MOLECULAR AND FUNCTIONAL CHARACTERIZATION OF ANTIMICROBIAL PEPTIDES IN CRUSTACEANS

Thesis submitted to

Cochin University of Science and Technology

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

UNDER THE FACULTY OF MARINE SCIENCES

by

AFSAL V.V.
Reg. No: 3324

DEPARTMENT OF MARINE BIOLOGY, MICROBIOLOGY AND BIOCHEMISTRY
SCHOOL OF MARINE SCIENCES
COCHIN UNIVERSITY OF SCIENCE AND TECHNOLOGY
KOCHI- 682016, INDIA.

June 2015

To my family…

Dr. Rosamma Philip
Associate Professor

Department of Marine Biology, Microbiology and Biochemistry
Cochin University of Science and Technology
Fine Arts Avenue, Kochi-16

Certificate

*This is to certify that the thesis entitled "**Molecular and Functional Characterization of Antimicrobial Peptides in Crustaceans"** is an authentic record of the research work carried out by **Shri. Afsal V.V.** under my supervision and guidance in the Department of the Marine Biology, Microbiology and Biochemistry, School of Marine Sciences, Cochin University of Science and Technology, in partial fulfillment of the requirements of the degree of **Doctor of Philosophy** under the Faculty of Marine Sciences of Cochin University of Science and Technology, and no part thereof has been presented for the award of any other degree, diploma or associateship in any university.*

Dr. Rosamma Philip

Kochi – 16

9th June 2015

Declaration

*I hereby declare that the thesis entitled "**Molecular and Functional Characterization of Antimicrobial Peptides in Crustaceans"** is a genuine record of research work done by me under the supervision and guidance of **Dr. Rosamma Philip,** Associate Professor**,** Department of Marine Biology, Microbiology and Biochemistry, Cochin University of Science and Technology. The work presented in this thesis has not been submitted for any other degree or diploma earlier.*

Afsal V.V.

Kochi-16

9th June 2015

Dr. Rosamma Philip
Associate Professor

Department of Marine Biology, Microbiology and Biochemistry
Cochin University of Science and Technology
Fine Arts Avenue, Kochi-16

Certificate

This is to certify that all relevant corrections and modifications suggested by the audience during the presynopsis seminar and recommended by the doctoral committee of ***Shri. Afsal V.V.*** *has been incorporated in the thesis entitled* ***"Molecular and Functional Characterization of Antimicrobial Peptides in Crustaceans".***

Dr. Rosamma Philip

Kochi-16

9th June 2015

Acknowledgements

I am deeply indebted to all those who supported me for the successful completion of this thesis.

Prima facea, I am grateful to God for blessing me with good health and wellbeing that were necessary to complete this work.

I would like to express my deepest gratitude to my mentor and guide Dr. Rosamma Philip. It has been an honor to be her Ph.D. student. The joy and enthusiasm she has for her research was contagious and motivational for me. I am extremely grateful to her, for her valuable guidance, scholarly inputs and consistent encouragement I received throughout the research period. I consider it as a great opportunity to do my doctoral programme under her guidance and to learn from her research expertise. Thank you Ma'm, for all your help and support.

I owe to Prof. I. S. Bright Singh, National Centre for Aquatic Animal Health for his encouragement, guidance and for permitting to use the facilities at the Centre.

I would like to express my sincere thanks to The Head, Department of Marine Biology, Microbiology and Biochemistry, for providing all kind of support and facilities required for this work. I also express my deep felt gratitude to the Dean, Faculty of Marine Sciences and Director, School of Marine Sciences, CUSAT for all the help rendered and facilities provided for research.

I am thankful to all my beloved teachers in the Department for their support and encouragement. I thank Dr. A.V. Saramma, Dr. A. A. Mohamed Hatha, Dr. Aneykutty Joseph, Dr. S. Bijoy Nandan, Dr. Babu Philip and Dr. C.

K. Radhakrishnan for their valuable advice, suggestion and constant support. I would like to thank the Administration staff, Technical staff and other Supporting staff of the Department for their timely help during the period.

I would also like to express my sincere thanks to Dr. Sajeevan T. P. and Dr. Valsamma Joseph for the valuable suggestions and support rendered to me during the period. I express my heartfelt gratitude to Dr. Joice V. Thomas, CE, NETFISH whose moral support and advice have helped me a lot for the successful completion of this work.

I am also grateful to the Ministry of Earth Sciences (MoES), Govt. of India for the research grants (MoES/10-MLR/2/2007) with which the work was carried out.

I am deeply indebted to Dr. Swapna P. Antony, my friend, my colleague and moreover my better half, for her valuable guidance, suggestions and encouragement. This feat was possible only because of the unconditional love and support provided by her.

I would like to express my sincere thanks to Naveen Sathyan, Chaithanya E.R, Anil Kumar P.R, Sruthy K.S and Aiswarya Ajith, the members of the AMP group who have contributed immensely to this work. The group has been a source of friendship as well as good advice and collaboration. Special gratitude is owed to Dr. Jayesh P. for rendering valuable help, support and suggestions in the completion of this work. This thesis would not have come to a successful completion without their support.

I express my sincere thanks to Divya T. Babu, Ramya K. D., Jimly Jacob, Deepthi Augustine, Solly Solomon, Jayanath G. and all my good friends and colleagues in the University for their love and support. I express

my gratefulness to Deepu A.V, K.S.S.M. Yousuf, Ratheesan K., Ambrose T.V., Anoop A, Anoop B, Hashim M, Rajool Shanis and all other good friends and colleagues in CMFRI, CMLRE, MPEDA & NETFISH for their love and support rendered throughout my professional and personal life. I would like to express special gratitude to Arun Hari for being there with me in all the ups and downs of my life. And thank you to all the wonderful people who have helped me enormously and who I did not mention here.

It is hard to express my gratitude in words to my parents; Mr. V.K. Veeravunni and Mrs. Mariyumbi. I am very grateful for their love, care and trust on me. I would like to specially mention my twin boys, Eshan and Ehzan. They are my inspiration and motivation to move ahead in life. I thank my sister, uncles, aunts, cousins and in-laws for their unconditional love and constant support. My special thanks goes to my bro, Shiraz Kareem whose deep affection and moral support has helped me in coping with the hard ships during my entire life.

I sincerely thank all those who have helped me in one way or other in the fulfillment of this cherished ambition.

Afsal V.V.

Table of Contents

1

General Introduction

1.1 Introduction

Most living organisms are constantly exposed to potentially harmful pathogens. It is the immune system of the organism that enables it to survive in an environment loaded with dangerous pathogenic microorganisms. The innate immunity provides organisms with a rapid and non-specific first line of defense against pathogens. It includes physical barriers such as skin and mucous membranes and chemical barriers including the high acidity of gastric juice, and specialized soluble molecules that possess antimicrobial activity. One of the well-known innate immune defense mechanisms is the production of antimicrobial substances by specific cells or tissues of the organisms. Antimicrobial peptides (AMPs) are such natural substances that possess the capacity to directly kill or inhibit the growth of microbes.

Antimicrobial peptides, also designated as host defense peptides, are small sized amphipathic molecules that play a crucial role in the innate immune response of all living organisms ranging from prokaryotes to multi-cellular eukaryotes including humans. AMPs are typically relatively short (12 to 100 amino acids), positively charged (net charge of +2 to +9), hydrophobic peptides with molecular mass less than 10 kDa, and are membrane active (Bowman, 2003; Yeaman and Yount, 2003). AMPs are

gene encoded, ribosomally synthesized and therefore evolutionarily conserved biomolecules, which serve as natural first-line of defense system in majority of living organisms against variety of pathogens (Hancock and Lehrer, 1998; Zasloff, 2002; Brogden et al., 2003; Gallo and Nizet, 2003). AMPs are better known as "natural antibiotics" due to their rapid and efficient antimicrobial action against a broad range of microorganisms, including Gram-positive and Gram-negative bacteria, yeast, filamentous fungi and, to a lesser extent, protozoans and enveloped viruses (Bulet et al., 2004; Yount et al., 2006; Guaní-Guerra et al., 2010). In addition to their antimicrobial activity, AMPs were proved to encompass a number of other diverse biological roles and are, indeed, multifunctional molecules. It was demonstrated that these peptides have antitumor effects, mitogenic activity and, most importantly, participate in immune regulatory mechanisms by modulating signal transduction and cytokine production and/or release (Bowdish et al., 2005; Brown and Hancock, 2006; Yount et al., 2006; Lai and Gallo, 2009; Guaní-Guerra et al., 2010). AMPs are promptly synthesized at low metabolic cost, easily stored in large amounts and readily available shortly after an infection, to rapidly neutralize a broad range of microbes. Some AMPs are produced constitutively whilst others are synthesized in response to microbial attack (Gallo et al., 2002) at rates, which are up to one hundred fold faster than those used for protein synthesis by the adaptive immune system (Boman, 2003). This ready availability of AMPs form a crucial component of innate immune systems making it a highly effective first line of defense in animals (Ganz, 2003). Based on the unique mode of action, broad spectrum of activity, non-cytotoxicity to eukaryotic cells and lower

chances of resistance development, AMPs are considered as potent molecules for therapeutic applications.

AMPs have been recognized in prokaryotic cells since 1939 when antimicrobial substances, named gramicidins, were isolated from *Bacillus brevis*, and were found to exhibit activity both *in vitro* and *in vivo* against a wide range of Gram-positive bacteria (Dubos, 1939). In case of eukaryotes, AMPs were initially found in invertebrates (Boman, 1991), and later also in vertebrates, including humans (Gudmundsson et al., 1996). Eukaryotic AMPs first became a research focus in the middle decades of the twentieth century with the description of cecropins from moths and magainins from frogs (Boman et al., 1985; Boman, 1995; Hancock and Lehrer, 1998). Since then, the number of reported AMPs has burgeoned to over 2300 with representatives in virtually almost all organisms, according to the online updated Antimicrobial Peptide Database (APD) (http://aps.unmc.edu/AP/main.php). Interestingly, irrespective of their origin, spectrum of activity and structure, most of these peptides share several common properties.

1.2 Characteristics of AMPs

For all conventional AMPs, characteristics such as size, charge, molecular weight, isoelectric point, hydrophobicity and amphipathicity are critical in determining antibacterial activity and modes of killing and have been well reviewed previously (Brogden, 2005).

Size: Basically AMPs are relatively short polypeptides with fewer than 100 amino acid residues, usually within a range between 12 and 50. The smaller size of AMPs facilitates the rapid diffusion and secretion of peptide outside the cells, which is required for eliciting immediate

defense response against pathogenic microbes (Nissen-Meyer and Nes, 1997). The size of an AMP may vary from six amino acid residues in case of anionic peptides to greater than 59 amino acid residues as in the case of Bac7. Even di- and tri-peptides with antimicrobial activity have been reported. It has been observed that AMPs active against Gram-negative bacteria exhibited the narrowest range of sequence lengths, 19 to 27 amino acids, and molecular masses, 1.7 to 2.5 kDa; whilst those AMPs that were active against both Gram-positive and Gram-negative bacteria, exhibited the widest variation in these parameter with sequence lengths between 10 and 68 residues and molecular masses between 1.1 and 7.4 kDa respectively (Dennison et al., 2003).

Charge: Many of the AMPs characterized to date display a net positive charge, ranging from +2 to +9 (Giangasper et al., 2001; Yeaman and Yount, 2003), due to the presence of few or no acidic residues, such as glutamate or aspartate and a high number of cationic amino acids such as lysine or arginine and/or histidine (Hancock and Diamond, 2000). Hence the term 'cationic' AMPs is used to describe these molecules, with at least two excess positive charges due to lysine and arginine residues and around 50 % hydrophobic amino acids. More than 500 such peptides have been discovered. They fit into at least four structural classes, namely α- helices, β-strands stabilized by disulphide bridges, extended structures, and loop structures. The cationicity of AMPs enables them to bind with phospholipid bilayers in negatively charged bacterial surfaces and membranes through electrostatic interactions (Matsuzaki et al., 1997; Huang, 2000; van 't Hof et al., 2001; Yeaman and Yount, 2003). These interactions can then lead to a range of effects, including membrane permeabilization, depolarization, leakage or lysis, resulting in cell death

(Wiesner and Vilcinskas, 2010; Fjell et al., 2012). It has been proved that anionic peptides that are complexed with zinc, or highly cationic peptides, are often more active than neutral peptides or those with a lower charge.

Molecular weight: The molecular weight is the sum of the masses of each atom constituting a molecule. The molecular weight is directly related to the length of the amino acid sequence and is expressed in units called daltons (Da). AMPs due to its short length are characterized by a molecular weight <10 kDa. It has been suggested that both sequence length and molecular weight, per se, are not important to either the antimicrobial efficacy or specificity of AMPs analyzed. However, the vast majority of AMPs possess a molecular weight of <10 kDa and <100 residues in length (Dennison et al., 2003) and the relative invariance of these factors across a diverse range of species implies biological relevance to the role of these peptides as defense agents. This importance may derive from the fact that short, low molecular weight peptides are metabolically economical to the host and can be more easily stored in large amounts (Gallo et al., 2002), thereby increasing the efficiency of host response to microbial attack.

Isoelectric point: The isoelectric point (*p*I) is the pH at which the net charge of the protein is equal to zero. It is a variable that affects the solubility of the peptides under certain conditions of pH. When the pH of the solvent is equal to the *p*I of the protein, it tends to precipitate and lose its biological function. AMPs exhibit a wide range in isoelectric points, varying from 12.7, shown by horse myeloid cathelicidin (Scocchi et al., 1999), to 4.2 shown by enkelytin (Goumon et al., 1996) and the Amoeba pore forming peptide isoform A (Leippe et al., 1991). Most of the AMPs described till date possess an isoelectric point close to 10 (Torrent et al.,

2011), which is consistent with the proposed mechanisms of action for these peptides. However, it has been proposed that there exhibit no discernable correlation between *p*I values and MICs of AMPs (Dennison et al., 2003)

Hydrophobicity: A hydrophobicity of a chemical compound is related to its transfer free energy from a polar medium (phase) to non-polar medium (phase). Hydrophobicity of a peptide, defined as the percentage of hydrophobic residues within a peptide, is usually evaluated as a sum of particular amino acid transfer free energies. The AMPs often possess nearly 50 % hydrophobic residues. Hydrophobicity describes how polypeptide chain forms and stabilizes their 3D structure in polar or non-polar environment. The hydrophobicity is an important stabilization force in protein folding; this force changes depending on the solvent in which the protein is found. It is considered as an essential feature for AMP-membrane interactions, as it governs the extent to which a peptide can partition into the lipid bilayer. Many AMPs are moderately hydrophobic, such that they optimize the activity against microbial cell membranes.

Also, the arrangement of hydrophobic amino acids is usually observed in the pattern of i+3 or i+4 (Pasupuleti et al., 2008). This arrangement ensures that all the hydrophobic and hydrophilic amino acids are on two different planes when the peptide assumes helical structures resulting in a perfect amphipathic molecule.

Amphiphathicity: One of the main features of AMPs is their amphipathicity (Yeaman and Yount, 2003). Amphiphilicity may be defined as a relative abundance and distribution of hydrophobic and hydrophilic residues or domains within a peptide (Yount and Yeaman, 2005; Yount et

al., 2006). It is an essential factor directing protein folding and polypeptide-membrane association process. Nearly all AMPs often exhibit spatially separated hydrophobic and hydrophilic regions and show amphipathic properties upon interaction with target membranes (Splith and Neundorf, 2011). Amphipathicity can be achieved via a multitude of protein conformations; however, one of the simplest and perhaps most elegant is the amphipathic helix. The amphipathic α-helix has a periodicity of three to four residues and is optimal for interaction with amphipathic biomembranes. While the extent of amphipathic helicity influences peptide activity against negatively charged membranes, it may have an even more pronounced effect in rendering peptides hemolytic against zwitter ionic or neutral membranes. Thus, a high degree of helicity and/or amphipathicity yielding a segregated hydrophobic domain is correlated with increased toxicity toward cells composed of neutral phospholipids.

1.3 Classification of AMPs

AMPs can be categorized into many subtypes based on different criteria such as origin, size, charge, length, structure, amino acid sequence, biological action and mechanism of action. Earlier, AMPs were classified based on the source organism as this type of classification helped to correlate its function and its habitat. However due to the discovery of large number of AMPs showing high degree of sequence dissimilarity, classification of AMPs became complex. At present, a grouping approach based on secondary structures and charge is mostly preferred as well as widely accepted.

