Peptides for inhibition of insect proteases

Thesis Submitted to AcSIR for the Award of
the Degree of
DOCTOR OF PHILOSOPHY
In **Biological Sciences**

By
Nidhi Saikhedkar

Registration Number : 10BB13J26035

Under the guidance of
Dr. Ashok P. Giri

Biochemical Sciences Division
CSIR-National Chemical Laboratory
Pune, India.

January, 2019

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

Certificate

This is to certify that the work incorporated in this Ph. D. thesis entitled "**Peptides for inhibition of insect proteases**" submitted by **Ms. Nidhi Saikhedkar** to **Academy of Scientific and Innovative Research (AcSIR)** in the fulfillment of the requirements for the award of the Degree of **Doctor of Philosophy in Biological Sciences** embodies original research work under my supervision. I further certify that this work has not been submitted to any other University or Institution in part or full for the award of any degree or diploma. Research material obtained from other sources has been duly acknowledged in the thesis. Any text, illustration, table etc. used in the thesis from other sources, have been duly cited and acknowledged.

It is also certified that this work done by the student, under my supervision, is plagiarism free.

Nidhi

Research Student
Nidhi Saikhedkar

Research Supervisor
Dr. Ashok P. Giri
Biochemical Sciences Division,
CSIR-National Chemical Laboratory,
Pune – 411008

Date: 2/1/19
Place: CSIR-NCL, Pune

Communication Channels
NCL Level DID : 2590
NCL Board No. : +91-20-25902000
EPABX : +91-20-25893300
: +91-20-25893400
FAX
Director's Office : +91-20-25902601
COA's Office : +91-20-25902660
SPO's Office : +91-20-25902664
WEBSITE
www.ncl-india.org

Declaration

I hereby declare that the thesis entitled **"Peptides for inhibition of insect proteases"** submitted for the degree of Doctor of Philosophy to the **Academy of Scientific and Innovative Research (AcSIR)** has been carried out by me at the Biochemical Sciences Division, National Chemical Laboratory, Pune – 411 008, India, under the supervision of **Dr. Ashok P. Giri**. Any material as has been obtained by other sources has been duly acknowledged in this thesis. I declare that the present work or any part thereof has not been submitted to any other University for the award of any other degree or diploma.

Nidhi

Date : 2/1/13

CSIR-National Chemical Laboratory

Pune-411008

Research Student

Nidhi Saikhedkar

Contents

Chapter 1: Introduction

No one can whistle a symphony. It takes a whole orchestra to play it

HE Luccock

This quote summarizes all the gratefulness that I wish to express for completion of my doctoral thesis. Without a doubt, the first person to be acknowledged and given credit for this work is my research supervisor, Dr. Ashok P. Giri. He has been a constant source of encouragement though all these years, and has been exceptionally supportive during my hard times. Specifically, I remember struggling through official matters (that's another long story) in my early days of Ph.D, when he was the only person standing by my side and giving a constant assurance that everything will be sorted at the end. I must say that it was his boost up that helped me stay focused on my work, and complete my Ph.D timely. My past five years at NCL have been a happy journey, because my guide was very dependable. He introduced me to Plant molecular biology, specifically to the interesting subject of plant-insect interactions. Our group works on diverse subject areas, and this exposure helped me in gaining vast knowledge at a single place. He also taught me the skills of being a researcher, most importantly, collaborative and presentation skills. He used to say, "At the end of your Ph.D, you should be able to call yourself an independent scientist", and he really made me one. I particularly appreciate that he gave me freedom to explore science, and never said no to any idea that I suggested. His constant encouragement to work in an inter-disciplinary environment, learn a variety of techniques, and opportunity to present my work wherever possible, are skills that have given me enormous confidence which shape my career. He also used to pressurize and put strict

deadlines, which I am now thankful for, because without that attitude I would have slipped into the “comfort zone”, which is deadly for a researcher. He has worked hard with me during manuscript preparations, my presentations, and I would always try to reciprocate all the things that I have learnt from him in my future. In all these years, we have made a bond for life, and I can say with pride that today I am able to write this acknowledgement, only because of him.

I am grateful to Dr. Moneesha Fernandes, who has been like my co-guide (though not officially), during my doctoral work. She made the work of peptide synthesis feasible for me, which was a completely new experience. Being a biotechnology student, I was inspired by her constant motivation and encouragement to work in a chemistry lab.

I am extremely thankful to Dr. Kiran Kulkarni, because he introduced me to the fascinating world of X-ray crystallography. He is an excellent teacher, and it is always exciting to work with him, because he made me actually “see” the peptides. It was only because of his efforts and encouragement that I was able to learn this new area and got a chance to work on crystals.

I express my heartfelt gratitude to Dr. Vidya Gupta, who has been a constant source of inspiration during my stay at NCL. I also thank Dr. Narendra Kadoo, who has always been available for discussion, and was continuously supportive of my research work.

I would also like to thank my DAC members, Dr. G.J. Sanjayan, Dr. Kiran Kulkarni, Dr. Moneesha Fernandes, and also my external examiners, Dr. Janesh Kumar (NCCS, Pune) and Dr. Kisan M. Kodam (SPPU, Pune) for their timely suggestions and evaluation of my research work. My gratitude towards scientists from NCL, Dr. Mahesh Kulkarni, for helping in Mass spectrometric analysis, Dr. Sushma Gaikwad, for facilitating the work on HPLC, Dr. B. Santhakumari, who allowed to work on HR-MS, Dr. UdaykiranMareilli who taught me NMR data analysis. Also, I would like to express my sincere thanks to Dr. Ashwini Kumar Nangia, Director, CSIR-NCL, and past directors, Dr. Sourav Pal, Dr. K. Vijayamohanan Pillai, for giving an opportunity to work at this esteemed institute. I am also grateful to HOD, Biochemical Sciences division, for providing the lab infrastructure and facilities.

I wish to express my gratitude to Dr. Rakesh Joshi, who has been my support system during all these years, not just scientifically, but as a friend also. Being my senior, I continued his doctoral work, and his constant piece of advice helped me stay on track of work. He has always been available for discussion, experimental design, manuscript corrections, and what not. He has been a motivator when I was filled with frustration, and a companion for party during my success. Together, we made a great team, and I will always cherish his guidance.

I would like to specifically thank Ms. Smrati Sanghi, who has been my first and best friend at NCL, and has been a companion for me in all my good and bad times. She was the person who was always there around me, for sharing the little details of our lives, and I will always treasure the memories that we made while staying and working together at NCL.

I was fortunate to work closely with wonderful people in my lab, Tejas Chirmade, Ranjit Barbole, Veena Devangan, Bhakti Dayama, Gopal Kallure, Santosh Lavhale, Monika Mathur, who were partners in everything, from outing to lab meetings to everyday lunch time. I am also thankful to my friends, Ashwini Rane, Md. Saleem Dar, Sucheta Patil, Deepika Somani, Vaishnavi Salunkhe, Pranjali Oak, Amol Kasodekar, Gayatri Salunke, with whom I share special memories since the time I joined in the lab. I am grateful to lab members Sonali Gaikwad, Amrita Ghosh, DivyaKarade, Uma Jadhav, Tanuja Gurav, Sagar Yadav, Gauri Pathak, Ravindra Vidhate, Dr. Saborni, Dr. Ashwathy Nair, Dr. Sachin Gharat for their direct or indirect help in conducting my research work. I also thank the project partners with whom I have worked and shared lab bench and experience, Radhika Mohandasan, Aarohi Summanwar, Preeti Ranawde, Trushna Wagh and Priyanka Jawale.

I am especially thankful to my seniors, Dr. Vishal Dawkar, Dr. Purushottam Lomate, Dr. Ashwini Rajwade, Dr. Sheon Mary, Dr. Rahul Tanpure, Dr. Neha Mahajan, Dr. Neha Khandelwal, Dr. Yojana Chikate, Dr. Amey Bhide, Dr. Ashish Deshpande for their guidance in handling the projects. It was because of these people that I learnt good laboratory practices and etiquettes of working together. I would like to specially mention Mr. Ajit Ranjekar, who glued me to all the PMB members, and whose greeting cards always made my special occasions very memorable.

I also had the opportunity to work in Organic chemistry lab for peptide synthesis. I am thankful to have worked with Dr. Amit Yadav, with whom I had one of the best collaborative experience. He has taught me the basics of working with peptides. I am also deeply thankful to Dr. Govind Bhosale and Harsha Chilukuri who taught me the skills needed to work in a chemistry lab. I also thank my friends Ragini Awachat, Komal, Atish Wagh, Manisha Aher who were so kind and generous, and with whom I share memories of endless discussions while working in the lab.

I would also like to thank people from BL21 beamline for protein crystallography, RRCAT, Indore, especially Dr. Biplab Ghosh and Dr. Ravindra Makde for hosting me and showing the exciting area of synchrotrons and X ray diffraction. Also, I have been warmly welcomed at GE Heathcare, Bengaluru, by Dr. Shubendu Seal, Dr. Abhishek Chaudhary, Dr. Mosami Galvankar and Mr. Bruno Thomas, where they trained me on the basics of SPR and were always available for designing and execution of experiments.

This thesis has also been a significant contribution from my trainees, Bhavana, Geeta Bora, Madhura, Komal, Madhuri, Sweta. They have worked immensely with me and generated important pieces of data. Also, I am thankful to students at NCL, Bhushan Burkul, for assisting in HRMS data acquisition; Rajeshwari Rathore, for helping in Synapt MS; Fayaz Mulani, for helping in HPLC runs; Anand Sandholu, for experiments with crystals. Without them, it would not have been possible to work in a friendly and collaborative environment.

Finally, my deepest gratitude for my parents, who had the courage to send me away from them to pursue my goals. It was their support and patience, because of which I was able to work hard here. They always encouraged me to be positive in all the difficult situations. Their cheerful face kept me motivated to achieve the best in life. They have been assertive of all my decisions, and havegiven me the strength to stay focused. I thank them for making me an independent person.

Certainly, several people have helped to complete this work. I thank the ones whose contribution I may not have mentioned here.

Nidhi Saikhedkar

List of tables

List of figures

AU	Absorbance Units
Ac	Acetate
ACN	Acetonitrile
BApNA	Benzoyl-DL-Arginyl-P-Nitroanilide
Boc	Di-tert butyl carbonate
Bt	Bacillus Thuringiensis
Can	Canavanine
CanPI	Capsicum Annuum Proteinase Inhibitor
DBMB	α,α′-Dibromo-m-xylene
DCM	Dichloromethane
DMF	N,N-dimethylformamide
DIPCDI	N,N'Diisopropylcarbodiimide
DIPEA/DIEA	Diisopropylethylamine
DMSO	N,N-Dimethyl sulfoxide
EDTA	Ethylenediaminetetraacetic acid
Fmoc	Fluorenylmethyloxycarbonyl
h	Hours
HGP	*Helicoverpa armigera* Gut Proteinases
HOBt	Hydroxy benztriazole
HPLC	High Performance Liquid Chromatography
hr	Homoarginine
HRMS	High Resolution Mass Spectrometry
IC_{50}	Inhibitor Concentration At 50% Inhibition
IRD	Inhibitory Repeat Domain
kDa/kD	Kilo Dalton
Ki	Inhibition Constant
M	Molar
MALDI-TOF	Matrix Assisted Laser Desorption Ionisation-Time of Flight
MBHA	4-(methyl)benzhydrylamine-resin
mg	milligram

min	Minutes
mL	Millilitre
mM	millimolar
mmol	millimoles
MS	Mass spectrometry
MW	Molecular weight
µL	Microliter
µM	Micromolar
nm	Nanometer
orn	Ornithine
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
PI	Proteinase Inhibitor
PIN-I/II	Potato Type Inihibitor I And II
RCL	Reactive Center Loop
RP-HPLC	Reversed Phase-HPLC
RT-PCR	Reverse Transcriptase-Polymerase Chain Reaction
TBTU	O-(Benzotriazol-1-yl)-N,N,N',N'-tetramethyluronium tetrafluoroborate
TBMB	1,3,5-Tris(bromomethyl)benzene
TFA	Trifluoroacetic Acid
TFMSA	Trifluoromethanesulfonic acid

Synopsis

Chapter 1: Introduction

Infestation by insects is one of the major biotic factor affecting crop productivity worldwide. Lepidopteran insects like *Helicoverpa armigera* are important agricultural pests. Characteristics like polyphagy (host plants), high fecundity (300-1000 eggs per female per life cycle) and high migratory ability (>100 km per insect) makes them one of the most devastating insect species. Currently, both chemical and biological methods are used for control of this pest. However, chemical insecticides are toxic to non-target organisms and the environment. On the other hand, use of biological methods in the form of formulation or transgenic plants has led to development of resistance in insects. Therefore, it is important to develop alternative methods of controlling this pest. One of the promising methods for biological control of insects is the targeted use of plant protease inhibitors (PIs) as antifeedant molecules. The mechanism of action of PIs is based on inhibition of proteases present in insect gut. The lepidopteran midgut consists of serine proteases, mainly trypsin and chymotrypsin-like enzymes. Therefore, inhibition of insect serine proteases by plant PIs will cause reduction in the availability of amino acids necessary for insect growth and development.

Among the various types of plant PIs, Potato inhibitor-II (Pin-II) family of serine PIs is well known for defense against plant pathogens and insects. Pin-II PIs are composed of multiple inhibitory domains joined together by linker regions. Proteolytic processing at the linker region generates 55 amino acid long functional units called "inhibitory repeat domains" (IRDs). Each IRD has a tripeptide reactive center loop (RCL), which extends as a solvent exposed region from the remaining IRD scaffold. The scaffold is characterized by presence of 6-8 conserved cysteines which form disulphide linkages, holding the RCL in proper orientation for interacting with the target protease. The tripeptide RCL region of IRDs forms crucial interactions with target proteases. Hence, we postulated that the RCL tripeptide sequence might be the minimum requirement for inhibitory activity. In this work, we have studied the inhibitory mechanism of RCL tripeptides independent of the IRD scaffold.

Chapter 2: Characterization of tripeptides derived from reactive center loops of potato type II protease inhibitors

Among the 389 Pin-II PI proteins analyzed, 237 unique IRD sequences were present. Further study of the reactive center loop region indicated that there are 21 RCL variants distributed across the

Pin-II family. Out of these 21 RCLs, six sequences (TRE, PRN, PRY, PKN, PLN and TLN) are predominant. These six peptides were synthesized by Fmoc-solid phase synthesis and evaluated for *in vitro* inhibition of serine proteases. The RCL peptides TRE, PRN and PRY showed remarkable inhibition of serine proteases in the *H. armigera* gut extract (up to 95% reduction) with IC_{50} values ranging from 50 to 200 µM for trypsin and chymotrypsin-like proteases. These peptides adversely affected the growth and development of *H. armigera*. They showed enhanced inhibitory activity at alkaline midgut pH, retention and stability in insect gut. Gene expression analysis of representative trypsin and chymotrypsin genes by qRT-PCR revealed that majority of these genes were downregulated in response to insect feeding. Further, the binding mechanism and differential affinity of the RCL peptides with serine proteases was delineated by crystal structures of complexes of the RCL peptides with trypsin. Residues P1 and P2 of the inhibitors play a crucial role in the interaction and specificity of these inhibitors. P1 Arg was primarily involved in H-bond with active site pocket of trypsin, while P2 residue stabilized the interaction by Van der Waals interaction with Leu89 in trypsin. These attributes of RCL peptides make them suitable candidates for the engineering of new molecules for crop protection.

Chapter 3: Tailoring of reactive loop peptides by cyclization for inhibition of lepidopteran serine proteases

Linear RCL peptides are known to inhibit insect midgut proteases and cause growth reduction in lepidopteran pests. We speculated that since the RCL in native Pin-II protein is held by disulphide bridges, it is possible that cyclization of tripeptides using the flanking Cys residues might enhance their inhibitory potential. Bi-cyclization with two RCL peptides was performed on a mesitylene scaffold, envisaging an enhancement in potency. Similarly, monocyclic peptides were prepared using dimethyl benzene scaffold, or disulphide linkages. Selected peptides were synthesized and screened for in vitro inhibition of insect trypsins. Bicyclic peptides were ten times more potent than linear RCLs, which was validated by surface plasmon resonance based affinity kinetics. Also, bicyclic peptide treated leaves were less prone to herbivory by lepidopteran pests, namely, *H. armigera* and *Spodoptera litura*. Besides, severe reduction in larval growth was observed upon obligatory feeding. Target identification by pull down assays and gene expression analysis revealed that, in addition to serine proteases, peptides bind to antioxidant enzymes, which supplements their *in vivo* effectiveness. Binding modes predicted by docking and molecular

dynamics simulations showed that bicyclic peptides interact with the active site of proteases, as well as specificity determining residues, which might be the reason for enhanced activity.

Chapter 4: Design and evaluation of RCL peptidomimetics

Structural analysis has revealed that P1 Arg in RCL peptides makes maximum number of contacts with active site residues in target protease. Thus, Arg was selected for development of modified of RCL peptidomimetics. Most potent tripeptide, PRN was modified in order to improve potency and functionality against insect gut proteases. Several modifications were incorporated, like (a) the peptide bond was replaced with carbamate bond in order to enhance stability, (b) Arginine side chain was elongated by addition of -CH_2, in order to increase penetration into the active site. Docking of modified peptides in the active site of *H. armigera* trypsin indicated that the peptides show enhanced binding. The modified peptides showed better reactivity than the parental peptides, as evaluated by biochemical and biophysical studies. The modified peptides were stable to proteolysis, suggesting that the modifications have been helpful in enhancing the affinity and stability of peptides with proteases.

Chapter 5: Summary and future prospects

This work is the first report demonstrating that the RCL regions of Pin-II type PIs inhibit serine proteases independent of the parent IRD scaffold. Crystal structures of tripeptides in complex with trypsin unraveled their binding modes, highlighting the basis of differential reactivity towards proteases. RCL tripeptides were characterized for their in vitro and in vitro efficacy against lepidopteran serine proteases. Further, grafting of the RCLs onto mesitylene scaffold generated bicyclic peptides, which showed enhanced inhibition potency as compared to linear RCL tripeptides. Apart from protease inhibition, these peptides show insect deterrent activity and inhibition of antioxidant enzymes. Furthermore, design of peptidomimetics based on RCL regions led to generation of more potent and functionally diverse molecules.

This study shows that diverse peptide combinations can be generated to tackle with diverse insect proteases, making RCL peptides as potential pest control agents. Being small peptides, they are east to synthesize and formulate, making them promising molecules for agricultural applications.

Organization of thesis

The thesis is organized into five chapters, the contents of which are as described here:

Chapter 1: Introduction

This chapter gives detailed description of the topics underlying the genesis of thesis. The mechanism of plant-insect interaction is explained, followed by various defence strategies adapted by plants, with specific emphasis on plant protease inhibitors. It is followed by description of plant protease inhibitor families, and their use for control of insect pests. Specifically, a detailed account of structure and function of Pin-II family of inhibitors is presented, along with their effects on lepidopteran pests. Further, the role of reactive center loops of Pin-II inhibitors in interaction with target protease is emphasized, and the use of small peptides as pest control molecules is proposed. Thereafter, various natural and synthetic scaffolds are evaluated for grafting of peptides. Finally, some of the methods used for development of peptidomimetics are described, especially the use of plant derived un-natural amino acids and peptide bond modifications. These techniques could be used for development of dietary protease inhibitors based on the reactive loop of Pin-II type inhibitors.

Chapter 2: Characterization of tripeptides derived from reactive centre loops of potato type II protease inhibitors

In this chapter, the biochemical and structural characterization of tripeptide reactive centre loops derived from Pin-II type protease inhibitors is shown. The synthesis and purification of six reactive loop tripeptides, followed by *in vitro* biochemical assays for determination of protease inhibition is described. Further, the negative effects on growth and development upon feeding of tripeptides on insect, *Helicoverpa armigera* is demonstrated. The mode of binding of peptides to protease is elucidated by the structural determination of protease-peptide complex by X-ray crystallography.

Chapter 3: Tailoring of reactive loop peptides by cyclization for inhibition of lepidopteran serine proteases

In continuation with Chapter 2, the design of cyclic peptides based on the reactive loop peptides is presented. In order to enhance the potency of linear tripeptides, bicyclic peptides were generated by cyclization of two reactive loop tripeptides on mesitylene scaffold. Virtual screening and

synthesis of bicyclic peptides is shown, followed by the biochemical activity screening. Further, affinity and kinetic screening of the peptides by surface plasmon resonance is shown. *In vivo* target identification and feeding choice assays were performed with two lepidopteran pest species. The binding modes of peptides were predicted by molecular dynamics simulations, which showed the possible mechanism for increase in activity.

Chapter 4: Design and evaluation of RCL peptidomimetics

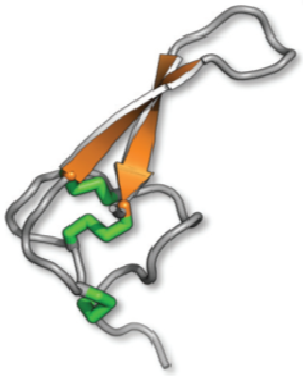
This chapter describes the design of modified peptides based on the reactive center loop tripeptides. Arginine analogue, homoarginine is used to replace the arginine of tripeptide. Also, replacement of amide backbone by carbamate groups is shown. Evaluation of these peptidominetics is described by means of biochemical activity assays and biophysical characterization by surface plasmon resonance. Further, the proteolytic stability of modified peptides is studied in comparison with parent tripeptides.

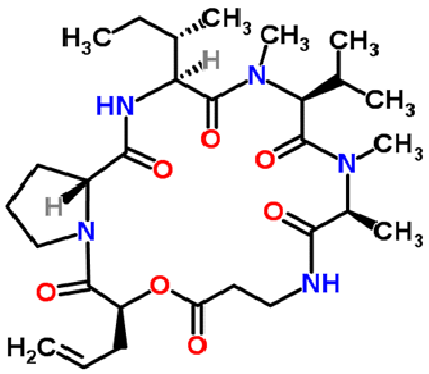
Chapter 5: Summary and future prospects

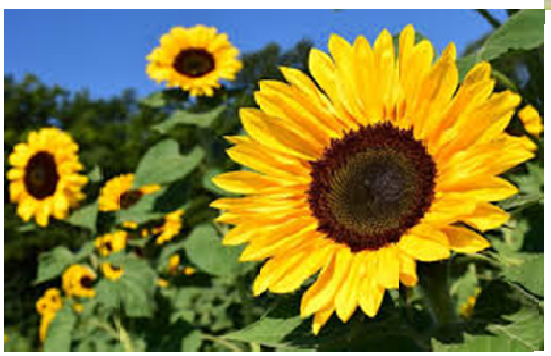
This chapter highlights the important findings of this work, and the possible future avenues generated by this research.

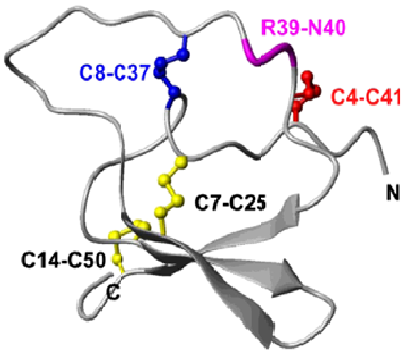
Chapter 1

Introduction









"God's Experiments Created Nature, Natural Variants Lets Explore Its Perfection…"

1.1 Plant-insect interaction

Plants and insects are known to interact with each other since millions of years. During this association, they have co-evolved depending on the surrounding conditions like climate, geography and species abundance/distribution(Mishra et al., 2015). They have been sharing a mutualistic association by maintaining a service-resource relationship(Figure 1.1). Several such interactions exist in nature, for example, plants are pollinated by bees, which in return receive food in form of nectar. Other plant products like resin, gum and wax are used by the insects to build their hives(Koch, 2008). Plant parts also serve as shelters for reproduction of insects. These associations have resulted in development of specific morphological and physiological features in plants, and correspondingly with insects, which indicates their co-evolution. However, deception behaviours are also exhibited by both the partners, which tend to exploit the mutualism. Like, certain carpenter bees (called nectar thieves) enter the flower and feed on nectar, but do not pollinate in return(Bronstein et al., 2006). Likewise, orchid flowers exhibit fake floral signals like scent, inflorescence and color in absence of nectar to attract the insects for pollination(Jersáková et al., 2006). Thus, not all plant-insect interactions are mutualistic. Herbivorous insects are known to feed on plant parts, like leaves and fruits, which are beneficial to the plants as well as for human consumption. Such insects have been termed as "pests" owing to the huge agricultural yield losses caused by them. In parallel, plants have also evolved to defend themselves against pests by means of morphological and physiological adaptations. Altogether, there exists a 'molecular arms race' between the two taxa, which governs the plant-insect interactions.

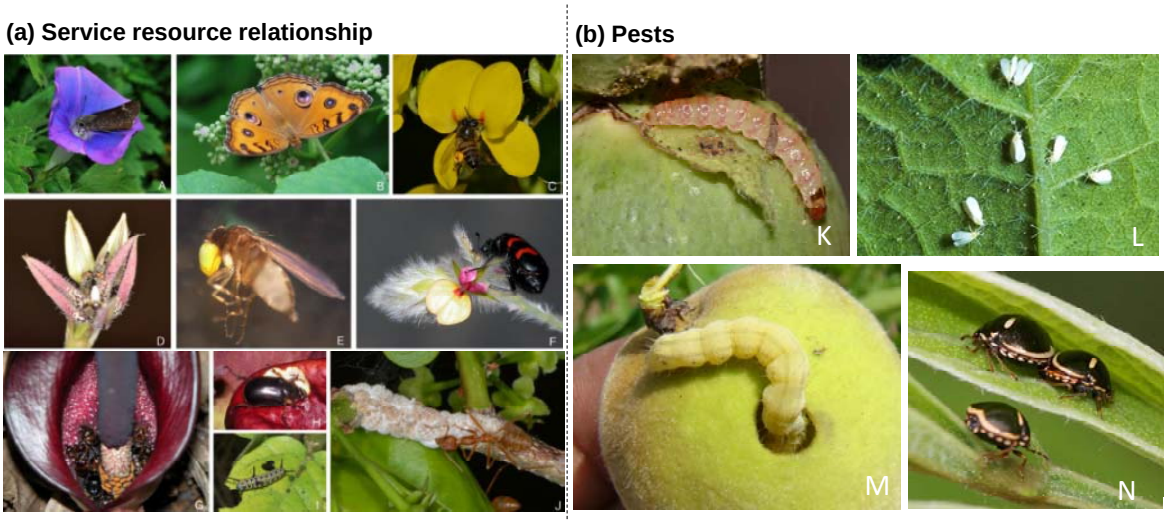


Figure 1.1: **Myriad of plant insect interactions**. (a) Mutualistic associations between plants

and insects (Mishra et al., 2015): A. Rice Swift (*Borbo cinnara*), a skipper butterfly feeding on Ipomoea; B. Peacock pansy (*Junonia almana*) butterfly feeding on Leea indica; C. Honey bee (*Apiscerana indica*) feeding on *Smithia setulosa*; D. Fly pollination in *Brachystelma malwanense*; E. Fly with *Ceropegia pubescens* pollinarium; F. Banded blister beetle (*Mylabris pustulata*) feeding on *Alysicarpus pubescens*; G. Scarab Beetles (Onthophagus sp.) pollination in *Amorphophallus commutatus* var. anshiensis; H. Root grub beetle (Rutelinae) feeding on stinkyappendage; I. Plain tiger (*Danaus chrysippus*) butterfly caterpillar feeding on leaves of *Ceropegia maharashtrensis*; J. Red tree ants (*Oecophylla smaragdina*) harvesting honey dew from mealybugs; (b) Pest-plant associations (source: http://www.nbair.res.in/insectpests/): K. Pink bollworm (*Pectinophora gossypiella*) larvae feeding on cotton boll; L: Whitefly (*Trialeurodes vaporariorum*) feeding of lettuce leaves; M: Cotton bollworm (*Helicoverpa armigera*) infestation on peach; N: Bruchid (*Coptosoma variegate*) infesting red gram leaves

1.1.1 Plant defense mechanisms

To survive in presence of herbivores, plants have evolved a range of mechanisms, which involve morphological and biochemical adaptations. Physical barriers, like thorns, protect the plants by causing injury to the predator. Other morphological changes include secretion of chemicals on plant surface like resins, lignins and wax, which cause feeding avoidance in insects(Fernandes, 1994). An interesting mechanism adapted by plants is “calling enemy of the enemy”. Many plants produce volatiles to attract the natural enemies of herbivore, which may be predators or parasitoids. For example, herbivory induced plant volatile, indole, increases the recruitment of parasitoid, *Micropitis rufiventris* on maize plants infected with *Spodoptera litura* larvae(Ye et al., 2018). The emission of plant volatiles is specific and governs the complex tritrophic interactions with the insect species(Clavijo McCormick et al., 2012).

The second line of defense comprises of phytochemicals or secondary metabolites targeting the insect pests. These compounds act by causing feeding aversion or by reducing the feeding of insects by affecting the physiological processes. Plant secondary metabolites, for example terpeniods (β-farnesine), alkaloids (Nicotine), glucosinolates (Sinigrin) act as repellents or deterrents for insects, and thus protect the plants. For instance, the volatiles emitted by plants attract insect predators, and also alert neighbouring tissues(Aartsma et al., 2017). Similarly, upon biotic stress, plant divert defensive compounds towards damaged tissues, and modify the flux of metabolites(Huot et al., 2014; Moghe and Last, 2015). Trichomes possess a large number of defense related secondary metabolites. In addition to these, plants produce secondary metabolites in form of polypeptides (Chitinases, protease inhibitors) or amino acids (canavanine, γ-aminobutyric acid) which are toxic to the insects(Mithöfer and Boland, 2012). These proteins

adversely affect the physiological processes of the insect, which hamper their growth and reproduction. Hence, a diverse and complex range of plant defense compounds have evolved in response to insect pests(Schuman and Baldwin, 2016). However, insects have also simultaneously devised adaptations to resist the plant defense mechanisms.

1.1.2 Insect adaptation to plant defense

In order to deal with the plant defense tactics, insects also have evolved mechanisms to avoid or detoxify the plant metabolites(Després et al., 2007; Zhu-Salzman and Zeng, 2015). The adaptive mechanisms include manipulation of host plants to gain access to resources, for example, caterpillars roll mature leaves surrounding the growing buds of *Psychotria horizontalis*(Sagers, 1992). Gall formation is another strategy, which ensures nutrition from tissues inside the gall, and also protection from predators(Weis and Kapelinski, 1994; Zhao et al., 2015). Insects also show behavioral avoidance towards plants containing toxic compounds, thus leading to diversification of host plant range(Janz et al., 2006; Rashid War et al., 2018). Also, insects tend to sequester the plant defense compounds and later use them against predators(Després et al., 2007). Cytochrome p450 oxidase mediated the sequestration of nicotine in *Manduca sexta* larvae(Morris, 1983; Stevens et al., 2000). Other enzymes such as glutathione S transferases or esterases detoxify plant metabolites(Matthews et al., 1990; Z.-B. Xu et al., 2015; Yu and Powles, 2014). These chemical adaptations often depend on the type of plant defense molecule and its mode of action(Mithöfer and Boland, 2012). Like, insects have evolved egg laying as a mechanism to suppress plant defense. Egg laying leads to salicylic acid accumulation, which interferes with the jasmonic assay pathway upon herbivory(Bruessow et al., 2010). Furthermore, several biochemical strategies have also been developed by insects to cope up with plant defense molecules. One of the most important strategies is the regulation of enzymes, which counteract or degrade the plant metabolites. For example, in response to plant protease inhibitors, insects regulate the expression of their protease isoforms(Lomate et al., 2018; Zhu-Salzman and Zeng, 2015). They are also known to adapt to dietary protease inhibitors by expressing isoforms with altered specificity or ones which degrade the ingested inhibitors(Dunse et al., 2010a; Sarate et al., 2012; Srinivasan et al., 2005). Thus, there is a continuous cycle of coevolution between plants and insects, resulting in generation of enormous diversity of adaptive mechanisms and secondary metabolites(Becerra, 2015)(Figure 1.2). This adaptive nature of the insects has enabled them to gain resistance to various pest control agents(Dawkar et al., 2013). It

is therefore necessary to understand the physiology of target pests in order to develop novel insect control molecules.

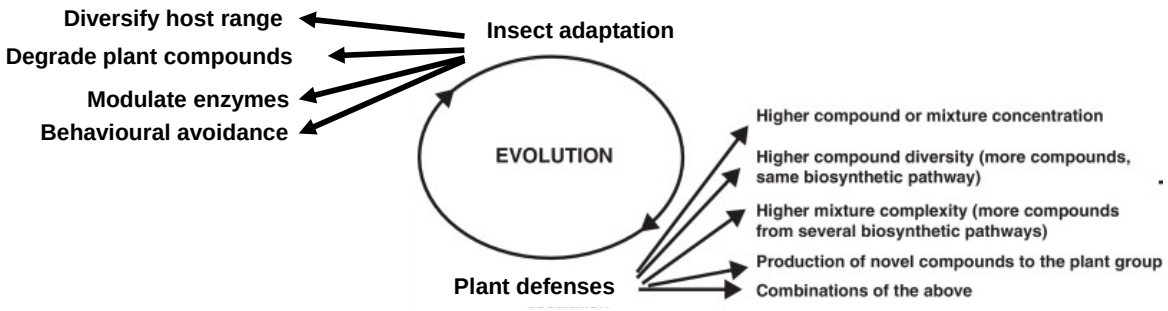


Figure 1.2: Co-evolution of insect adaptations with plant defenses. (Becerra, 2015)

1.2 Lepidopteran pests: Major threat to agriculture

The order lepidoptera is composed of 126 families of insects (moths and butterflies). These insects are associated with plants by herbivory, which, in some cases, is mutualistic, while parasitic in others. Of the 126 families, insects from Tortricidae, Noctuidiae and Pyralidae families are considered to be major pests worldwide. Lepidopteran pests possess characteristics such as polyphagy (diverse host plant range), high migratory ability, and high fecundity, which makes them severe pests of agricultural crops(Dawkar et al., 2013). Since herbivory is their principal feeding method, the digestive system is the most important interface between the insect and plants. The development of insects takes place upon breakdown of the ingested complex plant material. For this purpose, several enzymes like amylases, lipases and proteases are secreted in the insect gut(Capinera, 2008). Protein digestion in the insect gut is carried out by serine endoproteases, mainly, trypsin and chymotrypsin like enzymes, along with several exoproteases(Srinivasan et al., 2006). The number and type of isoforms expressed in the insect gut varies according to diet, developmental stage of insects and also in response to plant protease inhibitors (PIs) or plant metabolites(Chikate et al., 2013; Dawkar et al., 2013; Jongsma and Bolter, 1997). Therefore, a dynamic interaction occurs between plant proteins and insect proteases which leads to generation of numerous isoforms of insect proteases. This is one of the adaptive mechanisms, which allows lepidopteran insects to digest their complex diets. However, this highly flexible nature of lepidopteran gut physiology has also led to the emergence of

resistance to insect control agents. In this work, we have focused on one of the important lepidopteran pest species, *Helicoverpa armigera*, which is responsible for huge economic losses, owing to all the adaptive characteristics mentioned above.

1.2.1 *Helicoverpa armigera*

Helicoverpa armigera, commonly known as cotton bollworm, belongs to the Noctuidiae family of lepidoptera. It is a polyphagous insect, which infests more than 300 plant species(Akanksha, 2018; Zehr, 2004). It is prevalent on cotton, legumes, tomato, okra, etc. The larvae of this insect feed on the reproductive structures and economically important plant parts like seeds, fruits, pods which is the main cause of reduction in crop yield. The characteristics like polyphagy, high fecundity, excellent migratory capability and facultative diapause makes it one of the most devastating pests. The life cycle of *H.armigera* completes in 4-6 weeks from egg to adult during summer, and 8-12 weeks in autumn(Figure 1.3).Egg hatching takes place in about three days during warm weather (25°C average) or 6-10 days in cold weather. The eggs change color from white to black, and then produce a neonate. Larvae are 1-1.5 mm long in neonate stage, with a brownish black head and yellowish-white body. From this stage, there are six growth stages (instars) of the larvae, post to which they are fully grown. This takes 2-3 weeks to complete. At higher temperatures (38°C), the development is faster. Fully grown sixth instar larvae are 4-5 cm long and show variable colors and markings.

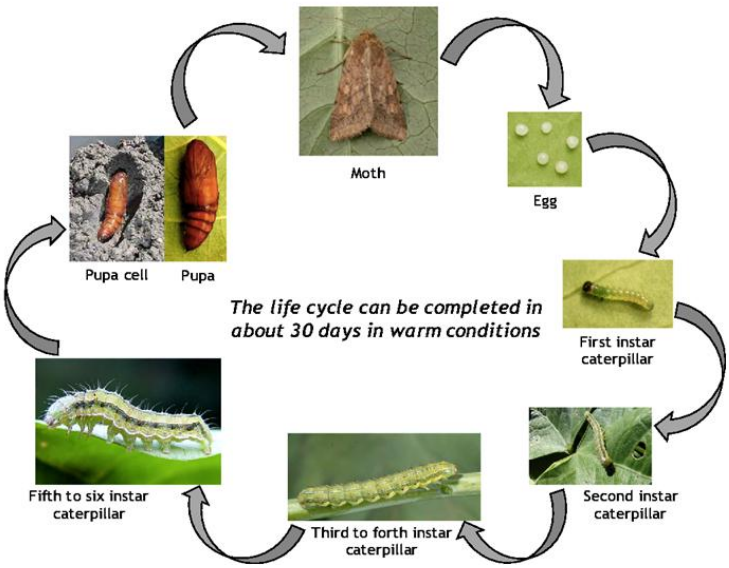


Figure 1.3: Life cycle of *H. armigera* (extracted from http:// helicoverpaaspests.weebly.com/ life-cycles.html)

Third instar larvae are the most voracious feeders, and cause majority of the damage. Therefore, it is best to target the larvae when they are in initial stages of development. The small larvae also infest the reproductive structures (flowers, squares), just upon hatching, for example in mung beans and cotton. Upon establishment in these concealed locations, larvae less prone to control with insecticides.

After feeding, the sixth instar larvae undergo pupation, by forming a tunnel up to 10 cm into the soil. Pupation takes two weeks in summer and more time in spring and autumn. This insect also shows facultative diapause (overwintering) upon exposure to adverse conditions and takes much longer to emerge. After pupation, the insect undergoes metamorphosis to form adult (moth). Adult moth wingspan is 4-5 cm. In females, the forewings are reddish-brown while they are yellow or light brown in males. Moths feed on nectar and have lifespan of around 10 days. An adult female lays about 1000 eggs on host plant parts, like leaves, flowers and developing fruits.

1.2.2. Digestive physiology of *H. armigera*

The digestive system of *H. armigera* is divided into foregut, midgut and hindgut (Figure 1.4). Majority of digestion takes place in the midgut, where the digestive enzymes are secreted into the lumen. For digestion of complex food materials, various enzymes, like lipases, amylases and proteases are secreted(Chikate et al., 2013; Kotkar et al., 2012; Terra and Ferreira, 1994). Physiology of the lepidopteran midgut suggests that serine proteases, mainly trypsin and chymotrypsin, comprise 95% of the total midgut protease complement. These proteases digest the ingested proteins into oligopeptides, which are then cleaved by exopeptidases to yield free amino acids. Till date, 32 trypsins, 22 chymotrypsin and few elastase genes have been identified(Chikate et al., 2013). These enzymes belong to the S1 family of proteases and are active at alkaline midgut pH of ~12. These proteases are differentially expressed in response to diet composition, explaining the polyphagous behavior of *H.armigera*. The insect is able to upregulate several protease isoforms in response to plant protease inhibitors, showing their adaptability(Bown et al., 1997; Jongsma et al., 1995; Kuwar et al., 2015). Moreover, several novel isoforms which could degrade the plant defense molecules are also expressed. Studies on the midgut protease properties have revealed that the surface hydrophobicity of proteases is altered in the course of evolution, suggesting it as an adaptive mechanism to inhibitors(Tamaki and Terra, 2015).

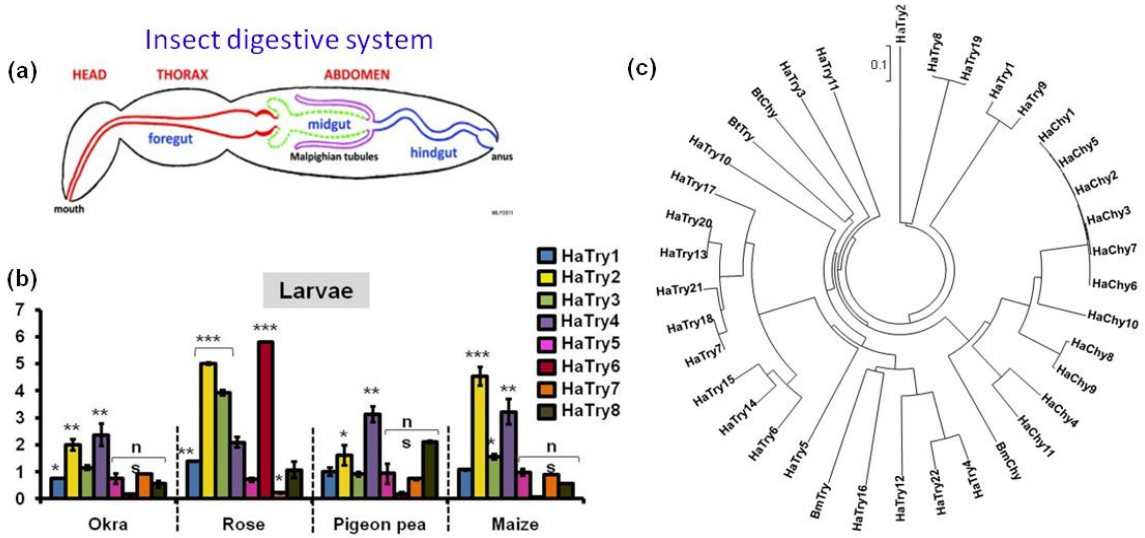


Figure 1.4: Digestive physiology of *H. armigera*. (a) The digestive system of *H. armigera*; (b) Differential expression of trypsin-like proteases in larval gut (c) Trypsin and chymotrypsin diversity in *H. armigera*

1.3. Control of *H. armigera*

There are several strategies used for the management of *H. armigera,* which have shown promising results for decreasing the yield losses. Various established practices as well as newly emerging methods are listed below:

1.3.1 Cultural practices

These practices involve procedures like deep ploughing of soil, hand picking of larvae, shaking of plants etc. Also, good agricultural methods like weeding, intercropping, use of trap crops, appropriate sowing time and use of fertilizers are considered as beneficial for control of insect pests(Sharma and et al, 2005). These practices are environment friendly and economical, however, they are labor intensive.

1.3.2 Chemical methods

Among the chemical methods, insecticides like organophosphates, pyrethroids and carbamates have been predominantly used. Their mode of action is based on paralysis of insects by acting on the nervous system (knock down effect)(Dias et al., 2015; International Program on Chemical Safety., 1986). However, the strong selection pressure exerted by these chemicals led

to development of resistance in very short span of time(Hussain et al., 2015; Yang et al., 2013). Thus, the alternate application of these insecticides in every season was suggested. A major drawback associated with chemical methods is that they are toxic to the environment as well as to other beneficial organisms(Barr and Buckley, 2011; Dias et al., 2015). Hence, the focus has now shifted to natural insect control agents.

1.3.3 Biological methods

These methods involve several approaches based on the natural plant defense mechanisms. For example, use of the natural predators of the target pest (calling enemy of the enemy). These include bio pesticides, in form of fungi, viruses, nematodes or predators(Lacey et al., 2015). These methods are gaining popularity because of their eco-friendly nature. For instance, entomopathogenic fungi such as *Beauveria bassiana* and *Metarhizium anisopliae*, are applied as natural enemies of *H. armigera*(Kulkarni et al., 2008; Qayyum et al., 2015; Younas et al., 2017). Similarly, Nuclear Polyhedrosis Virus (NPV) sprays are used as control agents(Mensah et al., 2007; Rabindra et al., 1992). Another approach includes formulations based on plant secondary metabolites like neem oil for repelling insects(Packiam et al., 2014; War et al., 2014). Similarly, insecticidal proteins from bacteria, mainly *Bacillus thuringiensis* (Bt) have also been very popular(Singh et al., 2007). However, the effectiveness of these methods is limited, due to concerns with the stability of active molecules. Also, the production of biologically active molecules on a large scale is difficult.

1.3.4 Biotechnological methods

These methods are an advancement of biological methods, which use recombinant DNA technology to improve the resistance of plants to herbivorous insects. In this direction, Bt toxin gene is most well-known example for generation of transgenic plants which are resistant to lepidopteran pests(Barton et al., 1987; Trtikova et al., 2015). Bt produces insecticidal crystal proteins (Cry) which leads to pore formation in the insect midgut epithelium and therefore death of insects. Bt transgenics were first developed for maize, potato and cotton plants and was credited for remarkable increase in the productivity of these crops(Downes et al., 2007; Naimov et al., 2002). However, the lepidopteran insects have evolved resistance to the Bt toxin by undergoing a mutation in the Bt toxin receptor site. This has led to search of other insecticidal

proteins like Bt genes, which could be used for commercial transgenic crops(Dovrat and Aharoni, 2016; Downes et al., 2007). Examples of such proteins include vegetative insecticidal proteins (Vip), which act in a similar way as Bt toxin, and have shown to be effective against lepidopteran insects like *Spodoptera frugiperida*(Palma et al., 2012). Transgenics containing other insect control molecules have also been developed, like biotin binding molecules (avidin)(Burgess et al., 2002; Murray et al., 2010), which act by depriving the insects of essential vitamins and hence hampering the development. Also, transgenic plants containing enzymes like chitinase(Fitches et al., 2004; Osman et al., 2016), cholesterol oxidase(Corbin et al., 2001) or lipoxygenase(Yan et al., 2013) have been developed and proved to be effective in control of lepidopteran pests. Transgenic rapeseed plants expressing chitinase from *M. sexta* reduced the growth of *Plutella maculipenis* larvae(Wang et al., 2005). Furthermore, plant defense proteins like amylase and protease inhibitors have been extensively studied for use as biocontrol agents(Hartl et al., 2010; . Tamhane et al., 2012a; Tanpure et al., 2017). Alpha amylase inhibitor from *Withania somnifera* seeds was shown to adversely affect the growth of *Tribolium castaneum* larvae(Kasar et al., 2017). These methods target the insect digestive physiology, since most of the feeding is the primary route of administration. Moreover, the use of RNAi in control of insect pests has also been documented. The insect genes can be downregulated by injection or feeding of dsRNA, causing alteration in their reproductive cycle(Mamta and Rajam, 2017; Zhang et al., 2017). In spite of the lab scale success of these approaches, their applicability is limited due to lack of appropriate delivery methods (for dsRNA or unstable proteins). Also, the difficulties in acceptance of transgenics in food crops in India have pointed towards the use of alternative methods. One promising method relies on targeting the digestive serine proteases of the pests by means of inhibitors, which will cause indigestion and hence growth reduction of insects. The development of stable and effective PIs for exogenous application could open new arenas of insect control molecules.

1.4 Protease inhibitors as insect control molecules

Protease inhibitors (PIs) are well known as plant defense compounds. These are proteinaceous molecules which are found in various plant parts, and highly expressed upon wounding or herbivore attack(Jamal et al., 2013; Mithöfer and Boland, 2012). PIs bind to proteases present in the insect midgut, preventing proteolysis, hence limiting the availability of amino acids, which

are essential for growth and reproduction of the insects. The mode of action of PIs does not involve a wipe-out effect like chemical insecticides. Thus, the use of PIs will not impose strong selection pressure on insects, leading to delay in the development of resistance(Stevens et al., 2012).PIs from plants have been classified based on their specificity towards target proteases, namely, serine, cysteine, aspartic and metallo-protease. Within these classes, there are several inhibitor families which are sub classified on the basis of their sequence similarity, structural homology and expression patterns(Leo et al., 2002; Rawlings et al., 2013). Up till now, ten protease inhibitor families have been identified from plants. Detailed classification of inhibitors is provided in the MEROPS database(Rawlings et al., 2016). Since serine proteases are the predominantly occurring digestive proteases in lepidopteran insects, we will be discussing in depth about serine protease inhibitor families from plants.

1.4.1 Serine protease inhibitors from plants

In lepidopteran pests, serine proteases are the primary proteolytic enzymes that work at an alkaline pH of 9-11(Srinivasan et al., 2006). Mainly, these are trypsin and chymotrypsin like enzymes. In response to insect attack, plants induce or over express protease inhibitors, which bind to and inactivate serine proteases in the insect gut. Yet, continuous exposure to any one type of PI turns on the expression of diverse protease isoforms in the insect(Zhu-Salzman and Zeng, 2015). Thus, there is a huge diversity of PIs in plants, owing to the co-evolution with proteases(Jongsma and Beekwilder, 2011). Further, all serine protease families from plants are competitive inhibitors and work by the Laskowski mechanism of inhibition(Laskowski and Kato, 1980). In this mechanism, the inhibitor acts as a ‘limited proteolysis substrate’, and mimics the substrate by binding to the active site of protease. This conformation is called the “canonical conformation”. This interaction is mediated by a constrained loop, which projects from the surface of the inhibitor, and contains a scissile bond (reactive site). Binding of the inhibitor leads to the formation of an acyl intermediate with a high association constant. But, even if the peptide bond of the inhibitor is hydrolyzed, it does not allow completion of the reaction, since the bound inhibitor does not dissociate from the active site. The dissociation rate of product from the protease active site is very low. This results in an apparent equilibrium between the free enzyme and inhibitor on the one side and the complex on the other(Grosse-Holz and van der Hoorn, 2016). A brief description of plant serine PI families is given below:

- **Serpins (MEROPS :I4)**

  The serpin family is the most widespread and largest family of PIs. Serpin-like genes are found in most of the organisms, like viruses, bacteria, plants and animals. These inhibitors interact with serine proteases of pathogens, and are also known to aid in the immune response of plants(Fluhr et al., 2012; Silverman et al., 2001; Vercammen et al., 2006). In addition to serine proteases, serpins also show mixed specificity towards caspases and papain like proteases. For instance, wheat serpins inhibit trypsin as well as cathepsin G at overlapping reactive sites(Østergaard et al., 2000). Serpins are composed of three beta sheets, nine alpha helices and a protruding reactive loop. The mode of action of serpins involves binding to protease active site and undergoing a conformational change leading to irreversible alteration in the structure of protease(Jamal et al., 2013). Many serpins have shown effectiveness for pest control. Like, serpins from winged bean and tobacco are effective against lepidopteran pests like *Helicoverpa spp*(Roy et al., 2009). Also, trypsin and chymotrypsin inhibitors from *Solanum americanum* showed reduction in growth of *H. armigera* and *S. litura* at nanomolar concentrations(Luo et al., 2009).

- **Kunitz type inhibitors (MEROPS: I3)**

  These inhibitors are found in legumes, cereals, in solanaceous species (Jamal et al., 2013; Wang and Ng, 2006). Along with defense, they are also expressed in stress conditions. These proteins have molecular mass of 18–22 kDa; and contain a three-fold β-trefoil, accompanied by twelve anti-parallel β-strands connected with thirteen loops with Reactive Site Loops in between β-strands held by two disulfide bridges. Their mechanism of inhibition involves forming a tight complex with the target protease, which dissociates very slowly(Otlewski et al., 2001).These PIs are potential candidate for generation of pest control molecules. For example, Kunitz type inhibitors from chickpea are expressed at high levels during herbivore attack, and carry out inhibition of midgut proteases in *H. armigera*(Harsulkar et al., 1999) Similarly, trypsin inhibitor found in the seeds of *Poecilanthe parviflora*inhibits trypsinlike proteases in midguts of *Diatraea saccharalis*,*Anagasta kuehniella*, *Spodoptera frugiperda*, and *Corcyra cephalonica* larvae(Garcia et al., 2004).

- **Bowman-Birk inhibitors(MEROPS: I12)**

  The family is named after D.E. Bowman and Y. Birk, who first identified and characterized a member of this family from soybean(Birk, 1985). The inhibitors of this family are mostly found in seeds, but are also wound-inducible in leaves. BBI exhibit a characteristic 'bowtie motif', which is comprised of two-stranded antiparallel β-sheet separated by the reactive loop in each of the domains. The 30 residues forms two domains, each with an independent reactive loop(Chen et al., 1992). In dicots, these inhibitors are double headed inhibitors, which interact simultaneously with two proteases(Grosse-Holz and van der Hoorn, 2016; Qi et al., 2005). The N-terminal domain has four disulfide bridges, whereas the C-terminal domain has three. Bowman-birk type inhibitors from *Acacia senegal* have shown growth inhibitory effects on the larvae of *H. armigera*(Babu and Subrahmanyam, 2010). Being double headed and bifunctional inhibitors, this family of PIs provides a versatile scaffold for generation of potent PIs against insects.

- **Squash type inhibitors (MEROPS: I7)**

  This family of PIs is found exclusively in seeds of cucurbitaceae. They composed of 27–33 amino-acid residues and cross-linked with three disulfide bridges. The major structural motif is a distorted, triple-stranded antiparallel beta-sheet(Heitz et al., 2001; Hernandez et al., 2000). They show inhibition of trypsin, plasmin, and kallikrein, blood clotting factors: Xa and XIIa, cathepsin G, by means of canonical protease inhibition mechanism(Mahatmanto, 2015). Representative of this type of inhibitor (MCoTI-I) has been used as a scaffold for protein engineering(D'Souza et al., 2016).

- **Cereal trypsin inhibitors (MEROPS: I6)**

  These inhibitors have serine protease and/or alpha amylase activity, and are found mostly in cereals(Jamal et al., 2013). Representative of this type of inhibitors is from Ragi(Neelima Alam et al., 2001). Also, Barley trypsin/amylase inhibitor shows dual specificity and is capable of inhibiting the midgut proteases of *S. frugiperida* larvae(Alfonso et al., 1997). Transgenic expression of this gene in wheat seeds provided resistance to the lepidopteran insect, *Sitotoga cereallela*(Altpeter et al., 1999). A few other studies have also showed the potential of this type of inhibitors in generation of multifunctional inhibitors for plant

protection, like the bifunctional inhibitor from pigeonpea shows bioefficacy against *H. armigera*(Gadge et al., 2015).

- **Potato type I inhibitors (MEROPS: I13)**

  The inhibitors of this family are prevalent in plants. These are overexpressed upon wounding and follow a systemic route of transmission. These inhibitors have molecular mass of ~8 kDa and are primarily monomeric. The inhibitors in this family do not contain any disulphide bridges. These inhibitors are active against chymotrypsin and elastase like proteases(Cleveland et al., 1987). Example of this family of PIs include an 8kDa chymotrypsin inhibitor from *Amaranthus hypochondriacus* active against the insect *Prostephanus truncates*(Sanchez-Hernandez et al., 2004; Valdes-Rodriguez et al., 1993).

