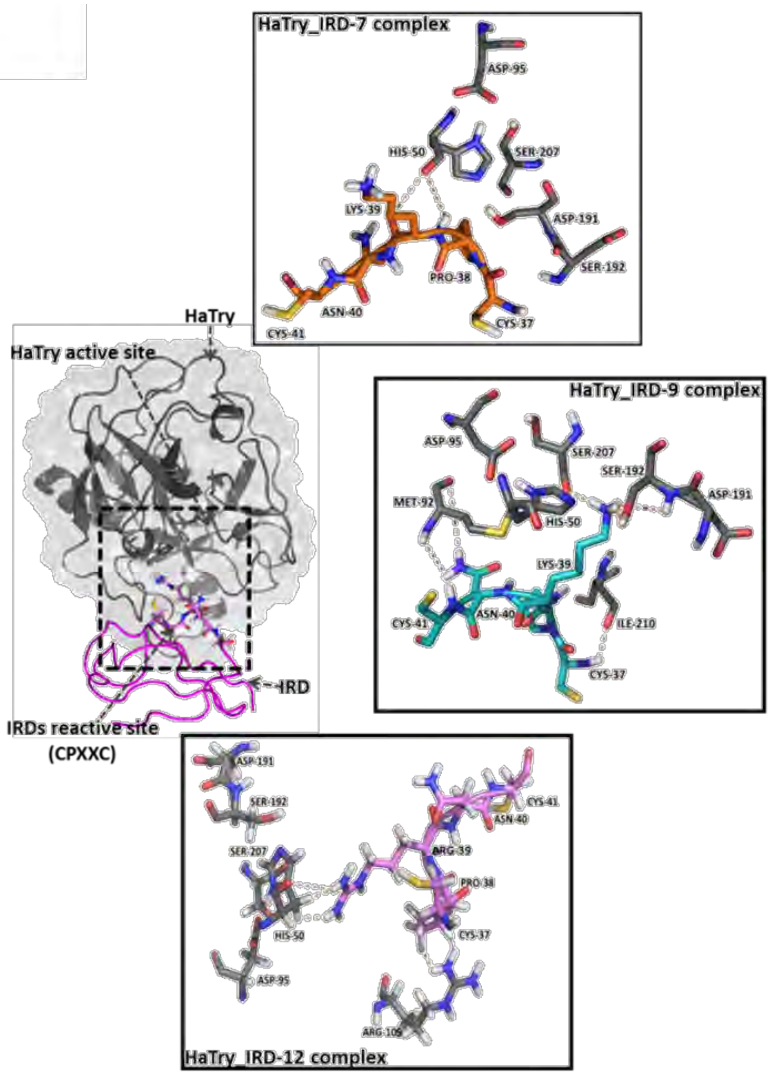


Fig. 3.5: Molecular interaction of IRD-*H. armigera* trypsin. The modeled *H. armigera* trypsin (gray) in complex with the predicted structures of the IRDs. The important residues at the interface of IRDs and trypsin in complexes of IRD-7 (orange), -9 (cyan), -12 (violet) are shown separately in boxes. The models were obtained using a combination of homology modeling, loop prediction, and MDs. Thin dotted wheat colored lines represent hydrogen bonds. Lys39H, Asn40H in IRD-7, -9, and Arg40H in IRD-12 form a number of important contacts with active site of HaTry.

IRD-9 forms the maximum number of stable hydrogen bonds with the active site residues (His50, Met92 and Ser192) of the HaTry and which were maintained for longer duration **(Fig. 3.6)**. Regardless of the fact that IRD-12 forms relatively more hydrogen bonds, however, they are very unstable as reflected by their

fluctuating nature. MD simulations provide structural insight into an importance of inter/intra molecular hydrogen bonds and its effect on the interaction between protease and PIs. The results of this analysis were supported with previous reports (Hansen et al., 2007).

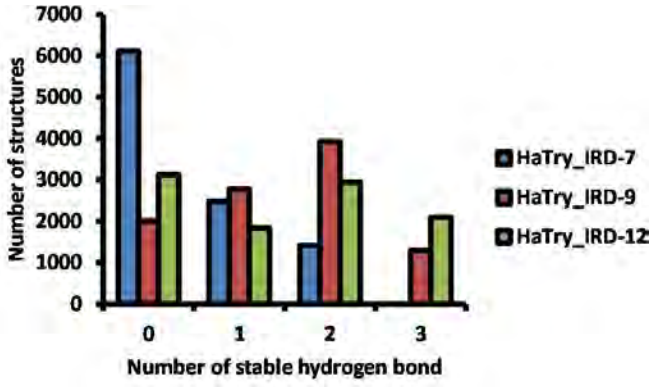


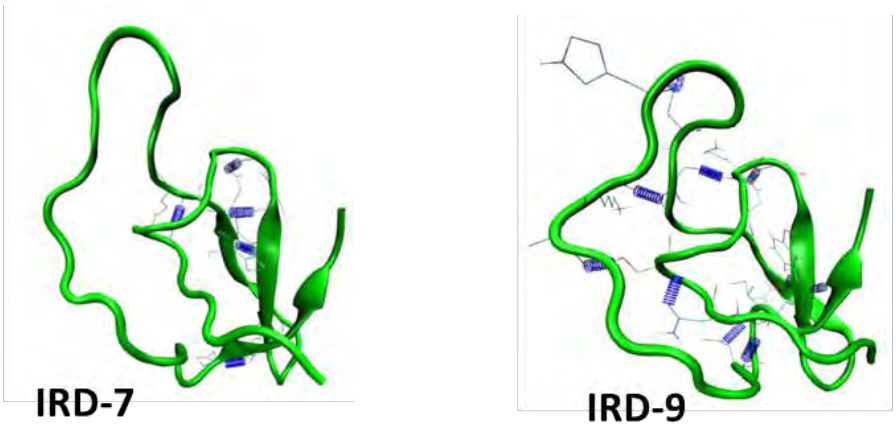
Fig. 3.6: Number of structures having stable intermolecular hydrogen bond in inhibitor's reactive loop and active site of HaTry during 20 ns of simulation time.

Post-simulation analysis also explained the molecular mechanisms behind the enhanced activity of IRD-9 towards proteases. Molecular dynamic simulations revealed that serine residues in place of cysteines at seventh and eighth positions of IRD-9 resulted in an increase in the density of intramolecular hydrogen bonds and reactive site loop flexibility **(Fig. 3.7 A and B)**. Previous reports suggested the role of C4–C41 disulfide bond in maintaining flexibility of the reactive loop and that of C8–C37 in holding a reactive loop of the inhibitor in active and stable form. Interestingly in our study, IRD-9 was found to be a good inhibitor although it lacked a C8–C37 disulfide bond. *In silico*, analysis of a series of mutations at the seventh and eighth positions could provide insights into the significance of C7S and C8S variation on IRD-9 inhibitory activity.

Three variants and mutations tried are IRD-9A: S7A & S8A; IRD-9B: C28S & C37S and IRD-9C: S7A, S8A, C28A, and C37A. IRD-9A (1 hydrogen bond) and -9C (2 hydrogen bond) showed less number of intermolecular and intramolecular hydrogen bonds as compared to IRD-9B (3 hydrogen bonds), in complex with HaTry **(Table 3.4)**. This analysis showed that replacement of cysteine with a hydrophilic

residue, serine can prevent the hydrophobic collapse of the inhibitor molecule and might provide better flexibility and active conformation to the reactive loop and hence enhancement in the inhibitory potential (Schirra et al., 2010). Calculations of the free energy of binding between IRDs and HaTry (ΔG_{bind}) pointed to a comparatively more stable complex formed by IRD-9 with the lowest ΔG value of -68.63 kcal/mol, as compared to IRD-7 (-40.03 kcal/mol) and IRD-12 (-54.01 kcal/mol), a trend similar to that observed in inhibition assays. The free energy of binding was further calculated for HaTry_IRD-9 variants complexes, in which the binding of IRD-9B (-74.14 kcal/mol) was found more stable as compared to IRD-9A (-39.88 kcal/mol) and IRD-9C (-38.04 kcal/mol), respectively.

(A)



(B)

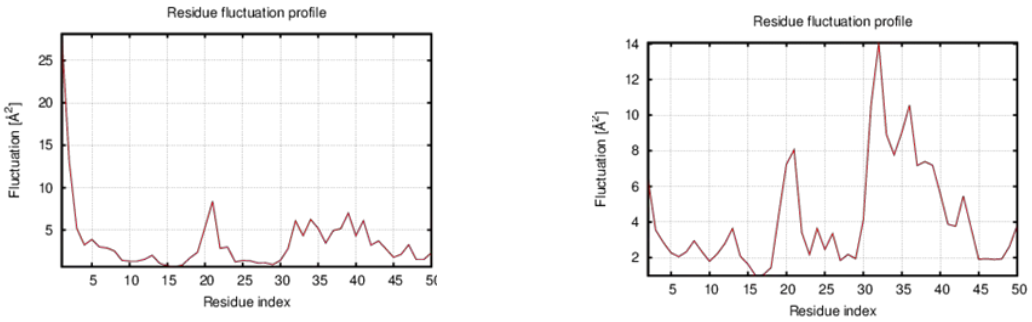


Fig. 3.7: (A) Post simulation analysis of intramolecular hydrogen bond density in IRD-7 and -9. Blue spring like structures indicates the hydrogen bonding in the side chain of IRD-7 and -9 residues. (**B)** Loop flexibility analysis of IRD-7 and -9. IRD-9 showed enhanced flexibility in reactive site loop region as compared to IRD-7.

This analysis of the variants has provided valuable insight for carrying out potential site directed mutations of IRDs for higher stability and adaptability. Higher conformations flexibility of IRD-9 by the loss of two disulfide bonds has helped it to spatially adopt a better complementary shape suited to the active site of HaTry compared to the more rigid four disulfide containing IRD-7 and IRD-12.

Table 3.4: Prediction of various intramolecular and intermolecular interactions of IRDs with respect to H bonds and disulfide bonds formed.

Inhibitor	Variation	Number of disulfide bonds	Number of free cysteine	Number of intra-molecular H bonds	HaTry_IRD complex	
					Number of intermolecular H bonds	$\Delta G_{binding}$ (Kcal/mol)
IRD-7	Null	4	0	10	2	- 40.03
IRD-9#	C7S, C8S	2	2	20	6	-68.63
IRD-9A*	S7A, S8A	2	2	8	2	-39.88
IRD-9B*	C28S, C37S	2	2	12	5	-74.14
IRD-9C*	S7A, S8A, C28A,C37A	2	2	6	1	-38.04
IRD-12	Null	4	0	10	5	-54.27

Natural variant
*Signifies *in silico* variants

III.1.3.6 Solution structure of CanPI-7 and its interaction with multiple proteases

SAXS data were collected from the CanPI-7 in different concentration solutions. Observations obtained for different concentration of protein are listed in **Table 3.5**. SAXS structure of CanPI-7 showed significant fit with the predicted model of CanPI-7 (**Fig. 3.8**).

Table 3.5: Different physical parameters of CanPI-7 measured using SAX data.

Sample	Conc., mg/ml	R_g, nm (autoRg)	R_g, nm (GNOM)	I(0)	D max, nm	MM, kDa	Volume (nm^3)
1	5	3.64	3.76	65.5	12.7	68.39	109
2	4	3.64	3.72	56.6	12.7	59.09	106
3	2	3.48	3.55	52.5	11.9	54.81	100
4	1	3.52	3.52	49.8	11.8	51.99	98
5	0.5	3.61	3.63	54.1	12.4	56.48	95

β-sheets and unordered loops were the major secondary structures of CanPI-7. This model represented the four solvent exposed reactive sites (RL1–RL4) lying on unordered loops as shown in **Fig. 3.8**. RL1 and RL2 were in close proximity with a distance of 21 Å while RL3 and RL4 appeared distant with 48 Å distance.

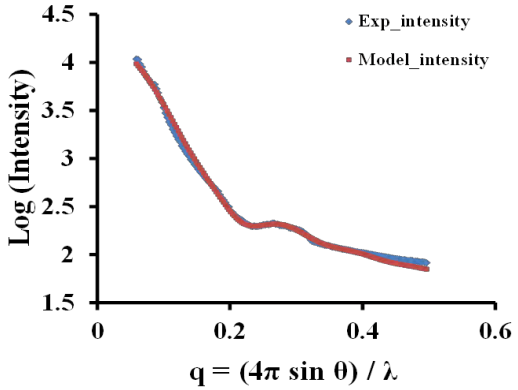


Fig. 3.8: Experimental data (blue dot) with the fit of the predicted model scattering pattern (in red dot) and the theoretical scattering pattern (in red) calculated using FoXS server. The SAXS patterns are displayed as the logarithm of the scattering intensity versus the momentum transfer s = (4Лsinθ)/ λ, where λ is the X-ray wavelength (1.5 Å) and 2θ is the scattering angle. For clarity, only one of every average of all experimental points is represented in the figure.

Earlier structural studies on Pin-II PIs had been limited to two-domain precursors from tomato (Barrett et al., 2003a) and have shown the orientation of the two domains directly facing each other leading to binding with two protease molecules independently. In case of multi domain PIs, the role of inter-domain interactions has been speculated as a key orientations determinant of the domains is the relative distance to each other (Nelson et al., 1999). The presence of 2–6 TI domains in many Pin-II PIs has suggested the functional significance of the combination of IRDs within a single PI. It was intriguing to speculate how the orientation and binding of the four-domain inhibitor occur in case of CanPI-7.

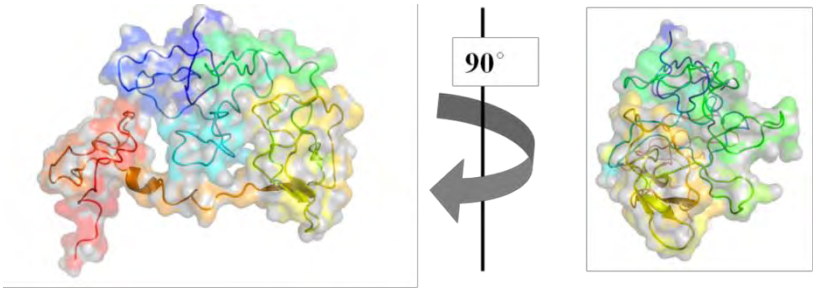
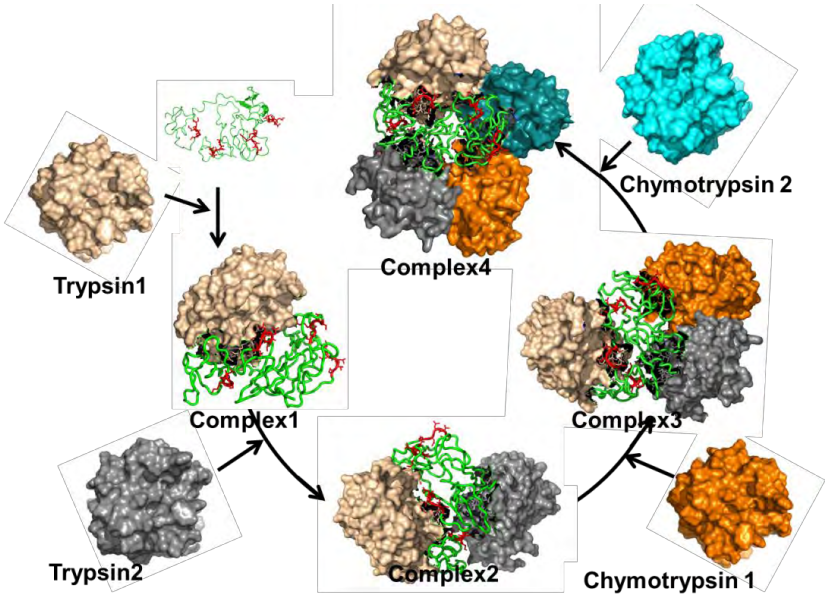


Fig. 3.9: SAXS model of CanPI-7 superimposed with its predicted structure and other view by turned 90°.

In order to understand the inhibitory mechanism of CanPI-7 against target proteases, primary docking simulations were performed with ZDOCK. The goal of the initial stage of docking was to generate as many near-native complex structures (hits) as possible. Four complexes, Complex 1 to 4, were obtained from multiple docking analyses with two trypsin and two chymotrypsin molecules, respectively **(Fig. 3.10)**. Complex1 showed the close proximity of trypsin molecule with the first reactive loop (CTLNC) with distance of 4.6 Å between Cα atom of central residue of the reactive loop and Ser195 of trypsin. Binding of the first trypsin molecule at RL1, caused the steric hindrance for binding of the next trypsin molecule at second reactive loop (RL2) which was only 21Å away from RL1.

However, further two chymotrypsin molecules showed close proximity with the third (RL3) and the fourth (RL4) reactive loops with intermolecular distances of 8.1 and 10.4 Å, respectively. The static complex of CanPI-7 representing at least three protease molecules in close proximity to the RSLs, supports its higher potency corroborating with the enzymatic assays. Earlier studies have suggested inter-domain flexibility in the unbound form of two-domain tomato-inhibitor-II as compared to the bound form of the inhibitor, mediated mostly by slight changes in the inter-domain surface (Barette-Ng et al., 2003; Schirra et al., 2005). Thus, it is possible that both, the bound and the unbound forms of CanPI-7 could exhibit conformational flexibility within the domains relative to each other or in reactive site loops leading to either strong interaction with more number of protease molecules or less than those observed in the static model of the complex (**Fig. 3.11**). Molecular dynamic simulations of each of the complexes or crystal structures would deliver the dynamic picture of this interaction. Cross reactivity in binding of trypsin or chymotrypsin molecules to either TI or CI sites was also observed which suggested that the reactive site loops retained adequate conformational flexibility to allow recognition by a variety of proteinase molecules. PIs being an innate part of the plant defense system, for protecting them from insects, fungi and bacteria, must have enough variability to interact with a wide range of proteases they come across.

(A)



(B)

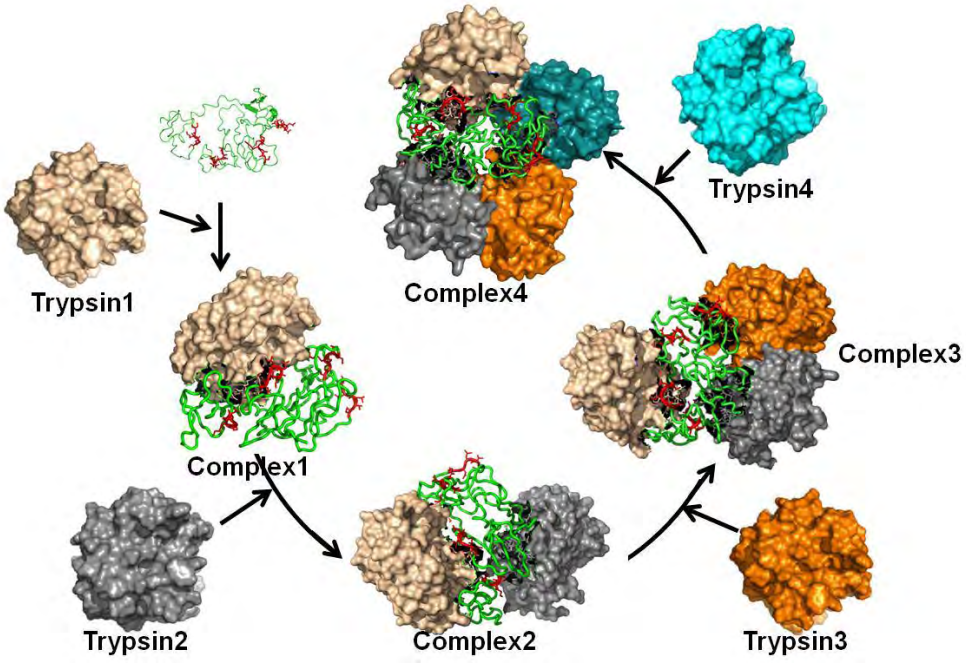


Fig 3.10: **(A)** Docking of CanPI-7 with two molecules of trypsin and two molecules of chymotrypsin. Close proximity to RSL 1, 3 and 4 was observed with protease molecules while binding to RSL2 appeared obstructed because of steric hindrance. **(B)** Docking of CanPI-7 with four molecules of trypsin. Cross reactivity in binding to either TI or CI site was evident.

III.2 Heterologous expression and characterization of *Helicoverpa armigera* protease (HaTry4) and IRD-HaTry4 complex

III.2.1 INTRODUCTION

In insects, proteases are a major group of hydrolytic enzymes categorized as endo- and exo-proteases, and are predominantly involved in digestive processes, pro-enzyme activation, metamorphosis, the release of physiologically active peptides and complement activation (Terra, 1988; Terra and Ferreira, 1994). Serine proteases namely, trypsin and chymotrypsin are abundant in the digestive tract of *H. armigera* (Srinivasan *et al.*, 2006). Twenty-one trypsin, 14 chymotrypsin, 2 elastase-like, several amino-peptidases and carboxypeptidases genes were found in the gut tissue of *H. armigera* that had been reared on a high-protein diet free of inhibitors (Bown *et al.,* 1997; 1998; Gatehouse *et al.,* 1998; Bown *et al.*, 2004; Chougule *et al.*, 2005; Angelucci *et al.,* 2008). These enzymes are released extracellularly into the gut lumen and are active at an alkaline pH (Johnston *et al.,* 1991; Purcell *et al.,* 1992). The super-family of serine proteases formed a catalytic triad with serine, histidine, and aspartic acid. In the case of trypsins, the Asp (189), Gly (216) and Gly (226) residues contribute to a negatively charged S1 (ratio of residues flanking to the catalytic site) site, thus it is highly specific for the positively charged side chain of arginine or lysine in the substrate. Similarly, residues of S1 sites form a deep hydrophobic pocket of chymotrypsin and make it more preferable for phenylalanine, tryptophan, and tyrosine at P1 on the substrate (Srinivasan *et al.,* 2006).

From the diverse *H. armigera* gut trypsins and chymotrypsins, we have selected fewfor obtaining detailed insights of their functional attributes. We cloned, expressed and purified one of the most prevalent *H. armigera* trypsin 4. We have performed kinetic study, biochemical and biophysical characterization of the purified recombinant trypsin 4 (rHaTry4).

III.2.2 MATERIALS AND METHODS

III.2.2.1 Cloning of *H. armigera* trypsin (HaTry4) and HaTry+IRDs complex

Based on sequence analysis and previous expression analysis *H. armigera* trypsin 4 (HaTry4; Accession No. EF600059) was selected for the expression. Mature peptide region for HaTry4 was amplified by PCR from a *H. armigera* cDNA library using gene specific primers (FP: CGGGGTTUCATCACCATCATCATCACGGGAGCAGCA GCCCGAGCCGTATTGTTG; RP: CCCCCCGTTUTTATTTCTTCTCGAACTGAG GGTGAGACCAAGAACCAACTGCGGTCTGAATAATCCAATCGGTA) and cloned into the pOECDS vector, Shuttle with an N-terminal His X 6 tag as described (Yang et al., 2009).The diagrammatic skeleton of clone is shown in the **Fig.3.11.** The USER methodology was used for all the cloning practice. Clones were confirmed by sequencing and further used for expression.

Fig. 3.11: Diagrammatic representation of the clone for pOECDS_HaTry4 clone.

To obtain the HaTry4 and IRDs complex, we performed coexpression cloning in pOECDS vector. We have used two set of primers PCR1 and PCR2 (PCR1_FP: CG GGGTTUGAACCCATATGCACCAATTGTTGTGCAG; PCR1_RP: AGCCGTCGU CATATGGGCATATGTTGCCAAACTCTAAACC; PCR2_FP ACGACGGCU CCGG GACCTTTAATTCAACCCAACAC; PCR2_RP: CCCCCCGTTUTTAGACGGCGG TTTGGATGATCCAGT). PCR1 forward primer includes P10 promoter and starting region of IRDs construct, while reverse primer contains sequence complementary to terminal region of IRD and linker region. In PCR2 primers, forward primer contain sequence complementary to linker region and start region of HaTry4 construct, while reverse primer contain complementary sequence to HaTry4 gene terminal region. Schematic of coexpression cloning is depicted in **Fig. 3.12.**

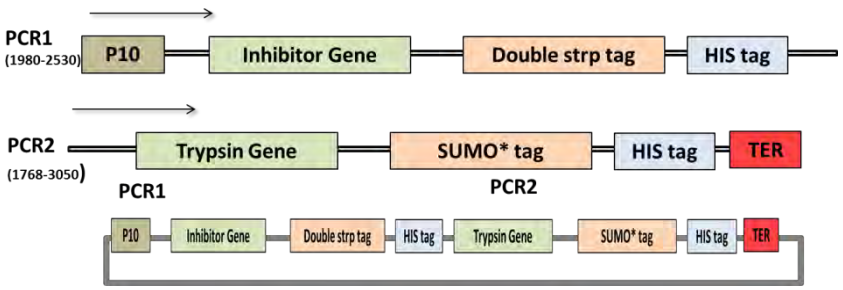


Fig 3.12: Diagrammatic representation of the clone for pOECDS_HaTry4+IRD clone. PCR1 covers the P10 promoter and IRD gene start region, whereas PCR2 contain the HaTry4 end region and termination site.

III.2.2.2 Expression and purification of recombinant HaTry4

Recombinant HaTry4 and rHaTry4+IRDs complex were expressed and purified. All the recombinant proteins were expressed in *Escherichia coli* C41 cells were initially grown overnight at 37ºC in 10 ml LB medium supplemented with antibiotics (Ampicillin-100 µg/ml). The pre-culture was used to inoculate 1 lit LB broth medium with appropriate antibiotics and allowed to grow until the OD (600 nm) reached 0.5 to 0.6. Cells were induced with IPTG (0.5 mM) overnight at 16˚C and harvested by centrifugation. The cell pellet was solubilized in ice cold cell lysis buffer A (50 mM Tris-HCl, pH 8.0; 300 mM NaCl; 5% glycerol) and disrupted by sonication (10 s pulse with 10 s intervals for 10 min) using an Ultrasonic Disruptor UD-201. The supernatant was collected by centrifugation for 60 min at 14,000g, 4˚C (Sorvall) and loaded on Ni-NTA resin (Qiagen, Valencia, California, USA) for purification using standard manufacturer's protocol. The HaTry4 protein eluted with buffer B (50 mM Tris-HCl, pH 8.0; 300 mM NaCl; 5% glycerol; 250 mM imidazole). This was applied on pre-equilibrated (buffer: 50 mMTris-HCl, pH 8.0; 200 mM NaCl; 5% glycerol) Sephacryl S-200 (size exclusion) for further purification.

III.2.2.3 MALDI-TOF MS analysis

Recombinant rHaTry4 and rHaTry4+IRD complex (2 µg) were thoroughly mixed with sinapic acid matrix and analyzed on MALDI-TOF. Detailed procedure is described in section III 1.2.4.

III.2.2.4 Enzyme kinetic study

Trypsin kinetic activity assays for rHaTry4 was performed using BA*p*NA as a substrate. Michaelis-Menton constant (Km) for trypsin was calculated by using various concentrations of BA*p*NA substrate (1 to 5 mM), and then plotting double reciprocal curve with 1/ (v) and 1/ (s). Slope and the line equation were used to determine Km and Vmax for HaTry4.

III.2.2.5 CD spectroscopy analysis

The CD spectra of rHaTry4 and bovine trypsin were recorded on a J-815 spectropolarimeter (Jasco) at 28°C in quartz cuvettes. Detailed procedure is described in section **III 1.2.5**.