1.3.1 Classification of AMPs based on structure

Based on their secondary structure, AMPs are classified as α-helices, β-strands with one or more disulphide bridges, loop structures and extended structures (van't Hof et al., 2001; Jenssen et al., 2006a; Nguyen et al., 2011; Pushpanathan et al., 2013). The details of various structural classes of AMPs are shown in Table 1.1 and Figure 1.1. Among these the first two groups including α-helix and β-sheet are more abundant in nature (Powers and Hancock, 2003) and α-helical peptides are the most studied AMPs to date. In addition to the natural peptides, thousands of synthetic variants have also been produced which can also be grouped into these structural classes. The fundamental structural principle is the ability of the AMP to adopt a shape in which clusters of hydrophobic and cationic amino acids are spatially organized in discrete parts of the molecule. The existences of such diverse structural forms of AMPs are highly essential for their broad spectrum antimicrobial activity (Hancock, 2001).

Linear peptides with α-helical structure

One of the largest and deeply studied classes of AMPs are those forming α-helices, such as magainin, cecropin A and temporins (Zasloff et al., 1987; Boman, 1995; Mangoni et al., 2000). Majority of this group are cationic and are highly amphipathic thereby promoting selective interaction with the negatively charged microbial membrane by either absorbing onto the membrane surface or inserting into the membrane as a cluster of helical bundles. In aqueous solution, these peptides adopt disordered structures, and fold into an α-helical conformation upon interaction with hydrophobic solvents or lipid surfaces. There does exist hydrophobic or

slightly anionic α-helical peptides. However, peptides that are not cationic exhibit less selectivity towards microbes compared with mammalian cells. An example of a well-studied hydrophobic and negatively charged cytotoxic peptide is alamethicin (Duclohier and Wroblewski, 2001; Kikukawa and Araiso, 2002). Studies have proved a direct correlation between α-helical conformation and antibacterial activity of these AMPs (Park et al., 2000).

Cysteine rich AMPs possessing β-pleated structure

β-sheet peptides are cyclic peptides consisting of 17 to 88 amino acid residues constrained either by disulfide bonds, as in the case of human β-defensin-2 (Hancock, 2001), tachyplesins (Matsuzaki, 1999), protegrins (Harwig et al., 1995), and lactoferricin (Jones et al., 1994) or by cyclization of the peptide backbone, as in the case of gramicidin S (Prenner et al., 1999), polymyxin B (Zaltash et al., 2000), and tyrocidines (Bu et al., 2002). They largely exist in the β-sheet conformation in aqueous solution that may be further stabilized upon interactions with lipid surfaces. Larger peptides within this family may also contain minor helical segments. The overall structure and activity of this group of AMPs depend on the number of disulfide bridges. Also, the cyclic structure has been shown to be essential for its antibacterial activity (Matsuzaki et al., 1997; Rao, 1999). Defensins and tachyplesins are among the most characterized β-sheet-forming AMPs (Matsuzaki et al., 1997; Tamamura et al., 1998; Rao, 1999).

Peptides containing a looped structure

In contrast to other AMPs, this group of AMPs are characterized by their looped structure because of the presence of a single bond such as

disulfide or amide or isopeptide bond. Proline-arginine-rich peptides cannot form amphipathic structures due to the incompatibility of high concentration of proline residues in such structures and have been proposed to adopt a polyproline helical type-II structure (Boman et al., 1993; Cabiaux et al., 1994). Because of their short size, easy to synthesize and being proteolytically stable, this class of peptides hold considerable potential in fighting emerging infectious diseases.

Linear peptides with an extended structure

This group of AMPs include those possessing an unusual amino acid composition, having a sequence that is rich in one or more specific amino acids. Most of them are linear in shape and do not possess any secondary structure either in α-helix or in β-sheet. Examples of this group of AMPs include histatin, which is highly rich in histide residues (Brewer et al., 1998; Tsai and Bobek, 1998; Helmerhorst et al., 1999); cathelicidin, rich in proline and arginine or proline and phenylalanine (Zhao et al., 1995; Linde et al., 2001); tripticin (Lawyer et al., 1996) and indolicidin (Selsted et al., 1992), that are rich in tryptophan.

While most AMPs belong to one of the above four classes, some AMPs do not belong to any of these groups (McManus et al., 1999). Some AMPs have shown to possess two different structural components (Uteng et al., 2003). Whereas, many peptides have shown to form their active structure only when they interact with the target cell membrane. For example, indolicin shows globular and amphipathic conformation in aqueous solutions while it is wedge-shaped in lipid bilayer mimicking environments (Rozek et al., 2000). This AMP also changes its conformation during interaction with DNA evidenced with decreased

fluorescence intensity and a slight shift in the wavelength of maximum emission (Hsu et al., 2005).

Table 1.1 Various classes of AMPs based on its structure

Structure	AMP	Source organism
α-helix	Cecropins, Mellitin	Insect
	Magainin, LL-37, PGLa, Brevinin-1, Temporin	Amphibian
	PMAP-23	Porcine
β-Sheet	α and β-Defensin	Human, Rabbit
	Dermaseptin	Frog
	Tachyplesin, Polyphemusin	Horseshoe crab
	Protegrin	Human, Porcine
	Androctonin	Scorpion
	Pn-AMP 1	Plant
Cyclic β-sheet	θ-Defensin	Primate, Human
β-Turn	Lactoferricin	Bovine, Human
Linear with repeating motifs	Bactenecins 5 and 7, PR-39, Indolicidin	Mammals
	Diptericin, Apidaecin	Insects
Extended	Indolicidin, PR-39	Bovine, Porcine
Extended turn	Tenecin-3	Insect
α-helix / extended	Buforin II	Toad
Looped peptide	Bactenecin 1	Bovine
	Ranalexin	Frog
	Thanatin	Insect
Mixed structure (possessing both α-helix and β-sheet)	Defensins	Plants
	Drosomycin	Arthropods
	α- and β-defensins	Mammals
	Defensins	Arthropods, Molluscs
	γ-thionins	Plants

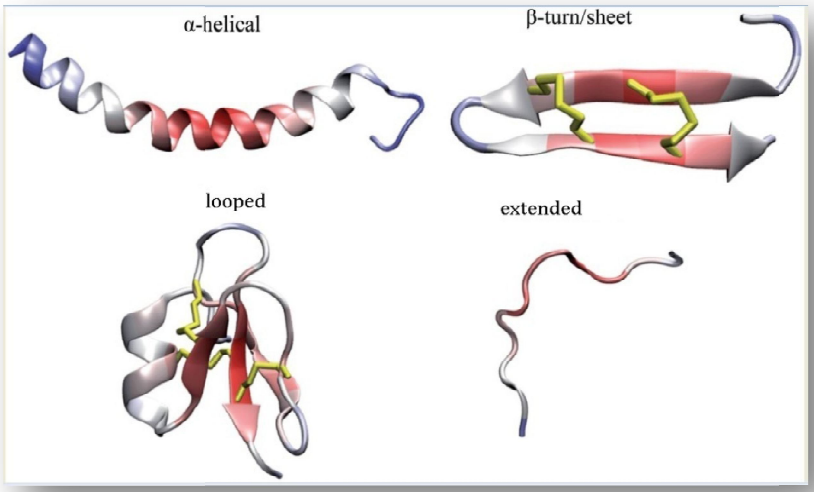


Fig. 1.1 Structural Classes of Antimicrobial Peptides

1.3.2 Classification of AMPs based on charge

Though majority of AMPs are cationic, amphipathic molecules, possessing a net positive charge of +2 to +9 (Giangasper et al., 2001; Yeaman and Yount, 2003); there does exist smaller groups of AMPs composed of anionic peptides, aromatic dipeptides, processed forms of oxygen-binding proteins, and peptides derived from neuropeptide precursors (Stefano et al., 1998; Brogden et al., 2003). The detailed list of various classes of AMPs based on charge / nature are as given in Table 1.2.

Table1.2 Various Classes of AMPs based on its charge / nature

Nature of AMP	AMP	Source organism
Anionic peptides	Maximin H5	Amphibian
	Dermcidin	Humans
Linear cationic α-helical peptides	Cecropin-A, Andropin, Moricin, Ceratotoxin, Melittin, Cecropin P1	Insects
	Magainin, Dermaseptin, Bombinin, Brevinin-1, Esculentins, Buforin	Amphibian
	Pleurocidin	Fish
	Cathelicidins, Seminalplasmin, ovispirin	Cattle, Sheep and Pig
	Cap18	Rabbit
	LI37	Human
Cationic peptides enriched for specific amino acids	Abaecin, Apidaecins, Hymenoptaecin, Drosocin, Pyrrhocoricin, Coleoptericin, Holotricin	Insect
	Bactenecins, PR-39, Prophenin, Indolicidin	Cattle, sheep, goat, pig
	Histatins	Humans
Anionic and cationic peptides that contain cysteine and form disulphide bonds	Brevinins	Amphibian
	Protegrin, β-defensins, Np-1	Pigs, Rabbit, rat, Cattle, goat and poultry
	Tachyplesins	Horseshoe crab
	HNP-1, -2, β-defensins	Humans
	θ-defensin	Rhesus monkey
	Defensin A, Drosomycin	Insect, plants
Anionic and cationic peptide fragments of larger proteins	Casocidin I, Lactoferricin	Human casein, Bovine milk
	Antimicrobial domains	Bovine α-lactalbumin, Human haemoglobin, Human lysozyme, Human ovalbumin

1.4 Mode of action

The activities of AMPs are generally dependent upon their interaction with bacterial cell membranes (Shai, 2002; Dawson and Liu, 2008). The mechanisms of action of AMPs have been widely studied and several models have been proposed to explain peptide insertion and membrane permeability. The molecular mechanism of membrane permeation and disruption by the AMPs depend on several parameters such as the amino acid sequence, membrane lipids and the concentration of the AMP. As per the structural model established by Shai-Matzusaki-Huang (Hancock, 1997; Matsuzaki, 1999; Shai 1999; Yang et al., 2000; Zasloff 2002; Brogden 2005), the cationic AMPs initially binds to the anionic components on the outer bacterial envelope, such as the phosphate groups within the lipopolysaccharides (LPS) of Gram-negative bacteria or the lipoteichoic acids on the surfaces of Gram-positive bacteria. This is followed by displacement of lipids, alteration of membrane structure, and, in certain cases, entry of the peptide into the target cell. Once peptides have reached the cytoplasmic membrane, they can interact with lipid bilayers. At low peptide / lipids ratios, peptides are bound parallel to the lipid bilayer. After attaining certain threshold concentration (Melo et al., 2009), the peptide molecules are oriented perpendicularly to the membrane and inserted into the lipid bilayer, forming transmembrane pores. AMP mediated cell killing can be rapid. Some linear α-helical peptides kill bacteria so quickly that it is technically challenging to characterize the steps (if there are any) preceding cell death (Boman, 1995). AMPs such as magainin 2 (Zasloff, 1987), cecropin P1 (Boman et al., 1993), PR-39 (Boman et al., 1993) and SMAP29 (Kalfa et al., 2001) has been found to kill bacteria in 15–90 min. Regardless of the time required,

or the specific antimicrobial mechanism, specific steps must occur to induce bacterial killing (Matsuzaki et al., 1995). Obviously first step is the attraction of the AMPs to bacterial surfaces which is mainly aided by electrostatic bonding between anionic or cationic peptides and the structures on the bacterial surface. Once close to the microbial surface, the next important step is attachment of the peptides to the bacterial membrane. Peptides must traverse capsular polysaccharides before they can interact with the outer membrane, which contains LPS in Gram-negative bacteria, and traverse capsular polysaccharides, teichoic acids and lipoteichoic acids before they can interact with the cytoplasmic membrane in Gram-positive bacteria. Once peptides have gained access to the cytoplasmic membrane they can interact with lipid bilayers.

Membrane disruptive models

Among the several models describing the mechanism of membrane disruption by AMPs, the most widely accepted ones are the barrel-stave model, the carpet model, the toroidal model and the aggregate channel model (Oren et al., 1999; van 't Hof et al., 2001; Huang et al., 2010). The pictorial representation of these models are illustrated in Figure 1.2.

Barrel-stave model: According to this model, after initial electrostatic binding to the outer membrane in bacteria, the helical amphipathic peptides reorient perpendicularly to the membrane and align like the staves in a barrel, lining the amphipathic transmembrane pores. The non-polar side chains associate with the hydrophobic fatty acid tails at the inside of the phospholipid bilayer, and the hydrophilic side-chains are pointed inward an aqueous pore or channel (Powers and Hancock 2003;

Brogden, 2005). These pores thus formed will allow leakage of cytoplasmic components and also disrupt the membrane potential.

Carpet Model: The carpet model mechanism describes that, peptides are electrostatically attracted to the anionic phospholipid head groups and align parallel to the surface of the bilayer, covering the surface in a carpet-like manner. Once a saturation point is reached, peptides are thought to disrupt the bilayer in a detergent-like manner, eventually leading to the formation of micelles. This local disturbance in membrane stability will cause the formation of cracks, leakage of cytoplasmic components, disruption of the membrane potential and, ultimately, membrane disintegration.

Toroidal-pore model: In the toroidal-pore model, peptides at a low concentration are reoriented parallel to the plane of the bilayer and insert perpendicularly into the membrane. Then they cluster into unstructured bundles that induce the lipid monolayers to bend continuously through the pore so that the water core is lined by the inserted peptides and the lipid head groups. The pores created will be responsible for leakage of ions and possibly larger molecules throughout the membrane.

Aggregate channel model: In the aggregate channel model, after binding to the phospholipid head groups, the peptides insert into the membrane and then cluster into unstructured daggregates that span the membrane. These aggregates provide channels for ion leakage through the membrane.

Non-membrane disruptive model

However, not all AMPs seem to exert their action on membranes. Several studies have shown that cell killing may also proceed with relatively little

membrane disruption because of the interaction of AMPs with putative key intracellular targets. An increasing number of peptides have been described as acting on intracellular targets in bacteria such as altering cytoplasm membrane septum formation, and inhibiting protein, cell wall, or nucleic acid synthesis (Hale and Hancock, 2007).

Certain peptides have been shown to bind with heat- shock protein DNA K and inhibit the associated ATPase activity (Kragol et al., 2001). Yet others target DNA gyrase, and interfere with normal DNA function. AMPs such as apidaecin, alters the cytoplasmic membrane septum formation, inhibit cell-wall synthesis, inhibit nucleic-acid synthesis, inhibit protein synthesis or inhibit enzymatic activity (Casteels et al., 1993); PR-39 induces cell filamentation thereby blocking cell division (Shi et al., 1996; Subbalakshmi and Sitaram, 1998); Lantibiotics inhibits peptidoglycan biosynthesis (Brotz et al., 1998); defensins (HNP-1) inhibit DNA, RNA and protein synthesis in *E. coli* (Lehrer et al., 1989; Boman et al., 1993; Subbalakshmi and Sitaram, 1998; Patrzykat et al., 2002); pleurocidin and dermaseptin inhibit nucleic acid and protein synthesis (Patrzykat et al., 2002); PR-39 hinders protein synthesis and induces degradation of some proteins that are required for DNA replication (Boman et al., 1993); Indolicidin completely inhibits DNA and RNA synthesis (Subbalakshmi and Sitaram, 1998) and histatins disrupt the cell cycle and lead to the generation of reactive oxygen species (Andreu and Rivas, 1998). Other non-membrane external targets such as autolysins and phospholipases are also activated by AMPs. These alternative mechanisms of action may act independently or synergistically with membrane permeabilization.

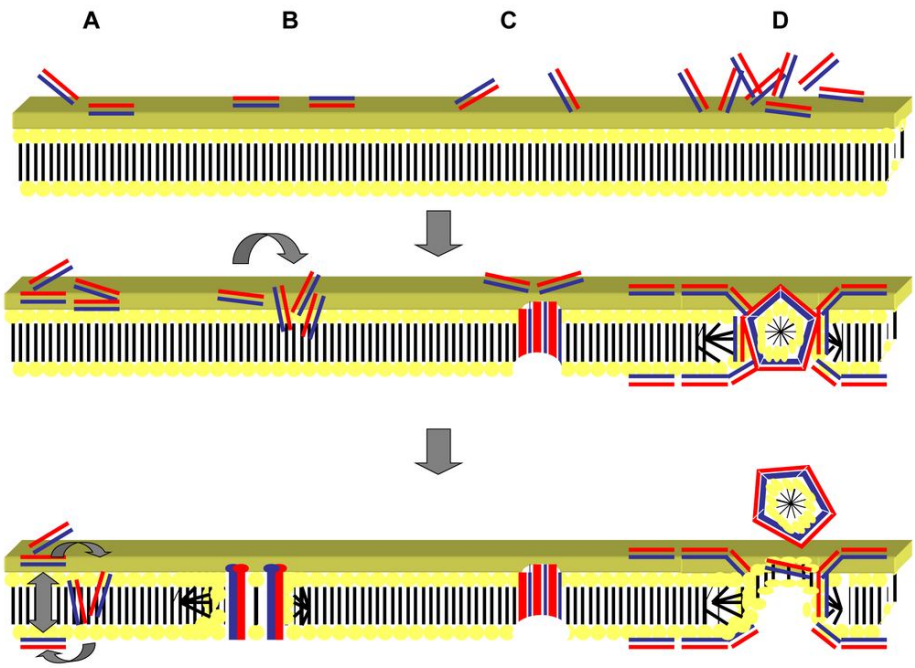


Fig.1.2 Proposed mechanisms of action of antimicrobial peptides. Antimicrobial peptides (cylinders) with the charged hydrophilic regions (red) and hydrophobic regions (blue). (**A**) The "aggregate" model: (**B**) The "toroidal pore" model (**C**) The "barrel-stave" (**D**) The "carpet" model. (Adopted from Jenssen et al., 2006a).