- **Potato type II inhibitors (MEROPS: I20)**

  Predominantly found in Solanaceae, the characteristic feature of Pin-II PIs is presence of variable number of inhibitory repeat domains (IRDs)which form multi-domain precursor proteins(Grosse-Holz and van der Hoorn, 2016). A Pin-II PI protein consists of an endoplasmic reticulum signal peptide of 25 amino acids followed by 1 to 8 IRDs of 50 aa which are separated by 5 aa long linkers(Tamhane et al., 2012b). A IRD shows three anti-parallel β-sheets joined by a flexible loop containing the reactive site and stabilized by four disulfide bonds(Isabelle H Barrette-Ng et al., 2003). These PIs are wound inducible proteins, and accumulate in the aerial tissues. Expression of Pin-II PI genes from potato in transgenic tobacco plants has shown to increase in their resistance towards insect pest, *Manduca sexta*(Stevens et al., 2012). Similarly, transgenic expression of Pin-II PI from *Capsicum annuum* in tomato leaves imparted resistance to *H. armigera* larvae(Tanpure et al., 2017). Transgenic cotton expressing individual or combination of inhibitors (NaPI-StPin1A) showed promising results in field trials with *H. punctigera*(Dunse et al., 2010b).

1.4.2 Current status of PIs in control of insect pests

Protease inhibitors have been used for control of lepidopteran pest species by means of expression as transgenic plants, or by external application as recombinant proteins. For example, the expression of Giant taro proteinase inhibitor (Kunitz type) as transgene in tobacco inhibited

the trypsin activity reduced the growth of *H. armigera* larvae(Wu et al., 1997). Similarly, expression of Soybean Bowman-Birk trypsin inhibitor in Sugarcane led to severe reduction in growth of *Diatraea saccharalis*(Falco and Silva-Filho, 2003). Pin-II type proteases have been more widely used for development of transgenic plants. Pin-II type inhibitors from tomato, potato, nicotiana and capsicum have shown significant adverse effects on growth of lepidopteran insects when expressed as transgenics(Dunse et al., 2010b; Stevens et al., 2012; Tanpure et al., 2017).However, the use of transgenic plants containing PIs has been limited. This is because of two reasons. Firstly, the level of expression of PIs in transgenes is very low, like in tobacco plants expressing low levels of mustard trypsin inhibitor (MTI 2) consumed more leaves in order to compensate for reduction in gut protease activity, and ultimately maintained the same population growth rates as non-transgenic plants(De Leo F et al., 1998). Secondly, the success of PIs as transgenics is limited due to the highly adaptive response of pests. Insect pests which suffer from severe reduction in proteolytic activity tend to switch their substrate specificity or express proteases which are insensitive to PIs(Dunse et al., 2010b; Tamaki and Terra, 2015). For instance, upon feeding with tobacco plants transformed with soyabean kunitz trypsin inhibitor, *H. armigera* larvae expressed resistant proteases, and was therefore unaffected by the PI(Broadway, 1996).

Therefore, in order to overcome the resistance of pests to one kind of PI, it is necessary to use proteins which have a wider specificity and activity. Also, gene stacking or pyramiding is an interesting strategy, which uses two proteins with different targets. Transgenic tobacco containing both Bt and cowpea trypsin inhibitor were more resistant to *H. armigera* compared to plants expressing only Bt toxin(Fan et al., 1999). Also, synergistic effect of *N. alata* trypsin inhibitor with Pin-I inhibitor was observed on *H. armigera* larvae(Dunse et al., 2010b). To develop novel insect control agents, one has to look into the diversity of PI families or use multiple inhibitors to show broad spectrum of protease inhibition. Here, we have focused on the Pin-II family of PIs, because of their stable structure and enormous diversity in the plant kingdom, these inhibitors are attractive candidates to engineer protease inhibitors for insect control.

1.5 Pin-II family of protease inhibitors

Pin-II PIs wound inducible proteins that are involved in herbivory related defence. Though primarily found in Solanaceae, identification of many Pin-II homologs throughout the mono- and

dicotyledonous plants has pointed towards a more widespread occurrence of this family(Tamhane et al., 2012b). Pin-II family of PIs are competitive inhibitors of serine proteases and act by the standard Laskowski mode of inhibition(Barrette-Ng et al., 2003).

1.5.1 Structure of Pin II PIs

Pin II PI consists of a signal peptide (25 amino acids) followed by multiple inhibitory domains (55 amino acids) which are connected to each other by linkers (5 amino acids). The inhibitory domains (IRDs) have a conserved structure comprising of 6-8 cysteine residues held by 3-4 disulphide bridges(Joshi et al., 2013). Each IRD is a trypsin or chymotrypsin inhibitor, whose specificity depends on the sequence of its reactive loop of 3 amino acids(Isabelle H. Barrette-Ng et al., 2003; Beekwilder et al., 2000a) (Figure 1.5). The precursor Pin-II PI can simultaneously inhibit multiple protease molecules(Mishra et al., 2010).

For multi-domain Pin II proteins, the IRDs can be organized into one of the two possible domain organizations(Grosse-Holz and van der Hoorn, 2016; Tamhane et al., 2012b) (Figure 1.5):

1) Clasped Bracelet: In the precursor Pin-II protein, there are partial IRDs at the N- and C-terminus, which upon processing form a covalent bond, resulting in a functional IRD. This is called as circularly permuted domain organization. The formation of disulphide bridges between partial repeat regions at the N- and C- terminus of the precursor.
2) Bead on string: Tandem repeats of domains are arranged as bead on string manner.

The IRDs are joined by linker regions of 5 amino acids. The sequence of linker regions is conserved in Nicotiana sp. (EEKKN), whereas they show sequence variation in other genera(Mishra et al., 2013; Tamhane et al., 2012b). In *C. annuum*, the sequences are QRNAK, EENAE, EASAE, EGNAE and EETQK. This region is sensitive to proteases and is proteolytically cleavable by endogenous plant proteases, releasing individual functional units, or IRDs.

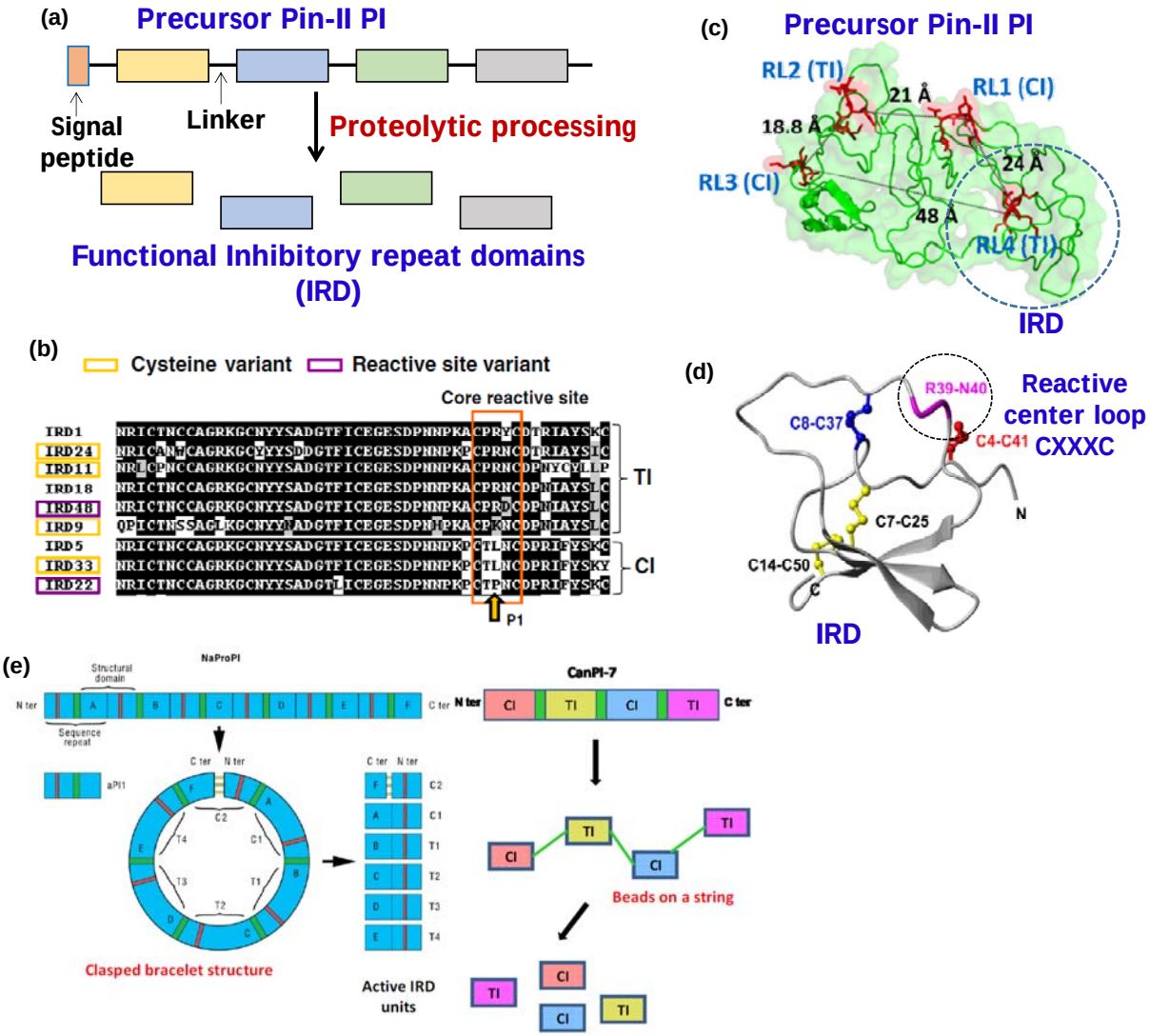


Figure 1.5: Features of Pin-II PIs. (a) Proteolytic processing of precursor Pin-II PI, (b) Variation in sequences of IRDs with core reactive sites highlighted(Mishra et al., 2013) (c) Structure of precursor Pin-II PI(Mishra et al., 2010), (d) Structure of individual IRD, with core reactive loop circled(Joshi et al., 2014b), (e) Possible domain orientations in Pin-II PIs(Tamhane et al., 2012b)

The three dimensional structure of Pin-II PIs has been revealed by X-ray crystallography and NMR(Barrette-Ng et al., 2003; Schirra et al., 2008). The X-ray crystal structure of two domain Pin-II PI from *N. alata* was solved in complex with serine protease Subtilisin Carlsberg. The Pin-II Protein contained two IRDs, which formed a ternary complex with two molecules of the protease(Isabelle H Barrette-Ng et al., 2003). It was seen that the IRD is composed of 3 beta

strands which is stabilized by four disulphide bonds. The primary interaction with protease is mediated by an extended loop, which is held by four conserved disulphide bonds. This loop, called the reactive center loop (RCL) is held in the "canonical conformation" by the network of disulphide bridges, which limit its flexibility. NMR studies on disulphide bond variants of trypsin inhibitor from *N. alata* showed that the cysteines flanking the RCL are conserved, and are indispensable for the activity of IRDs(Schirra et al., 2008). Specifically, the loss of disulphide bridges holding the RCL resulted in a disordered and flexible RCL, which led to difficulty in binding to trypsin. The selective substitution of Cys residues by Ala in RCLs resulted in reduction of inhibitory potency. However, studies have also demonstrated the role of cysteine variants in enhancement of inhibition potential of IRDs. Li et al., 2011 identified six naturally occurring variants of Pin-II PIs from potato, with missing cysteine residues. This loss of cysteines was specific, and associated with functional differentiation of the protein(Li et al., 2011). Further analysis of cysteine variants revealed that the replacement of Cys by Ser leads to enhanced flexibility of the RCL. Due to this reason, the RCL is able to adopt a highly favourable conformation for trypsin binding, enhancing its inhibitory potential(Joshi et al., 2014). Also, RCL region determines the specificity of the IRD for trypsin or chymotrypsin, depending on whether a Lys/Arg or Leu is present at P1 position. The interaction between protease and PI is an entropically driven process, which is mediated by several H-bond and Van der Waals interactions with the RCL(Isabelle H. Barrette-Ng et al., 2003)(Figure 1.6). The RCL is present in an extended flexible conformation in the unbound form, but attains a rigid conformation upon binding to protease. This helps the PI to ensure balance between the broad specificity and tight protease binding. Further, grafting of RCL sequences within two domains of Pin-II type potato inhibitor led to change in specificity of the two domains(Beekwilder et al., 2000b). Hence, RCL regions are the main elements in protease-PI interaction.

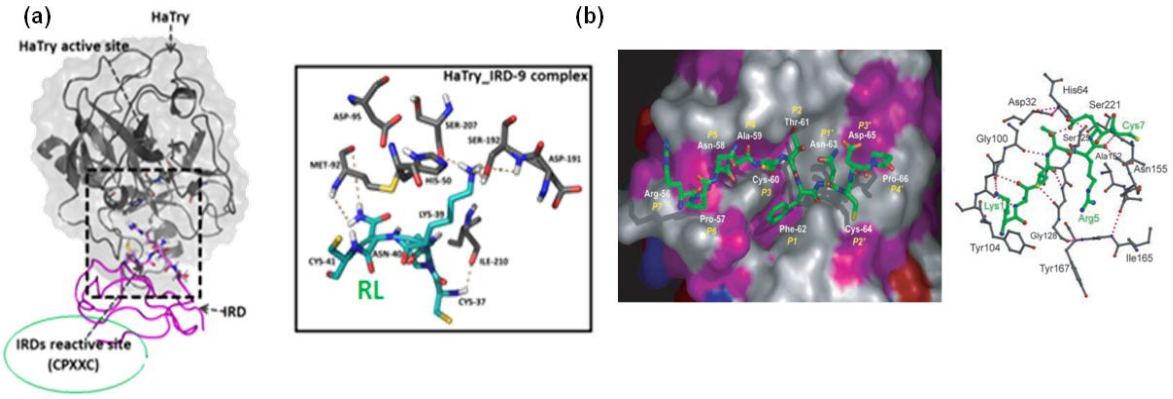


Figure 1.6: Mode of IRD-protease binding. (a) IRD from *C. annum* in complex with *H.armigera* trypsin (predicted model)(Joshi et al., 2014), (b) Tomato IRD in complex with Subtilisin Carlsberg (X-ray structure)(Isabelle H Barrette-Ng et al., 2003)

1.5.2 Effect of dietary Pin II PI on insects

Pin-II PIs act in plant defence against lepidopteran insects, which rely on serine proteases for digestion of the complex diet components. These PIs are known to be competitive inhibitors, i.e. they bind in the active site of proteases. They mimic the substrates and block the proteases in insect midgut, essentially causing indigestion, thereby limiting the supply of amino acids to the insects. As a result, there is a negative effect on the larval physiology, which retards the growth and development of insect. In a long term, this leads to the reduction in population size of the insects, therefore reducing their impact on the crops.

The effect of ingestion of Pin-II PIs on lepidopteran insects has been demonstrated by several studies. Pin-II PIs from tomato and potato were found to cause adverse effects on the larval growth and fecundity of *H. armigera*(Stevens et al., 2012). Pin-II PIs from *C. annuum* fed by artificial diet showed reduction in larval weight as well as in vitro inhibition of insect gut proteases for *H. armigera*(Joshi et al., 2014a; Tamhane et al., 2007). Further, recombinantly expressed single domain IRDs also showed antibiosis against H. *armigera* larvae, and were also effective when used as transgenic tomato plants(Tanpure et al., 2017). An increase of 30% in productivity was observed in transgenic plants with Pin-II from *Solanum tuberosum* and *Nicotiana alata*. Similarly, including *N. alata* PIs in the artificial diet resulted in significant growth reduction of *H. punctigera*, and also arrest of molting(Dunse et al., 2010b).

However, feeding on transgenic Pin-II PIs led to increase in foliage by *Plutella xylostella* larvae for overcoming the reduced protease efficiency. Similarly, the growth of *Spodoptera exigua* larvae was not affected when fed on tobacco transformed leaved containing Pin-II PI(Jongsma et al., 1995). The reason for this was production of insensitive proteases(Dunse et al., 2010a). Thus, new molecules with broad specificity are needed. One approach in this direction is to search the natural diversity of plant PIs. Another approach could be the use of synthetic or mutant inhibitors. It will be interesting to develop PIs based on the sequences involved primarily in the protease-PI interaction, i.e. the reactive center loop.

1.5.3 Reactive center loop (RCL) of Pin-II PIs as insect protease inhibitor

RCL is a region of IRD that interacts with target proteases and it is found to be highly variable in sequence. The RCL binds in the active site of the target enzyme in a particular conformation which blocks the access of substrate in the active site, thus inhibiting its activity. The P1 residue of the IRD loop makes the most contacts with the protease in the S1 pocket and is the primary determinant of the inhibitory specificity(Isabelle H Barrette-Ng et al., 2003). As in many of the PI families, disulphide bonds provide stability to the exposed reactive loops by holding it through covalent attachments. The remaining molecule is known as scaffold and plays a role in holding in the reactive loop. Thus it is capable of interacting with the protease when bound in the appropriate confirmation. Since the RCL region is the main interacting loop in Pin-II PI, it is speculated that these RCL regions are capable of protease inhibition devoid of the remaining protein scaffold. Similar studies with the reactive loop regions of Bowman-Birk family have shown that a 9 residue monocyclic peptide of the trypsin binding domain was adequate to retain inhibitory properties of the parent inhibitor. Starting from 90 amino acids, similar proteins were progressively reduced to 11 amino acid reactive loop regions, which show potent protease inhibition(McBride et al., 1996). Similarly, eight amino acid reactive loop region of serpins showed enhanced immunomodulatory activity(Ambadapadi et al., 2016). Further, grafting of reactive loop peptides on various scaffolds led to enhancement of protease inhibition. Thus, RCL peptides could be engineered to develop more potent protease inhibitors.

1.6 Small peptides as lead molecules for generation of insect control agents

Plants produce defense related proteins like PIs and lectins in response to insect attack. However, the limited success of these proteins as transgenic has pointed towards the need for their external application. Yet, the application of these proteins as formulations is also difficult, because of their unstable and proteolytically susceptible nature. Thus, there is always a need for new molecules which could be used for insect control. In this direction, peptide based small molecules have shown promising results (Figure 1.7).

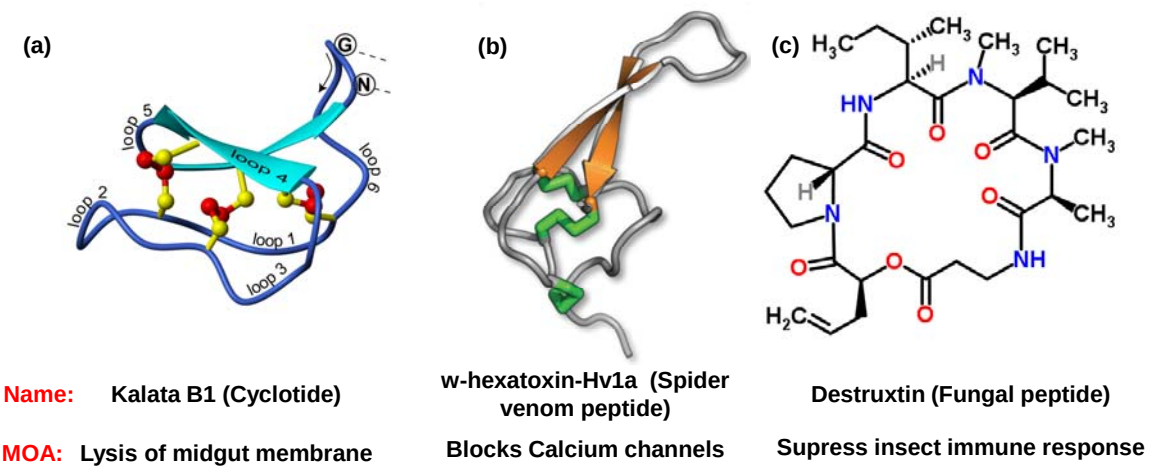


Figure 1.7: Representative examples of insect control peptides. (a) KalataB1 from *Oldenlandia affinis*(Barbeta et al., 2008) (b) Spider venom toxin from *Hadronyche versut*(King and Hardy, 2013) (c) Destruxin from fungus *Metarhizium robertsii*(Wang et al., 2012). (MOA:Mode of action)

Small peptides serve as excellent scaffolds for design of protease inhibitors, owing to their high stability towards pH, temperature and proteases. Being in a constrained conformation, they are entropically favorable than the flexible linear peptides, hence also show improved binding affinity. Furthermore, the sequences of these peptides could be replaced by other sequences, which might lead to enhancement in efficacy, specificity and also functional differentiation of the peptides. This concept, called "grafting" of linear peptides is illustrated in figure 1.8. This method could be used for grafting the reactive loops from different plant PI families on the small peptide scaffolds, which opens up an interesting arena for design of novel pest control molecules.

1.6.1 Natural scaffolds

Among the naturally occurring small peptides with stable architecture, peptides from various sources have been used as scaffolds for design of novel bioactive compounds. A brief description of some of the examples is given below:

1.6.1.1 Peptides from plant sources

- **Cyclotides**

  Cyclotides form a predominant family of plant cyclic peptides which act as host defense agents against pests and pathogens, and also have other bioactivities with potential pharmaceutical or agricultural applications(Craik et al., 2013; Troeira Henriques and Craik, 2017).These are cyclic proteins of 28-37 amino acids containing N to C cyclized peptide backbone and six disulphide bridges. These disulphide bridges show a knotted topology, which combined with the cyclic peptide backbone form a "cysteine knot"(Craik et al., 1999). This architecture confers high stability to these peptides. More than 300 cyclotide sequences have been identified in 60 plant species, comprising the Rubiaceae, Violaceae, Cucurbitaceae, Solanaceae, and Fabaceae families(Troeira Henriques and Craik, 2017). Cyclotides are synthesized in plants by ribosomally produced precursor proteins(Poth et al., 2011). The cyclotide loops show variations in sequence and size between successive Cys residues. Interestingly, cyclotides are exclusive to a particular plant species, with the exception of kalata B1, which is found in six different plant species(Troeira Henriques and Craik, 2017). Cyclotides are classified into two subfamilies, Mobius and bracelet, based on the presence or absence of Pro residue preceded by a cis-peptide bond in loop 5, respectively(Craik et al., 1999). Another subfamily, the trypsin inhibitor cyclotides, contain the cysteine knot motif, but lack any of the other features. The trypsin inhibitor cyclotides are found in the seeds of cucurbitaceae family and share sequence homology with acyclic trypsin inhibitor family, knottins(Craik et al., 2004).

  Mobius and bracelet type cyclotides have shown efficacy against insect pests. For example, kalata B1 showed potent insecticidal activity of *H. armigera*, by causing lysis of midgut epithelial membrane(Barbeta et al., 2008). The insecticidal activity of cyclotides was also reported against other insects, like *Diatrea saccharalis* and

Drosophila melanogaster(Pinto et al., 2012; Simonsen et al., 2008). Whereas, trypsin inhibitor cyclotides do not bind to the midgut epithelial membrane, but are potent trypsin inhibitors(Hernandez et al., 2000).

Cyclotides are amiable for chemical synthesis and can tolerate substitutions in their loop regions. This property has been used for grafting of linear peptides into the cyclotide scaffold to enhance the cellular permeability(Craik et al., 2018). Specifically, kalata b1 (Mobius type cyclotide), and McoTI-I and II (trypsin inhibitory cyclotide) have been extensively used for grafting of peptide sequences to improve stability(D'Souza et al., 2016; Ji et al., 2013). Small linear peptides mimicking the N-terminal fragment of p53 were engineered into McoTI-I scaffold, for enhancing stability and cell permeability for activation of p53 tumor suppressor pathway(Ji et al., 2013). Another interesting application of cyclotide scaffolds is for achieving dual functionality of peptides. This was shown by grafting of two different anti-angiogenesis sequences into McoTI-II scaffold, which targeted two angiogenesis pathways. These bifunctional molecules offer advantage over single function scaffolds(Chan et al., 2016).

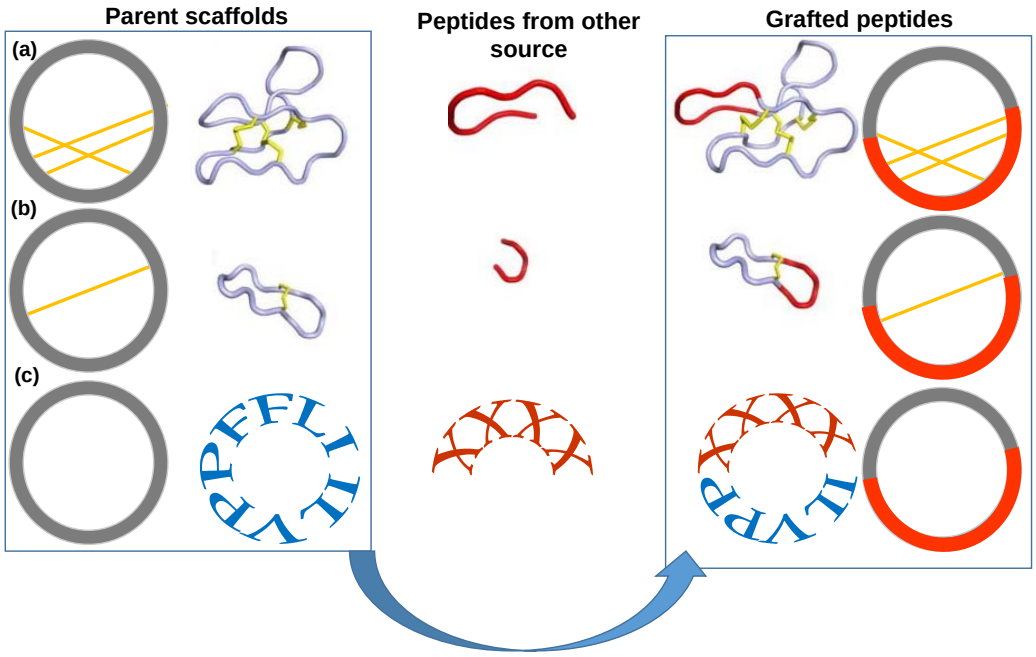


Figure 1.8: Grafting of peptides. Natural scaffolds (a) Cyclotide/knottin (b) Sunflower trypsin inhibitor(Craik et al., 2018) (c) orbitides are shown in cartoon representations, with simplified diagrams shown as circle (for peptide backbone) and yellow lines (for disulphide linkage).

- **Knottins**

  These are small proteins of ~30 amino acids, and possess the characteristic cysteine knot motif. These also include the cyclotides and conotoxins. Knottins are found in phylogenetically diverse sources, like plants, spiders, cone shells and fungi, and display a myriad of pharmacological activities. These peptides are characterized by an acyclic backbone, and three disulphide bridges, which form a "disulphide within disulphide" knot (one disulphide bond is weaved through the macrocycle formed by the remaining two disulphides). These proteins display a variety of functions, including pharmacological and agricultural applications. Also, they show diverse mode of action, involving protease inhibition, ion channel disruption and neurotoxicity(Moore et al., 2012). Recently, a 3.8 kDa protein (PA1b) from *Pisum sativum* has shown potent insecticidal activity against cereal weevils like *S. oryzae*, *S. granaries* and *S. zeamays*(Jouvensal et al., 2003; Rahioui et al., 2014). However, this protein was ineffective against lepidopteran insects, showing a specific mode of action which needs to be explored further.

  The knottin fold is a useful scaffold for design of PIs, because of its small size, ease of synthesis, stability and high sequence tolerance(Kolmar, 2008; Moore et al., 2012). For example, squash miniproteins displayed plasma stability and stability to proteolytic enzymes. Loop grafting on the trypsin inhibitor EETI-II from *Ecballium elaterium* by epitope sequences resulted in novel functionalities (i.e. antibody binding)(Christmann et al., 1999). Moreover, directed evolution of EETI-II knottin lead to generation of novel functionalities(Souriau et al., 2005).

- **Sunflower trypsin inhibitor**

  Sunflower trypsin inhibitor (SFTI-1) is found in the seeds of sunflower, *Helianthus annuus*, and forms a tight complex with trypsin. SFTI-1 is a 14-amino acid long peptide, which is head-to-tail cyclized. It is a bicyclic peptide due to the presence of a single disulfide bond between Cys3 and Cys11, forming an extended 7-residue bioactive loop containing active site Lys residue, and a 5-residue cyclisation loop. The rigid structure is optimized by an extensive hydrogen bond network within the SFTI framework, which is the basis for its high potency(de Veer et al., 2015). The biosynthesis of these proteins

takes place from a seed storage protein encoding a storage albumin, called Preproalbumin with SFTI-1 (PawS1)(Mylne et al., 2011). The peptide sequence for SFTI is processed by albumin processing machinery. The screening of Asteraceae family for SFTI-1 like peptides led to the discovery of 36 bicyclic peptides, making them part of a separate family, called the PawS-Derived Peptides (PDPs)(Elliott et al., 2014).

SFTI-1 has been used as a model scaffold since a long time for engineering of protease inhibitors. By replacing the cyclization loop of this inhibitor with another reactive loop, or by replacing the catalytic Lys by other residue lead to generation of bifunctional protease inhibitors(Franke et al., 2018). Also, replacement of P1 residue by un-natural amino acids led to altered specificity of SFTI-1(Łegowska et al., 2009). For pharmaceutical applications, tetrapeptide sequence FVQR has been grafted into SFTI-1 to yield a selective and potent inhibitor of kallikrein-related peptidase 4 (KLK4), which is overexpressed in malignant prostate tumors(Swedberg et al., 2009).

- **Orbitides**

  These are small cyclic peptides of 5-12 amino acids, which contain N-to-C terminus cyclization of the peptide backbone. Cyclolinopeptide A, was the first orbitide isolated from flaxseed (Linum usitatissimum), in 1959(Kaufmann and Tobschirbel, 1959) and had the sequence ILVPPFFLI(Kaneda et al., 2016). Up to 2017, over 168 orbitides have been discovered in plant families Caryophyllaceae, Annonaceae, Rutaceae, Euphorbiaceae, and Linaceae(Chekan et al., 2017). These are ribosomally synthesized and posttranslationally modified peptides (RiPPs)(Chekan et al., 2017)and are interesting molecules, both chemically and structurally. Though the functions of these small cyclic peptides are yet to be elucidated, many of these orbitides demonstrate antimalarial, vasodilatory, immunomodulating and protease inhibitor activities(Craik et al., 2018). A number of orbitides have been identified from *Saponaria vaccaria*. These are named assegetalins (A to K), and possess vasorelaxant properties(Morita et al., n.d.). The potential of orbitide backbone in improvement of physicochemical properties has been demonstrated by a systematic study of the backbone N-methylation of cyclic peptides(White et al., 2011). Although orbitides have not been used extensively for the generation of PIs, they might be promising scaffolds for grafting studies, because they are

similar to small cyclic peptides, and thus can improve the bioavailability of these peptides. These peptides provide substantial clues for the generation of modified peptides with enhanced properties.

1.6.1.2 Peptides from arthropods

Insect toxic peptides have been identified in venoms of arthropods like bees, wasp, spider, scorpion etc(Smith et al., 2013). Most of these peptides possess the inhibitory cysteine knot as the structural scaffold. About 116 insect selective peptides have been discovered from spider venom alone(Herzig et al., 2011). These are small peptides, and show varied mode of action on the lepidopteran insects. These peptides are active on Na^+ channels, voltage-gated calcium channels, calcium-activated potassium channels, and glutamate receptors. These peptides act as neurotoxins, and paralyze the lepidopteran insects by disrupting the sodium ion channels. Transgenic plants expressing the venom peptide w-hexatoxin-Hv1a from the Australian spider *Hadronyche versuta* were developed for cotton, tobacco, and poplar plants. *H. armigera* larvae feeding on these transgenic tobacco leaves showed 75-100% mortality as compared to non-transgenic plants(Fletcher et al., 1997; King and Hardy, 2013; Smith et al., 2013).

1.6.1.3 Peptides from microbial sources

Potent insect control peptides are also reported from microbial sources, mainly bacteria and fungi. One type of insecticidal peptides are depsipeptides. These are small cyclic peptides in which one or more amino acid is replaced by a hydroxylated carboxylic acid. This leads to the formation of at least one lactone bond in the core ring(Wang et al., 2018). Their biosynthesis occurs by means of non-ribosomal peptide synthetases, along with either polyketide synthase or fatty acid synthase enzyme systems(Sivanathan and Scherkenbeck, 2014). Fungal depsipeptides, like beauvericins, destruxins and enniatins have been studied for insecticidal properties. Entomopathogenic fungus, *Beauveria felina* is known to produce toxic peptides, named Isarfelin A-D, which showed potent insecticidal activity against lepidopteran insect, *Leucania separata*(Guo et al., 2005; Langenfeld et al., 2011). Recently, Schellenberger et al. identified an insecticidal peptide from a soil-dwelling bacterium, *Pseudomonas chlororaphis*. This peptide

was expressed in corn plants, which were protected from attack by western corn rootworm(Schellenberger et al., 2016).

1.6.2 Synthetic scaffolds

In addition to natural scaffold based optimization of peptides, synthetic scaffolds have also been studied to generate peptides with enhanced binding affinities. These small molecule scaffolds benefit from the fact that they maximize the enthalpic interactions of the binding affinity such as hydrogen bonding and salt bridges. Also, constraining a peptide into a small cyclic framework enhances its stability towards endopeptidades. An important reason for this stability is that their structure does not fit into the endopeptidase active sites, which generally require an extended conformation for binding. Therefore, the small cyclic scaffolds are promising molecules for development of novel peptides(Yudin, 2015) (Figure 1.9).

A classic example of small scaffold is cyclosporine A, which is a non ribosomally synthesized cyclic peptide isolated from *Tolypocladium inflatum*. This peptide optimizes its network of intramolecular hydrogen bonds while passing through the lipid bilayer, and later with its cellular target by making numerous intermolecular interactions. Thus, cyclosporine is an attractive candidate for drug development(Beck et al., 2012; Yudin, 2015). Further, peptides have been constrained using the properties of cysteines to form disulphide linkages. Intramolecular disulphide bonds have been used to increase the helical character of small peptides(Jackson et al., 1991). Later, α-aminoisobutyric acid was used for this purpose(Christian E. Schafmeister et al., 2000). Furthermore, the use of α, α-disubstituted amino acids and ring closure metathesis led to development of "stapled peptides", which are macrocyclic peptides(Blackwell et al., 2001).

Additionally, combinatorial peptide screening has emerged as a novel method for generation of new peptide architectures. Methods like phage display, aptamer, SELEX and mRNA display have given several lead molecules for various applications(Yudin, 2015). These methods use large populations of bacteriophages, which are then selected on the basis of binding to the desired biological target. Libraries containing 10^{13}members have been developed in order to select macrocyclic ligands binding to serine protease inhibitors(Angelini et al., 2012).

Further, an interesting approach used for cyclization of cysteine containing linear peptides is the use of bromomethyl functionalized aromatic scaffolds. This chemistry causes

rapid and high yielding cyclization of linear peptides of 2-30 amino acids, and is exclusive for free cysteines(Timmerman et al., 2005). Further, optimization of this scaffold with additional functional groups led to increased stability and potency for protease inhibition. This was possible due to enhancement in intramolecular interactions of the cyclized peptide, which imparted stability and bonding potential(Chen et al., 2014). Screening of peptide libraries using phage display, and cyclized by this chemistry led to the selection of potent inhibitor of human plasma kallikrein, a serine protease(Chen et al., 2012).

Other methods are also reported for grafting of peptide epitopes on pre-existing scaffolds. β-strand mimics were synthesized by replacement of two amino acids of a peptide macrocycle by a planar pyrrole. This molecule retained the geometry needed for active site binding, and were potent inhibitors of serine proteases, like chymotrypsin, Cathepsin L and S(Chua et al., 2014). In another example, five out of ten amino acids were replaced in the inhibitor scaffold of mupain-1, which inhibited serine proteasemurine urokinase-type plasminogen activator. This led to its functional differentiation by showing specific inhibition of plasma kallikrein with high potency(Xu et al., 2015).

Thus, peptide sequences could be grafted on natural as well as synthetic scaffolds, which might lead to enhancement in their stability as well as potency. Further, specific modifications could also be incorporated into peptide backbone or amino acid residues in order to engineer novel molecules.

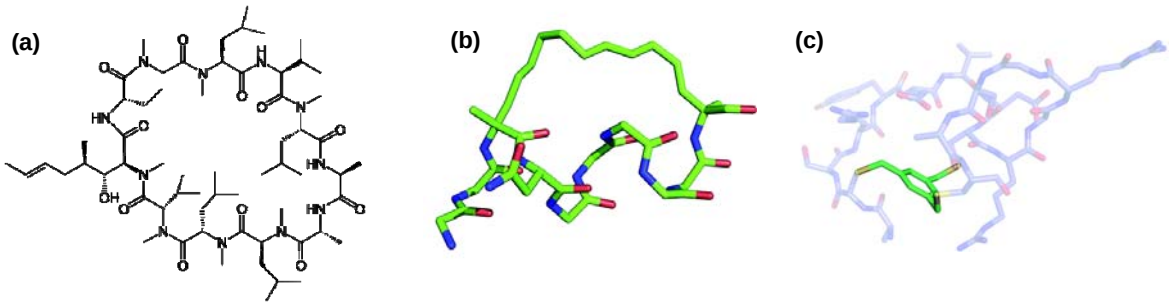


Figure 1.9: Examples of synthetic scaffolds used for peptide grafting (a) Cyclosporine A(Chung and Yudin, 2015), (b) Stapled peptides(Stewart et al., 2010), (c) Bicyclic peptide on bromomethylated scaffold(Angelini et al., 2012)

1.7 Use of peptidomimetics for generation of potent protease inhibitors

1.7.1 Un-natural amino acids in plant defence

In addition to secondary metabolites, plants also produce un-natural amino acids, which serve as anti-nutrition compounds for insects. These amino acids also possess antimicrobial, allelochemical, signalling and nitrogen storage activities(Huang et al., 2011). A few examples of unnatural amino acids which help in plant defense against herbivores are given here (Figure 1.10):

L-Canavanine*:* It is an arginine analogue found in many leguminous plants and is highly toxic to insects and other organisms, including mammals. It was first isolated from jack bean (*Canavalia ensiformis*)(Rosenthal, 2001). It incorporates into insect proteins, making "canavanyl" proteins that are structurally unorganized and hence, useless to the insects(Staszek et al., 2017). However, some insects, like *H. virescens*(Melangeli et al., 1997) and *Caryedes brasiliensis*have also shown the ability to avoid or detoxify canavanine(Rosenthal et al., 1978). Nevertheless, this amino acid is an antimetabolite and shows strong toxic effects on the insect survival.

γ-aminobutyric acid*:* γ-aminobutyric acid (GABA) is found in most prokaryotic and eukaryotic organisms, and is reported to increase in level upon insect attack. The negative effect of GABA on the growth of leaf roller larvae (*Choristoneura rosaceana*)(Bown et al., 2002) and tobacco budworms (*Heliothis virescens*)(MacGregor et al., 2003)was demonstrated by feeding assays on GABA containing diets. GABA acts on the peripheral nervous system of the insects(Kinnersley and Turano, 2000). A structurally related analogue of GABA is DL-β-Aminobutyric acid (BABA), and it is shown to reduce the growth of lepidopteran insects like *Trichoplusia*and *Plutella xylostella*(Hodge et al., n.d.).

Others*:* In leguminous plants, several toxic amino acids like L-DOPA (L-3,4-dihydroxyphenylalanine), 5-HTP (5-hydroxy-L-tryptophan), L-mimosine ((2S)-2-amino-3-(3-hydroxy-4-oxopyridin-1(4H)-yl)propanoic acid), p-aminophenylalanine, and L-azetidine-2-carboxylic acid are found. L-DOPA and 5-HTP, are derivativesof tyrosine and tryptophan, respectively, and mimic the insect neurotransmitters. Incorporation of 5-HTP in artificial diet causes toxicity in *Spodoptera* spp(Huang et al., 2011; Vranova et al., 2011).

1.7.2 Peptide bond modification

Peptides are important molecules for discovery of PIs due to their high affinity and specificity. However, they are limited in use because of poor stability and low bioavailability. Hence, one of strategies to develop peptide mimics for increased stability is by replacement of peptide bonds by un-natural linkages(Avan et al., 2014), some of which are mentioned below (Figure 1.10):

Modification in amino functionality*:* Backbone amino group can be isosterically replaced by other atoms, like oxygen (Depsipeptides) and sulphur (Thiodepsipeptides). Depsipeptides have been isolated from fungi, bacteria and marine organisms, and show a vast area of biological activities. These replacements change the secondary structure of peptides by altering the H-bonding pattern(Scheike et al., 2007).

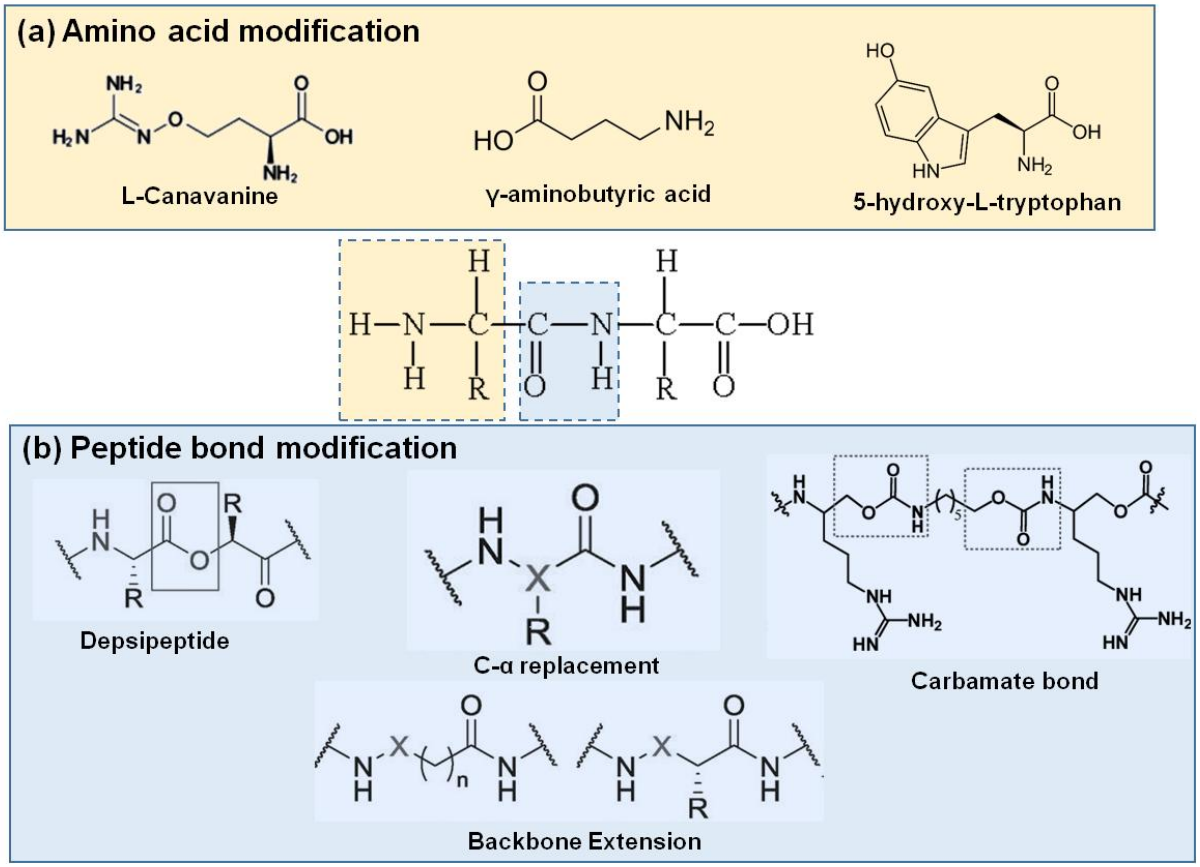


Figure 1.10: Strategies for development of peptidomimetics

Replacement of α carbon*:* Replacing the α carbon of peptide backbone by nitrogen leads to formation of azapeptides, which differ significantly from the parent peptide in terms of chemical and biological activity(Boeglin and Lubell, 2005). This is due to loss of chirality, ultimately leading to β-turn conformation. Azapeptides have been used as inhibitors of serine and cysteine proteases, like human neutrophil protease, hepatitis A virus protease, and HIV protease(Avan et al., 2014).

Extension of peptide backbone*:* Addition of extra atoms in the peptide backbone can lead to novel peptides, for example, backbone extended peptides derived from β or γ amino acids can adopt secondary structures similar to those found in native proteins, like helices, sheets and turns(Avan et al., 2014).

Carbamate linkages: These are amide-ester hybrid structures and show very good chemical and proteolytic stability. Peptides with carbamate linkages show good cellular permeability(Ghosh and Brindisi, 2015). Organic carbamates are valuable backbones for design of therapeutic agents. Also, carbamate derivatives are used as pesticides and fungicides. Inhibitors of serine proteases, like kallikrein, thrombin, and elastase have been designed which contain carbamate functionality(Ghosh and Brindisi, 2015).

1.8 Objectives of work

Pin-II family of plant protease inhibitors are interesting candidates to develop dietary pest control agents for lepidopteran pests, like *H. armigera*. These inhibitors bind to serine proteases in the midgut of insect, and interfere with protein digestion, causing adverse effects on insect growth. The main interaction of Pin-II protein with protease is by means of reactive center loop (RCL), which makes numerous contacts with protease active site. The RCL region is a tripeptide sequence flanked by cysteine residues and is held in a defined orientation by disulphide bonds made by the protein scaffold. The Pin-II family consists of 21 conserved RCL sequences, which show mixed representation in the family. We speculated that the tripeptide RCL region might be the minimum sequence required for inhibitory activity of the Pin-II PIs. Therefore, the RCL regions could be used for designing of small peptide based pest control agents by grafting on a

cyclic scaffold, or by developing peptidomimetics. On this background, we have defined the objectives of thesis work as follows:

1. Characterization of tripeptides derived from reactive center loops of potato type II protease inhibitors

- Synthesis of RCL peptides
- Biochemical activity analysis of peptides
- *In vitro* evaluation of peptides against *H. armigera*
- Determination of mode of peptide binding with trypsin

2: Tailoring of reactive loop peptides by cyclization for inhibition of lepidopteran serine proteases

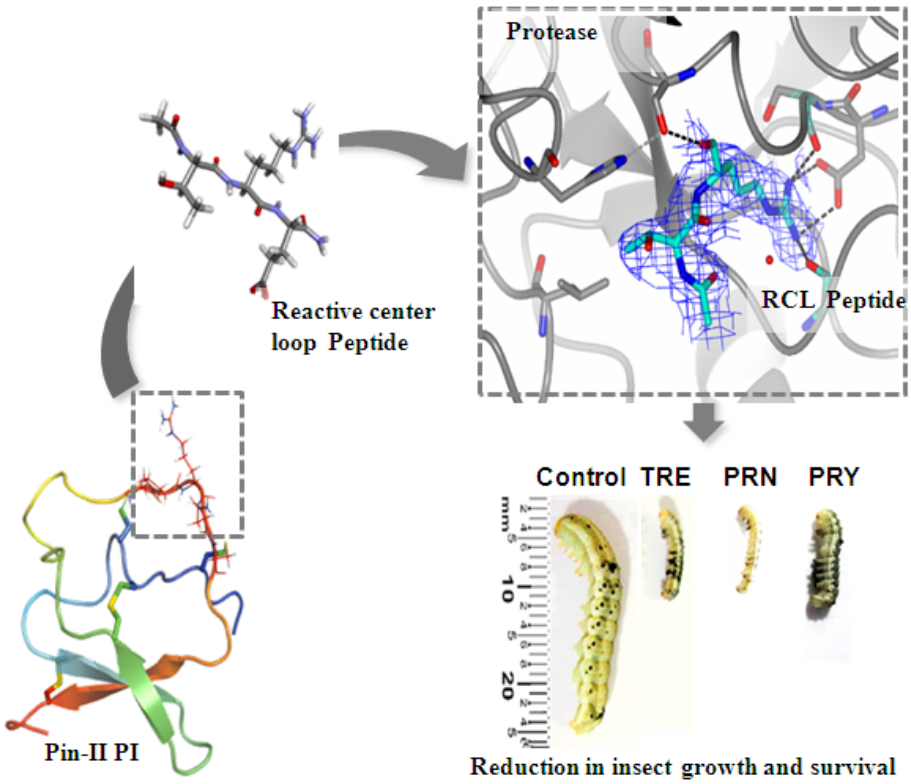
- Design of cyclic peptides
- Synthesis and purification of cyclic peptides
- Screening and selection of peptides by biophysical and biochemical activity assays
- Study of effects on lepidopteran insects

3: Design and evaluation of RCL peptidomimetics

- Design of Arg-analog and peptide backbone modified containing peptides
- Estimation of binding modes and selection by virtual screening
- Synthesis of RCL peptidomimetics
- Biochemical and biophysical evaluation for serine protease inhibition

Chapter 2

Characterization of tripeptides derived from reactive center loops of potato type II protease inhibitors



Contents of Chapter 2 have been published in the research article…

Saikhedkar, N. S., Joshi, R. S., Bhoite, A. S., Mohandasan, R., Yadav, A. K., Fernandes, M., Kulkarni, K. A. & Giri, A. P. (2018). Tripeptides derived from reactive centre loop of potato type II protease inhibitors preferentially inhibit midgut proteases of *Helicoverpa armigera*. *Insect Biochemistry and Molecular Biology*, 95, 17-25.

2.1 Introduction

Potato type II protease inhibitors (Pin-II PIs) are wound-inducible proteins which are involved in plant defence against insect pests. Pin-II PIs are considered as potential alternatives to insect toxins and deterrents(Dunse et al., 2010), because they inhibit trypsin- and chymotrypsin-like enzymes (Sarate et al., 2012; Srinivasan et al., 2006) in the insect midgut, specifically for lepidopteran pests like *Helicoverpa armigera*. Previous reports have shown that Pin-II PIs cause adverse effects on the growth and development of lepidopteran insects. The mechanism of action of Pin-II PIs involves mimicking the substrate and thus binding to active site of protease, ultimately forming a non-dissociable complex (Czapinska and Otlewski, 1999; Turra and Lorito, 2011). Pin-II PIs are multidomain proteins, which are composed of multiple inhibitory repeat domains (IRDs) joined by linker regions of 5-6 amino acids. Upon proteolytic cleavage at the linker regions, IRDs are released which are ~50 amino acid long subunits with 6-8 disulphide bridges. These IRDs in turn, are involved in complex formation with target protease (Mahajan et al., 2015; Mishra et al., 2013; Schirra and Craik, 2005). Recombinant IRDs are shown to have growth inhibitory effect on lepidopteran pests. However, the use of proteinaceous PIs as insect control agents is limited, because of the need for a properly folded structure, large size, proteolytic susceptibility and environmental instability. Therefore, by using RCL sequences of Pin-II PIs, it is possible to develop peptide based pest control agents, which could overcome the drawbacks associated with plant PIs.

Study of the complex between protease and IRD shows that it is mediated by the reactive centre loop (RCL). The RCL is a tripeptide sequence, which is projected in a specific conformation with the help of IRD scaffold(Joshi et al., 2014; Schirra et al., 2008) (Figure 2.1). Previous reports have suggested that the P1 residue of RCL forms a scissile bond with the S_1 pocket of the proteases along with polar contacts and salt bridges from both the interacting partners (Barrette-Ng et al., 2003; Otlewski et al., 2001). Motif or pattern search analysis of 389 Pin-II PIs from MEROPS database showed that there are 237 unique IRDs in the Pin-II family. These IRDs are clustered into lineage-specific clades as shown in the phylogram (Fig. 2.2a). In spite of the distribution of 237 IRDs into nine clades, there are only 21 distinct RCL variants, which show mixed representation in the lineage-specific clades. Furthermore, six variants with sequences, TRE, PRN, PRY, PKN, PLN and TLN show a higher preponderance with ~81% of total occurrence (Fig. 2.2b).

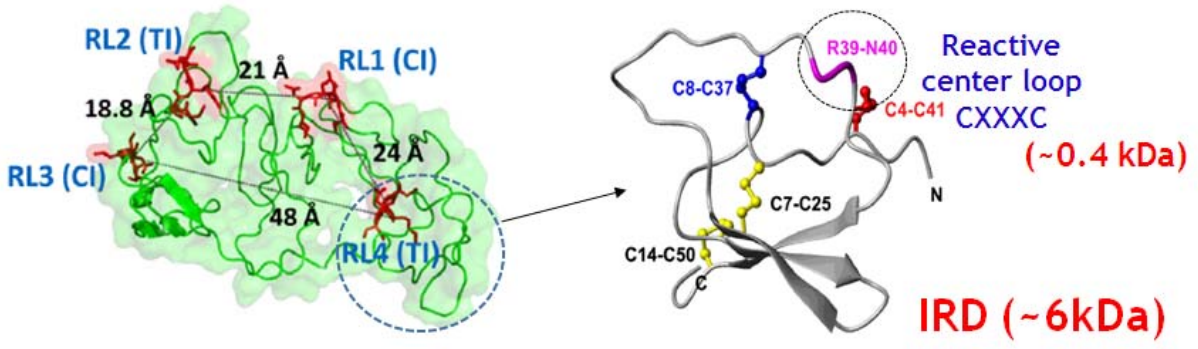


Figure 2.1: From Pin-II PI to RCL peptide. Representative molecules are shown with constituent subunit encircled. Approximate molecular weights are given in brackets.

