TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

1.1 Introduction to Bombyx mori

Bombyx mori (commonly known as Mulberry Silkworm) is one of the major domesticated silkworms for producing commercial silk. This belongs to family *Bombycidae*. Besides the production of good quality silk, the silkworm has been exploited for its different bioactive properties, particularly, antimicrobial properties. Silkworms go through four stages of development, egg, larva (silkworm caterpillar), pupa and adult (silkworm moth). Silkworm has received extensive attention as a model for human antimicrobial drug screening, Kaito *et al.* (2002), human disease models for diabetes, phenylketonuria, Parkinson's disease and Hermansky–Pudlak syndrome, (Zhang *et al.* 2015). Silkworm is sensitive to environmental pollutions, especially to pesticides, heavy metals, and other harmful chemicals, which is why it is used as a model animal used in environment monitoring (Sekimura 2005, Hamamoto *et al.* 2009).

1.1.1 Immune system of *Bombyx mori*

Bombyx mori possesses an effective innate immune system against foreign microorganisms. Innate immunity is divided into two major reaction types: humoral and cellular reactions. Humoral reactions involve soluble proteins within the hemolymph like phenoloxidase, antimicrobial proteins (AMPs), lysozymes, and lectins, whereas hemocytes mediate cellular reactions like phagocytosis, encapsulation and nodule formation. (Tanaka, H & Yamakawa, M. (2011).

1.2 Antimicrobial peptides

Antimicrobial peptides (AMPs), conjointly referred to as Host Defense Peptides (HDPs) are a part of the innate immunological response found among all categories of life. These peptides are considered as a probable candidate for forthcoming drugs, because of their

broad range of activity, lesser toxicity and decreased resistance development by target cells. The smaller size of AMPs helps in the rapid diffusion and secretion outside the cells, which is mandatory for evoking immediate response against pathogenic microbes. Antimicrobial peptides are incontestable to kill Gram-negative and Gram-positive bacteria, enveloped viruses, fungi and even transformed or cancerous cells. However, antibiotic resistance is projected as one of the greatest threats to human health in the future and hence alternatives are being explored to combat resistance. Antimicrobial peptides have shown great promise because the use of AMPs leads bacteria to develop no or low resistance.

1.2.1 Antimicrobial peptides in *Bombyx mori*

In *Bombyx mori*, a large number of antibacterial proteins are classified into seven major groups namely, Gloverin, Cecropin, Attacin, Lebocin, Moricin, Defensin and Enbocin. The detergent properties of these antimicrobial proteins disrupt the cell membranes of the invading microbes and enzymatically attack bacteria by hydrolyzing their peptidoglycan cell walls. Antimicrobial peptides are secreted from the fat body of *Bombyx mori* and play a crucial role in eliminating invaders.

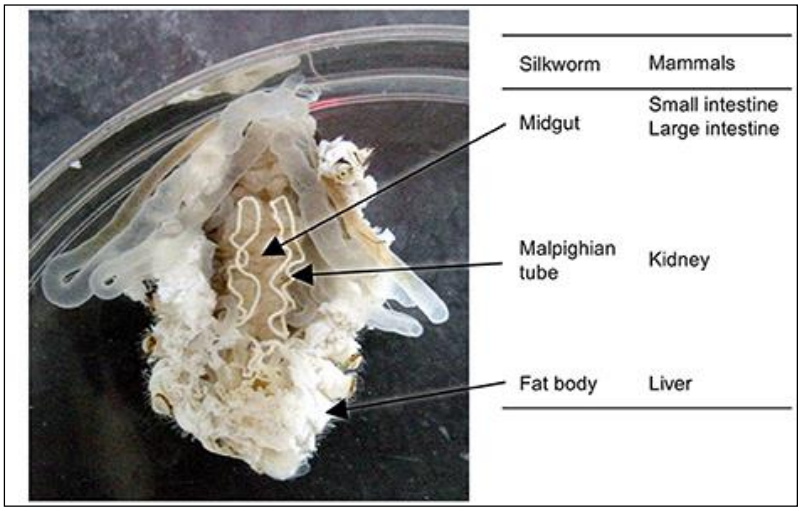


Figure 1.2. 1: Comparison of organs involved in absorption, metabolism and elimination between silkworm and mammals (Panthee et. al 2017)

1.2.2 Applications of antimicrobial peptides

There are several potential applications of AMPs to develop new strategies or substitute the existing ones for more sustainable and efficient activities. Though not yet released as drugs, many AMPs have been assessed as potential therapeutics which could replace the abolished antibiotics (Gordon *et al.* 2005; Schro¨der and Harder 2006). These new antibiotics would revolutionize the health situation of human and domesticated animals. Recombinant DNA techniques have also facilitated new strategies based on transgenic plants with AMP-expressing genes (van der Biezen 2001). Some of the AMP-based products have been introduced to markets and are used in a broad range of agricultural activities. Use of AMPs is considered as a solution for concerns about antibiotic resistance development and antibiotic residues in aquaculture products (Akinbowale *et. al* 2006). AMPs have numerous potentials to develop the quality of conventional animal husbandry and to resolve some of the existing problems which are consequences of the extensive use of antibiotics. (Zhang,G *et. al* 2000) Agera antibacterial peptide cream is currently in the market used as an anti- acne cream which is a combination of lysozyme and stimulysin.

1.3 Silkworm as a model organism

As an important economic insect, silkworm *Bombyx mori* has numerous advantages in life science, such as low breeding cost, large progeny size, short generation time, and clear genetic background. The completion of the silkworm genome has further accelerated it to be a modern model organism in life science. Genomic studies showed that some silkworm genes are highly homologous to certain genes related to human hereditary disease and, therefore, are a candidate model for studying human disease. Silkworm models for bacterial infection, fungal infection, virus infection, and natural immune stimulation have been established (Kaito and Sekimizu 2007, Ishii *et al.* 2015a). Thus, the use of silkworm as a model organism for studying human tumour, degenerative disease, and metabolic disease has become a current research focus.

1.4 *in-silico* approach

in silico biology refers to computational models of biology. Due to the huge amounts of data that is being generated by molecular and cell experimental biologists, computational biology is increasingly necessary to manage it. *in silico* biology draws from the vast amounts of biological information available, and applies sophisticated algorithms or simulations to advance scientific understanding. The vast majority of known proteins have not yet been characterized experimentally, and there is very little that is known about their function. New unannotated sequences are added to the databases at a pace that far exceeds the one in which they are annotated in the lab. Computational biology offers tools that can provide insight into the function of proteins based on their sequence, their structure, their evolutionary history, and their association with other proteins. The results of these simulations can then be tested experimentally or serve as a guide for future physical experimentation.

1.4.1 Primary Sequence Analysis

The analysis of primary sequence (amino acid sequence) of AMP's in *Bombyx mori* shows logical insights into understanding the structural characteristics and variations, in order to give us a better understanding of their function and analyzed a wide range of basic information including a prediction of repetitive sequences, the motif, domain, or active sites, and the ability to form a coiled-coiled structure by performing primary sequence analysis.

ProtParam (Gasteiger *et al*., 2005) is a tool which allows the computation of various physical and chemical parameters for a given protein stored in Swiss-Prot or TrEMBLor for a user entered protein sequence. The computed parameters include the molecular weight, theoretical pI, amino acid composition, atomic composition, extinction coefficient, estimated half-life, instability index, aliphatic index and grand average of hydropathicity (GRAVY). ExPASy ProtParam, the primary sequence analysis tool was utilized to analyze the physicochemical properties of the selected protein.

1.4.2. Secondary Sequence Analysis

Secondary structure analysis aims to predict the secondary structure of proteins and peptides based only on the knowledge of their amino acid sequence. For proteins, a prediction consists of assigning regions of the amino acid sequence as alpha helices, beta strands, or turns.

SSPro (Cheng *et. al*, 2005), GOR4 (Garnier *et al.,* 1996), SOPMA (Geourjon *et al.,* 1995) and Porter (Mirabello & Pollastri, 2013) are the tools that are used in order to predict the secondary structure of the AMPs.

1.4.3. Tertiary and Quaternary Sequence Analysis

The protein structure prediction remains an extremely difficult and unresolved task. This is due to two problems which are - calculation of the protein free energy, and finding the global minimum of this protein free energy. In case of quaternary structure prediction, which is the case of complexes formed due to two or more proteins, protein-protein docking methods can be only used if the structure is known, or if the structure can be predicted with high accuracy. As we are working closely with AMPs and not proteins, further problems arise with respect to the structure of peptides as they are known to be less well defined in structure due to the formation of complex secondary, tertiary and quaternary structures and are governed by short-range interactions, unlike proteins.

Correlated motions in proteins can mediate fundamental biochemical processes such as signal transduction and allostery. The mechanisms that underlie these processes remain largely unknown due mainly to limitations in their direct detection. Here, based on a detailed analysis of protein structures deposited in the protein data bank, as well as molecular simulation we provide general evidence for the transfer of structural information by correlated backbone motions, mediated by hydrogen bonds, across β-sheets. We also show that the observed local and long-range correlated motions are mediated by the collective motions of β-sheets and investigate their role in large-scale conformational changes. Correlated motions represent a fundamental property of β-sheets that contributes to protein function. (Fenwick *et. al* 2014)

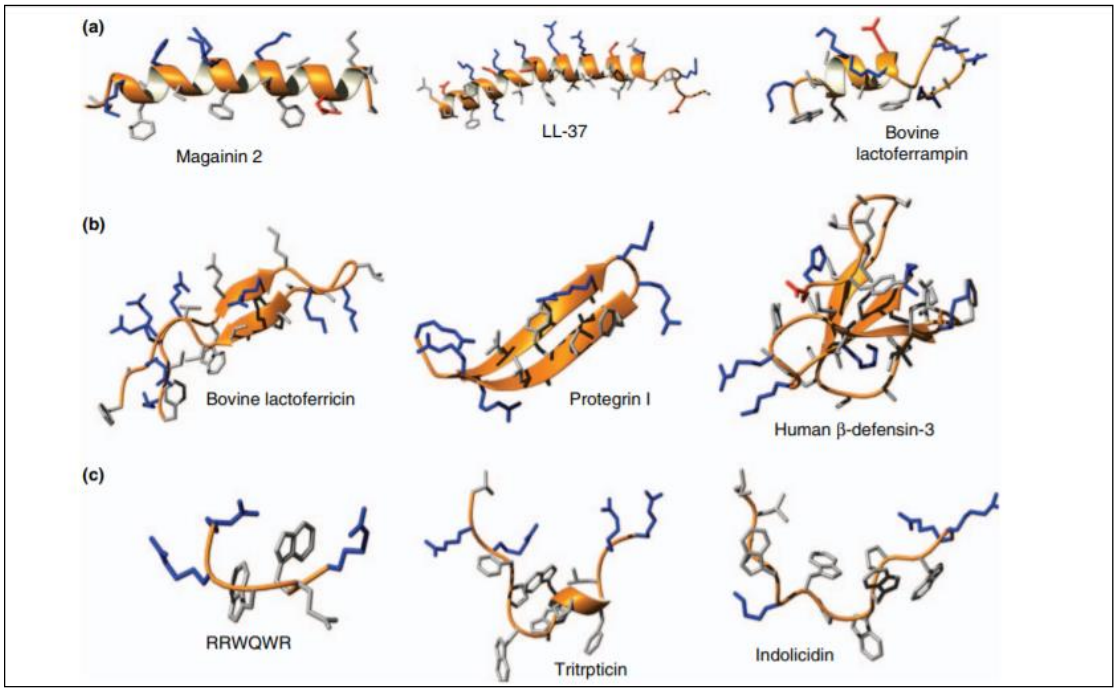


Figure 1.4.3: An overview of the major structural classes of antimicrobial peptides (AMPs). (a) a-Helical peptides, (b) b-sheet peptides and (c) extended peptides. All of these structures were solved by solution NMR spectroscopy in the presence of detergent micelles, except for the b-sheet peptides, which were studied in aqueous solution. Positively charged side chains are colored in blue, negatively charged side chains in red and remaining side chains in grey. PDB IDs: magainin 2, 2MAG; LL-37, 2K6O; bovine lactoferricin, 1LFC; protegrin 1, 1PG1; human b-defensin-3, 1KJ5; tritrpticin, 1D6X; indolicidin, 1G89. (Nguyen *et al.,* 2011)

CHAPTER 2: REVIEW OF LITERATURE

2.1 *Bombyx mori* as a model organism

Silkworm *Bombyx mori* (Lepidoptera: *Bombycidae*) has big progeny size and short generation time with larval stage lasting about 25–30. It has 28 pairs of chromosomes, which is far more than those of *Drosophila melanogaster* (four pairs), rich in genetic traits. Silkworm genome is about 450 million bp, which is 1/6 of the human genome and four times more than the genome of *D. melanogaster*. In addition, silkworm has diverse mutant strains and morphological mutations. Its death, either genetic or imposed, does not involve any bioethical issues (Chen et al. 2014). In addition, silkworm has moderate body size and can be easily dissected to obtain many tissues and organs such as midgut, fat body, silk gland, and hemolymph. Furthermore, silkworm can be used to perform oral administration and intravenous injection experiments similar to those of mammals. With the completion of the silkworm genome project and the establishment of silkworm genomic database and protein database, silkworm begin to stand out as a valuable model in scientific research.

Silkworm has received extensive attention as a model for human antimicrobial drug screening. Kaito et al. (2002) used silkworm to the study of antimicrobial drugs and demonstrated that silkworm could replace mammals for bacterial pathogenicity experiments. Silkworm is used as a human disease model where diseases like phenylketonuria, Sepiapterin reductase deficiency (SRD), Parkinson's disease and Hermansky–Pudlak syndrome. The insulin-like peptide encoded by the silkworm gene was found to have about 40% similarity to human insulin, based on which, a silkworm diabetes model was established (Zhang et al. 2015). Silkworm is sensitive to environmental pollutions, especially to pesticides, heavy metals, and other harmful chemicals, which is why it is used as a model animal used in environment monitoring (Sekimura 2005, Hamamoto et al. 2009). The silkworm model can be used to monitor toxic substances that have negative effect on soil, water quality and medical environment. *Bombyx mori* has been used as an Bacterial infection model (Hamamoto et al. (2004), Fujiyuki et al. (2012), Miyazaki et al. (2012)), fungal (Matsumoto et al. (2012), Ueno et al. (2011)) and viral infection model (Wang and Schweizer (2008)).

2.2 Experimental work done on AMPs

Cecropin B isolated from H. cecropia and B.mori was shown to have antibacterial activity in transgenic rice and antiviral effect on the viral pathogens of fish. (Sharma et al. (2000) and Chiou et al. (2002)). Lactoferricin isolated from cow was shown to control mastitis in goat (Zhang et al. (2007)).

Recently, two novel antimicrobial peptides have been identified from *B. mori.* First, an Enbocin gene was cloned using a partial cDNA fragment detected by differential hybridization as a probe. Although the amino acid sequence deduced from the nucleotide sequence suggests that enbocin belongs to the cecropin family, a recombinant enbocin has an antibacterial spectrum opposite to cecropin, but no full-size Enbocin cDNA has yet been reported, and no extensive analysis of Enbocin gene expression focusing on induction mechanisms has been conducted. Another novel antibacterial peptide from B. mori is gloverin. Four Gloverin genes have been reported (GenBank accession numbers AB190863, AB190864, AB190865, and AB190866) and found.

cDNAs encoding enbocin2 and enbocin3 were cloned from the fat body of *B. mori* larvae. The Enbocin2 and 3 cDNAs contained an open reading frame consisting of 59 amino acid residues. A putative cleavage site for the signal peptide was predicted between 20-Ala and 21-Lys of these peptides. Further, a cleavage site between 22-Pro and 23-Trp was also predicted by alignment of the amino acid sequence of this peptide with that of Hyalophora cecropia cecropin D. Expression of Enbocin and Bmgloverin isoform genes stimulated by bacteria and bacterial cell wall components was analyzed by real-time RT-PCR using specific primers. Enbocin2 and 3 genes were strongly activated by E. coli and B. subtilis and weakly activated by *S. aureus.* Bmgloverin1, 2, 3, and 4 genes were strongly activated by E. coli and *B. subtilis* and weakly activated by *S. aureus.*

2.3 *in-silico* approach towards AMPs

BmLeb5 sequence obtained was analyzed using the SeqBuilder program in DNASTAR software package. The signal peptide was predicted with SignalP 4.1 (http://www.cbs.dtu.dk/services/SignalP/). The potential O-glycosylation sites were predicted with NetOGlyc 4.0 (http://www.cbs.dtu.dk/services/NetOGlyc/). The homologous sequences were found by BLASTp (https://blast.ncbi.nlm.nih.gov/Blast.cgi). Multiple protein sequences were aligned using Clustalx 2.1 software. The phylogenetic tree was constructed in MEGA 7 (S. Kumar, G. Stecher, and K. Tamura (2016)) using neighbor-joining method.

Ten AMPs of *Bombyx mori* were retrieved from the APD2 database. All ten sequences were aligned using ClustalW2 multiple sequence alignment tool providing FASTA format sequences and PHYLIP format option was selected as result output. Multiple alignment result were transfer to BoxShade tool as other input sequence format and RTF_new option were selected for output results. Best four homologous sequences of each AMP were also retrieved from APD2 database. Sequence homology were observed by using ClustalW2 multiple alignment tool and pretty printing as well as shading of multiple aligned files were carried out by BoxShade tool. All AMP sequences were analyzed by using ProtParam tool of ExPASy for determining length, amino acid composition, Isoelectric point, instability index, aliphatic index, GRAVY and in vivo half-life. Net charge, % hydrophobicity and Boman index were retrieved from APD database search tool. Cell penetrating ability of multiple AMPs were determined using CellPPD tool with its Support Vector Machine (SVM) values, molecular weights, isoelectric points and net charge. Antigenic peptide prediction tool of Immunomedicine Group was used for prediction of antigenic determinants within AMP. EVALLER™ (also referred to as EVALLER 2) web-tool were used for electronically test (eTesting) allergic potential of AMP on the basis of amino acid sequence. ToxinPred tool were used for predicting regions in AMP, which contributes in toxicity. Half-life of AMP in the intestine were predicted by using HLP tool. Helical Wheel Projections tool were used for helix wheel diagram prediction. Sequence Annotated by Structure (SAS) tool of EMBL-EBI were

used for prediction of secondary structures of AMPs. (Bajare Jitendra Shahaji , Oli Ajay Kumar, Jalkute Chidambar Balbhim, Dulange Sanjay Mallikarjun and Thorat Prakash Ramrao)

CHAPTER 3: OBJECTIVES

in silico approach: (7 families - 35 AMPs)

1. Retrieval and sequence analysis of AMPs of *Bombyx mori*
    a) Analysis of full-length and mature forms
    b) Amino acid composition
    c) Multiple Sequence Alignment
    d) Physico-chemical- pI, MW, activity, net charge, hydropathicity, aliphatic and instability index
2. Secondary Structure Prediction of the AMPs

CHAPTER 4: METHODOLOGY

4.1. Retrieval of sequences

CAMP, APD3 and Pfam databases were queried for sequence of AMPs related to silkworm using "*Bombyx mori*" for source organism in the Search tool. The number of hits obtained were 26, 14 and 6 from CAMP, APD3 and Pfam, respectively. These were filtered to remove the duplicates and 35 unique AMP sequences were obtained. The FASTA format of these sequences was retrieved from UniProt.

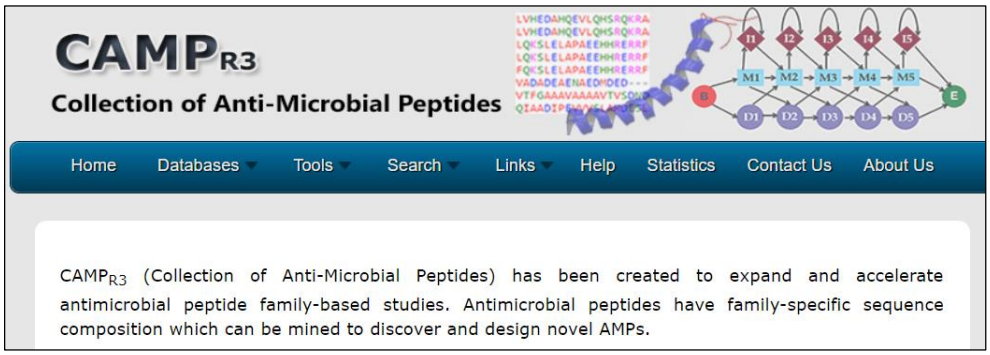


Figure 4.1 a: Snapshot of the CAMP database

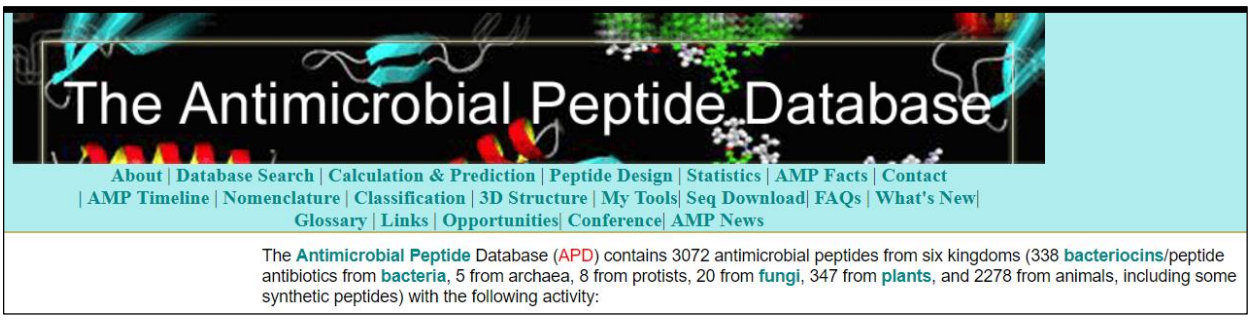


Figure 4.1 b: Snapshot of APD3 database

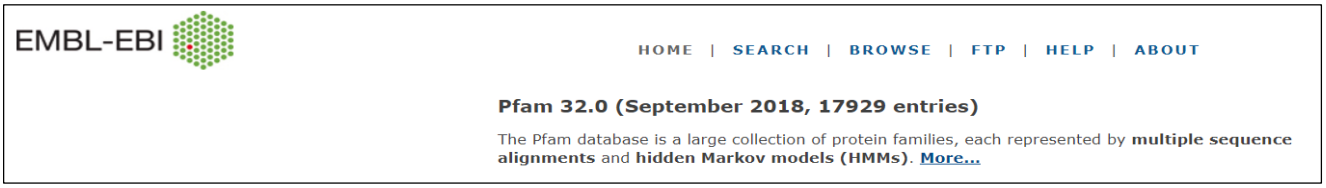


Figure 4.1 c: Snapshot of Pfam

The full length sequences of the AMPs, which were found to have signal peptide, propeptide and mature peptide, were retrieved from UniProt. The sequences for signal peptide were predicted using Phobius and SignalP and the data for propeptide and mature peptides were obtained from UniProt.