1.5 Biological activity of AMPs

AMPs display a broad spectrum of biological activity. They are found to be active against various microorganisms, including Gram-negative and Gram-positive bacteria, fungi, protozoa and enveloped viruses such as HIV, herpes simplex virus, WSSV and vesicular stomatitis virus as well as malignant cells (Cruciani et al., 1991; Baker et al., 1993; Fehlbaum et al., 1996; Tamamura et al., 1998; Arrighi et al., 2002; Lindholm et al., 2002; Kieffer et al., 2003).

Antibacterial activity: AMPs possess broad-spectrum activity against Gram-negative bacteria and Gram-positive bacteria (Miyasaki and Lehrer, 1998). To date, antibacterial activities of AMPs are the most studied among its biological activity. Most of these are cationic, amphipathic AMPs, which target bacterial cell membranes and cause disintegration of the lipid bilayer structure (Zhang et al., 2001; Shai, 2002).

In some cases, certain AMPs show activity against the several multi drug resistant bacteria such as *Staphyloccocus aureus, Pseudomonas aeruginosa* and *Escherichia coli* (van't Hof et al., 2001; Mygind et al, 2005; Marr et al., 2006). For example, both nisin (an AMP) and vancomycin (an antibiotic), possess the ability to block cell wall synthesis. However, *S. aureus* strain was reported to be resistant to vancomycin, while it is still sensitive to nisin (Brumfitt et al., 2002). The most active AMPs present a minimum inhibitory concentration (MIC) of 1 to 4 µg/ml, corresponding to the minimum concentration of AMPs that completely prevents bacterial growth (Hancock, 2001). Magainins 1 and 2 present a broad antimicrobial activity spectrum against both Gram-positive and Gram-negative bacteria (Zasloff, 1987). Whereas, AMPs such as androPin (Samakovlis et al., 1991) and insect defensins (Meister et al., 1997) preferentially kill Gram-positive bacteria, while apidaecin (Casteels and Tempst, 1994), drosocin (Bulet et al., 1996), and cecropin P1 (Boman et al., 1991) are known to be active against Gram-negative bacteria. Binding of AMPs to anionic lipopolysaccharides and teichoic acids, is crucial for their activity, since modification of the bacterial envelope leading to charge reduction is a common mechanism of bacterial resistance against cationic AMPs. Recently, researchers have demonstrated that some AMPs can kill bacteria even at low concentrations without changing the membrane

integrity. Instead of directly interacting with the membrane, these AMPs kill bacteria by inhibiting some important pathways inside the cell such as DNA replication and protein synthesis (Brogden, 2005). Drosocin, pyrrhocoricin, apidaecin and buforin II are examples of AMPs with an active site for their intracellular target (Park et al., 1998; Otvos et al., 2000; Kragol et al., 2001).

Antifungal activity: Several AMPs belonging to different structural classes such as α-helical, extended and β-sheet have been proved to possess antifungal properties which kill fungi by targeting either the cell wall or intracellular components (De Lucca and Walsh, 1999; Barbault et al., 2003; Lee et al., 2003, 2004; Jiang et al., 2008). Cell wall targeting AMPs kill the target cells by disrupting the integrity of fungal membranes, by increasing permeabilization of the plasma membrane, or by forming pores directly (Lehrer, et al., 1985; Terras, et al., 1992; Moerman, et al., 2002; Van der Weerden et al., 2010). AMPs such as drosomycin possess the ability to preferentially kill fungi (Meister et al., 1997). As per the database till 2010, 483 of the 'Antimicrobial Peptide Database' (APD) peptides and 570 of the 'Collection of Antimicrobial Peptides' (CAMP) peptides were listed as having antifungal activity. These peptides range from large molecules such as histone H2A of the rainbow trout, *Oncorhynchus mykiss* (13.6 kDa) (Fernandes et al., 2002), to much smaller peptides such as the jelleines (8–9 amino acids) found in honeybee (*Apis melliferia*) royal jelly (Fontana et al., 2004). Some AMPs such as penaeidin has shown to possess chitin binding ability. Chitin is one of the major components of fungal cell walls and the chitin-binding ability is most often related to an anti-fungal activity (Fujimura et al., 2004; Cuthbertson et al., 2006; Yokoyama,et al., 2009; Pushpanathan et al., 2012). Such

binding ability helps AMPs to target fungal cells efficiently. The fungal mammalian outer membrane is enriched largely with neutral phospholipids making the surface less negatively charged and hence the basis of selective targeting of peptides for fungal over mammalian membranes is much less well understood (Yeaman and Yount, 2003). The presence of ergosterol in fungal membranes and differences in electrochemical gradients across the outer membrane could play important roles in the initial selection of fungal cells (Yeaman and Yount, 2003). Although the majority of antifungal AMPs have polar and neutral amino acids in their structures, there does not appear to be a clear correlation between the structure of an AMP and the type of cells that it targets (Jenssen et al., 2006a, 2006b).

Antiparasitic activity: A very small number of AMPs have been shown to possess activity against parasites. Some examples of AMPs possessing antiparasitic activity include magainin, which is able to kill *Paramecium caudatum* (Zasloff, 1987); a synthetic AMP possessing activity against *Leishmania* parasite (Alberola et al., 2004) and cathelicidin, which is able to kill *Caernohabditis elegans* by forming pores in the cell membrane (Park, et al., 2004). Even though some parasitic microorganisms are multicellular, the mode of action of antiparasitic peptides is believed to be the same as other AMPs, i.e. by direct interaction with cell membrane (Park et al., 2004).

Antiviral activity: Antiviral activities has been proved for several cationic AMPs such as defensins against herpes simplex virus, vesicular stomatitis virus and influenza virus (Daher et al., 1986; Ganz and Lehrer, 1995); tachyplesins and polyphemusins against vesicular stomatitis virus,

influenza A virus and HIV (Tamamura et al., 1996) and melittins, cecropins and indolicidin against HIV (Wachinger et al., 1998).

Antitumor activity: AMPs have been found to kill malignant cells also. For example, magainin have been found to lyse haematopoietic tumor and solid tumor cells with little toxic effect on normal blood lymphocytes by targeting cell membrane by a non-receptor pathway (Cruciani et al., 1991; Baker et al., 1993); tachyplesin, has been showed to inhibit the proliferation of both cultured tumor and endothelia cells by disrupting their membranes and inducing apoptosis (Hirakura et al., 2002) and temporin L has been shown to induce necrosis of tumor cells (Rinaldi et al., 2001, 2002). Other AMPs, such as defensins, cecropin, lactoferricin and lactoferrin and cyclotides also possess similar antitumor activity (Kagan et al., 1990; Moore et al., 1994; Vogel et al., 2002). The activity profile of cyclotides differed significantly from those of antitumor drugs in clinical use, which may indicate a new mode of anticancer action (Lindholm et al., 2002).

Unlike bacterial, fungal or tumor cell membranes, normal mammalian membranes are rich in sterols and zwitterionic phospholipids with neutral net charge including phosphatidyl ethanol amine, phosphatidyl choline, or sphingomyelin. Apart from that, presence of significant amounts of cholesterol in mammalian membranes reduce the activity of AMPs by affecting the fluidity and dipole potential of phospholipids, in addition to stabilizing the lipid bilayers and delaying the binding of peptides to the membranes (Tytler et al., 1995; Matsuzaki, 1999). Thus, a higher proportion of negatively charged lipids on the surface monolayer of the microbial cytoplasmic membrane play an important role in the selectivity of AMPs for bacterial cells over eukaryotic cells. Though,

normal human cells are found to be relatively resistant; certain cationic AMPs, such as melittin from bees, mastoparan from wasps, charybdotoxin from scorpions and temporin L from frogs are found to be potent toxins (Perez-Paya et al., 1994; Tenenholz et al., 2000; Delatorre et al., 2001; Rinaldi et al., 2002).

1.6 Therapeutic potential of AMPs

The evolution of pathogenic organisms has resulted in increasing resistance by several against conventional antibiotics. In the present scenario, there is no question that, with the increasing antibiotic resistance problem, there is a need to develop new classes of antibiotics (Bonomo, 2000). AMPs, due their broad spectrum of activity against several species of bacteria, fungi, protozoa, enveloped virus and malignant cells, have gained increased attention as a promising therapeutic alternative against pathogenic microorganisms and are hence on the brink of a breakthrough. AMPs are found to be efficient towards multi-resistant bacteria and are not hindered by resistance. The unique mode of action and therefore the least resistance by pathogenic organisms might make them potential replacement for conventional antibiotics. The intriguing idea of developing AMPs as innovative antibiotics has been followed up by several biotechnological companies. Compared to conventional antibiotics, AMPs can kill bacteria rapidly even at low concentrations. Employing solid phase synthesis as well as recombinant DNA technology, the structures of naturally occurring peptides serve as starting points for the development of new therapeutic agents.

Advantages of AMP as therapeutic agent: As described by Altman et al. (2006) the important features that make AMPs promising candidates for clinical applications and potential alternatives to conventional antibiotics are as follows:

- Activity even in a very low concentration
- Rapid and unique mechanisms of action
- The ability to discriminate between host and microbial cells (cell selectivity)
- Activity against a broad spectrum of microorganisms, including resistant and multidrug-resistant strains
- A low propensity for developing microbial resistance

A better comprehension of AMP's mode of action and counterpart resistance mechanisms is fundamental for the design of optimized AMPs that could be efficiently used as therapeutic drugs.

Several derivatives of AMPs have been through the pharmaceutical process, including human phase I–III studies. The use of human AMPs as drug is restricted so far because of the unknown biological function of these molecules and the high cost for the generation of sufficient amount (Bals, 2000). A number of naturally occurring peptides and their derivatives have been developed as novel anti-infective therapies for conditions as diverse as oral mucositis, lung infections associated with cystic fibrosis, cancer, and topical skin infections. Pexiganan (or MSI-78) derived from magainin-2 has entered Phase III trials and proved to be effective in wound healings and did not show any notable toxicity or side-effects (Lamb and Wiseman, 1998). But its efficiency towards infected diabetic foot ulcers did not offer any improvement over the conventional

treatment with ofloxacin, a fluoroquinolone antibiotic and hence rejected by the Food and Drug Administration (FDA) in 1999 (Maloy and Kari, 1995).

Most clinical trials proposed or underway involve topical therapy. Such treatments are likely to be effective and safe because the more toxic cationic peptides and lipopeptides, including gramicidin S and polymyxin B, could be successfully included in skin creams. A more advanced form of topical treatment would be aerosol therapy into the lung. Oral therapy may also be possible for gastro intestinal infections; an example is nisin which is being developed through to clinical trial in *Helicobacter pylori* infection (Hancock, 1997).

Many AMPs are currently being tested in clinical trials. MX-226 and MX-594NA, bovine indolicidin-based AMPs have showed efficiency in Phase III clinical trials. MX-226 was developed for the prevention of catheter related infections, whereas MX-594NA was developed for the treatment of acnea vulgaris. XOMA629, an AMP deriving from the human BPI protein, which is under preclinical studies has showed promising activities against skin bacteria. P113, developed by Demegen (USA), derived from histatin has showed excellent *in vitro* activities against Gram-positive and Gram-negative bacteria and is to be used as a mouth rinse product, to fight gingivitis (Giuliani et al., 2007). Plectasin, a fungal defensin, is currently under preclinical development and was shown to be active against *Streptococcus pneumonia* (Mygind et al., 2005).

Due to the increasing importance and wide acceptance of AMPs as therapeutic agents, several companies are making efforts to introduce the AMP products to the market. Natural AMPs have potential application in

food preservation as they possess the ability to specifically kill microbial cells by destroying their unique membranes. Example for AMPs that have already been commercialized include bacteriocins against food-borne pathogens and spoilage microorganisms (Cleveland et al., 2001); nisin as a food preservative (Delves-Broughton, 1990) and pediocin PA-1 with applications in dairy and canned products (Vandenbergh et al., 1989).

The future of AMPs as potential therapeutic agents appear to be great and as mentioned above. The major considerations that will determine the therapeutic potential of AMP include toxicity, stability, immunogenicity, route of application, and formulation. However, very little information on these questions has been published. There still remain several issues that remain to be solved. AMPs possess relatively high molecular weights compared with most antibiotics and need to be produced recombinantly to keep the prices down (Hancock and Lehrer, 1998). Another important issue is the toxicity of these AMPs. Though AMPs are generally considered to be highly selective antimicrobial agents, an absolute discrimination between eukaryotes and prokaryotes, still remains uncleared. Furthermore, some cationic AMPs have been proved to be very toxic for mammalian cells (e.g. bee venom melittin), whereas others show little or no acute cytotoxicity. Another issue would be their lability to proteases produced by the human body. In this regard, there are strategies for protecting the peptides from proteases, including liposomal incorporation or chemical modification.

Technical difficulties and high production costs have made the pharmaceutical industry reluctant to invest much effort in the development of AMPs as therapeutics so far. Hence, the biggest challenge

of the near future will be to overcome the pharmacological limitations of these interesting molecules and to develop them into therapeutics.

1.7 Significance of AMPs in crustaceans

Invertebrate animals, with poor adaptive immune systems, possess a well developed innate immunity system that responds to common antigens on the cell surfaces of potential pathogens. The major defense molecules of the innate immune system include phenoloxidases, clotting factors, complement factors, lectins, protease inhibitors, antimicrobial peptides, toll receptors, and other humoral factors found mainly in haemolymph plasma and haemocytes. The innate immune system is the first line of inducible host defense against bacterial, fungal, and viral pathogens (Hoebe et al., 2004). This defense system is essential for the survival and perpetuation of all multi cellular organisms (Hoffmann et al., 1999; Salzet 2001). In invertebrates, toll-like receptor-mediated AMP production (Lemaitre et al., 1996; Krutziket al., 2001; Underhill and Orinsky, 2002), hemolymph coagulation (Iwanaga et al., 1978), melanin formation (Sugumaran, 2002), and lectin mediated complement activation are prominent immune responses. In addition to these enzyme cascades, a variety of agglutinin-lectins and reactive oxygen producing and phagocytic systems cooperate with immune reactions to kill invading pathogens (Bogdan et al., 2000).

Crustacea is the largest, most conspicuous and, arguably, the most important group of marine or aquatic arthropods in terms of their biomass and ecological or economic value. Crustaceans represent one of the most abundant animals inhabiting both aquatic and terrestrial habitats. By virtue of their diversity and abundance they have earned

considerable attention as a potential source of bioactive compounds. Crustaceans have been popular experimental animals in nearly all aspects of biology, but it is decapods that attract most attention in relation to their immune responses because of their huge commercial importance and the need to control disease outbreaks in shellfish aquaculture.

Crustaceans live in an environment where they are exposed to a large number of micro-organisms causing health hazards. The hard rigid exoskeleton present in crustaceans normally acts as a physicochemical barrier protecting the organism from microbial injury and invading pathogen. Chitinous membranes of gastro intestinal tracts also functions as protective barrier (Jiravanichpaisal et al., 2006). During adverse conditions, sometimes the natural barriers are penetrated by the pathogens and they enter in to the circulating system of the host. Such situation triggers the internal immune response of the crustaceans. Crustaceans lack the highly efficient adaptive immune system as in vertebrates, and hence solely depend on their innate immune responses to fend off invading pathogens. To survive in a potentially hostile and microbe-enriched environment crustaceans have evolved efficient innate immune system. Crustaceans such as crabs and shrimps possess a simpler and more basic immune system to protect themselves against disease-causing microorganisms.

The innate immune system of crustaceans is primarily related to their blood or haemolymph and is comprised of cellular and humoral responses (Rosa and Barracco, 2010). The haemolymph plasma of crustaceans contains many soluble defense molecules, such as hemocyanins, various lectins, and C-reactive proteins, and thioester bond containing proteins. Haemocytes, which compose more than 99 % of circulating cells, contain a

variety of defense molecules, including clotting factors, a clottable protein coagulogen, proteinase inhibitors, lectins, and antimicrobial proteins.

AMPs are one of the major components of the humoral part of innate immune defense of crustaceans. In recent years, scientists have discovered an abundance of AMPs in the body fluids of marine crustaceans. Due to the natural antibiotic property of these molecules, substantial research has been carried out on investigating their potential as therapeutic alternatives to current antibiotics. Also, AMPs are promising measures for the prevention and cure of diseases that marine crustaceans are susceptible to. This holds significant importance for fisheries and aquaculture all around the world, since marine crustaceans like crabs and shrimps are important resources for these industries.

