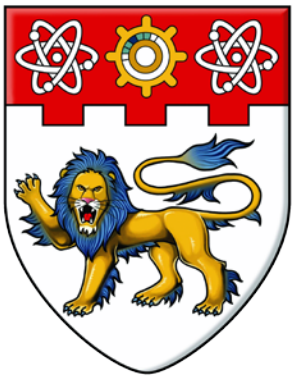


NANYANG TECHNOLOGICAL UNIVERSITY

DISCOVERY AND APPLICATION OF CYSTEINE-RICH PEPTIDES FROM MEDICINAL PLANTS IN DRUG DEVELOPMENT

KUMARI GEETA

SCHOOL OF BIOLOGICAL SCIENCES

2016

Discovery and Application of Cysteine-Rich Peptides from Medicinal Plants in Drug Development

KUMARI GEETA

SCHOOL OF BIOLOGICAL SCIENCES

A thesis submitted to the Nanyang Technological University in partial fulfillment of the requirement for the degree of Doctor of Philosophy

2016

Acknowledgments

At the outset, I thank the Almighty for his blessings. First and foremost, I would like to express my deepest gratitude to my supervisor, Professor James P. Tam whose guidance was invariably a beacon of light. This thesis has been possible only because of his help and support. I feel fortunate for to be a part of his lab and shall remain ever grateful to him for providing a conducive environment for growing as a student of science.

I am very thankful to each member of Tam's lab for all the help and support. Special thanks to Phuong for enlightening me the first glance of research. Many special thanks to Misako Tiachi for helping me understand peptide synthesis. I owe a great deal to Giang and Ka Ho for all enlightening scientific conversations, special thanks to Antony, Kim and Wei Liang for their help and assistance regarding experimental work. My heartfelt gratitude to Shruthi Kini for being a wonderful friend and supporting me through all times.

I gratefully acknowledge the financial support rendered by the Nanyang Technological University of Singapore in the form of Research Scholarship. I would like to thank my thesis committee: Prof. Liu Chuan Fa and Prof. Kathy Luo Qian for their insightful comments and encouragement during our meetings. I am also grateful to the academic and technical staffs at the School of Biological Sciences, who have helped me in one way or other in my research work.

Functional assays cannot be performed without healthy experimental models. I am grateful for the kind assistant of Prof. Ding Xiang, Prof. Kathy Luo Qian, Prof. Thirumaran, Prof. Xavier Roca, and Prof. Tan Suet Mien for giving us all the cell lines necessary for carrying out the assays.

I am extremely thankful to all my friends all over the world for their support and encouragement. I am ever indebted to the immense care and support offered by Anjali Sengar. She has been more than a good friend to me throughout my Ph.D. and provided valuable emotional support. Needless to say, I owe heartfelt thanks to Malini, Nitin, Areetha and Ipsita for their love, care, and undying support. I am thankful for all the happy times that we have shared and that are yet to come.

I wish to acknowledge gratefully the assistance offered by Dr. Aida for teaching me proteomic analysis, and Dr. Shin Joon for NMR experiments. They have been extremely helpful.

It would not have been possible for me to be pursuing my dream here, if not for beautiful mentoring done by my teachers from my earlier years of education and I wish to take this opportunity to express my most sincere gratitude to my teachers, Prof. Abdullah Dayo and Dr. Faryal Almani for recommending me to NTU.

With all my heart I dedicate great appreciation to my better half, Jiteendar Kumar. No extent of words would do justice for his support, patience, love, and encouragement. He was always there for me to provide emotional guidance and without which I believe my research would not have been completed successfully.

And the most of all, my heartfelt gratitude to my parents, my sister and my brother for their undying selfless support, love, and understanding and for being the greatest source of inspiration. I also thank my extended family members who have encouraged and supported me.

Table of Contents

Chapter 2

Materials and Methods

Chapter 3

Discovery and Characterization of a New CRP Family with Novel Disulfide Connectivity from *Jasminum sambac*

Chapter 4

Designing of Metabolically Stable Substance P Antagonists

List of Figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

List of Tables

Chapter 1

Chapter 3

Chapter 4

Abbreviations

Standard abbreviations are used for the amino acids and protecting groups [IUPAC-IUB Commission for Biochemical Nomenclature

ACN	Acetonitrile
AEP	Asparaginyl endopeptidase
AMP	Antimicrobial peptide
BLAST	Basic local alignment search tool
Boc	Tert-butoxycarbonyl
BOP	Benzytrizol-1-yl-oxy-tris(dimethylamino)phosphonium hexafluorophosphate
BrEA	Bromoethylamine
cDNA	Complementary deoxyribonucleic acid
CHCA	Alpha-cyano-4-hydroxycinnamic acid
CKAI	Cystine-knot α-amylase inhibitor
CRP	Cysteine-rich peptide
CTAB	Cetyl trimethylammonium bromide
DCM	Dichloromethane
DIC	*N,N*-diisopropylcarbodiimide
DIEA	*N,N*-diisopropylethylamine
DMEM	Dulbecco's modified Eagle's medium

DMF	*N, N*-dimethylformaide
DMSO	Dimethylsulfoxide
dNTP	Deoxyribonucleotide triphosphate
DTT	Dithiolthreitol
ELISA	Enzyme-linked immunosorbent assay
EndoGlu-C	Endoproteinase Glu-C
ER	Endoplasmic reticulum
ESI-MS	Electrospray ionization mass spectrometry
EtOH	Ethanol
FBS	Foetal bovine serum
Fmoc	9-fluroenylmethyloxycarbonyl
HATU	(1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid -hexafluorophosphate)
HBTU	*N,N,N′,N′*-Tetramethyl-O-(1H-benzotriazol-1-yl)uronium hexafluorophosphate
HF	Hydrofluoric acid
HIV	Human immunodeficiency virus
HOBt	*N*-hydroxybenzotriazole
hpi	Hour post infection
HPLC	High-performance liquid chromatography

IAM	Iodoacetamide
IBV	Infectious bronchitis virus
IC_{50}	Half-maximal inhibitory concentration
IL	Interleukin
jS	Jasmintide
kDa	Kilo Dalton
LB	Luria-Bertani broth
LC	Liquid chromatography
LC-MS	Liquid chromatography mass spectrometry
LPS	Lipopolysaccharide
MALDI	Matrix-assisted laser desorption ionization
MeCys	*N*-methylated cysteine
MeOH	Methanol
MESNa	Sodium 2-mercaptoethane sulfonate
MMA	Methyl mercaptoacetate
MOI	Multiplicity of infection
mRNA	Messenger ribonucleic acid
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry

MW	Molecular weight
NaCl	Sodium chloride
NEM	*N*-ethylmaleimide
NMR	Nuclear magnetic resonance
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
phDs	Petunia hybrida defensins
r.t.	Room temperature
RP	Reverse phase
rpm	Runs per minute
RPMI	Roswell Park Memorial Institute medium
RT	Retention time
SCX	Strong cation exchange
SD	Standard deviation
SFTI-1	Sunflower seed Trypsin Inhibitor
SP	Substance P
SPaseI	Membrane-bound type I signal peptidases
SPPS	Solid phase peptide synthesis
SS	Disulfide bond

TCEP	Tris[2-carboxyethyl]phosphine
TEBA	Thioethylbutyl amido thioester surrogate
TFA	Trifluoroacetic acid
TI	Trypsin inhibition
TIS	Triisopropylsilane
TNF-α	Tumor necrosis factor-alpha
TOF	Time-of-fly
UPLC	Ultra-performance liquid chromatography
UTR	Untranslated region
VEGF	Vascular endothelial growth factor

Abstract

Cysteine-rich peptides (CRPs) are mini-proteins, highly diverse in sequence and possess great thermal and enzymatic stability. They can be exploited as scaffolds for engineering peptidyl drugs. The focus of my thesis is to isolate and characterize CRPs from *Jasminum sambac* of Oleaceae family. Using a peptidomic approach we have identified 15 jasmintides with a novel disulfide connectivity of CysI-CysV, CysII-CysIV and CysIII-CysVI. The discovery of a new CRP family expands our knowledge of cystine motifs in plant CRPs. Understanding the structure and stability of jasmintides could lead to their use as a scaffold for designing peptide drugs. My thesis also describes the application of a CRP as a scaffold by grafting substance P antagonists into sunflower trypsin inhibitor-1. The grafted analogs are bifunctional and showed improved stability against endo- and exopeptidases. Taken together, my thesis demonstrates that CRPs could be useful for developing peptide therapeutics.

Chapter 1

Introduction

1. Bioactive Peptides in Herbal Medicine

Herbal medicine has been a mainstream in health care approach due to their pivotal role in medicinal history and its effectiveness for treatment of human diseases. The history of herbal medicine is as old as human civilization. Many cultures depend on a traditional system of medicine and medicinal plants for their primary health care. Due to limited understanding of the constituents present in plants, many traditional medicines were used based on its association and experience such as the plants resembling a snake's tounge were used to treat snake bites (e.g. *Hedyotis diffusa*) [1]. Through the practice of thousands of years, experts have determined the most effective traditional medicines and have omitted out the poisonous and fatal medicines [2, 3].

The use of plants as medicine dates back from prehistoric times. It is estimated that about 25% of pharmaceuticals prescribed worldwide come from plant source [4]. Their ethnopharmacological studies showed that the biological active principles are found to be small molecules such as flavonoids, alkaloids, tannins, volatile oils and polyphenols [5], which are plant's secondary metabolites with a molecular weight (MW) of <1 kDa. For example, morphine, one of the most popular analgesic drug, is an alkaloid found in opium (*Papaver somniferum*) [6], quinine, an antimalarial drug, is originally isolated from the bark of cinchona trees (*Cinchona officinalis*) [7] and aspirin, a salicylate drug, with anti-pyretic, anti-inflammatory and analgesic properties, is isolated from willow bark and spirea species [8]. These small chemical entities have been the focus of the most pharmaceutical companies due to the advantage of the ease and low cost of production and their oral bioavailability. Despite these advantages, there has been a shift in interest in recent decades from small molecule drugs to

biopharmaceuticals or biologics. This move is due to low specificity of small molecules owing to their small footprints which often cause off-target effects and long term-toxicity [9].

Bioactive peptides are mini-proteinaceous drugs with therapeutic values. They exist in nature along with the primary and secondary metabolites. Compared with the small metabolites which are generally <1 kDa in size, peptide bioactives are considerable large and have MWs between 2 to 8 kDa. With a large molecular weight, a peptide bioactive also would have a larger footprint than metabolites, conferring a peptide bioactive with low toxicity and high specificity as a therapeutic. Currently, peptide bioactives and protein biologics represent one of the largest group of therapeutics which include insulin, vaccine, growth hormones, enzyme inhibitors and engineered antibodies. Some of them are under the different phase of clinical trials [10]. At present, the majority of peptide bioactives are derived or inspired by animal sources [11-14]. Despite a long history of herbal usage in traditional medicines, only a few peptide bioactives from plants are known, a feature underscoring that peptide bioactives from plants are still an under-explored area in peptide- or protein-based drug development.

In the past, all the documented plant-derived compounds with known or unknown medicinal properties are non-peptide drugs. This neglect is due to the common perception about their susceptibility to denaturation by heat during decoction preparation or degradation by harsh conditions in the gut, i.e. acidic and enzymatic hydrolysis during ingestion. In addition, peptide bioactives are not readily available for absorption in the gut, irrespective of their source of origin [15]. There is still a big unexplored chemical space lying between small molecules and large proteins, which includes therapeutic mini-proteins with a molecular weight of >1 kDa but <5 kDa.

The discovery of cysteine-rich peptide (CRP) has changed this perception. CRPs are mini-proteins ranging from 2-6 kDa with a high content of cysteine-residues, in an even number (2, 4, 6 or 8). Cysteine residues are involved in forming intramolecular disulfide bonds with different disulfide connectivity patterns [16]. These disulfide bonds provide conformational constraints which lead to structural and thermodynamic stability [17, 18]. Consequently, CRPs have the size of peptides, but they possess protein-like characteristics including a stable and a well-defined tertiary structure which usually relate to their biological activities.

2. Plant CRPs

Plants produce CRPs with diversity in amino acids sequence structure and biological function. There are several families of CRPs in plants such as defensins, thionins, cyclotides, knottins and hevein-like peptides. These are gene-encoded peptides and are present throughout the plant kingdom (Figure 1.1).

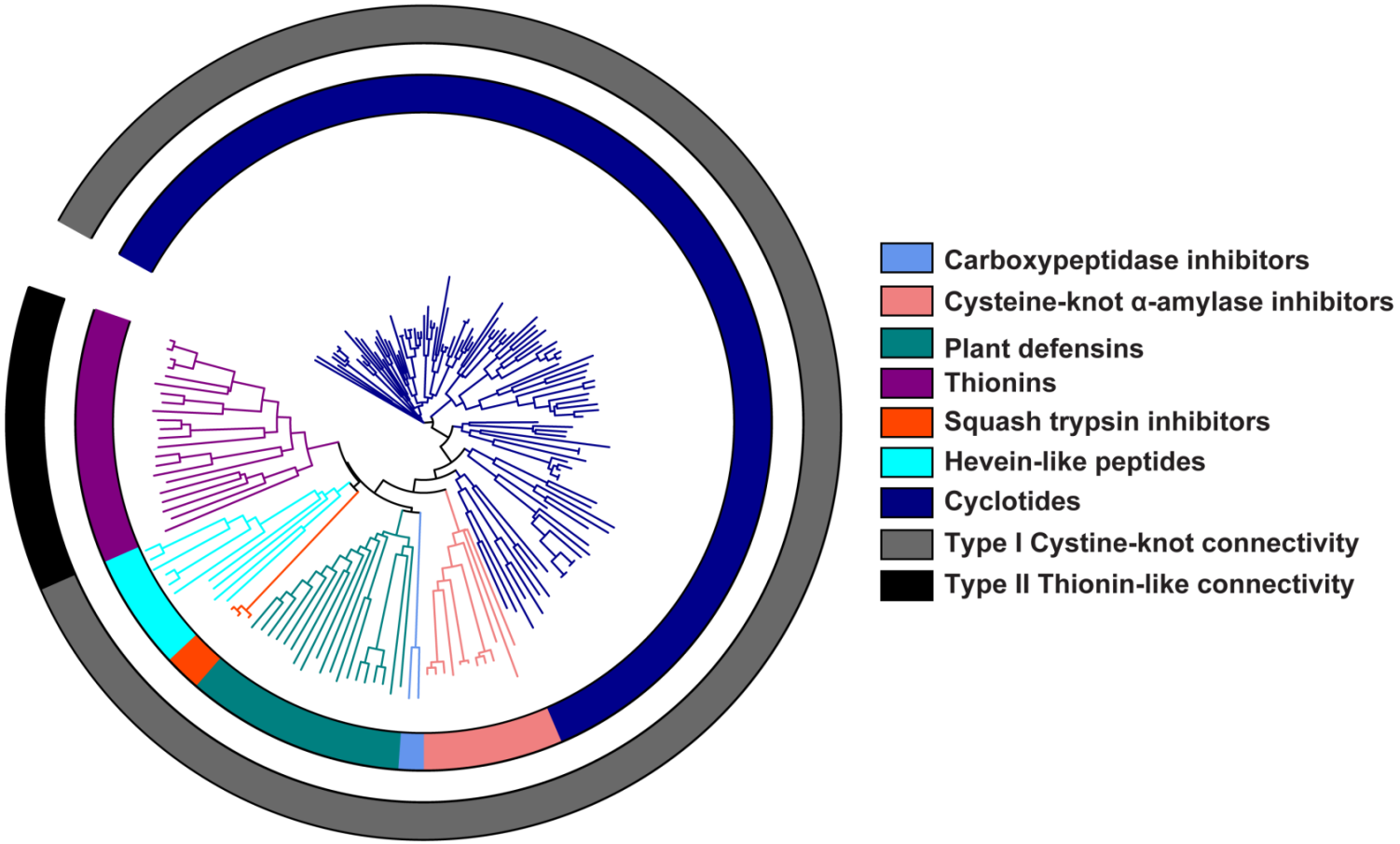


Figure 1.1. Phylogenetic tree of plant CRPs. Multiple alignment of full-length precursors of plant CRPs extracted from PhytAMP was performed using Clustral Omega (unrelated sequences were eliminated). Phylogenetic tree in Newick format was generated using Clustral Omega and drawn using ITOL [19].

A sequence motif search through the genome of dicot *Arabidopsis thaliana* and monocot *Oriza sativa* revealed that CPRs account for up to 2-3% of gene repertoire of each species [20]. Moreover, CRPs display a wide range of biological activities. Most of them play defense-related roles in plants, especially in peripheral cells of vulnerable, nutrient-rich tissues such as flower and seed. The most extensively studied function of CRPs is antimicrobial activity; hence those CRPs with antimicrobial activity are often known as antimicrobial peptides. Some CRPs are toxic to mammalian cells [21] and insect larvae [22], inhibit α-amylase activity [23], or play non-defensive roles, ranging from reproductive regulation to growth and development [24, 25]. Hereafter, CRPs can be considered as an excellent candidate for novel peptide therapeutics discovery. Taking into account the diverse pharmacological effects displayed by CRPs, as well as their extraordinary stability against thermal and enzymatic degradation, we hypothesize that they constitute a part of the active principles responsible for various therapeutic effects of herbal medicine.

3. Classification of CRPs in Plants Based on Disulfide Connectivity

CRPs contain a high content of cysteine residues in their sequences which are involved in the disulfide bond formation. CRPs with six cysteine residues would have 15 possibilities for disulfide connectivity. Thus far, out of 15 possible connectivities, only two types are known in plant CRPs. Here, CRPs containing six cysteine residues are classified into two classes based on their known disulfide connectivity.

Type I: Cystine-knot disulfide connectivity bonds are formed between CysI-CysIV, CysII-CysV and CysIII-CysVI.

Type II: Thionin-like disulfide connectivity bonds are formed between CysI-CysVI, CysII-CysV, and CysIII-CysIV.

Plant CRPs with 8 to 10 cysteines in their sequence, forming intramolecular disulfide bonds, are variations of either type I or type II disulfide connectivity.

3.1. Type I: Cystine-knot Disulfide Connectivity

This is the most common cystine connectivity found in plant CRPs and known as a cystine-knot. A cystine-knot is a structural motif found in CRPs comprising three disulfide bonds. The knot is formed when one disulfide bond pierces through the ring formed by other two disulfide bonds and interconnects the backbone (Figure 1.2). The conserved disulfide bonds are formed between CysI-CysIV, CysII-CysV, and CysIII-CysVI, but the cysteine spacing differs among the different classes. This knotted motif confers high thermal, chemical and proteolytic stability [20].

Cystine-knot peptides are a superfamily of CRPs with members having the molecular weight between 3-5 kDa and contain six to ten cysteine residues, which connect to form intramolecular disulfide bonds. Three disulfide bonds are involved in a cystine-knot arrangement whereas additional disulfide bonds do not disturb the knot-topology. Here we describe plants CRPs with three disulfide bonds under Type Ia class and CRPs with additional one or two disulfide bond are discussed under Type Ib. Type Ia class includes knottins and Type Ib includes hevein-like peptides and plant defensins. Figure 1.3 summarizes the classification of plant CRPs.

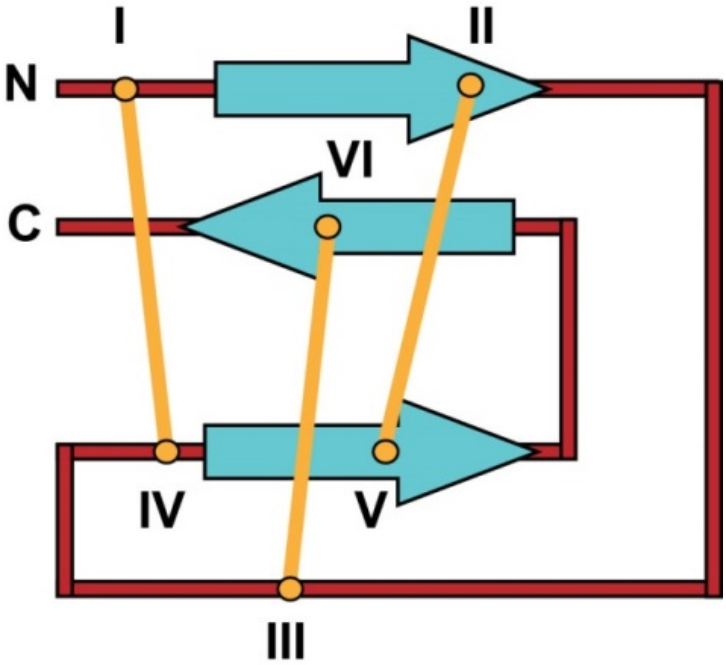


Figure 1.2. Schematic illustration of cystine-knot connectivity. The arrows represent β-strands. Cysteines are represented with small circle and numbered I to VI from N- to C-terminus. The straight yellow lines represent the disulfide linkage and red line represents the backbone.

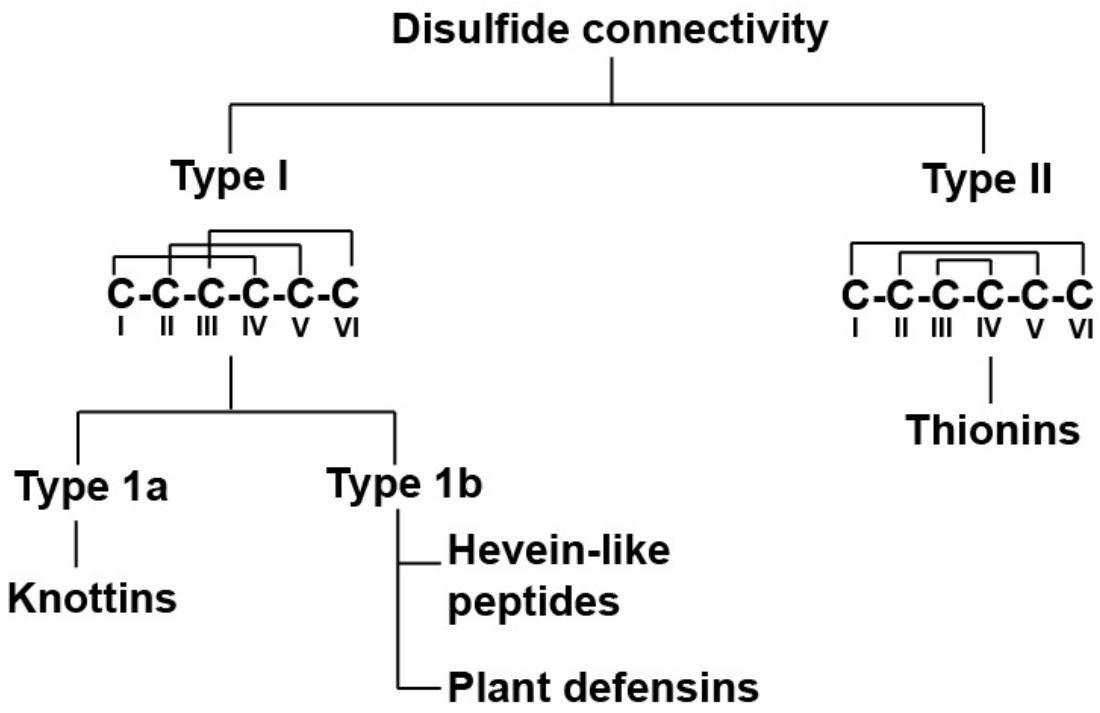


Figure 1.3. Classification of plants CRPs based on disulfide connectivity. All the CRPs are categorized based on specific type of disulfide bond connectivity. Type I reflects the cystine-knot connectivity and Type II contains thionin-like connectivity.

3.1.1. Knottins

Plant knottins are a family comprising approximately 30 amino acids residues in primary sequence with 6 cysteines. They share a common fold where cysteines are connected to form cystine-knot, but the cysteine spacing varies among sub-families. Disulfide bonds provide resistance against thermal and enzymatic stability. They include cyclotides, protease (trypsin and carboxypeptidase and α-amylase) inhibitors and hevein-like peptides. Knottins display a wide-range of biological activities, such as antimicrobial, insecticidal, cytotoxic, and enzyme inhibition [26].

3.1.1.1. Protease Inhibitors

A sub-class of cystine-knot peptides is protease inhibitors which act as a natural defense against pests and pathogens. However, not all the proteinase inhibitors are knottin-type, but here the discussion is limited to plant CRPs with cystine-knot connectivity only.

Squash proteinase inhibitors are knottin-type peptides and comprise of 27-32 residues. These are the inhibitors of trypsin and Hageman factor (β-factor XIIa) [27]. The first squash trypsin inhibitor was found in Cucurbitaceae family [28]. Except MCoTI-1 and MCoTI-II all identified squash trypsin inhibitors have a linear backbone [29]. The structures of squash inhibitors, and inhibitor and proteinase complexes have been determined by X-ray crystallography and NMR spectroscopy [30].

The other sub-class of knottin-protease inhibitors includes carboxypeptidase A inhibitors from potatoes and tomatoes [31, 32]. The plant carboxypeptidase A inhibitors comprise of 34-39 residues. Potato carboxypeptidase inhibitor (PCI) has long loops instead of typical β-strands observed in other knottins (Figure 1.4) [33]. It binds to the active site groove of carboxypeptidase A with the C-terminal

tail and utilizes some aromatic residues as secondary binding sites to inhibit the substrate binding [33].

Cystine-knot α-amylase inhibitors (CKAIs) are cysteine-rich and proline-rich peptides found in Apocynaceae and Amaranthaceae families [34, 35]. CKAIs typically consist of 30-32 residues including six cysteine and 2-4 prolines. The six cysteines are connected in knot topology. CKAIs are also referred as pseudocyclic peptides because their disulfide bond located at the penultimate cysteine residue lock the backbone, which provides resistance against exopeptidases [35, 36]. NMR and X-ray crystallographic studies of the structure of CKAIs show that they consist of three short anti-parallel β-strands [35-37] (Figure 1.4). In addition, they have a conserved proline in loop 3 (loop referred as inter cysteinyl space from N- to C-terminal) which appears in cis-configuration. Wr-AI1 and Wr-AI2, CKAIs isolated from plant *W. religosa* have shown to inhibit the α-amylase activity from *T. molitor*, but does not inhibit fungal or mammalian α-amylase [35], CKAIs can be an attractive candidates to provide pest resistance to transgenic plants because they target insect α-amylase and do not interfere with α-amylase from mammalian gut.

3.1.1.2. Cyclotides

Cyclotides are characterized by 29-37 amino acid residues with N- to C-terminal cyclized backbone. They consist of 6 cysteine residues which are involved to form three disulfide bonds in a cyclic cystine-knot (Figure 1.4), resulting in the ultrastable peptide. They have been detected in a number of plant families such as Rubiaceae, Violaceae, Solanaceae, Cucurbitaceae and Fabaceae [38]. Individual plants express suits of 10-100 cyclotides from different tissues of the plants. They were originally discovered from *Oldenlandia affinis* from Rubiaceae family, where it was used as a medicinal tea to accelerate childbirth [39].

NMR structure of cyclotides highlights some important features, including the cystine-knot, a small β-sheet structure and a series of loops and turns that project from the molecular core [40]. Cyclotide structures possess three anti-parallel β-strands with a loop connecting the first and second β-strand. Cyclotides are highly diverse in sequence but have conserved cystine-knot motif. They are of two types: Möbius and Bracelet. Möbius type has one cis-proline in loop 5 and a twist in the cyclic backbone, whereas bracelet type does not [41]. Structure scaffold is similar in both types. Currently, both cyclic and acyclic (linear) variants of cyclotides are found in monocot [42] and dicot plants [38, 43]. A wide range of activities have been reported for cyclotides including antibacterial, antifungal, antiviral, hemolytic, cytotoxic, insecticidal against molluscs, barnacles and nematodes as well as trypsin inhibition [44].

3.1.3. Hevein-like Peptides

Hevein-like peptides are chitin binding peptides ranging 29-45 amino acids. They are rich in Gly and contain conserved aromatic residues, which are known as chitin binding domain. Hevein was the first identified from the latex of rubber tree *Hevea brasiliensis*. They displayed strong antifungal activity *in vitro* [45]. Their primary sequences differ extensively with varying number of three to five disulfide bonds [46, 47]. Based on the number of cysteine residues hevein-like peptides can be divided into three sub-groups such as 6C-, 8C-, and 10C-hevein-like peptides [48]. 6C-hevein-like peptides are classified under type Ia and 8C- and 10C-hevein-like peptides are classified in type Ib. All hevein-like peptides possess cystine-knot connectivity; however, the additional disulfide bond is formed outside the cystine-knot in tail region (C-terminal) of 8C-hevein-like peptides. The structure is represented by α-helix and three to four antiparallel β-strands [49, 50]. Two anti-parallel β-strands are the main central element of hevein motif, along with the two helical turns located on each side, which are stabilized by disulfide bonds (Figure 1.4). Based on their chitin-binding capability, they were proposed to have a plant defense role. Indeed hevein-like peptides

display antimicrobial activity against fungi [51, 52], although the exact mode of action is not known yet.

3.1.4. Plant Defensins

Plant defensins are small, highly stable, cationic CRPs consisting of 45-54 residues that constitute a part of the innate immune system primarily directed against fungal pathogens [53-55]. They are generally characterized by an α-helix and a triple stranded β-sheet, stabilized by four disulfide bonds (Figure 1.4) [56]. The exterior disulfide bond is an end-to-end connected to the inner disulfide bridge and the inner three disulfide bonds are arranged in cystine-knot with the exception for phDs, isolated from *Petunia hydrida,* which contains five disulfide bonds [57]. Plant defensins have shown various biological activities including antifungal, antibacterial, insect amylase inhibitory activity and proteinase inhibitory activity [54, 58]. In addition, they show structural and functional homology to human defensins which describe their role in treating human diseases [18]. Their function is not limited to antimicrobial only, but they are also involved in growth regulation and cellular signaling. Various defensins genes have been successfully transformed into tomato, tobacco, rice, oilseed rape, and papaya [18, 53, 58-60]. Increased resistance against *Botrytis cinerea, Erwinia carotovora* and *Magnaporthe grisea* was observed when wasabi defensin (WT1) was overexpressed in orchid, potato and rice [18].

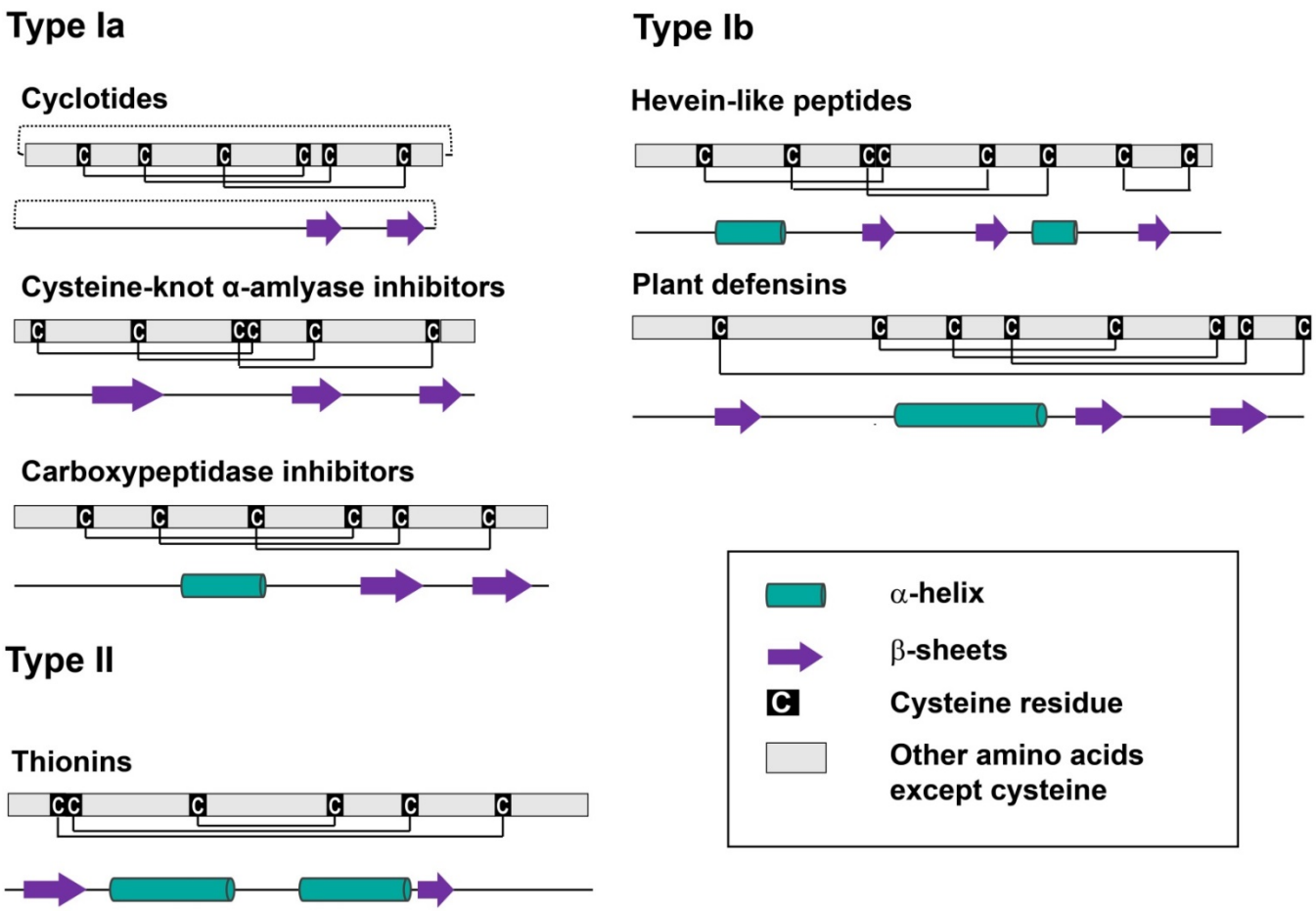


Figure 1.4. The secondary and tertiary structure of plant CRPs. A representative of each CRP family is shown with their cystine motif, disulfide connectivity and their tertiary structure. Dotted line represents the head-to-tail cyclized backbone. Cystine connectivity is shown with bold lines.

3.2. Type II: Thionin-like Disulfide Connectivity

3.2.1. Thionins

Thionins are cationic peptides of 45-48 amino acids with 6-8 cysteines. The first thionin, α-purothionin was isolated from wheat endosperm and possess antimicrobial activity [61]. Later, the subsequent thionins identified from plants were named with Greek alphabet with descending letters, for example, α-thionins, β-thionins, γ1/ γ1-thionins). Thionins are classified based on α-purothionin and include thionins of crambin, hordothionins, viscotoxins, purothionins and phoratoxin. The γ-thionins are known as a part of plant defensins because of the structural features resembling defensins found in plants and animals, but differ from α/β-thionins. Despite variation in length of amino acids (45-48), the structural fold for α/β-thionins is similar. They share a conserved $\beta 1-\alpha 1-\alpha 2-\beta 2-$coil secondary structural motif, which forms a distinct shape resembling capital letter Γ with N-terminus forming the first strand in β1-sheet.

This family of 5 kD is characterized by two antiparallel α-helices and a double stranded β-sheet, stabilized by three to four disulfide bonds [62]. The conserved four disulfide bonds are formed between CysI-CysVIII, CysII-CysVII, CysIII-CysVI and CysIV-CysV referred as 8C-thionins [48]. The disulfide bond between CysI-CysVIII connects β1-sheet to the C-terminal coil and giving the pseudocyclic structural topology. However, they are not true pseudocyclics because of the presence of additional non-cysteine residues at C-and N-termini. Conversely, in thionins with three disulfide bond, the bond between CysII-CysVII is absent [63] and referred as 6-C thionins [48].

Thionins are identified in various monocots and dicots [64] and are expressed in different plant tissues such as roots, leaves and seeds. The expression of thionins is also inducible via microbe infection [65, 66]. Thionins are ribosomally synthesized and expressed as propeptides, which are activated upon pathogenic attack [67, 68]. The gene architecture of thionins is similar to other CRPs containing three domains; a signal peptide, thionin domain and C-terminal acidic domain. The mature thionin domain is less conserved than N- and C-terminal domains in the propeptide [69, 70].

Thionins are known as plant toxins because of a broad range of toxicity against microbes, mammalian cells, yeast, and insect larvae [62], suggesting their role in plant defense. It has been postulated that most of the observed biological effects of thionins result from their interaction with the target cell membrane, but the mechanism is still not confirmed [66, 71]. However, *Stec at el*., has proposed a model to explain the interaction of thionin to anionic phospholipids leading to cell membrane permeabilization and cell membrane lysis [72].

Several experiments suggest an important role in defense against pathogenic invaders [37–39]. Thionin synthesis in response to bacterial invasion [40–42], its accumulation in vulnerable tissues [15, 43–45], its toxicity to different organisms and cell lines [37, 46–49], and the improved resistance observed when expressed transgenetically [50–52], all strongly support its antimicrobial function.

4. Biosynthesis of CRPs

CRPs are gene-encoded, ribosomally synthesized mini-proteins. In general, a precursor contains mature CRP domain which is found with N- or C-terminus propeptides along with the N-terminal signal peptide destined to the endoplasmic reticulum (ER). After being synthesized as a linear precursor, signal peptide directs the preprotein to the targeted destination before being excised by the membrane-bound type I signal peptidases (SPase I) during translocation across the cell membrane [73]. Signal peptides usually contain 16-30 amino acids comprised of a hydrophilic N-terminal region, central hydrophobic region and C-terminal region with SPase I cleavage site [74].

The cellular processing to produce mature CRP involves multiple steps. The study suggests that signal peptide cleaves first possibly with SPase I and targets the precursor to ER lumen co- or post-translationally. In this highly oxidative environment, the folding and disulfide bond formation occur with the assistance of protein disulfide isomerase and chaperone to produce correctly folded prepropeptides [75]. Misfolded peptides remain inside the ER lumen for subsequent ER-associated degradation [76, 77]. After folding prepropeptides leave the endoplasmic reticulum and transported to the pre-vacuolar compartment and the vacuole for further proteolytic processing. Propeptides are cleaved by proteases to release a mature CRP, however much remains unknown regarding the enzymes involved in the biosynthesis of CRPs (Figure 1.5 A).

Various gene arrangements have been reported for CRPs. Gene organizations for plant CRPs discussed in this thesis are described below.

Cyclotide precursor is composed of an ER signal peptide, pro-domain, one (or more) mature cyclotide domain(s), and a short C-terminal tail [42, 78] (Figure 1.5 B). The synthesis of mature cyclotide requires an additional step of transpeptidation before releasing a mature cyclic cyclotide. The studies suggest the asparaginyl endopeptidases (AEP) could be responsible for backbone cyclization [79]. Linear variants of cyclotides share the high sequence similarity but, its gene sequence revealed a premature stop codon at the position of ultimate Asp/Asn residue, which is essential for backbone cyclization [43]. Carboxypeptidase inhibitors share the same four-domain precursor as cyclotides from Rubiaceae and Violaceae [78]. Instead of multiple mature domains as in cyclotides, carboxypeptidase inhibitors consist of one mature CRP domain.

Plant defensins are processed from two types for precursors. The majorities of plant defensins comprise of two-domain precursor N-terminal signal peptides, a mature defensin domain [80]. However, the minor group consists of an additional C-terminal tail. Plant defensins synthesized from later type precursor are found in vacuoles suggesting the role of C-terminal tail in vacuolar sorting mechanism [81]. Hevein-like peptides [82] and thionins [69, 70] share the three-domain precursor similar to the later type of plant defensin precursor. The genetic structure of cystine-knot α-amylase inhibitors also share three-domain precursor, but it consists of N-terminal pro-domain instead of C-terminal domain [35, 36] (Figure 1.5 B). This high diversity in gene architectures suggests that plants advance with different mechanisms of gene expression modulation, to facilitate better adaptation and survival.

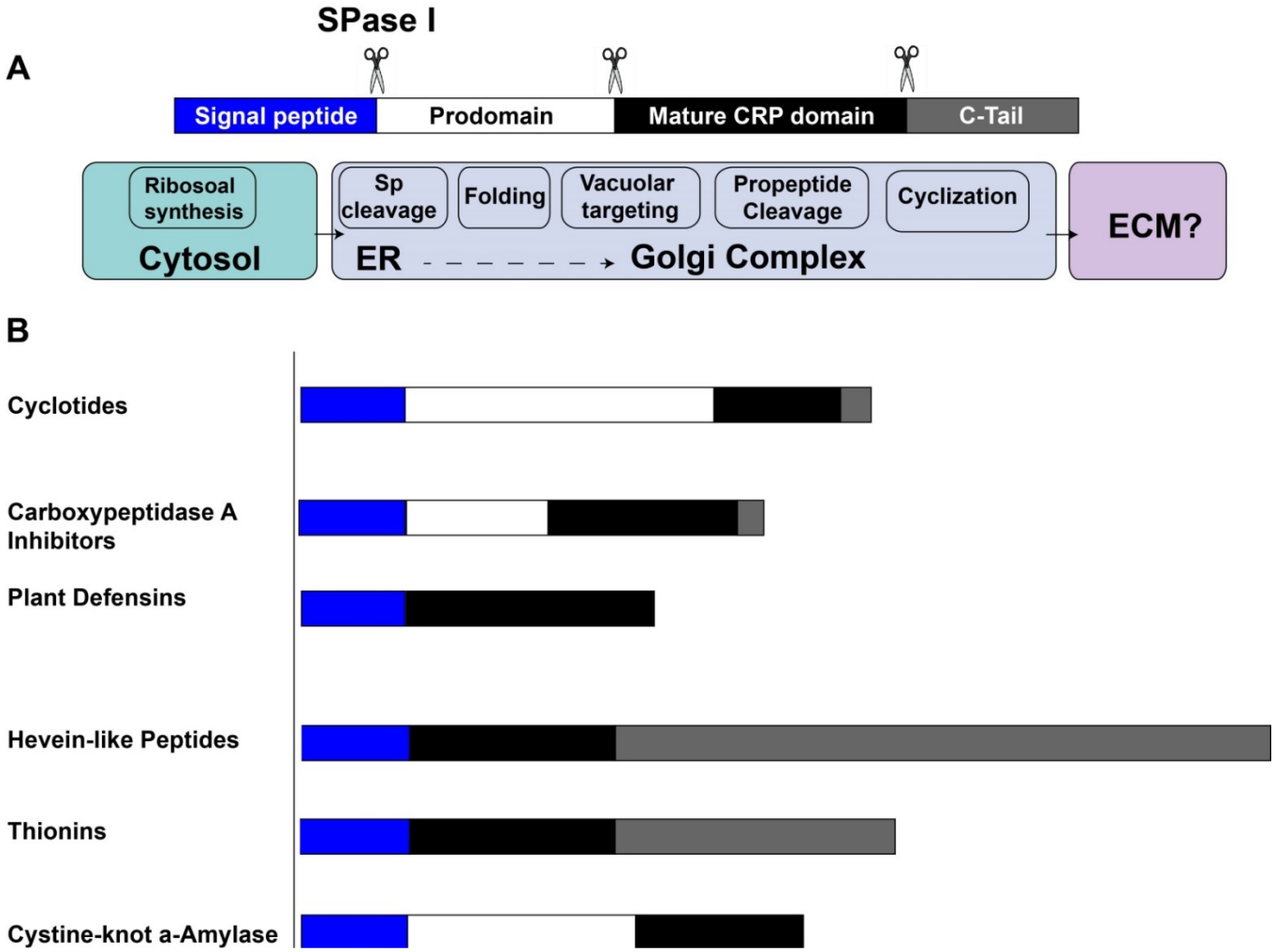


Figure 1.5. Proposed biosynthesis of CRPs. A) showing the cellular process to produce mature CRP. ER, Endoplasmic reticulum; ECM, extracellular matrix. B) various types of gene arrangements of plant CRPs.

5. Biological Activities

CRPs have also been screened for various biological activities including antimicrobial, anti-viral, uterotonic, hemolytic, cytotoxic, neurotensin antagonism activities. To gain a better understanding into the mechanism of each function, further studies are required. These activities have interesting pharmaceutical prospective but are unlikely to be the primary function of CRPs.

With the discovery of plant CRPs the first question arises about their natural role in plants. CRPs are expressed in different tissues of the plant in various isoforms and their seasonal and geographical variations suggest they could be involved in host-defense mechanism to protect which was later depicted by cyclotides, α-amylase and trypsin inhibitors. Bioactivies of peptides are listed in Table 1.1 and some are described here in detail.

5.1. Antimicrobial Activity

Plants produce antimicrobial peptides (AMPs) as a component of plant defense mechanism. They are usually small, positively charged and rich in disulfide bonds [16]. They are active against phytopathogens as well as against organisms pathogenic to bacteria, virus, human, parasite, protozoa and neoplastic cells [83]. The main families which show antimicrobial activities are plant defensins, thionins, hevein-like peptides, cyclotides and protease inhibitors [48].

Most of the AMPs exhibit toxicity to bacteria, fungi, plant and animals by interaction and forming pores in the membrane which results in ion and metabolite leakage, depolarization, interruption of the respiratory processes, and cell death. Hydrophobic patches and positives charged residues allow them to interact with membrane lipids [84]. The best characterized members of this family are AMPs extracted from the seeds of *Mirabilis jalapa* and *Amaranthus caudatus*,

which exhibit a broad spectrum of antimicrobial activity against fungi and Gram-positive bacteria [85].

Tam and colleagues reported the antimicrobial activity of four synthetically produced cyclotides; Kalata B1, circulin A, circulin B and cyclopsychotride [86]. Unlike other knottins, cyclotides have a shorter spectrum of antimicrobial activities. Ec-AMP-1 isolated from the seeds of baryard grass *Echinochloa crus-gali*, a member of α-hairpinin family, shows antifungal activity against phytopathogenic fungi. It binds the fungal conidia surface and then internalizes and accumulates in the cytoplasm without disturbing membrane integrity [87].

AMPs are encoded by small genes with conserved sequences; therefore, gene amplification and transgenesis are one of the feasible ways to increase production and enhance the specific activity of selected peptides [88]. Hence, AMPs have a high potential for therapeutic use in healthcare and agriculture, and can be used as natural antibiotics as an alternative to their chemical counterparts, for the protection of plants, and/or animals against diseases.

5.2. Anti-viral Activity

Along with the defensive properties, CRPs have been reported for antiviral activity. For instance, cyclotides exhibit anti-HIV activity by inhibiting the cytopathic effects of HIV in the infected cell. Both subfamilies of cyclotides bracelet and Möbius retain anti-HIV activity. Structural studies of kB1 showed that its linear mutants lack the anti-HIV activity, speculating the role of cyclic backbone in anti-HIV activity. Furthermore, the hydrophobic patch in loop 2 and 3 of bracelet family and loop 5 and 6 of Möbius family was found to be responsible for HIV inhibitory properties probably via cell membrane interaction [89]. Another peptide Luffin P1, isolated from the seeds of sponge gourd *Luffa cylindrical*, has been reported to possess anti-HIV activity in HIV-infected C8166 T cells *in vitro*. This study proposed that Luffin P1 displays a novel inhibitory mechanism owing to its charge complementation with viral and cellular proteins [90].

α-Amylase inhibitors from *Alstonia scholaris* (As1 and As3) have shown to possess antiviral activity against the infectious bronchitis virus (IBV), a highly infectious disease of fowl. It inhibits the early phase of IBV infection via binding to viral S proteins [91]. Thus, the anti-IBV activity and non-cytotoxic nature of As1 makes it a potential candidate for anti-IBV drugs.

5.3. Enzyme Inhibition Activity

Plants express mini-proteins that specifically confer resistance against insects, pathogen and pests by forming complexes with enzymes which blocks the active sites or alter enzyme conformation, ultimately reducing enzyme function. These mini-proteins are CRPs which include plant defensins and protease inhibitors. These peptides are naturally produced in high amount in the plant to propagate the plant defense mechanism.

Sunflower trypsin Inhibitor-1 (SFTI-1) is a 14 amino acid cyclic peptide, isolated from the seeds of sunflower, *Helianthus annus* [92]. It is the smallest potent trypsin inhibitor known. SFTI-1 has high sequence homology to the trypsin-reactive loop region from Bowman-Birk Inhibitors (BBIs), a class of serine protease inhibitors [93]. Structure analysis of solution structure of SFTI-1 and in complex with trypsin enzyme does not show any difference [94], which suggests SFTI-1 structure is rigid and in pre-defined for potent binding. This features of stability and pre-organized form makes it stronger and potent inhibitors than other BBIs.

5.4. Insecticidal Activity

The control of insect pests in plants is of great social and economic importance. CRPs have been reported to minimize the crop loss from insect predation. Jennings and colleagues have reported the insecticidal activity of cyclotides, Kalata B1 and B2, against *Helicoverpa species*. Kalata B1 induces disruption of the microvilli, blebbing, swelling, and ultimately rupture of the cells of the gut epithelium similar to *Bacillus thuringiensis* delta-endotoxin. Plant defensin (PA1b) extracted from the seeds of *P. sativum* induces mortality into cereal pests *Sitophilus sp.* [95-97]. *Gressent et al*., speculates this inhibitory effect could be attributed to the binding of peptides to protein binding sites, further investigations are needed [98]. Inhibition of endogeneous digestive proteases is one of the mechanisms by which small CRPs exerts their effects, for example, Amaranthus α-amylase inhibitor which inhibits a-amylase from *Periplaneta americana*, *Tribolium castaneum*, *Tenebrio molitor*, and *Prostaphanus truncates* but, does not inhibits from humans [34].

Table 1.1. Bioactivities of plant CRPs

	Representative member			
Peptide family	**Name**	**Plant source**	**Bioactivity**	**Reference**
Cyclotides	Katala B1	*Oldenlandia affinis*	Urotensin	[99]
			Antimicrobial	[100]
			Insecticidal	[78]
	Violapeptide I	*Viola tricolor*	Hemolysis	[101]
	Cyclopsychotride A	*Psychotria longipes*	Neurotensin antagonism	[102]
α-Amylase inhibitors	Alstotide S1	*Alstonia scholaris*	α-Amylase inhibition	[91]
			Anti-IBV	[91]
	SFTI-1	*Helianthus annus*	Trypsin inhibition	[92]
	MoCTI-I	*Momordica cochinchinensis*	Trypsin inhibition	[29]
	MCEI-1	*Momordica cliarantia*	Elastase inhibition	[103]
	PCI	*Solanum tuberosum*	Carboxypeptidase A inhibition	[104]

Cont'd.

	Representative member			
Peptide family	**Name**	**Plant source**	**Bioactivity**	**Reference**
Hevein-like peptides	Hevein	*Hevea brasiliensis*	Antifungal	[45]
	Ac-AMP	*Amaranthus caudatus*	Antimicrobial	[51]
Plant defensins	AlfAFP	*Medicago sativa*	Antifungal	[105]
	Cp-thionin II	*Vigna unguiculata*	Antibacterial	[106]
	Cp-thionin	*Vigna unguiculata*	Trypsin inhibition	[107]
	S1α1	*Sorghum bicolor*	α-amylase inhibition	[108]
	γ1-zeathionin	*Zea mays*	Na^+ channel blocker	[109]
Thionins	TuAMP	*Tulipa gensneriana*	Antimicrobial	[110]

6. Peptides as Drug Leads and Scaffolds

Peptides play an important role as drug leads and pharmaceutical discovery tools because of their high specificity and low toxicity. However, the utility of peptides as therapeutics has limitations due to their poor stability and limited bioavailability, which results in short half-lives, poor membrane permeability and difficulties for oral delivery. Linear peptides degrade rapidly into inactive fragments by proteolytic enzymes in the gastrointestinal tract (GIT) and by serum enzymes, which are then filtered out from the blood stream via kidney within minutes. As a result, the desired therapeutic effects are not produced. Because of these limitations, injections remained the most common method for peptides administration and limiting the therapeutic applications of peptides. An approach to develop peptidyl drugs with the stability of small molecule and specificity of protein drug could fill this gap in between two therapeutic groups and can be a promising drug candidate.

In response to this challenge, several approaches are utilized such as cyclization [111], side chain cross-linking [112], retro-inverso compounds [113], introducing unnatural amino acids [114] and engineering of bioactive peptides into stable natural scaffold [115-118] . The last approach is termed as ‘grafting’, i.e. insertion of the bioactive peptide into stably folded protein such as CRPs.

The concept of peptide grafting was first introduced by *Vita et al.,* in 1995 on scorpion toxins called charybdotoxin by engineering metal-binding site on it [119] In 2008, *Li et al*., showed the capability of a cystine-knot scaffold by turning a scorpion toxin named BmBKTx1 into an anti-tumor mini-potein by grafting a p53 inhibition and substitution of multiple cationic residues to allow cell penetration [120]. The results showed that the cystine-knot structure was not affected by the grafting of an external peptide. It was able to form the disulfide bonds into the knotted structure and most importantly, it had the ability to pass through the cell membrane and inhibit the p53 to process the anti-tumor activity. In the same year, Gunasekera and colleagues studied the peptide engineering concept by

grafting an anti-angiogenic peptide into loop 2, 3, 5 or 6 of kalata B1 (Figure 1.6). The most active analog was found to be one with grafted epitope into loop 3 (Cpr3) and exhibited antagonism against VEGF-A with an IC_{50} of 12 μM [116].

Our laboratory has developed an orally active bradykinin B1 receptor antagonist by grafting a linear B1-receptor antagonist DALK (nine residues) and DAK (seven residues) into the stable kB1 scaffold (Figure 1.7). Bradykinin B1 receptor is involved in stimulating chronic inflammation and its antagonism could treat inflammatory pain. The linear antagonists have limited clinical applications because of their susceptibility against peptidases. The engineered bradykinin antagonists became not only stable but have superior bioavailability. These antagonists also displayed significant pain inhibition in writhing assay done on mice with oral administration and compared to its linear analogs [118]. Few examples of engineering peptides are described here from many applications listed in Table 1.2.

These studies suggest that CRPs as scaffold confers impressive stability to linear peptide epitopes and maintain their biological activity as well. With the emergence of more peptide base therapeutics in the market, their applications as scaffold empower more prominent advances in drug design.

Table 1.2. Applications of CRPs as molecular scaffold

Scaffold	Epitope	Engineered site	Application	References
Conotoxin*	GGAAGG linker	N-C termini	Neuropathic pain	[115]
Kalata B1	VEGF-A antagonist	Loop 3	Anti-Cancer	[116]
	Bradykinin antagonist	Loop 6	Analgesic	[118]
	Fragments of Myelin Oligodendrocyte glycoprotein	Loop 5	Multiple Sclerosis	[121]
		Loop 6		
		Loop 5 & 6		
	Melanocortin agonists	Loop 6	Obesity	[122]
MCoTI-I	CXCR4 antagonist	Loop 6	Anti-HIV	[123]
	Hdm2/HdmX antagonist	Loop 6	Antitumor	[124]
SFTI-1	Osteopontin antagonist	TI-loop	Angiogenesis (Cardiovascular or wound healing)	[117]
	VEGF-A			
	Laminin α 1			

Cont'd.

Scaffold	Epitope	Engineered site	Application	References
MCoTI-II	Osteopontin antagonist	Loop 6	Angiogenesis (Cardiovascular or wound healing)	[117]
	VEGF-A			
	Laminin α 1	Loop 1	Foot-and-mouth disease	[125]
	3C protease inhibitor	Loop 1 & 6	Chronic myeloid leukemia	[126]
	BCR-ABL inhibitor			
θ-defensin RTD-1**	Integrin-binding domain (RDG)	β-turns	Anti-platelet aggregation	[127]

* Peptide from cone snail venom

** Mammalian antimicrobial peptide

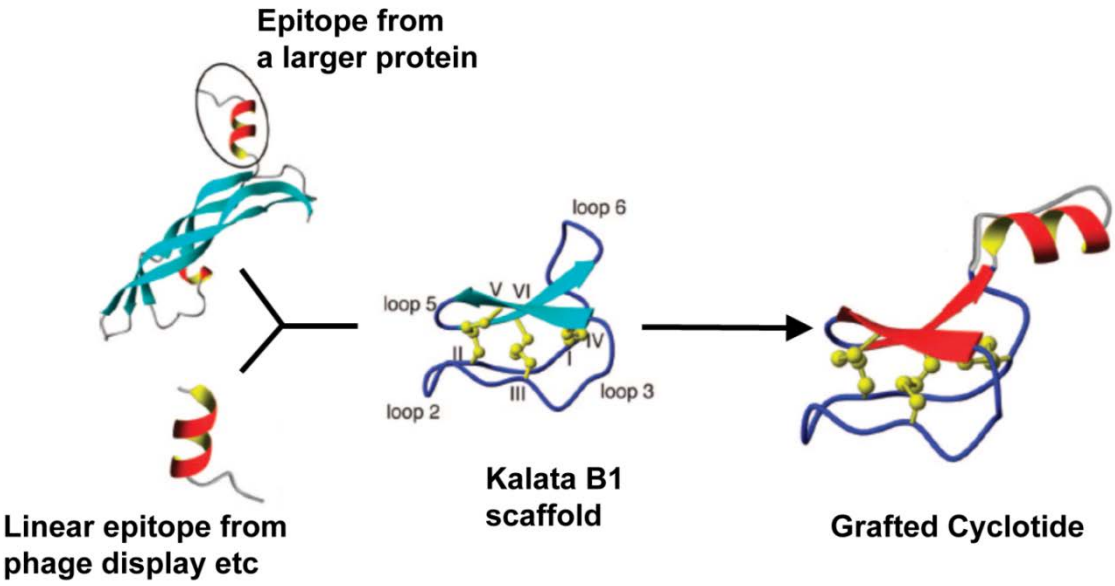


Figure 1.6. Scheme representing the concept of grafting of VEGF antagonist on the cyclotide kalata B1 scaffold. VEGF-antagonist was grafted into each loop 2, 3, 5 and 6. The figure is adapted from reference [116].

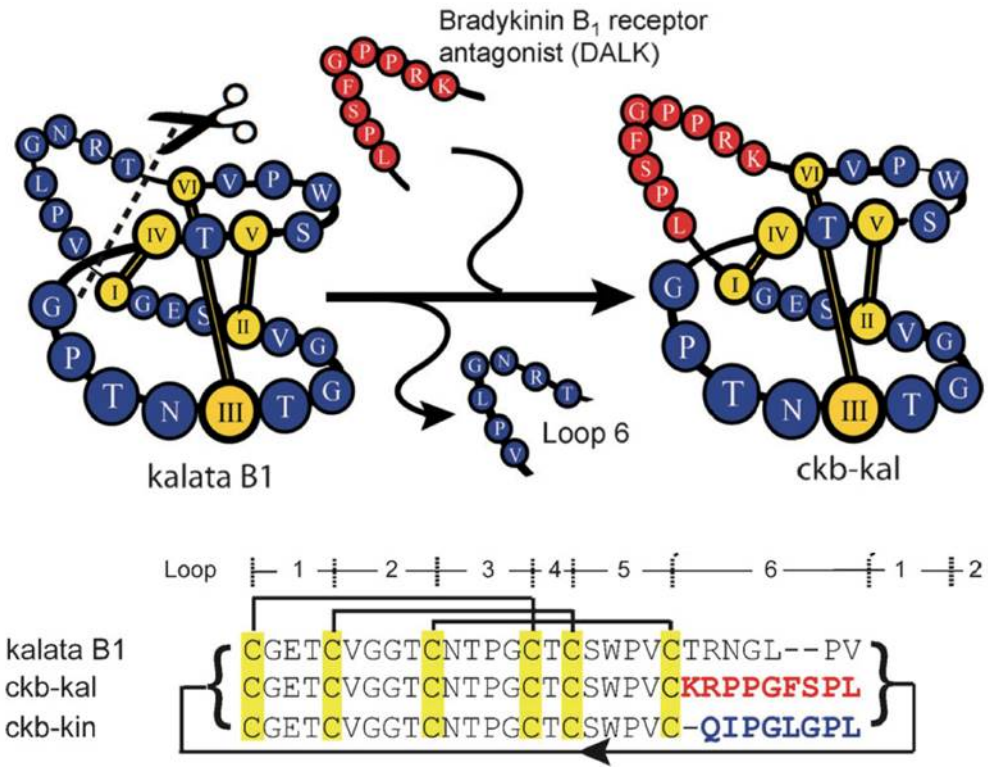


Figure 1.7. Schematic representation of engineering of bradykinin antagonist using kalata B1 scaffold. Two bradykinin antagonist ckb-kal and ckb-kin were grafted in loop 6 of kalata B1. The sequences of linear antagonists are highlighted in red and blue. Cysteines are highlighted in yellow. The figure is adapted from reference [118].

7. Aims and Objectives

CRPs are mini-proteins, highly diverse in sequence and possess great stability against thermal and enzymatic degradation. They also play a defensive role and have the potential for novel protein-based therapeutic development. CRPs are a under-explored class of active principle in medicinal plants, therefore, insight into their diverse sequence and bioactivities will greatly benefit their applications in drug discovery.

The specific aims of my thesis are:

1. To isolate and characterize novel CRPs from *Jasminum sambac* by genomic, proteomics and transcriptomic approaches

2. To determine the biological function of CRPs isolated from *Jasminum sambac*

3. To design, synthesize and engineer cyclic CRPs as peptide bioactives

Research in this thesis focuses on the discovery of novel plant CRPs from *Jasminum sambac*. It was shown that these CRPs display hypervariable sequences and most interestingly, they do not show any homology to known plant CRPs. In addition, the disulfide connectivity pattern is novel in plant CRPs. In-depth characterizations of jasmintides have been studied. Biological, structural and genetic characterizations have also been described in detail. These studies reveal a new scaffold from the plant with unique cystine connectivity for designing of peptide drugs.

My thesis also describes the engineering and synthesis of a cyclic CRP SFTI-1 by chemical synthesis approaches. Enhanced stability against thermal and enzymatic degradation was obtained and the engineered peptide showed bifunctional property possessing both the trypsin inhibition activity and the substance P antagonistic activity. This study provides the insight for developing new peptidyl drugs using the concept of protein engineering.

Chapter 2

Materials and Methods

1. Materials

1.1. Chemicals and Reagents

All the chemicals and reagents used in this study were of analytical or molecular biology grade and purchased from the following companies:

Acetic acid	Merck
Acetonitrile (ACN)	Fisher
Agarose	Bio-Rad
Ammonium bicarbonate (NH_4HCO_3)	Sigma-Aldrich
Dichloromethane (DCM)	Merck
Dithiothreitol (DTT)	Sigma-Aldrich
Ethanol (EtOH)	Merck
Iodoacetamide (IAA)	Sigma-Aldrich
Isopropanol	Fisher
Methanol	Merck
Sodium chloride	Sigma-Aldrich
Sulfuric acid (H_2SO_4)	Sigma-Aldrich
Trifluoroacetic acid (TFA)	Merck
Triton-X	Bio-Rad

1.2. Enzymes

The enzymes used for MS sequencing (trypsin, chymotrypsin, and EndoGlu-C) were of sequencing grade and purchased from Roche (Switzerland). The enzymes for proteolytic stability tests were purchased from Sigma-Aldrich (USA). The enzymes used in molecular cloning were of molecular biology grade and were bought from Fermentas (Canada, USA), Promega (USA) and NEB (UK).