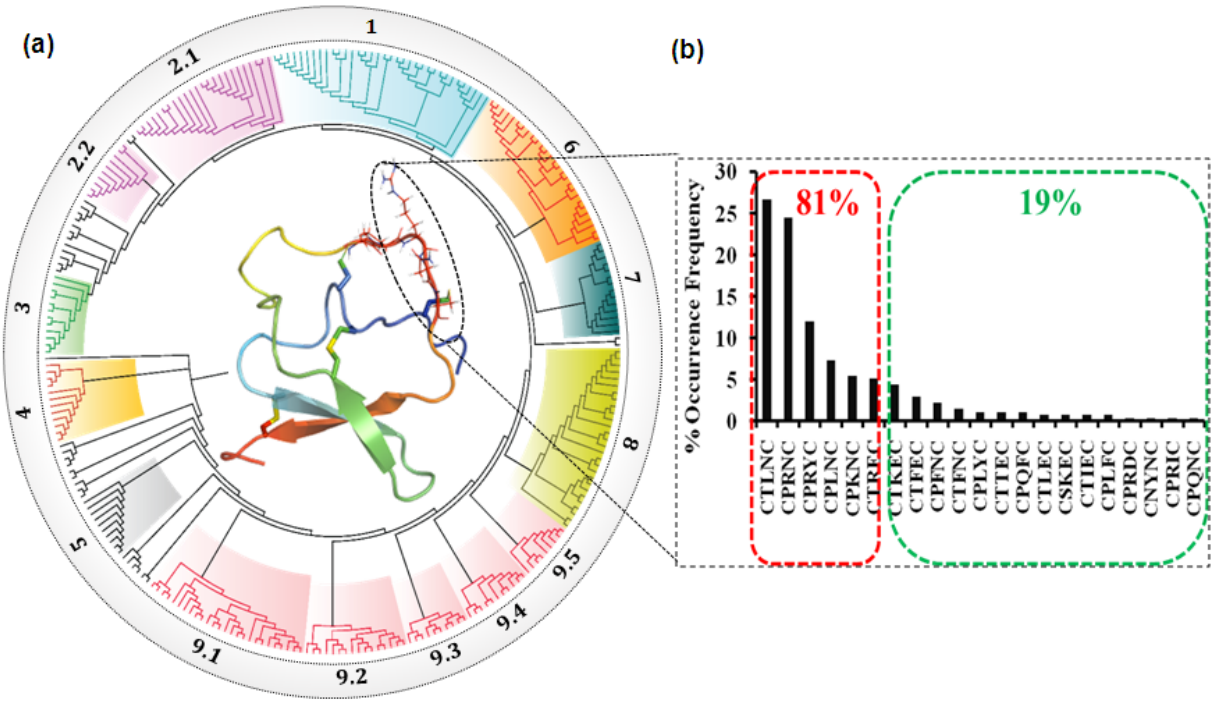


Figure 2.2: Selection of RCL peptides. (a) Phylogenetic tree of Pin-II Inhibitory repeat domain (IRD) protein sequences in which clad 1 and 2 majorly constitute *Capsicum annuum* IRDs, clades 3 to 6 constitute *Nicotiana alata* IRDs, and clades 7-9 includes IRDs from *Solanum tuberosum, Solanum lycopersicum* and other IRDs. (b) 21 RCL variants obtained from motif analysis by MEME. Recurrent variants are shown in red box, which constitute 81% of the total RCL diversity.

This signifies that the RCL sequences have been conserved in the evolutionary process. Co-evolution of RCL with their target protease indicates its crucial role in plant-insect interaction. Also, it was observed that Thr or Pro is conserved at P1 position of the RCL peptides, suggesting that these residues might be responsible for high inhibition activity, as reported earlier for cystatins (Rasoolizadeh et al., 2016). It was also proposed that the P2 Pro helps in rigidity of the RCL and P2 Thr makes additional H bond with P1' residue, hence stabilizing the RCL(Isabelle H. Barrette-Ng et al., 2003). Since the RCL region is the main interacting loop in Pin-II PI, it is an intriguing aspect to study that whether these RCL regions are capable of protease inhibition devoid of the remaining protein scaffold.

Therefore, the RCL peptides were selected for this study with the following objectives:

- Synthesis of RCL tripeptides
- Study of the inhibitory potential of RCL peptides against serine proteases
- Evaluation of *in vitro* and *in vivo* effects of RCL peptides on lepidopteran pest, *H. armigera*
- Study of binding modes of RCL peptides with serine proteases

2.2 Methods

2.2.1 Synthesis of RCL tripeptides

Six tripeptides, namely TRE, TLN, PRY, PRN, PKN and PLN were synthesized by solid-phase method using Fmoc-chemistry on MBHA resin (substitution 1.75mmol/g). The following amino acid derivatives were used: Fmoc-Arg(Pbf), Fmoc-Thr(*t*Bu), Fmoc-Pro, Fmoc-Glu(O*t*Bu), Fmoc-Tyr(*t*Bu), Fmoc-Asn(trt), Fmoc-Leu, Fmoc-Lys(Boc) from Novabiochem, Germany. Peptide chains were elongated in the consecutive cycles of deprotection and coupling. 5 equiv. of TBTU was used as coupling agent along with HOBt and DIPEA. Deprotection of the Fmoc-protected amino groups was achieved by treating with 20% piperidine in DMF. Coupling and deprotection steps were monitored by the Kaiser Test. *N*-terminal acetylation was then carried out, followed by cleavage from the resin using TFA/thioanisole/EDT/TFMSA procedure. N-terminal tagging with carboxyfluorescein was performed by coupling with 10 equiv. of 5/6-carboxyfluorescein using HOBt and diisopropylcarbodiimide (DIPCDI) as coupling agents. The crude tripeptides were purified by reverse phase HPLC using an Agilent TC-2 C18 column. The solvent systems were 5% Acetonitrile/water/0.1% TFA (A) and 50% acetonitrile/water/0.1% TFA (B). A linear gradient from A to B for 10 min, hold at B for 10 min, flow rate 1.5 mL/min, monitored at 220 nm. For fluorescent peptide, ex/em wavelength of 490/520 nm was used. The mass spectrometry analysis was carried out on a MALDI MS (AB Sciex TOF/TOF 5800 spectrometer) using DHB (2,5-dihydroxybenzoic acid) as the matrix.

Tripeptides were quantified by spectrophotometric measurement at 205 nm(Anthis and Clore, 2013). Molar extinction coefficients (M^{-1} cm^{-1}) were calculated using online tool (spin.niddk.nih.gov/clore): TRE (6910), PRY (12290), PRN (7310), PKN (5960), PLN (5960), TLN (5960).

2.2.2 Protease inhibition assays

Concentration-dependent reduction in activity of serine proteases was estimated through specific chromogenic substrate BA*p*NA (Benzoyl-L-arginyl-*p*-nitroanilide) for trypsin and SAApNA (*N*-Succinyl-Ala-Ala-Pro-Phe-*p*-nitroanilide) for chymotrypsin assays. Protease assays were performed as previously described (Tamhane et al., 2007, 2005). 0.3µM of bovine trypsin (15µl) was added to 100µl of 1 mM substrate solution and incubated at 37°C for 10 min. The reaction was terminated by the addition of 40 µl of 30% acetic acid and absorbance was checked

at 410 nm. For enzyme inhibitor assay, the inhibitor was mixed with the enzyme and the premix was incubated at 37°C for 10 min. The residual enzyme activity was then estimated as above. For HGP extraction, fourth instar larvae were dissected and snap frozen larval guts were ground to a fine powder, followed by extraction in the 3x volume of 0.2 M Glycine-NaOH buffer (pH 10) for 2 h at 4 °C. The extract was then centrifuged at 12,000xg (4 °C; 10 min) and the clear supernatant was used as a source of enzyme for protease assays. Assays were performed at pH 7.8 and 10 using 0.1 M Tris-Cl and 0.2 M Glycine-NaOH buffers, respectively. A minimum of three replicates of each experiment was performed. IC_{50} values for each inhibitor were calculated from the sigmoid curve, indicating the best fit for the percentage inhibition data obtained. Type of inhibition and inhibition constant (Ki) for trypsin and chymotrypsin was determined by using various concentrations of substrate BApNA (0.5-5mM) and SAApNA (0.1-2.5mM), with increasing peptide concentration (0.2-0.8mM). Results were analysed by plotting double reciprocal plots and data fitting into competitive inhibition model of GraphPad Prism version 6.04, GraphPad Software, La Jolla California USA. Statistical analyses were performed by t-tests assuming equal variances, and one way ANOVA. Asterisks indicate significant differences (*$p < 0.05$; **$p<0.01$;***$p < 0.001$).

Total HGP activity was measured by azocaseinolytic assay. For azocaseinolytic assay 60 µl of diluted enzyme was added to 200 µl of 1% azocasein (in 0.2 M glycine–NaOH, pH 10.0) and incubated at 37 °C for 30 min. The reaction was terminated by the addition of 300 µl of 5% trichloroacetic acid. After centrifugation at 14,230xg for 10 min, an equal volume of 1 M NaOH was added to the supernatant and absorbance was measured at 450 nm. A minimum of three replicates of each experiment was performed. Statistical analyses were performed by t-tests assuming equal variances, asterisks indicate significant differences (*$p < 0.05$; **$p<0.01$;***$p < 0.001$).

2.2.3 Crystallization and structure determination of trypsin-RCL peptide complex

Porcine pancreatic trypsin (PPT) (Sigma, T4799) was dissolved in 10mM Tris pH 8, 250mM NaCl, with 5% glycerol at 20mg/ml. The inhibitor was dissolved in water at 10mg/ml. Trypsin and inhibitor were mixed in 1:10 molar ratios and incubated for 10 min at room temperature for complex formation. Initial crystallization screening was performed by sitting drop vapour diffusion method at 20°C, using commercial screens. Crystals were obtained with 0.1M HEPES pH 7, 70% MPD. Further, diffraction quality crystals were grown by hanging drop

method at 20°C. To grow crystals at pH 10, trypsin was dissolved in 10mM glycine NaOH buffer with 5% glycerol. These crystals were obtained by hanging drop vapour diffusion method, with 0.1 M CHES pH 10, 70% MPD as precipitant.

Diffraction data were collected at RRCAT Indore, PX-BL21 beamline fitted with MARCCD 225 detector (Kumar et al., 2016). Data were processed with XDS (Kabsch et al., 2010) and scaled with Scala (Evans, 2006). Structures were solved using the PHENIX (Adams et al., 2010) program from CCP4 (Winn et al., 2011) suite of programs. Coot was employed in manual model building. Structures were refined with Refmac(Vagin et al., 2004) and validated with molprobity(Chen et al., 2010). Final data processing and structure refinement statistics are given in table 2.1.

2.2.4 Accession numbers

The atomic coordinates and structure factors for the structure of porcine trypsin in complex with peptides have been deposited in the Protein Data Bank (PDB) (*http://www.pdb.org/*), codes (5XW1:PRN pH10; 5XW8:PRN pH7; 5XWA:PRY pH10; 5XW9:PRY pH7; 5XWL:TRE pH10; 5XWJ:TRE pH7)

2.2.5 Feeding assays and nutritional indexing

Bioassays were conducted by feeding *H. armigera* larvae on an artificial diet containing 50, 100 and 200 ppm of the tripeptides. Three independent sets of n = 30 first instar larvae were used for feeding on control and peptide containing AD. Each larva was placed in a separate vial containing 0.5 g of diet. Insects were maintained at 25 ± 1°C, 16:8 h (light/dark). On alternate days, the larvae were weighed; the amount of fecal pellets produced and diet remaining was recorded. The assay was continued till completion of the life cycle. Further, deformities in pupa and moth formation were observed. Statistical analysis of growth reduction pattern was performed by two-way ANOVA. Asterisks indicate significant differences (*$p < 0.05$; **$p<0.01$;***$p < 0.001$).

Nutritional parameters, namely efficiency of conversion of ingested food (ECI), efficiency of conversion of digested food (ECD) and approximate digestibility (AD) were calculated as described previously(Farrar et al., 1989; Rakesh S. Joshi et al., 2014). Estimation of residual protease activity post feeding was performed by using insect gut extract of *H. armigera* prepared as described above.

Table 2.1: Data collection and refinement statistics

	PRN pH 7	PRN pH 10	TRE pH 7	TRE pH 10	PRY pH 7	PRY pH 10
Data collection and processing						
Beam line	BL21	BL21	BL21	BL21	BL21	BL21
No. of Crystals used	1	1	1	1	1	1
Space Group	$P6_122$	$P6_122$	$P2_12_12_1$	$P2_12_12_1$	$P6_122$	$P6_122$
a (Å)	82.881	83.165	61.435	61.10	82.66	83.165
b (Å)	82.881	83.165	101.311	100.21	82.66	83.165
c (Å)	135.643	135.232	116.585	116.65	134.67	135.23
α(°)	90	90	90	90	90	90
β(°)	90	90	90	90	90	90
γ(°)	120	120	90	90	120	120
Z	1	1	1	1	1	1
Solvent fraction	0.56	0.57	0.67	0.67	0.56	0.57
Resolution limits (Å)[a]	41.50-2.00 (2.06-2.01)	45.05-1.90 (1.94-1.90)	47.89-1.80 (1.84-1.80)	47.62-2.10 (2.16-2.10)	71.5-2.00 (2.05-2.00)	45.08-1.90 (1.94-1.90)
R_{merge}	0.134 (0.460)	0.141 (0.660)	0.078 (1.110)	0.114 (0.722)	0.103 (0.738)	0.114 (0.524)
I/σ (I)	22.4 (7.6)	20.8 (5.7)	13.6 (1.4)	11.3 (2.3)	18.2 (3.1)	17.3 (5.2)
Total number unique	19102 (1868)	22450 (2176)	67984 (6707)	42534 (4201)	19062 (1872)	22450 (2176)
Completeness (%)	99.9 (98.8)	100 (100)	99.8 (99.6)	99.9 (100)	100 (100)	99.9 (100)
Multiplicity	21.3 (20.9)	21.4 (21.7)	4.9 (4.9)	4.9 (4.9)	16.4 (7.3)	10.9 (11.1)
Refinement						
Resolution limits (Å)	39.63- 2.002 (2.074-2.00)	39.75-1.9 (1.968 - 1.9)	39.08- 1.8 (1.864- 1.8)	42.19- 2.1 (2.175- 2.1)	63.21- 2.0 (2.072- 2.0)	39.75 - 1.9 (1.968- 1.9)
Working set	19094	22435	67954	42512	19056	22438
Test set	978	1150	3365	2137	976	1150
R_{work}/R_{free}	0.149/0.178	0.144/0.173	0.162/0.181	0.167/0.194	0.163/0.194	0.152/0.182
No. of Protein atoms	1652	1662	3333	3325	1652	1658
Average B-factor	17.31	17.57	30.14	32.96	25.48	17.72
macromolecules	15.21	15.36	28.64	32.03	24.34	15.58
ligands	24.23	33.48	46.19	54.93	15.59	12.99
solvent	30.74	32.3	42.1	41.22	36.96	32.67
RMSD						
Bond length (Å)	0.006	0.007	0.008	0.007	0.007	0.007
Bond angles (°)	0.78	0.84	0.96	0.85	0.81	0.82
All atom contacts and geometry analysis[c]						
Residues						
in favoured regions	211 (96.97%)	213 (97.26%)	430 (97.95%)	427 (97.27%)	211 (96.79%)	213 (97.71%)
with poor rotamers/bad angles/bad bonds	0/0/0	0/0/0	0/2/2	1/0/0	0/0/0	0/0/0
With Cβ deviation>0.25A/ clashscore	0/ 3.04 (99^{th} percentile)	0/ 4.17 (98^{th} percentile)	0/3.94 (98^{th} percentile)	0/6.36 (97^{th} percentile)	0/2.15 (99^{th} percentile)	0/2.45 (99^{th} percentile)
Molprobity score	1.29 (99^{th} percentile)	1.34 (98^{th} percentile)	1.19 (99^{th} percentile)	1.48 (98^{th} percentile)	1.19 (100^{th} percentile)	1.09 (100^{th} percentile)
PDB ID	**5XW8**	**5XW1**	**5XWJ**	**5XWL**	**5XW9**	**5XWA**

[a]Data processing values in parentheses refer to the outermost resolution shell
[b] Ligand atom is MPD
[c] Calculated by MOLPROBITY

For estimating retention time of fluorescent tripeptides in insect gut, fourth instar larvae (n=5) in three independent sets were starved for 12 h, and then fed carboxyfluorescein tagged tripeptides (200 ppm each) by applying them on agar blocks (0.5 g each). Carboxyfluorescein dye was also fed as the control. Larvae were allowed to feed on peptide-containing blocks, after which their diets were replaced by 0.5 g of agar block. Frass was collected at 1, 3 and 6 h and diluted with equal volume of 0.2 M Glycine-NaOH buffer pH 10. Fluorescence intensities were detected using Promega Glomax spectrophotometer at excitation/emission wavelength of 490/520nm.

2.2.6 Quantitative Real-Time PCR

Total RNA was isolated from the 200 ppm peptide-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 (Table 2.2) was assessed(Dawkar et al., 2013; Mahajan et al., 2013). For each gene, amplification efficiency was assessed by constructing a standard plot using 5 serial dilutions of cDNA pool which were prepared by combining aliquots from all the cDNA samples under study. *H. armigera* GAPDH (Accession No.: JF417983) was used as a reference gene for normalization. Quantitative real-time PCR was carried out in 10µl reactions containing 5 µl of 2x concentrate SYBR mix, 0.5 µl of forward and reverse primer each (10µM) and 1 µl of cDNA (10 µg). Thermal cycler conditions used were 95 °C for 10 min, 40 cycles of 95 °C for 3 s and 60 °C for 30 s. At the end of each run, dissociation curve analysis of the amplified product was analysed. Relative gene expression ratios of protease genes were calculated using $\Delta\Delta C_t$ method(Chikate et al., 2013; Livak and Schmittgen, 2001). For each treatment, two biological replicates (each constituting three technical replicates), the average transcript abundance and sub-sequent fold difference with respect to the control were calculated. Statistical analysis was performed using unpaired t-tests (Livak and Schmittgen, 2001). Asterisks indicate significant differences (*$p < 0.05$; **$p<0.01$;***$p < 0.001$).

Table 2.2: Primer sequences used for qRT-PCR

Name	Genbank ID	Forward primer	Reverse primer
HaTry1	EU982841	GAGGACACAGATGTGGAGGGG	GAACACACGGAATTCAGCCACG
HaTry2	EU770391	GCGTAAAGGATGCGGTTGG	CAGGATGGCAACCATCCATG
HaTry3	EU325548	CGACCACACTGACGCGAG	GCACGCCACTGGACATGG
HaTry4	EF600059	GTGCTACCCCTTCTGATTC	AACTTGTCGATGGAGGTGAC
HaTry5	EF600054	GGTCTCTGCTAACCTCCACC	CTGGATGCCAGGGACGTGC
HaTry6	Y12276	TGGCTGGGGTGACACTTTCT	GTCTCCCTGGCACTGGTC
HaTry7	Y12271	CAGAGGATTGTGGGTGGTTCG	GCGGTGAGGATAGCCCTGTT
HaTry8	Y12286	GGGCTACTGGTGCCTTCAACG	CAGAGTCATACACGTCACCGACG
HaChy1	HM209422.1	CGACTTGTCAGGTGGTCAGGCTG	GCGATTCTGGTACCGCCGGAGAAC
HaChy2	EU325550.1	GACTTGTCAGGTGGCCAGGCTG	GCGATTCTGGTACCGCCGGAGAAC
HaChy3	GU323796.1	TGACTTGTCAGGTGGCCAAGCTG	GCGATTCTGGTACCGCCGGAGAAC
HaChy4	Y12273	CACCATCTTCATCTTCCAATCCGTGTGC	GTGTTGATACGAGTACCACCGAAGAAC
HaGAPDH	JF417983	TGCTGAATACGTCGTTGAATCC	TTCTTAGCACCACCCTCTAAATGAG

2.2.7 MALDI TOF based assays

Stability of RCL tripeptides in presence of HGP was monitored by MALDI-TOF analysis(Mishra et al., 2010). Reduction in the intensity of the RCL peptide is monitored by MALDI-TOF-MS on the addition of target protease. RCL tripeptides were incubated with HGP for 1h, 3h and 6h at 37°C. The reaction mixture of 5 µl volumes was mixed with 20 µl of freshly prepared DHB matrix for tripeptides and 2 µl aliquots in three replicates were spotted on the stainless steel plate and MALDI-TOF profiles were acquired. Also, serine protease activities were tested at respective times to check the stability of the protease-inhibitor complex.

2.3 Results and discussion

2.3.1 Synthesis of reactive loop peptides

Six peptides, namely, TRE, PRN, PRY, PKN, PLN and TLN were synthesized using F-moc solid phase peptide synthesis, with N-terminal acetylation and C-terminal amide. Peptides were >90% pure as indicated by the HPLC spectra (Figure 2.3). Identity of the peptides was confirmed by molecular weights obtained from MALDI-TOF spectra (Figure 2.4). The 2D structures of the synthesized peptides are shown in figure 2.5.

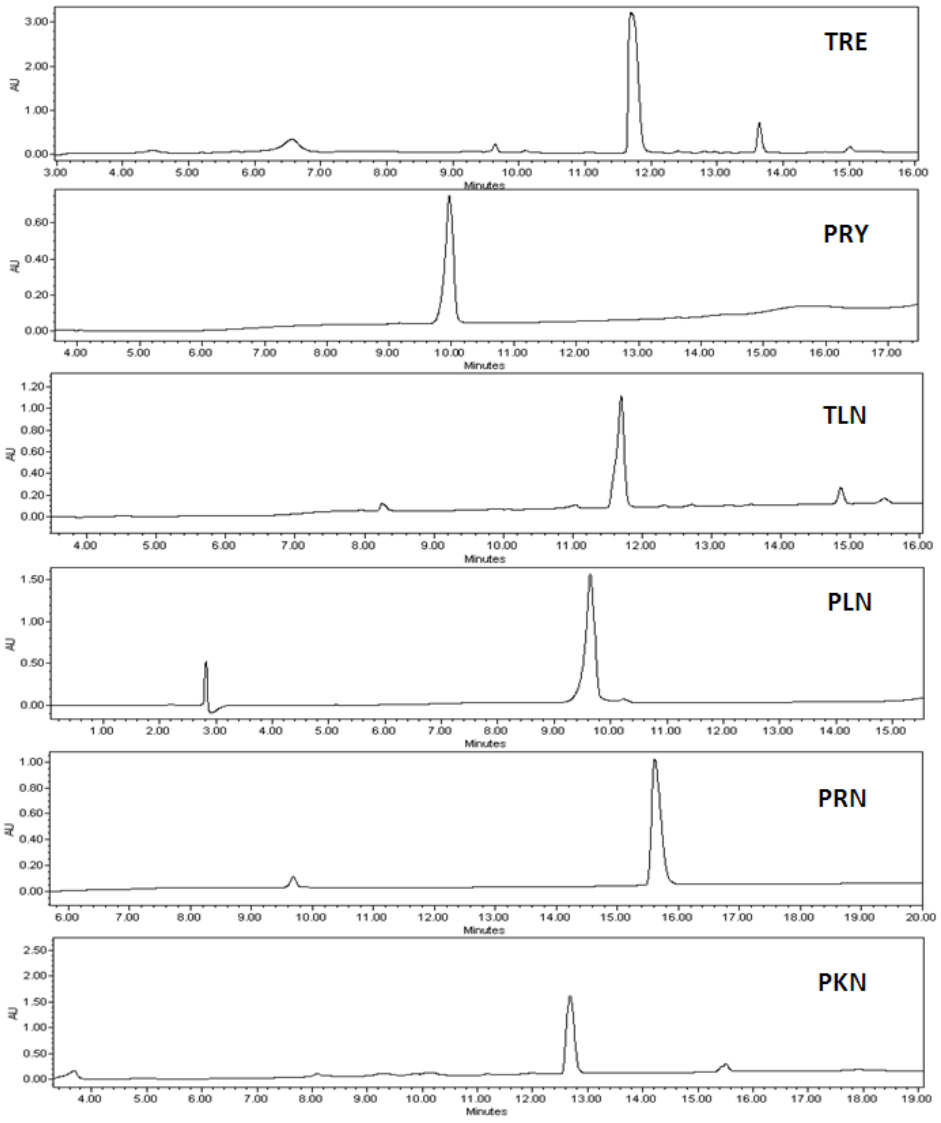


Figure 2.3: Purification of synthesized RCL tripeptides. Synthesis was carried out by F-moc chemistry, and purified by RP- HPLC using linear gradient of 5% ACN/water to 50% ACN/water.

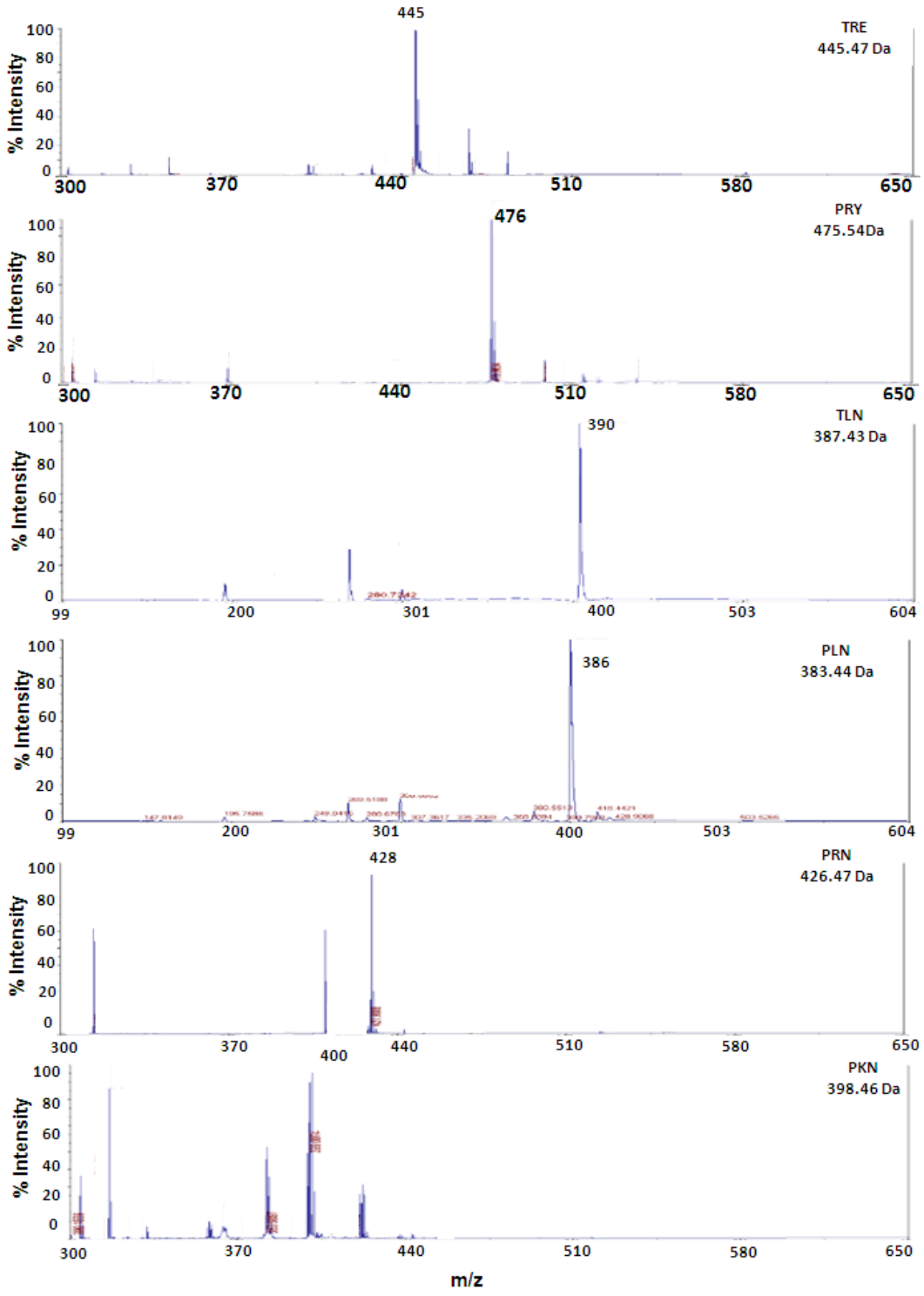


Figure 2.4: Molecular weight confirmation of synthesized peptides by MALDI-TOF MS. Observed and calculated molecular weights are mentioned alongside the peaks.

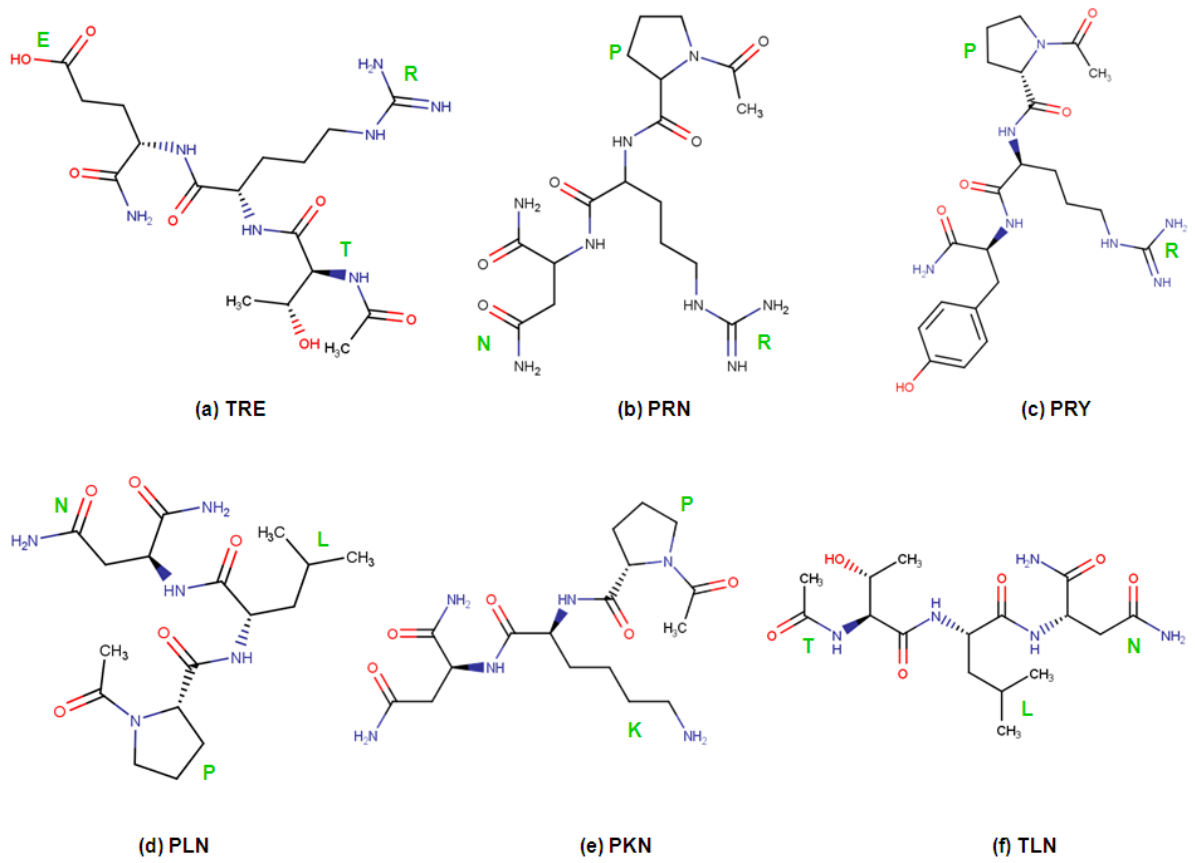


Figure 2.5: 2D structures of synthesized peptides

2.3.2 Biochemical assessment of inhibition potential

Biochemical evaluation of inhibitory activity was performed using chromogenic substrates for trypsin and chymotrypsin, namely, BA*p*NA and SAA*p*NA. It was observed that the RCL peptides are effective protease inhibitors (Figure 2.6). Among the six RCL peptides, TRE, PRN and PRY exhibited remarkable inhibition of trypsin and chymotrypsin like proteases in the *H. armigera* gut extract (up to 95% reduction) with IC_{50} values ranging from 50 to 200 µM (Table 2.3). In contrast to the results obtained for insect proteases, the RCL peptides were significantly less effective in inhibiting bovine and porcine enzymes (Figure 2.6). We observed that the peptides are 5 to 10 times less effective against mammalian enzymes. This showed that the RCL peptides preferentially bind to HGPs with higher affinity. This preferential inhibition is beneficial characteristic with respect to development to dietary pest control molecules.

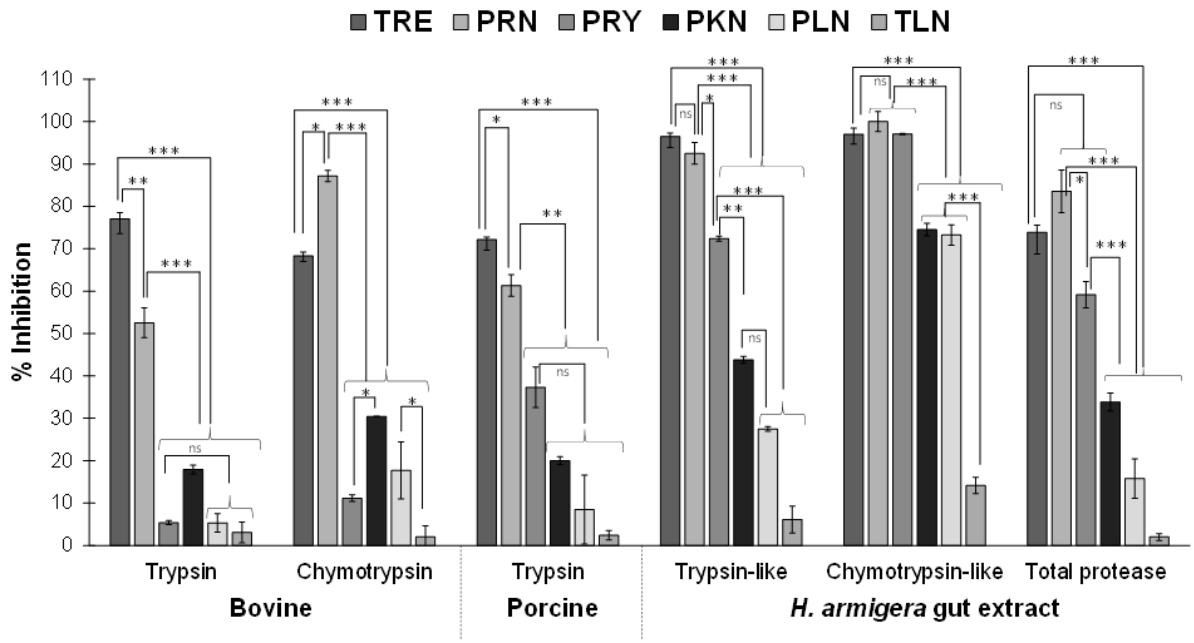


Figure 2.6: Inhibition of serine proteases by RCL tripeptides. Screening of RCL peptides for inhibition of bovine trypsin, bovine chymotrypsin, porcine trypsin and *H. armigera* gut extract was performed at effective peptide concentration of 1mM, at 37°C in pH 10. Asterisks indicate significant inhibition as compared to other peptides, analysed by one way ANOVA and unpaired t-test from three independent assays (*p<0.05, **p<0.01, ***p<0.001), selected peptides are indicated in box.

Table 2.3: IC_{50} values (mM) determined for inhibition of bovine and *H. armigera* gut protease activities, and Ki values (mM) determined for inhibition of bovine proteases at pH 10.

IC_{50} (mM)				
	Bovine		***H. armigera* gut extract**	
	Trypsin	**Chymotrypsin**	**Trypsin-like**	**Chymotrypsin-like**
TRE	0.8	0.9	0.2	0.05
PRN	1	0.5	0.04	0.05
PRY	5	4	0.2	0.2
Ki (mM)				
TRE	0.4	0.2		
PRN	0.7	0.2		

Because it is unamiable to produce purified trypsin(s) or chymotrypsin(s) from *H. armigera*, thus, in order to study the binding kinetics of the selected RCL peptides, trypsin and chymotrypsin from bovine source was used as surrogate system. TRE and PRN exhibited

limited inhibition of bovine trypsin while PRY did not show substantial inhibition. Therefore, TRE and PRN were studied for analysis of binding kinetics. The type of inhibition was determined from saturation curves with increasing substrate concentrations and then plotting double reciprocal plots. As shown in figure 2.7, both peptides are competitive inhibitors of trypsin and chymotrypsin. The Ki values obtained for TRE were 0.4 and 0.2 mM (trypsin and chymotrypsin), while those for PRN were 0.7 and 0.2 mM (trypsin and chymotrypsin).

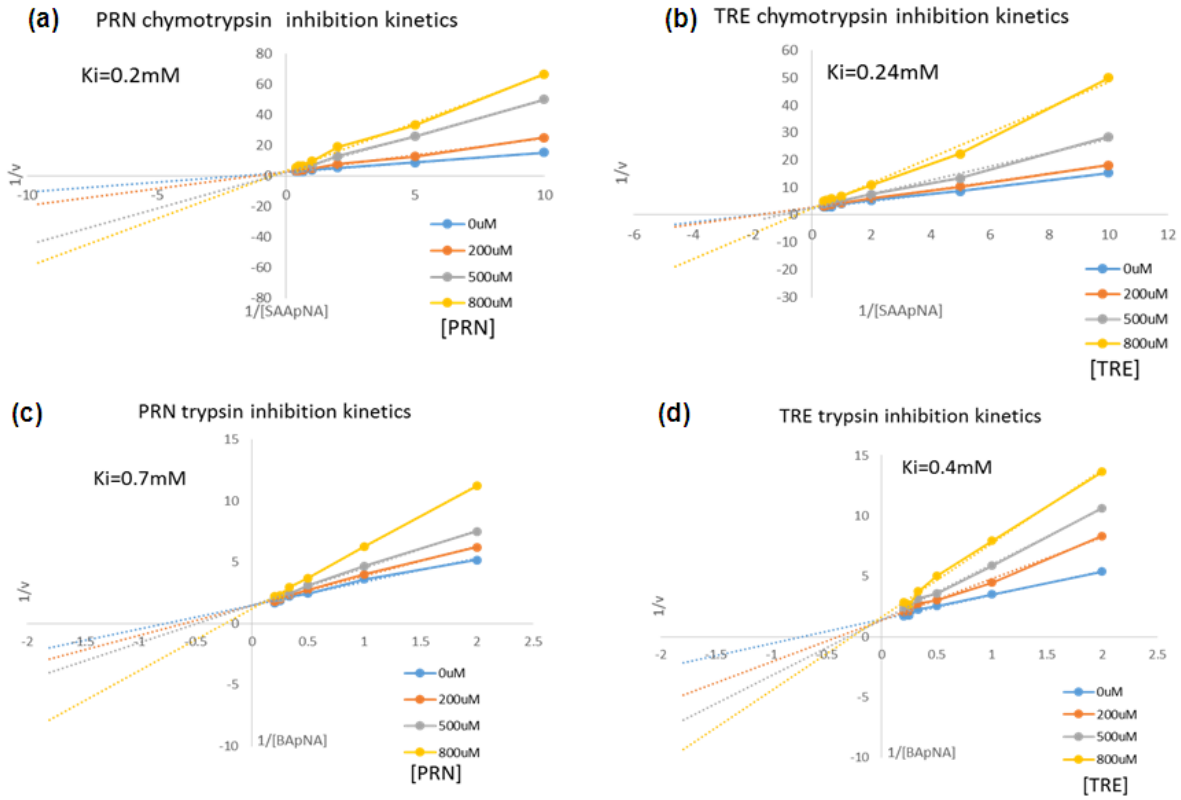


Figure 2.7: Double reciprocal plots for inhibition kinetics of peptides with bovine trypsin and chymotrypsin. Peptides (0, 200, 500, 800 µM) were assayed with increasing concentrations of substrates (BApNA: 0.5-5 mM, SAApNA: 0.2-2.5 mM). Ki values were calculated by fitting the data was into non linear fit for competitive inhibition in Graph Pad Prism 6.

Furthermore, we evaluated the inhibitory activity of RCL peptides at both neutral and alkaline pH, to check whether pH affects the potency of these peptides. Biochemical assays were performed at pH 7.8 (optimum pH for bovine and porcine trypsins), as well as pH 10 (pH of insect midgut). Surprisingly, we found that except for TRE, all the RCL peptides showed remarkably lower inhibition potential at pH 7.8, highlighting that the RCL peptides are not only

selective towards midgut proteases but also sensitive to change in the environmental pH (Figure 2.8).

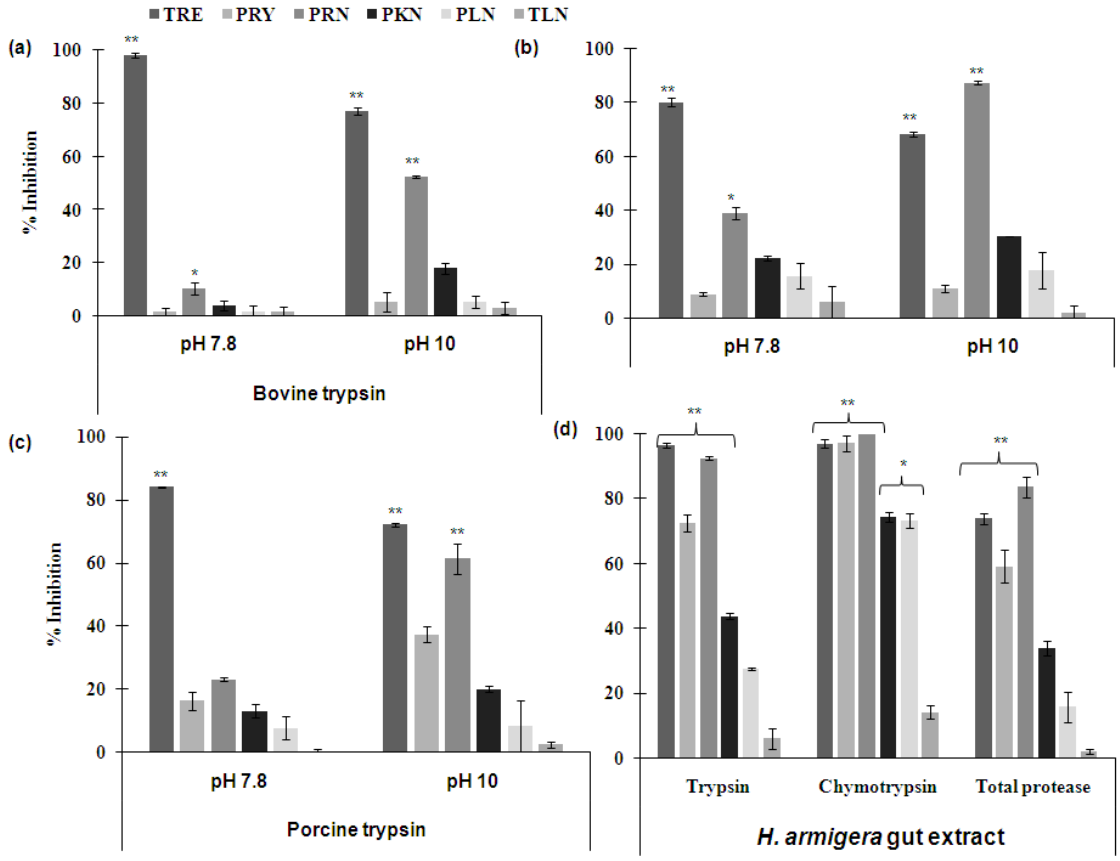


Figure 2.8: Inhibition of serine proteases by RCL tripeptides. Bar graph representing protease inhibition assessed at pH 7.8 and 10 for (a) bovine trypsin, (b) bovine chymotrypsin, (c) porcine trypsin, and (d) *H. armigera* gut extract using specific chromogenic substrate BApNA and SAApNA for trypsin and chymotrypsin respectively. Total protease activity was estimated by azocasein as substrate. All tripeptides were used at effective concentration of 1mM. Asterisks indicate significant inhibition with respect to other tripeptides as analysed by unpaired t-tests from three independent assays (*$p<0.05$, **$p<0.001$).

2.3.3 Structural basis of protease inhibition by RCL tripeptides

For unravelling the mechanism of binding of the RCL tripeptides, we analysed the 3D crystal structures of peptides (TRE, PRN and PRY) in complex with porcine pancreatic trypsin (PPT). Crystallization was performed using hanging drop method, and it was observed that the crystals of PPT-PRY and PPT-PRN complexes lied in hexagonal space group with $P6_122$ symmetry, while the PPT-TRE complex crystallized with $P2_12_12_1$ orthorhombic space-group (Figure 2.9). We crystallized the complexes at both pH 7 as well as pH 10 to understand the

difference in interactions. All of these crystals diffracted with resolution in the range of 1.8 to 2Å.

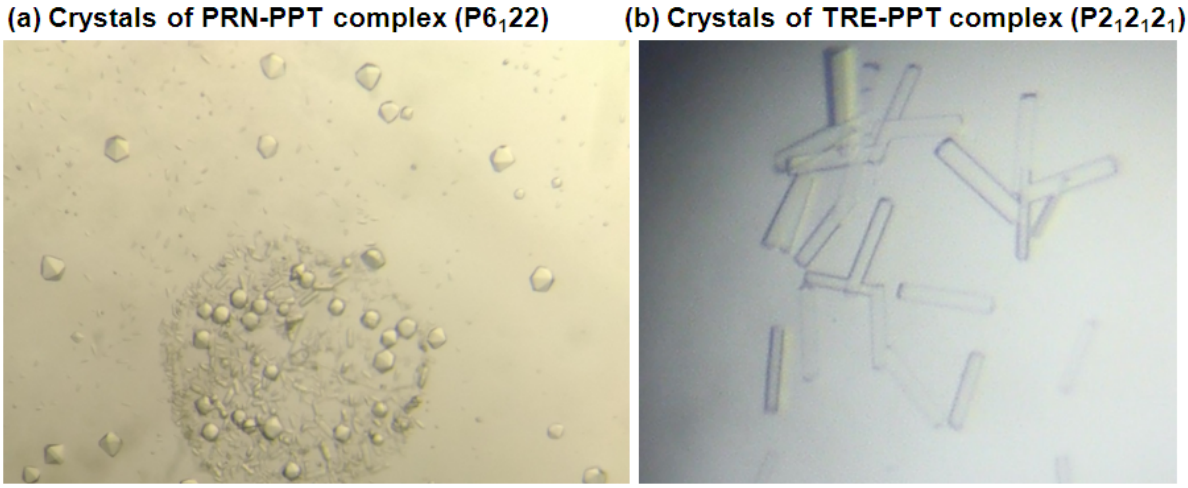


Figure 2.9: Crystals obtained for (a) PPT-PRN and (b) PPT-TRE complexes. Crystallization was performed using hanging drop vapor diffusion method with 0.1M HEPES pH 7, 70% MPD as precipitant.

Upon superposition of the structures, it was seen that the conformation of trypsin remains identical except for the loop comprising residues 131 to 141 (Figure 2.10). These variations might be because of crystal packing as these residues are involved in crystal contacts. For the peptides, we observed clear electron density for P2 and P1 residues in all of the structures, i.e TR in case of TRE and PR in case of PRN and PRY. However, the electron density for the P1' residue (E in TRE, N in PRN and Y in PRY) as well as the scissile bond was uninterpretable (Figure 2.11). Analysis of binding mode of the RCL peptides revealed that they interacted with PPT similar to the canonical Pin-II type protease inhibitors(Isabelle H Barrette-Ng et al., 2003; Bode and Huber, 2000). The side chain of P1 residue, Arg403, was deeply seated into the S1 pocket of PPT and made contacts with specificity determining residue Asp 179 (Figure 2.12a). The P2 and P1' residues of the peptide were placed nearby the active site (His48, Asp92, Ser185) and oxyanion hole (Gly 183).

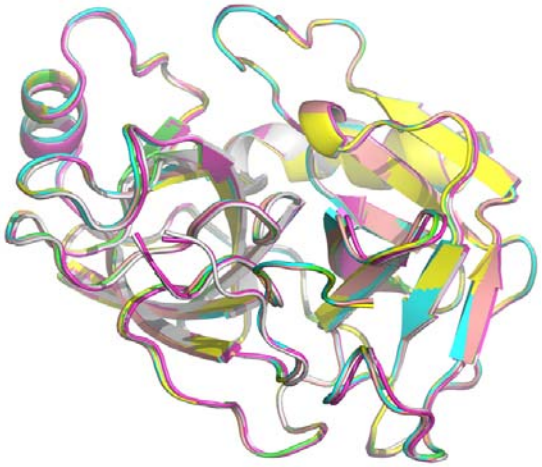


Figure 2.10: Superimposition of PPT-peptide complex structures (PPT-PRN:Yellow, PPT-PRY:cyan, PPT-TRE:magenta, PPT:Green). The overall conformation of PPT remains same for all structures. RMSD values for alignments are mentioned in table 2.4.

Table 2.4: RMSD values for alignment of PPT-peptide structures, performed using PyMol. Porcine trypsin structure (PPT) was downloaded from PDB (4doq).

Protein-1	Protein-2	RMSD (Å^2)
TRE-PPT	PPT	0.2
PRN-PPT	PPT	0.287
PRY-PPT	PPT	0.286
PRY-PPT	PRN-PPT	0.023
PRY-PPT	TRE-PPT	0.249
PRN-PPT	TRE-PPT	0.251

The interaction with Arg side chain was held by a network of hydrogen bonds with Asp179, Gly204 and Ser180 of PPT. Also, the amide groups in the polypeptide backbone were involved in hydrogen bonding with Gln182, Gly183, Ser 185, and Ser200, which lie in close proximity to active site (Figure 2.12b). Also, the N-terminus acetyl caps of the RCL peptides made hydrogen bonds with backbone nitrogen of Gly202 (Figure 2.12c). The P1 residue, Arg403 interacts with Gln182 and Gly216 of PPT through a water molecule. Thus, the protease-tripeptide interaction is stabilised mainly by the P1 residue (Arg403). Also, if this Arg residue is replaced by a hydrophobic residue like Leu, it might reduce the crucial interactions, and hence, the potency of the inhibitors for trypsin-like enzymes. Further, comparison of peptide bound structures at pH 7 and pH 10 showed that the RCL peptides adopt a similar orientation at both

pH (Figure 2.13). Detailed analysis of peptide binding modes also gave similar interacting residues with PPT, indicating that the bound conformation of RCL peptides is not dependent on pH.

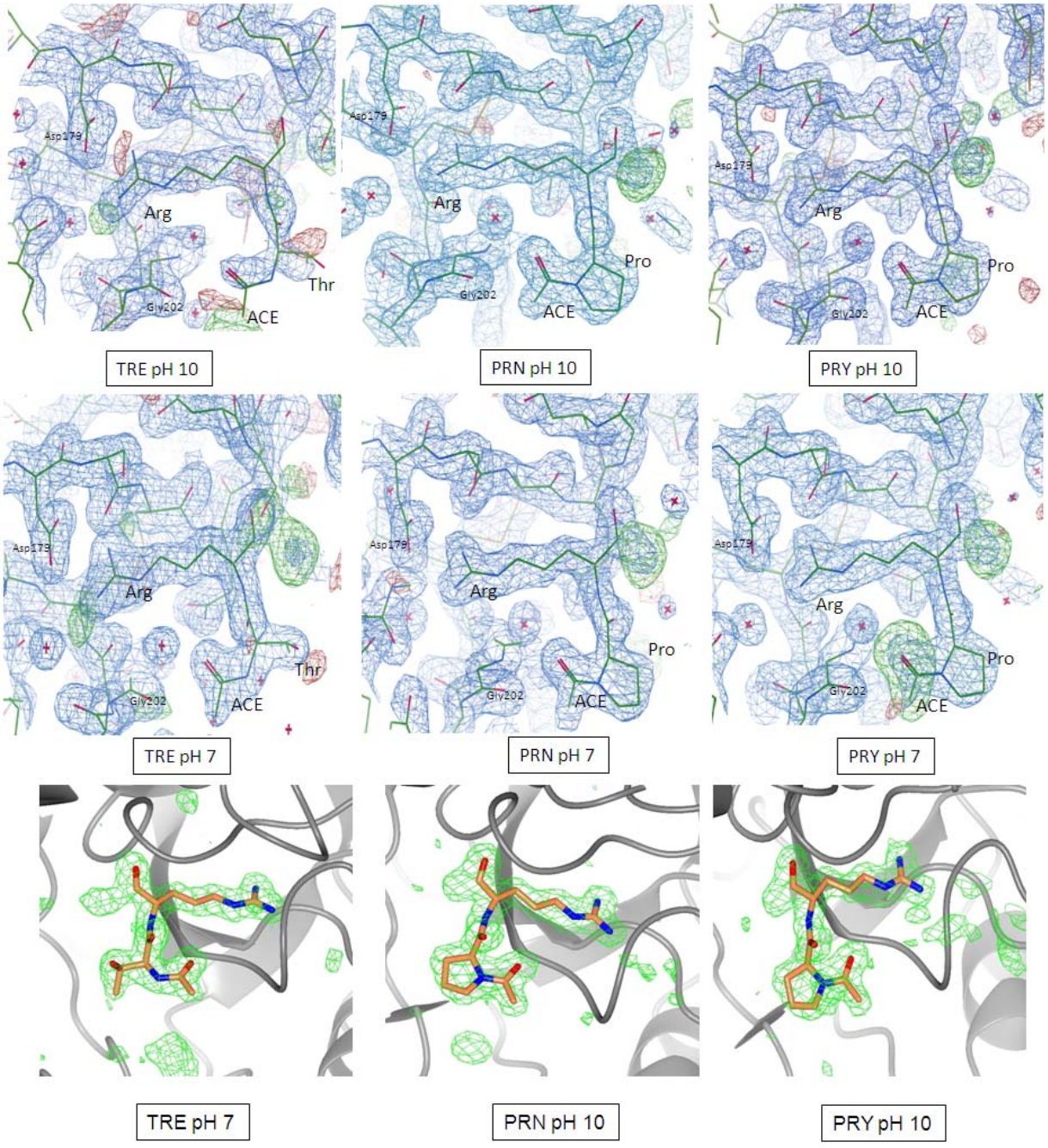


Figure 2.11: Electron density maps for RCL peptides in complex with PPT. 2Fo-Fc density (blue), Fo-Fc positive density (green) and negative density (red) contoured at 1σ.