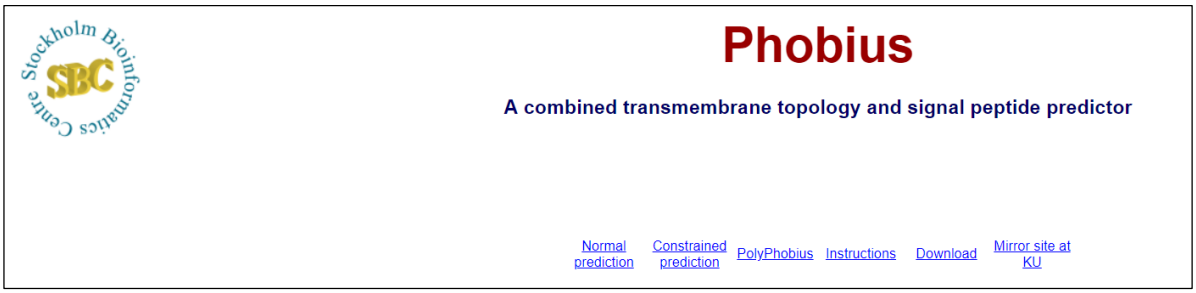


Figure 4.1 d: Snapshot of the Phobius tool

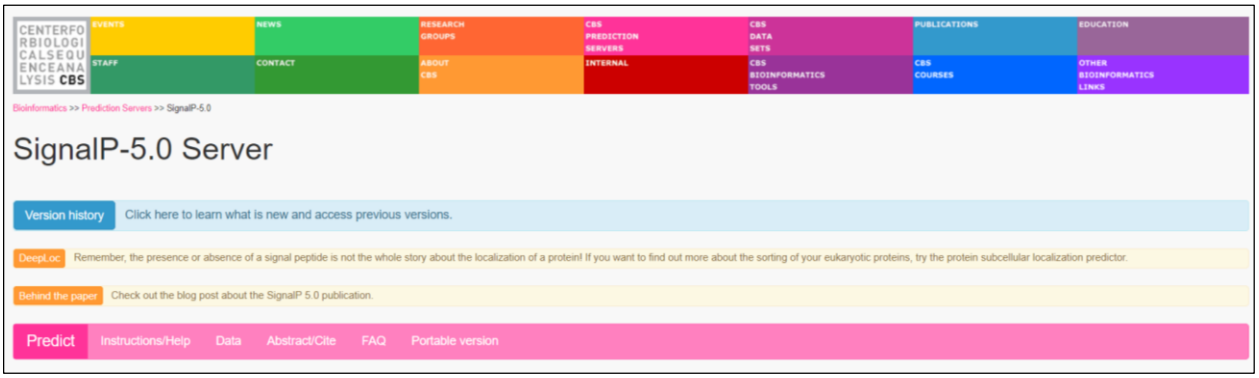


Figure 4.1 e: Snapshot of the SignalP tool

All the sequences with their data were tabulated, as shown in Table 4.2(a). A schematic representation of the full length sequences of all 35 AMPs was created using Illustrator for Biological Sciences (IBS).

4.2. Primary sequence analysis

All AMP sequences were analyzed by using ProtParam tool of ExPASy for determining length, amino acid composition, isoelectric point, instability index, aliphatic index, and GRAVY (grand average of hydropathicity). These values are noted in Table 5.2. The net charge at pH 7.0 was obtained using the Protein Calculator v3.4 (protcalc.sourceforge.net/). ModPred was used to predict the post translational modification sites on the sequences and only the ones with a score of greater than 0.75 (high) were considered. A schematic representation of the predicted post translational modifications for all 35 AMPs was created using Illustrator for Biological Sciences (IBS).

4.3. Secondary Structure Prediction

The FASTA sequences were submitted to Secondary structure prediction tools like PORTER, GOR 4, SOPMA and SSPRO (available on PORTER 4.0, NPS@ , NPS@ and SCRATCH servers respectively). The prediction results were analysed by giving them weighted values based on the percentage accuracy of the respective tools. SSPRO has an accuracy of 92%, PORTER has an accuracy of 82.2%, SOPMA has an accuracy of 69.5% and GOR 4 has an accuracy of 64.4%. Percentiles were considered and SSPRO was given the highest weightage of 1 while the other 3 servers are given weighted values of 0.89, 0.75 and 0.7 respectively. This was done using Microsoft Excel Macros.

Table 4. 3: Secondary structure tabulated using the mentioned tools for Cecropin A (Q27239)

Amino Acids	GOR4	SOPMA	SSPRO	PORTER
R	C	H	C	C
W	C	H	H	C
K	C	H	H	C
L	C	H	H	H
F	E	H	H	H
K	C	H	H	H
K	C	H	H	H
I	C	H	H	H
E	C	H	H	H
K	C	H	H	H
V	C	H	H	H
G	C	H	C	H
R	C	H	H	H
N	C	H	H	H
V	C	H	H	H
R	H	H	H	H
D	H	H	H	H
G	H	H	H	H
L	H	H	H	H
I	H	H	H	H
K	C	H	C	H

A	C	T	C	H
G	C	C	C	H
P	C	C	C	H
A	H	C	H	H
I	H	E	H	H
A	H	E	H	H
V	H	E	H	H
I	H	E	H	H
G	H	E	H	H
Q	C	C	H	H
A	C	C	H	H
K	E	C	C	H
S	E	C	C	C
L	C	H	C	C

CHAPTER 5: RESULTS AND DISCUSSIONS

5.1 Retrieval of sequences

$CAMP_{R3}$ (Collection of Anti-Microbial Peptides) has been created to expand and accelerate antimicrobial peptide family-based studies. Antimicrobial peptides have family-specific sequence composition which can be mined to discover and design novel AMPs. The antimicrobial peptide database (APD, http://aps.unmc.edu/AP/) is an original database initially online in 2003. The APD2 (2009 version) has been regularly updated and further expanded into the APD3. This database currently focuses on natural antimicrobial peptides (AMPs) with defined sequence and activity. Pfam version 32.0 was produced at the European Bioinformatics Institute using a sequence database called *Pfamseq*, which is based on UniProt. The Pfam database is a large collection of protein families, each represented by multiple sequence alignments and hidden Markov models (HMMs). A total of 35 AMPs were derived after eliminating the duplicates. Of the 35 sequences retrieved, only BM Moricin (P82818) has a solved structure (1KV4) in the Protein Data Bank. These 35 AMPs were further categorized into 7 families namely Gloverin, Cecropin, Attacin, Lebocin, Moricin, Defensin and Attacin based on literature.

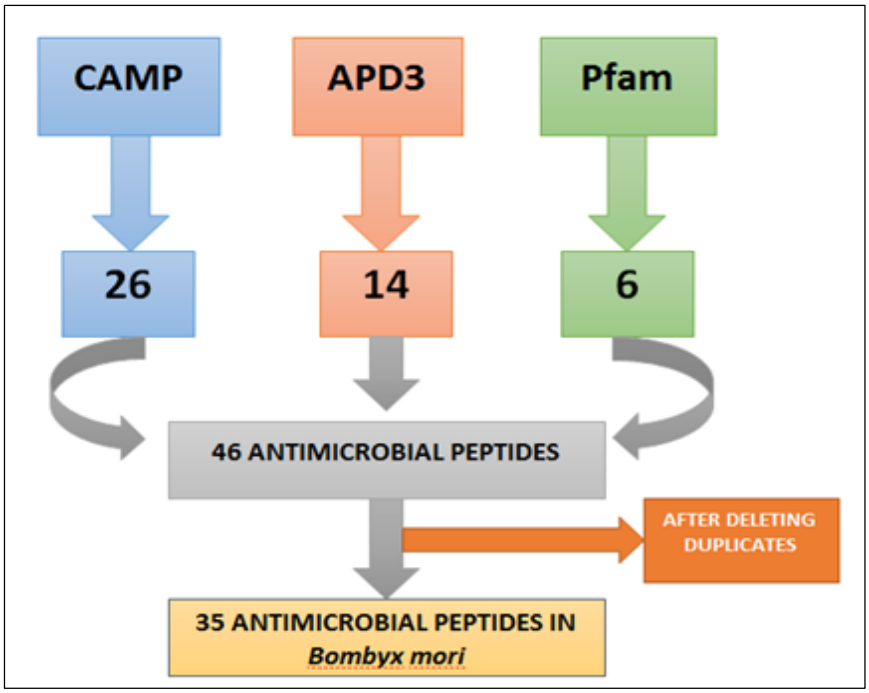


Figure 5.1 a: Workflow followed for retrieval of sequences

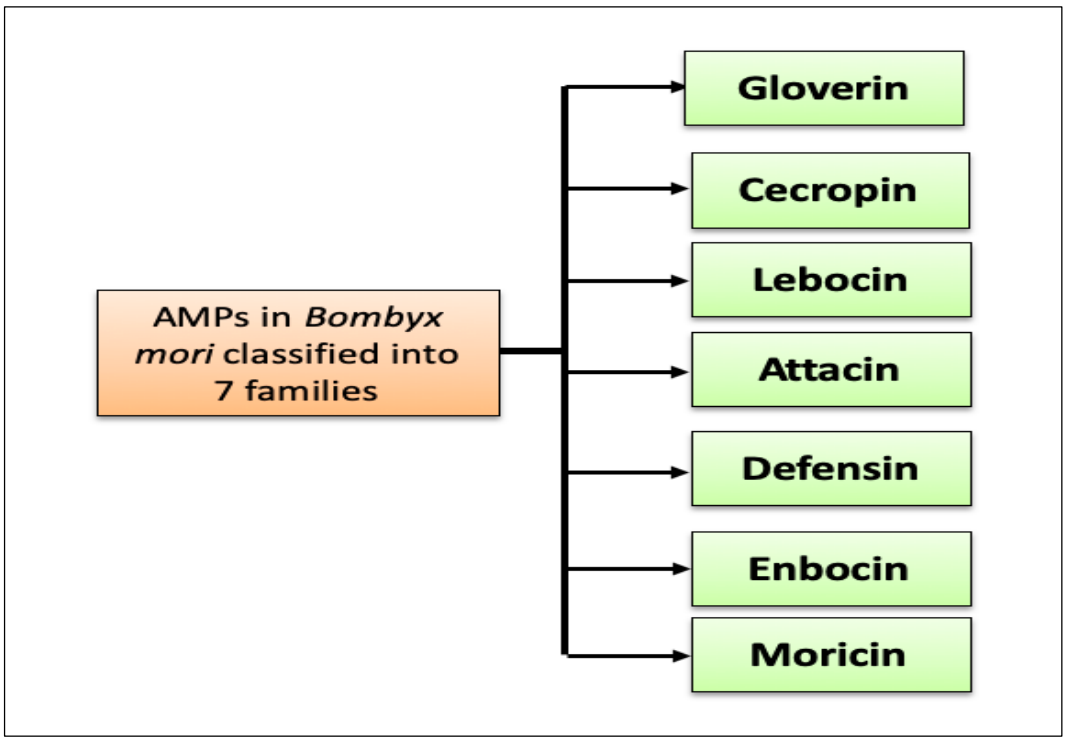


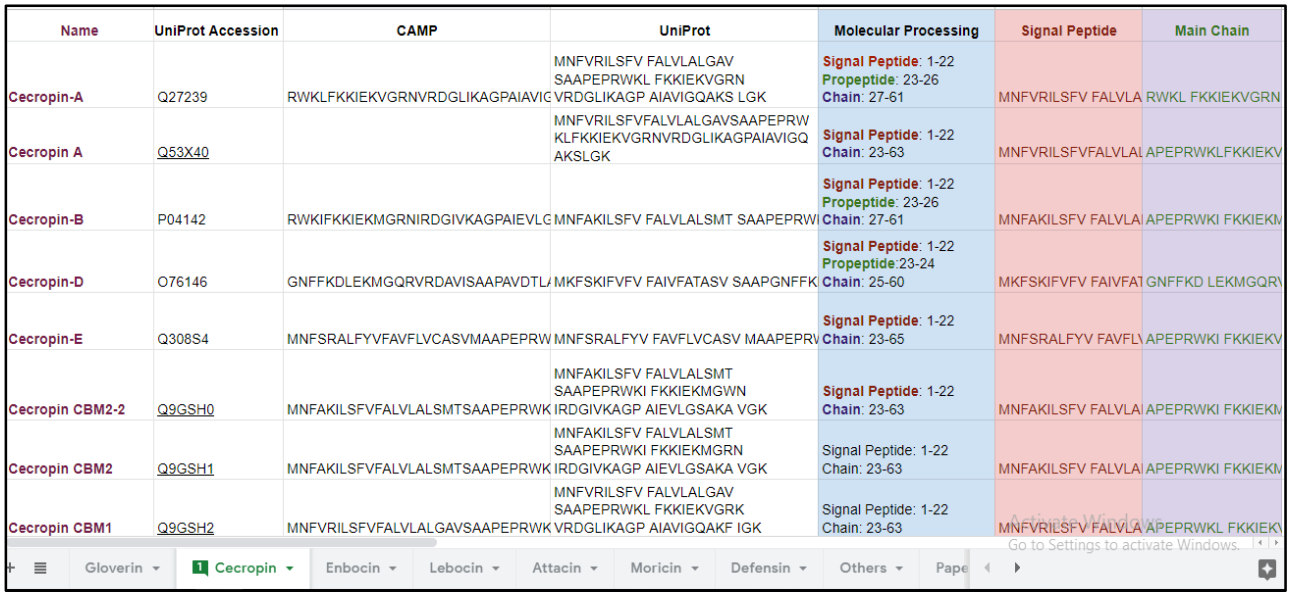
Figure 5.1 b: Classification of AMPs

Table 5. 1: Table showing all 35 AMPs with their UniProt accession number

Sl. No.	AMP	Accession number
1	Gloverin 3	A5LHW4
2	Uncharacterised protein	H9JWE9
3	Gloverin-4	Q2WGL0
4	Gloverin-2	Q2WGL1
5	Cecropin A	Q53X40
6	Cecropin-B	P04142
7	Cecropin-D	O76146
8	Cecropin-E	Q308S4
9	Cecropin	P14666
10	Cecropin-A	Q27239
11	Cecropin CBM2-2	Q9GSH0
12	Cecropin CBM2	Q9GSH1
13	Cecropin CBM1	Q9GSH2
14	Cecropin-like	H9IS00
15	Antibacterial peptide	Q2WGL2
16	Antibacterial peptide enbocin precursor	P48821
17	Enbocin2	A4PBJ7
18	Enbocin3	A5LHW2
19	Lebocin-4 precursor	O15946
20	Lebocin-3 precursor	P55796
21	Lebocin-1/2 precursor	P54684
22	Moricin B3	G3D628

23	Moricin 2	O96059
24	BM Moricin	P82818
25	Putative defense protein	Q008X1
26	Defensin like protein 2	B5MF85
27	Defensin A	G0T497
28	Defensin-like protein	Q45RF8
29	Attacin	Q26431
30	Attacin	H9IZR2
31	Attacin	D2XRA5
32	Attacin like protein	D0PWZ4
33	Uncharacterised protein	H9IZQ4
34	Transferrin	O97158
35	Lysozyme	P48816

The propeptide and mature chain lengths were obtained from Uniprot database under the option ‘PTM/processing’. The segregation of the sequence was conducted and stored systematically in an Excel worksheet.



Name	UniProt Accession	CAMP	UniProt	Molecular Processing	Signal Peptide	Main Chain
Cecropin-A	Q27239	RWKLFKKIEKVGRNVRDGLIKAGPAIAVIG	MNFVRILSFV FALVLALGAV SAAPEPRWKL FKKIEKVGRN VRDGLIKAGP AIAVIGQAKS LGK	Signal Peptide: 1-22 Propeptide: 23-26 Chain: 27-61	MNFVRILSFV FALVLA	RWKL FKKIEKVGRN
Cecropin A	Q53X40		MNFVRILSFVFALVLALGAVSAAPEPRW KLFKKIEKVGRNVRDGLIKAGPAIAVIGQ AKSLGK	Signal Peptide: 1-22 Chain: 23-63	MNFVRILSFVFALVLAL	APEPRWKLFKKIEKV
Cecropin-B	P04142	RWKIFKKIEKMGRNIRDGIVKAGPAIEVLG	MNFAKILSFV FALVLALSMT SAAPEPRWI	Signal Peptide: 1-22 Propeptide: 23-26 Chain: 27-61	MNFAKILSFV FALVLA	APEPRWKI FKKIEKM
Cecropin-D	O76146	GNFFKDLEKMGQRVRDAVISAAPAVDTLA	MKFSKIFVFV FAIVFATASV SAAPGNFFK	Signal Peptide: 1-22 Propeptide:23-24 Chain: 25-60	MKFSKIFVFV FAIVFAT	GNFFKD LEKMGQRV
Cecropin-E	Q308S4	MNFSRALFYVFAVFLVCASVMAAPEPRW	MNFSRALFYV FAVFLVCASV MAAPEPRV	Signal Peptide: 1-22 Chain: 23-65	MNFSRALFYV FAVFLV	APEPRWKI FKKIEKV
Cecropin CBM2-2	Q9GSH0	MNFAKILSFVFALVLALSMTSAAPEPRWK	MNFAKILSFV FALVLALSMT SAAPEPRWKI FKKIEKMGWN IRDGIVKAGP AIEVLGSAKA VGK	Signal Peptide: 1-22 Chain: 23-63	MNFAKILSFV FALVLA	APEPRWKI FKKIEKM
Cecropin CBM2	Q9GSH1	MNFAKILSFVFALVLALSMTSAAPEPRWK	MNFAKILSFV FALVLALSMT SAAPEPRWKI FKKIEKMGRN IRDGIVKAGP AIEVLGSAKA VGK	Signal Peptide: 1-22 Chain: 23-63	MNFAKILSFV FALVLA	APEPRWKI FKKIEKM
Cecropin CBM1	Q9GSH2	MNFVRILSFVFALVLALGAVSAAPEPRWK	MNFVRILSFV FALVLALGAV SAAPEPRWKL FKKIEKVGRK VRDGLIKAGP AIAVIGQAKF IGK	Signal Peptide: 1-22 Chain: 23-63	MNFVRILSFV FALVLA	APEPRWKL FKKIEKV

Gloverin | 1 Cecropin | Enbocin | Lebocin | Attacin | Moricin | Defensin | Others | Pape

Figure 5.1 c: Snapshot of Excel worksheet showing sorted sequence of AMPs

5.1.1 Full length sequence of AMPs

All 35 AMPs have signal peptide and mature AMP. Cecropin A, Cecropin B, Cecropin D, Attacin and Enbocin 1 show the presence of propeptides of one to eight amino acids in length whereas Lebocins have propeptides on either side of the mature AMP.