1.8 AMPs identified from crustaceans

So far less than 10 % of all known animal AMPs are from crustaceans, with nearly all of these found in decapods (Smith et al., 2008). Different families of AMPs have been identified and characterized from crustaceans. In decapods they have been isolated from crabs (Schnapp et al., 1996), lobsters (Hauton et al., 2006), crayfish (Jiravanichpaisal et al., 2007) and several shrimp species (Destoumieux et al., 1997). From the available reports, 15 AMP families or single peptides sharing common molecular features with the currently known AMP families have been recognized in crustaceans (Rosa and Barracco, 2010). Based on amino acid composition and structure, Rosa and Barracco (2010) grouped the families of AMPs found in crustacean species into four main groups: (i) single-domain linear α-helical AMPs and peptides enriched in certain amino acids, (ii) single-domain peptides containing cysteine residues engaged in disulfide

bonds, (iii) multi-domain or chimeric AMPs and (iv) unconventional AMPs including multifunctional proteins and protein-derived fragments that exhibit antimicrobial functions.

(i) Single-domain linear α-helical AMPs and peptides enriched in certain amino acids

In crustaceans five linear AMPs have been described, including homarin from the haemocytes of the American lobster, *Homarus americanus* (Battison et al., 2008), glycine rich armadillidin from the haemocytes of the non-decapod terrestrial isopod, *Armadillidium vulgare* (Herbinière et al., 2005) and three proline/arginine-rich peptides: Bac-like from the haemocytes of the shore crab, *Carcinus maenas* (Schnapp et al., 1996), callinectin from the haemocytes of the blue crab, *Callinectes sapidus* (Khoo et al., 1999) and astacidin 2 from the haemocytes of the freshwater crayfish, *Pacifastacus leniusculus* (Jiravanichpaisal et al., 2007). These peptides lack cysteine residues engaged in disulfide linkages and some of them are enriched with a high proportion of arginine, proline, glycine, tryptophan or histidine (Tossi and Sandri, 2002).

(ii) Single-domain peptides containing cysteine residues engaged in disulfide bonds

This group is characterized by the presence of 2- 12 cysteine residues that are capable of forming intramolecular disulfide bridges and result in the formation of cyclic or open-ended cyclic stabilized peptides. Only three families of AMPs containing cysteine residues have been characterized from decapods which comprises of defensins, anti-lipopolysaccharide factors and scygonadins.

(iii) Multi-domain or chimeric AMPs

This group includes AMPs with at least two distinct domains and each domain may exhibit particular characteristics of classical single-domain AMPs, such as PRP- or cysteine-rich peptides. The chimeric structure is essential for establishing cationicity and amphipathicity of AMP. The structure is also responsible for the multifunctional and/or synergic properties in addition to its antimicrobial activities. This group includes multi-domain penaeidins, crustins, hyastatin, arasins and stylicins.

(iv) Unconventional AMPs

According to Smith et al. (2010), unconventional AMPs are multifunctional proteins that primarily serve other functions and protein fragments that display antimicrobial activity and are generated by the processing of larger proteins. It has important molecular elements that are found in the structure of classical AMPs, such as charge, hydrophobicity and/or amphipathicity (Brogden, 2005). This group includes histone-derived AMPs and haemocyanin-derived AMPs. In crustaceans, antimicrobial activities were reported for histones and histone fragments and also for peptides derived from the crustacean oxygen carrier hemocyanin.

Table 1.3 List of reported crustacean AMPs and its first description

AMP	Mol. Wt. (kDa)	Characteristics	Crustacean order	First description
Bac-like	6.5	Cationic, PRP rich	Decapoda (crab)	*Carcinus maenas*- Schnapp et al., 1996.
Penaeidin	5.5-6.6	Cationic, N-terminal PRP-rich domain and a C-terminal region with six cysteine	Decapoda (penaeid shrimp)	*Litopenaeus vannamei*- Destoumieux et al., 1997

Callinectin	3.7	Cationic, PRP rich	Decapoda (crab)	*Callinectes sapidus*- Khoo et al., 1999.
Crustin	7-14	Cationic, N-terminal WAP domain	Decapoda, Amphipoda	*C. maenas* - Relf et al., 1999
Anti-lipopolysaccharide factor	7-11	Cationic, cysteine containing	Decapoda (various)	*Litopenaeus setiferus*- Gross et al., 2001 *Penaeus monodon*- Supungul et al., 2002
Hemocyanin-derived peptides	1.9-8.3	Anionic, Cationic	Decapoda (shrimp, crayfish)	Destoumieux-Garzon et al., 2001 Lee et al., 2003
Histone derived peptides	11-15	Cationic	Decapoda (penaeid shrimp)	Patat et al., 2004
Armadillidin	5.2	Cationic	Isopoda (woodlouse)	*Armadillidium vulgare*- Herbiniere et al., 2005
Scygonadin	10.8-11.4	Anionic, cysteine containing	Decapoda (crab)	*Scylla serrata*- Huang et al., 2006
Astacidin 2	1.8	Cationic, PRP rich	Decapoda (crayfish)	*Pacifastacus leniusculus*- Jiravanichpaisal et al., 2007
Arasin	4.3-4.8	Cationic, N-terminal PRP : C- terminal 4 cysteine	Decapoda (crab)	*Hyas araneus*- Stensvag et al., 2008
Homarin	4-6	Cationic, Glycine rich	Decapoda (lobster)	*Homarus americanus*- Battison et al., 2008
Hyastatin	11.7	Cationic, N- terminal Gly: PRP : C terminal-6Cys	Decapoda (crab)	*Hyas araneus*- Sperstad et al., 2009
Defensin	6.7-7.1	Cationic, cysteine containing	Decapoda (lobster)	*Panulirus japonicus*- Pisuttharachai et al., 2009
Stylicin	8.9	Anionic, N-terminal P rich : C-terminal 13 cysteine	Decapoda (penaeid shrimp)	*Litopenaeus stylirostris*- Rolland et al., 2010

1.9 Relevance of the present study

Aquaculture industry is facing constraints in production due to severe health problems resulting in large scale mortality. Also, the emergence of pathogenic bacterial resistance to conventional antibiotics calls for an increased focus on the isolation of antimicrobials with new mechanisms of actions. AMPs are promising candidates, because their initial interaction with microbes is through binding to lipids and the interference with such a fundamental cell structure is assumed to hamper resistance development. Marine invertebrates, representing an enormous genetic and biological diversity, have proven to be a rich source for discovering potent AMPs with novel and unique structural motifs.

However, despite significant efforts in the identification and characterization of a large number of pathogens affecting crustacean aquaculture systems and an increasing number of research groups studying crustacean immunity, host defense mechanisms in shrimps remains relatively poorly understood. Crustaceans live in an environment where they are exposed to a large number of microorganisms causing health hazards. To survive in such a potentially hostile and microbe-enriched environment they have evolved efficient immune system and AMPs are considered as an important component of their innate immune system. A better understanding of the haemolymph defense system, especially with respect to AMPs, will facilitate a further development of health parameters. Therefore, the aim of the present research was to contribute to the knowledge on the functioning of the haemolymph defense system of crustaceans in terms of AMPs and its function. The present study was focused on identification of new AMP isoforms from crustaceans and to study its physical, chemical and biological

characteristics. The identification of new AMPs in crustaceans will bring interesting insight into the crustacean defense mechanisms as well as disease control in aquaculture systems.

1.10 Objectives of the present study

The present study was undertaken with the following objectives:

- Screening of antimicrobial peptides in crustaceans
- Molecular characterization and phylogenetic analysis of AMPs in crustaceans
- Heterologous expression of an AMP and its functional characterization
- Structural and functional characterization of a synthetic AMP

1.11 Outline of the thesis

The thesis is presented in seven chapters. Chapter 1 gives a general introduction on the topic of research. Chapter 2, 3 and 4 deals with the molecular characterization and phylogenetic analysis of AMPs belonging to anti-lipopolysaccharide factor (ALF), crustin and penaeidin family identified from crustaceans respectively. Chapter 5 deals with the recombinant expression of an AMP belonging to ALF family viz. SsALF2, an ALF identified from *Scylla serrata* in *E. coli* expression host system and testing its antimicrobial property. Chapter 6 deals with synthesis of an ALF identified from *Portunus pelagicus* viz. PpALF1 and testing the antibacterial and anticancer activities and cytotoxicity of the synthesized peptide. The work is summarized in Chapter 7 with special emphasis on salient findings of the study. This is followed by a list of references, GenBank accessions and publications.

2

Molecular Characterization and Phylogenetic Analysis of Anti-lipopolysaccharide Factors (ALFs) in Crustaceans

2.1 Introduction

Anti-lipopolysaccharide factor (anti-LPS factor or ALF) is one of the major families of antimicrobial peptides (AMPs), that can bind and neutralize lipopolysaccharide (LPS), thereby forming a key effector molecule of the crustacean innate immune system. ALFs were first isolated and characterized from large granular haemocytes of horseshoe crabs, *Tachypleus tridentatus* (*T*ALF) and *Limulus polyphemus* (*L*ALF) and they consisted of basic proteins of around 100 amino acids (Tanaka et al., 1982; Morita et al., 1985; Aketagawa et al., 1986). In crustaceans, ALFs were first identified from the haemocytes of the penaeid shrimp, *Litopenaeus setiferus* (Gross et al., 2001) based on EST approach. Thereafter, many ALF isoforms were identified and characterized from crustaceans including shrimps and prawns viz. *Penaeus monodon* (Supungul et al., 2002; Somboonwiwat et al., 2005), *Fenneropenaeus chinensis* (Liu et al., 2005), *Marsupenaeus japonicus* (Nagoshi et al., 2006), *Macrobrachium olfersi, Litopenaeus schmitti* and *Farfantepenaeus paulensis* (Rosa et al., 2008); crayfishes like *Pacifastacus leniusculus* (Liu

et al., 2006); crabs including *Carcinus maenas, Callinectes sapidus* (Towle and Smith, 2006), *Scylla paramamosain* (Imjongjirak et al., 2007), *Eriocheir sinensis* (Li et al., 2008; Zhang et al., 2010a; Wang et al., 2011), *Scylla serrata* (Yedery and Reddy, 2009) and *Portunus trituberculatus* (Yue et al., 2010; Xu and Liu, 2011; 2012a,b; 2013a,b) as well as lobsters viz. *Homarus americanus* (Beale et al., 2008).

Basically, ALFs are amphipathic peptides comprising of 114–124 amino acid residues with a short signal peptide sequence ranging between 16–26 residues at the N-terminal region. Following the signal peptide, within the mature peptide region, the ALFs possess a conserved LPS-binding domain, and hence are classified among single domain AMPs. The LPS-binding domain, the characteristic feature of ALFs are formed between two conserved cysteine residues which form a disulphide loop, and contain a cluster of positively charged residues within it (Hoess et al., 1993). Structure of ALF was elucidated using NMR and found that it consists of three α-helices packed against a four-stranded β-sheet (Hoess et al., 1993; Yang et al., 2009). The β-hairpin structure linked by a conserved disulfide bridge is the LPS-binding domain. This typical structure of ALFs make it capable of binding and neutralizing LPS (Aketagawa et al., 1986).

The mature peptide of ALF possesses a molecular weight of ~ 11 kDa. The theoretical isoelectric points (*p*I) of the mature peptide and LPS domain of ALFs ranges from 5–11 and 4–11, respectively. According to their *p*I, ALFs were thought to be cationic peptides, but growing evidence has revealed the existence of anionic ALFs also (Zhang et al., 2010; Rosa et al., 2013). In penaeid shrimps, ALFs have been classified into four groups (A–D) according to the sequences and range of calculated *p*I, i.e. anionic and

cationic (Group A), highly cationic (Group B), cationic (Group C) and very anionic (Group D) ALFs (Rosa et al., 2013). Cationic LPS-binding domains have been found in both basic and anionic ALFs, while anionic LPS-binding domains have been found only in anionic ALFs.

The haemocyte, which is a key cell involved in the innate immune mechanism of crustaceans, was identified as the production and storage site of ALFs (Somboonwiwat et al., 2005; Nagoshi et al., 2006; Burge et al., 2007; Ponprateep et al., 2012; Liu et al., 2012; Arockiaraj et al., 2014). But in some cases, ALF transcripts are found in many tissues and the level of expression of ALFs is found to vary with the tissues (Mekata et al., 2010; Ponprateep et al., 2012; Antony et al., 2012). Also, some ALF isoforms are predominantly expressed in other immune-related tissues, such as the gills, lymphoid organ, intestine and hepatopancreas (Antony et al., 2012; Liu et al., 2013). However, the transcriptional levels of most identified ALFs in crustaceans were usually up-regulated in response to bacterial or viral infection, suggesting that ALFs might play important roles in the defense against bacterial and viral pathogens. Expression of ALF transcripts have been demonstrated to respond rapidly upon pathogen infections or immune stimulations by LPS, lipoteichoic acid (LTA) and β-glucan injections (Antony et al., 2011). Though ALFs are found to be constitutively expressed in haemocytes and has been found to be up-regulated following a viral or bacterial challenge (Agoshi et al., 2006; Antony et al., 2011). It has been reported that the expression of ALF*Pm*3 transcripts were up-regulated by *Vibrio harveyi* infection (Somboonwiwat et al., 2005). Recently, transcription elements in the promoter of ALF gene was also studied which will be useful to know about the ALF gene regulation mechanism involved in the gene expression (Tang et al., 2014).

Distinct ALF isoforms usually coexist in one species and exhibit different expression patterns. For instance, six isoforms of ALF (ALF*Pm*1-6) could be identified from *P. monodon* which was constitutively expressed in the haemocytes, heart, gill, intestine and lymphoid organ (Supungul et al., 2004; Ponprateep et al., 2012). Seven ALF isoforms (*Pt*ALF1-7) were isolated from *P. trituberculatus* which showed distinct expression profiles. *Pt*ALF1-3, *Pt*ALF5, and *Pt*ALF7 were mainly expressed in haemocytes, while *Pt*ALF4 and *Pt*ALF6 showed high expression levels in eyestalk and gill (Xu and Liu., 2011; Liu et al., 2012a, 2012b, 2013). Although different isoforms have been reported in different crustacean species, their similarities in the same species or among different species were very low (Li et al., 2013).

ALF possess the ability to bind to microbial polysaccharides including LPS, LTS and β-glucan, which is an evidence for its antimicrobial activity (Sun et al., 2011). Studies on *in vivo* and *in vitro* activities of synthetic LPS domain of ALFs and recombinantly produced ALFs have been carried out against various pathogens (Liu et al., 2006; de la Vega et al., 2008; Pan et al., 2009; Zhang et al., 2010). The synthetic disulfide loops of ALF isoforms from *P. monodon* (ALF*Pm*3) and *S. paramamosain* (ALF*Sp*) have been clearly shown to be active against broad range of microbes (Somboonwiwat et al., 2005; Imjongjirak et al., 2007). ALFs have the ability to inhibit the endotoxin or LPS mediated coagulation system, and thus exhibits strong antibacterial activity against the Gram-negative bacteria in particular (Morita et al., 1985).

The antimicrobial activity of highly cationic ALFs against Gram-negative bacteria was first characterized for the *L. polyphemus* ALF, (*L*ALF or *Lp*ALF) (Morita et al., 1985). Wang et al. (2002) have proved the strong antibacterial activity of *L*ALF, especially on the growth of Gram-negative R type bacteria. Highly cationic ALFs such as ALF*Pm*3 from *P. monodon* exhibited a broad spectrum of activities against Gram-negative and Gram-positive bacteria, fungi and viruses (Somboonwiwat et al., 2005; Liu et al., 2006; Carriel-Gomes et al., 2007; de la Vega et al., 2008; Tharntada et al., 2009; Liu et al., 2012a). Cationic ALFs, such as *Pc*ALF1 (*P. clarkii*), *Sp*ALF1 and *Sp*ALF2 (*S. paramamosain*), and anionic ALFs, such as *Es*ALF2 (*E. sinensis*), *Pt*ALF6 (*P. trituberculatus*), *Mr*ALF(*M. rosenbergii*) and *Sp*ALF4 (*S. paramamosain*), have been shown to exhibit lower antibacterial activity against Gram-negative and Gram-positive bacteria (Zhang et al., 2010; Sun et al., 2011; Liu et al., 2013, 2014; Zhu et al., 2014). On the other hand, very anionic ALFs (shrimp ALF group D) have an impaired LPS binding domain and hence are devoid of antimicrobial activity (Rosa et al., 2013). The absence of a positively charged amino acid cluster in the anionic LPS binding domain is considered as the reason for the deficient LPS binding ability of this group of ALFs.