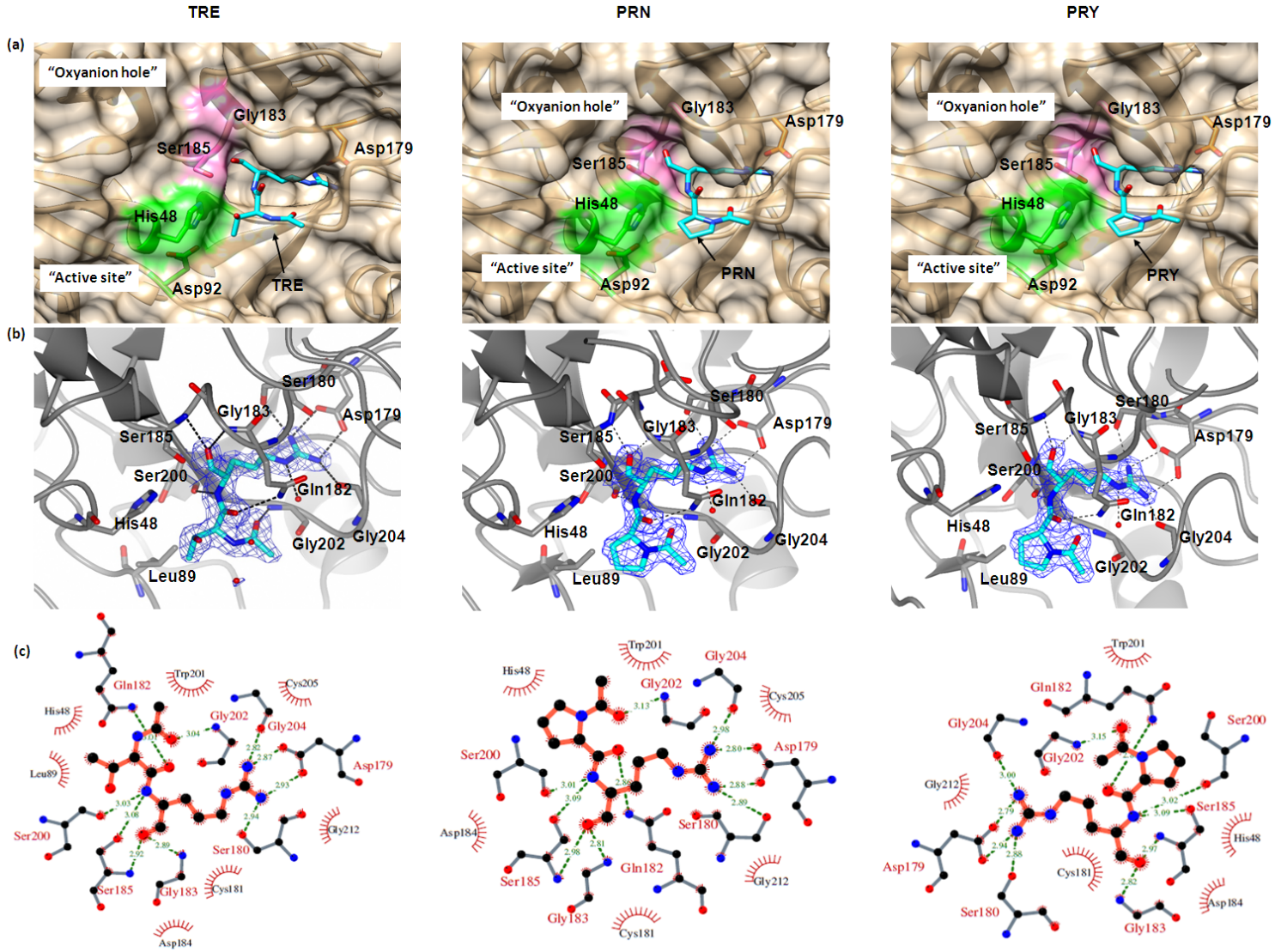
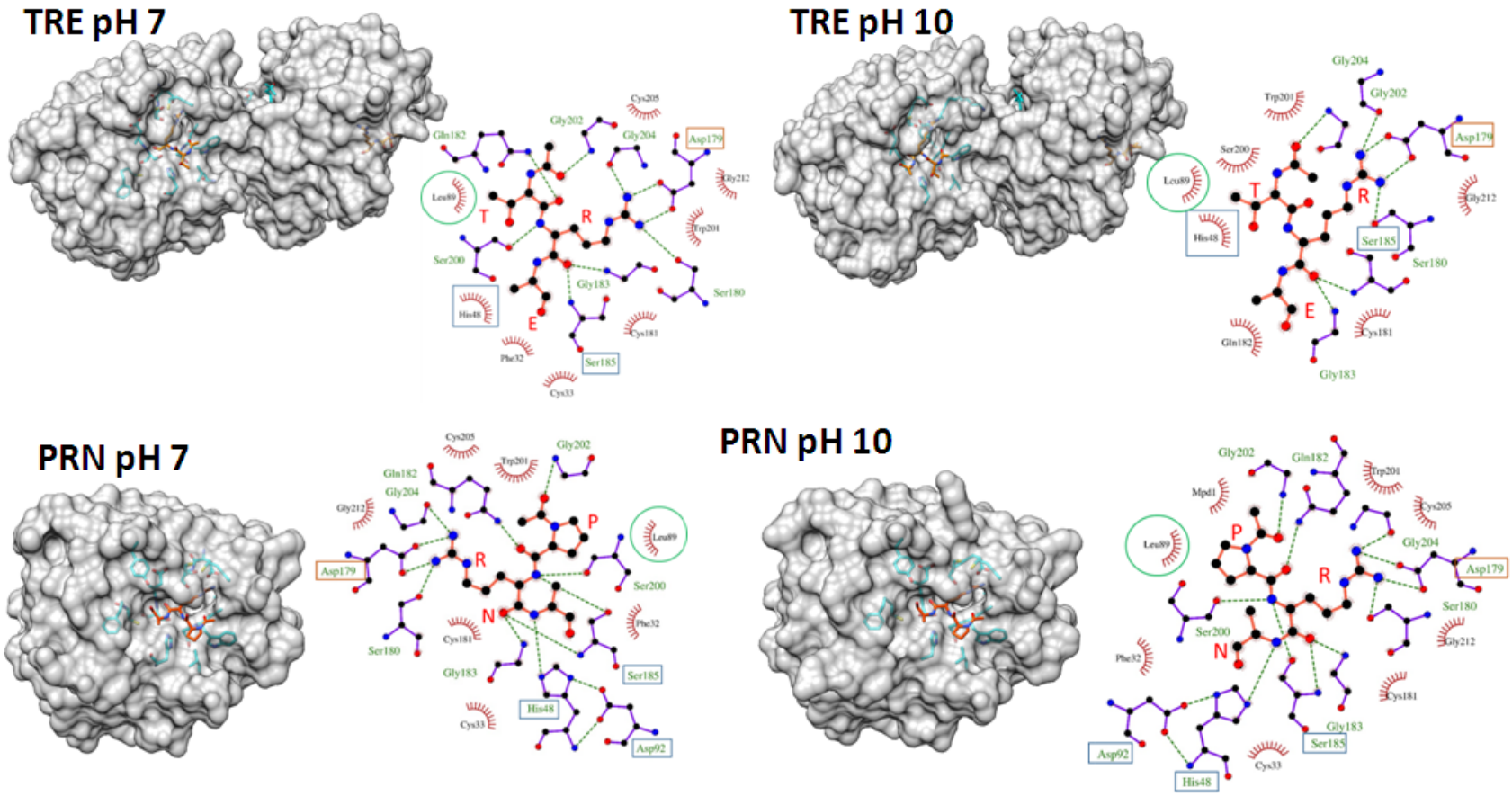


Figure 2.12: Structural basis of binding RCL-protease inhibition (a) View of the binding pocket of PPT in complex with RCL peptides (cyan) with active site and oxyanion hole indicated in green and pink, respectively. Structures ofTRE-PPT at pH 7, PRN-PPT

and PRY-PPT at pH 10 are shown(b) Representation of the RCL peptide binding site showing 2Fo-Fc electron density (blue mesh) contoured at 1σ around the RCL region. RCL peptides and interacting PPT residues are represented as sticks. Hydrogen bonding interactions are indicated (black dotted lines). Water molecules involved in complex stabilization are shown as red spheres. (c) Simplified depiction of the interactions between PPT and RCL peptides generated using Ligplot+. Red circles represent hydrophobic interactions whereas green lines represent polar interactions.



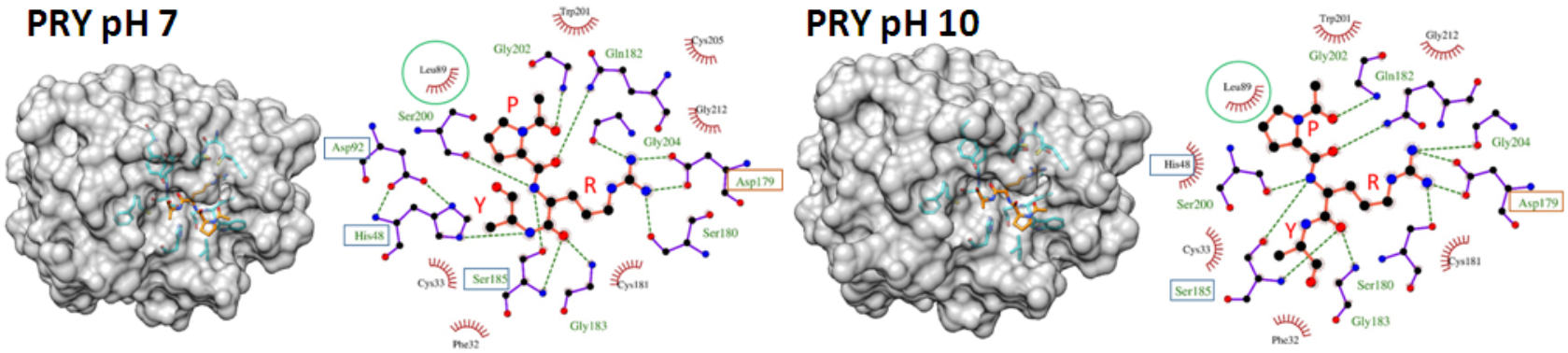


Figure 2.13: Binding modes of RCL peptides at pH 7 and 10. Surface representation of PPT bound to peptides. Peptides and binding site residues are shown as orange and cyan sticks respectively. 2D interactions of peptides is shown alongside the surface. Active site residues are marked in box, Leu89 hydrophobic residue is circled.

2.3.4 Crucial role of P2 residue in estimating the relative potency of RCL peptides

Since TRE, PRN and PRY differ in their potency towards trypsin, we identified the basis for this differential potency by superposition of the PPT-RCL peptides. It was observed that, in addition to the interactions made by P1 (Arg403) residue, the side-chain methyl group of threonine (Thr402) at the P2 position makes hydrophobic interactions with Leu89, which is crucial in enhancing the potency (Figure 2.14a). Whereas, in case of PRN and PRY, substitution of Thr402 with proline leads to increase in the distance between the side chains of P2 and Leu89 by ~1Å, thereby reducing the strength of P2-Leu89 hydrophobic interaction significantly (Figure 2.14b). Possibly, threonine is the optimal residue at the P2 position for efficient inhibition of trypsin-like proteases, because it stabilizes the hydrophobic interactions with protease and simultaneously makes hydrophilic interaction with the solvent.

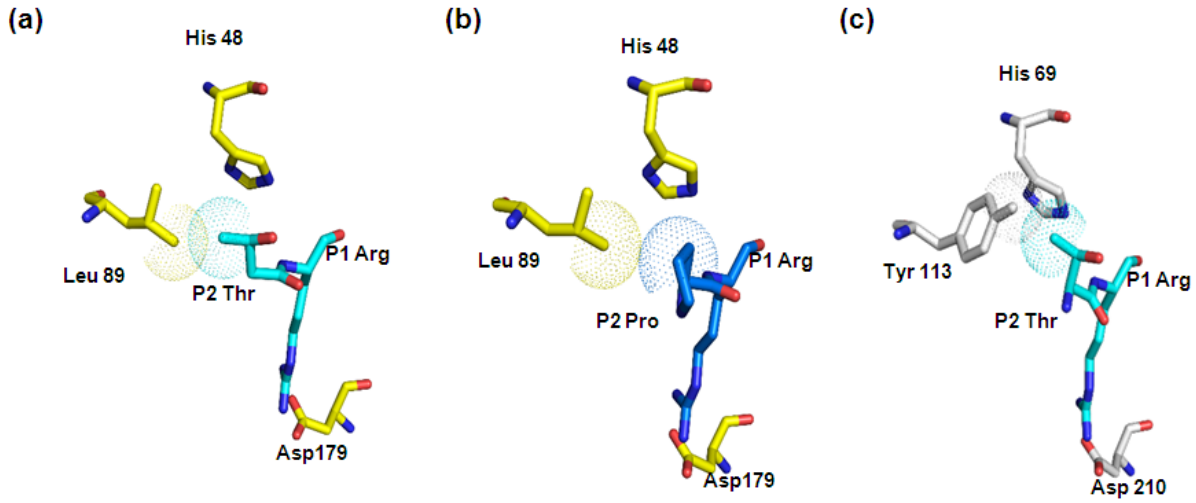


Figure 2.14: Selectivity of RCL peptides for *H. armigera* trypsin (a) Hydrophobic interaction made by P2 residue, Thr402 in TRE and (b) Pro402 in PRN and PRY with Leu89 of PPT, which enhanced in H. armigera trypsin 4 model (c) due to substitution of Leu89 with Tyr113.

Further, as indicated by biochemical inhibition assays, the RCL peptides are more active against trypsin-like proteases from insect gut. Thus, to study the structural basis for this preferential inhibition for *H. armigera* trypsin, we modelled the structure of *H. armigera* trypsin4 (HaTry4) (Uniprot ID: B1NLE4), which is known to be active and highly expressed in the insect gut (Chikate et al., 2013). Upon superimposition of this structure with PPT-TRE and PPT-PRN structures, we observed that, in case of HaTry4, Leu89 is substituted with a bulkier hydrophobic amino acid, Tyr 113 (Figure 2.14c). This might lead to enhancement in the potency of the inhibitor due to more stable hydrophobic interaction, even if a shorter hydrophobic amino acid residue like Pro is present at the P2 position. Multiple sequence alignment of putative trypsin-like proteins from *H. armigera* shows conservation of hydrophobic residues at the position Leu89 (Figure 2.15). Thus, it shows that plant PIs might be evolved in such a way, so that their interaction with insect proteases are stabilized with a finer balance of hydrophobic and hydrophilic interactions with P2 and P1 residues.

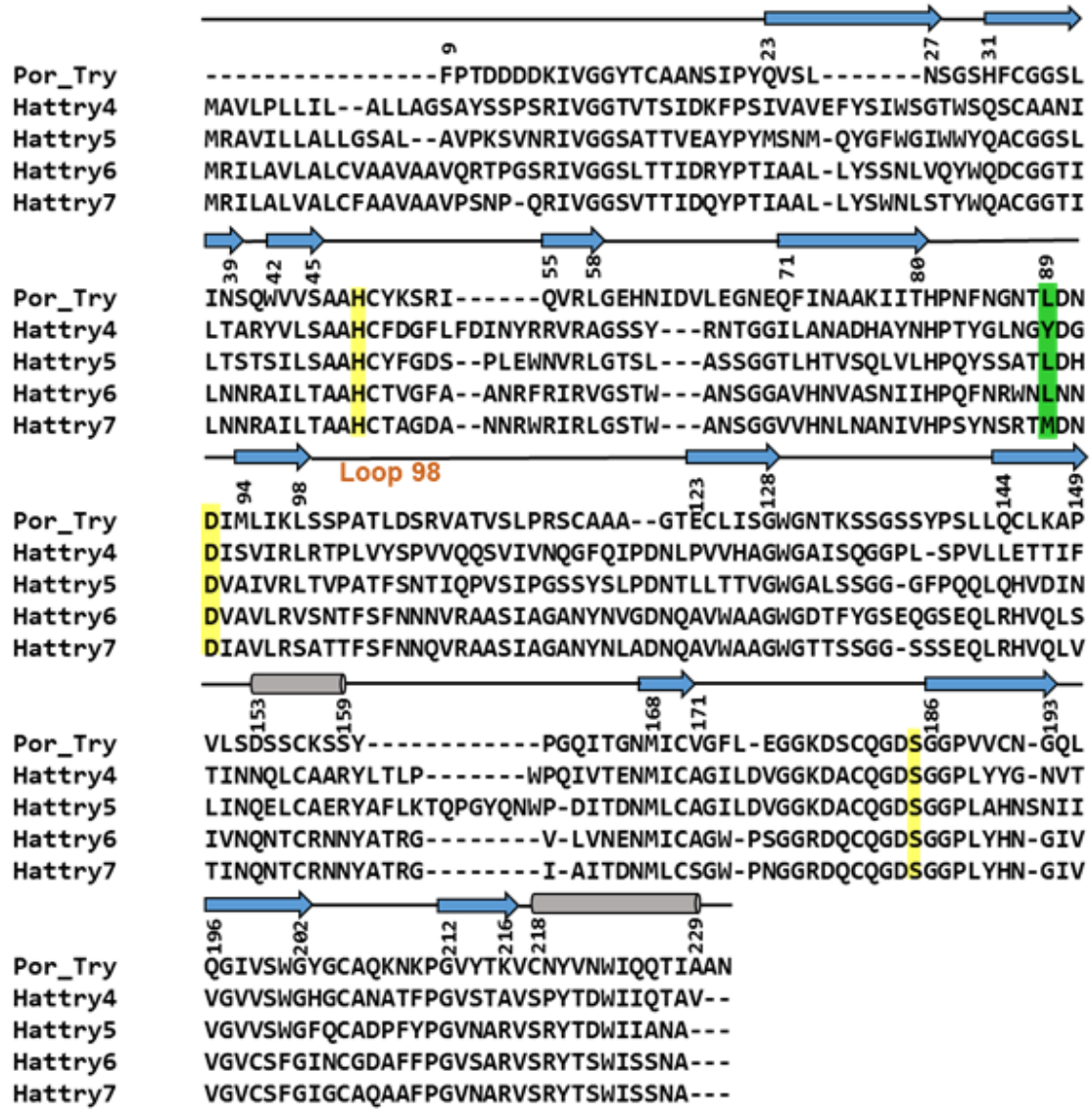


Figure 2.15: Sequence alignment of *H. armigera* trypsins with Porcine trypsin. Secondary structure elements are shown as arrows (β-sheets), cylinder (α-helices). Active site (His48, Asp 92, Ser 185) is marked in yellow, Leu89 and corresponding hydrophobic residues are highlighted in green. Accession nos. Hattry4: ABU98624.1, Hattry5: ABU98619.1, Hattry6: CAA72955.1, Hattry7: CAA72950.1

2.3.5 Reduction in *H. armigera* growth upon feeding with RCL tripeptides

In vivo feeding experiments on *H. armigera* larvae showed that starting from first instar, there was a significant reduction in the growth of insects which were fed with artificial diet

containing RCL tripeptides (TRE, PRN and PRY). The growth reduction was prominently observed on the 10th day of RCL feeding, wherein the insects showed growth reduction of 40% upon feeding with 200 ppm peptides. A dose dependent reduction in larval mass was observed from 50 to 200 ppm (Figure 2.16). Further, a reduction in survival rate of the larvae by ~10 to 15% was caused by the RCL peptides. Between the selected tripeptides, PRN showed the highest negative effect on larval growth, followed by TRE and PRY. Interestingly, when the peptides were fed in combinations, it led to higher deleterious effect (~10% increase in growth reduction) as compared to peptides fed individually (Figure 2.17). Also, an adverse effect on insect moulting and metamorphosis was observed upon RCL ingestion as indicated by 10 to 15% pupal deformities and delayed eclosion. Approximately 25% moths were deformed, as identified by the presence of curled wings, deformed appendages and small body size.

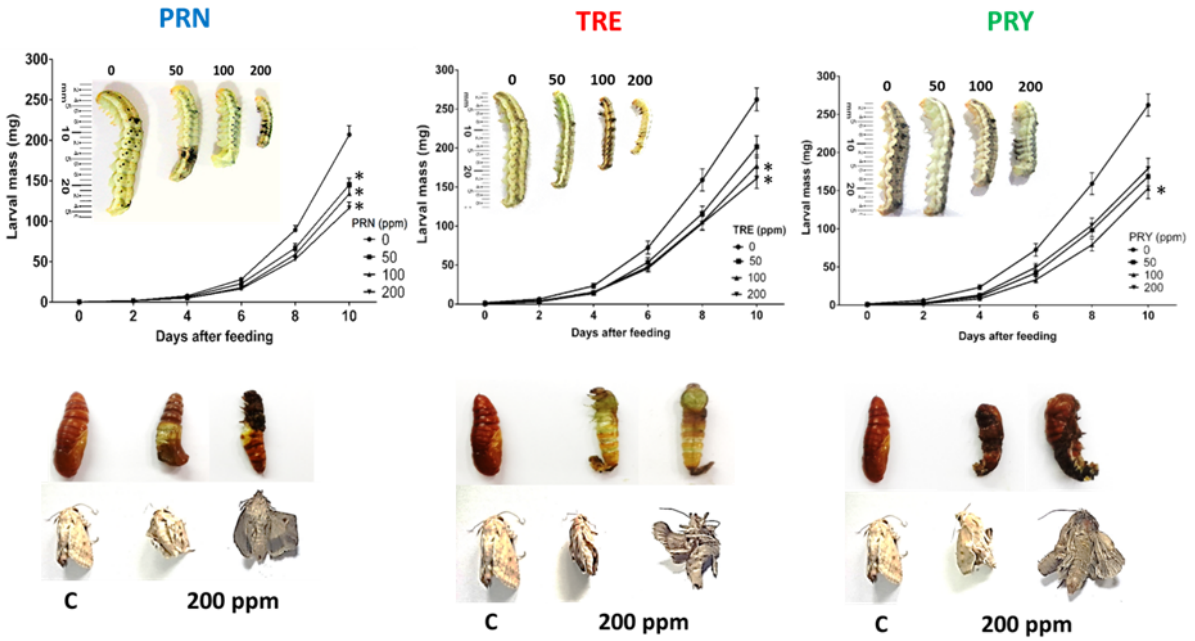


Figure 2.16: Effect of RCL peptides on growth and development of *H. armigera*. Decrease in mass upon feeding larvae (n=90) on artificial diet containing 50, 100 and 200 ppm of TRE, PRN and PRY individually, also deformities observed during development.

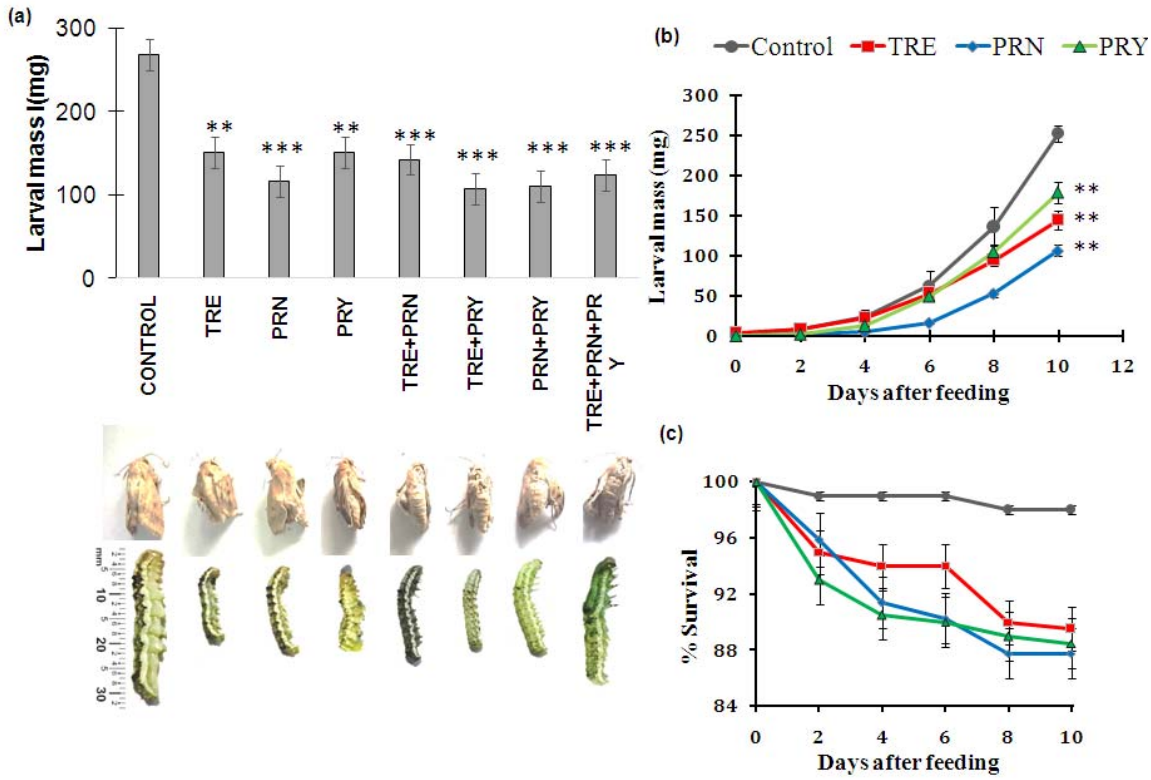


Figure 2.17: Physiological effect of RCL peptides on *H. armigera* larvae (a) Decrease in mass upon feeding larvae (n=90) on artificial diet containing 200 ppm of TRE, PRN and PRY individually and in combinations, also deformities observed during development (b) Difference in growth observed on feeding first instar H. armigera larvae (n=90) on artificial diet containing 50,100 and 200 ppm of TRE, PRN and PRY for 10 days (c) Reduction in insect survival during feeding experiment

Further, calculation of nutritional parameters, which are indicative of digestion efficiency was done. The values obtained for efficiency of Conversion of Ingested Food (ECI), Efficiency of Conversion of Digested Food (ECD) and Approximate Digestibility (AD) showed that these parameters were reduced by 10 to 20% upon peptide treatment. These observations clearly indicate a negative impact of RCL tripeptides on insect growth and digestive metabolism (Farrar et al., 1989) (Table 2.5). The growth inhibitory effect of RCL peptides on *H. armigera* larvae is comparable to that observed for feeding with whole Pin-II PI or IRDs, indicating that RCL region of Pin-II PI is primarily responsible for the protease inhibition in larval gut.

Table 2.5: Effect of tripeptides on dietary utilization of *H. armigera*. ECI: Efficiency of conversion of ingested food; ECD: Efficiency of conversion of digested food; AD: Approximate digestibility

	ECI	ECD	AD
Control	24.9 (±3.24)	48.0 (±2.0)	70.1 (±1.5)
TRE	11.4 (±2.8)	21.7 (±5.9)	52.6 (±5.0)
PRN	18.6 (±1.8)	25.3 (±1.8)	41.2 (±3.4)
PRY	17.3 (±2.1)	34.0 (±5.2)	50.6 (±9.0)
TRE+PRY	13.4 (±2.5)	25.2 (±2.0)	53.3 (±5.0)
PRN+PRY	18.6 (±2.1)	33.9 (±1.8)	42.5 (±3.4)
PRN+TRE	16.5 (±2.2)	35.7 (±4.5)	36.2 (±9.0)
TRE+PRN+PRY	16.8 (±3.2)	34.0 (±2.3)	49.3 (±1.5)

The inhibition of insect gut proteases by RCL peptides was evident in the biochemical activity assessment for serine proteases. It was observed that the residual protease activity showed up to ~40% reduction in trypsin, chymotrypsin and total protease activity (Figure 2.18). These observations show that TRE, PRY and PRN bind to insect proteases *in vivo*, resulting in the inhibition of proteolytic digestion required for the growth and development of *H. armigera*. Furthermore, the peptides were tagged with fluorescent dye, carboxyfluorescein to trace their retention in insect gut. Time dependent fluorescence observed in the insect faeces was considered indicative of retention time in the insect gut. Carboxyfluorescein tagged TRE was detected in the faeces of third instar larvae up to 6h, followed by PRN, which was retained for more than 3h. This indicates that the peptides are retained in the insect midgut (Figure 2.18). Whereas, carboxyfluorescein dye was excreted within 1h. Thus, the changes in developmental cycle of the insects may lead to reduction in their fecundity in long term(Ryan, 1990).

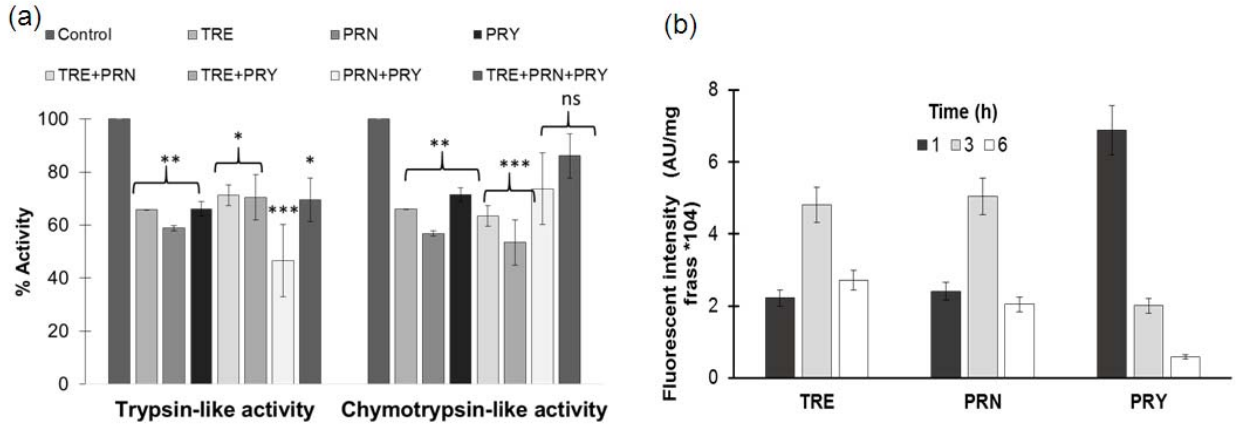


Figure 2.18: In vivo interaction of RCL peptides with insect gut. Reduction in residual gut serine protease activity (c) Retention time of tripeptides in insect gut, as measured by residual fluorescence intensity in frass after 1, 3 and 6 h after feeding carboxyfluorescein-tagged tripeptides

2.3.6 Effect of RCL tripeptides ingestion on *H. armigera* midgut physiology

Insects are known to regulate the expression of proteases upon feeding with protease inhibitors. Therefore, in order to understand the effect of RCL peptide feeding on the midgut protease expression, we performed gene expression analysis of representative trypsin and chymotrypsin genes by q-RT-PCR. It was observed that majority of the trypsin genes were downregulated in response to insect feeding, except for HaTry1, which was upregulated upon exposure to all three RCL tripeptides. Also, chymotrypsin genes showed reduced expression upon RCL peptide feeding (Figure 2.19). It implies that the tripeptides cause inhibition of major serine proteases, in response to which insects alter their expression. The results obtained for downregulation of protease transcripts correlate with the reduction in residual gut protease activity, thus validating the *in vivo* potential of the RCL peptides.

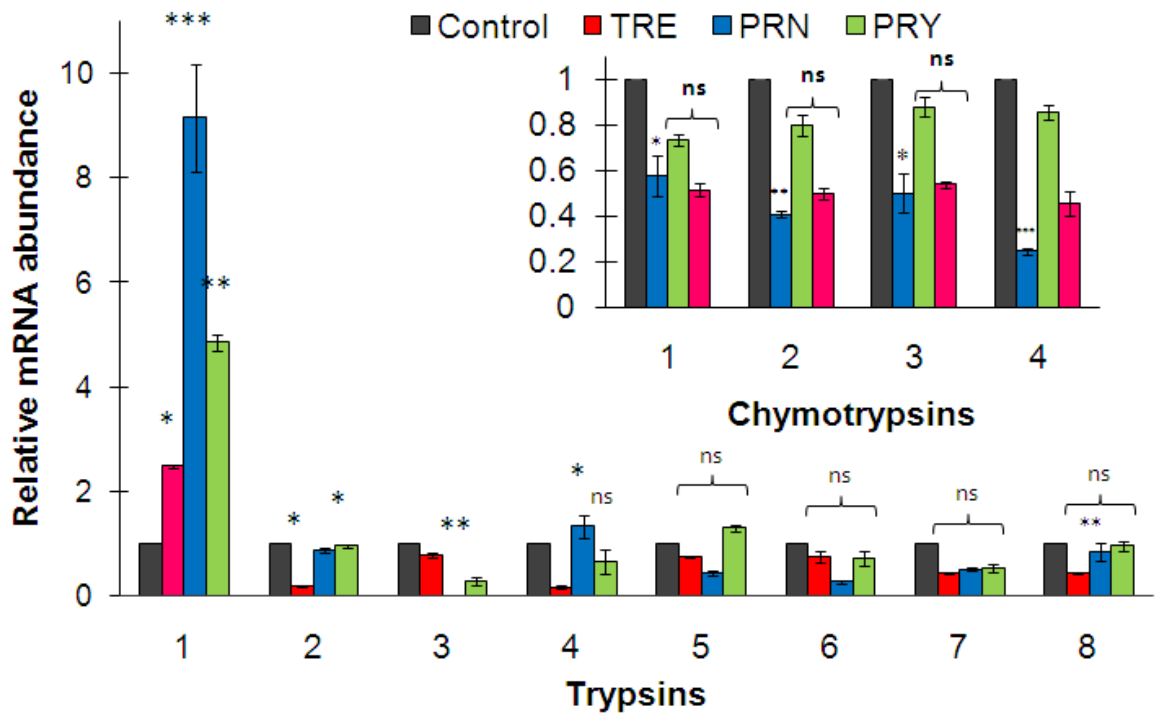


Figure 2.19: Relative protease expression upon peptide feeding. qRT-PCR profile of *H. armigera* trypsin and chymotrypsin genes upon peptide feeding. Asterisks indicate significant difference compared to control, analyzed by one way ANOVA for three independent experiments (*p<0.05, **p<0.01, ***p<0.001).

2.3.7 Stability of RCL peptides in insect gut

Proteolytic stability of RCL peptides was studied by incubation with HGP and analysing the mass spectra by MALDI-TOF MS at different time intervals. In the presence of HGP, TRE was highly resistant to proteolytic degradation up to 6h. Whereas, PRN showed the presence of smaller fragment peaks within 1h, suggesting that it was susceptible to initial degradation by HGP, after which no further degradation was observed. On the other hand, continuous decrease in intensity of intact PRY, as well as the appearance of smaller peaks, which increased in intensity with time, indicated that it was more prone to digestion by insect gut proteases (Figure 2.20). Thus, TRE shows enhanced proteolytic stability as compared to Pin-II PI from *C. annuum*, which shows a reduction in availability when incubated with HGP for long time(Mishra et al., 2010). Correspondingly, the revival of HGP activity in presence of RCL was monitored by in *vitro* biochemical assays. TRE and PRN in complex with HGP retained the inhibition up to 6h while PRY-HGP complex regained proteolytic activity with time. The stability of TRE and PRN

in presence of HGP highlights that RCL tripeptides can insert and freeze into the protease active site, forming a stable complex for a long time.

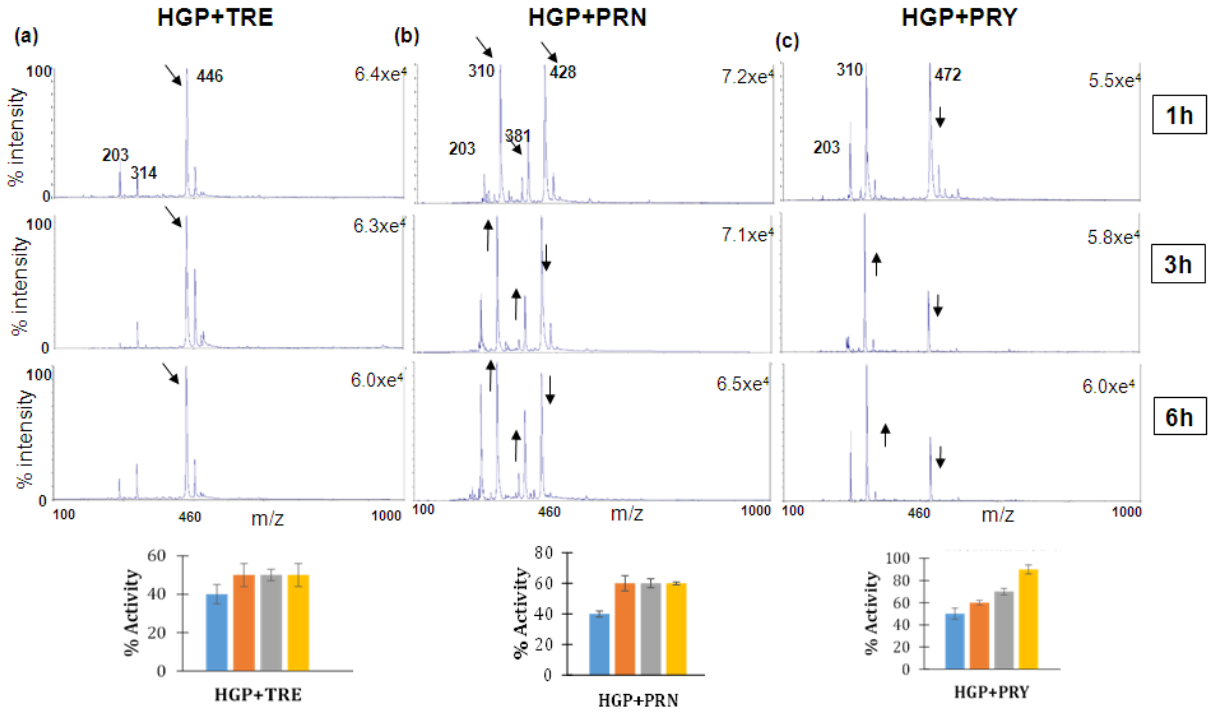


Figure 2.20: Stability of RCL peptides. MALDI-TOF spectra of (a) TRE-HGP (b) PRN-HGP and (c) PRY-HGP complex at 1, 3 and 6 h after incubation with HGP at 37°C. Arrows indicate the increase or decrease in intensity of peptides. Bar graphs underneath the spectra represent restoration of *H. armigera* gut protease activity after incubation with peptides at 1, 3 and 6h (HGP+ TRE/PRN/PRY).

2.4Conclusion

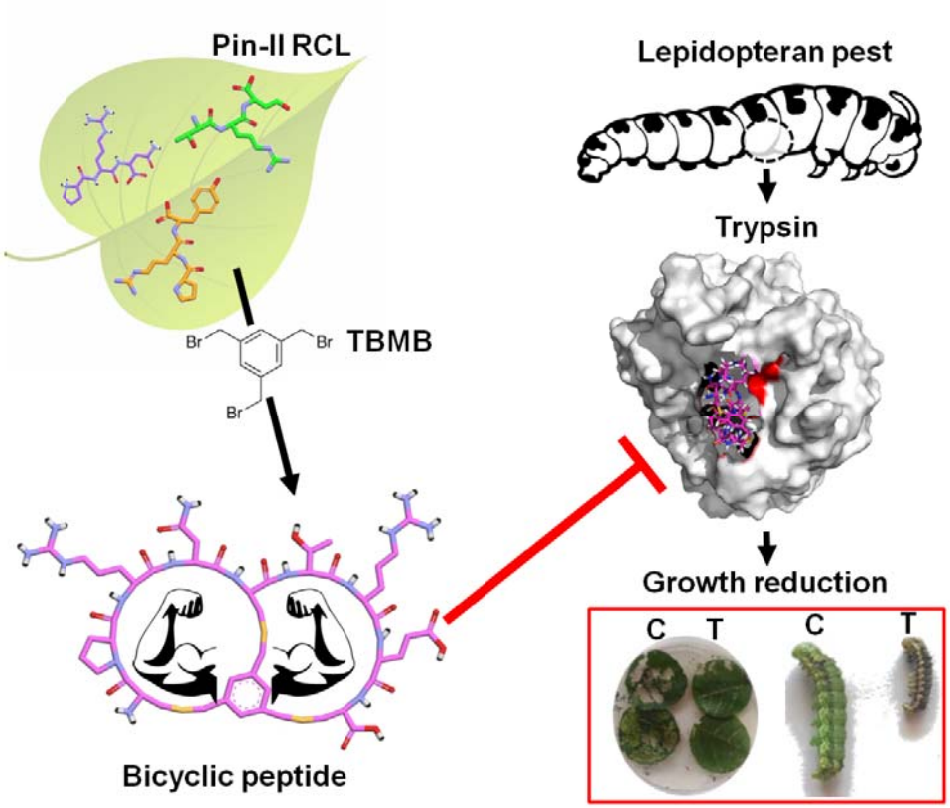
Potato type II protease inhibitors (Pin-II PIs) impede the growth of lepidopteran insects by inhibiting serine protease-like enzymes in the larval gut. The three amino acid reactive centre loop (RCL) of these proteinaceous inhibitors is crucial for protease binding and is conserved across the Pin-II family. However, the molecular mechanism and inhibitory potential of the RCL tripeptides in isolation of the native protein has remained elusive. In this study, six peptides corresponding to the predominant RCLs of the Pin-II PIs were synthesized and evaluated for *in vitro* and *in vivo* inhibitory activity against serine proteases of the polyphagous insect, *H. armigera.* RCL peptides with sequences PRN, PRY and TRE were found to be potent inhibitors that adversely affected the growth and development of *H. armigera* at sub-millimolar concentration. They showed higher inhibition of insect proteases compared to bovine trypsin, and enhanced efficacy at alkaline gut pH. Further, the RCL peptides showed retention in insect gut upto 6h and proteolytic stability towards insect proteases. The binding mechanism and differential affinity of the RCL peptides with serine proteases was delineated by crystal structures of complexes of the RCL peptides with trypsin. RCL peptides bind to serine proteases by deeply inserting into the active site pocket. Residues P1 (Arg) and P2 (Thr or Pro) of the inhibitors play a crucial role in the interaction and specificity of these inhibitors. Molecular responses of the insect upon ingestion of the RCL tripeptides were analysed by enzymatic assays and gene expression analysis, showing that the protease activity is reduced upon peptide ingestion. The current study provides new avenues to explore the potential of the RCL tripeptides for application in crop protection.

Key highlights of the study:

- RCL tripeptides inhibit serine proteases without parent IRD protein
- RCL tripeptides are deeply seated in active site pocket of trypsin, as shown by crystal structures
- They show greater activity at alkaline pH, and preference towards insect proteases than mammalian enzymes
- Interaction of peptide P2 residue with Tyr113 of trypsin is crucial for determining relative potency of RCL peptides
- They cause adverse effects on growth and life cycle of *H. armigera*
- RCL tripeptides are retained and stable in insect gut upto 6 h post feeding

Chapter 3

Tailoring of reactive loop peptides by cyclization for inhibition of lepidopteran serine proteases



Contents of Chapter 3 have been submitted as manuscript…

Saikhedkar, N. S., Joshi, R. S., Yadav, A. K., Seal, S., Fernandes, M., & Giri, A. P. Tailoring cyclic reactive loop peptides for augmentation of inhibition potential and functional differentiation. Manuscript submitted

3.1 Introduction

For the control of lepidopteran pests, several chemical and biological methods are available. However, the use of chemical pesticides is toxic to the environment and to non-target species. The biological methods, on the other hand, are limited by difficulty in synthesis and formulation applications, as well as lead to development of resistance(Dawkar et al., 2013). Therefore, alternate pest control methods are proposed based on the inactivation of insect digestive proteases by means of protease inhibitors(Dunaevsky et al., 2005). In this direction, we have shown the potential of peptide based protease inhibitors as insect control molecules. The tripeptide reactive centre loop (RCL) of Potato type-II serine protease inhibitors is capable of inhibiting trypsin-like enzymes without the native protein scaffold. Crystal structures of RCL peptides TRE, PRN and PRY in complex with porcine trypsin showed that the P1 Arg is deeply seated in the active site, and stabilized by number of H-bonds. These tripeptides also showed *in vitro* and *in vivo* inhibition of serine proteases in the lepidopteran pest, *H.armigera*(Saikhedkar et al., 2018). However, their potency was limited, since sub-millimolar concentrations of the peptide (~200μM) were required for effective protease inhibition. This necessitated for modification of linear tripeptides, so that more potent dietary insect control agents can be developed.

In native Pin-II protein, this tripeptide loop is flanked by conserved cysteine residues and held in the “canonical conformation” by means of a protein scaffold of 47 amino acids, featuring 3-4 disulphide bonds(Joshi et al., 2014; Schirra et al., 2008, 2010) (Figure 3.1). Therefore, we speculated that the flexibility of linear RCL tripeptide might be one of the reasons for its low binding affinity. Also, since the RCL in native Pin-II protein is held by disulphide bridges, it is possible that cyclization of tripeptides using the conserved Cys residues might enhance their inhibitory potential. Yet, in order to develop small molecule protease inhibitors, it was essential to replace the large protein scaffold by a small linker molecule.

A number of methods are available for cyclization of cysteine rich peptides(Hill et al., 2014; Northfield et al., 2014), most common being the formation of disulphide bridges(Chung and Yudin, 2015; Zha et al., 2018). But, the lower stability of disulphide bonds in reducing conditions, and their isomerization were not the desirable characteristics.In recent years, synthetic scaffolds have been used for cyclization of cysteine-containing peptides. First reported in 2005 by Timmerman et al., reaction of bromomethyl benzene derivatives with thiol group of

cysteine leads to generation of mono, bi or tricyclic peptides(Timmerman et al., 2005) (Figure 3.2).

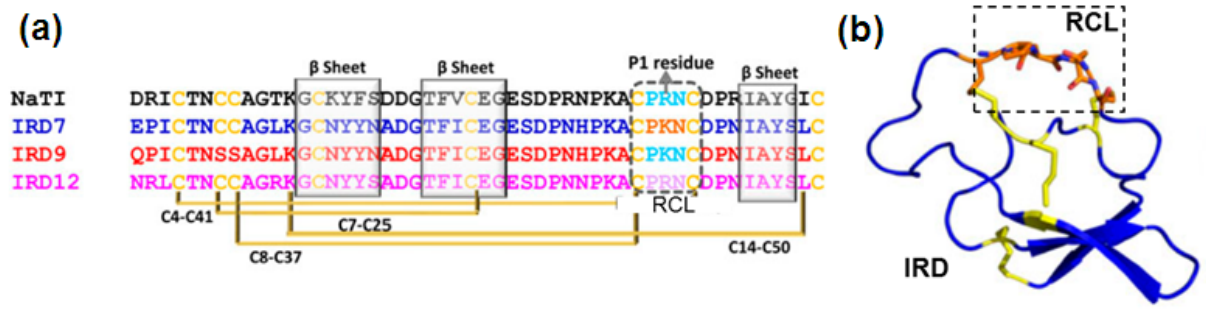


Figure 3.1: Reactive center loop of Pin-II protease inhibitors (Taylor et al., 2014)**.** (a) Representative sequences of IRDs from *Nicotianaalata* (NaTI), and *Capsicum annuum* (IRD7, 9 and 12) with disulphide connectivity marked by yellow lines, (b) Structure of IRD7 with highlighted RCL region held by disulphide bridges shown as yellow lines

Screening of phage displayed libraries of peptides cyclized by tris(bromomethyl)benzene (TBMB), led to selection of potent inhibitors of urokinase-type plasminogen activator, a serine protease(Angelini et al., 2012). Thereafter, TBMB scaffold has been optimized to generate linker molecules, which direct the folding and affinity of bicyclic peptides(Chen et al., 2014; Liu et al., 2017; van de Langemheen et al., 2017). The reaction with TBMB is efficient with small peptides, and specific towards Cys residues.

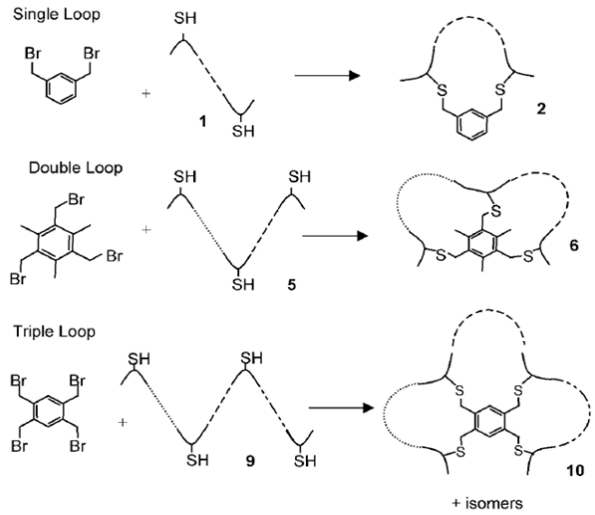


Figure 3.2: Cyclization of peptides by treatment with bromomethyl benzene derivatives(Timmerman et al., 2005)

Thus, in order to immobilize the tripeptide RCL sequences, TBMB was chosen as the favoured scaffold, with the following objectives:

- Design of peptide library by using combinations of RCL sequences and TBMB
- Synthesis of cyclic peptides
- Evaluation of in vitro protease inhibition by cyclic peptides
- Study of in vivo effects of cyclic peptides on lepidopteran pests, *H. armigera* and *Spodoptera litura*
- Identification of target proteins in insect gut
- Determination of binding modes of peptides with proteases

3.2 Methods

3.2.1 Synthesis of cyclic peptides

Linear peptides, as mentioned in table 3.1, were synthesized by Fmoc-Solid phase peptide synthesis on MBHA resin. Cleavage was performed by TFA/thioanisole/EDT/TFMSA procedure as mentioned in section 2.2.1. Peptides were precipitated in cold diethyl ether, and washed twice. After air drying, the crude peptide was dissolved in 1:1 mixture of ACN: H_2O. 1,3,5-Tris(bromomethyl)benzene (TBMB) (5mM) (for bicyclic peptides) or 1,3-Bis(bromomethyl)benzene (DBMB) (for monocyclic peptides) was dissolved in ACN. Reaction buffer was prepared by adding 5% DMSO and 45% 10mM NH_4HCO_3. Peptides and TBMB were added to the reaction mixture in equal volumes, and the reaction was incubated at 30 °C for 3 hr. After lyophilisation, purification was performed by RP-HPLC on a semi-prep C-18 column using a linear gradient from solvent A (95% H_2O, 0.1% TFA) to solvent B (50% ACN, 0.1% TFA). The mass of peptides was measured by MALDI-TOF mass spectrometry (Applied Biosystems, 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 250 single laser shots over each sample spot in the range of 200-2000 Da. They were processed for baseline correction and noise removal using Data Explorer software (Applied Biosystems). Confirmation of mass of P1 and P2 was performed by LC-HRMS analysis. MS acquisition was performed on the Q-Exactiveorbitrap mass spectrometer (ThermoFisher, MA, USA) operated in positive electrospray ionization (ESI) mode.

Disulphide bridged peptides were synthesized using air oxidation(Tam et al., 1991). 0.5mM peptide in 5% DMSO/water was vigorously mixed at room temperature for 12 h to incorporate atmospheric oxygen. Disulphide bridge formation was monitored by HPLC and MALDI-TOF MS.

Table 3.1 Sequences of synthesized peptide library

Linear peptide sequence	Scaffold used	Product peptide
CTRECTREC	1,3,5-Tris(bromomethyl)benzene	Bicyclic
CTRECPRNC	1,3,5-Tris(bromomethyl)benzene	Bicyclic
CTRECPRYC	1,3,5-Tris(bromomethyl)benzene	Bicyclic

CPRNCPRNC	1,3,5-Tris(bromomethyl)benzene	Bicyclic
CPRNCPRYC	1,3,5-Tris(bromomethyl)benzene	Bicyclic
CPRNCTREC	1,3,5-Tris(bromomethyl)benzene	Bicyclic
CPRYCPRYC	1,3,5-Tris(bromomethyl)benzene	Bicyclic
CPRYCPRNC	1,3,5-Tris(bromomethyl)benzene	Bicyclic
CPRYCTREC	1,3,5-Tris(bromomethyl)benzene	Bicyclic
CTREC	1,3-Bis(bromomethyl)benzene	Monocyclic
CPRNC	1,3-Bis(bromomethyl)benzene	Monocyclic
CPRYC	1,3-Bis(bromomethyl)benzene	Monocyclic
CTREC	Disulphide bridge	Monocyclic
CPRNC	Disulphide bridge	Monocyclic
CPRYC	Disulphide bridge	Monocyclic
TRE	None	Linear
PRN	None	Linear
PRY	None	Linear

3.2.2 Surface Plasmon Resonance

All SPR kinetics experiments were carried out on Biacore T200 (GE Healthcare, Bengaluru). Bovine trypsin was purchased from Sigma Aldrich (T8003). Bovine trypsin powder was resuspended in acetate buffer pH 5.5 and immobilized by the amine coupling method on a CM5 sensor chip according to the manufacturer's protocol (GE Healthcare). The immobilization level obtained was 4500 RU. Peptides were diluted in HBS-EP+ with 0.1 M $CaCl_2$at concentrations ranging from 0.5 to 50 µM, with a series of 10 fold escalations for kinetics screening using multi-cycle kinetics format (three concentrations injected as separate cycles), and 2 to 32 µM with 2 fold escalations for kinetics characterization of selected molecules using single cycle kinetics format (5 concentrations injected serially in one cycle). The association and dissociation times were 180 seconds each for multi cycle kinetics, and 60 seconds each for single cycle kinetics. The sensor surface was regenerated after each injection cycle with 10 mM Glycine-HCl pH 2.5 (GE Healthcare) for 60 seconds. The sensorgrams of test flow cells were subtracted from the reference flow cell. The kinetic fitting was carried out with Biacore T200

evaluation software by global fitting using 1:1 binding model. The kinetics data were described as ka (association rate or On-rate) and kd (Dissociation rate or Off-rate) and K_D (affinity constant or Equilibrium constant of Dissociation). Each SPR run was evaluated based on the statistical measurements provided by the Biacore T200 evaluation software, like Chi^2.

3.2.3 Biochemical activity assays

Trypsin activity was estimated through chromogenic substrate BA*p*NA (Benzoyl-L-arginyl-*p*-nitroanilide), as mentioned in section 2.2.2. Total HGP activity was measured by azocaseinolytic assay, as described in section 2.2.2.

Superoxide dismutase assay was performed according to the method of Winterbourn et al.(Winterbourn et al., 1975) 10µl of HGP was incubated with 30 µl 0.12mM riboflavin, 10 µl 1.5mM Nitroblue tetrazolium (NBT), 250µl of 12 mM methionine, which was followed by illumination for 10 min in 16W fluorescent light. Absorbance was measured immediately at 560nm.

3.2.4 Docking and Molecular Dynamics simulation

HaTry (Uniprot ID: B1NLE4) was modelled using Modeller 9.17(Šali and Blundell, 1993) and porcine trypsin (PDB: 4DOQ) as template. Bovine trypsin coordinates were obtained from crystal structure (PDB: 4I8H). 3D structures of peptides were prepared by Marvin Sketch v 16.10.17.0, and energy minimized by PRODRG online tool(Schüttelkopf and van Aalten, 2004). Docking was performed by AutoDockVina(Trott and Olson, 2009) using grid boxes of 25x25x25 points and 1Å spacing centred on the catalytic triad. Analysis of docking poses was carried out by using Autodock tools: out of the eight possible binding modes, the one with greatest proximity to the active site and most negative free energy of binding was selected. 2D interactions were mapped by Discovery studio Visualizer (Dassault Systèmes BIOVIA, Discovery studio Visualizer, v 16.1.0.15350, San Diego: Dassault Systèmes, 2015.

The best docked pose was used for molecular dynamics simulations. All MD simulations were performed with AMBER99SB forcefield using GROMACS 4.6.3 package (Van Der Spoel et al., 2005). Topologies for the peptides were prepared by acpype server(Sousa da Silva and Vranken, 2012). During the MD simulations, all the protein atoms were surrounded by a triclinic water box of TIP3P water molecules. The systems were neutralized with Na+ and Cl− counter ions replacing the water molecules and energy minimization was performed using steepest

descent algorithm. NVT and NPT equilibrations were done for 100ps each using PME electrostatics and Parrinello-Rahman pressure coupling. 200ns production MD simulations with a time step of 10 ns were performed. Upon completion, the most representative structures of the peptide-protease complexes were selected using the GROMACS cluster tool with a root mean square deviation (RMSD) threshold of 2.0 Å. The RMSD, RMSF values were calculated using GROMACS, and given as graphs in Figure 3.3 and 3.4. 3D structure representations were prepared by PyMol (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.), Chimera (Chimera is developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIGMS P41-GM103311)(Pettersen et al., 2004) and Discovery Studio Visualizer programs.