III.2.2.6 Determination of pH optima and stability of the rHaTry4

The activity of purified rHaTry4 in various buffers was evaluated by adding 20 µg of purified recombinant protein in pre-incubated 500 µl buffer of pH ranging from pH 3 to 12 (Glycine–HCl buffer pH 3, acetate buffer pH 4, 5, phosphate buffer pH 6, 7 Tris–HCl buffer pH 8, 9, and glycine–NaOH buffer pH 10, 11 and 12) and this premix was used for subsequent BA*p*NA assay. For the determination of pH stability, the purified enzyme was incubated in buffers having different pH ranging from pH 7 to 12 for 1 h at 37ºC and the enzyme activity was determined under standard assay conditions as described earlier (Sanatan et al., 2013). The same sample was used for the CD spectroscopy analysis to understand the structural stability of HaTry4 under different pH conditions.

III.2.2.7 Effect of temperature on activity and stability of the rHaTry4

The stability of purified protease at different temperatures (30 to 100°C) was evaluated by adding 10 µg purified enzyme in Tris HCl pH 7.8. The reaction mixture was pre-incubated for 30 min at respective temperatures before activity measurement. After pre-incubation the standard BA*p*NA assay was carried out. To examine the temperature stability, the purified enzyme was incubated at different temperature ranges from 40 to 90ºC for 1 h and the residual enzyme activity was measured under standard assay conditions (Sanatan et al., 2013). The same sample

was used for the CD spectroscopy analysis to determine the structural stability of rHaTry4 under different temperature conditions.

III.2.2.8 Effects of inhibitors and metal ions on the rHaTry4 activity

The effect of enzyme inhibitors on purified rHaTry4 was studied using PMSF (1 mM), EDTA (5 mM), DTNB (5 mM) and β-mercaptoethanol (5 mM). The reaction mixture was prepared by pre-incubating the purified rHaTry4 with inhibitors for 10 min at 40ºC. The residual protease activity was determined under standard assay conditions. The activity of the purified rHaTry4 determined without inhibitor was considered as 100%.

The effects of monovalent (Na^+ and K^+) and divalent ($Cd2^+$, $Zn2^+$, $Cu2^+$, $Ba2^+$ and $Hg2^+$) metal ions on purified rHaTry4 activity at a concentration of 1, 5 and 10 mM were investigated by using BA*p*NA as substrate. The reaction mixture was prepared by pre-incubating the purified rHaTry4 with metal ions at each concentration for 10 min at 37 °C. The residual protease activity was measured under standard assay conditions. The activity of the enzyme without metallic ions was considered as 100%.

III.2.2.9 Site directed mutagenesis of rHaTry4

The quick change lightning site-directed mutagenesis kit (Stratagene, Edinburgh, UK) was used to introduce the K2111A, S218A, S218C mutation into the rHaTry4. Synthesize two complimentary oligonucleotides containing the desired mutation, flanked by unmodified nucleotide sequence. Cycle each reaction using the standard cycling parameters. After completion of amplification, add 2 µl of the provided DpnI restriction enzyme directly to each amplification reaction. Incubate at 37°C for 5 min to digest the parental (i.e. non-mutated) supercoiled dsDNA. Gently thaw the XL10-Gold ultra competent cells on ice. For each control and sample reaction to be transformed, aliquot 45 µl of the ultra-competent cells to a prechilled 1.5 ml tube. Transfer 2 µl of the DpnI treated DNA from each control and sample reaction to separate aliquots of the ultra-competent cells. Incubate the reactions on ice for 30 minutes followed Heat-pulse the tubes in a 42°C water bath for 30 s. Plate the

appropriate volume of each transformation reaction on agar plates containing the 100 µl /ml of ampicilin.

III.2.2.10 Preparation and purification of IRD- Porcine pancreatic trypsin complex

For preparation of the complex, 5:1 molar amounts of IRDs and porcine pancreatic trypsin (Sigma) were mixed and incubated for 30 min at 25˚C in 20 mM Tris pH 8.0, 200 mM NaCl and 5% glycerol. The mixture was then applied onto a SephacrylS-200 HR (GE Healthcare) column pre-equilibrated with the same buffer. Fractions of 2 ml were collected at a flow rate of 0.2 ml/min. Samples pooled from different peaks were verified by SDS-PAGE. The fractions corresponding to IRDs-trypsin complex were selected and concentrated to around 10 mg/ml. The freshly prepared complex sample was aliquoted and immediately used for crystallization trials. The remaining aliquots were stored at -80^{o}Cfor future optimization.

III.2.2.11 Crystallization setup of the rHaTry4, IRDs-trypsin complex

Initial crystallization trials of the HaTry4 and IRDs-trypsin complex at 15 to 20 mg/ml were performed by using a mosquito LCP crystallization robot. Different crystallization screens Classic, JCSG+, JCSG Core II and Procomplex (Qiagen, Valencia, CA, USA) were tested. Purified protein was loaded on one side of The Crystal Former (Microlytic) and 0.35 µl of precipitant were loaded on the other side. Plates were incubated 20°C for several days.

III.2.3 RESULTS AND DISCUSSION

III.2.3.1 Cloning and purification and molecular weight determination of HaTry4

HaTry4 was cloned and confirmed sequence of HaTry4 is shown in **Fig 3.11A**.Active site of HaTry4 is highlighted and marked in the **Fig. 3.13A**. Furthermore, we expressed and purified HaTry4 to obtain homogenous protein, which was depicted in 12% SDS-PAGE (**Fig. 3.11B**). Expressed rHaTry4 was characterized by MALDI-TOF mass spectrometry analysis. The molecular mass for recombinant HaTry4 determined by MALDI-TOF mass spectrometry is 29.8 kD, which is almost similar to

theoretical mass (29.4 kD) calculated using ExPASy -ProtParam. In case of HaTry4+IRD complex, we observed two overexpressed band one corresponds to HaTry4 (43 kD) and IRD (10 kD).

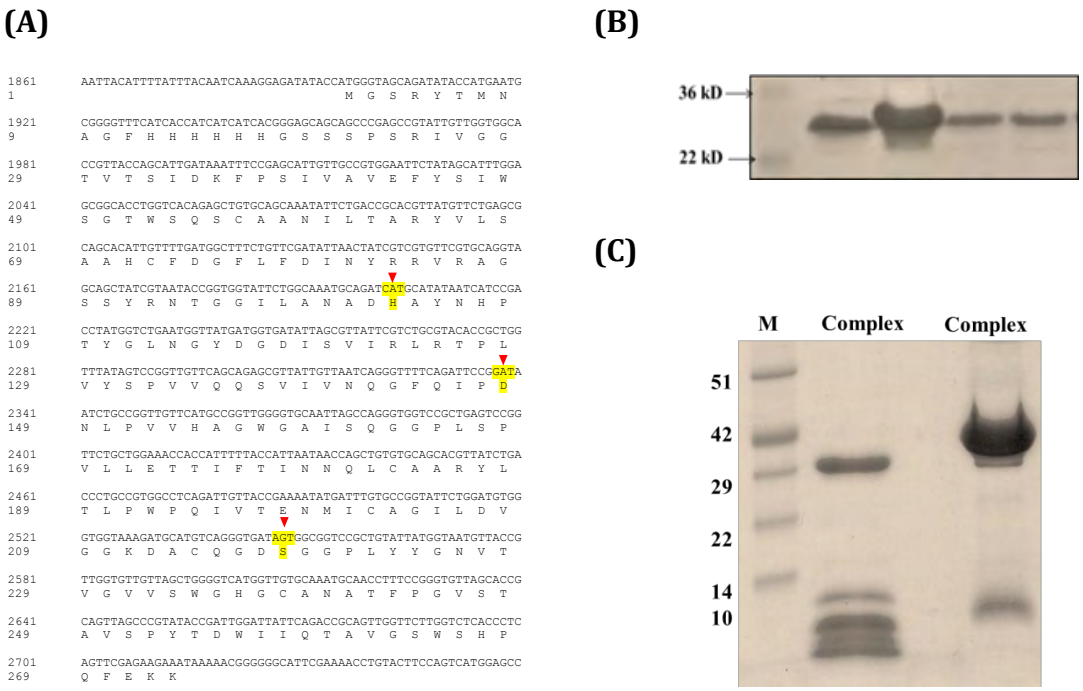


Fig. 3.13: Expression and purification of recombinant HaTry4. **(A)** Cloned sequence of HaTry4 aligned with its translated sequence **(B)** Purification of recombinant HaTry4. Protein was separated on 12% SDS-PAGE and stained with Coomassie Blue R-250 **(C)** Expression of recombinant HaTry4+IRD complex. Protein was separated on 12% SDS-PAGE and stained with Coomassie Blue R-250

III.2.3.2 Enzyme assays and kinetic parameter assessment

Activity of expressed rHaTry4 was first estimated by using azocasein as substrate. Different kinetic parameter for the rHaTry4 was estimated using BA*p*NA as trypsin specific substrate. Enzyme activity at different concentrations showed in **Fig. 3.14A**, it is observed HaTry4 activity increased with substrate concentration and reach to saturation phase at 4 to 5 mM BA*p*NA concentration. In **Fig. 3.14B**, double reciprocal curve with 1/ (v) and 1/[s], which is used to estimate Vmax and Km. Calculations for different kinetics parameters are shown below, which depicted that rHaTry4 had Km = 36.79 µM and Vmax = 18.39.

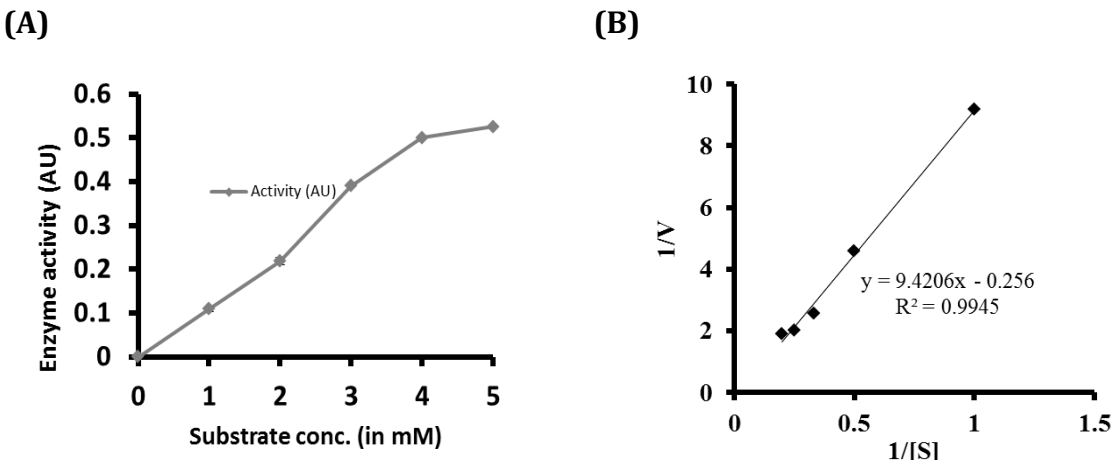


Fig. 3.14: Kinetic properties of HaTry4. **(A)** Saturation curve for an enzyme showing the relation between the concentration of substrate and rate **(B)** Lineweaver–Burk or double-reciprocal plot for the calculation of Km and Vmax. The data represent the average of three independent assays.

III.2.3.3 Effect of pH on activity and stability of the purified rHaTry4

The effect of pH on purified rHaTry4 activity was determined over a pH range of 3 to12. The enzyme was highly active in the pH range of 8 to 11 with an optimum at pH 9 (**Fig. 3.15A**). The pH stability profile of the enzyme is shown in **Fig. 3.15B**. It observed that the purified enzyme was stable at a pH range between 6.0 and 11.0, maintaining 90% of its original activity after 1 h incubation at 37ºC. CD spectroscopy analysis showed secondary structure of HaTry4 was maintained through wide range of pH from 5to12. These observations indicated that rHaTry4 is active at alkaline pH and exhibited structural stability at variable pH.

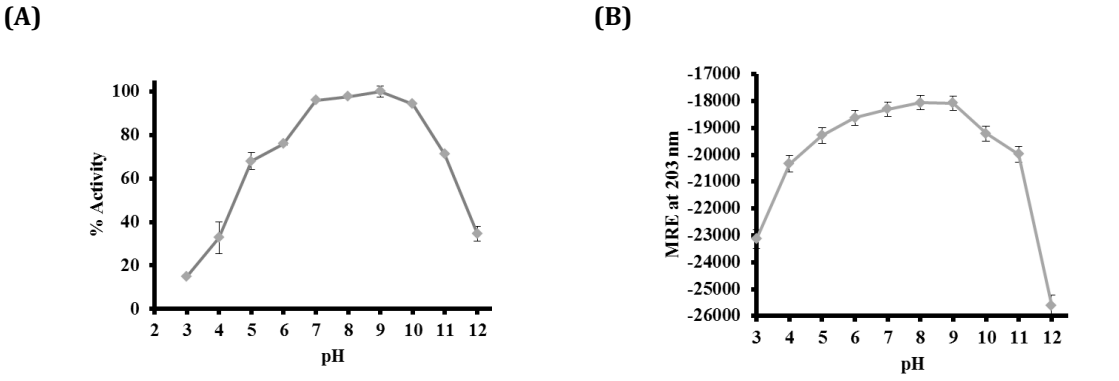


Fig 3.15: Effect of pH on (A) activity and (B) structural stability of the rHaTry4

III.2.3.4 Effect of temperature on activity and stability of the rHaTry4

The effect of temperature on purified rHaTry4 activity was determined by using BA*p*NA as substrate. The enzyme was active at temperatures from 30 to 60 °C with an optimum at 40ºC (**Fig. 3.16A**). After 70ºC, the enzyme activity was decreased rapidly. The thermal stability profile showed that the enzyme is fully active for at least 100 min between the temperatures range 40 to 60ºC. The protease activity relative to control at 50, 60, and 70 °C were about 84, 78 and 62%, respectively (**Fig. 3.16A**). CD spectroscopy analysis showed secondary structure of HaTry4 was maintained through a wide range of temperature from 40 to 80ºC. These observations indicated that HaTry4 is active at 30 to 50ºC and exhibited structural stability at variable temperature (**Fig. 3.16B**).

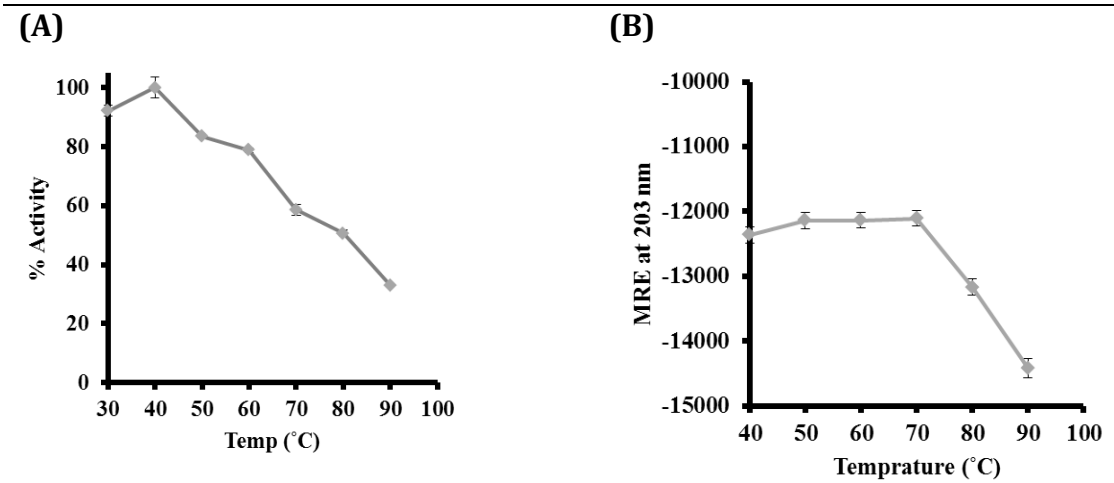


Fig 3.16: Effect of temperature on (A) activity and (B) stability of the purified rHaTry4. The temperature optimum was determined by assaying protease activity at different temperature ranges between 30 to 90 °C, while the structural stability was checked at temperatures ranges between 40 to 90 °C for 1 hr. The purified protease was found to be stable at higher temperature. The error bars show the standard deviation of at least tree replicates.

III.2.3.5. Effect of inhibitors and metal ions on the rHaTry4 activity

In order to determine the nature of the purified HaTry4, the effects of various inhibitors such as PMSF, EDTA, DTNB and β-mercaptoethanol were investigated. The enzyme activity was completely inhibited by the serine protease inhibitor PMSF **(Fig. 3.17)**. The rHaTry was completely deactivated in presence of PMSF, which

confirms that serine plays an important role in enzyme activity. In presence of thiol reagent DTNB, chelating agent EDTA and β-mercaptoethanol protease activity was slightly reduced as compared to control.

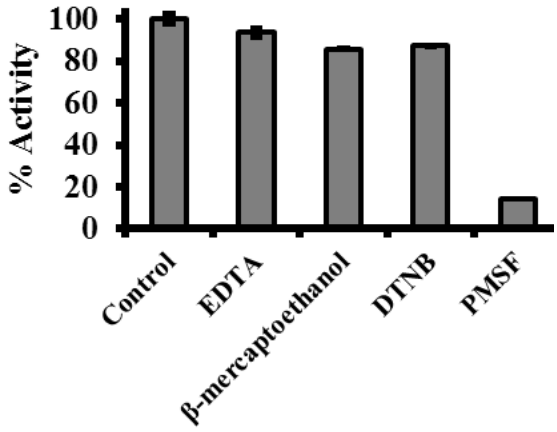


Fig. 3.17: Effect of inhibitors on the activity of purified HaTry4 protease. Different inhibitors were pre-incubated with the purified enzyme for 30 min at 37°C and assay was carried out using casein as substrate. The activity of enzyme without inhibitor was considered as 100%. The error bars show the standard deviation of at least tree replicates.

The effect of various metal ions (at three different concentrations; 1, 5 and 10 mM) on the activity of HaTry4 was investigated at pH 9.0 and 40ºC by addition of the respective metal ions to the reaction mixture. Results are shown in **Table 3.6**. It was noted that at a concentration of 5 and 10 mM of Ca2+, Zn2+, Ba2+, Hg2+, Cu2+, Na+ and K+ enzyme activity was slightly decreased. While at 20 mM concentrations of each metal ion significant inhibition of the enzyme was observed.

Table 3.6: Effect of metal ions on the activity of purified HaTry4

Metal ion	Relative activity (%)		
	1 mM	5 mM	10 mM
Control	100	100	100
Ca2+	96	92	80
Zn2+	98	95	80
Cu2+	98	93	85
Ba2+	95	95	89
Hg2+	95	92	81
Na+	98	94	85
K+	95	92	82

III.2.3.6 Site-directed mutagenesis and purification of mutants

We performed site-directed mutagenesis to individually replace K211A, S218A and S218C residues. Positive clones were selected and sequenced. Mutations were

incorporated in respective sites and mutant plasmids were further cloned for their expression **(Fig. 3.18).**

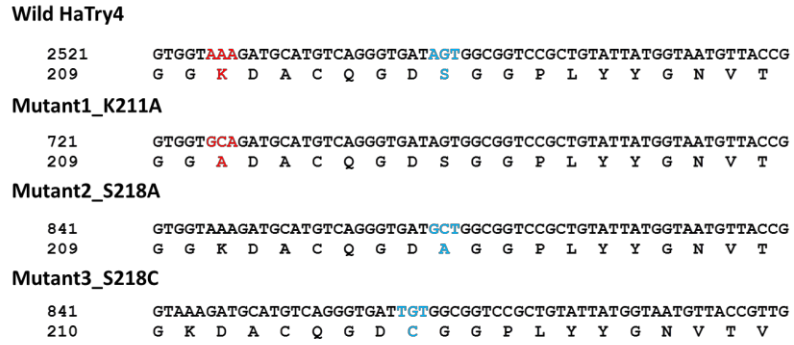


Fig. 3.18: Sequence of HaTry4 mutant with mutation at position K211A, S218A, S218C

III.2.3.7 Characterization of IRDs-rHaTry4 complex

IRDs exhibits strong binding with porcine pancreatic trypsin and a number of other serine proteases. A larger amount of the IRD-trypsin complex was prepared with a Sephacryl S-200 HR column. A reasonably large single peak corresponding to IRD-trypsin complex was observed in chromatogram (**Fig. 3.17A**). Formation of the IRD-trypsincomplex was tested on 15% SDS-PAGE as shown in **Fig. 3.17B.**

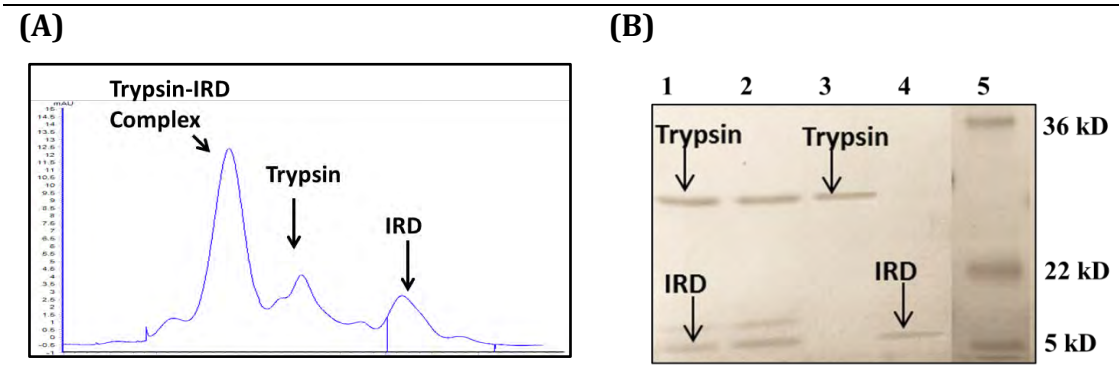


Fig. 3.19: Preparation and purification of IRD-trypsin complex (A) Free IRD, trypsin and IRD-trypsin complex were run on the same Sephacryl S-200 HR column, with the same buffer. The three curves were observed, with first representing IRD-trypsin complex and remain two are representing free trypsin and IRD, respectively. **(B)** Purification of IRD-trypsin complex using Sephacryl S-200 HR column. Proteins were separated on 15% SDS-PAGE and stained with Coomassie Blue R-250. Lane 1 and 2, IRD-trypsin complex; Lane 3, free trypsin; Lane 4, free IRD; Lane 5, Marker.

III.2.3.8. Protein crystallization

In Initial crystallization trials of the IRD-trypsin complex, some hits were obtained from standard crystallization screens Pro-complex and JCSG Core II. Rhombohedra shape crystals were obtained, we have used additive screen for the improvement of crystals. Here we have shown on crystals of IRD7-trypsin complex as representative crystals (**Fig. 3.20**).

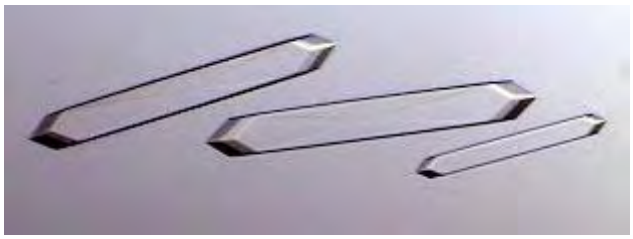
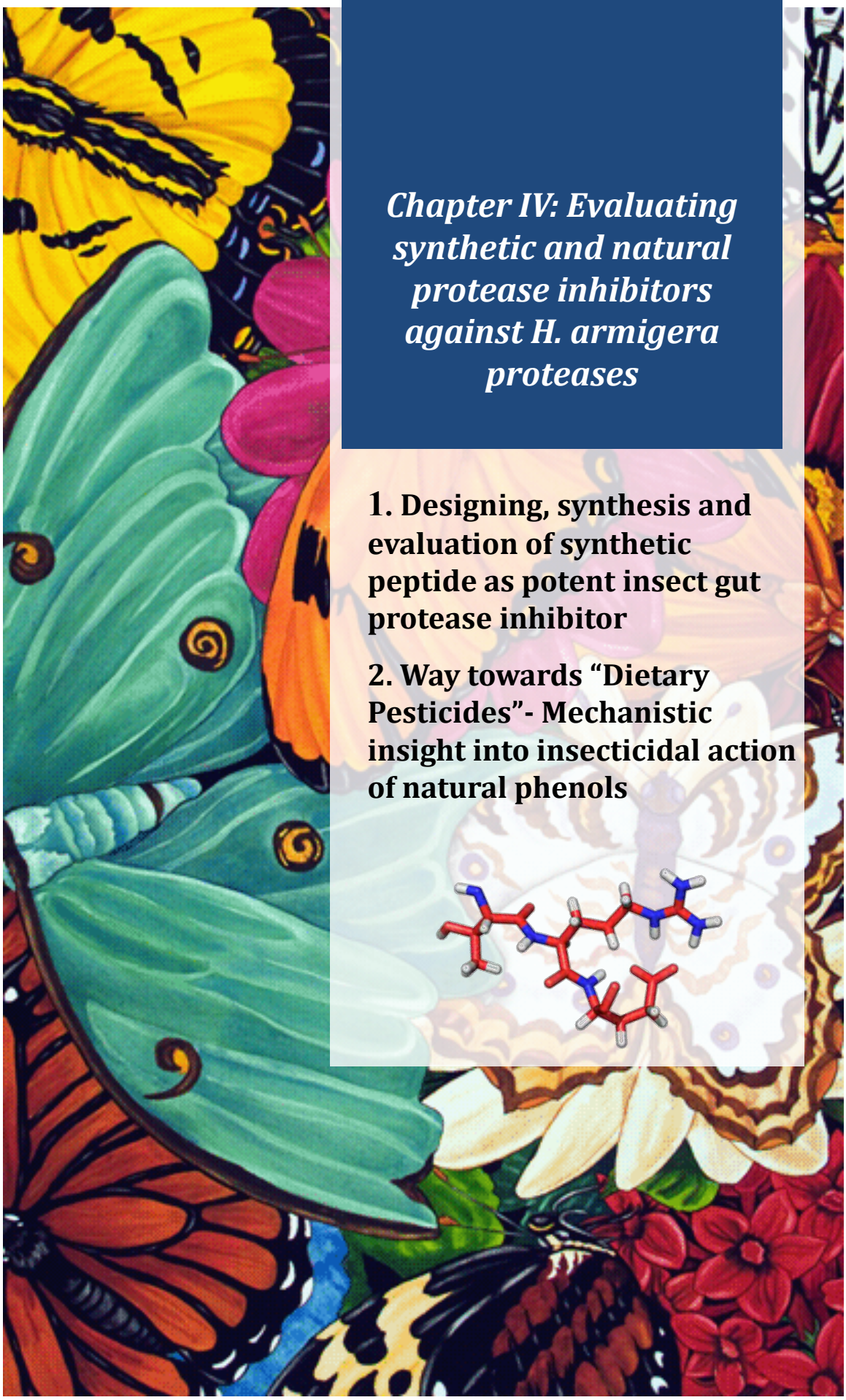


Fig. 3.20: Crystals of IRD7-trypsin complex (need bouquet shape). These crystals were grown in 0.1 M HEPES pH 7.5 and 70 % MPD.