In addition to antibacterial activity, ALFs were known to posses antifungal (de la Vega et al., 2008), antiviral (Antony et al., 2011), immunomodulatory (Montero et al., 2003), anti-tumour (Lin et al., 2013) and anti-inflammatory (Pan et al., 2009) activities. Another possible role of the ALF as an antiviral molecule has been shown in *P. leniusculus* as it interferes with the replication of white spot syndrome virus (Liu et al., 2006). Also, ALFs have been proved to enhance phagocytic activity of haemocytes. As described by Liu et al. (2014), *Mr*ALF in the plasma and

haemocytes at a lower concentration than the MIC, can function as an opsonin, thereby enhancing phagocytosis. ALF also has the ability to bind and neutralize LPS-induced gelation of amebocyte lysate in *Limulus* (Alpert et al., 1992).

ALF*Pm*3, one of the most abundantly expressed isoform of ALF identified from *P. monodon*, was cloned and expressed in a *Pichia pastoris* expression system and its antimicrobial activity was investigated by Somboonwiwat and co-workers in 2005. It was found to possess broad spectrum of activity against several classes of microorganisms, including Gram-positive and Gram-negative bacteria and filamentous fungi. Moreover, the recombinant ALF*Pm*3 has also been demonstrated to possess an antiviral effect against human herpes virus (HSV-1) (Carriel-Gomes et al., 2007).

In fact, increased resistance of bacteria towards antibiotic drugs has stimulated intensive effort for discovery and characterization of AMPs as sources or templates for the design of new therapeutic antibiotics. With strong and broad spectrum antimicrobial activity, the ALFs and their derivatives are becoming potential therapeutic agents for prophylactic treatment of viral and bacterial infectious diseases, as well as for septic shocks (Vallespi et al., 2000; Pan et al., 2007). As most of the infectious bacterial pathogens in aquatic organisms are Gram-negative bacteria, studies on ALFs will be more useful in disease control in the aquaculture systems. Most of the cultured crustaceans are prone to infection by microbes at various stages of its growth; the losses due to disease can be enormous. As a result, a considerable research has been focused in isolation and characterization of proteins involved in immune regulation. Hence, the present study was focused on the discovery of new ALF

isoforms from various crustacean species through molecular approach and to investigate their characteristics and phylogeny so as to understand the role of ALFs in the innate immune defense of crustaceans. Since crustaceans are important from economic point of view, understanding their innate immunity will go a long way in better exploitation and management of these organisms.

2.2 Materials and Methods

2.2.1 Experimental animals

The experimental organisms used in the present study include the mud crab, *Scylla serrata* and the blue swimmer crab, *Portunus pelagicus* belonging to decapod crustaceans of the brachyuran family, Portunidae. Both of these organisms are edible and demand a high economic value in the international market. Live specimens of *S. serrata* and *P. pelagicus* were collected from Vypeen and Munambam area in Kochi respectively and were brought to the laboratory (Fig. 2.1 a and b).

2.2.2 RNase control

Following were the basic precautions taken when working with RNA:

a) Using diethyl pyrocarbonate (DEPC) treated water, glass wares, homogenizers, scissors, forceps and gloves. DEPC treatment involved overnight incubation of the utensils and solutions with 0.1 % DEPC at room temperature, to make them RNase free, followed by autoclaving at 15 lb pressure for 1 h, to remove residual DEPC.

b) Wearing gloves throughout the experiments to prevent contamination from RNases found on human hands

c) Having a dedicated set of pipettes that are used solely for RNA work

d) Using filter tips and tubes that are tested and guaranteed to be RNase-free and

e) Using RNase-free chemicals and reagents

These precautionary measures helped to minimize RNase contamination problems.

2.2.3 Haemolymph collection

Haemolymph was collected from the base of abdominal appendages of *S. serrata* and *P. pelagicus* using specially designed capillary tubes (RNase-free) rinsed in pre-cooled anticoagulant solution (RNase free, 10 % sodium citrate, pH 7.0). Haemolymph was suspended in TRI reagent (Sigma) for total RNA isolation.

2.2.4 RNA isolation

Total RNA was extracted from the haemolymph of the experimental organisms using TRI™ reagent (Sigma) following manufacturer's protocol. In brief, about 1 ml haemolymph was homogenized in 1 ml TRI Reagent in a tissue homogenizer. To ensure complete dissociation of the nucleoprotein complexes, samples were allowed to stand for 5 min at room temperature. Following that, 0.2 ml chloroform per 1 ml TRI reagent was added to the homogenate, shaken vigorously for 15 sec and allowed to stand for 2-15 min at room temperature. The resulting mixture was centrifuged at 12,000 xg for 15 min at 4 °C. Centrifugation separates the mixture into three phases: a red organic phase (containing protein), an interphase (containing DNA) and a colourless upper aqueous phase

(containing RNA). The aqueous phase was then transferred to a fresh tube and 0.5 ml isopropanol per ml TRI Reagent was added. The samples were allowed to stand for 5-10 min at room temperature and was further centrifuged at 12,000 xg for 10 min at 4 °C. The RNA precipitated out in the side and bottom of the tube was washed by adding 1 ml 75 % ethanol per 1 ml of TRI Reagent. The samples were then vortexed and centrifuged at 7500 xg for 5 min at 4 °C. RNA pellets were dried for 5-10 min and dissolved in RNase free water by repeated pipetting with a micropipette at 55-60 °C for 10-15 min.

2.2.5 Determination of the quantity and quality of RNA

RNA was quantified and qualified by measuring optical density (OD) at 260 and 280 nm in a UV spectrophotometer and visualizing RNA by agarose gel electrophoresis. The ratio of absorbance at 260 nm and 280 nm is an indication of RNA quality. Only RNAs with absorbance ratios (A_{260}:A_{280}) greater than 1.8 were used for cDNA synthesis. For quantification of RNA, the OD at 260 nm was taken and the concentration of RNA was calculated as follows:

1 OD of RNA = 40 µg/ml

RNA concentration (µg/ml) = OD at 260 nm x Dilution factor x 40

2.2.6 cDNA synthesis

Complementary DNA (cDNA) synthesis was performed using good quality RNA in a reaction catalyzed by the enzyme, reverse transcriptase. First strand cDNA was produced in a 20 µl reaction volume containing 5 µg total RNA, 1x RT buffer, 2µM dNTP, 2µM oligo d(T_{20}), 20 U of RNase inhibitor and 100 U of MMLV reverse transcriptase. The reaction was

carried out at 42 °C for 1 h followed by an inactivation step at 85 °C for 15 min. The cDNA synthesized was stored at -20 °C until use.

2.2.7 PCR amplification

PCR amplification of cDNA were carried out using gene specific primers, designed using GeneTool software based on consensus sequences of ALFs retrieved from GenBank. β-actin was used as the internal control to verify reverse transcription reaction. The sequences of the primers used to amplify β-actin and ALF genes are given in Table. 2.1.

PCR amplification of 1 µl cDNA was performed in a 25 µl reaction volume containing 1x standard Taq buffer (10 mM Tris-HCl, 50 mM KCl, pH 8.3), 200 µM dNTPs, 0.4 µM each primer and 1U Taq DNA polymerase. The thermal profile used was an initial denaturation at 94 °C for 2 min followed by 35 cycles of denaturation at 94 °C for 15 sec, and extension at 60 °C for 30 sec and a final extension at 68 °C for 10 min, for the target genes. Annealing temperature varied for the different genes as given in Table 2.1. The PCR products were visualized by electrophoresis in 1.5 % agarose gel.

2.2.8 Agarose gel electrophoresis

Electrophoresis was performed in 1.5 % agarose gel prepared in 1 x TBE buffer (Tris-base- 10.8 g, 0.5 M EDTA– 4ml, Boric acid– 5.5 g, Double distilled water– 100 ml, pH– 8.0). To the melted agarose, 2 µl ethidium bromide (1 mg/ml stock stored in dark) was added. After cooling to 45 °C, the agarose was poured on to gel tray and was allowed to solidify. The gel tray with solidified agarose was then submerged in 1x TBE buffer, filled in a buffer tank. PCR product (10 µl) was mixed with 2 µl of 6x gel loading buffer (1 % bromophenol blue– 250 µl, 1 % xylene cyanol– 250 µl,

glycerol– 300 µl, double distilled water– 200 µl) and loaded onto the well. Electrophoresis was done at a voltage of 3-5 volt/cm till the bromophenol blue dye front migrate to the middle of the gel. The gel was visualized on a UV transilluminator using the Gel Doc XR system and documented using Quantity One software (Bio-Rad Hercules, Ca).

2.2.9 Cloning of the PCR product

2.2.9.1 Ligation

The PCR products were cloned onto the pGEM-T Easy vector (Promega, USA). The ligation mix (10 µl) consisted of 5 µl ligation buffer (2x), 0.5 µl of the vector (50 ng/ µl), 3.5 µl PCR product (600 ng/ µl) and 1 µl of T4 DNA ligase (3 U/ µl). The ligation mix was incubated at 22 °C for 1 h.

2.2.9.2 Preparation of competent cells

E. coli DH5α cells were used as the cloning host for transformation of pGEM-T Easy vector with the insert. Competent cells of *E. coli* DH5α were prepared as follows. Briefly, a single colony of *E. coli* DH5α was inoculated to 10 ml LB media and grown overnight at 37 °C at 150 rpm. An aliquot of 5 ml of overnight culture was inoculated into 50 ml LB and incubated at 37 °C for 2 h at 250 rpm. The cells in their log phase were centrifuged at 6000 rpm for 20 min at 4 °C. The supernatant was decanted and the pelletized cells were then re-suspended by gentle vortexing with 100 mM $CaCl_2$. The cells were placed on ice for 45 min with intermittent swirling and mixing. The cells were then centrifuged at 6000 rpm for 20 min at 4 °C. The supernatant was decanted and cell pellet was again resuspended in 1 ml of 100 mM $CaCl_2$. The resulting competent cells were then stored at -80 °C supplemented with 15 % glycerol.

2.2.9.3 Transformation

E. coli DH5α competent cells were used for transforming the ligated cloning vector. In brief, 10 µl of the ligated product was mixed with 50 µl of competent cells in a polypropylene tube and was incubated on ice for 20 min. The cells were given a heat-shock for 45 sec in a water bath at exactly 42 °C without shaking, to facilitate the entry of recombinant DNA to the host cells. Then the tubes were immediately returned to ice for 2 min. To this, 600 µl SOC medium was added at room-temperature and incubated for 1.5 h at 37 °C with shaking at 250 rpm. The transformation mixture (200 µl) was then spread onto LB agar plates supplemented with ampicillin (100 µg/ml), IPTG (100 mM) and X-gal (80 µg/ml). The plates were then incubated at 37 °C overnight. The transformants were selected based on blue/white colony screening. The white colonies were selected and streaked on LB+Ampicillin+X-gal+IPTG plates and incubated overnight at 37 °C.

2.2.9.4 Confirmation of the presence of insert by colony PCR

To confirm the presence of the insert DNA in the transformed cells, colony PCR was performed for all the individually streaked colonies using vector specific primers viz. T7 and SP6 as well as gene specific primers (Table 2.1). White colonies (template) picked from the transformed plate were dispensed into the PCR reaction mix (25 µl) containing 1x standard Taq buffer (10 mM Tris-HCl, 50 mM KCl, pH 8.3), 200 µM dNTPs, 0.4 µM each primer and 1U Taq DNA polymerase. The thermal profile used was an initial denaturation at 95 °C for 5 min followed by 35 cycles of denaturation at 94 °C for 15 sec, annealing at 57 °C for 20 sec and extension at 72 °C for 1 min and a final extension at 72 °C for 10 min. PCR

products were then visualized by electrophoresis and documented using Gel documentation system (Gel Doc™ XR+ imaging system, Bio-Rad, USA). The clones with the required size inserts were segregated for further use.

2.2.9.5 Plasmid extraction and purification

Positive colonies with the insert were propagated in LB broth supplemented with ampicillin (100 µg μl^{-1}) by incubating at 37 °C with shaking at 250 rpm. Plasmid with the insert was then extracted and purified using GenElute HP' plasmid MiniPrep kit (Sigma) following manufacturer's protocol. Briefly, cells were harvested by centrifugation of 2 ml overnight recombinant *E. coli* culture at 12000 xg. The pellet was then resuspended in 200 µl resuspension solution supplemented with RNase. The resuspended cells were then lysed by adding 200 µl of the lysis buffer supplied with the kit. The contents were then mixed immediately by gentle inversion until the mixture becomes clear and viscous. The cell debris was then precipitated by adding 350 µl of the neutralization buffer and was pelletized by centrifugation at 12000 xg for 10 min. Columns were prepared by inserting a GenElute HP MiniPrep binding column into the provided microcentrifuge tubes. About 500 µl of the column preparation solution was then added to this MiniPrep column and centrifuged at 12000 xg for 1 min. The flow through liquid was then discarded and the cleared lysate was transferred to the column provided with the kit and subjected to washing steps twice with wash solution I and II respectively. Plasmid DNA was then eluted with 100 µl of elution solution (10 mM Tris-HCl) by centrifuging at 12000 xg for 1 min. The DNA present in the eluate was then stored at -20 °C. Plasmid DNA obtained was analyzed and confirmed by agarose gel electrophoresis for the presence of the target insert and were sent for sequencing at SciGenom, Kochi, India.

2.2.9.6 Sequencing and sequence analysis

The plasmid DNA were sequenced using ABI Prism Big Dye Terminator Cycle Sequencing Ready Reaction kit using T7 and SP6 primers on an ABI Prism 377 DNA sequencer (Applied Biosystem) at SciGenom, Kochi, India. The nucleotide sequences obtained were analyzed using BioEdit and GeneTool software. Open reading frame of the sequences were found out using GeneTool software. Gene translation and prediction of deduced protein were performed with ExPASy (http://www.au.ExPASy.org/) and GeneTool software. Molecular weight and isoelectric point of the target AMP were predicted using PROTPARAM tool of ExPASy program. Conserved domains/motifs/pattern search were also found out using MotifScan and ScanProsite tool of ExPASy program. Homology searches of the nucleotide and deduced amino acid sequences were performed using BLASTn and BLASTp tools (Altschul et al., 1990) at the National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/blast). The signal peptide was predicted by SignalP program (http://www.au.ExPASy.org/). Similar sequences were retrieved from the GenBank database and multiple sequence alignment of the peptides was performed with ClustalW, BioEdit and Genedoc computer programs. Amino acid sequences of the target genes were retrieved from the NCBI GenBank and phylogenetic tree was constructed by the Neighbor-Joining (NJ) method using MEGA version 5.05 (Tamura et al., 2007). Bootstrap analysis was based on 1000 replicates. The nucleotide sequence and the deduced amino acid sequences were submitted to GenBank.

The physico-chemical properties and amino acid composition of the peptide were predicted and confirmed using PROTPARAM, ExPASy and APD programs (http://www.au. expasy.org/; http: //aps.unmc.edu /AP/main.php). The hydrophobicity (<H>) and hydrophobic moment (<μH>) were calculated by using HELIQUEST program (http://heliquest.ipmc.cnrs.fr). To find out the amphipathicity of the peptide, a helical wheel representation of the peptide was also generated using Schiffer-Edmundson helical wheel analysis program. Data for secondary structure and spatial organization of the obtained amino acid sequence was generated using SWISS – MODEL, a homology based protein modelling server (Guex and Peitsch, 1997; Schwede et al., 2003 and Arnold et al., 2006). Using the data generated by SWISS-MODEL, 3-dimensional structure of obtained amino acid sequence was predicted using PyMOL software. The secondary structure analysis of the peptides was also carried out using STRIDE (http://webclu.bio.wzw.tum.de/cgi-bin/stride/stridecgi.py).

2.3 Results

In the present study, four different isoforms of ALFs could be identified, cloned and characterized from *S. serrata* (SsALF1 and SsALF2) and *P. pelagicus* (PpALF1 and PpALF2) (Table 2.2).

2.3.1 Molecular characterization of ALF genes in *S. serrata*

From the haemocytes of the mud crab, *S. serrata*, two new isoforms of AMPs belonging to ALF family could be identified and characterized. The ALFs are designated as SsALF1 and SsALF2.

2.3.1.1 SsALF1

2.3.1.1.1 Cloning and sequencing of SsALF1 cDNA

An initial RT-PCR reaction with *S. serrata* haemocyte RNA and gene specific primer ALF-1, a product ~ 477 bp could be amplified which was subsequently cloned and sequenced. The full-length cDNA of SsALF1 obtained consisted of 477 bp containing an ORF of 369 bp (Fig. 2.2). Based on the deduced polypeptide sequence, SsALF1 was found to consist of 123 amino acids, including a 26 amino acid signal sequence predicted by SignalP software (ExPASy) in the N-terminal region of the polypeptide chain (Fig. 2.3). Therefore, the mature peptide comprised of 97 amino acid residues with a predicted molecular weight (MW) of 11.2 kDa. The SsALF1 was found to be highly cationic and the isoelectric point (p*I*) was estimated to be 9.95 as predicted by the PROTPARAM software. The SsALF1 cDNA sequence and deduced amino acid sequence has been submitted to the NCBI GenBank under the accession number **HQ638024**.