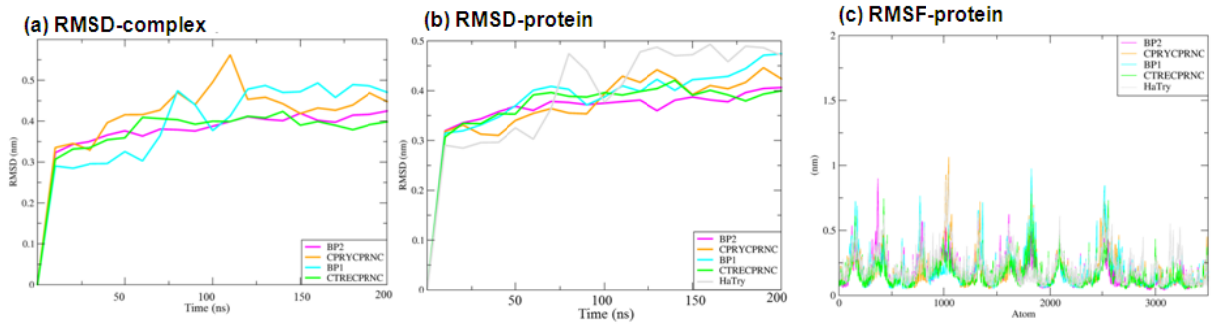


Figure 3.3: RMSD and RMSF plots for selected ligands after 200ns MD simulation with *H. armigera* trypsin

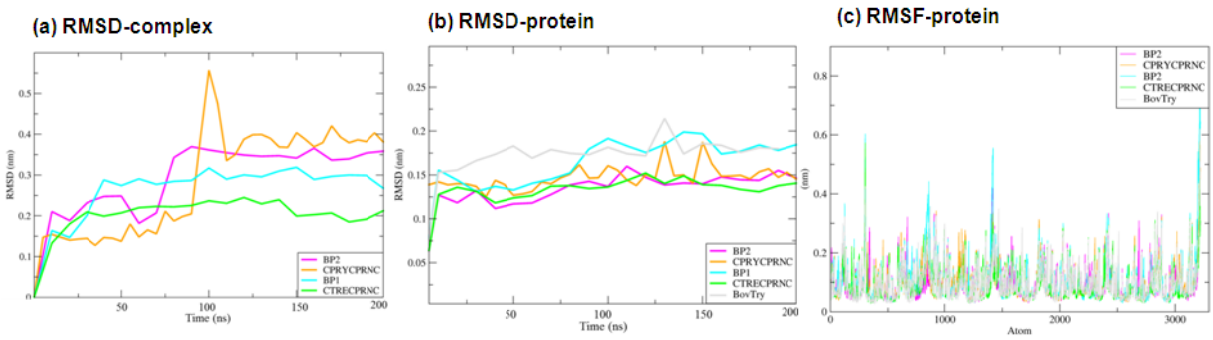


Figure 3.4: RMSD and RMSF plots for ligands after 200ns MD simulation with bovine trypsin

3.2.5 Feeding assays

3.2.5.1 Choice assay

Two independent sets of 30 insects each of *H. armigera* and *S. litura* were subjected to open choice between untreated and peptide treated castor leaves (CPRNCPRYC or CPRNCTREC).The leaves were painted with peptides on both sides of leaves, 10 neonates of *S. litura*and *H. armigera* were placed in each dish, having two leaves of untreated (control) and two leaves of peptide painted leaves (test). Moist tissue paper was used to provide moisture to overcome early dryness of leaves, on alternate days leaves were replaced. Images of eaten area were captured on each day, followed by calculation of eaten area by ImageJ software(Rueden et al., 2017).

3.2.5.2 No choice assay

Obligatory feeding assay was done independently with two Lepidopteran insects *H. armigera* and *S. litura* on castor leaf and cotton leaves. Three independent sets of 100 neonates were fed on untreated and peptide treated (CPRNCPRYC and CPRNCTREC) leaves. 20 insects were placed in five replicates on one peptide painted leaf in a petri dish. Moist tissue paper was used to provide moisture to overcome early dryness of leaves, on alternate days leaves were replaced. Insects were counted daily to check their mortality, after insects reached in to second instar; their weights were measured on every day.

3.2.6 Pull down assay

For this assay, biotinylated cyclic peptide CPRNCPRYC, CPRNCTREC, and linear peptide PRN were synthesized as mentioned above. Streptavidin slurry was incubated at 25°C with biotinylated peptide. Unbound peptide was washed off. To this slurry, HGP was added and incubated at 25°C for 5 min. The flow through was collected and slurry and washed three times with 0.1M glycine NaOH buffer at pH10. For elution, Sodium citrate buffer at pH 5 was added. Eluted proteins were collected and total proteins were estimated by Bradford assay. 1µg/µl of eluted proteins was used for in solution trypsin digestion. The proteins were diluted in 100mM ammonium bicarbonate, followed by denaturation at 80°C for 15min. Thereafter, to reduce disulphide bonds, dithiothreitol was added at 65°C for 20min, followed by iodoacetamide for alkylation. After incubation in dark for 30min, added trypsin at 1mg/ml concentration and

incubated for 16 hours at 37°C. To stop the reaction, formic acid was added followed by centrifugation at 4°C for 15min at 15000rpm. The supernatant was used for peptide purification by zip tip. Ziptip eluted peptide samples were reconstituted in 3%ACN+ 0.1% formic acid and used for identification by MS/MS.

Peptide samples were subjected to a NanoAcquity ultra performance liquid chromatography (UPLC) coupled to MALDI-SYNAPT HDMS (Waters Corporation, Milford, MA, USA). The nano-LC separation was performed using a bridged-ethyl hybrid (BEH) C18 reversed phase column (1.7µm particle size) with an internal diameter of 75µm and length of 150mm. Each sample of total digested protein was injected into the trapping column and flushed with 0.1% solvent A (0.1% formic acid containing water) for 3min at a flow rate 15µL/min. Upon each injection, peptides were eluted into the NanoLockSpray ion source at a flow rate of 300nL/min using a gradient of 2%–40% B (0.1% formic acid containing acetonitrile) over 110 min. The data was processed and analyzed using Protein Lynx Global Server 2.5.2 software (PLGS; Waters Corporation, Milford, MA, USA). Protein identifications were obtained by searching Helicoverpa database (www.uniprot.org). LC–MSE data was analysed with a fixed carbamidomethyl modification for Cys residues, along with a variable modification for oxidized Met residues. Ion Accounting search parameters used for data independent analysis (DIA) were precursor and product ion tolerance (automatic setting), minimum number of peptide matches (3), minimum number of product ion matches per peptide (5), minimum number of product ion matches per protein (7), and maximum number of missed tryptic cleavage sites (2). The false positive rate was 4%. The results of DIA (proteins and the individual MS/MS spectra) with confidence level ≥95% were accepted.

3.2.7 Quantitative Real-Time PCR

Total RNA was isolated from the 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 (Table 3.2) was assessed as described in section 2.2.6. (Dawkar et al., 2013; Mahajan et al., 2013). For each treatment, two biological replicates (each constituting three technical replicates), the average transcript abundance and sub-sequent fold difference with respect to the control were calculated.

Table 3.2: Primers used for qRT-PCR analysis

Target gene	Genbank acc.	Forward Primer	Reverse Primer
HaTry4	EF600059	GTGCTACCCCTTCTGATTC	AACTTGTCGATGGAGGTGAC
HaTry7	Y12271	CAGAGGATTGTGGGTGGTTCG	GCGGTGAGGATAGCCCTGTT
HaTry6	Y12283	TGGCTGGGGTGACACTTTCT	GTCTCCCTGGCACTGGTC
HaChy4	Y12273	CACCATCTTCATCTTCCAATCCGTGTGC	GTGTTGATACGAGTACCACCGAAGAAC
HaChy9	Y12287	TGCTGGTTCGATGGTCG	TTACCAGAGGGCAGGGC
CuSOD	JQ009331	CTGACCCTGATGACCTTGGAG	GATAACACCGCAGGCAATACG
MnSOD	GU115810.1	TGCCCTATGAGTACAG	GGGAGCTAAACTGATG
HaGAPDH	JF417983	TGCTGAATACGTCGTTGAATCC	TTCTTAGCACCACCCTCTAAATGAG

3.3 Results and discussion

3.3.1 Design and docking of cyclic RCL peptides

Based upon our previous study, we have used the most abundant Pin-II RCL sequences, namely, **TRE**, **PRN**, **PKN**, **PLN**, **TLN** and **PRY** for design of bicyclic peptides. Since in nature, the RCL sequence is flanked by two cysteine residues, we used these cysteines to cyclize the peptides. Scaffolds TBMB and DBMB were used to generate bicyclic and monocyclic peptides, respectively. Also, disulphide linkages and N- to C- terminus cyclized peptides were designed. A virtual library of 142 combinations of peptides was screened in silico by docking on bovine trypsin and five representative *H. armigera* trypsins. A comparison of binding energies (ΔG) revealed that cyclic peptides were effective binders of trypsins as compared to their linear counterparts (Figure 3.4, table 3.3). Among these peptides, TBMB linked peptides were the strongest interacting partners with bovine as well as insect trypsins. A few peptides with N- to C-terminus cyclization also showed potential binding as indicated by a highly negative ΔG value.

Interaction analysis of the best docking pose with bovine trypsin showed that bicyclic peptide covered higher interface surface area with trypsin than linear peptides, thus blocking the active site (Figure 3.5). Bicyclic peptides showed H-bonding interactions as well as hydrophobic interactions with the active site and nearby amino acids of bovine trypsin. Also, both Arg of the two loops were interacting with trypsin, which might be the reason for higher binding affinity of bicyclic peptides (Figure 3.6) Moreover, the extended structure of bicyclic peptide allowed the reactive loop peptides to stay in a rigid confirmation, thereby may enhance the favourable H-bond interactions with active site residues. Hence, scaffold attachment leading to stabilization of peptide structure may increase the potency of the RCL peptide inhibitors.

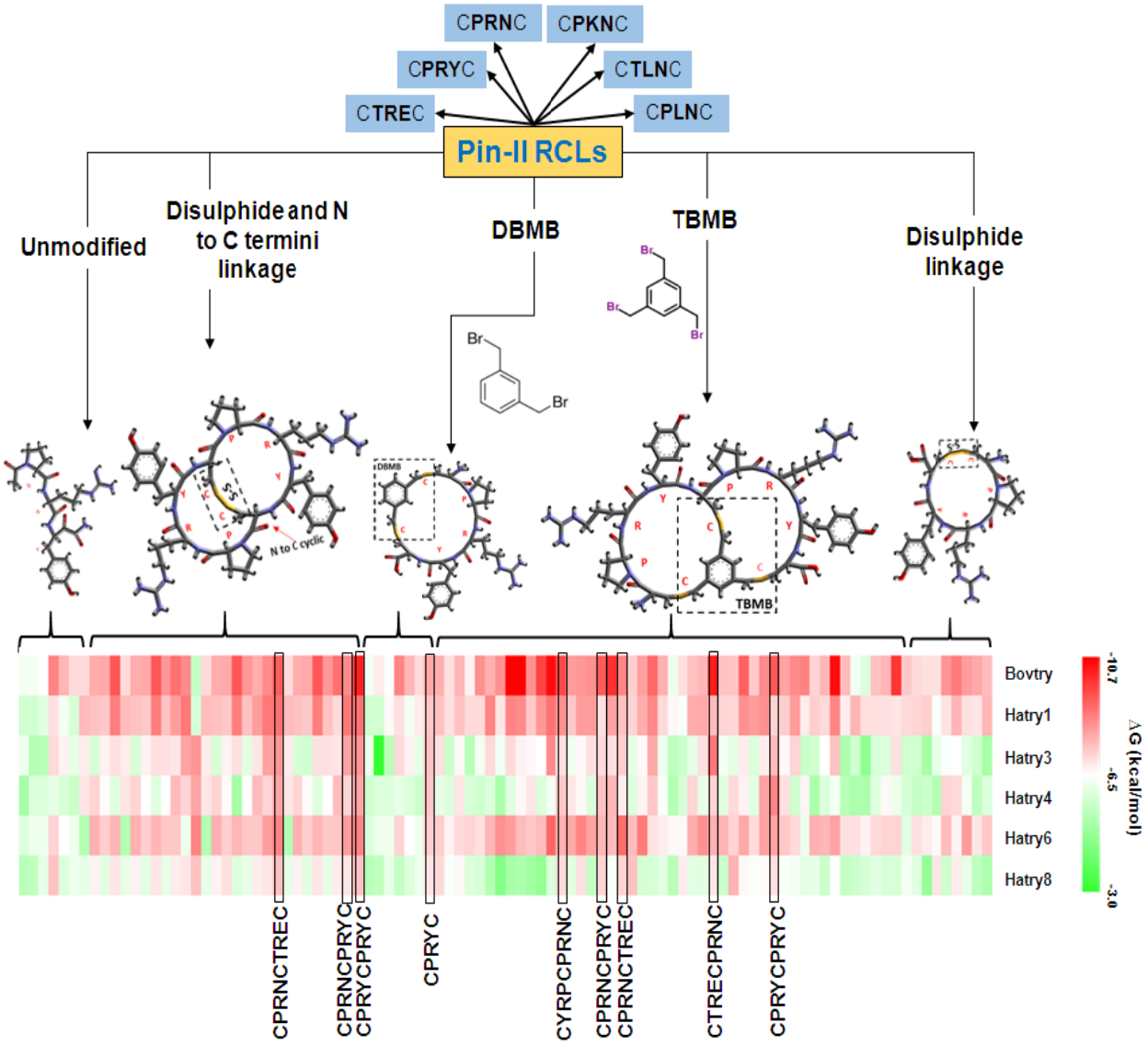
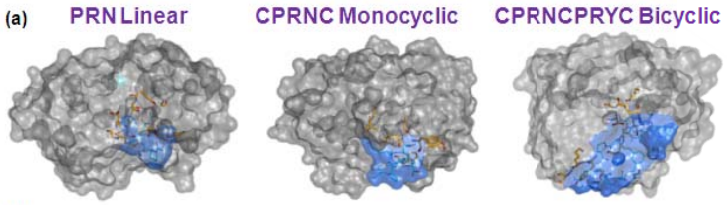


Figure 3.4: Representative structures of ligands designed for virtual screening. Peptides are shown as sticks,Tris-bromomethyl benzene (TBMB), Di-bromomethyl benzene (DBMB) and disulphide linkages (S-S) are marked by boxes. Amino acid residues are marked as single letter codes. Heat map showing relative binding affinities of peptides with different trypsins obtained after docking studies: red-strong binding; green-poor binding.

Table 3.3: Binding energies (kcal/mol) obtained for peptides upon docking using AutoDockVina. **B:** Biyclic peptide with TBMB Sacffold; **M:** Monocyclic peptide with DBMB scaffold; **N-C:** Bicyclic peptide with N to C cyclization and disulphide bridge; **SS:** Monocyclic peptide with disulphide bridge

Name	Bovtry	Try1	Try3	Try4	Try6	Try8
CTRECTREC_B	-7.6	-7.7	-5	-5.9	-7.4	-5.1
CTRECTLNC_B	-7.9	-6.9	-5.7	-5.4	-6.9	-5.4
CTRECPKNC_B	-8.1	-7	-6	-6.3	-8.1	-6.1
CTRECPRYC_B	-8.65	-7.7	-6.5	-7.1	-6.7	-6.8
CTRECPRNC_B	-8.25	-6.95	-6.4	-7.2	-7.6	-5.7
CTRECPLNC_B	-7	-6.4	-6.1	-5.2	-6.6	-6.9
CTLNCTREC_B	-7	-7.3	-5	-5.6	-6.8	-4.7
CTLNCTLNC_B	-7.1	-7.2	-5.7	-5.5	-7	-6
CTLNCPKNC_B	-7.4	-7.1	-5.7	-5.8	-6	-6.3
CTLNCPRYC_B	-9.8	-6.9	-7	-6.2	-7.2	-5.8
CTLNCPRNC_B	-7.7	-7	-5.9	-6.6	-7.2	-5.8
CTLNCPLNC_B	-7.5	-7	-5.8	-6	-7.1	-6
CPKNCTREC_B	-5.9	-7	-4.6	-5.1	-6.7	-5.8
CPKNCTLNC_B	-6.3	-7.7	-5.7	-5.1	-7	-5.6
CPKNCPKNC_B	-7.6	-6.8	-5.1	-5.3	-7	-5.5
CPKNCPRYC_B	-10.5	-8.5	-7.5	-6.3	-8	-6.5
CPKNCPRNC_B	-7.7	-7.1	-6.5	-5.9	-7.7	-5.6
CPKNCPLNC_B	-7.2	-7.3	-5.7	-5.1	-7.8	-6.1
CPRYCTREC_B	-8.55	-6.8	-5.55	-6.65	-6.2	-6.05
CPRYCTLNC_B	-8.4	-8.1	-7.2	-5.7	-7.1	-6.3
CPRYCPKNC_B	-8	-7.3	-6.6	-7	-7.4	-6.4
CPRYCPRYC_B	-9.25	-7.2	-7.6	-8.5	-8	-7
CPRYCPRNC_B	-8.4	-8.3	-6.8	-8	-8	-6.55
CPRYCPLNC_B	-7.4	-8.1	-5.9	-6	-7.1	-6.4
CPRNCTREC_B	-9.1	-8.3	-6.8	-6.7	-7.9	-6.6
CPRNCTLNC_B	-7.45	-7.15	-6.05	-7.1	-7.2	-7.75
CPRNCPKNC_B	-7.5	-7.2	-6.3	-6	-7.7	-5.6
CPRNCPRYC_B	-10.2	-7.6	-8.7	-7.2	-7.8	-6.9
CPRNCPRNC_B	-7.7	-7.9	-7	-6.1	-8.3	-6
CPRNCPLNC_B	-7.5	-7.85	-6.1	-7.15	-7.85	-6.05
CPLNCTREC_B	-7	-6.4	-5.6	-5.8	-6.6	-6.8
CPLNCTLNC_B	-6.2	-6.6	-5.4	-5.2	-6.6	-5.3
CPLNCPKNC_B	-7.8	-7.1	-6.2	-6.4	-6.8	-6
CPLNCPRYC_B	-8.95	-8.45	-7.9	-7.4	-6.95	-5.8
CPLNCPRNC_B	-8.05	-7.85	-6.6	-6.7	-8.65	-5.75
CPLNCPLNC_B	-7.3	-7.2	-6.1	-5.9	-7.6	-7.5

Name	Bovtry	Try1	Try3	Try4	Try6	Try8
CYRPCPRYC_B	-8.5	-7.5	-7.3	-6.8	-8.9	-7.2
CYRPCPRNC_B	-9.9	-8	-7	-7.7	-7.6	-6.4
CNRPCPRYC_B	-9.2	-7	-7.1	-7.5	-7.7	-7.2
CERTCPRYC_B	-8.3	-8.1	-7	-5.9	-8	-5.5
CNLPCPRYC_B	-8.2	-8.1	-6.6	-6.6	-8.5	-7
CNRPCTLNC_B	-8.1	-7.4	-6.2	-6.9	-7.9	-6
CNLPCPRNC_B	-9.5	-8.3	-7	-7.4	-7.8	-7.1
CPRYCYRPC_B	-10.5	-8.8	-8.4	-6.3	-9	-5.9
CPRYCNRPC_B	-9.6	-8.2	-6.5	-5.9	-7.6	-4.7
CPRNCYRPC_B	-8.3	-7.4	-6.6	-6.9	-7.4	-4.9
CTRECYRPC_B	-10.7	-8.4	-6.7	-6.6	-7.6	-5.1
CPLNCYRPC_B	-10.7	-8.5	-7.1	-6	-8.1	-5
CPRNCNLTC_B	-8.3	-6.9	-6.3	-5.1	-8.3	-4.4
CPLNCNRPC_B	-7.8	-8.5	-7.1	-5.9	-7.7	-5.2
CPRYCPRY_N-C	-10.15	-8.8	-8.2	-7.95	-8.5	-7.5
CPRYCPRN_N-C	-8.4	-8.7	-8	-7.2	-8.1	-6.9
CPRNCPRY_N-C	-8.85	-8.1	-6.9	-7.6	-7.4	-6.8
CTRECPRY_N-C	-8.2	-7.5	-7	-6.5	-7.4	-6.1
CPLNCPRY_N-C	-9.25	-8.6	-7	-5.95	-7.7	-6.9
CPRNCTLN_N-C	-7.9	-7.4	-6.7	-6.55	-7.3	-6
CPLNCPRN_N-C	-7.85	-7.5	-5.8	-7.2	-7.8	-6.7
CPRNCTREC_N-C	-7.1	-7.6	-6.8	-6.9	-5.2	-6.1
CTREC_M	-8.3	-7.1	-6.3	-5.7	-7	-5.8
CTLNC_M	-6.7	-7.1	-5.3	-5.8	-7.2	-5.9
CPRNC_M	-7.7	-7.5	-6.6	-6.9	-7.1	-6.6
CPKNC_M	-7	-7	-5.6	-5.9	-6.7	-6.4
CPLNC_M	-7.8	-7.5	-7	-6.1	-7.3	-7.1
CPRYC_M	-7.9	-7.4	-7.5	-7.4	-7.3	-6.9
CTREC_SS	-7.1	-6.8	-6.1	-5.6	-6.2	-6.2
CTLNC_SS	-7.7	-6.3	-6.7	-5.9	-6.5	-5.9
CPRNC_SS	-8.5	-7	-7.1	-6.1	-7.2	-5.5
CPKNC_SS	-6.5	-6.6	-5.3	-5.9	-6.2	-6.1
CPLNC_SS	-6.8	-5.7	-2.2	-5.9	-6.2	-5.4
CPRYC_SS	-6.3	-5.7	-6.2	-5.7	-5.9	-5.3
TLN	-7.1	-5.8	-6.1	-5.6	-6.2	-6.2
PRN	-7.7	-6.3	-6.7	-5.8	-6.5	-5.9
PRY	-8.5	-7	-7.1	-6.1	-7.2	-7.5
PLN	-6.5	-6	-5.3	-5.9	-6.2	-6.1
PKN	-6.3	-5.7	-5.2	-5.9	-6.2	-5.4
TRE	-6.3	-5.7	-6.2	-5	-5.9	-5.3



(b)

	Bovine trypsin		*H. armigera* trypsin 4	
Peptide	Area (Å²)	Affinity (kcal/mol)	Area (Å²)	Affinity (kcal/mol)
PRN	375	-7.7	250	-6.7
PRN S-S	358	-8.5	385	-7.1
PRN-monocyclic	246	-7.7	344	-6.9
CPRNCPRYC N-C	486	-8.5	521	-7.6
CPRNPRYC-bicyclic	536	-10.2	627	-8.7

Figure 3.5: Docking of peptides. (a) Docking poses of linear, monocyclic and bicyclic peptides

bound to bovine trypsin, (b) Comparison of interface surface area and binding affinities obtained after docking with bovine trypsin and *H. armigera* trypsin 4.

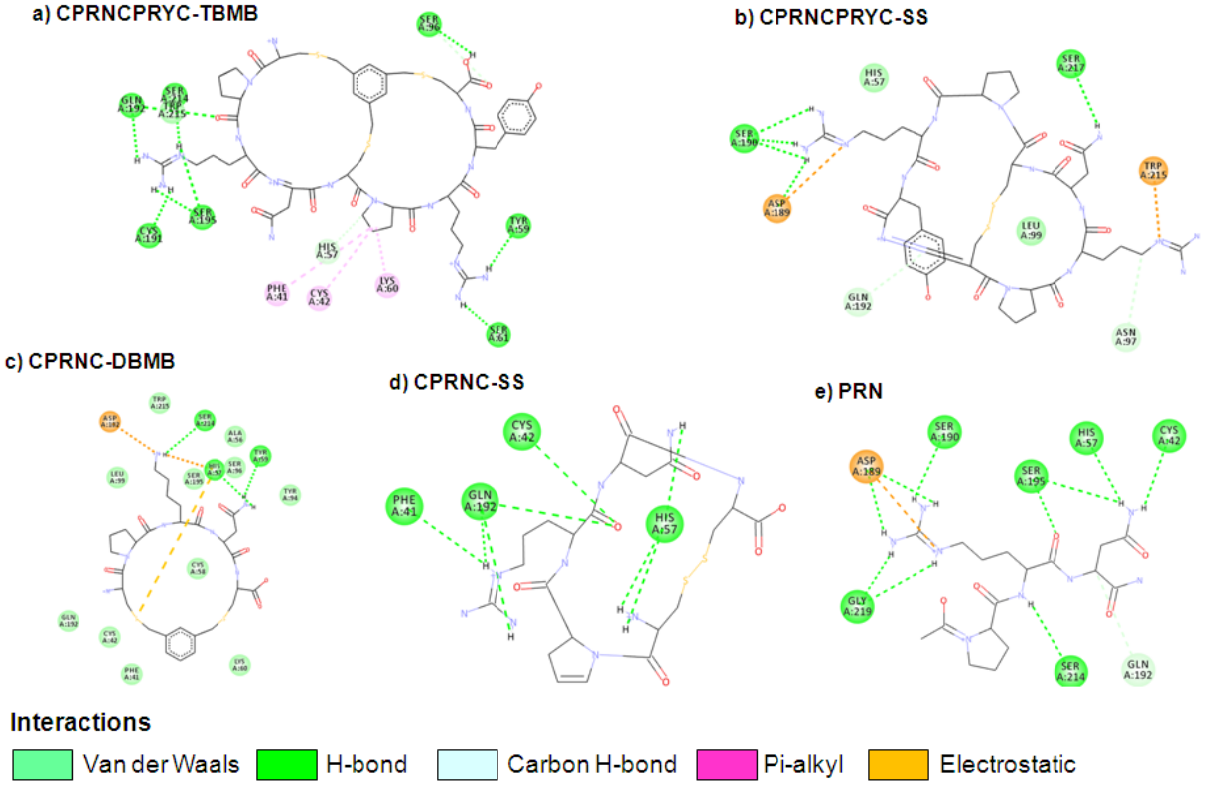


Figure 3.6: Comparison of docked poses of representative peptides with bovine trypsin. Docking was performed by AutoDockVina, and 2D representations prepared by Discovery Studio Visualizer

3.3.2 Synthesis of cyclic peptides

Linear peptides with flanking cysteine residues were synthesized by Fmoc-solid phase synthesis, after which, the crude peptides cleaved from the resin were used for cyclization reaction with TBMB. We synthesized a collection of eighteen peptides, comprising of nine bicyclic peptide combinations from TRE, PRN and PRY, six monocyclic peptides (disulphide and scaffold linked), and three linear peptides. The synthesis reaction with TBMB scaffold was optimized as given in table 3.4. Completion of reaction was monitored by observing change in molecular weight by MALDI-TOF MS (Figure 3.7). Thereafter, the peptides were purified by HPLC and molecular weights were confirmed by MALDI TOF MS (Figure 3.8, 3.9). Also, the peptides CPRNCTREC and CPRNCPRYC were confirmed by LC-HR-MS analysis (Figure 3.10).

Table 3.4: Optimization of reaction conditions with TBMB

Peptide	Buffer	Solvent	Reducing agent	TBMB conc.	Time	Peptide obtained
On resin	-	DMF/ACN	-	5mM	0-4h	Linear
Crude	40mM NH_4HCO_3	ACN	-	5mM	0-4h	Linear
Pure (1mM)	20mM NH_4HCO_3	45% ACN; 5%DMSO	ß-Me	5mM	1h	Linear
Pure (1mM)	20mM NH_4HCO_3	45% ACN; 5%DMSO	ß-Me	5mM	12h	**Bicyclic**
Pure (1mM)	20mM NH_4HCO_3	45% ACN; 5%DMSO	DTT	5mM	12h	Linear
Crude	20mM NH_4HCO_3	45% ACN; 5%DMSO	ß-Me	5mM	12h	**Bicyclic**

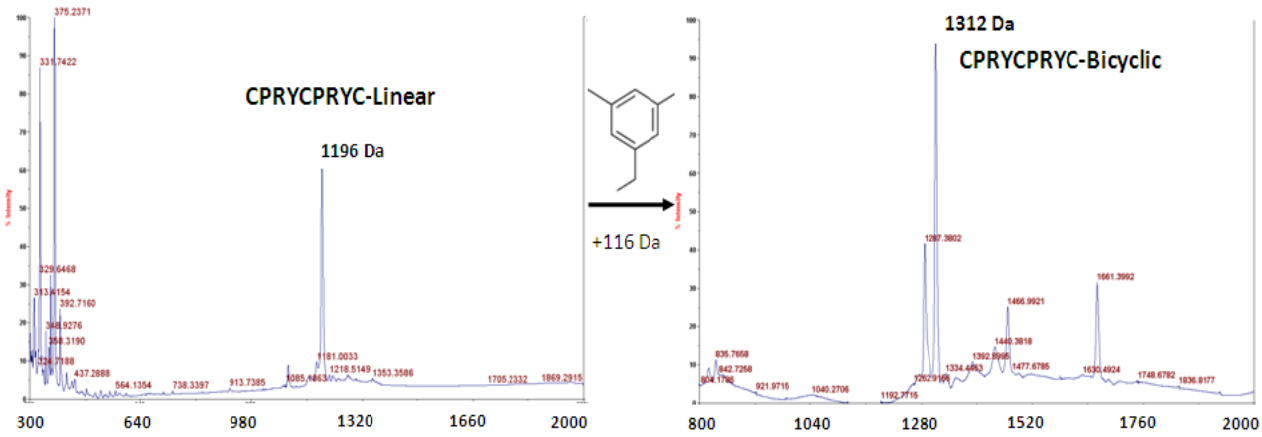


Figure 3.7: Confirmation of cyclization reaction by MALDI-TOF MS.

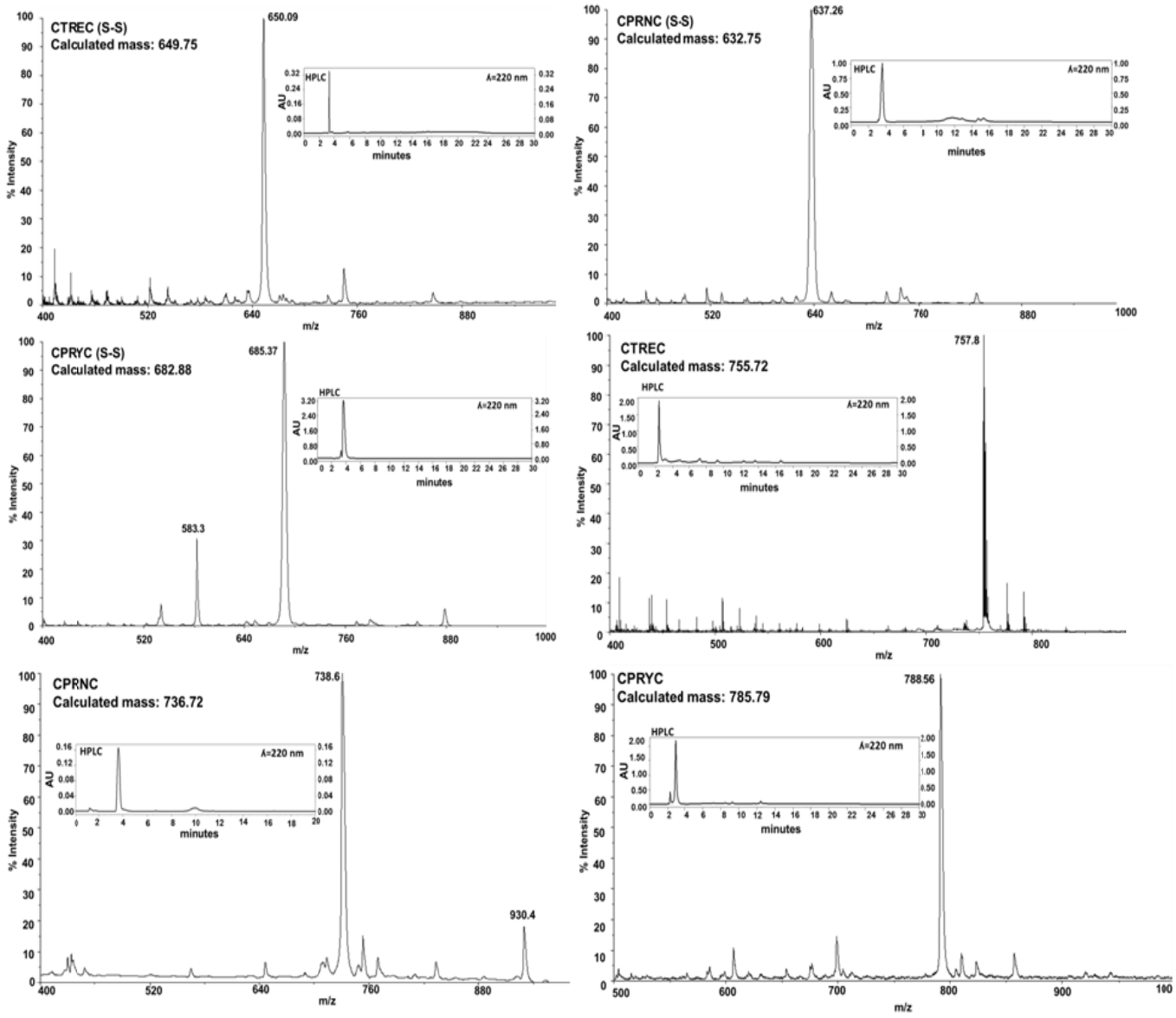


Figure 3.8: Purification of synthesized monocyclic peptides by HPLC and confirmation of molecular weights by MALDI-TOF MS

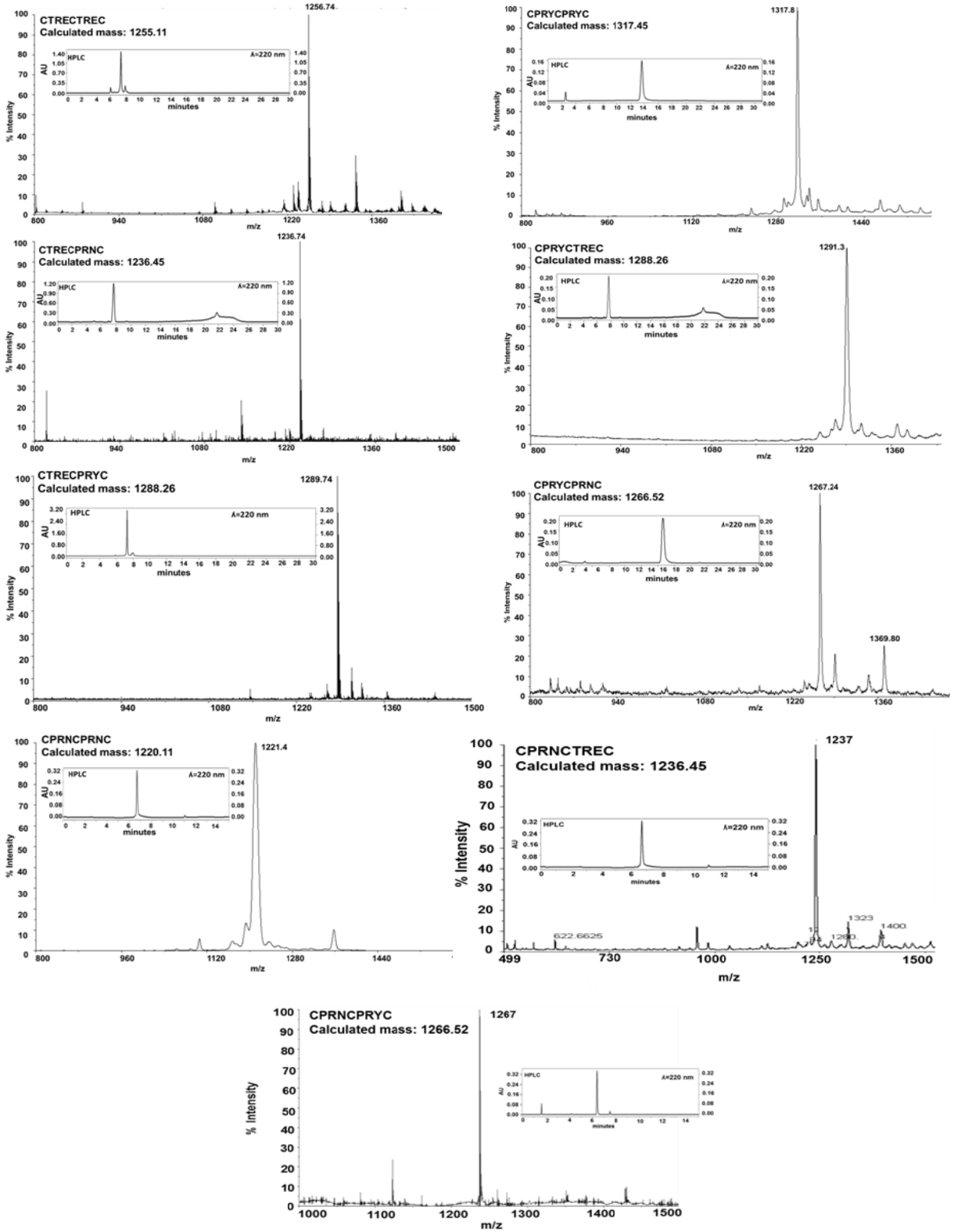


Figure 3.9: Purification of synthesized bicyclic peptides by HPLC and confirmation of molecular weights by MALDI-TOF MS

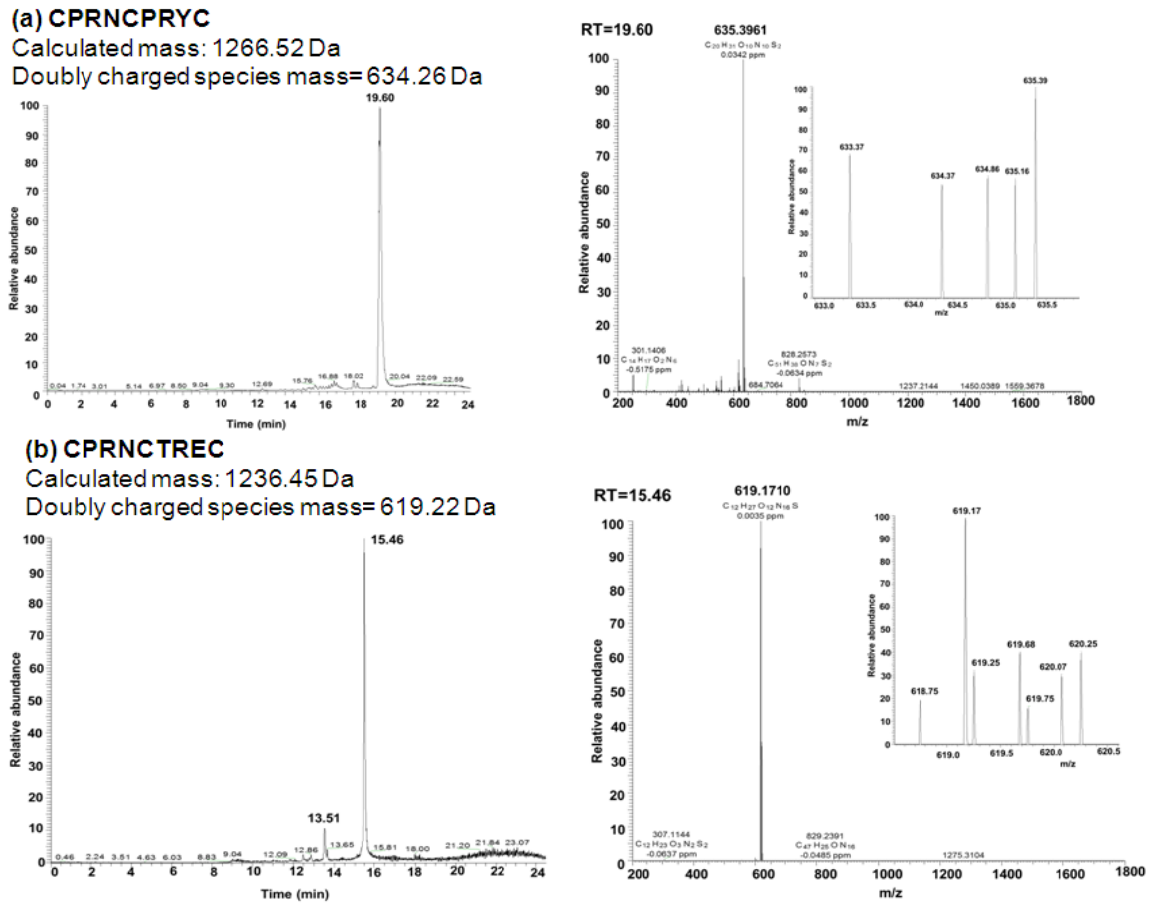


Figure 3.10: Confirmation of molecular weights of peptides by LC-HR MS

3.3.3 *In vitro* screening of cyclic peptides

We screened the collection of 18 peptides for *in vitro* inhibition of trypsin-like midgut proteases of *H. armigera*. The inhibitory potential was compared with linear RCL peptides. It was observed that bicyclic peptides showed strongest inhibition of insect proteases. At the studied concentration of 20µM, bicyclic peptides were ten times more potent than linear peptides. Out of the bicyclic peptides, CPRNCTREC and CPRNCPRYC were the most potent inhibitors which showed more than 70% reduction in trypsin activity. The IC_{50} values obtained were significantly lesser than the previously reported inhibition constants for linear RCL peptides (Table 3.5, Figure 3.11). Further, cyclic peptides showed higher inhibitory activity against *H. armigera* trypsin-like proteases than bovine trypsin. Cyclic peptides showed ~80% inhibition at 20 µM, while similar inhibition was achieved at 1mM for bovine trypsin. Therefore, the selectivity of RCL peptides towards insect proteases is retained upon cyclization

with TBMB (Saikhedkar et al., 2018). Whereas, disulphide linked RCL peptides, or peptides attached to dimethylbenzene scaffold were not effective inhibitors at 20µM, which states that cyclization of two RCL peptides using TBMB is a better strategy to enhance their potency.

Table 3.5: IC_{50} values (µM) obtained for bicyclic and linear RCL peptides

Peptide	*H.armigera* trypsin-like	Bovine trypsin
CPRNCTREC	15	250
CPRNCPRYC	20	400
TRE	200	900
PRN	40	500
PRY	200	400

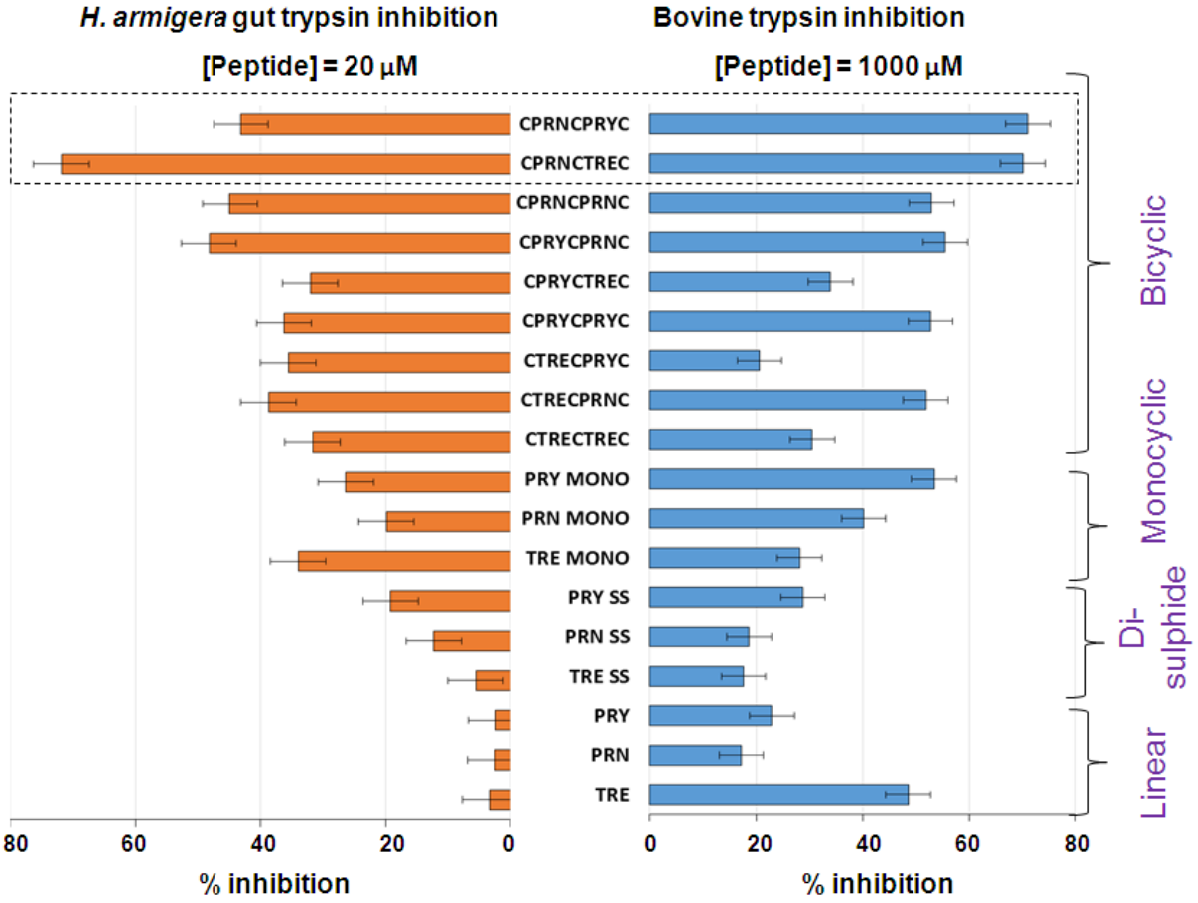


Figure 3.11: Screening of peptides for protease inhibition. Peptide library comprising of bicyclic, monocyclic and linear peptides was screened for in vitro inhibition of bovine trypsin and *H.armigera* trypsin-like proteases. Selected peptides are highlighted in box.

Interestingly, the sequence orientation of tripeptides around TBMB scaffold also affected the inhibition potency. For instance, CPRNCTREC showed higher inhibition than CTRECPRNC at the same concentration. This suggested that the scaffold might help the RCL to adopt a specific conformation, which could be optimized depending on the sequence of first and second loop.

3.3.4 Peptide-trypsin interaction analysis by Surface Plasmon Resonance (SPR)

Interaction of the synthesized peptides with trypsin was studied by SPR. Since purified insect midgut trypsin is challenging to express, bovine trypsin was used as an alternate system. Initial screening of interacting peptides was performed by multi-cycle kinetics assay. Increasing concentrations of the peptides (0.5, 5 and 50 μM) were injected onto immobilized trypsin. Based on the kinetic screening results, we rejected the peptides which did not show significant interaction, as indicated by sensorgramshape and kinetics parameters (Figure 3.12, table 3.6). Among the peptides that showed binding, there was a variation in the binding constant (K_D), which is a measure of binding affinity. Bicyclic peptides CPRNCTREC and CTRECTREC had K_D< 5μM, CPRNCPRYC and CTRECPRNC had K_D of 11 and 16 μM respectively. While, other peptides showed K_D above 50μM. But, in addition to high affinity, we desired a peptide which shows fast binding, as well as releases from the trypsin at a slow rate. Analysis of kinetic parameters (on-rate, ka for association and off-rate, kd for dissociation), highlighted bicyclic peptides CPRNCTREC, CPRNCPRYC and CTRECTREC as potent binders. As seen from figure 3.13, CPRNCTREC was having lowest kd (in the range of 10^{-4}), hence, would dissociate from trypsin slowly. However, its ka value (10^2) was also low, meaning it shows slow binding with trypsin. CPRNCPRYC and CTRECTREC were selected because of high ka (~10^3) and slightly high kd value (~10^{-3}), indicating fast complex formation with trypsin, and slow rate of dissociation. Other bicyclic peptides like CTRECPRNC and CPRYCPRYC had high ka (~10^3), but also high kd values (~10^{-3}), indicating that they form an unstable complex with trypsin, which rapidly dissociates. Comparatively, only linear peptide TRE was potent to bind to trypsin, while none of the disulphide linked peptides showed strong interaction. But, DBMB linked monocyclic peptides, CPRNC and CPRYC displayed binding with low ka (~10^1) values, signifying that they form a slow complex. Thus, we based our selection on kd, as the stability of trypsin-inhibitor complex is dependent on slow dissociation rate of peptides. We further evaluated kinetic parameters from single cycle kinetics assay for CPRNCTREC and

CPRNCPRYC (Figure 3.13, table 3.7). The values obtained for K_D for both these peptides were below 10μM. Thus, we selected these two peptides as best hits for further studies.

Table 3.6: Kinetic parameters obtained from multicycle kinetics in SPR.

Peptide	ka (1/Ms)	kd (1/s)	KD (M)	Rmax (RU)	Chi² (RU²)
CPRNC	3.04	7.96e-03	2.62e-03	46.4	0.0265
CPRNC(S-S)	2.78e+04	2.14e-03	7.69e-08	0	0.0375
CPRNCPRNC	30.3	2.98e-03	9.87e-05	80.2	0.198
CPRNCPRYC	**244**	**2.75e-03**	**1.13e-05**	**3.9**	**0.0736**
CPRNCTREC	**50.5**	**1.45e-04**	**2.88e-06**	**9.2**	**0.00871**
CPRYC	1.5	7.56e-04	5.02e-04	537.8	0.0693
CPRYC(S-S)	9.60e+04	1.72e-03	1.79e-08	0	0.0184
CPRYCPRYC	119	9.24e-02	7.76e-05	6.4	0.0859
CPRYCTREC	1.74e+05	2.13e-03	1.22e-08	0	0.101
CPRYPRNC	9.31e+04	2.21e-03	2.38e-08	0	0.544
CTREC	8.00e+04	2.27e-03	2.84e-08	0	0.0721
CTREC(S-S)	562	2.48e-02	4.41e-05	1.4	0.0178
CTRECPRNC	1.80e+03	2.90e-02	1.62e-05	1.6	0.0422
CTRECPRYC	1.09e+05	2.60e-03	2.38e-08	0	0.275
CTRECTREC	1.16e+03	4.51e-03	3.88e-06	10.5	0.0802
PRN	2.54e+05	4.72e-02	1.85e-07	0	0.218
PRY	1.18e+05	1.25e-02	1.06e-07	0	0.0264
TRE	246	4.00e-03	1.63e-05	2.4	0.0986

Table 3.7: Kinetic parameters for peptide-trypsin interaction obtained by SPR single cycle kinetics

Peptide	ka (1/Ms)	kd (1/s)	KD (μM)	Chi² (RU²)
CPRNCTREC	2.32e+02	6.94e-04	2.7	0.023
CPRNCPRYC	3.24+02	2.25e-03	6.9	0.054

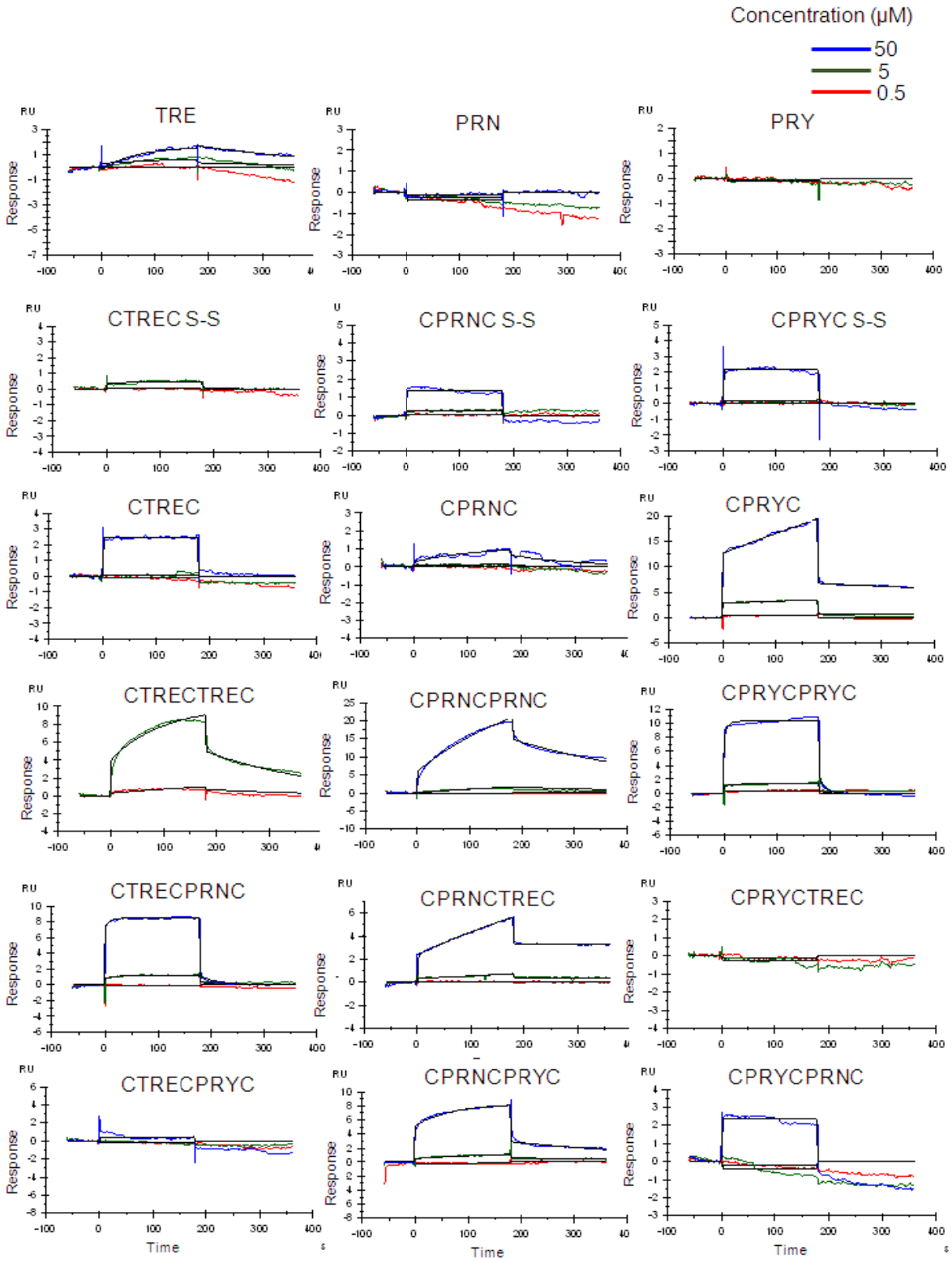


Figure 3.12: Sensorgrams of peptides screened by single cycle SPR kinetics. Colored lines represent original sensorgram, and black lines indicate fitted curves.