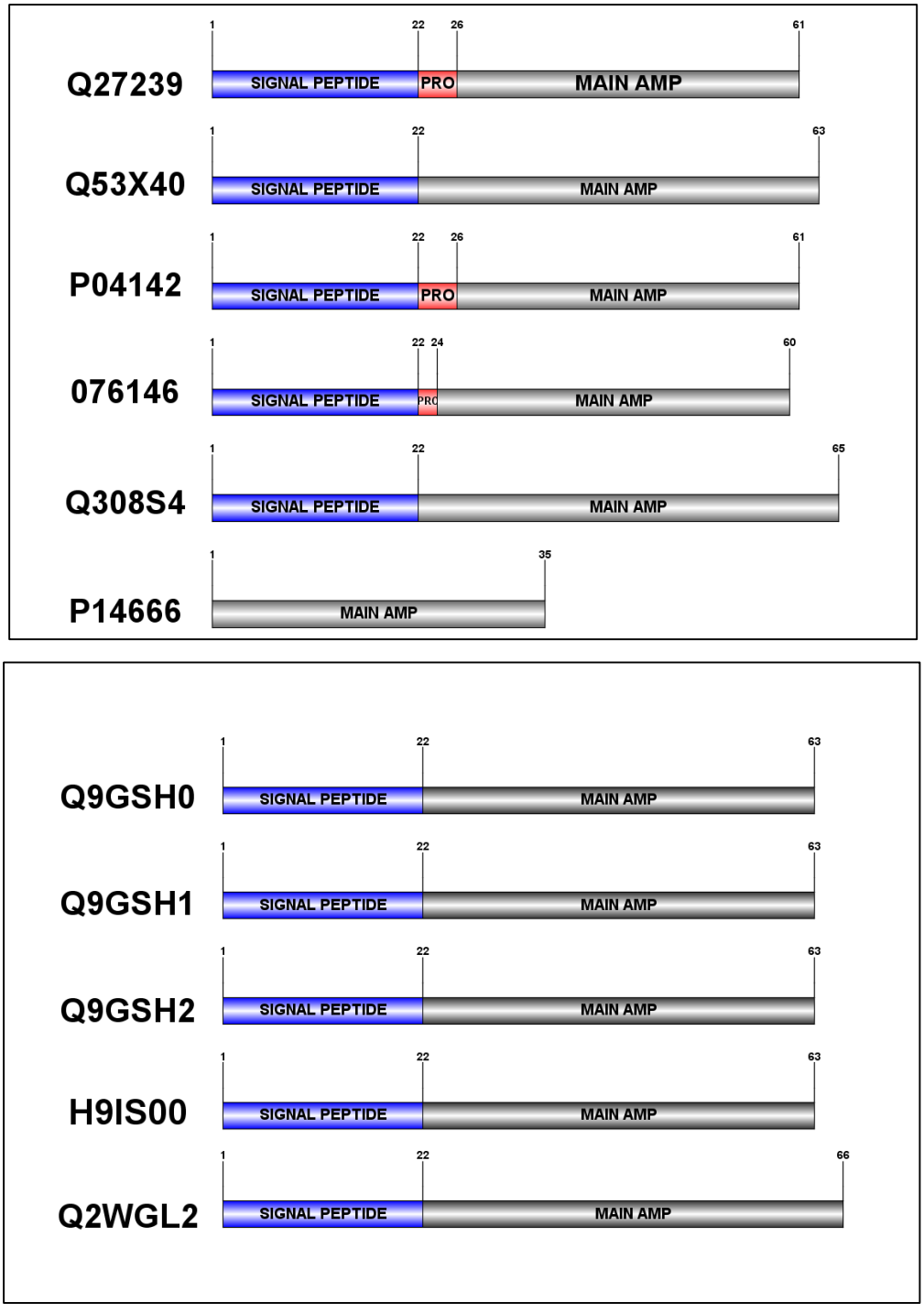


Figure 5.1.1. 1: Schematic diagram representing the signal peptide, propeptide and mature AMP in Cecropin

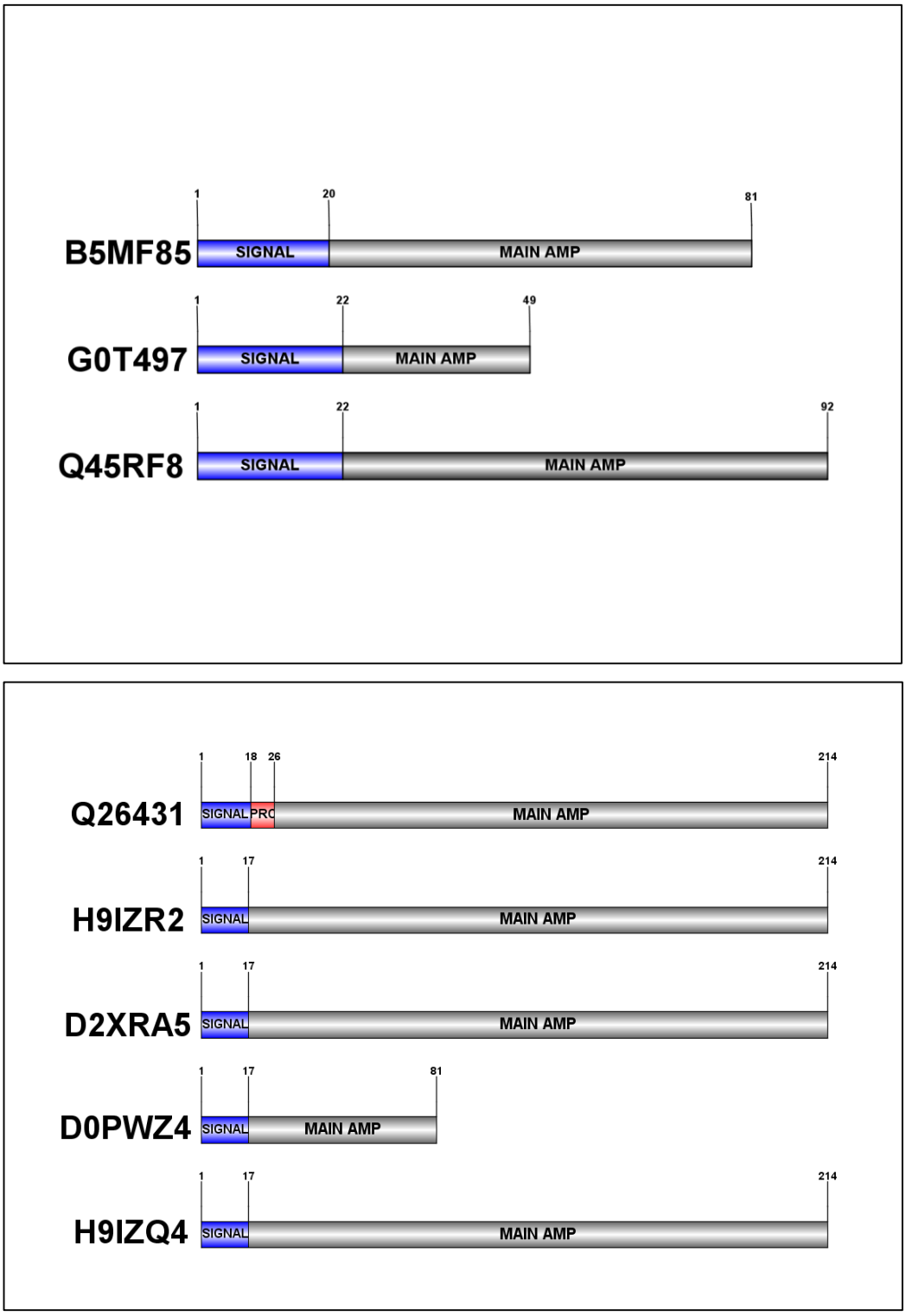


Figure 5.1.1. 2: Schematic diagram representing the signal peptide, propeptide and mature AMP in Defensin and Attacin respectively

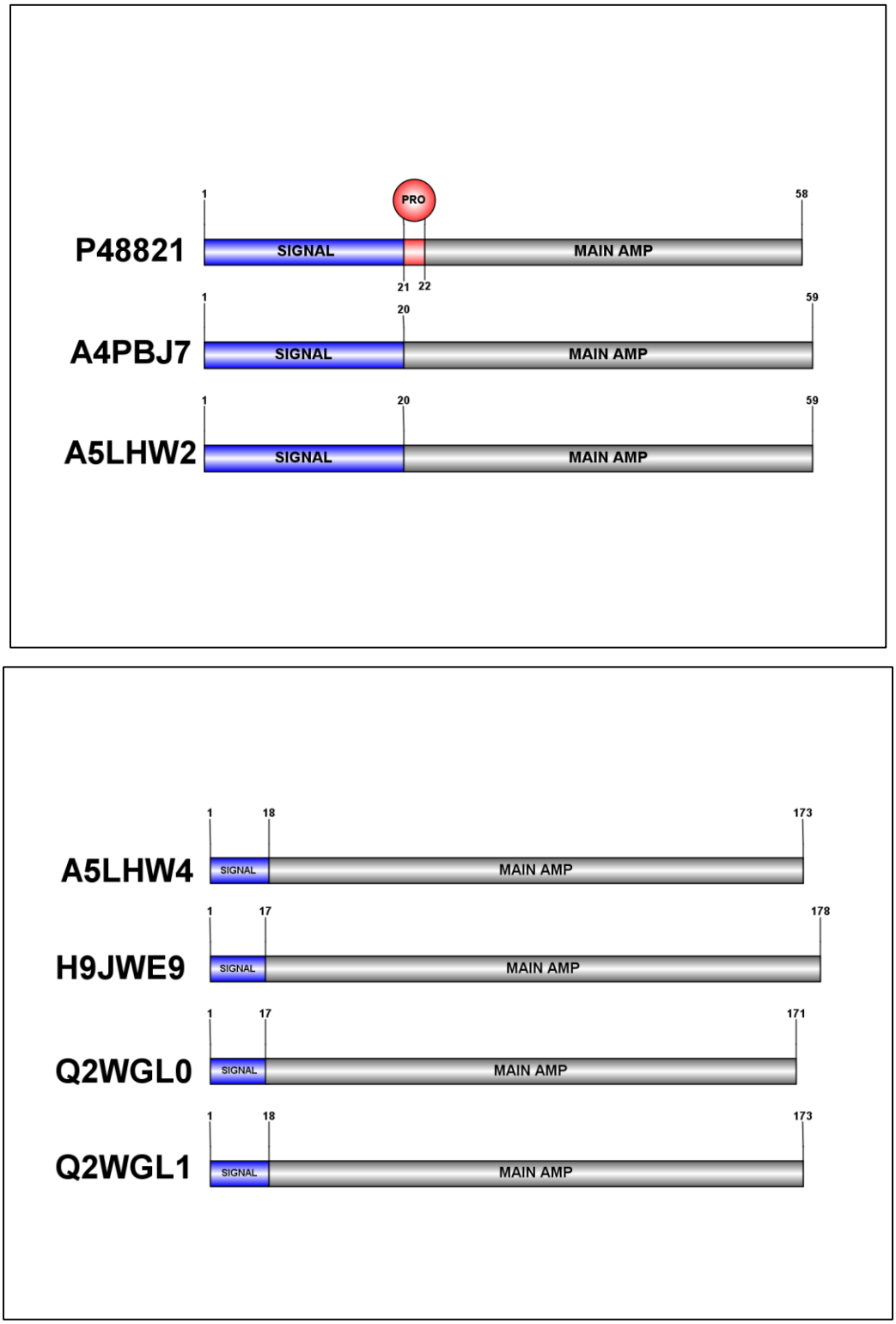


Figure 5.1.1. 3: Schematic diagram representing the signal peptide, propeptide and mature AMP in Enbocin and Gloverin respectively

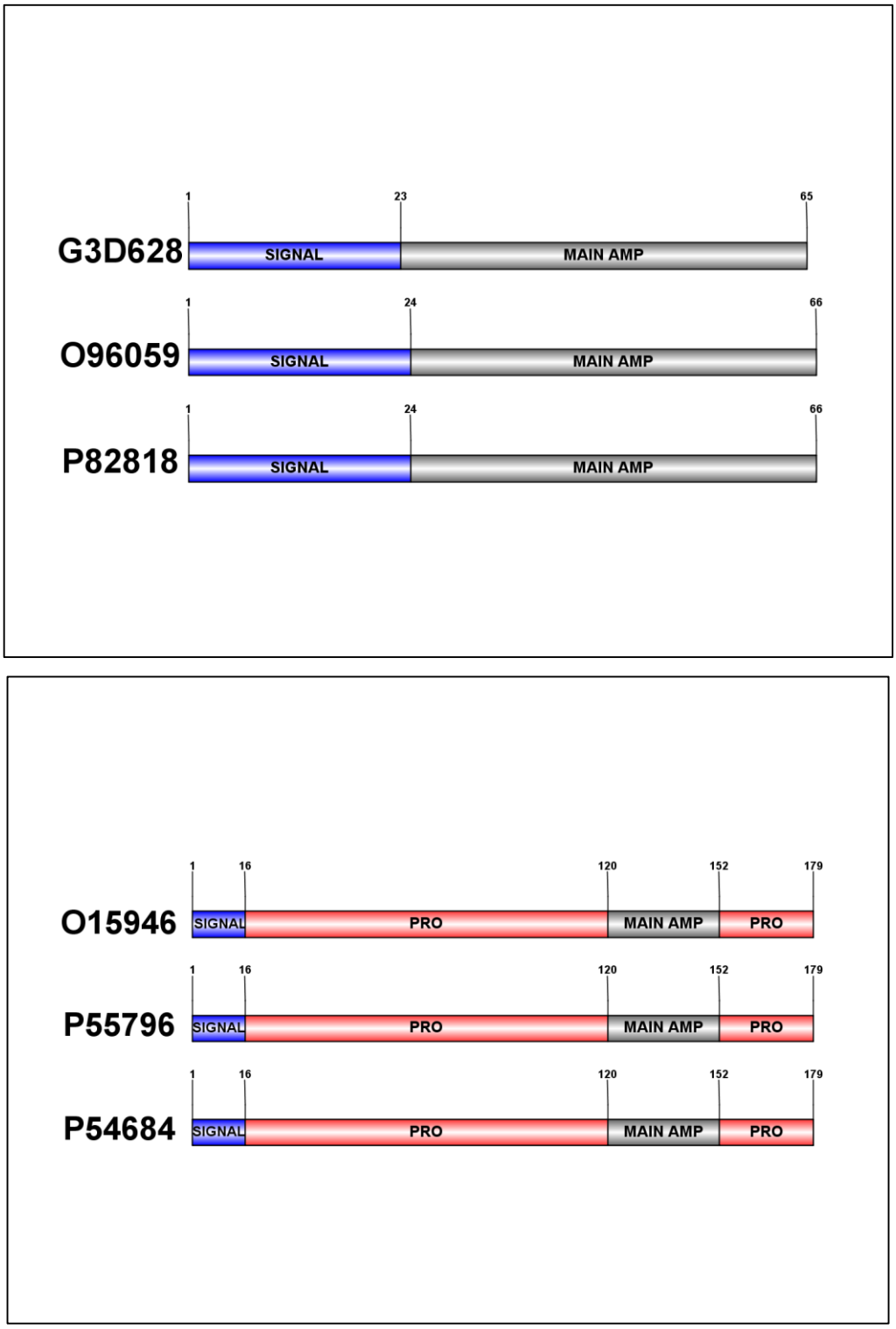


Figure 5.1.1. 4: Schematic diagram representing the signal peptide, propeptide and mature AMP in Moricin and Lebocin respectively

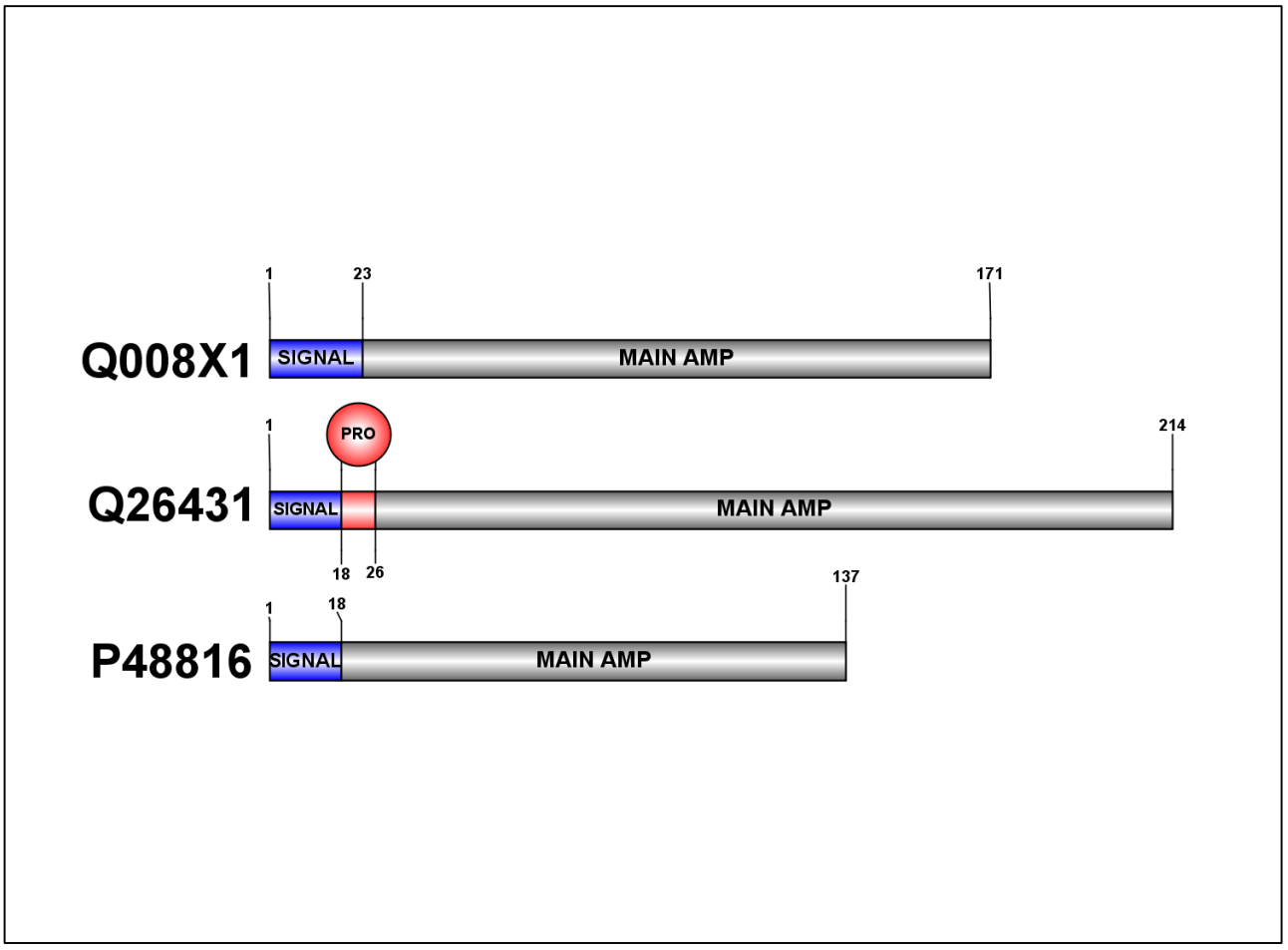


Figure 5.1.1. 5: Schematic diagram representing the signal peptide, propeptide and mature AMP grouped under others

5.2 Primary sequence analysis

ExPASy is the SIB Bioinformatics tool which provides access to scientific databases and software tools in different areas such as proteomics, genomics, phylogeny etc. ProtParam is a tool which allows the computation of various physical and chemical parameters for a given protein stored in Swiss-Prot or TrEMBL. The computed parameters include molecular weight, theoretical pI, amino acid composition, instability index, aliphatic index, grand average of hydropathicity (GRAVY).

Table 5. 2: Physicochemical properties of all 35 AMPs

Sl. No.	Name of the AMP	UniProt Accession number	No. of AA in main	Molecular weight	Theoretical pI	Total no. of negatively charged residues (Asp + Glu)	Total no. of positively charged residues (Arg + Lys)	Net Charge	Instability Index	Aliphatic Index	Grand average of hydropathicity (GRAVY)
1	Gloverin 3	A5LHW4	155	17080.76	8.26	20	21	1.9	26.42	56.65	-0.877
2	Uncharacterised protein	H9JWE9	161	17345.1	6.57	19	18	0.1	20.92	59.38	-0.665
3	Gloverin-4	Q2WGL0	153	16761.33	6.36	21	20	-0.3	20.95	54.84	-0.861
4	Gloverin-2	Q2WGL1	155	16817.32	6.48	20	19	-0.1	33.84	51.61	-0.905
5	Cecropin A	Q53X40	41	4441.34	10.71	3	10	6.9	37.86	100	-0.317
6	Cecropin-B	P04142	41	4474.38	10.38	4	10	5.9	68.61	92.93	-0.339
7	Cecropin-D	O76146	37	3874.47	9.52	4	6	1.9	10.41	84.59	-0.141
8	Cecropin-E	Q308S4	43	4564.4	10.29	3	8	5.2	35.57	97.67	-0.135
9	Cecropin	P14666	35	3762.5	10.56	2	7	4.9	34.59	111.43	0.129

10	Cecropin-A	Q27239	35	3861.68	11.17	2	9	6.9	31.24	114.29	-0.109
11	Cecropin CBM2-2	Q9GSH0	41	4490.38	10.12	4	9	4.9	69.55	90.49	-0.259
12	Cecropin CBM2	Q9GSH1	41	4460.35	10.38	4	10	5.9	66.54	90.49	-0.346
13	Cecropin CBM1	Q9GSH2	41	4515.5	10.77	3	11	7.9	32.78	100	-0.222
14	Cecropin-like	H9IS00	41	4217.09	11.07	2	7	4.9	27.75	116.34	0.461
15	Antibacterial peptide	Q2WGL2	44	4850.37	4.46	10	6	-4.1	35.42	86.36	-0.5
16	Antibacterial peptide enbocin precursor	P48821	36	3795.36	10.67	3	5	1.9	43.29	97.78	0.261
17	Enbocin2	A4PBJ7	38	3899.47	9.98	3	5	1.9	19.56	87.63	0.321
18	Enbocin3	A5LHW2	38	3929.49	9.98	3	5	1.9	24.03	87.63	0.311
19	Lebocin-4 precursor	O15946	32	3899.57	9.52	3	5	1.9	35.48	73.12	-0.825
20	Lebocin-3 precursor	P55796	32	3789.5	9.82	2	5	2.9	37.09	82.19	-0.5
21	Lebocin-1/2 precursor	P54684	32	3773.46	9.82	2	5	2.9	37.09	70	
22	Moricin B3	G3D628	42	4239.89	9.87	3	7	4.4	10.7	85.95	-0.383
23	Moricin 2	O96059	42	4543.52	11.36	1	11	10.1	8.32	100	-0.21
24	BM Moricin	P82818	70	7755.69	4.74	11	6	10.1	34.52	72.43	-0.274
25	Putative defense protein	Q008X1	148	16024.22	8.64	15	17	2.6	54.05	81.62	-0.335

26	Defensin like protein 2	B5MF85	61	7128.22	8.88	8	12	3.7	69.76	57.54	-0.675
27	Defensin A	G0T497	27	2975.22	4.83	5	3	-1.8	50.28	72.22	-0.719
28	Defensin-like protein	Q45RF8	70	7755.69	4.74	11	6	-4.8	34.52	72.43	-0.274
29	Attacin	Q26431	187	19635.77	9.44	13	17	5.1	24.79	70	-0.247
30	Attacin	H9IZR2	197	20908.25	9.99	13	21	9.4	32.11	68.93	-0.34
31	Attacin	D2XRA5	197	20942.27	9.99	13	21	9.4	31.77	66.95	-0.352
32	Attacin like protein	D0PWZ4	64	6575.28	10.25	5	8	3.4	23.25	80.94	-0.331
33	Uncharacterised protein	H9IZQ4	197	20894.22	9.87	13	20	8.4	32.87	69.44	-0.337
34	Transferrin	O97158	452	50714.18	7.85	56	58	0.7	34.09	79.58	-0.346
35	Lysozyme	P48816	119	13751.52	9.06	14	21	7.9	22.54	53.28	-0.903