2.3.1.1.2 Analysis of SsALF1 sequence and peptide characteristics

Sequence comparison of SsALF1 amino acid revealed a 24 amino acid conserved domain from amino acid residues 54 through 77 in the region of LPS binding domain (Fig. 2.2). The SsALF1 molecule also showed the conservation of two cysteine residues at positions Cys^{55} and Cys^{76}, important for one disulfide bond (loop) formation in the peptide (Fig. 2.2). The deduced amino acid sequence of SsALF1 was found to be rich in positively charged amino acid residues, arginine (10.3 %) and lysine (7.2 %) forming a cluster within the disulfide loop regarded as the functional domain of ALF. The physico-chemical properties of SsALF1 including its molecular weight, net charge, isoelectric point and hydrophobicity

predicted using PROTPARAM, ExPASy, APD and HELIQUEST programs are described in Table 2.3. The SsALF1 was found to be highly cationic with a net charge of +9. Hydrophobicity was found to be 37 % as predicted by the PROTPARAM programs. Among the residues taken for analysis, 37 % of amino acids were found to be hydrophobic residues and its total net charge was determined to be +9, confirming the antimicrobial nature of the peptide. Also, the peptide was predicted to interact with membranes and possess a chance to be a potential AMP (result based on APD program http://aps.unmc.edu/AP/main.php). The α-helix wheel projections for SsALF1 constructed using Schiffer-Edmundson helical wheel analysis clearly showed a perfect amphipathic nature of SsALF1 (Fig. 2.4). Analysis of the mature SsALF1 sequence using the Kyte-Doolittle plot also demonstrated a significant presence of hydrophobic amino acids concentrated in the first 20 residues of the protein (Fig. 2.5). The analysis confirmed the peptide to be amphipathic in nature due to the presence of hydrophobic residues.

Peptide model of SsALF1 created using SWISS-MODEL server showed that SsALF1 consisted of two α-helices crowded against a four-strand β-sheet. Two of the β-strands are in turn linked by a disulfide bond to form an amphipathic loop rich in cationic amino acid side chains. Secondary structure of SsALF1 was created in PyMOL software using data generated in pdb format by SWISS-MODEL (Fig. 2.6).

2.3.1.1.3 Sequence alignment and phylogenetic analysis of SsALF1

When compared with the complete non-redundant GenBank database using BLAST program, the translated SsALF1 sequence demonstrated similarities to the ALFs present in decapod crustaceans (Table 2.4).

SsALF1 was found to be 93 % similar to an ALF isoform characterized from *S. serrata*. However, a 100 % similar ALF molecule could be found in *S. paramamosain*. SsALF1 also shared similarity to ALFs of *P. trituberculatus* (76 %), *P. leniusculus* (52 %) and *F. indicus* (41 %). Multiple alignment performed for SsALF1 with other ALFs revealed the presence of conserved regions within the sequence. Sequence comparison of amino acids of SsALF1 revealed conserved residues in the region of LPS binding domain (Fig. 2.7).

The phylogenetic relationship between SsALF1 and other ALFs of decapod crustaceans were analyzed using neighbor-joining (NJ) method (Fig. 2.8). Molecular phylogenetic tree based on amino acid sequences suggests that all the ALF members possess the same ancestral origin, which has subsequently diverged at different phases of evolution. The tree could be broadly divided into two major groups, Group 1 included ALFs from shrimps and lobsters and Group 2 consisted of ALFs from crabs and crayfishes respectively. The bootstrap distance tree clearly indicated that the SsALF1 possessed great similarity to ALFs of other crabs.

2.3.1.2 SsALF2

2.3.1.2.1 Cloning and sequencing of SsALF2 cDNA

A 522 bp nucleotide sequence containing the complete coding sequence of SsALF2 could be cloned and sequenced from *S. serrata* (Fig. 2.9). The sequence possessed an ORF of 372 bp encoding a polypeptide of 123 amino acid residues. The obtained ALF sequence encoded for precursor molecules, starting with a signal peptide containing 26 amino acid residues, followed by a highly cationic mature peptide with 97 amino acid residues (Fig. 2.10). The predicted molecular mass of SsALF2 was 11.2

kDa with an estimated *p*I of 10.0. The SsALF2 cDNA sequence and deduced amino acid sequence has been submitted to the NCBI GenBank under accession number **JQ899453**.

2.3.1.2.2 Analysis of SsALF2 sequence and peptide characteristics

The N-terminus of SsALF2 had the consistent features of a signal peptide as defined by SignalP program analysis, with a putative cleavage site located after position 26 (CEA-QY). The mature peptides of SsALF2 contained the two characteristic conserved cysteine residues forming a disulphide loop, which possess a cluster of positive charges within it. These conserved cysteine residues were found at the positions C^{29} and C^{50} in SsALF2.

In SsALF2, the lysine and arginine percentages were 12.4 % and 6.2 % respectively. In addition, this molecule contained the highly hydrophobic N-terminal region and a consensus amino acid sequence (W/T)CP(G/S)W, but contained only 9 hydrophobic residues in a total of 35 residues. SsALF2 showed the typical pattern of alternating hydrophobic and hydrophilic residues in their putative disulphide loop. Out of the 20 positively charged amino acids, 10 positively charged amino acids and a tryptophan (Trp^{75}) were clustered within the disulfide loop consisting of 22 residues. These features suggest that SsALF2 possess an amphipathic factor.

The physico-chemical properties of SsALF2 including its molecular weight, net charge, isoelectric point and hydrophobicity predicted using PROTPARAM, ExPASy, APD and HELIQUEST programs are described in Table 2.5. SsALF2 was found to be highly cationic with a net charge of +10. Hydrophobicity was found to be 34 % as predicted by the

PROTPARAM and HELIQUEST programs. The antimicrobial property of SsALF2 was predicted using APD and HELIQUEST programs. Among the residues taken for analysis, 34 % of amino acids were found to be hydrophobic residues and its total net charge was determined to be +10, thereby confirming the antimicrobial nature of the peptide. The peptide was predicted to interact with membranes and possess a chance to be a potential AMP (result based on APD program http://aps. unmc. edu/ AP/ main.php). The α-helix wheel projections for SsALF2 constructed using Schiffer-Edmundson helical wheel analysis clearly showed a perfect amphipathic nature of SsALF2 (Fig. 2.11). The analysis confirmed the peptide to be amphipathic in nature due to the presence of hydrophobic face. Analysis of the mature SsALF2 sequence using the Kyte-Doolittle plot demonstrated a significant presence of hydrophobic amino acids concentrated in the first 20 residues of the protein (Fig. 2.12).

The spatial structure of SsALF2 was created in PyMOL software using data generated in pdb format by SWISS-MODEL (Fig. 2.13). Analysis of the results showed the presence of four β-strands and two α-helices in the spatial structure of SsALF2.

2.3.1.2.3 Sequence alignment and phylogenetic analysis of SsALF2

BLAST analysis of the deduced amino acid sequences of SsALF2 revealed significant identities only with the ALFs of crustaceans and not with any other groups including Limulids (Table 2.6). SsALF2 shared a significant identity with portunid ALFs (76–93 %) and shared low homology with the ALFs of *E. sinensis*, crayfishes and shrimps (less than 61 %).

Multiple alignment of SsALF2 with members of ALF family revealed that two cysteine residues (C^{29} and C^{50}) were highly conserved in portunid

ALFs (Fig. 2.14). Twenty residues (H^{30}, R^{33}, K^{34}, P^{35}, K^{36}, F^{37}, R^{38}, K^{39}, F^{40}, K^{41}, L^{42}, Y^{43}, H^{44}, E^{45}, G^{46}, K^{47}, F^{48}, W^{49}) found in between the cysteine residues (C^{29} and C^{50}) in the β-strand of portunid ALFs were also found to be conserved for the SsALF2, except in case of I^{31} and R^{32}.

To evaluate the molecular evolutional relationships SsALF2 against other ALF family members, a phylogenetic tree was constructed based on the 45 amino acid sequences of ALF members by the NJ method (Fig. 2.15). Phylogenetic analysis of ALFs further revealed that ALF sequences were clustered according to species. Phylogenetic trees also showed that SsALF2 was phylogenetically ancient immune effector molecule which may play an essential role in the host defense mechanism. SsALF2 was closely related to other crab ALFs rather than to the other groups such as shrimps, lobsters and crafishes. There were three distinct groups in the phylogenetic tree. The first and third groups were shrimp and crab ALFs respectively. In the second group, ALFs from all crustaceans viz. crabs, shrimps, lobsters and crayfishes were clustered together. Within this cluster, there were three distinct subgroups of sequences: a subgroup of lobster and shrimp ALFs, a second subgroup of shrimp, crayfish and crab ALFs to which the SsALF2 belonged to and a third subgroup containing ALFs from the crabs alone. Phylogenetic tree showed that ALFs from crabs including SsALF2 and ALFs from *P. trituberculatus, S. paramamosain, S. serrata* and *E. sinensis* were clustered together, and then had a closer relationship with crayfish *P. leniusculus* and shrimp ALFs (*P. monodon* and *M. japonicus*). This convergence between crustaceans might imply the complex evolution of ALFs in these groups and potential similarity in biological functions.

2.3.2 Molecular characterization of ALF genes in *P. pelagicus*

The present study reported the first ALF isoforms from *P. pelagicus*, named as PpALF1 and PpALF2 respectively.

2.3.2.1 PpALF1

2.3.2.1.1 Cloning and sequencing of PpALF1 cDNA

An 880 bp fragment cDNA encoding 293 amino acids and an ORF of 123 amino acids was obtained and successfully cloned from the mRNA of *P. pelagicus* haemocyte by RT-PCR (Fig. 2.16). Analysis with the SignalP software revealed the presence of a putative signal peptide with 24 amino acid residues at the N-terminal region of the PpALF1 (Fig. 2.17). The mature peptide consisted of 123 amino acid residues with a predicted molecular weight (MW) of 13.86 kDa. The PpALF1 was highly cationic and the isoelectric point (p*I*) was estimated to be 8.49 as predicted by the PROTPARAM software. The PpALF1 cDNA sequence and deduced amino acid sequence has been submitted to the NCBI GenBank under accession number **JQ745295**.

2.3.2.1.2 Analysis of PpALF1 sequence and peptide characteristics

The deduced amino acid sequence of PpALF1 was found to be rich in positively charged amino acid residues viz. arginine (10.6 %) forming a cluster within the disulfide loop. PpALF1 was found to be cationic with a net charge of +2. Hydrophobicity was found to be 39 % as predicted by the PROTPARAM and HELIQUEST programs. Among the residues taken for analysis, 39 % of amino acids were found to be hydrophobic residues and its total net charge was determined to be +2, thereby confirming the antimicrobial nature of the peptide. The deduced amino acid sequence of

PpALF1 showed a 24 amino acid domain from residue 54 through 77 that was necessary for LPS binding and neutralization.

The physico-chemical properties of PpALF1 including its molecular weight, net charge, isoelectric point and hydrophobicity predicted using PROTPARAM, ExPASy, APD and HELIQUEST programs are described in Table 2.7. The antimicrobial property of PpALF1 was predicted using APD and HELIQUEST programs. The results clearly indicated that the synthetic peptide taken for analysis possesses antimicrobial property. The peptide was predicted to interact with membranes and possess a chance to be a potential AMP (result based on APD program http://aps.unmc.edu/AP/main.php). The α-helix wheel projections for PpALF1 constructed using Schiffer-Edmundson helical wheel analysis clearly showed a perfect amphipathic nature of PpALF1 (Fig. 2.18). The analysis confirmed the peptide to be amphipathic in nature due to the presence of hydrophobic face. Analysis of the mature PpALF1 sequence using the Kyte-Doolittle plot demonstrated a significant presence of hydrophobic amino acids concentrated in the first 20 residues of the protein (Fig. 2.19).

Secondary structure of PpALF1 was created in PyMOL software using data generated in pdb format by SWISS-MODEL (Fig. 2.20). Analysis of the results showed that PpALF1 consisted of two α-helices crowded against a four-strand β-sheet. Two of the β-strands are in turn linked by a disulfide bond to form an amphipathic loop rich in cationic amino acid side chains.

2.3.2.1.3 Sequence alignment and phylogenetic analysis of PpALF1

BLAST analysis of the nucleotide sequences revealed the relation of PpALF1 to other ALFs present in decapod crustaceans (Table 2.8). PpALF1 showed only 46 % similarity to ALF isoforms present in *P.*

trituberculatus followed by 39 % similarity to ALF of *E. sinensis* and 38 % similarity to ALFs of *S. paramamosain* and *S. serrata.*

Multiple alignment performed for PpALF1 with other ALFs revealed the presence of conserved regions within the sequence. Sequence comparison of PpALF1 peptide revealed that it possessed a signal peptide and conserved amino acid residues in the region of LPS binding domain (Fig. 2.21). The PpALF1 molecule also showed the conservation of two cysteine residues at positions Cys^{55} and Cys^{76}, important for one disulfide bond (loop) formation in the peptide.

The phylogenetic relationship between PpALF1 and other ALFs of decapod crustaceans were analyzed using NJ method (Fig. 2.22). The phylogenetic tree clearly showed that the ALF sequences clustered according to species. The crab and shrimp ALFs were found to cluster separately. The tree could be broadly divided into three major groups. Group 1 consisted of ALFs from shrimps; whereas Group 2 consisted of ALFs from all crustaceans including shrimps, crabs, crayfishes and lobsters. Group 3 was found to consist of ALFs found in crabs alone and PpALF1 belonged to this group (Group 3). The analysis showed that PpALF1 was more closely related to crab ALFs rather than to the shrimp, crayfish and lobster ALFs. Also, within the crab cluster, ALFs of *Scylla* sp. grouped together and were found to be different from that of the *Portunus* and *Eriocher* sp. Also, PpALF1 was found to be different from the other ALF isoforms of *P. trituberculatus*. The bootstrap distance tree calculated clearly indicated PpALF1 to be a new isoform of the ALF family. Also, molecular phylogenetic tree based on amino acid sequence suggests that all the ALF members are of same ancestral origin, which has subsequently diverged at different phases of evolution.

2.3.2.2 PpALF2

2.3.2.2.1 Cloning and sequencing of PpALF2 cDNA

A 417 bp nucleotide sequence representing the complete cDNA sequence of PpALF2 with an open reading frame (ORF) of 372 bp encoding a polypeptide of 123 amino acid residues was obtained by RT-PCR and was cloned to a cloning vector (Fig. 2.23). The obtained ALF sequence encoded for precursor molecules, starting with a signal peptide containing 26 amino acid residues, followed by a highly cationic mature peptide containing 97 amino acid residues (Fig. 2.24). The PpALF2 cDNA sequence and deduced amino acid sequence has been submitted to the NCBI GenBank under the accession number **JQ899452**.

2.3.2.2.2 Analysis of PpALF2 sequence and peptide characteristics

The N-terminus of PpALF2 had the consistent features of a signal peptide as defined by SignalP program analysis, with a putative cleavage site located after position 26 (CEA-QY). The mature peptides of PpALF2 contained the two characteristic conserved cysteine residues forming a disulphide loop, which possess a cluster of positive charges within it. These conserved cysteine residues were found at the positions C^{29} and C^{50} in PpALF2. The deduced amino acid sequence of PpALF2 was found to be rich in positively charged amino acid residues, lysine (12.4 %) and arginine (5.2 %). In addition, it contained the highly hydrophobic N-terminal region and a consensus amino acid sequence (W/T)CP(G/S)W. The predicted molecular mass of the mature ALF was found to be 11.2 kDa with an estimated *p*I of 10.0. PpALF2 also showed the typical pattern of alternating hydrophobic and hydrophilic residues in their putative disulphide loop, suggesting that they comprise the same functional

domain. PpALF2 contained a highly hydrophobic N-terminal region including 10 hydrophobic residues in a total of 38 residues. Moreover, out of the 17 positively charged amino acids, seven positively charged amino acids and a tryptophan (Trp^{75}) were clustered within the disulfide loop consisting of 22 residues.

The physico-chemical properties of PpALF2 including its molecular weight, net charge, isoelectric point and hydrophobicity predicted using PROTPARAM, ExPASy, APD and HELIQUEST programs are described in Table 2.9. The PpALF2 was found to be cationic with a net charge of +10. Hydrophobicity was found to be 37 % as predicted by the PROTPARAM and HELIQUEST programs. The antimicrobial property of PpALF2 was predicted using APD and HELIQUEST programs. The results clearly indicated that the synthetic peptide taken for analysis possesses antimicrobial property. Among the residues taken for analysis, 37 % of amino acids were found to be hydrophobic residues and its total net charge was determined to be +10, thereby confirming the antimicrobial nature of the peptide. The peptide was predicted to interact with membranes and possess a chance to be a potential AMP (result based on APD program http://aps.unmc.edu/AP/main.php). The α-helix wheel projections for PpALF2 constructed using Schiffer-Edmundson helical wheel analysis clearly showed a perfect amphipathic nature of PpALF2 (Fig. 2.25). The analysis confirmed the peptide to be amphipathic in nature due to the presence of hydrophobic face. Analysis of the mature PpALF2 sequence using the Kyte-Doolittle plot demonstrated a significant presence of hydrophobic amino acids concentrated in the first 20 residues of the protein (Fig. 2.26).