III.3 CONCLUSION

Interaction of IRDs and CanPIs with protease is highly exothermic and thus spontaneous in nature. In case of IRD9 flexibility of reactive site loop is determining factor in enhanced inhibitory activity. Multidimensional analysis exhibited that although in case of CanPI-7 although there are 4 reactive sites, but due to steric hindrance only three reactive sites can be functional simultaneously. We have successfully purified the *H. armigera* and it resembles with porcine trypsin in term of secondary structure and kinetic parameters. Crystals of IRDs-porcine trypsin complex were obtained which can be further use for structural analysis of this interaction and the information obtained can be used for structure-based designing of the insect protease inhibitors.



Chapter IV: Evaluating synthetic and natural protease inhibitors against H. armigera proteases

1. Designing, synthesis and evaluation of synthetic peptide as potent insect gut protease inhibitor

2. Way towards "Dietary Pesticides"- Mechanistic insight into insecticidal action of natural phenols

Chapter IV: Evaluating synthetic and natural protease inhibitors against *H. armigera* proteases

IV.1 Designing, synthesis and evaluation of synthetic peptide as potent insect gut protease inhibitor

Reactive site region of 387 Pin-II PIs was used to design tri-peptides for its potential application as insect protease inhibitor. Out of 23 tripeptides 6 peptides (TLN, PRN, TRE, PKN, PLN, PRY) were predominantly distributed (81%) among the Pin II PIs population, and also showed strong binding score against various insect proteases. These peptides were further synthesized on solid phase platform using F-Moc chemistry. Furthermore, these peptides were characterized and assessed for their *in vitro* protease inhibition potential. Most of the selected peptides showed significant protease activity inhibition; while TRE was the most potent protease inhibitor with Ki~24µM. Feeding bioassay showed that 50 to 200 ppm of TRE had significant negative impact on growth and survival of *H. armigera*. Docking analysis of TRE and other RSL tripeptides against various insect protease exhibited variation in the molecular interaction. Differential molecular responses at translational and transcriptional level were observed in TRE fed larvae. These results shed light on the potential of small tailored peptides for further selective and multi-targeted inhibitor design for effective pest control.

IV. 1.1 INTRODUCTION

Basic structure of IRD consists of 50 aa including eight conserved cysteine residues. Reactive site loop (RSL) is a part of IRD that interacts with target proteases and it is found to be highly variable. Co-evolution of RSL with their target protease indicates its crucial role in plant-insect interaction (Jongsma & Beekwilder, 2011). Variation in RSL results in deviation in their affinity and specificity toward target protease, thus this information can be explored to tailor small peptide inhibitors against *H. armigera* gut proteases. In second chapter of this thesis, we have shows effect of structural and functional variations in three IRDs namely, IRD-7, -9 and -12 resulted in differential antibiosis effect on *H. armigera* growth. These findings indicate that

IRD-9 exhibits enhanced protease inhibition due to lack of disulfide bond and flexibility in reactive loop as compared other IRDs.

Here, we have *in silico* screened 23 RSL peptides and analysed for their occurrence frequency. Six peptides were selected, synthesized by F-Moc chemistry on solid phase. Different inhibitory kinetics and *in vitro* studies of synthesized peptides was performed to understand inhibitory specificities of these inhibitors against trypsin, chymotrypsin and HGPs. *In vivo* efficacy of these tripeptides was analysed by monitoring growth performance and nutritional parameters. Molecular response of insect digestive proteinases after ingestion of peptide was evaluated by proteinase gene expression, activity and zymography studies. This report demonstrates the approach of exploring RSL sequence variations for designing a potent inhibitor for effective control of insect pests.

IV. 1.2 MATERIALS AND METHODS

IV.1.2.1 Materials

The protected amino acids were purchased from NovaBiochem (Merck Milipore, Darmstadt, Germany). All solvents used during the synthesis of peptides were of peptide synthesis grade and for HPLC, of HPLC grade. Enzymes and substrates were obtained from Sigma Chemical Co., St. Louis, MO, USA. X-ray films and Kodak163 DA developer were purchased from Kodak (Chennai, India). Highly pure chemicals for *AD* (AD), natural diets and the rest of the insect rearing materials of the highest purity were purchased locally. *H. armigera* larvae were maintained in laboratory as described in earlier section **II.2.1**.

IV.1.2.2 Synthesis of peptide

Our interest in synthesis of these peptides is exploration of RSL variation for obtaining effective protease inhibitor, thus we eliminated the cysteine residues from both end and only three central residues of RSL were considered for further synthesis. The tripeptides obtained from MEME and occurrence frequency analysis were synthesized on MBHA resin as the solid support (1.75 mmol/g). The synthesis

was carried out manually using 3 equivalent of Fmoc-protected amino acids, and HOBt and TBTU as the coupling agents, in the presence of N,N-Diisopropylethylamine. Successive deprotection and coupling steps were carried out as iterative cycles until the desired tripeptide was synthesized. The deprotection of the Fmoc-protected amino group and the coupling reaction were monitored by the Kaiser test. The N-terminal amino group, after Fmoc-deprotection, was capped using acetic anhydride in the presence of dry pyridine. The synthesized tripeptides were cleaved from the solid support using TFA-TFMSA to yield tripeptides with N-terminal acetates and C-terminal amides. They were purified by RP-HPLC using an increasing gradient of acetonitrile in water containing 0.1% TFA, and characterized by MALDI-TOF mass spectrometric analysis.

IV.1.2.3 Motif and distribution analysis

To characterize reactive loop region, we first investigated features represented by conserved regions or motifs. From several algorithms available, we chose the expectation maximization method MEME (Bailey *et al.*, 2009). For MEME, a fixed minimum motif length of 5 and a maximum of 10 was set and 20 motifs were requested using the zero or one occurrence per sequence model. The results obtained with MEME were further used for distribution as well as frequency analysis. The obtained reactive loop motifs with their occurrence frequency in Pin II PIs population were plotted in bar graph.

IV.1.2.4 Virtual screening

Structures for predominant tripeptides were predicted using Chimera software. The structure was energy minimized for 10000 steps with steepest gradient. Structures of *H. armigera* trypsins, chymotrypsins, cathepsins were predicted using CPH model server. A docking study was performed to determine the binding energy and interaction of tripeptides with *H. armigera* proteases. Predicted structures of proteases were refined by energy minimization and restraint relaxation using Swiss PDB-Viewer (v4.1.0) (Guex *et al.*, 1997). In order to perform molecular docking,

models of *H. armigera* proteases and inhibitors were submitted to Patchdock online server (http://bioinfo3d.cs.tau.ac.il/PatchDock/) following the standard package protocols and further refined by FireDock online server (Schneidman-Duhovny *et al.*, 2005). Binding energy obtained for each complex was normalized by mean values and represented in heat map format using MeV software packages (http://www.tm4.org/mev/). The gradient ruler from -5 to 5 is an indicator of interaction strength. Data was clustered using hierarchical clustering method (Saeed *et al.*, 2006).

IV.1.2.5 Inhibition kinetics

Concentration dependent reduction in trypsin activity by synthesized tripeptides was also estimated through a BA*p*NA assay. Detailed methodology is given in section **II.3.4**.

IV.1.2.6 Feeding assay

Bioassays were conducted by feeding *H. armigera* larvae on AD containing PIs (Tamhane *et al.*, 2005). AD was supplemented with the tripeptides in 50 to 200 ppm/g (25 to 100 μM) of AD. The neonates obtained lab-reared moths were reared for the first 2 days on control diet and then transferred to peptide-containing. Detailed methodology is given in section **II.3.7**.

IV.1.2.7 Assessment of nutritional parameters

Various nutritional parameters of fourth-instar larvae exposed to control and tripeptide-containing diet were compared. Detailed methodology is given in section **II.3.8**.

IV.1.2.8 *In vivo* inhibition of protease activity

Inhibition of total gut protease, trypsin and chymotrypsin activities of larvae fed on 50 to 200 ppm tripeptide were determined using azocasein and other specific

substrates. Assay procedures were already discussed in earlier sections of thesis (**section II.3.4**). Minimum three replicates of each experiment were performed.

IV.1.2.9 Semi-quantitative PCR

Total RNA was isolated from the TRE-fed insect gut tissues using Trizol reagent (Invitrogen, Carlsbad, CA, USA) and synthesis of the first strand cDNA was carried out with High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Foster, CA, USA) using random primers based on the manufacturer's protocol. Detailed methodology is given in section **II.3.10**.

IV.1.2.10 Statistical analysis

All data were statistically analyzed by independent sample t-test. Asterisks indicate significant differences (* p<0.05; ** p<0.01).

IV.1.3 RESULTS AND DISCUSSION

IV.1.3.1 Identification of conserved reactive site loop motifs

To identify reactive site loop motifs in 389 Pin II PIs, we first searched for conserved in regions of 5 residues with cysteines at both terminals. The search was carried using MEME. Twenty-three motifs correspond to RSL were retrieved per PIs sequence. A comprehensive list and sequence of the all motifs is provided in **Fig.4.1**.

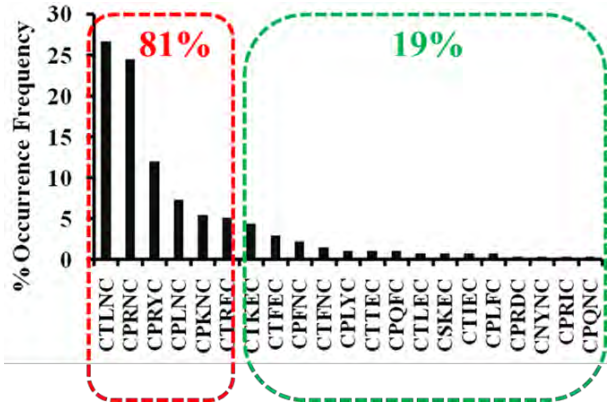


Fig. 4.1: Occurrence frequency of 23 RSL variants in Pin II PIs population. RSL marked with red block constitute 81%, while remaining 17 RSL contribute to 19% of total population.

Occurrence frequency analysis of 23 unique RSL in 389 Pin II PIs showed that 6 RSL with sequence CTLNC, CPRNC, CPRYC, CPLNC, CPKNC and CTREC were found to be predominant with 81% of total population. These analyses showed that RSL region in Pin II PIs is most prone to natural variations and thus engaged in generation of diversity of PIs against various target proteases. Occurrence analysis indicated that nature promotes specific sequence to propagate in population might be due to their superior functional attributes.

IV.1.3.2 *In vitro* assay indicates inhibition of trypsin activity by RSL tripeptides

Six tripeptides with sequence TLN, PRN, PRY, PLN, PKN and TRE were synthesized and purified on RP-HPLC. They were characterized by MALDI-TOF mass spectrometric analysis. The synthesized tripeptides along with their MALDI-TOF mass characterization data are listed in **Table 4.1**.

Table 4.1: MALDI-TOF mass characterization data of the synthesized tripeptides of the study

Sr. No.	Peptide sequence	MALDI-TOF mass (Da or m/z)	
		Calculated	Observed
1.	*Ac*-TRE-NH_2	445.23	446.89
2.	*Ac*-TLN-NH_2	387.21	388.12
3.	*Ac*-PRY-NH_2	475.25	476.66
4.	*Ac*-PKN-NH_2	398.23	400.55
5.	*Ac*-PRN-NH_2	426.23	426.99
6.	*Ac*-PLN-NH_2	383.21	384.32

Inhibition kinetic studies displayed a sigmoidal pattern with increasing concentrations of the tripeptides suggesting reversible and competitive inhibition with tight binding. TRE turned out to be a potent inhibitor of bovine trypsin (IC50 ~24.05 μM) compared to other 5 tripeptides from this study **(Fig. 4.2)**. The inhibition constant Ki determined directly from IC_{50} by using the Cheng-Prusoff's equation. Ki values for all 6 tripeptides were enlisted in **Table 4.2**. Differences in inhibition kinetics of tripeptides exhibited that amino acid variation of RSL tripeptide account for its differential protease binding and inhibition efficiency.

These results suggest that inhibitory property of RSL of Pin II PIs is sequence and conformation dependent.

Table 4.2: Ki values for synthesized RSL tripeptides

RSL Tripeptide	~Ki (in µM)
TRE	24.05
PRN	126.66
PRY	203.12
PLN	326.96
PKN	327.32
TRE	60.9

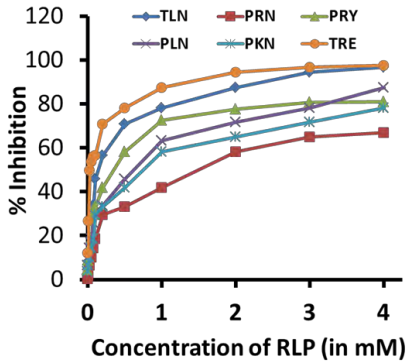


Fig. 4.2: Estimation of IC_{50} and Ki values by using inhibition of bovine trypsin with various concentrations of RSL tripeptide and with substrate BA*p*NA (1mM).

IV.1.3.3 TRE exhibit higher binding affinity and broader specificity against various insect proteases

Docking and relative analysis displayed significant differences in binding energies suggesting that synthesized tripeptides had variable interaction with *H. armigera* proteases. Among the 6 tripeptides, TRE and PRY showed strong interaction with the lowest binding energy with various *H. armigera* proteases supporting our *in vitro* results **(Fig. 4.3)**. Docking studies revealed broad specificity of TRE with *H. armigera* serine proteases. Strong binding of TRE with trypsin and chymotrypsin among all the tripeptides motivated us to access it's *in vivo* effect on *H. armiger* digestive physiology. Interaction pattern of TRE and PRY with most of the proteases was similar, which led to their clustering for all the analyzed proteases. Binding energy comparison and hierarchical clustering analysis provides wide overview of specific interaction of inhibitor with various proteases. Furthermore, this analysis might give us global overview about effect tripeptide ingestion on *H. armigera* digestive physiology.

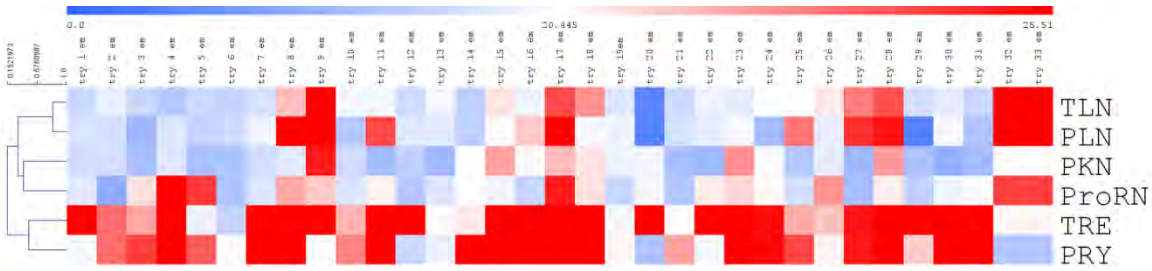


Fig 4.3: Interaction energy of binding of all three IRDs with *H. armigera* proteases is compared by using Heatmap analysis. Heat-map with hierarchical clustering of relative free binding energy (obtained from docking study) normalized by mean values of cumulative free energy obtained from *H. armigera* serine proteases binding with 6 RSL tripeptides. The gradient ruler is an indicator of interaction strength, where blue colour indicates weak binding and red indicate strong binding.

IV.1.3.4 TRE exerts negative impact on the growth and development of *H. armigera* larvae

To understand the *in vivo* effect of TRE on the development of *H. armigera* larvae, feeding experiments were conducted with appropriate controls. Various concentrations (50 to 200 ppm, 25 to 100 µM) of TRE were incorporated into diet to examine their *in vivo* potential against *H. armigera*. Development of larvae reared on a control and TRE containing diets is presented in **Fig. 4.4A**. Feeding of insects on TRE-containing diet caused concentration dependent reduction in larval mass gain and survival rate. On day 11, larvae fed on diets TRE containing 50, 100 and 200 ppm weighed ~10, 15 and 25 % less, than the larvae fed on AD.

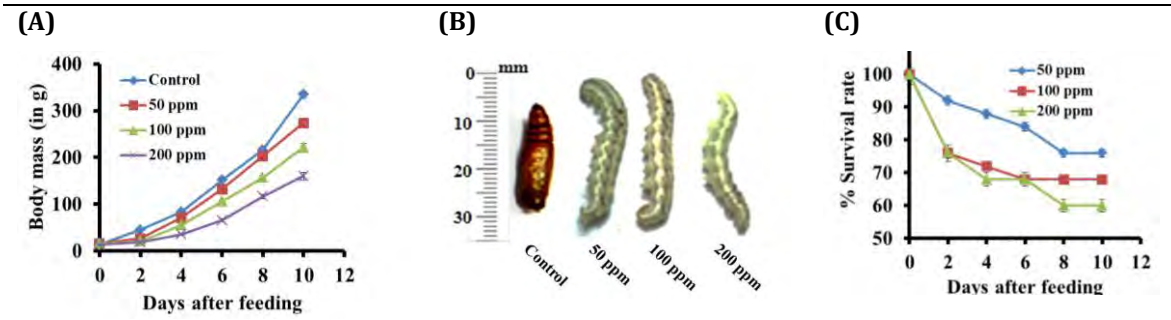


Fig 4.4: Growth **(A)** average size of larvae recorded on day 10 **(B)** and **(C)** Survival rate of *H. armigera* larvae raised on AD containing 50 to 200 ppm of TRE. Eggs were hatched, and neonates were transferred to AD containing of TRE.

In comparison, larvae fed on AD containing TRE were ~50 to 60% smaller in body size than control larvae **(Fig. 4.4B)**. Furthermore, larvae fed on inhibitor containing diet showed significant ($p \leq 0.05$) reduction in survival rate. At day 11, there was 30 to 40% reduction in survival rate of larvae fed on 100 and 200 ppm TRE containing AD as compared to control larvae **(Fig. 4.4C)**. TRE fed larvae displays early and sharp decrease in larval survival rates, followed by partial recovery as the feeding period extends. This is might be due to expression of PIs insensitive proteases and overexpression of proteases, which might help insect to overcome the lethal and detrimental effect of inhibitors (Dunse *et al.*, 2010; de Oliveira *et al.*, 2013).

Evaluation of nutritional parameters like Efficiency of Conversion of Ingested Food (ECI), Efficiency of Conversion of Digested Food (ECD) and Approximate Digestibility (Ad) revealed that the ingestion of TRE had deleterious effect on growth and rudimentary metabolism of the insect **(Table 4.3)**. There was direct correlation in the inhibitory potential and reduction in ECI, ECD and Ad. Assessment of these parameters showed that TRE negatively affected the digestive physiology of insect and thus impedes insect growth and survival **(Table 4.3)**. Inhibition of serine protease activities also obstructed normal developmental pathways leading to delay in pupation and molting, which was also evident from data **(Fig. 4.4B)**. Our results indicated that TRE could serve as potent inhibitor molecules against gut proteases from *H. armigera*.

Table 4.3: Effect of IRDs ingestion on the *H. armigera* feeding behavior and dietary utilization.

	50 ppm	**100 ppm**	**200 ppm**
ECI	12.31 (±1.5)	8.35 (±1.12)	6.86(±1.82)
ECD	29.09 (±2.31)	18.89(±1.87)	12.42(±2.3)
Ad	56.15 (±1.05)	46.55(±1.36)	40.05(±1.5)

Abbreviations: ECI = Efficiency of Conversion of Ingested Food; ECD = Efficiency of Conversion of Digested Food; Ad = Approximate Digestibility.

IV. 1.3.5 *In vivo* inhibition of gut proteinases in *H. armigera* larvae reared on TRE containing diet

In comparison with control HGP activity, HGP of larvae fed on TRE showed concentration dependent reduction. Larvae fed on 50 and 100 ppm TRE showed ~50 to 70% reduction of total protease activity. Larvae fed on 200 ppm of TRE showed significant inhibition (~80%) of HGP activity as compared to control larvae **(Fig. 4.5).** In case of trypsin and chymotrypsin activity of HGP from larvae fed on TRE showed ~60 to 80% and 50 to 70% reduction, respectively. Protease activity provides quantitative account about *in vivo* proteases activity upon TRE ingestion.

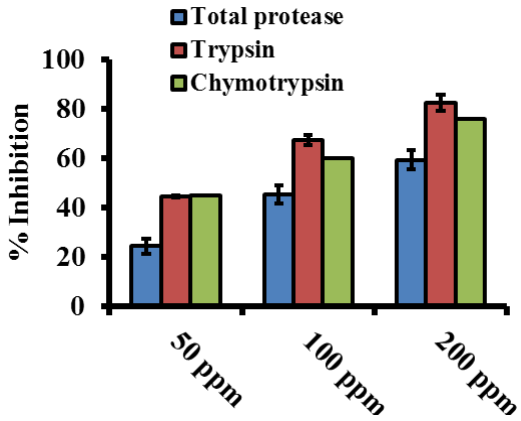


Fig. 4.5: *In vivo* inhibition of total proteolytic, trypsin and chymotrypsin activity of larvae fed on 50 to 200 ppm TRE was assessed by azocasein, BA*p*NA and SA*p*NA assay.

IV.1.3.6 Ingestion of TRE triggers the differential gene expression of the gut proteases

Expression analysis of eight major gut trypsin genes further provides the quantitative evidence of altered digestive physiology of *H. armigera* **(Fig. 4.6A)**. Real time PCR analysis showed that trypsins namely HaTry4, 5 and 6 were found to be downregulated in TRE fed larvae. Downregulation of these trypsins might also correlate with drop in proteolytic activity of numerous HGP isoforms when countered with TRE. HaTry1 and 8were highly upregulated in TRE fed larvae. HaTry2, 3 and 7 were exclusively upregulated in 200 ppm TRE fed larvae. Overexpression of protease transcripts could help the insect to overcome the

inhibitory effect of PIs, which might in turn help in growth and development of insects.

Among four major chymotrypsins, three transcripts HaChy1, 2 and 3 showed down regulation in TRE-fed larvae **(Fig. 4.6B)**. While HaChy4 showed distinct behavior with overexpression in 50 and 100 ppm TRE fed larvae. Expression dynamics of gut protease genes in an attempt to overcome the effect of inhibition and to acquire optimum nutritional requirements for growth and development provides an excellent survival tactics for the insects.

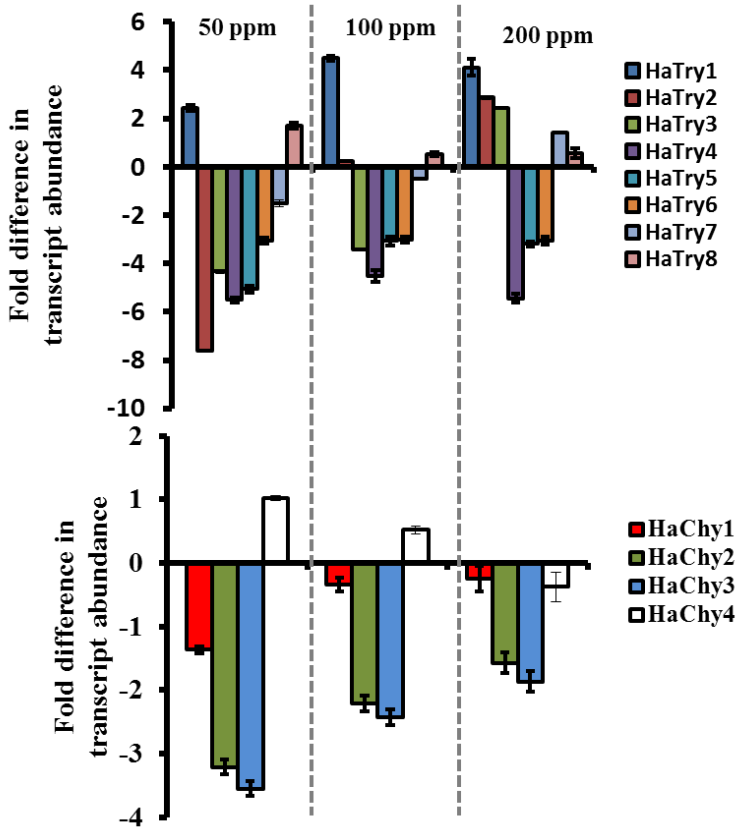


Fig. 4.6: Fold difference in transcript abundance of major digestive trypsins and chymotrypsins for larvae fed on AD containing 50 to 200 ppm of TRE on the 10th day after feeding. (A) Fold difference in transcript abundance of trypsin genes (B) Fold difference in transcript abundance of chymotrypsin genes. Most of trypsin and chymotrypsin gene showed down regulation; while few of them showed upregulation in TRE fed insects tissues.

IV.1.4 CONCLUSION

Exposure of RSL tripeptide 'TRE' to neonatal stage of larvae exhibited negative impact on *H. armigera* growth and development. This investigation provided insight

in to the potential of RSL amino acid sequence variations on efficacy and specificity against target proteases. Exposure of larvae to TRE offered variable molecular response and resulted in alteration in digestive physiology, particularly in protease expression. Current study suggested small peptides with 3 aa i.e. TRE proved to be effective growth inhibitors to *H. armigera* larvae. This phenomenon of antibiosis by RSL peptides can be explored for designing effective inhibitor candidate against *H. armigera*.

IV.2 Way towards "Dietary pesticides"- Mechanistic insight into insecticidal action of natural phenols

Insect infestation on various crops leads to reduce food productivity. Existing pest management strategies have raised serious environmental issues thereby resulting into the emergence of pesticide resistance in insects. Present investigation provides mechanistic insights into insecticidal activity of a natural phenol, caffeic acid (CA). *In silico* and subsequent *in vitro* screening of several natural phenols indicated CA as an effective inhibitor of *H. armigera* gut serine proteases. Furthermore, CA was found to be responsible fo*r in vivo* inhibition of the gut protease activity that led to reduced growth, survival and development of the insects. Upon CA ingestion dynamic alternation in protease expression and activity were apparent in *H. armigera* larvae. In addition, structure-activity relationship of CA highlighted the significance of all functional groups for its potency against target proteases. Various biophysical evidences suggested that binding of CA caused conformational changes in the target enzymes and thus decrease enzyme activity. Furthermore, molecular dynamic simulations and isothermal titration calorimetry results revealed that binding of first CA molecule at active site of trypsin provides a nucleation center for sequential binding of multiple CA and thus disrupted the function. In addition, along with the inhibition of digestive activity, CA showed significant reduction in detoxifying enzyme activities upon ingestion, intensifying the detrimental effect on overall insect physiology. In conclusion, our findings suggested that the natural phenols especially CA could be implicated as highly potent insecticide.

IV.2.1 INTRODUCTION

Modern agricultural practices are dependent on the extensive use of chemical pesticides which leads to resistance and resurgence in insects and also a reason for substantial environmental and human health problems (Cheng, 1990; Abdollahi *et al.*, 2004). This scenario highlights a need of natural and eco-friendly pesticides. Plants own diverse pool of secondary metabolites possessing insecticidal activity could provide lead molecules for developing ecofriendly insect pest management strategies. Recently, remarkable interest in potential of plants species with insecticidal activity, initiates application of these molecules in biopesticides development. Monomeric and polymeric phenolic compounds constitute one of the most widespread groups in plant secondary metabolites and they are generated biogenetically from the shikimate-phenylpropanoids-flavonoids pathways (Lattanzio *et al.*, 2008). Plants produce phenolic compounds for pigmentation, growth, reproduction and to cope with multiple stresses, etc. Pivotal role of natural phenols in plant-environment interactions especially in plants defense is evidenced from their differential accumulation in response to abiotic stresses (Borejan *et al.*, 2003; Asakawa *et al.*, 2013; Bi *et al.*, 1995; Horvath *et al.*, 2007; Ingersoll *et al.*, 2010; Mehmood *et al.*, 2013). Furthermore, toxic or deterrent activity of phenols against insect pests facilitates the direct defense; while in case of indirect defense they attract natural enemies of insect pests (Green *et al.*, 2003; Magalhaes *et al.*, 2010; Rani *et al.*, 2013; War *et al.*, 2012; War *et al.*, 2013). In phenolic acids, hydroxybenzoic and hydroxycinnamic acid (HCA) serve as basic scaffold for evolution of population of various compounds (Borejan *et al.*, 2003).