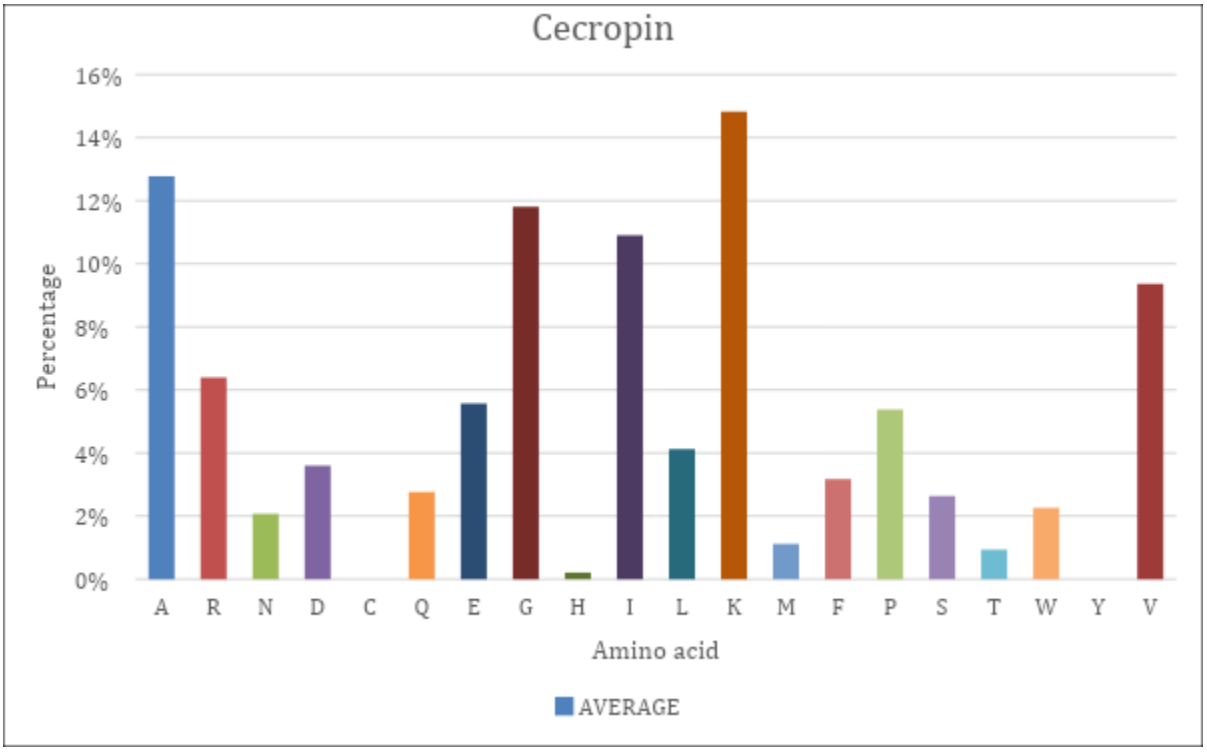


Figure 5.2 a: Graphical representation of amino acid propensity for Cecropin

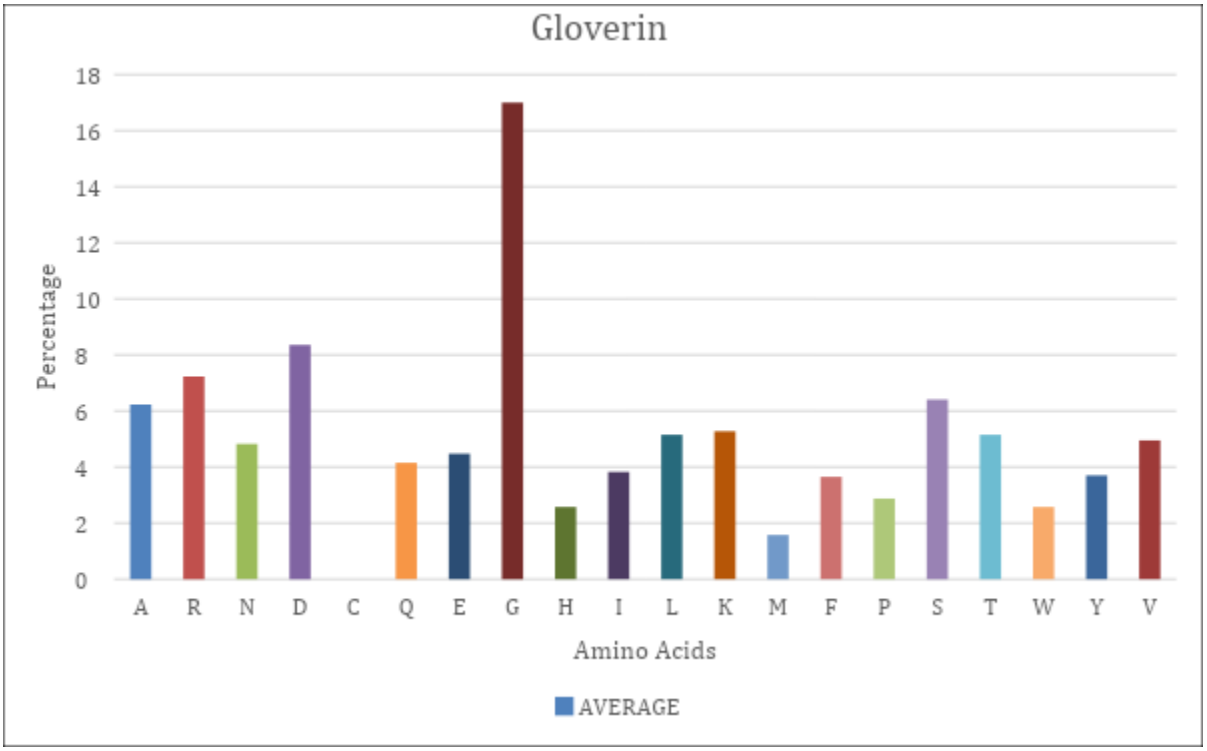


Figure 5.2 b: Graphical representation of amino acid propensity for Gloverin

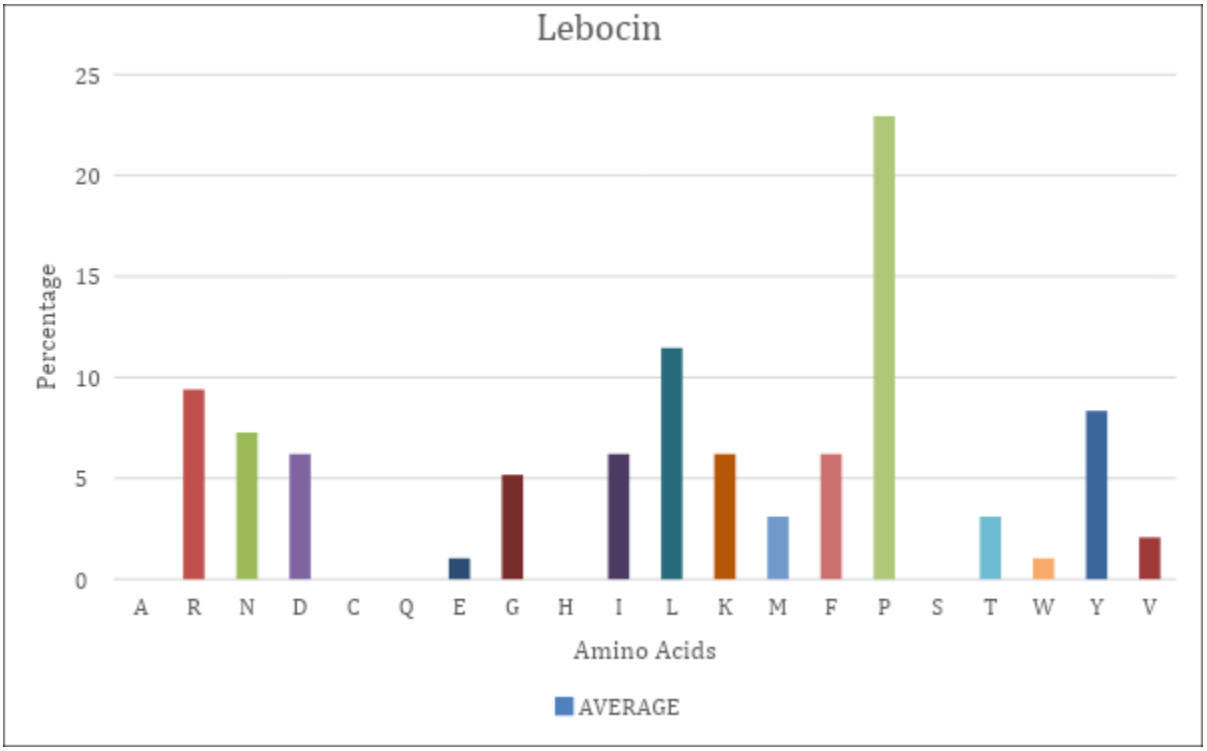


Figure 5.2 c: Graphical representation of amino acid propensity for Lebocin

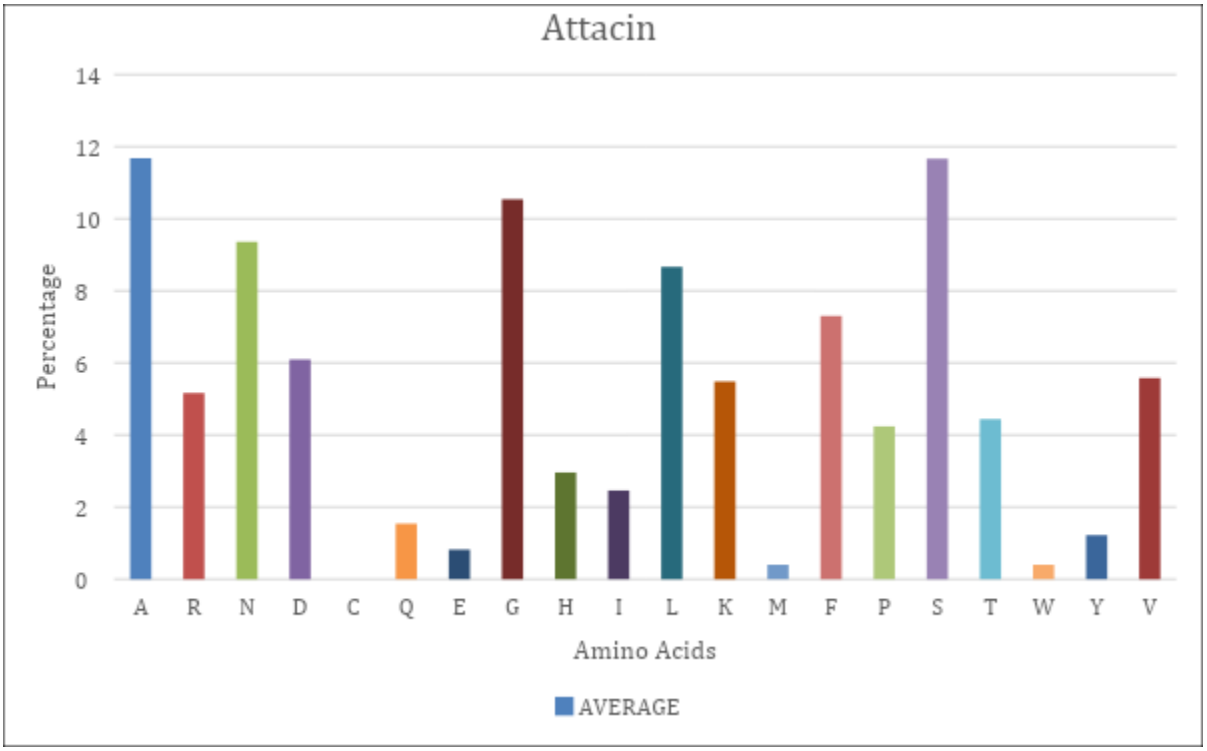


Figure 5.2 d: Graphical representation of amino acid propensity for Attacin

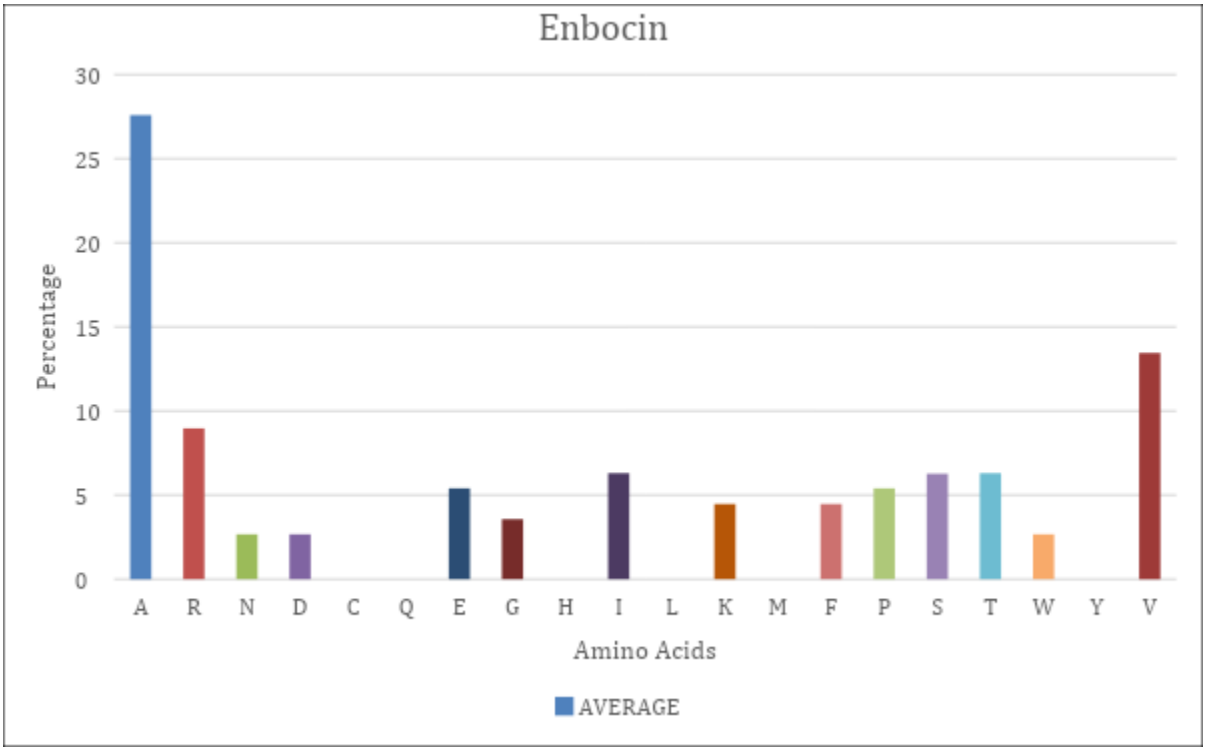


Figure 5.2 e: Graphical representation of amino acid propensity for Enbocin

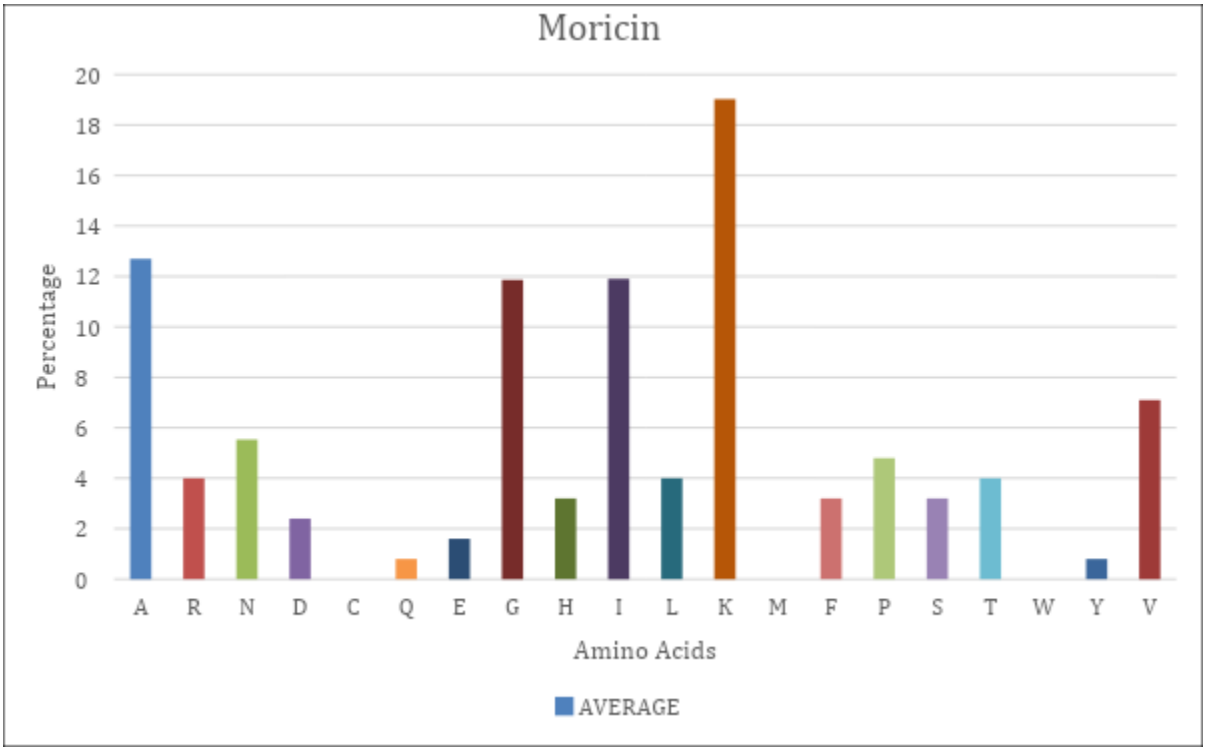


Figure 5.2 f: Graphical representation of amino acid propensity for Moricin

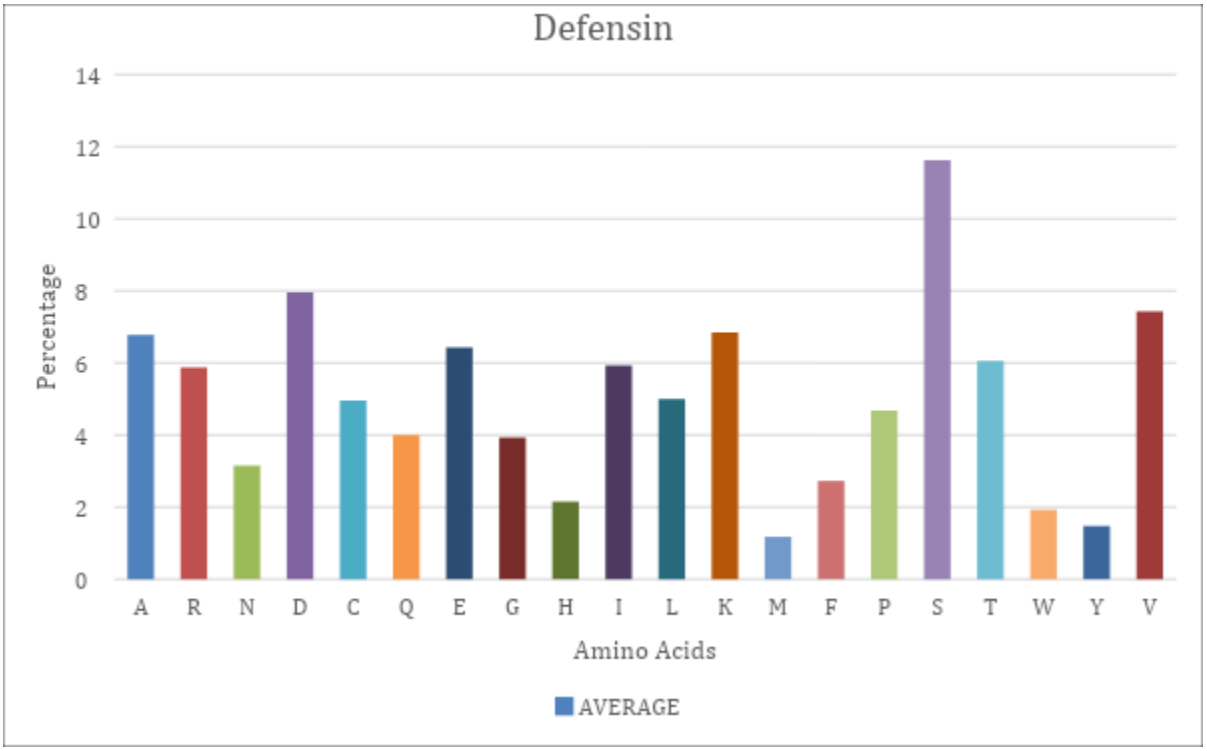


Figure 5.2 g: Graphical representation of amino acid propensity for Defensin

5.2.1 Amino acid patterns

5.2.1.1 Cecropin

The lengths of the mature peptide vary from 35 to 44 amino acid residues, with the molecular weight ranging from 3.7 kDa to 4.8 kDa. The composition of Lysine was found to be 14.82%, which was the highest followed by Alanine (13%) and Glycine (11.81%). No Cysteine, Methionine, Histidine, Threonine and Tyrosine residues are present in Cecropins. Absence of disulphide bonds are observed due to the absence of cysteine residues.

5.2.1.2 Gloverin

The lengths of the mature peptide vary from 153 to 161 amino acid residues, with the molecular weight ranging from 16.8 kDa to 17.3 kDa. The composition of Glycine was found to be 17%, which was the highest followed by Aspartic acid (8.35%) and Arginine (7.2%).

No Cysteine residues were found to be present in gloverins indicating an absence of disulphide bonds.

5.2.1.3 Lebocin

The mature peptide contains 32 amino acid residues, with the molecular weight ranging from 3.7 kDa to 3.8 kDa. Lebocin can be classified as Pro-rich family because it has 23% Proline followed by Leucine (11.5%) and Arginine (9.4%). No Alanine, Cysteine, Glutamine, Histidine and Serine residues were found to be present in Lebocins. Absence of disulphide bonds are observed due to the absence of cysteine residues.

5.2.1.4 Attacin

The lengths of the mature peptide vary from 185 to 200 amino acid residues, with the molecular weight ranging from 19.6 kDa to 21 kDa, with the exception of Attacin like protein (D0PWZ4) that has 64 amino acid residues and weighs 6.5kDa. The composition of Alanine and Serine was found to be 11.7%, which was the highest followed by Glycine (10.5%). No Cysteine residues were found to be present in attacins indicating an absence of disulphide bonds.