1.3. Kits

The kits for molecular cloning were purchased from Invitrogen (Life Technologies, USA), Qiagen (USA), Clontech (Takara Bio, Japan), and Promega (USA).

1.4. Cell Lines, Viruses and Antibodies

All cell lines used in this thesis were received as kind gifts from various laboratories in School of Biological Sciences, Nanyang Technological University (NTU), Singapore. African green monkey kidney (Vero) cells, p53-deficient human non-small lung cancer (H1299) cells were received from Dr. Ding Xiang Liu, Nanyang Technological University, Singapore. Human umbilical vein endothelial cells (Huvec-CS) cells were gifted from Professor Kathy Luo Qian, Human Foreskin Fibroblast (HFF) cells were gifted from Professor Thirumaran, Nanyang Technological University, Singapore and Mouse macrophages derived from Abelson murine leukemia virus transformed (Raw 264.7) were gifted from Professor Tan Suet Mien, NTU, Singapore. These cell lines were recovered from previously cryopreserved vials and grown in culture flasks containing 10 mL of DMEM medium (Thermo Scientific Hyclone) supplemented with 10% foetal bovine serum (FBS) and 1% Penicillin /Streptomycin (PAA Laboratories) for 3-5 days until the desired density was reached. THP-1 (Human acute monocytic leukemic cells) was a gift from Professor Xavier Roca, Nanyang Technological Univeristy, Singapore. These cells were cultured in culture flasks containing 10

mL of RPMI-1640 medium (Thermo Scientific Hyclone) supplemented with 10% foetal bovine serum (FBS) and 1% Penicillin /Streptomycin. All the cultured cells were incubated at 37°C, 5% CO_2 humidified incubator.

The virus stock of infectious bronchitis virus (IBV) used in this study and the primary antibody targeting IBV protein were gifted by the laboratory of Dr. Ding Xiang Liu. Secondary antibodies conjugated with horseradish peroxidase (HRP) were bought form Santa Cruz (USA).

1.5. Plant Materials

All the plant materials included in this study for extraction were collected throughout the Singapore whereas jasmine flowers were purchased from flower shops, at Little India, Singapore. Different parts of the plants such as leaves, roots and flowers used for RNA extraction were immediately used after harvesting. The plant was identified by Paul Leong at the Singapore Botanic Garden. A voucher specimen (SING 2015-202) was deposited at Singapore Botanic Garden Herbarium.

1.6. Insects

In this study, *Tenebrio molitor* (yellow mealworms) used were purchased from a fish shop in Singapore. The colony was kept in glass containers on a diet consisting of oat flakes. The colony was kept in a rearing chamber at +29°C in the dark.

2. Instrumentation

2.1. MALDI-TOF MS and MS/MS

Mass spectrometry was performed on the ABI 4800 MALDI-TOF/TOF system (Applied Biosystems, Framingham, MA, USA). The instrument was equipped with a solid-state laser (diode pumped Nd:YAG laser) pulsing at a repetition rate of 200 Hz. An 8 mg/mL solution of CHCA in 60% acetonitrile, 0.05% TFA was used as a MALDI matrix. Samples were mixed thoroughly with the matrix solution at the ratio of 1:1 (v/v) and 0.5 µL of the mixture was spotted onto a target plate. The instrument was calibrated externally using a mixture of peptide standards obtained from Sigma- Aldrich (MSCal1).

Both MS and MS/MS spectra were acquired using dual-stage reflectron mirror. The laser intensity was set between 3000-4000. Accelerating voltages applied for MS and MS/MS measurements were 20 and 8 kV, respectively. The mass spectra were accumulated up to 1000 and 5000 shots in MS and MS/MS mode, respectively. In MS/MS mode, collision energy of 1 kV was applied and nitrogen was used as a collision gas in collision-induced dissociation experiments.

2.2. High-Performance Liquid Chromatography (HPLC)

HPLC was run on a Shimadzu system equipped with UV detector at 220, 254 and 280 nm. For RP-HPLC, a Grace Vydac 250 $\times$ 21 mm C18 column (5 µm particle size and 300 Å pore size) was used for preparative purifications at a flow rate of 5 mL/min; a 250 $\times$ 10 mm C18 column was used for semi-preparative and analytical purifications at a flow rate of 3 mL/min and 1 mL/min, respectively. For ion exchange, PolyLC 250 $\times$ 9.4 mm and 250 $\times$ 4.6 mm polysulfoethyl A column were used for semi-preparative and analytical purifications at a flow rate of 3 mL/min and 1 mL/min, respectively.

2.3. LC-MS/MS

LC-MS/MS was performed in an Orbitrap Elite mass spectrometer (Thermo Scientific Inc., Bremen, Germany) coupled with a Dionex UltiMate 3000 UHPLC system (Thermo Scientific Inc.) into an Acclaim PepMap RSL column (75 µm × 15 cm; 2 µm particles; Thermo Scientific Inc., Bremen, Germany).

3. Genomics

3.1. RNA Extraction

Total RNA was extracted from fresh leaves, flowers and root of *J. sambac* grown in NTU Herb Garden using Trizol reagent (Life Technologies, California, USA) using manufacture's protocol with slight modification. Briefly, plant material was ground using liquid nitrogen and homogenized, followed by addition of 1 mL trizol per 100 mg of plant material. 200 µL of chloroform was added and centrifuged at 12000 rpm for 15 min. The aqueous layer was collected and RNA was precipitated using 500 µL of isopropanol and 125 µL of high salt solution (1.2 M NaCl and 0.8 M sodium acetate in RNase free water) and incubating at -20°C for 30 min followed by two washes of 75% ethanol in water. Concentration and purity of extract was measured using NanoDrop 2000 UV-Vis spectrophotometer (Thermo Scientific, Waltham, Massachusetts, USA) and quality of RNA was assessed by gel electrophoresis. Purified RNA of high quality (A260/280 and A260/230 not less than 1.8) was stored at -80°C. The sequencing and analysis were performed by Macrogen (Seoul, Korea).

3.2. DNA Extraction

Samples were homogenized in liquid nitrogen. CTAB buffer (500 µl) was added to 50-60 mg of homogenized sample and incubated in a water bath at 55°C for 1 h with intermittent shaking. After incubation, chloroform (500 µl) was added to the samples followed by centrifugation at 16000g for 7 minutes. The aqueous layer

was collected and 0.08 volumes of cold 7.5 M ammonium acetate and 0.54 volumes of cold isopropanol were added and incubated on ice for 30 minutes. The sample was centrifuged at 16000g for 3 minutes and the pellet was washed twice with 75% ethanol. After the pellet was dried it was resuspended in DEPC water (30 µl) and DNA concentration and integrity were determined using NanoDrop 2000 UV-Vis spectrophotometer (Thermo Scientific, Waltham, Massachusetts, USA) and 1.0% agarose gel electrophoresis.

3.3. Rapid Amplification of cDNA Ends (RACE) and PCR Analysis

The RNA extract served as the template to generate 3′ RACE (Rapid Amplification of cDNA Ends) cDNA population with GeneRacerTM Kit (Invitrogen, Carlsbad, CA, USA) following the manufacturer‟s instruction. Degenerate forward primers were designed via a combination of CODEHOP [128, 129] or iCODEHOP [130] program and manual modification to achieve optimum degeneracy. The obtained PCR products amplified from degenerate primers and universal primer UAP/AUAP on a PTC-100 Peltier Thermal Cycler were run on 1.0% agarose gel electrophoresis. The target fragments were excised, purified with Wizard® SV Gel and PCR Clean-up System (Promega, Madison, WI, USA) and cloned with pGEM®-T Easy Vector System (Promega, Madison, WI, USA) into JM109 high-efficiency competent cells, following the manufacturer‟s instruction. The Insert-containing plasmids were sequenced via 1st Base Company‟s service. Specific primers to walk upstream were derived from the newly found partial sequences and primed to 5′ RACE cDNA prepared with SMARTer™ RACE cDNA Amplification Kit (Clontech Takara Biotechnology, Dalian, China) (Figure 2.1). The whole procedure from the PCR reaction to the final sequencing step was repeated to complete the transcripts of interest.

In order to determine intron position and sequence on DNA template, specific primers were designed from obtained cDNA sequences. Forward primers were

derived from the 5′ UTR or ER signal peptide, and reverse primers were obtained from 3′ UTR or peptide domain (Figure 2.2).

A typical PCR setting on a PTC-100 Peltier Thermal Cycler was as follows: initial denaturation at 94°C for 5 min, main amplification for 35 cycles (denaturation at 94°C for 30 sec, annealing at Tm - 3°C for 30 sec, elongation at 72°C for 45 – 60 sec), final elongation at 72°C for 10 min, and cooling to 4°C. PCR products were tested on a 1.2% agarose electrophoresis, purified, cloned into pGEM-T Easy Vector, transformed into JM109 high-efficiency competent cells, and sequenced as previously described.

3.4. Sequence Analysis

Sequencing results were analyzed using the BioEdit software. ExPasy translate tool (http://www.expasy.ch/tools/dna.html) was used to predict the amino acid sequences of the clones. The ER signal peptides were predicted with SIGNALP 4.1 server [131]. The nucleotide sequences reported for jS1 and jS2 have been deposited in GeneBank database under accession numbers KT438682 and KT438683, respectively.

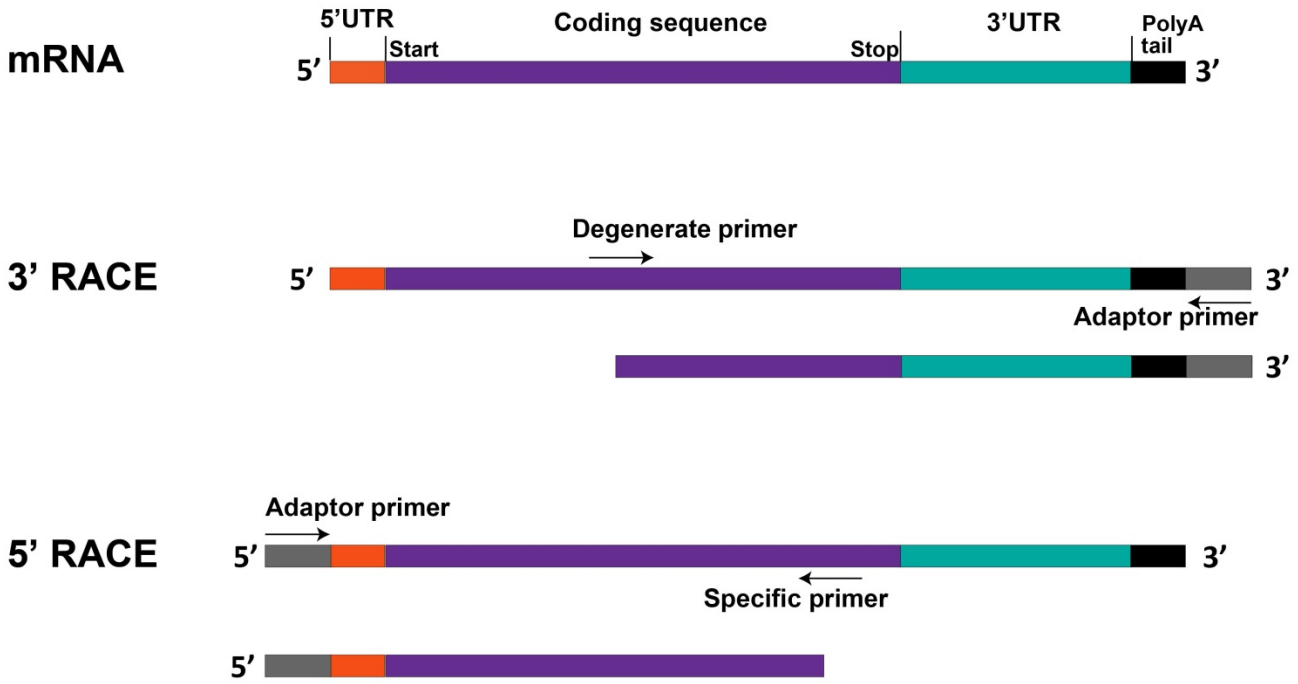


Figure 2.1. Strategy for genetic cloning at mRNA level. The general organization of mRNA is depicted, starting with 5′ untranslated region (5′ UTR), continuing with the coding sequence, and ending with 3′ untranslated region (3′ UTR) to which a poly A tail is added at the end of transcription process. The 3′ RACE cDNA is the template on which a combination of a degenerate primer and an adaptor were employed to first get the 3′ end partial sequence. The newly found segment became the base to design a specific primer to walk upstream 5′ RACE cDNA in order to achieve a complete transcript.

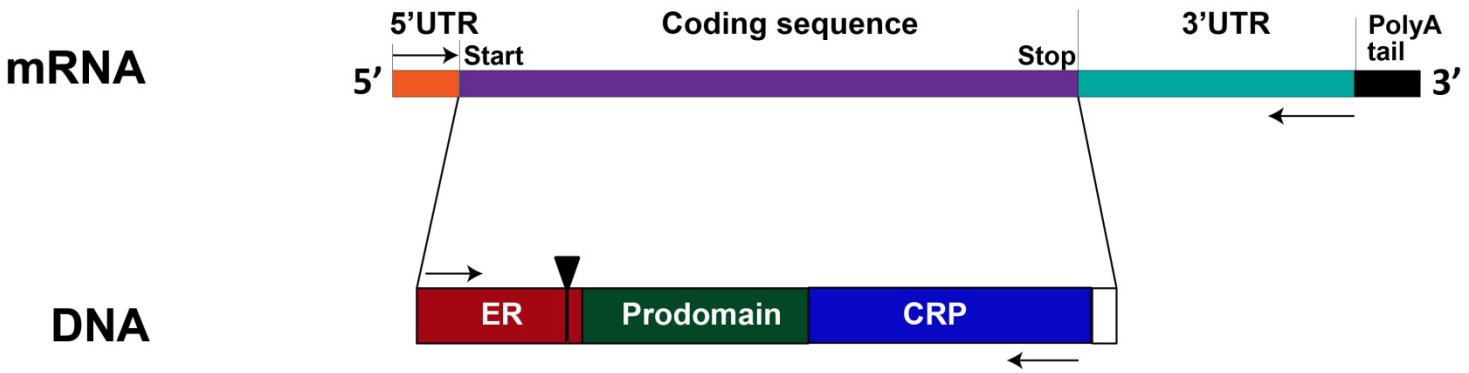


Figure 2.2. Strategy for genetic cloning at DNA level. To determine intron locations and sequences in the open reading frame (ORF), a pair of specific primers, including a forward primer designed either in 5′ UTR or ER region and a reverse primer in 3′ UTR or CRP domain, are utilized in a polymerase chain reaction. Intron information is later deduced from the difference between DNA and mRNA levels.

3.5. High-throughput Transcriptome Sequencing

The sequencing and analysis were performed by Macrogen Inn (Seoul, Korea). Briefly, poly (A) mRNA was isolated from total RNA using magnetic oligo (dT) beads. Following purification, the mRNA is fragmented into small pieces using divalent cations under elevated temperature. The cleaved RNA fragments are copied into first strand cDNA using reverse transcriptase and random primers. This is followed by the second strand cDNA synthesis using DNA Polymerase I and RNase H. These cDNA fragments then go through an end repair process, the addition of a single 'A' base, and then ligation of the adapters. The products are then purified and enriched with PCR to create the final cDNA library. cDNA library was examined on an Aligent Bio analyzer High Sensitivity DNA chip to assess the library quality. Finally, the library was sequenced using Illumina HiSeq™ 2000. Raw data were analyzed and assembled by Macrogen using Trinity software (2011-11-26 ver).

4. Proteomics

4.1. Screening of Oleaceae Family Plants

Leaves of seven different plants from the Oleaceae family (*Osmanthus fragrans, J. multiflorum, J. sambac, J. biflorum*, *J. angustifolium, J. rex*, and *Nyctanthes arbor-tristis;* 300 mg of each) were homogenized in 900 µL of 50% EtOH and centrifuged at 1000g for 10 min at 4 °C. The extracts were diluted twice in order to reduce the percentage of EtOH.

4.2. Isolation and Purification of the CRPs from *J. sambac*

Fresh *J. sambac* leaves (4 kg) or flowers (1kg) were homogenized and extracted twice in 50% (v/v) EtOH. After centrifugation (8500g, 10 min, 4 °C), the supernatant was filtered through 1-µm and 0.45-µm pore-size filters, diluted to 20% in EtOH, and purified in a C_{18} flash column (Grace Davison, Columbia, MD, USA) by washing with 20% EtOH and eluting with 70% EtOH.

For purification of the most abundant jasmintides several dimensions of SCX- and RP-HPLC were performed. Removal of small molecules by SCX-HPLC was conducted with a linear gradient from 0-100% with buffer A (20% MeCN and 10 mM phosphate buffer; pH 3), and buffer B (20% MeCN, 10 mM phosphate buffer, and 1M NaCl; pH 3). CRP-containing fractions were subsequently purified by RP-HPLC using a linear gradient from 25-55% with buffer A (0.1% TFA) and buffer B (0.1% TFA in MeCN).

4.3. Reduction and Alkylation

The number of disulfide bonds in each peptide was determined from the mass difference after reduction-alkylation treatment of crude extracts. Reduction and alkylation were performed as previously described [43]. Briefly, isolated peptides were incubated at 37 °C with 20 mM DTT in 20 mM Tris-HCl buffer (pH 7.2) for 1 h. Subsequently, peptides were alkylated by incubating the reduced extract with 40 mM IAM for 45 min in the dark.

4.4. Sequence Characterization of Jasmintide jS1

Purified jS1 (40 µg) dissolved in 100 µL of 50 mM NH_4HCO_3 buffer (pH 7.8) was reduced in 20 mM DTT for 1 h at 37 °C. Jasmintide jS**1** was digested with trypsin, chymotrypsin or EndoGLu-C at 1:10 (protein:enzyme). Enzymatic digestions were performed at 37 °C for 10 min, and then sequenced by MALDI-TOF MS/MS as previously described [43]. Ile/Leu and Lys/Gln were assigned based on gene

sequence.

4.5. Disulfide Mapping

Partial Reduction and alkylation: Jasmintide jS1 (150 µg) was partially reduced in 100 mM citrate buffer (pH 3.0) with 15 mM tris(2-carboxyethyl)phosphine at 37 °C for 30 min. Subsequently, 30 mM NEM was added and incubated at room temperature for 2 h. The reaction was quenched by immediate injection of sample into analytical RP-HPLC.

HPLC Fractionation: Reduced and alkylated species were separated with a linear gradient of 20-60% buffer B (0.1% TFA in MeCN). Fractions were collected and analyzed by MALDI-TOF MS to verify NEM-alkylated intermediate species. The two-NEM-alkylated intermediate specie **2SS** was lyophilized and redissolved in 2 M urea. For determining the other two disulfide connectivities, **2SS** was enzymatically digested with trypsin for 2 h at 37 °C. Disulfide bond-containing digested fragments were fully reduced with 20 mM DTT at 37 °C for 30 min and alkylated with 40 mM IAM for 1 h at room temperature.

Confirmation by LC-MS/MS: The sequence of reduced and alkylated digested peptides was confirmed by LC-MS/MS. LC separation was performed with 0.1% formic acid (eluent A) and 90% MeCN in 0.1% formic acid (eluent B), respectively, in a 60-min gradient starting at 3% eluent B for 1 min, 3-35% eluent B over 47 min, 35-50% eluent B over 4 min, 50-80% eluent B over 6 s, 80% eluent B for 78 s, and then reverted to initial conditions in 6 s and maintained isocratically for 6.5 min. Spray was generated with a Michrom Thermo Captive Spray nanoelectrospray ion source (Bruker-Michrom Inc, Auburn, CA, USA) with a source voltage of 1.5 kV and a capillary temperature of 250 °C. Data was collected in positive ion mode using LTQ Tune Plus software (Thermo Scientific Inc.) alternating between full scan-MS (350–1600 *m/z*, 60.000 resolution at 400 *m/z*, 1 microscan per spectrum) and full scan-MS/MS (150–2000 *m/z*, 15.000 resolution at 400 *m/z*, 1 microscan averaged per spectrum). High-energy

collisional dissociation fragmentation was performed for the 10 most intense ions with a 500-count threshold using 32% normalized collision energy per spectrum and a 2-Da isolation window. Automatic gain control for full scan-MS and MS/MS was set to 1×10^6 ions.

4.6. Data Analysis

Data from LC-MS/MS were analyzed using PEAKS Studio [132] (version 7.0, Bioinformatics Solutions, Waterloo, Canada). A precursor mass tolerance of 10 ppm and a fragment mass error tolerance of 0.05 Da were applied.

4.7. Sample Preparation for Characterization of Jasmintides at Proteomic Level

J. sambac was collected from NTU herb garden. Leaves, flowers and roots (800 mg) were washed and individually blended with 50% EtOH (1:3 w/v). After centrifugation at 1000 rpm for 10 min, the supernatant was collected and diluted to 20% in ethanol. SCX-HPLC was run at 0% and 100% until no peak was observed. The sample was then desalted in C_{18} Sep-pack column by washing with 5% ACN and eluted with 80% ACN (1 mL). After desalting, the sample was concentrated to 200 uL using Speed Vac.

Reduction and alkylation was performed as previously described by Aida *et al*,. [133]. Briefly, in fractionated jS-leaves, jS-flowers and jS-root 30 mM DTT and 60 mM bromoethylamine (BrEA) in 0.2 M Tris buffer (pH 8.6) was added and incubated at 55°C for 1 h. The reduced, alkylated peptide samples were desalted and peptide solutions were stored at -20 °C.

Sample purification, reduction and alkylation reactions for each sample were monitored by MALDI-TOF mass spectrometry using an ABI 4800 apparatus, with reflectron acquisition mode. Peptides were focused in the range of 1000-6000 Da mass. Peptide samples were desalted using C_{18} zip-tip before spotting. Each

MALDI spot contained 0.5 µL desalted peptide solution and 0.5 µL matrix solution (α-cyano-4-hydroxycinnamic acid (CHCA) in 75% ACN with 0.1% TFA).

4.8. LC-MS/MS Spectrometry

The samples of jasmine leaves, flower and roots were injected to LC-MS/MS. LC separation was performed with 0.1% formic acid (eluent A) and 90% MeCN in 0.1% formic acid (eluent B), respectively, in a 60-min gradient starting at 3% eluent B for 1 min, 3– 35% eluent B over 47 min, 35–50% eluent B over 4 min, 50–80% eluent B over 6 s, 80% eluent B for 78 s, and then reverted to initial conditions in 6 s and maintained isocratically for 6.5 min. Spray was generated with a Michrom Thermo Captive Spray nanoelectrospray ion source (Bruker-Michrom Inc, Auburn, CA, USA) with a source voltage of 1.5 kV and a capillary temperature of 250 °C. Data was collected in positive ion mode using LTQ Tune Plus software (Thermo Scientific Inc.) alternating between full scan-MS (350–1600 *m/z*, 60.000 resolution at 400 *m/z*, 1 microscan per spectrum) and full scan-MS/MS (150–2000 *m/z*, 15.000 resolution at 400 *m/z*, 1 microscan averaged per spectrum). High-energy collisional dissociation fragmentation was performed for the 10 most intense ions with a 500-count threshold using 32% normalized collision energy per spectrum and a 2-Da isolation window. Automatic gain control for full scan-MS and MS/MS was set to 1×10^6 ions.

4.9. Data Analysis (Database construction)

The transcriptome of *J. sambac* was analyzed using the in-house built software Protein Analyzer 1.6, and open reading frames (ORFs) containing at least six Cys residues were exported. The database search and post-translational modifications analysis were assessed in PEAKS studio (version 7.0, Bioinformatics Solutions, Waterloo, Canada) using 10 ppm MS and 0.05 Da MS/MS tolerances. The PEAKS PTM function [132] was used to search our data for >650 PTMs and mutations listed in the Unimod database [134]. We also used

PEAKS studio to apply the SPIDER algorithm to the identification of peptide mutations and homology [135]. All identified sequences were validated manually.

4.10. Spectrophotometric Determination of Protein Concentration

The concentrations of purified peptides were calculated from the Beer-Lambert law:

$A = \varepsilon \cdot l \cdot c$

where A: the absorbance at 280 nm of peptide solution in miliQ water measured on the Nanophotometer (Implemen, Germany)

ε : molar absorption coefficient ($M^{-1}cm^{-1}$)

l : cell path length (cm)

The thereotical ε value of a protein at 280 nm was calculated as follows [159]:

$\varepsilon 280 = (5500 \cdot n\text{Trp}) + (1490 \cdot n\text{Tyr}) + (125 \cdot n\text{SS})$

where

*n*Trp : the number of Trp residues

*n*Tyr : the number of Tyr residues

*n*SS : the number of disulfide bonds in the protein

4.11. NMR Spectroscopy

NMR samples were prepared by dissolving 1 mg of lyophilized peptide in 500 µL 90% H_2O/10 % D_2O or 99.99% D_2O at pH 7.0 in a 20 mM Na_3PO_4 buffer containing 50 mM NaCl and 0.01% NaN_3. The concentration of the NMR sample was approximately 0.7-1 mM. Homonuclear 2D NOESY experiments were performed with mixing times of 200 and 300 ms in collecting NOE spectra for jS1 and jS3 [136, 137]. COSY [138] data were also recorded in both H_2O and D_2O solutions with a mixing time of 78 ms using MLEV17 spin-lock pulses [139]. Vicinal coupling constants were determined using DQF COSY [140] and ^1H NMR experiments. All 2D-NMR data were recorded in the phase-sensitive model using the time-proportional phase increment method [141] with 2048 data points in the t2 domain and 512 points in the t1 domain. Slowly exchanging amide protons were identified by lyophilization of a fully protonated sample in H_2O solution to dryness, resuspension in a 99.99% D_2O solution, and immediate acquisition of a series of 1D-spectra. The strong solvent resonance was suppressed by water-gated pulse sequences [142] or excitation sculpting [143] combined with pulsed-field gradients. All NMR data were processed using a Bruker TOPSPIN 2.1 (Bruker Instruments) or NMRPipe [144] programs on a Linux workstation and analyzed by using Sparky 3.12 software. The DQF-COSY spectra were processed to 8192 x 1024 data matrices to obtain a maximum digital resolution for coupling constant measurements. The proton chemical shifts were referenced to internal sodium-4,4-dimethyl-4-silapentane-1-sulfonate.

4.12. Structure Calculations

Solution structures of jS1 and jS3 were calculated by hybrid distance geometry and simulated annealing simulated annealing protocols in torsion angle space with CNS 1.2 [145]. The three disulfide bonds were restrained in accordance with the disulfide bonding patterns based on the observed Hβ-Hβ NOEs from NOESY spectra by generation of covalent disulfide linkages during the initial molecular topology file generation stage using a CNS script. A total of 656 distance and 22

torsion angle constraints were used for structure calculations. NOE distance restraints were classified strong (1.8-3.0 Å), medium (1.8-3.5 Å), weak (1.8-5.0 Å), or very weak (1.8-6.0 Å) based on the intensities derived from NOESY spectra recorded in 90% H_2O/10% D_2O or 99.99% D_2O at 25 °C. Corrections for pseudo atom representations were used for non stereospecifically-assigned methylene, methyl group, and aromatic ring protons [141]. Backbone dihedral angle restraints were derived from $^3J_{HN-H\alpha}$ coupling constants in DQF-COSY or ^1H NMR spectra in H_2O solution [146, 147]. Backbone dihedral restraints were used as -55° ± 45° ($^3J_{HN-H\alpha}$ < 6 Hz) and -120° ± 50° ($^3J_{HN-H\alpha}$ > 8 Hz). Hydrogen bond donors were identified from proton-deuterium exchange experiments and hydrogen bond acceptors were determined at the preliminary structure calculation stage. The 200 starting structures were generated and refined using a hybrid distance geometry-simulated annealing protocol by the CNS 1.2 program [148-150]. Finally, 20 final structures were selected by their total energy values for display and structural analysis. MOLMOL [151] and PyMOL [152] programs were used for structure visualization, and PROCHECK-NMR [153] were used for structure validation. The jS1 solution structure solved for 20 ensembles is available in the Protein Data Bank under accession number 2N5Q.

5. Stability Assay

5.1. Heat Stability

Peptides were heated in boiling H_2O for 1 h and then subjected to UPLC. A linear peptide (RV) of 14 residues (RLYRRGRLYRRNHV) was used as the control. Peaks were collected from UPLC and monitored by MALDI-TOF MS. Peak areas in UPLC were compared to evaluate stability.

5.2. Proteolytic Stability Assay

Peptides were incubated with or without trypsin in 100 mM NH_4HCO_3 buffer (pH 7.8) and pepsin in 100 mM sodium citrate buffer (pH 2.5) at a final peptide:enzyme ratio of 20:1 (w/w) at 37 °C for 8 h. A nine-residue (KL) peptide (KRPPGFSPL) treated in the same manner was used as a positive control. Treated and control peptides were subjected to UPLC at intervals, and collected peaks were monitored by MALDI-TOF.

In the carboxypeptidase A stability assay, the peptides were incubated with or without enzyme [50:1 (w/w)] in Tris buffer (50 mM Tris and 100 mM NaCl; pH 7.5) at 37 °C for up to 6 h. A 14-residue RV peptide (RLYRRGRLYRRNHV) was subject to the same treatment and served as the control. Collected peaks were monitored by MALDI-TOF. Peak areas in UPLC were compared to evaluate their stability.

5.3. Serum Stability Assay

Purified peptides 0.2 mg were incubated with 25% Human male serum AB type in phenol red-free Dulbecco's modified Eagle medium (DMEM) at 37°C for 24 h. Peptides not subjected to serum treatment functioned as the controls. Aliquots at particular intervals were collected and quenched with 96% EtOH. Samples were kept at 4°C for 15 min, and centrifuged at 18000 rpm for 10 min. Collected supernatants were analyzed using UPLC and MALDI-TOF. DALK peptide was subjected to the same treatment, serving as the positive control for jasmintides.

6. Bioassays

6.1. Cytotoxicity Assay

The cytotoxicity of peptides on Vero, HUVEC-CS, Raw 264.7, and THP-1 was measured using standard PrestoBlue TM reagent (Invitrogen, Carlsbad, California, USA). Briefly, cells were seeded in 96-well plate at a concentration of 8 X 10^3 cells per well in DMEM or RPMI 1640 containing 10% FBS and 1% penicillin and streptomycin and incubated at 37 °C with 5% CO_2 in the air. After 24 h, varying concentrations of peptides ranging 0.01-100 µM were added in triplicate and then incubated for 24 h. Presto Blue (10 µL) was added to each well and maintained for 1 h at 37 °C. The absorbance was read using the Tecan plate reader at an absorbance of 570 nm. 1% Triton X-100 and PBS served as positive and negative control, respectively.

6.2. Hemolysis Assay

Fresh blood type +O was donated by a healthy volunteer. To obtain erythrocytes, blood was centrifuged (1000 rpm, 15 min). The pellet was isolated and washed with phosphate buffer saline (PBS) thrice. A stock dispersion of erythrocytes was prepared via 100X dilution in fresh PBS buffer. Various concentrations of peptides were added, with an equal volume of treated blood sample into a 96-well plate. After incubation at 37 °C for 4 h, the plate was centrifuged (1000 rpm, 5 min) and supernatants were transferred into a new plate. After transferring supernatants into a new plate, absorbance was measured at 415 nm using Tecan Infinite® M200 plate reader. 0.1% Triton X-100 solution and PBS solution served as the negative and positive controls, respectively.

6.3. Antimicrobial Activity using Radial Diffusion Assay

Two bacterial strains and one fungal strain from ATCC were used including *Staphylococcus aureus* ATCC 12600 *Escherichia coli* ATCC 25922 and *Candida albicans* 11006. All strains were cultured in trypticase soy broth. The antimicrobial activity of peptides was examined using radial diffusion assay as described previously under low salt condition (10 mM Na_3PO_4) [154]. D4R was used as positive control [155].

6.4. Carboxypeptidase Inhibition Assay

The assay was performed as reported previously [104]. Briefly, various concentrations of peptides (0.1 to 100 µM) were added to the 96 well plate with 0.1 mM of N-(4-methoxyphenylazoformyl) phenylalanine substrate. The reaction was started by adding 25 nM of bovine carboxypeptidase A (bCPA). The substrate hydrolysis was monitored at 350 nm in 50 mM Tris HCL, 1M NaCl buffer using end point measurement, quenched with 1 M of Na_2CO_3 after 5 min.

6.5. Determination of Trypsin Inhibition Activity

To test the trypsin inhibition effect of peptides, we conducted a trypsin inhibition assay in 96-well plate with a total volume of 100 µl per well. Each well contained 94 µl of 0.05 M Tris-HCl containing 0.025 M $CaCl_2$ (pH 8) and 1 µl 0.5 µg/ml bovine β-trypsin (T1426, Sigma). SFTI-1, TIRF7 and TINF7 were seriously diluted to a concentration of 0.01-1 µM. Then peptide and trypsin was incubated at 25 °C for 15 min. Just before monitoring by the spectrometer, 5 µl of 20 mM BAPNA was added to each well by a multi-channel pipette and the mixture was mixed well quickly. The process of the reaction was monitored by an ELISA microplate reader (Tecan-2) with an excitation wavelength of 405 nm. IC_{50} values were determined using Graph Pad Prism 5.

6.6. Platelet Aggregation Assay

To obtain platelets, fresh A+ blood from a volunteer student was centrifuged at 1000 rpm for 15 min. Supernatant was collected and centrifuged (2310 g, 4 min), then washed with sequestine buffer (4.69 g/L NaH2PO4, 9 g/L NaCl, 3.31 g/L Na2EDTA, 0.1% w/v BSA) twice and centrifuged. Platelets (50x106 platelets/mL) were re-suspended in HEPES buffer (7.714 g/L NaCl, 0.447 g/L KCl, 0.2465 g/L MgSO4, 0.163 g/L KH_2PO_4, 4.766 g/L HEPES, 0.9 g/L glucose). In a 96-well plate, platelets (100 µl) with $CaCl_2$ (1.8 mM), and peptides (1 µM) were incubated. Collagen (1 mg/mL) will serve as the positive control. Absorbance at 405 nm was measured every minute for 1 h.

6.7. Collagen Synthesis Assay

With reference to the paper Park et al., [156] collagen synthesis in Human Foreskin Fibroblast (HFF) cells was performed according to the protocols, to investigate the effect of jS1, jS3 and jS4 on collagen synthesis.

6.8. Angiogenesis Assay (Endothelial Tube Formation Assay)

BD Matrigel^{TM} Basement Membrane Matrix acted as a basement membrane for effective attachment of Huvec-CS cells (Biosciences, Singapore).The experiment was conducted in triplicate, using different concentrations of peptide stock. 18 µl of Matrigel was added in µ-slide, upon solidification, $3.5x10^3$ Huvec-CS cells were seeded in 45 ul of DMEM, and 5 µl of peptide was added. The negative control used was DMEM with 10% Fetal Bovine Serum (FBS). The positive control was Apelin. After incubation at 37°C for 6 h, samples were observed under Nikon Eclipse Ti-s inverted microscope at 4X magnification. Microscopic Images were analyzed using ImageJ software [157] and the number of junctions in each well were deduced.

6.9. Migration Assay

Culture-inserts (Ibidi, Martinsried, Germany) were placed in a 24-well plate, Huvec-CS cells ($5x10^5$ cells/mL, 70 µl) were seeded into each well of the culture-inserts, and incubated at 37 °C overnight. Following, culture-inserts were removed using sterile tweezers and various concentrations of peptides jS1, jS3 and jS4 (0.01 µM, 0.1 µM, 1 µM, 10 µM) in DMEM (300 µl) were added into corresponding wells, following incubation at 37°C. DMEM served as the negative control. The gap was monitored under Nikon Eclipse Ti-s inverted microscope at 4X magnification. Images were taken at 4.5 h and analyzed using T-Scratch software [158].

6.10. Cytokines Release Assay

6.10.1. Preparation of Cells and Treatment with Peptides

THP-1 Cells were counted under the light microscope using a hematocytometer. Dilutions of cell culture were made with the same medium to achieve 10^6 cells/mL per well in a 24-well plate. The cells were either treated with 0.1 µg/mL lipopolysaccharide (LPS) only, with different concentrations of peptides or both LPS and peptides. The cells were incubated for 6 h at 37 °C. The supernatant was collected by centrifugation at 1000 rpm for 10 min and stored at -20°C until assayed.

6.10.2. Enzyme-Linked Immunosorbent Assay (ELISA)

This assay was used to monitor the levels of secretion of cytokines by the THP-1 cells after the treatments. Levels of secreted Interleukin-6 (IL-6), Interleukin-1β (IL-1β) and Tumor necrosis factor-alpha (TNF-α) were assessed. The experimental procedures were executed according to the protocol described in the manufacturer's information booklet enclosed in the ELISA MAX™ Deluxe Sets (Biolegend, USA).

6.11. Nitric Oxide Release Assay

Raw 264.7 cells were grown in DMEM with 10% FBS and 1% penicillin and streptomycin at 37 °C in 5% CO_2. The cells were plated in to the cell culture plates and incubated for 24 h. Next day the cells were treated with 0.1 µg/mL LPS, 100 nM SP and peptides (0µM, 1µM, 10µM and 20µM) for 24 h. cells treated with 100 nM SP alone or with 0.1 µg/ mL LPS or both served as controls. NO, quantified by measuring its oxidation product nitrite, was determined by the Griess reaction by a micro plate assay method [159]. 50 µl of the macrophages supernatant were removed and incubated with an equal volume of Griess reagent (1% sulfanilamide, 0.1% *N*-1 naphthyleneamine dihydrochloride, 2.5% phosphoric acid) at room temperature for 10 min in the dark. The absorbance at 540 nm was determined with a micro plate reader. Concentrations were calculated from a standard sodium nitrite curve.

6.12. Effect of Jasmintides on the Growth of *Tenebrio molitor* Larvae

Same weight larvae of *T. molitor* were sorted by the weighing method. The body weight of larvae was between 40.0-43.06mg. The test diet was supplemented with jasmintides (1mg/g of diet). The control diet contained only the solvent in place of the inhibitor. Oats flakes were used as artificial diet. Four larvae were added to each 11 mm dish with 200 mg of diet and weight gain was recorded with the interval of 3 days. In total 20 larvae were used for each concentration. There were 5 -replicates for each test. The diet was replaced as required to provide a continuous supply.

6.13. Antifeedant No-Choice Assay

Antifeedant activity of jasmintides was assayed using larvae of same weight as done in larval growth bioassay using the same test and controlled diet. The oat flakes weighed and placed in tissue culture dish dishes (11 mm dia) together with 4 larvae. Five replicates of tests were done. The dishes were transferred into a rearing chamber and kept at 29 ± 1°C in the dark. After 3-days of feeding, the remaining uneaten oat flakes were re-weighed and the mean weight of food eaten was calculated.

6.14. Anti-viral Assay

Vero cells (8×10^4 cells) were seeded to 48-well plates and incubated for 24 h. Various concentrations of jamintides (10 μM, 1 μM, 0.1 μM, 0.05 μM, 0.01 μM) were diluted in DMEM and mixed with IBV (equivalent amount to MOI=0.1) and incubated at 37 °C for 30 min. The mixtures were then inoculated onto the cell monolayers for 2 h at 37 °C with occasional shaking. The supernatants were removed and the media were replaced with DMEM containing peptides. The cells treated with virus only served as the negative control. After 24-hour incubation, cytopathic effect was observed under a light microscope to ensure complete infection of the negative control. The virus supernatants were collected for plaque assay. Each concentration of jasmintide was assayed in triplicate.

6.14.1. Plaque assay

H1299 cells grown to confluency in 6-well plates were washed once with PBS and infected with 250 μl of serial dilutions of the virus stocks in DMEM at 37 °C for 2 h with occasional shaking. The inoculum was aspirated, the cells washed with PBS and overlayed with 2 mL of 0.3% agarose in water. The plates were incubated in humidified 5% CO2 incubator at 37 °C for 48 h. The cells were subsequently fixed with 4% formaldehyde at RT for 0.5 h and virus plaques were visualized and counted by staining with 0.2% crystal violet in 20% MeOH.

6.14.2. Determination of TCID50

Vero cells ($1.5x10^4$ cells/well) were plated onto the 96-well plate 1 day before infection. Cell cultures were washed once with PBS and infected with serial dilutions of virus samples in DMEM. 10 replicates were prepared for each dilution. Non-infected cells dispensed with DMEM served as controls. The cultures were incubated for 3-5 days at 37°C till the cytopathic effects observed. TCID50 titers were calculated using the Excel spreadsheet from Yale School of Medicine available at www.med.yale.edu/micropath/pdf/Infectivity%20calculator.xls.

7. Phylogenetic Tree Construction

Full precursor sequences of CRP were pulled from the database PhytAMP [160]. The sequence without signal peptides and α-hairpinins were excluded. Sequences were aligned using Clustral Omega multiple alignment tool [161] and the phylogenetic tree was generated and visualized using Interactive tree of life (ITOL) [19].

8. Chemical Synthesis

8.1. Synthesis of Thiolethylamidobutyl (TEBA) Resin using Cl-Trt (2-Cl) Resin

Cl-Trt (2-Cl) resin **1** (1 g, 1.14 mmol) was swollen for half an hour in DCM (10 ml) and then filtered and washed with DCM three times for 1 min. (2-butylamino) Ethanethiol **2** (0.8 mmol, 118.3 µl) in DCM (20 ml) was added, and the suspension was shaken for 1 h at room temperature. DIEA (3.42 mmol, 595.68 µl) and MeOH (3.42 mmol, 138.22 µl) was added to the reaction mixture and shaken for 10 min to quench the unreacted resin. The resin was washed with DCM, DMF, MeOH and DMF three times, respectively to give TEBA resin **3**. With this resin **3** Fmoc- Xaa-OH was coupled using a mixture of Fmoc-AA, HOBt, and DIC (4/4/4 eq) in DMF. The suspension was shaken for 1 h. This step was

repeated twice to give Fmoc-Xaa-TEBA resin **4**. It was washed with DMF 6 times. The coupling reaction was monitored by acetaldehyde/chloranil test for detection of the secondary amino group. The resin was washed with DCM and diethyl ether three times for 1 min and dried under vacuum for later use. The substitution of resin **4** was calculated by weight.

Peptides **TI-RF7, TI-NF7, RF7-T** and **NF7-T** were synthesized using Fmoc-Phe-TEBA **4a** and SFTI-1 using Fmoc-Arg-TEBA **4b.**

8.2. General Method for SPPS Synthesis

Initially, the desired resin was swelled in DMF for 30 min. Amino acid residue was coupled by adding a pre-activated solution which contains Fmoc AA, HBTU, HOBt and DIEA (4/4/4/6 eq) in DMF. Coupling was done for 30 min. After that resin was washed three times with DMF. Fmoc-D-Arg-OH was coupled using Fmoc AA, DIC and HOAt (4/4/4 eq) for an hour three times. De-protection was done using 20% piperidine in DMF three times for 3 min. After that resin was washed again with DMF three times. The subsequent sequence was synthesized using the same steps. Finally, after complete synthesis and de-protection of Fmoc protecting group it was washed three times with DMF, DCM and diethyl ether and dried under vacuum. The completion of the reaction was monitored by acetaldehyde and chloranil test (2% acetaldehyde, 2% chloranil in DMF, 5 min at rt) for secondary amines whereas TNBS test (20ul of picryl sulfonic acid and 20ul of 20% DIEA in DMF) for primary amine. The peptide was cleaved using a cocktail mixture of TFA/TIS/H_2O/thioanisole (89/5/5/1 v/v). After 2 h the reaction mixture was added drop by drop in chilled diethyl ether and centrifuged for 10 min at 4000 rpm. This step was repeated three times. The precipitates of crude peptide were dried under vacuum. Two different ester forms of peptide were obtained N-form **5** and S-ester form **6** (Figure 4.8). To confirm the presence of peptide analytical HPLC was run at linear gradient 10-60 % MeCN/0.1% TFA for 25 min. Peaks were collected and confirmed by MALDI-TOF.

Peptides **RF7** and **NF7** were synthesized using Rink amide resin. For peptides **TI-RF7, TI-NF7, RF7-T, NF7-T** and **SFTI-1** Fmoc-Xaa-TEBA (**4a, 4b)** were synthesized manually. The remaining peptide sequences were synthesized by automated microwave peptide synthesizer (CEM, Liberty1) using a solution of 0.5 M PyBOP in DMF as activator and a solution of 55% DMF, 35 % DIEA and 10% DCM as activator base for coupling and 20% piperidine with 0.1 M HOBt in DMF for deprotection.

8.3. Thioesterification and Cyclization

Compound **5** and **6** were dissolved in 0.1 M sodium phosphate buffer (pH 3) to a final peptide concentration of 2.5 mM. To this solution, MESNa (50 eq) was added and the reaction mixture was incubated at 40°C. Reaction completion was monitored by injecting 10ul sample in analytical RP-HPLC using linear gradient 10-60% of ACN/ 0.1%TFA for 25 min at different time intervals. Peaks were collected and confirmed by MALDI-TOF. Along with thioester **7** we got thiolactones **8**.

To this solution, TCEP (10 eq) was added and pH was adjusted to 7 with 2 N NaOH. The reaction mixture was shaken for 1 h and then quenched with 1 M HCL. It was purified by RP-HPLC using preparative column using linear gradient between 10-40% of ACN/ 0.1%TFA for 80 min to get cyclized compound **9**. MS of compound **9** was confirmed by MALDI-TOF**.**

Compound **9** was oxidized using DMSO. It was dissolved in 0.1 M sodium phosphate buffer pH 7.5 to make the final concentration of 1 mM. To that solution, 10 % DMSO was added and incubated at room temperature. After that, it was purified by RP-HPLC using preparative column get the compound **10**.

8.4. Purification

Crude peptide obtained after chemical synthesis, cyclization and oxidation were purified with HPLC SHIMADZU (Prominence LC2AD) using preparative column (Grace Vydac Prosphere 300 C18 300A, 250 x22 mm, 5-micrometer particle size, flow rate 6ml/min) at a gradient of 10-40% of buffer B for 80 min. Buffer A was 0.1% TFA in water and buffer B was 0.1% TFA in ACN. Eluents were monitored at 220 nm and 260 nm. The purity of peptide was examined by analytical RP-HPLC on a Phemnomenex Aeris PEPTIDE XB-C18, 250 x 4.6 mm, 3.6 µ column, with a flow rate of 1.0 ml/min at 50° C. Masses were determined by MALDI-TOF.

Chapter 3

Discovery and Characterization of a New CRP Family with Novel Disulfide Connectivity from *Jasminum sambac*

1. Introduction

Jasminum sambac (L.) Ait. belongs to the Oleaceae family and is commonly known as Motia or Arabian jasmine. It is known to originate from India and then in the 3rd Century *J. sambac* was brought to tropical areas of South-East Asia. After that, it is widely cultivated throughout the tropical and sub-tropical parts of the world [162]. Jasmine is a small shrub with young pubescent branches, broadly ovate or elliptic, opposite leaves and fragrant white flowers [163] (Figure 3.1). Since ancient times, jasmine has been thought as “queen of flowers”. The name “Jasmine” is derived from Persian word “Yasmin” which means fragrant flower. In India and Indonesia, flowers are well known for making garland and bouquets for honored guests. Jasmine flowers are a rich source of essential oil which is used as a flavoring. In India, flowers are used in religious offering as a symbol of divine hope. In China, the fragrant flowers are processed and used for the preparation of jasmine tea [164].

Traditionally, flowers are used as a tonic to the brain, allay fever, purgative, anti-emetic, for skin diseases, cure itching sensation, leprosy and ulcers [163]. Aroma-therapists find it relaxing herb for the dry and sensitive skin. Dried jasmine flowers decoction is applied as an eyewash for swollen and painful eyes [163, 165]. Jasmine leaves are reported to contain antipyretic properties, whereas their juice is useful for the treatment of ulcers. Leaves extract exhibit analgesic and anti-inflammatory activity in rats [166, 167].

Its roots have been used as a purgative, analgesic, and anthelmintic against ringworm and tapeworm [165-167]. In Asia and India, folk medicine practitioners recommend jasmine for liver complaints, dysentery and various types of pain including painful menstruation. Also, jasmine oil is applied externally as calming agent to soothe stress, pain and anxiety. Other medicinal properties are listed in Table 3.1.

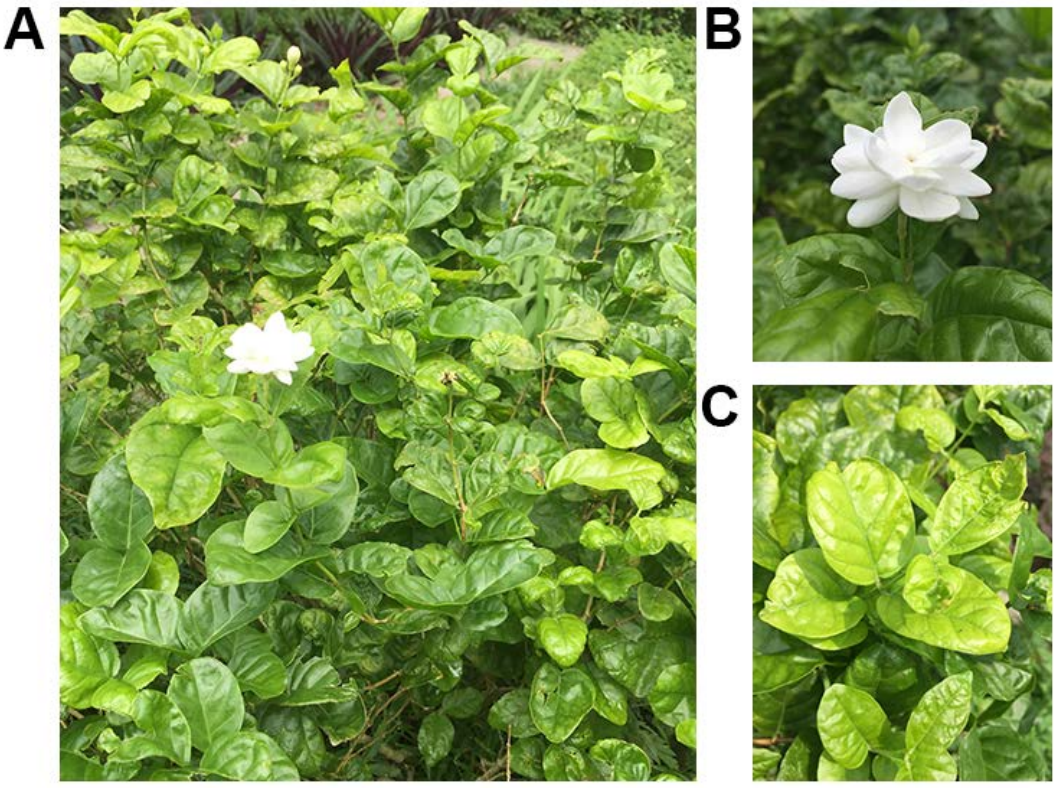


Figure 3.1. *Jasminum sambac* in the herb garden of Nanyang Technological University, Singapore. A) *J. sambac* grows into a small shrub, B) with white flowers and C) with ovate and opposite leaves.

Table 3.1. Bioactivities of different parts of *Jasminum sambac*

Bioactivity	Plant parts	References
Antioxidant	Leaf, flower	[168, 169]
Anticancer	Leaf, flower	[170, 171]
Analgesic	Leaf	[171]
Antiviral	Flower	[172]
Antimicrobial	Leaf, flower and essential oil	[173-175]
Anti-inflammatory	Leaf	[166, 167]
Anti-depressant	Jasmine oil	[176]
Gastroprotective effect	Leaf	[177]
Wound healing	Leaf	[178]
Vasodilation	Flower	[179]
Sedative	Jasmine tea	[180]

Pharmacological and phytochemical studies of jasmine plant have led to the identification of numerous secondary metabolites including a wide spectrum of flavonoids, alkaloids, saponins, terpenoids, steroids and glycosides [165, 181, 182]. Despite these extensive researches, active compounds characterized from jasmine plant are mostly restricted to small molecules. Proteins and peptides also exist in nature along with the primary and secondary metabolites, which also possess therapeutic value. Compared with the small metabolites which are >1 kDa in size, peptides have a molecular weight in between 2-8 kDa. Peptides, especially CRPs with MW between 2-6 kDa, have not been considered as a source of bioactive compounds in traditional medicine, because of their thermal and proteolytic degradation and low bioavailability.

This chapter describes the discovery and characterization of a new CRP family, jasmintides (*Jasminum sambac*; peptide) from Oleaceae family. A combination of genomic, proteomic and transcriptomic approaches was used to identify a suite of jasmintides from all parts of *J. sambac* including leaves, flowers, and roots. Jasmintides with an *N*-terminal pyroglutamic acid residue were shown to be resistant not only to heat and endopeptidase degradation but also to exopeptidases. The precursor of jasmintides contained three domains similar to many CRPs. Importantly; jasmintides display novel disulfide connectivity between CysI-CysV, CysII-CysIV, and CysIII-Cys-VI which is different from other known plant CRPs. The discovery of novel Oleaceae jasmintides, the corresponding precursor sequences, and their characteristics will be discussed in this chapter.

2. Results

2.1. Screening for CRPs from the Oleaceae Family

We investigated the presence of CRPs in Oleaceae family by a small-scale extraction using RP-HPLC. Briefly, we extracted 300 mg of plant material with 50% EtOH and loaded to C_{18} cartridge. The column was prior washed and balanced with 20% acetonitrile and eluted with 70% acetonitrile. The occurrence of putative CRPs was profiled using MALDI-TOF MS. The screening of CRPs was based on a mass range of 2-6 kDa and presence of three disulfide bonds. Mass shift before and after reduction of their disulfide bond with DTT and *S*-alkylation with IAM was observed with an increment of 58 Da for each cysteine [183].

Leaves of seven plants were collected from Singapore Botanic Garden namely *Osmanthus fragrans, J. multiflorum, J. sambac, J. biflorum*, *J. angustifolium, J. rex*, and *Nyctanthes arbor-tristis* and screened by mass spectrometry. Only *J. sambac* showed strong signals in the mass range of 2-6 kDa. The other six plant extracts showed a low abundance of CRPs under our screening conditions. After reduction with DTT and *S*-alkylation with IAM, a mass shift of 348 Da was observed, indicating the presence of putative CRPs in *J. sambac* with three disulfide bonds.

2.2. Isolation of Jasmintides from *J. sambac*

To investigate the presence of CRPs in different tissues of jasmine, we did preliminary screening of various plant parts, including leaf, root, and flower. All samples were extracted with 50% EtOH and extracts were spotted on MALDI plate and analyzed. The resulting mass spectra (Figure 3.2) revealed a group of peaks in the range of 3-4 kDa.

We performed a scale-up extraction of the fresh leaves of *J. sambac* and isolated two putative CRPs, named as jasmintides. The major CRP from leaves with a MW of 3105 Da was named jasmintide 1 (jS1), and the minor CRP of 3181 Da as jasmintide 2 (jS2). The most abundant jasmintide (jS1) was purified sequentially by SCX-HPLC and then by RP-HPLC. The sequence of jS1 was determined by mass spectrometry shown in Table 3.2.

We also extracted fresh flowers of *J. sambac* with 50% EtOH and isolated two putative CRPs, jasmintide jS3 and jS4 with MW of 3199 and 3198 Da. These two jasmintides were purified sequentially by SCX-HPLC and RP-HPLC. Jasmintides jS3 and jS4 differ by 1 Da only but elute at different retention time. The sequences of jS3 and jS4 shown in Table 3.2 were determined from transcriptome data, described in later part.

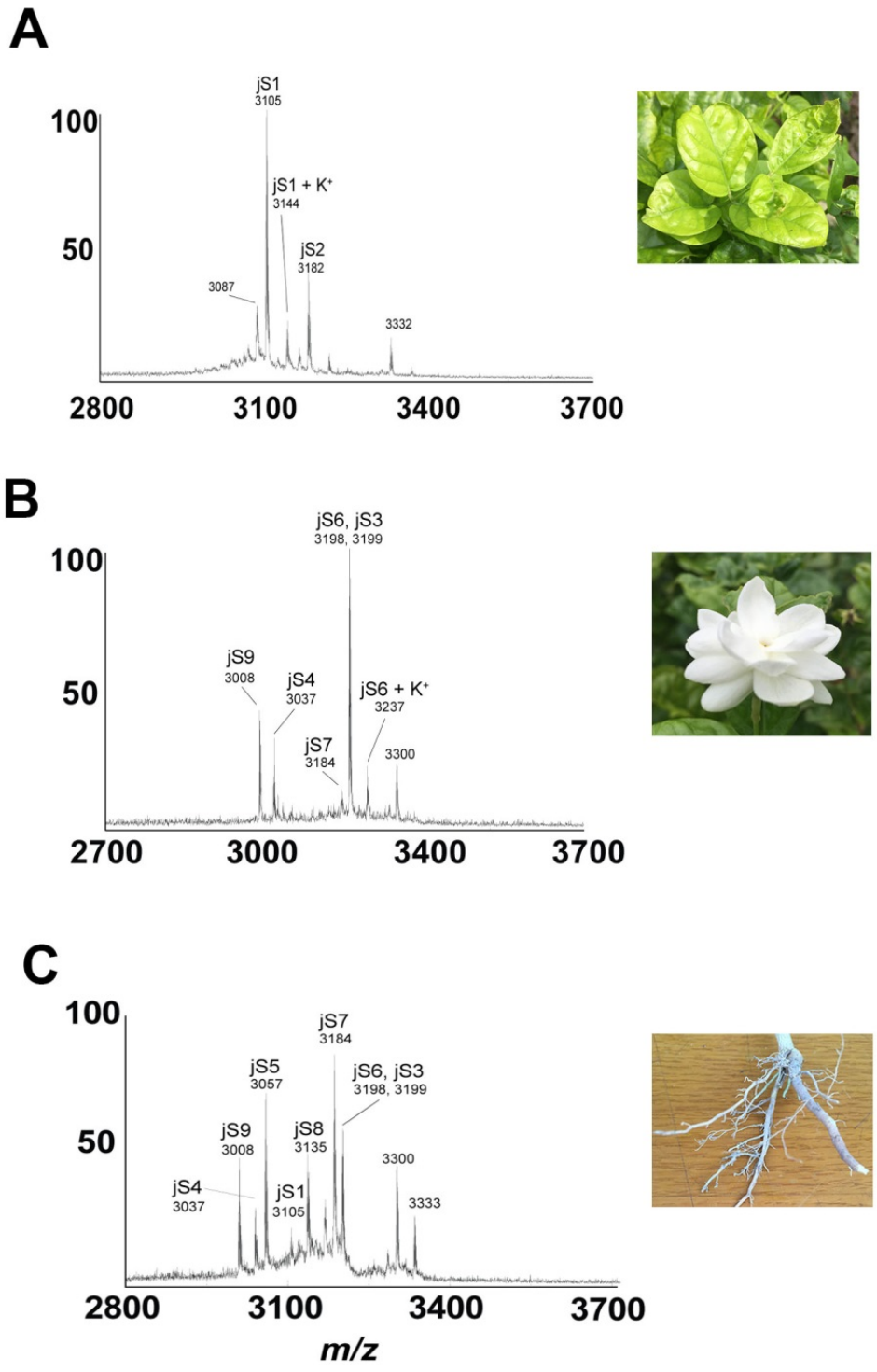


Figure 3.2. Tissue-specific expression of putative CRPs from *J. sambac*. A) shows MS profile of CRPs from leaves, B) shows MS profile of flowers and C) shows MS profile of roots. All parts were extracted with 50% ethanol and analyzed by MALDI-TOF.

2.3. *De Novo* Sequencing of Jasmintide jS1 Isolated from Leaves

To determine the primary sequence, isolated and purified jasmintide jS1 from leaves was reduced by DTT and digested with trypsin and chymotrypsin. The resulted fragments were sequenced by tandem mass spectrometry MALDI-TOF MS/MS by analyzing *b*- and *y*-ions in their profiles. Figure 3.3 showed the sequencing of jasmintide jS1 where *S*-reduced jasmintide jS1 had the *m/z* value of 3111, indicating a mass increase of 6 Da in comparison to native peptide. Trypsin digestion of *S*-reduced jasmintide jS1 produced three fragments with *m/z* values of 846, 1207 and 1095, respectively. The total m/z value of these three digested fragments was 37 units larger than the m/z of whole *S*-reduced jS1, suggesting two tryptic cleavages of the backbone. MALDI-TOF MS/MS gave a full sequence of jS1 (ZLCLQCRSNSDCNIIWRICRDGCCNVI) with *N*-terminal pyroglutamic acid (Z; Figure 3.3). *De novo* sequencing by MS/MS limits the identification of isobaric amino acids such as Leu/Ile and Lys/Gln. These ambiguities were resolved by gene cloning.

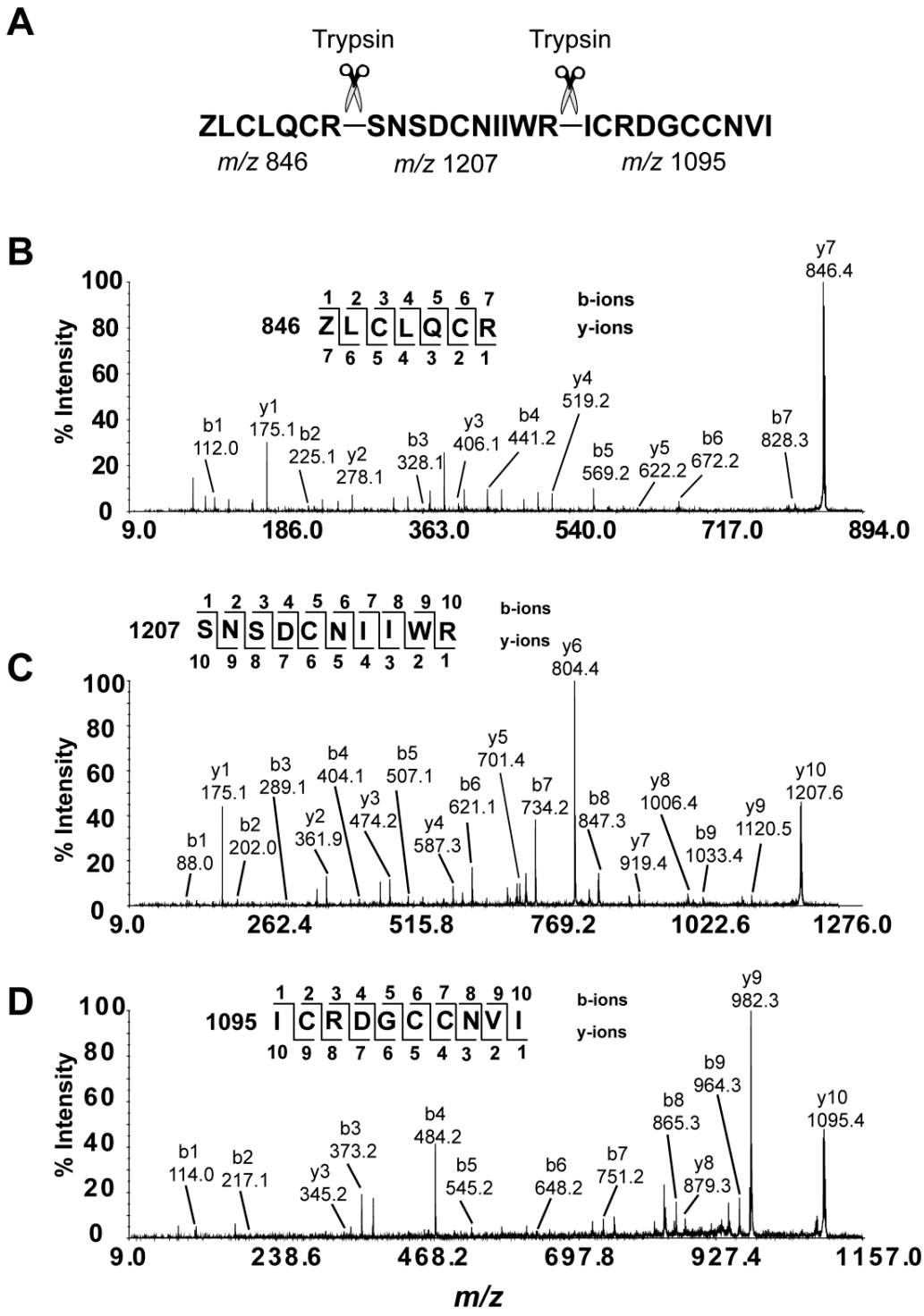


Figure 3.3. MS/MS sequencing of jS1. A) Enzymatic digestion of reduced jasmintide jS1 by trypsin yielded three fragments with *m/z* values 846, 1207 and 1095 Da; Z= pyroglutamic acid. B-D) shows MS/MS spectra of digested fragments. Ile/Leu and Lys/Gln assignments were determined by jasmintide gene.