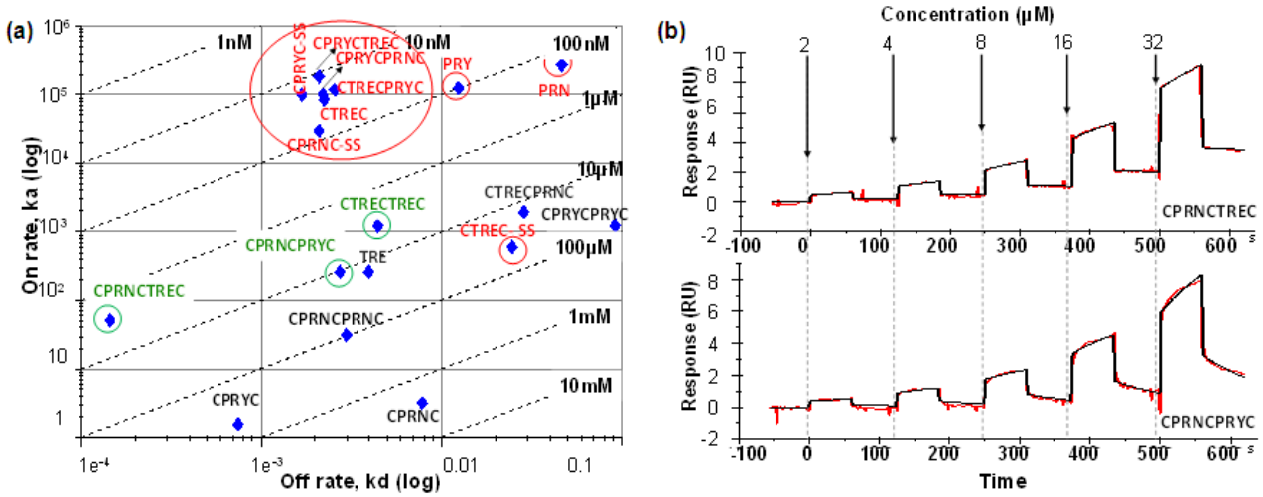


Figure 3.13: Screening of 18 peptides by SPR. (a) On/off plot representing the ka (vertical) and kd (horizontal) values is shown, with dotted diagonal lines showing the dissociation constant, KD. Peptides in red circle were rejected, green circles are best hits, (b)Single cycle SPR kinetics of the selected peptides

3.3.5 Effect of RCL peptides on feeding of lepidopteran insects

In vivo evaluation of RCL peptides was performed on two major lepidopteran pest species, *H. armigera* and *Spodoptera litura*. To test whether cyclic peptides can be used as pest deterrent molecules, a preferential feeding choice was provided to the insects from neonate stage. The peptides CPRNCTREC and CPRNCPRYC were sprayed on castor leaves, which is a general host plant for both insects. Thereafter, the insects were allowed to choose between untreated and peptide treated leaves for feeding. Very clearly, the insects preferred to feed on untreated leaves, whereas peptide treated leaves were consumed in minimal amount by the end of four days (Figure 3.14). The insects consumed ~50% lesser leaves when treated with cyclic peptides. However, upon complete foliage of untreated leaves, the insects also consumed the peptide treated leaves on day 6 of the assay. Therefore, although the cyclic peptides were not the first choice of insects, they were still prone to be consumed by the pests if there are limited untreated leaves. It is thus necessary to evaluate the effect of obligatory peptide feeding on the development of insects. For this, neonates of *H. armigera* and *S. litura* were fed on peptide treated castor leaves for 10 days. It was evident that feeding on peptide treated leaves led to

severe reduction in growth of insects, which led to delay in their developmental cycle. Feeding on BP1 led to more than 60% reduction in larval weight of both insect species.

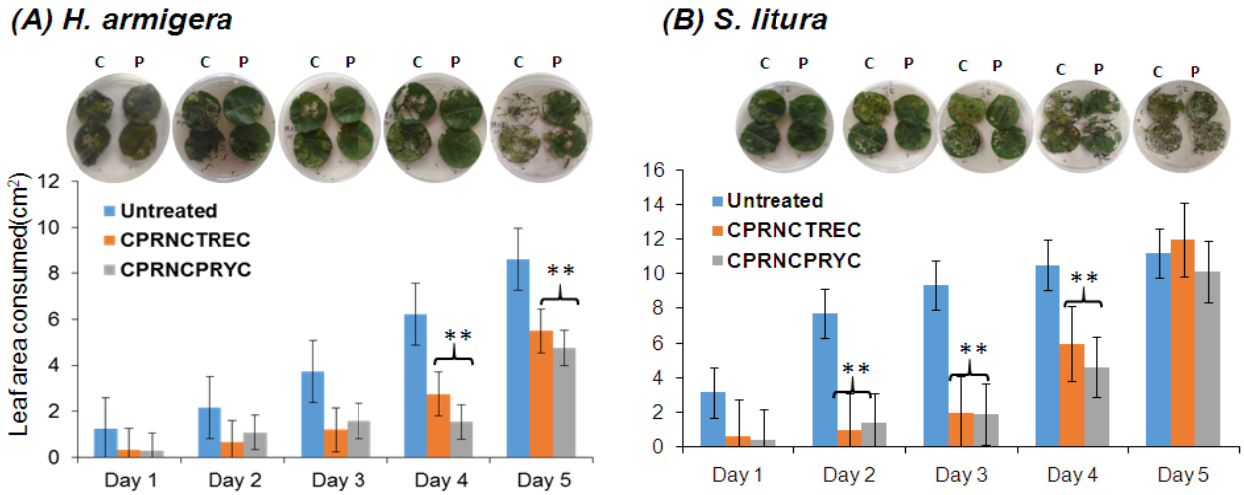


Figure 3.14: Feeding choice assay. *H. armigera* and *S. litura* were subjected open choice assay with castor leaves treated with peptides at 10µM/cm^2.

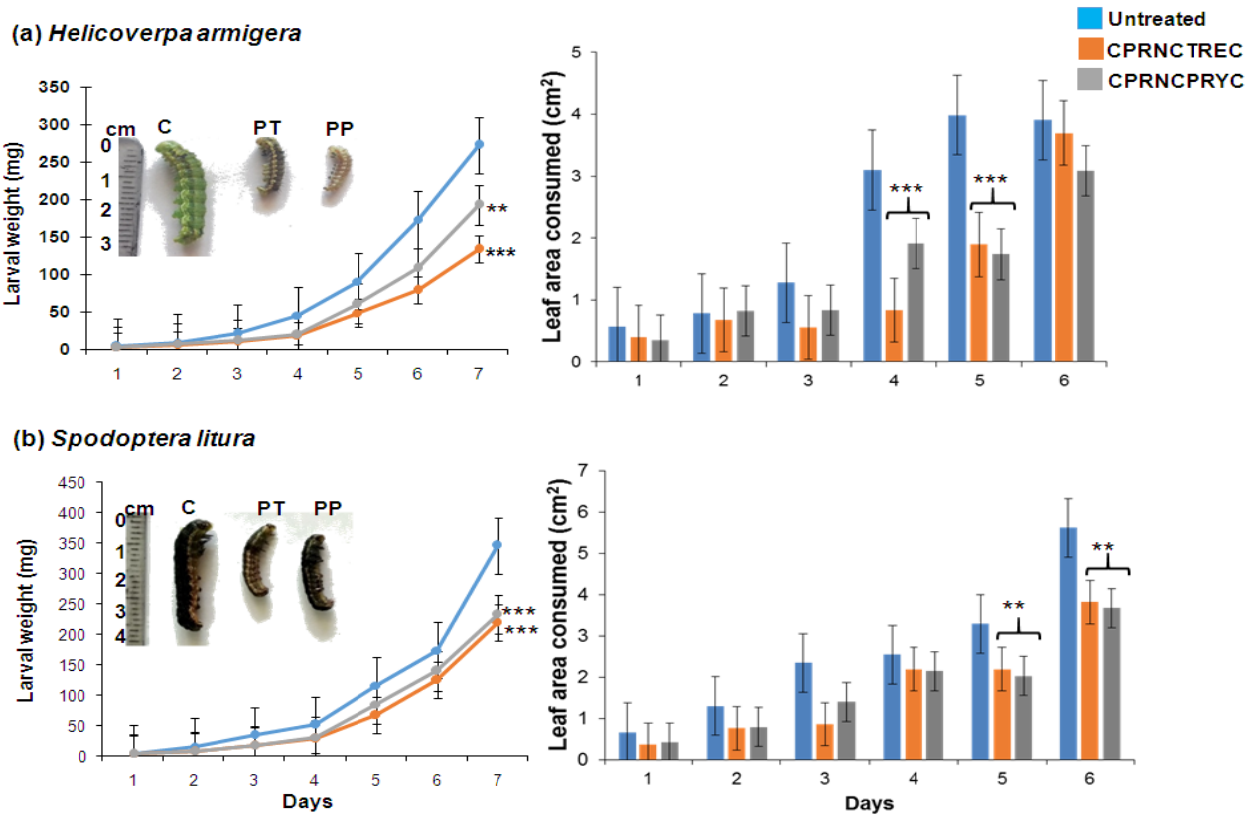


Figure 3.15: Effect of peptides on growth of insects. *H. armigera* and *S. litura* were exposed to

peptide treated leaves (10µM/cm^2) for 7 days. Significant difference in foliage and larval weights is represented by asterisks (***p<0.001; **p<0.05; *p<0.01).

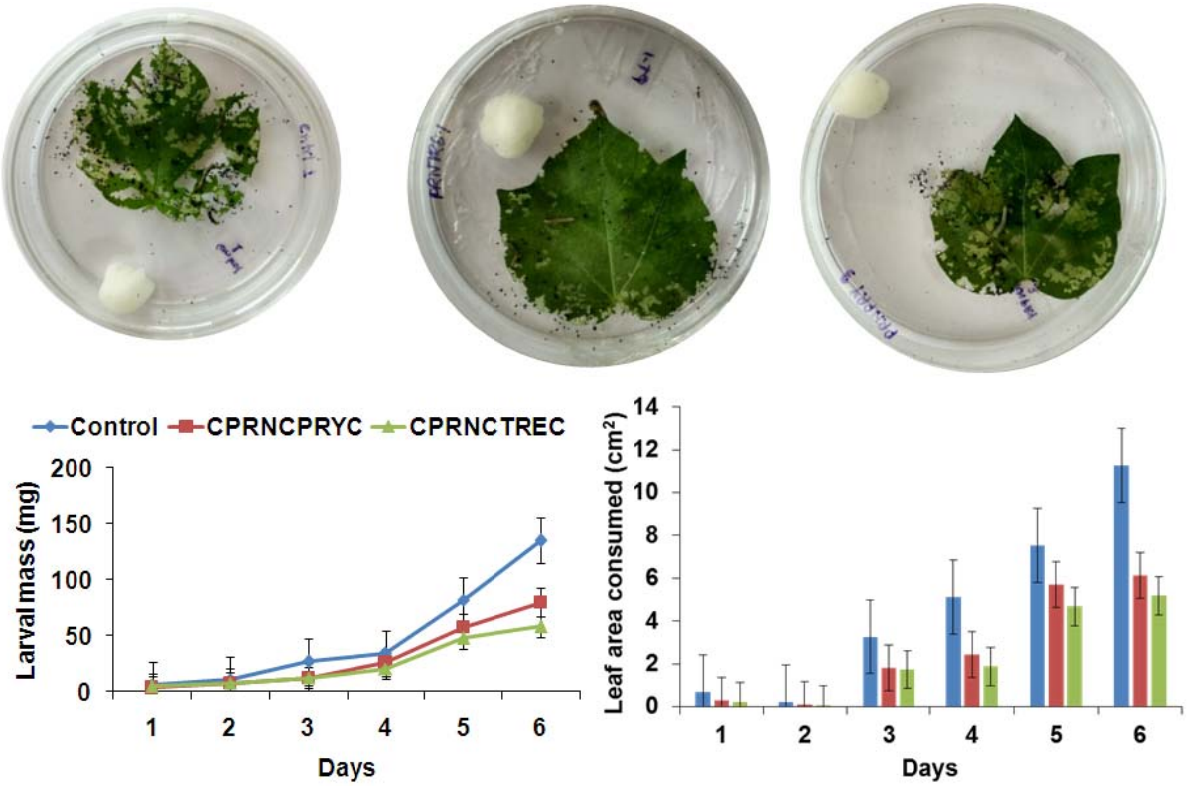


Figure 3.16: Effect of peptide treated cotton leaves on insects. (a) *H. armigera* and *S. litura* were exposed to peptide treated cotton leaves (10µM/cm^2) for 7 days. Difference in foliage (b) and larval weights (c) is shown for six consecutive days.

Similarly, feeding with CPRNCPRYC resulted in 40 to 50% reduction in larval weight (Figure 3.15). However, we could not see a significant difference in the mortality of pests, even when the insects were exposed to peptides from neonate stage. Similar results were obtained when the first instar larvae of *H. armigera* were fed on another host plant, cotton, treated with peptides CPRNCTREC or CPRNCPRYC. There was growth reduction of 50% in peptide fed insects relative to untreated leaves (Figure 3.16). Correspondingly, the leaf area consumed by insects was significantly higher in untreated leaves than peptide treated cotton leaves. This clearly showed that the peptides causing distaste in the insects, which leads to negative effect on growth of insects. Furthermore, the peptides were found to be stable on leaf surface for 10 days in environmental conditions, as analysed by MALDI-TOF spectra (Figure 3.17). Hence, RCL

cyclic peptides could serve as excellent molecules for development of alternative pest control strategies.

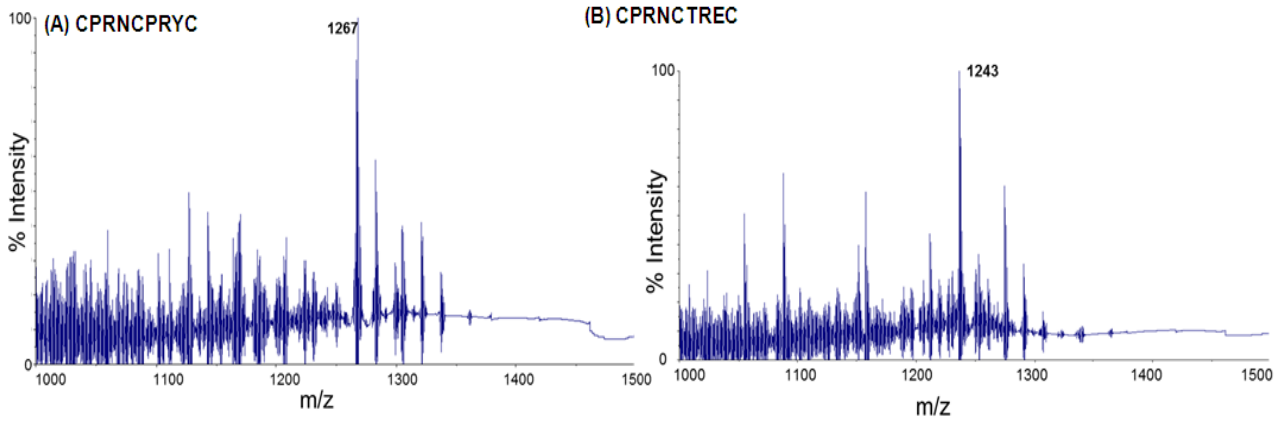


Figure 3.17: Stability of peptides on leaf surface. MALDI-TOF MS of leaf washed with water after 10 days of feeding assay.

3.3.6 Identification of target proteins in insect gut

Based on the *in vivo* inhibitory activity, we asked that whether cyclization of RCL peptides hints to a possibility of interaction with other proteins in the insect midgut. Therefore, we identified the target proteins in the gut of *H. armigera* by affinity based pull down assay using biotin tagged peptides. As control, insect gut extract treated with biotin was also used. MS/MS based identification of eluted proteins showed that the cyclic peptides interact majorly with serine proteases (Figure 3.18). We identified both trypsin and chymotrypsin-like enzymes (table 3.8). We also identified antioxidant enzyme, superoxide dismutase (SOD) as one of the hit proteins. This points out that the peptides might interfere in the antioxidant mechanism of insects by binding to SOD. Remarkably, SOD was found to bind only to cyclic peptides, and not for linear peptide PRN. The possibility of interaction with biotin was also ruled out, since SOD was not found in the proteins eluted with biotin treated midgut. Therefore, the cyclic peptides may bind to SOD, which might mead to reduction in the availability of antioxidant enzyme, ultimately leading to accumulation of free radicals. This mode of action is opposite to that of the insect response to stress, wherein the antioxidant enzymes are known to upregulate under stress conditions. We also validated these results by *in vitro* biochemical activity assays. It was seen that both the residual protease as well as SOD activity was reduced in the insects fed with RCL

cyclic peptides (Figure 3.18). Thus, the cyclization of peptides leads to functional differentiation of the RCL, which might be reason for enhancement in their in vivo adverse effects.

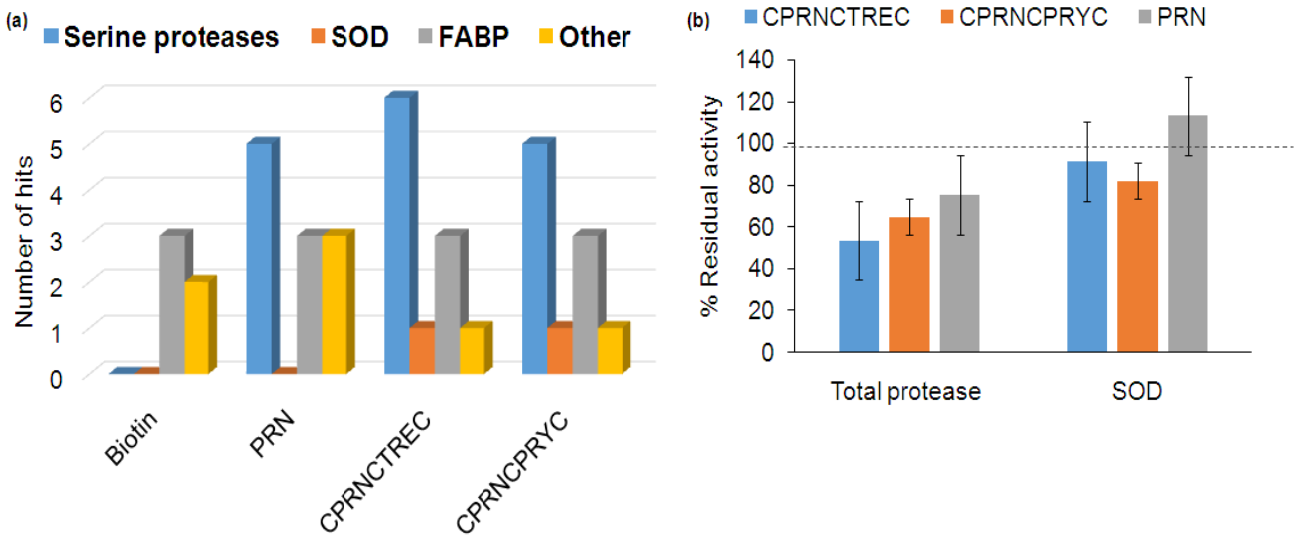


Figure 3.18: In vivo targets of cyclic peptides. (a) Identifiedproteins from *H. armigera* gut extract by MS/MS, (b) Residual proteases and SOD activity in midgut of *H. armigera* larvae (fourth instar) upon feeding with peptides

Table 3.8: Proteins identified by MS/MS analysis after pull down assay.

Accession	Description	PLGS Score	Peptides	Coverage (%)
CPRNCPRYC-Biotin				
J7IF09	Trypsinogen OS Helicoverpa punctigera PE 2 SV 1	477.2182	11	21.2
O18436	Serine protease 5 OS Helicoverpa armigera PE 2 SV 1	1360.533	10	44.0
O18444	Chymotrypsin like protease Fragment OS Helicoverpa armigera PE 2 SV 1	874.7203	11	31.5
O18447	Serine protease OS Helicoverpa armigera PE 2 SV 1	1348.475	18	37.0
O18450	Chymotrypsin like protease OS Helicoverpa armigera PE 2 SV 1	5925.833	20	51.5
O76515	Fatty acid binding protein OS Helicoverpa zea PE 2 SV 1	2855.326	20	63.1
B6CMF9	Fatty acid binding protein 2 OS Helicoverpa armigera PE 2 SV 1	3739.546	31	79.8
B6CMG0	Fatty acid binding protein 3 OS Helicoverpa armigera PE 2 SV 1	4611.932	32	64.3
H9BEW2	Superoxide dismutase Cu Zn OS Helicoverpa armigera PE 2 SV 1	1661.042	12	64.0
O76336	Cytochrome b5 OS Helicoverpa armigera PE 2 SV 1	3971.662	18	71.6
CPRNCTREC-Biotin				
A0A1B0RHP3	Serine protease 3 Fragment OS Helicoverpa zea GN SP3 PE 2 SV 1	295.0183	5	19.3
O18447	Serine protease OS Helicoverpa armigera PE 2 SV 1	2480.579	26	55.9
O18450	Chymotrypsin like protease OS Helicoverpa armigera PE	2053.362	10	21.3

	2 SV 1			
O18436	Serine protease 5 OS Helicoverpa armigera PE 2 SV 1	2239.99	22	45.2
Q9N6C6	HzC4 chymotrypsinogen OS Helicoverpa zea PE 2 SV 1	1144.139	10	35.5
Q9NH07	HzC20 chymotrypsinogen OS Helicoverpa zea PE 2 SV 1	335.319	5	23.3
B6CMF9	Fatty acid binding protein 2 OS Helicoverpa armigera PE 2 SV 1	2121.107	10	65.6
B6CMG0	Fatty acid binding protein 3 OS Helicoverpa armigera PE 2 SV 1	6009.566	27	72.7
O76515	Fatty acid binding protein OS Helicoverpa zea PE 2 SV 1	2984.685	8	45.1
H9BEW2	Superoxide dismutase Cu Zn OS Helicoverpa armigera PE 2 SV 1	2275.891	7	49.6
CPRYCTREC-Biotin				
O18434	Trypsin like protease OS Helicoverpa armigera PE 2 SV 1	5225.509	17	48.0
O18436	Serine protease 5 OS Helicoverpa armigera PE 2 SV 1	3402.905	21	38.5
O18450	Chymotrypsin like protease OS Helicoverpa armigera PE 2 SV 1	1902.373	6	17.9
Q9N6C6	HzC4 chymotrypsinogen OS Helicoverpa zea PE 2 SV 1	3527.733	20	37.2
B6CMF8	Fatty acid binding protein 1 OS Helicoverpa armigera PE 2 SV 1	1894.092	21	51.4
B6CMF9	Fatty acid binding protein 2 OS Helicoverpa armigera PE 2 SV 1	1009.689	3	33.5
B6CMG0	Fatty acid binding protein 3 OS Helicoverpa armigera PE 2 SV 1	3405.637	19	65.9
O76515	Fatty acid binding protein OS Helicoverpa zea PE 2 SV 1	1859.93	6	31.5
A0A291ARU4	Glutathione S transferase OS Helicoverpa armigera GN GST8 PE 2 SV 1	524.4631	13	58.1
PRN-Biotin				
O18434	Trypsin like protease OS Helicoverpa armigera PE 2 SV 1	1082.954	7	22.4
O18436	Serine protease 5 OS Helicoverpa armigera PE 2 SV 1	800.7201	7	24.0
O18436	Serine protease 5 OS Helicoverpa armigera PE 2 SV 1	2485.677	8	33.0
O18447	Serine protease OS Helicoverpa armigera PE 2 SV 1	992.827	8	27.1
O18448	Chymotrypsin like protease Fragment OS Helicoverpa armigera PE 2 SV 1	1635.499	7	28.5
O18450	Chymotrypsin like protease OS Helicoverpa armigera PE 2 SV 1	1758.566	10	34.2
O76515	Fatty acid binding protein OS Helicoverpa zea PE 2 SV 1	2791.753	2	23.3
O76515	Fatty acid binding protein OS Helicoverpa zea PE 2 SV 1	2788.983	4	31.5
B6CMG0	Fatty acid binding protein 3 OS Helicoverpa armigera PE 2 SV 1	772.1895	7	34.0
B9UCQ5	Ultraspiracle isoform 1 OS Helicoverpa armigera PE 2 SV 1	585.9171	10	17.8
I4ZSI7	Ribosome recycling factor OS Acinetobacter sp HA GN frr PE 3 SV 1	373.3872	7	21.7
I4ZW46	Uncharacterized protein OS Acinetobacter sp HA GN HADU 01907 PE 4 SV 1	448.5259	2	20.0
O76336	Cytochrome b5 OS Helicoverpa armigera PE 2 SV 1	365.6363	1	12.5
Biotin				
B6CMF9	Fatty acid binding protein 2 OS Helicoverpa armigera PE 2 SV 1	695.3282	5	47.0
B6CMG0	Fatty acid binding protein 3 OS Helicoverpa armigera PE 2 SV 1	645.2405	5	28.7
L7R1Y9	Putative enolase protein Fragment OS Helicoverpa zea PE 2 SV 1	460.2393	12	32.2
Q8T7V0	Cytoplasmic actin A3a2 OS Helicoverpa zea PE 3 SV 1	560.1713	4	11.7

3.3.7 Insect response to RCL cyclic peptide feeding

The effect of RCL peptide feeding on the midgut physiology of insects was studied by relative gene expression analysis using qRT-PCR (Figure3.19). Representative trypsin, chymotrypsin genes and SOD, as identified in pull down assay were selected for gene expression analysis. The results indicated differential expression as compared to control (untreated) insects. HaTry7 was downregulated upon bicyclic peptide feeding, whereas it was upregulated in PRN fed insects. Similarly, HaChy4 was downregulated upon bicyclic peptide ingestion. Similar to the biochemical activity results, Cu/Zn SOD was downregulated in bicyclic peptide fed insects, while feeding with linear peptide PRN caused upregulation of SOD expression. Thus, cyclic RCL peptides might bind to several proteins in the insect gut, which causes alteration in the midgut physiology.

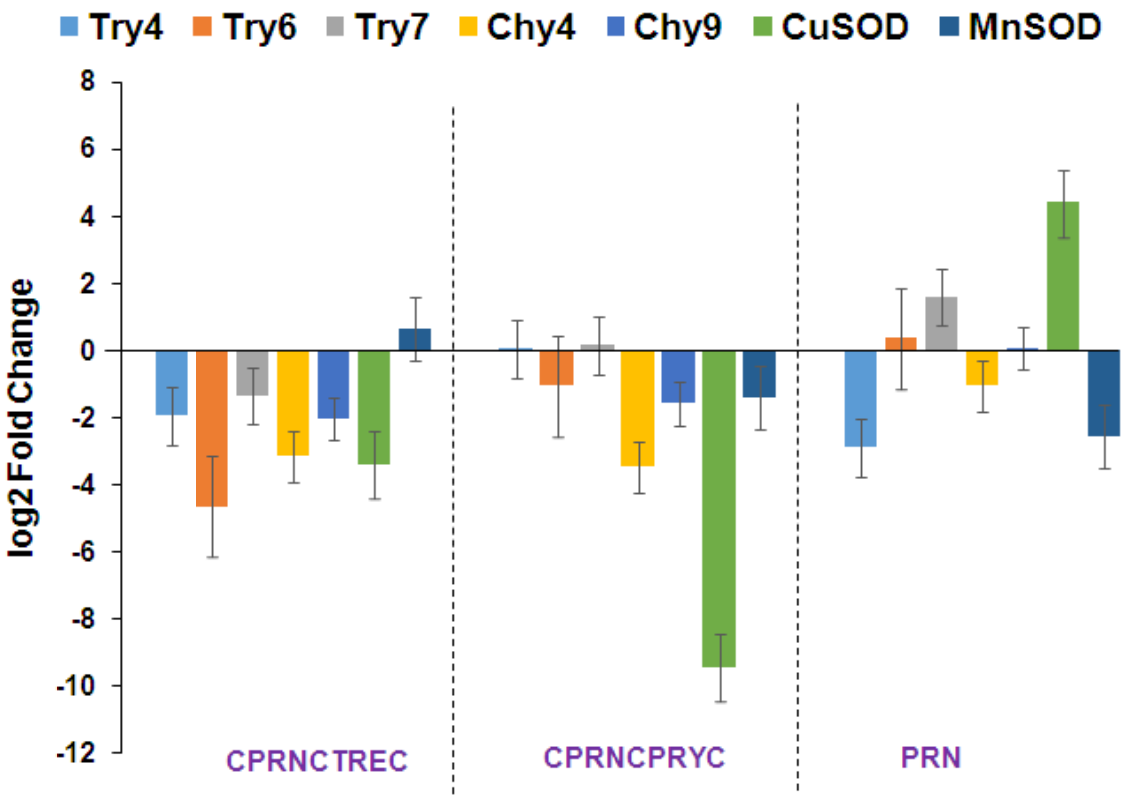


Figure 3.19: Relative gene expression of protease and SOD genes upon exposure to peptides

3.3.8 Binding modes of cyclic peptides with proteases

We explored the possible binding modes of bicyclic peptides CPRNCTREC and CPRNCPRYC with trypsin like enzymes by molecular dynamics simulation with bovine and *H. armigera* trypsin (HaTry, Uniprot ID: B1NLE4). Comparison of the "most populated cluster" obtained after 200 ns MD simulation with HaTry showed that cyclization enhances the interaction of RCL peptides with active site residues His69 and Ser211, specificity determining residue Gln192, and other residues which line the active site cavity (Figure 3.20, 3.21). Hydrophobic interactions also increased as compared to linear peptides, thus resulting in higher binding affinity. Being larger in size, they cover a higher interface surface area on the trypsin molecule. In particular, Arg3 of first loop in both CPRNCTEC and CPRNCPRYC, lies in close proximity to the active site His69 of HaTry. However, the poses of peptides differ in other interactions, such as, Asn4 in CPRNCTREC forms H-bond with Gln213 of HaTry, whereas in CPRNCPRYC, it forms H-bond with Gln192. Also, Pro2 of PRN is positioned near Tyr113, where it makes Van der Waals interaction in CPRNCTREC, but it is involved in H-bond in CPRNCPRYC. It is important to note that this residue determines the relative potency of linear RCL peptides towards trypsin(Saikhedkar et al., 2018). It helps to stabilize the complex by hydrophobic interaction with P2 Pro residue of the inhibitor, when simultaneously P1 Arg is in involved in hydrophilic contact at active site.

Also, the second loop of bicyclic peptide CPRNCTREC, TRE, is positioned near nonspecific substrate binding pocket of HaTry(Chen et al., 2015; Schmidt et al., 2005), lined by residues Trp46-Ser51. Arg7 and Thr6 of the peptide make H-bond with Gly214 (oxyanion hole), and His69 (active site). In addition to it, this loop contributes in numerous hydrophobic contacts with residues close to the active site, like Gly112, Tyr113, Tyr108, Phe71 etc. Similarly, in CPRNCPRYC, the second loop, PRY is placed near specificity determining residues, Gly235, Tyr185, and Gln160 of HaTry, where it participates in hydrophobic interactions.

Switching of sequences on the two loops led to significant decrease in the number of interactions. CTRECPRNC was placed in a similar orientation as CPRNCTREC, but could not interact with HaTry, which might be leading to reduction in binding potency (Figure 3.20B, 3.21B). Crucial interactions with HaTry, like His69, Tyr113, Gly235 were missing, however, H-bonds with Gly233, Gln190, Gln160 and Ser47 were present. Similarly, reversing the sequence of CPRNCPRYC to CPRYCPRNC led to shifting of the peptide far from HaTry active site, with

TBMB facing towards HaTry and the P1 Arg3 sticking out from the complex. Therefore, none of the Arg was making interaction with HaTry.

Study of intramolecular H-bond patterns for bicyclic peptides showed that the peptides fold around TBMB scaffold, and these interactions help in stabilizing the peptide as well as its interactions with trypsin(Angelini et al., 2012; Chen et al., 2014). This observation is supported by previous reports, which show that the TBMB scaffold stabilizes the peptide structure by formation of intramolecular interactions, thereby optimizing the conformation of peptides.

Comparison of binding poses for bovine trypsin showed that the bicyclic peptides showed lesser interactions with the active site residues of protease. This might be one of the reasons for lower affinity of the peptides with bovine trypsin as compared to HaTry. Specifically, the peptide CPRNCTREC and CPRNCPRYC lied close to active site of BovTry, and made H-bond contact with active site Ser195, oxyanion hole Gly193, and other residues close to substrate binding pocket. Further, the second loop of the bicyclic peptides was not significantly involved in interactions with the BovTry. Thus, the sequence of peptide greatly affects the resulting conformation of bicyclic peptide, and thus the interaction with target protease (Figure 3.22, 3.23).

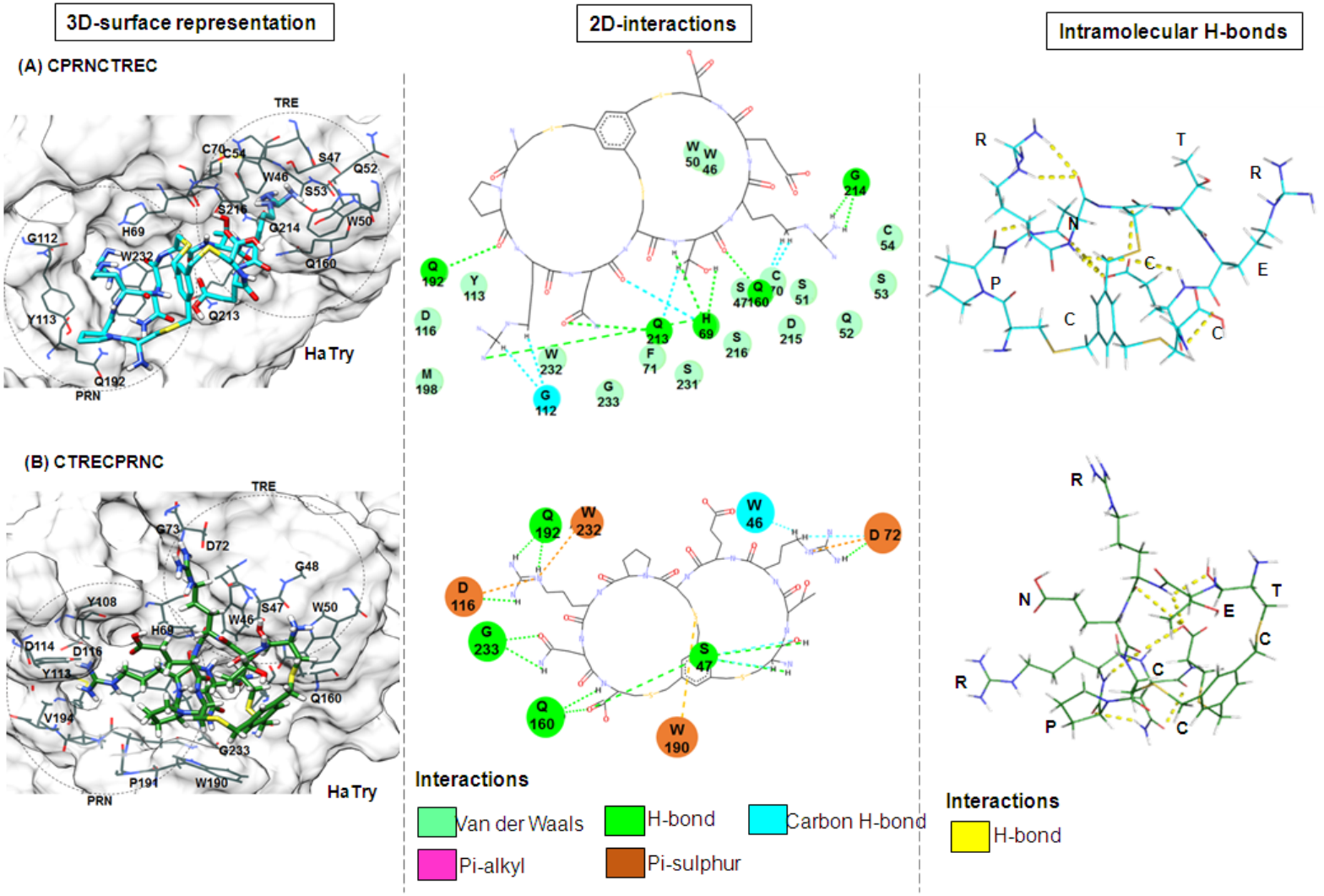


Figure 3.20: Binding poses of cyclic peptides CPRNCTREC and CTRECPRNC with HaTry. HaTry is shown in surface representation, with interacting residues as shown as wires. Peptides are shown as sticks. Also shown 2D interaction diagrams of peptides with HaTry, colour coded according to legend. Intramolecular H-bond pattern of peptides in bound pose is shown as wire, H-bonds are represented as yellow dotted lines

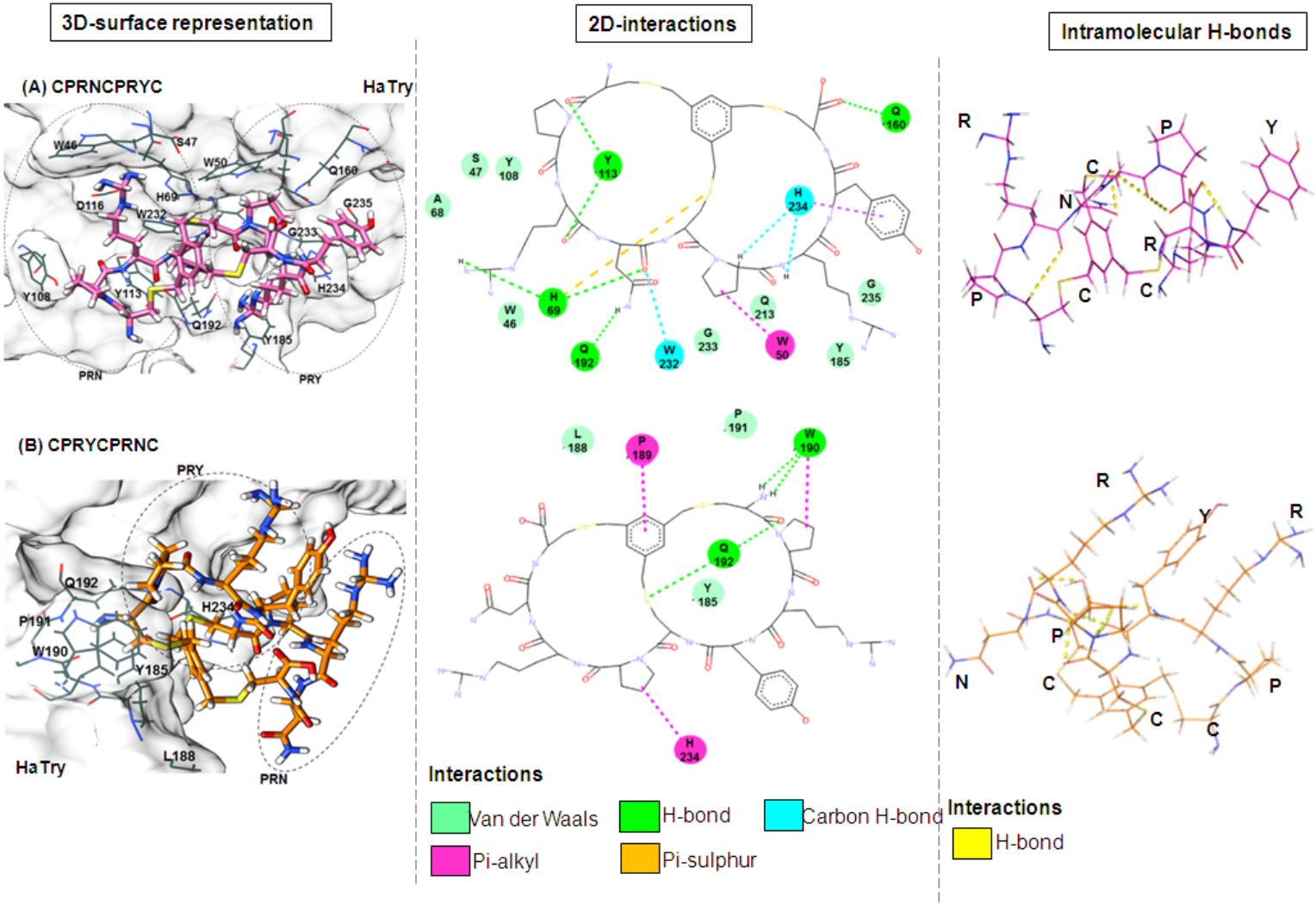


Figure 3.21: Binding poses of cyclic peptides CPRNCPRYC and CPRYCPRNC with HaTry. HaTry is shown in surface representation, with interacting residues as shown as wires. Peptides are shown as sticks. Also shown 2D interaction diagrams of peptides with HaTry, colour coded according to legend. Intramolecular H-bond pattern of peptides in bound pose is shown as wire, H-bonds are represented as yellow dotted lines

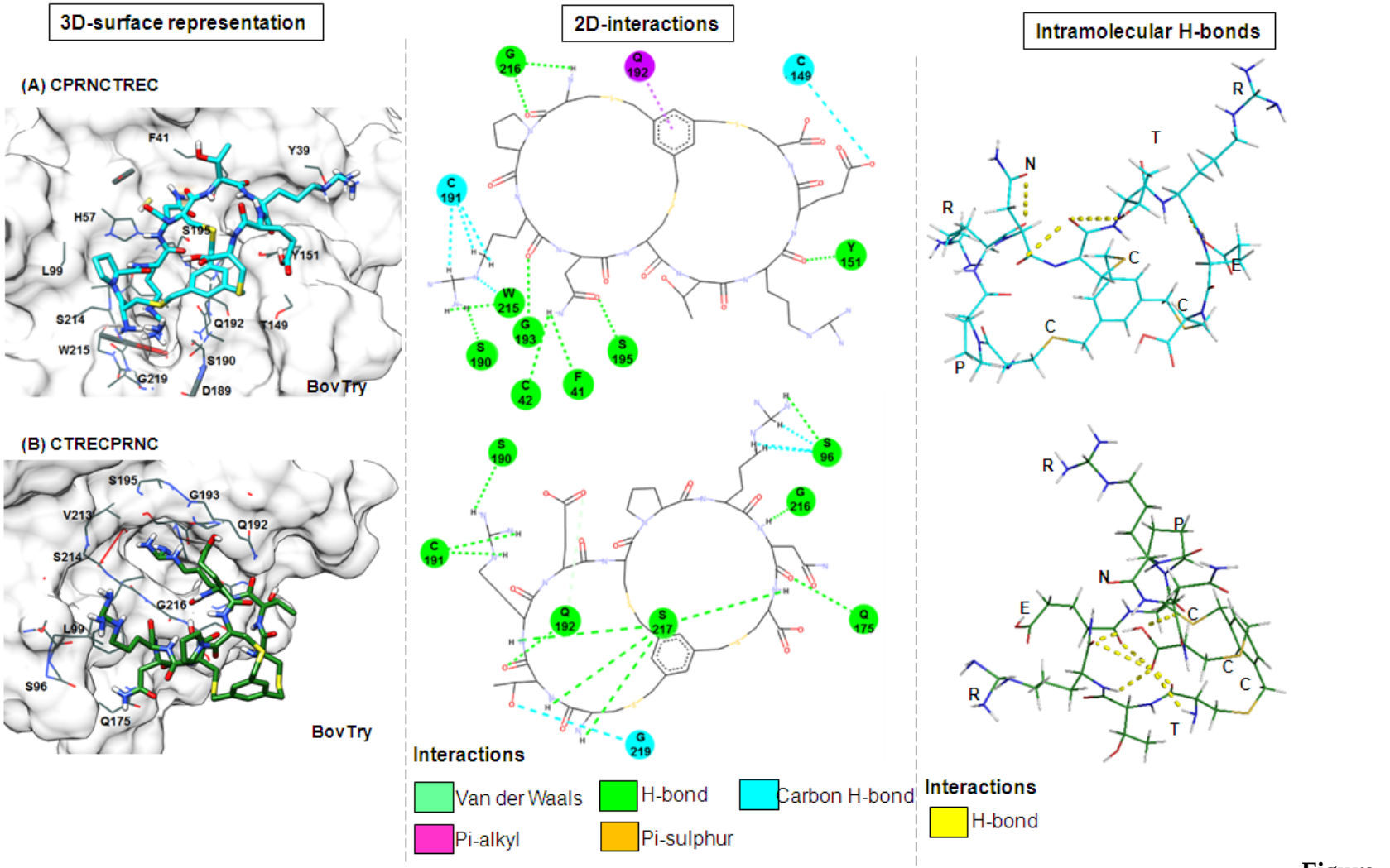


Figure 3.22: Binding poses of cyclic peptides CPRNCTREC and CTRECPRNC with Bovine trypsin. BovTry is shown in surface representation, with interacting residues as shown as wires. Peptides are shown as sticks. Also shown 2D interaction diagrams of peptides with BovTry, colour coded according to legend. Intramolecular H-bond pattern of peptides in bound pose is shown as wire, H-bonds are represented as yellow dotted lines

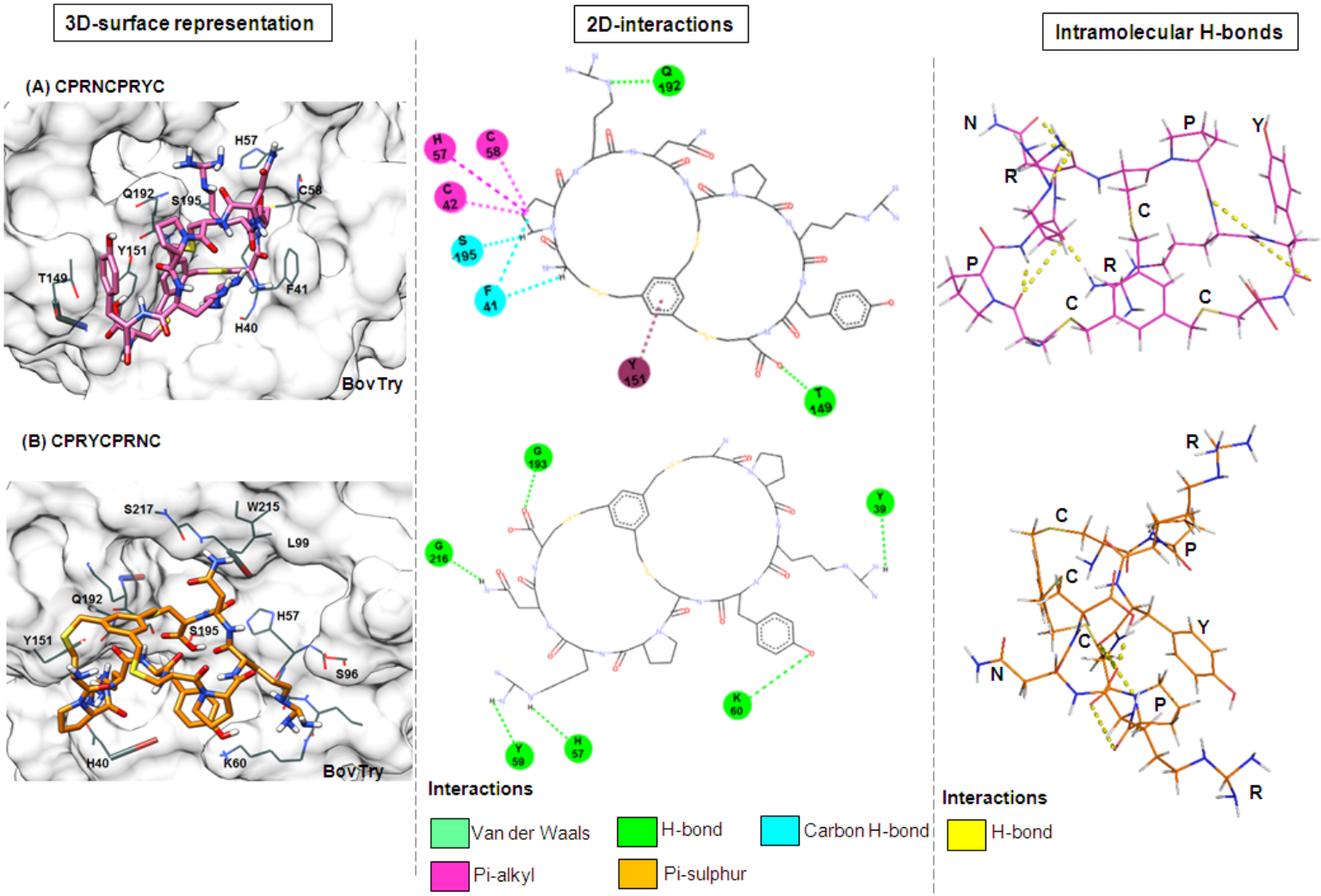


Figure 3.23: Binding poses of cyclic peptidesCPRNCPRYC and CPRYCPRNC with Bovine trypsin. BovTry is shown in surface representation, with interacting residues as shown as wires. Peptides are shown as sticks. Also shown 2D interaction diagrams of peptides with BovTry, colour coded according to legend. Intramolecular H-bond pattern of peptides in bound pose is shown as wire, H-bonds are represented as yellow dotted lines

3.4 Conclusion

In this study, we have demonstrated that tripeptide RCL regions of Pin-II type protease inhibitors can be grafted on a small scaffold TBMB, which enhances their effectiveness compared to linear peptides. The bicyclic peptides provide a fixed conformation to the RCL peptide, which might be leading to similar potency as the native Pin-II protein(Taylor et al., 2014) (Table 3.9). In comparison to their linear counterparts, bicyclic RCLs showed enhanced inhibitory potency and cause feeding avoidance in lepidopteran pests, *H. armigera* and *S. litura*. They not only inhibit the serine proteases, but also might interfere with the antioxidant mechanism of insects, as indicated by affinity based pull down assay. Further, molecular dynamics simulation suggested that the probable mode of inhibition by bicyclic peptides involves concurrent interaction with the active site and specificity-determining residues of trypsin-like insect proteases. Hence, characteristics like environmental stability, and specificity and potency towards lepidopteran insects suggests that they can be used as peptide based pest control agents. In order to tackle the adaptive nature of lepidopteran pests, combinations of bicyclic RCL peptides could be generated using the natural Pin-II RCL sequences, making them value-added molecules in pest management strategies.

Table 3.9: Comparison of inhibitory potential of Pin-II PIs and RCL peptides

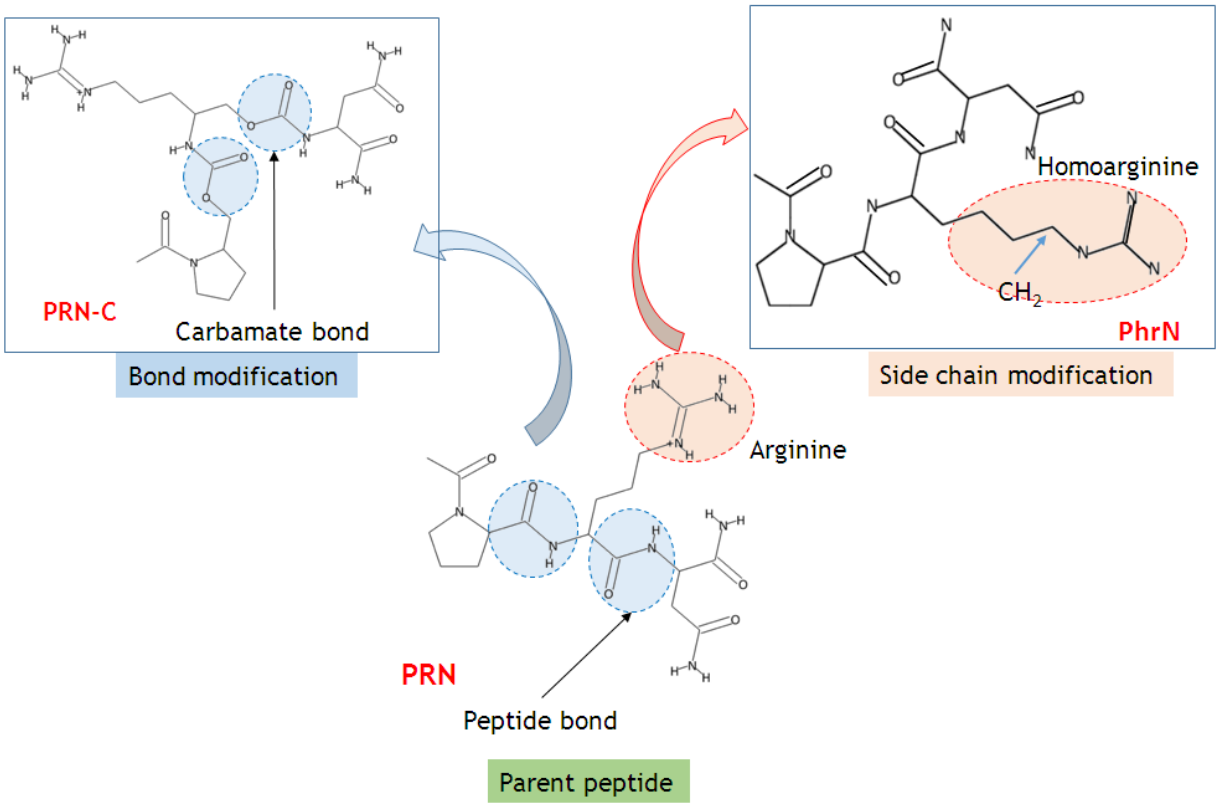
	CanPI(Tamhane et al., 2005)	**IRD**(Joshi et al., 2014)	**RCL**(Saikhedkar et al., 2018)	**RCL-cyclic**
IC_{50} (µM)	300	60	200	10
% Insect growth reduction	30	50	40	70

Key highlights of this study:

- Cyclic RCL peptides are more potent as compared to parental linear tripeptides
- Cyclic RCL peptides show feeding deterrent and growth inhibitory activity against insect pests, *H. armigera* and *S. litura*
- Cyclic RCL peptides interact with serine proteases and also with antioxidant enzymes in insects
- They show proteolytic stability in insect midgut and are stable in ambient conditions.
- Mode of interaction of cyclic peptides involves interaction of P1 Arg with trypsin active site and additional contacts with substrate binding residues.

Chapter 4

Design and evaluation of RCL peptidomimetics



4.1 Introduction

In vitro and *in vivo* studies on the RCL regions of Pin-II PIs highlighted PRN as one of the selective inhibitor of insect trypsin. Analysis of three dimensional crystal structure of PRN in complex with trypsin showed that the P1 Arginine was deeply seated in the active site of trypsin and forming hydrogen bonds with the active site residues His 48, Asp 92 and Ser 185 (Figure 4.1). Also, it formed crucial interactions with Asp 179. Hence, the P1 residue determined the potency of the tripeptide.