In this study we investigated the potential of natural phenols as insect protease inhibitor and molecule(s) as dietary pesticide for insect pest control in agriculture. Protease inhibition potential of selected molecules was evaluated by *in silico* study, *in vitro* inhibition kinetics and feeding bioassay. Various growth and development parameters were assessed in support of antibiosis of these natural phenols to *H. armigera*. Biophysical characterization and multi-scale molecular

dynamic simulations have been performed to elucidate binding mode of natural phenols to proteases and resultant structural changes in vital insect proteins.

IV.2.2 MATERIALS AND METHODS

IV.2.2.1 Virtual screening of natural phenols as protease inhibitor

Structures of various hydroxycinnamic acid derivatives were obtained from PubChem database and they were optimized for their 3D coordinates using Marvin Sketch Tool (http://www.chemaxon.com). Three dimensional structures of bovine trypsin (BtTry; PDB ID: 4I8G) and chymotrypsin (BtChy; PDB ID: 1YPH) were accessed from RCSB PDB. Also, *H. armigera* trypsin (HaTry) and chymotrypsin (HaChy) structures were modeled by using CPH 3.0 model server and processed using SwissPDB viewer(Kaplan *et al.*, 2001). Conversion of receptor and ligand from .pdb to .pdbqt format, also grid map and other docking parameters were set using AutoDock 4.2 software (Morris *et al.*, 2009). Virtual screening was carried out using AutoDock Vina software and the Lamarckian genetic algorithm as a searching procedure (Trott *et al.*, 2010). Binding energy obtained for each complex was represented in heat map format using MeV software packages (http://www.tm4.org/mev/). The gradient ruler is an indicator of interaction strength. Represented molecules showing strong binding against all the selected targets, represented by lower free energy were selected for further *in vitro* and simulation study.

IV.2.2.2 Inhibition kinetics study

Concentration dependent reduction in trypsin activity by CA was also estimated through a BA*p*NA assay using chromogenic substrate Benzoyl-L-arginyl p-nitroanilide (BA*p*NA). BA*p*NA assays were performed as described previously (**section II.3.4**).

IV.2.2.3 Feeding assay and assessment of nutritional parameters

Bioassays were conducted by feeding *H. armiger* larvae on *AD* containing 5 to 200 ppm/g of CA. Detailed method is as described previously (**section II.3.7 and II.3.8**).

The insect growth, development and survival of insect fed on control and CA containing diet were analyzed on each alternate day for 10 days.

IV.2.2.4 Semi-quantitative gene expression analysis

Total RNA was extracted from whole larval body using Trizol reagent (Invitrogen, CA, USA), followed by RQ1 DNase treatment (Promega, Fitchburg, Wisconsin, USA). CDNA was synthesized from 2µg of the DNA-free RNA samples by reverse transcription using oligo dT primers and reverse transcriptase (Promega) following the manufacturer's recommendations. Primers pairs were designed for Acetylcholinesterase, Amylase, Chymotrypsin, CYP450, GSTs and Trypsin genes of *H. armigera*. cDNA was diluted (1: 10) before use in a PCR. Semi-quantitative RT-PCR performed under the following conditions: initial denaturation at 95 °C for 2 min; 30 cycles at 9 5°C for 30 s; 60 °C for 30 s; 72 °C for 30 s and a final extension at 72 °C for 5 min (Kotkar *et al.*, 2012).

IV.2.2.5 Fluorescence quenching assay

Fluorometric experiments were carried out on a fluorescence spectrometer (Thermo Scientific, Waltham, Massachusetts, USA). Stock solutions of CA (1 mM) in ethanol and Bovine Trypsin (1 mM) in 100 mM Tris-HCl (pH 7.8) was prepared at room temperature. The final concentration of CA was from 0.2 to 30.0 µM with a constant bovine trypsin content of 1 µM. The fluorescence spectra were recorded at λ_{exc} = 280 nm and λ_{em} from 300 to 500 nm (Jiang *et al.*, 2004; Kang *et al.*, 2004; Bian *et al.*, 2007; Jin *et al.*, 2012).

IV.2.2.6 MALDI-TOF MS analysis

Different molar concentration of CA and trypsin were mixed and incubated for 20 min at 37 °C. The mass spectral analysis of the reaction mix was done on Q-TOF-MALDI-TOF-MS with a standard instrumental protocol (ABSCIX). Sample preparation, spectral acquisition and processing were done as described earlier (Mishra *et al.*, 2013).

IV.2.2.7 Isothermal titration calorimetry (ITC)

ITC was performed using Microcal Auto-iTC instrument (GE Healthcare, Buckinghamshire, UK). 19 injections of 2 µl bovine trypsin (Stock =2.0 mM) was

titrated against 0.1 mM solution of CA. Experiments were carried out at 37 ^{0}C in a Tris buffer, pH 7.8. Reference titration was carried out by injecting the same concentration of trypsin into buffer. Reference titration was subtracted from experimental titration. Origin 6.0 software was used to derive affinity constants (Kd), the molar reaction enthalpy (CH) and the stoichiometry of binding (N), by fitting the integrated titration peaks (Weber *et al.*, 2003).

IV.2.2.8 Circular dichorism spectroscopy

A Jasco J-810 spectro-polarimeter (Jasco, Easton, MD, USA) running the software Jasco J815CD was used to collect spectra at wavelengths 190-240 nm (far-UV) and 240-300 nm (near-UV). Quartz cuvettes of path lengths 5 mm was used for spectral scan. The scanning speed was 20nm/min, the response was 4 s, and the bandwidth was 1 nm. Trypsin (10 μM) was incubated with CA in 1:2, 1:1 and 2:1 ratio for 4 h at room temperature, similar process was done for chymotrypsin. Baseline corrections were performed by subtracting the spectra of the buffer and phenolic compounds from the sample spectra. The data are reported as mean residue ellipticity ([Θ]) in deg^{3}cm^{2} 3dmol^{1}. Details about spectra acquisition and analysis were described previously in **section III.1.2.5**.

IV.2.2.9. Molecular dynamic simulation

The structures with 27 CA molecules (aggregation study) and bovine trypsin-CA complex were relaxed using all-atom MD simulation in explicit water with GROMACS software package using the AMBER99SB force field and the simple point charge (SPC) water model (Ghosh *et al.*, 2012). The time step used in the simulation was 0.002 pswhile list of neighbors was updated every 0.01 ps with the grid method and a cut off radius of 14 Å. The coordinates of all atoms in the simulation were saved every 2 ps. The initial velocities were chosen randomly. NPT ensemble was used with the user-defined box dimension. The temperature and pressure were kept at the desired value using the Berendsen method and an isotropic coupling for the pressure (T=300 K, ζT=0.1 ps, P0=1 bar, coupling time = ζT=1 ps). The electrostatic term was calculated using the Particle Mesh Ewald (PME) algorithm with the radius of 16 Å with Fast Fourier Transform (FTT) optimization (with an order of 4 for the

cubic interpolation). The cut off algorithm was applied for the non-coulomb potential with a radius of 9 Å. Initial minimization was done for a total of 50,000 steps, followed by volume and pressure equilibration each for 1 ns. The production MD run for the complexes were then carried out.

Both of these systems were then simulated using the simulation parameters mentioned in the previous section. The parameters for CA were retrieved from the SWISSPARAM server (http://www.swissparam.ch/; Zoete *et al.*, 2011). Both of these simulations were performed and analyzed using the supercomputing facility provided by BRAF cluster using GROMACS (Ghosh *et al.*, 2012).

IV.2.2.10 Enzymatic assays

H. armigera larvae (500 mg) were ground to a fine powder and extracted in 500 mL of 200 mM Glycine-NaOH buffer (pH 10) for 2 h at 4˚C. The extract was then centrifuged at 12000 X g (4˚C; 10 min) and the clear supernatant was used as a source of enzyme for all solution assays. Control and CA fed insects protein lysate (50 µg) was used to carry out the following functional assays (a) total protease activity (b) Trypsin-like activity (c) Chymotrypsin like activity (d) Amylase activity (e) Lipase activity (f) Acetylcholinesterase activity (g) GSTs activity and (h) CYP450 activity (Kotkar *et al.*, 2012; Sarate *et al.*, 2012; Ellman *et al.*, 1961; Dwakar *et al.*, 2011). The detailed methodology is described in corresponding references.

IV.2.2.11. Statistical analysis

All the experiments were performed three times independently. Student's t-test was used for statistical analysis. Data were expressed as mean ± SD. A $p < 0.05$ was considered as statistically significant.

IV.2.3 RESULTS AND DISCUSSION

IV.2.3.1 Protease inhibition potential of hydroxycinnamic acid derivatives

Comparison of binding energies of various HCA derivatives against bovine and *H. armigera* proteases displayed variability in binding affinity **(Fig.4.7A)**. Caffeic, ferulic, sinapic and *p*-coumaric acid emerged as strong inhibitors of candidate

proteases **(Fig. 4.7B)**. It was observed that caffeic acid (CA) showed strong competitive inhibition of bovine (*Bos tarus*) trypsin (Ki ~22 µM) and bovine chymotrypsin (Ki~32 µM), while ferulic acid (FA) showed exclusive inhibition of bovine chymotrypsin with Ki ~32 µM. CA and FA showed significant inhibition of *H. armigera* trypsin and chymotrypsin activity **(Fig. 4.7B)**.

Molecular docking analysis and binding energy comparison indicated that natural phenols have affinity towards the insect proteases and the interaction amongst the proteases from different insects is conserved (Polyphagous: *H. armigera* and *Plutella xylostella*; Monophagous: *Manduca sexta*) **(Fig. 4.7C and D)**. Analysis of 130 docked complexes revealed that CA had strong interaction (low binding energy) with various insect serine proteases, especially with trypsin like proteases. Elevated concentration of CA and other similar natural phenols in plant tissues upon insect infestation might have direct role in inhibition of digestive function of the insects (Green *et al.*, 2003; Magalhaes *et al.*, 2010; Rani *et al.*, 2013; War *et al.*, 2012; War *et al.*, 2013).

IV.2.3.2 Molecular response of *H. armigera* to caffeic acid (CA) ingestion

Delayed pupation indicates the retardation of development in larvae fed on CA **(Fig.4.8A)**. CA appeared to have significant and concentration dependent negative impact on growth and survival of *H. armigera*. Larvae fed on diet containing 50 to 200 ppm of CA showed ~ 20 to 50% less body mass (g) as compared to control **(Fig. 4.8B)**.

Furthermore, CA fed leave had ~ 50 to 80% reductions in survival rate **(Fig. 4.8C)**. Evaluation of nutritional parameters like Efficiency of Conversion of Ingested Food (ECI), Efficiency of Conversion of Digested Food (ECD) and Approximate Digestibility (AD) revealed that the ingestion of CA had deleterious effect on growth and rudimentary metabolism of the insect (**Fig. 4.8D**). Protease activity of insect fed on CA exhibit substantial reduction in trypsin like (~ 20 to 60%) and thus total protease (~15 to 50%) activity **(Fig. 4.8E)**. Inhibition of digestive proteases by CA probably causes starvation in insects. It might led to nutritional scarcity, less pool of free amino acids and also energy required for metabolism and causes the growth

retardation followed by insect death (Green *et al.*, 2003; Magalhaes *et al.*, 2010; War *et al.*, 2012; Rani *et al.*, 2013; War *et al.*, 2013).

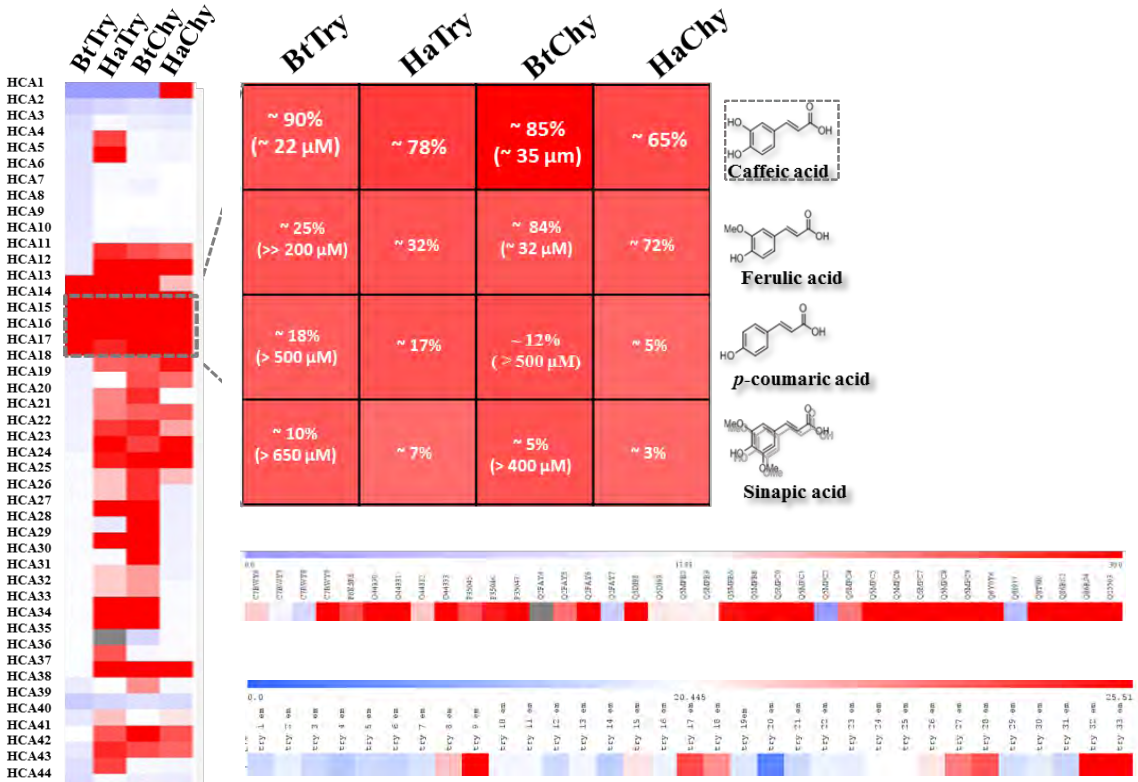


Fig. 4.7: (A) Interaction energy of binding of various HCA derivatives with bovine and *H. armigera* proteases is compared by using Heatmap analysis. Relative free binding energy (obtained from docking study) normalized by mean values of cumulative free energy was obtained from *Bos taurus* (Bovine) trypsins, bovine chymotrypsin, *H. armigera* trypsin and chymotrypsin binding with different derivatives. The gradient ruler is an indicator of interaction strength, where blue colour indicates weak binding and red indicates strong binding. Parenthesis values indicate Ki **(B)** Derivatives with strong binding affinities are marked with dash board. Interaction energy of binding of various HCA derivatives with **(C)** *M. sexta* and **(D)***P. xylostella* proteases is compared by using Heatmap analysis.

Candidate protease gene expression analysis indicated that ingestion of CA also had differential response. Several of the *H. armigera* trypsin's and chymotrypsin's (HaTry1 to 8; HaChy1 to 4) were down regulated except HaTry4

(Fig. 4.8F). Overexpression of HaTry4 and HaChy4 in CA fed insects might have been resulted in response to compensate for reduced the protease activity. This overexpression might be also in response to rescuing the insect from growth retardation by producing inhibitor resistant/insensitive proteases, which could hijack the insect digestive system and fulfill the insect growth requirement (Bown *et al.*, 1997; Broadway *et al.*, 1997; Chikate *et al.*, 2013).

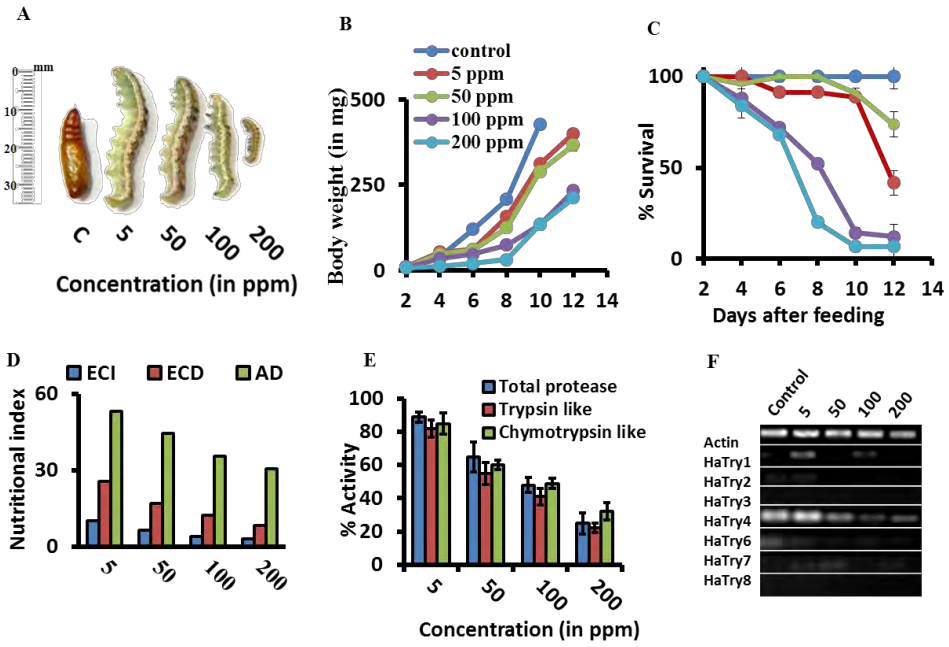


Fig. 4.8: *H. armigera* larvae were raised on AD containing 5 to 200 ppm/g of CA **(A)** Average size of larvae recorded on day 10 **(B)** Insect growth and **(C)** Survival rate normalized with control **(D)** Nutritional parameters ECI, ECD and Ad and (E) *In vivo* proteolytic activity of insect fed on control and CA containing diet were assessed. Bars represent means (±) S.E. from three independent experiments at *p 0.05 **(F)** Protease gene expression analysis at the transcript level by performing semi-quantitative RT-PCR using total RNA isolated from control and CA fed larvae. All samples were analyzed on 2% agarose gels containing gel red. This analysis showed differential expression of protease gene in CA fed larvae.

IV.2.3.3 Fate of CA in insect gut environment

Stability of CA in insect gut environment and probability of formation of subsequent metabolites were investigated by LC-ESI(-)-HRMS analysis **(Fig. 4.9)**. Chromatograms of gut metabolite extract of CAfed larvae in 1:1 acetonitrile and

methanol showed the presence of intact CA molecule and no additional peaks in comparison to control. Thus, it excluded the possibility of the formation of degraded products of CA in digestive track of insect. Stability of CA in insect gut led to maximum availability and hence caused to be constitutively active as protease inhibitor. *In vivo* stability of CA made it a lucrative molecule to be used as an insecticidal agent (Yang *et al.*, 2013).

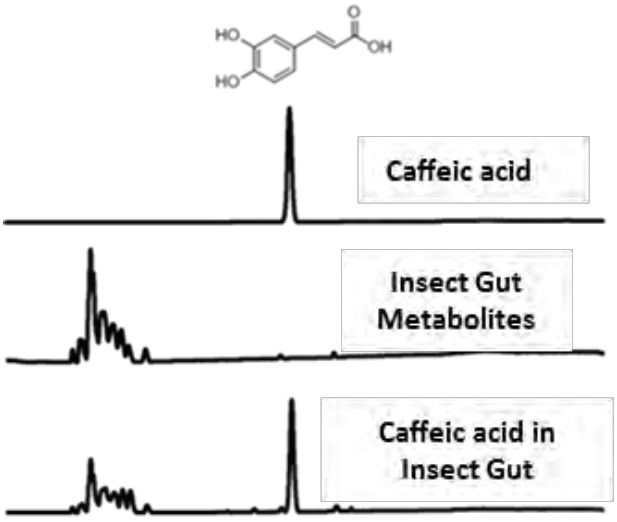


Fig. 4.9: HRMS analysis of insect gut metabolites indicated the fate of CA in insect gut environment.

IV.2.3.4 Structure-activity relationship of CA as protease inhibitor

Functional group-activity relationships provide a detailed insight into the essentiality of individual functional groups to maintain the inhibitory potency **(Table 4.4)**. Semi-synthesized derivatives were structurally characterized by NMR and mass spectrometric data which were in well agreement with the previous reports (Barontini *et al.*, 2014; Percec *et al.*, 2006; Roche *et al.*, 2005; Takahashi *et al.*, 2010; Uwai *et al.*, 2008; Zhu *et al.*, 2010). Masking of 3-hydroxyl group as methoxy on the phenolic ring (FA, **1b**) resulted in the considerable reduction in gut protease inhibition (40 to 60%). Even the addition of another methoxy group at 5-position (**1c**, sinapic acid) on the phenolic ring led to further reduction in inhibitory

potencies (20 to 30%). Removal of the 3-hydroxyl group (**1d**, cinnamic acid) from the CA structure resulted in drastic reduction in the inhibitory activity (15 to 25%). Masking of the terminal carboxylic acid as a methyl (**1e**) ester also caused a significant lowering in inhibitory activities (35 to 50%). Further increment in the chain length as an ethyl ester (**1f**) showed far more reduction in inhibitory potency (25 to 30%). Hydrogenation (reduction) of the double bond in the side chain of CA (**1g**) reduced the activity extensively (12 to 20%). Similar effect was observed when methyl ester of CA was subjected to hydrogenation (**1h**) (inhibition 25 to 50%). Protection of both the hydroxyl groups on the phenolic ring by acetylation (**1i** and **1j**) drastically attenuated its inhibitory potential (10 to 20%). In brief, modification in any of the functional groups (phenolic hydroxyl, carboxylic acid and double bond) on CA structure led to reduction in the inhibitory activity against proteases with variable extent and all of them were found to be essentially responsible for the activity. Therefore, the parent natural product (**1a**) found to be most potent among all the derivatives screened.

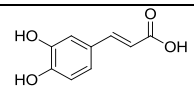
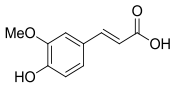
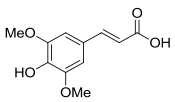
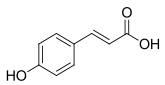
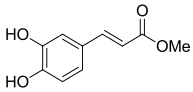
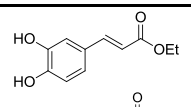
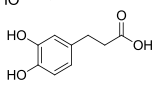
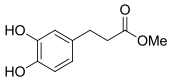
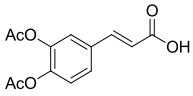
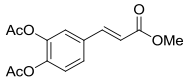
Compound		Inhibition		
		Trypsin (~ IC_{50} in µM)	**Chymotrypsin (~ IC_{50} in µM)**	**% Inhibition (~) of HGP activity (100 µM)**
1a		22	35	70-80
1b		120	32	40-60
1c		650	460	20-30
1d		780	525	15-25
1e		80	125	35-50

Table 4.4: Various derivatives of CA and their inhibition values against Trypsin, Chymotrypsin and HGP

1f		165	205	25-30
1g		225	320	12-20
1h		109	165	25-50
1i		380	520	10-20
1j		350	451	10-20

IV.2.3.5 Binding of CA causes structural and conformational change in target protein

Bovine trypsin showed reduction in tryptophan fluorescence in response to titration with CA suggesting the change of the molecular environment of tryptophan **(Fig. 4.10A)**. Red shift is observed in this case indicated the conformational change in protease on ligand binding. Furthermore, it was found that multiple CA molecules were interacting non-covalently with proteases. It was validated by mass spectrometric analysis, where shift in mass (*m/z*) was observed in comparison to native enzymes due to binding of multiple CA molecules to the proteases **(Fig. 4.10B)**. Circular Dichorism (CD) spectroscopy of proteins in near and far UV regions suggested the secondary and tertiary structure alteration in the protease respectively, on CA binding **(Fig. 4.10C and D)**. Fluorescence quenching and CD Spectroscopy analysis together indicated that binding of CA caused environmental alternation around tryptophan residue resulting in structural and conformational change in the protease, which might further be attributed to the reduced enzyme activity (Jiang *et al.*, 2004; Kang *et al.*, 2004;Bian *et al.*, 2007; Jin *et al.*, 2012).

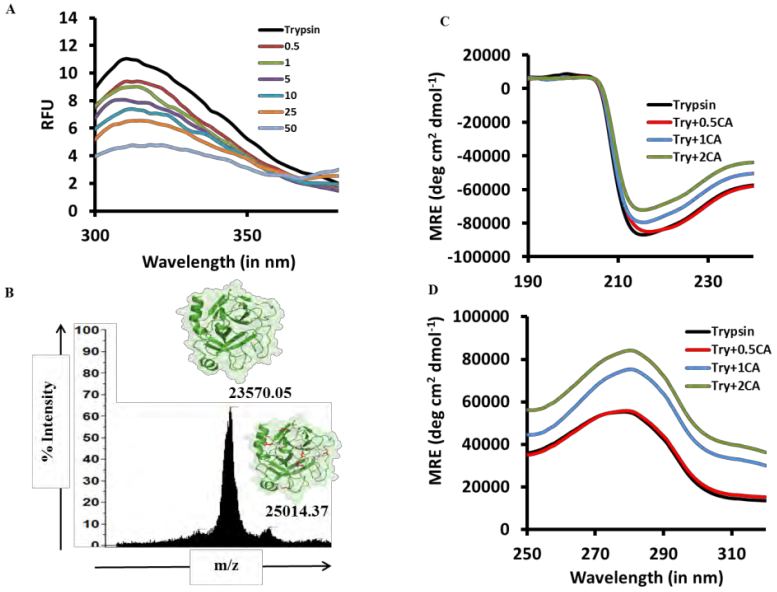


Fig. 4.10: Biophysical characterization of CA and protease interaction: (A) Fluorescence emission of bovine trypsin in presence of various concentration of CA (B) Binding of multiple CA to protease molecule was analyzed by MALDI TOF analysis. Structural change in protease on CA binding was examined at secondary level by (C) near UV CD spectra and at tertiary level by (D) far UV CD spectra.

IV.2.3.6 CA shows sequential binding to protease and exhibit heterogeneous mode of inhibition

In set of simulation system to study the impact of CA binding on the structure of protease, system containing *H. armigera* trypsin 4 (HaTry4, predominant trypsin from *H. armigera* gut) with multiple CA in vicinity was simulated for 1µS and trajectory was analyzed. The CA molecules in this simulation were initially restrained at discrete position on the protease molecule and the restrain was released till 10 ns of the simulation. Out of these 5 CA molecules, one was found to remain in active site for the rest of simulation, and interacting with Ser200 of HaTry4 **(Fig. 4.11A)**. Hydroxyl groups of CA can form multiple transient interactions with protease active site, however, H-bond between OH group of CA and Tyr45 and Ser200 was found to be stable throughout simulation time **(Fig. 4.11B)**. Further, after around 208 and 300 ns of the simulation, two more CA molecules were found to be located near mouth of active site **(Fig. 4.11C)**. These three CA clubbed near active site to form stacked triad *via* Pi-Pi interactions and

they maintains their position through another set of Pi-Pi interaction with lining residues of active site mouth i.e. Tyr45, Trp144 and Tyr154 **(Fig. 4.11D)**. In this way CA binding to *H. armigera* trypsin might cause heterogeneous type of inhibition by (i) competitively binding to the active site residue i.e. Ser200 (ii) blocking the mouth of the active site, leads to inaccessibility to the substrate. This suggests that CA might be having some crucial interactions with the active site residues and longer time-scale simulation might be required to explore major structural changes in the protein.