2.4. Gene Cloning of Jasmintides

The full-length gene of jS1 using 3´- and 5´-rapid amplification of cDNA ends (RACE) polymerase chain reaction (PCR) was obtained to gain insight into its biosynthesis. A degenerate primer targeting the LCRDGCC sequence was designed for 3´-RACE PCR, which resulted in the 3´-end partial gene sequence. Based on this partial gene sequence, a specific primer from the 3´-untranslated region (UTR) was designed and used for 5´-RACE PCR. After assembling the 3´- and 5´-RACE clones, two full-length cDNA gene sequences were obtained that encoded jS1 and jS2 and named *jsac1* and *jsac2*, respectively. To better understand the DNA sequence of jasmintides, gene-specific primers targeting 3´- and 5´-UTRs of *jsac1* and *jsac2* were designed and used in PCR amplification of genomic DNA extracts. Comparison between mRNA and DNA revealed that these jasmintide genes contained a single intron located in the endoplasmic reticulum (ER) signal peptide region.

Gene sequences showed that jasmintide precursors contained a 27-residue ER signal peptide region followed by a 46-residue pro-domain and a 27-residue jasmintide domain at the C-terminus (Figure 3.4 and 3.5). The signal peptide sequence and pro-domain of jS1 and jS2 precursors were highly similar, with only a single residue difference, and had a stop codon immediately after the mature peptide domain (Figure 3.6). Jasmintides jS1 and jS2 shared high sequence homology (88%), differing from each other by only three residues (Table 3.2); they contained six cysteine, four isoleucine, three arginine, two leucine, and two asparagine residues.

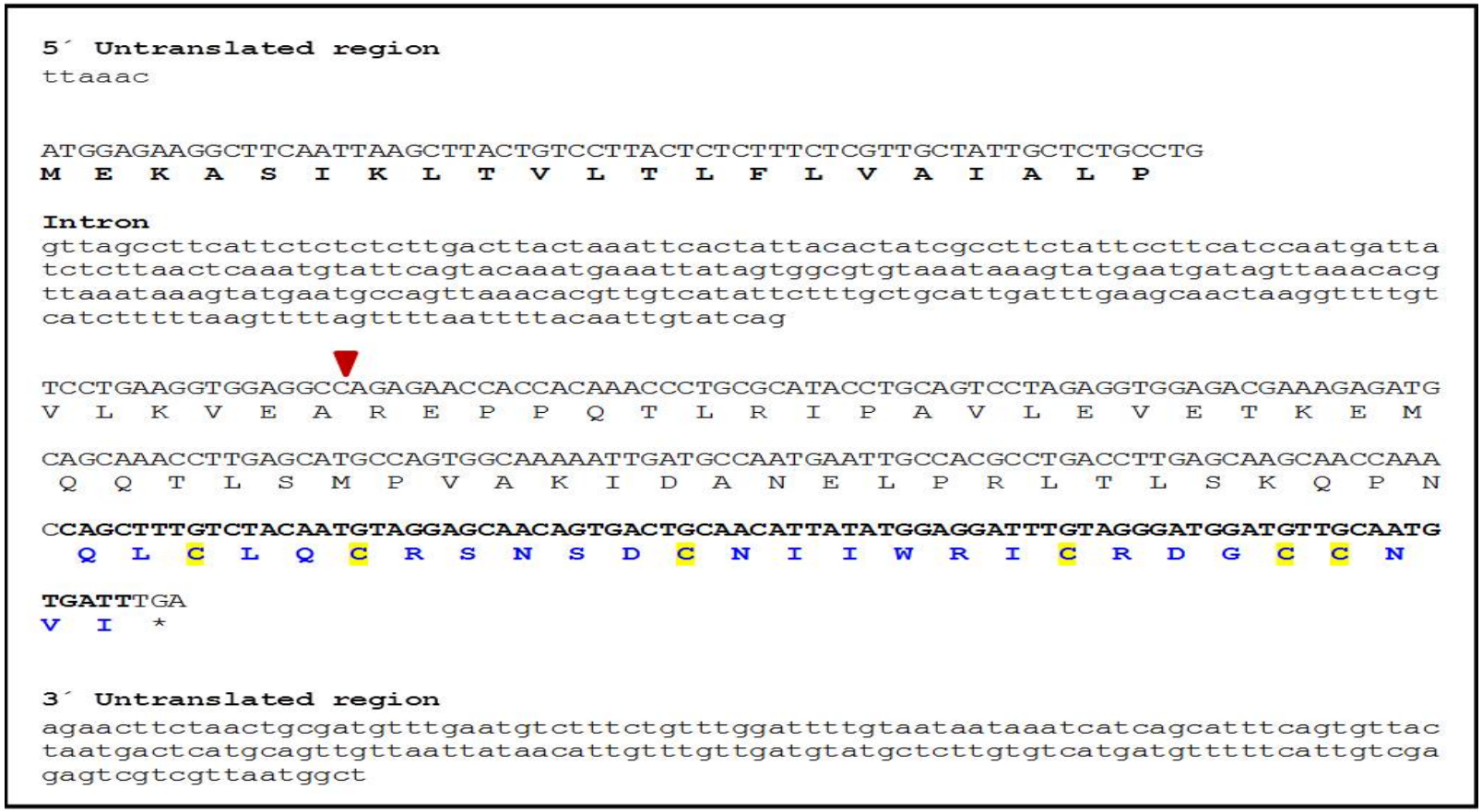


Figure 3.4. The nucleotide and deduced the amino acid sequence of jS1 from *J. sambac*. Untranslated regions and intron are presented in low case whereas the ORF in upper case, the signal peptide is shown in the bold and jasmintide domain in blue with highlighted cysteine residues. The red triangle marks the signal peptide cleavage side which is predicted with SignalP 4.1 using hidden Markov models (HMM) [131]. The asterisk is put beneath the stop codon.

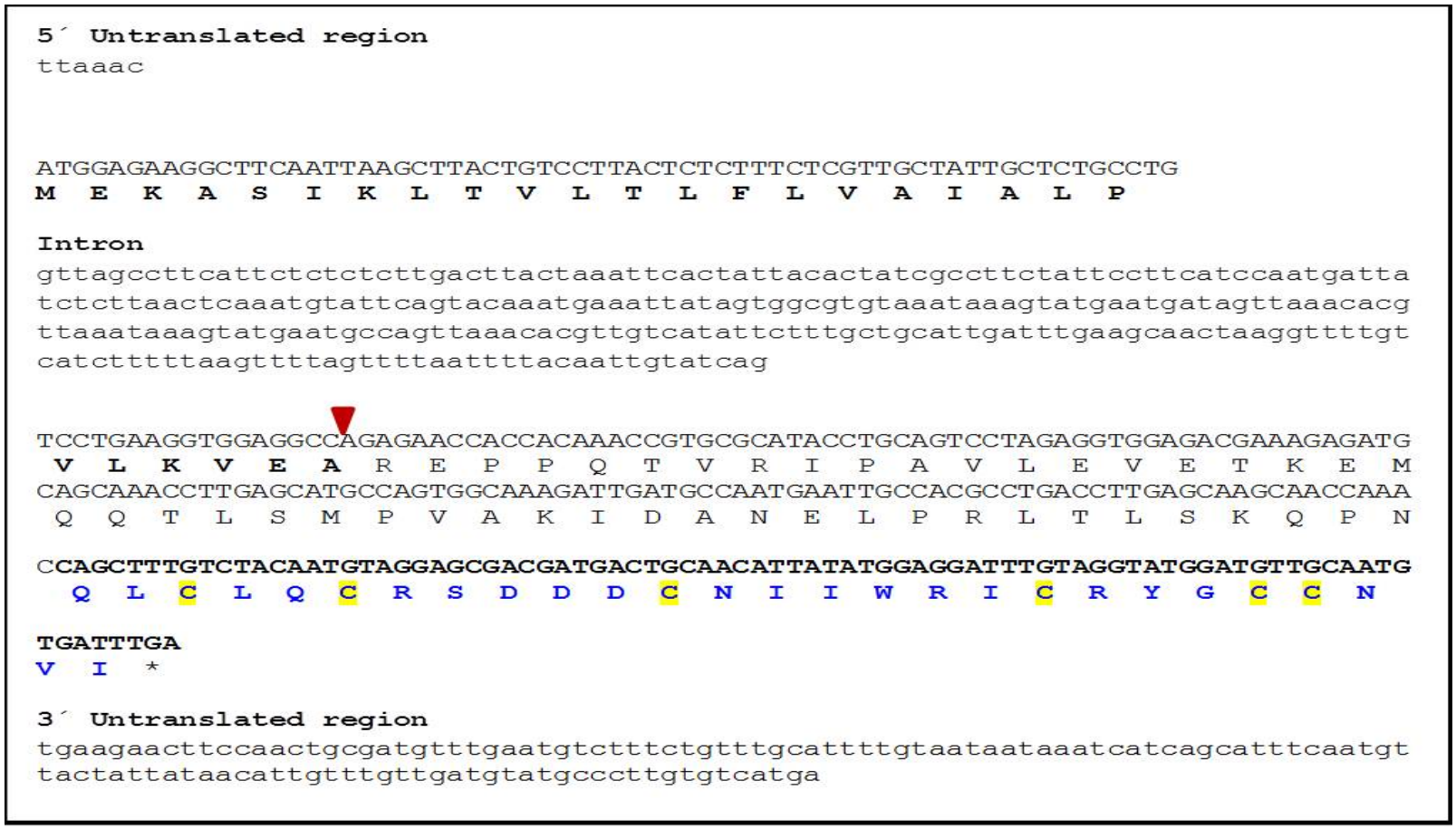


Figure 3.5. The nucleotide and deduced the amino acid sequence of jS2 from *J. sambac*. Untranslated regions and intron are presented in low case whereas the ORF in upper case, the signal peptide is shown in the bold and jasmintide domain in blue with highlighted cysteine residues. The red triangle marks the signal peptide cleavage side which is predicted with SignalP 4.1 using hidden Markov models (HMM) [131]. The asterisk is put beneath the stop codon.

2.5. Analysis of Jasmintides at Nucleic Acid Level

To obtain a more comprehensive library of jasmintides produced by *J. sambac*, we performed transcriptome profiling of the leaf, flower and root. RNA was extracted using TRIZOL reagent from the fresh leaf, flower and root samples. It was purified with oligo (dT) magnetic beads to enrich for mRNA and reverse transcribed into cDNAs, which were then sequenced with Illumina Hiseq 2000 instrument, resulting in 284.2 million clean reads for leaf and flower (referred as LF_JS sample) and 242.1 million reads for root (referred as R_JS sample). The reads were assembled using Trinity to produce a transcriptome of 157,889 and 138,127 contigs for LF_JS sample and R_JS sample, respectively with a minimum length of 201 nucleotides. Using in-house built software tool Protein analyzer 1.6 we extracted 12 contigs, with similar cystine spacing to jS1 sequence, from the root, leaf and flower. The sequences were aligned using ClustalW (Figure 3.6). The work flow of transcriptome analysis is shown in Figure 3.7.

All novel sequences share the same precursor architecture with identified jasmintide (jS1) precursors with three major domains consisting of endoplasmic reticulum signal sequence, prodomain and mature jasmintide domain. The most common N-terminal cleavage site of jasmintide precursor occurs between Asn and Gln/Gly residue. Surprisingly, three precursors jS10, jS12 and jS13 differ in from the common processing site. Interestingly, loop 4 (between Cys 4 and Cys 5) in these three jasmintides is more diverse with all different residues in comparison to other identified jasmintides. Notably, jS10 produce a jasmintide with three residues within loop 1, which is one residue longer compared to the identified jasmintides (jS1-jS9, jS11-jS13). Among all jasmintides, loop 2 consists of five residues and loop 4 consists of three residues, but loop 3 is the most diverse loop with four to seven residues.

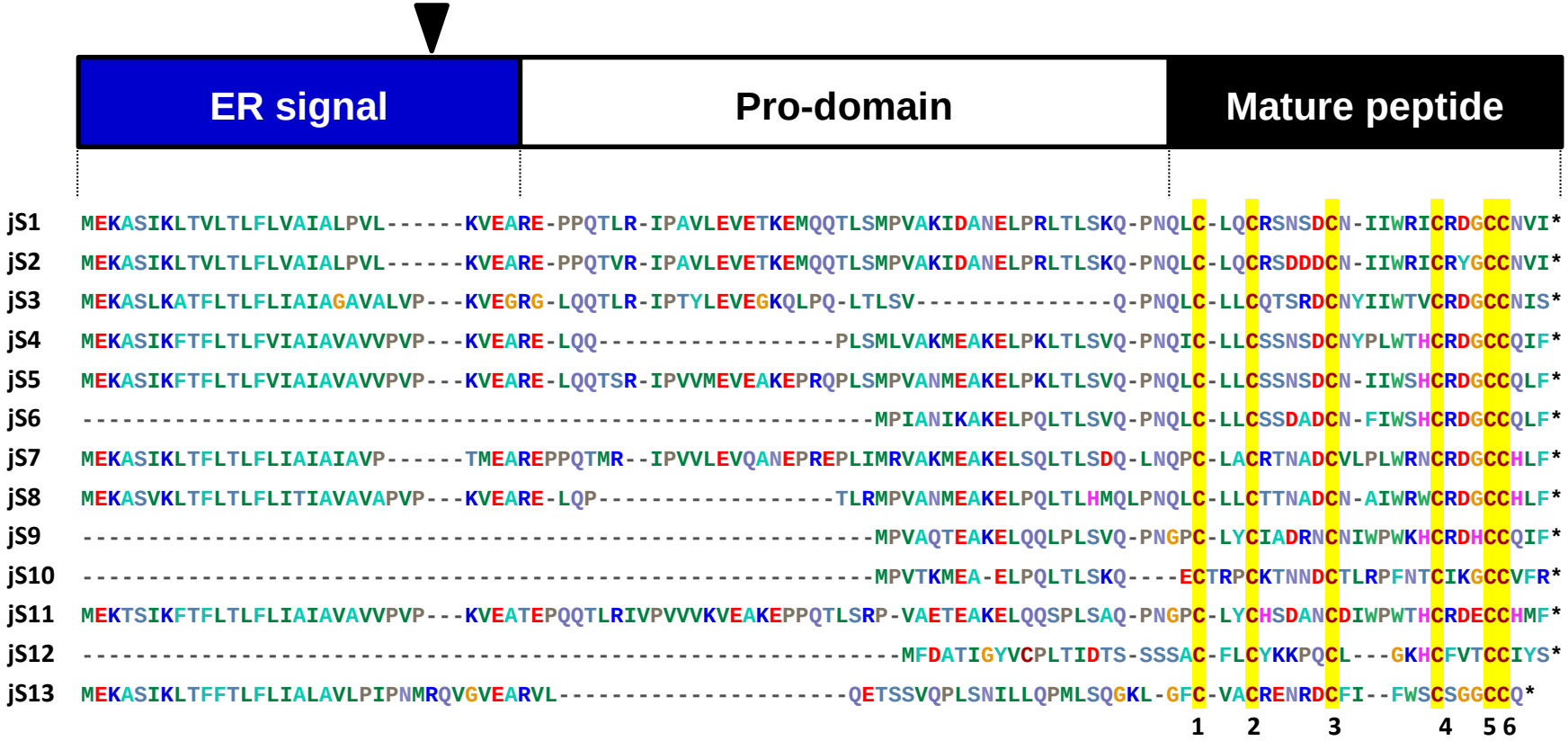


Figure 3.6. Schematic (top) and deduced amino acid sequences of jasmintide precursors (bottom). A jasmintide precursor consists of an endoplasmic reticulum (ER) signal sequence (white) followed by a prodomain (grey), and a jasmintide domain (black). All sequences are obtained from high-throughput transcriptome sequencing of *J. sambac* leaf, flower and root. Strictly conserved Cys residues in the jasmintide domain are shaded with yellow and labeled in Arabic numerals 1-6. The asterisks denote the positions of the stop codon. Intro is represented by a black triangle. The sequence of jS2 was confirmed from the gene.

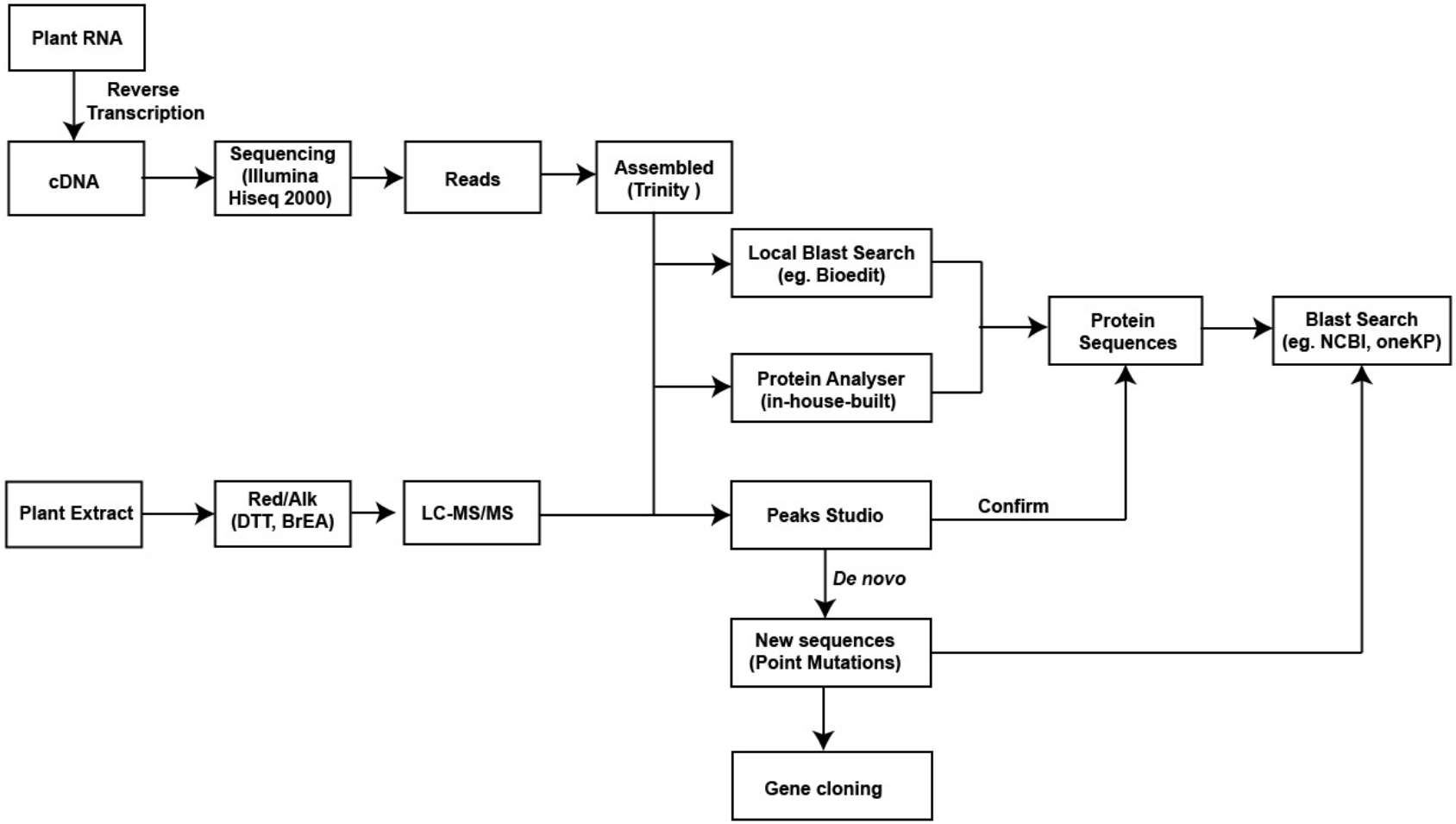


Figure 3.7. Schematic diagram for the work flow of transcriptomic and proteomic analysis.

The transcriptome data also revealed molecular diversity regarding net charge. Precursors of jS1, jS3, jS8 and jS9 encode jasmintides with no net charge, whereas jS10 encodes for a highly positively charged jasmintide with +4 net charge. In addition, four jasmintides jS4, jS5, jS6 and jS11 were found to be negatively charged with -1 or -2 net charge, respectively.

It is known that peptides expression level is affected by environmental changes [184] and some peptide precursors are exclusively expressed in certain plant tissue only [185], these all factors have an influence on peptide transcription. Most importantly, the major limitation in data mining is the erroneous prediction of processing site, which is based on sequence homology only. So, we performed peptide analysis at proteome level using LC-MS/MS to verify the transcriptomic analysis.

2.6. Analysis of Jasmintide Expression at Proteome Level

To elucidate the presence of jasmintides in *J. sambac* at proteome level we used LC-MS/MS in collaboration with Professor Newman Sze's laboratory. Samples including leaves, flowers and roots were extracted with 50% ethanol in water. Disulfide bonds were reduced using DTT and alkylated with 2-bromoethylamine (BrEA) using the method described previously [133]. A mass increment of 264 Da was observed after reduction and alkylation of CRPs with three disulfide bonds. This sample was desalted using C_{18} ziptip and sequenced by using ETD mode in LC-MS/MS. The work flow of proteomic analysis is shown in Figure 3.7.

The data was analyzed in PEAK Studio using the database obtained from jasmine transcriptome mining. PTM were analyzed using PEAKS PTM function [132] in PEAKS Studio and data was searched for >650 PTMs listed in Unimod database [186]. Single point mutations were identified using SPIDER algorithm of PEAKS [135].

We have identified seven jasmintides from leaves named as jS1 to jS7 and nine jasmintides from flowers (jS1-7 and jS14 and jS15) by LC-MS/MS. Jasmintides were named based on the sequences found in transcriptome data. In flowers, seven jasmintides (jS1 to jS7) were found to be identical with those in leaves, but two jasmintide sequences were new which were named as jS14 and jS15, and these two jasmintides were not detected in transcriptome data. However, from the root, we also identified seven jasmintides (jS1 to jS3 and jS5 to jS8). Jasmintide jS4 is not expressed in roots; in contrast, jS8 is only detected in roots. The jasmintide sequences are summarized in Table 3.2.

Table 3.2. Jasmintide from *Jasminum sambac*

Peptide		Sequence	MW[a] (Da.)	Charge[b]	Approaches	Part
Intercysteine segment		**1 2 3 4**				
Jasmintide 1	**jS1**	ZLCLQ-CRSNSDCN-IIWRICRDGCCNVI	3105.62	+1	G, P, T	L, F, R
Jasmintide 2	**jS2**	ZLCLQ-CRSDDDCN-IIWRICRYGCCNVI	3182.70	0	G, P	L, F, R
Jasmintide 3	**jS3**	ZLCLL-CQTSRDCNYIIWTVCRDGCCNIS	3199.42	0	P, T	L, F, R
Jasmintide 4	**jS4**	ZICLL-CSSNSDCNYPLWTHCRDGCCQIF	3198.31	-1	P, T	L, F
Jasmintide 5	**jS5**	ZLCLL-CSSNSDCN-IIWSHCRDGCCQLF	3037.48	-1	P, T	L, F, R
Jasmintide 6	**jS6**	ZLCLL-CSSDADCNF-IWSHCRDGCCQLF	3057.48	-2	P, T	L, F, R
Jasmintide 7	**jS7**	ZPCLA-CRTNADCVLPLWRNCRDGCCHLF	3184.72	+1	P, T	L, F, R
Jasmintide 8	**jS8**	ZLCLL-CTTNADCNA-IWRWCRDGCCHLF	3135.64	0	P, T	R

Cont'd.

Peptide		Sequence	MW[a] (Da.)	Charge[b]	Approaches	Part
Intercysteine segment		**1 2 3 4**				
Jasmintide 9	**jS9**	GPCLY-CIADRNCNIWPWKHCRDHCCQIF	3388.94	0	T	L, F, R
Jasmintide 10	**jS10**	-ECTRPCKTNNDCTLRPFNTCIKGCCVFR	3216.76	+4	T	L, F
Jasmintide 11	**jS11**	GPCLY-CHSDANCDIWPWTHCRDECCHMF	3336.76	-2	T	R
Jasmintide 12	**jS12**	SACFL-CYKKPQCL---GKHCFVTCCIYS	2839.44	+3	T	L, R, F
Jasmintide 13	**jS13**	GFCVA-CRENRDC-FIFWS-CSGGCCQ	2685.07	+1	T	R
Jasmintide 14	**jS14**	ZLCLL-CSSNADCNV-IWSHCRDGCCQLF	3008.44	-1	G, P	L, F, R
Jasmintide 15	**jS15**	QPCLA-CRTNPDCVLPLWRNCRDGCCHLF	3226.45	+1	P	F

[a]Molecular weights are reported as monoisotopic mass. [b]Arg and Lys are taken as positively charged, Asp and Glu as negatively charged and His as neutral. Cysteine connectivity is shown at the bottom with bold lines. (Z) represents the pyroglutamic acid. In approach G=Genomic, P=Proteomic and T=Transcriptomic approaches. In parts L=leaf and F=Flower.

All spectra of sequences identified by proteomics jS1-jS8, jS14 and jS15 are shown in Appendix A, Figure A1. All identified jasmintides sequences from leaves, flowers and roots (Appendix A, Table A1, A2 and A3) were combined in one excel sheet and duplicates were removed. In total, we identified 82 jasmintides by LC-MS/MS analysis and confirmed ten jasmintides, representing 71% coverage of a total number of jasmintides genes present in *J. sambac*. Additionally, 12 truncated sequences with six cysteine residues and 54 truncated sequences with five or less cysteine residues are also reported here (Table 3.3). In total 32 jasmintides were identified with *N*-terminal pyroglutamic acid. This modification is spontaneous or occurs due to the enzymes is still unknown.

Table 3.3. Jasmintides sequences identified from leaves, flowers and roots of *J. sambac* using proteomic analysis

Jasmint--ides	Sequence	-10lgP[a]	Mass	ppm	z	No. (c)	#Spec	MOD[b]
jS1	I.IWRIC(+43.04)RDGC(+43.04)C(+43.04)NVI	153.12	1678.858	-0.4	4	3	17	-
jS1	N.IIWRIC(+43.04)RDGC(+43.04)C(+43.04)NVI	200	1791.942	-0.2	4	3	27	-
jS1	D.C(+43.04)NIIWRIC(+43.04)RDGC(+43.04)C(+43.04)NVI	108.48	2052.037	-0.3	4	4	13	-
jS1	N.SDC(+43.04)NIIWRIC(+43.04)RDGC(+43.04)C(+43.04)NVI	110.5	2254.096	-0.3	4	4	3	-
jS1	Q.C(+43.04)RSNSDC(+43.04)NIIWRIC(+43.04)RDGC(+43.04)C(+43.04)NVI	150.64	2757.323	-1.8	5	5	30	-
jS1	C.LQC(+43.04)RSNSDC(+43.04)NIIWRIC(+43.04)RDGC(+43.04)C(+43.04)NVI	200	2998.466	0.3	5	5	71	-
jS1	L.C(+43.04)LQC(+43.04)RSNSDC(+43.04)NIIWRIC(+43.04)RDGC(+43.04)C(+43.04)NVI	117.6	3144.517	0.2	5	6	3	-
jS1	Q(-17.03)LC(+43.04)LQC(+43.04)RSNSDC(+43.04)NIIWRIC(+43.04)RDGC(+43.04)C(+43.04)N.V	186.11	3156.481	0.7	5	6	13	Pyro-glu from Q
jS1	Q(-17.03)LC(+43.04)LQC(+43.04)RSNSDC(+43.04)NIIWRIC(+43.04)RDGC(+43.04)C(+43.04)NV.I	165.43	3255.549	0.6	6	6	9	Pyro-glu from Q
jS1	Q.LC(+43.04)LQC(+43.04)RSNSDC(+43.04)NIIWRIC(+43.04)RDGC(+43.04)C(+43.04)NVI	200	3257.601	0.4	6	6	25	-
jS1	N.Q(-17.03)LC(+43.04)LQC(+43.04)RSNSDC(+43.04)NIIWRIC(+43.04)RDGC(+43.04)C(+43.04)NVI	172.71	3368.633	1.2	6	6	252	Pyro-glu from Q
jS2	C.LQC(+43.04)RSDDDC(+43.04)NIIWRIC(+43.04)RYGC(+43.04)C(+43.04)N.V	200	2863.328	-0.3	5	5	30	-
jS2	C.LQC(+43.04)RSDDDC(+43.04)NIIWRIC(+43.04)RYGC(+43.04)C(+43.04)NV.I	200	2962.397	-0.5	5	5	17	-
jS2	C.LQC(+43.04)RSDDDC(+43.04)NIIWRIC(+43.04)RYGC(+43.04)C(+43.04)NVI	200	3075.481	0.8	5	5	46	-
jS2	Q(-17.03)LC(+43.04)LQC(+43.04)RSDDDC(+43.04)NIIWRIC(+43.04)RYGC(+43.04)C(+43.04)NV.I	164.75	3332.565	0.1	5	6	1	Pyro-glu from Q
jS2	Q(-17.03)LC(+43.04)LQC(+43.04)RSDDDC(+43.04)NIIWRIC(+43.04)RYGC(+43.04)C(+43.04)NVI	193.97	3445.648	0.8	6	6	98	Pyro-glu from Q
jS3	N.Q(-17.03)LC(+43.04)LLC(+43.04)QTSRD.C	111.19	1347.664	-0.2	3	2	13	Pyro-glu from Q
jS3	N.YIIWTVC(+43.04)RDGC(+43.04)C(+43.04).N	76.12	1559.741	1	3	3	2	-
jS3	I.IWTVC(+43.04)RDGC(+43.04)C(+43.04)NIS	44.85	1597.753	-0.4	3	3	4	-
jS3	N.Q(-17.03)LC(+43.04)LLC(+43.04)QTSRDC(+43.04)N.Y	133.47	1607.758	0.6	4	3	37	Pyro-glu from Q
jS3	Y.IIWTVC(+43.04)RDGC(+43.04)C(+43.04)NIS	167.9	1710.837	-0.6	4	3	10	-

Jasmint--ides	Sequence	-10lgP[a]	Mass	ppm	z	No. (c)	#Spec	MOD[b]
jS3	N.Q(-17.03)LC(+43.04)LLC(+43.04)QTSRDC(+43.04)NY.I	74.91	1770.822	0.3	3	3	1	Pyro-glu from Q
jS3	D.C(+43.04)NYIIWTVC(+43.04)RDGC(+43.04)C(+43.04)NIS	117.04	2133.994	-0.4	4	4	18	-
jS3	N.Q(-17.03)LC(+43.04)LLC(+43.04)QTSRDC(+43.04)NYIIW.T	163.06	2183.069	-0.8	4	3	9	Pyro-glu from Q
jS3	C.QTSRDC(+43.04)NYIIWTVC(+43.04)RDGC(+43.04)C(+43.04)NIS	64.09	2721.261	0.1	5	4	3	-
jS3	L.C(+43.04)QTSRDC(+43.04)NYIIWTVC(+43.04)RDGC(+43.04)C(+43.04)NIS	120.33	2867.312	-0.3	5	5	16	-
jS3	L.LC(+43.04)QTSRDC(+43.04)NYIIWTVC(+43.04)RDGC(+43.04)C(+43.04)NIS	42.73	2980.396	1.4	5	5	4	-
jS3	N.Q(-17.03)LC(+43.04)LLC(+43.04)QTSRDC(+43.04)NYIIWTVC(+43.04)RDGC(+43.04)C(+43.04)NIS	153.61	3463.648	4.7	5	6	15	Pyro-glu from Q
jS3	N.QLC(+43.04)LLC(+43.04)QTSRDC(+43.04)NYIIWTVC(+43.04)RDGC(+43.04)C(+43.04)NIS	41.89	3480.674	-8	5	6	8	-
jS4	Q(-17.03)IC(+43.04)LLC(+43.04)SSNSDC(+43.04)NYPL.W	33.38	1883.858	0.6	3	3	1	Pyro-glu from Q
jS4	D.C(+43.04)NYPLWTHC(+43.04)RDGC(+43.04)C(+43.04)QIF	151.69	2230.006	-0.5	4	4	10	-
jS4	N.SDC(+43.04)NYPLWTHC(+43.04)RDGC(+43.04)C(+43.04)QIF	90.58	2432.065	-0.7	5	4	5	-
jS4	C.SSNSDC(+43.04)NYPLWTHC(+43.04)RDGC(+43.04)C(+43.04)QIF	133.19	2720.172	-0.3	5	4	26	-
jS4	L.C(+43.04)SSNSDC(+43.04)NYPLWTHC(+43.04)RDGC(+43.04)C(+43.04)QIF	154.64	2866.223	0	5	5	57	-
jS4	L.LC(+43.04)SSNSDC(+43.04)NYPLWTHC(+43.04)RDGC(+43.04)C(+43.04)QIF	109.64	2979.307	-0.7	5	5	14	-
jS4	Q(-17.03)IC(+43.04)LLC(+43.04)SSNSDC(+43.04)NYPLWTHC(+43.04)RDGC(+43.04)C(+43.04)QI.F	79.59	3315.49	0	5	6	2	Pyro-glu from Q
JjS4	Q(-17.03)IC(+43.04)LLC(+43.04)SSNSDC(+43.04)NYPLWTHC(+43.04)RDGC(+43.04)C(+43.04)QIF	153.62	3462.559	2	5	6	67	Pyro-glu from Q
jS4	QIC(+43.04)LLC(+43.04)SSNSDC(+43.04)NYPLWTHC(+43.04)RDGC(+43.04)C(+43.04)QIF	58.47	3479.585	-8.9	5	6	10	-
jS4/jS5	N.Q(-17.03)LC(+43.04)LLC(+43.04)SSNSDC(+43.04).N	97.07	1396.615	-0.9	3	3	4	Pyro-glu from Q
jS4/jS5	N.Q(-17.03)LC(+43.04)LLC(+43.04)SSNSDC(+43.04)N.I	77.57	1510.658	0.8	3	3	7	Pyro-glu from Q
jS5	D.C(+43.04)NIIWSHC(+43.04)RDGC(+43.04)C(+43.04)QLF	138.02	2068.958	-0.4	4	4	2	
jS5	N.SDC(+43.04)NIIWSHC(+43.04)RDGC(+43.04)C(+43.04)QLF	91.96	2271.017	4.1	4	4	2	-
jS5	C.SSNSDC(+43.04)NIIWSHC(+43.04)RDGC(+43.04)C(+43.04)QLF	51.01	2559.124	0.4	5	4	4	-
jS5	L.C(+43.04)SSNSDC(+43.04)NIIWSHC(+43.04)RDGC(+43.04)C(+43.04)QLF	116.53	2705.175	-0.5	4	5	8	-

Jasmint--ides	Sequence	-10lgP[a]	Mass	ppm	z	No. (c)	#Spec	MOD[b]
jS5	L.LC(+43.04)SSNSDC(+43.04)NIIWSHC(+43.04)RDGC(+43.04)C(+43.04)QLF	45.77	2818.259	0	5	5	2	-
jS5	N.Q(-17.03)LC(+43.04)LLC(+43.04)SSNSDC(+43.04)NIIWSHC(+43.04)RDGC(+43.04)C(+43.04)Q.L	137.26	3041.358	0.2	5	6	1	Pyro-glu from Q
jS5	L.C(+43.04)LLC(+43.04)SSNSDC(+43.04)NIIWSHC(+43.04)RDGC(+43.04)C(+43.04)QLF	114.15	3077.395	1.8	5	6	2	-
jS5	N.Q(-17.03)LC(+43.04)LLC(+43.04)SSNSDC(+43.04)NIIWSHC(+43.04)RDGC(+43.04)C(+43.04)QL.F	144	3154.443	-0.5	5	6	7	Pyro-glu from Q
jS5	Q(-17.03)LC(+43.04)LLC(+43.04)SSNSDC(+43.04)NIIWSHC(+43.04)RDGC(+43.04)C(+43.04)QLF	200	3301.511	0.2	5	6	15	Pyro-glu from Q
jS5	T.LSVQPNQLC(+43.04)LLC(+43.04)SSNSDC(+43.04)NIIWSHC(+43.04)RDGC(+43.04)C(+43.04)QLF	38.67	3956.876	-8.4	5	6	8	-
jS5/jS6	F.IWSHC(+43.04)RDGC(+43.04)C(+43.04)QLF	116.95	1695.78	0.6	4	3	25	-
jS6	N.Q(-17.03)LC(+43.04)LLC(+43.04)SSDADC(+43.04)N.F	90.03	1495.647	-0.1	3	3	6	Pyro-glu from Q
jS6	D.C(+43.04)NFIWSHC(+43.04)RDGC(+43.04)C(+43.04)QLF	108.67	2102.942	-0.5	4	4	15	-
jS6	Q(-17.03)LC(+43.04)LLC(+43.04)SSDADC(+43.04)NFIWSHC(+43.04)RDGC(+43.04)C(+43.04)QLF	200	3320.484	0.3	5	6	100	Pyro-glu from Q
jJS6	QLC(+43.04)LLC(+43.04)SSDADC(+43.04)NFIWSHC(+43.04)RDGC(+43.04)C(+43.04)QLF	120.44	3337.511	-4.3	5	6	4	-
jS7	N.Q(-17.03)PC(+43.04)LAC(+43.04)RTNADC(+43.04)VL.P	159.69	1617.779	0.3	4	3	56	Pyro-glu from Q
jS7	L.WRNC(+43.04)RDGC(+43.04)C(+43.04)HLF	96.7	1637.749	0	4	3	3	-
jS7	N.Q(-17.03)PC(+43.04)LAC(+43.04)RTNADC(+43.04)VLP.L	110.95	1714.832	-0.5	4	3	13	Pyro-glu from Q
jS7	P.LWRNC(+43.04)RDGC(+43.04)C(+43.04)HLF	49.74	1750.833	0.7	4	3	5	-
jS7	D.C(+43.04)VLPLWRNC(+43.04)RDGC(+43.04)C(+43.04)HL.F	70.17	2059.021	0.7	4	4	6	-
jS7	D.C(+43.04)VLPLWRNC(+43.04)RDGC(+43.04)C(+43.04)HLF	158.47	2206.09	1.6	4	4	4	-
jS7	R.TNADC(+43.04)VLPLWRNC(+43.04)RDGC(+43.04)C(+43.04).H	135.94	2210.033	-0.4	4	4	5	-
JS7	R.TNADC(+43.04)VLPLWRNC(+43.04)RDGC(+43.04)C(+43.04)H.L	91.74	2347.092	-0.1	5	4	8	-
jS7	C.RTNADC(+43.04)VLPLWRNC(+43.04)RDGC(+43.04)C(+43.04).H	59.1	2366.134	-0.7	5	4	4	-
jS7	A.C(+43.04)RTNADC(+43.04)VLPLWRNC(+43.04)RDGC(+43.04)C(+43.04).H	47.17	2512.186	-1.3	5	5	3	-
jS7	C.RTNADC(+43.04)VLPLWRNC(+43.04)RDGC(+43.04)C(+43.04)HLF	66.88	2763.346	0.4	5	4	4	-
jS7	Q(-17.03)PC(+43.04)LAC(+43.04)RTNADC(+43.04)VLPLWRNC(+43.04)RDGC(+43.04)C(+43.04).H	104.52	3050.443	0.1	5	6	5	Pyro-glu from Q

Jasmint--ides	Sequence	-10lgP[a]	Mass	ppm	z	No. (c)	#Spec	MOD[b]
jS7	Q(-17.03)PC(+43.04)LAC(+43.04)RTNADC(+43.04)VLPLWRNC(+43.04)RDGC(+43.04)C(+43.04)H.L	150.62	3187.502	-0.1	5	6	9	Pyro-glu from Q
jS7	N.Q(-17.03)PC(+43.04)LAC(+43.04)RTNADC(+43.04)VLPLWRNC(+43.04)RDGC(+43.04)C(+43.04)HL.F	200	3300.586	0.3	5	6	34	Pyro-glu from Q
jS7	N.Q(-17.03)PC(+43.04)LAC(+43.04)RTNADC(+43.04)VLPLWRNC(+43.04)RDGC(+43.04)C(+43.04)HLF	190.75	3447.654	0.7	5	6	85	Pyro-glu from Q
jS7	N.QPC(+43.04)LAC(+43.04)RTNADC(+43.04)VLPLWRNC(+43.04)RDGC(+43.04)C(+43.04)HLF	89.43	3464.681	-5.1	6	6	8	-
jS8	N.Q(-17.03)LC(+43.04)LLC(+43.04)TTNADC(+43.04)N.A	94.24	1522.694	-0.5	3	3	2	Pyro-glu from Q
jS8	L.C(+43.04)TTNADC(+43.04)NAIWRWC(+43.04)RDGC(+43.04)C(+43.04)HLF	167.41	2802.255	-1.8	5	5	7	-
jS8	N.Q(-17.03)LC(+43.04)LLC(+43.04)TTNADC(+43.04)NAIWRWC(+43.04)RDGC(+43.04)C(+43.04)HLF	182.32	3398.59	0.6	5	6	53	Pyro-glu from Q
jS14	Q(-17.03)LC(+43.04)LLC(+43.04)SSNADC(+43.04).N	82.23	1380.62	0.4	3	3	3	Pyro-glu from Q
jS14	Q(-17.03)LC(+43.04)LLC(+43.04)SSNADC(+43.04)N.V	78.98	1494.663	0	3	3	10	Pyro-glu from Q
jS14	Q(-17.03)LC(+43.04)LLC(+43.04)SSNADC(+43.04)NVI.W	58.49	1706.815	0.6	3	3	1	Pyro-glu from Q
jS14	N.VIWSHC(+43.04)RDGC(+43.04)C(+43.04)QLF	88.73	1794.848	-1.3	4	2	1	-
jS14	L.C(+43.04)SSNADC(+43.04)NVIWSHC(+43.04)RDGC(+43.04)C(+43.04)QLF	107.86	2675.165	1	5	4	24	-
jS14	Q(-17.03)LC(+43.04)LLC(+43.04)SSNADC(+43.04)NVIWSHC(+43.04)RDGC(+43.04)C(+43.04)QL.F	200	3124.432	-0.3	5	6	9	Pyro-glu from Q
jS14	Q(-17.03)LC(+43.04)LLC(+43.04)SSNADC(+43.04)NVIWSHC(+43.04)RDGC(+43.04)C(+43.04)QLF	179.32	3271.501	0.1	5	6	28	Pyro-glu from Q
jS15	QPC(+43.04)LAC(+43.04)RTNPDC(+43.04)VLPLWRNC(+43.04)RDGC(+43.04)C(+43.04)HLF	145.47	3490.696	0.7	5	6	175	-

[a]In Peaks a higher -10logP value indicates a more confident sequencing result. Sequences are identified using false discovery rate (FDR) 0.01%.
[b]MOD: side chain modifications and chemical derivatization. All cysteine residues are alkylated with ethylamine with the addition of 43.04 Da.

We identified three new jasmintides from *J. sambac* that are not found at transcriptome level, Jasmintide jS2, jS14 and jS15. Jasmintide jS2 was identified by gene and is described earlier whereas Jasmintide jS15 differed from jS7 at A10P. This point mutation was manually validated and annotated as shown in Figure 3.8. Jasmintide jS14 was further validated by gene (Figure 3.9). These jasmintides do not seem to be derived from post-translational modifications or misidentification.

Tissue-specific expression of jasmintides is also observed such as jS8 was identified only in roots and jS15 only in flowers, In contrast, leaves do not express jS7 and roots do not express jS4. Jasmintides jS9 to jS13 expressed at transcriptome level were not identified at proteome level, this could be because the expression of CRPs is affected by climate change or the processing of these jasmintides could be using alternate pathway which is unknown and hence could not be detected at proteomic level [184, 185].

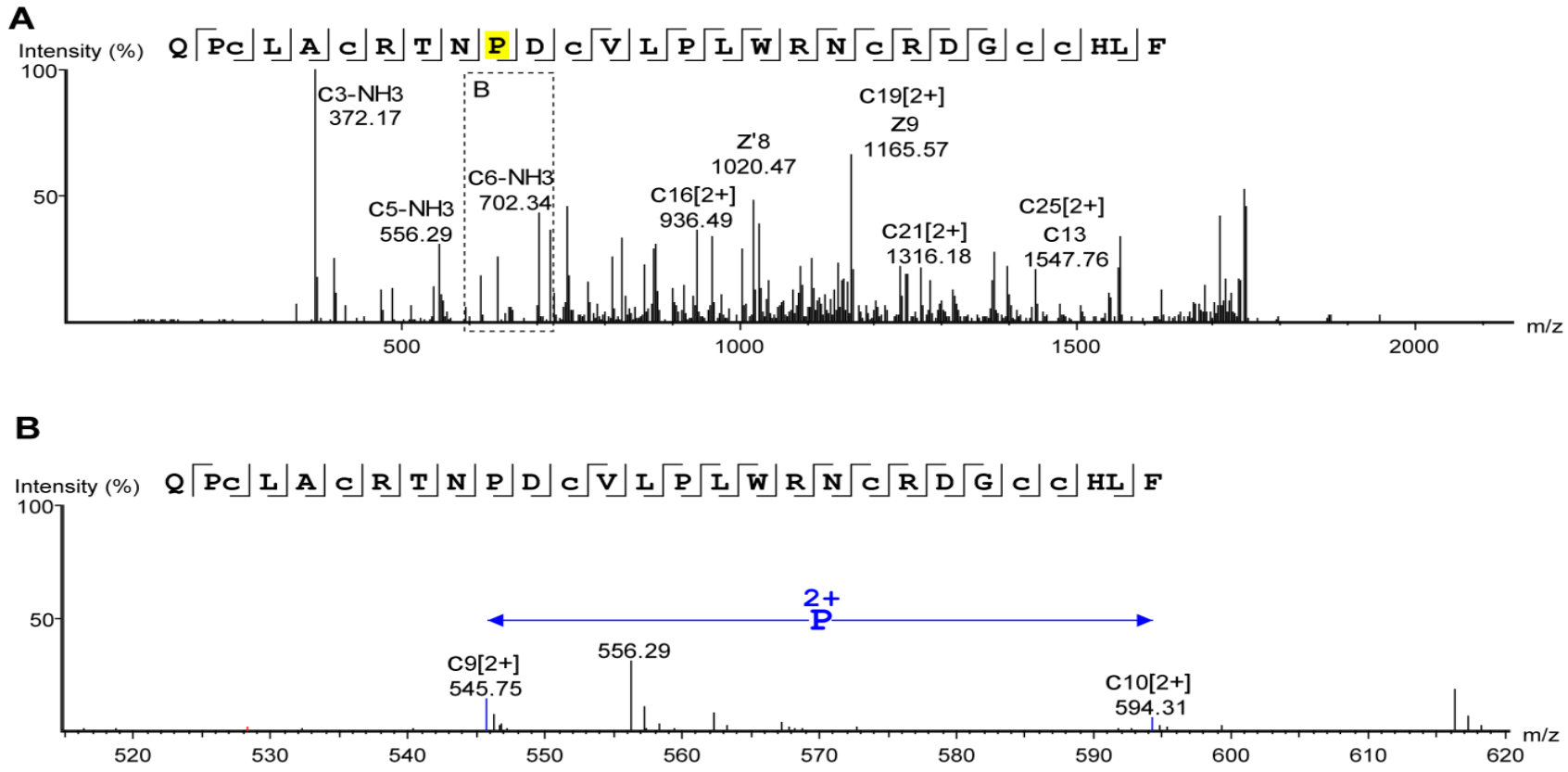


Figure 3.8. MS/MS spectra of the new jS15 sequence. A) shows full annotated spectrum of jS15 which exhibit single amino acid mutation (A to P) compared to jS7.B) shows the zoomed view of spectra of jS15 at P10.

Figure 3.9. The nucleotide and deduced the amino acid sequence of jS9 from *J. sambac*. Untranslated regions are presented in low case whereas the ORF in upper case, the signal peptide is shown in the bold and jasmintide domain in blue with highlighted cysteine residues. The red triangle marks the signal peptide cleavage side which is predicted with SignalP 4.1 using hidden Markov models (HMM) [131]. The asterisk is put beneath the stop codon.

2.7. Disulfide Connectivity Analysis of jS1 shows a New Motif

To obtain partially reduced intermediates of jasmintides in a significant amount, we tested various reduction conditions at room temperature, 37°C and 65°C (Table 3.4). In our strategy, purified jS1 was first partially reduced with a limited amount of TCEP to generate a series of fragments with one, two, or three reduced disulfide bonds. The reduced fragments were immediately alkylated with NEM under an acidic pH to avoid scrambling of disulfide linkages and then purified by RP-HPLC. The number of NEM-tagged cysteines (*S*-NEM) was determined by MALDI-TOF MS. *S*-NEM fragments caused a mass increase of 126 Da. Based on the mass shift, the four purified HPLC fractions were identified. Fraction one consisted of native jS1 with no reduced and alkylated cysteines. Fraction two corresponded to peptide 2*SS* with two *S*-NEM tagged cysteines and two intact disulfide bonds. Fraction three contained peptide 1*SS* with four *S*-NEM tagged cysteine and one intact disulfide bond. The fourth fraction corresponded to peptide 0*SS* with six *S*-NEM tagged cysteine and no intact disulfide bonds (Figure 3.10 A). Peptide 2*SS* was then fully reduced and *S*-alkylated with IAM and subjected to LC-MS/MS analysis to reveal its *S*-NEM connectivity at CysI-CysV (Figure 3.10 B).

Table 3.4. Different conditions for partial reduction

Condition #	TCEP concentration (mM)	Temperature (°C)	Time (min)
1	15	25	1, 5, 10, 15, 30
6	15	37	1, 5, 10, 15, 30
11	15	65	1, 5, 10, 15, 30
16	30	25	1, 5, 10, 15, 30
21	30	37	1, 5, 10, 15, 30
26	30	65	1, 5, 10, 15, 30
32	50	25	1, 5, 10, 15, 30
44	50	35	1, 5, 10, 15, 30
45	50	65	1, 5, 10, 15, 30

Jasmintide jS1 was incubated at pH3. Samples were collected at the indicated time points and spotted on MALDI plate and analyzed by MALDI-TOF MS.

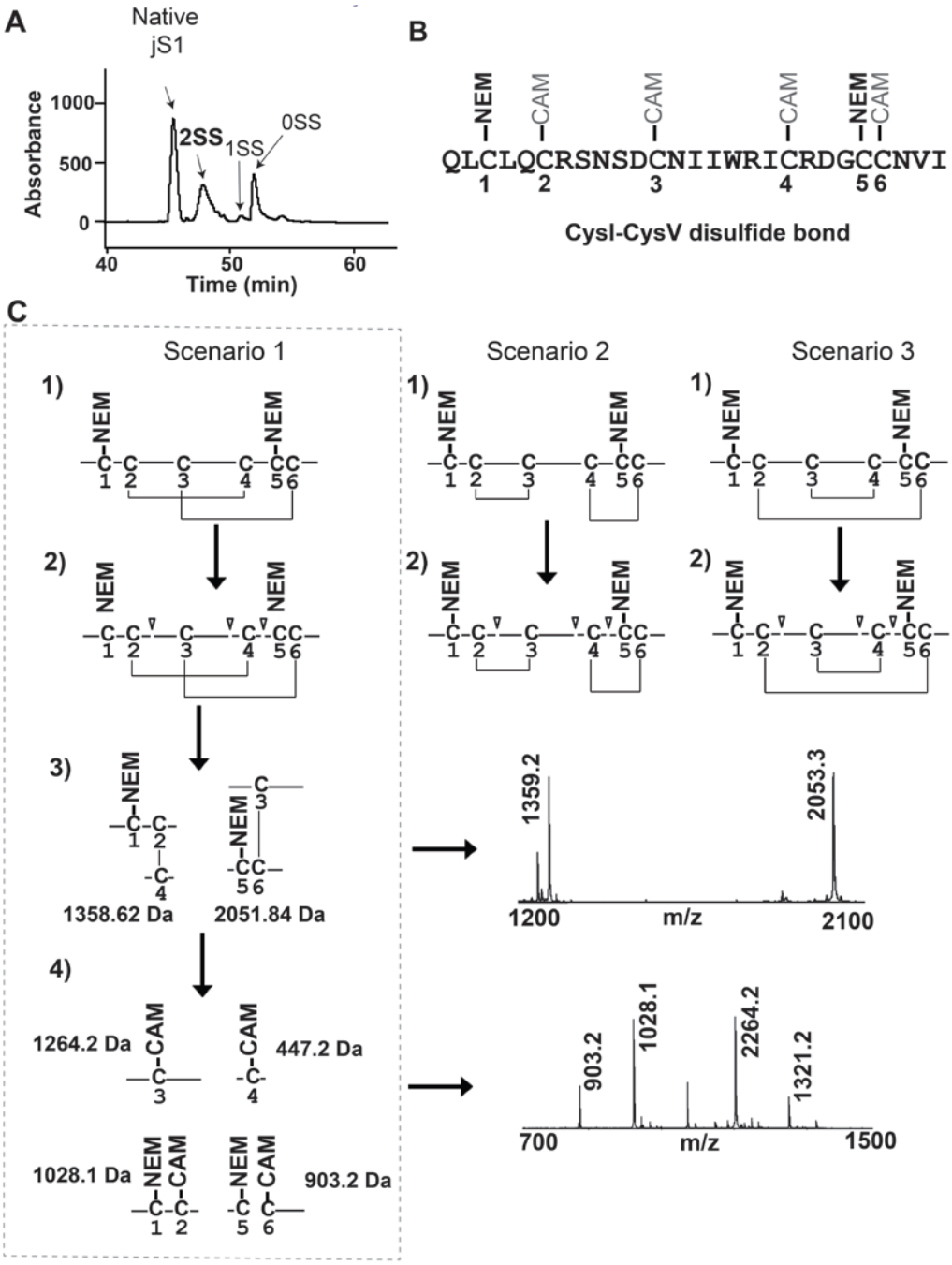


Figure 3.10. Disulfide mapping of Jasmintide jS1. A) RP-HPLC separation of partially *S*-reduced and *S*-NEM labeled jasmintide jS1 derivatives. B) Interpreted sequence of 2*SS* shows the disulfide connectivity CysI-CysV established after the second alkylation with iodoacetamide (IAM denoted as CAM). C) Schematic deconvolution of three possible scenarios by tryptic digestion patterns to arrive CysII-CysIV and CysIII-CysVI connectivity. MS showing two fragments of scenario 1 with *m/z* 1359.2 and 2053.3 generated after tryptic digestion of 2*SS* performed under 2 M urea. Four small fragments generated after reduction and second alkylation using iodoacetamide (IAM) as an alkylating agent.

To map the connectivity of the other two remaining disulfide bonds, peptide 2*SS* was enzymatically digested with trypsin in the presence of 2 M urea at pH 7 for 2 h. From that digestion, three possible combinations were calculated based on the possibilities of remaining two disulfide bonds (Figure 3.10 C). After digestion, two fragments with *m/z* of 1359 and 2053 Da, indicating the possibility of CysII-CysIV and CysIII-CysVI pairs (Figure 3.10 C, Scenario 1), were detected. These tryptic digested samples were fully reduced with DTT and alkylated with IAM, giving rise to four different short peptides which were further identified by LC-MS/MS (Figure 3.11). Based on the two detected tryptic digested peptides and four short peptides released after reduction and alkylation, disulfide linkages between CysII-CysIV and CysIII-CysVI were identified (Figure 3.10 C). Thus, the disulfide connectivity of jS1 was found to be CysI-CysV, CysII-CysIV, and CysIII-CysVI (Figure 3.12 A). It provides unambiguous evidence that jasmintides do not share the Type I connectivity CysI-CysIV, CysII-CysV, and CysIII-CysVI (cystine-knot arrangement) which is commonly found in many CRP families, such as plant defensins, knottins, and heveins [29, 35, 36, 42, 51, 78, 86, 187, 188]. Jasmintide jS1 neither share Type II connectivity CysI-CysVI, CysII-CysV, and CysIII-CysIVI, found in thionins. Thus, we named this novel disulfide connectivity as Type III connectivity.

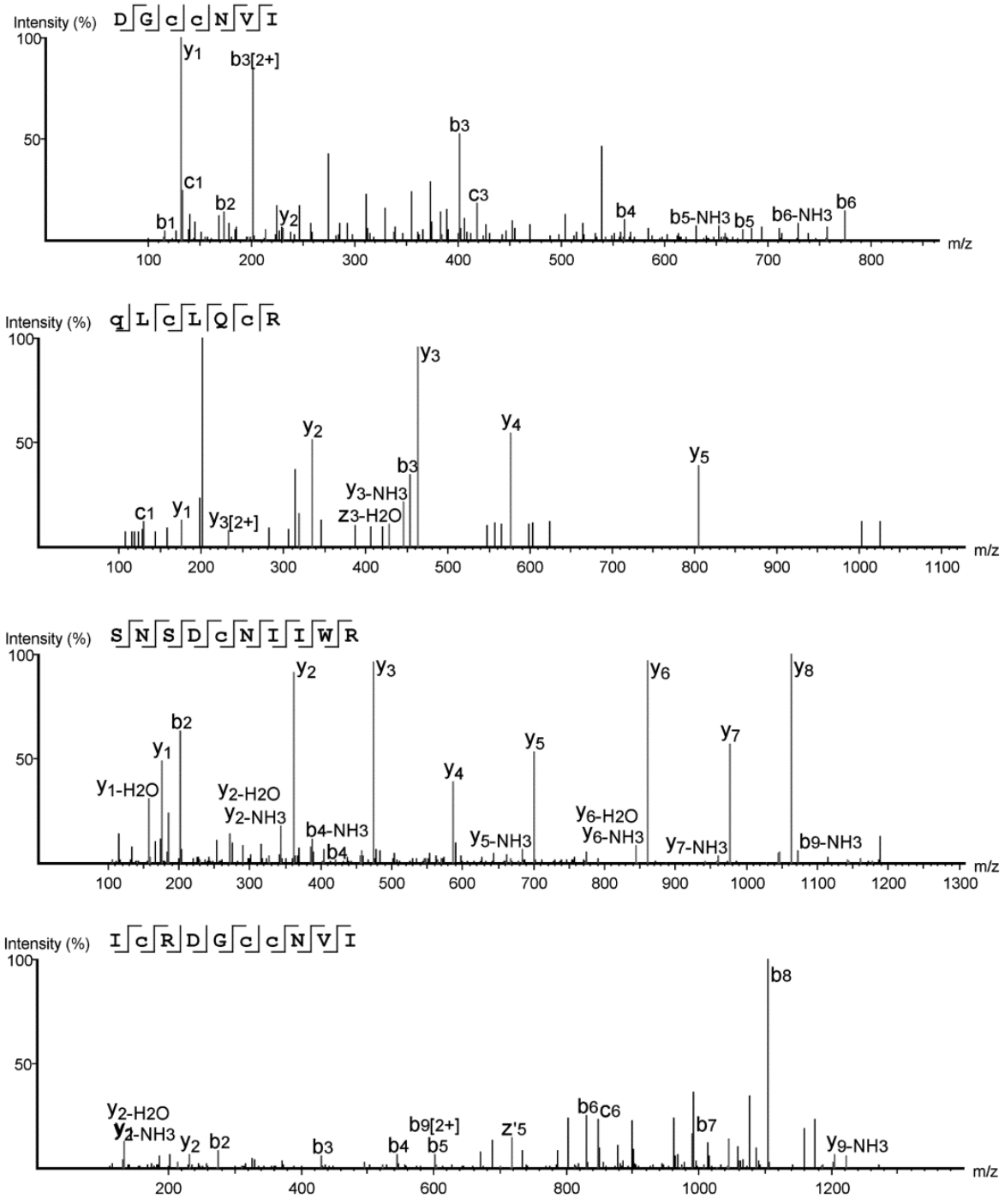


Figure 3.11. MS/MS spectra of each tryptic digested fragment of jS1. Tryptic fragments were identified after fully alkylating with iodoacetamide. MS/MS spectrums showed were pre-processed by PEAKS studio showing only single-charged peaks.

2.8. Solution Structure of the Jasmintide 1 (jS1)

In order to confirm that finding, the disulfide connectivity pattern of jS1 was further validated by NMR spectroscopy. The NMR spectrum of jS1 was characterized by well-dispersed resonances in the amide region signifying a stable folded structure. All spin-spin systems of jS1 were identified, and most proton resonances were unambiguously assigned as ~98% completeness (Appendix B, Figures B1, B2, and Table B1). The solution structure of jS1 was determined based on a total of 656 NMR-derived distance restraints and 22 dihedral angle restraints. The ensemble of 20 low-energy structures is shown in Figure 3.12 C. Root-mean-square deviation values relative to the mean coordinate of 20 conformers were 0.18 Å for backbone atoms and 0.90 Å for all heavy atoms (Table 3.5). Based on the strong NOE connectivities, sequential $d_{\alpha N}(i,i+1)$ NOEs together with hydrogen bond patterns determined by amide-hydrogen exchange experiments, jS1 was determined to be mainly shaped by a combination of three short extended *β*-strands and several loop elements. In addition, several long-range d_{NN}, $d_{\alpha N}$, and $d_{\alpha\alpha}$ NOEs were also observed, indicating a folded *β*-sheet conformation for these regions. The structure of jS1 was found to consist of three short antiparallel *β*-strands (*β*1: L4-Q5; *β*2: I18-R20; *β*3: C23-N25) and two loops (L1:C6-R17 between *β*1 and *β*2; L2: R20-C24 between *β*2 and *β*3), and its molecular shape is well-defined by a number of medium- and long-range NOEs (Figure 3.12 D, Table 3.5). PROCHECK analysis indicated that all residues were distributed in the allowed region of the Ramachandran map (Table 3.5). The position of the disulfide bridges was further confirmed by characteristic $d_{\beta\beta}$ (i,j) and $d_{\alpha\beta}$ (i,j) NOEs between cysteine pairs. An NOE signal between H*β*-H*β* from Cys3 and Cys23 (Figure 3.12 B), which corresponds to the CysI-CysV cysteine pair, was also identified, while the CysII-CysIV pair was identified based on an NOE between H*β*-H*β* from Cys6 and Cys19 (Figure 3.12 B).