5.2.1.5 Enbocin

The lengths of the mature peptide vary from 36 to 38 amino acid residues, with the molecular weight ranging from 3.8 kDa to 4 kDa. Enbocin can be classified as Ala-rich family because it has 27.6% Alanine followed by Valine (13.4%) and Arginine (9%). No Cysteine, Glutamine, Histidine, Leucine, Methionine and Tyrosine residues are present in Enbocins. Absence of disulphide bonds are observed due to the absence of cysteine residues.

5.2.1.6 Moricin

The mature peptide contains 42 amino acid residues, with the molecular weight ranging from 4.2 kDa to 4.5 kDa except BM Moricin (P82818) which has 70 amino acid residues and weighs 7.7kDa. The composition of Lysine was found to be 19%, which was the highest followed by Alanine (12.7%) and Glycine and Isoleucine (11.9%). No Cysteine,

Methionine and Tryptophan residues were found to be present in moricins. Absence of disulphide bonds are observed due to the absence of cysteine residues.

5.2.1.7 Defensin

The lengths of the mature peptide vary from 30 to 70 amino acid residues, with the molecular weight ranging from 3 kDa to 7.7 kDa except Putative defense protein (Q008X1) which has 148 amino acid residues and weighs 16 kDa. The composition of Serine was found to be 11.6%, which was the highest followed by Aspartic acid (7.95%) and Valine (7.5%). Presence of disulphide bonds are observed due to the presence of cysteine residues.

5.2.2 Isoelectric point (pI)

Isoelectric point (pI) is the pH where there is zero net charge on an AMP; hence it will affect the solubility of AMP as it precipitates and also loses its biological functions. Isoelectric point of majority of AMPs is found to be near to pH 10, similar to detergents or emulsifying agents like soaps, which assists the mechanism of action for interacting with lipid bilayer of biological membrane. But, Gloverin 2 (Q2WGL1), Gloverin 4 (Q2WGL0), and uncharacterized protein (H9JWE9) have ~ pH 6.5; Cecropin antibacterial peptide (Q2WGL2), BM Moricin (P82818), Defensin A (G0T497) and Defensin like protein (Q45RF8) have ~ pH 4.5.

5.2.3 Instability Index

The instability index is the estimate of stability of AMP in a test tube. Instability index <40 is an indication of the stability of AMPs. Of the 35 AMPs, Cecropin B (P04142), Cecropin CBM2-2 (Q9GSH0), Cecropin CBM2 (Q9GSH1), Antibacterial peptide enbocin (P48821), Putative defense protein (Q008X1), Defensin- like protein 2 (B5MF85) and Defensin A (G0T497) are unstable.

5.2.4 Aliphatic Index

Relative volume occupied towards aliphatic side chains (Alanine, Valine, Isoleucine and Leucine) of AMP is aliphatic index. Aliphatic index plays positive role in the great thermal stability of globular protein. Positive aliphatic index indicates enhanced thermostability of globular protein. Aliphatic index >70 were showed in most AMPs and indicates stability for wider range of temperature, except Gloverin 2 (Q2WGL1), Gloverin 3 (A5LHW4), Gloverin 4 (Q2WGL0), Uncharacterized protein (H9JWE9), Defensin- like protein 2 (B5MF85), Attacin (H9IZR2), Attacin (D2XRA5), Uncharacterized protein (H9IZQ4) and Lysozyme (P48816).

5.2.5 GRAVY

Grand Average of Hydropathy (GRAVY) value of AMP is calculated as the sum of hydropathy values of all the amino acids, divided by the number of amino acid residues in the complete sequence. Positive and negative GRAVY is indication of hydrophobicity and hydrophilicity respectively. Of the 35 AMPs, most are hydrophilic, except Cecropin (P14666), Cecropin-like (H9IS00), Antibacterial peptide enbocin (P48821), Enbocin 2 (A4PBJ7), Enbocin 3 (A5LHW2) are found to be hydrophobic.

5.2.6 Net Charge

Sum of ionizable amino acid residues at particular pH will give rise to anionic or cationic net surface charge to AMPs. All AMPs are cationic except Cecropin antibacterial peptide (Q2WGL2), Defensin A (G0T497) and Defensin like protein (Q45RF8) which are anionic and Gloverin 2 (Q2WGL1), Uncharacterized protein (H9JWE9) , Gloverin 4 (Q2WGL0) which are neutral.

5.2.7 Post translational modifications

From literature, it is found that O- Glycosylation takes place at the 15th position Threonine of the mature peptide of all lebocins, which plays a crucial role in its antibacterial activity. All AMPs undergo Amidation,

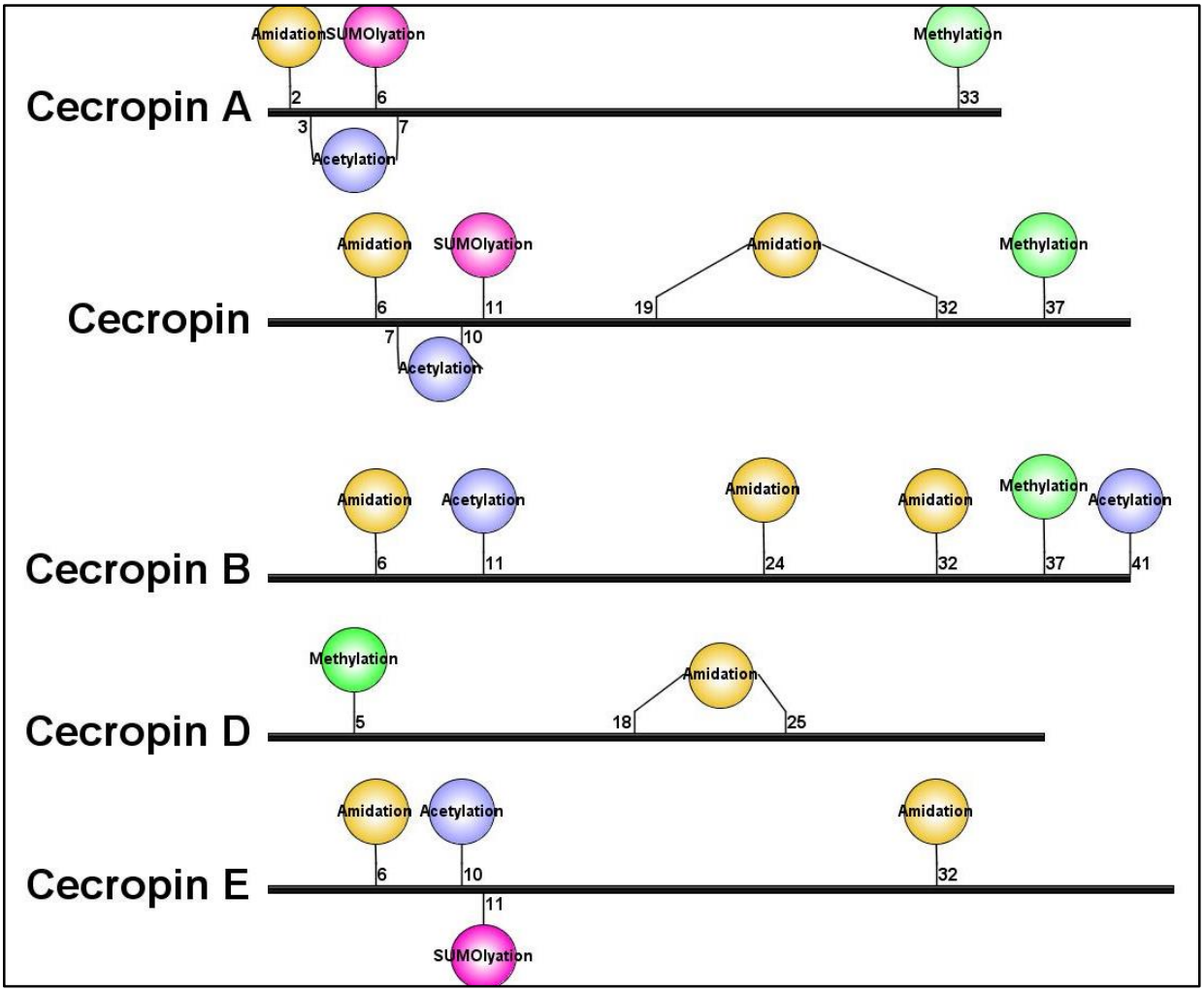


Figure 5.2.7. 1(a): Schematic representation of PTMs in Cecropin family

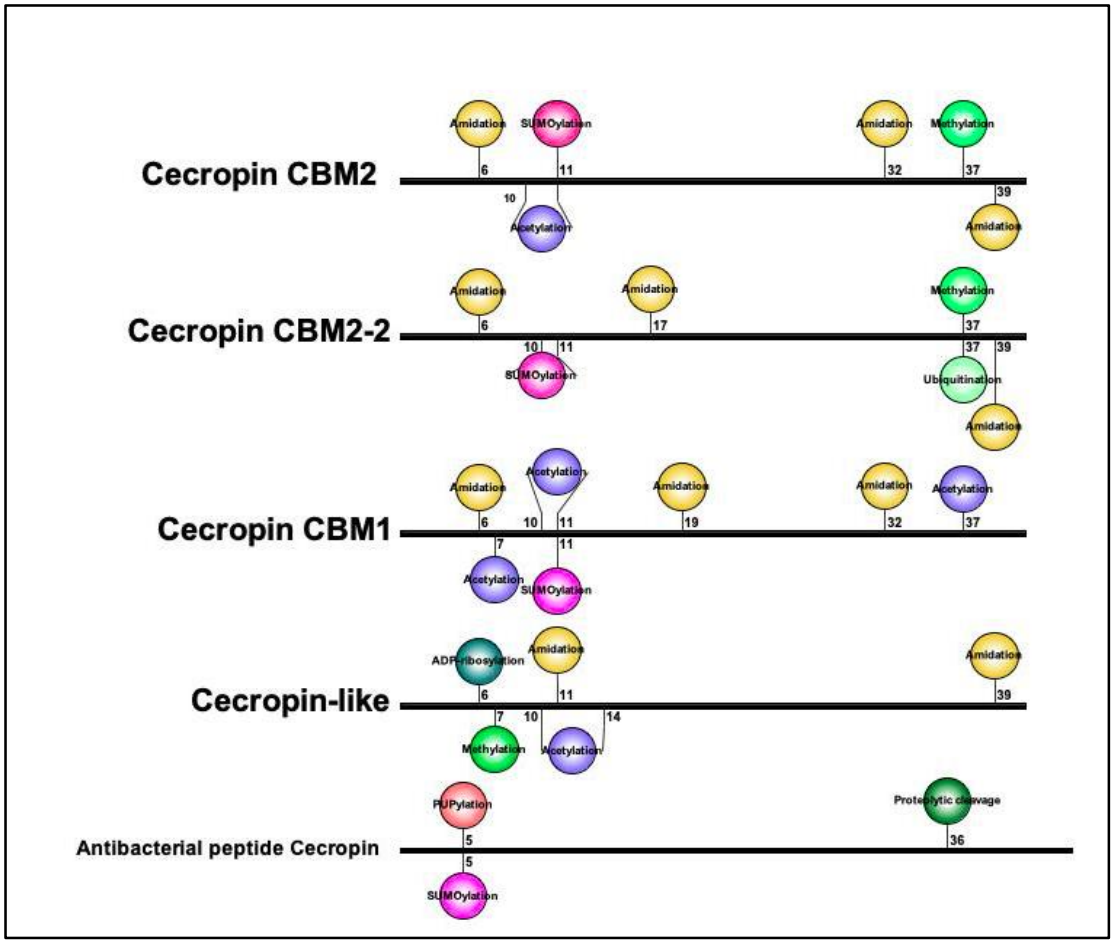


Figure 5.2.7.1 (b): Schematic representation of PTMs in Cecropin family

Most of the Antimicrobial peptides of Cecropin family undergo Amidation at positions 6, 39 and 32 and Acetylation at positions 10 and 11. Most of the AMPs undergo Amidation at more than one site.

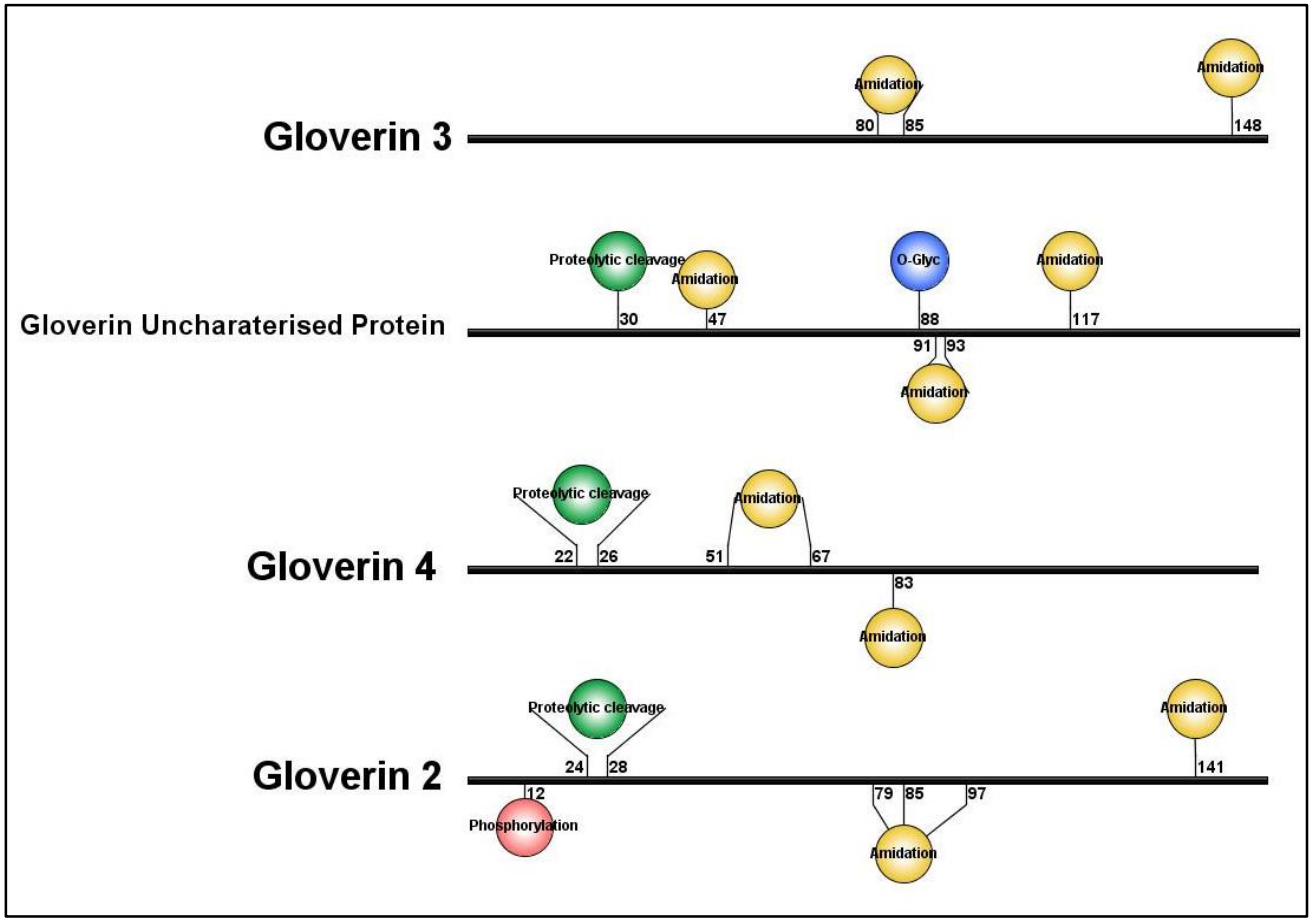


Figure 5.2.7. 2: Schematic representation of PTMs in Gloverin family

AMPs in Gloverin family, except for Gloverin 3, undergo Proteolytic cleavage between the range 22-30 amino acids and all the AMPs undergo Amidation atleast twice.

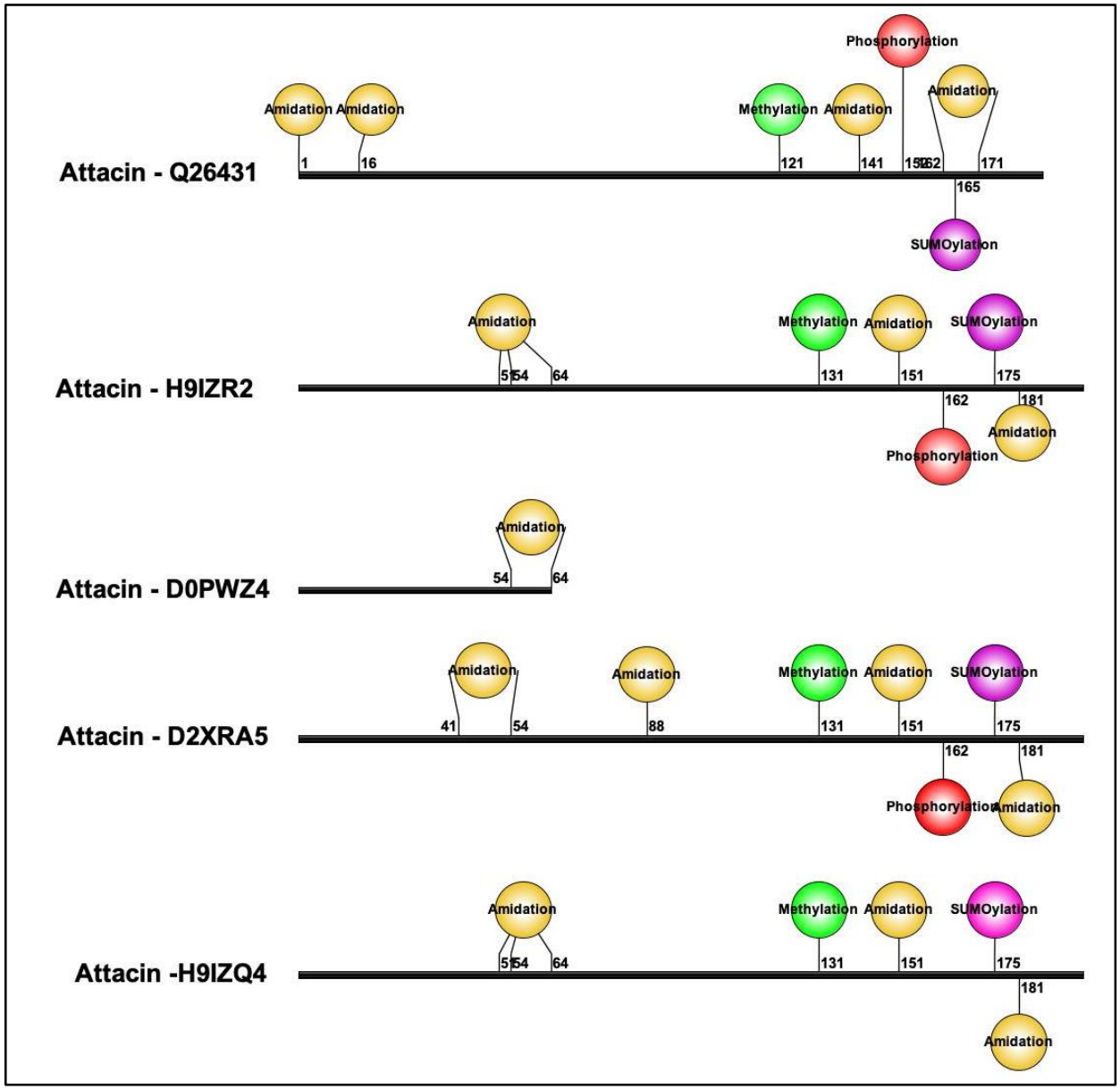


Figure 5.2.7. 3: Schematic representation of PTMs in Attacin family

Amidation is observed at the end of the main peptide for all AMPs. AMPs except Attacin (D0PWZ4) and Attacin (Q26431) undergo amidation at the site 181. SUMOlyation is seen in all the AMPs at the site 175 except for Attacin (Q26431) and is completely absent in Attacin (D0PWZ4). Attacin (H9IZR2) and Attacin (D2XRA5) undergo phosphorylation at position 162 whereas no phosphorylation is seen in Attacin (D0PWZ4) and Attacin (H9IZQ4).