Thermodynamic studies of trypsin-CA interaction was carried out using ITC as it is one of the most widely used quantitative technique for direct measurement of the enthalpy change when two species interact, allowing the determination of heat of association, stoichiometry, and binding affinity from a single experiment. The raw data and corresponding to the thermogram of the binding experiment is depicted in **Fig. 4.11E**. Binding was strongly exothermic and showed 1:3 stoichiometry for trypsin and CA. The spontaneity of the process was evidenced by a negative change in the enthalpy, CH, and a positive change in the entropy, CS. Binding of first CA molecule showed lower ΔG as compared to the second and third molecule binding. This thermodynamic pattern of binding indicates that binding of first CA molecule was strong as compared to the second and third CA molecule.

To investigate the aggregation propensity of CA in solution and its effect on protein binding, 27 CA molecules (aggregation study) were relaxed using all-atom MD simulation in explicit water with GROMACS software package using the AMBER99SB force field and the simple point charge (SPC) water model (Ghosh *et al.*, 2012). The 50 ns trajectory of this system was analyzed to see how individual CA molecules display the phenomenon of aggregation. At around 1 to 2 ns of simulation, formation of aggregates was initiated **(Fig. 4.11F)**. These aggregates were observed in different clusters with varying number of CA molecules in individual clusters. Initially after 2 ns of the simulation, two major clusters were observed having 12 CA molecules each and remaining 3 CA in none of them.

Likewise a big cluster of 21 CA molecules was observed after 6.5 ns and just after 30 ns of the simulation, a cluster of 27 CA molecules was also observed.

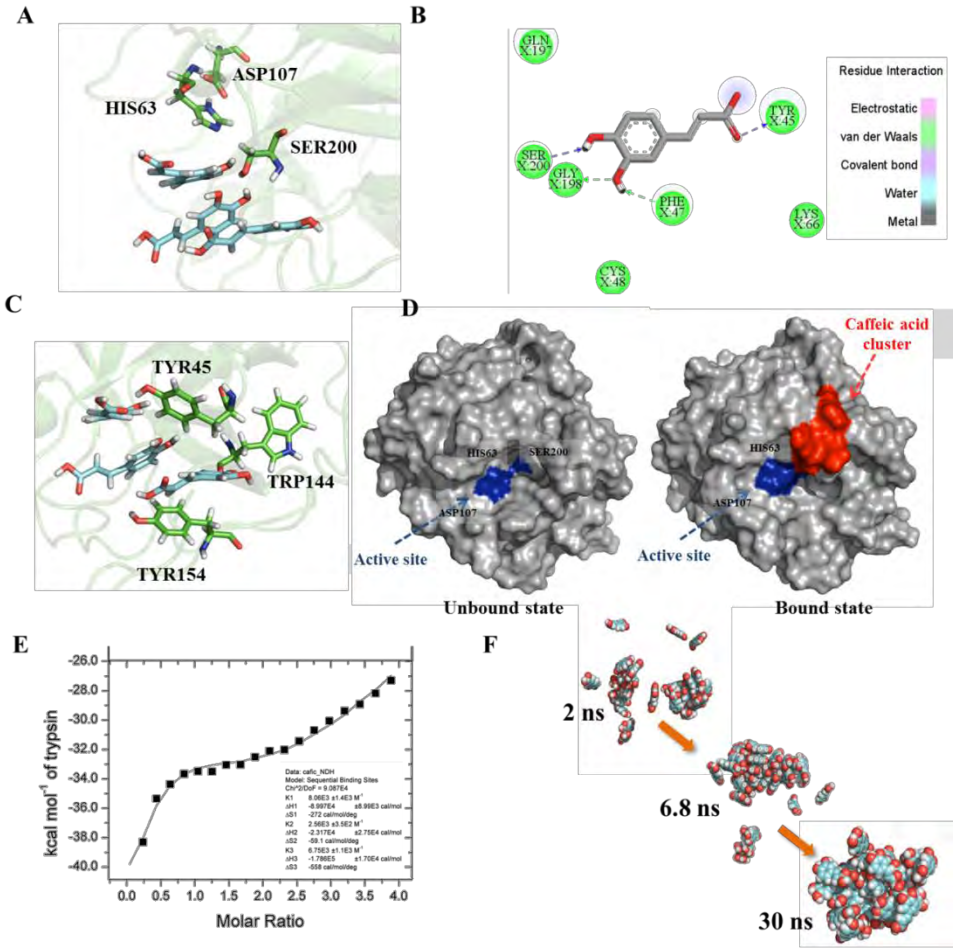


Fig.4.11: Simulation of bovine trypsin and CA in one system illustrates (A) binding of single CA molecule at the active site residue SER200. (B) 2D interaction map showed establishment of multiple interactions of CA with trypsin binding pocket. Strong binding of CA at active site leads to (C) formation of nucleation center for aggregation of other two CA molecules in the active site. (D) Aggregate formation in binding pocket further blocks complete access to the active site, which is showed in surface model. (E) Simulation of only CA in solvation box showed that CA has intrinsic propensity to form aggregate and this might enhance its chances to interact with the proteases.

Some of these aggregated CA molecules also formed stacked orientation with one another during the simulation. This suggested that the CA had an intrinsic

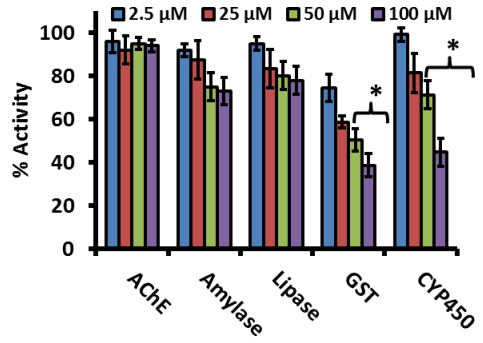
property of displaying stacking interactions as well as self-aggregation. Formation of aggregates might enhance the probability of CA to interact with proteases. Kinetic, Simulation and ITC data demonstrated the binding of single molecule of CA to the active site of trypsin initiates the sequential binding of CA molecules at the active site.

IV.2.3.7 CA inhibits the insect detoxification machinery

Estimation of different enzyme activities and semi-quantitative gene expression analysis indicated that other than protease inhibition activity, CA also had adverse effects on various enzymes primarily involved in detoxification mechanism. It was observed that activity of Glutathion S-tranferase (GST) and Cytochrome P450 (CYP450) reduced drastically in insects fed with CA containing diet**(Fig. 4.12A)**, and insect attempted to compensate this activity reduction by overexpression of these genes at transcript level **(Fig. 4.12B)**.

Reduction in these enzyme activities was concentration dependent and almost 60% GST and CYP450 enzyme activity was reduced at 200 ppm concentration. Other than these enzymes, various digestive enzymes activities were found to be maintained in presence of CA even at concentration of 200 ppm. Inhibition of detoxification enzymes by natural phenols might intensify the detrimental effect on insect growth and development.

(A)



(B)

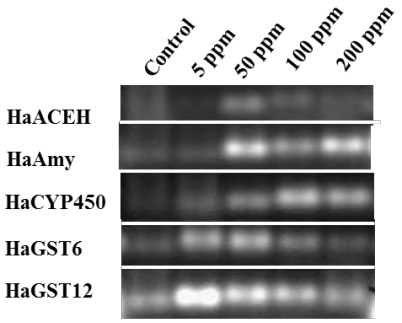


Fig. 4.12: The enzymatic activities of (A) AChE, Amylase, lipase, GST and CYP450 were analyzed for control and CA treated insects using protocols described in corresponding references. In CA fed insects, significant and concentration dependent reduction in GST and CYP450 activity was observed. Bars represent means (±) S.E. from three independent experiments at *p 0.05 (B) Gene expression analysis of AChE, Amylase, lipase, GST and CYP450 genes at the transcript level by performing semi-quantitative RT-PCR using total RNA isolated from control and CA fed larvae.

IV.2.4 CONCLUSION

Natural phenol, CA has been found out and validated as a potential insecticidal molecule. All the functional groups contributed to its protease inhibitor property and it also remain stable in insect gut environment. Ingestion of CA causes differential molecular responses at transcriptional and translational level. Structural investigation showed that CA showed sequential binding at the active site of major insect protease, thus inhibits its activity. Although overexpression of these proteases (e.g. HaTry4) were observed in CA fed larvae, but strong inhibition potential of CA overcome the effect of enhanced expression of insect proteases. All these finding suggests that exploration of natural phenols could be effective approach to develop "Dietary pesticide" against *H. armiger* infestation.

Chapter V: General discussion and Future prospects

V. General discussion and future directions

Plant-insect interaction is a dynamic system which is a subject to continual variation and change. They have co-evolved from past millions of years and developed distinctive strategies to overcome each other's defense or resistance mechanism. The study of this interaction has become very interesting owing to the new discoveries of plant defense mechanisms and also the pest adaptations or aggressions. Plants have evolved with, physical barriers, secondary metabolites and defensive proteins to resist insect herbivores. On the other hand, insects have also developed several strategies to overcome plant defense making them capable of feeding, growing and reproducing on their host plants. Among different defense strategies, causing "indigestion" in insects remains the plants best defense for resisting insects through an array of protease inhibitors as defensive proteins thereby depriving insects of nutrients and affecting their growth and development. Several proteinaceous, small molecules and peptide PIs have been studied and characterized for their antibiosis effects on insect pests. Pros and cons of these interactions have been discussed in detail in the first chapter. We have analysed the effect of non-host PIs on insect growth and survival and found that PIs from *C. annuum* (*Solanaceae*) showed adverse effect on *H. armigera* (Lepidoptera) growth and development. Further, several diverse PI genes (CanPIs) were identified from various *C. annuum* tissues and the PI proteins showed differential inhibition of gut proteases of *H. armigera*. Feeding bioassays with PI proteins had shown not only reduction in larval and pupal weights but also dramatic reduction in fertility and fecundity in a dose dependent manner. Moreover, significantly high expression levels of CanPIs under naturally infested conditions had indicated the involvement of CanPIs in plant defense. On this background information we have designed our objectives as follows (i) assessment of distinct IRDs for their activity against *H. armigera* proteases *in vitro* and *in vivo*; (ii) interaction studies of single and multidomain CanPIs with target proteases; (iii) recombinant expression and biochemical characterization of *H. armigera* proteases; (iv) evaluation of synthetic peptides and small molecules for their potential against *H. armigera* proteases, *in vitro and in vivo*. Altogether, these findings suggest that the structural information

obtained from interaction insect proteases with (i) different IRDs, (ii) small molecules and (iii) synthetic peptides could be used for developing tailor-made inhibitors for their effective application in insect pest control in agriculture.

V.1.1Functional differentiation of CanPIs IRDs governed by sequence variation

From the previous reports, a total of 55 unique IRDs were identified from *C. annuum* contributing to the diverse inhibitory pool (Tamhane *et al.*, 2009; Mishra *et al.*, 2013). *C. annuum* expresses diverse Pin-II PIs comprising IRD as basic functional unit. Most IRDs contain eight conserved cysteine forming four disulfide bonds, which are indispensible for their stability and activity. We investigated the functional significance of evolutionary variations in IRDs and their role in mediating interactions between the inhibitor and cognate protease. Among the 55 IRDs encoded by *C. annuum*, IRD-7, -9, and -12 were selected for further characterization on the basis of variation in their reactive site loop, number of conserved cysteine residues, and higher theoretical ΔG_{bind} for interaction with *H. armigera* trypsins. Individual recombinant IRDs displayed variation in their inhibitory potential against proteases like, trypsin, chymotrypsin or HGP. IRDs showed quantitative as well as qualitative variations with respect to the interaction with proteases. Moreover, inhibition kinetics showed that IRD-9, despite loss of two disulfide bonds, was more potent among the three selected IRDs. Furthermore, it showed special attributes like stability to proteolysis and synergistic inhibitory effect with other IRDs. The results suggested that selective loss of disulfide bond in IRD-9 leads to its functional superiority/differentiation. Post simulation analysis of the intramolecular hydrogen bonds illustrated that IRD-9 with two disulfide bonds (C7-C25 and C8-C37) less, has a relatively higher density of intra-molecular hydrogen bonds as compared to IRD-7 and -12. These intramolecular hydrogen bonds might be substituting the two lost disulfide bonds of IRD-9 to stabilize the protein structure in the active conformation and also might be protecting the molecules from a hydrophobic collapse (Hansen *et al.*, 2007). The replaced serine residues in the place of two cysteines C7 and C8 in IRD-9 may be contributing to the increased number of hydrogen bonds. This might

be a positive natural selection and led to functional differentiation of the inhibitor (Li *et al.*, 2011).

Analysis of interaction between RSL of IRDs and active site of proteases showed that the higher conformational flexibility of IRD-9 by the loss of two disulfide bonds has helped it to spatially adapt a better complementary shape suited to the active site of HaTry's compared to the more rigid four disulfide containing IRD-7 and IRD-12. Activity assay against various proteases indicated that sequence variation in the IRDs might affect its inhibitory efficiency and target specificity. Although the selected IRDs were primarily trypsin inhibitors, they exhibited activity against chymotrypsin, elastase and cathepsin. Binding energy comparison and hierarchical clustering analysis provides wide overview of specific interaction of inhibitor with various proteases. Furthermore, *H. armigera* fed on recombinant IRD proteins exhibited negative impact on larval growth, survival rate and other nutritional parameters. Major digestive gut trypsin and chymotrypsin genes were down regulated in the rIRD fed larvae, while few of them were up-regulated, this indicate alterations in insect digestive physiology. Inhibition of serine protease activities also obstructs normal developmental pathways leading to delay in pupation and molting, which was also evident from data. PIs-fed larvae displays early and sharp decrease in larval survival rates, in case of IRD-7 and -12, it is followed by partial recovery as the feeding period extends. This is might be due to expression of PIs insensitive proteases and overexpression of proteases, which might help insect to overcome the lethal and detrimental effect of inhibitors (Dunse *et al.*, 2010a, b; de Oliveira *et al.*, 2013). The expression of PI-insensitive proteases assists in obtaining optimal nutritional requirements by utilizing diet protein sources as well as using protease inhibitors as source of amino acids under PI stress (Jongsma *et al.*, 1995). Expression dynamics of gut protease genes in an attempt to overcome the effect of inhibition and to acquire optimum nutritional requirements for growth and development provides an excellent survival benefit to *H. armigera*. In general, differential translational and transcriptional response of insect suggests that this performance variability of IRDs depends on sequence variations. Altogether, these findings suggest that the sequence variations among IRDs reflect

in their efficacy against proteases in vitro and in vivo, which also could be used for developing tailor-made multi-domain inhibitor gene(s).

V. 1. 2 Solvent exposed and spatially distributed reactive site loops are prerequisite for strong interaction with proteases

Structural analysis and molecular dynamic simulation of protease-PIs interaction led to illustration of variation in spatial interaction of CanPIs with target protease.Structures of single domain (IRD-7, -9,-12) and multidomain (CanPI-7) inhibitorwere predicted by homology modeling and validated by NMR and SAX analysis. Single domain and multidomain inhibitor shared structural feature like predominant disorder loops with RSL on it and confined with disulfide bonds. RSL mostly contain polar amino acids which contribute in interaction with active site of proteases. Due to polar nature of reactive site loops they are solvent exposed and engaged in multiple hydrogen bond formation with surrounding water molecules. In case of multidomain inhibitor, individual domains are well dispersed to maximize the interaction of multiple proteases. Spatial distribution of RSL avoids the steric clashes in the interacting molecules. Although, in case of CanPI-7 there are 4 RSL present, but due to spatial distribution it can interact with only 3 target proteases.

V. 1. 3 *H. armigera* protease shares structural-functional similarities with alkaline proteases

H. armigera have serine proteases as major digestive enzymes. A trypsin like protease from *H. armigera* was expressed, purified and characterized for its biochemical properties (HaTry4). The purified protease was found to be active at alkaline pH (>8). The enzyme also showed good stability in broad pH range i.e. from pH 7 to 12. It has been well known that a large number of proteases present in the insect gut act in high alkaline pH range. Our findings are in accordance with the properties of proteases reported from previous studies; cockroach, *Spilosoma obliqua* (pH 11.0), *S. lit*ura (pH 9.0, 10.5, and 11.0), *Heliothis zea* (pH 11.0), *Galleria mellonella* (pH 10.5 and 11.2) and *Tenebrio molitor* (pH 8.5). Most of the insects have midgut pH in the range of 6 to 10. Gut pH conditions are likely to have a major

influence on the efficiency of nutrient extraction in insects. The high pH of many insect guts has been attributed to an adaptation of their leaf-eating ancestors for extracting hemicelluloses from plant cell walls. Furthermore, there are some correlations between the midgut enzymes and surrounding symbiotic microflora. Optimum temperature for activity of HaTry4 was determined to evaluate its physiological relevance. The purified alkaline protease exhibited a temperature optimum in range of40 to 50°C and stability in the temperature range 50 to 70°C. The alkaline proteases from other insects also found to be thermostable and their temperature optima were between 50 to 60°C.

Studies on enzyme inhibitors provided further insights into biochemical properties of the *H. armigera* protease (HaTry4). The purified enzyme was characterized as a serine protease because its activity was completely inhibited by a specific serine protease inhibitor PMSF. PMSF binds specifically to a serine residue in the active site of serine protease. The chelating agents DTNB, EDTA and β-mercaptoethanol had no influence on the activity of purified enzyme. The purified protease from *H. armigera* showed considerable activity at high concentration of various metal ions. CD spectroscopy and molecular modeling depicted that HaTry4 shares structural similarity with bovine as well as porcine trypsin. HaTry4 showed presence of two juxtaposed β-barrels joined with disorder loop. Active site residues (His63, Asp121 and Ser218) are present on this loop. Initial crystallization of HaTry4 using standard screen showed crystallization in 0.1 M sinapic acid and 0.1M HEPES pH 7.5. We are in process to obtain the diffraction data and solve the structure. We believe that the structure of *H. armigera* trypsin will provide excellent template for structure based inhibitor designing for effective pest management and crop protection.

V. 1. 4 RSL tripeptides can serve as next generation insect controlling molecule

Reactive site loop (RSL) is part of IRD that interacts with target proteases and it is found to be highly variable. Co-evolution of RSL with their target protease indicates its crucial role in plant-insect interaction (Jongsma & Beekwilder, 2011). Reactive site region of 387 Pin-II PIs was used to design tri-peptides for its potential

application as insect protease inhibitor. Occurrence frequency analysis of 23 unique RSL in 389 Pin II PIs showed that 6 RSL with sequence CTLNC, CPRNC, CPRYC, CPLNC, CPKNC and CTREC were found to be predominant with 81% of total population. These analyses showed that RSL region in Pin II PIs is most prone to natural variations and thus engaged in generation of diversity of PIs against various target proteases. Occurrence analysis indicated that nature promotes specific sequence to propagate in population probably due to their superior functional attributes. The inhibition constant Ki determined directly from IC_{50} by using the Cheng-Prusoff's equation. Ki values for all 6 tripeptides were found to variable. Differences in inhibition kinetics of tripeptides exhibited that amino acid variation of RSL tripeptides account for its differential protease binding and inhibition efficiency. These results suggest that inhibitory property of RSL of Pin II PIs is sequence and conformation dependent.Most of the selected peptides showed significant protease activity inhibition; while TRE was most potent protease inhibitor with Ki ~24 µM. Docking studies revealed broad specificity of TRE with *H. armigera*serine proteases. Strongbinding of TRE with trypsin and chymotrypsin among all thetripeptides motivated us to access it's*in vivo* effect on *H. armigera*digestive physiology. Feeding of insects on TRE-containing diet caused concentration dependent reduction in larval mass gain and survival rate. On day 11, larvae fed on diets TRE containing 50, 100 and 200 ppm weighed ~10, 15 and 25% less, respectively than the larvae fed on control diet. Assessment of nutritional parameters showed that TRE negatively affected the digestive physiology of insect and thus impedes insect growth and development. Inhibition of serine protease activities also obstructs normal developmental pathways leading to delay in pupation and molting, which was also evident from data. Our results indicate that TRE could serve as potent inhibitor molecules against gut proteases of *H. armigera*.

V. 1. 5 Way towards "Dietary Pesticides"

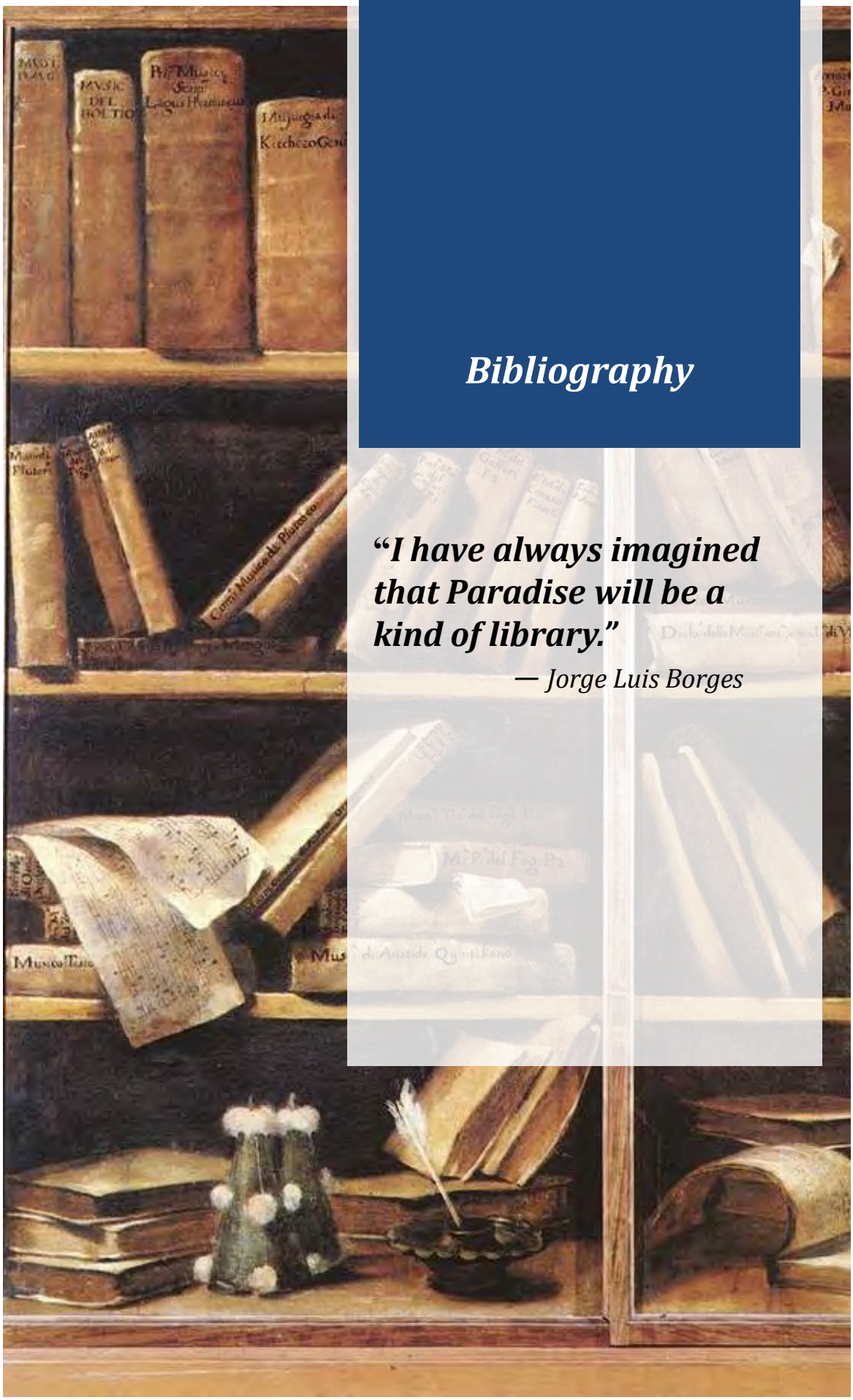
Natural phenol, CA was identified and validated as a potential insecticidal molecule. Comparison of binding energies of various HCA derivatives against bovine and *H. armigera* proteases displayed variability in binding affinity. Molecular docking

analysis and binding energy comparison indicated that natural phenols have affinity towards the insect proteases and the interaction amongst the proteases from different insects is conserved (Polyphagous: *H. armigera* and *P. xylostella*; Monophagous: *M. sexta*). It was observed that CA showed strong competitive inhibition of bovine trypsin (Ki ~22 μM) and bovine chymotrypsin (Ki~32μM). Inhibition of digestive proteases by CA probably causes starvation in insects. It might led to nutritional scarcity, less pool of free amino acids and also energy required for metabolism and causes the growth retardation followed by insect death (Green *et al.*, 2003; Magalhaes *et al.*, 2010; War *et al.*, 2012; Rani *et al.*, 2013; War *et al.*, 2013). Overexpression of digestive protease genes might be also in response to rescuing the insect from growth retardation by producing inhibitor resistant/insensitive proteases, which could hijack the insect digestive system and fulfill the insect growth requirement (Bown *et al.*, 1997; Broadway *et al.*, 1997; Chikate *et al.*, 2013). Chemical modification in any of the functional groups (phenolic hydroxyl, carboxylic acid and double bond) on CA structure led to reduction in the inhibitory activity against proteases with variable extent and all of them were found to be essentially responsible for the activity. Therefore, the parent natural product i.e. CA found to be most potent among all the derivatives. Stability of CA in insect gut led to maximum availability and hence caused to be constitutively active as protease inhibitor. Fluorescence quenching and CD spectroscopy analysis together indicated that binding of CA caused environmental alternation around tryptophan residue resulting in structural and conformational change in the protease, which might further be attributed to the reduced enzyme activity. Inhibition kinetic, simulation and ITC data of CA binding to proteases demonstrated the binding of single molecule of CA to the active site of trypsin initiates the sequential binding of CA molecules at the active site, thus inhibits its activity. Inhibition of protease activity is compensated by overexpression of these proteases (e.g. HaTry4) in CA fed larvae. Furthermore, CA also leads to inhibition of detoxification enzymes, which might intensify the detrimental effect of CA on insect growth and development. All these finding suggests that exploration of natural phenols could be effective approach to develop “Dietary pesticide” against *H. armigera* infestation.

In brief this study comprises of computational and biochemical analysis of insect gut proteases *a*nd their interactions with group of proteins/peptides and small molecules. This information has provided the basis for selecting the candidates for further investigations. We believe that the generation of structural insights of insect gutproteases and protease-inhibitor complexes using X-ray crystallography and NMR techniques could provide basis for developing new generation of insect growth retardant molecules and toformulate effective and sustainable pest management strategy in agriculture.

Future directions

1) Structural information of interaction of different range of protease inhibitor i.e. small molecules/peptides (Caffeic acid, TRE) and natural peptides (IRDs, CanPIs) to target insect proteases (HaTry4 and other) using X-ray crystallography and NMR spectroscopy
2) Lead inhibitor molecules optimization by means by QSAR or protein engineering to increase its potential against the target proteases
3) Efficacy of newly identified/designed inhibitor molecules in green house/fields for assessment of its insect controlling potential



Bibliography

"I have always imagined that Paradise will be a kind of library."