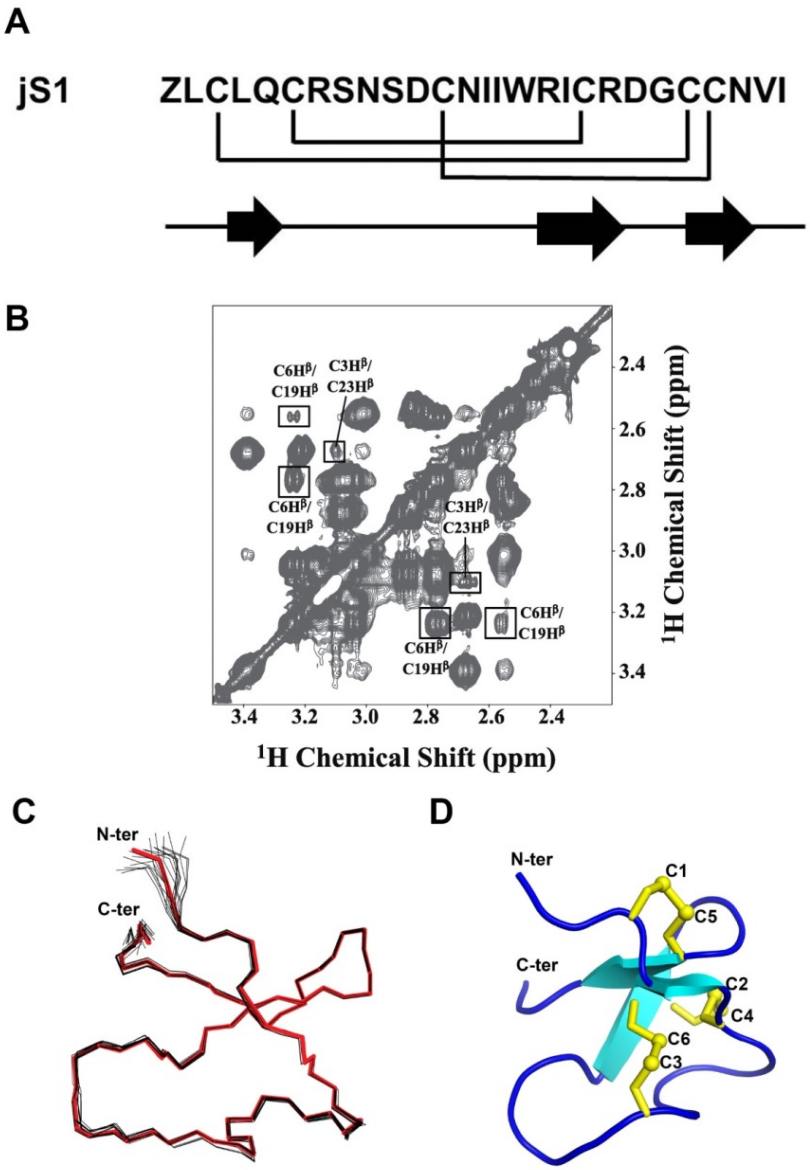


Figure 3.12. The primary, secondary and tertiary structure of the jS1. A) Amino acid sequence of jS1. Disulfide connectivity is shown by bold solid lines. The presence of β-strands is shown by thick arrows. B) Two-dimensional NOESY spectra are showing the expanded aliphatic regions of jS1 obtained with 300-ms mixing in D_2O at 25 °C. Correlations between cysteine-forming disulfide bridges of jS1 (Cys3-H^{β}/Cys23-H^{β}, Cys6-H^{β}/Cys19-H^{β}) are represented by rectangular boxes. C) Superposition of the backbone traces from the final 20 ensembles of solution structures and restrained energy minimized (REM) structures of jS1 by NMR spectroscopy. The REM structure of jS1 is highlighted in red. D) Ribbon representation of jS1 REM structure. Three disulfide bridges (CysI-CysV, CysII-CysIV and CysIII-CysVI) are shown by ball-and-stick representation.

Table 3.5. Structural statistics of NMR structures of jS1

Number of NOE constraints	
All	656
Intra residues |*i-j*|=0	165
Sequential, |*i-j*|=1	195
Medium-range, 1<|*i-j*|<5	86
Long-range, |*i-j*|>=5	210
Number of Hydrogen Bond Constraints	10
Number of Dihedral Angle Constraints	22
Number of Constraint Violations (>0.5Å)	0
Number of Angle Violations (>5°)	0
Energies (kcal/mol)	
E_{NOE}	5.31 ± 1.10
E_{cdih}	0.04 ± 0.05
$E_{bond} + E_{angle} + E_{improper}$	23.57 ± 0.98
E_{VDW}	16.70 ± 1.39
RMS Deviation of the structural segment(Cys^{3}-Ile^{27}) for final 20 structures to REM structure	
Backbone (N,C*α*,C')	0.18 ± 0.08Å
Heavy Atoms	0.90 ± 0.15Å

Additionally, several long-range NOE correlations were observed between HN/H*α* atoms of Cys3 and H*α*/H*β* of Cys23. Similar long-range NOE correlations were identified for Cys6-Cys19 (CysII-CysIV); all of these observations are in agreement with the formation of CysI-CysV and CysII-CysIV disulfide bridges. These results confirm a new cystine motif relative to conventional cystine-knot peptides with disulfide connectivity between CysI-CysV, CysII-CysIV, and CysIII-CysVI. The three disulfide bonds might play an important role in generating and stabilizing the overall structure of jS1.

Analysis of the solvent-exposed surface residues shows that the six cysteine residues were buried into the peptide core like other known CRPs such as cyclotides. However, hydrophobic residues (Ile14, Ile15, Trp 16) in the loop between β-strand 1 and β-strand 2, form a continuous hydrophobic surface combined with *N*- and *C*-terminal hydrophobic residues (Leu2, Leu4, Val26 and Ile27). The surface representation of the jS1 structure reveals the forming of these hydrophobic patches as shown in Figure 3.24. Additionally, most of the charged and polar residues are mainly exposed on the molecular surface.

2.9. Solution Structure of the Jasmintide 3 (jS3)

The NMR spectrum of jS3 showed well-dispersed resonances in the amide region signifying stable folded populations. Based on the combination of 2D-NOESY and TOCSY spectra, all spin−spin systems of jS3 were identified, and most proton resonances were unambiguously assigned as ~98% completeness (Appendix B, Figures B3 and B4, Table B2). The solution structure of jS3 was determined based on a total of 517 NMR-derived distance restraints and 23 dihedral angle restraints. The ensemble of 20 low-energy structures is shown in Figure 3.13 C. Root-mean-square deviation values relative to the mean coordinate of 20 conformers were 0.29 Å for backbone atoms and 1.00 Å for all heavy atoms (Table 3.6). On the basis of the strong $d_{\alpha N(i,i+1)}$ NOE connectivities, coupling constants measured by DQF-COSY and 1H-1D experiments, and hydrogen bond patterns determined by amide−hydrogen exchange experiments, jS3 was determined to be mainly shaped by a combination of three short extended *β*-strands and several loop elements, which resembles with jS1 structure. RMSD value between jS1 and jS3 is 1.16 Å for 24 Cα atoms. The overall structure of jS3 was determined to consist of three short antiparallel *β*-strands (*β*1: C3−L5; *β*2:T18−R21; β3: C24−N26) and two loop segments (L1:C6−W17 between *β*1 and *β*2; L2: D22−G23 between *β*2 and *β*3), and a number of medium and long-range NOEs were observed to prove a well-folded structure as shown in Figure 3.13 D. PROCHECK analysis indicated that all residues were distributed in the allowed region of the Ramachandran map (Table 3.6). The position of three disulfide bridges was further confirmed by characteristic $d_{\beta\beta\ (i,j)}$ and $d_{\alpha\beta\ (i,j)}$ NOEs between cysteine pairs. NOE signals between H*β*−H*β* from Cys3/Cys23 and Cys6/Cys20 (Figure 3.13 B) were observed in NOESY spectra, which correspond to the CI-CV and CII-CIV cysteine pairs, respectively. However, no NOE cross peaks were observed for CIII-CVI pair (Cys12/Cys24) due to spectral overlap. Additionally, several long-range NOE correlations were observed between HN/H*α* atoms of Cys3 and H*α*/H*β* of Cys23. Similar long-range NOE correlations were identified for

Cys6−Cys20 (CysII-CysIV) and Cys12-Cys25 (CysIII-CysVI); all of these observations are in agreement with the formation of CysI-CysV, CysII-CysIV and CysIII-CysVI disulfide bridges such as previously reported jS1. Like jS1, the three disulfide bonds might play an important role in generating and stabilizing the overall structure of jS3. Analysis of the solvent-exposed surface residues shows that the six cysteine residues were buried into the peptide core like jS1 and other known CRPs. And also, hydrophobic residues (Tyr14,Ile15, Ile16, Trp 17) in the loop between *β*-strand 1 and *β*-strand 2 form a continuous hydrophobic surface combined with N- terminal hydrophobic residues (Leu2, Leu4, Leu5). Additionally, most of the charged and polar residues are mainly exposed on the molecular surface.

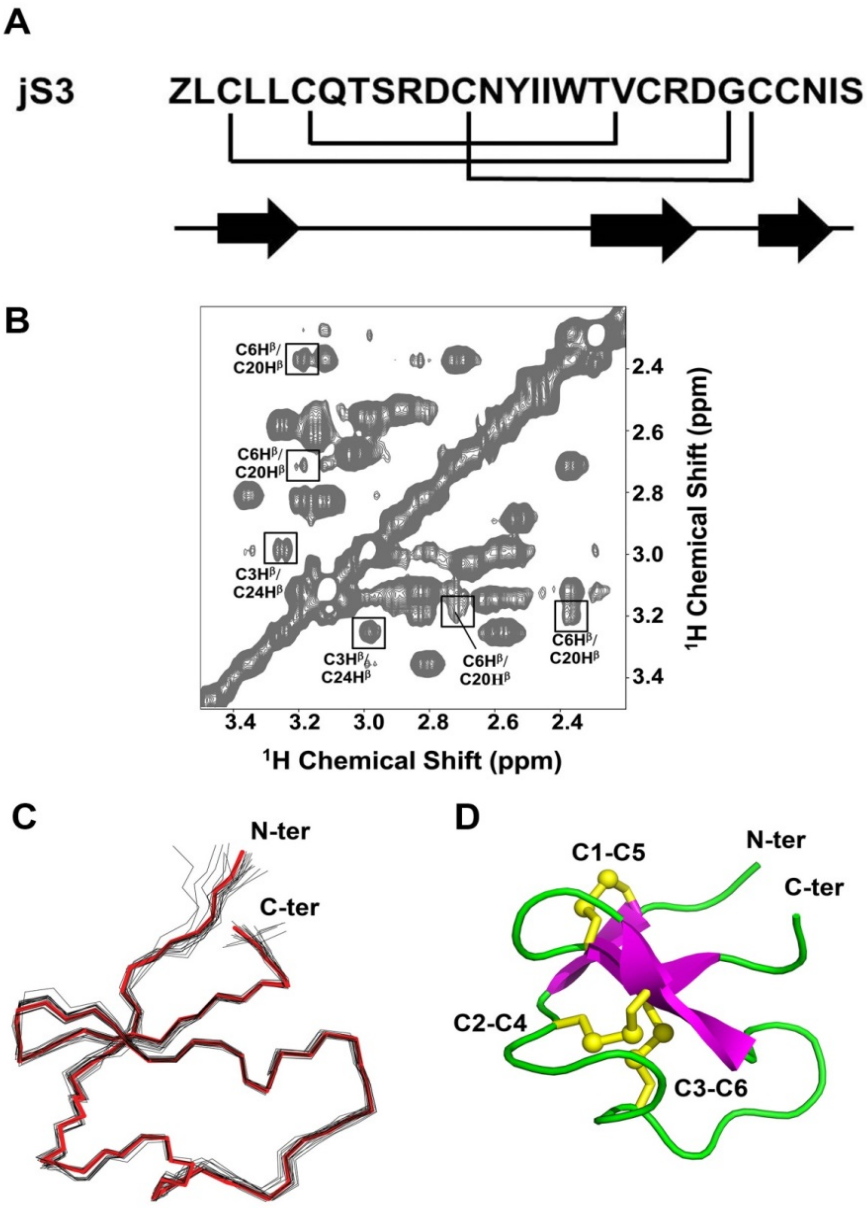


Figure 3.13. The primary, secondary and tertiary structure of the jS3. (A) The primary sequence of jS3. Disulfide connectivity is shown by bold solid lines. The presence of β-strands is shown by thick arrows. (B) Two-dimensional NOESY spectra showing the expanded aliphatic regions of jS3 obtained with 300-ms mixing in D_2O at 25 °C. Correlations between cysteine-forming disulfide bridges of jS3 (Cys3-H^{β}/Cys24-H^{β}, Cys6-H^{β}/Cys20-H^{β}) are represented by rectangular boxes. (C) Superposition of the backbone traces from the final 20 ensembles of solution structures and restrained energy minimized (REM) structures of jS3 by NMR spectroscopy. The REM structure of jS3 is highlighted in red. (D) Ribbon representation of jS3 REM structure. Three disulfide bridges (CysI-CysV, CysII-CysIV and CysIII-CysVI) are shown by ball-and-stick representation.

Table 3.6. Structural statistics of NMR structures of jS3

Number of NOE constraints	
All	503
Intra residues |i-j|=0	139
Sequential, |i-j|=1	141
Medium-range, 1<|i-j|<5	48
Long-range, |i-j|>=5	175
Number of Hydrogen Bond Constraints	14
Number of Dihedral Angle Constraints	23
Number of Constraint Violations (>0.5Å)	0
Number of Angle Violations (>5°)	0
Energies (kcal/mol)	
E_{NOE}	0.74 ± 0.33
E_{cdih}	0.02 ± 0.01
$E_{bond}+E_{angle}+E_{improper}$	18.53 ± 0.78
E_{VDW}	7.27 ± 1.20
RMS Deviation of the structural segment(Cys3-Ser28) for final 20 structures to REM structure	
Backbone(N,Cα,C')	0.29 ± 0.06Å
Heavy Atoms	1.00 ± 0.11Å

2.10. Thermal and Enzymatic Stability of Jasmintides

CRPs are well-known for their proteolytic and thermal stability [17, 35, 36]. This high resistance is an important feature for peptides in herbal medicine. To investigate how tolerant these stabilities are to sequence variation we heated jS1, jS3 and jS4 at 100 °C for 1 h or incubated with trypsin, pepsin or carboxypeptidase A for 8 h. Peptide degradation was monitored using MALDI-TOF and RP-UPLC. The result showed that > 95% of jasmintides survived the heat treatment. In enzyme stability assay, 73% of jS1 and >90% of the jS3 and jS4 remained intact with trypsin treatment and maintaining >91% of the initial concentration of jasmintides with pepsin treatment. Incubation of jasmintides with carboxypeptidase A for eight hours showed that jS4 was 78% stable compared to jS1 and jS3 (>90%). However, the control peptides KL and RV were completely degraded after 1 h of incubation with trypsin or 4 h treatment with pepsin or carboxypeptidase A at 37 °C (Figure 3.14).

Jasmintides were also subjected to protease degradation in 25% human male serum AB type in RPMI medium, and their stability was monitored by RP-HPLC and MALDI-TOF analysis. Linear RF7, a control peptide, was degraded 100% within 4 h. In contrast, more than 75% of jasmintides jS1, jS2 and jS3 remained intact over a period of 24 h in human serum (Figure 3.14).

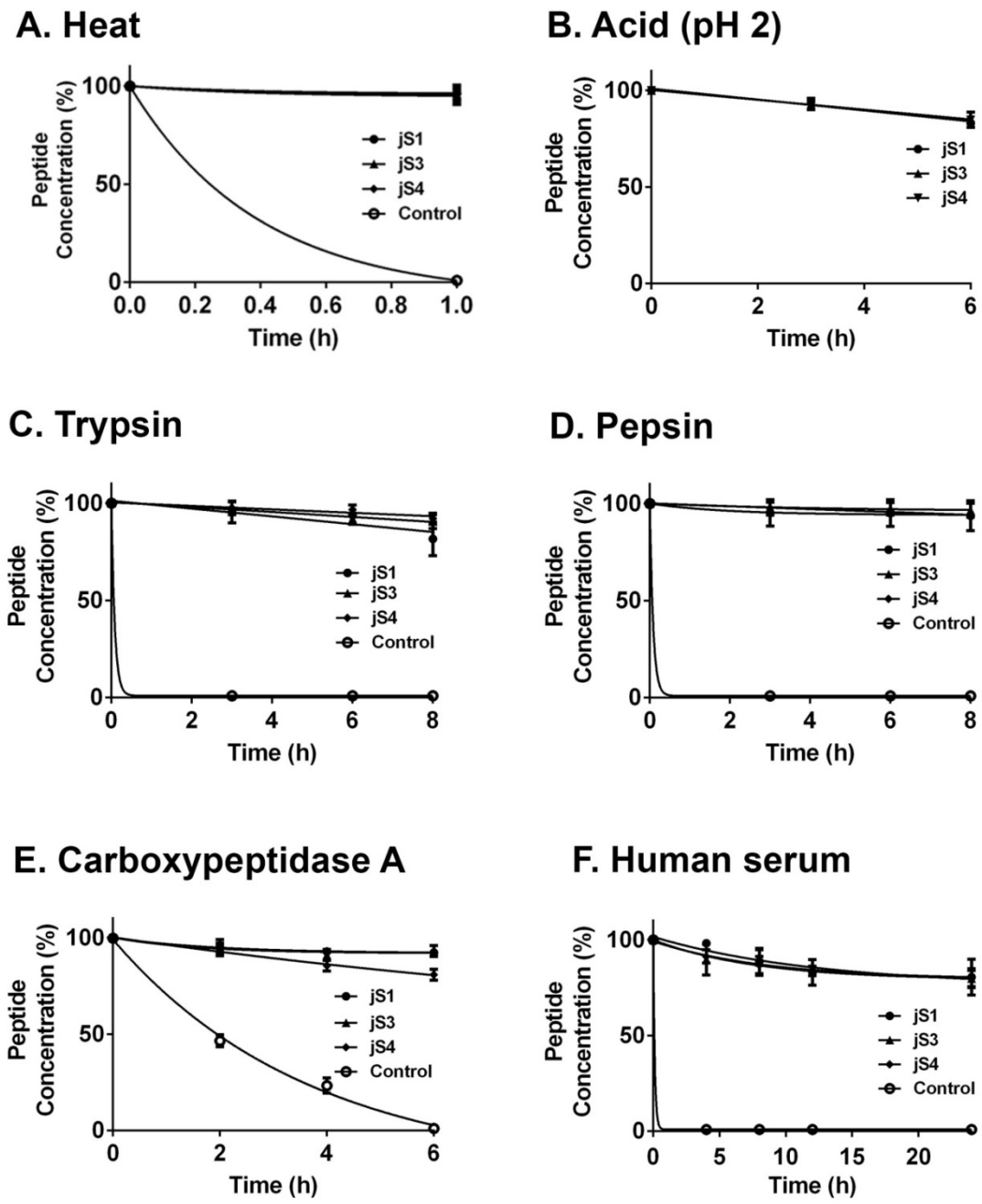


Figure 3.14. Stabilities of jasmintides. A) Thermal stability of at 100°C. B) Acid stability at pH2. C) Trypsin stability. D) Pepsin stability. E) Carboxypeptidase A stability. F) Human serum stability. Error bars represent standard deviation.

2.11. Biological Activity of Jasmintides

Next, we examined the biological activities of jasmintides including cytotoxicity, hemolysis, antimicrobial activity, wound-healing activity, angiogenic activity and collagenase synthesis activity of jS1, jS3 and jS4. To evaluate the cytotoxicity effect of jasmintides on Vero, H1299, Huvec and THP-1 cell lines, cell viability was measured using Presto blue reagent as an indicator. Jasmintides were nontoxic to the four tested cell lines at the concentration of 50 µM. The hemolytic effects were investigated on human type +O erythrocytes. Triton X-100 (0.1%) was used as a positive control. Jasmintides jS1, jS3 and jS4 did not show any hemolytic effect up to 10 µM. Next, we examined antimicrobial activity against *Escherichia coli*, *Staphylococcus aureus* and *Candida albicans* using the radial diffusion assay. The peptide dendrimer, D4R with broad antimicrobial activity, was used as a positive control [155]. Jasmintides did not exhibit antimicrobial activity against the tested strains at concentrations up to 100 µM. Further, we tested jasmintides for anti-inflammatory activity using scratch assay and tube formation assay, based on the traditional use of jasmine for inflammation. Jasmintides did not exhibit any significant change up to 25 uM. Finally, we also investigated the role of jasmintides for collagenase synthesis due to its usage in cosmetics, but jasmintides were not found to show any increase or decrease in collagenase synthesis.

2.12. Antiviral Activity against Infectious Bronchitis Virus (IBV)

We screened jasmintides for antiviral activities against the infectious bronchitis virus (IBV). Cytotoxicity studies demonstrated that jasmintides jS1 and jS4 were not cytotoxic to Vero cells up to the concentration of 50 μM. Thus, subsequent in vitro studies were performed with concentrations in the range of 0.01-10 μM.

Jasmintides jS1 and jS4 were evaluated for the ability to inhibit cytopathic effect caused by IBV. Briefly, Vero cells were inoculated with mock (medium only), jasmintide and IBV or IBV only (0.1= MOI) and viral replication was evaluated at 24 hpi (Figures 3.15). IBV replication in the absence of jasmintides was scored 100%, which was used to normalize the treated samples. The results showed that co-treatment with IBV and jasmintide could inhibit plaque formation in a dose-dependent manner with an IC_{50} of 70 nM, and 86 nM for jS1 and jS4, respectively. Jasmintide jS1 was found to be more potent than jS4.

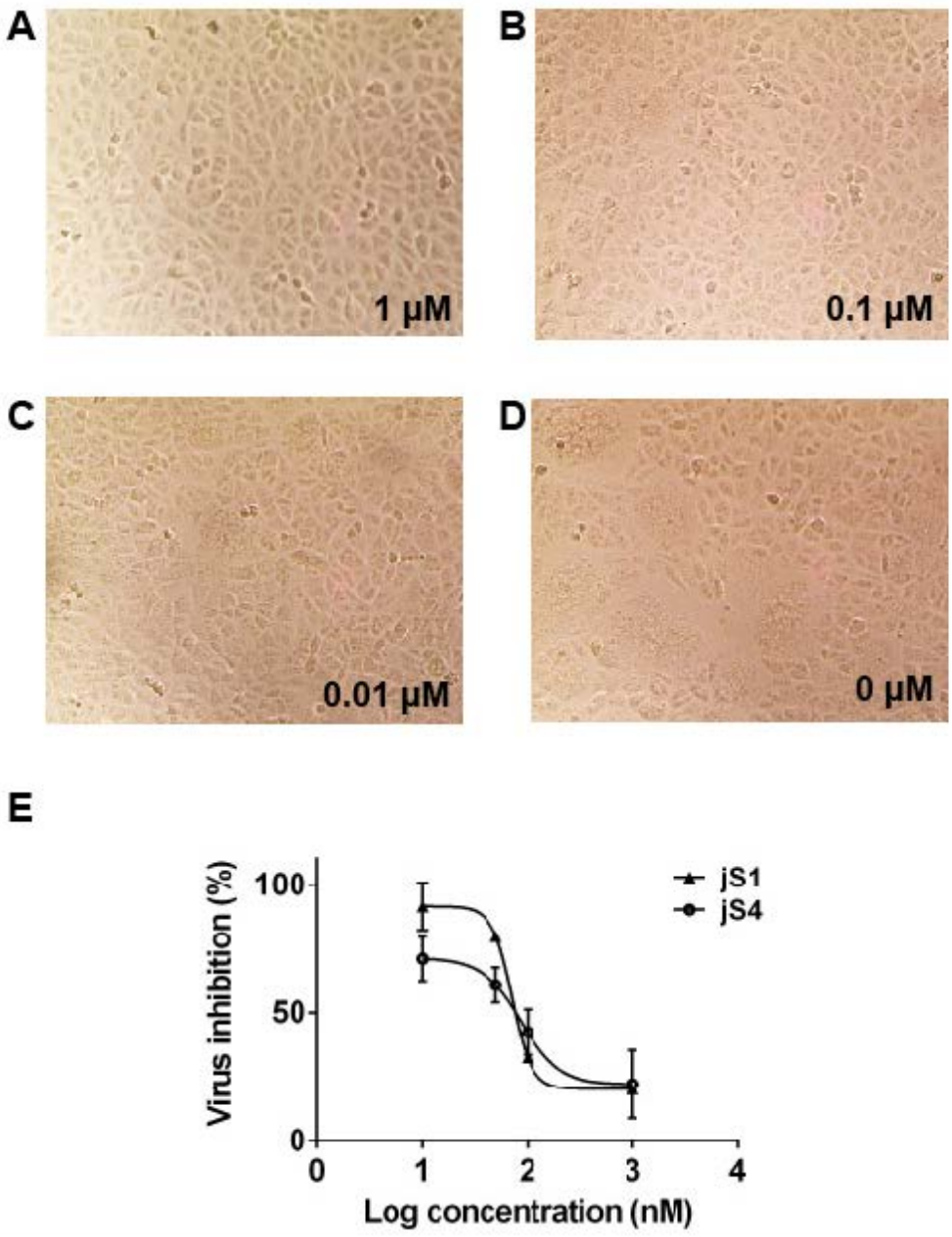


Figure 3.15. Evaluation of the Antiviral effect of jasmintides against IBV infection. Microscopic Images A to D corresponds to different concentrations of jasmintide jS1. Sample containing 0uM jS3 represents negative control. (E) The graph represents the dose-dependent inhibition of jasmintides jS1 and jS4 determined at a concentration ranging between 0.01 to 10 µM. Both jasmintides exhibit antiviral effect against IBV infection of Vero cells with estimated IC_{50} of 70, and 86 nM, respectively, calculated using Graph Pad prism 5. Error bars represent standard deviation.

2.13. Effect of Jasmintides on Growth of Larvae of *Tenebrio molitor*

The control of pest in plants is of great socio- and economic importance. Hence, we tested the role of jasmintides in plant defense. Purified jasmintides jS1, jS3, jS4 and cyclotide cT9 at the concentration of 1 mg/mL were incorporated into the artificial diet and fed to mealworm larvae for nine days. Jasmintides jS1, jS3 and jS4 showed a significant effect on the growth of larvae compared with larvae fed on jasmintide free diet. No mortality was observed with jS1, jS4 and cT9. However, only one larva was found dead with jS3 on day 3. Remaining all larvae survived until the end of the experiment. After nine days the control worms larva weighed 79 mg whereas jasmintide fed larva weighed between 58 to 62 mg and larva fed on cT9 diet weighed 54 mg as shown in Figure 3.16 and Table 3.7. The insecticidal potency of jasmintides jS1, jS3, jS4 and cT9 appeared to be very similar, despite the lack of sequence similarity between them.

2.14. Antifeedant Activity of Jasmintides in No-choice Assay against Larvae of *Tenebrio molitor*

In no-choice tests, we fed larvae with a diet containing jasmintides (jS1, jS3 or jS4) or cyclotide (cT9) or control (no peptide) for three days. After that, remaining diet was weighed and the difference of diet eaten by worms was calculated. Figure 3.17 showed that jasmintides jS1, jS3 and jS4 consumed only half the diet compare to the control (no peptide), which reflects similar result as by cyclotide cT9, a positive control. This data suggests that mealworm larvae do not prefer to consume jasmintides containing diet.

Table 3.7. Weight and mortality rates of *T. molitor* fed a diet containing jasmintides

Peptide	Weight (mg)± SEM		Weight increase (%)	Mortality (%)
	Day 0[a]	Day 9[a]		
Control	47.4±0.8	71.2±1.5	50	0
jS1	44.3±1.0	58.9±1.3	33	0
jS3	44.3±0.7	62.2±1.3	40	5
jS4	45.7±0.9	61.1±1.2	33	0
cT9	44.5±0.8	54.6±1.5	29	0

[a]Each value represents the mean of five replicates, each set with 4 larvae (n=40). For comparison of mean with the control, the *t*-test was used. Differences are statistically significant at p*<0.05

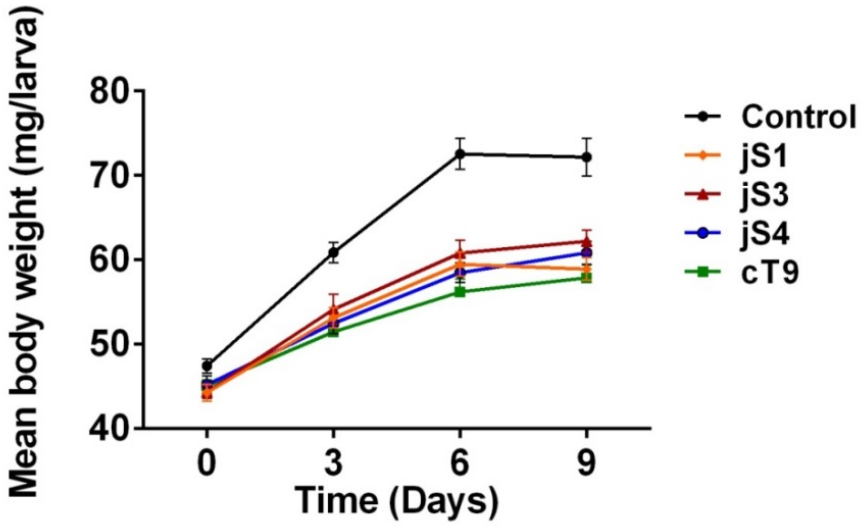


Figure 3.16. Effect of jasmintides on the growth of *T. molitor* larvae. Each set of 4 larvae (n=40) were fed with diet containing jasmintides (jS1, jS3, and jS4) or cyclotide (cT9) or no peptides as a control. The body weight of larvae was monitored after every 3 days. Error bars represent standard error mean. For comparison of mean with the control, the t-test was used. Differences are statistically significant at p*<0.05.

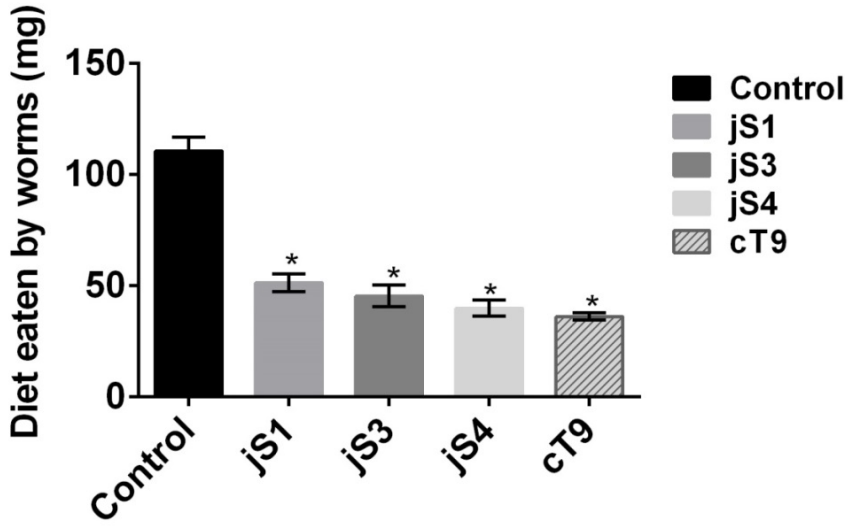


Figure 3.17. Antifeedant Activity of Jasmintides in No-choice Assay against Larvae of Tenebrio molitor. Each set of 4 larvae (n=40) were fed with a diet containing jasmintides (jS1, jS3, and jS4) or cyclotide (cT9) or no peptides as a control. The diet remaining in petri dishes was weighed after 3 days. Error bars represent standard deviation. For comparison of mean with the control, the t-test was used. Differences are statistically significant at p*<0.05.

3. Discussion

In this study, we report the discovery and characterization of 15 CRPs from *J. sambac,* jS1-jS15 (jasmintides). With the discovery of jasmintides, a new family of CRPs in plants is identified with Type III disulfide connectivity. These peptides contain 25-28 amino acid residues and adopt a novel disulfide connectivity pattern which has not been found in plant CRPs till date. The peptidomic studies provide further understanding about the molecular diversity of jasmintides regarding their sequence, structure and function. We have also shown that jasmintides possess an insecticidal activity against yellow mealworms, which suggests their role in plant defense. In addition, the antiviral effect discovered in this study suggests that similar to other CRPs such as defensins, α-amylase inhibitors and cyclotides, jasmintides may possess a promiscuous array of biological functions [55, 89, 91]. Thus, it is of interest to explore their therapeutic potential for applications as a drug lead in developing peptidyl drugs.

3.1. Sequence Conservation and Molecular Diversity of Jasmintides

Jasmintides are moderately conserved within specie; they share 22-85% sequence identity with each other. The consensus sequence for loop 2 to loop 4 (based on the inter-cysteinyl residues from the N-terminus) throughout jasmintides jS1-jS15 is $X_{1\text{-}2}CX_{2\text{-}3}CX_5CX_{4\text{-}7}CX_3CCX_3$.

The 15 jasmintide sequences reported here from *J. sambac* show extensive variations in the size and composition of their loops. Sequence logo of jasmintides (Figure 3.18) summarizes the sequence variations and also highlights the regions of conservation. In general, it would be expected that the highly conserved residues play a role in maintaining the structure integrity of the framework while other residues which show variations might be associated with the adaptation to target different pests or for different biological functions. The residues which are conserved throughout the jasmintide suite are the cysteine residues, a hydrophobic residue at the fouth position and an aromatic residue at

the fifth position in loop 3. For example, in jS3 Ile16 and Trp17 are at 4^{th} and 5^{th} position in loop 3. NMR structural characterization of jS1 and jS3 shows that these residues are involved in structural stabilization through hydrogen bonding or hydrophobic interactions. Overall, only one-third residues of the scaffold are highly conserved; hence, there are many possibilities that plants evolve the remainder of sequences for optimizing the biological functions.

One of the jasmintides, jS10 contains unusual sequence element such as three residues in loop 1 (TRP) whereas all other identified sequences contain only two residues. Along with the one additional residue, the amino acid content of this loop stands out (Table 3.2). With the presence of a basic residue in each loop, jS10 has a net charge of +4 and is thus; more cationic compare to all other jasmintides. Together with the hydrophobic residues, this makes jS10 more amphipathic. So far, in all jasmintides loop 4 contains basic and acidic residue, except jS12 and jS13 which do not contain any acidic or basic amino acid (Figure 3.18). Sequence diversity wheel of jasmintides in Figure 3.19 shows that the framework of jasmintide is capable of tolerating a range of sequences of varying length and amino acid contents.

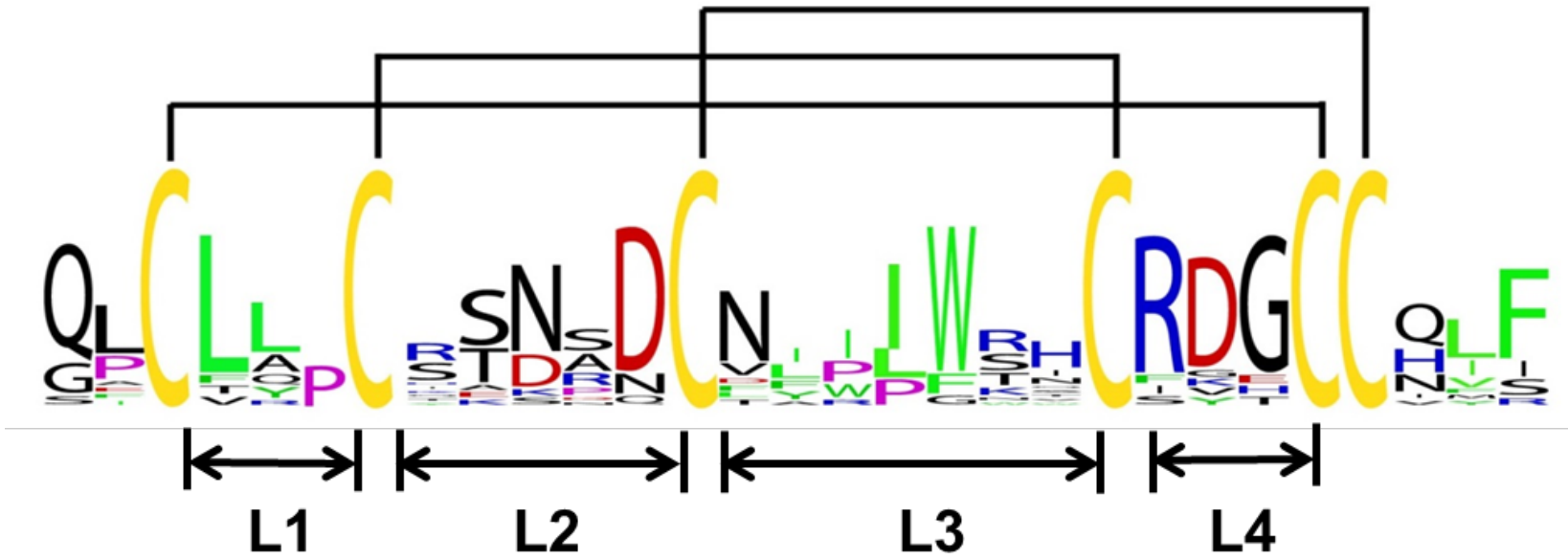


Figure 3.18. Sequence logo of jasmintides. The peptides were first aligned using ClustalW program. Logo sequences were built with WebLogo application, version 3 (http://weblogo.berkeley.edu/) [189]. The total height of each column reflects the conservation whereas the size of each letter demonstrates the relative occurrence of that residue at a given position. Disulfide connectivity is shown in bold lines. The sequence between inter cysteine is referred as a loop and represented as L1-L4.

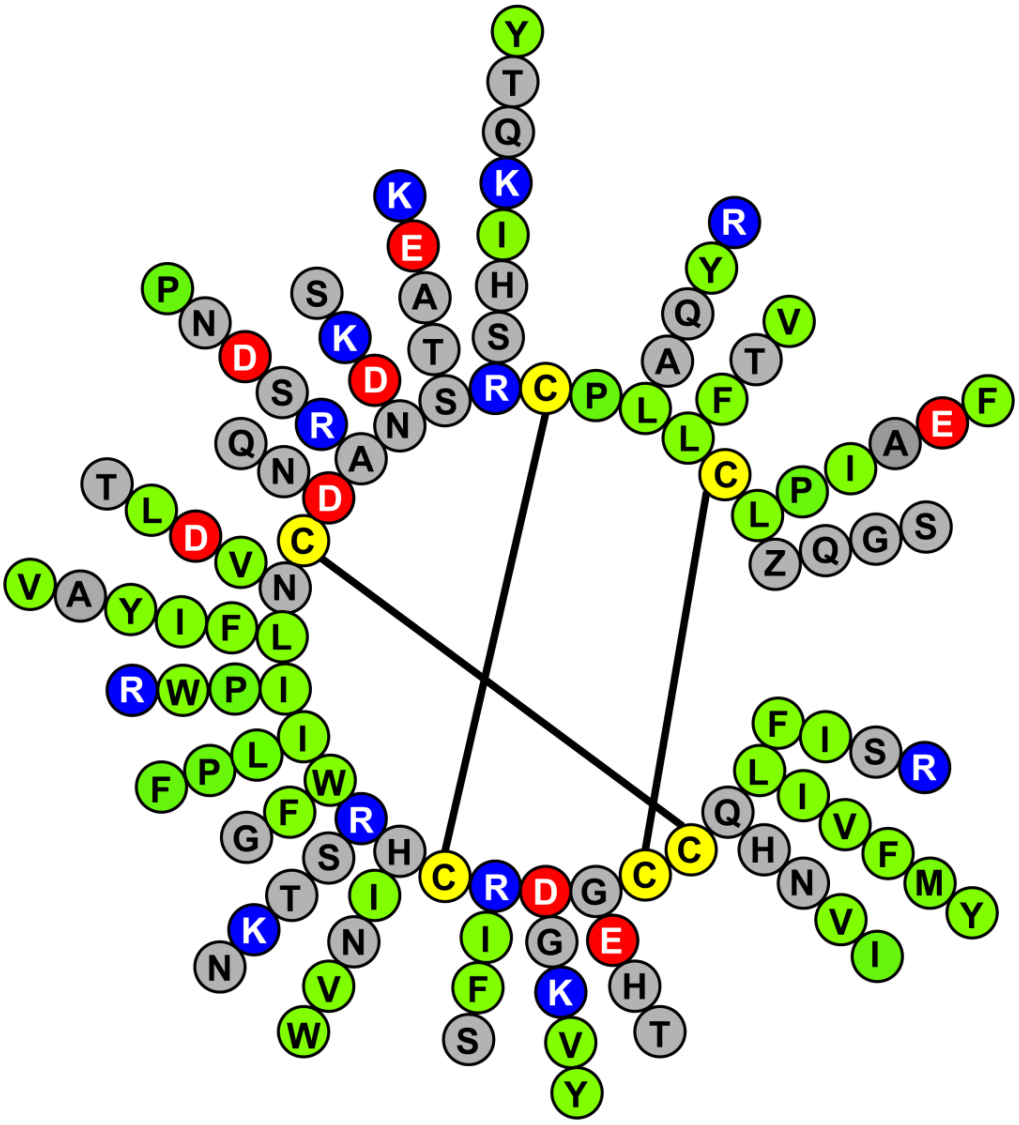


Figure 3.19. Diversity wheel of jasmintides. Diversity wheel represents the sequence diversity of jasmintides drawn from a multiple sequence alignment, where the consensus sequence is positioned in the inner circle and the spike protruding from each position represents the amino acid variation observed at that position.

Some of the jasmintides such as jS9-jS13 are not detected at the proteomic level, which could be because their expression is affected by the environmental change or they are processed using unknown pathways. However, those expressed at proteomic level showed the molecular diversity of jasmintides in terms of truncated peptides. The existence of truncated peptides is also a mean of diversification. For example, >100 truncated conotoxins are generated from one single gene MrIA [14, 190]. Similarly, partial sequences of cyclotides from *Clitoria ternatea* also confirm the natural existence of truncated peptides [133].

The reason for such molecular diversity is not known. Although, we also do not fully understand the physiological role of these jasmintides in plants, but this diversity suggests that plants evolve for better adaptation to the environmental stress. We also speculate that the production of less abundant jasmintides in plants might indicate that they have increased biological function for which plants need them in small amount only, being so bioactive peptide.

3.2. Biosynthesis of Jasmintides

Genetic characterization revealed that a novel jasmintide gene encodes three major domains. The ER signal sequence plays a role in directing the protein to the secretory pathway [191]. Signal peptidase I (SPase I) is involved in cleaving the signal peptide from the precursor to release the pro-peptide. The knottin family of *α*-amylase inhibitors is cleaved between a Leu/Ile/Val-Xaa, where Xaa is a small amino acid. In comparison, the *N*-terminal cleavage site of a jasmintide precursor occurs between Asn and Gln/Gly residues (Figure 3.20) which could be mediated by an asparaginyl endopeptidase enzyme (AEP) [192]. These characteristics show that jasmintides are ribosomally-synthesized peptides [78].

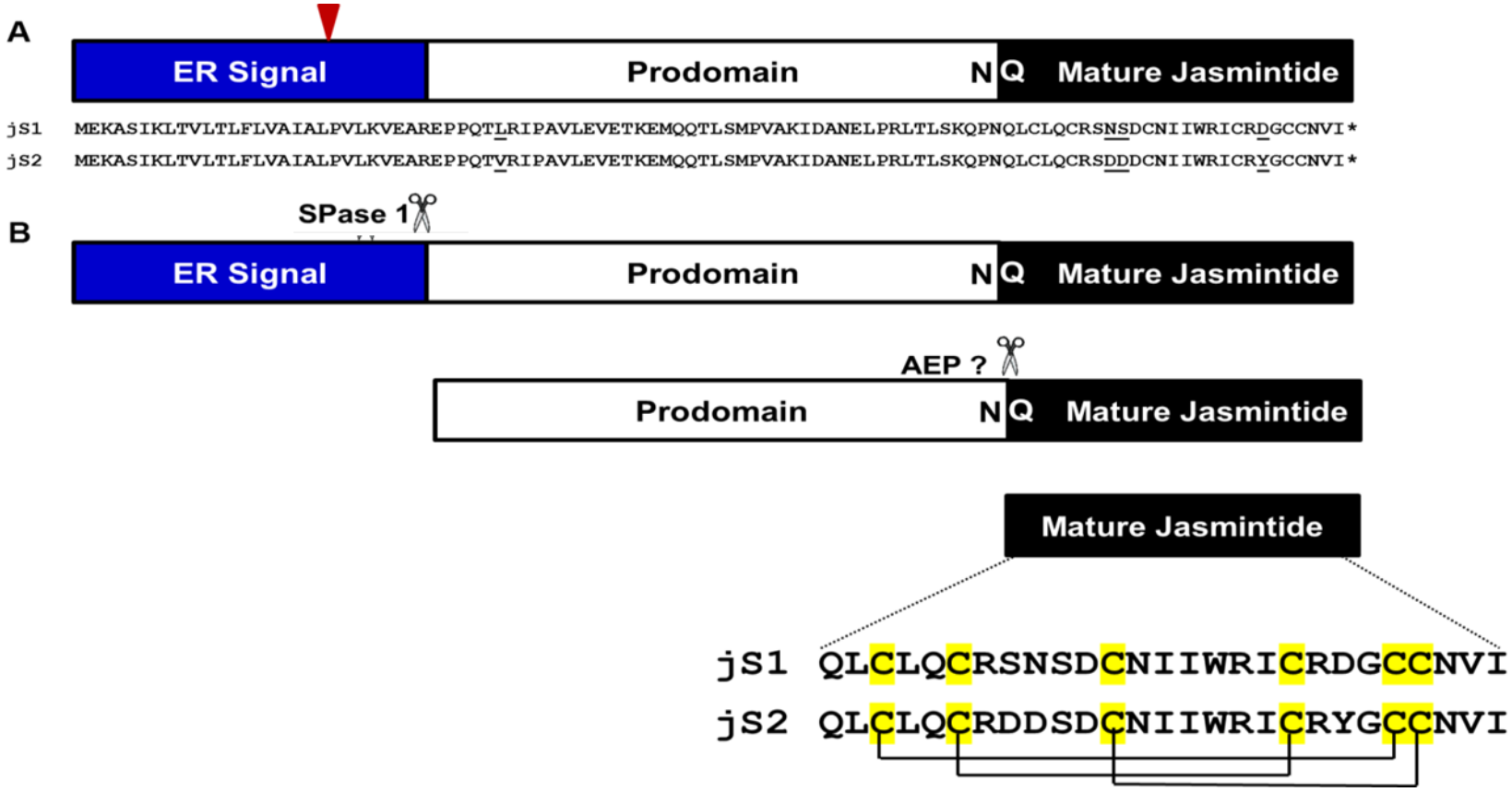


Figure 3.20. Precursor and putative biosynthesis of jasmintides. A) Alignments of deduced amino acid sequences of jasmintides precursors. The endoplasmic reticulum signal peptide was predicted using SignalP 4.1. The intron is depicted by a triangle within the endoplasmic reticulum signal peptide. B) Biosynthetic pathway of jasmintides.

The three-domain precursor is commonly found in many plants families producing cystine-knot peptides in the Rubiaceae, Violaceae, Solanaceae, Poaceae, and Apocynaceae families [35, 42, 78, 187, 188]. Jasmintides are found to contain a pyroglutamic acid at their *N*-terminus. It remains to be determined whether the conversion of glutamine to pyroglutamic acid occurs spontaneously or is catalyzed by a specific glutamyl cyclase. A similar modification can be found in the squash trypsin inhibitor from *Momordica cochinchinesnsi* [193], Phyb M isolated from *Petunia hybrid* [194], and Ib-Amp from *Impatiens balsamin* [195]. The incorporation of an *N*-terminal pyroglutamic acid unit could be an alternate strategy for backbone cyclization to protect jasmintides against exopeptidase degradation.

Among jasmintides, remarkable homology is found between jS1 and jS2 precursors (Figure 3.20) which suggest jasmintides have evolved from gene duplication and mutation events within the species through a process known as evolutionary divergence. Plants adopt the divergence process to increase the gene population of important molecules for better adaption to climate changes. Such homologous genes have also been reported previously. For example, α-amylase inhibitors As1 and As2 from *A. scholaris* which share 97% sequence homology [91].

3.3. Disulfide Connectivity

Disulfide bonds play a critical role in stabilizing the peptide structures. They aid in the correct folding of proteins by forming disulfide bridges [196]. The primary sequence of jasmintide jS1 showed that it contains six cysteine residues, two of which are adjacent with each other forming a CC motif. The disulfide connectivity of a CC motif in CRPs is a challenge to unravel because there is no enzyme known to cleave the CC bond for fragmenting a reduced peptide to be sequenced by a proteomic method. Thus, a method involving partial reduction coupled with differential alkylation and enzymatic digestion was adapted to map the disulfide bond of jasmintide. Previously, disulfide bond mapping of CRPs have been assessed by regional partial reduction and differential alkylation, where the fully-reduced and *S*-alkylated fragments are further digested with enzymatic treatment and sequenced [188]. This strategy often meets with difficulty when there is a CC motif with an adjacent cysteine.

Our early attempts to map the CC motif-containing jS1 failed as all disulfide bonds of the CC motif were reduced simultaneously. Later, with a modified method involving careful partial reduction and alkylation, we obtained intermediate 2*SS* with two disulfide bonds intact; suggesting that *S*-*S* bridge between CysI-CysV was the first to be broken. Next, two cystine linkages were determined by enzymatic digestion of 2*SS* intermediate to generate multiple disulfide bond-containing tryptic peptides in the presence of urea. The use of urea during enzymatic hydrolysis was necessary as trypsin was not able to cleave the partially alkylated two-NEM-tagged specie (2*SS*) in the urea-free solution.

With this method, the connectivity pattern depicted for jS1 was CysI-CysV, CysII-CysIV and CysIII-CysVI. Further validation by NMR, confirms this cysteine pairing due to the presence of $d_{\beta\beta}$ (*i.j*) NOEs between them (Figure 3.12 and 3.13).

In plants, for CRPs containing 6 cysteines (6C-) only two type of disulfide connectivity has been reported so far (discussed in detail in Chapter 1). Peptides with additional disulfide bonds (8C- or 10C-) are extensions of either Type I (Cystine-knot) or Type II (Thionin-like) connectivity. For example, Hevein-like peptides exist as 6C-, 8C-, or 10C- in nature. The additional bond does not disturb cystine-knot in their structure. Jasmintide jS1 does not share Type I (cystine-knot) connectivity: CysI-CysIV, CysII-CysV and CysIII-CysVI (cystine-knot arrangement) neither this disulfide connectivity matches with Type II (Thionine-like) connectivity (CysI-CysVI, CysII-CysV and CysIII-CysIVI) [197, 198]. This disulfide linkage has not been reported in plants; hence, revealing a new CRP family with Type III (Jasmintide-like) connectivity in plants. It still needs to be determined whether the members of this family also exist as 8C- and 10C-jasmintide or are limited to 6C- only (Figure 3.21).

Apart from new disulfide connectivity pattern in jasmintides, they also have different cystine motif in comparison to other plant CRPs. For example, the CC motif found in thionins is located at the N-terminal CysI and CysII, whereas in α-amylase inhibitors and hevein-like peptides CC motif is found in the middle of the sequence as CysIII and CysIV. In contrast, jasmintides have CC motif towards their C-terminus CysV and CysVI (Figure 3.22).

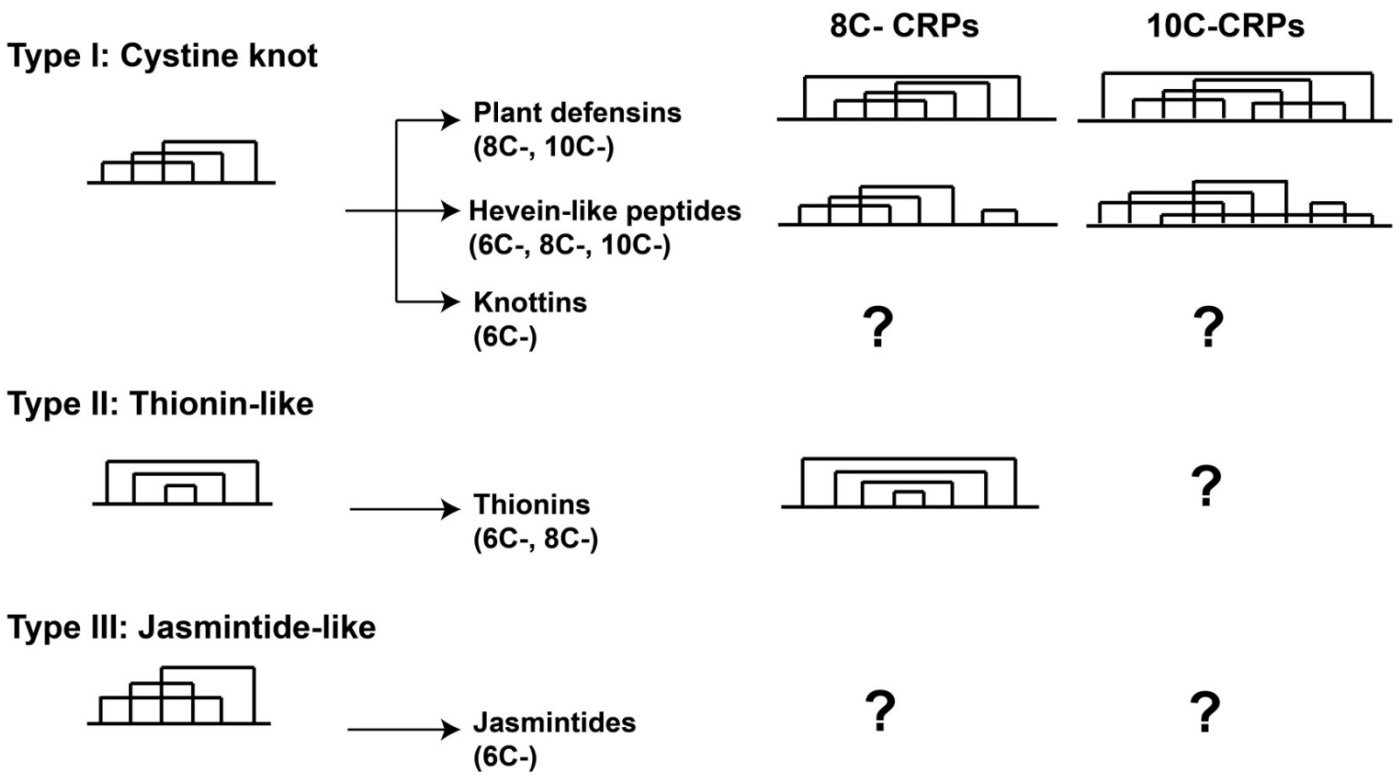
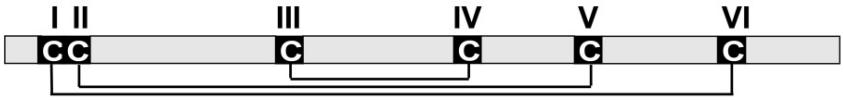


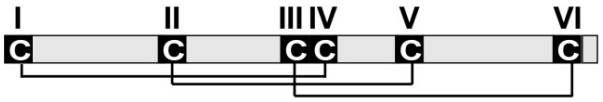
Figure 3.21. Disulfide connectivity of plant CRPs. The figure shows three types of cystine mapping. Type I and Type II are known in plant CRPs whereas Type III is a new discovered CRP family.

A. Thionins




```
Viscotoxin-A3    KSCCPNTTGRNIYNACRLTGAPRPTCAKLSGCKIISGSTCPSDYPK
Viscotoxin-B     KSCCPNTTGRNIYNTCRLGGGSRERCASLSGCKIISASTCPSDYPK
Br-thionin       KICCPRTIDRNIYNACRLTGASMTNCANLSGCKIVSGTTCPPGYTH
```

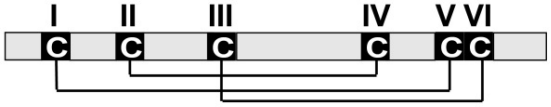
B. Cystine knot α-amlyase inhibitors & Hevein-like peptides




```
Allatide aC1     CIAHYG-KCDGIINQCCDPWLCTPPIIG-ICI
Alsotide aS1     CRP-YGYRCDGVINQCCDPYHCTPPLIG-ICL
Wrightide w1     CAQ-KGETCS-VYLQCCDPYHCTQPVIGGICA

aSG1             APGQCNHGRCPSGLCCSQYGYCGTGPAYCG-
aSR1             -VGECVQGRCPPGLCCSRFGYCGTGPAYCG-
Ar-AMP           -AGECVQGRCPSGMCCSQFGYCGRGPKYCGR
```

C. Jasmintides



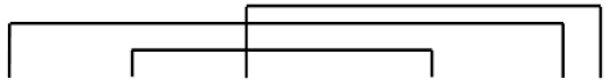

```
Jasmintide jS1   ZLCLQ-CRSNSDCN-IIWRICRDGCCNVI
Jasmintide jS2   ZLCLQ-CRSDDDCN-IIWRICRYGCCNVI
Jasmintide jS3   ZLCLL-CQTSRDCNYIIWTVCRDGCCNIS
```

Figure 3.22. Cystine motif and disulfide connectivity of CRPs. A) shows the cystine of thionins having CC motif at CysI and CysII. B) shows cystine motif of cystine-knot α-amylase inhibitors and hevein-like peptides with CC motif at CysIII and CysIV. C) shows cystine motif of jasmintides with CC at CysV and CysVI.

Searches from data banks from National Center for Biotechnology Information, DNA Data Bank of Japan homology and European Bioinformatics Institute showed jS1 shared 43.5% homology with a carboxypeptidase inhibitor from Ascaris *suum* (ACI). Importantly, these searches failed to uncover the cystine motif of jasmintides in plant CRPs. However, similar disulfide connectivity was found in human *β*-defensins [199] and in two scorpion toxins [200] (Table 3.8), all of which are substantially larger in size than jasmintides. None of the further 14 jasmintides (jS2-jS15) showed sequence homology to the other known plant CRPs.

Carboxypeptidase inhibitor ACI inhibits carboxypeptidase A (CPA). Based on the homology detected between jS1 and ACI the carboxypeptidase inhibitor activity of jS1 was tested against bovine carboxypeptidase A (bCPA). Nonetheless, jS1 does not inhibit bCPA up to 100 µM.

Table 3.8. Sequence alignment of peptides

Peptide	Sequence	References
Plant		
jS1	QLCLQ----CRSNSDCN---IIWRIC---RDGC-CNVI-	This work
jS2	QLCLQ----CRSDDDCN---IIWRIC---RYGC-CNVI-	This work
H. Defensins		
HBD1	DHYNCVSSGGQCLYS-ACPIFTKIQGTCYRGKAKC-CK---	[201, 202]
HBD2	VFGGIGDPVTCLKSGAICHPV-FCPRRYKQIGTCGLPGTKC-CKKP-	[203]
HBD3	GIINTLQKYYCRVRGGRCAVL-SCLPKEEQIGKCSTRGRKC-CRRKK	[204]
S. toxins		
k-BUTX-Tt2	GCMP--EYCAG--QCRGKVSQ-DYC-LKN--CRCIR--	[200]
Ts16	GCMK--EYCAG--QCRGKVSQ-DYC-LKH--CKCIPR-	[200]

The jasmintide jS1 and jS2 have been aligned against three human beta defensins (HBD1-3) and two scorpion toxins (κ-BUTX-Tt2, and Ts16) based on similar disulfide connectivitypattern.

3.4. Structural Features of Jasmintides

We also determined the three-dimensional structure of jS3 in collaboration with Professor Yoon Ho Sup. Figure 3.23 shows a comparison of the structures of jS1 and jS3. The backbones are superimposed with an RMSD of 1.16 Å. The main structural feature of jasmintides is that they have a compact core which comprises type III connectivity (CysI-CysV, CysII-CysIV and CysIII-CysVI) and three short extended β-strands which are linked by turns and disulfide bonds.

The structure of jS1 and jS3 are almost similar, but there are some interesting differences. JS3 elutes later than jS1 on RP-HPLC and it has less aqueous solubility. These both findings suggest that jS3 has more hydrophobic residues compare to jS1. Primary sequence analysis of jasmintides shows that the surface-exposed hydrophobic patch is conserved in jasmintides, mainly in loop 3. In jS3 the hydrophobic patch is extended with the presence of more hydrophobic residues Lue5, Thr21 in place of Glu5 and Arg20 and an additional Tyr17 in loop 3 (Figure 3.24).

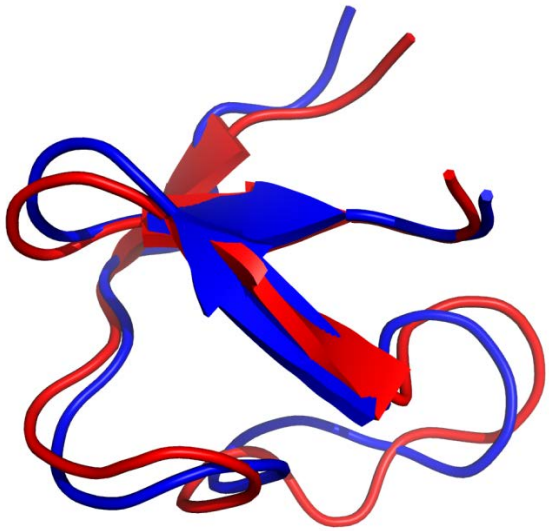


Figure 3.23. Structure overlapping of jS1 and jS3. The backbones of jS1 and jS3 are overlapped with RMSD of 1.16 Å.