The spatial structure of PpALF2 was established using the SWISS-MODEL prediction algorithm. There were four β-strands and two α-helices in the spatial structure of PpALF2. Secondary structure of SsALF1 was created in PyMOL software using data generated in pdb format by SWISS-MODEL (Fig. 2.27).

2.3.2.2.3 Sequence alignment and phylogenetic analysis of PpALF2

BLAST analysis revealed that PpALF2 exhibited significant similarity with crustacean ALF sequences (Table 2.10). BLAST analysis of the deduced amino acid sequences of PpALF2 revealed significant identities only with the ALFs of crustaceans and not with any other groups including Limulids. PpALF2 shared a significant identity with portunid ALFs (75–87 %) and shared low homology with the ALFs of *E. sinensis*, crayfishes and shrimps (less than 61 %).

Multiple alignment of PpALF2 with members of ALF family revealed that two cysteine residues (C^{29} and C^{50}) were highly conserved in portunid ALFs (Fig. 2.28). Twenty residues (H^{30}, R^{33}, K^{34}, P^{35}, K^{36}, F^{37}, R^{38}, K^{39}, F^{40}, K^{41}, L^{42}, Y^{43}, H^{44}, E^{45}, G^{46}, K^{47}, F^{48}, W^{49}) found in between the cysteine residues (C^{29} and C^{50}) in the β-strand of portunid ALFs were also found to be conserved for the PpALF2, except in case of F^{31}, F^{32}.

To evaluate the molecular evolutional relationships of PpALF2 against other ALF family members, a phylogenetic tree was constructed based on the 45 amino acid sequences of ALF members by the NJ method (Fig. 2.29). Phylogenetic analysis of ALFs further revealed that ALF sequences clustered according to species. PpALF2 was closely related to crab ALFs rather than to the other groups. There were three distinct groups in the phylogenetic tree. The first and third groups were shrimp and crab ALFs

respectively. In the second group, ALFs from all crustaceans viz. crabs, shrimps, lobsters and crayfishes were clustered together. Within this cluster, there were three distinct subgroups of sequences: a subgroup of lobster and shrimp ALFs, a second subgroup of shrimp, crayfish and crab ALFs to which the PpALF2 belonged to and a third subgroup containing ALFs from the crabs. Phylogenetic tree showed that ALFs from crabs including PpALF2 and ALFs from *P. trituberculatus, S. paramamosain, S. serrata* and *E. sinensis* were clustered together, and then had a closer relationship with crayfish *P. leniusculus* and shrimp ALFs (*P. monodon* and *M. japonicus*).

2.4 Discussion

AMPs are important components of the innate defense mechanism of most living organisms against microbial invasion. In crab species, several AMPs have been identified and characterized such as the 6.5 kDa proline-rich cationic protein from the shore crab, *C. maenas* (Schnapp et al., 1996), 3.7 kDa callinectin from the blue crab *C. sapidus* (Khoo et al., 1999), a cysteine-rich 11.5 kDa carcinin from the shore crab *C. maenas* (Relf et al., 1999) and scygonadin, from the mud crab, *S. serrata* (Wang et al., 2007) and ALFs from *C. maenas, C. sapidus* (Towle and Smith, 2006), *S. paramamosain* (Imjongjirak et al., 2007), *E. sinensis* (Li et al., 2008; Zhang et al., 2010a; Wang et al., 2011), *S. serrata* (Yedery and Reddy, 2009) and *P. trituberculatus* (Yue et al., 2010; Xu and Liu, 2011; 2012a, 2012b; 2013a, 2013b).

ALFs are small proteins that can bind and neutralize LPS. Nevertheless, a number of ALFs from horseshoe crabs and shrimps are well characterized, studies regarding crab ALFs are limited. Also, to date, no

ALF sequences have been identified in the portunid crab, *P. pelagicus*. In the present study, four ALF isoforms were identified and characterized from the haemocytes of the Indian mud crab *S. serrata*, and blue swimmer crab, *P. pelagicus*. Haemocytes were chosen as the starting material as the role of these circulating cells in protection of the invertebrates against invading microorganisms has been well documented (Lavine and Strand, 2002; Iwanaga and Lee, 2005). In the present study, complete ORFs of four ALF isoforms could be successfully identified and characterized at the molecular level. This study reports for the first time the presence of ALF isoforms in *P. pelagicus* as well as the second ALF isoform from *S. serrata*.

All the four ALF isoforms possessed 369 bp encoding 123 amino acids in the ORF. The ORF of obtained ALF sequences encoded for precursor molecules, starting with a signal peptide, followed by a highly cationic mature peptide possessing the LPS binding domain. The conserved characteristics and high similarity with known ALFs proved the identity of these new isoforms belonging to the ALF family. Signal peptide region consisted of either 24 amino acid residues as in the case of PpALF1 or 26 amino acid residues as in SsALF1, SsALF2 and PpALF2 respectively. Likewise the length of the mature peptide was found to be 99 in case of PpALF1 and 97 for the rest; 95-99 being the characteristic range of amino acids reported for ALFs of crustaceans. The molecular weight of all the four ALFs varied with that of each other and was found to range between 10 – 14 kDa. PpALF2 possessed the lowest molecular weight of 10 kDa followed by SsALF1 (11.17 kDa), SsALF2 (11.2 kDa) and PpALF1 (13.86 kDa). Isoelectric point of the four ALF isoforms ranged between 8.49 to 11.2. PpALF1 possessed the lowest *p*I of 8.49 followed by SsALF1 (9.95),

SsALF2 (10) and PpALF2 (11.2). All the four ALF isoforms were found to be highly cationic. Considering the net charge of these isoforms, PpALF1 possessed the lowest net charge of +2 followed by +9 of SsALF1. Whereas, SsALF2 and PpALF2 possessed the highest net charge of +10. However, not much variation in the hydrophobicity could be noticed. All the above features are in agreement with the required features of ALFs (Imjongjirak et al., 2007).

Multiple alignment performed for these ALF isoforms with other ALFs revealed the presence of conserved regions within the sequence. The deduced amino acid sequence of ALFs showed the presence of characteristic 24 amino acid domain, i.e. the LPS domain that was necessary for LPS binding and neutralization. Synthetic peptides, corresponding to the ALF putative LPS-binding domain of *Limulus* and decapod crustaceans have proved to possess an efficient LPS neutralizing activity as well as potent antimicrobial activity (Nagoshi et al., 2006; Imjongjirak et al., 2007). Multiple alignment of PpALF2 and SsALF2 revealed that two cysteine residues (C^{29} and C^{50}) were highly conserved in portunid ALFs. Whereas in case of *E. sinensis*, the two conserved cysteine residues were found at C^{27} and C^{48} (Zhang et al., 2010). Twenty residues (H^{30}, R^{33}, K^{34}, P^{35}, K^{36}, F^{37}, R^{38}, K^{39}, F^{40}, K^{41}, L^{42}, Y^{43}, H^{44}, E^{45}, G^{46}, K^{47}, F^{48}, W^{49}) found in between the cysteine residues (C^{29} and C^{50}) in the β-strand of portunid ALFs were also found to be conserved for the two ALF isoforms, except in case of F^{31}, F^{32} of PpALF2 and I^{31}, R^{32} in case of SsALF2. The biological functions of residue replacement between C^{29} and C^{50} need to be further investigated. The differences in the sequences of the LPS binding site might indicate the ability of these two ALF isoforms to bind to different microbial cell wall components.

The mature peptides of all the four isoforms contained the two characteristic conserved cysteine residues forming a disulphide loop, which possess a cluster of positive charges within it. This region of the mature peptide is defined as the LPS-binding domain (Imjongjirak et al., 2007). These conserved cysteine residues were found at the positions C^{29} and C^{50} in case of PpALF2 and SsALF2. Whereas, in case of SsALF1 and PpALF1 these conserved cysteine residues were found at the positions Cys^{55} and Cys^{76} respectively (Fig. 2.7 and 2.21). The conserved region is believed to be essential for the antimicrobial activity and for the stability of the 3D structure of ALFs (Yang et al., 2009). All the four isoforms possessed an amphipathic structure with clustering of hydrophobic amino acids at the amino terminal region and a highly conserved sequence of basic amino acids within the disulfide-bonded loop region. In all the cases, secondary structure consisted of two α-helices crowded against a four-strand β-sheet. Two of the β-strands are in turn linked by a disulfide bond to form an amphipathic loop rich in cationic amino acid side chains. Kyte-Doolittle plot also clearly demonstrated a significant presence of hydrophobic amino acids concentrated in the first 20 residues of all the isoforms. This clearly indicates that these amino acids may be part of α-helix and might be able to span across a lipid bilayer or membrane. Analysis of the putative LPS binding region of the protein situated between the two conserved cysteine residues, on the AMP Database program (http://aps.unmc.edu/AP/main.php), also explains the possibility of the LPS binding domain's potential to bind to the highly negative charged lipid A. Overall the newly identified ALFs demonstrated the presence of amphipathic regions which is an important characteristic of well studied AMPs.

BLAST analysis results revealed the similarity of these four ALF isoforms to other ALFs already identified in other decapod crustaceans. Both the ALF isoforms identified in *P. pelagicus* was found to be more diverse isoform than those identified from *S. serrata*. This result was well expected as PpALF1 and PpALF2 were the first ALFs to be reported from the species. All the four ALF isoforms identified in the present study were found to be more similar to ALFs from *Scylla* except PpALF1 which was more similar to ALF from *Portunus*. SsALF1 was found to be 100 % similar to an ALF isoform present in *S. paramamosain* (Imjongjirak et al., 2007). However, it showed only 93 % similarity to an already identified isoform from the same species i.e. *S. serrata*. This was followed by 76 % similarity to ALF from *P. trituberculatus*. SsALF2 was found to 93 % similar to *S. paramamosain* and *S. serrata* followed by 76 % similarity to other crab ALFs such as *P. trituberculatus* and 61 % to *E. sinensis*. PpALF1, the first isoform to be reported from *P. pelagicus* showed only 46 % similarity to ALF of *P. trituberculatus,* followed by 38-39 % similarity to other crab ALFs such as those from *E. sinensis, S. paramamosain* and *S. serrata.* Whereas, PpALF2, the second isoform to be reported from *P. pelagicus* showed 87 % similarity to ALFs already reported from *S. paramamosain* and *S. serrata* followed by 75 % to an ALF isoform from *P. trituberculatus* and 61 % to *E. sinensis* ALF respectively. All the four ALFs were found to be diverse from the ALFs already reported in shrimps and crayfishes. Only a maximum similarity of up to 53 % could be observed in comparison with shrimp and crayfish ALFs such as the ALFs from *P. monodon*, *M. japonicus* and *P. leniusculus* respectively.

To evaluate the molecular evolutional relationships of the newly identified ALF isoforms with other known ALFs of decapod crustaceans, a

phylogenetic tree was constructed based on the amino acid sequences of ALF isoforms employing NJ method. Phylogenetic analysis of ALFs suggested that all the ALF members are of the same ancestral origin, which has subsequently diverged at different phases of evolution. Phylogenetic trees showed that ALFs possess a similar evolutionary status and they were phylogenetically ancient immune effector molecules which may play an essential role in the host defense mechanism. Also, the analysis further revealed that ALF sequences were clustered according to species. The newly identified ALFs were closely related to other crab ALFs rather than to the other groups such as shrimps or crayfishes. Also the phylogenetic tree showed that ALFs from crabs including ALFs from *P. pelagicus*, *P. trituberculatus, S. paramamosain, S. serrata* and *E. sinensis* were clustered together, and had a closer relationship with crayfish *P. leniusculus* and shrimp ALFs (*P. monodon* and *M. japonicus*). This convergence between crustaceans might imply the complex evolution of ALFs in these groups and potential similarity in biological functions.

Generally, hydrohobicity varied from 34 to 39 percentage. Highest hydrophobicity was predicted for PpALF1 (39 %), followed by SsALF1 (37 %), PpALF2 (37 %) and SsALF2 (34 %). Previous reports have proved that the hydrophobicity is important for the peptide to bind LPS (Wasiluk et al., 2004). Like other ALF molecules, the newly reported ALFs of the present study, also showed the typical pattern of alternating hydrophobic and hydrophilic residues in their putative disulphide loop, suggesting that they comprise the same functional domain. The deduced amino acid sequence of SsALF1 showed that 37 % of amino acids were hydrophobic residues and its total net charge was determined to be +9, thereby confirming the antimicrobial nature of the peptide. Likewise in case of

PpALF1, 39 % of amino acids were found to be hydrophobic residues and its total net charge was determined to be +2, confirming the antimicrobial nature of the peptide. PpALF2 contained a highly hydrophobic amino-terminal region including 10 hydrophobic residues in a total of 38 residues. Moreover, out of the 17 positively charged amino acids, seven positively charged amino acids and a tryptophan (Trp^{75}) were clustered within the disulfide loop consisting of 22 residues. In contrast, SsALF2 also contained a highly hydrophobic N-terminal region as found in PpALF2, but contained only nine hydrophobic residues in a total of 35 residues. Out of the 20 positively charged amino acids, 10 positively charged amino acids and a tryptophan (Trp^{75}) were clustered within the disulfide loop consisting of 22 residues. This characteristic of Portunid ALFs was found to be a bit dissimilar to the horseshoe crab and shrimp ALFs, where, out of the 28 residues 12 hydrophobic residues have been reported in the LPS domain and of the 17 positively charged amino acids, 9 positively charged amino acids and a tryptophan (Trp^{75}) were found to be clustered within the disulfide loop consisting of 22 residues (de la Vega et al., 2008). These features suggest the newly identified ALFs to have an amphipathic factor.

All the identified ALF isoforms were found to be rich in positively charged amino acid residues such as arginine and lysine forming a cluster within the disulfide loop regarded as the functional domain of ALF as described by Imjongjirak et al. (2007). The peptide was predicted to interact with membranes and possess a chance to be a potential AMP (result based on APD program http://aps.unmc.edu/AP/main.php). The α-helix wheel projections for the newly identified ALFs constructed using Schiffer-Edmundson helical wheel analysis also clearly showed a perfect

amphipathic nature of these AMPs (Fig. 2.4; 2.11; 2.18; 2.25). Analysis of the mature ALFs using the Kyte-Doolittle plot also demonstrated a significant presence of hydrophobic amino acids concentrated in the first 20 residues of the protein (Fig. 2.5; 2.12; 2.19; 2.26). The analysis confirmed the peptide to be amphipathic in nature due to the presence of this hydrophobic face. These results clearly indicated that these peptides possess antimicrobial property. Also, the similarities in the *p*I of these AMPs indicate that antibacterial function or mode of action of these molecules may also be similar.

Based on the similarity with other ALFs, the spatial structures of the newly identified ALF isoforms were established using the SWISS-MODEL prediction algorithm based on the template 2jobA, and they were found to be similar to other known ALFs. There were four β-strands and two α-helices in the spatial structure. The two conserved-cysteine residues in the β-strands formed a disulfide bond that constrained a β-hairpin loop, in which there were an alternating series of hydrophilic and hydrophobic residues between the conserved cysteine residues (Hoess et al., 1993). The spatial structural similarity between these isoforms implied that they might represent similar biological functions in Portunid crabs. The primary structures indicate that all have amphipathic factors which are identical to other ALFs (Zhang et al., 2010). The amphipathic loop structures, believed to be the LPS binding sites (Nagoshi et al., 2006), were identified in the potential tertiary structures also. The presence of LPS binding site in these ALFs suggest the conservation of LPS binding activity and inhibitory activity to Gram-negative bacteria. All these isoforms are believed to involve in the defense responses of portunids.

Thus, from the characteristics of these ALF isoforms it is possible to predict that these peptides could function as a broad spectrum antimicrobial peptide against both Gram-negative and Gram-positive bacteria. Also there is possibly more than one ALF involved in crab immunity against various pathogens. Discovery of novel AMPs and its antimicrobial spectrum might pave way to unravel the obscurity of crustacean immunity. Further research on the expression profile of these molecules in response to various environmental conditions and microbial infection would reveal their role in the protection of the animals from the onslaught of diseases. Also, the functional characterization of these molecules in terms of its antimicrobial activities against various pathogens will provide a promising drug in crustacean health management.