— Jorge Luis Borges

BIBLIOGRAPHY

Abdeen A, Virgos A, Olivella E, Villanueva J, Avile´s X, Gabarra R, *et al.* (2005) Multiple insect resistance in transgenic tomato plants over-expressing two families of plant proteinase inhibitors. Plant Molecular Biology 57: 189–202.

Abdollahi M, Ranjbar A, Shadnia S, Nikfar S, Rezaiee A (2004) Pesticides and oxidative stress: a review. Medicinal Science Monitor 10: 141-147.

Aitken A, Learmonth M (1996) Estimation of Disulfide Bonds Using Ellman's Reagent. The Protein Protocols Handbook V: 487-488.

Ali JG, Agrawal AA (2012) Specialist versus generalist insect herbivores and plant defense. Trends in Plant Sciences 17: 293-302.

Altieri MA, Letourneau DK, Stephen J (1984) Vegetation diversity and insect pest outbreaks. Critical Reviews in Plant Sciences 2: 131-169.

Ames BN, Margie P, Gold LS (1990) Dietary pesticides (99.99% all natural). Proceedings of the National Academy of Sciences, USA 87: 7777-7781.

Angelucci C, Barrett-Wilt GA, Hunt DF, Akhurst RJ, East PD, Gordon KHJ, *et al.* (2008) Diversity of aminopeptidases, derived from four lepidopteran gene duplications, and polycalins expressed in the midgut of Helicoverpa armigera: Identification of proteins binding the delta-endotoxin, Cry1Ac of *Bacillus thuringiensis*. Journal of Insect Physiology 38: 685-696.

Antcheva N, Pintar A, Patthy AS, Simoncsits A (2001) Proteins of circularly permuted sequence present within the same organism: the major serine proteinase inhibitor from *Capsicum annuum* seeds. Protein Science 10: 2280-2290.

Antor RJ, García MB (1995) A new mite-plant association: mites living amidst the adhesive traps of a carnivorous plant. Oecologia 101: 51-54.

Appel HM (1993) Phenolics in ecological interactions: the importance of oxidation. Journal of Chemical Ecology 19: 1521-1552.

Asakawa Y, Ludwiczuk A, Nagashima F (2013) Phytochemical and biological studies of bryophytes. Phytochemistry 91: 52-80.

Atkinson AH, Heath RL, Simpson RJ, Clarke AE, Anderson MA (1993) Proteinase inhibitors in *Nicotiana alata* stigmas are derived from a precursor protein which is processed into five homologous inhibitors. Plant Cell 5: 203-213.

Azad MS, Hossan MA (2004) Agricultural biotechnology for green revolution–perceived expectations and potential risks. Journal of Food Technology 2: 8-17.

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.

Baldwin IT, Halitschke R, Kessler A, Schittko U (2001) Merging molecular and ecological approaches in plant–insect interactions. Current Opinion in Plant Biology 4: 351–358.

Baldwin IT, Halitschke R, Paschold A, Von dahl CC, Preston CA (2006) Volatile signaling in plant-plant interactions: "Talking Trees" in the genomics era. Science 311: 812-815.

Barontini M, Bernini R, Carastro I, Gentili P, Romani A (2014) Synthesis and DPPH radical scavenging activity of novel compounds obtained from tyrosol and cinnamic acid derivatives. New Journal of Chemistry 38: 809-816.

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

Barrette-Ng IH, Ng KK, Cherney MM, Pearce G, Ghani U, Ryan CA, *et al.* (2003b) Unbound form of tomato inhibitor-II reveals interdomain flexibility and conformational variability in the reactive site loops. Journal of Biological Chemistry 278: 31391-31400.

Barrette-Ng IH, Ng KK, Cherney MM, Pearce G, Ryan CA, James MN (2003a) Structural basis of inhibition revealed by a 1:2 complex of the two-headed tomato inhibitor-II and subtilisin Carlsberg. Journal of Biological Chemistry 278: 24062-24071.

Barta E, Pintar A, Pongor S (2002) Repeats with variations: accelerated evolution of the Pin2 family of proteinase inhibitors. Trends in Genetics 18: 600-603.

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.

Bateman KS, James MNG (2011) Plant protein proteinase inhibitors: Structure and mechanism of inhibition. Current Protein and Peptide Science 12: 341-347.

Bax AD, Grzesiek S (1993) Methodological advances in protein NMR. Accounts of Chemical Research 26: 131-138.

Bax AD, Ikura. M (1991) An efficient 3D NMR technique for correlating the proton and ^{15}N backbone amide resonances with the α-carbon of the preceding residue in uniformly $^{15}N/^{13}C$ enriched proteins. Journal of Biomolecular NMR 1: 99-104.

Bayes A, Sonnenschein A, Daura X, Vendrell J, Aviles FX (2003) Procarboxypeptidase A from the insect pest Helicoverpa armigera and its derived enzyme. European Journal of Biochemistry 270: 3026–3035.

Becerra JX, Noge K, Venable DL (2009) Macroevolutionary chemical escalation in an ancient plant-herbivore arms race. Proceedings of National Academy of Science, USA 106: 18062–18066.

Beekwilder J, Schipper B, Bakker P, Bosch D, Jongsma M (2000) Characterization of potato proteinase inhibitor II reactive site mutants. European Journal of Biochemistry 267: 1975-1984.

Bennett RN, Wallsgrove RM (1994) Secondary metabolites in plant defence mechanisms. New Phytologist 127: 617-633.

Bergey DR, Orozco-Cardenas M, de Moura DS, Ryan CA (1999) A wound-and systemin-inducible polygalacturonase in tomato leaves. Proceedings of National Academy of Sciences, USA 96: 1756–1760.

Bezzi S, Kessler D, Diezel C, Muck A, Anssour S, Baldwin IT (2010) Silencing NaTPI expression increases nectar germin, nectarins, and hydrogen peroxide levels and inhibits nectar removal from plants in nature. Plant Physiology 152: 2232-2242.

Bi J, Felton G (1995) Foliar oxidative stress and insect herbivory: primary compounds, secondary metabolites, and reactive oxygen species as components of induced resistance. Journal of Chemical Ecology 21: 1511-1530.

Bian H, Zhang H, Yu Q, Chen Z, Liang H (2007) Studies on the interaction of cinnamic acid with bovine serum albumin. Chemical and Pharmaceutical Bulletin 55: 871-875.

Bode W, Huber R (1992) Natural protein proteinase inhibitors and their interaction with proteinases. European Journal of Biochemistry 204: 433-451.

Boerjan W, Ralph J, Baucher M (2003) Lignin biosynthesis. Annual Review of Plant Biology 54: 519-546.

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.

Bravo A, Likitvivatanavong S, Gill SS, Soberón M (2011) *Bacillus thuringiensis*: A story of a successful bioinsecticide. Insect Biochemistry and Molecular Biology 41: 423-431.

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, Duffey SS (1986a) 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.

Broadway RM and Duffey SS (1986b) The effect of dietary protein on the growth and digestive physiology of larval *Heliothis zea* and *Spodoptera exigua*. Journal of Insect Physiology 32: 673-680.

Bronsoms S, Pantoja-Uceda D, Gabrijelcic-Geiger D, Sanglas L, Aviles FX, Santoro J, *et al.* (2011) Oxidative folding and structural analyses of a kunitz-related inhibitor and its disulfide intermediates, functional implications. Journal of Molecular Biology 414: 427-441.

Bronstein JL (1994) Our current understanding of mutualism. The Quarterly Review of Biology 69:31–51.

Bronstein JL, Alarcon R, Geber M (2006) The evolution of plant-insect mutualisms. New Phytologist 172: 412-428.

Brown TA, Jones MK, Powell W, Allaby RG (2009) The complex origins of domesticated crops in the Fertile Crescent. Trends in Ecology and Evolution 24, 103-109.

Bruessow F, Gouhier-darimont C, Buchala A, Metraux J, Reymond P (2010) Insect eggs suppress plant defence against chewing herbivores. The Plant Journal 62: 876–885.

Bryant J, Green TR, Gurusaddaiah T, Ryan CA (1976) Proteinase inhibitor II from potatoes: isolation and characterization of its promoter components. Biochemistry 15: 3418–3424.

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

Charity JA, Anderson MA, Bittisnich DJ, Whitecross M, Higgins TJV (1999) Transgenic tobacco and peas expressing a proteinase inhibitor from Nicotiana alata have increased insect resistance. Molecular Breeding 5: 357–365.

Chen CC, Hwang JK, Yang JM (2006) (PS)2: protein structure prediction server. Nucleic Acids Research 34: 152-157.

Chen H, Wilkerson CG, Kuchar JA, Phinney BS, Howe GA (2005) Jasmonate-inducible plant enzymes degrade essential amino acids in the herbivore midgut. Proceedings of National Academy of Sciences, USA 102:19237-19242.

Chen R, Li L, Weng Z (2003) ZDOCK Predictions for the CAPRI Challenge. Proteins 52: 80-87.

Chen VB, Arendall WB3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, *et al.* (2010) MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallography D Biological Crystallography 66:12-21.

Cheng YC, Prusoff WH (1973) Relationship between the inhibition constant (Ki) and the concentration of inhibitor which causes 50 per cent inhibition (I50) of an enzymatic reaction. Biochemical Pharmacology 22: 3099-3108.

Chikate YR, Tamhane VA, Joshi RS, Gupta VS, Giri AP (2013) Differential protease activity augments polyphagy in *Helicoverpa armigera*. Insect Molecular Biology 22: 258-272.

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.

Chougule NP, Hivrale VK, Chhabda PJ, Giri AP, Kachole MS (2003) Differential inhibition of *Helicoverpa armigera* gut proteinases by proteinase inhibitors of pigeonpea (*Cajanus cajan*) and its wild relatives. Phytochemistry 64: 681-687.

Christeller JT (2005) Evolutionary mechanisms acting on proteinase inhibitor variability. FEBS Journal 272: 5710-5722.

Christeller JT, Laing WA, Markwick NP, Burgess EPJ (1992) Midgut protease activities in 12 phytophagous lepidopteran larvae: Dietary and protease inhibitor interactions. Insect Biochemistry and Molecular Biology 22: 735-746.

Christou P, Twyman RM (2004) The potential of genetically enhanced plants to address food insecurity. Nutrition Research Reviews 17: 23-42.

Chye ML, Sin SF, Xu ZF, Yeung EC (2006) Serine proteinase inhibitor proteins: Exogenous and endogenous functions. In Vitro Cellular and Developmental Biology-Plant 42: 100-108.

Cleal CJ, Thomas BA (2009) An Introduction to Plant Fossils. Cambridge University Press, Cambridge.

Copeland RA, Lombardo D, Giannaras J, Decicco CP (1995) Estimating Ki values for tight binding inhibitors from dose response plot. Bioorganic and Medicinal Chemistry Letters 5: 1947-1952.

Costa JR, Yaliraki SN (2006) Role of rigidity on the activity of proteinase inhibitors and their peptide mimics. The Journal of Physical Chemistry 110: 18981-18988.

Craik DJ, Mylne JS, Daly NL (2010) Cyclotides: macrocyclic peptides with applications in drug design and agriculture. Cellular and molecular life sciences 67: 9-16.

Crane PR, Friis EM, Pedersen KR (1995) The origin and early diversification of angiosperms. Nature 374: 27-33.

Crepet WL (1979) Insect pollination: a paleontological perspective. BioScience 29: 102–108.

Crepet WL (2008) The fossil record of angiosperms: requiem or renaissance? Annals of the Missouri Botanical Garden 95: 3-33.

Currano ED, Wilf P, Wing SL, Labandeira CC, Lovelock EC, Royer DL (2008) Sharply increased insect herbivory during the paleocene-eocene thermal maximum. Proceedings of National Academy of Science, USA 105: 1960-1964.

Czapinska H, Otlewski J (1999) Structural and energetic determinants of the S1-site specificity in serine proteases. European Journal of Biochemistry 260: 671-695.

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.

Dawkar VV, Chikate YR, Lomate PR, Dholakia BB, Gupta VS, Giri AP (2013) Molecular insights into resistance mechanisms of lepidopteran insect pests against toxicants. Journal of Proteome Research 12: 4727-4737.

Dawkar VV, Chikate YR, Gupta VS, Slade SE, Giri AP (2011) Assimilatory potential of *Helicoverpa armigera* reared on host (Chickpea) and nonhost (*Cassia tora*) diets. Journal of Proteome Research 10: 5128-5138.

De Leo 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.

de Oliveira CF, de Paula Souza T, Parra JR, Marangoni S, Silva-FilhoMde C, Macedo ML (2013) Insensitive trypsins are differentially transcribed during Spodoptera frugiperda adaptation against plant protease inhibitors. Comparative Biochemistry and Physiology - Part B 165: 19-25.

Diezel C, Kessler D, Baldwin IT (2011) Pithy protection: *Nicotiana attenuata's* jasmonic acid-mediated defenses are required to resist stem-boring weevil larvae. Plant Physiology 155: 1936-1946.

Dobler S, Dalla S, Wagschal V, Agrawal AA (2012) Community-wide convergent evolution in insect adaptation to toxic cardenolides by substitutions in the Na, K-ATPase. Proceedings of National Academy of Science, USA 109: 13040-13045.

Down RE, Gatehouse MR, Hamilton W, Gatehouse JA (1996) Snowdrop lectin inhibits development and decreases fecundity of the Glasshouse Potato Aphid *Aulacorthum solani* when administered *in vitro* and via transgenic plants both in laboratory and glasshouse trials. Journal of insect physiology 42: 1035-1045.

Duan X, Li X, Xue Q, Abo el Saad M, Xu D, Wu R (1996) Transgenic rice plants harbouring an introduced potato proteinase inhibitor II gene are insect resistant. Nature Biotechnology 14: 494-498.

Duan JJ, Lundgren JG, Naranjo S, Marvier M (2010) Extrapolating non-target risk of Bt crops from laboratory to field. Biology Letters 6: 74-77.

Dunse K, Anderson M (2011) Towards the next generation of pest resistant plants. ISB News Report, June 2011.

Dunse KM, Kaas Q, Guarino RF, Barton PA, Craik DJ, Anderson MA (2010b) Molecular basis for the resistance of an insect chymotrypsin to a potato type II proteinase inhibitor. Proceedings of the National Academy of Sciences, USA 107: 15016-15021.

Dunse KM, Stevens JA, Lay FT, Gaspar YM, Heath RL, Anderson MA (2010a) Co-expression of potato type I and II proteinase inhibitors gives cotton plants protection against insect damage in the field. Proceedings of the National Academy of Sciences, USA 107: 15011-15015.

Eddy JL, Derr JE, Hass GM (1980) Chymotrypsin inhibitor from potatoes: interaction with target enzymes. Phytochemistry 19: 757-761.

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

Ellman GL, Courtney D, Andres V, Featherston RM (1961) A new and rapid colorimetric determination of acetylcholinesterase activity. Biochemical Pharmacology 7: 88-95.

Eran P, Gang DR (2000) Genetics and biochemistry of secondary metabolites in plants: an evolutionary perspective. Trends in Plant Science 5: 439-445.

Eschenfeldt WH, Stols L, Millard CS, Joachimiak A, Donnelly MI (2009) A family of LIC vectors for high-throughput cloning and purification of proteins. Methods Molecular Biology 498: 105–115.

Farrar RR, Barbour JD, Kennedy GG (1989) Quantifying food consumption and growth in insects. Annals of the Entomology Society of America 82: 593–598.

Farrell B, Mitter C (1993) Phylogenetic determinants of insect/plant community diversity. Historical and geographical determinants of community diversity (Chicago University Press): 253-266.

Fass D (2012) Disulfide bonding in protein biophysics. Annual Review of Biophysics 41: 43–59.

Fefelova Y, Frolov A (2008) Distribution and mortality of corn earworm (*Helicoverpa armigera*, Lepidoptera, Noctuidae) on maize plants in Krasnodar territory. Entomological Review 88: 480-484.

Felton GW, Korth KL, Wesley SV, Huhman DV, Mathews MC, Murphy JB, *et al.* (1999) Inverse relationship between systemic resistance of plants to microorganisms and to insect herbivory. Current Biology 9: 317–320.

Fernandes GW (1994) Plant mechanical defenses against insect herbivory. Revista Brasileira de Entomologia 38: 421–433.

Ferry N, Edwards MG, Gatehouse J, Capell T, Christou P, Gatehouse AMR (2006) Transgenic plants for insect pest control: a forward looking scientific perspective. Transgenic Research 15:13–19.

Feyereisen R (1999). Insect P450 enzymes. Annual Review of Entomology 44: 507-533.

Fitt GP (1989) The ecology of Heliothis in relation to agroecosystems. Annual Review of Entomology 34: 17-52.

Fraenkel GS (1959) The raison d'être of secondary plant substances. Science 129: 1466–1470.

Frampton GK (1999) Spatial variation in non-target effects of the insecticides chlorpyrifos, cypermethrin and pirimicarb on Collembola in winter wheat. Pesticide Science 55: 875-886.

Freeman BC, Beattie GA (2008) An overview of plant defenses against pathogens and herbivores. The Plant Health Instructor. DOI: 10.1094/PHI-I-2008-0226-01.

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, Pauchet Y, Vogel H, Heckel DG (2010) An ABC transporter mutation is correlated with insect resistance to *Bacillus thuringiensis* Cry1Ac toxin. PLoS Genetics 6, e1001248.

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 AMR, Ferry N, Raemaekers RJM (2002) The case of the monarch butterfly: a verdict is returned. Trends in Genetics 18: 249–251.

Gatehouse AMR, Norton E, Davison GM, Babbe´ SM, Newell CA, Gatehouse JA (1999) Digestive proteolytic activity in larvae of tomato moth, *Lacanobia oleracea*, effects of plant protease inhibitors *in vitro* and *in vivo*. Journal of Insect Physiology 45: 545-558.

Gatehouse JA (2002) Plant resistance towards insect herbivores: a dynamic interaction. 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.

Gavrilescu M (2005) Fate of pesticides in the environment and its bioremediation. Engineering in Life Sciences 5: 497-526.

Gepts PA (2002) Comparison between crop domestication, classical plant breeding and genetic engineering. Crop Science 42: 1780-1790.

Ghosh A, Sonavane U, Andhirka S, Aradhyam G, Joshi R (2012) Structural insights into human GPCR protein OA1, a computational perspective. Journal of Molecular Modelling 18: 2117-33.

Giladi I (2006) Choosing benefits or partners: a review of the evidence for the evolution of myrmecochory. Oikos 112: 481-492.

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, *et al.* (2003) Identification of potent inhibitors of Helicoverpa armigera gut proteinases from winged bean seeds. Phytochemistry 63: 523–532.

Glenn KF, Hardy MC (2013) Spider-venom peptides: structure, pharmacology, and potential for control of insect pests. Annual review of entomology 58: 475-496.

Gowda CL, Sharma HC (2005) Legume pod borer/cotton bollworm, Heliothis/Helicoverpa - the global problem. In souvenir national symposium on *Helicoverpa* management, National challenge, Kanpur, India.

Graham JS, Pearce G, Merryweather J, Titani K, Ericsson LH, Ryan CA (1985) Wound-induced proteinase inhibitors from tomato leaves. II. The cDNA-deduced primary structure of pre-inhibitor II. Journal of Biological Chemistry 260: 6561-6564.

Green PW, Stevenson PC, Simmonds MS, Sharma HC (2003) Phenolic compounds on the pod-surface of pigeonpea, *Cajanus cajan*, mediate feeding behavior of *Helicoverpa armigera* larvae. Journal of Chemical Ecology 29: 811-821.

Green TR and Ryan CA (1972) Wound induced proteinase inhibitor in plant leaves: A possible defence mechanism. Science 175: 776-777.

Greenblatt HM, Ryan CA, James MNG (1989) Structure of the complex of *Streptomyces griseus* proteinase B and polypeptide chymotrypsin inhibitor-1 from Russet Burbank potato tubers at 2.1 Å resolution. Journal of Molecular Biology 205: 201-228.

Guex N, Peitsch MC (1997) SWISS-MODEL and the Swiss-Pdb Viewer: An environment for comparative protein modeling. Electrophoresis 18: 2714-2723.

Gunning RV, Moores GD, Devonshire AL (1998) Insensitive acetycholinesterase causes resistance to organophosphates in Australian *Helicoverpa armigera* (Hübner) (Lepidoptera: Noctuidae). Pesticide Science 54: 319-320.

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.

Hames BD, Hooper NM (2000) Instant notes in Biochemistry. Bios Scientific Publishers Ltd.; 2nd revised edition.

Hansen D, Macedo-Ribeiro S, Veríssimo P, Yoo IS, Sampaio MU, Oliva MLV (2007) Crystal structure of a novel cysteinless plant Kunitz-type protease inhibitor. Biochemical and Biophysical Research Communications 360: 735-740.

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.

Hardy MC, Daly NL, Mobli M, Morales RAV, King GF (2013) Isolation of an orally active insecticidal toxin from the venom of an *Australian tarantula*. PloS one 9: e73136.

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

Harter AV, Gardner KA, Falush D, Lentz DL, Bye RA, Rieseberg LH (2004) Origin of extant domesticated sunflowers in eastern North America. Nature 430: 201–205.

Hartl M, Giri AP, Kaur H, Baldwin IT (2010) Serine protease inhibitors specifically defend *Solanum nigrum* against generalist herbivores but do not influence plant growth and development. Plant Cell 22: 4158-4175.

Hartl M, Giri AP, Kaur H, Baldwin IT (2011) The multiple functions of plant serine protease inhibitors: Defense against herbivores and beyond. Plant Signalling and Behaviour 6: 1-3.

Haruta M, Major IT, Christopher ME, Patton JJ, Constabel CP (2001) A Kunitz trypsin inhibitor gene family from trembling aspen (*Populus tremuloides* Michx.): cloning, functional expression, and induction by wounding and herbivory. Plant Molecular Biology 46: 347–359.

Hause B, Hause G, Kutter C, Miersch O, Wasternack C (2003) Enzymes of jasmonate biosynthesis occur in tomato sieve elements. Plant Cell Physiology 44: 643-648.

Heath RL, Barton PA, Simpson RJ, Reid GE, Lim G, Anderson MA (1995) Characterization of the protease processing sites in a multidomain proteinase inhibitor precursor from *Nicotiana alata*. European Journal of Biochemistry 230: 250-257.

Heath RL, McDonald G, Christeller JT, Lee M, Bateman K, West J, *et al.* (1997) Proteinase inhibitors from *Nicotiana alata* enhance plant resistance to insect pests. Journal of Insect Physiology 43: 833-842.

Heckel DG (2012) Insecticide resistance after silent spring. Science 337: 1612–1614.

Hegland SJ, Nielsen A, Lazaro A, Bjerknes A, Totland O (2009) How does climate warming affect plant-pollinator interactions? Ecology Letter 12: 184-195.

Hegyi H, Gerstein M (2002) The relationship between protein structure and function: a comprehensive survey with application to the yeast genome. Journal of Molecular Biology 288: 147-164.

Heil M, Fiala B, Linsenmair KE, Zotz G, Menke P, Maschwitz U (1997). Food body production in Macaranga triloba (*Euphorbi aceae*): A plant investment in anti-herbivore defence via symbiotic ant partners. Journal of Ecology 85: 847–861.

Heimhofer U, Hochuli PA, Burla S, Dinis JML, Weissert H (2005) Timing of Early Cretaceous angiosperm diversification and possible links to major paleoenvironmental change. Geology 33: 141–144.

Heimhofer U, Hochuli PA, Burla S, Weissert H (2007) New records of Early Cretaceous angiosperm pollen from Portuguese coastal deposits: Implications for the timing of the early angiosperm radiation. Review of Palaeobotany and Palynology 44: 39–76.

Hering EM (1951). Biology of the leaf miners. Springer Netherlands

Hermsmeier D, Schittko U, Baldwin IT (2001) Molecular interactions between the specialist herbivore *Manduca sexta* (Lepidoptera, Sphingidae) and its natural host *Nicotiana attenuata*. I. Large-scale changes in the accumulation of growth-and defense-related plant mRNAs. Plant Physiology 125: 683–700.

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

Hilder VA, Gatehouse AMR, Boulter D (1993) Transgenic plants conferring insect tolerance. Proteinase inhibitor approach. In: Transgenic plants vol. 1. Engineering and utilization. (Eds. SD Kung and R Wu). Academic Press, Inc. California, 317-338.

Hilder VA, Gatehouse AMR, Sherman SE, Barker RF, Boulter D (1987) A novel mechanism for insect resistance. Nature 330: 160–163.

Hivrale VK, Lomate PR, Basaiyye SS, Kalve ND (2013) Compensatory proteolytic responses to dietary proteinase inhibitors from *Albizia lebbeck* seeds in the *Helicoverpa armigera* larvae. Arthropod-Plant Interaction 7: 259-266.

Höfte H, Whiteley HR (1989) Insecticidal crystal proteins of *Bacillus thuringiensis*. Microbiological Reviews 53: 242-255.

Hogg PJ (2003) Disulfide bonds as switches for protein function. Trends in Biochemical Sciences 28: 210-214.

Horn M, Patankar AG, Zavala JA, Wu J, Doleckova-Maresova L, Vujtechova M, *et al.* (2005) Differential elicitation of two processing proteases controls the processing pattern of the trypsin proteinase inhibitor precursor in *Nicotiana attenuata*. Plant Physiology 139: 375-388.

Horváth E, Szalai G, Janda T (2007) Induction of abiotic stress tolerance by salicylic acid signaling. Journal of Plant Growth Regulators 26: 290-300.

Hosokawa T, Kikuchi Y, Shimada M, Fukatsu T (2007) Obligate symbiont involved in pest status of host insect. Proceedings of the Royal Society B: Biological Sciences 22: 1979-1984.

http:// www.ebi.ac.uk

http://www.rcsb.org/pdb

http://zhang-lab.ccmb.med.umich.edu/I-TASSER

Huesgen PF, Overall CM (2011) N-and C-terminal degradomics: new approaches to reveal biological roles for plant proteases from substrate identification. Physiologia Plantarum 145: 5-17.

Hummel GM, Naumann M, Schurr U, Walter A (2007) Root growth dynamics of *Nicotiana attenuate* seedlings are affected by simulated herbivore attack. Plant, Cell, Environment 30: 1326-1336.

Iannuzzi R, Labandeira CC (2008) The oldest record and early history of insect folivory. Annals of the Entomological Society of America 10: 179–194.

Ingersoll CM, Niesenbaum RA, Weigle CE, Lehman JH (2010) Total phenolics and individual phenolic acids vary with light environment in Lindera benzoin. Botany 88: 1007-1010.

Jaenike J (1990) Factors maintaining genetic variation for host preference in Drosophila. Ecological and Evolutionary Genetics of Drosophila (New York: Plenum): 195-207.

Jamieson MA, Trowbridge AM, Raffa KF, Lindroth R L (2012) Consequences of climate warming and altered precipitation patterns for plant-insect and multitrophic interactions. Plant Physiology 160: 1719-1727.

Jersakova J, Johnson SD, Kindlmann P (2006) Mechanisms and evolution of deceptive pollination in orchids. Biological Reviews 81: 219-235.

Jiang M, Xie MX, Zheng D, Liu Z, Li XY Chen X (2004) Spectroscopic studies on the interaction of cinnamic acid and its hydroxyl derivatives with human serum albumin. Journal of Molecular Structure 692: 71-80.

Jin XL, Wei X, Qi FM, Yu SS, Zhou B, Bai S (2012) Characterization of hydroxycinnamic acid derivatives binding to bovine serum albumin. Organic & Biomolecular Chemistry 10: 3424-3431

Johnson ED, Miller EA, Anderson MA (2007) Dual location of a family of proteinase inhibitors within the stigmas of *Nicotiana alata.* Planta 225: 1265-1276.