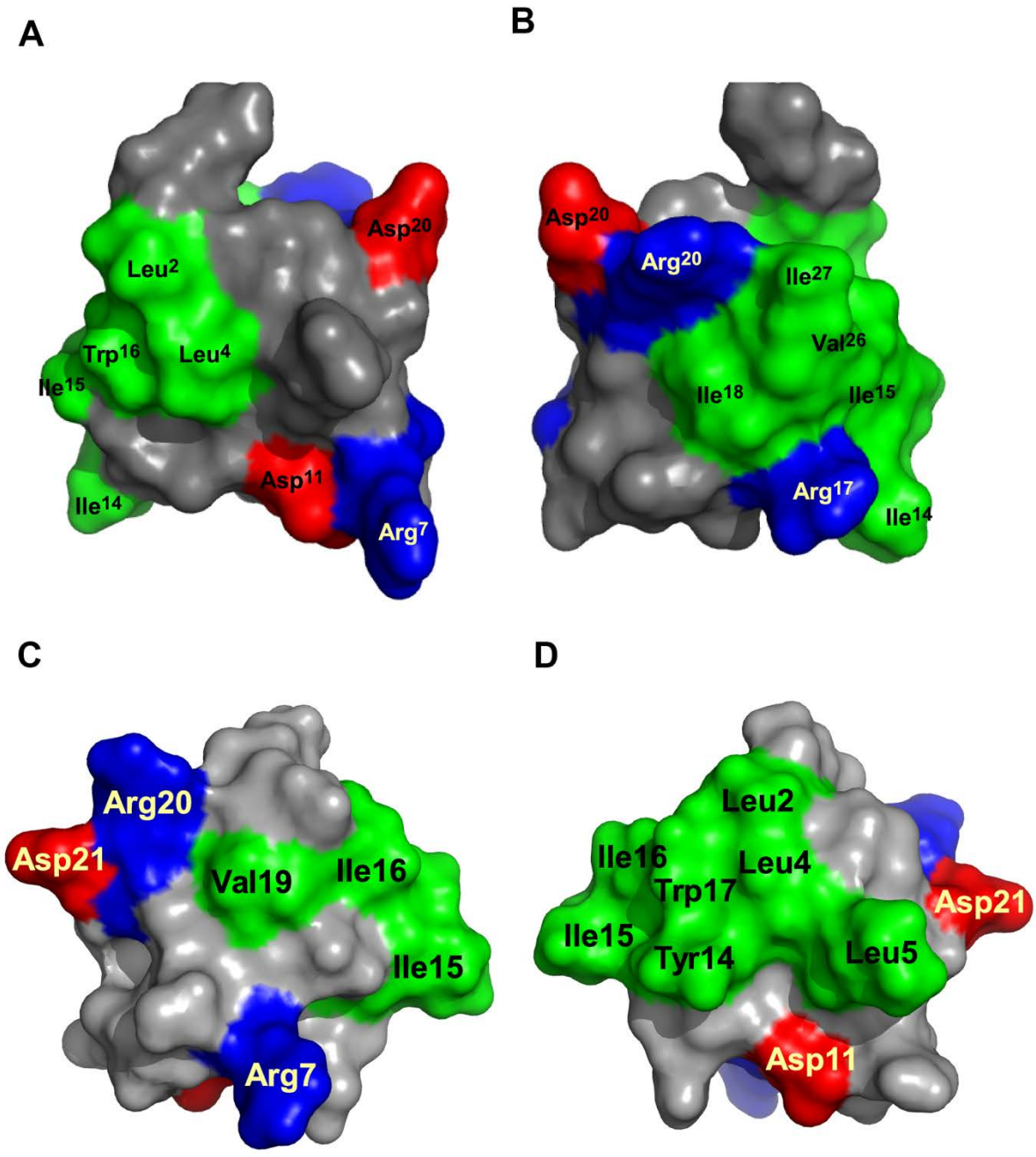


Figure 3.24. Stereo view of jS1 and jS3. A and C) showing the distribution of hydrophobic residues colored green, negatively charged residues colored red, positively charged residues colored blue, all other residues colored grey for jS1 and jS3. B and D), the views are rotated 180° about the vertical axis for jS1 and jS3, respectively.

3.5. Insecticidal Activity of Jasmintides

Insect pest control in plants is of great socio and economic importance. With increasing world population and decreasing availability of farmable land, there is urgent need to expand food production efficiently in crop plants. Molecular approaches to combat insect infestation include direct contact of the chemical or biological agent by spraying or their delivery via transgenic plants. Chemical agents are mostly small molecules ranging between few hundreds to thousand Daltons and biological agents are typically large proteins. CRPs are an intermediate class between chemical and biological insecticides and are not much focused on their insecticidal activity.

In this study, we have isolated a suite of jasmintides from *J. sambac* which raise the question of their natural role in plants. Defensive role of plants could be attributed to their antimicrobial activity against bacterial or fungal species, or it can be associated with digestive proteinase inhibition. Jasmintides did not show any antimicrobial activity against tested human and plant pathogenic strains. However, to further understand the role of jasmintides in a defense mechanism, we fed it to yellow mealworms. This study showed that insect larvae do not prefer to consume a diet containing jasmintides. The molecular basis of this insecticidal activity by jasmintides is not fully understood. There are few possible mechanisms for insecticidal activity by CRPs. First, it could be from inhibition of major digestive enzymes such as trypsin and chymotrypsin [78]. In particular, MCoTI-I and MCoTI-II from *Momordica cochinchinensis,* cysteine-knot trypsin inhibitors retard insect growth by impairing digestion of proteins [29].

The second possible mechanism is membrane disruption; however, it is unclear whether it will be a result of the formation of pores or simply a generalized disruption of the membrane. Many peptidic toxins, particularly those with surface exposed hydrophobic patches, have been shown to form pores in membranes, including melittin from bee venom [205]. NMR structure of jasmintides showed that the hydrophobic residues are exposed on the surface, but it is difficult to

speculate about their role for an insecticidal activity unless further studies are performed. Larvae fed on jasmintides containing diet have shown body weight reduction, also suggesting the possibility of irritation and necrosis of intestinal epithelium cells, which occurs by the continuous administration of jasmintides in the diet.

The function of jasmintides in plants has not been fully elucidated yet, but the high level of expression of jasmintides in leaves and flowers, as well as the production of several isoforms within a single plant, is consistent with a role in recognition or defense. Furthermore, plant defense molecules are often small cysteine-rich proteins (33) that are trafficked through the secretory pathway. These defense molecules include insecticidal molecules such as proteinase and α-amylase inhibitors as well as potent antimicrobial molecules such as thionins and defensins.

3.6. Antiviral Activity of Jasmintides against IBV

The emergence of viral resistance has posed a significant challenge. Many efforts have been laid to develop alternative antiviral chemotherapeutics. However, a big proportion of antiviral drugs are plant sourced and small molecules being the active component. The utility of peptidyl drugs has limitations due to poor bioavailability, cytotoxicity, and off-target effects. Our results describe the discovery of metabolic stable and non-cytotoxic jasmintides, with antiviral activity against IBV on Vero cells.

To date, only alostides from *A. scholaris* have been reported as an anti-IBV drug for infectious bronchitis virus among domestic fowl. The potent antiviral activity of jasmintides characterizes them as the second member of peptidyl drugs effective against IBV disease with an IC_{50} of 70 nM, and 86 nM for jS1 and jS4, respectively. The IC_{50} of As1 (6 µM) is almost 85 times higher than IC_{50} of jS1, which showed that jasmintides are more potent than alostides. Jasmintide jS1 and jS4 possess high sequence diversity in loop 2 and loop3. Despite such high

sequence diversity in jasmintides, they still retain the function which suggests antiviral activity could be associated with the conserved residues. However, further studies are needed to get insight into the mechanism involved in inhibition of activity.

3.7. Stability of Jasmintides

CRPs are well-known for their stability against thermal and enzymatic degradation [17, 35, 36]. This high resistance is an important feature for CRP peptide bioactives in traditional medicine preparations. Interestingly, the cross linked structure of jasmintides, conferred by three cystine linkages, are resistant to both thermal denaturation and proteolytic degradation, including exopeptidase treatment. High thermal stability means jasmintides will still be present in decoction after prolonged heat treatment for brewing, and high enzymatic stability shows that jasmintides are able to diffuse across our gastrointestinal tract. In addition, their stability in human serum showed that these do not degrade in human blood. These stability features make jasmintides comparable to cystine-knot peptides, such as α-amylase inhibitors [35, 36].

3.8. Potential Applications of Jasmitides

CRPs have been employed in nature as a scaffold for a variety of protein families found from animals, plants and microbes. Having the advantage of small size, stability to thermal and proteolytic degradation and high tolerance to sequence variation make them an appealing candidate for designing peptidyl drugs. The discovery of jasmintides which contain a unique cystine connectivity of CysI-CysV, CysII-IV, and CysIII-CysVI but their structural and molecular stability features are comparable to other CRP families such knottins, defensins, and heveins. Along with these features, jasmintides are notably smaller CRPs with 27-28 residues, which make them more accessible to chemical synthesis. Further, the peptidomic analysis suggests the high tolerance of jasmintides to sequence variations. Thus, CRPs, such as jasmintides with a molecular weight of

3 kDa possess remarkable characteristics as a scaffold for protein engineering in order to attain new function. A successful example includes grafting of bradykinin antagonist DALK or DAK in cyclotide kalata B1 scaffold [118].

Jasmintides also demonstrate insecticidal activity, which suggests their application as a natural insecticide to crop plants. Insight into the biosynthesis of jasmintides reveals that jasmintides are gene-encoded which lead to the possibility of production of transgenic plants in the same way as the *Bacillus thuringinesis* (Bt) have been employed for the protection of corn, soybean and cotton.

Today, no specific treatment is available for treating IBV infection; however, vaccination has remained the prime method to control the spread of infection. In addition, antibiotics are given to treat secondary infections, but not the acute bronchitis disease itself. Prior vaccination of chickens is an expensive and labor-intensive process and still provides moderate prophylaxis. Jasmintides with an antiviral function can be potentially used in poultry. They have two major advantages, firstly they are stable to proteolytic enzymes, thus, can be given orally. Secondly, they are isolated from the natural source, i.e., plant. Hence, the cost of production of jasmintides is much lower than the cost of the vaccine.

In summary, there are no six-cysteine-containing linear plant CRPs with a MW of 3 kDa and a cystine connectivity motif of CysI-CysV, CysII-CysIV, and CysIII-CysVI reported. Currently, the CysI-CysIV, CysII-CysIV, and CysIII-CysVI cystine-knot arrangement in the knottin family is the most common. The proline-rich, CC-bond-containing *α*-amylase inhibitors with 30-32 amino acids in the knottin family represent the closest relatives to jasmintides [36]. Other six-cysteine-containing knottin CRPs include both linear and cyclic cyclotides which contain 29-37 amino acids [42, 43, 188]. The isolation and characterization of a new family of CRPs with an unusual cystine motif from *J. sambac* of the Oleaceae family in this thesis serve to expand our knowledge of the occurrence

of cystine motifs in plant CRPs. Understanding the structure and stability of jasmintides could lead to their use as a scaffold for designing new peptide drugs.

Chapter 4

Designing of Metabolically Stable Substance P Antagonists

1. Introduction

Peptide therapeutics plays an important role as drug leads and pharmaceutical discovery tools because of their high specificity to target and less side-effects as compared to conventional medicines, which was dominated by small molecules (<1 kDa). They are emerging as a new class besides small molecules and proteins. They are potentially valuable therapeutics in the chemical space between 1-8 kDa, combining the stability of small molecules and specificity of proteins. These advantages enhance the application of peptides in drug development.

Developing peptide therapeutics has always been promising but challenging. However, the utility of peptides as therapeutics in the field of the pharmaceutical industry has limitations due to their poor stability and limited bioavailability. Linear peptides degrade rapidly into inactive fragments by proteolytic enzymes in the gastrointestinal tract (GIT) and by serum enzymes. As a result, the desired therapeutic effects are not exerted. This is particularly true for peptides when administered orally. Because of these limitations, intravenous injections remained the most common method for peptides administration. In order to improve their stability and bioavailability, several approaches are utilized such as conventional chemical modifications [206-208], introducing non-natural amino acid, cyclization [111] retro-inverso compounds [113], incorporating lanthionine modifications [209] and engineering of bioactive peptides into the stable natural scaffold [115-118].

Grafting bioactive epitopes into a stable scaffold is one of the novel approaches for increasing the peptide stability. CRPs are stable peptides found in nature from various species i.e., plants and animals. The presence of intramolecular disulfide bonds stabilize these against chemical, thermal and enzymatic environment [17]. CRPs have also shown tolerance for the wide range of sequence substitutions. These properties make CRPs a suitable scaffold for drug designing and protein grafting [210, 211].

1.1. Sunflower Trypsin Inhibitor

Sunflower trypsin inhibitor SFTI-1 is a cyclic peptide containing 14 amino acid residues and has been isolated from sunflower seeds, *Helianthus annus*. It is the most potent and the smallest trypsin inhibitor known (Ki=0.1 nM) [92]. SFTI-1 has high sequence homology to the trypsin reactive site from Bowman-Birk Inhibitors (BBIs), a class of serine protease inhibitors [93]. The tertiary structure consists of two anti-parallel short β-strands which are connected by two turns at each end and braced centrally with a single disulfide bond between Cys3 and Cys11. The hydrogen bonds in the crystal structure (PDB: 1SFI) are responsible for its well-defined structure. SFTI-1 has two loops, trypsin inhibition (TI) loop and the secondary loop [212]. The reactive sites of binding loop comprise reactive site TI loop between Lys5 and Ser6 [212], and are responsible for its trypsin inhibition activity (Figure 4.1). The cyclic backbone structure of SFTI-1 prevents its enzymatic degradation by exopeptidases [212] and its ability to form β-turns within a cysteine-rich framework demonstrates that it can be used as a scaffold for drug design application [212]. *Chan et al*., reported the grafting of angiogenic VEGF peptide into SFTI-1 scaffold by replacing trypsin binding loop. This grafted peptide has shown stability in human serum and produced angiogenesis at nano-molar concentrations [117]. Thus, extensive stability and biopharmaceutical property of SFTI-1 promise it as an excellent scaffold for stabilization of foreign bioactive peptides.

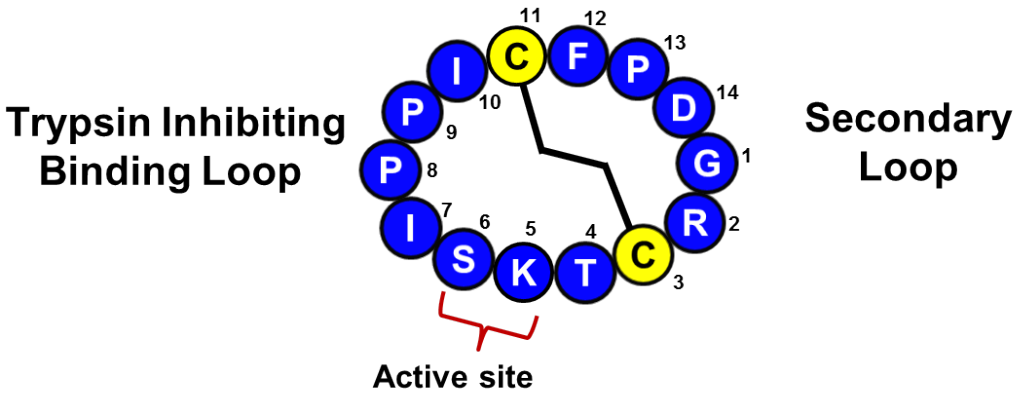


Figure 4.1. Schematic structure and amino acid sequence of SFTI-1. The sequence is represented with one letter amino acid code. The disulfide bond is shown with a bold line. Amino acids are numbered from N- to C-terminal as reported by *Luckett et al.* [92].

1.2. Substance P (SP)

SP was first discovered by Ulf von Euler and John. H. Gaddum in 1931. Initially, it was isolated from the horse's small intestine and brain and found to induces smooth muscle contractions and reduces hypertension [213]. Later in 1934, *Gaddum and Schild et al*., named this preparation as 'Substance P' and 'P' refers to the crude extract powder acquired after the extraction [214]. In 1953, *Lambeck and Pernow et al.,* suggested that SP is a primary neuronal sensory transmitter which induces pain and its higher concentration was found in the spinal cord [215, 216].

In 1971, Leeman and coworkers confirmed the amino acid sequence of undecapeptide SP (RPKPQQFFGLM) in the bovine hypothalamus [217]. After a decade, SP was recognized as a member of tachykinin family [215]. Tachykinin family shares a common C-terminal sequence FXGLM-NH_2 (X is Phe or Val) [218] and varying N-terminal sequences. Many of the tachykinins have been discovered, but only those specific to mammals are known as neurokinins. Along with SP, Neurokinin A (NKA) and Neurokinin B (NKB) also exist in nature. The varying N-terminal sequences are responsible for recognition of the specific neurokinin receptor [215]. Three different neurokinin receptors have been characterized. SP preferentially binds with NK1 receptor, whereas NKA and NKB bind to NK2 and NK3 receptors, respectively. All these receptors are G-protein coupled receptors [219, 220].

Primarily, SP was considered to be the nociceptive transmitter in afferent sensory fibers [215] which is released in response to noxious stimuli and participate in conduction across sensory afferent nerves (C-fibers). Later, the role of SP in transmitting noxious sensory information to the brain was confirmed by identification of NK1 receptors [221, 222]. SP and NK1 are present in small sensory neurons throughout the neuraxis responsible for the integration of pain, stress, and anxiety [220].

1.3. Substance P in Inflammation and Nociception

It is believed that SP together with other tachykinins is responsible for nociception. The structural basis for such hypothesis was the fact that SP is synthesized in small sensory neurons. Injury or tissue damage causes the release of inflammatory mediators such as bradykinin and prostaglandins, which activate nociceptors. Activation of nociceptors promotes the release of SP and CGRP (calcitonin gene related peptide) from C-fibers [223, 224]. In the vicinity of sensory endings, SP acts on mast cells to induce degranulation to release histamine, which directly excites nociceptors. In response, SP causes plasma extravasation and CGRP causes dilation of peripheral blood vessels.(Figure 4.2) The resultant edema enhances the bradykinin release [225]. More than one mechanism is involved in the slow or prolonged release of SP. It has been suggested that in response to prolong noxious stimuli SP and excitatory amino acids are released which bind to the NK1 and NMDA (N-Methyl-D-aspartate) receptors. Excitations of receptors cause penetration of calcium ions into the neuronal cells. Nitric oxide synthases is activated and thus nitric oxide (NO) is produced (Figure 4.3). Nitric oxide plays a role in increasing nociceptive transmission. Long term excitation of receptors activates secondary messenger system involving changes in G-protein, which leads to alterations in neuron excitability [226, 227].

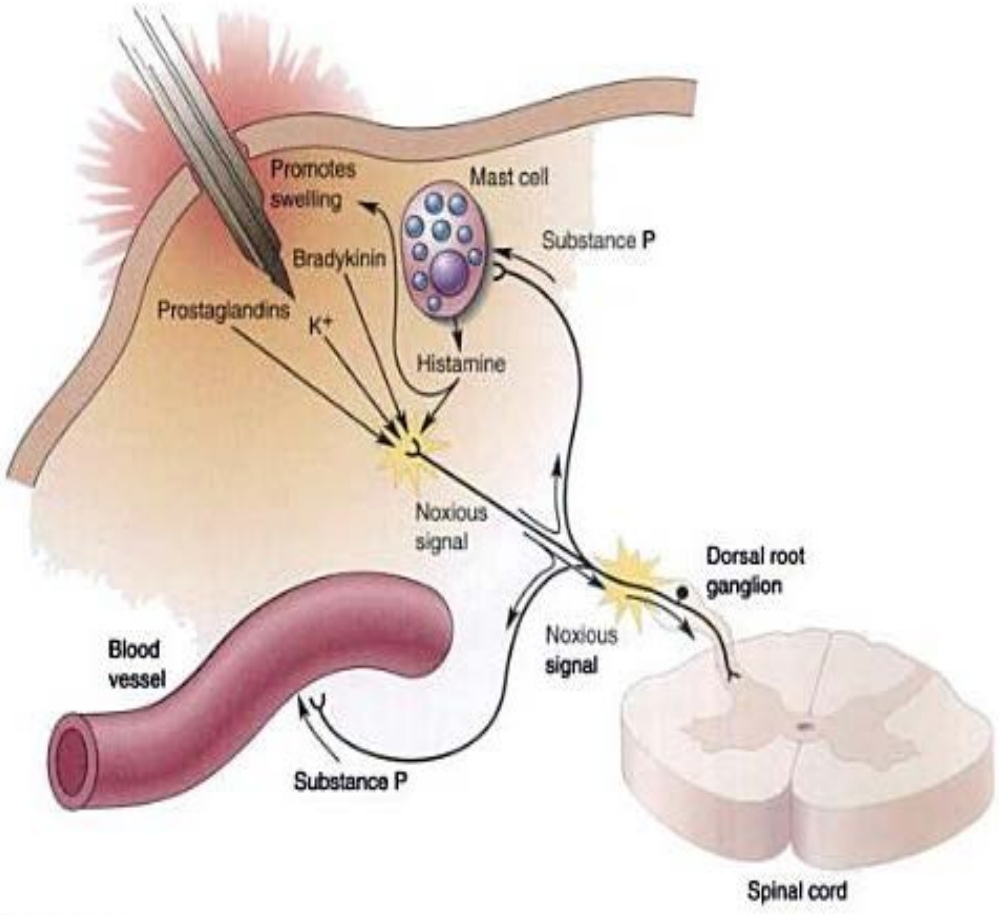


Figure 4.2. **The release of substance P and nociception.** In response to Injury or tissue damage inflammatory mediators such as bradykinin and prostaglandins are release, which activates nociceptors. In turn, activation of nociceptors promotes the release of SP and CGRP (calcitonin gene related peptide) from C-fibers. In the vicinity of sensory endings, SP acts on mast cells to induce degranulation to release histamine, which directly excites nociceptors. In response, SP causes plasma extravasation and CGRP causes dilation of peripheral blood vessels. The resultant edema enhances the bradykinin release. Figure adapted from reference [225].

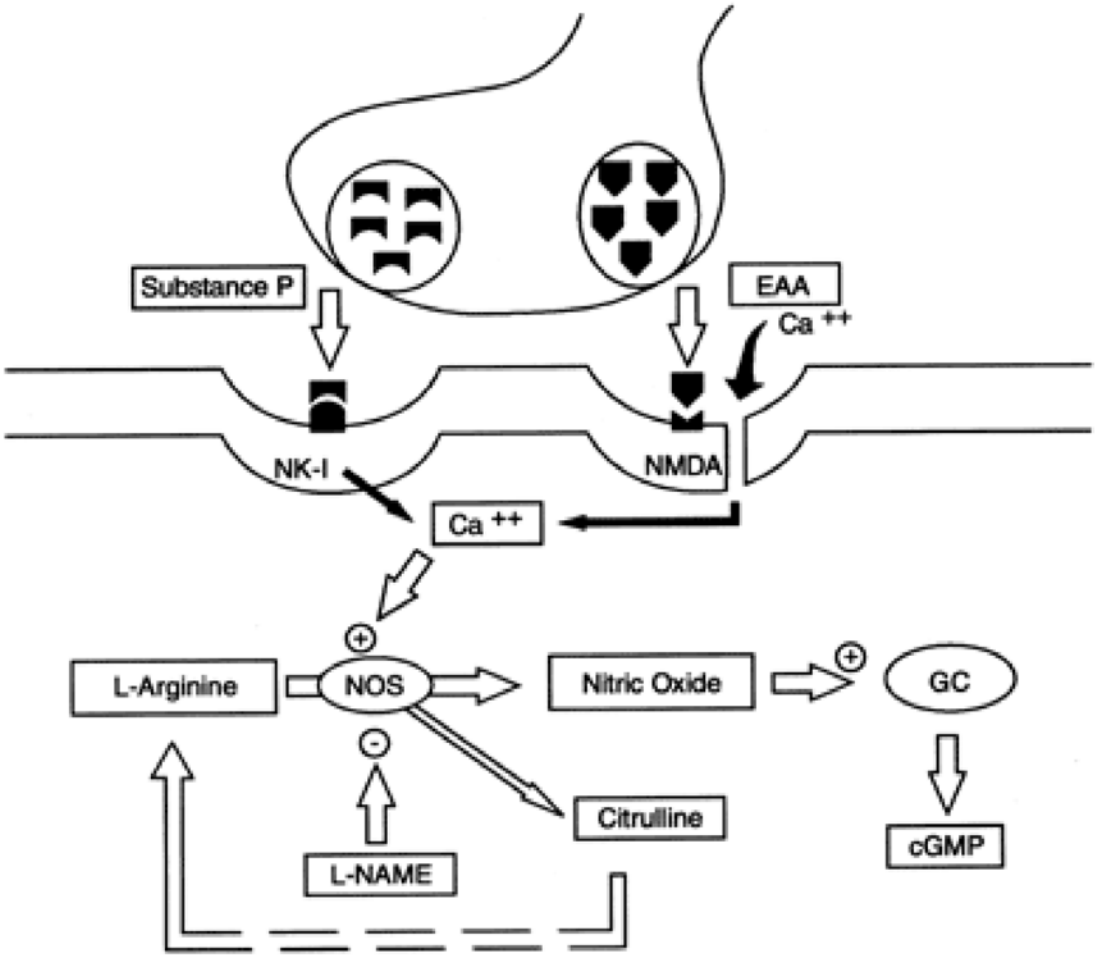


Figure 4.3. Schematic representation of steps involved in the release of Nitric oxide followed by the release of SP and excitatory amino acids (EEA). NMDA (N-methyl-D-aspartate), NOS (Nitric oxide synthase), NK1 (Neurokinin-1 receptor). Figure adapted from reference [227].

1.4. Substance P Antagonists

Efforts have been taken in past years to develop SP receptor antagonists to understand the physiological role of endogenous ligand SP. Various peptidyl and non-peptidyl SP antagonist have been developed. *Folker et al*., was the first to synthesize many SP analogs as SP antagonists using the approach of substitution of a natural amino acid (L-amino acid) by an unnatural amino acid (D-amino acid) [228]. Some of the examples are mentioned in Table 4.1. Later in 1991, *Sinder et al,* reported the first non-peptide SP antagonist (CP96.345) [229]. Afterward, many non-peptidyl SP receptor antagonists have been developed.

1.5. N-Terminal SP_{1-7}

N-terminal SP_{1-7}, hexapeptide (RPKPQQF), is the major bioactive fragment [230], which is released from the undecapeptide SP after enzymatic degradation [231, 232]. The N-terminal fragment SP_{1-7} is widely distributed in central nervous system (CNS) and possess opposed biological effects of the SP [230, 233, 234]. For instance, SP_{1-7} is responsible for antinociception in CNS [235-237], anti-inflammatory response [238] and antihyperalgesic effects in diabetic mice reflecting its involvement in neuropathic pain [239]. It was also shown to inhibit opioid withdrawal effects stimulated by NK-1 receptor agonists [240, 241].

Table 4.1. List of SP Antagonists

S.#	SP Antagonists	References
1	N-Fragment SP (1-7)	[230]
2	(D-Arg^{1}, D-Trp$^{7, 9}$, Leu^{11}) SP	[242]
3	[N$^{\alpha}$ - Z-Arg^{1}, N^{E}-Z-Lys^{3}, D-Trp$^{7, 8}$, D-Met^{11}] SP-OMe	[243]
4	[N$^{\alpha}$ - Z-Arg^{1}, N^{E}-Z-Lys^{3}, D-Pro$^{9, 10}$] SP-OMe	[243]
5	(Arg^{6}, D-Trp^{10}) SP (6-11)	[244]
6	(D-Pro^{4}, D-Trp$^{7, 9}$) SP (4-11)	[244]
7	[D-Arg^{1}, D-Nal^{5},D- Trp$^{7, 9}$ Nle^{11}] SP	[245]
8	[D-Arg^{1}, D-Cl$_{2}$Phe^{5}, Asp^{6}, d-Trp$^{7, 9}$, Nle^{11}] SP	[245]
9	(D-Arg^{1}, D-Pro^{2}, D-Trp$^{7, 9}$, Leu^{11}) SP	[246]
10	(D-Pro^{2}, D-Trp$^{7, 9}$)SP	[246]
11	(D-Pro^{2}, D-Phe^{7}, D-Trp^{9,})SP	[246]
12	(D-Pro^{2}, D-Phe^{7}) SP (1-7)	[247]

The mechanism for N-terminal SP_{1-7} is not fully understood. *Stewart et al.*, has reported that the antinociceptive effect of SP_{1-7} is blocked by naloxone [235]. Initially, SP_{1-7} was thought to exert effects through μ-opioid receptor [238]. However, *Botros and Igwe et al*., suggested that there are specific binding sites for SP_{1-7} in the CNS, which are distinct from the known opioid receptors [248-250]. For instance, these binding sites are present in the ventral tegmental area (VTA) in the spinal cord [250] and rat brain [248]. Recently in 2010, *Carlsson and Oshawa et al*., reported that the effects produced by SP_{1-7} were thought to be intervened through indirect activation of naloxone-sensitive sigma σ1 receptors [239]. However, SP_{1-7} is not an active ligand for the σ1 receptor. There has been no report of SP_{1-7} biological effects which are produced by any of the known tachykinin or opioid receptors [239]. Hence, the effects of SP_{1-7} are generated through an undefined receptor present at the specific binding sites found in the brain and spinal cord of rat and mouse [248, 250]. No successful attempt for cloned SP_{1-7} receptor has been reported so far. Here, we report the development of four SP antagonists by grafting SP_{1-7} analogs into the SFTI-1 scaffold. The engineered peptide was expected to be bifunctional, possessing both trypsin inhibition and substance P antagonistic activity.

1.6. Chemical Synthesis of Peptides

1.6.1. Solid Phase Peptide Synthesis

In 1963 Merrifield introduced solid phase peptide synthesis (SPPS) [251]. In SPPS, a fully protected peptide chain is assembled stepwise on an insoluble polymer support. Initially, an N^{α} protected amino acid is coupled through its C-terminus to the solid support of beads [251]. The advantage of this method is that the excess amino acid and reagents are flushed through filtration and washing. The N^{α} protecting group of first amino acid is deprotected, and the consecutive N-terminal protected amino acid is then reacted with existing amino acid bound on the resin support. The amino acid coupling cycles are repeated till the desired peptide sequence is attained. At the end, the peptide chain bound to the resin

(solid support), N-terminal protecting and orthogonal side-chain protecting groups are cleaved to afford full-length crude peptide [252] (Figure 4.4). SPPS assembles peptides in C→N terminus direction, unlike *in vivo*, where protein synthesis proceeds in opposite N→C terminus direction.

Two different approaches of the α-amino protecting group are employed in SPPS including 9-flourenylmethoxycarbonyl (Fmoc) and Tert-butyloxycarbonyl (Boc). Both approaches use different methods for deprotection of protecting groups and cleavage. In Boc strategy, 50% TFA is used to deprotect Boc protection group and HF or TMFSA is used for cleavage [253] whereas in Fmoc strategy 20% piperidine in DMF is used to remove Fmoc protection group and TFA is used for the final cleavage [254, 255].

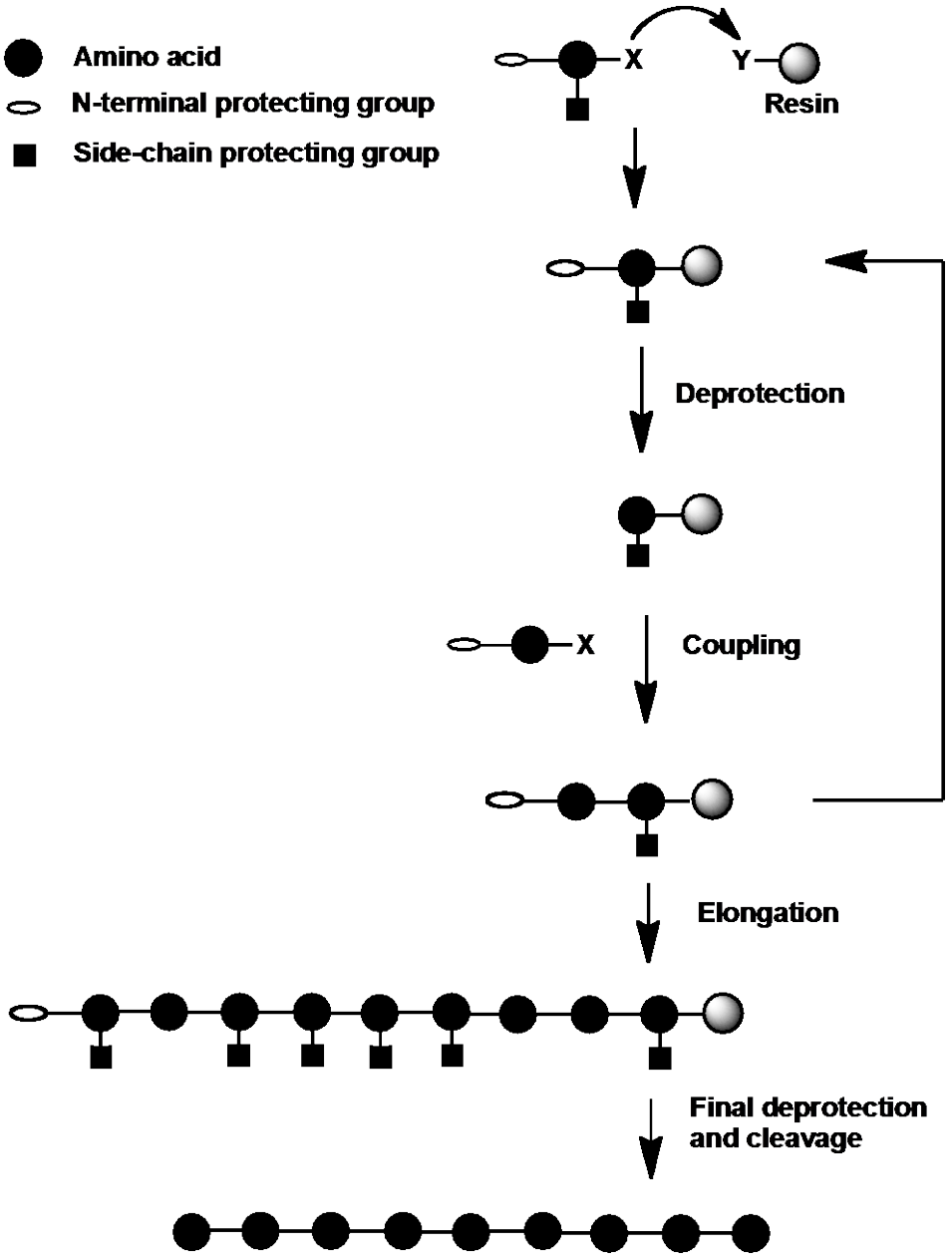


Figure 4.4. Schematic representation of procedure of solid phase peptide synthesis.

1.6.2. Cyclization

Cyclic peptides are commonly prepared using an approach of unprotected peptides containing N-terminal cysteine and C-terminal thioester [256-258]. This approach is commonly known as native chemical ligation or cysteine thioester ligation [258, 259] (Figure 4.5). In the past years, different methodologies have been encountered for preparing thioesters using SPPS using the Boc and Fmoc chemistries. In Fmoc chemistry, thioesters are susceptible to piperidine deprotection. In Boc chemistry, a repetitive acid deprotection step is an effective strategy for peptide thioester preparation, but it is not suitable for synthesis of peptides which are sensitive to acidic pH such as phospho or glycopeptides. Furthermore, a specialized apparatus is required for final cleavage step using HF, hence, limiting the use of Boc chemistry for preparing peptide thioesters.

To overcome these challenges, Fmoc-compatible methods for preparing thioesters have been reported which includes the use of less basic Fmoc-deprotection cocktail [260], direct conversion to thioester from C-terminal acid [261], “safety-catch” auxiliaries [262], *N*-acylurea [263], acid hydrazide [264], and O/N-S acyl shift methods such as N-4,5-dimethoxy-2-mercapto-benzyl (Dmmb) [265], N-alkylated cysteine [266], bis-(sulfanylethylamino) (SEA) [267], SEA-anilide [268], thio-proline [269] and thiazolidine [270]. A common functional group in these approaches is the thioethylamido moiety (TEA), which facilitates proximity-driven N-S acyl shift reaction via five-member ring expansion to obtain thioester after thiol-thioester exchange reaction with an external thiol. An example of TEA-thioester surrogate is MeCys-thioester surrogate [266, 271], which has several limitations such as coupling of C-terminal amino acid with MeCys attached resin is slow due to the low reactivity of the secondary amine on MeCys residue and other disadvantage is β-elimination side reaction of a protected MeCys residue due to piperidine addition during peptide elongation [272]. To avoid this side reaction, Tam’s laboratory has proposed a new approach for preparing TEA-thioester surrogate using thioethylbutylamido (TEBA) moiety [273] which is a simplified form of MeCys residue and retains TEA

group instead of MeCys residue and also eliminating electron-withdrawing carbonyl group of MeCys to minimize β-elimination side reaction completely. This methodology is adopted in this thesis for the synthesis of grafted peptides. The presence of TEBA-thioester surrogate at C-terminus with N-terminal cysteine makes it an ideal candidate for peptide cyclization. Cyclization of peptides involves series of N-S, S-S and S-N acyl shift reactions. However, this method involves two purification steps which make it laborious and time-consuming. A more efficient synthesis using one-pot synthesis scheme would facilitate high yield in less time.

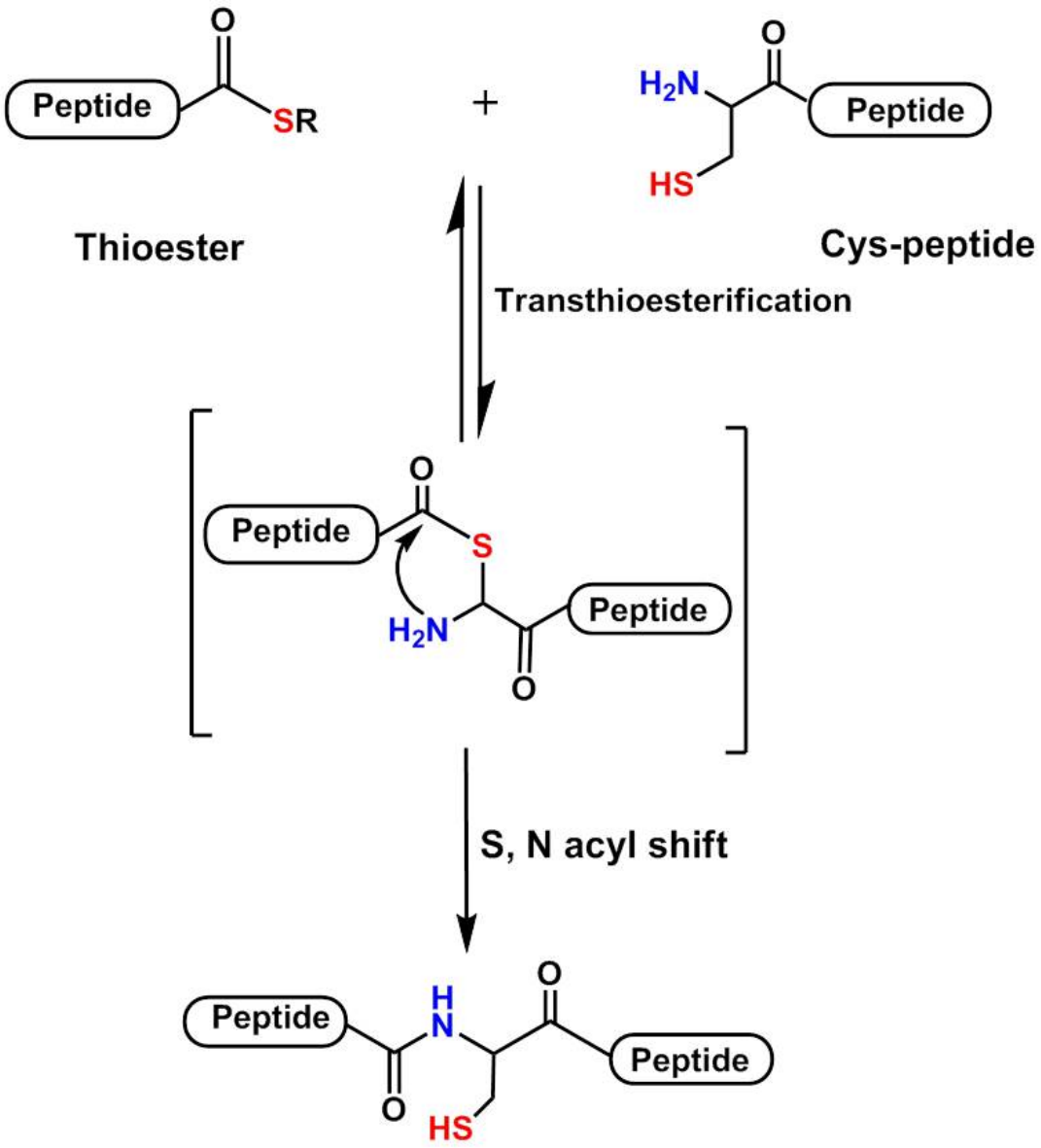


Figure 4.5. Schematic representation of native chemical ligation or Cys-thioester ligation.

In 1999, *Tam et al*., proposed the theory of thia-zip mechanism involved in macrocyclization of CRPs [257], which require N^{α} Cysteine, C-terminal thioester and at least one free internal thiol group. Cyclization of CRPs containing TEBA-thioester surrogate fulfill the needs of proximity-driven thia-zip cyclization and the side chain thiols of CRPs replace the role of external thiols in S-S acyl shift reaction and gives series of thiolactone intermediates instead of C-terminal thioester. The N-terminus cysteine joins the C-terminus of the peptide to form the amide bond by expansions of thialoactone rings via reversible S-S acyl shift reactions, followed by irreversible S-N acyl shift reaction (Figure 4.6) [257].

Here, we also report a one-pot synthesis of cyclic peptides using TEBA- thioester surrogate without utilizing external thiol. Peptides are cyclized using thia-zip cyclization strategy and directly subjected to oxidation, avoiding the purification or lyophilization of intermediate products. The resulting synthetic peptide is obtained in high yield compared to conventional method.

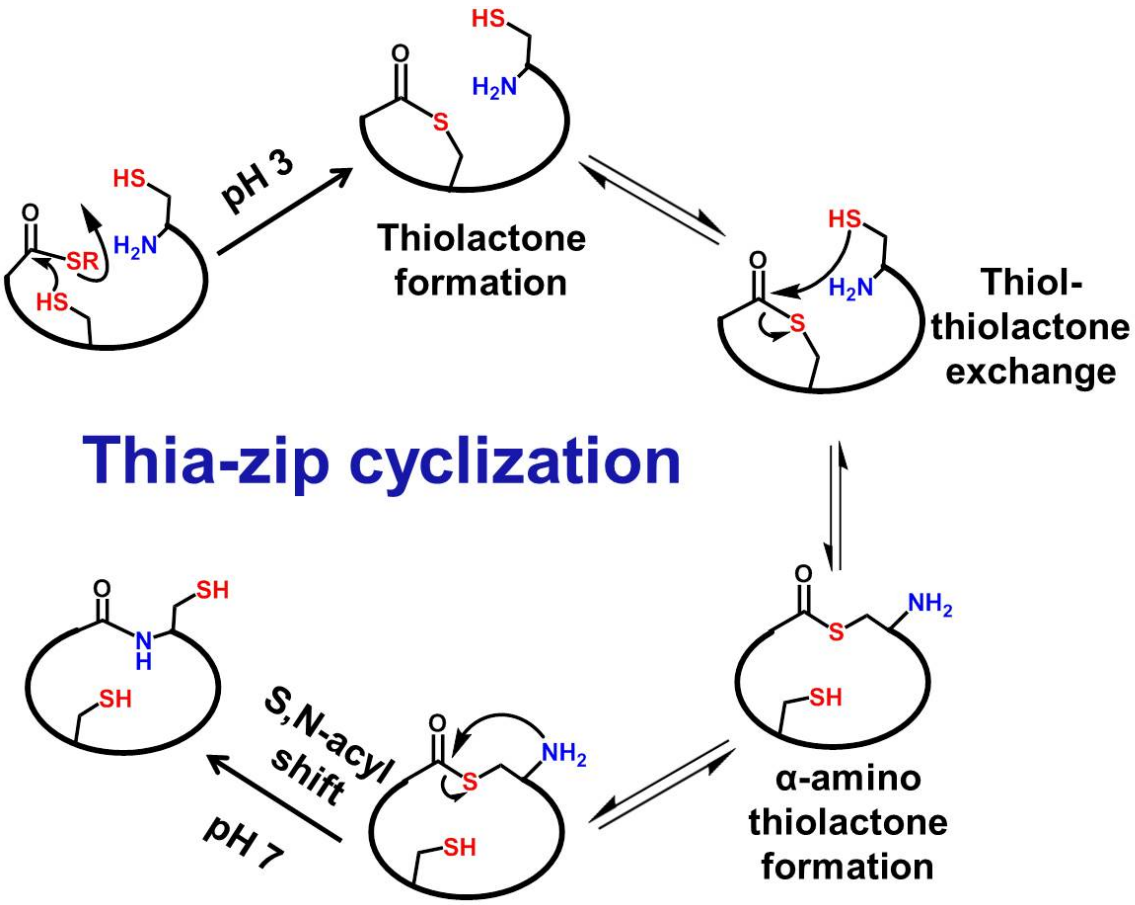


Figure 4.6. Schematic illustration of thia-zip cyclization. Thiolactone intermediates were formed and the ring structure expanded through the thiol-thiolactone exchange. The N-terminus cysteine joins the C-terminus of the peptide to form the amide bond via a proximity-driven S, N-acyl shift. Since the peptide bond formation was irreversible, the last expansion reaction was driven towards the product side efficiently.

2. Design and Synthesis of Grafted Peptides

2.1. Design of Grafted Substance P Antagonists

SFTI-1 consists of two loops. One is the trypsin inhibition binding loop and the other is a secondary loop. We replace either loop of SFTI-1 with SP_{1-7} and its analogs (Figure 4.7A and 4.7B). The disulfide bridge between the two cysteine residues is conserved since Korsinczky and coworkers have reported that SFTI-1 loses its internal hydrogen bonding network and secondary structure when the disulfide bond is removed and hence, reducing its proteolytic stability [212]. Here we report SP_{1-7} N-terminal sequence as **RF7** (RPKPQQF) and its analog used for grafting in SFTI-1 as **NF7** (rPKPQNF)**.**

In 1989, Ljungqvist and coworkers had reported that by substituting Gln with Asn at position 6 shortens the side chain by one CH_2 and increases antagonistic potency pA2 value 7.4 in GPI (guniea pig ileum) [245]. Substitution of natural amino acids (L-amino acid) with unnatural amino acids (D-amino acid) can significantly alter physiological properties. Therefore, in peptide **RF7,** Gln6 is substituted with Asn6 and further to determine physiological properties natural amino acid Arg1 is substituted with un natural D-Arg1. **TI-RF7** and **TI-NF7** peptides are grafted sequence of **RF7** and **NF7,** respectively in SFTI-1, in which secondary loop of SFTI-1 is substituted whereas in peptides **RF7-T** and **NF7-T** trypsin inhibition loop of SFTI-1 is substituted shown in Table .4.2.

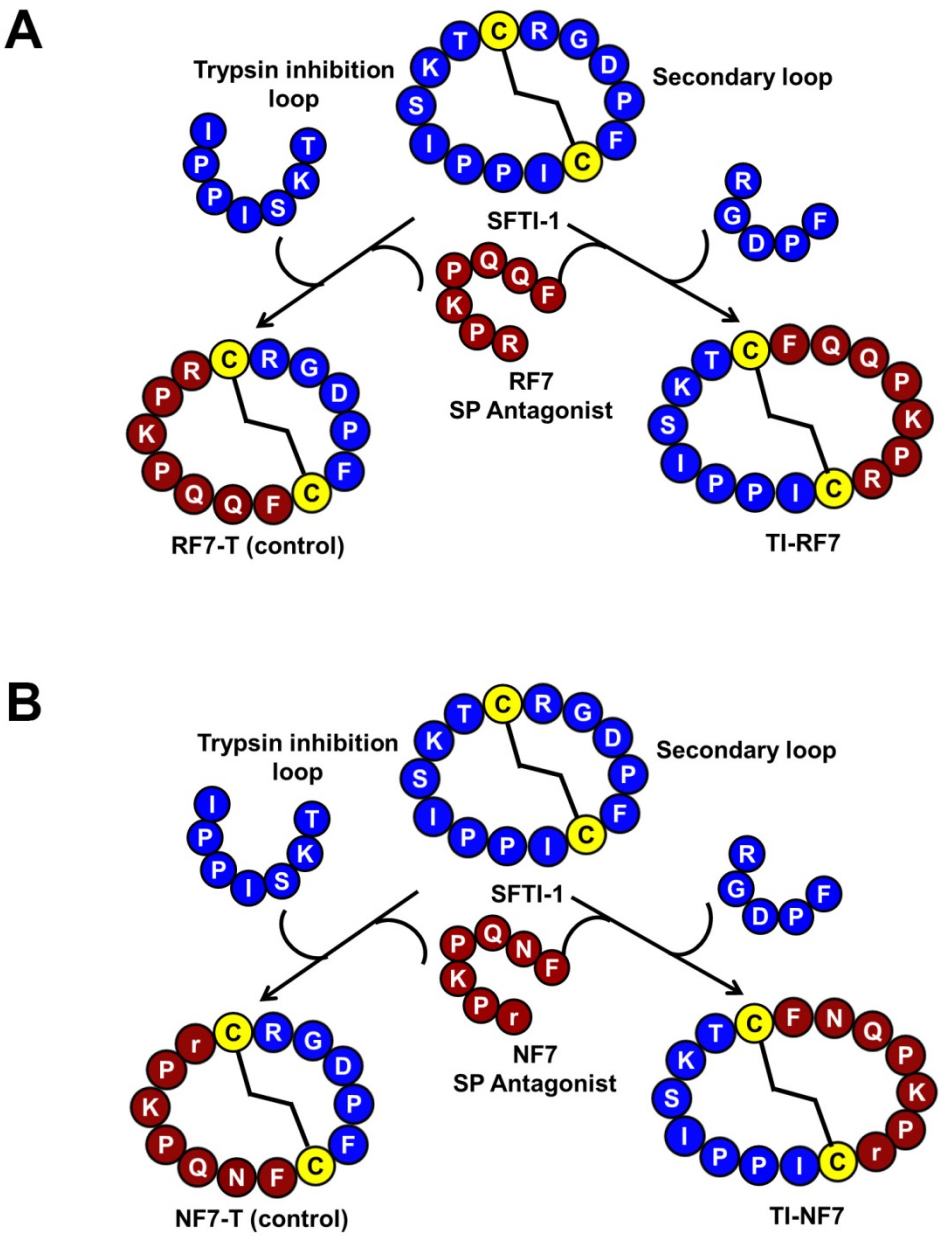
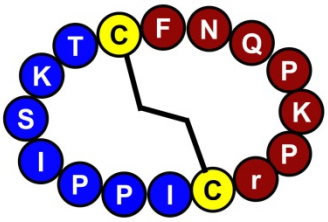
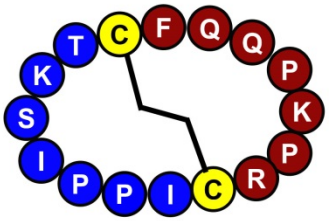
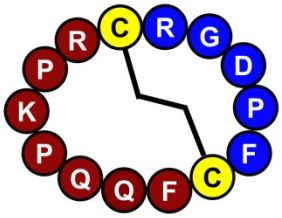
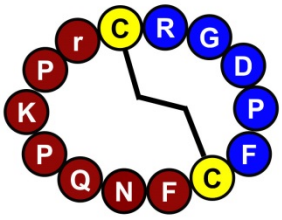


Figure 4.7. Schematic overview of the design for grafting of Substance P antagonist in SFTI-1 scaffold. Each loop of SFTI-1 is substituted with SP antagonist alternatively. A) showing the scheme for grafting of substance P antagonist RF7 in the SFTI-1 scaffold. B) showing the scheme for grafting of substance P antagonist NF7 in SFTI-1 scaffold Beads in blue color show the half loop of SFTI-1 and red color shows SP antagonists. Disulfide bond is shown with bold line

Table 4.2. Linear Substance P antagonists and its cyclic analogs grafted on SFTI-1

Peptide	Grafting	Sequence	Residues
Linear substance P anatagonists			
RF7	SP-N-terminal (SP_{1-7})		7
NF7	Susbstituted SP_{1-7}		7
Cyclic substance P anatagonists grafted using SFTI-1 scaffold			
TI-NF7	TI loop + NF7		16
TI-RF7	TI loop + RF7		16
RF7-T	Sec. loop + RF7		14
NF7-T	Sec. loop + NF7		14

3. Synthesis of Grafted Peptides using Fmoc-SPPS

Peptides **TI-RF7, TI-NF7, RF7-T** and **NF7-T** were synthesized using TEBA linker **3** prepared by incubating commercially available TEBA [(2-butylamino) ethanethiol] **2** with Cl-Trt (2-Cl) resin **1** in DCM. The C-terminal Fmoc-Phe-OH was coupled to secondary amine **3** using HATU/DIEA for 1 h to give Fmoc-Phe-TEBA resin **4a** (Figure 4.8). The coupling of amino acid Fmoc-Phe-OH to the 2°-amine on the TEBA linker was monitored by acetaldehyde/chloronil test. Peptide elongation for **TI-NF7, TI-RF7, RF7-T** and **NF7-T** precursors was carried out by automated microwave peptide synthesizer (CEM, Liberty1). For peptides **RF7-T** and **NF7-T** the last three amino acids (CFP) were coupled manually using Fmoc-AA/HBTU/HOBt/DIEA in DMF for 30 min for each amino acid and Fmoc-deprotection was done using 20% morpholine in DMF instead of piperidine to avoid aspartamide formation.

The peptides were cleaved using a mixture of TFA/H_2O/TIS/thioanisole (89/5/5/1, v/v). HPLC traces showed the major products were amide form (N-forms) (**5a, 5b, 5c** and **5d**) for peptides **TI-NF7, TI-RF7, RF7-T** and **NF7-T** as shown in Figure 4.9 and about 5% of N-form rearranged to its thioester isomer S-forms (**6a, 6b, 6c** and **6d**). This conversion occurred due to acidic pH during peptide cleavage and analytical RP-HPLC with 0.1% TFA containing an aqueous buffer. MS of peptides were confirmed by MALDI-TOF and shown in Table 4.4.

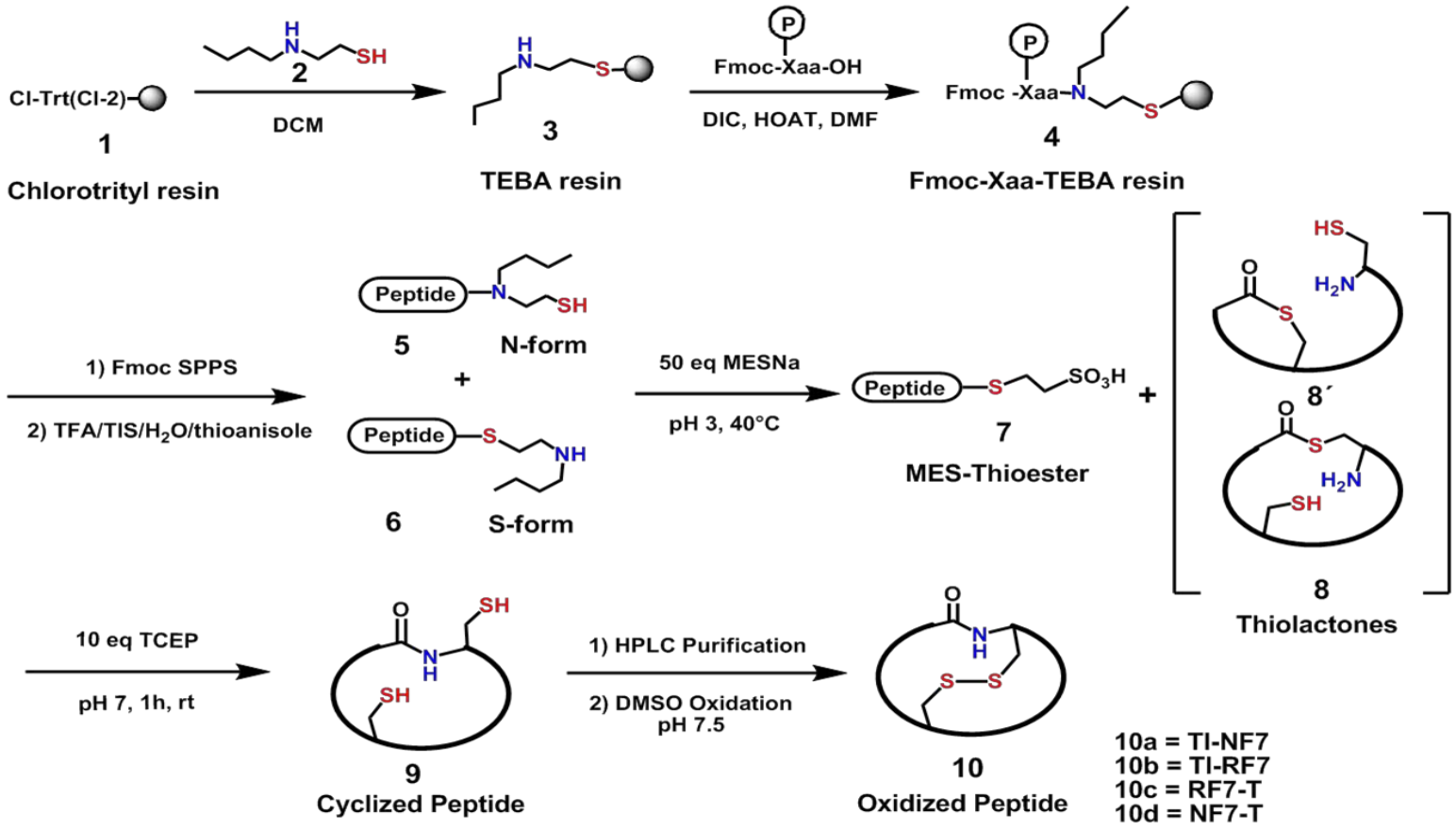


Figure 4.8. General synthesis scheme showing the process of peptide synthesis using Fmoc-Xaa-TEBA resin in SPPS which involves the conversion of peptide TEBA **5** (N-form) and its thioester form **6** (S-form) to MES-thioester **7** followed by thia-zip cyclization and oxidation to get final peptide with disulfide bond **10**. Thiolactones **8** are the intermediate products. **4a**; Xaa = Phe, **4b**; Xaa = Arg

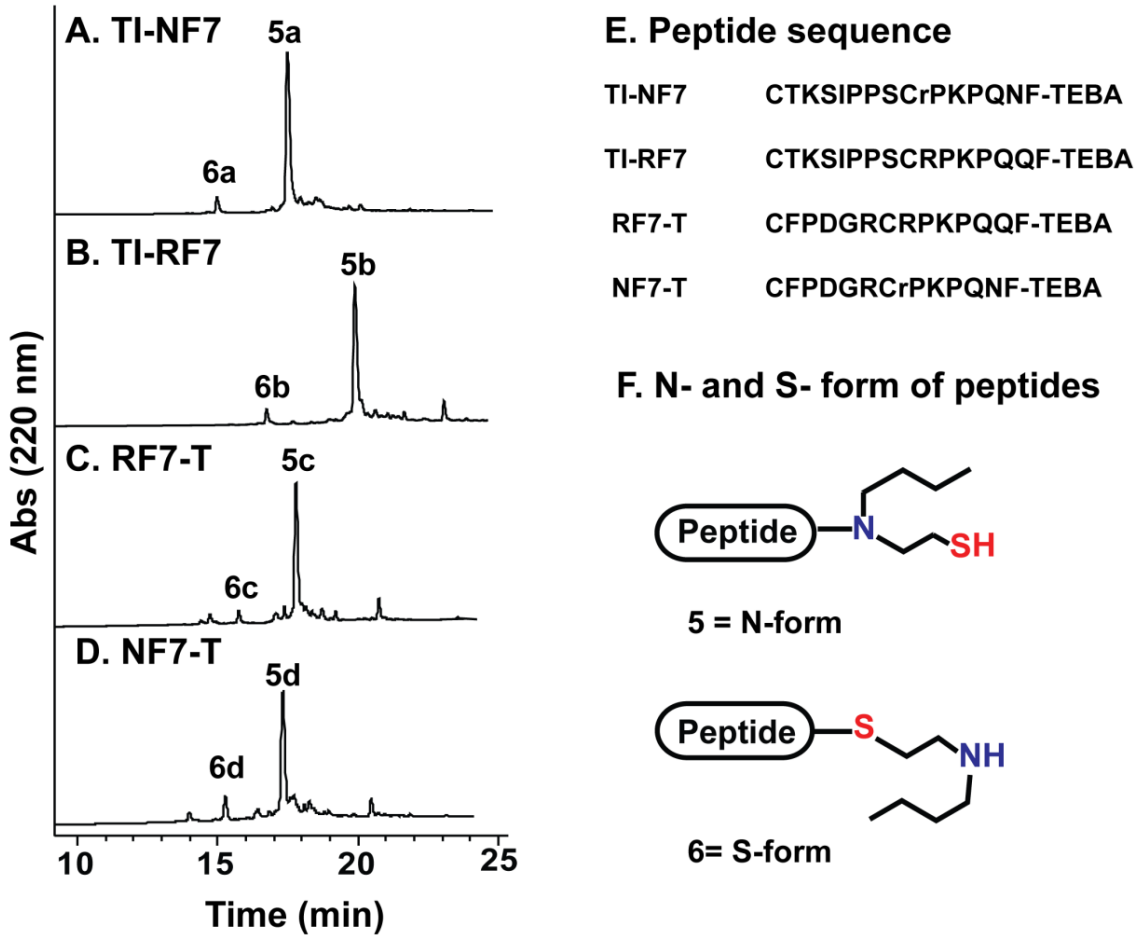


Figure 4.9. RP-HPLC profiles of linear peptides. A) shows the **5a** (N-form) and **6a** (S-form) of TI-NF7. B) shows the **5b** (N-form) and **6b** (S-form) of TI-RF7. C) shows the **5c** (N-form) and **6c** (S-form) of RF7-T. D) shows the **5d** (N-form) and **6d** (S-form) of NF7-T. E) shows the amino acid sequence of each peptide and F) shows thioester (S-) and amide (N-) form of the peptide.

3.1. Thioesterification and Cyclization of Peptides using 2-mercaptoethane sulfonate sodium (MESNa)

Thioesterification and cyclization of the peptide **TI-NF7 (5a, 6a), TI-RF7 (5b, 6b), RF7-T (5c, 6c)** and **NF7-T (5d, 6d)** was carried out in the presence of external thiol 2-mercaptoethane sulfonate sodium (MESNa). Here, we report thioesterification and cyclization of **TI-NF7** as an example to demonstrate the methodology.

Thioesterification of linear peptide **TI-NF7-TEBA (5a** and **6a)** was performed in the presence of external thiol MESNa. Compound **5a** and **6a** were dissolved in pH 3 sodium phosphate buffer (0.1 M) to make a final peptide concentration 2.5 mM. To this solution 50 times more MESNa was added and incubated at 40 °C (Figure 4.8). The thioester **6a** was derived from N-S acyl shift of **5a** and then converted to MES-thioester **7a** in 47 h which is suitable for thia-zip cyclization reaction. Along with MES-thioester **7a** thiolactones **8a** and **8a'** were also obtained by intramolecular thiol-thioester exchange (Figure 4.10B). Both MES-thioester **7a** and thiolactones **8a** and **8a'** were directly subjected to thia-zip cyclization by increasing the pH of the reaction mixture to 7 with 2 N NaOH in the presence of TCEP (10 eq) to obtain cyclized peptide **cTI-NF7 9a.** The reaction mixture was shaken for 1 h and then quenched with 1 N HCl. These thiolactones **8a** and **8a´** can be easily distinguished from end-to-end cyclic peptide **9a** by monitoring the absorbance at 260 nm in RP-HPLC. The thioester and thiolactones are strongly absorbed at 260 nm wavelength. In contrast, amide bond is strongly absorbed at 220 nm wavelength. The cyclized peptide **9a** was further purified by RP-HPLC using a preparative column with a linear gradient of 15-40% buffer B for 80 min. MALDI-TOF MS of compound **9a** was 1810.98 Da $[M+H]^+$; theoretical value is 1810.94 Da. Figure 4.10C shows the analytical HPLC profile of cyclized peptide **cTI-NF7 9a** with a gradient of 10-60% buffer B for 25 min at RT 16.7 min.