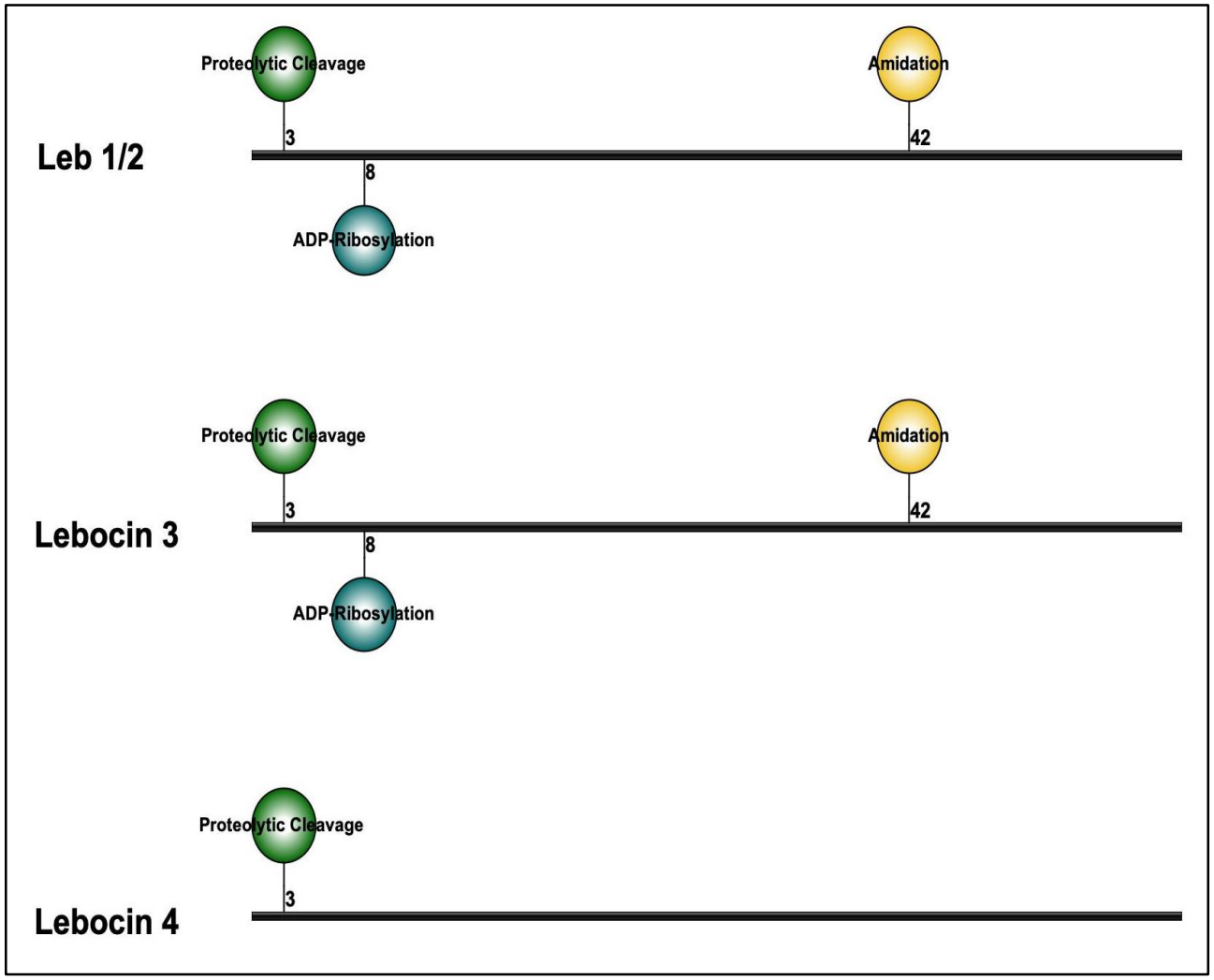


Figure 5.2.7. 4: Schematic representation of PTMs in Lebocin family

Proteolytic cleavage is seen in all the AMPs of Lebocin at position 3, amidation at position 42 for Lebocin 3 and Lebocin 4 and ADP-Ribosylation at position 8 for Lebocin ½ and Lebocin 3.

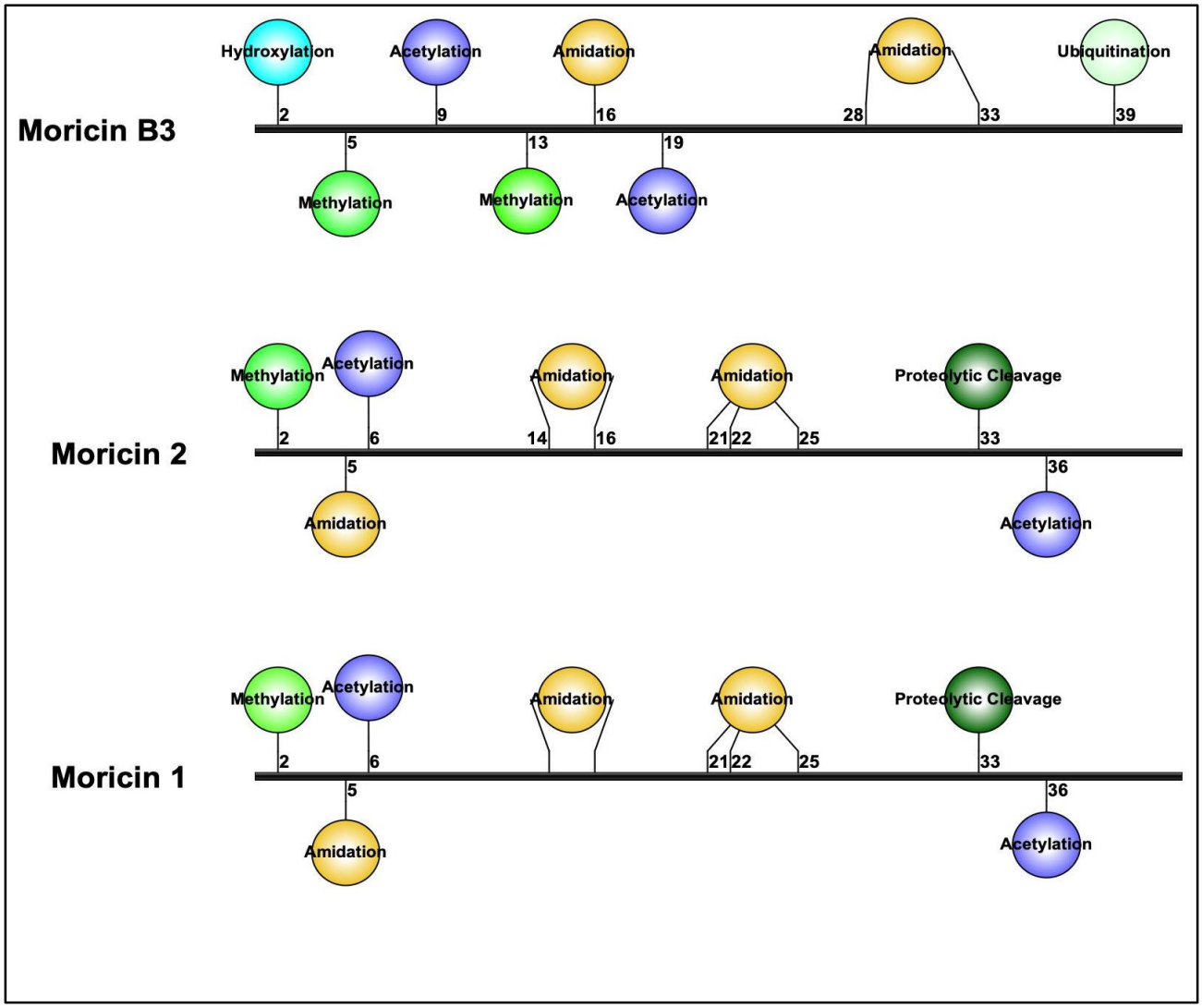


Figure 5.2.7. 5: Schematic representation of PTMs in Moricin family

Identical PTMs were observed for Moricin 1 and Moricin 2. In Moricin B3, hydroxylation and ubiquitination is observed.

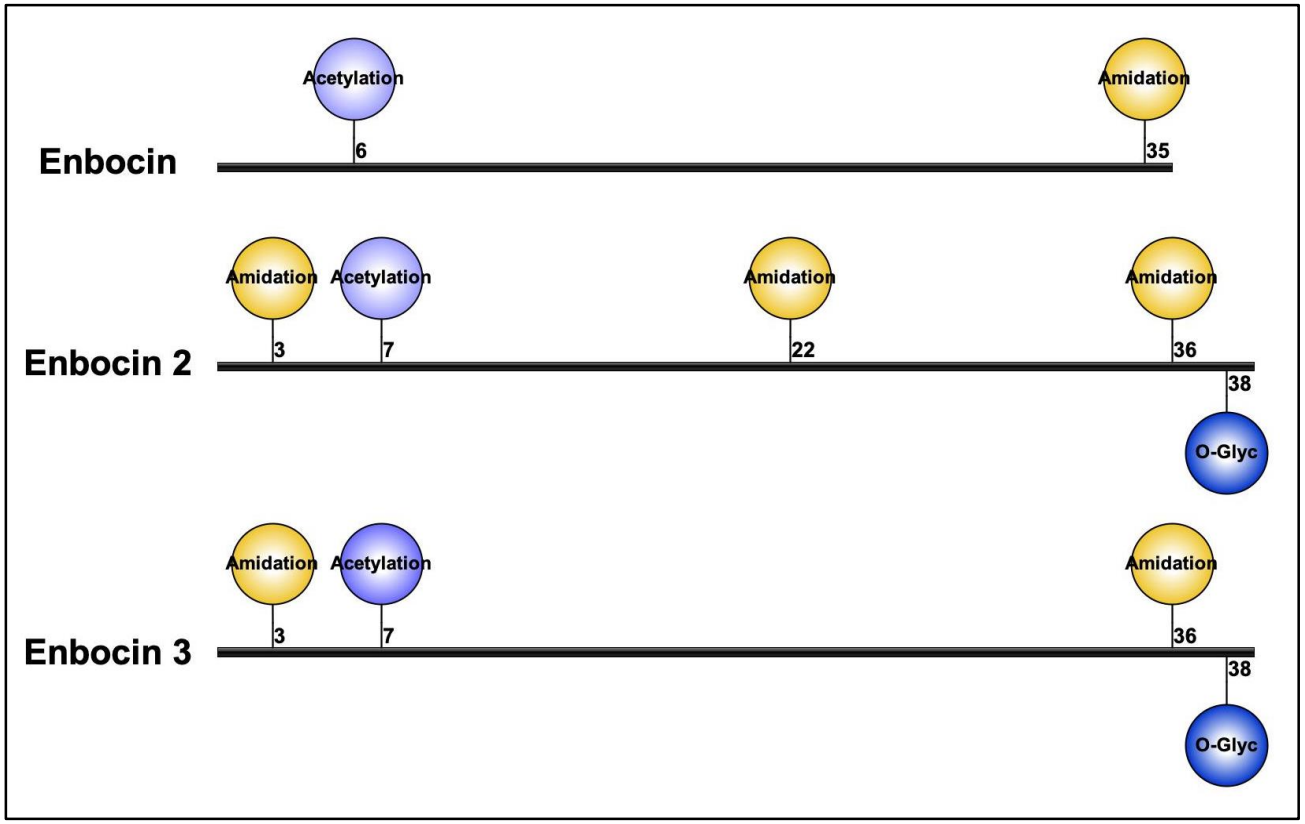


Figure 5.2.7. 6: Schematic representation of PTMs in Enbocin family

Acetylation at position 6 was observed for Enbocin and position 7 for Enbocin 2 and 3. Enbocin 2 and 3 have identical PTMs with an exception of Amidation occuring at position 22 for Enbocin 2.

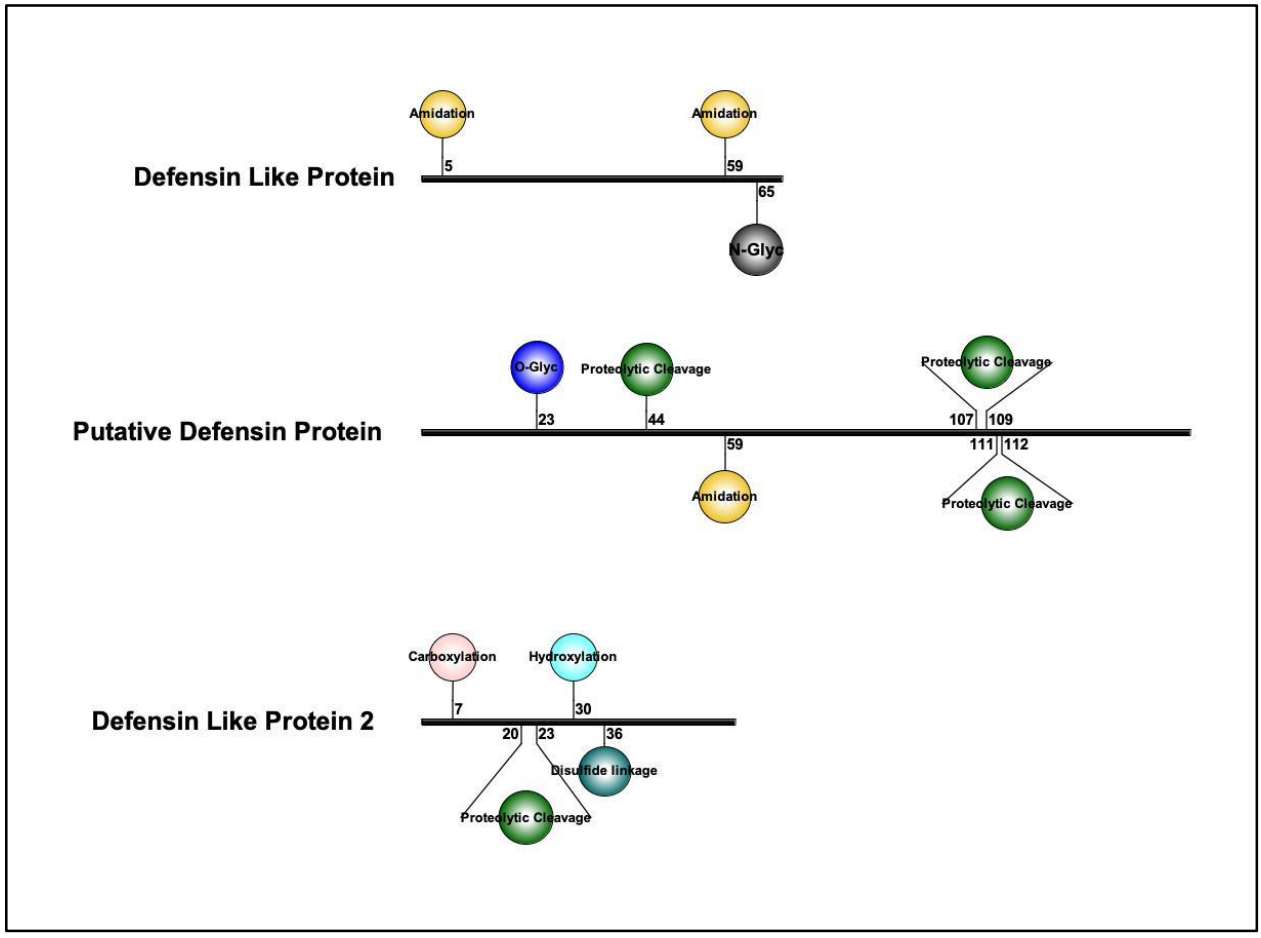


Figure 5.2.7. 7: Schematic representation of PTMs in Defensin family

Amidation was observed at position 59 for both putative and defensin like protein

5.3 Secondary Structure Prediction

The secondary structure of proteins can be broadly classified into alpha helices, beta sheets and random coils. In the result obtained from GOR 4, SOPMA, SSPro and PORTER, as shown in Table 4.3, C stands for Random coils, H is for alpha helices and E is for extended strand which is beta sheet. In cases where all tools predicted differently, weighted average method was applied to predict the final secondary structure. The predicted secondary structures were visually represented using PolyView where it was observed that Cecropins, Defensins, Lebocins and Attacins have highly coiled structures whereas Enbocins, Moricins

and Cecropins are mostly helical in nature. Transferrin is found to have helices as well as beta sheets to the same degree.

Table 5. 3: Final secondary structure tabulated using weighted average method for Cecropin A (Q27239)

Amino Acids	GOR4	SOPMA	SSPRO	PORTER	**FINAL**
R	C	H	C	C	**C**
W	C	H	H	C	**H**
K	C	H	H	C	**H**
L	C	H	H	H	**H**
F	E	H	H	H	**H**
K	C	H	H	H	**H**
K	C	H	H	H	**H**
I	C	H	H	H	**H**
E	C	H	H	H	**H**
K	C	H	H	H	**H**
V	C	H	H	H	**H**
G	C	H	C	H	**C**
R	C	H	H	H	**H**
N	C	H	H	H	**H**
V	C	H	H	H	**H**
R	H	H	H	H	**H**
D	H	H	H	H	**H**
G	H	H	H	H	**H**
L	H	H	H	H	**H**
I	H	H	H	H	**H**

K	C	H	C	H	**C**
A	C	T	C	H	**C**
G	C	C	C	H	**C**
P	C	C	C	H	**C**
A	H	C	H	H	**H**
I	H	E	H	H	**H**
A	H	E	H	H	**H**
V	H	E	H	H	**H**
I	H	E	H	H	**H**
G	H	E	H	H	**H**
Q	C	C	H	H	**H**
A	C	C	H	H	**H**
K	E	C	C	H	**C**
S	E	C	C	C	**C**
L	C	H	C	C	**C**