Johnson ED, Miller EA, Anderson MA (2007) Dual location of a family of proteinase inhibitors within the stigmas of *Nicotiana alata*. Planta 225: 1265-1276.

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

Johnston KA, Gatehouse JA, Anstee JH (1993) Effect of soybean protease inhibitors on the growth and development of larval *Helicoverpa armigera*. Journal of Insect Physiology 39: 657-664.

Johnston KA, Lee MJ, Gatehouse JA, Anstee JH (1991) The partial purification and characterization of serine protease activity in midgut of larval *Helicoverpa armigera.* Insect Biochemistry 21: 389-397.

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

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, Stiekema WJ, Bosch D (1996) Combatting inhibitor-insensitive proteases of insect pests. Trends in Biotechnology 14: 331–333.

Jongsma MA, Beekwilder J (2011) Co-evolution of insect proteases and plant protease inhibitors. Current Protein Peptide Science 12: 437-447.

Kang J, Liu Y, Xie MX, Li S, Jiang M, Wang YD (2004) Interactions of human serum albumin with chlorogenic acid and ferulic acid. Biochimica et Biophysica Acta (BBA) - General Subjects 1674: 205-214.

Kaplan W, Littlejohn TG (2001) Swiss-PDB viewer (deep view). Briefings in Bioinformatics 2: 195-197.

Karban R (1989) Induced plant response to herbivory. Annual Review of Ecology, Evolution, and Systematics 20: 331-48.

Karban R, Baldwin IT (1997) Induced responses to herbivory. Chicago: University of Chicago Press.

Kasting JF, Catling D (2003) Evolution of a habitable planet. Annual Review of Astronomy and Astrophysics 41: 429–436.

Kessler A, Baldwin IT (2001) Defensive function of herbivore-induced plant volatile emissions in nature. Science 1291: 2141–2144.

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

Kim S, Hong Y-N, An CS, Lee K-W (2001) Expression characteristics of serine proteinase inhibitor II under variable environmental stresses in hot pepper (*Capsicum annuum* L.). Plant Science 161: 27-33.

Koller W, Kolattukudy PE (1982) Mechanism of action of cutinase, chemical modification of the catalytic triad characteristic for serine hydrolases. Biochemistry 21: 3083-3090.

Komiyama T, VanderLugt B, Fugère 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: 8205-10.

Konarev, Petr V., Volkov VV, Sokolova AV, Koch MHJ, Svergun DI (2003) PRIMUS: a Windows PC-based system for small-angle scattering data analysis. Journal of Applied Crystallography 36: 1277-1282.

Kong L, Ranganathan S (2008) Tandem duplication, circular permutation, molecular adaptation: How Solanaceae resist pests via inhibitors. BMC Bioinformatics 9 (Suppl 1): S22.

Koo AJK, Gao X, Jones AD, Howe GA (2009) A rapid wound signal activates the systemic synthesis of bioactive jasmonates in Arabidopsis. The Plant Journal 59: 974–986.

Korth K, Dixon RA (1997a) Evidence for chewing insect-specific molecular events distinct from a general wound response in leaves. Plant Physiology 115: 1299-1305.

Korth KL, Dixon RA (1997b) Differential transcript accumulation near wound sites in response to insect or mechanical damage. Plant Physiology 11: 4-26.

Kotkar HM, Sarate PJ, Tamhane VA, Gupta VS, Giri AP (2009) Responses of midgut amylases of Helicoverpa armigera to feeding on various host plants. Journal of Insect Physiology 55(8): 663-670.

Kotkar HM, Bhide AJ, Gupta VS, Giri AP (2012) Amylase gene expression patterns in *Helicoverpa armigera* upon feeding on a range of host plants. Gene 501: 1-7.

Krieger RI, Feeny PP, Wilkinson CF (1971. Detoxification enzymes in the guts of caterpillars: An evolutionary answer to plant defenses? Science 172: 579–581.

Krowarsch D, Cierpicki T, Jelen F, Otlewski J (2003) Canonical protein inhibitors of serine proteases. Cellular and Molecular Life Sciences 60: 2427–2444.

Kumar P, Pandit SS, Steppuhn A, Baldwin IT (2014) Natural history-driven, plant-mediated RNAi-based study reveals CYP6B46's role in a nicotine-mediated antipredator herbivore defense. Proceedings of National Academy of Science, USA 111: 1245-1252.

Kwan KM, Fujimoto W, Grabher C, Mangum BD, Hardy ME, Campbell DS, *et al.* (2007) The Tol2kit: A multisite gateway-based construction kit for Tol2 transposon transgenesis constructs. Developmental Dynamics 236: 3088-3099.

Labandeira CC (1997) Insect mouthparts: ascertaining the paleobiology of insect feeding strategies. Annual Review of Ecology, Evolution, and Systematics 28: 153–193.

Labandeira CC (1998) Early history of arthropod and vascular plant associations. Annual Review of Earth and Planetary Sciences 26: 329–377.

Labandeira CC (2013) A paleobiological perspective on plant–insect interactions. Current Opinion in Plant Biology 16: 414–421.

Labandeira CC, Beall BS (1990) Arthropod terrestriality. Arthropods: notes for a short course (Knoxville: University of Tennessee Press): 214–256.

Labandeira CC, Currano ED (2013) The fossil record of plant-insect dynamics. Annual Review of Earth and Planetary Sciences 41: 287–311.

Labandeira CC, Dilcher DL, Davis DR Wagner DL (1994). Ninety-seven million years of angiosperm-insect association: paleobiological insights into the meaning of coevolution. Proceedings of National Academy of Science, USA 91: 12278–12282.

Labandeira CC, Johnson KR, Wilf P (2002) Impact of the terminal Cretaceous event on plant–insect associations. Proceedings of National Academy of Science, USA 99: 2061–2066.

Labandeira CC, Sepkoski JJ (1993) Insect diversity in the fossil record. Science 261: 310–315.

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

Laskowski M Jr, Kato I, Ardelt W, Cook J, Denton A, Empie MW et al. (1987) Ovomucoid third domains from 100 avian species: Isolation, sequences, and hypervariability of enzyme inhibitor contact residues. Biochemistry 26: 202–221.

Lattanzio V, Kroon PA, Quideau S, Treutter D (2008) Plant phenolics—secondary metabolites with diverse functions. Recent Advancement in Polymer Research 1: 1-35.

Lawrence PK, Koundal KR (2002) Plant proteinase inhibitors in control of phytophagous insects. Electronic Journal of Biotechnology 5: 93-109.

Lay FT, Anderson MA (2005) Defensins-components of the innate immune system in plants. Current Protein and Peptide Science 6: 85-101.

Lee MC, Scanlon MJ, Craik DJ, Anderson MA (1999) A novel two-chain proteinase inhibitor generated by circularization of a multidomain precursor protein. Nature Structural and Molecular Biology 6: 526-530.

Lesner A, Legowska A, Wysocka M, Rolka K (2011) Sunflower trypsin inhibitor 1 as a molecular scaffold for drug discovery. Current Pharmaceutical Design 17: 4308-4317.

Lewis WJ, van Lenteren JC, Phatak SC, Tumlinson III JH (1997) A total system approach to sustainable pest management. Proceedings of the National Academy of Sciences, USA 94: 12243-12248.

Li L, Li C, Lee GI, Howe GA (2002) Distinct roles for jasmonate synthesis and action in the systemic wound response of tomato. Proceedings of National Academy of Sciences, USA 99: 6416-6421.

Li VI, Melati MR (1999) Examples of seed dispersal by entomochory. Acta botanica Gallica 146: 145-156.

Li XQ, Zhang T, Donnelly D (2011) Selective loss of cysteine residues and disulphide bonds in a potato proteinase II family. PLoS ONE 6(4): e18615.

Lindroth RL (2010). Impacts of elevated atmospheric CO2 and O3 on forests: Phytochemistry, trophic interactions and ecosystem dynamics. Journal of Chemical Ecology 36: 2-21.

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.

Liu J, Xia KF, Zhu JC, Deng YG, Huang XL, Hu BL, *et al.* (2006) The nightshade proteinase inhibitor IIb gene is constitutively expressed in glandular trichomes. Plant and Cell Physiology 47: 1274-1284.

Logan JA, Regniere J, Powell JA (2003) Assessing the impacts of global warming on forest pest dynamics. Frontiers in Ecology and the Environment 1: 130–137.

Lomate PR, Hivrale VK (2010) Partial purification and characterization of *Helicoverpa armigera* (Lepidoptera: Noctuidae) active aminopeptidase secreted in midgut. Comparative Biochemistry and Physiology - Part B 155: 164-170.

Lomate PR, Hivrale VK (2011) Differential responses of midgut soluble aminopeptidases of Helicoverpa armigera to feeding on various host and non-host plant diets. Arthropod-Plant Interaction 5: 359–368.

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

Lucas WJ, Lee JY (2004) Plant cell biology-plasmodesmata as a supracellular control network in plants. Nature Reviews Molecular Cell Biology 5: 712-726.

Luckey M, Ling R, Dose A, Malloy B (1991) Role of a disulfide bond in the thermal stability of the lam b protein trimer in escherichia coli outer membrane. Journal of Biological Chemistry 266: 1866-1871.

Luo M, Wang Z, Li H, Xia KF, Cai Y, Xu ZF (2009) Overexpression of a weed (*Solanum americanum*) proteinase inhibitor in transgenic tobacco results in increased glandular trichome density and enhanced resistance to *Helicoverpa armigera* and *Spodoptera litura*. International Journal of Molecular Sciences 10: 1896-1910.

Magalhães ST, Fernandes FL, Demuner AJ, Picanço MC, Guedes RNC (2010) Leaf alkaloids, phenolics, and coffee resistance to the leaf miner *Leucoptera coffeella* (Lepidoptera: Lyonetiidae). Journal of Economical Entomology 103: 1438-1443.

Mahajan NS, Mishra M, Tamhane VA, Gupta VS, Giri AP (2013) Plasticity of protease gene expression in *Helicoverpa armigera* upon exposure to multi-domain *Capsicum annuum* protease inhibitor. Biochimica et Biophysica Acta (BBA) - General Subjects 1830: 3414-3420.

Maheswaran G, Pridmore L, Franz P, Anderson MA (2007) A proteinase inhibitor from *Nicotiana alata* inhibits the normal development of light-brown apple moth, *Epiphyas postvittana* in transgenic apple plants. Plant Cell Reports 26: 773–782.

Mahmood K, Khan MB, Song YY, Ijaz M, Luo SM, Zeng RS (2013) UV-irradiation enhances rice allelopathic potential in rhizosphere soil. Journal of Plant Growth Regulators 71: 21-29.

Malone LA, Burgess EPJ (2000) Interference of protease inhibitors on non-target organisms. Recombinant protease inhibitors in plants (Georgetown TX: Landes Bioscience): 89–106.

Malone M (1996) Rapid, long-distance signal transmission in higher plants. Advances in botanical research (Academic Press, San Diego): 163–228.

Mazumdar-Leighton S, Broadway RM (2001b) Identification of six chymotrypsin cDNAs from larval mid guts of *Helicoverpa zea* and *Agrtis ipsilon* feeding on the soybean (Kunitz) trypsin inhibitor. Insect Biochemistry and Molecular Biology 31: 633-644.

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.

Mazumdar-Leighton S, Broadway RM (2001a) 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.

McCafferty HR, Moore PH, Zhu YJ (2008) Papaya transformed with the *Galanthus nivalis* GNA gene produces a biologically active lectin with spider mite control activity. Plant science 175: 385-393.

McCaffery AR (1998) Resistance to heliothine Lepidoptera: A global view. Philosophical Transactions of the Royal Society of London B 353: 1735-1750.

Mcelwain JC, Punyasena S (2007) Mass extinction events and the plant fossil record. Trends in Ecology, Evolution 22: 548–557.

McGaughey WH, Gould F, Gelernter W (1998) Bt resistance management: a plan for reconciling the needs of the many stakeholders in Bt–based products. Nature Biotechnology 16: 144–146.

McGurl B, Orozco-Cardenas M, Pearce G, Ryan CA (1994) Overexpression of the prosystemin gene in transgenic tomato plants generates a systemic signal that constitutively induces proteinase inhibitor synthesis. Proceedings of National Academy of Sciences, USA 91: 9799-9802.

McGurl B, Pearce G, Orozco-Cardenas M, Ryan CA (1992) Structure, expression and antisense inhibition of the systemin precursor gene. Science 255: 1570-1573.

McManus MT, White DWR, McGregor PG (1994) Accumulation of the chymotrypsin inhibitor in transgenic tobacco can affect the growth of insect pests. Transgenic Research 3: 50-58.

Melissa P, Grossi-de-Sá MF, Barcellos GBS, Carlini, Moraes MG (2003) Characterization and expression of a novel member (JBURE-II) of the urease gene family from jackbean [*Canavalia ensiformis* (L.) DC]. Plant and Cell Physiology 44: 139-145.

Mello MO, Silva-Filho MC (2002) Plant-insect interactions: an evolutionary arms race between two distinct defense mechanisms. Brazilian Journal of Plant Physiology 14: 71-81.

Memmott J, Craze PG, Waser NM, Price MV (2007) Global warming and the disruption of plant–pollinator interactions. Ecology Letters 10: 710–717.

Michaud D (1997) Avoiding protease-mediated resistance in herbivorous pests. Trends in Biotechnology 7799: 10072-X

Michener CD (2007) The bees of the world. 2nd ed. Baltimore, M. D.: The John Hopkins University Press.

Mickel CE, Sytandish 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

Miller EA, Lee MC, Atkinson AH, Anderson MA (2000) Identification of a novel four-domain member of the proteinase inhibitor II family from the stigmas of *Nicotiana alata.* Plant Molecular Biology 42: 329-333.

Mishra M, Joshi RS, Gaikwad S, Gupta VS, Giri AP (2013) Structural-functional insights of single and multi-domain Capsicum annuum protease inhibitors. Biochemical and Biophysical Research Communications 430: 1060-1065.

Mishra M, Tamhane VA, Khandelwal N, Kulkarni MJ, Gupta VS, Giri AP (2010) Interaction of recombinant CanPIs with *Helicoverpa armigera* gut proteases reveals their processing patterns, stability and efficiency. Proteomics 10: 2845-2857.

Mithofer A, Boland W (2008) Recognition of herbivory-associated molecular patterns. Plant Physiology 146: 825-31.

Mithofer A, Boland W (2012) Plant defense against herbivores: Chemical aspects. Annual Review of Plant Biology 63: 431-450.

Mithofer A, Boland W, Maffei ME (2009) Chemical ecology of plant-insect interactions. Plant disease resistance (Chichester: Wiley-Blackwell): 261-291.

Mitter C, Farrell B (1991) Macroevolutionary aspects of insect- plant relationships. Insect-plant interactions. (Boca Raton: CRC Press): 35-78.

Mitterboeck TF, Adamowicz SJ (2013) Flight loss linked to faster molecular evolution in insects. Proceedings of the Royal Society B: Biological Sciences 280: 11-28.

Morris G, Goodsell D, Pique M, Lindstrom W, Huey R, Forli S, *et al.* (2009) AutoDock 4.2 with AutoDockTools

Mosolov VV, Valueva TA (2008) Proteinase Inhibitors in Plant Biotechnology: A Review. Applied Biochemistry and Microbiology 44: 233–240.

Mosolov VV, Loginova MD, Fedurkina NV, Benken II (1976) The biological significance of proteinase inhibitors in plants. Plant Science Letters 7: 77-80.

Moura DS, Ryan CA (2001) Wound-inducible proteinase inhibitors in pepper.Differential regulation upon wounding, systemin and methyl jasmonate. Plant Physiology 126: 289-298.

Moyen C, Hammond-Kosack KE, Jones J, Knight MR, Johannes E (1998) Systemin triggers an increase of cytoplasmic calcium in tomato mesophyll cells: Ca2+ mobilization from intra-and extracellular compartments. Plant Cell Environ 21: 1101–1111.

Mulinari F, Staniscuaski F, Bertholdo-Vargas LR, Postal M, Oliveira-Neto OB, Rigden DJ *et al.* (2007). Jaburetox-2Ec: An insecticidal peptide derived from an isoform of urease from the plant *Canavalia ensiformis*. Peptides 28: 2042-2050.

Narv´aez-V´asquez J, Florin-Christensen J, Ryan CA (1999) Positional specificity of a phospholipase a activity induced by wounding, systemin, and oligosaccharide elicitors in tomato leaves. Plant Cell 11: 2249–2260.

Navazio L, Moscatiello R, Bellincampi D, Baldan B, Meggio F, Brini M, *et al.* (2002) The role of calcium in oligogalacturonide-activated signalling in soybean cells. Planta 215: 596–605.

Nelson WA, Bjornstad ON, Yamanaka T (2013) Recurrent insect outbreaks caused by temperature-driven changes in system stability. Science 341: 796–799.

Nielsen KJ, Heath RL, Anderson MA, Craik DJ (1994) The Three-dimensional solution structure by H NMR of a 6-kDa proteinase inhibitor isolated from the stigma of *Nicotiana alata.* Journal of Molecular Biology 242: 231-243.

Nielsen KJ, Heath RL, Anderson MA, Craik DJ (1995) Structures of a series of 6-kDa trypsin inhibitors isolated from the stigma of *Nicotiana alata*. Biochemistry 34: 14304–14311.

Niklas KJ, Tiffney BH, Knoll AH (1983) Pattern in vascular land plant diversification. Nature 303: 614-616.

Nisbet EG, Sleep NH (2001) The habitat and nature of early life. Nature 409: 1083–1091.

O'callaghan M, Glare TR, Burgess EPJ, Malone LA (2005) Effects of plants genetically modified for insect resistance on non-target organisms. Annual Review of Entomology 50: 271–292.

O"Donnell PJ, Calvert C, Atzorn R, Wasternack C, Leyser HMO, Bowles DJ (1996) Ethylene as a signal mediating the wound response of tomato plants. Science 274: 1914–1917.

Oganesyan N, Kim SH, Kim R (2004) On-column Chemical Refolding of Proteins. Pharmagenomics 22-25.

Orona-tamayo D, Wielsch N, Escalante-pe´rez M, Svatos A, Molina-torres J, Muck A, *et al.* (2013) Short-term proteomic dynamics reveal metabolic factory for active extrafloral nectar secretion by *Acacia cornigera* ant-plants. The Plant Journal 73: 546–554.

Orozco-C´ardenas M, McGurl B, Ryan CA (1993) Expression of an antisense prosystemin gene in tomato plants reduces the resistance toward *Manduca sexta* larvae. Proceedings of National Academy of Sciences, USA 90: 8273-8276.

Otlewski J, Jaskolski M, Buczek O, Cicerpicki 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.

Patankar AG, Giri AP, Harsulkar AM, Sainani MN, Deshpande VV, Ranjekar PK, *et al.* (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 MN, Ranjekar PK, *et al.* (1999) Diversity in inhibitors of trypsin and *Helicoverpa armigera* gut proteinases in chickpea (*Cicer arietinum*) and its wild relatives. Theoretical and Applied Genetics 99: 719–726.

Pearce G, Bhattacharya R, Chen Y, Barona G, Yamaguchi Y, Ryan CA (2009) Isolation and characterization of Hydroxyproline-rich glycopeptides signals in black nightshade leaves. Plant Physiology 150: 1422-1433.

Pearce G, Johnson S, Ryan CA (1993) Purification and characterization from tobacco (*Nicotiana tabacum*) leaves of six small, wound-inducible, proteinase isoinhibitors of the potato inhibitor II family. Plant Physiology 102: 639-644.

Pearce G, Ryan CA (2003) Systemic signaling in tomato plants for defense against herbivores: Isolation and characterization of three novel defense-signaling glycopeptide hormones coded in a single precursor gene. Journal of Biological Chemistry 278: 30044–30050.

Pearce G, Strydom D, Johnson S, Ryan CA (1991) A polypeptide from tomato leaves induces wound-inducible proteinase inhibitor proteins. Science 253: 895–898.

Pearson DE (2009) Invasive plant architecture alters trophic interactions by changing predator abundance and behavior. Oecologia 159: 549–558.

Pelhate M, Zlotkin E (1981) Voltage-dependent slowing of the turn off of Na^+ current in the cockroach giant-axon induced by the scorpion-venom insect toxin. Journal of Physiology-London 319: 30-31.

Pelton JT, McLean LR (2000) Spectroscopic methods for analysis of protein secondary structure. Analytical Biochemistry 277: 167-176.

Pena-Cortes H, Albrecht T, Prat S, Weiler E Willmitzer L (1993) Aspirin prevents wound-induced gene expression in tomato leaves by blocking jasmonic acid biosynthesis. Planta 191: 123-128.

Peña-Cortés H, Sanchez-Serrano J, Mertens R, Willmitzer L, Prat S (1989) Abscisic acid is involved in the wound-induced expression of the proteinase inhibitor II gene in potato and tomato. Proceedings of National Academy of Sciences, USA 86: 9851-9855.

Pena-Cortes H, Willmitzer L, Sánchez-Serrano JJ (1991) Abscisic acid mediates wound induction but not developmental-specific expression of the proteinase inhibitor II gene family. Plant Cell 3: 963-972.

Percec V, Peterca M, Sienkowska MJ, Ilies MA, Aqad E, Smidrkal J, *et al.* (2006) Synthesis and retrostructural analysis of libraries of AB3 and constitutional isomeric AB2 phenylpropyl ether-based supramolecular dendrimers. Journal of the American Chemical Society 128: 3324-3334.

Philipp S, Kim YM, Durr I, Wenzl G, Vogt M, Flecker P (1998) Mutational analysis of disulfide bonds in the trypsin-reactive subdomain of a Bowman-Birk-type inhibitor of trypsin and chymotrypsin-cooperative versus autonomous refolding of subdomains. European Journal of Biochemistry 251: 854-862.

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

Pieterse CMJ, Ton J, Van Loon LC (2001) Cross-talk between plant defence signalling pathways: boost or burden? AgBiotechNet 3: ABN 068.

Pimentel D, Acquay H, Biltonen M, Rice P, Silva M, Nelson J, *et al.* (1992) Environmental and economic costs of pesticide use. BioScience 42: 750-760.

Pimentel D, Edwards CA (1982) Pesticides and ecosystems. BioScience 32: 595-600.

Pimentel D, Pimentel M (1990) Comment: Adverse environmental consequences of the Green Revolution. Population and Development Review 16: 329-332.

Pires-Alves M, Grossi-de-Sá MF, Barcellos GB, Carlini CR, Moraes MG (2003) Characterization and expression of a novel member (JBURE-II) of the urease gene family from jackbean [*Canavalia ensiformis* (L.) DC]. Plant and cell physiology 44: 139-145.

Plunkett G, Senear DF, Zuroske G, Ryan CA (1982) Proteinase inhibitors I and II from leaves of wounded tomato plants: Purification and properties. Archives of Biochemistry and Biophysics 213: 463-472.

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

Powell JA, Bentz BJ (2009) Connecting phenological predictions with population growth rates for mountain pine beetle, an outbreak insect. Landscape Ecology 24: 657–672.

Price PW, Westoby M, Rice B, Atsatt PR, Fritz RS, Thompson JN, *et al.* (1986) Parasite mediation in ecological interactions. Annual Review of Ecology, Evolution, and Systematics 17: 487-505.

Purcell JP, Greenplate JT, Douglas Sammons R (1992) Examination of midgut luminal proteinase activities in six economically important insects. Insect Biochemistry and Molecular Biology 22: 41-47.

Qi RF, Song ZW, Chi CE (2005) Structural features and molecular evolution of Bowman-Birk protease inhibitors and their potential application. Acta Biochimica et Biophysica Sinica 37: 283-292.

Raffa KF, Aukema BH, Bentz BJ, Carroll AL, Hicke JA, Turner MG, *et al.* (2008) Cross-scale drivers of natural disturbances prone to anthropogenic amplification: The dynamics of bark beetle eruptions. BioScience 58: 501-517.

Rajapakse CNK, Walter GH (2007) Polyphagy and primary host plants: Oviposition preference versus larval performance in the lepidopteran pest *Helicoverpa armigera*. Arthropod-Plant Interactions 1: 17-26

Rani UP, Devanand P (2013) Bioactivities of caffeic acid methyl ester (methyl-(E)-3-(3, 4-dihydroxyphenyl) prop-2-enoate): A hydroxycinnamic acid derivative from *Solanum melongena* L. fruits. Journal of Pest Science 86: 579-589.

Rattan ES (2010) Mechanism of action of insecticidal secondary metabolites of plant origin. Crop Protection 29: 913-920.

Reddy KVS, Zehr UB (2004) In New directions for a diverse planet. Proceedings of the 4th International Crop Science Congress. Brisbane, Australia, 1-8.

Reeck 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, 157-183.

Reymond P, Weber H, Damond M, Farmer EE (2000) Differential gene expression in response to mechanical wounding and insect feeding in Arabidopsis. Plant Cell 12: 707–719.

Richardson M (1979) The complete amino acid sequence and trypsin reactive (inhibitory) site of the major proteinase inhibitor from the fruits of aubergine (*Solanum melongena* L.). FEBS Letters 104: 322-326.

Roche M, Dufour C, Mora N, Dangles O (2005) Antioxidant activity of olive phenols: mechanistic investigation and characterization of oxidation products by mass spectrometry. Organic and Biomolecular Chemistry 3: 423-430.

Rolfe WDI (1985) Aspects of the Carboniferous terrestrial arthropod community. Congres International de Stratigraphie et de Geologie du Carbonifere Compte Rendu 9: 303-316.

Romero GQ, Souza JC, Vasconcellos-neto J (2008) Anti-herbivore protection by mutualistic spiders and the role of plant glandular trichomes. Ecology 89: 3105–3115.

Roy A, Kucukural A, Zhang Y (2010) I-TASSER: A unified platform for automated protein structure and function prediction. Nature Protocols 5: 725-738.

Roy A, Xu D, Poisson J, Zhang Y (2011) A protocol for computer-based protein structure and function prediction. journal of visualized experiments 57: e3259.

Royo J, Leo J, Vancanneyt G, Albar JP, Rosahl S, Ortego F, *et al.* (1999) Antisense -mediated depletion of a potato lipoxygenase reduces wound induction of proteinase inhibitors and increases weight gain of insect pests. Proceedings of the National Academy of Sciences, USA 96: 1146-1151.

Ryan CA (1990) Proteinase inhibitors in plants, genes for improving defenses against insects and pathogens. Annual Reviews on Phytopathology 28: 425-449.

Ryan CA (2000) The systemin signaling pathway: Differential activation of plant defensive genes. Biochimia et Biophysicia Acta 1477: 112-121.

Ryan CA, Moura DS (2002) Systemic wound signaling in plants: A new perception. Proceedings of National Academy of Sciences, USA 99: 6519-6520.

Saeed AI, Bhagabati NK, Braisted JC, Liang W, Sharov V, Howe EA, *et al.* (2006). TM4 microarray software suite. Methods in Enzymology 411: 134-193.

Sagers CL (1992) Manipulating host plant quality: Herbivores keep leaves in the dark. Functional Ecology 6: 741-743.

Saha P, Majumder P, Dutta I, Ray T, Roy SC, Das S (2006) Transgenic rice expressing *Allium sativum* leaf lectin with enhanced resistance against sap-sucking insect pests. Planta 223: 1329-1343.