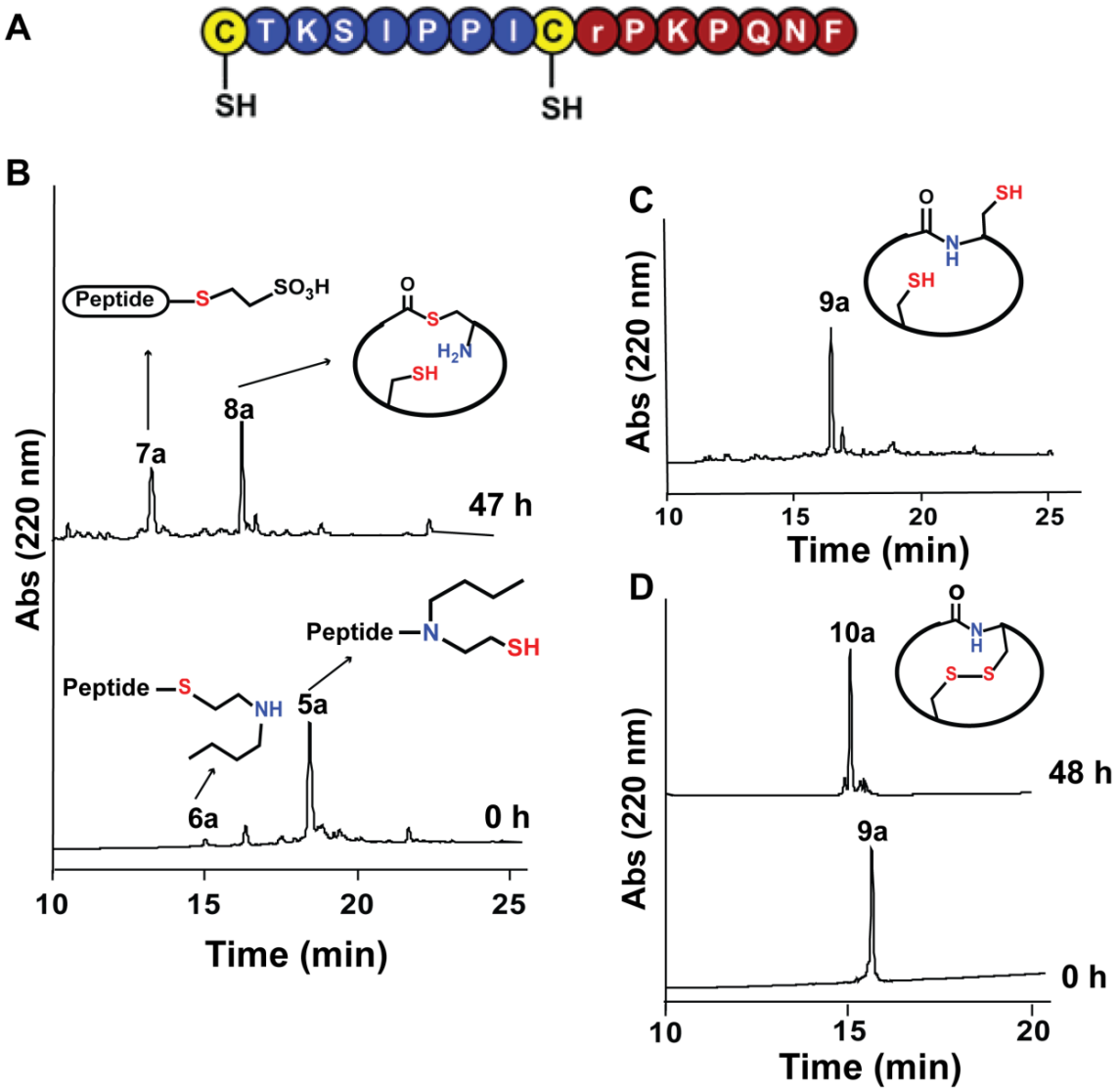


Figure 4.10. RP-HPLC profile of synthesis of TI-NF7. A) Amino acid sequence of TI-NF7. B) Thioesterification is completed in 47 h with 50 eq MESNa, pH 3 and incubated at 40°C. **5a** is amide form (N-form) and **6a** is the isomer thioester form (S-form) of the peptide, **7a** is MES-thioester and **8a** is thiolactone. C) Cyclized peptide **9a** obtained in 1 h after adding 10 eq TCEP and adjusting pH to 7. D) Oxidation step of the peptide at pH 8 with 10% DMSO. Oxidized TI-NF7 **10a** obtained in 24 h.

3.2. Oxidation of TI-NF7

Purified cyclic peptide **cTI-NF7 (9a)** was subjected to oxidation with 10% DMSO in sodium phosphate buffer (0.1 M, pH 8) to 1 mM final concentration. At different time intervals, reaction was monitored by RP-HPLC and peaks were confirmed by MALDI-TOF. After 48 h disulfide bond was formed and cyclic peptide **9a** was completely converted into **10a** (Figure 4.10D). The folded cyclic peptide **TI-NF7 (10a)** was purified using preparative RP-HPLC with a linear gradient of 10-40% of buffer B for 80 min to get the purified compound **TI-NF7 (10a)** in 9% yield compared to the crude peptide. MALDI-TOF MS is 1808.85 Da and theoretical value is 1808.92 Da. Figure 4.10D shows the analytical HPLC profiles for conversion of **c-TI-NF7 9a** to **TI-NF7 10a** at a RT of 16.1 min with a gradient of 10-60% buffer B for 25 min.

Remaining, all other peptides **TI-RF7 (10b), RF7-T (10c)** and **NF7-T (10d)** were cyclized and oxidized in a similar manner. The time specifications for each step and the final yield of purified peptides are shown in Table 4.3 and m/z is mentioned in Table 4.4.

Table 4.3. Time specifications and yield of peptides using thiol MESNa

	Time (h)			
Peptide	**Thioesterification[ab]**	**Cyclization[c]**	**Oxidation[c]**	**Yield (%)[d]**
TI-RF7	24	1	24	9.4
TI-NF7	47	1	48	9.0
RF7-T	24	1	14	12.8
NF7-T	24	1	18	14.8

[a] Thioesterification is performed using external thiol MESNa, [b]Temperature 40 °C
[c] Temperature rt., [d] **Y**ield is calculated in comparison to crude peptide

Table 4.4. Calculated and found m/z of grafted peptides for each step from peptide-TEBA precursor

	Crude Peptide		**Thioesterification (m/z)**		**Cyclization (m/z)**		**Oxidation (m/z)**	
	Peptide-TEBA		**MES-Thioester**		**Cyclized Peptide**		**Oxidized Peptide**	
Peptide	**Cald.**	**Found**	**Cald.**	**Found**	**Cald.**	**Found**	**Cald.**	**Found**
TI-RF7	1944.03	1944.21	1966.93	1967.73	1824.95	1825.14	1822.94	1823.10
TI-NF7	1958.05	1958.15	1952.91	1952.83	1810.94	1810.98	1808.92	1808.85
RF7-T	1793.87	1793.52	1802.75	1802.83	1660.78	1660.74	1658.78	1658.55
NF7-T	1779.85	1779.89	1788.74	1788.78	1646.78	1646.56	1644.74	1644.59

3.3. One-pot Synthesis

One-pot synthesis of **TI-NF7 10a** included three steps: thioesterification, cyclization and oxidation as shown in Figure 4.11. For thioesterification, the peptide was dissolved in sodium phosphate buffer (0.1 M, pH 3) to a final concentration of 2.5 mM. 1eq TCEP was added to this solution avoid the formation of disulfide bonds and reaction mixture was incubated at 40 °C for 66 h. After 66 h, pH of reaction mixture was increased to 7 with 2 N NaOH and incubated at rt. for 1 h for cyclization. Later, for oxidation, the crude reaction mixture was diluted with phosphate buffer (0.1 M, pH 8) to a final concentration of 1 mM and 10% DMSO was added and incubated at rt. for 13 h. Intermediates at each step were monitored by RP-HPLC and MALDI-TOF (Figure 4.12). This one pot synthesis involves only single purification step after oxidative folding for the production of high purity cyclized and oxidized peptide.

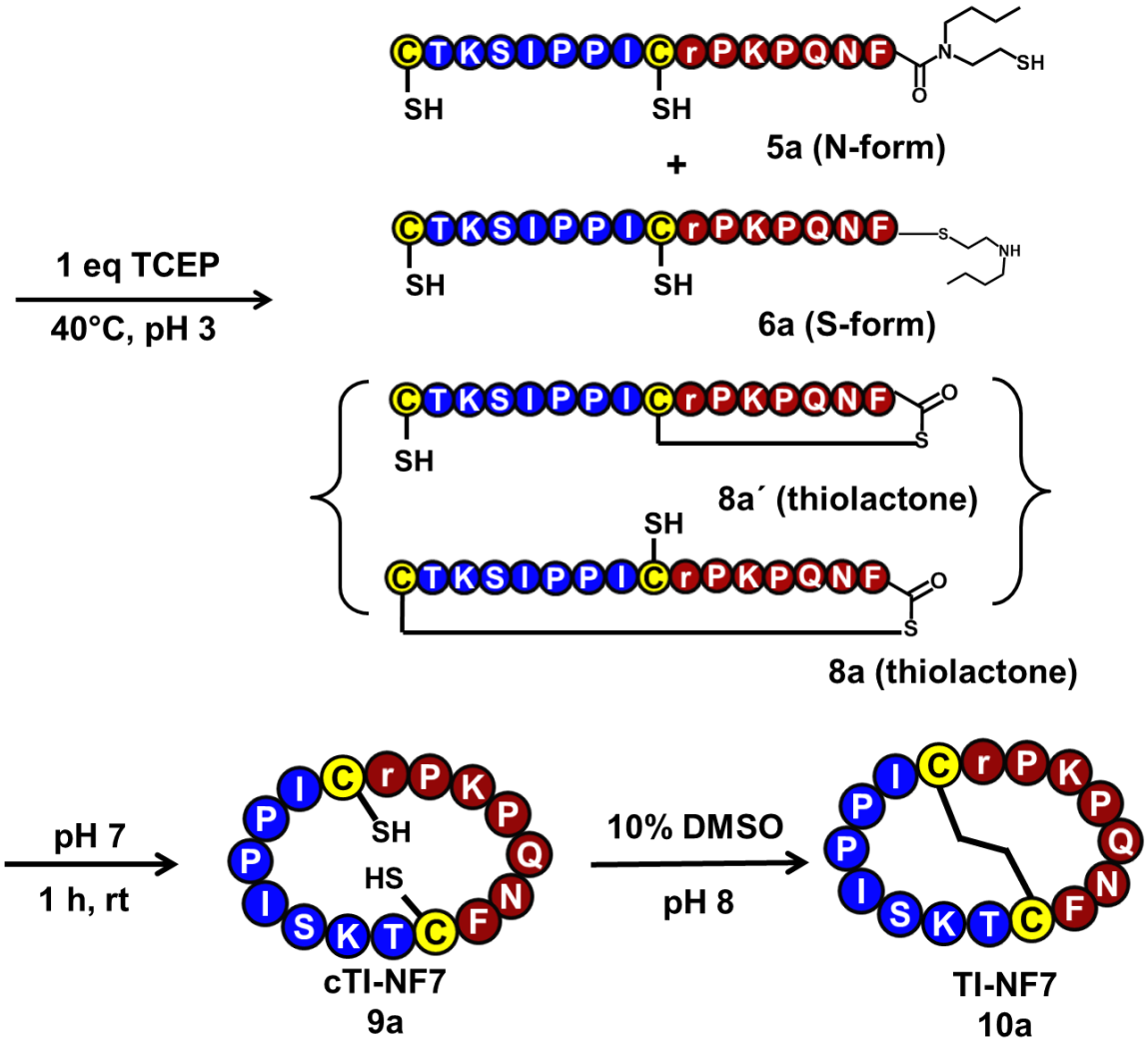


Figure 4.11. Schematic diagram showing one-pot synthesis. N-form **5a** and isomer **6a** (S-form) of peptide TI-NF7 are subjected to thioesterification using 1 eq TCEP at pH 3 and incubated at 40°C for 66 hours. Intermediate products are thiolactones **8a** and **8a´**. After adjusting pH to 7 **9a** is obtained in 1h and with 10% DMSO at pH8 oxidized TI-NF7 (**10a**) is obtained in 13 h.

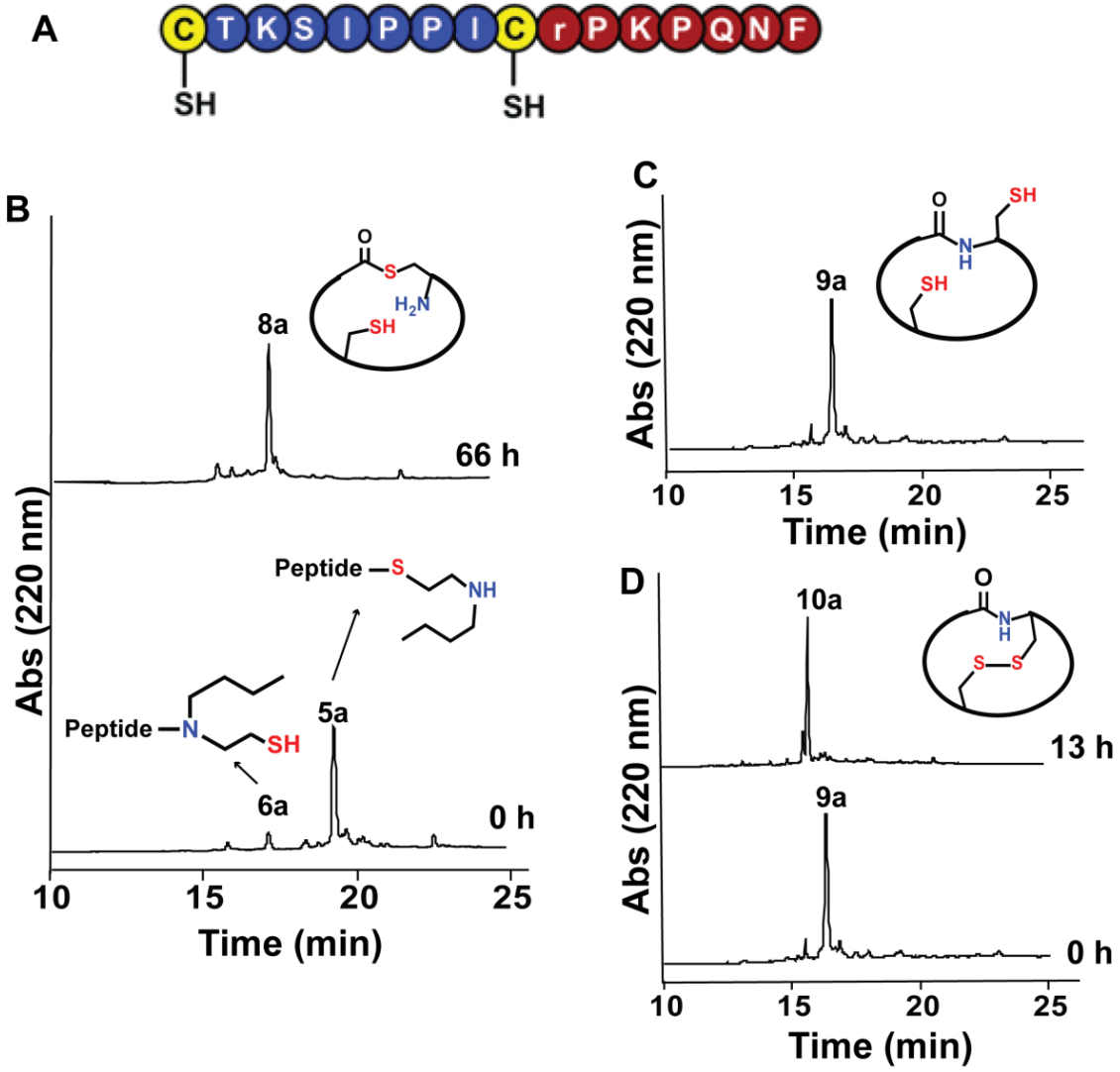


Figure 4.12. RP-HPLC profile of synthesis of TI-NF7 with one-pot synthesis method. A) Amino acid sequence of TI-NF7. B) Thioesterification of TI-NF7 is completed in 66 h with 1eq TCEP, pH 3 and incubated at 40°C.**5a** is N-form and **6a** is the isomer (S-form) of peptide, Intermediate products are **8a** and **8a´** thiolactones. C) Cyclized peptide **9a** obtained in 1 after adjusting pH to 7. D) Oxidation of peptide, the reaction mixture is diluted to final conc. 1 mM at pH 8 with 10% DMSO. Oxidized TI-NF7 **10a** obtained in 13 h.

In order to make one pot synthesis possible, we investigated and optimized the necessary reaction conditions for direct oxidation step after cyclization to avoid purification and lyophilization in between. Initially, all the conditions of the previous method were kept same such as thioesterification was carried out using 50 eq MESNa at pH 3 and incubated at 40°C. Once thioesterification was completed, it was cyclized in a similar manner by adding 10 eq TCEP at pH 7. Later, the reaction mixture was directly subjected to oxidation by diluting the cyclization mixture with sodium phosphate buffer (0.1 M, pH 8) to the final concentration of 1 mM. 10% DMSO was added and incubated reaction was monitored using RP-HPLC and MALDI-TOF.

After 18 h the HPLC profile showed the formation of MES mixed disulfide. The peak **'a'** at rt 15.7 min with *m/z* = 2090.65 and peak **'b'** at rt 16.17 min with *m/z* = 1950.68 corresponded with binding of two MES and one MES group with internal thiols of cysteine, respectively (Figure 4.13), resulting blocked the availability of thiol groups for formation of disulfide bond whereas peak **'c'** at RT 16.5 min with *m/z* = 1810.68 corresponded cyclic peptide **9a**. The reaction mixture was further diluted to 0.1 mM final concentration but still after 48 h incubation, internal thiols groups of cysteine were aggregated with MES and oxidation reaction remained incomplete.

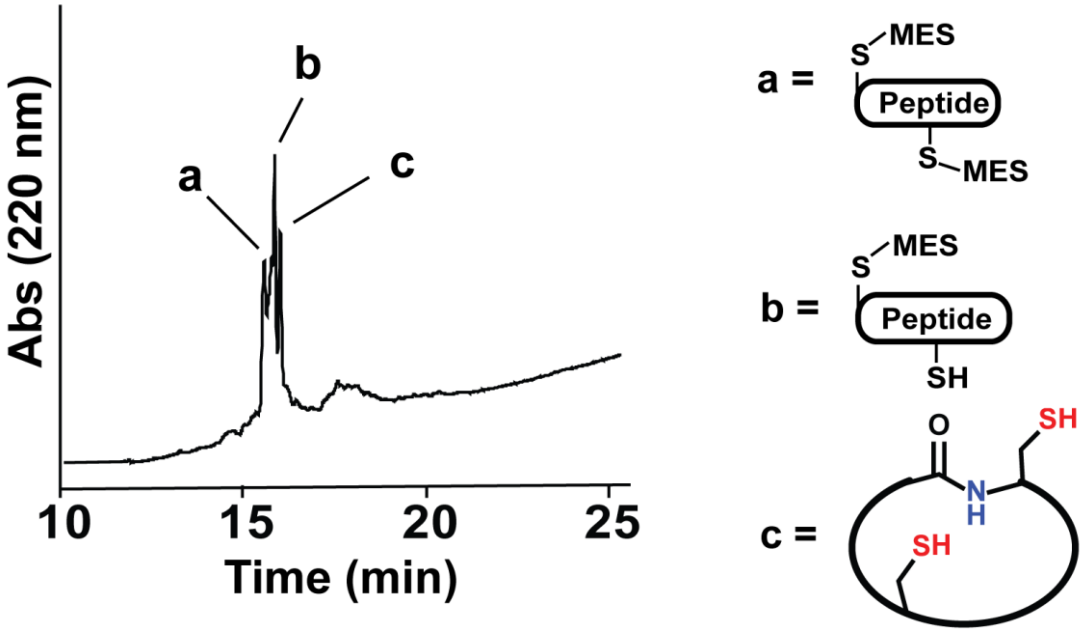


Figure 4.13. RP-HPLC profile of oxidation step for one-pot synthesis using condition 1. Intermediates confirmed by MALDI-TOF after 18 h shows MES mixed-disulfide peaks. The peak **'a'** at RT 15.7 min with *m/z* = 2090.65 and peak 'b' at RT 16.17 min with *m/z* = 1950.68 corresponded with the binding of two MES and one MES group with internal thiols of cysteine, respectively. peak **'c'** at RT 16.5 min with *m/z* = 1810.68 corresponded to cyclic peptide **9a**

In the next step, we reduced the concentration of external thiol (MESNa) to 10 eq, instead of 50 eq, keeping all other conditions same. Thioesterifictaion and cyclization of **TI-NF7-TEBA** was completed very smoothly in 48 h. As the reaction mixture was subjected to oxidation at pH 8 to a final concentration of 1 mM with 10% DMSO and incubated for 48 h, aggregation of MES with internal thiols of cysteine was observed.

The peptide **TI-NF7** has two internal thiol groups. Using the thia-zip cyclization approach, we omitted the use of external thiol MESNa and TCEP completely. **TI-NF7-TEBA** was subjected to thioesterification in phosphate buffer (0.1 M, pH 3). HPLC profiles showed that the thioesterification was completed in 66 h with the formation of two thiolactone isomers **8a** and **8a´** (Figure 4.14A)**.** The thiolactones gave *m/z* 1810.98 in MALDI-TOF. Addition of 2 N NaOH to the crude reaction mixture for increasing pH to 7 resulted in the conversion of **8a** and **8a´** to **9a** in 1 h at rt. Along with cyclized product **9a**, HPLC profile showed a peak at RT 18.3 min for starting material with disulfide bond formation (*m/z* = 1941.71), in the absence of TCEP (Figure 4.14B). Further, the reaction mixture was subjected to direct oxidation with 10% DMSO in 1 mM final concentration. HPLC profile showed the conversion of **9a to 10a** in 13 h (Figure 4.14C). The final yield of **TI-NF7** was 20%.

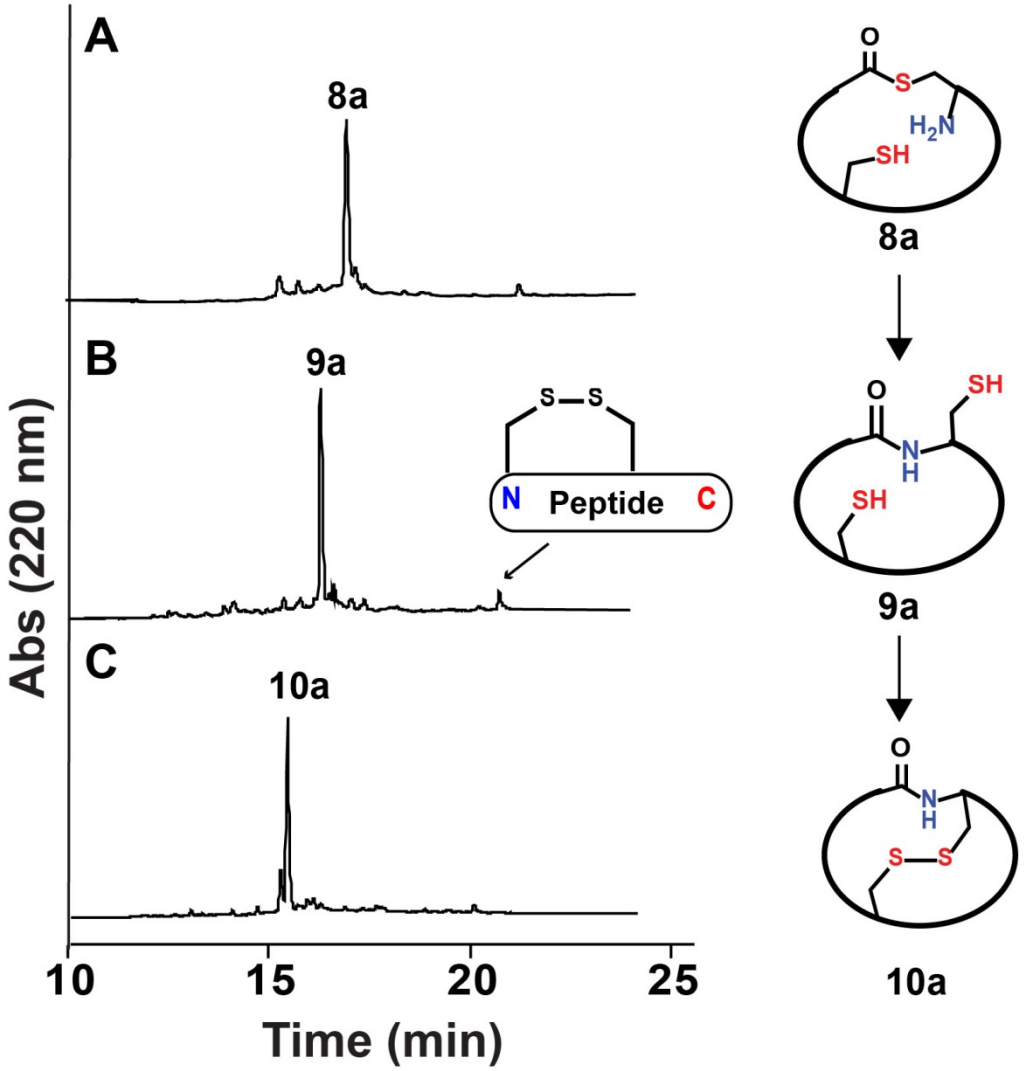


Figure 4.14. RP-HPLC profile of oxidation for one-pot synthesis using condition 3. A) shows thioesterification at 66 h without using MESNa. B) shows cyclization at 1 h. C) shows Oxidation with 10% DMSO at 13 h. (⇾ shows accumulation of starting material with disulfide bond.

Finally, to maintain all the thiols in the reduced form we included 1 eq TCEP in the thioesterification step. All other conditions were kept optimal. RP-HPLC showed completion of thioesterification, cyclization and oxidation in 66 h, 1 h and 13 h, respectively (Figure 4.12). The reaction mixture was purified by RP-HPLC with a gradient of 10-40% buffer B for 80 min. After lyophilization **TI-NF7** was obtained in a yield of 28% (16 mg). Different conditions used in optimization of one pot synthesis are summarized in Table 4.5.

Peptides **TI-RF7**, **RF7-T** and **NF7-T** were also synthesized using one-pot synthetic scheme. The time duration and yields for purified peptides are mentioned in Table 4.6. Peptide **TI-RF7** showed >50% hydrolysis. The rate of hydrolysis is increased with time and temperature.

Table 4.5. Optimization for one-pot synthetic scheme

Reaction	MESNa conc. (Eq)	TCEP (Eq)	Yield (%)
1	50	10	ND
2	10	10	ND
3	-	-	20
4	-	01	28

* DMSO 10% and Temperature 40°C
ND: Not determined

Table 4.6. Time specifications and yield of peptides using one-pot synthetic scheme without using external thiol

	Time (h)			
Peptide	Thioesterification (without MESNa)	Cyclization	Oxidation	Yield (%)
TI-NF7	66	1	13	28.0
TI-RF7	32	1	13	28.7
RF7-T	30	1	08	24.4
NF7-T	>66	-	-	-

4. Synthesis of RF7 and NF7

Linear peptides **RF7 (11)** and **NF7 (12)** were synthesized on Rink amide resin at 0.1 mmol scale manually by Fmoc chemistry. Coupling mixture included Fmoc-AA/HBTU/HOBt/DIEA (4/4/4/6 eq). Fmoc-D-Arg-OH was coupled using Fmoc-AA/DIC/HOAt (4/4/4 eq) for an hour. Fmoc was deprotected by using 20% piperidine. The purity of crude peptides was monitored by analytical RP-HPLC. Purification was performed by preparative RP-HPLC with a gradient of 10-40% using buffer B for 80 min. The yield of peptides **RF7** and **NF7** were 48% (43 mg) and 27% (23.9 mg), respectively. MALDI-TOF MS for **RF7** is *m/z* = 900.02 and theoretical value is *m/z* = 899.51. MALDI-TOF MS for **NF7** is *m/z* = 886.42 and theoretical value is *m/z* = 885.51 (Figure 4.15).

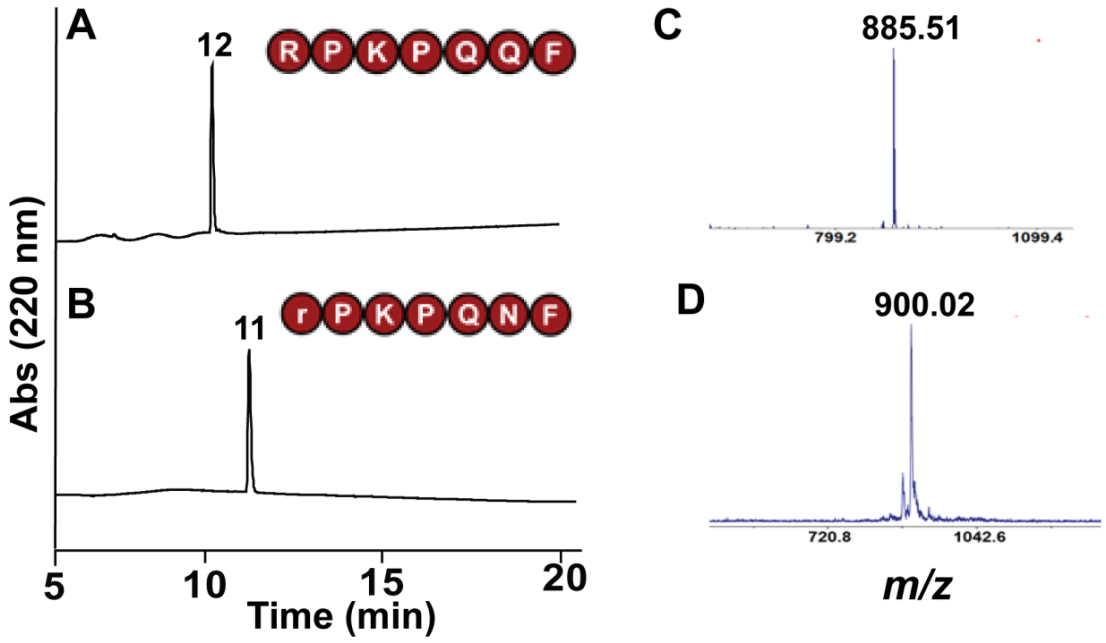


Figure 4.15. **RP-HPLC and MS profiles of RF7 and NF7**. A) and C) show the purity of RF7 **(11)** and NF7 **(12)** with a gradient of 10-60% buffer B for 25 min and B) and D) show their MS profiles using MALDI-TOF.

5. Characterization of Grafted Peptides

To prove our hypothesis that grafted peptides would be more stable and act bifunctionally, a series of stability assays and bioassays were performed on the synthetic grafted peptides. *In vitro* assays which reflect the gastro-intestinal environment such as, acidic pH stability, endopeptidase stability (trypsin and pepsin) assays were performed and monitored by RP-HPLC analysis at various intervals. Trypsin inhibition assay and the nitric oxide release assay were carried out to investigate their bifunctional property.

5.1. Acid Stability Assay

We tested the stability of peptides in the presence of 0.2 M HCl to simulate the acidic environment of the stomach. All the peptides **TIRF7, TI-NF7, RF7-T, NF7-T** and **SFTI-1** were subjected to the 0.2 M HCl at room temperature for 6 h. RP-HPLC profiles of peptides have shown that all the grafted peptides were resistant to acidic pH of the stomach (Figure 4.16A). More than 90% of peptides remained intact over a period of 6 h.

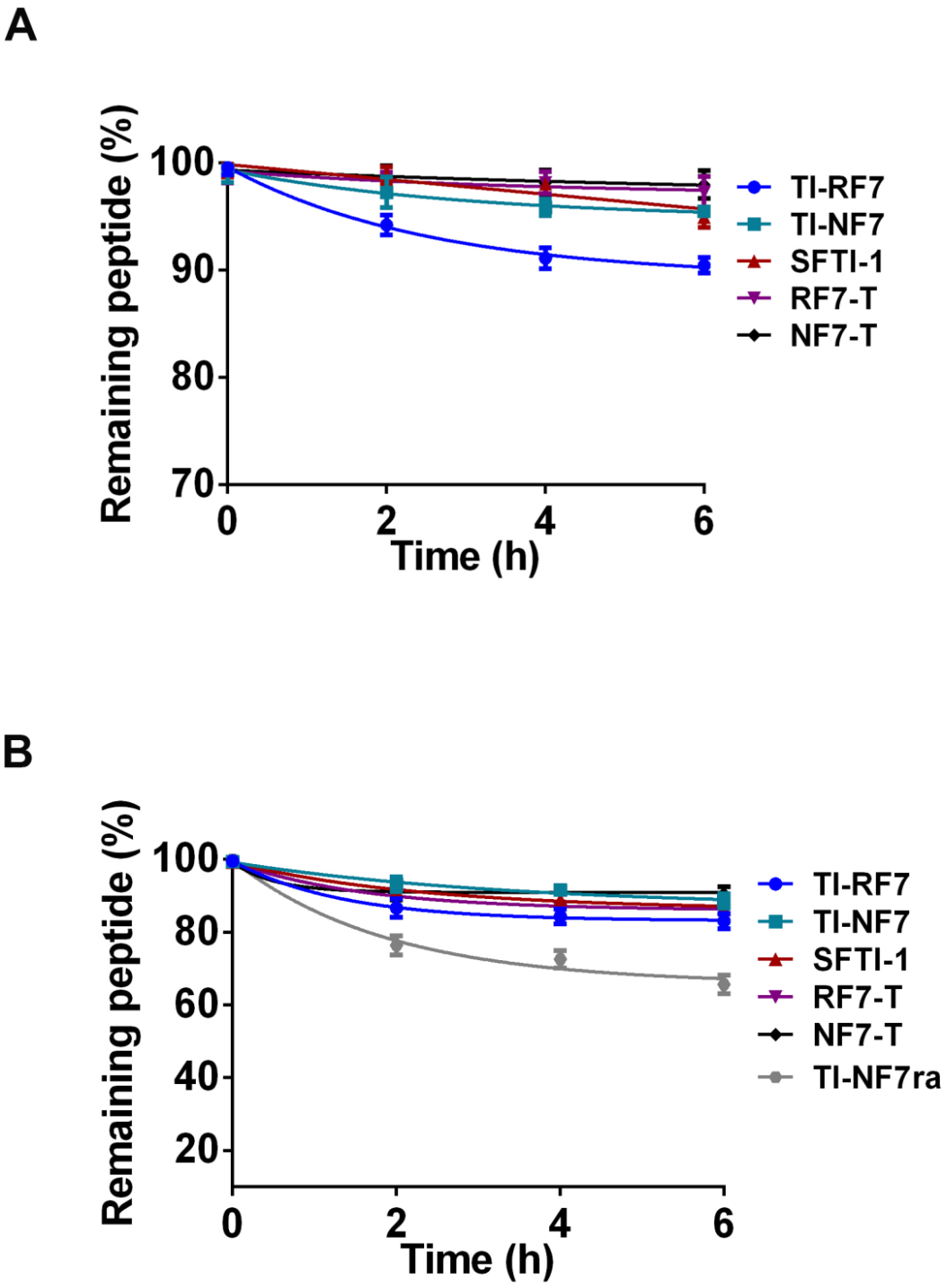


Figure 4.16. Stability of peptides. A) shows in acidic conditions. Peptides were incubated at room temperature with 0.2M HCl for 6 h. B) shows against pepsin enzyme. Peptides were incubated at 37°C with 1:50 ratio of pepsin enzyme for 6 h.

5.2. Endopeptidase Stability Assay

To evaluate the stability of grafted peptides against endopeptidase such as pepsin and trypsin, peptides were subjected to enzyme digestion at 37 °C. The digestion processes were monitored by analytical RP-HPLC and MALDI-TOF analysis. **SFTI-1** and *S*-reduced and *S*-alkylated **TI-NF7-ra** were used as the control peptides.

5.2.1. Pepsin Stability Assay

All the peptides and the controls were subjected to pepsin digestion for 6 h. The results showed that 90% of all the grafted peptides were resistant against pepsin enzyme whereas, 30% of the *S*-reduced and *S*-alkylated **TI-NF7**-ra was digested with pepsin enzyme in 6 h (Figure 4.16B). The result suggests that cysteine bond conferred pepsin resistance to grafted cyclic peptides.

5.2.2. Trypsin Stability Assay

All grafted peptides and the control peptides were subjected to trypsin digestion at pH 7.2. The results were compared to the control peptides **SFTI-1** and *S*-reduced and *S*-alkylated **TI-NF7-ra.** Peptides **TI-RF7, RF7-T, TI-NF7** and **NF7-TI** have shown stability against trypsin degradation for 12 hours (Figure 4.17A). **RF7-T** and **NF7-T** lack trypsin inhibition loop but they have shown stability against trypsin due to the presence of RP, KP and RC sites, which are not major proteolytic sites for trypsin digestion. In contrast, the *S*-reduced and *S*-alkylated **TI-NF7-ra** (m/z = 2061.18) was completely converted to linear analog in 1 h. The peak (*m/z* = 2079.18) corresponds to the addition of H_2O to the peptide (Figure 4.17B). However, **TI-NF7** and **NF7-T** showed 50% degradation in 12 h which may be due to some distortion in the structure after grafting.

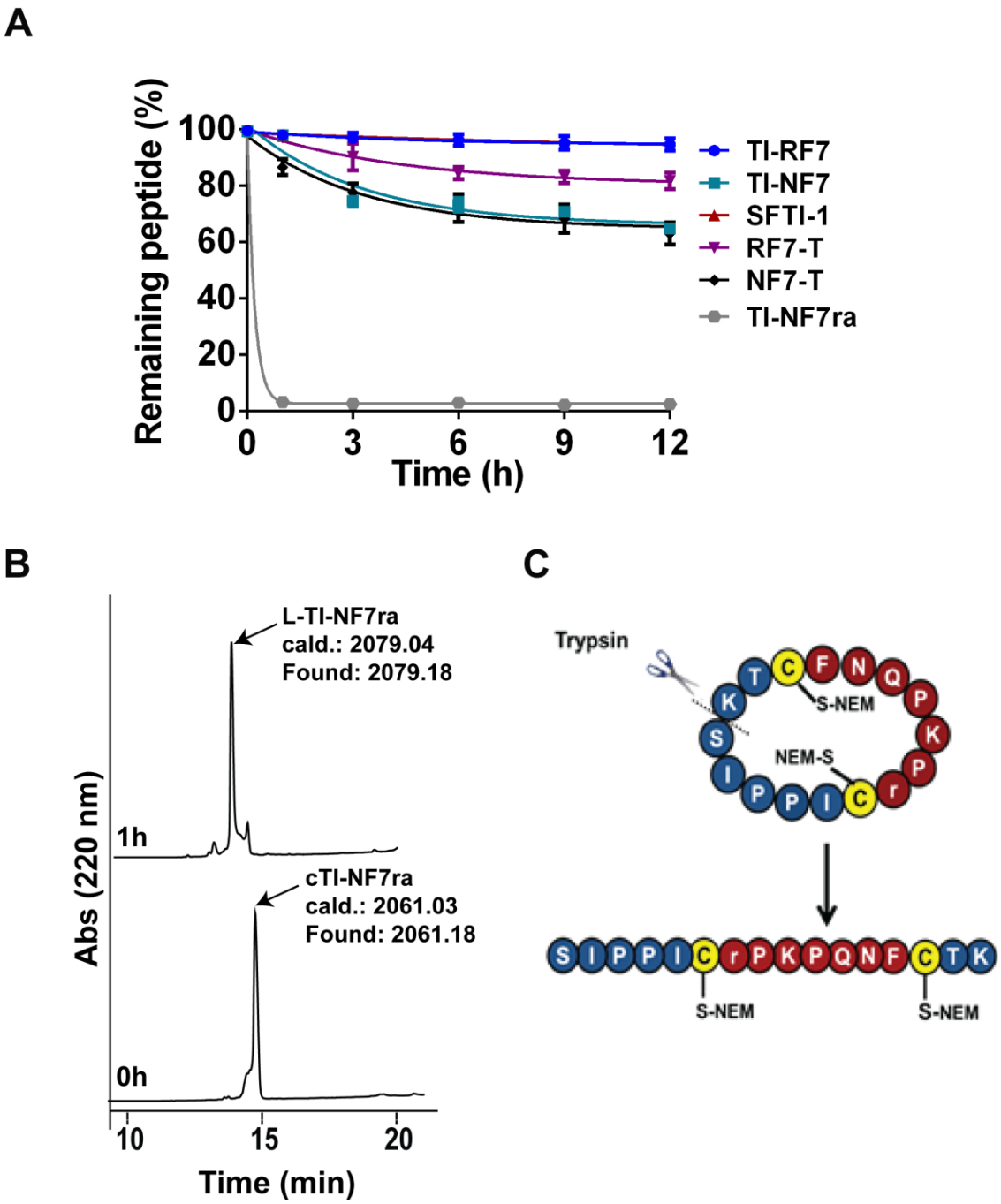


Figure 4.17. Stability of peptides against trypsin enzyme. A) Peptides were incubated at 37°C with 1:50 ratio of trypsin enzyme for 12 h. S-Reduced and S-alkylated peptide **TI-NF7-ra** has shown complete degradation in 1 h. B) shows the HPLC profile for complete tryptic digestion of TI-NF7ra in 1 h. C) shows the cleavage site of trypsin in S-reduced and S-alkylated TI-NF7ra.

5.3. Serum Stability Assay

Peptides were subjected to protease degradation in 25% human male serum AB type in RPMI-1640, phenol free medium, and their stability was monitored by analytical RP-HPLC and MALDI-TOF analysis. Linear **RF7** was degraded 100% whereas **NF7** 60% in 6 h, showing the half-life of <1 h and <7 h, respectively. In contrast, **RF7-T**, **TI-RF7, TI-NF7, and NF7-T** showed only 10%, 20%, 30% and 30% degradation over a period of 24 h, respectively. **SFTI-1**, the control peptide, showed no degradation over 24 h. (Figure 4.18).

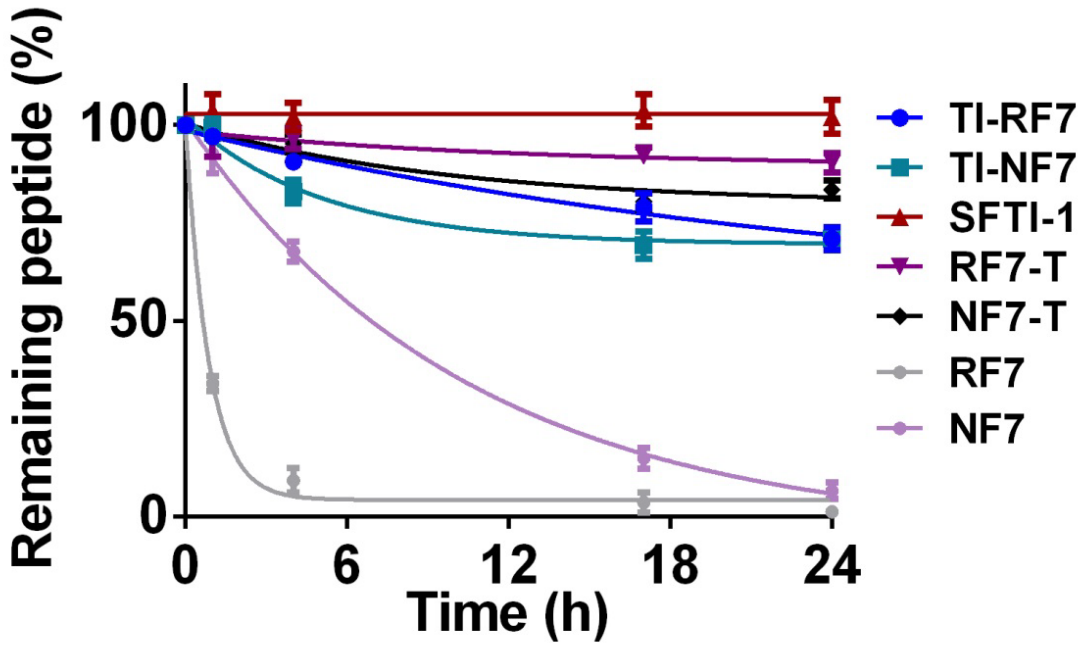


Figure 4.18. Stability of peptides in human serum. Peptides were incubated at 37°C with 25% human serum with RPMI-1640 medium without phenol for 24 h.

5.4. Trypsin Inhibition Assay

Engineered peptides contained the trypsin inhibition loop from SFTI-1, we would like to investigate whether they retain the trypsin inhibition activity. All the peptides were subjected to trypsin treatment in the presence of BAPNA substrate at 37°C. The activity of trypsin was determined through measurement of absorbance. As shown in Figure 4.19 peptides **TI-RF7 and TI-NF7** exhibited almost similar effect as SFTI-1 with IC_{50} of 54.3 nM and 42.9 µM, respectively. These results reflect that **TI-RF7** and **TI-NF7** are able to inhibit trypsin with the same potency as SFTI-1. The other RF7-T and NF7-T analogs without trypsin inhibition loop were completely inactive in the trypsin inhibition assay, suggesting that trypsin inhibition activity depends on the presence of trypsin inhibition loop.

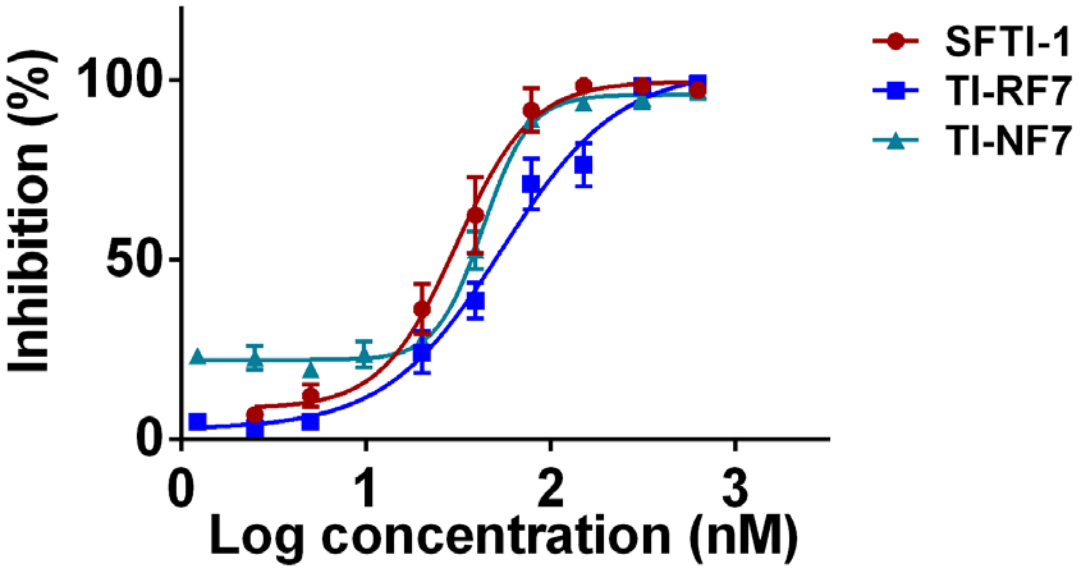


Figure 4.19. Trypsin Inhibition activity of grafted peptides. Peptides were incubated with 100 nM trypsin enzyme in 50 mM Tris-HCl buffer (25 mM $CaCl_2$ at pH 8.2) and 20 mM of the substrate (DL-BAPNA). The result showed the data from three individual experiments. Error bars represent the standard deviation. IC_{50} of peptides were calculated using Graphpad Prism 5.

5.5. Substance P Antagonistic Activity

5.5.1. Effect of SP on Nitrite Production by LPS-stimulated Macrophages

Firstly, we studied the effect of SP on NO production by Raw 264.7 macrophages. Briefly, we incubated macrophages with different concentrations of substance P ranging between 10^{-6} M to 10^{-10} M for 24 h and stimulated with 0.1 µg/ml LPS. Figure 4.20 showed that nitrite production was increased by LPS stimulated cells and was more with 10 nM SP, though, substance P alone did not affect the basal level of nitrite production by macrophages as reported by *Jeon, H.K., et al.* [274].

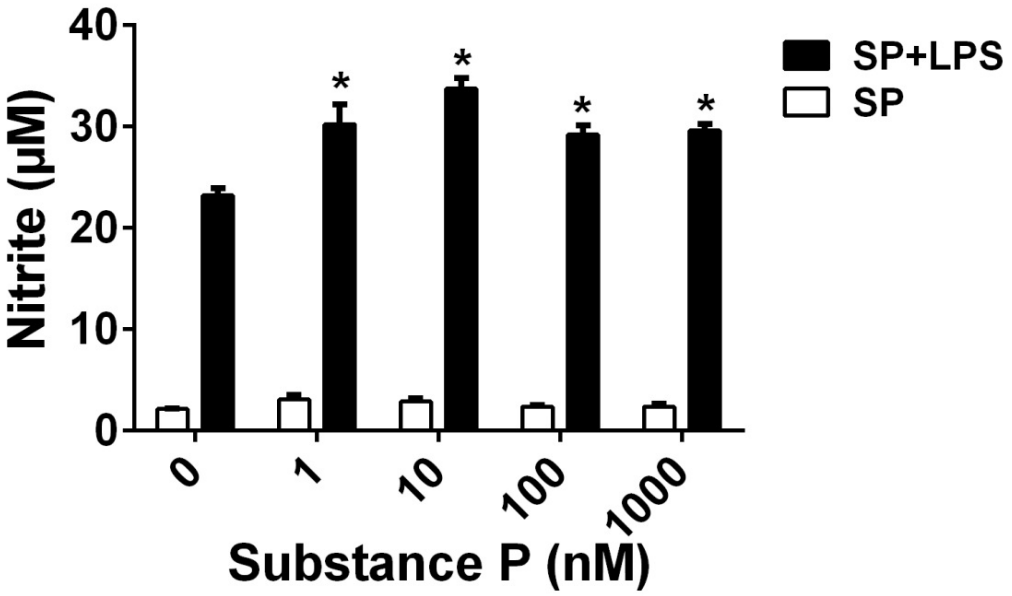


Figure 4.20. Effect of SP on nitrite production by LPS-stimulated macrophages. Raw 264.7 cells were incubated with substance P 1nM to 1000nM in presence or absence of LPS (100 ng/ml) for 24 h.

5.5.2. Effect of SP Antagonists on Nitrile Production

To determine whether our designed SP antagonist inhibit the LPS mediated NO production by SP, the cells were treated with different concentrations of peptides in the presence of LPS and SP or LPS only for 24 h. As shown in figure 4.21 A **TI-RF7** and **RF7** abolished the stimulatory effect of substance P in 10 µM concentration on LPS-induced nitrite production by macrophages. The other designed analogs TINF7 and NF7 (Fig 4.21 B) did not inhibit NO production with the tested concentrations. It could be because of the substitution of amino acid at Q6 to N6 in the primary sequence which may affect the structure conformation. Further NMR studies would be required to validate the structure conformation.

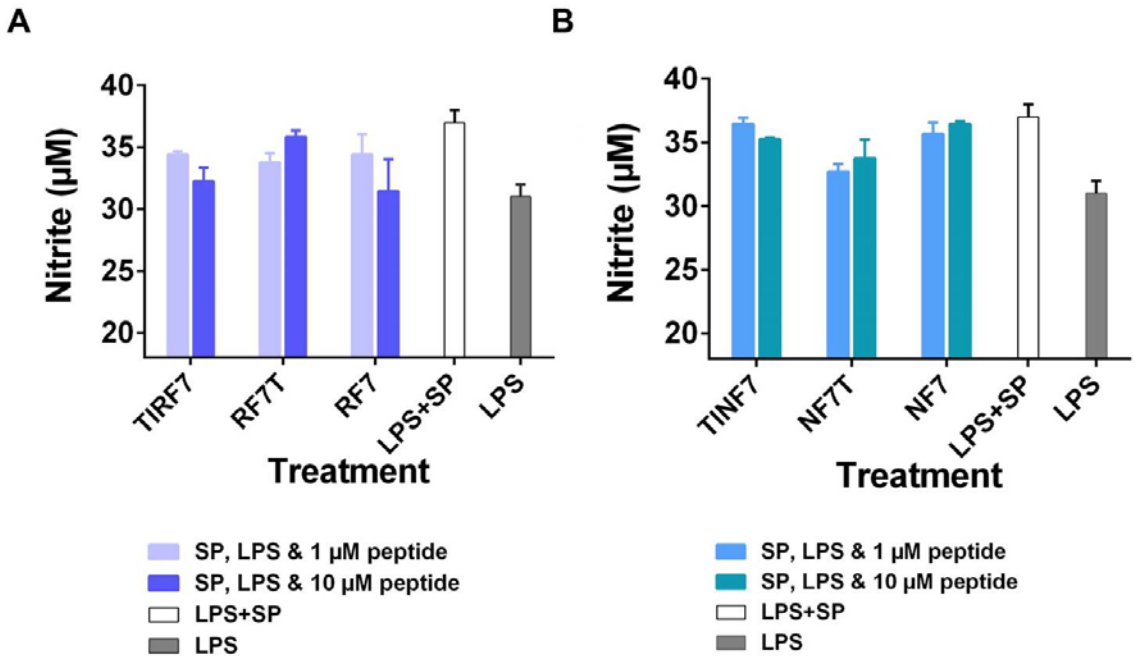


Figure 4.21. Effect of SP Antagonists on Nitrile Production. Raw 264.7 cells were treated with different concentrations of peptide in the presence of SP (100 nM) and LPS (100 ng/ml) for 24 h. A) Macrophages treated with RF7 and its grafted analogs. B) Macrophages treated with NF7 and its grafted analogs.

6. Discussion

Cyclic CRPs have shown to be stable scaffolds [111, 115, 116, 211, 275, 276] for grafting of bioactive peptides. Sunflower trypsin inhibitor was selected for designing of bi-functional analogs of SP antagonist due to their end-to-end cyclized backbone, disulfide bond and the extensive network of hydrogen bonds. SP_{1-7} have shown to possess anti-nociceptive (26, 27), anti-inflammatory (28) and anti-hyperalgesic activities (29). Recently, SP_{1-7} has also been reported to be effective for neuropathic pain in diabetic patients, but the data does not report its use as an oral drug candidate. The role of SP_{1-7} in pain motivated us to graft it in SFTI-1 and design orally active substance P antagonist.

6.1. Design and Synthesis of Substance P Antagonists

All peptides were synthesized using TEBA-thioester surrogate [273], thioesterification was carried out in the presence of external thiol MESNa. Under acidic pH, N-S acyl shift reaction occurred and N-form of peptide **5a** was rearranged its isomer S-form **6a** first and then converted to MES thioester **7a** and thiolactones via S-S acyl shifts. Thia-zip cyclization proceeded through ring expansion of thiolactones, ultimately leading to the macrolactam formation by the N-terminal Cys residue via an S-N acyl shift [257]. This method involves two purifications and lyophilization, one before disulfide bond formation to remove excess thiol and second after disulfide bond formation, which reduces the overall yield of final product and utilizes more time. Therefore, despite the use of this method, we aimed to design one-pot synthetic scheme with a single purification step after oxidation of cyclic peptide for the production of high purity cyclic peptides. In 1999, *Tam et al*. proposed the theory of thia-zip mechanism involved in macrocyclization of CRPs [257], which require N^{α} Cysteine, C-terminal thioester and at least one free internal thiol group. Cyclization of CRPs containing TEBA-thioester surrogate fulfill the needs of proximity-driven thia-zip cyclization and the side chain thiols of CRPs replace the role of external thiols in S-S acyl

shift reaction and gives series of thiolactone intermediates instead of C-terminal thioester. The N-terminal of cysteine joins with the C-terminus of the peptide to form the amide bond by expansions of thiolactone rings via reversible S-S acyl shift reactions, followed by irreversible S-N acyl shift reaction [257].

Based on thia-zip cyclization mechanism, thioesterification was conducted at acidic pH in the absence of external thiol. At this acidic pH, N-S acyl shift occurred and N-form of peptide **5a** rearranged its isomer S-form **6a**. The internal thiols of **TI-NF7-TEBA** spontaneously underwent a series of S-S acyl shift reactions to form TI-NF7-thiolactones **8a** and **8a´**. Two isomers of thiolactones were obtained because of the presence of two side-chain internal thiols of two cysteines. The formation of thiolactones suggests that internal side-chain thiols participate actively in trans thioesterification. With the expansion of thiolactone rings via reversible S-S acyl shift, the N-terminus cysteine joins the C-terminal of the peptide by proximity-driven S-N acyl shift to give backbone amide bond. The cyclization mixture of peptide **cTI-NF7 9a** was directly subjected to oxidation with 10% DMSO without purification, side-chain thiols of internal cysteine interacted smoothly to form a disulfide bond. In the last step, after disulfide bond formation it was purified to obtain **10a.**

This one-pot synthetic scheme has a significant advantage over the method of synthesis using external thiol. The synthesis using external thiol involves two purifications and lyophilizations, which affects overall yield and provided ~9% of purified **TI-NF7 10a**. However, the current one-pot synthesis with only one purification step gave cyclized and oxidized peptide **TI-NF7 10a** in identical purity and with 3 folds improved yield of ~28%, in overall less time.

6.2. Stability

To determine the stability of synthesized grafted peptides **TI-NF7, TI-RF7, RF7-T** and **NF7-T** we performed the assays against heat, acid and proteolytic enzymes simulating the *in vivo* GI tract environment**.** Our designed peptides showed remarkable stability under acidic and thermal degradation for 6 h. Most of the peptides are prone to denature and cannot maintain their secondary structure under these harsh conditions.

Peptides have to survive over the proteolytic degradation in order to be available for absorption in the gut. Thus, we also examined the stability of our grafted peptides against trypsin and pepsin, a major protease in the stomach and small intestine, and compared against the control peptide RF7 and NF7. **TI-RF7** and **TI-NF7** showed significant stability against proteolytic enzymes compare to the control peptides RF7 and NF7. This showed that our grafting strategy improves the stability against endopeptidases. In addition, the results were also compared with the *S*-reduced and *S*-alkylated **(ra-TI-NF7),** which is end-to-end cyclized but does not contain any disulfide bond. This **ra-TI-NF7** was completely digested in 1 h, suggesting the role of disulfide bond in stability against enzymatic digestion. The major cleavage sites for trypsin are carboxyl ends of lysine and arginine [277]. Data showed the stability of **RF7-T** and **NF7-T** against trypsin digestion in a similar manner as **TI-NF7** and **TI-RF7**. It was hypothesized that peptide **RF7-T** and **NF7-T** grafted by replacing secondary loop of SFTI-1, lack trypsin inhibition loop and were expected to degrade in the presence of trypsin. The stability of these two peptides **RF7-T** and **NF7-T** was because of presence RP, KP, and CP sites. The existence of proline at the carboxyl terminus of arginine or lysine decreases the rate of digestion by trypsin. Rodriguez and Gupta had also reported that trypsin activity is reduced with the presence of cysteine at C-terminal [278].

In addition to proteolytic stability, stability in human serum is also important in the design of novel drug leads. Plasma stability is an essential factor that affects the half-life of peptides, and therefore all the grafted peptides were tested in human serum stability assay. Both the linear sequences **RF7** and **NF7** showed improvement in stability after grafting into the cyclic scaffold, SFTI-1, and > 60% remained intact at 37°C for over 24 h period. The remarkable stability of the grafted Substance P antagonists is one of the important factors that allow the peptides to retain in the blood for producing biological effects after oral administration.

6.3. Trypsin Inhibition Activity

Engineered peptides **TI-RF7** and **TI-NF7** also retained trypsin inhibition loop from SFTI-1, thus, beside their stability over trypsin digestion, they also possess trypsin inhibition activity. The inhibition activity was mainly because of TI loop, the peptides lacking TI loop (**RF7-T** and **NF7-T)** did not exhibit trypsin inhibition, although they possessed stability against trypsin enzyme. However, the engineered peptides (**TI-RF7** and **TI-NF7**), containing both a trypsin inhibition loop and substance P antagonist loop retained activity against trypsin in nano-molar concentrations, comparable to SFTI-1. Our results displayed that SFTI-1 scaffold can tolerate the foreign sequence without losing its trypsin inhibition activity.

6.4. Antagonistic Activity of Engineered Peptides

Previous studies showed that SP_{1-7} is a major bioactive metabolite of substance P often has opposite effects from its parent compound [234, 279, 280]. SP_{1-7} is responsible for antinociception [235-237], anti-inflammatory response [238] and antihyperalgesic effects in diabetic mice [239]. The binding sites for SP1-7 are found in spinal cord and brain but have not been cloned yet [281]. These binding sites have neither affinity for neurokinin receptor ligands nor μ-opioid receptor ligands [250]. Due to the lack of the known receptor, it was difficult to test the

functional activity of our synthesized peptides in the cell-based assay and limiting the option to only animal model study. However, due to time constraint we were not able to conduct animal model studies. Therefore, we used the nitric oxide measurement assay system.

NO is an important neurotransmitter which is involved in acute [282] and chronic [283] pain in both CNS [284] and PNS [285]. NO is one of the effector molecules produced by activated mouse macrophages [286] and substance P can increase the NO production in LPS-activated macrophages [287]. Our results also showed that the grafted analogs **TI-RF7** and the linear analog **RF7** inhibited the NO production in concentration dependent manner. However, **TI-NF7** and **NF7** do not show significant antagonistic activity. It could be because of the substitution of amino acid at Q6 to N6 in the primary sequence which may affect the structure conformation. Further NMR studies would be required to validate the structure conformation.

Furthermore, studies have shown that SFTI-1 can penetrate the mammalian cell membrane and later in 2014 *D'Souza et al*. confirmed that a non-specific endocytosis mechanism is responsible for cell permeability, regardless of mutations in the secondary loop. These findings suggest that the grafted analogs are likely to be cell permeable, an essential characteristic for oral bioavailability of peptides. In addition, the hydrophobicity of peptides also plays a role for cell penetration. Tam *et al*. have demonstrated that peptides with the cluster of hydrophobic residues are able to enter the cell [290]. Our designed substance P antagonists **TI-RF7** and **TI-NF7**, after grafting, contained 4 proline, 2 isoleucine and 1 phenylalanine, which made up 43% in comparison to full sequence. Thus, we speculate that these antagonists would possess cell penetrating property. In summary, our engineered peptide **TI-RF7** was proved to be bifunctional possessing both the trypsin inhibition activity and the substance P antagonistic activity. Moreover, to understand its role in nociception animal studies would be conducted in future.