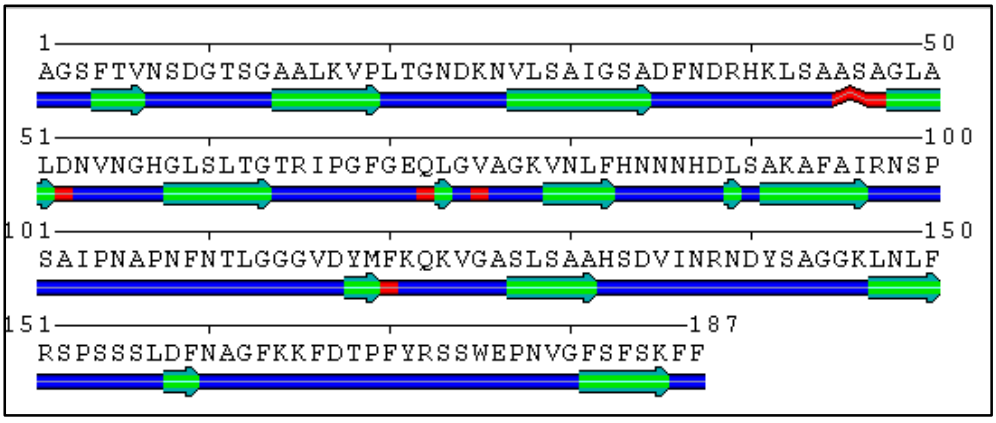


Figure 5.3.(aa): Secondary structure of Attacin-Q26431

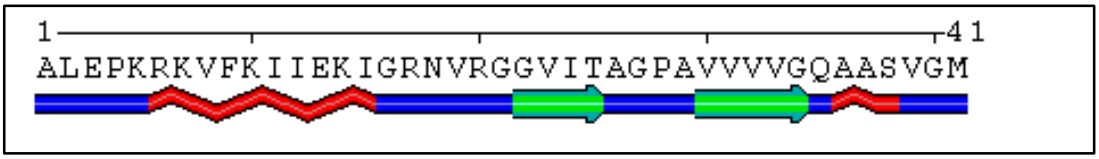


Figure 5.3(ab): Secondary structure of Attacin- H9IZR2

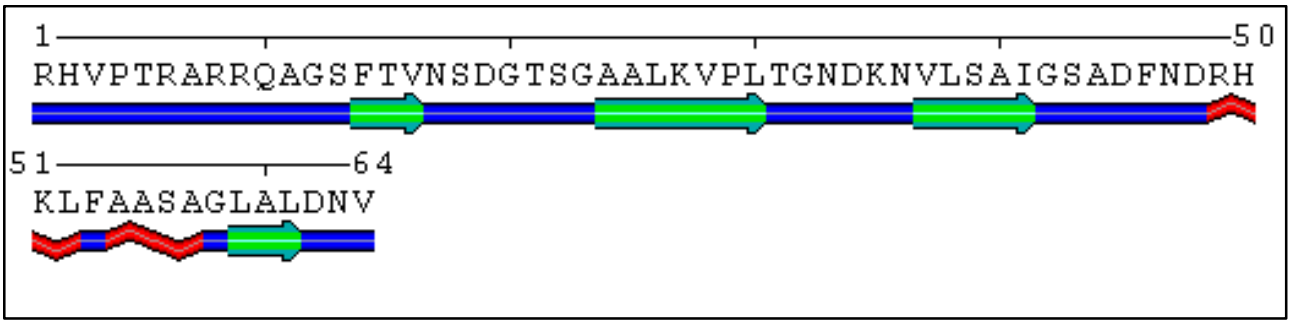


Figure 5.3 (ac): Secondary structure of Attacin- D0PWZ4

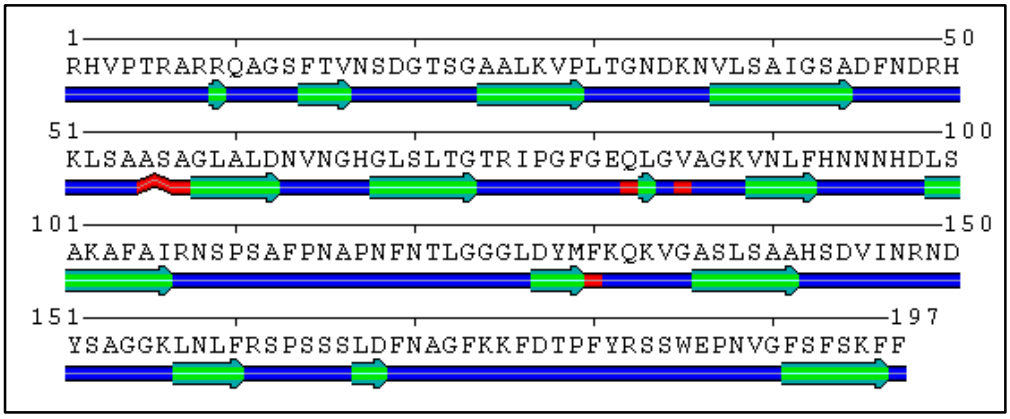


Figure 5.3 (ad): Secondary structure of Attacin- D2XRA5

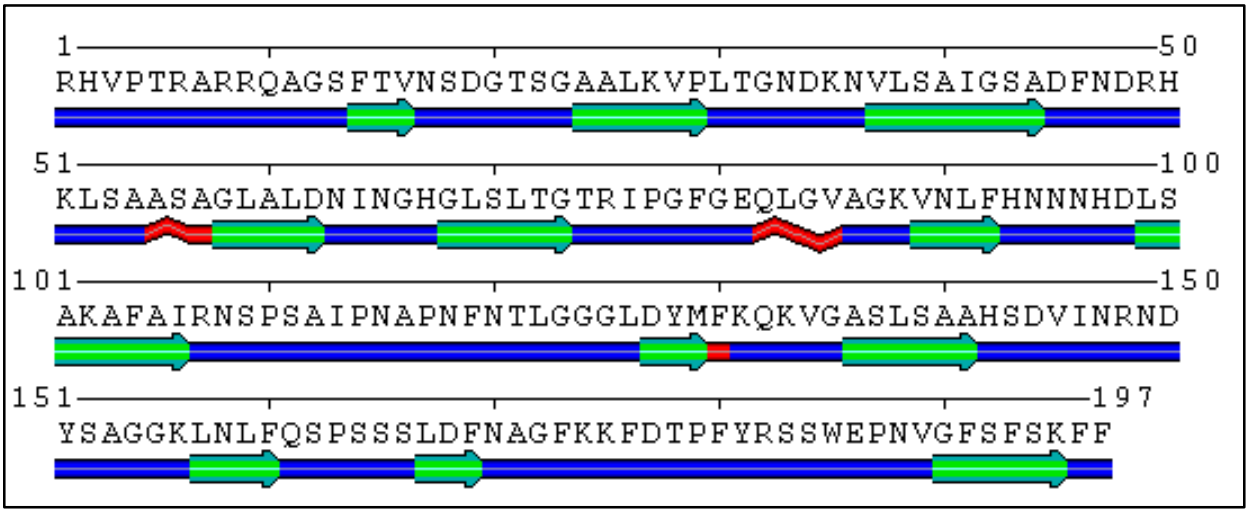


Figure 5.3 (ae): Secondary structure of Attacin- H9IZQ4

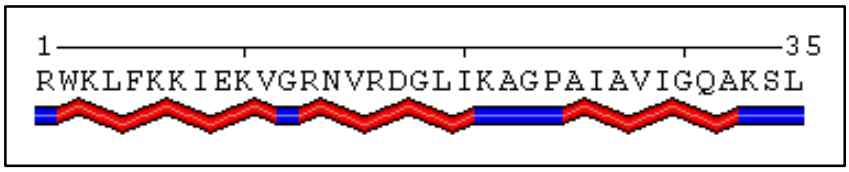


Figure 5.3 (ba): Secondary structure of Cecropin – Q27239

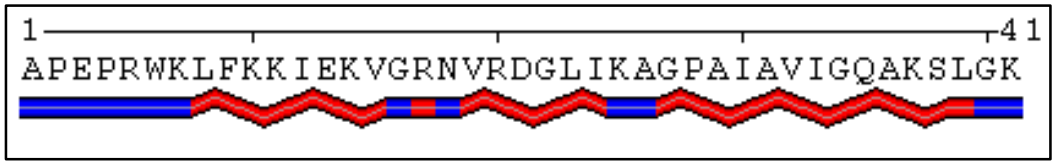


Figure 5.3 (bb): Secondary structure of Cecropin – Q53X40

Figure 5.3 (bc): Secondary structure of Cecropin – P04142

Figure 5.3 (bd): Secondary structure of Cecropin – O76146

Figure 5.3 (be): Secondary structure of Cecropin – Q308S4

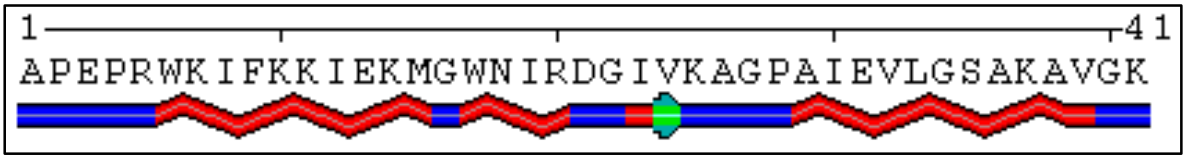


Figure 5.3 (bf): Secondary structure of Cecropin – Q53X40

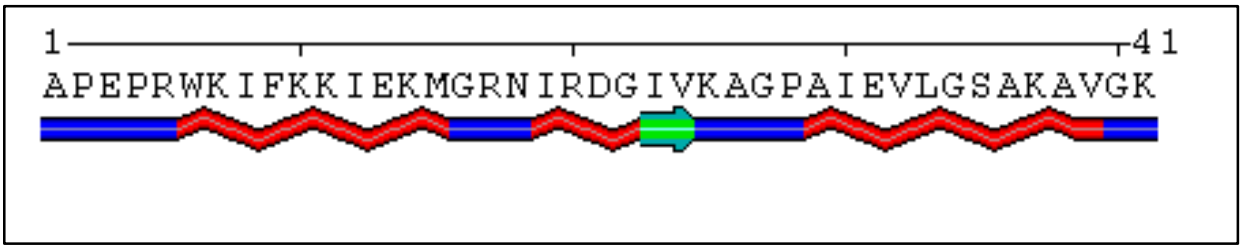


Figure 5.3 (bg): Secondary structure of Cecropin – Q53X40

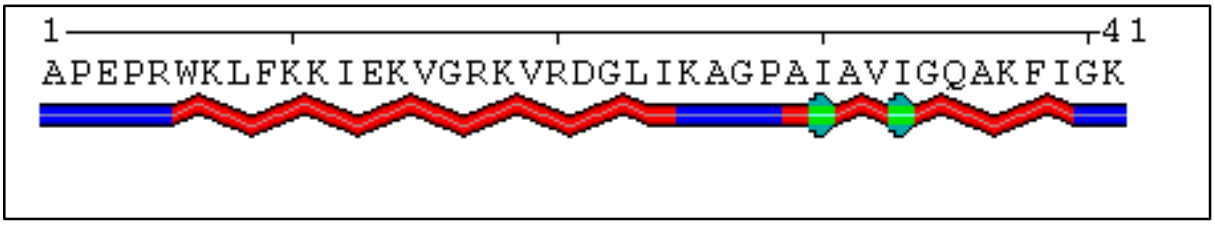


Figure 5.3 (bh): Secondary structure of Cecropin – Q53X40

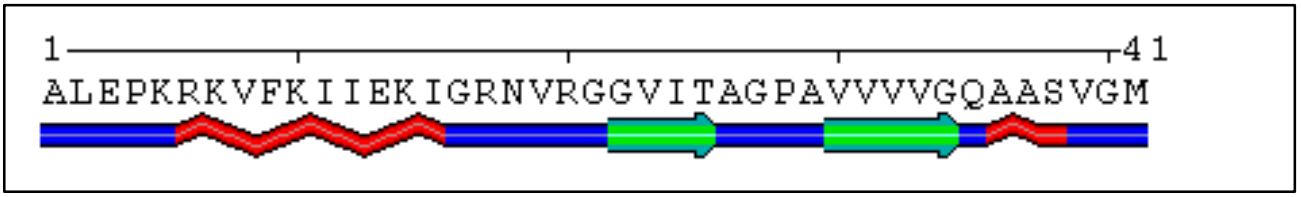


Figure 5.3 (bi): Secondary structure of Cecropin – Q53X40

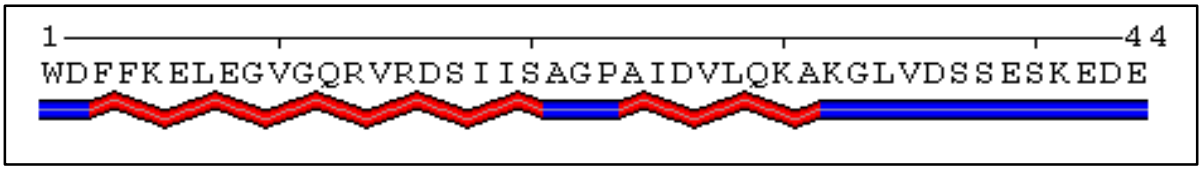


Figure 5.3 (bj): Secondary structure of Cecropin – Q53X40

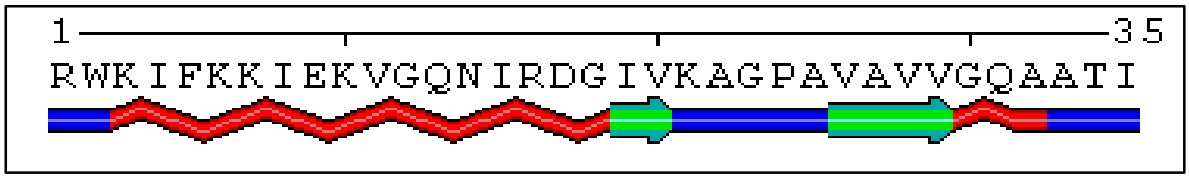


Figure 5.3 (bk): Secondary structure of Cecropin – Q53X40

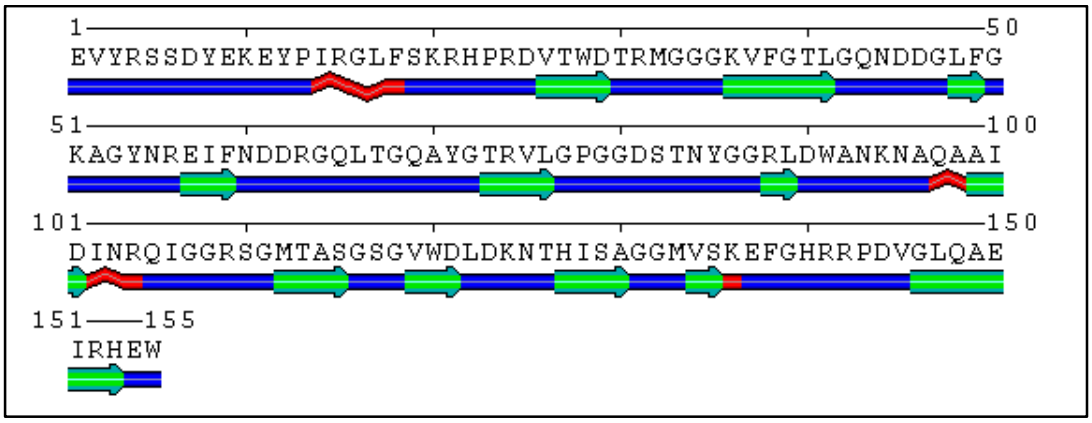


Figure 5.3 (ca): Secondary structure of Gloverin-A5LHW4

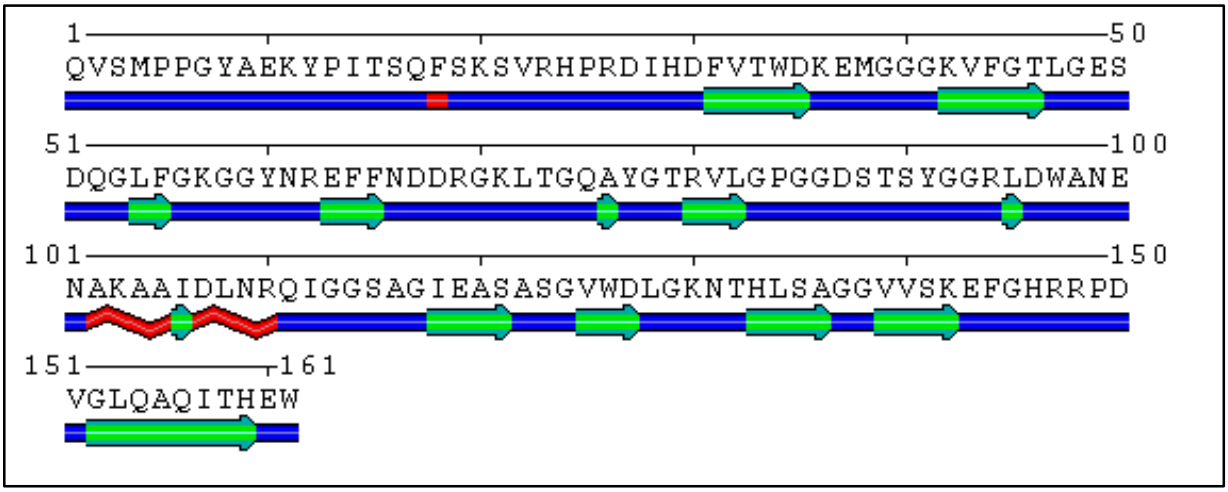


Figure 5.3 (cb): Secondary structure of Gloverin-H9JWE9

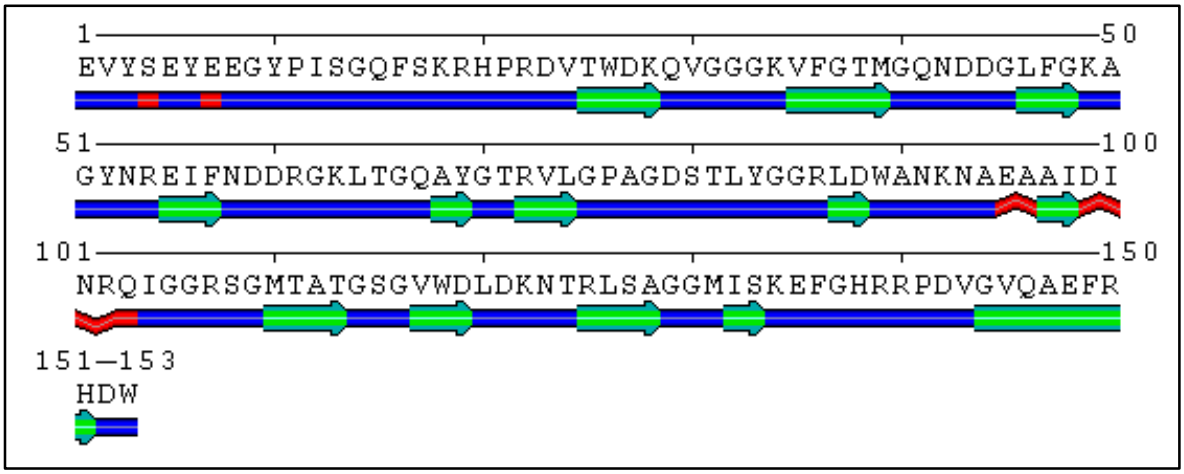


Figure 5.3 (cd): Secondary structure of Gloverin4-Q2WGL0

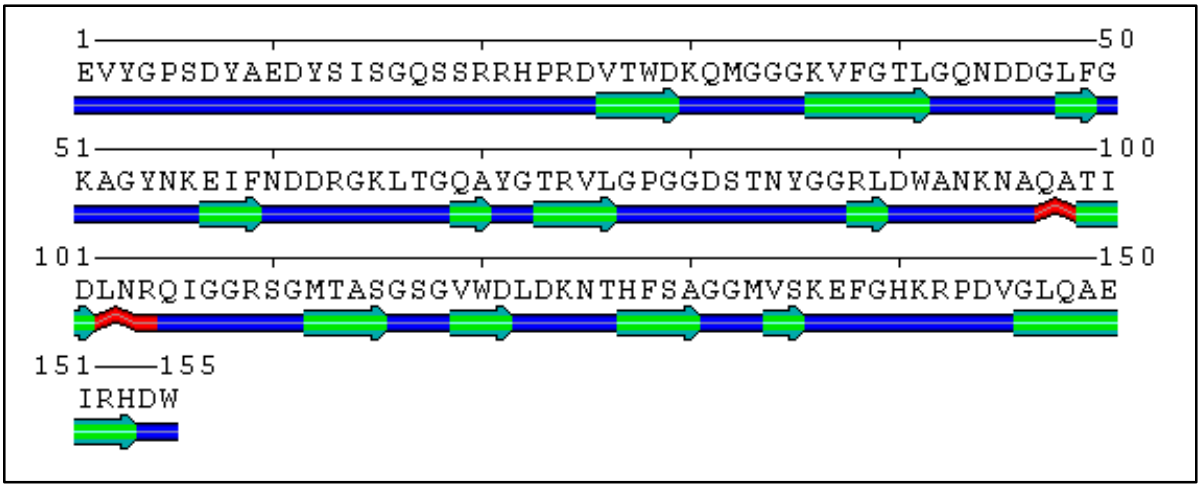


Figure 5.3 (ce): Secondary structure of Gloverin-Q2WGL1

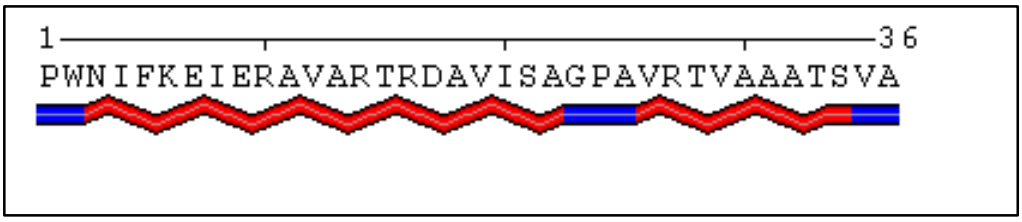


Figure 5.3 (da): Secondary structure of Enbocin 1

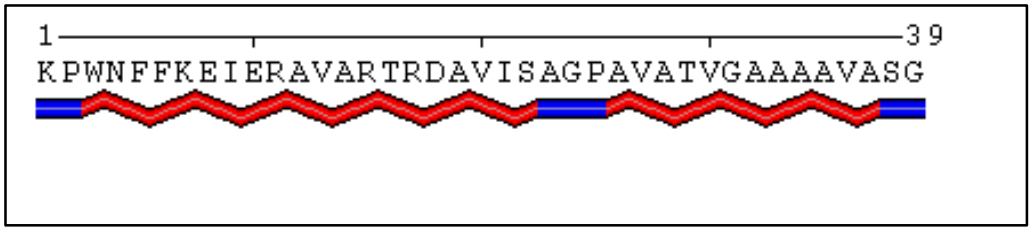


Figure 5.3 (db): Secondary structure of Enbocin 2

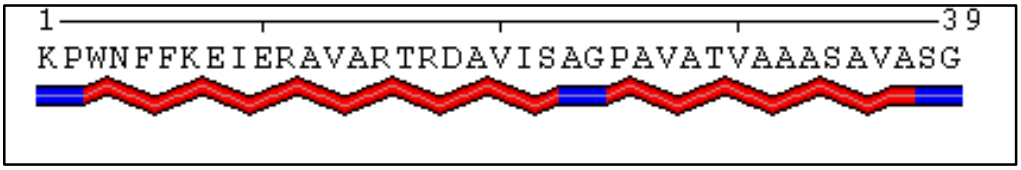


Figure 5.3 (dc): Secondary structure of Enbocin-3

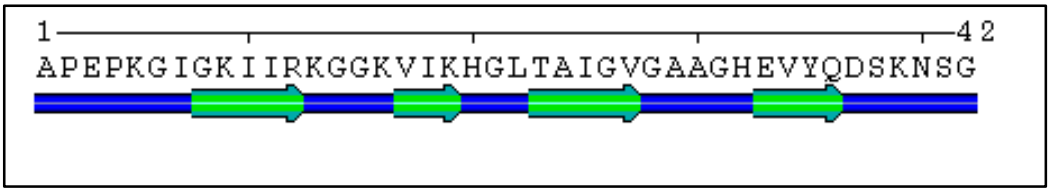


Figure 5.3 (ea): Secondary structure of Moricin-B3

Figure 5.3 (eb): Secondary structure of Moricin-2

Figure 5.3 (ec): Secondary structure of Moricin-1

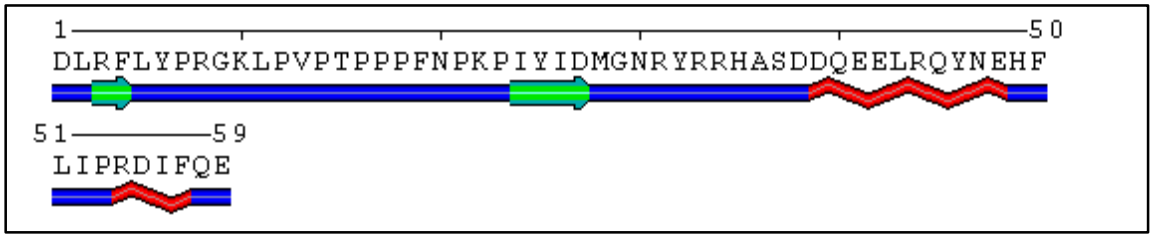


Figure 5.3 (fa): Secondary structure of Lebocin 1/2

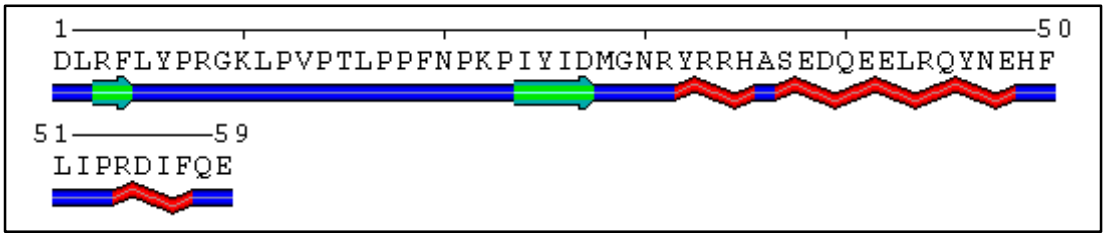


Figure 5.3 (fb): Secondary structure of Lebocin-3

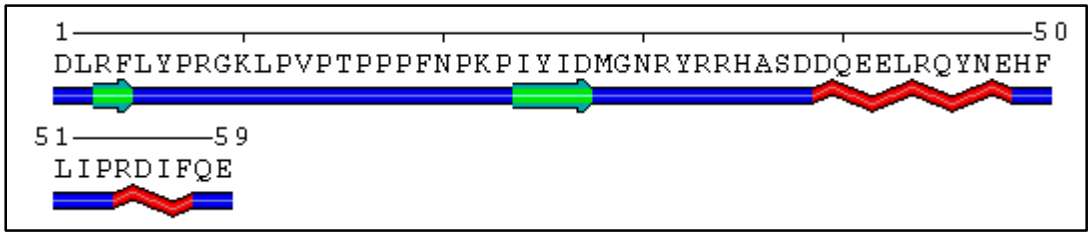


Figure 5.3 (fc): Secondary structure of Lebocin-4

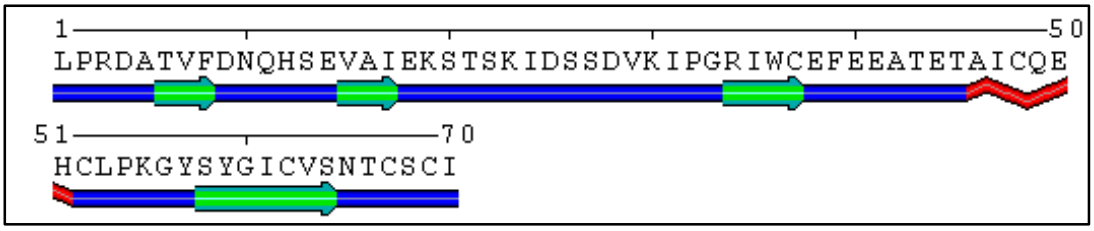


Figure 5.3 (ga): Secondary structure of Defensin like protein

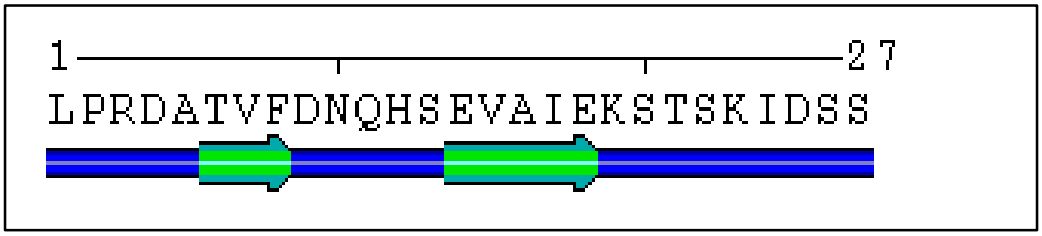


Figure 5.3 (gb): Secondary structure of Defensin A

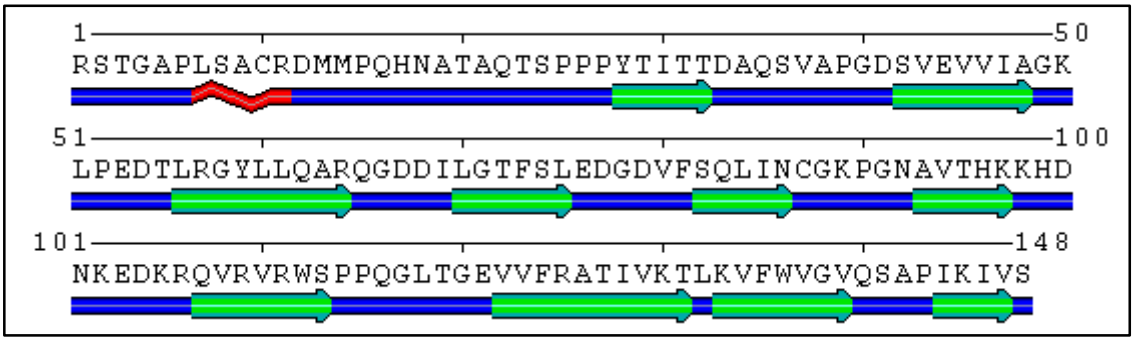


Figure 5.3 (gc): Secondary structure of Putative defensin protein

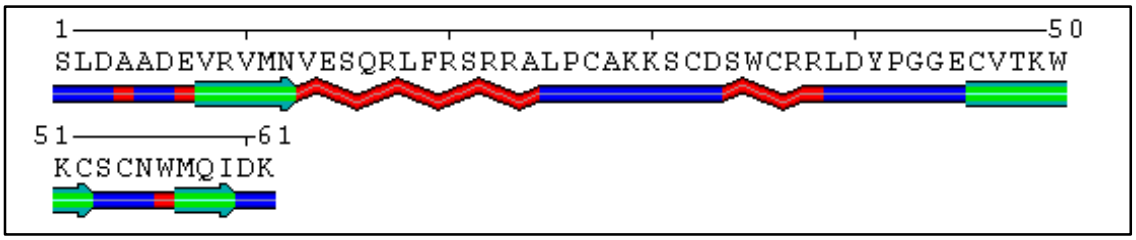


Figure 5.3 (gd): Secondary structure of Defensin like protein-2

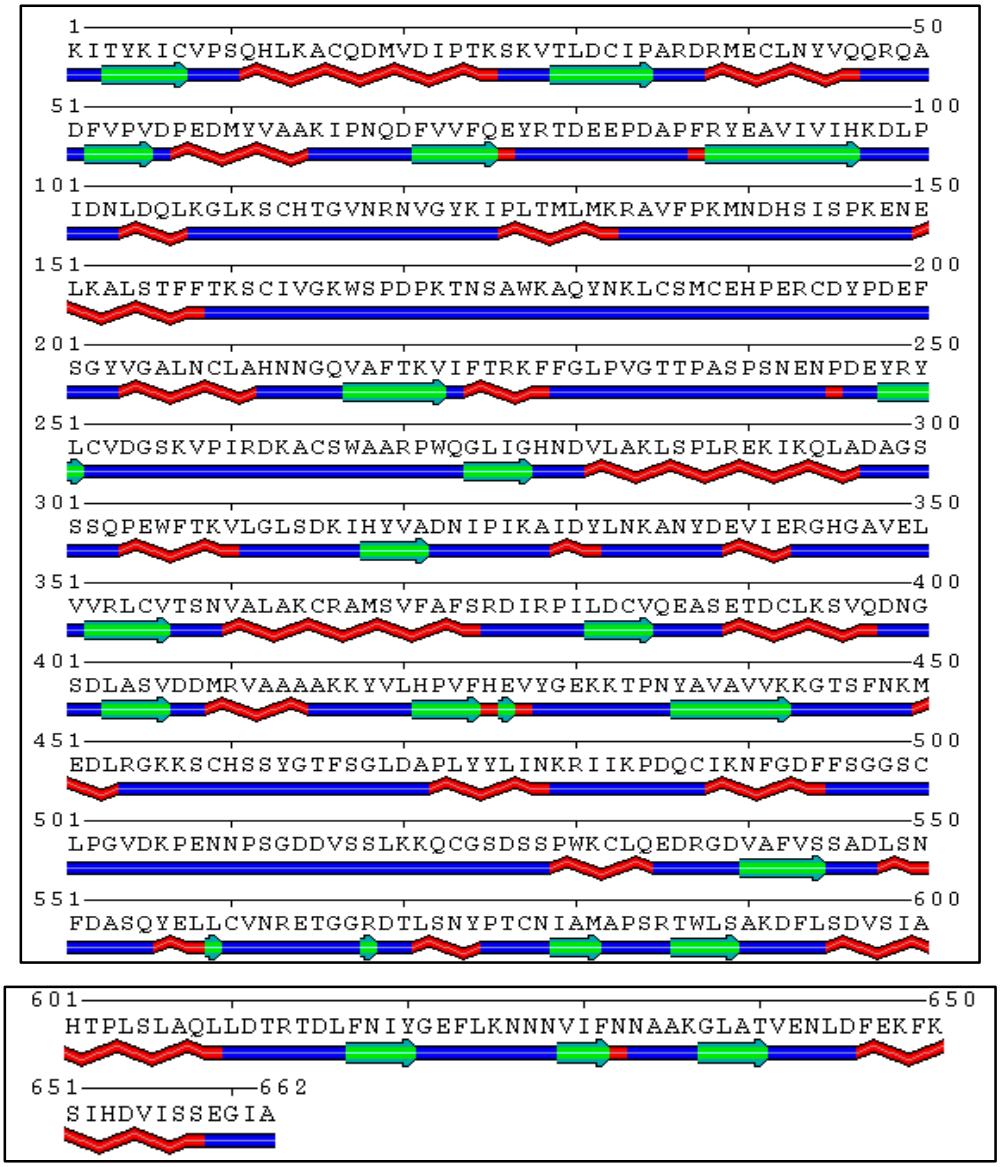


Figure 5. 3 (h): Secondary structure of Transferrin

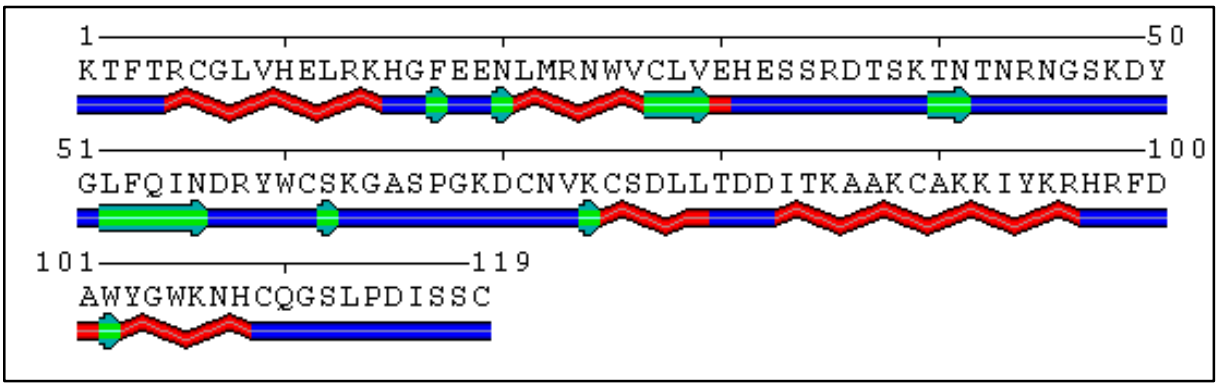


Figure 5.3 (i): Secondary structure of Lysozyme

5.4 Validation of the results

The solution NMR structure of BM Moricin (Hemmi, Ishibashi, Hara, & Yamakawa, 2002) showed 8 helices which validates our predicted structure.

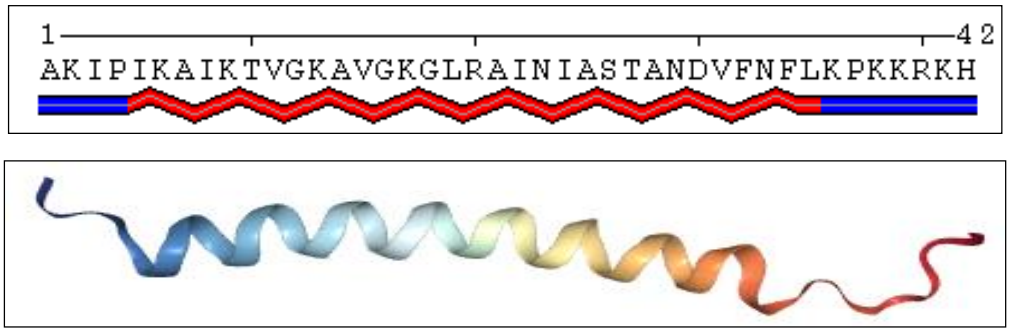


Figure 5.4. 1: Figure showing correlation between predicted and solved structures respectively

The structure of Cecropin A predicted in the study concurs with the solution NMR studies carried out. With the use of combination of two-dimensional NMR, Cecropin A (Holak et al., 1988) was shown to have two helical regions with a GP sequence in the middle.

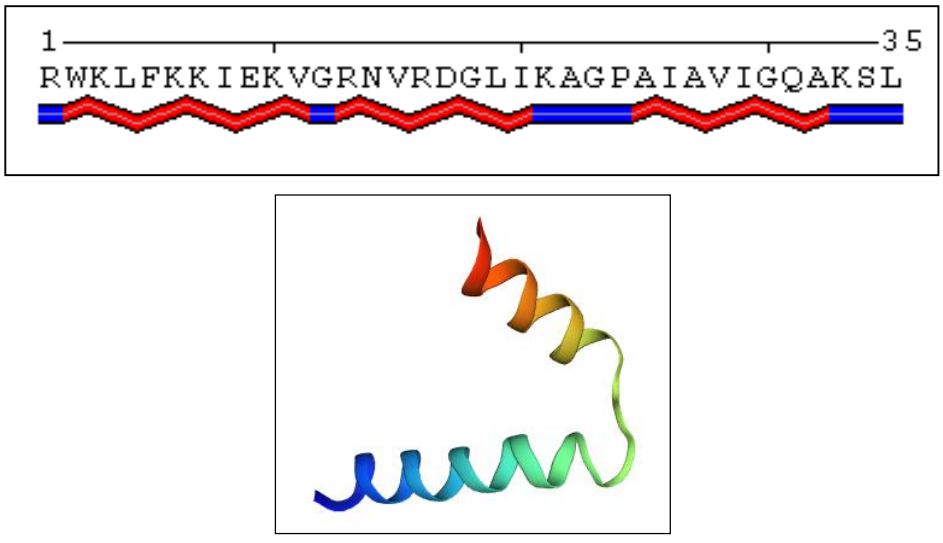


Figure 5.4. 2: Figure showing correlation between predicted and solved structures respectively

CHAPTER 6: SUMMARY

From CAMP, APD3 and Pfam, 35 AMPs were collected and classified into 7 families (Cecropin, Attacin, Defensin, Gloverin, Moricin, Lebocin and Enbocin). The highest frequency of amino acid in the families of AMPs was Glycine (17%) for Gloverin, Lysine (14.82%) for Cecropin, Serine (11.6%) and Alanine (11.7%) in Attacin, Alanine (27.6%) in Enbocin, Lysine (19%) in Moricin, Serine (11.6%) in Defensin and Proline (23%) in Lebocin. It was found that all families of AMP except Defensin lack the amino acid Cysteine, which indicates the absence of di-sulphide bonds. Methionine is the least frequent amino acid (1-3%). Isoelectric point (pI) of majority of AMPs was found to be ≈ 10, similar to detergents or emulsifying agents which assists the mechanism of action for interacting with lipid bilayer of biological membrane. Instability index of less than 40 were shown in 29 AMPs which indicates that they are stable. Aliphatic index of more than 70 was shown in 25 AMPs which indicates stability over a wide range of temperature. 30 AMPs studied were found to be hydrophilic. It was observed that off 35 AMPs, 29 are cationic, 3 are neutral and 3 are anionic. Cecropin, Enbocin and Moricin fall under the alpha family of secondary structures. Lebocin can be classified as Pro-rich; Proline is known to be a helix breaker, therefore, very few helices and they are highly coiled. Enbocin can be classified as Ala-rich; and therefore, more helices are found in the secondary structure.

CHAPTER 7: CONCLUSIONS

The severe problems related with multi drug resistant microorganisms have created crucial demand for the development of alternative therapeutics. Concurrently with the increase in resistance to commercially available antibiotics, there is a serious need for novel, effective therapeutics with lesser or no side effects. However, AMPs are the promising candidate for the production of new generation antibiotics. AMPs have various mechanisms through which they act and majority have a theoretical pI ~10, that is similar to detergents or emulsifying agents which assists the mechanism of action for interacting with lipid bilayer of biological membrane. Most AMPs are hydrophilic and cationic, which promotes selectivity for negatively charged microbial cytoplasmic membranes. They also show stability over a wide range of temperatures as indicated by Aliphatic and instability index. All AMPs have signal peptide and propeptides which are cleaved during activation or maturation. Amidated peptides are less sensitive to proteolytic degradation, extending their half-life in the bloodstream. Secondary structures of 35 AMPs were solved and they were classified into non-αβ and α family. Cecropin, Enbocin and Moricin fall under the α family of secondary structures. Gloverin, Defensin, Lebocin and Attacin fall under non-αβ family of secondary structures. Lebocins are Pro-rich; Proline- helix breaker amino acid, there is a perfect correlation between amino acid composition and it's predicted secondary structure of Lebocins (highly coiled). Enbocins are Ala-rich; Alanine- helix maker amino acid, there is a perfect correlation between amino acid composition and it's predicted secondary structure of Enbocins (highly helical). The predicted secondary structure of BM Moricin agrees with the experimentally proven structure. Moricins have alternating charged and hydrophobic residues which is a characteristic of alpha helices. The predicted structure of Cecropin A concurs with the structure determined by H-NMR. Almost all Cecropins have 2 helical regions with a GP sequence in the middle.