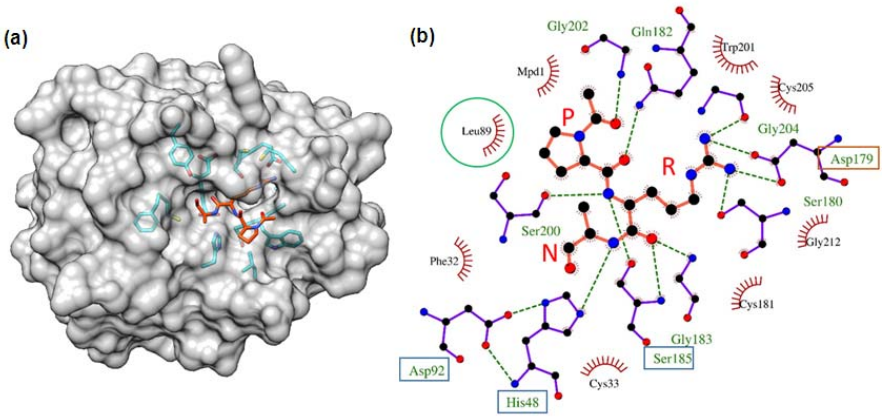


Figure 4.1. Reactive loop peptide -protease interactions. (a) Reactive loop peptide PRN in active site of trypsin, trypsin is represented as surface with active site residues highlighted in blue sticks; (b) 2D interaction diagram of PRN with trypsin

In order to enhance the reactive potential of these peptides, arginine residues can be targeted for modification. Peptidomimetics have been recently popular for design of molecules with enhanced functional properties and stability towards proteases. The underlying reason is that the proteases are unable to identify the modified residues, and thus cannot cleave in the scissile bond. Particularly, peptide bond modification results in enhanced stability of peptides. For example, the replacement of amide linkages by carbamate bonds leads to enhanced stability and potency, as shown in the case of R-X-R motifs for cell penetrating properties(Patil et al., 2012). The carbamates have also been popular as pest control molecules. Recently, the carbamate linkages have been used in the design of protease inhibitors. Carbamate-containing kallikrein, thrombin, and elastase inhibitors showed high potency and stability to degradation(Ghosh and Brindisi, 2015).

Further, the use of modified amino acids for altering the inhibitory potential of peptides has also been well documented. Plants are reported to produce unnatural amino acids as a defence mechanism upon insect attack(Huang et al., 2011). The arginine analogue, canavanine is a non-protein amino acid produced in plants for defence against insects. L-canavanine is used a nitrogen storage compound in the seeds of Leguminosae family. It is an analogue and antimetabolite of l-arginine, and shows structural similarity to the natural amino acid. Thus, it disincorporates into proteins in place of arginine. This is considered as the major mode of action of this amino acid. Also, due to this property, l-canavanine is shown to be highly toxic to other organisms including bacteria, fungi, yeast, algae, plants, insects, and mammals(Staszek et al., 2017) (Figure 4.2).

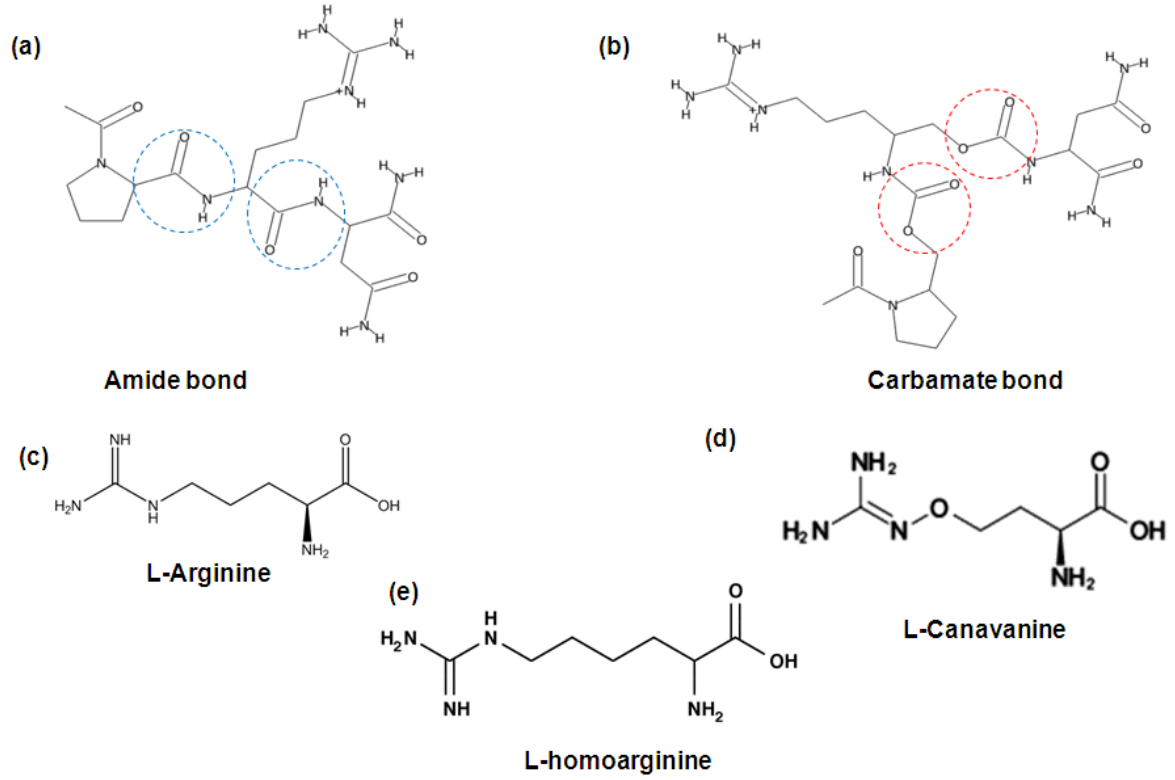


Figure 4.2: Peptide modifications. (a) Amide bond in the peptide backbone is replaced by carbamate bond (b); (c) Structure of naturally occurring arginine and its un-natural analogcanavanine (d), and homoarginine (e).

Thus, PRN was modified in order to improve the potency, which will be evaluated for its activity against *H. armigera* gut proteases. The following modifications were incorporated:

- The peptide bond was replaced with carbamate bond
- Unnatural amino acid canavanine was used as an anti-metabolite in place of arginine
- Arginine side chain was elongated by addition of methyl group, in order to increase penetration into the active site [this modified version of the arginine residue will be represented as 'hr' (homoarginine)].

Based on the above background, the following objectives were defined:

- To design modified peptides based on reactive centre loops of IRDs
- In silico binding analysis of peptides with *H. armigera* and bovine trypsin
- Synthesis and purification of peptides
- Evaluation of protease inhibitory activity by *in vitro* assays
- Biophysical assessment of peptide interactions
- Assessment of proteolytic stability of peptides

4.2 Methods

4.2.1 In silico analysis

a. Design of peptides

Eight peptides were designed using Marvin Sketch and Discovery Studio (Table 4.1). The 2D structures were cleaned in 3D with hydrogenation to remove steric clashes using the inbuilt tool from Marvin sketch. The 3D structures were then saved as PDB files.

Table 4.1: List of Peptides designed for in silico analysis

Peptide	Sequence	Modification
PcnN	Pro-Canavanine-Asn	Arg was substituted by canavanine
PhrN	Pro-M-Arg-Asn	Arg side chain elongated with -CH_2 (Homoarginine)
PorN	Pro-Ornithine-Asn	Arg was substituted by Ornithine
TcnE	Thr-Canavanine-Glu	Arg was substituted by canavanine
ThrE	Thr-M-Arg-Glu	Arg side chain elongated with -CH_2 (Homoarginine)
TorE	Thr-Ornithine-Glu	Arg was substituted by Ornithine
TRE-C	Thr-Cab-Arg-Cab-Glu	The amide bond was replaced by a carbamate bond
PRN-C	Pro-Cab-Arg-Cab-Asn	The amide bond was replaced by a carbamate bond

The structures of peptides are given in figure 4.3.

a. Docking

Energy minimization of peptides was performed by Prodrg server. The peptide 3D structures were prepared for docking with the help of AutoDock tools 1.5.6. The structure of bovine trypsin was downloaded from RCSB-PDB (PDB-ID: 4I8H), while *H. armigera* trypsins were used as modelled structures. The ligand molecules were docked into the active site of bovine trypsin and *H. armigera* trypsins by AutoDockVina. All dockings were performed by running virtual screening shell script on the NCL-HPC Linux cluster. Grid boxes were designed using 25x25x25 points and 1Å spacing centred on the catalytic triad. Analysis of docking results was carried out manually by using Autodock tools: out of the eight possible binding modes, the one with greatest proximity to the active site and most negative free energy of binding was selected. All the solutions were catalogued and represented as heat maps.

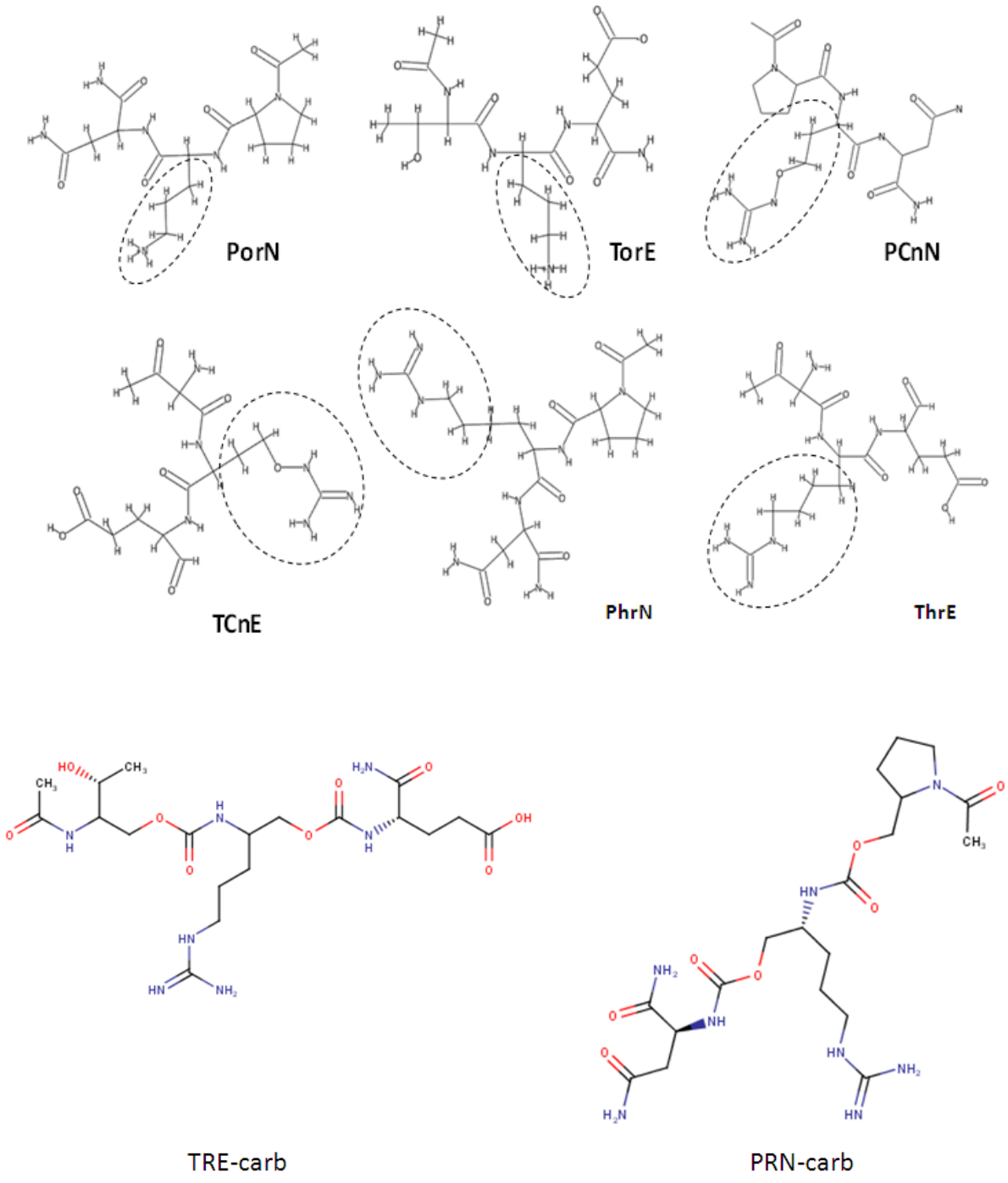


Figure 4.3: Structures of modified peptides designed for in silico analysis

4.2.2. Synthesis of peptides

Modifications were performed for the RCL sequence, PRN. PRN-carbamate, PhrN, PKN-carbamate were synthesized. As controls, PRN and PKN were used. Activated monomers for carbamate bond containing peptides were synthesized as mentioned in (Patil et al., 2012). Oligomeric peptides were then synthesized by Fmoc- solid phase synthesis on MBHA resin. For homoarginine containing peptide, lysine (Fmoc-Lys(boc) OH) was used during the synthesis of oligomer. The lysine side chain amine group was deprotected in 50% TFA/DCM and then neutralization with 5% DIPEA/DCM. Conversion to homoarginine was achieved by treatment with 10 eq. of 3,5-Dimethylpyrazole-1-carboximidamide in DMF. Cleavage of the peptides from resin was performed by TFA/thioanisole/EDT/TFMSA procedure as mentioned in section 2.2.1. Peptide purification was performed by RP-HPLC using Agilent TC-2 C18 column, and linear gradient of 5% ACN/water to 50%ACN/water. Peaks were collected and dried in rotary evaporator. Molecular weights were confirmed by MALDI-TOF spectra obtained on AB Sciex TOF/TOF 5800 spectrometer using DHB (2, 5 dihydroxybenzoic acid) as matrix.

4.2.3. Biochemical assays

Trypsin activity was estimated through chromogenic substrate BA*p*NA (Benzoyl-L-arginyl-*p*-nitroanilide), as mentioned in section 2.2.2. IC_{50} values for inhibition of *H.armigera* gut trypsin were determined. The HGP was incubated with increasing concentrations of peptides (10µM to 2mM) and residual protease activities were determined as mentioned above.

4.2.4 Surface plasmon resonance studies

SPR kinetics experiments were carried out on Biacore X100 (GE Healthcare, Bengaluru). Bovine trypsin was purchased from Sigma Aldrich (T8003). Bovine trypsin powder was resuspended in acetate buffer pH 5.5 and immobilized by the amine coupling method on a CM5 sensor chip according to the manufacturer's protocol (GE Healthcare). Peptides were diluted in HBS-EP+ with 0.1 M $CaCl_2$ at concentrations ranging from 4 to 32 µM, with 2 fold escalations for kinetics characterization of selected molecules using single cycle kinetics format (5 concentrations injected serially in one cycle). The association and dissociation times were 60 seconds each for single cycle kinetics. The sensor surface was regenerated after each injection cycle with 10 mM Glycine-HCl pH 2.5 (GE Healthcare) for 60 seconds. The sensorgrams of test

flow cells were subtracted from the reference flow cell. The kinetic fitting was carried out with Biacore X100 evaluation software by global fitting using 1:1 binding model. The kinetics data were described as aska (association rate or On-rate) and kd (Dissociation rate or Off-rate) and K_D (affinity constant or Equilibrium constant of Dissociation). Each SPR run was evaluated based on the statistical measurements provided by the Biacore X100 evaluation software, like Chi^2.

4.2.5 Proteolytic stability assays

Stability of peptides in presence of HGP was monitored by MALDI-TOF analysis(Mishra et al., 2010). Reduction in the intensity of the RCL peptide is monitored by MALDI-TOF-MS on the addition of target protease. Peptides were incubated with HGP for 1h, 3h and 6h at 37°C. The reaction mixture of 5 µl volumes was mixed with 20 µl of freshly prepared DHB matrix for peptides and 2 µl aliquots in three replicates were spotted on the stainless steel plate and MALDI-TOF profiles were acquired.

4.3 Results and discussion

4.3.1 In silico designing and docking

Docking of modified peptides in the active site of *H. armigera* trypsin 4 and bovine trypsin indicated that the peptides show differential binding with proteases. Also, their increased binding energy with respect to tripeptides suggested that modification of P1 Arg might increase the potency of peptides. The binding affinities generated by AutodockVina are shown in Table 4.2.

Table 4.2: Binding affinities of peptides with insect and bovine trypsin

Sr. No	Peptide	*H. armigera* trypsin 4 (kcal/mol)	Bovine trypsin (kcal/mol)
1	PcnN	-6.3	-6.6
2	PhrN	-5.3	-4.7
3	PorN	-5.3	-5.5
4	PRN	-5.2	-6.5
5	TcnE	-6.1	-6.5
6	ThrE	-5.1	-5.8
7	TorE	-5.3	-5
8	TRE	-5.3	-6.1
Carbamate linked peptides			
9	PKN-C	-5.5	-6.1
10	PRN-C	-6.8	-6.3
11	TRE-C	-6.3	-6.5

The results showed that few modifications showed positive effects on the binding of modified peptides as compared to parental peptides. On the basis of in-silico analysis, PglN, PRN-C and PKN-C were selected for further studies.

4.3.2 Interaction between *H. armigera* Try4 (HaTry4) and peptides

Study of binding modes generated by AutoDockVina highlighted the possible orientation of peptides in the active site of HaTry4. The detailed interactions visualized by Discovery studio indicated that these peptides interact by means of hydrogen bonds with the active site residues of protease (His 69, Ser211). This suggests that the peptides are deeply seated in the protease active

site and are held by strong H-bonds. In addition, various hydrophobic contacts like Pi-alkyl interactions and Pi-donor hydrogen bond were observed in PRN-C, which might help in better placement of PRN-C in protease catalytic site (Table 4.3). Furthermore, PKN-C was oriented by means of additional Carbon hydrogen bonds, Pi-sigma and alkyl hydrophobic interactions. These differences in bonding might reflect in the inhibition potential of these peptides. On the other hand, it was seen that PglN was penetrating deep inside the HaTry4 active site and formed three hydrogen bonds with active site His69. These analyses provided an insight into the probable mechanism of peptide-protease binding (Figure 4.4, 4.5).

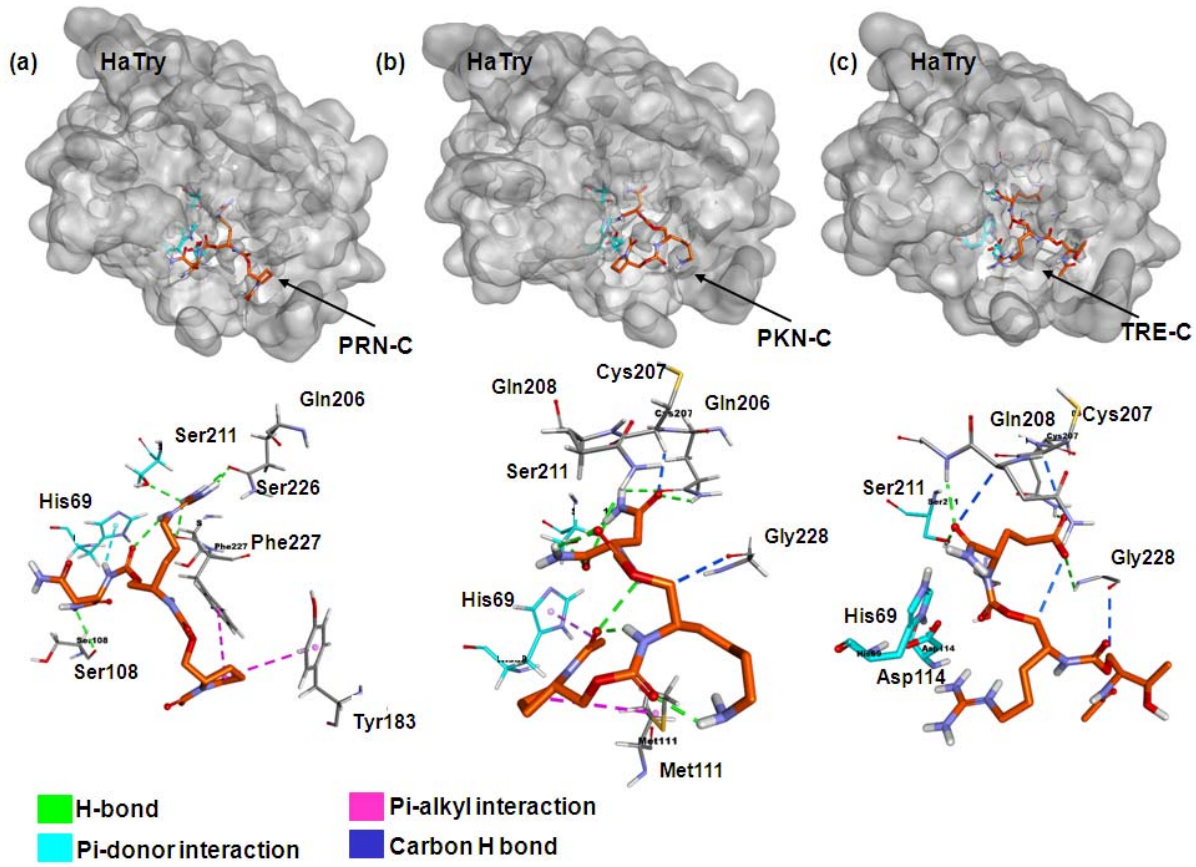


Figure 4.4: Modified peptides in interaction with *H. armigera* Trypsin. Trypsin is shown as surface, with active site represented as cyan sticks. Peptides are shown as orange sticks (a) PRN-carbamate, (b) PKN-carbamate, (c) TRE-carbamate, interactions are shown as dotted lines, color coded according to legend.

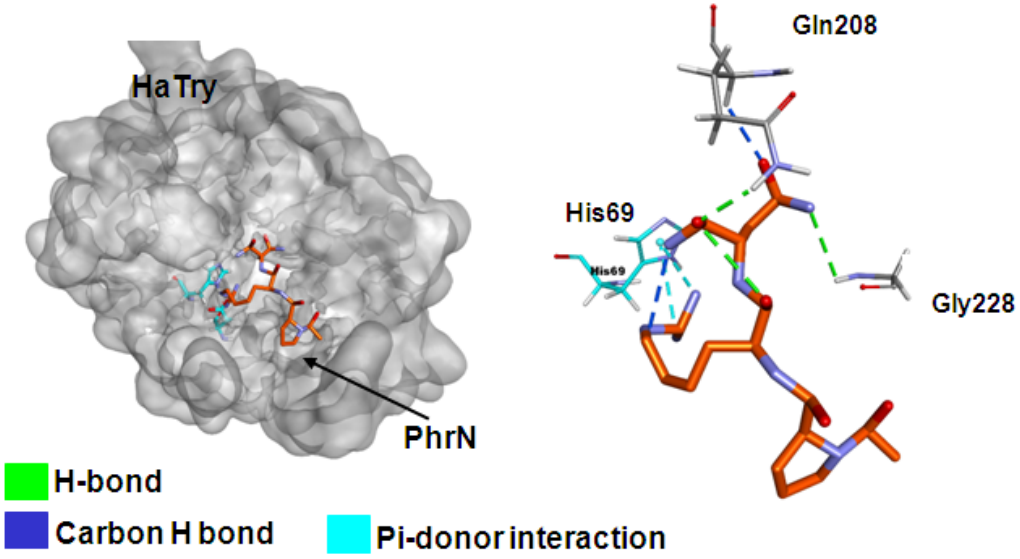


Figure 4.5: PhrN interactionswith *H. armigera* Trypsin. Trypsin is shown as surface, with active site represented as cyan sticks. Peptides are shown as orange sticks. Interactions are shown as dotted lines, color coded according to legend.

4.3.3. Purification of peptides

The crude extract of synthesized peptides were purified to 90% by RP-HPLC. The retention time of PRN-C and PKN-C was similar at 30% ACN and 35% ACN respectively. Whereas, PglN eluted at 10% ACN. This can be attributed to the difference in polarity of these compounds. As carbamate bond is more hydrophobic as compared to amide bond, this results in slower elution on C18 column. The collected HPLC fractions were checked for presence of desired peptide by MALDI-MS. The m/z peaks obtained were compared with expected molecular weights and the aliquots containing the expected mass were selected for further analysis (Figure 4.6, 4.7 and 4.8).

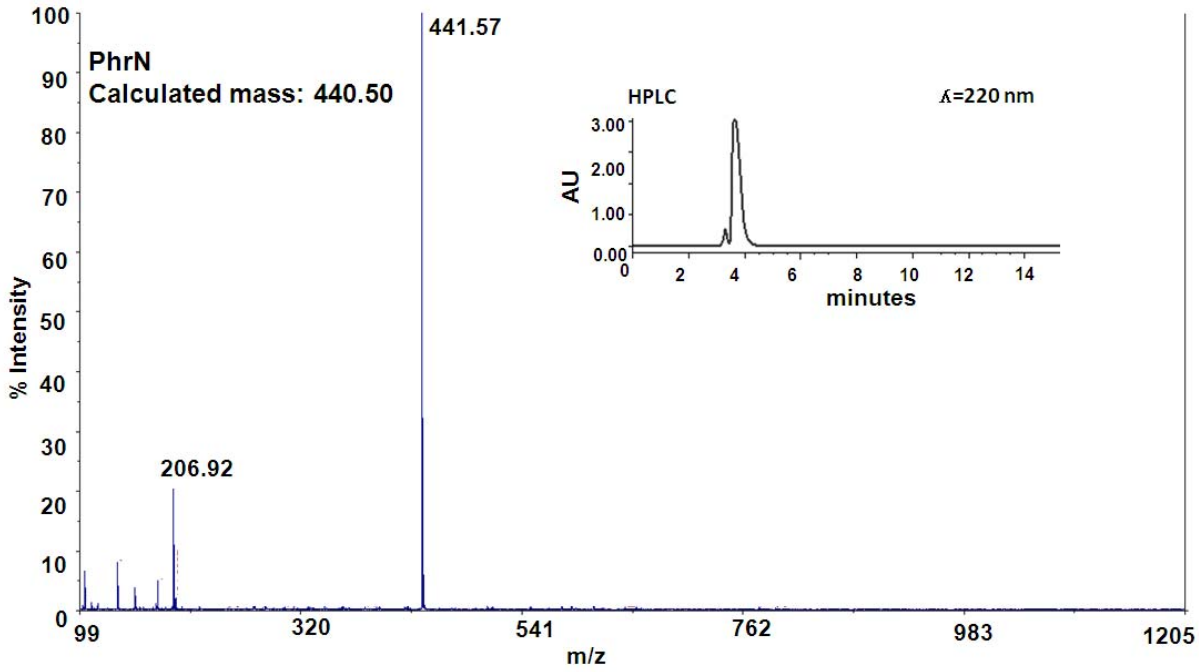


Figure 4.6: MALDI MS spectrum and HPLC chromatogram for PhrN

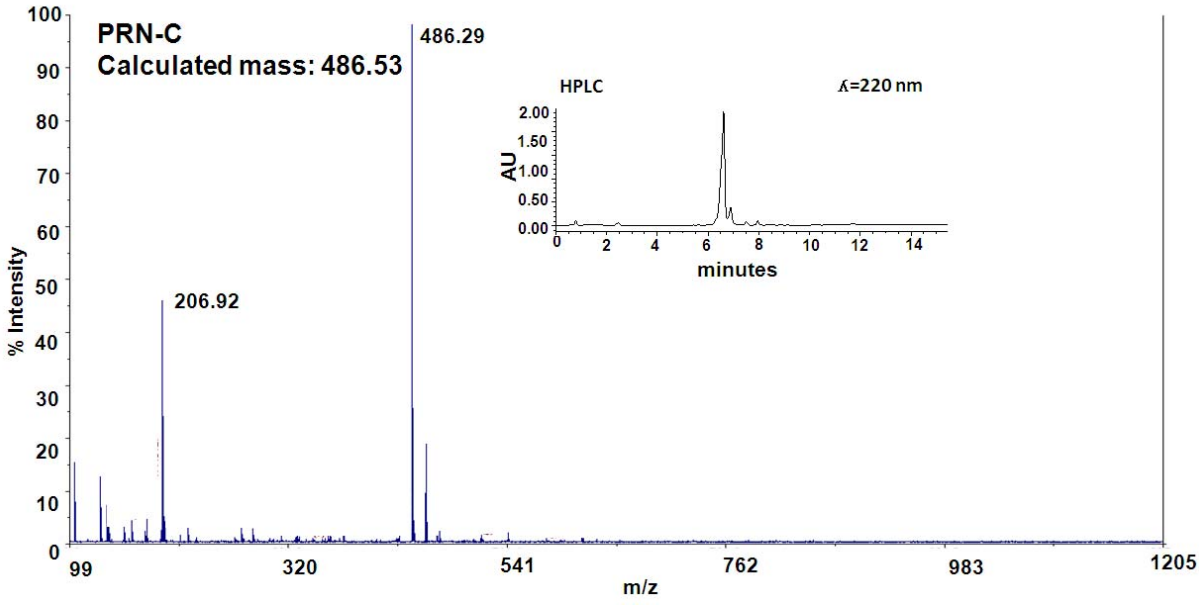


Figure 4.7: MALDI MS spectrum and HPLC chromatogram for PRN-carbamate

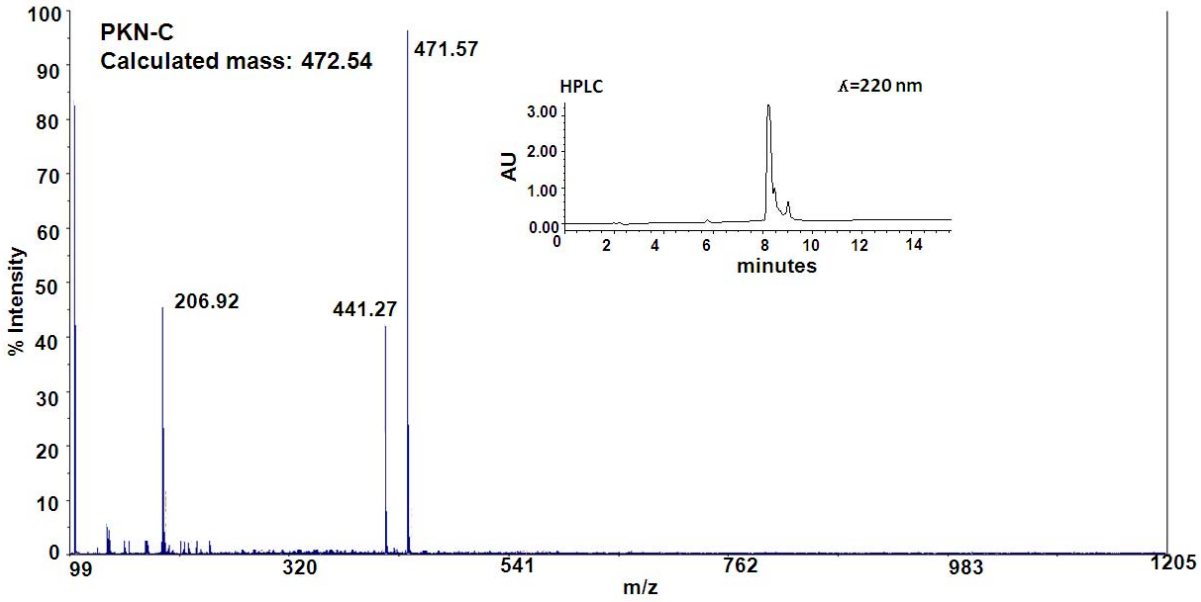


Figure 4.8: MALDI MS spectrum and HPLC chromatogram for PKN-carbamate

4.3.4 Biochemical activity assay

The reactive potential of peptides was exhibited in the protease inhibitory assays, highlighting PRN-C as the most potent inhibitor. The modified peptides showed better reactivity than the native peptides, suggesting that the modifications have been helpful in enhancing the affinity of peptides with proteases (Figure 4.9). Similar to the parent peptides, the mimetic peptides showed differential activity towards insect trypsin-like enzymes and bovine proteases. At 200µM concentration, the peptides were less active against bovine trypsin, showing that they have retained the selectivity of parent peptides. IC_{50} values of peptide binding further explained the differential inhibitory activity of peptides. It was observed that PglN showed 50% inhibition at 105µM, which was lower than the other peptides. Among the carbamate peptides, PKN-C showed $IC_{50}$of 110µM, while PRN-C showed a lower IC_{50} at 90µM (Figure 4.10). Therefore, PRN-C and PglN are selected as the best hits among the modified peptides.

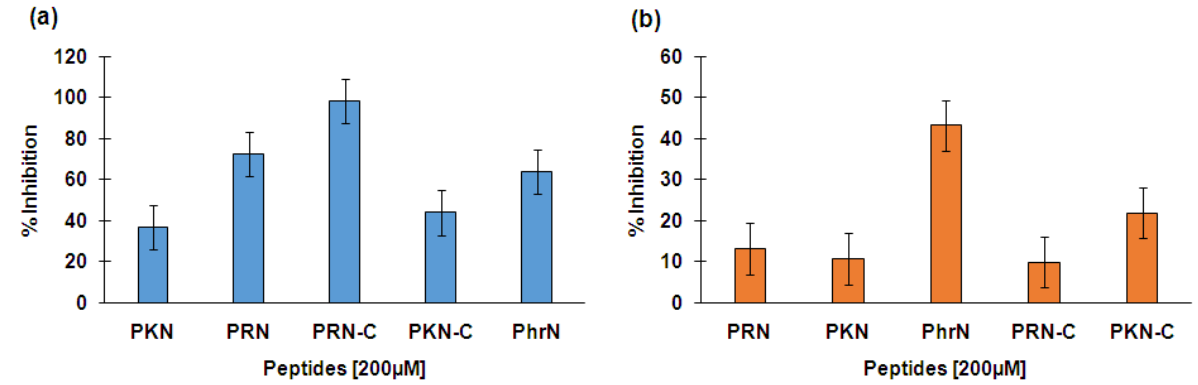


Figure 4.9: Inhibition potential of RCL peptidomimetics. (a) Activity of peptides against *H. armigera* gut proteases and (b) Bovine trypsin. Assay was performed using chromogenic substrate BApNA, at 200µM peptide concentration.

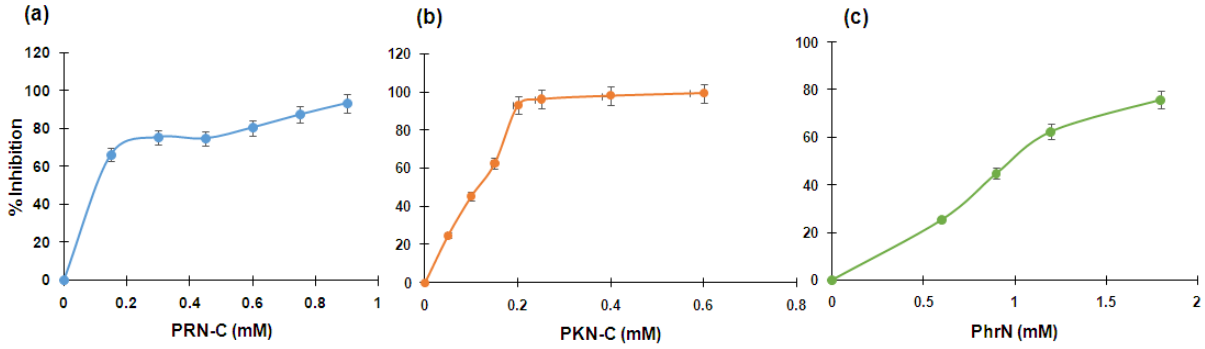


Figure 4.10: Dose dependent inhibition of *H. armigera* gut proteases. Inhibition of trypsin-like proteases was measured by BApNA assay, with increasing peptide concentrations till saturation was achieved.

4.3.5 Surface plasmon resonance studies

The binding kinetics for the peptides PRN-Carbamate and PhrN was studied by surface plasmon resonance, and compared with their unmodified peptide sequence, PRN. Bovine trypsin was immobilized on a CM5 chip as the ligand (Figure 4.11). The peptides were used as analytes from concentrations of 4 to 32µM, with 2 fold escalaltions. Kinetic constants obtained for the peptides are given in Table 4.3. It was found that out of the three peptides, PRN-Carbamate showed slow association with trypsin, as indicated by sensorgram shape. PRN and PhrN showed fast association. Comparison of values for dissociation constant (k_d) showed that PRN-carbamate dissociated slower from the trypsin molecule, followed by PRN and PhrN. Thus, the mimetic

peptide PRN-C shows advantage over PRN in terms of stable binding and low binding constant (KD= 1.2µM).

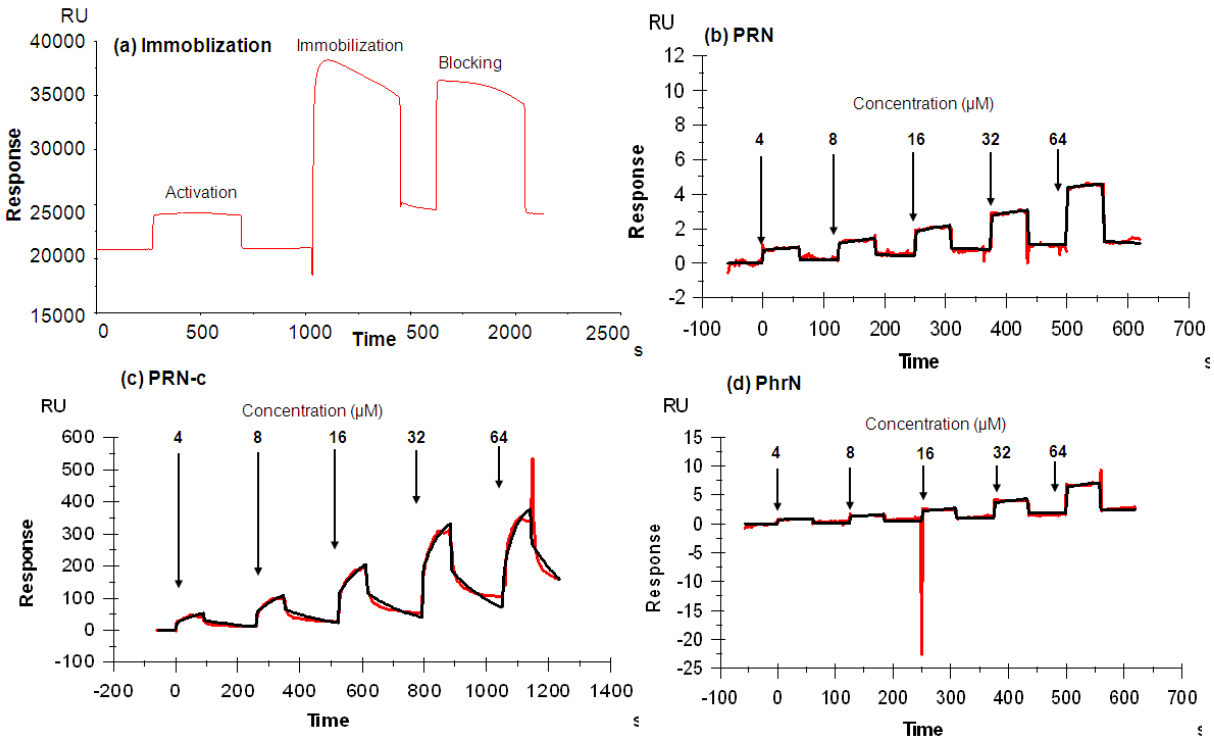


Figure 4.11: Binding study of RCL peptidomimetics by SPR. Bovine trypsin was used as ligand for immobilization (a) Immobilization curve (b) Sensorgram for binding of PRN, (c) PRN-carbamate and (d) PhrN are shown. Red lines represent original sensorgram, and black line shows fitted data.

Table 4.3: Kinetic parameters obtained by SPR

Peptide	ka (1/Ms)	kd (1/s)	KD (M)	Chi² (RU²)	U-value
PRN	316.7	0.005965	1.88E-05	368	5
PRN-C	955.4	0.001205	1.26E-06	0.0449	26
PhrN	3190	0.01828	5.73E-06	2.57	73

4.3.6 Proteolytic stability assay

The stability of peptides in presence of *H. armigera* gut proteases was evaluated by MLDI-TOF based assays. It was found that PRN-carbamate was highly stable to degradation by proteases, possibly because of the presence of carbamate linkage, as reported in earlier studies(Patil et al., 2012). Whereas, PRN and PhrN showed degradation after 6h of treatment with proteases. Hence, carbamate linked peptides could be used as highly stable and active molecules for *H. armigera* control.

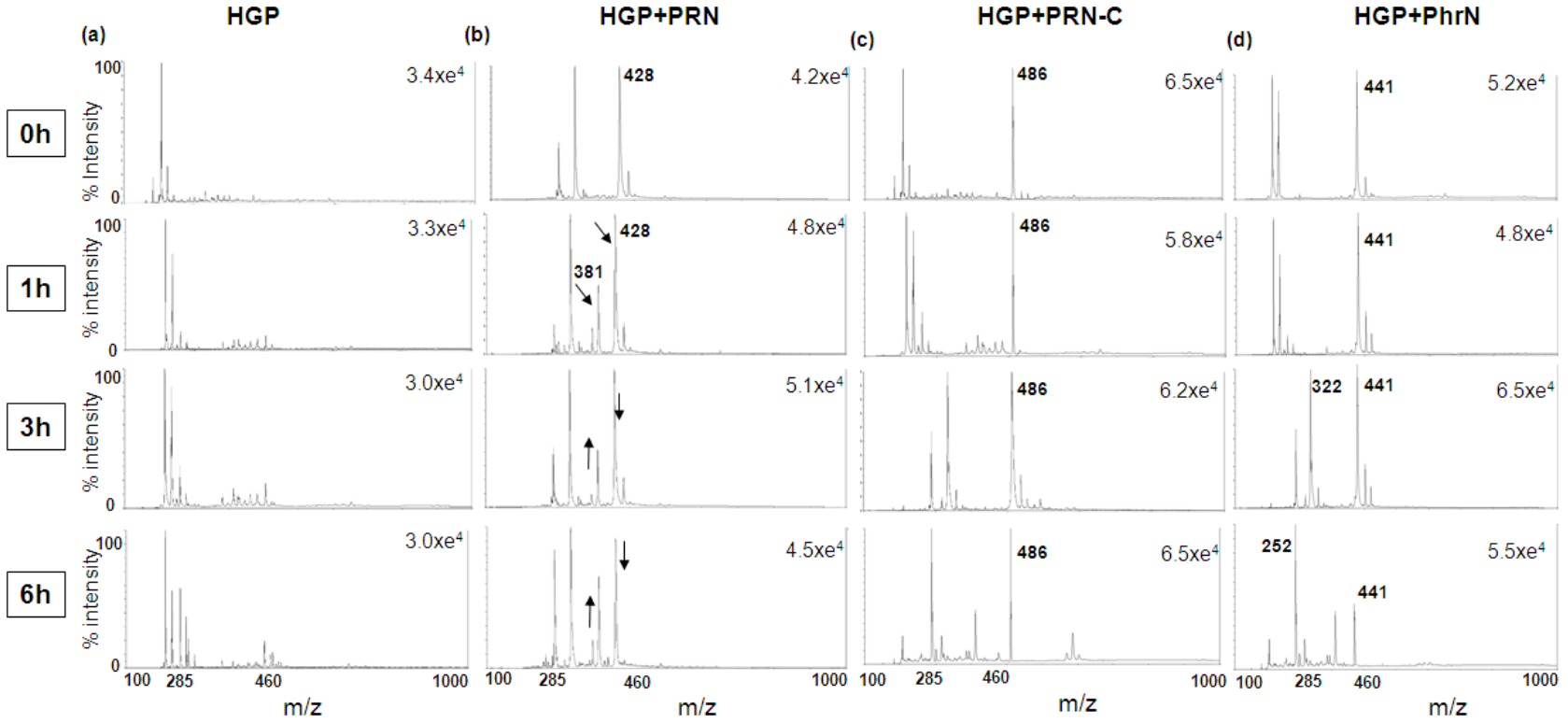


Figure 4.12: Proteolytic stability of peptides. MALDI-TOF spectra of (a) HGP (b) PRN-HGP (c) PRN-Carb-HGP and (d) PhrN-HGP complex at 1, 3 and 6 h after incubation with HGP at 37°C.

4.4 Conclusion

Peptidomimetic modifications are an interesting approach to develop novel molecules, which show enhanced functional attributes, like stability and potency. Also, RCL peptides from Pin-II type inhibitors are excellent lead molecules for generation of pest control agents, owing to their small size and ability to preferentially inhibit proteases. In this study, we have attempted to augment the proteolytic stability and potency of RCL peptides by incorporating modifications in the side chains as well as the peptide backbone. Based on the crystal structures of RCL peptides, we envisaged a deeper penetration of RCL into the active site pocket of trypsin, and thus, increased the side chain length of Arginine, by converting to homoarginine. Also, we replaced the peptide backbone with carbamate bonds, with the aim of improving stability and functional differentiation. Modified RCL peptides were synthesized, and assayed for in vitro inhibition of proteases. Interestingly, they showed higher potency towards insect trypsin like enzymes compared to linear peptides. Also, they were proteolytically stable than their parental peptide, PRN. Thus, peptidomimetic RCL peptides could be used for inhibition of insect serine proteases.

Key highlights:

- RCL peptides were modified by incorporation of homoarginine and replacement of amide backbone by carbamate bond
- In vitro assays indicated higher potency of modified peptides for inhibition of insect proteases
- Binding kinetics by SPR showed affinity with trypsin
- Carbamate peptides are more stable than amide linked peptides in presence of proteases

Chapter 5

Summary and future prospects

“Science of today is the technology of tomorrow”

-Edward Teller

5.1 Summary

Protease inhibitors are of interest as pest control molecules, since they interfere with the digestive proteases of the insects, and lead to reduction in their growth. In contrast to the "wipe out" effect, caused by toxic insecticidal molecules, the protease inhibitors affect the growth and development of insects. Thus, the development of resistance is delayed, making dietary inhibitors more suitable pest control agents from ecological perspective. But, application of these proteins is limited due to lack of environmental stability because of large molecular weight and complexity in formulation due to their labile nature. In this direction, various small molecules and peptides have been used for control of insect pests. In this work, we have tried to develop small peptide-based pest control agents derived from Pin-II type protease inhibitor family from plants.

The Pin-II type inhibitors are wound inducible, multidomain proteins which inactivate serine proteases. This protease-inhibitor interaction is governed by tripeptide reactive center loops (RCL), which fits into the active site of proteases. In native Pin-II inhibitor, this RCL is held by 50 amino acid scaffold, which is stabilized by disulphide bonds. We have identified the RCL of Pin-II type protease inhibitors as potential candidate for development of peptide-based insect control agents. Among the 387 Pin-II inhibitor sequences reported till date, only 21 RCL sequences are present. Further, out of these 21, only 6 sequences are predominant. Thus, we synthesized these 6 RCL tripeptides by solid phase peptide synthesis, and evaluated their activity against serine proteases by in vitro biochemical assays. All the six peptides exhibited inhibition of trypsin and chymotrypsin. However, the RCL peptides were less potent in comparison to their parent IRD. Three peptides with sequences, TRE, PRN and PRY were selected as best inhibitors, with Ki values of ~200μM for trypsin like proteases. Interestingly, the RCL peptides showed higher inhibition of midgut proteases for the lepidopteran pest, *H. armigera* in *in vitro* experiments. Also, the peptides were more potent at pH 10 as compared to pH 7, which indicated that the peptides are specific towards insect serine proteases, and show optimum inhibitory activity at the alkaline physiological pH of insect gut. These are beneficial characteristics for use of RCL peptides as targeted insect control molecules. Therefore, the *in vivo* effect of RCL peptides was evaluated on *H. armigera.* Feeding on tripeptides containing artificial diet showed adverse effects on the growth of *H. armigera* larvae. At 200ppm concentration, PRN exhibited 40% decrease in the larval mass. Also, upon continuous exposure to the peptides, the larvae

showed alteration in the life cycle, as indicated by deformities in pupa and moth formation. In response to the peptide feeding, there was a differential regulation of trypsin and chymotrypsin genes in the insect gut, specifically, there was downregulation of several protease isoforms. Overall, there was a decrease in the residual protease activity post feeding, showing that the peptides are causing inhibition of proteases in the insect gut, ultimately leading to reduction of growth. Also, the peptides TRE and PRN were retained in the insect gut for upto 6h. TRE was stable in the presence of midgut proteases for upto 6h, while PRN and PRY showed degradation after 3h. We analyzed the binding modes of peptides in complex with porcine trypsin by means of X ray crystallography, and found that the peptides bind in the "canonical conformation" of Laskowskiinhibiton. The P1 residue, Arginine was deeply seated into the active site pocket, and showed multiple interactions with Asp179 of the trypsin. Also, the peptide backbone was involved in numerous H-bond and hydrophobic interactions with active site residues of trypsin. The reason for differential affinity of peptides for mammalian and insect proteases was delineated by the P2 residue of the peptide. P2 Thr of TRE was making hydrophobic interaction with Leu89 of porcine trypsin, and simultaneously hydrophilic interaction with the solvent. This led to stabilization of the peptide-protease complex. Whereas, P2 Pro of PRN and PRY did not make hydrophobic contact with Leu89, and hence were less active against porcine trypsin. However, PRN and PRY were more active against insect trypsins, because Leu89 was replaced by Tyr in insect trypsin, allowing the hydrophobic interaction with P2 Pro. Thus, P1 Arg was critical in determining the potency and P2 Pro was important for specificity of the RCL peptides. Hence, our study is the first report establishing that the RCL regions of Pin-II type PIs are able to inhibit serine proteases independent of the native Pin-II scaffold.

Further, with the aim of enhancing potency of the linear RCL peptides, we proposed that the linear peptides could be constrained in a way similar to that in the native Pin-II inhibitor. The cysteine residues flanking the tripeptide RCL sequence were used for this purpose. We performed grafting of two RCL tripeptides on a chemical scaffold (mesitylene) to generate bicyclic peptides. The cysteine containing linear peptides were synthesized and cyclized by means of 1,3,5-Tris(bromomethyl)benzene to yield bicyclic peptides, and with 1,3-Bis(bromomethyl)benzene to yield monocyclic peptides. In silico screening of peptide combinations highlighted CPRNCTREC and CPRNCPRYC bicyclic peptides as best hits. Screening for affinity with bovine trypsin by surface plasmon resonance gave the same peptides

as most potential hits, which showed rapid association and slow dissociation with trypsin. Further, a library of 18 peptides, comprising of bicyclic, monocyclic and linear peptides was tested for *in vitro* inhibition of insect trypsin like enzymes. The bicyclic peptides showed tenfold enhanced inhibition of insect trypsin like proteases as compared to linear RCL tripeptides, with IC_{50} of ~20μM. Thereafter, the peptides were evaluated for feeding preference and antibiosis against two lepidopteran pests, *H. armigera* and *S. litura*. Interestingly, upon cyclization, the peptides showed insect deterrent activity, wherein the larvae did not prefer to feed on peptide treated leaves. Also, upon obligatory feeding, there was a severe growth reduction of the larvae, with ~60% reduction at 100ppm peptide concentration. Target identification by pull down assays showed that in addition to protease inhibition, these peptides interfere with the antioxidant enzymes, which was also reflected by decrease in residual protease and antioxidant enzyme activities. The prediction of binding modes by molecular dynamics simulation studies showed that the bicyclic peptides cover a larger interfacial surface area on trypsin, and hence simultaneously bind to the active site as well as the substrate selective pocket of trypsin. Furthermore, the cyclic peptides are stabilized by a network of intramolecular H-bonds, which might also be a reason for enhanced potency.

Thereafter, peptidomimetics were designed based on RCL regions. Since in parent RCL peptide, the P1 arginine is the major binding partner, we hypothesized enhanced binding interactions by addition of homoarginine (elongated side chain of arginine by $–CH_2$) in place of P1 Arg. Also, replacement of peptide bonds with carbamate linkages was performed envisaging higher proteolytic stability over parent peptides and functional differentiation as carbamate linkages have been used in pesticides. The modified RCL tripeptides were designed in and in silico analyzed for their affinity with bovine and insect proteases. The mimetic peptides showed slightly higher binding affinity towards proteases than their parent peptides. Then, they were synthesized by solid phase synthesis, and assayed for *in vitro* inhibition of bovine and insect proteases. The RCL mimetic peptides were twofold more potent compared to linear peptides, and retained the specificity towards insect proteases. Also, biophysical characterization by surface plasmon resonance studies showed that the peptides bind to trypsins with high association rate. Remarkably, the carbamate peptides were stable to proteolysis for more than 6h, compared to their linear peptide, PRN.

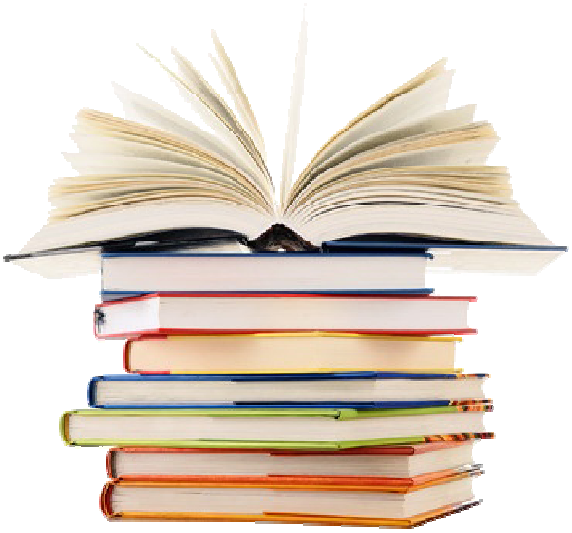
Hence, this work shows that an arena of peptide combinations can be designed for coping with diverse insect proteases, making RCL peptides as potential pest control agents. Our results highlight the importance of targeted reactive loop engineering, as opposed to random screening of peptides. Being small peptides, there is an ease of formulation and application by targeted delivery systems. Our work will be of interest to develop next generation of bio-inspired peptides for agriculture or medicinal applications.

5.2 Future prospects

Tripeptide reactive loop regions are promising candidates for engineering of peptide based pest control molecules. The work presented in this thesis opens up new avenues in this direction:

- Cyclic peptide combinations could be generated for all 21 RCL regions, making a library of varied molecules.
- Field trials of peptides and its study for successive generations will enable commercialization of these peptides.
- Formulations and large scale synthesis of peptides can be optimized.
- Mimetic peptides could be studied for in vivo effects on insects.
- The incorporation of toxic amino acids or arginine analogs in the RCL region can help in enhancing the potency.
- Off-target effects of mimetic peptides can be studied.
- The expression of RCL peptides in tandem can be done to generate a plethora of RCLs, and their effect on the insect digestive physiology can be studied.