Summary, Conclusion and Future Outlook

CRPs are mini-proteins, highly diverse in sequence and possess great stability against thermal and enzymatic degradation. They display a wide range of biological activities such as antimicrobial, immunomodulation, insecticidal and hormonal functions. During the past few years, this family of plant mini-proteins has attracted increasing interest as a rich source for peptide therapeutics development. In this thesis, the molecular and functional diversity of plant mini-proteins have been studied using genomic, proteomic and transcriptomic methods.

Chapter 3 describes the isolation and characterization of a family of CRPs from *Jasminum sambac* of Oleaceae family. These CRPs are named as jasmintides and consist of 27-29 amino acid residues. Jasmintides are expressed in a tissue-specific manner with unique profiles in all plant tissues examined, including flowers, leaves and roots. Using transcriptomic method, a total of 12 jasmintides has been identified. The expression of jasmintides in plants was further confirmed by proteome analysis. Proteomic analysis further expanded the suite of jasmintides to 82. Three sequences (jS1, jS2 and jS14) were confirmed by a genomic approach. In total 87 jasmintides have been discovered in *J. sambac*. The high tolerance for sequence substitution makes jasmintides hypervariable. This suggests the robustness and suitability of jasmintides as a scaffold for engineering peptide therapeutics.

My thesis also revealed that jasmintides are ribosomally synthesized and their precursor contains a three-domain arrangement. It comprises an ER signal peptide, a pro-domain followed by the mature peptide domain. This arrangement is very common in other CRPs. The understanding of the biosynthesis is useful for future application of jasmintides in crop protection using transgenic plants.

An important finding of my thesis is the novel cystine connectivity of jasmintides. Prior to my thesis only two types of cystine motifs were known which are cystine-knot and thionin-like arrangements. In my thesis, we discovered a novel disulfide connectivity of CysI-CysV, CysII-CysIV and CysIII-CysVI by disulfide mapping together with mass spectrometry and NMR.

Jasmintides exhibited exceptional thermal and proteolytic stability. In addition, jasmintides play a role in plant-host defense by affecting the growth of insect larvae and they also inhibit cytopathic effect of Infectious bronchitis virus in Vero cells. Taken together, jasmintides are a new family of plant CRPs with novel cystine connectivity and multiple bioactivities.

Chapter four describes the application of a CRP as a scaffold for grafting linear bioactive epitopes. We have grafted four substance P antagonists using the simplest CRP scaffold, the sunflower trypsin inhibitor-1. Sunflower trypsin inhibitor-1 is the smallest trypsin inhibitor with a cyclic backbone, conferring it stability against enzymatic degradation. The designed grafted substance P antagonists were synthesized by solid phase methodology using C-terminal thioethylbutyl amido (TEBA) thioester surrogate. The peptide-TEBA was transformed into thioester or thiolactones via *N-S* and *S-S* acyl shifts in the presence of internal or external thiol under acidic condition. Under the basic condition, the thioester/thiolactones underwent *S-S* and *S-N* acyl shift reactions to afford the cyclic peptide. DMSO (10% v/v) was applied to the cyclization mixture for disulfide formation in a one-pot synthetic scheme.

We showed that grafted substance P analogs (TI-RF7 and TI-NF7) improved stability against endo- and exopeptidases. End-to-end cyclized backbone made them metabolic stable in human serum and about 80% of TI-RF7 and TI-NF7 remained intact over 12 h. TI-RF7 and TI-NF7 also retained trypsin inhibition activity from SFTI-1 and showed trypsin inhibition in nano-molar concentrations with IC_{50} of 54.3 nM and 42.9 nM, respectively. Furthermore, our results also showed that the grafted analogs TI-RF7 and the linear analog RF7 inhibited the

NO production in concentration-dependent manner. Taken together, our results revealed that the grafted analogs are bifunctional, possessing both trypsin inhibition activity from SFTI-1 and substance P antagonistic activity.

In conclusion, this thesis uncovers the new CRP family with a novel disulfide connectivity of CysI-CysV, CysII-CysIV and CysIII-CysVI, which expands our knowledge of the occurrence of cystine motifs in plant CRPs. This scaffold could be useful for developing therapeutics and for designing peptide drugs.

My thesis supports the hypothesis that CRPs are an under-explored class of bioactive principle in medicinal plants. The outlook of further exploration of CRPs in this chemical space could provide promising therapeutics.

Publications and Poster Presentations

Kumari, G.; Serra, A.; Shin, J.; Nguyen, P. Q.; Tam. J. P.; (2015) Cysteine-rich peptide family with unusual disulfide connectivity from Jasminum sambac. *Journal of Natural Products*. 78(11): 2791-2799

Kumari, G.; Tam, J. P.; (2016) A bifunctional substance P antagonist using sunflower trypsin inhibitor as scaffold (In preparation).

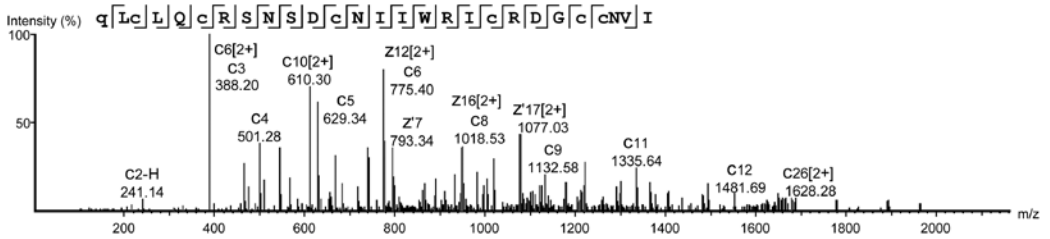
Kumari, G.; Serra, A.; Shin, J.; Tam, J.P.; (2016) A novel suit of jasmintides from Jasminum sambac: sequence variation and the implications for structure, function and stability (In preparation).

Kumari, G.; Tam, J. P.; (2015) Towards the goal in designing of metabolic stable bioactive substance P antagonist, *7th International Peptide Symposim,* Singapore

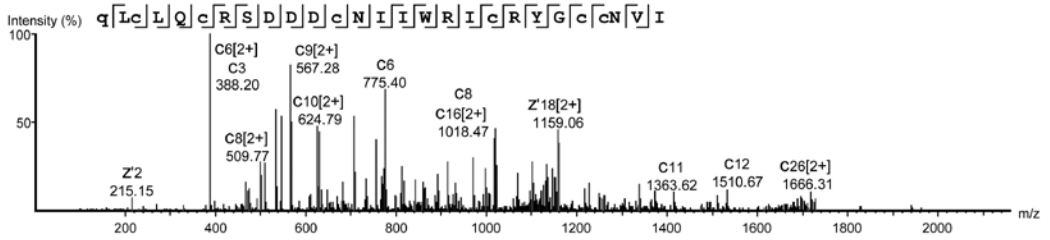
Kumari, G.; Serra, A.; Shin, J.; Tam, J.P. ; (2015) Discover of new scaffold for drug design from Jasminum sambac, *The 10th Annual Peptide Therapeutic Symposium*, Sandiago, USA

Appendix A

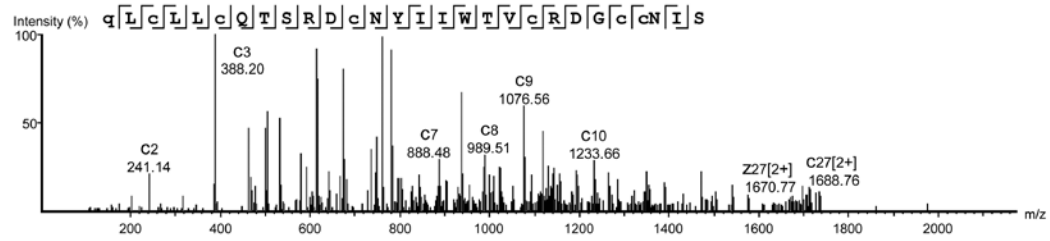
Jasmintide 1



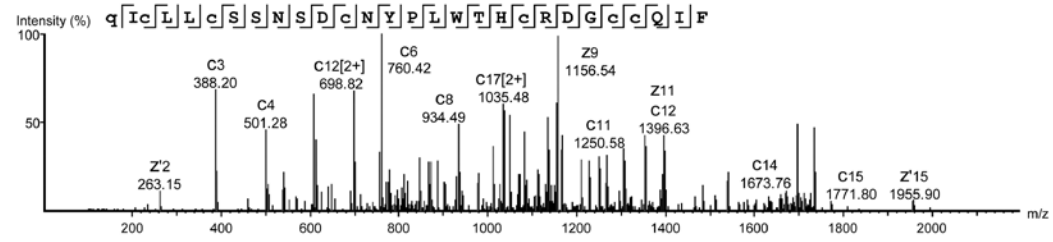
Jasmintide 2



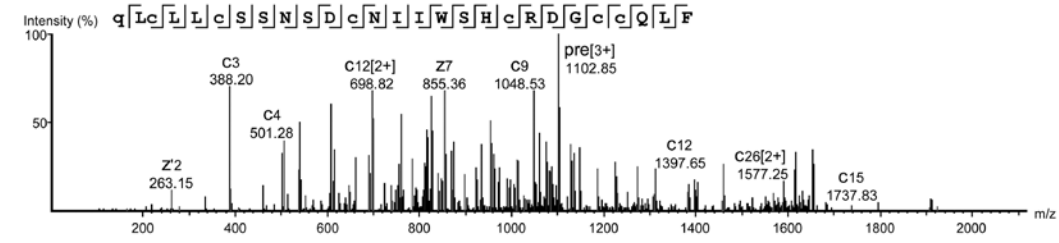
Jasmintide 3



Jasmintide 4



Jasmintide 5



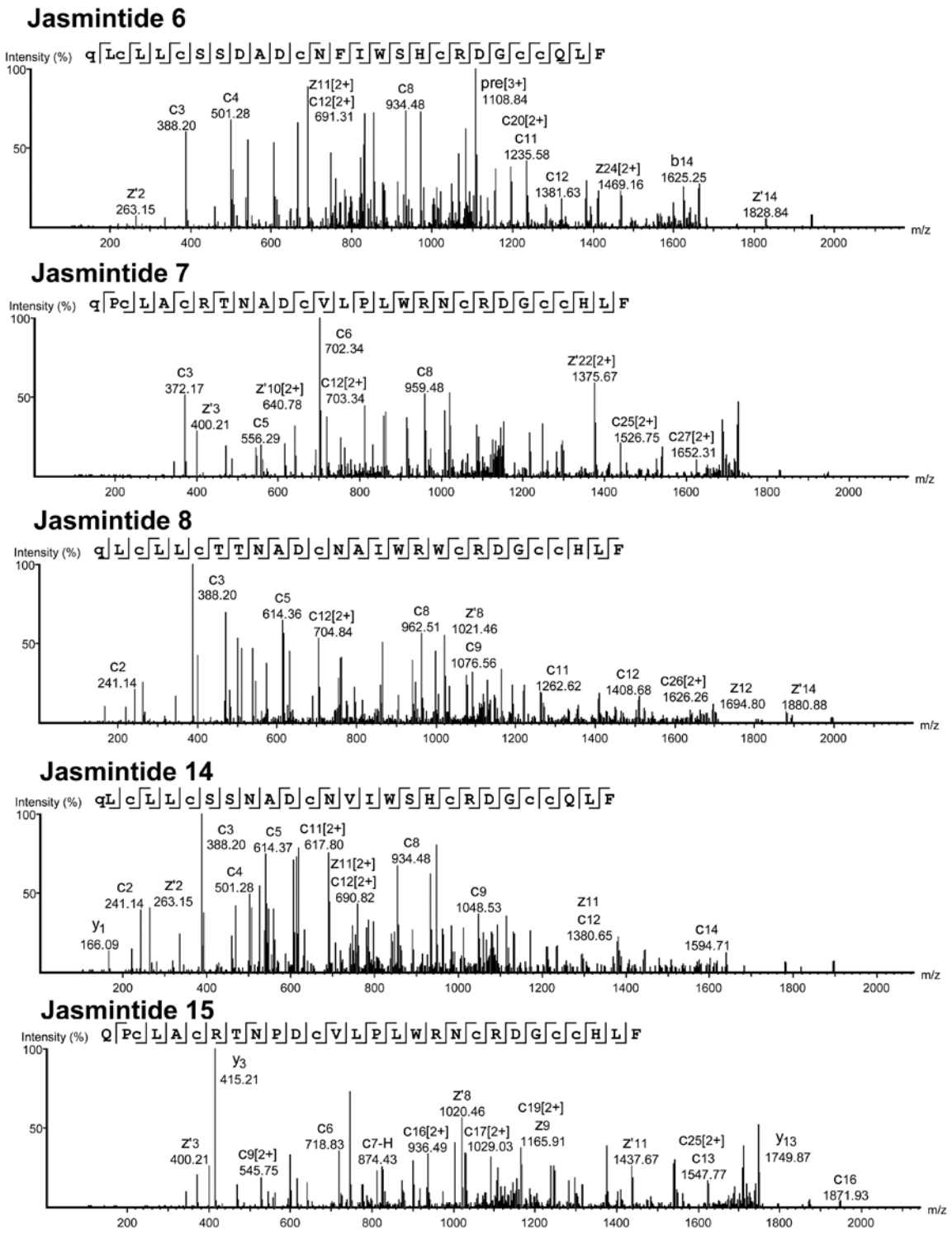


Figure A1. Annotated MS/MS spectra of jasmintides identified using proteomics approach. Fragments are labeled with *c*-, *z*-, *z*+1 (*z*'), *z*+2 (*z*(+2)), *b*- and *y*- ions.

Table A1. CRPs identified from leaf of *J. sambac* extract using LC-MS/MS

Jasmintide	Sequence	-10lgP[a]	Mass	ppm	z	#Spec	MOD[b]
JS1	C.LQC(+43.04)RSNSDC(+43.04)NIIWRIC(+43.04)RDGC(+43.04)C(+43.04)NVI	200	2998.4656	0.3	5	71	-
JS1	Q.LC(+43.04)LQC(+43.04)RSNSDC(+43.04)NIIWRIC(+43.04)RDGC(+43.04)C(+43.04)NVI	200	3257.6011	0.4	6	25	-
JS1	N.IIWRIC(+43.04)RDGC(+43.04)C(+43.04)NVI	200	1791.9423	-0.2	4	27	-
JS1	Q(-17.03)LC(+43.04)LQC(+43.04)RSNSDC(+43.04)NIIWRIC(+43.04)RDGC(+43.04)C(+43.04)N.V	186.11	3156.4807	0.7	5	13	Pyro-glu from Q
JS1	N.Q(-17.03)LC(+43.04)LQC(+43.04)RSNSDC(+43.04)NIIWRIC(+43.04)RDGC(+43.04)C(+43.04)NVI	172.71	3368.6331	1.2	6	252	Pyro-glu from Q
JS1	Q(-17.03)LC(+43.04)LQC(+43.04)RSNSDC(+43.04)NIIWRIC(+43.04)RDGC(+43.04)C(+43.04)NV.I	165.43	3255.5491	0.6	6	9	Pyro-glu from Q
JS1	I.IWRIC(+43.04)RDGC(+43.04)C(+43.04)NVI	153.12	1678.8582	-0.4	4	17	-
JS1	Q.C(+43.04)RSNSDC(+43.04)NIIWRIC(+43.04)RDGC(+43.04)C(+43.04)NVI	150.64	2757.323	-1.8	5	30	-
JS1	N.SDC(+43.04)NIIWRIC(+43.04)RDGC(+43.04)C(+43.04)NVI	110.5	2254.0955	-0.3	4	3	-
JS1	D.C(+43.04)NIIWRIC(+43.04)RDGC(+43.04)C(+43.04)NVI	108.48	2052.0366	-0.3	4	13	-
JS1	L.C(+43.04)LQC(+43.04)RSNSDC(+43.04)NIIWRIC(+43.04)RDGC(+43.04)C(+43.04)NVI	54.79	3144.5171	1	5	3	-
JS2	C.LQC(+43.04)RSDDDC(+43.04)NIIWRIC(+43.04)RYGC(+43.04)C(+43.04)N.V	200	2863.3284	-0.3	5	30	-
JS2	C.LQC(+43.04)RSDDDC(+43.04)NIIWRIC(+43.04)RYGC(+43.04)C(+43.04)NV.I	200	2962.397	-0.5	5	17	-
JS2	C.LQC(+43.04)RSDDDC(+43.04)NIIWRIC(+43.04)RYGC(+43.04)C(+43.04)NVI	200	3075.481	0.8	5	46	-
JS2	Q(-17.03)LC(+43.04)LQC(+43.04)RSDDDC(+43.04)NIIWRIC(+43.04)RYGC(+43.04)C(+43.04)NVI	193.97	3445.6484	0.8	6	98	Pyro-glu

							from Q
JS2	Q(-17.03)LC(+43.04)LQC(+43.04)RSDDDC(+43.04)NIIWRIC(+43.04)RYGC(+43.04)C(+43.04)NV.I	164.75	3332.5645	0.1	5	1	Pyro-glu from Q
JS3	N.Q(-17.03)LC(+43.04)LLC(+43.04)QTSRDC(+43.04)NYIIWTVC(+43.04)RDGC(+43.04)C(+43.04)NIS	163.12	3463.6477	1.2	6	5	Pyro-glu from Q
JS3	D.C(+43.04)NYIIWTVC(+43.04)RDGC(+43.04)C(+43.04)NIS	144.61	2133.9944	-0.9	4	2	-
JS4	L.C(+43.04)SSNSDC(+43.04)NYPLWTHC(+43.04)RDGC(+43.04)C(+43.04)QIF	165.22	2866.2229	-0.8	5	8	-
JS4	Q(-17.03)IC(+43.04)LLC(+43.04)SSNSDC(+43.04)NYPLWTHC(+43.04)RDGC(+43.04)C(+43.04)QIF	161.86	3462.5586	6.6	5	19	Pyro-glu from Q
JS4	D.C(+43.04)NYPLWTHC(+43.04)RDGC(+43.04)C(+43.04)QIF	151.69	2230.0056	-0.5	4	10	-
JS4	C.SSNSDC(+43.04)NYPLWTHC(+43.04)RDGC(+43.04)C(+43.04)QIF	132.32	2720.1716	-0.4	4	4	-
JS4	N.SDC(+43.04)NYPLWTHC(+43.04)RDGC(+43.04)C(+43.04)QIF	93.98	2432.0647	-0.4	4	2	-
JS5	Q(-17.03)LC(+43.04)LLC(+43.04)SSNSDC(+43.04)NIIWSHC(+43.04)RDGC(+43.04)C(+43.04)QLF	199.73	3301.511	0.3	5	6	Pyro-glu from Q
JS5	N.Q(-17.03)LC(+43.04)LLC(+43.04)SSNSDC(+43.04)NIIWSHC(+43.04)RDGC(+43.04)C(+43.04)QL.F	141.32	3154.4426	-1	5	1	Pyro-glu from Q
JS5	D.C(+43.04)NIIWSHC(+43.04)RDGC(+43.04)C(+43.04)QLF	138.02	2068.958	-0.4	4	2	-
JS5	L.C(+43.04)SSNSDC(+43.04)NIIWSHC(+43.04)RDGC(+43.04)C(+43.04)QLF	116.53	2705.1753	-0.5	4	8	-
JS5	N.Q(-17.03)LC(+43.04)LLC(+43.04)SSNSDC(+43.04)NIIWSHC(+43.04)RDGC(+43.04)C(+43.04)Q.L	109.85	3041.3584	0.4	5	1	Pyro-glu from Q

JS5	N.SDC(+43.04)NIIWSHC(+43.04)RDGC(+43.04)C(+43.04)QLF	91.96	2271.0168	4.1	4	2	-
JS6	Q(-17.03)LC(+43.04)LLC(+43.04)SSDADC(+43.04)NFIWSHC(+43.04)RDGC(+43.04)C(+43.04)QLF	181.17	3320.4844	6.8	5	5	Pyro-glu from Q
JS6	D.C(+43.04)NFIWSHC(+43.04)RDGC(+43.04)C(+43.04)QLF	200	2102.9424	-0.7	4	2	-
JS14	Q(-17.03)LC(+43.04)LLC(+43.04)SSNADC(+43.04)NVIWSHC(+43.04)RDGC(+43.04)C(+43.04)QLF	149.12	3271.5005	7.4	5	8	Pyro-glu from Q
JS14	Q(-17.03)LC(+43.04)LLC(+43.04)SSNADC(+43.04)NVIWSHC(+43.04)RDGC(+43.04)C(+43.04)QL.F	90.54	3124.4319	-0.2	5	1	Pyro-glu from Q
JS14	L.C(+43.04)SSNADC(+43.04)NVIWSHC(+43.04)RDGC(+43.04)C(+43.04)QLF	68.5	2675.1648	0.4	4	2	Pyro-glu from Q

[a]In Peaks a higher -10logP value indicates a more confident sequencing result. Sequences are identified using false discovery rate (FDR) 0.01%.
[b]MOD: side chain modifications and chemical derivatization. All cysteine residues are alkylated with ethylamine and labelled as EA.

Table A2. CRPs identified from flower of *J. sambac* extract using LC-MS/MS

Jasmintide	Sequence	-10lgP[a]	Mass	ppm	z	#Spec	MOD[b]
JS1	C.LQC(+43.04)RSNSDC(+43.04)NIIWRIC(+43.04)RDGC(+43.04)C(+43.04)NVI	200	2998.4656	-0.1	5	19	-
JS1	N.Q(-17.03)LC(+43.04)LQC(+43.04)RSNSDC(+43.04)NIIWRIC(+43.04)RDGC(+43.04)C(+43.04)NVI	178.52	3368.6331	-0.5	6	55	Pyro-glu from Q; EA
JS1	Q.C(+43.04)RSNSDC(+43.04)NIIWRIC(+43.04)RDGC(+43.04)C(+43.04)NVI	160.47	2757.323	0.1	5	11	-
JS1	N.Q(-17.03)LC(+43.04)LQC(+43.04)RSNSDC(+43.04)NIIWRIC(+43.04)RDGC(+43.04)C(+43.04)N.V	154.03	3156.4807	-0.7	5	5	Pyro-glu from Q
JS1	N.Q(-17.03)LC(+43.04)LQC(+43.04)RSNSDC(+43.04)NIIWRIC(+43.04)RDGC(+43.04)C(+43.04)NV.I	145.8	3255.5491	0	5	2	Pyro-glu from Q
JS1	D.C(+43.04)NIIWRIC(+43.04)RDGC(+43.04)C(+43.04)NVI	121.83	2052.0366	-0.7	4	3	-
JS1	L.C(+43.04)LQC(+43.04)RSNSDC(+43.04)NIIWRIC(+43.04)RDGC(+43.04)C(+43.04)NVI	117.6	3144.5171	0.2	5	3	-
JS2	C.LQC(+43.04)RSDDDC(+43.04)NIIWRIC(+43.04)RYGC(+43.04)C(+43.04)N.V	200	2863.3284	0.3	5	6	-
JS2	C.LQC(+43.04)RSDDDC(+43.04)NIIWRIC(+43.04)RYGC(+43.04)C(+43.04)NVI	176.89	3075.481	-0.1	5	4	-
JS2	Q(-17.03)LC(+43.04)LQC(+43.04)RSDDDC(+43.04)NIIWRIC(+43.04)RYGC(+43.04)C(+43.04)NVI	163.18	3445.6484	-0.4	5	19	Pyro-glu from Q
JS2	Q(-17.03)LC(+43.04)LQC(+43.04)RSDDDC(+43.04)NIIWRIC(+43.04)RYGC(+43.04)C(+43.04)NV.I	155.73	3332.5645	-0.4	5	3	Pyro-glu from Q
JS3	Y.IIWTVC(+43.04)RDGC(+43.04)C(+43.04)NIS	167.9	1710.8368	-0.6	4	10	-
JS3	N.Q(-17.03)LC(+43.04)LLC(+43.04)QTSRDC(+43.04)NYIIW.T	163.06	2183.0691	-0.8	4	9	Pyro-glu from Q
JS3	N.Q(-17.03)LC(+43.04)LLC(+43.04)QTSRDC(+43.04)NYIIWTVC(+43.04)RDGC(+43.04)C(+43.04)NIS	153.61	3463.6477	4.7	5	15	Pyro-glu from Q
JS3	N.Q(-17.03)LC(+43.04)LLC(+43.04)QTSRDC(+43.04)N.Y	133.47	1607.7582	0.6	4	37	Pyro-glu

							from Q
JS3	L.C(+43.04)QTSRDC(+43.04)NYIIWTVC(+43.04)RDGC(+43.04)C(+43.04)NIS	120.33	2867.3123	-0.3	5	16	-
JS3	D.C(+43.04)NYIIWTVC(+43.04)RDGC(+43.04)C(+43.04)NIS	117.04	2133.9944	-0.4	4	18	-
JS3	N.Q(-17.03)LC(+43.04)LLC(+43.04)QTSRD.C	111.19	1347.6639	-0.2	3	13	Pyro-glu
							from Q
JS3	N.YIIWTVC(+43.04)RDGC(+43.04)C(+43.04).N	76.12	1559.7411	1	3	2	-
JS3	N.Q(-17.03)LC(+43.04)LLC(+43.04)QTSRDC(+43.04)NY.I	74.91	1770.8215	0.3	3	1	Pyro-glu
							from Q; EA
JS3	C.QTSRDC(+43.04)NYIIWTVC(+43.04)RDGC(+43.04)C(+43.04)NIS	64.09	2721.2607	0.1	5	3	-
JS3	Q(-17.03)LC(+43.04)LLC(+43.04)QTSRDC(+43.04)NY.I	49.81	1770.8215	0.3	3	1	Pyro-glu
							from Q
JS3	I.IWTVC(+43.04)RDGC(+43.04)C(+43.04)NIS	44.85	1597.7527	-0.4	3	4	-
JS3	L.LC(+43.04)QTSRDC(+43.04)NYIIWTVC(+43.04)RDGC(+43.04)C(+43.04)NIS	42.73	2980.3962	1.4	5	4	-
JS3	N.QLC(+43.04)LLC(+43.04)QTSRDC(+43.04)NYIIWTVC(+43.04)RDGC(+43.04)C(+43.04)NIS	41.89	3480.6743	-8	5	8	-
JS4	L.C(+43.04)SSNSDC(+43.04)NYPLWTHC(+43.04)RDGC(+43.04)C(+43.04)QIF	154.64	2866.2229	0	5	57	-
JS4	Q(-17.03)IC(+43.04)LLC(+43.04)SSNSDC(+43.04)NYPLWTHC(+43.04)RDGC(+43.04)C(+43.04)QIF	153.62	3462.5586	2	5	67	Pyro-glu
							from Q
JS4	C.SSNSDC(+43.04)NYPLWTHC(+43.04)RDGC(+43.04)C(+43.04)QIF	133.19	2720.1716	-0.3	5	26	-
JS4	D.C(+43.04)NYPLWTHC(+43.04)RDGC(+43.04)C(+43.04)QIF	121.19	2230.0056	-0.6	4	46	-
JS4	L.LC(+43.04)SSNSDC(+43.04)NYPLWTHC(+43.04)RDGC(+43.04)C(+43.04)QIF	109.64	2979.3071	-0.7	5	14	-
JS4	N.SDC(+43.04)NYPLWTHC(+43.04)RDGC(+43.04)C(+43.04)QIF	90.58	2432.0647	-0.7	5	5	-
JS4	Q(-17.03)IC(+43.04)LLC(+43.04)SSNSDC(+43.04)NYPLWTHC(+43.04)RDGC(+43.04)C(+43.04)QI.F	79.59	3315.4902	0	5	2	Pyro-glu
							from Q
JS4	QIC(+43.04)LLC(+43.04)SSNSDC(+43.04)NYPLWTHC(+43.04)RDGC(+43.04)C(+43.04)QIF	58.47	3479.5852	-8.9	5	10	-
JS4	Q(-17.03)IC(+43.04)LLC(+43.04)SSNSDC(+43.04)NYPL.W	33.38	1883.858	0.6	3	1	Pyro-glu
							from Q

JS5	Q(-17.03)LC(+43.04)LLC(+43.04)SSNSDC(+43.04)NIIWSHC(+43.04)RDGC(+43.04)C(+43.04)QLF	200	3301.511	0.2	5	15	Pyro-glu from Q
JS5	N.Q(-17.03)LC(+43.04)LLC(+43.04)SSNSDC(+43.04)NIIWSHC(+43.04)RDGC(+43.04)C(+43.04)QL.F	144	3154.4426	-0.5	5	7	Pyro-glu from Q
JS5	N.Q(-17.03)LC(+43.04)LLC(+43.04)SSNSDC(+43.04)NIIWSHC(+43.04)RDGC(+43.04)C(+43.04)Q.L	137.26	3041.3584	0.2	5	1	Pyro-glu from Q
JS5	D.C(+43.04)NIIWSHC(+43.04)RDGC(+43.04)C(+43.04)QLF	106.77	2068.958	-0.7	4	2	-
JS5	L.C(+43.04)SSNSDC(+43.04)NIIWSHC(+43.04)RDGC(+43.04)C(+43.04)QLF	98.29	2705.1753	-0.1	5	18	-
JS5	I.IWSHC(+43.04)RDGC(+43.04)C(+43.04)QLF	56.19	1695.7795	-1	4	3	-
JS5	C.SSNSDC(+43.04)NIIWSHC(+43.04)RDGC(+43.04)C(+43.04)QLF	51.01	2559.124	0.4	5	4	-
JS5	L.LC(+43.04)SSNSDC(+43.04)NIIWSHC(+43.04)RDGC(+43.04)C(+43.04)QLF	45.77	2818.2593	0	5	2	-
JS5	T.LSVQPNQLC(+43.04)LLC(+43.04)SSNSDC(+43.04)NIIWSHC(+43.04)RDGC(+43.04)C(+43.04)QLF	38.67	3956.8762	-8.4	5	8	-
JS4/JS5	N.Q(-17.03)LC(+43.04)LLC(+43.04)SSNSDC(+43.04).N	97.07	1396.6149	-0.9	3	4	Pyro-glu from Q
JS4/JS5	N.Q(-17.03)LC(+43.04)LLC(+43.04)SSNSDC(+43.04)N.I	77.57	1510.6578	0.8	3	7	Pyro-glu from Q
JS6	Q(-17.03)LC(+43.04)LLC(+43.04)SSDADC(+43.04)NFIWSHC(+43.04)RDGC(+43.04)C(+43.04)QLF	200	3320.4844	1	5	49	Pyro-glu from Q
JS6	QLC(+43.04)LLC(+43.04)SSDADC(+43.04)NFIWSHC(+43.04)RDGC(+43.04)C(+43.04)QLF	120.44	3337.511	-4.3	5	4	-
JS6	D.C(+43.04)NFIWSHC(+43.04)RDGC(+43.04)C(+43.04)QLF	108.67	2102.9424	-0.5	4	15	-
JS6	F.IWSHC(+43.04)RDGC(+43.04)C(+43.04)QLF	37.53	1695.7795	-1	4	3	-
JS7	Q(-17.03)PC(+43.04)LAC(+43.04)RTNADC(+43.04)VLPLWRNC(+43.04)RDGC(+43.04)C(+43.04)HL.F	176.56	3300.5857	-1.1	5	7	Pyro-glu from Q
JS7	Q(-17.03)PC(+43.04)LAC(+43.04)RTNADC(+43.04)VLPLWRNC(+43.04)RDGC(+43.04)C(+43.04)H.L	150.62	3187.5017	-0.1	5	9	Pyro-glu from Q
JS7	Q(-	138.39	3447.6541	0.4	5	30	Pyro-glu

	17.03)PC(+43.04)LAC(+43.04)RTNADC(+43.04)VLPLWRNC(+43.04)RDGC(+43.04)C(+43.04)HLF						from Q
JS7	R.TNADC(+43.04)VLPLWRNC(+43.04)RDGC(+43.04)C(+43.04).H	135.94	2210.033	-0.4	4	5	-
JS7	Q(-17.03)PC(+43.04)LAC(+43.04)RTNADC(+43.04)VLPLWRNC(+43.04)RDGC(+43.04)C(+43.04).H	104.52	3050.4429	0.1	5	5	Pyro-glu from Q
JS7	R.TNADC(+43.04)VLPLWRNC(+43.04)RDGC(+43.04)C(+43.04)H.L	91.74	2347.0918	-0.1	5	8	-
JS7	C.RTNADC(+43.04)VLPLWRNC(+43.04)RDGC(+43.04)C(+43.04)HLF	66.88	2763.3455	0.4	5	4	-
JS7	C.RTNADC(+43.04)VLPLWRNC(+43.04)RDGC(+43.04)C(+43.04).H	59.1	2366.134	-0.7	5	4	-
JS7	P.LWRNC(+43.04)RDGC(+43.04)C(+43.04)HLF	49.74	1750.833	0.7	4	5	-
JS7	A.C(+43.04)RTNADC(+43.04)VLPLWRNC(+43.04)RDGC(+43.04)C(+43.04).H	47.17	2512.1855	-1.3	5	3	-
JS14	Q(-17.03)LC(+43.04)LLC(+43.04)SSNADC(+43.04)NVIWSHC(+43.04)RDGC(+43.04)C(+43.04)QL.F	200	3124.4319	-0.3	5	9	Pyro-glu from Q
JS14	Q(-17.03)LC(+43.04)LLC(+43.04)SSNADC(+43.04)NVIWSHC(+43.04)RDGC(+43.04)C(+43.04)QLF	179.32	3271.5005	0.1	5	28	Pyro-glu from Q
JS14	L.C(+43.04)SSNADC(+43.04)NVIWSHC(+43.04)RDGC(+43.04)C(+43.04)QLF	107.86	2675.1648	1	5	24	-
JS14	N.VIWSHC(+43.04)RDGC(+43.04)C(+43.04)QLF	88.73	1794.848	-1.3	4	1	-
JS14	Q(-17.03)LC(+43.04)LLC(+43.04)SSNADC(+43.04).N	82.23	1380.62	0.4	3	3	Pyro-glu from Q
JS14	Q(-17.03)LC(+43.04)LLC(+43.04)SSNADC(+43.04)N.V	78.98	1494.663	0	3	10	Pyro-glu from Q
JS14	Q(-17.03)LC(+43.04)LLC(+43.04)SSNADC(+43.04)NVI.W	58.49	1706.8154	0.6	3	1	Pyro-glu from Q
JS15	QPC(+43.04)LAC(+43.04)RTNPDC(+43.04)VLPLWRNC(+43.04)RDGC(+43.04)C(+43.04)HLF	145.47	3490.6963	0.7	5	175	-

[a]In Peaks a higher -10logP value indicates a more confident sequencing result. Sequences are identified using false discovery rate (FDR) 0.01%.
[b]MOD: sidechain modifications and chemical derivatization. All cysteine residues are alkylated with ethylamine and labelled as EA.

Table A3. CRPs identified from root of *J. sambac* extract using LC-MS/MS

Jasmintide	Sequence	-10lgP[a]	Mass	ppm	z	#Spec	MOD[b]
JS1	N.Q(-17.03)LC(+43.04)LQC(+43.04)RSNSDC(+43.04)NIIWRIC(+43.04)RDGC(+43.04)C(+43.04)NVI	119.44	3368.6331	0.2	5	2	Pyro-glu from Q;
JS1	L.C(+43.04)LQC(+43.04)RSNSDC(+43.04)NIIWRIC(+43.04)RDGC(+43.04)C(+43.04)NVI	55.96	3144.5171	-1.4	5	3	-
JS2	Q(-17.03)LC(+43.04)LQC(+43.04)RSDDDC(+43.04)NIIWRIC(+43.04)RYGC(+43.04)C(+43.04)NVI	105.33	3446.6323	4.1	6	6	Pyro-glu from Q
JS3	Q(-17.03)LC(+43.04)LLC(+43.04)QTSRDC(+43.04)N.C	91.17	1607.7582	-0.1	3	9	Pyro-glu from Q
JS5	N.Q(-17.03)LC(+43.04)LLC(+43.04)SSNSDC(+43.04)NIIWSHC(+43.04)RDGC(+43.04)C(+43.04)QLF	200	3301.511	-0.1	5	12	Pyro-glu from Q
JS5	L.C(+43.04)LLC(+43.04)SSNSDC(+43.04)NIIWSHC(+43.04)RDGC(+43.04)C(+43.04)QLF	114.15	3077.3948	1.8	5	2	-
JS5/JS6	F.IWSHC(+43.04)RDGC(+43.04)C(+43.04)QLF	116.95	1695.7795	0.6	4	25	-
JS6	Q(-17.03)LC(+43.04)LLC(+43.04)SSDADC(+43.04)NFIWSHC(+43.04)RDGC(+43.04)C(+43.04)QLF	200	3320.4844	0.3	5	100	Pyro-glu from Q
JS6	N.Q(-17.03)LC(+43.04)LLC(+43.04)SSDADC(+43.04)N.F	90.03	1495.647	-0.1	3	6	Pyro-glu from Q
JS7	N.Q(-17.03)PC(+43.04)LAC(+43.04)RTNADC(+43.04)VLPLWRNC(+43.04)RDGC(+43.04)C(+43.04)HL.F	200	3300.5857	0.3	5	34	Pyro-glu from Q
JS7	N.Q(-17.03)PC(+43.04)LAC(+43.04)RTNADC(+43.04)VLPLWRNC(+43.04)RDGC(+43.04)C(+43.04)HLF	190.75	3447.6541	0.7	5	85	Pyro-glu from Q
JS7	N.Q(-17.03)PC(+43.04)LAC(+43.04)RTNADC(+43.04)VL.P	159.69	1617.7789	0.3	4	56	Pyro-glu from Q
JS7	D.C(+43.04)VLPLWRNC(+43.04)RDGC(+43.04)C(+43.04)HLF	158.47	2206.0896	1.6	4	4	-

JS7	N.Q(-17.03)PC(+43.04)LAC(+43.04)RTNADC(+43.04)VLPLWRNC(+43.04)RDGC(+43.04)C(+43.04)H.L	158.2	3187.5017	0.5	5	5	Pyro-glu from Q
JS7	Q(-17.03)PC(+43.04)LAC(+43.04)RTNADC(+43.04)VL.P	115.8	1617.7789	0.3	4	54	Pyro-glu from Q
JS7	N.Q(-17.03)PC(+43.04)LAC(+43.04)RTNADC(+43.04)VLP.L	110.95	1714.8317	-0.5	4	13	Pyro-glu from Q
JS7	L.WRNC(+43.04)RDGC(+43.04)C(+43.04)HLF	96.7	1637.749	0	4	3	-
JS7	N.QPC(+43.04)LAC(+43.04)RTNADC(+43.04)VLPLWRNC(+43.04)RDGC(+43.04)C(+43.04)HLF	89.43	3464.6807	-5.1	6	8	-
JS7	QPC(+43.04)LAC(+43.04)RTNADC(+43.04)VLPLWRNC(+43.04)RDGC(+43.04)C(+43.04)HLF	76	3464.6807	-5.1	6	8	-
JS7	D.C(+43.04)VLPLWRNC(+43.04)RDGC(+43.04)C(+43.04)HL.F	70.17	2059.0212	0.7	4	6	-
JS8	N.Q(-17.03)LC(+43.04)LLC(+43.04)TTNADC(+43.04)NAIWRWC(+43.04)RDGC(+43.04)C(+43.04)HLF	182.32	3398.5901	0.6	5	53	Pyro-glu from Q
JS8	L.C(+43.04)TTNADC(+43.04)NAIWRWC(+43.04)RDGC(+43.04)C(+43.04)HLF	167.41	2802.2546	-1.8	5	7	-
JS8	N.Q(-17.03)LC(+43.04)LLC(+43.04)TTNADC(+43.04)N.A	94.24	1522.6942	-0.5	3	2	Pyro-glu from Q
jS14	Q(-17.03)LC(+43.04)LLC(+43.04)SSNADC(+43.04)NVIWSHC(+43.04)RDGC(+43.04)C(+43.04)QLF	200	3271.5005	-4.6	5	6	-

[a]In Peaks a higher -10logP value indicates a more confident sequencing result. Sequences are identified using false discovery rate (FDR) 0.01%.
[b]MOD: side chain modifications and chemical derivatization. All cysteine residues are alkylated with ethylamine and labelled as EA.

Appendix B

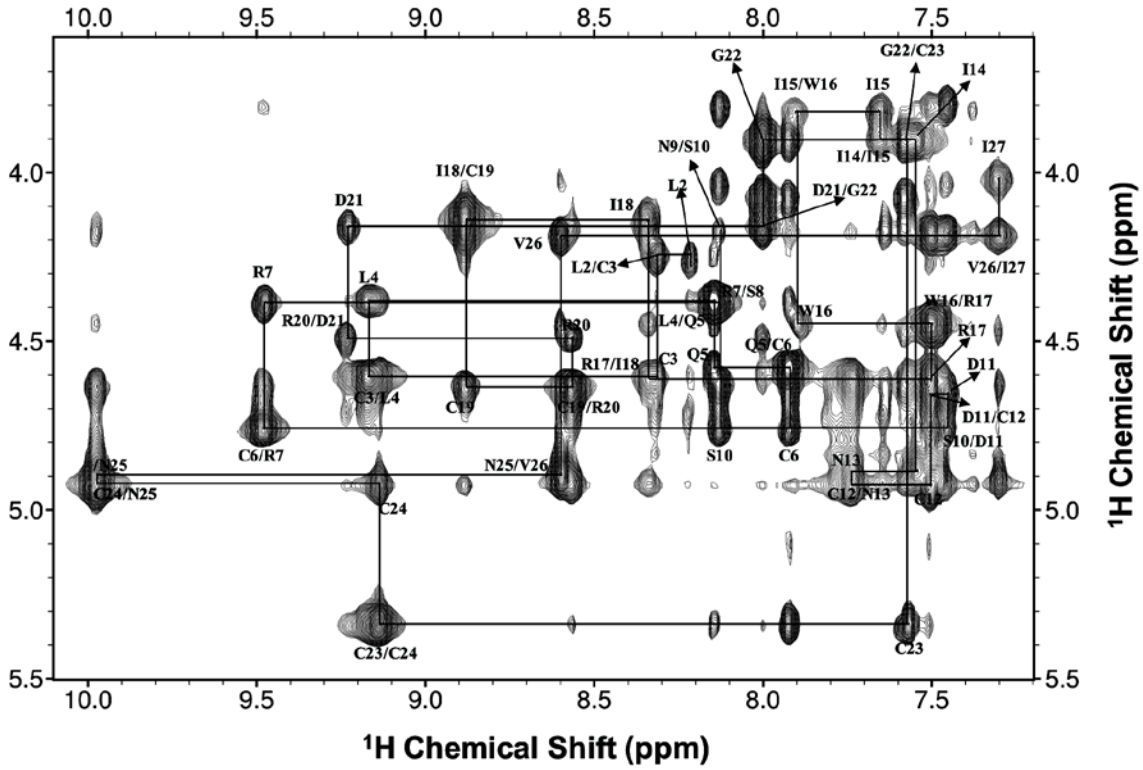


Figure B1. NH/CαH fingerprint region of 2D-NOESY spectrum (mixing time, 300 ms) of jS1 recorded in 90%H_2O/10%D_2O at 298K. Sequential connectivities of each amino acid residue are shown by solid lines.

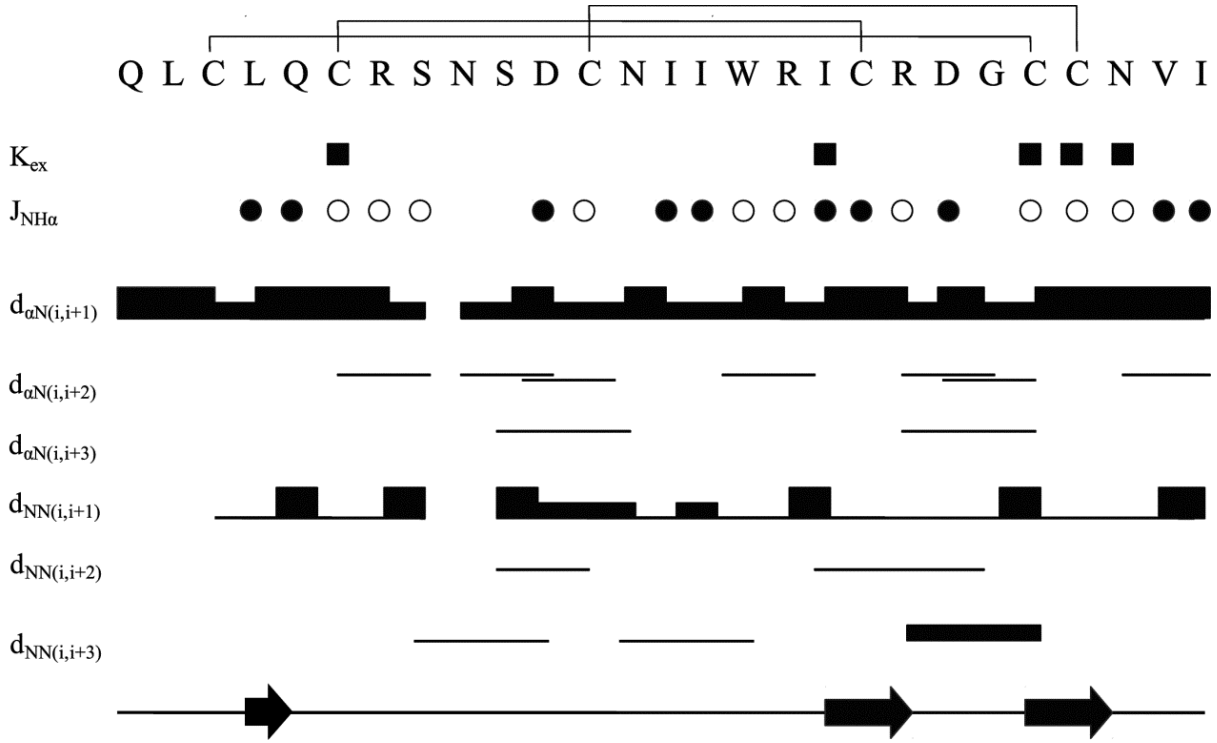


Figure B2. Summary of sequential and medium NOE connectivity. Slowly exchanging amide protons (filled box) and backbone vicinal coupling constants (open circle: $^{3}J_{HN\alpha} < 6Hz$, filled circle: $^{3}J_{HN\alpha} > 8Hz$) are indicated.

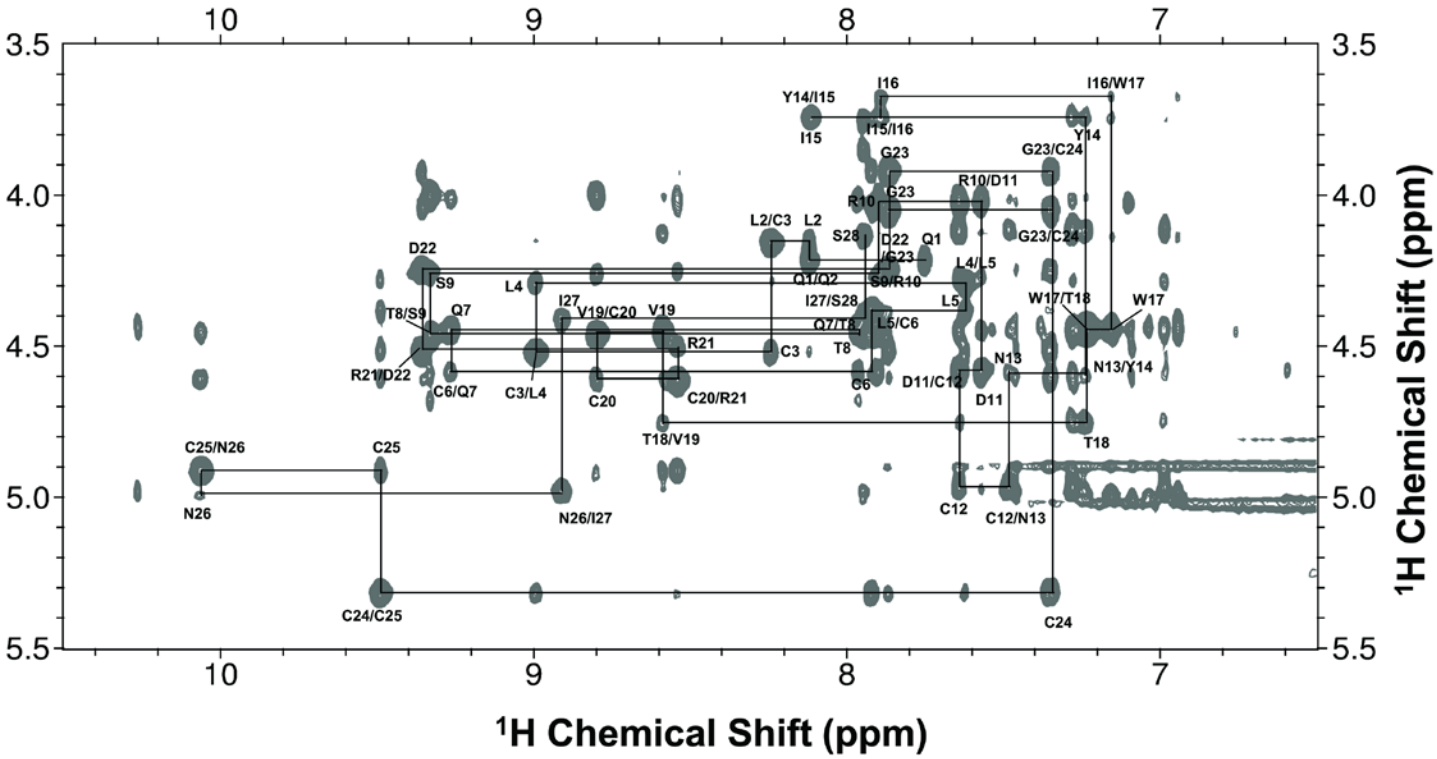


Figure B3. NH/CαH fingerprint region of 2D-NOESY spectrum (mixing time, 300 ms) of jS3 recorded in 90%H2O/10%D2O at 298K. Sequential connectivities of each amino acid residue are shown by solid lines.

.

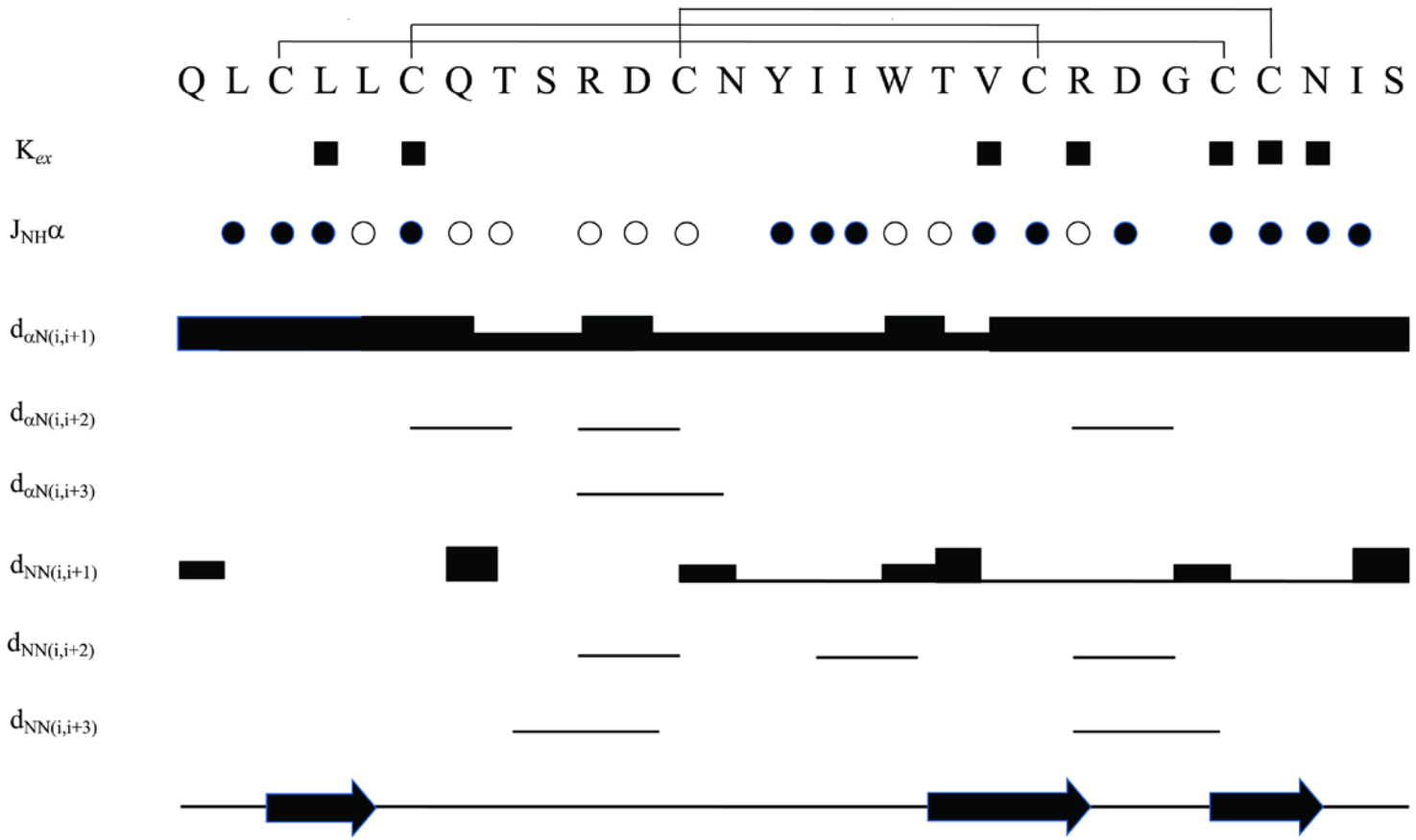


Figure B4. Summary of sequential and medium NOE connectivity. Slowly exchanging amide protons (filled box) and backbone vicinal coupling constants (open circle: ^{3}JHNα < 6Hz, filled circle: ^{3}JHNα > 8Hz) are indicated.

Table B1. Proton NMR chemical shift assignments for jS1

Residues	NH	CαH	CβH	Others
Gln1	-	4.27	1.96, 1.94	CγH 2.44, 2.33
Leu2	8.22	4.25	1.50	CδH 0.82, 0.77
Cys3	8.31	4.60	3.10	
Leu4	9.17	4.38	1.52, 1.40	CδH 0.83, 0.78
Gln5	8.15	4.58	1.94,1.91	CγH 2.39, 2.34 $N\varepsilon H^2$ 7.51, 6.74
Cys6	7.92	4.77	3.22, 3.03	
Arg7	9.48	4.38	1.86, 1.75	CγH 1.61 CδH 3.13 NεH 7.23
Ser8	8.13	4.63	4.03, 3.81	
Asn9	-	4.18	3.06, 2.76	$N\delta H^2$ 7.64, 7.00
Ser10	8.13	4.76	4.18, 3.80	
Asp11	7.45	4.66	3.05, 2.76	
Cys12	7.51	4.94	2.77	
Asn13	7.74	4.89	3.07, 2.77	$N\delta H^2$ 7.77, 6.33
Ile14	7.55	3.90	1.85	$C\gamma H^1$ 1.33, 1.16 $C\gamma H^2$ 0.85 $C\delta H^1$ 0.79
Ile15	7.66	3.82	1.38	$C\gamma H^1$ 1.09, 0.82 $C\gamma H^2$ 0.12 $C\delta H^1$ 0.53
Trp16	7.90	4.45	3.00, 2.54	$N\varepsilon H^1$ 10.18 $C\delta H^1$ 7.08 $C\varepsilon H^3$ 7.48
Arg17	7.50	4.60	1.78, 1.71	CγH 1.68, 1.53 CδH 3.10, 3.80
Ile18	8.43	4.13	1.64	$C\gamma H^1$ 1.24, 1.18 $C\gamma H^2$ 0.68 $C\gamma H^1$ 0.64
Cys19	8.82	4.63	2.76, 2.56	
Arg20	8.57	4.49	1.61, 1.55	CγH 1.46, 1.38 CδH 3.16, 3.04 NεH 7.80
Asp21	9.23	4.16	2.81, 2.54	
Gly22	8.00	4.07, 3.91		
Cys23	7.58	5.35	3.21, 2.66	
Cys24	9.14	4.93	3.38, 2.67	
Asn25	9.98	4.90	2.84, 2.52	$N\delta H^2$ 7.30, 7.07
Val26	8.60	4.19	1.98	CγH 0.89, 0.83
Ile27	7.3	4.02	1.68	$C\gamma H^1$ 1.38, 1.10 $C\gamma H^2$ 0.79 $C\delta H^1$ 0.76

Table B2. Proton NMR chemical shift assignments for jS3

Residues	NH	CαH	CβH	Others
Gln1	7.75	4.22	1.88	CγH 2.29, 2.39
Leu2	8.12	4.15	1.40,1.35	CγH 0.87 CδH 0.74, 0.66
Cys3	8.25	4.52	2.99	
Leu4	8.99	4.29	1.33, 1.17	CγH 1.09 CδH 0.48, -0.07
Leu5	7.62	4.38	1.58,1.41	CδH 0.85, 0.88
Cys6	7.92	4.58	3.19, 3.12	
Gln7	9.27	4.45	2.08	CγH 2.34, 2.28 NHε^2 7.53, 6.76
Thr8	7.97	4.46	-	CγH^2 1.15
Ser9	9.33	4.26	4.00	
Arg10	7.90	4.02	1.69	CγH 1.63, 1.58 CδH 3.11 NHε 7.10
Asp11	7.57	4.58	3.11, 2.83	
Cys12	7.64	4.97	2.90	
Asn13	7.48	4.76	2.50, 2.46	NHδ^2 7.35, 6.59
Tyr14	7.25	3.75	3.36,2.81	CδH 7.28, CεH 6.99
Ile15	8.12	3.74	1.73	CγH^1 1.40, 1.09 CγH^2 0.77 CδH^1 0.75
Ile16	7.89	3.68	0.82	CγH^1 0.63 CγH^2 0.46 CδH^1 0.08
Trp17	7.16	4.44	2.99, 2.55	CδH^1 6.94, NεH^1 10.27
Thr18	7.24	4.76	4.45	CγH^2 1.09
Val19	8.59	4.46	1.69	CγH 0.79, 0.71
Cys20	8.80	4.61	2.71, 2.37	
Arg21	8.54	4.51	1.68, 1.60	CγH 1.51, 1.34 CδH 3.15, 3.05 NHε 7.64
Asp22	9.36	4.25	3.03, 2.68	
Gly23	7.87	4.05, 3.92	-	
Cys24	7.35	5.32	3.25, 2.58	
Cys25	9.49	4.91	3.14, 2.61	
Asn26	10.06	4.99	2.89, 2.53	NHδ^2 7.46, 7.09
Ilel27	8.91	4.41	1.82	CγH^1 1.45, 1.04 CγH^2 0.89 CδH^1 0.79
Ser28	7.95	4.14	3.85,3.76	

References

1. Su C: **Snake needle grass (Hedyotis diffusa): a mini review**. *Journal of Medicinal Food Plants* 2010, **2**(1).

2. Kamboj V: **Herbal medicine**. *Current Science* 2000, **78**(1):35-38.

3. Al-Achi A: **An introduction to botanical medicines: history, science, uses, and dangers**: ABC-CLIO; 2008.

4. Rates SMK: **Plants as source of drugs**. *Toxicon* 2001, **39**(5):603-613.

5. Gurib-Fakim A: **Medicinal plants: traditions of yesterday and drugs of tomorrow**. *Molecular aspects of Medicine* 2006, **27**(1):1-93.

6. Norn S, Kruse P, Kruse E: **[History of opium poppy and morphine]**. *Dansk Medicinhist Arbog* 2004, **33**:171-184.

7. Ball P: **History of science: quinine steps back in time**. *Nature* 2008, **451**(7182):1065-1066.

8. Wick J: **Aspirin: a history, a love story**. *The Consultant Pharmacist®* 2012, **27**(5):322-329.

9. Wong G: **Biotech scientists bank on big pharma's biologics push**. *Nature Biotechnology* 2009, **27**(3):293-295.

10. Obradovic M, Mrhar A, Kos M: **Market uptake of biologic and small-molecule - Targeted oncology drugs in Europe**. *Clinical Therapeutics* 2009, **31**(12):2940-2952.

11. Ryan JT, Ross RP, Bolton D, Fitzgerald GF, Stanton C: **Bioactive peptides from muscle sources: meat and fish**. *Nutrients* 2011, **3**(9):765-791.

12. Suarez-Jimenez G-M, Burgos-Hernandez A, Ezquerra-Brauer J-M: **Bioactive peptides and depsipeptides with anticancer potential: Sources from marine animals**. *Marine Drugs* 2012, **10**(5):963-986.

13. Adams ME: **Agatoxins: ion channel specific toxins from the American funnel web spider, Agelenopsis aperta**. *Toxicon* 2004, **43**(5):509-525.

14. Han TS, Teichert RW, Olivera BM, Bulaj G: **Conus venoms-a rich source of peptide-based therapeutics**. *Current Pharmaceutical Design* 2008, **14**(24):2462-2479.

15. Baumann A: **Early development of therapeutic biologics-pharmacokinetics**. *Current Drug Metabolism* 2006, **7**(1):15-21.

16. Broekaert WF, Cammue BP, De Bolle MF, Thevissen K, De Samblanx GW, Osborn RW, Nielson K: **Antimicrobial peptides from plants**. *Critical Reviews in Plant Sciences* 1997, **16**(3):297-323.

17. Colgrave ML, Craik DJ: **Thermal, chemical, and enzymatic stability of the cyclotide kalata B1: the importance of the cyclic cystine knot**. *Biochemistry* 2004, **43**(20):5965-5975.

18. Lay F, Anderson M: **Defensins-components of the innate immune system in plants**. *Current Protein and Peptide Science* 2005, **6**(1):85-101.

19. Letunic I, Bork P: **Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees**. *Nucleic Acids Research* 2016:gkw290.

20. Silverstein KA, Moskal WA, Wu HC, Underwood BA, Graham MA, Town CD, VandenBosch KA: **Small cysteine-rich peptides resembling antimicrobial peptides have been under-predicted in plants**. *The Plant Journal* 2007, **51**(2):262-280.

21. Carrasco L, Vazquez D, Hernandezlucas C, Carbonero P, Garciaolmedo F: **Thionins - Plant Peptides That Modify Membrane-Permeability in Cultured Mammalian-Cells**. *European Journal of Biochemistry* 1981, **116**(1):185-189.

22. Kramer KJ, Klassen LW, Jones BL, Speirs RD, Kammer AE: **Toxicity of Purothionin and Its Homologs to the Tobacco Hornworm, Manduca-Sexta (L) (Lepidoptera-Sphingidae)**. *Toxicology and Applied Pharmacology* 1979, **48**(1):179-183.

23. Bloch C, Jr., Richardson M: **A new family of small (5 kDa) protein inhibitors of insect alpha-amylases from seeds or sorghum (Sorghum bicolar (L) Moench) have sequence homologies with wheat gamma-purothionins**. *FEBS letters* 1991, **279**(1):101-104.

24. Goldman MH, Goldberg RB, Mariani C: **Female sterile tobacco plants are produced by stigma-specific cell ablation**. *The EMBO journal* 1994, **13**(13):2976-2984.