CHAPTER 8: REFERENCES

1. Akinbowale, O. L., Peng, H., & Barton, M. D. (2006). Antimicrobial resistance in bacteria isolated from aquaculture sources in Australia. *Journal of Applied Microbiology*. https://doi.org/10.1111/j.1365-2672.2006.02812.x
2. Buhroo, Z. I., M. A. Bhat, G. K. Bali, A. S. Kamili, and N. A. Ganai. 2018. "Antimicrobial Peptides from Insects with Special Reference to Silkworm *Bombyx mori* L : A Review." Journal of Entomology and Zoology Studies 6(4):752–59.
3. Bulet, Phillipe, Charles Hetru, Jean Luc Dimarcq, and Daniéle Hoffmann. 1999. "Antimicrobial Peptides in Insects; Structure and Function." Developmental and Comparative Immunology.
4. Cheng, J., A. Z. Randall, M. J. Sweredoski, and P. Baldi. 2005. "SCRATCH: A Protein Structure and Structural Feature Prediction Server." Nucleic Acids Research.
5. Fenwick, R. B., Orellana, L., Esteban-Martín, S., Orozco, M., & Salvatella, X. (2014). Correlated motions are a fundamental property of β-sheets. *Nature Communications*. https://doi.org/10.1038/ncomms5070
6. Garnier, J., J. Gibrat, and B. Robson. 1996. "GOR Secondary Structure Prediction Method Version IV." Methods in Enzymology.
7. Gasteiger, E., C. Hoogland, A. Gattiker, S. Duvaud, Wilkins M, RD Apel, and A. Bairoch. 2005. "ProteinIdentification and Analysis Tool on the ExPASy Server." in The Proteomic Protocols Handbook.
8. Geourjon, C. and G. Deléage. 1995. "Sopma: Significant Improvements in Protein Secondary Structure Prediction by Consensus Prediction from Multiple Alignments." Bioinformatics.
9. Gordon, Gavín J., Graham N. Rockwell, Roderick V. Jensen, James G. Rheinwald, Jonathan N. Glickman, Joshua P. Aronson, Brian J. Pottorf, Matthew D. Nitz, William G. Richards, David J.
10. Hamamoto, Hiroshi, Akiko Tonoike, Kazuya Narushima, Ryo Horie, and Kazuhisa Sekimizu. 2009. "Silkworm as a Model Animal to Evaluate Drug Candidate Toxicity and Metabolism." Comparative Biochemistry and Physiology - C Toxicology and Pharmacology.

11. Hemmi, H., Ishibashi, J., Hara, S., & Yamakawa, M. (2002). Solution structure of moricin, an antibacterial peptide, isolated from the silkworm Bombyx mori. *FEBS Letters*. https://doi.org/10.1016/S0014-5793(02)02637-6
12. Holak, T. A., Engström, A., Kraulis, P. J., Lindeberg, G., Bennich, H., Jones, T. A., … Gronenborn, A. M. (1988). The Solution Conformation of the Antibacterial Peptide Cecropin A: A Nuclear Magnetic Resonance and Dynamical Simulated Annealing Study. *Biochemistry*. https://doi.org/10.1021/bi00420a008
13. Ishii, K., K. -i. Kubo, T. Endo, K. Yoshida, S. Benner, Y. Ito, H. Aizawa, M. Aramaki, A. Yamanaka, K.Tanaka, N. Takata, K. F. Tanaka, M. Mimura, C. Tohyama, M. Kakeyama, and K. Nakajima. 2015. "Neuronal Heterotopias Affect the Activities of Distant Brain Areas and Lead to Behavioral Deficits." Journal of Neuroscience.
14. Islam, Sebrin, Sukanya Bezbaruah, and Jatin Kalita. 2016. "A Review on Antimicrobial Peptides from *Bombyx mori* L and Their Application in Plant and Animal Disease Control." Journal of Advances in Biology & Biotechnology 9(3):1–15.
15. Jitendra Shahaji, Bajare, Oli Ajay Kumar, Jalkute Chidambar Balbhim, Dulange Sanjay Mallikarjun, and Thorat Prakash Ramrao. 2015. "Comparative Study of *Bombyx mori* Antimicrobial Peptides (AMPs)Retrieved from APD2 Database." Journal of Advanced Bioinformatics Applications and ResearchOnline ISSN.
16. Kaito, C. and K. Sekimizu. 2007. "A Silkworm Model of Pathogenic Bacterial Infection." Drug Discoveries & Therapeutics.
17. Kaito, Chikara, Nobuyoshi Akimitsu, Haruo Watanabe, and Kazuhisa Sekimizu. 2002. "Silkworm Larvae as an Animal Model of Bacterial Infection Pathogenic to Humans." Microbial Pathogenesis.
18. KANEKO, Yoichi, Seiichi FURUKAWA, Hiromitsu TANAKA, and Minoru YAMAKAWA. 2007. " Expression of Antimicrobial Peptide Genes Encoding Enbocin and Gloverin Isoforms in the Silkworm, *Bombyx mori* ." Bioscience, Biotechnology, and Biochemistry 71(9):2233–41.
19. Lata, Sneh, B. K. Sharma, and G. P. S. Raghava. 2007. "Analysis and Prediction of Antibacterial Peptides." BMC Bioinformatics.

20. Meng, Xu, Feifei Zhu, and Keping Chen. 2017. “Silkworm: A Promising Model Organism in Life Science.” Journal of Insect Science.
21. Mirabello, Claudio and Gianluca Pollastri. 2013. “Porter, PaleAle 4.0: High-Accuracy Prediction of Protein Secondary Structure and Relative Solvent Accessibility.” Bioinformatics.
22. Nguyen, L. T., Haney, E. F., & Vogel, H. J. (2011). The expanding scope of antimicrobial peptide structures and their modes of action. *Trends in Biotechnology*, *29*(9), 464–472. https://doi.org/10.1016/j.tibtech.2011.05.001
23. Panthee, S., Paudel, A., Hamamoto, H., & Sekimizu, K. (2017). Advantages of the silkworm as an animal model for developing novel antimicrobial agents. *Frontiers in Microbiology*. https://doi.org/10.3389/fmicb.2017.00373
24. Sugarbaker, and Raphael Bueno. 2005. “Identification of Novel Candidate Oncogenes and Tumor Suppressors in Malignant Pleural Mesothelioma Using Large-Scale Transcriptional Profiling.” American Journal of Pathology.
25. Tanaka, H. and M. Yamakawa. 2011. “Regulation of the Innate Immune Responses in the Silkworm, *Bombyx mori*.” ISJ.
26. Van Der Biezen, Erik A. 2001. “Quest for Antimicrobial Genes to Engineer Disease-Resistant Crops.” Trends in Plant Science.
27. Zhang, G., Ross, C. R., & Blecha, F. (2000). Porcine antimicrobial peptides: New prospects for ancient molecules of host defense. *Veterinary Research*. https://doi.org/10.1051/vetres:2000121
28. Zhang, Xiaoli, Renyu Xue, Guangli Cao, Zhonghua Pan, Xiaojian Zheng, and Chengliang Gong. 2012. “Silkworms Can Be Used as an Animal Model to Screen and Evaluate Gouty Therapeutic Drugs.” Journal of Insect Science.