Bibliography



Aartsma, Y. et al., 2017. Herbivore-induced plant volatiles and tritrophic interactions across spatial scales. *New Phytologist*, 216(4), pp.1054–1063.

Adams, P.D. et al., 2010. *PHENIX*□: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D Biological Crystallography*, 66(2), pp.213–221.

Akanksha, 2018. Avoidable crop losses, Economic Threshold Level and management of *Helicoverpa armigera* (Hübner) in chickpea. (Doctoral dissertation, Punjab Agricultural University, Ludhiana)

Alfonso, J. et al., 1997. Wheat and Barley Inhibitors Active Towards -Amylase and Trypsin-like Activities from *Spodoptera frugiperda*. *Journal of Chemical Ecology*, 23(7), pp.1729–1741.

Altpeter, F. et al., 1999. Increased insect resistance in transgenic wheat stably expressing trypsin inhibitor CMe. *Molecular Breeding*, 5(1), pp.53–63.

Ambadapadi, S. et al., 2016. Reactive Center Loop (RCL) peptides derived from serpins display independent coagulation and immune modulating activities. *Journal of Biological Chemistry*, 291(6), pp.2874–2887.

Angelini, A. et al., 2012. Bicyclic peptide inhibitor reveals large contact interface with a protease target. *ACS Chemical Biology*, 7(5), pp.817–821.

Anthis, N.J. & Clore, G.M., 2013. Sequence-specific determination of protein and peptide concentrations by absorbance at 205 nm. *Protein Science*, 22(6), pp.851–858.

Avan, I., Dennis Hall, C. & Katritzky, A.R., 2014. Peptidomimetics via modifications of amino acids and peptide bonds. *Chemical Society Reviews*, 43(10), pp.3575–3594.

Babu, S.R. & Subrahmanyam, B., 2010. Bio-potency of serine proteinase inhibitors from Acacia senegal seeds on digestive proteinases, larval growth and development of *Helicoverpa armigera* (Hübner). *Pesticide Biochemistry and Physiology*, 98(3), pp.349–358.

Barbeta, B.L. et al., 2008. Plant cyclotides disrupt epithelial cells in the midgut of lepidopteran larvae. *Proceedings of the National Academy of Sciences*, 105(4), pp.1221–1225.

Barr, D.B. & Buckley, B., 2011. In vivo biomarkers and biomonitoring in reproductive and developmental toxicity. *Reproductive and Developmental Toxicology*, pp.253–265.

Barrette-Ng, I.H. et al., 2003a. Structural basis of inhibition revealed by a 1:2 complex of the two-headed tomato inhibitor-II and subtilisin Carlsberg. *The Journal of biological chemistry*, 278(26), pp.24062–71.

Barrette-Ng, I.H. 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(33), pp.31391–31400.

Barton, K.A., Whiteley, H.R. & Yang, N.S., 1987. Bacillus thuringiensis section sign-Endotoxin Expressed in Transgenic *Nicotiana tabacum* Provides Resistance to Lepidopteran Insects. *Plant physiology*, 85(4), pp.1103–9.

Becerra, J.X., 2015. Macroevolutionary and geographical intensification of chemical defense in plants driven by insect herbivore selection pressure. *Current Opinion in Insect Science*, 8, pp.15–21.

Beck, J.G. et al., 2012. Intestinal permeability of cyclic peptides: common key backbone motifs identified. *Journal of the American Chemical Society*, 134(29), pp.12125–33.

Beekwilder, J. et al., 2000. Characterization of potato proteinase inhibitor II reactive site mutants.

European Journal of Biochemistry, 267(7), pp.1975–1984.

Birk, Y., 1985. The Bowman-Birk inhibitor. Trypsin- and chymotrypsin-inhibitor from soybeans. *International journal of peptide and protein research*, 25(2), pp.113–31.

Blackwell, H.E. et al., 2001. Ring-Closing Metathesis of Olefinic Peptides: Design, Synthesis, and Structural Characterization of Macrocyclic Helical Peptides. *The Journal of Organic Chemistry*, 66(16), pp.5291–5302.

Bode, W. & Huber, R., 2000. Structural basis of the endoproteinase-protein inhibitor interaction. *Biochimica et biophysica acta*, 1477(1–2), pp.241–52.

Boeglin, D. & Lubell, W.D., 2005. Aza-Amino Acid Scanning of Secondary Structure Suited for Solid-Phase Peptide Synthesis with Fmoc Chemistry and Aza-Amino Acids with Heteroatomic Side Chains. *Journal of Combinatorial Chemistry*, 7(6), pp.864–878.

Bown, A.W., Hall, D.E. & MacGregor, K.B., 2002. Insect footsteps on leaves stimulate the accumulation of 4-aminobutyrate and can be visualized through increased chlorophyll fluorescence and superoxide production. *Plant physiology*, 129(4), pp.1430–4.

Bown, D.P., Wilkinson, H.S. & Gatehouse, J.A., 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(7), pp.625–638.

Broadway, R.M., 1996. Dietary proteinase inhibitors alter complement of midgut proteases. *Archives of Insect Biochemistry and Physiology*, 32(1), pp.39–53.

Bronstein, J.L., Alarcón, R. & Geber, M., 2006. The evolution of plant insect mutualisms. *New Phytologist*, 172(3), pp.412–428.

Bruessow, F. et al., 2010. Insect eggs suppress plant defence against chewing herbivores. *The Plant Journal*, 62(5), pp.876–885.

Burgess, E.P.J. et al., 2002. Avidin Expressed in Transgenic Tobacco Leaves Confers Resistance to Two Noctuid Pests, *Helicoverpa Armigera* and *Spodoptera Litura*. *Transgenic Research*, 11(2), pp.185–198.

Capinera, J.L., 2008. *Encyclopedia of entomology* 2nd ed., New York: Springer.

Chan, L.Y., Craik, D.J. & Daly, N.L., 2016. Dual-targeting anti-angiogenic cyclic peptides as potential drug leads for cancer therapy. *Scientific Reports*, 6(1), p.35347.

Chekan, J.R. et al., 2017. Characterization of the macrocyclase involved in the biosynthesis of RiPP cyclic peptides in plants. *Proceedings of the National Academy of Sciences of the United States of America*, 114(25), pp.6551–6556.

Chen, J. et al., 2015. A comparative study of trypsin specificity based on QM/MM molecular dynamics simulation and QM/MM GBSA calculation. *Journal of Biomolecular Structure and Dynamics*, 33(12), pp.2606–2618.

Chen, P. et al., 1992. Reactive sites of an anticarcinogenic Bowman-Birk proteinase inhibitor are similar to other trypsin inhibitors. *The Journal of biological chemistry*, 267(3), pp.1990–4.

Chen, S. et al., 2014. Peptide ligands stabilized by small molecules. *Angewandte Chemie - International Edition*, 53(6), pp.1602–1606.

Chen, S. et al., 2012. Structurally Diverse Cyclisation Linkers Impose Different Backbone Conformations

in Bicyclic Peptides. *ChemBioChem*, 13(7), pp.1032–1038.

Chen, V.B. et al., 2010. *MolProbity*: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D Biological Crystallography*, 66(1), pp.12–21.

Chikate, Y.R. et al., 2013. Differential protease activity augments polyphagy in *Helicoverpa armigera*. *Insect Molecular Biology*, 22(3), pp.258–272.

Christian E. Schafmeister, Julia Po, and & Verdine, G.L., 2000. An All-Hydrocarbon Cross-Linking System for Enhancing the Helicity and Metabolic Stability of Peptides. *Journal of the American Chemical Society*, 122(24), pp.5891-5892

Christmann, A. et al., 1999. The cystine knot of a squash-type protease inhibitor as a structural scaffold for *Escherichia coli* cell surface display of conformationally constrained peptides. *Protein engineering*, 12(9), pp.797–806.

Chua, K.C.H. et al., 2014. Macrocyclic protease inhibitors with reduced peptide character. *Angewandte Chemie - International Edition*, 53(30), pp.7828–7831.

Chung, B. & Yudin, A.K., 2015. Disulfide-Bridged Peptide Macrobicycles from Nature. *Org. Biomol. Chem.*13(33), pp.8768-8779.

Clavijo McCormick, A., Unsicker, S.B. & Gershenzon, J., 2012. The specificity of herbivore-induced plant volatiles in attracting herbivore enemies. *Trends in plant science*, 17(5), pp.303–10.

Cleveland, T.E., Thornburg, R.W. & Ryan, C.A., 1987. Molecular characterization of a wound-inducible inhibitor I gene from potato and the processing of its mRNA and protein. *Plant Molecular Biology*, 8(3), pp.199–207.

Corbin, D.R. et al., 2001. Expression and chloroplast targeting of cholesterol oxidase in transgenic tobacco plants. *Plant physiology*, 126(3), pp.1116–28.

Craik, D.J. et al., 2004. Discovery, structure and biological activities of the cyclotides. *Current protein & peptide science*, 5(5), pp.297–315.

Craik, D.J. et al., 1999. Plant cyclotides: A unique family of cyclic and knotted proteins that defines the cyclic cystine knot structural motif. *Journal of Molecular Biology*, 294(5), pp.1327–1336.

Craik, D.J. et al., 2018. Ribosomally-synthesised cyclic peptides from plants as drug leads and pharmaceutical scaffolds. *Bioorganic & Medicinal Chemistry*, 26(10), pp.2727–2737.

Craik, D.J. et al., 2013. The Future of Peptide-based Drugs. *Chemical Biology & Drug Design*, 81(1), pp.136–147.

Czapinska, H. & Otlewski, J., 1999. Structural and energetic determinants of the S1-site specificity in serine proteases. *European Journal of Biochemistry*, 260(3), pp.571–595.

D'Souza, C. et al., 2016. Using the MCoTI-II Cyclotide Scaffold To Design a Stable Cyclic Peptide Antagonist of SET, a Protein Overexpressed in Human Cancer. *Biochemistry*, 55(2), pp.396–405.

Dawkar, V. V. et al., 2013. Molecular insights into resistance mechanisms of lepidopteran insect pests against toxicants. *Journal of Proteome Research*, 12(11), pp.4727–4737.

Després, L., David, J.-P. & Gallet, C., 2007. The evolutionary ecology of insect resistance to plant chemicals. *Trends in Ecology & Evolution*, 22(6), pp.298–307.

Dias, E. et al., 2015. A Review on the Assessment of the Potential Adverse Health Impacts of Carbamate

Pesticides. In *Topics in Public Health*. InTech.

Dovrat, D. & Aharoni, A., 2016. Evolved to overcome Bt-toxin resistance. *Nature*, 533(7601), pp.39–40.

Downes, S., Mahon, R. & Olsen, K., 2007. Monitoring and adaptive resistance management in Australia for Bt-cotton: Current status and future challenges. *Journal of Invertebrate Pathology*, 95(3), pp.208–213.

Dunaevsky, Y.E. et al., 2005. Protease Inhibitors in Improvement of Plant Resistance to Pathogens and Insects. *Molecular Biology*, 39(4), pp.608–613.

Dunse, K.M., Stevens, J.A., et al., 2010. Coexpression 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 of the United States of America*, 107(34), pp.15011–5.

Dunse, K.M., Kaas, Q., et al., 2010. Molecular basis for the resistance of an insect chymotrypsin to a potato type II proteinase inhibitor. *Proceedings of the National Academy of Sciences of the United States of America*, 107(34), pp.15016–21.

Elliott, A.G. et al., 2014. Evolutionary origins of a bioactive peptide buried within Preproalbumin. *The Plant cell*, 26(3), pp.981–95.

Evans, P., 2006. Scaling and assessment of data quality. *Acta Cryst*, 62, pp.72–82.

Falco, M.C. & Silva-Filho, M.C., 2003. Expression of soybean proteinase inhibitors in transgenic sugarcane plants: effects on natural defense against *Diatraea saccharalis*. *Plant Physiology and Biochemistry*, 41(8), pp.761–766.

Fan, X. et al., 1999. Insecticidal activity of transgenic tobacco plants expressing both Bt and CpTI genes on cotton bollworm (*Helicoverpa armigera*). *Chinese journal of biotechnology*, 15(1), pp.1–5.

Farrar, R.R., Barbour, J.D. & Kennedy, G.G., 1989. Quantifying Food Consumption and Growth in Insects. *Annals of the Entomological Society of America*, 82(5), pp.593–598.

Fernandes, G.W., 1994. Plant mechanical defenses against insect herbivory. *Revista Brasileira de Entomologia*, 38(2), pp.421–433.

Fitches, E. et al., 2004. Cloning, expression and functional characterisation of chitinase from larvae of tomato moth (*Lacanobia oleracea*): a demonstration of the insecticidal activity of insect chitinase. *Insect Biochemistry and Molecular Biology*, 34(10), pp.1037–1050.

Fletcher, J.I. et al., 1997. The structure of a novel insecticidal neurotoxin, omega-atracotoxin-HV1, from the venom of an Australian funnel web spider. *Nature structural biology*, 4(7), pp.559–66.

Fluhr, R., Lampl, N. & Roberts, T.H., 2012. Serpin protease inhibitors in plant biology. *Physiologia Plantarum*, 145(1), pp.95–102.

Franke, B., Mylne, J.S. & Rosengren, K.J., 2018. Buried treasure: biosynthesis, structures and applications of cyclic peptides hidden in seed storage albumins. *Natural Product Reports*, 35(2), pp.137–146.

Gadge, P.P. et al., 2015. A bifunctional α-amylase/trypsin inhibitor from pigeonpea seeds: Purification, biochemical characterization and its bio-efficacy against *Helicoverpa armigera.Pesticide Biochemistry and Physiology*, 125, pp.17–25.

Garcia, V.A. et al., 2004. Trypsin inhibitor from *Poecilanthe parviflora* seeds: purification, characterization, and activity against pest proteases. *The protein journal*, 23(5), pp.343–50.

Ghosh, A.K. & Brindisi, M., 2015. Organic Carbamates in Drug Design and Medicinal Chemistry. *Journal of Medicinal Chemistry*, 58(7), pp.2895–2940.

Grosse-Holz, F.M. & van der Hoorn, R.A.L., 2016. Juggling jobs: Roles and mechanisms of multifunctional protease inhibitors in plants. *New Phytologist*, 210(3), pp.794–807.

Guo, Y.X. et al., 2005. Isarfelin, a peptide with antifungal and insecticidal activities from *Isaria felina*. *Peptides*, 26(12), pp.2384–2391.

Harsulkar, A.M. et al., 1999. Successive Use of Non-Host Plant Proteinase Inhibitors Required for Effective Inhibition of *Helicoverpa armigera* Gut Proteinases and Larval Growth1. *Plant physiology*, 121(2), pp.497–506.

Hartl, M. et al., 2010. Serine protease inhibitors specifically defend *Solanum nigrum* against generalist herbivores but do not influence plant growth and development. *The Plant cell*, 22(12), pp.4158–75.

Heitz, A. et al., 2001. Solution structure of the squash trypsin inhibitor MCoTI-II. A new family for cyclic knottins. *Biochemistry*, 40(27), pp.7973–83.

Hernandez, J.F. et al., 2000. Squash trypsin inhibitors from *Momordica cochinchinensis* exhibit an atypical macrocyclic structure. *Biochemistry*, 39(19), pp.5722–30.

Herzig, V. et al., 2011. ArachnoServer 2.0, an updated online resource for spider toxin sequences and structures. *Nucleic Acids Research*, 39(Database), pp.D653–D657.

Hill, T.A. et al., 2014. Constraining cyclic peptides to mimic protein structure motifs. *Angewandte Chemie - International Edition*, 53(48), pp.13020–13041.

Hodge, S. et al., 2006. The effect of b-aminobutyric acid on the growth of herbivorous insects feeding on Brassicaceae. *Annals of Applied Biology*, 148(3), pp.223-229.

Huang, T., Jander, G. & De Vos, M., 2011. Non-protein amino acids in plant defense against insect herbivores: Representative cases and opportunities for further functional analysis. *Phytochemistry*, 72(13), pp.1531–1537.

Huot, B. et al., 2014. Growth–Defense Tradeoffs in Plants: A Balancing Act to Optimize Fitness. *Molecular Plant*, 7(8), pp.1267–1287.

Hussain, D. et al., 2015. Insecticide Resistance in Field Populations of *Helicoverpa armigera* (Hübner) (Lepidoptera: Noctuidae). *Journal of Entomological Science*, 50(2), pp.119–128.

International Program on Chemical Safety., 1986. *Carbamate pesticides□: a general introduction.*, World Health Organization.

Jackson, D.Y. et al., 1991. General approach to the synthesis of short .alpha.-helical peptides. *Journal of the American Chemical Society*, 113(24), pp.9391–9392.

Jamal, F. et al., 2013. Serine protease inhibitors in plants: Nature's arsenal crafted for insect predators. *Phytochemistry Reviews*, 12(1), pp.1–34.

Janz, N., Nylin, S. & Wahlberg, N., 2006. Diversity begets diversity: host expansions and the diversification of plant-feeding insects. *BMC Evolutionary Biology*, 6(1), p.4.

Jersáková, J., Johnson, S.D. & Kindlmann, P., 2006. Mechanisms and evolution of deceptive pollination in orchids. *Biological Reviews*, 81(02), p.219.

Ji, Y. et al., 2013. In Vivo Activation of the p53 Tumor Suppressor Pathway by an Engineered Cyclotide.

Journal of the American Chemical Society, 135(31), pp.11623–11633.

Jongsma, M.A. et al., 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 of the United States of America*, 92(17), pp.8041–5.

Jongsma, M.A. & Beekwilder, J., 2011. Co-evolution of insect proteases and plant protease inhibitors. *Current protein & peptide science*, 12(5), pp.437–447.

Jongsma, M.A. & Bolter, C., 1997. The adaptation of insects to plant protease inhibitors. *Journal of Insect Physiology*, 43(10), pp.885–895.

Joshi, R.S. et al., 2013. Complementation of intramolecular interactions for structural–functional stability of plant serine proteinase inhibitors. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 1830(11), pp.5087–5094.

Joshi, R.S. et al., 2014. 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. *Journal of biomolecular structure & dynamics*, 32(1), pp.13–26.

Joshi, R.S., Gupta, V.S. & Giri, A.P., 2014. Differential antibiosis against *Helicoverpa armigera* exerted by distinct inhibitory repeat domains of *Capsicum annuum* proteinase inhibitors. *Phytochemistry*, 101, pp.16–22.

Jouvensal, L. et al., 2003. PA1b, an Insecticidal Protein Extracted from Pea Seeds (*Pisum sativum*): ^{1}H-2-D NMR Study and Molecular Modeling. *Biochemistry*, 42(41), pp.11915–11923.

Kabsch, W., K., W. & G., R.R.B., 2010. *XDS*. *Acta Crystallographica Section D Biological Crystallography*, 66(2), pp.125–132.

Kaneda, T. et al., 2016. Cyclolinopeptides, cyclic peptides from flaxseed with osteoclast differentiation inhibitory activity. *Bioorganic & Medicinal Chemistry Letters*, 26(7), pp.1760–1761.

Kasar, S.S. et al., 2017. A glycoprotein α-amylase inhibitor from *Withania somnifera* differentially inhibits various α-amylases and affects the growth and development of *Tribolium castaneum*. *Pest management science*, 73(7), pp.1382–1390.

Kaufmann, H.P. & Tobschirbel, A., 1959. Über ein Oligopeptid aus Leinsamen. *Chemische Berichte*, 92(11), pp.2805–2809.

King, G.F. & Hardy, M.C., 2013. Spider-Venom Peptides: Structure, Pharmacology, and Potential for Control of Insect Pests. *Annual Review of Entomology*, 58(1), pp.475–496.

Kinnersley, A.M. & Turano, F.J., 2000. Gamma Aminobutyric Acid (GABA) and Plant Responses to Stress. *Critical Reviews in Plant Sciences*, 19(6), pp.479–509.

Koch, F., 2008. Charles D. Michener (2000). The Bees of the World. - Baltimore and London (The John Hopkins University Press). - 913 S., 48 Farbfotos und zahlr. s/w Illustrationen. ISBN 0-8018-6133-0 (hardcover). US$ 165.-. *Zoosystematics and Evolution*, 78(2), pp.353–353.

Kolmar, H., 2008. Alternative binding proteins: Biological activity and therapeutic potential of cystine-knot miniproteins. *FEBS Journal*, 275(11), pp.2684–2690.

Kotkar, H.M. et al., 2012. Amylase gene expression patterns in *Helicoverpa armigera* upon feeding on a range of host plants. *Gene*, 501(1), pp.1–7.

Kulkarni, S.A. et al., 2008. Comparison of *Metarhizium* isolates for biocontrol of *Helicoverpa*

armigera(Lepidoptera: Noctuidae) in chickpea. *Biocontrol Science and Technology*, 18(8), pp.809–828.

Kumar, A. et al., 2016. Protein crystallography beamline (PX-BL21) at Indus-2 synchrotron. *Journal of Synchrotron Radiation*, 23(2), pp.629–634.

Kuwar, S.S. et al., 2015. Adaptive regulation of digestive serine proteases in the larval midgut of *Helicoverpa armigera* in response to a plant protease inhibitor. *Insect Biochemistry and Molecular Biology*, 59, pp.18–29.

Lacey, L.A. et al., 2015. Insect pathogens as biological control agents: Back to the future. *Journal of Invertebrate Pathology*, 132, pp.1–41.

van de Langemheen, H. et al., 2017. Polar Hinges as Functionalized Conformational Constraints in (Bi)cyclic Peptides. *ChemBioChem*, 18(4), pp.387–395.

Langenfeld, A. et al., 2011. Insecticidal Cyclodepsipeptides from *Beauveria felina*. *Journal of Natural Products*, 74(4), pp.825–830.

Laskowski, M. & Kato, I., 1980. Protein Inhibitors of Proteinases. *Annual Review of Biochemistry*, 49(1), pp.593–626.

Łegowska, A. et al., 2009. Introduction of non-natural amino acid residues into the substrate-specific P1 position of trypsin inhibitor SFTI-1 yields potent chymotrypsin and cathepsin G inhibitors. *Bioorganic and Medicinal Chemistry*, 17(9), pp.3302–3307.

Leo, F.D. et al., 2002. PLANT-PIs: a database for plant protease inhibitors and their genes. *Nucleic Acids Research*, 30(1), pp.347–348.

De Leo F, F. De et al., 1998. Opposite effects on *spodoptera littoralis* larvae of high expression level of a trypsin proteinase inhibitor in transgenic plants. *Plant physiology*, 118(3), pp.997–1004.

Li, X.-Q., Zhang, T. & Donnelly, D., 2011. Selective Loss of Cysteine Residues and Disulphide Bonds in a Potato Proteinase Inhibitor II Family D. Liberles, ed. *PLoS ONE*, 6(4), p.e18615.

Liu, W. et al., 2017. Precisely Regulated and Efficient Locking of Linear Peptides into Stable Multicyclic Topologies through a One-Pot Reaction. *Angewandte Chemie - International Edition*, 56(16), pp.4458–4463.

Livak, K.J. & Schmittgen, T.D., 2001. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. *Methods*, 25(4), pp.402–408.

Lomate, P.R. et al., 2018. Integrated Transcriptomic and Proteomic Analyses Suggest the Participation of Endogenous Protease Inhibitors in the Regulation of Protease Gene Expression in *Helicoverpa armigera*. *Molecular & cellular proteomics : MCP*, 17(7), pp.1324–1336.

Luo, M. et al., 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(4), pp.1896–1910.

MacGregor, K.B. et al., 2003. Overexpression of glutamate decarboxylase in transgenic tobacco plants deters feeding by phytophagous insect larvae. *Journal of chemical ecology*, 29(9), pp.2177–82.

Mahajan, N.S. et al., 2013. Plasticity of protease gene expression in *Helicoverpa armigera* upon exposure to multi-domain *Capsicum annuum* protease inhibitor. *Biochimica et Biophysica Acta - General Subjects*, 1830(6), pp.3414–3420.

Mahajan, N.S. et al., 2015. Structural features of diverse Pin-II proteinase inhibitor genes from C*apsicum annuum*. *Planta*, 241(2), pp.319–331.

Mahatmanto, T., 2015. Review seed biopharmaceutical cyclic peptides: From discovery to applications. *Biopolymers*, 104(6), pp.804–814.

Mamta, B. & Rajam, M. V., 2017. RNAi technology: a new platform for crop pest control. *Physiology and Molecular Biology of Plants*, 23(3), pp.487–501.

Matthews, J.M. et al., 1990. Cross-Resistance to Herbicides in Annual Ryegrass (*Lolium rigidum*): I. Properties of the Herbicide Target Enzymes Acetyl-Coenzyme A Carboxylase and Acetolactate Synthase. *PLANT PHYSIOLOGY*, 94(3), pp.1180–1186.

McBride, J.D. et al., 1996. Selection of chymotrypsin inhibitors from a conformationally-constrained combinatorial peptide library. *Journal of molecular biology*, 259(4), pp.819–827.

Melangeli, C., Rosenthal, G.A. & Dalman, D.L., 1997. The biochemical basis for L-canavanine tolerance by the tobacco budworm *Heliothis virescens* (Noctuidae). *Proceedings of the National Academy of Sciences of the United States of America*, 94(6), pp.2255–60.

Mensah, R.K. et al., 2007. International Journal of Pest Management Improving the efficacy of nuclear polyhedrosis virus and *Bacillus thuringiensis* against *Helicoverpa* spp. with ultra-violet light protected petroleum spray oils on cotton crops in Australia. *International Journal of Pest Management*, 51(2), pp.101–109.

Mishra, M. et al., 2015. Ecological turmoil in evolutionary dynamics of plant–insect interactions: defense to offence. *Planta*, 242(4), pp.761–771.

Mishra, M. et al., 2010. Interaction of recombinant CanPIs with *Helicoverpa armigera* gut proteases reveals their processing patterns, stability and efficiency. *Proteomics*, 10(15), pp.2845–2857.

Mishra, M. et al., 2013. Structural-functional insights of single and multi-domain *Capsicum annuum* protease inhibitors. *Biochemical and Biophysical Research Communications*, 430(3), pp.1060–1065.

Mithöfer, A. & Boland, W., 2012. Plant Defense Against Herbivores: Chemical Aspects. *Annual Review of Plant Biology*, 63(1), pp.431–450.

Moghe, G.D. & Last, R.L., 2015. Something Old, Something New: Conserved Enzymes and the Evolution of Novelty in Plant Specialized Metabolism. *Plant physiology*, 169(3), pp.1512–23.

Moore, S.J., Leung, C.L. & Cochran, J.R., 2012. Knottins: disulfide-bonded therapeutic and diagnostic peptides. *Drug Discovery Today: Technologies*, 9(1), pp.e3–e11.

Morita, H. et al., 2006. Structure of a new cyclic nonapeptide, segetalin F, and vasorelaxant activity of segetalins from *Vaccaria segetalis*. *Bioorganic & medicinal chemistry letters,* 16(17), pp.4458-4461.

Morris, C.E., 1983. Uptake and metabolism of nicotine by the CNS of a nicotine-resistant insect, the tobacco hornworm (*Manduca sexta*). *Journal of Insect Physiology*, 29(11), pp.807–817.

Murray, C. et al., 2010. Expression of various biotin-binding proteins in transgenic tobacco confers resistance to potato tuber moth, *Phthorimaea operculella* (Zeller) (fam. Gelechiidae). *Transgenic Research*, 19(6), pp.1041–1051.

Mylne, J.S. et al., 2011. Albumins and their processing machinery are hijacked for cyclic peptides in sunflower. *Nature chemical biology*, 7(5), pp.257–9.

Naimov, S., Dukiandjiev, S. & De Maagd, R.A., 2002. A hybrid *Bacillus thuringiensis* delta-endotoxin gives resistance against a coleopteran and a lepidopteran pest in transgenic potato. *Plant Biotechnology Journal*, 1(1), pp.51–57.

Neelima Alam et al., 2001. Substrate−Inhibitor Interactions in the Kinetics of α-Amylase Inhibition by Ragi α-amylase/Trypsin Inhibitor (RATI) and Its Various N-Terminal Fragments.*Biochemistry*, 40(14), pp.4229-4233.

Northfield, S.E. et al., 2014. Disulfide-rich macrocyclic peptides as templates in drug design. *European Journal of Medicinal Chemistry*, 77, pp.248–257.

Osman, G.H. et al., 2016. Development of insect resistant maize plants expressing a chitinase gene from the cotton leaf worm, *Spodoptera littoralis*. *Scientific Reports*, 5(1), p.18067.

Østergaard, H. et al., 2000. Inhibitory Serpins from Wheat Grain with Reactive Centers Resembling Glutamine-rich Repeats of Prolamin Storage Proteins cloning and characterization of five major molecular forms. *Journal of Biological Chemistry*, 275(43), pp.33272-33279.

Otlewski, J. et al., 2001. Structure-function relationship of serine protease-protein inhibitor interaction. *Acta biochimica Polonica*, 48(2), pp.419–28.

Packiam, S.M., Baskar, K. & Ignacimuthu, S., 2014. Feeding deterrent and growth inhibitory activities of PONNEEM, a newly developed phytopesticidal formulation against *Helicoverpa armigera* (Hubner). *Asian Pacific Journal of Tropical Biomedicine*, 4, pp.S323–S328.

Palma, L. et al., 2012. Vip3C, a novel class of vegetative insecticidal proteins from *Bacillus thuringiensis.Applied and environmental microbiology*, 78(19), pp.7163–5.

Patil, K.M. et al., 2012. Highly efficient (R-X-R)-type carbamates as molecular transporters for cellular delivery. *Journal of the American Chemical Society*, 134(17), pp.7196–7199.

Pettersen, E.F. et al., 2004. UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of computational chemistry*, 25(13), pp.1605–12.

Pinto, M.F.S. et al., 2012. Identification and Structural Characterization of Novel Cyclotide with Activity against an Insect Pest of Sugar Cane. *Journal of Biological Chemistry*, 287(1), pp.134–147.

Poth, A.G. et al., 2011. Discovery of an unusual biosynthetic origin for circular proteins in legumes. *Proceedings of the National Academy of Sciences*, 108(25), pp.10127–10132.

Qayyum, M.A. et al., 2015. Infection of *Helicoverpa armigera* by endophytic *Beauveria bassiana* colonizing tomato plants. *Biological Control*, 90, pp.200–207.

Qi, R.-F., Song, Z.-W. & Chi, C.-W., 2005. Structural features and molecular evolution of Bowman-Birk protease inhibitors and their potential application. *Acta biochimica et biophysica Sinica*, 37(5), pp.283–292.

Rabindra, R.J., Sathiah, N. & Jayaraj, S., 1992. Efficacy of nuclear polyhedrosis virus against *Helicoverpa armigera* (Hbn.) on Helicoverpa-resistant and susceptible varieties of chickpea. *Crop Protection*, 11(4), pp.320–322.

Rahioui, I. et al., 2014. Host range of the potential biopesticide Pea Albumin 1b (PA1b) is limited to insects. *Toxicon*, 89, pp.67–76.

Rashid War, A. et al., 2018. Plant Defense Against Herbivory and Insect Adaptations. *AoB PLANTS*, 10(4).

Rasoolizadeh, A. et al., 2016. Single substitutions to closely related amino acids contribute to the functional diversification of an insect-inducible, positively selected plant cystatin. *The FEBS Journal*, 283(7), pp.1323–1335.

Rawlings, N.D. et al., 2013. MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res*, p.gkt953.

Rawlings, N.D., Barrett, A.J. & Finn, R., 2016. Twenty years of the MEROPS database of proteolytic enzymes, their substrates and inhibitors. *Nucleic acids research*, 44(D1), pp.D343-50.

Rosenthal, G.A., 2001. L-Canavanine: A higher plant insecticidal allelochemical. *Amino Acids*, 21(3), pp.319–330.

Rosenthal, G.A., Dahlman, D.L. & Janzen, D.H., 1978. L-Canaline Detoxification: A Seed Predator's Biochemical Mechanism. *Science*, 202(4367), pp.528–529.

Roy, S. et al., 2009. Purification of Chymotrypsin-Trypsin Inhibitor from Winged Bean Seeds using Single Step Immunoaffinity Column. *American Journal of Biochemistry and Biotechnology*, 5(3), pp.142–146.

Rueden, C.T. et al., 2017. ImageJ2: ImageJ for the next generation of scientific image data. *BMC Bioinformatics*, 18(1), p.529.

Ryan, C.A., 1990. Protease Inhibitors in Plants: Genes for Improving Defenses Against Insects and Pathogens. *Annual Review of Phytopathology*, 28(1), pp.425–449.

Sagers, C.L., 1992. Manipulation of Host Plant Quality: Herbivores Keep Leaves in the Dark. *Functional Ecology*, 6(6), p.741.

Saikhedkar, N.S. et al., 2018. Tripeptides derived from reactive centre loop of potato type II protease inhibitors preferentially inhibit midgut proteases of *Helicoverpa armigera*. *Insect Biochemistry and Molecular Biology*, 95, pp.17–25.

Šali, A. & Blundell, T.L., 1993. Comparative Protein Modelling by Satisfaction of Spatial Restraints. *Journal of Molecular Biology*, 234(3), pp.779–815.

Sanchez-Hernandez, C. et al., 2004. Trypsin and alpha-amylase inhibitors are differentially induced in leaves of amaranth (*Amaranthus hypochondriacus*) in response to biotic and abiotic stress. *Physiologia Plantarum*, 122(2), pp.254–264.

Sarate, P.J., Tamhane, V. A, et al., 2012. Developmental and digestive flexibilities in the midgut of a polyphagous pest, the cotton bollworm, *Helicoverpa armigera*. *Journal of insect science (Online)*, 12, p.42.

Scheike, J.A. et al., 2007. Amide-to-Ester Substitution in Coiled Coils: The Effect of Removing Hydrogen Bonds on Protein Structure. *Angewandte Chemie International Edition*, 46(41), pp.7766–7769.

Schellenberger, U. et al., 2016. A selective insecticidal protein from Pseudomonas for controlling corn rootworms. *Science*, 354(6312), pp.634–637.

Schirra, H.J. et al., 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(3), pp.609–626.

Schirra, H.J., Anderson, M.A. & Craik, D.J., 2008a. Structural refinement of insecticidal plant proteinase inhibitors from *Nicotiana alata. Protein and peptide letters*, 15(9), pp.903–9.

Schirra, H.J., Anderson, M.A. & Craik, D.J., 2008b. Structural refinement of insecticidal plant proteinase inhibitors from *Nicotiana alata*. *Protein and peptide letters*, 15(9), pp.903–9.

Schirra, H.J. & Craik, D.J., 2005. Structure and folding of potato type II proteinase inhibitors: circular permutation and intramolecular domain swapping. *Protein and peptide letters*, 12(5), pp.421–431.

Schmidt, A.E. et al., 2005. Crystal structure of Kunitz domain 1 (KD1) of tissue factor pathway inhibitor-2 in complex with trypsin: Implications for KD1 specificity of inhibition. *Journal of Biological Chemistry*, 280(30), pp.27832–27838.

Schuman, M.C. & Baldwin, I.T., 2016. The Layers of Plant Responses to Insect Herbivores. *Annual Review of Entomology*, 61(1), pp.373–394.

Schüttelkopf, A.W. & van Aalten, D.M.F., 2004. *PRODRG*: a tool for high-throughput crystallography of protein–ligand complexes. *Acta Crystallographica Section D Biological Crystallography*, 60(8), pp.1355–1363.

Sharma, H.C. & et al, ., 2005. Insect pest management in food legumes: The future strategies.

Silverman, G.A. et al., 2001. The Serpins Are an Expanding Superfamily of Structurally Similar but Functionally Diverse Proteins. *Journal of Biological Chemistry*, 276(36), pp.33293–33296.

Simonsen, S.M. et al., 2008. Alanine Scanning Mutagenesis of the Prototypic Cyclotide Reveals a Cluster of Residues Essential for Bioactivity. *Journal of Biological Chemistry*, 283(15), pp.9805–9813.

Singh, G., Rup, P.J. & Koul, O., 2007. Acute, sublethal and combination effects of azadirachtin and *Bacillus thuringiensis* toxins on *Helicoverpa armigera* (Lepidoptera: Noctuidae) larvae. *Bulletin of Entomological Research*, 97(04), p.351.

Sivanathan, S. & Scherkenbeck, J., 2014. Cyclodepsipeptides: A Rich Source of Biologically Active Compounds for Drug Research. *Molecules*, 19(8), pp.12368–12420.

Smith, J.J. et al., 2013. The insecticidal potential of venom peptides. *Cellular and Molecular Life Sciences*, 70(19), pp.3665–3693.

Souriau, C. et al., 2005. New Binding Specificities Derived from Min-23, a Small Cystine-Stabilized Peptidic Scaffold †. *Biochemistry*, 44(19), pp.7143–7155.

Sousa da Silva, A.W. & Vranken, W.F., 2012. ACPYPE - AnteChamber PYthon Parser interfacE. *BMC Research Notes*, 5(1), p.367.

Van Der Spoel, D. et al., 2005. GROMACS: Fast, flexible, and free. *Journal of Computational Chemistry*, 26(16), pp.1701–1718.

Srinivasan, A. et al., 2005. Podborer (*Helicoverpa armigera*) does not show specific adaptations in gut proteinases to dietary *Cicer arietinum* Kunitz proteinase inhibitor. *Journal of Insect Physiology*, 51(11), pp.1268–1276.

Srinivasan, A., Giri, A.P. & Gupta, V.S., 2006a. Structural and functional diversities in lepidopteran serine proteases. *Cellular & molecular biology letters*, 11(1), pp.132–154.

Staszek, P. et al., 2017. l-Canavanine: How does a simple non-protein amino acid inhibit cellular function in a diverse living system? *Phytochemistry Reviews*, 16(6), pp.1269–1282.

Stevens, J. et al., 2012. Biotechnological Approaches for the Control of Insect Pests in Crop Plants. *Pesticides - Advances in Chemical and Botanical Pesticides*.

Stevens, J.L. et al., 2000. Inducible P450s of the CYP9 family from larval *Manduca sexta* midgut. *Insect Biochemistry and Molecular Biology*, 30(7), pp.559–568.

Stewart, M.L. et al., 2010. The MCL-1 BH3 helix is an exclusive MCL-1 inhibitor and apoptosis sensitizer. *Nature chemical biology*, 6(8), pp.595–601.

Swedberg, J.E. et al., 2009. Substrate-Guided Design of a Potent and Selective Kallikrein-Related Peptidase Inhibitor for Kallikrein 4. *Chemistry & Biology*, 16(6), pp.633–643.

Tam, J.P. et al., 1991. Disulfide bond formation in peptides by dimethyl sulfoxide. Scope and applications. *Journal of the American Chemical Society*, 113(17), pp.6657–6662.

Tamaki, F.K. & Terra, W.R., 2015. Molecular insights into mechanisms of lepidopteran serine proteinase resistance to natural plant defenses. *Biochemical and Biophysical Research Communications*, 467(4), pp.885–891.

Tamhane, V. a. et al., 2007. Diverse forms of Pin-II family proteinase inhibitors from *Capsicum annuum* adversely affect the growth and development of *Helicoverpa armigera*. *Gene*, 403(1–2), pp.29–38.

Tamhane, V. a. et al., 2005. In vivo and in vitro effect of *Capsicum annum* proteinase inhibitors on *Helicoverpa armigera* gut proteinases. *Biochimica et Biophysica Acta - General Subjects*, 1722(2), pp.156–167.

Tamhane, V.A. et al., 2012a. Enhanced permeation, leaf retention, and plant protease inhibitor activity with bicontinuous microemulsions. *Journal of Colloid and Interface Science*, 383(1), pp.177–183.

Tamhane, V.A. et al., 2012b. Plant Pin-II Family Proteinase Inhibitors□: Structural and Functional Diversity.*Funct Plant Sci Biotechnol*, 6, 42-58.

Tanpure, R.S. et al., 2017. Improved tolerance against *Helicoverpa armigera* in transgenic tomato over-expressing multi-domain proteinase inhibitor gene from *Capsicum annuum*. *Physiology and Molecular Biology of Plants*, 23(3), pp.597–604.

Terra, W.R. & Ferreira, C., 1994. Insect digestive enzymes: properties, compartmentalization and function. *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry*, 109(1), pp.1–62.

Timmerman, P. et al., 2005. Rapid and quantitative cyclization of multiple peptide loops onto synthetic scaffolds for structural mimicry of protein surfaces. *ChemBioChem*, 6(5), pp.821–824.

Troeira Henriques, S. & Craik, D.J., 2017. Cyclotide Structure and Function: The Role of Membrane Binding and Permeation. *Biochemistry*, 56(5), pp.669–682.

Trott, O. & Olson, A.J., 2009. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry*, 31(2), pp.455-461.

Trtikova, M. et al., 2015. Transgene expression and Bt protein content in transgenic Bt maize (MON810) under optimal and stressful environmental conditions. *PloS one*, 10(4), p.e0123011.

Turra, D. & Lorito, M., 2011. Potato type I and II proteinase inhibitors: modulating plant physiology and host resistance. *Current protein & peptide science*, 12(5), pp.374–385.

Vagin, A.A. et al., 2004. *REFMAC* 5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallographica Section D Biological Crystallography*, 60(12), pp.2184–2195.

Valdes-Rodriguez, S. et al., 1993. Purification, characterization, and complete amino acid sequence of a

trypsin inhibitor from amaranth (*Amaranthus hypochondriacus*) seeds. *Plant physiology*, 103(4), pp.1407–12.

de Veer, S.J. et al., 2015. Engineered protease inhibitors based on sunflower trypsin inhibitor-1 (SFTI-1) provide insights into the role of sequence and conformation in Laskowski mechanism inhibition. *The Biochemical journal*, 469(2), pp.243–53.

Vercammen, D. et al., 2006. Serpin1 of *Arabidopsis thaliana* is a Suicide Inhibitor for Metacaspase 9. *Journal of Molecular Biology*, 364(4), pp.625–636.

Vranova, V. et al., 2011. Non-protein amino acids: Plant, soil and ecosystem interactions. *Plant and Soil*, 342(1–2), pp.31–48.

Wang, B. et al., 2012. Unveiling the biosynthetic puzzle of destruxins in Metarhizium species. *Proceedings of the National Academy of Sciences, 109(4), 1287-1292*

Wang, H.X. & Ng, T.B., 2006. Concurrent isolation of a Kunitz-type trypsin inhibitor with antifungal activity and a novel lectin from *Pseudostellaria heterophylla* roots. *Biochemical and Biophysical Research Communications*, 342(1), pp.349–353.

Wang, J. et al., 2005. Novel insect resistance in *Brassica napus* developed by transformation of chitinase and scorpion toxin genes. *Plant Cell Reports*, 24(9), pp.549–555.

Wang, X. et al., 2018. Structural Diversity and Biological Activities of Cyclic Depsipeptides from Fungi. *Molecules (Basel, Switzerland)*, 23(1), 169

War, A.R. et al., 2014. Efficacy of a Combined Treatment of Neem Oil Formulation and Endosulfan against *Helicoverpa armigera* (Hub.) (Lepidoptera: Noctuidae). *International Journal of Insect Science*, 6, p.IJIS.S13608.

Weis, A.E. & Kapelinski, A., 1994. variable selection on *eurosta gall size. a path analysis of the ecological factors behind selection.Evolution*, 48(3), pp.734–745.

White, T.R. et al., 2011. On-resin N-methylation of cyclic peptides for discovery of orally bioavailable scaffolds. *Nature Chemical Biology*, 7(11), pp.810–817.

Winn, M.D. et al., 2011. Overview of the CCP4 suite and current developments. *Acta crystallographica. Section D, Biological crystallography*, 67(Pt 4), pp.235–42.

Winterbourn, C.C. et al., 1975. The estimation of red cell superoxide dismutase activity. *The Journal of laboratory and clinical medicine*, 85(2), pp.337–41.

Wu, Y. et al., 1997. Adaptation of *Helicoverpa armigera* (Lepidoptera: Noctuidae) to a proteinase inhibitor expressed in transgenic tobacco. *Molecular Breeding*, 3(5), pp.371–380.

Xu, P. et al., 2015. Design of Specific Serine Protease Inhibitors Based on a Versatile Peptide Scaffold: Conversion of a Urokinase Inhibitor to a Plasma Kallikrein Inhibitor. *Journal of Medicinal Chemistry*, 58(22), pp.8868–8876.

Xu, Z.-B. et al., 2015. Detoxification of insecticides, allechemicals and heavy metals by glutathione *S* -transferase SlGSTE1 in the gut of *Spodoptera litura*. *Insect Science*, 22(4), pp.503–511.

Yan, L. et al., 2013. Role of Tomato Lipoxygenase D in Wound-Induced Jasmonate Biosynthesis and Plant Immunity to Insect Herbivores. *PLoS Genetics*, 9(12), p.e1003964.

Yang, Y., Li, Y. & Wu, Y., 2013. Current status of insecticide resistance in *Helicoverpa armigera* after 15 years of Bt cotton planting in China. *Journal of economic entomology*, 106(1), pp.375–81.

Ye, M. et al., 2018. An herbivore-induced plant volatile reduces parasitoid attraction by changing the smell of caterpillars. *Science Advances*, 4(5), p.eaar4767.

Younas, A. et al., 2017. The efficacy of *Beauveria bassiana,* jasmonic acid and chlorantraniliprole on larval populations of *Helicoverpa armigera* in chickpea crop ecosystems. *Pest Management Science*, 73(2), pp.418–424.

Yu, Q. & Powles, S., 2014. Metabolism-Based Herbicide Resistance and Cross-Resistance in Crop Weeds: A Threat to Herbicide Sustainability and Global Crop Production. *Plant Physiology*, 166(3), pp.1106–1118.

Yudin, A.K., 2015. Macrocycles: lessons from the distant past, recent developments, and future directions. *Chemical Science*, 6(1), pp.30–49.

Zehr, U.B., 2004. Novel strategies for overcoming pests and diseases in India. *4th International Crop Science Congress,*, pp.1–8.

Zha, M. et al., 2018. A phage display-based strategy for the: De novo creation of disulfide-constrained and isomer-free bicyclic peptide affinity reagents. *Chemical Communications*, 54(32), pp.4029–4032.

Zhang, J. et al., 2017. Next-Generation Insect-Resistant Plants: RNAi-Mediated Crop Protection. *Trends in Biotechnology*, 35(9),pp.871-882.

Zhao, C. et al., 2015. A Massive Expansion of Effector Genes Underlies Gall-Formation in the Wheat Pest Mayetiola destructor. *Current Biology*, 25(5), pp.613–620.

Zhu-Salzman, K. & Zeng, R., 2015. Insect Response to Plant Defensive Protease Inhibitors. *Annual Review of Entomology*, 60(1), pp.233–252.

1. Saikhedkar, N., Summanwar, A., Joshi, R., & Giri, A. (2015). Cathepsins of lepidopteran insects: Aspects and prospects. *Insect Biochemistry and Molecular Biology*, 64, 51-59.
2. Saikhedkar, N. S., Joshi, R. S., Bhoite, A. S., Mohandasan, R., Yadav, A. K., Fernandes, M., Kulkarni, K. A. & Giri, A. P. (2018). Tripeptides derived from reactive centre loop of potato type II protease inhibitors preferentially inhibit midgut proteases of *Helicoverpa armigera*. *Insect Biochemistry and Molecular Biology*, 95, 17-25.
3. Saikhedkar, N. S., Joshi, R. S., Yadav, A. K., Seal, S., Fernandes, M., & Giri, A. P. Tailoring cyclic reactive loop peptides for augmentation of inhibition potential and functional differentiation. Manuscript submitted.

Patent:

"Synthetic Peptides As Inhibitors of Proteases for Effective Pest Control and Compositions thereof" (Application no. 201611019109, India). Published: February 2018.

1. Attended **Accelerating Biology 2015,** organized by Centre for Development of Advanced Computing,January 20-22, 2015, Pune, India.
2. Attended **National Symposium on Biophysics,** organized by Jamia Millia Islamia University, February 14-17, 2015, New Delhi, India.
3. Poster presentation on "Peptides of the future: Mining and design strategies" at **Insights in Biology 2015,**organized by CSIR-NCL and Maharashta Academy of Sciences, October 29, 2015, Pune, India.
4. Poster presentation on "From proteins to peptides: Development of insect control agents based on protease inhibitors from *Capsicum annum"* at **Peptide Engineering Meeting-7 (PEM-7)** organised by IISER Pune, December 5-7 2015, Pune, India.
5. Poster presentation on "Reactive centre loop tripeptides of Pin-II serine protease inhibitors as insect control agents against *Helicoverpa armigera"* at **44th National Seminar on Crystallography,** organized by NCCS, CSIR-NCL, IISER Pune and SP Pune University, July 10-13, 2016, Pune, India.
6. Best presentation award for "From proteins to peptides: Development of pest control agents based on plant protease inhibitors" at **Annual Students' Conference 2018** organised by NCL-Research Foundation, November 29-30, 2018, Pune, India.
7. Poster presentation on "From proteins to peptides: Development of pest control agents based on plant protease inhibitors" at **4th International Plant Physiology Congress,** organizedby CSIR-NBRI, Lucknow, December 2-5, 2018, Lucknow, India.
8. Poster presentation on **National Science Day Celebration,** held at CSIR-NCL in 2014, 2017 and 2018, Pune, India
9. Awarded with NCL RF-Agnimitra Memorial Poster Award 2018 as a part of **National Science Day Celebration**, 2018 at CSIR-NCL, Pune, India.

CURRICULUM VITAE

Nidhi Saikhedkar

E-mail: nidhsai@gmail.com
Contact Number: +91-9637840393

EDUCATION

Qualification	Board/University	Year	Aggregate
Ph.D. Biological Sciences	Academy of Scientific and Innovative Research (AcSIR) CSIR-National Chemical Laboratory, Pune, India	2018	
B. Tech Biotechnology	Amity University Rajasthan	2012	GPA 9.0
HSC (12th)	CBSE Kendriya Vidyalaya No. 1, Udaipur, Rajasthan	2008	89.6%
SSC (10th)	CBSE Kendriya Vidyalaya No. 1, Udaipur, Rajasthan	2006	90.6%

RESEARCH

- **CSIR-National Chemical Laboratory, Pune, India**

Researcher **2013-2018**

Full-time fellowship by CSIR, Government of India

Supervisor: Dr. Ashok P. Giri

Thesis title: ***Peptides for inhibition of insect proteases***

Brief synopsis:

Plants produce protease inhibitors in response to insect attack. These inhibitors inactivate the digestive proteases of insects which leads to reduction in their growth. In my doctoral work, I have:

- Analysed databases to identify reactive loop regions of plant protease inhibitors
- Designed and synthesized peptidomimetics against insect pest proteases
- Analysed peptide-protease interaction by X-ray crystallography and MD Simulations
- Biophysical studies by Surface Plasmon Resonance and Isothermal Calorimetry
- Conducted in vivo feeding experiments on agricultural pests
- Studied molecular responses of insects towards designed peptides

- **Mohanlal Sukhadia University, Udaipur, India**

Research Intern **2012-2013**

Supervisor: Dr. Maheep Bhatnagar

Project title: ***Effect of GSM mobile phone radiations on rat brain- behavioral, biochemical and histological studies***

- Developed protocols for behavioural assays on rats upon exposure to mobile phone radiation
- Performed histological examination of rat brain tissues
- Standardized procedures for biochemical estimation of oxidative stress in rat brain

- **International Centre for Genetic Engineering and Biotechnology, New Delhi, India**

Summer Research Fellow of Indian Academy of Sciences **June-July 2011**

Supervisor: Dr. Virendar S. Chauhan

Project title: ***Synthesis and characterization of small peptide nanoparticles for cancer therapy***

- Undertook training on solid phase peptide synthesis and peptide characterization
- Synthesized short peptides and purified by High Performance Liquid Chromatography (HPLC)

- **Mohanlal Sukhadia University, Udaipur, India**

Research Intern **May-July-2010**

Supervisor: Dr. S. D. Purohit

Project title: ***Modern techniques in molecular biology and Plant tissue culture***

- Trained in handling plant tissue culture
- Experience in Molecular Biology techniques

KEY AREAS OF EXPERIENCE AND TECHNICAL SKILLS

Chemistry:

- Peptide synthesis by solid phase method
- Liquid Chromatography: HPLC
- Mass spectrometry: MALDI-TOF, LC-MS/MS

Biophysics:

- Surface Plasmon Resonance: BiacoreX100
- Isothermal Calorimetry
- X-ray crystallography: CCP4 suite, Phenix, Scala, Phaser
- CD spectroscopy

Computational biology:

- Bioinformatics tools-Sequence and structural database analysis
- Molecular Dynamics simulation: GROMACS
- Docking softwares: AutoDock Vina, Z-Dock, HEX, online docking platforms

- Peptide design and protein interaction studies

Molecular biology:

- PCR
- Recombinant protein expression
- Real time PCR

Biochemistry:

- Enzyme kinetics

PUBLICATIONS

- **Saikhedkar, N**. **S.**, Joshi, R. S., Bhoite, A. S., Mohandasan, R., Yadav, A. K., Fernandes, M., Kulkarni, K. A. & Giri, A. P. (2018). Tripeptides derived from reactive centre loop of potato type II protease inhibitors preferentially inhibit midgut proteases of *Helicoverpa armigera*. *Insect Biochemistry and Molecular Biology*, 95, 17-25.
- **Saikhedkar, N.,** Summanwar, A., Joshi, R., & Giri, A. (2015). Cathepsins of lepidopteran insects: Aspects and prospects. *Insect Biochemistry and Molecular Biology*, 64, 51-59.
- **Saikhedkar, N**., Bhatnagar, M., Jain, A., Sukhwal, P., Sharma, C., & Jaiswal, N. (2014). Effects of mobile phone radiation (900 MHz radiofrequency) on structure and functions of rat brain. *Neurological Research*, 36(12), 1072-1079.
- Kelkar, S., Pandey, K., Agarkar, S., **Saikhedkar, N**., Tathavadekar, M., Agrawal, I., Gundloori, R.V. & Ogale, S. (2014). Functionally engineered egg albumen gel for quasi-solid dye sensitized solar cells. *ACS Sustainable Chemistry & Engineering*, 2(12), 2707-2714.

AWARDS AND ACHIEVEMENT

- NCL-RF Agnimitra Memorial Poster Award 2018, National Science Day Celebration, CSIR-National Chemical Laboratory, Pune
- Best presentation award for “From proteins to peptides: Development of pest control agents based on plant protease inhibitors” at Annual Students’ Conference 2018, CSIR-NCL, Pune.

CONFERENCES ATTENDED

- 44th National Seminar on Crystallography 2016, Indian Institute of Science Education and Research, Pune, India
- National Symposium on Biophysics 2015, Jamia Millia Islamia, New Delhi, India
- Peptide Engineering Meeting-7, 2015, Indian Institute of Science Education and Research, Pune, India