25. Pearce G, Moura DS, Stratmann J, Ryan CA, Jr.: **RALF, a 5-kDa ubiquitous polypeptide in plants, arrests root growth and development**. *Proceedings of the National Academy of Sciences of the United States of America* 2001, **98**(22):12843-12847.

26. Pallaghy PK, Norton RS, Nielsen KJ, Craik DJ: **A common structural motif incorporating a cystine knot and a triple-stranded β-sheet in toxic and inhibitory polypeptides**. *Protein Science* 1994, **3**(10):1833-1839.

27. Wieczorek M, Otlewski J, Cook J, Parks K, Leluk J, Wilimowska-Pelc A, Polanowski A, Wilusz T, Laskowski M: **The squash family of serine proteinase inhibitors. Amino acid sequences and association equilibrium constants of inhibitors from squash, summer squash, zucchini, and cucumber seeds**. *Biochemical and Biophysical Research Communications* 1985, **126**(2):646-652.

28. Polanowski A, Wilusz T, Nienartowicz B, Cieślar E, Słomińska A, Nowak K: **Isolation and partial amino acid sequence of the trypsin inhibitor from the seeds of Cucurbita maxima**. *Acta Biochimica Polonica* 1979, **27**(3-4):371-382.

29. Hernandez J-F, Gagnon J, Chiche L, Nguyen TM, Andrieu J-P, Heitz A, Trinh Hong T, Pham TTC, Le Nguyen D: **Squash trypsin inhibitors from Momordica cochinchinensis exhibit an atypical macrocyclic structure**. *Biochemistry* 2000, **39**(19):5722-5730.

30. Habib H, Fazili KM: **Plant protease inhibitors: a defense strategy in plants**. *Biotechnology and Molecular Biology Reviews* 2007, **2**(3):68-85.

31. Hass GM, Nau H, Biemann K, Grahn DT, Ericsson LH, Neurath H: **Amino acid sequence of a carboxypeptidase inhibitor from potatoes**. *Biochemistry* 1975, **14**(6):1334-1342.

32. Hass GM, Hermodson MA: **Amino acid sequence of a carboxypeptidase inhibitor from tomato fruit**. *Biochemistry* 1981, **20**(8):2256-2260.

33. Rees D, Lipscomb W: **Refined crystal structure of the potato inhibitor complex of carboxypeptidase A at 2.5 Å resolution**. *Journal of Molecular Biology* 1982, **160**(3):475-498.

34. Chagolla-Lopez A, Blanco-Labra A, Patthy A, Sánchez R, Pongor S: **A novel alpha-amylase inhibitor from amaranth (Amaranthus hypocondriacus) seeds**. *Journal of Biological Chemistry* 1994, **269**(38):23675-23680.

35. Nguyen PQ, Wang S, Kumar A, Yap LJ, Luu TT, Lescar J, Tam JP: **Discovery and characterization of pseudocyclic cystine-knot α-amylase inhibitors with high resistance to heat and proteolytic degradation**. *FEBS Journal* 2014, **281**(19):4351-4366.

36. Nguyen PQ, Luu TT, Bai Y, Nguyen GK, Pervushin K, Tam JP: **Allotides: Proline-Rich Cystine Knot alpha-Amylase Inhibitors from Allamanda cathartica**. *Journal of Natural Products* 2015, **78**(4):695-704.

37. Martins JC, Enassar M, Willem R, Wieruzeski JM, Lippens G, Wodak SJ: **Solution structure of the main α-amylase inhibitor from amaranth seeds**. *European Journal of Biochemistry* 2001, **268**(8):2379-2389.

38. Craik DJ, Malik U: **Cyclotide biosynthesis**. *Current Opinion in Chemical Biology* 2013, **17**(4):546-554.

39. Gran L: **On the effect of a polypeptide isolated from "Kalata-Kalata"(Oldenlandia affinis DC) on the oestrogen dominated uterus**. *Acta pharmacologica et toxicologica* 1973, **33**(5-6):400-408.

40. Craik DJ, Daly NL: **NMR as a tool for elucidating the structures of circular and knotted proteins**. *Molecular BioSystems* 2007, **3**(4):257-265.

41. Rosengren KJ, Daly NL, Plan MR, Waine C, Craik DJ: **Twists, knots, and rings in proteins structural definition of the cyclotide framework**. *Journal of Biological Chemistry* 2003, **278**(10):8606-8616.

42. Nguyen GKT, Lian Y, Pang EWH, Nguyen PQT, Tran TD, Tam JP: **Discovery of linear cyclotides in monocot plant Panicum laxum of Poaceae family provides new insights into evolution and distribution of cyclotides in plants**. *Journal of Biological Chemistry* 2013, **288**(5):3370-3380.

43. Nguyen GKT, Zhang S, Nguyen NTK, Nguyen PQT, Chiu MS, Hardjojo A, Tam JP: **Discovery and characterization of novel cyclotides originated from chimeric precursors consisting of albumin-1 chain a and cyclotide domains in the Fabaceae family**. *Journal of Biological Chemistry* 2011, **286**(27):24275-24287.

44. Craik DJ, Mylne JS, Daly NL: **Cyclotides: macrocyclic peptides with applications in drug design and agriculture**. *Cellular and Molecular Life Sciences* 2010, **67**(1):9-16.

45. Van Parijs J, Broekaert WF, Goldstein IJ, Peumans WJ: **Hevein: an antifungal protein from rubber-tree (Hevea brasiliensis) latex**. *Planta* 1991, **183**(2):258-264.

46. Andreev YA, Korostyleva TV, Slavokhotova AA, Rogozhin EA, Utkina LL, Vassilevski AA, Grishin EV, Egorov TA, Odintsova TI: **Genes encoding hevein-like defense peptides in wheat: distribution, evolution, and role in stress response**. *Biochimie* 2012, **94**(4):1009-1016.

47. Shukurov RR, Voblikova VD, Nikonorova AK, Komakhin RA, Komakhina VV, Egorov TA, Grishin EV, Babakov AV: **Transformation of tobacco and Arabidopsis plants with Stellaria media genes encoding novel hevein-like peptides increases their resistance to fungal pathogens**. *Transgenic Research* 2012, **21**(2):313-325.

48. Tam JP, Wang S, Wong KH, Tan WL: **Antimicrobial Peptides from Plants**. *Pharmaceuticals (Basel)* 2015, **8**(4):711-757.

49. Andersen NH, Cao B, Rodriguez-Romero A, Arreguin B: **Hevein: NMR assignment and assessment of solution-state folding for the agglutinin-toxin motif**. *Biochemistry* 1993, **32**(6):1407-1422.

50. Dubovskii PV, Vassilevski AA, Slavokhotova AA, Odintsova TI, Grishin EV, Egorov TA, Arseniev AS: **Solution structure of a defense peptide from wheat with a 10-cysteine motif**. *Biochemical and Biophysical Research Communications* 2011, **411**(1):14-18.

51. Broekaert WF, Marien W, Terras FR, De Bolle MF, Proost P, Van Damme J, Dillen L, Claeys M, Rees SB: **Antimicrobial peptides from Amaranthus caudatus seeds with sequence homology to the cysteine/glycine-rich domain of chitin-binding proteins**. *Biochemistry* 1992, **31**(17):4308-4314.

52. de Souza Cândido E, Sousa DA, Viana JC, de Oliveira-Júnior NG, Miranda V, Franco OL: **The use of versatile plant antimicrobial peptides in agribusiness and human health**. *Peptides* 2014, **55**:65-78.

53. Thomma BP, Cammue BP, Thevissen K: **Plant defensins**. *Planta* 2002, **216**(2):193-202.

54. Stotz HU, Thomson J, Wang Y: **Plant defensins: defense, development and application**. *Plant Signaling and Behavior* 2009, **4**(11):1010-1012.

55. Broekaert WF, Terras F, Cammue B, Osborn RW: **Plant defensins: novel antimicrobial peptides as components of the host defense system**. *Plant Physiology* 1995, **108**(4):1353.

56. Lay FT, Schirra HJ, Scanlon MJ, Anderson MA, Craik DJ: **The three-dimensional solution structure of NaD1, a new floral defensin from Nicotiana alata and its application to a homology model of the crop defense protein alfAFP**. *Journal of Molecular Biology* 2003, **325**(1):175-188.

57. Janssen BJ, Schirra HJ, Lay FT, Anderson MA, Craik DJ: **Structure of Petunia hybrida defensin 1, a novel plant defensin with five disulfide bonds**. *Biochemistry* 2003, **42**(27):8214-8222.

58. de Oliveira Carvalho A, Gomes VM: **Plant defensins—prospects for the biological functions and biotechnological properties**. *Peptides* 2009, **30**(5):1007-1020.

59. Stotz HU, Spence B, Wang Y: **A defensin from tomato with dual function in defense and development**. *Plant Molecular Biology* 2009, **71**(1-2):131-143.

60. Zhu YJ, Agbayani R, Moore PH: **Ectopic expression of Dahlia merckii defensin DmAMP1 improves papaya resistance to Phytophthora palmivora by reducing pathogen vigor**. *Planta* 2007, **226**(1):87-97.

61. De Caleya RF, Gonzalez-Pascual B, García-Olmedo F, Carbonero P: **Susceptibility of phytopathogenic bacteria to wheat purothionins in vitro**. *Applied Microbiology* 1972, **23**(5):998-1000.

62. Stec B: **Plant thionins–the structural perspective**. *Cellular and Molecular Life Sciences* 2006, **63**(12):1370-1385.

63. Bohlmann H, Apel K: **Thionins**. *Annual Review of Plant Biology* 1991, **42**(1):227-240.

64. Van Loon LC, Rep M, Pieterse C: **Significance of inducible defense-related proteins in infected plants**. *Annual Review of Phytopathology* 2006, **44**:135-162.

65. Epple P, Apel K, Bohlmann H: **An Arabidopsis thaliana thionin gene is inducible via a signal transduction pathway different from that for pathogenesis-related proteins**. *Plant Physiology* 1995, **109**(3):813-820.

66. Bohlmann H, Clausen S, Behnke S, Giese H, Hiller C, Reimann-Philipp U, Schrader G, Barkholt V, Apel K: **Leaf-specific thionins of barley—a novel class of cell wall proteins toxic to plant-pathogenic fungi and possibly involved in the defence mechanism of plants**. *The EMBO Journal* 1988, **7**(6):1559.

67. Romero A, Alamillo JM, García-Olmedo F: **Processing of thionin precursors in barley leaves by a vacuolar proteinase**. *European Journal of Biochemistry* 1997, **243**(1-2):202-208.

68. Stevens C, Titarenko E, Hargreaves JA, Gurr SJ: **Defence-related gene activation during an incompatible interaction between Stagonospora (Septoria) nodorum and barley (Hordeum vulgare L.) coleoptile cells**. *Plant Molecular Biology* 1996, **31**(4):741-749.

69. Schrader-Fischer G, Apel K: **Organ-specific expression of highly divergent thionin variants that are distinct from the seed-specific crambin in the crucifer Crambe abyssinica**. *Molecular and General Genetics* 1994, **245**(3):380-389.

70. Ponz F, Paz-Ares J, Hernández-Lucas C, Carbonero P, García-Olmedo F: **Synthesis and processing of thionin precursors in developing endosperm from barley (Hordeum vulgare L.)**. *The EMBO Journal* 1983, **2**(7):1035.

71. Oka T, Murata Y, Nakanishi T, Yoshizumi H, Hayashida H, Ohtsuki Y, Toyoshima K, Hakura A: **Similarity, in molecular structure and function, between the plant toxin purothionin and the mammalian pore-forming proteins**. *Molecular Biology and Evolution* 1992, **9**(4):707-715.

72. Stec B, Markman O, Rao U, Heffron G, Henderson S, Vernon L, Brumfeld V, Teeter M: **Proposal for molecular mechanism of thionins deduced**

from physico-chemical studies of plant toxins. *The Journal of Peptide Research* 2004, **64**(6):210-224.

73. Tuteja R: **Type I signal peptidase: an overview**. *Archives of Biochemistry and Biophysics* 2005, **441**(2):107-111.

74. Von Heijne G: **Signal sequences: the limits of variation**. *Journal of Molecular Biology* 1985, **184**(1):99-105.

75. Gruber CW, Čemažar M, Clark RJ, Horibe T, Renda RF, Anderson MA, Craik DJ: **A novel plant protein-disulfide isomerase involved in the oxidative folding of cystine knot defense proteins**. *Journal of Biological Chemistry* 2007, **282**(28):20435-20446.

76. Hebert DN, Molinari M: **In and out of the ER: protein folding, quality control, degradation, and related human diseases**. *Physiological Reviews* 2007, **87**(4):1377-1408.

77. Stevens FJ, Argon Y: **Protein folding in the ER**. *Seminars in Cell and Developmental Biology* 1999, **10**(5):443-454.

78. Jennings C, West J, Waine C, Craik D, Anderson M: **Biosynthesis and insecticidal properties of plant cyclotides: the cyclic knotted proteins from Oldenlandia affinis**. *Proceedings of the National Academy of Sciences of the United States of America* 2001, **98**(19):10614-10619.

79. Saska I, Gillon AD, Hatsugai N, Dietzgen RG, Hara-Nishimura I, Anderson MA, Craik DJ: **An asparaginyl endopeptidase mediates in vivo protein backbone cyclization**. *Journal of Biological Chemistry* 2007, **282**(40):29721-29728.

80. Finkina EI, Shramova EI, Tagaev AA, Ovchinnikova TV: **A novel defensin from the lentil Lens culinaris seeds**. *Biochemical and Biophysical Research Communications* 2008, **371**(4):860-865.

81. Lay FT, Brugliera F, Anderson MA: **Isolation and properties of floral defensins from ornamental tobacco and petunia**. *Plant Physiology* 2003, **131**(3):1283-1293.

82. Kini SG, Nguyen PQ, Weissbach S, Mallagaray A, Shin J, Yoon HS, Tam JP: **Studies on the Chitin Binding Property of Novel Cysteine-Rich Peptides from Alternanthera sessilis**. *Biochemistry* 2015, **54**(43):6639-6649.

83. Montesinos E: **Antimicrobial peptides and plant disease control**. *FEMS Microbiology Letters* 2007, **270**(1):1-11.

84. Barbosa Pelegrini P, del Sarto RP, Silva ON, Franco OL, Grossi-de-Sa MF: **Antibacterial peptides from plants: what they are and how they probably work**. *Biochemistry Research International* 2011, **2011**.

85. Cammue B, De Bolle M, Terras F, Proost P, Van Damme J, Rees SB, Vanderleyden J, Broekaert WF: **Isolation and characterization of a novel class of plant antimicrobial peptides form Mirabilis jalapa L. seeds**. *Journal of Biological Chemistry* 1992, **267**(4):2228-2233.

86. Tam JP, Lu Y-A, Yang J-L, Chiu K-W: **An unusual structural motif of antimicrobial peptides containing end-to-end macrocycle and cystine-knot disulfides**. *Proceedings of the National Academy of Sciences* 1999, **96**(16):8913-8918.

87. Marcus JP, Green JL, Goulter KC, Manners JM: **A family of antimicrobial peptides is produced by processing of a 7S globulin protein in Macadamia integrifolia kernels**. *The Plant Journal* 1999, **19**(6):699-710.

88. Nawrot R, Barylski J, Nowicki G, Broniarczyk J, Buchwald W, Goździcka-Józefiak A: **Plant antimicrobial peptides**. *Folia Microbiologica* 2014, **59**(3):181-196.

89. Ireland DC, Wang CKL, Wilson JA, Gustafson KR, Craik DJ: **Cyclotides as natural anti-HIV agents**. *Biopolymers* 2008, **90**(1):51-60.

90. Ng Y-M, Yang Y, Sze K-H, Zhang X, Zheng Y-T, Shaw P-C: **Structural characterization and anti-HIV-1 activities of arginine/glutamate-rich polypeptide Luffin P1 from the seeds of sponge gourd (Luffa cylindrica)**. *Journal of Structural Biology* 2011, **174**(1):164-172.

91. Nguyen QT, Ooi JSG, Nguyen NTK, Wang SJ, Huang M, Liu DX, Tam JP: **Antiviral Cystine Knot a-Amylase Inhibitors from Alstonia scholaris**. *Journal of Biological Chemistry* 2015, **290**(52):31138-31150.

92. Luckett S, Garcia RS, Barker JJ, Konarev AV, Shewry PR, Clarke AR, Brady RL: **High-resolution structure of a potent, cyclic proteinase inhibitor from sunflower seeds**. *Journal of Molecular Biology* 1999, **290**(2):525-533.

93. Chen P, Rose J, Love R, Wei C, Wang B: **Reactive sites of an anticarcinogenic Bowman-Birk proteinase inhibitor are similar to other trypsin inhibitors**. *Journal of Biological Chemistry* 1992, **267**(3):1990-1994.

94. Korsinczky ML, Schirra HJ, Rosengren KJ, West J, Condie BA, Otvos L, Anderson MA, Craik DJ: **Solution structures by 1 H NMR of the novel cyclic trypsin inhibitor SFTI-1 from sunflower seeds and an acyclic permutant**. *Journal of Molecular Biology* 2001, **311**(3):579-591.

95. Jouvensal L, Quillien L, Ferrasson E, Rahbé Y, Guéguen J, Vovelle F: **PA1b, an insecticidal protein extracted from pea seeds (Pisum sativum): 1H-2-D NMR study and molecular modeling**. *Biochemistry* 2003, **42**(41):11915-11923.

96. Louis S, Delobel B, Gressent F, Rahioui I, Quillien L, Vallier A, Rahbé Y: **Molecular and biological screening for insect-toxic seed albumins from four legume species**. *Plant Science* 2004, **167**(4):705-714.

97. Delobel B, Grenier A, Gueguen J, Ferrasson E, Mbaiguinam M: **Utilisation d'un polypeptide dérivé d'une albumine PA1b de légumineuse comme insecticide**. *French Patent* 1998, **98**:05877.

98. Gressent F, Rahioui I, Rahbe Y: **Characterization of a high-affinity binding site for the pea albumin 1b entomotoxin in the weevil Sitophilus**. *European Journal of Biochemistry* 2003, **270**(11):2429-2435.

99. Grain L: **Isolation of oxytocic peptides from Oldenlandia affinis by solvent extraction of tetraphenylborate complexes and chromatography on sephadex LH-20**. *Lloydia* 1973, **36**(2):207-208.

100. Tam JP, Lu YA, Yang JL, Chiu KW: **An unusual structural motif of antimicrobial peptides containing end-to-end macrocycle and cystine-knot disulfides**. *Proceedings of the National Academy of Sciences of the United States of America* 1999, **96**(16):8913-8918.

101. Schöpke T, Hasan Agha M, Kraft R, Otto A, Hiller K: **Hämolytisch aktive komponenten aus Viola tricolor L. und Viola arvensis Murray**. *Scientia Pharmaceutica* 1993, **61**:145-153.

102. Witherup KM, Bogusky MJ, Anderson PS, Ramjit H, Ransom RW, Wood T, Sardana M: **Cyclopsychotride A, a biologically active, 31-residue cyclic peptide isolated from Psychotria longipes**. *Journal of Natural Products* 1994, **57**(12):1619-1625.

103. Hara S, Makino J, Ikenaka T: **Amino acid sequences and disulfide bridges of serine proteinase inhibitors from bitter gourd (Momordica charantia LINN.) seeds**. *Journal of Biochemistry* 1989, **105**(1):88-92.

104. Ryan CA, Hass GM, Kuhn RW: **Purification and properties of a carboxypeptidase inhibitor from potatoes**. *Journal of Biological Chemistry* 1974, **249**(17):5495-5499.

105. Gao A-G, Hakimi SM, Mittanck CA, Wu Y, Woerner BM, Stark DM, Shah DM, Liang J, Rommens CM: **Fungal pathogen protection in potato by expression of a plant defensin peptide**. *Nature Biotechnology* 2000, **18**(12):1307-1310.

106. Franco OL, Murad AM, Leite JR, Mendes PA, Prates MV, Bloch C: **Identification of a cowpea γ-thionin with bactericidal activity**. *FEBS Journal* 2006, **273**(15):3489-3497.

107. Melo FR, Rigden DJ, Franco OL, Mello LV, Ary MB, Grossi de Sá MF, Bloch C: **Inhibition of trypsin by cowpea thionin: characterization, molecular modeling, and docking**. *Proteins: Structure, Function, and Bioinformatics* 2002, **48**(2):311-319.

108. Bloch C, Richardson M: **A new family of small (5 kDa) protein inhibitors of insect α-amylases from seeds or sorghum (Sorghum bicolor (L) Moench) have sequence homologies with wheat γ-purothionins**. *FEBS Letters* 1991, **279**(1):101-104.

109. Kushmerick C, de Souza Castro M, Santos Cruz J, Bloch C, Beirão PS: **Functional and structural features of γ-zeathionins, a new class of sodium channel blockers**. *FEBS Letters* 1998, **440**(3):302-306.

110. Fujimura M, Ideguchi M, Minami Y, Watanabe K, Tadera K: **Purification, characterization, and sequencing of novel antimicrobial peptides, Tu-AMP 1 and Tu-AMP 2, from bulbs of tulip (Tulipa gesneriana L.)**. *Bioscience, Biotechnology, and Biochemistry* 2004, **68**(3):571-577.

111. Clark RJ, Fischer H, Dempster L, Daly NL, Rosengren KJ, Nevin ST, Meunier FA, Adams DJ, Craik DJ: **Engineering stable peptide toxins by means of backbone cyclization: Stabilization of the α-conotoxin MII**. *Proceedings of the National Academy of Sciences of the United States of America* 2005, **102**(39):13767-13772.

112. Schafmeister CE, Po J, Verdine GL: **An all-hydrocarbon cross-linking system for enhancing the helicity and metabolic stability of peptides**. *Journal of the American Chemical Society* 2000, **122**(24):5891-5892.

113. Taylor EM, Otero DA, Banks WA, O'Brien JS: **Retro-inverso prosaptide peptides retain bioactivity, are stable In vivo, and are blood-brain barrier permeable**. *Journal of Pharmacology and Experimental Therapeutics* 2000, **295**(1):190-194.

114. Ward P, Ewan GB, Jordan CC, Ireland SJ, Hagan RM, Brown JR: **Potent and highly selective neurokinin antagonists**. *Journal of Medicinal Chemistry* 1990, **33**(7):1848-1851.

115. Clark RJ, Jensen J, Nevin ST, Callaghan BP, Adams DJ, Craik DJ: **The Engineering of an Orally Active Conotoxin for the Treatment of Neuropathic Pain**. *Angewandte Chemie International Edition* 2010, **49**(37):6545-6548.

116. Gunasekera S, Foley FM, Clark RJ, Sando L, Fabri LJ, Craik DJ, Daly NL: **Engineering stabilized vascular endothelial growth factor-A antagonists: synthesis, structural characterization, and bioactivity of grafted analogues of cyclotides**. *Journal of Medicinal Chemistry* 2008, **51**(24):7697-7704.

117. Chan LY, Gunasekera S, Henriques ST, Worth NF, Le S-J, Clark RJ, Campbell JH, Craik DJ, Daly NL: **Engineering pro-angiogenic peptides using stable, disulfide-rich cyclic scaffolds**. *Blood* 2011, **118**(25):6709-6717.

118. Wong CT, Rowlands DK, Wong CH, Lo TW, Nguyen GK, Li HY, Tam JP: **Orally active peptidic bradykinin B1 receptor antagonists engineered from a cyclotide scaffold for inflammatory pain treatment**. *Angewandte Chemie* 2012, **51**(23):5620-5624.

119. Vita C, Roumestand C, Toma F, Ménez A: **Scorpion toxins as natural scaffolds for protein engineering**. *Proc Natl Acad Sci U S A* 1995, **92**(14):6404-6408.

120. Li C, Liu M, Monbo J, Zou GZ, Li CQ, Yuan WR, Zella D, Lu WY, Lu WY: **Turning a scorpion toxin into an antitumor miniprotein**. *Journal of the American Chemical Society* 2008, **130**(41):13546-+.

121. Wang CK, Gruber CW, Cemazar M, Siatskas C, Tagore P, Payne N, Sun G, Wang S, Bernard CC, Craik DJ: **Molecular grafting onto a stable**

framework yields novel cyclic peptides for the treatment of multiple sclerosis. *ACS Chemical Biology* 2014, **9**(1):156-163.

122. Eliasen R, Daly NL, Wulff BS, Andresen TL, Conde-Frieboes KW, Craik DJ: **Design, synthesis, structural and functional characterization of novel melanocortin agonists based on the cyclotide kalata B1**. *Journal of Biological Chemistry* 2012, **287**(48):40493-40501.

123. Aboye TL, Ha H, Majumder S, Christ F, Debyser Z, Shekhtman A, Neamati N, Camarero JA: **Design of a Novel Cyclotide-Based CXCR4 Antagonist with Anti-Human Immunodeficiency Virus (HIV)-1 Activity**. *Journal of Medicinal Chemistry* 2012, **55**(23):10729-10734.

124. Ji Y, Majumder S, Millard M, Borra R, Bi T, Elnagar AY, Neamati N, Shekhtman A, Camarero JA: **In Vivo Activation of the p53 Tumor Suppressor Pathway by an Engineered Cyclotide**. *Journal of the American Chemical Society* 2013, **135**(31):11623-11633.

125. Thongyoo P, Roqué-Rosell N, Leatherbarrow RJ, Tate EW: **Chemical and biomimetic total syntheses of natural and engineered MCoTI cyclotides**. *Organic and Biomolecular Chemistry* 2008, **6**(8):1462-1470.

126. Huang Y-H, Henriques ST, Wang CK, Thorstholm L, Daly NL, Kaas Q, Craik DJ: **Design of substrate-based BCR-ABL kinase inhibitors using the cyclotide scaffold**. *Scientific Reports* 2015, **5**.

127. Conibear AC, Bochen A, Rosengren KJ, Stupar P, Wang C, Kessler H, Craik DJ: **The Cyclic Cystine Ladder of Theta-Defensins as a Stable, Bifunctional Scaffold: A Proof-of-Concept Study Using the Integrin-Binding RGD Motif**. *ChemBioChem* 2014, **15**(3):451-459.

128. Rose TM, Schultz ER, Henikoff JG, Pietrokovski S, McCallum CM, Henikoff S: **Consensus-degenerate hybrid oligonucleotide primers for amplification of distantly related sequences**. *Nucleic Acids Research* 1998, **26**(7):1628-1635.

129. Rose TM, Henikoff JG, Henikoff S: **CODEHOP (COnsensus-DEgenerate hybrid oligonucleotide primer) PCR primer design**. *Nucleic Acids Research* 2003, **31**(13):3763-3766.

130. Boyce R, Chilana P, Rose TM: **iCODEHOP: a new interactive program for designing COnsensus-DEgenerate Hybrid Oligonucleotide Primers from multiply aligned protein sequences**. *Nucleic Acids Research* 2009:gkp379.

131. Petersen TN, Brunak S, von Heijne G, Nielsen H: **SignalP 4.0: discriminating signal peptides from transmembrane regions**. *Nature Methods* 2011, **8**(10):785-786.

132. Han X, He L, Xin L, Shan B, Ma B: **PeaksPTM: Mass spectrometry-based identification of peptides with unspecified modifications**. *Journal of Proteome Research* 2011, **10**(7):2930-2936.

133. Aida Serra XH, Nguyen GK, Nguyen NT, Sze SK, Tam JP: **A high-throughput peptidomic strategy to decipher the molecular diversity of cyclic cysteine-rich peptides**. *Scientific Reports* 2016, **6**.

134. Creasy DM, Cottrell JS: **Unimod: Protein modifications for mass spectrometry**. *Proteomics* 2004, **4**(6):1534-1536.

135. Han Y, Ma B, Zhang K: **SPIDER: software for protein identification from sequence tags with de novo sequencing error**. *Journal of Bioinformatics and Computational Biology* 2005, **3**(03):697-716.

136. Jeener J, Meier B, Bachmann P, Ernst R: **Investigation of exchange processes by two-dimensional NMR spectroscopy**. *The Journal of Chemical Physics* 1979, **71**(11):4546-4553.

137. Kumar A, Ernst R, Wüthrich K: **A two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete proton-proton cross-relaxation networks in biological macromolecules**. *Biochemical and Biophysical Research Communications* 1980, **95**(1):1-6.

138. Davis DG, Bax A: **Assignment of complex proton NMR spectra via two-dimensional homonuclear Hartmann-Hahn spectroscopy**. *Journal of the American Chemical Society* 1985, **107**(9):2820-2821.

139. Bax A, Davis DG: **MLEV-17-based two-dimensional homonuclear magnetization transfer spectroscopy**. *Journal of Magnetic Resonance (1969)* 1985, **65**(2):355-360.

140. Rance M, Sørensen O, Bodenhausen G, Wagner G, Ernst R, Wüthrich K: **Improved spectral resolution in COSY 1 H NMR spectra of proteins via double quantum filtering**. *Biochemical and Biophysical Research Communications* 1983, **117**(2):479-485.

141. Wüthrich K, Billeter M, Braun W: **Pseudo-structures for the 20 common amino acids for use in studies of protein conformations by measurements of intramolecular proton-proton distance constraints with nuclear magnetic resonance**. *Journal of Molecular Biology* 1983, **169**(4):949-961.

142. Liu M, Mao X-a, Ye C, Huang H, Nicholson JK, Lindon JC: **Improved WATERGATE pulse sequences for solvent suppression in NMR spectroscopy**. *Journal of Magnetic Resonance* 1998, **132**(1):125-129.

143. Hwang T-L, Shaka A: **Water suppression that works. Excitation sculpting using arbitrary wave-forms and pulsed-field gradients**. *Journal of Magnetic Resonance, Series A* 1995, **112**(2):275-279.

144. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A: **NMRPipe: a multidimensional spectral processing system based on UNIX pipes**. *Journal of biomolecular NMR* 1995, **6**(3):277-293.

145. Brünger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang J-S, Kuszewski J, Nilges M, Pannu NS: **Crystallography & NMR system: A new software suite for macromolecular structure determination**. *Acta Crystallographica Section D: Biological Crystallography* 1998, **54**(5):905-921.

146. Pardi A, Billeter M, Wüthrich K: **Calibration of the angular dependence of the amide proton-C α proton coupling constants, 3 J HNα, in a globular protein: Use of 3JHNα for identification of helical secondary structure**. *Journal of Molecular Biology* 1984, **180**(3):741-751.

147. Driscoll PC, Gronenborn AM, Beress L, Clore GM: **Determination of the three-dimensional solution structure of the antihypertensive and antiviral protein BDS-I from the sea anemone Anemonia sulcata: a study using nuclear magnetic resonance and hybrid distance geometry-dynamical simulated annealing**. *Biochemistry* 1989, **28**(5):2188-2198.

148. Nilges M, Gronenborn AM, Brünger AT, Clore GM: **Determination of three-dimensional structures of proteins by simulated annealing with interproton distance restraints. Application to crambin, potato carboxypeptidase inhibitor and barley serine proteinase inhibitor 2**. *Protein Engineering* 1988, **2**(1):27-38.

149. Nilges M, Clore G, Gronenborn A: **Determination of the three-dimensional structure of proteins from interproton distance data by**

hybrid distance geometry-dynamical simulated annealing calculations. *FEBS Letters* 1988, **229**:317-324.

150. Nilges M: **A calculation strategy for the structure determination of symmetric demers by 1H NMR**. *Proteins: Structure, Function, and Bioinformatics* 1993, **17**(3):297-309.

151. Koradi R, Billeter M, Wüthrich K: **MOLMOL: a program for display and analysis of macromolecular structures**. *Journal of Molecular Graphics* 1996, **14**(1):51-55.

152. DeLano WL: **The PyMOL molecular graphics system Version 1.7.4 Schrödinger, LLC.** 2002.

153. Laskowski RA, MacArthur MW, Moss DS, Thornton JM: **PROCHECK: a program to check the stereochemical quality of protein structures**. *Journal of Applied Crystallography* 1993, **26**(2):283-291.

154. Lehrer RI, Rosenman M, Harwig SSSL, Jackson R, Eisenhauer P: **Ultrasensitive assays for endogenous antimicrobial polypeptides**. *Journal of Immunological Methods* 1991, **137**(2):167-173.

155. Tam JP, Lu Y-A, Yang J-L: **Antimicrobial dendrimeric peptides**. *European Journal of Biochemistry* 2002, **269**(3):923-932.

156. Park HJ, Cho DH, Kim HJ, Lee JY, Cho BK, Bang SI, Song SY, Yamasaki K, Di Nardo A, Gallo RL: **Collagen synthesis is suppressed in dermal fibroblasts by the human antimicrobial peptide LL-37**. *Journal of Investive Dermatology* 2009, **129**(4):843-850.

157. Schneider CA, Rasband WS, Eliceiri KW: **NIH Image to ImageJ: 25 years of image analysis**. *Nature Methods* 2012, **9**(7):671-675.

158. Geback T, Schulz MM, Koumoutsakos P, Detmar M: **TScratch: a novel and simple software tool for automated analysis of monolayer wound healing assays**. *Biotechniques* 2009, **46**(4):265-274.

159. Ding AH, Nathan CF, Stuehr DJ: **Release of reactive nitrogen intermediates and reactive oxygen intermediates from mouse peritoneal macrophages. Comparison of activating cytokines and evidence for independent production**. *The Journal of Immunology* 1988, **141**(7):2407-2412.

160. Hammami R, Ben Hamida J, Vergoten G, Fliss I: **PhytAMP: a database dedicated to antimicrobial plant peptides**. *Nucleic Acids Research* 2009, **37**(Database issue):D963-968.

161. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li WZ, Lopez R, McWilliam H, Remmert M, Soding J *et al*: **Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega**. *Molecular Systems Biology* 2011, **7**.

162. **The wealth of India, Raw material**. **vol 5: H-K**:289-290.

163. Kiritikar KRaB, B. D. : **Indian Medicinal Plants with illustrations. 2nd Edition**. *Oriental Enterprises* 2003, **Vol 7**:2093-2096.

164. Ito Y, Sugimoto A, Kakuda T, Kubota K: **Identification of potent odorants in Chinese jasmine green tea scented with flowers of Jasminum sambac**. *Journal of Agricultural and Food Chemistry* 2002, **50**(17):4878-4884.

165. Lim TK: **Jasminum sambac**. In: *Edible Medicinal and Non Medicinal Plants.* vol. 8: Springer Netherlands; 2014: 529-540.

166. Bhangale J, Patel R, Acharya S, Chaudhari K: **Preliminary studies on anti-inflammatory and analgesic activities of Jasminum sambac (L.) Aiton in experimental animal models**. *American Journal of PharmTech Research* 2012, **2**(4):804-812.

167. Bhagat A, Khainar A, Tenpe C, Upaganlawar A, Yeole P: **Anti-inflammatory activity of Jasminum sambac leaf extracts against carageenen induced rat paw edema**. *Indian Journal of Natural Product* 2007, **23**(3):25-28.

168. Abdoul-Latif F, Edou P, Eba F, Mohamed N, Ali A, Djama S, Obame L-C, DICKO MH: **Antimicrobial and antioxidant activities of essential oil and methanol extract of Jasminum sambac from Djibouti**. *African Journal of Plant Science* 2010, **4**(3):038-043.

169. Kalaiselvi M, Narmatha R, Ragavendran P, Ravikumar G, Sophia D, Gomathi D, Arulraj C, Kumar DG, Uma C, Kalaivani K: **In vitro free radical scavenging activity of Jasminum sambac (L.) Ait oleaceae flower**. *Asian Journal of Pharmaceutical & Biological Research (AJPBR)* 2011, **1**(3).

170. Talib WH, Mahasneh AM: **Antiproliferative activity of plant extracts used against cancer in traditional medicine**. *Scientia Pharmaceutica* 2010, **78**(1):33.

171. Rahman MA, Hasan MS, Hossain MA, Biswas N: **Analgesic and cytotoxic activities of Jasminum sambac (L.) Aiton**. *Pharmacologyonline* 2011, **1**:124-131.

172. Chiang L-C, Cheng H-Y, Liu M-C, Chiang W, Lin C-C: **In vitro anti-herpes simplex viruses and anti-adenoviruses activity of twelve traditionally used medicinal plants in Taiwan**. *Biological and Pharmaceutical Bulletin* 2003, **26**(11):1600-1604.

173. Ragasa CY, Tamboong B, Rideout J: **Secondary metabolites from Jasminum sambac and Cananga odorata**. 2003.

174. Rath C, Devi S, Dash S, Mishra R: **Antibacterial potential assessment of jasmine essential oil against E. coli**. *Indian Journal of Pharmaceutical Sciences* 2008, **70**(2):238.

175. Joy P, Raja DP: **Anti-bacterial activity studies of Jasminum grandiflorum and Jasminum sambac**. *Ethnobotanical Leaflets* 2008, **2008**(1):59.

176. Hongratanaworakit T: **Stimulating effect of aromatherapy massage with jasmine oil**. *Natural Product Communications* 2010, **5**(1):157-162.

177. Alrashdi AS, Salama SM, Alkiyumi SS, Abdulla MA, Hadi AH, Abdelwahab SI, Taha MM, Hussiani J, Asykin N: **Mechanisms of Gastroprotective Effects of Ethanolic Leaf Extract of Jasminum sambac against HCl/Ethanol-Induced Gastric Mucosal Injury in Rats**. *Evidence Based Complementary Alternative Medicine* 2012, **2012**:786426.

178. Sabharwal S, Aggarwal S, Vats M, Sardana S: **Preliminary phytochemical investigation and wound healing activity of Jasminum sambac (Linn) Ait.(Oleaceae) leaves**. *International Journal of Pharmacognosy and Phytochemical Research* 2012, **4**(3):146-150.

179. Kunhachan P, Banchonglikitkul C, Kajsongkram T, Khayungarnnawee A, Leelamanit W: **Chemical composition, toxicity and vasodilatation effect of the flowers extract of Jasminum sambac (L.) Ait.“G. Duke of Tuscany”**. *Evidence-Based Complementary and Alternative Medicine* 2012, **2012**.

180. Kuroda K, Inoue N, Ito Y, Kubota K, Sugimoto A, Kakuda T, Fushiki T: **Sedative effects of the jasmine tea odor and (R)-(−)-linalool, one of its major odor components, on autonomic nerve activity and mood states**. *European Journal of Applied Physiology* 2005, **95**(2-3):107-114.

181. Zhang Y-J, Liu Y-Q, Pu X-Y, Yang C-R: **Iridoidal glycosides from Jasminum sambac**. *Phytochemistry* 1995, **38**(4):899-903.

182. Tanahashi T, Nagakura N, Inoue K, Inouye H: **Sambacosides A, E and F, novel tetrameric iridoid glucosides from Jasminum sambac**. *Tetrahedron Letters* 1988, **29**(15):1793-1796.

183. Gruber CW, Elliott AG, Ireland DC, Delprete PG, Dessein S, Göransson U, Trabi M, Wang CK, Kinghorn AB, Robbrecht E: **Distribution and evolution of circular miniproteins in flowering plants**. *The Plant Cell Online* 2008, **20**(9):2471-2483.

184. Trabi M, Svangård E, Herrmann A, Göransson U, Claeson P, Craik DJ, Bohlin L: **Variations in cyclotide expression in viola species**. *Journal of Natural Products* 2004, **67**(5):806-810.

185. Trabi M, Craik DJ: **Tissue-specific expression of head-to-tail cyclized miniproteins in Violaceae and structure determination of the root cyclotide Viola hederacea root cyclotide1**. *The Plant Cell* 2004, **16**(8):2204-2216.

186. Creasy DM, Cottrell JS: **Unimod: Protein modifications for mass spectrometry**. *Proteomics* 2004, **4**(6):1534-1536.

187. Dutton JL, Renda RF, Waine C, Clark RJ, Daly NL, Jennings CV, Anderson MA, Craik DJ: **Conserved structural and sequence elements implicated in the processing of gene-encoded circular proteins**. *Journal of Biological Chemistry* 2004, **279**(45):46858-46867.

188. Nguyen GKT, Zhang S, Wang W, Wong CTT, Nguyen NTK, Tam JP: **Discovery of a linear cyclotide from the bracelet subfamily and its disulfide mapping by top-down mass spectrometry**. *Journal of Biological Chemistry* 2011, **286**(52):44833-44844.

189. Crooks GE, Hon G, Chandonia JM, Brenner SE: **WebLogo: A sequence logo generator**. *Genome Research* 2004, **14**(6):1188-1190.

190. Dutertre S, Jin A-h, Kaas Q, Jones A, Alewood PF, Lewis RJ: **Deep venomics reveals the mechanism for expanded peptide diversity in cone snail venom**. *Molecular & Cellular Proteomics* 2013, **12**(2):312-329.

191. Mergaert P, Nikovics K, Kelemen Z, Maunoury N, Vaubert D, Kondorosi A, Kondorosi E: **A novel family in Medicago truncatula consisting of more than 300 nodule-specific genes coding for small, secreted polypeptides with conserved cysteine motifs**. *Plant Physiology* 2003, **132**(1):161-173.

192. Nguyen GK, Wang S, Qiu Y, Hemu X, Lian Y, Tam JP: **Butelase 1 is an Asx-specific ligase enabling peptide macrocyclization and synthesis**. *Nature Chemical Biology* 2014, **10**(9):732-738.

193. Mylne JS, Chan LY, Chanson AH, Daly NL, Schaefer H, Bailey TL, Nguyencong P, Cascales L, Craik DJ: **Cyclic peptides arising by evolutionary parallelism via asparaginyl-endopeptidase–mediated biosynthesis**. *The Plant Cell Online* 2012, **24**(7):2765-2778.

194. Poth AG, Mylne JS, Grassl J, Lyons RE, Millar AH, Colgrave ML, Craik DJ: **Cyclotides associate with leaf vasculature and are the products of a novel precursor in Petunia (Solanaceae)**. *Journal of Biological Chemistry* 2012, **287**(32):27033-27046.

195. Tailor RH, Acland DP, Attenborough S, Cammue BP, Evans IJ, Osborn RW, Ray JA, Rees SB, Broekaert WF: **A novel family of small cysteine-rich antimicrobial peptides from seed of Impatiens balsamina is derived from a single precursor protein**. *Journal of Biological Chemistry* 1997, **272**(39):24480-24487.

196. Betz SF: **Disulfide bonds and the stability of globular proteins**. *Protein Science* 1993, **2**(10):1551-1558.

197. Cheek S, Krishna SS, Grishin NV: **Structural classification of small, disulfide-rich protein domains**. *Journal of Molecular Biology* 2006, **359**(1):215-237.

198. Lavergne V, J Taft R, F Alewood P: **Cysteine-rich mini-proteins in human biology**. *Current Topics in Medicinal Chemistry* 2012, **12**(14):1514-1533.

199. Selsted ME, Ouellette AJ: **Mammalian defensins in the antimicrobial immune response**. *Nature Immunology* 2005, **6**(6):551-557.

200. Saucedo AL, Flores-Solis D, Rodriguez de la Vega RC, Ramirez-Cordero B, Hernandez-Lopez R, Cano-Sanchez P, Noriega Navarro R, Garcia-Valdes J, Coronas-Valderrama F, de Roodt A *et al*: **New tricks of an old pattern: structural versatility of scorpion toxins with common cysteine spacing**. *The Journal of biological chemistry* 2012, **287**(15):12321-12330.

201. Bensch KW, Raida M, Magert HJ, Schulz-Knappe P, Forssmann WG: **hBD-1: a novel beta-defensin from human plasma**. *FEBS Letters* 1995, **368**(2):331-335.

202. Liu L, Zhao C, Heng HHQ, Ganz T: **The Human β-Defensin-1 and α-Defensins Are Encoded by Adjacent Genes: Two Peptide Families with Differing Disulfide Topology Share a Common Ancestry**. *Genomics* 1997, **43**(3):316-320.

203. Harder J, Bartels J, Christophers E, Schroder JM: **A peptide antibiotic from human skin**. *Nature* 1997, **387**(6636):861.

204. Jia HP, Schutte BC, Schudy A, Linzmeier R, Guthmiller JM, Johnson GK, Tack BF, Mitros JP, Rosenthal A, Ganz T *et al*: **Discovery of new human beta-defensins using a genomics-based approach**. *Gene* 2001, **263**(1-2):211-218.

205. Dempsey CE: **The actions of melittin on membranes**. *Biochimica et Biophysica Acta (BBA)-Reviews on Biomembranes* 1990, **1031**(2):143-161.

206. Schiller PW, Weltrowska G, Nguyen TM, Wilkes BC, Chung NN, Lemieux C: **TIPP [. psi.]: a highly potent and stable pseudopeptide. delta. opioid receptor antagonist with extraordinary. delta. selectivity**. *Journal of Medicinal Chemistry* 1993, **36**(21):3182-3187.

207. Schiller PW, Weltrowska G, Berezowska I, Nguyen TMD, Wilkes BC, Lemieux C, Chung NN: **The TIPP opioid peptide family: development of δ antagonists, δ agonists, and mixed μ agonist/δ antagonists**. *Peptide Science* 1999, **51**(6):411-425.

208. Schiller PW, Nguyen TM-D, Berezowska I, Dupuis S, Weltrowska G, Chung NN, Lemieux C: **Synthesis and in vitro opioid activity profiles of DALDA analogues**. *European Journal of Medicinal Chemistry* 2000, **35**(10):895-901.

209. Rew Y, Malkmus S, Svensson C, Yaksh TL, Chung NN, Schiller PW, Cassel JA, DeHaven RN, Goodman M: **Synthesis and biological activities of cyclic lanthionine enkephalin analogues: δ-opioid receptor selective ligands**. *Journal of Medicinal Chemistry* 2002, **45**(17):3746-3754.

210. Craik DJ, Clark RJ, Daly NL: **Potential therapeutic applications of the cyclotides and related cystine knot mini-proteins**. *Expert Opin Investig Drugs* 2007, **16**(5):595-604.

211. Jagadish K, Camarero JA: **Cyclotides, a promising molecular scaffold for peptide-based therapeutics**. *Peptide Science* 2010, **94**(5):611-616.

212. Korsinczky ML, Schirra HJ, Craik DJ: **Sunflower trypsin inhibitor-1**. *Current Protein and Peptide Science* 2004, **5**(5):351-364.

213. US VE, Gaddum JH: **An unidentified depressor substance in certain tissue extracts**. *The Journal of Physiology* 1931, **72**(1):74-87.

214. Gaddum JH, Schild H: **Depressor substances in extracts of intestine**. *The Journal of Physiology* 1934, **83**(1):1-14.

215. Vaught JL: **Substance P antagonists and analgesia: A review of the hypothesis**. *Life Sciences* 1988, **43**(18):1419-1431.

216. Pernow B: **Substance P**. *Pharmacological Reviews* 1983, **35**(2):85-141.

217. Chang MM, Leeman SE, Niall HD: **Amino-acid sequence of substance P**. *Nature* 1971, **232**(29):86-87.

218. Otsuka M, Yoshioka K: **Neurotransmitter functions of mammalian tachykinins**. *Physiological Reviews* 1993, **73**(2):229-308.

219. Kerdelhué B, Gordon K, Williams R, Lenoir V, Fardin V, Chevalier P, Garret C, Duval P, Kolm P, Hodgen G: **Stimulatory effect of a specific substance P antagonist (RPR 100893) of the human NK1 receptor on the estradiol-induced LH and FSH surges in the ovariectomized cynomolgus monkey**. *Journal of Neuroscience Research* 1997, **50**(1):94-103.

220. Stahl SM: **Peptides and psychiatry, Part 3: Substance P and serendipity: Novel psychotropics are a possibility**. *Journal of Clinical Psychiatry* 1999.

221. Henry JL: **Substance P and inflammatory pain: potential of substance P antagonists as analgesics**. *Agents and Actions Supplements* 1992, **41**:75-87.

222. Cao YQ, Mantyh PW, Carlson EJ, Gillespie A-M, Epstein CJ, Basbaum AI: **Primary afferent tachykinins are required to experience moderate to intense pain**. *Nature* 1998, **392**(6674):390-394.

223. Oku R, Satoh M, Takagi H: **Release of substance P from the spinal dorsal horn is enhanced in polyarthritic rats**. *Neuroscience Letters* 1987, **74**(3):315-319.

224. Duggan A, Hendry I, Morton C, Hutchinson W, Zhao Z: **Cutaneous stimuli releasing immunoreactive substance P in the dorsal horn of the cat**. *Brain research* 1988, **451**(1):261-273.

225. Bear MF, Connors BW, Paradiso MA: **Neuroscience, Exploring the Brain**, 3rd edn: Wolters Kluwer Health; 2007.

226. DeVane CL: **Substance P: a new era, a new role**. *Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy* 2001, **21**(9):1061-1069.

227. Radhakrishnan V, Henry J: **Electrophysiology of neuropeptides in the sensory spinal cord**. *Progress in Brain Research* 1995, **104**:175-195.

228. Folkers K, Horig J, ROSELL S, BJÖRKROTH U: **Chemical design of antagonists of substance P**. *Acta Physiologica Scandinavica* 1981, **111**(4):505-506.

229. Snider RM, Constantine JW, Lowe Jd, Longo KP, Lebel WS, Woody HA, Drozda SE, Desai MC, Vinick FJ, Spencer RW: **A potent nonpeptide antagonist of the substance P (NK1) receptor**. *Science* 1991, **251**(4992):435-437.

230. Sakurada T, Greves P, Stewart J, Terenius L: **Measurement of substance P metabolites in rat CNS**. *Journal of Neurochemistry* 1985, **44**(3):718-722.

231. Lee CM, Sandberg BE, Hanley MR, Iversen LL: **Purification and Characterisation of a Membrane-Bound Substance-P-Degrading Enzyme from Human Brain**. *European Journal of Biochemistry* 1981, **114**(2):315-327.

232. Nyberg F, Terenius L, Henriksen J: **Enzymatic inactivation of neuropeptides**. *Degradation of Bioactive Substances, Physiology and Pathophysiology* 1991.

233. Hall ME, Stewart JM: **Substance P and behavior: opposite effects of N-terminal and C-terminal fragments**. *Peptides* 1983, **4**(5):763-768.

234. Hallberg M, Nyberg F: **Neuropeptide conversion to bioactive fragments-an important pathway in neuromodulation**. *Current Protein and Peptide Science* 2003, **4**(1):31-44.

235. Stewart JM, Hall ME, Harkins J, Frederickson RC, Terenius L, Hökfelt T, Krivoy WA: **A fragment of substance P with specific central activity: SP (1–7)**. *Peptides* 1982, **3**(5):851-857.

236. Sakurada C, Watanabe C, Sakurada T: **Occurrence of substance P (1-7) in the metabolism of substance P and its antinociceptive activity at the mouse spinal cord level**. *Methods and Findings in Experimental and Clinical Pharmacology* 2004, **26**(3):171-176.

237. Ohsawa M, Carlsson A, Asato M, Koizumi T, Nakanishi Y, Fransson R, Sandström A, Hallberg M, Nyberg F, Kamei J: **The effect of substance P 1–7 amide on nociceptive threshold in diabetic mice**. *Peptides* 2011, **32**(1):93-98.

238. Wiktelius D, Khalil Z, Nyberg F: **Modulation of peripheral inflammation by the substance P N-terminal metabolite substance P 1–7**. *Peptides* 2006, **27**(6):1490-1497.

239. Carlsson A, Ohsawa M, Hallberg M, Nyberg F, Kamei J: **Substance P1–7 induces antihyperalgesia in diabetic mice through a mechanism involving the naloxone-sensitive sigma receptors**. *European Journal of Pharmacology* 2010, **626**(2–3):250-255.

240. Kreeger JS, Larson AA: **The substance P amino-terminal metabolite substance P (1-7), administered peripherally, prevents the development of acute morphine tolerance and attenuates the expression of withdrawal in mice**. *Journal of Pharmacology and Experimental Therapeutics* 1996, **279**(2):662-667.

241. Zhou Q, Frändberg P-A, Kindlundh A, Le Grevès P, Nyberg F: **Substance P (1–7) affects the expression of dopamine D2 receptor mRNA in male rat brain during morphine withdrawal**. *Peptides* 2003, **24**(1):147-153.

242. Folkers K, Feng D-M, Asano N, Håkanson R, Weisenfeld-Hallin Z, Leander S: **Spantide II, an effective tachykinin antagonist having high potency and negligible neurotoxicity**. *Proceedings of the National Academy of Sciences* 1990, **87**(12):4833-4835.

243. Dutta AS, Gormley JJ, Graham AS, Briggs I, Growcott JW, Jamieson A: **Analogs of substance P. Peptides containing D-amino acid residues in various positions of substance P and displaying agonist or receptor selective antagonist effects**. *Journal of Medicinal Chemistry* 1986, **29**(7):1163-1171.

244. Caranikas S, Mizrahi J, Escher E, Regoli D: **Synthesis and biological activities of substance P antagonists**. *Journal of Medicinal Chemistry* 1982, **25**(11):1313-1316.

245. Ljungqvist A, Chu J-Y, Tang P-FL, Bender W, Håkanson R, Leander S, Rosell S, Folkers K: **Increased potency of antagonists of substance P having asparagine in position 6**. *Regulatory peptides* 1989, **24**(3):283-291.

246. Jensen RT, Jones SW, Lu Y-A, Xu J-C, Folkers K, Gardner JD: **Interaction of substance P antagonists with substance P receptors on dispersed pancreatic acini**. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* 1984, **804**(2):181-191.

247. Regoli D, Boudon A, Fauchere J-L: **Receptors and antagonists for substance P and related peptides**. *Pharmacological Reviews* 1994, **46**(4):551-599.

248. Botros M, Johansson T, Zhou Q, Lindeberg G, Tomboly C, Toth G, Le Greves P, Nyberg F, Hallberg M: **Endomorphins interact with the substance P (SP) aminoterminal SP(1-7) binding in the ventral tegmental area of the rat brain**. *Peptides* 2008, **29**(10):1820-1824.

249. Igwe OJ, Sun X, Larson AA: **Correlation of substance P-induced desensitization with substance P amino terminal metabolites in the mouse spinal cord**. *Peptides* 1990, **11**(4):817-825.

250. Botros M, Hallberg M, Johansson T, Zhou Q, Lindeberg G, Frandberg PA, Tomboly C, Toth G, Le Greves P, Nyberg F: **Endomorphin-1 and endomorphin-2 differentially interact with specific binding sites for substance P (SP) aminoterminal SP1-7 in the rat spinal cord**. *Peptides* 2006, **27**(4):753-759.

251. Merrifield RB: **Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide**. *Journal of the American Chemical Society* 1963, **85**(14):2149-2154.

252. Merrifield B: **Concept and early development of solid-phase peptide synthesis**. In: *Methods in Enzymology.* vol. 289; 1997: 3-13.

253. Stewart JM: **Cleavage methods following Boc-based solid-phase peptide synthesis**. In: *Methods in Enzymology.* vol. 289; 1997: 29-44.

254. Wellings DA, Atherton E: **Standard Fmoc protocols**. In: *Methods in Enzymology.* vol. 289; 1997: 44-67.

255. Chan WC, White PD: **Fmoc solid phase peptide synthesis**: Oxford University Press; 2000.

256. Tam JP, Lu YA, Liu CF, Shao J: **Peptide synthesis using unprotected peptides through orthogonal coupling methods**. *Proceedings of the National Academy of Sciences* 1995, **92**(26):12485-12489.

257. Tam JP, Lu Y-A, Yu Q: **Thia Zip Reaction for Synthesis of Large Cyclic Peptides: Mechanisms and Applications†**. *Journal of the American Chemical Society* 1999, **121**(18):4316-4324.

258. Tam JP, Lu Y-A: **Synthesis of large cyclic cystine-knot peptide by orthogonal coupling strategy using unprotected peptide precursor**. *Tetrahedron Letters* 1997, **38**(32):5599-5602.

259. Dawson PE, Muir TW, Clark-Lewis I, Kent SB: **Synthesis of proteins by native chemical ligation**. *Science* 1994:776-776.

260. Li X, Kawakami T, Aimoto S: **Direct preparation of peptide thioesters using an Fmoc solid-phase method**. *Tetrahedron Letters* 1998, **39**(47):8669-8672.

261. Futaki S, Sogawa K, Maruyama J, Asahara T, Niwa M, Hojo H: **Preparation of Peptide Thioesters using Fmoc-Solid-Phase Peptide Synthesis and its Application to the Construction of a Template-Assembled Synthetic Protein (TASP)**. *Tetrahedron Letters* 1997, **38**(35):6237-6240.

262. Backes BJ, Virgilio AA, Ellman JA: **Activation Method to Prepare a Highly Reactive Acylsulfonamide "Safety-Catch" Linker for Solid-**

Phase Synthesis1. *Journal of the American Chemical Society* 1996, **118**(12):3055-3056.

263. Blanco-Canosa JB, Dawson PE: **An Efficient Fmoc-SPPS Approach for the Generation of Thioester Peptide Precursors for Use in Native Chemical Ligation**. *Angewandte Chemie International Edition* 2008, **47**(36):6851-6855.

264. Fang G-M, Li Y-M, Shen F, Huang Y-C, Li J-B, Lin Y, Cui H-K, Liu L: **Protein Chemical Synthesis by Ligation of Peptide Hydrazides**. *Angewandte Chemie International Edition* 2011, **50**(33):7645-7649.

265. Nakamura Ki, Kanao T, Uesugi T, Hara T, Sato T, Kawakami T, Aimoto S: **Synthesis of peptide thioesters via an N–S acyl shift reaction under mild acidic conditions on an N-4,5-dimethoxy-2-mercaptobenzyl auxiliary group**. *Journal of Peptide Science* 2009, **15**(11):731-737.

266. Hojo H, Onuma Y, Akimoto Y, Nakahara Y, Nakahara Y: **N-Alkyl cysteine-assisted thioesterification of peptides**. *Tetrahedron Letters* 2007, **48**(1):25-28.

267. Ollivier N, Dheur J, Mhidia R, Blanpain A, Melnyk O: **Bis(2-sulfanylethyl)amino Native Peptide Ligation**. *Organic Letters* 2010, **12**(22):5238-5241.

268. Tsuda S, Shigenaga A, Bando K, Otaka A: **N→S Acyl-Transfer-Mediated Synthesis of Peptide Thioesters Using Anilide Derivatives**. *Organic Letters* 2009, **11**(4):823-826.

269. Nagaike F, Onuma Y, Kanazawa C, Hojo H, Ueki A, Nakahara Y, Nakahara Y: **Efficient Microwave-Assisted Tandem N- to S-Acyl Transfer and Thioester Exchange for the Preparation of a Glycosylated Peptide Thioester**. *Organic Letters* 2006, **8**(20):4465-4468.

270. Sharma RK, Tam JP: **Tandem Thiol Switch Synthesis of Peptide Thioesters via N–S Acyl Shift on Thiazolidine**. *Organic Letters* 2011, **13**(19):5176-5179.

271. Hemu X, Taichi M, Qiu Y, Liu DX, Tam JP: **Biomimetic synthesis of cyclic peptides using novel thioester surrogates**. *Peptide Science* 2013, **100**(5):492-501.

272. Lukszo J, Patterson D, Albericio F, Kates S: **3-(1-Piperidinyl)alanine formation during the preparation ofC-terminal cysteine peptides with the Fmoc/t-Bu strategy**. *Letters in Peptide Science* 1996, **3**(3):157-166.

273. Taichi M, Hemu X, Qiu Y, Tam JP: **A thioethylalkylamido (TEA) thioester surrogate in the synthesis of a cyclic peptide via a tandem acyl shift**. *Organic Letters* 2013, **15**(11):2620-2623.

274. Jeon HK, Jung NP, Choi IH, Oh YK, Shin HC, Gwag BJ: **Substance P augments nitric oxide production and gene expression in murine macrophages**. *Immunopharmacology* 1999, **41**(3):219-226.

275. Henriques ST, Craik DJ: **Cyclotides as templates in drug design**. *Drug Discovery Today* 2010, **15**(1):57-64.

276. Wang CK, Gruber CW, Cemazar M, Siatskas C, Tagore P, Payne N, Sun G, Wang S, Bernard CC, Craik DJ: **Molecular Grafting onto a Stable Framework Yields Novel Cyclic Peptides for the Treatment of Multiple Sclerosis**. *ACS Chemical Biology* 2013.

277. Olsen JV, Ong S-E, Mann M: **Trypsin cleaves exclusively C-terminal to arginine and lysine residues**. *Molecular & Cellular Proteomics* 2004, **3**(6):608-614.

278. Rodriguez J, Gupta N, Smith RD, Pevzner PA: **Does trypsin cut before proline?** *Journal of Proteome Research* 2007, **7**(01):300-305.

279. Zhou Q, Liu Z, Ray A, Huang W, Karlsson K, Nyberg F: **Alteration in the brain content of substance P (1–7) during withdrawal in morphine-dependent rats**. *Neuropharmacology* 1998, **37**(12):1545-1552.

280. Sakurada T, Le Greves P, Stewart J, Terenius L: **Measurement of substance P metabolites in rat CNS**. *Journal of Neurochemistry* 1985, **44**(3):718-722.

281. Igwe OJ, Kim DC, Seybold VS, Larson AA: **Specific binding of substance P aminoterminal heptapeptide [SP(1-7)] to mouse brain and spinal cord membranes**. *The Journal of Neuroscience* 1990, **10**(11):3653-3663.

282. Toriyabe M, Omote K, Kawamata T, Namiki A: **Contribution of interaction between nitric oxide and cyclooxygenases to the production of prostaglandins in carrageenan-induced inflammation**.

The Journal of the American Society of Anesthesiologists 2004, **101**(4):983-990.

283. Chen Y, Boettger MK, Reif A, Schmitt A, Üçeyler N, Sommer C: **Nitric oxide synthase modulates CFA-induced thermal hyperalgesia through cytokine regulation in mice**. *Molecular Pain* 2010, **6**(1):1.

284. Freire MAM, Guimarães JS, Gomes-Leal W, Pereira A: **Pain modulation by nitric oxide in the spinal cord**. *Frontiers in Neuroscience* 2009, **3**:24.

285. Miyamoto T, Dubin AE, Petrus MJ, Patapoutian A: **TRPV1 and TRPA1 mediate peripheral nitric oxide-induced nociception in mice**. *PloS One* 2009, **4**(10):e7596.

286. Lorsbach RB, Murphy WJ, Lowenstein CJ, Snyder SH, Russell SW: **Expression of the nitric oxide synthase gene in mouse macrophages activated for tumor cell killing. Molecular basis for the synergy between interferon-gamma and lipopolysaccharide**. *Journal of Biological Chemistry* 1993, **268**(3):1908-1913.

287. Jeon HK, Jung NP, Choi IH, Oh YK, Shin HC, Gwag BJ: **Substance P augments nitric oxide production and gene expression in murine macrophages**. *Immunopharmacology* 1999, **41**(3):219-226.

288. Mustazza C, Bastanzio G: **Development of nociceptin receptor (NOP) agonists and antagonists**. *Medicinal Research Reviews* 2011, **31**(4):605-648.

289. Mika J, Obara I, Przewlocka B: **The role of nociceptin and dynorphin in chronic pain: Implications of neuro–glial interaction**. *Neuropeptides* 2011, **45**(4):247-261.

290. Sadler K, Eom KD, Yang JL, Dimitrova Y, Tam JP: **Translocating proline-rich peptides from the antimicrobial peptide bactenecin 7**. *Biochemistry* 2002, **41**(48):14150-14157.