Sánchez-Serrano JJ, Schmidt R, Schell J, Willmitzer L (1986) Nucleotide sequence of proteinase inhibitor II encoding cDNA of potato (*Solanum tuberosum*) and its mode of expression. Molecular and General Genetics 203: 15-20.

Sanchis V, Bourguet D (2008) *Bacillus thuringiensis*: applications in agriculture and insect resistance management: A review. Agronomy for Sustainable Development 28: 11-20.

Sarate PJ, Tamhane VA, Kotkar HM, Ratnakaran N, Susan N, Gupta VS, *et al.* (2012) Developmental and digestive flexibilities in the midgut of a polyphagous pest, the cotton bollworm, *Helicoverpa armigera*. Journal of Insect Science 12: 1-16.

Scanlon MJ, Lee MC, Anderson MA, Craik DJ (1999) Structure of a putative ancestral protein encoded by a single sequence repeat from a multidomain proteinase inhibitor gene from *Nicotiana alata*. Structure, Folding and Designing 7: 793- 802.

Schägger H (2006) Tricine–SDS-PAGE. Nature Protocols 1: 16-22.

Schirra HJ, Craik DJ (2005) Structure and folding of potato type II proteinase inhibitors: Circular permutation and intramolecular domain swapping. Protein and Peptide Letters 12: 421-431.

Schirra HJ, Anderson MA, Craik DJ (2008) Structural refinement of insecticidal plant proteinase inhibitors from *Nicotiana alata.* Protein and Peptide Letters 15: 903-909.

Schirra HJ, Guarino RF, Anderson MA, Craik DJ (2010) Selective removal of individual disulfide bonds within a potato type II serine proteinase inhibitor from *Nicotiana alata* reveals differential stabilization of the reactive-site loop. Journal of Molecular Biology 395: 609-626.

Schirra HJ, Scanlon MJ, Lee MC, Anderson MA, Craik DJ (2001) The solution structure of C1-T1, a two-domain proteinase inhibitor derived from a circular precursor protein from *Nicotiana alata.* Journal of Molecular Biology 306: 69-79.

Schlüter U, Benchabane M, Munger A, Kiggundu A, Vorster L, Goulet M, *et al.* (2010) Recombinant protease inhibitors for herbivore pest control: A multitrophic perspective. Journal of Experimental Botany 61: 4169-4183.

Schneidman-Duhovny D, Hammel M, Sali A (2010) FoXS: A web server for rapid computation and fitting of SAXS profiles. Nucleic Acids Research 38: 540-544.

Schneidman-Duhovny D, Inbar Y, Nussinov R Wolfson HJ (2005) PatchDock and SymmDock: Servers for rigid and symmetric docking. Nucleic Acids Research 33: 363-367.

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

Scott AC, Chaloner WG, Paterson S (1985). Evidence of pteridophyte-arthropod interactions in the fossil record. Proceedings of the Royal Society B: Biological Sciences 86: 133–140.

Scott AC, Stephenson J, Chaloner WG (1992) Interaction and coevolution of plants and arthropods during the Palaeozoic and Mesozoic. Philosophical Transactions of the Royal Society B: Biological Sciences 335: 129–165.

Semisotnov GV, Rodionova NA, Razgulyaev OI, Uversky VN, Gripas AF, Gilmanshin RI (1991) Study of the "molten globule" intermediate state in protein folding by a hydrophobic fluorescent probe. Biopolymers 31: 119-128.

Sharma HC, Sharma KK, Seetharam N, Ortiz R (2000) Prospect for using transgenic resistance to insects in crop improvement. Electronic Journal of Biotechnology 3: 76-95.

Shear WA (1991) The early development of terrestrial ecosystems. Nature 351: 183–189.

Shin R, Lee GJ, Park CJ, Kim TY, You JS, Nam YW, *et al.* (2001) Isolation of pepper mRNAs differentially expressed during the hypersensitive response to tobacco mosaic virus and characterization of a proteinase inhibitor gene. Plant Science 161: 727-737.

Sin SF, Chye ML (2004) Expression of proteinase inhibitor II proteins during floral development in *Solanum americanum*. Planta 219: 1010–1022.

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

Skibbe M, Qu N, Galis I, Baldwin IT (2008) Induced plant defenses in the natural environment: *Nicotiana attenuata* WRKY3 and WRKY6 coordinate responses to herbivory. Plant Cell 20: 1984–2000.

Slack A, Gate J (2000) Carnivorous Plants. MIT Press.

Smyth E, Syme CD, Blanch EW, Hecht L, Vasak M, Barron LD (2001) Solution structure of native proteins with irregular folds from raman optical activity. Biopolymers 58: 138–151.

Snyder MS, Glendinning JI (1996) Causal connection between detoxification enzyme activity and consumption of a toxic plant compound. Journal of Comparative Physiology A 179: 255–261.

Spafford RD, Lortie CJ (2013) Sweeping beauty: Is grassland arthropod community composition effectively estimated by sweep netting? Ecology and Evolution 3: 3347–3358.

Srinivasan A, Giri AP, Gupta VS (2006) Structural and functional diversities in Lepidopteran serine proteases. Cellular and Molecular Biology Letters 11: 132–154.

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.

Stone GN, Hernandez-lopez A, Nicholls JA, Di pierro E, Pujade-villar J, Melika G, *et al.* (2009) Extreme host plant conservatism during at least 20 million years of host plant pursuit by oak gallwasps. Evolution 63: 854–869.

Stotz, HU, Thomson JG, Wang Y (2009) Plant defensins: Defense, development and application. Plant Signaling and Behavior 4: 1010-1012.

Strauss AS, Peters S, Boland W, Burse A (2013) ABC transporter functions as a pacemaker for sequestration of plant glucosides in leaf beetles. elife 2: e01096.

Swedberg JE, de Veer SJ, Sit KC, Reboul CF, Buckle AM, Harris JM (2011) Mastering the canonical loop of serine protease inhibitors: Enhancing potency by optimizing the internal hydrogen bond network. PLoS One 6: e19302.

Tabashnik BE, Gassmann AJ, Crowder DW, Carriere Y (2008) Insect resistance to Bt crops: Evidence versus theory. Nature Biotechnology 26: 199-202.

Tahvanainen J, Niemela P (1987). Biogeographical and evolutionary aspects of insect herbivory. Annales Zoologici Fennici 24: 239-247.

Takahashi S, Itino T (2012) Larger seeds are dispersed farther: The long-distance seed disperser ant *Aphaenogaster famelica* prefers larger seeds. Sociobiology 59: 1401-1411.

Takahashi T, Miyazawa M (2010) Tyrosinase inhibitory activities of cinnamic acid analogues. Pharmazie 65: 913-918.

Takhtajan A (1991) Evolutionary trends in flowering plants. New York: Columbia University Press.

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 -General Subjects 1722: 156-167.

Tamhane VA, Giri AP, Gupta VS (2007) *Helicoverpa armigera*: Ecology and Control Using Novel Biotechnological Approaches. Encyclopedia of Pest Management (CRC Press) 2: 232-236.

Tamhane VA, Giri AP, Kumar P, Gupta VS (2009) Spatial and temporal expression patterns of diverse Pin-II proteinase inhibitor genes in *Capsicum annuum* Linn. Gene 442: 88-98.

Tamhane VA, Giri AP, Sainani MN, Gupta VS (2007) Diverse forms of Pin-II family proteinase inhibitors from Capsicum annuum adversely affect the growth and development of *Helicoverpa armigera*. Gene 403: 29-38.

Taylor BH, Young RJ, Scheuring CF (1993) Induction of a proteinase inhibitor II class gene by auxin in tomato roots. Plant Molecular Biology 23: 1005-1014.

Taylor EL, Taylor TN (1992) Reproductive biology of the permian glossopteridales and their suggested relationship to the flowering plants. Proceedings of National Academy of Science, USA 89: 11495-11497.

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

Telang MA, Srinivasan A, Patankar A, Harsulkar A, Joshi V, Damle A, *et al.* (2003) Bitter gourd proteinase inhibitor: Potential growth inhibitors of *Helicoverpa armigera* and *Spodoptera litura*. Phytochemistry 63: 643-652.

Terenius O, Papanicolaou A, Garbutt JS, Eleftherianos I, Huvenne H, Kanginakudru S, et al. (2011) RNA interference in Lepidoptera: An overview of successful and unsuccessful studies and implications for experimental design. Journal of Insect Physiology 57: 231-245.

Terra WR (1988) Physiology and biochemistry of insect digestion: an evolutionary perspective. Brazilian Journal of Medical and Biological Research 21: 675-734.

Terra WR, Ferreira, C (1994) Insect digestive enzymes: Properties, compartmentalization and function. Comparative Biochemistry and Physiology Part B: Comparative Biochemistry 109: 1-62.

Thaler JS, Bostock RM (2004) Interactions between abscisic-acid-mediated responses and plant resistance to pathogens and insects. Ecology 85: 48–58.

Thaler JS, Owen B, Higgins VJ (2004) The role of the jasmonate response in plant susceptibility to diverse pathogens with a range of lifestyles. Plant Physiology 135: 530-538.

Theiling KM, Croft BA (1988) Pesticide side-effects on arthropod natural enemies: A database summary. Agriculture, Ecosystems, Environment 21: 191-218.

Treutter D (2006) Significance of flavonoids in plant resistance: A review. Environmental Chemistry Letters 4: 147-157.

Trott O, Olson AJ (2010) AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. Journal of Computational Chemistry 31: 455-461.

Tsui V, Case DA (2000) Theory and applications of the generalized born solvation model in macromolecular simulations. Biopolymers 56: 275-291.

Turra D, Lorito M (2011) Potato type I and II proteinase inhibitors: Modulating plant physiology and host resistance. Current Protein Peptide Science 12: 374-385.

Tylianakis JM, Didham RK, Bascompte J, Wardle DA (2008) Global change and species interactions in terrestrial ecosystems. Ecology Letters 11: 1351–1363.

Ulijaszek SJ (2002) Human eating behavior in an evolutionary ecological context. Proceedings of the Nutrition Society 61: 517-526.

Upchurch GR, Dilcher DL (1990) Cenomanian angiosperm leaf mega fossils, Dakota Formation, Rose Creek locality, Jefferson County, southeastern Nebraska. US Geological Survey Bulletin 1915: 1-52.

Usha Rani P, Pratyusha S (2013) Defensive role of *Gossypium hirsutum* L. anti-oxidative enzymes and phenolic acids in response to *Spodoptera litura* F. feeding. Journal of Asia Pacific Entomology 16: 131-136.

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.

Uwai K, Osanai Y, Imaizumi T, Kanno S, Takeshita M, Ishikawa M (2008) Inhibitory effect of the alkyl side chain of caffeic acid analogues on lipopolysaccharide-induced nitric oxide production in RAW264.7 macrophages. Bioorganic & Medicinal Chemistry 16: 7795-7803.

Valueva TA, Mosolov VV (2004) Role of inhibitors of proteolytic enzymes in plant defense against phytopathogenic microorganisms. Biokhimiya 69: 1600-1606.

Van der putten WH, Macel M, Visser ME (2010) Predicting species distribution and abundance responses to climate change: Why it is essential to include biotic interactions across trophic levels. Philosophical Transactions of the Royal Society B: Biological Science 365: 2025-2034.

Venturi M, Seifert C, Hunte C (2002) High level production of functional antibody fab fragments in an oxidizing bacterial cytoplasm. Journal of Molecular Biology 315: 1–8.

Visser ME, Holleman LJ (2001) Warmer springs disrupt the synchrony of oak and winter moth phenology. Proceedings of the Royal Society B: Biological Sciences 268: 289–294.

Voet D, Voet JG (2010) Biochemistry, 4th edition. John Wiley and Sons Inc., Corporate Headquarters, 111 River street, Hoboken, NJ 07030-5774.

Vogel S, Martens J (2000) A survey of the function of the lethal kettle traps of Arisaema (Araceae), with records of pollinating fungus gnats from Nepal. Botanical Journal of the Linnean Society 133: 61–100.

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

Walling LL (2000) The myriad plant responses to herbivores. Journal of Plant Growth Regulation 19: 195-216.

Walther CE Zlotkin, Rathmayer W (1976) Action of different toxins from the scorpion *Androctonus australis* on a locust nerve-muscle preparation. Journal of Insect Physiology 22: 1187-1194.

Wappler T, Currano ED, Wilf P, Rust J, Labandeira CC (2009) No post-Cretaceous ecosystem depression in European forests? Rich insect-feeding damage on diverse middle Palaeocene plants, Menat, France. Proceedings of the Royal Society B: Biological Sciences 276: 4271–4277.

War AR, Paulraj MG, Hussain B, Buhroo AA, Ignacimuthu S, Sharma HC (2013) Effect of plant secondary metabolites on legume pod borer, *Helicoverpa armigera*. Journal of Pest Science 86: 399-408.

War AR, Paulraj MG, Ahmad T, Buhroo AA, Hussain B, Ignacimuthu S, *et al.* (2012) Mechanisms of plant defense against insect herbivores. Plant Signal Behaviour 7: 1306-1320.

Weber PC, Salemme FR (2003) Applications of calorimetric methods to drug discovery and the study of protein interaction. Current Opinion in Structural Biology 13: 115-121.

Weinhold A, Baldwin IT (2011) Trichome-derived O-acyl sugars are a first meal for caterpillars that tags them for predation. Proceedings of National Academy of Science, USA 108: 7855-7859.

Weis AE, Kapelinski AD (1994) Variable selection on Eurosta's gall size, II: A path analysis of the ecological factors behind selection. Evolution 48: 734-745.

Wetzel R, Perry LJ, Baase WA, Becktel WJ (1988) Disulfide bonds and thermal stability in T4 lysozyme. Proceedings of National Academy of Sciences, USA 85: 401-405.

Wildon DC, Thain JF, Minchin PEH, Gubb IR, Reilly AJ, Skipper YD, *et al.* (1992) Electrical signalling and systemic proteinase inhibitor induction in the wounded plant. Nature 360: 362-65.

Wilf P (2008) Insect-damaged fossil leaves record food web response to ancient climate change and extinction. New Phytologist 178: 486-502.

Wilf P, Labandeira CC (1999) Response of plant-insect associations to paleocene-eocene warming. Science 284: 2153-2155.

Winkler IS, Labandeira CC, Wappler T, Wilf P (2010) Diptera (Agromyzidae) leaf mines from the Paleogene of North America and Germany: Implications for host use evolution and an early origin for the Agromyzidae. Journal of Paleontology 84: 935–954.

Wu J, Baldwin IT (2010) New insights into plant responses to the attack from insect herbivores. Annual Review of Genetics 44: 1-24.

Wu J, Hettenhausen C, Baldwin IT (2006) Evolution of proteinase inhibitor defences in North American allopolyploid species of *Nicotiana*. Planta 224: 750–760.

Wu Y, Llewellyn D, Mathews A, Dennis ES (1997) Adaptation of *Helicoverpa armigera* (Lepidoptera, Noctuidae) to a proteinase inhibitor expressed in transgenic tobacco. Molecular Breeding 3: 371-380.

Xia Q, Zhou Z, Lu C, Cheng D, Dai F, Li B, *et al.* (2004) A draft sequence for the genome of the domesticated silkworm (*Bombyx mori*). Science 306: 1937-1940.

Xie J, Ouyang XZ, Xia KF, Huang YF, Pan WB, Cai YP, *et al.* (2007) Chloroplast-like organelles were found in enucleate sieve elements of transgenic plants overexpressing a proteinase inhibitor. Bioscience, Biotechnology and Biochemistry 71: 2759-2765.

Xiong S, Wang Y, Rong Ren X, Li B, Zhang MY, Luo Y, *et al.* (2005) Solubility of disulfide-bonded proteins in the cytoplasm of *Escherichia coli* and its "oxidizing" mutant. World Journal of Gastroenterology 11: 1077-1082.

Xu L, Li Y, Zhou S, Hou Y (2012) Understanding microscopic binding of macrophage migration inhibitory factor with phenolic hydrazones by molecular docking, molecular dynamics simulations and free energy calculations. Molecular BioSystem 8: 2260-2273

Xu ZF, Qi WQ, Ouyang XZ, Yeung E, Chye ML (2001) A proteinase inhibitor II of *Solanum americanum* is expressed in phloem. Plant Molecular Biology 47: 727–738.

Xu ZF, Teng WL, Chye ML (2004) Inhibition of endogenous trypsin-and chymotrypsin-like activities in transgenic lettuce expressing heterogeneous proteinase inhibitor SaPIN2a. Planta 218: 623–629.

Yanes O, Aviles FX, Roepstorff P, Jørgensen TJD (2007) Exploring the "IntensityFading" Phenomenon in the Study of Noncovalent Interactions by MALDI-TOF Mass Spectrometry. American Society for Mass Spectrometry 18: 359-367.

Yang W, Fang DM, He HP, Hao XJ, Wu ZJ, Zhang GL (2013) Analysis of mexicanolide-and phragmalin-type limonoids from Heynea trijuga using high-performance liquid chromatography/electrospray tandem mass spectrometry. Rapid Communications in Mass Spectrometry 27: 1203-1212.

Yang Z (2008) I-TASSER server for protein 3D structure prediction. BMC Bioinformatics 9: 40.

Young MA, Pysh ES (1975) Vacuum ultraviolet circular dichroism of Poly (L-proline) I and II. Journal of the American chemical Society 97: 5100-5103.

Zalucki M, Daglish G, Firempong S, Twine P (1986) The biology and ecology of *Heliothis armigera* (Hubner) and *Heliothis punctigera* Wallengren (Lepidoptera, Noctuidae) in Australia - what do we know. Australian Journal of Zoology 34: 779-814.

Zavada MS (1984) Angiosperm origins and evolution based on dispersed fossil pollen ultrastructure. Annals of the Missouri Botanical Garden 71: 444–463.

Zavala JA, Giri AP, Jongsma MA, Baldwin IT (2008) Digestive duet: Midgut digestive proteinases of Manduca sexta ingesting Nicotiana attenuata with manipulated trypsin proteinase inhibitor expression. PLoS ONE 3, e2008.

Zavala JA, Patankar AG, Gase K, Baldwin IT (2004a) Constitutive and inducible trypsin proteinase inhibitor production incurs large fitness costs in *Nicotiana attenuata.* Proceedings of the National Academy of Sciences, USA 101: 1607-1612.

Zavala JA, Patankar AG, Gase K, Hui D, Baldwin IT (2004b) Manipulation of endogenous trypsin proteinase inhibitor production in *Nicotiana attenuata* demonstrates their function as anti herbivore defenses. Plant Physiology 134: 1181-1190.

Zavala JA, Nabity PD, Delucia EH (2013) An emerging understanding of mechanisms governing insect herbivory under elevated CO2. Annual Review of Entomology 58: 79–97.

Zhu Y, Zhang LX, Zhao Y, Huang GD (2010) Unusual sesquiterpene lactones with a new carbon skeleton and new acetylenes from *Ajania przewalskii*. Food Chemistry 118: 228-238.

Zlotkin ED, Kadouri D, Gordon M, Pelhate MF, Martin, Rochat H (1985) An excitatory and a depressant insect toxin from scorpion venom both affect sodium conductance and possess a common binding site. Archives of Biochemistry and Biophysics 240: 877-887.

Zoete V, Cuendet MA, Grosdidier A, Michielin O (2011) SwissParam: A fast force field generation tool for small organic molecules. Journal of Computational Chemistry 32: 2359-2368.

Rakesh Shamsunder Joshi
Senior Research Fellow
Division of Biochemical Sciences
CSIR-National Chemical Laboratory,
Pune 411008, India
Phone No. : +91-20-2590-2245/2247
Mobile No. : +91 960-433-0691
Email ID : rakeshjoshi687@gmail.com; rs.joshi@ncl.res.in
Website : http://rakeshjoshi.webs.com/

Educational Qualifications

Ph. D. (Biotechnology)	2010 to 2014	University of Pune **Working Place:** Biochemical Sciences Division, CSIR-NCL, Pune **Guide:** Dr. Ashok Giri; **Co-guides:** Dr. C. G. Suresh and Dr. Vidya Gupta **Thesis title:** Tailoring potent plant protease inhibitor against *Helicoverpa armigera* gut proteases
M. Sc. (Biotechnology)	2007 to 2009	**University of Pune; First class**
B. Sc. (Biotechnology)	2004 to 2007	**University of Pune; First class**

Work Experience

1. **UGC Junior and Senior Research Fellow:** January 2010 to Current; **Working Place:** Biochemical sciences Division, CSIR-NCL, Pune, India
2. **Teaching assistant:** 2011-2013; **Working Place:** Biotechnology Department, Modern College of ACS, Shivajinagar, Pune
3. **Post-Graduate Intern:** 2008- 2009; **Working Place:** Insect and Microbial Molecular Biology Unit, National Centre for Cell Sciences, India

Awards and Achievements

1. **Appointed as Assistant Professor in Institute of Bioinformatics & Biotechnology (IBB) and Department of Biotechnology in University of Pune in August 2014** (Success rate > 2%)
2. **Represented University of Pune at state level in " Avishkar-2013" held at NMU under Research category**
3. **Best Poster and oral presentation prize in International Symposium** on Proteomics Beyond IDs in November 2012 (In Top 3 amongst the 400 posters)
4. **Nominated for Dr. Shama Prasad Mukhrji (SPM) Fellowship**, CSIR, India in 2010 (Top 50 student Nationwide)
5. Hold All India **31st Rank in CSIR-NET** and **168th Rank in GATE** in year 2009
6. Qualified highly competitive **National Eligibly Test (CSIR-NET)** for JRF in **Dec 2008 and 2009** (Success rate ~ 5%)

Publications [Total impact: 32; Total citations: 26; H-index: 4; I_{10} index: 0; http://scholar.google.co.in/citations?hl=en&user=S2suMmsAAAA]

Articles

1. **R. S. Joshi**, R. S. Tanpure, R. K. Singh, V. S. Gupta, A. P. Giri. *Resistance through inhibition: Ectopic expression of serine protease inhibitor offers stress tolerance via delayed senescence in yeast cell,* , **Biochem. Biophys. Res. Commun. (2014)** Accepted
2. S.B. Bansode, A. K. Jana, K. B. Batkulwar, S. Warkad, **R.S. Joshi** et al. *Molecular investigations of protriptyline as a multi-target directed ligand in Alzheimer's disease* **PloS one (2014)** 9 (8), e105196
3. **R.S. Joshi**, M. Mishra, V.A. Tamhane, A. Ghosh, U. Sonavane et al. *The remarkable efficiency of a Pin-II proteinase inhibitor sans two conserved disulfide bonds is due to enhanced flexibility and hydrogen bond density in the reactive site loop***, J. Biomol. Struct. Dyn.** 32 **(2014)** 13-26.
4. **R.S. Joshi**, V.S. Gupta, A.P. Giri. *Differential antibiosis against Helicoverpa armigera exerted by distinct inhibitory repeat domains of Capsicum annuum proteinase inhibitors*, **Phytochemistry** 101 **(2014)** 16-22
5. M. Mishra, **R.S. Joshi**, S. Gaikwad, V.S. Gupta, A.P. Giri. *Structural-functional insights of single and multi-domain Capsicum annuum protease inhibitors*, **Biochem. Biophys. Res. Commun.** 430 **(2013)** 1060-1065.
6. **R.S. Joshi***, M.D. Jamdhade, M.S. Sonawane, A.P. Giri. *Resistome analysis of Mycobacterium tuberculosis: Identification of aminoglycoside 2'-Nacetyltransferase (AAC) as co-target for drug desigining*, **Bioinformation** 9 **(2013)** 174-181. **[* Corresponding Author]**
7. Y.R. Chikate, V.A. Tamhane**, R.S. Joshi**, V.S. Gupta, A.P. Giri*, Differential protease activity augments polyphagy in Helicoverpa armigera*, **Insect Mol. Biol.** 22 **(2013)** 258-272. **[IF: 2.976; Citation: 5]**
8. S.B. Bansode, A.D. Chougale, **R.S. Joshi**, A.P. Giri, S.L. Bodhankar et al. *Proteomic analysis of protease resistant proteins in the diabetic rat kidney,* **Mol. Cell Proteomics** 12 **(2013)** 228-236.
9. N.J. Pawar, V.S. Parihar, S.T. Chavan, **R. Joshi**, P.V. Joshi et al. *alpha-Geminal dihydroxymethyl piperidine and pyrrolidine iminosugars: synthesis, conformational analysis, glycosidase inhibitory activity, and molecular docking studies*, **J Org. Chem.** 77 **(2012)** 7873-7882.
10. **R. S. Joshi**, T. P. Wagh, N. Sharama, F. A. Mulani, U. Sonavane *et al.,* Way towards "Dietary Pesticides": Mechanistic insight into insecticidal action of natural phenols, **J. Agri. Food Chem.** (under review**)**
11. A. Anand, H. J. Ramesha, S. Bidekar, P. A. Singh, **R. S. Joshi** et al., *Phenylpropanoid content of Ocimum spp. is regulated by upstream hydroxylase and acyltransferase,* **Phytochemistry** (Under review)

12. N. S. Mahajan, V. Deewagan, P. R. Lomate, R. S. Joshi, M. Mishra *et al*., *Structural features of diverse Pin-II proteinase inhibitor genes from Capsicum annuum,* **Planta** (Revision submitted)
13. Y. Kolekar, **R. S. Joshi**, K. Kodam, *Biotransformation of 3,5-diaryl pyrazoline and 3,5-diaryl pyrazole by Bacillus fusiformis*, **Med. Chem. Letters** (under review)

Reviews

1. **R.S. Joshi**, M. Mishra, C.G. Suresh, V.S. Gupta, A.P. Giri. *Complementation of intramolecular interactions for structural-functional stability of plant serine proteinase inhibitors*, **Biochim. Biophys. Acta** 1830 **(2013)** 5087-5094.
2. M. Mishra*, P. R. Lomate*, **R. S. Joshi***, S. A. Punekar, V. S. Gupta *et al*., *Ecological turmoil in evolutionary dynamics of plant-insect interactions: Defense to offence*, **Oecologia** (Under review, * Equal contribution)

Book Chapter

1. N. Khandelwal, **R.S. Joshi**, V.S. Gupta, A.P. Giri. *Protease inhibitors as biopesticides: Potential and constraints,* **Biopesticides in Environment and Food Security, issues and strategies**, Editors: Koul O, Dhaliwal GS, Khokhar S, Singh R, Scientific Publishers (India), Chapter #9, pp 146-181.

Patents

1. **R.S. Joshi**, M. Mishra, V.A. Tamhane, A. Ghosh, U. Sonavane, C.G. Suresh, R. Joshi, V.S. Gupta, A.P. Giri. *Effective management of Helicoverpa armigera involves cloning of inhibitory repeat domain 9 in Pichia pastoris in specific vector, feeding inhibitory repeat domain 9 proteins to Helicoverpa armigera and calculating specific growth parameters*. **Patent Number: WO2013102937-A2; IN201200035-I1; WO2013102937-A3**

References

1.	**Dr. Ashok P. Giri**	Senior Scientist	CSIR-NCL	ap.giri@ncl.res.in
2.	**Dr. Vidya S. Gupta**	Chief Scientist and Head	CSIR-NCL	vs.gupta@ncl.res.in
3.	**Dr. C. G. Suresh**	Principal Scientist	CSIR-NCL	cg.suresh@ncl.res.in
4.	**Prof. KVR Chary**	Professor	TIFR	chary@tifr.res.in
5.	**Dr. Rajesh Kumar Singh**	Assistant General Manager and Head of Research and Development	Gennova Biopharma Ltd.	rajeshkumar.singh@gennova.co.in