(a) **(b)**

Fig. 2.1 Experimental organisms used for the study (a) *Scylla serrata* (b) *Portunus pelagicus*

Fig. 2.2 Nucleic acid (above) and deduced amino acid (below) sequence of SsALF1. The signal peptide sequence (green), mature peptide region (bright blue & light blue) and putative LPS domain (ight blue) have been highlighted.

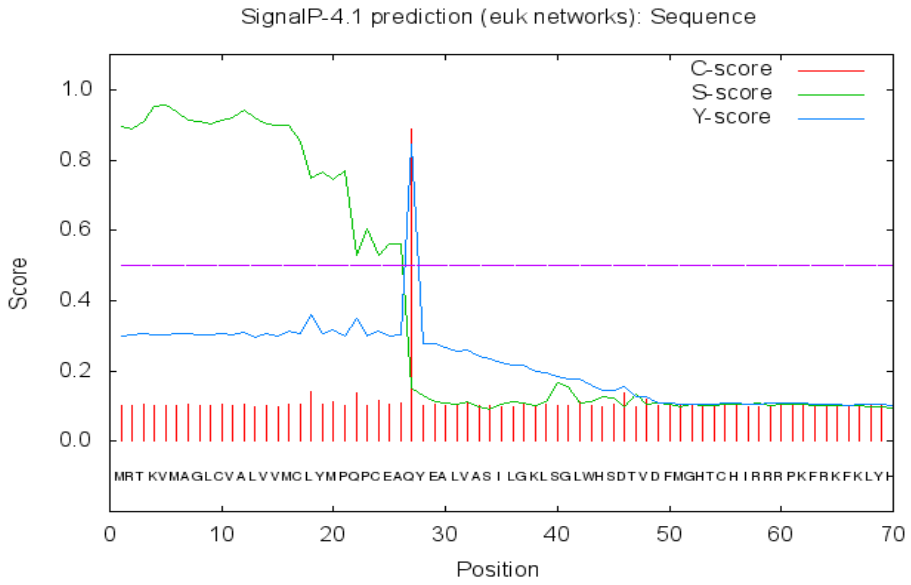


Fig.2.3 Signal peptide analysis of SsALF1 as predicted by the SignalP 3.0 server.

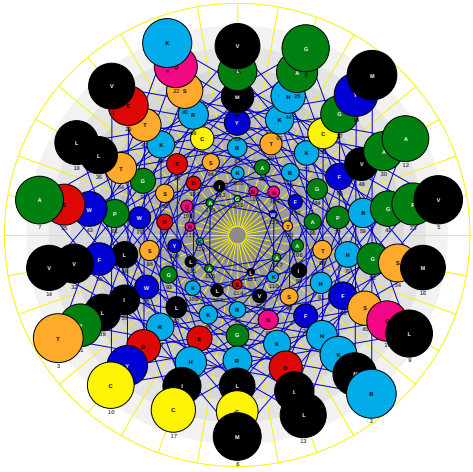


Fig.2.4 Schiffer-Edmundson Helical wheel representation of SsALF1 generated using DNASTAR software. Hydrophobic and hydrophilic amino acids are seen occupying opposite positions.

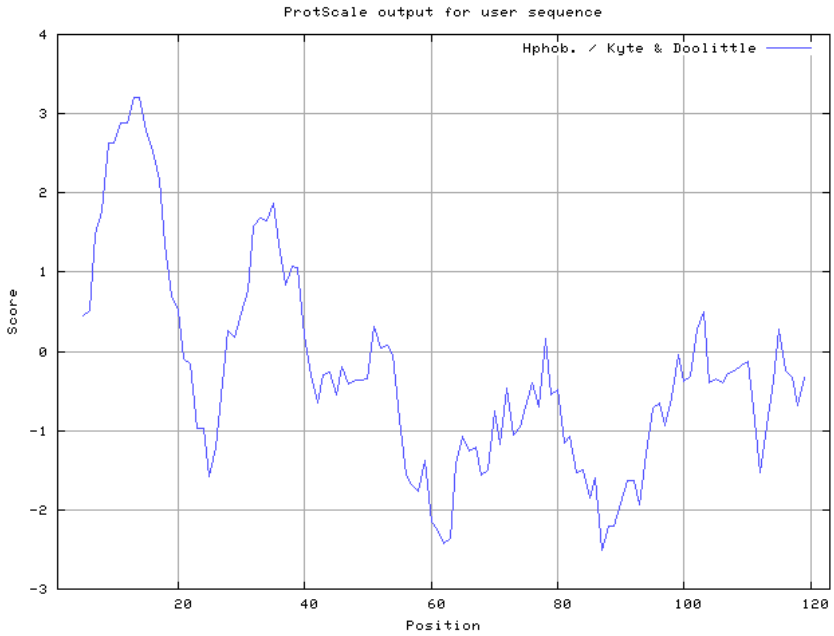


Fig. 2.5 Kyte-Doolittle plot showing hydrophobicity of SsALF1. The peaks above the mean (0.0) indicate the hydrophobic nature of the predicted protein.

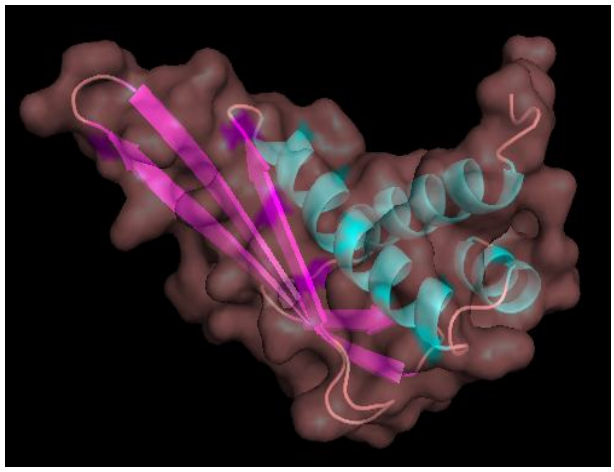


Fig.2.6 Secondary structure of SsALF1 created using PyMOL showing α-helix, β-sheet and extended structures

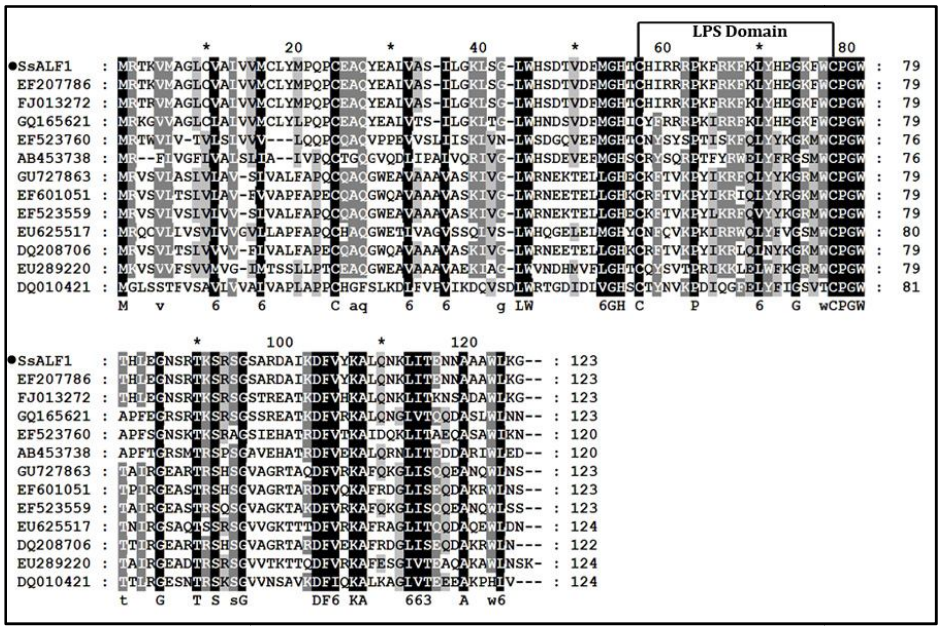


Fig.2.7 Multiple alignment of nucleotide sequence of SsALF1 with other crustacean ALFs obtained using GeneDoc program. Black and grey indicates conserved sequences.

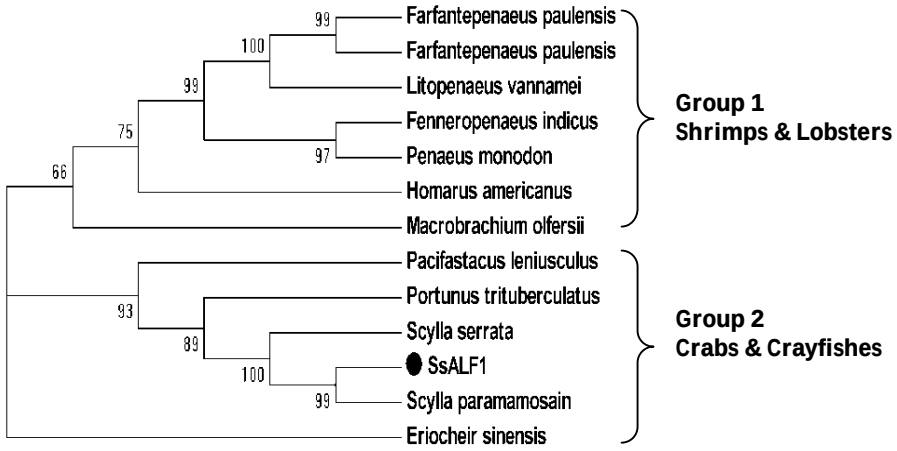


Fig.2.8 A bootstrapped neighbour-joining tree obtained using MEGA version 5.0 illustrating relationships between the deduced amino acid sequence of SsALF1 with other crustacean ALFs. Values at the node indicate the percentage of times that the particular node occurred in 1000 trees generated by bootstrapping the original deduced protein sequences.

```
ggacagaagaaacattgaggacgacgcaaccaagcttccctcaagatgcggaccagggtg
                                                M  R  T  R  V
atggccggcctgtgcgtggcgctggtggtgatgtgcctgtacatgccccagccgtgcgag
M  A  G  L  C  V  A  L  V  V  M  C  L  Y  M  P  Q  P  C  E
gctcagtatgaagctctggtagcttccattcttggaaagctgtcgggactgtggcacagc
A  Q  Y  E  A  L  V  A  S  I  L  G  K  L  S  G  L  W  H  S
gacacagtggacttcatgggacacacctgccacatccgccgcaagccgaagttcaggaaa
D  T  V  D  F  M  G  H  T  C  H  I  R  R  K  P  K  F  R  K
tttaagctgtaccacgagggcaagttttggtgtcctggctggacacatctcgagggcaat
F  K  L  Y  H  E  G  K  F  W  C  P  G  W  T  H  L  E  G  N
tcgaggaccaagagcaggtcggggtcaaccagggaagccaccaaggacttcgtgcacaaa
S  R  T  K  S  R  S  G  S  T  R  E  A  T  K  D  F  V  H  K
gctttacaaaacaaactcatcacgaagaatagcgcggacgcctggctgaaggggtgaggc
A  L  Q  N  K  L  I  T  K  N  S  A  D  A  W  L  K  G  *
```

Fig. 2.9 Nucleic acid (above) and deduced amino acid (below) sequence of SsALF2. The signal peptide sequence (green), mature peptide region (bright blue & light blue) and putative LPS domain (ight blue) have been highlighted.

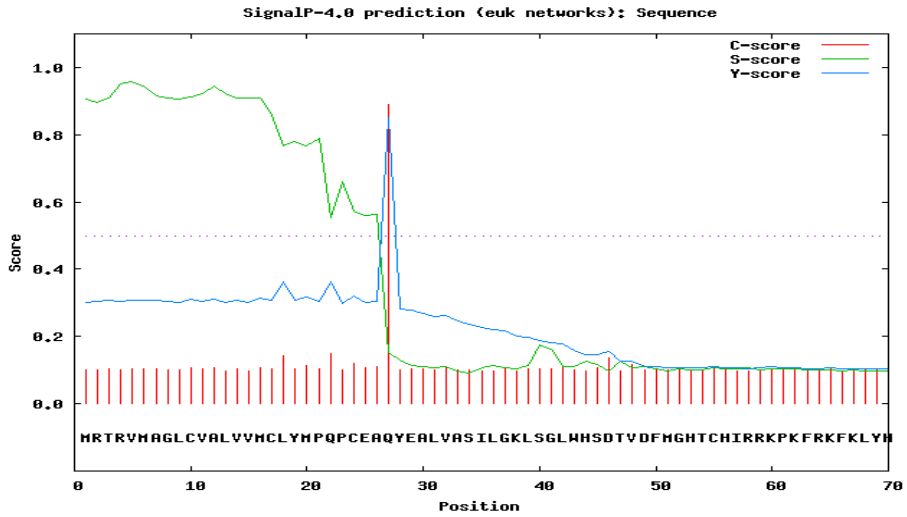


Fig.2.10 Signal peptide analysis of SsALF2 as predicted by the SignalP 3.0 server. The underlined amino acid residues indicate the putative signal sequence.

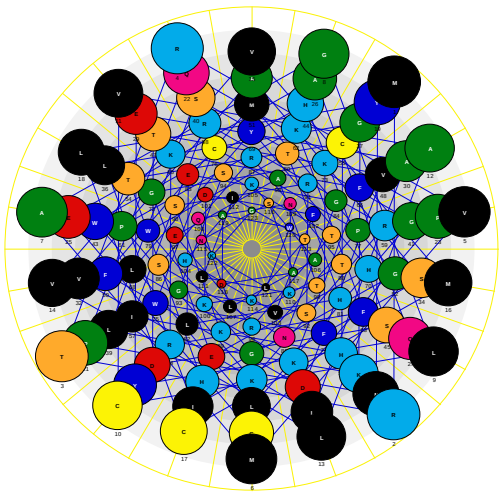


Fig.2.11 Schiffer-Edmundson Helical wheel representation of SsALF2 generated using DNASTAR software. Hydrophobic and hydrophilic amino acids are seen occupying opposite positions.

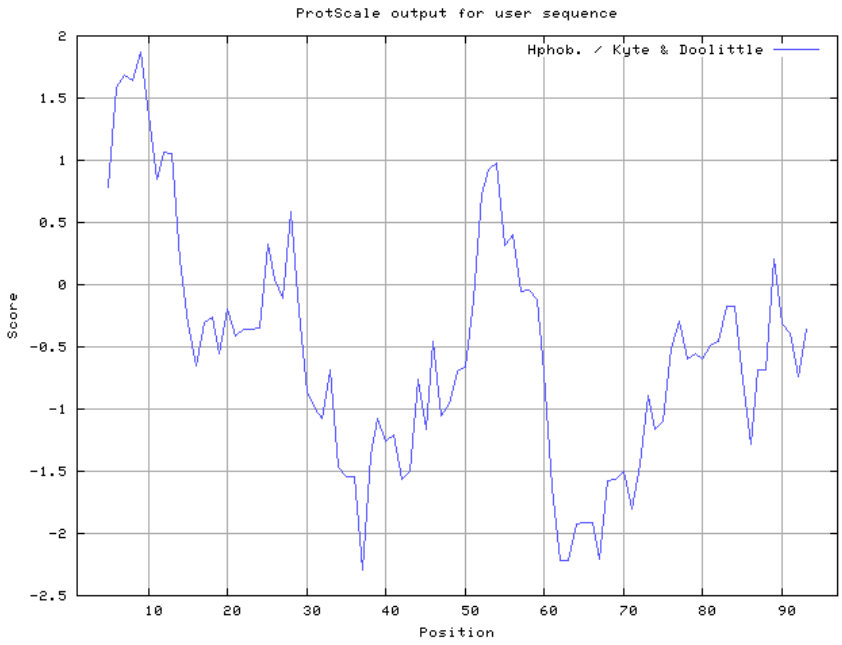


Fig. 2.12. Kyte-Doolittle plot showing hydrophobicity of SsALF2. The peaks above the mean (0.0) indicate the hydrophobic nature of the predicted protein.

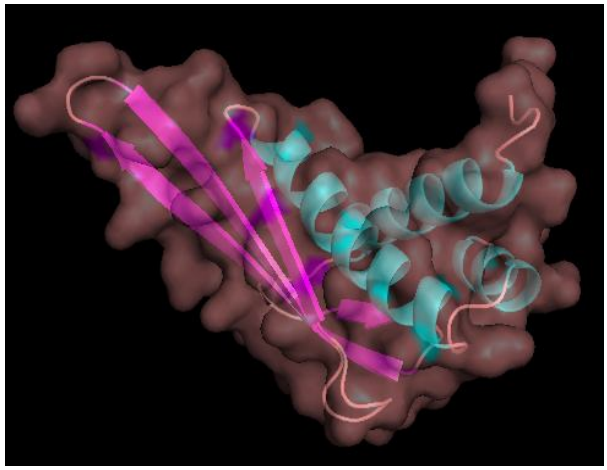


Fig.2.13 Secondary structure of SsALF2 created using PyMOL showing α-helix, β-sheet and extended structures

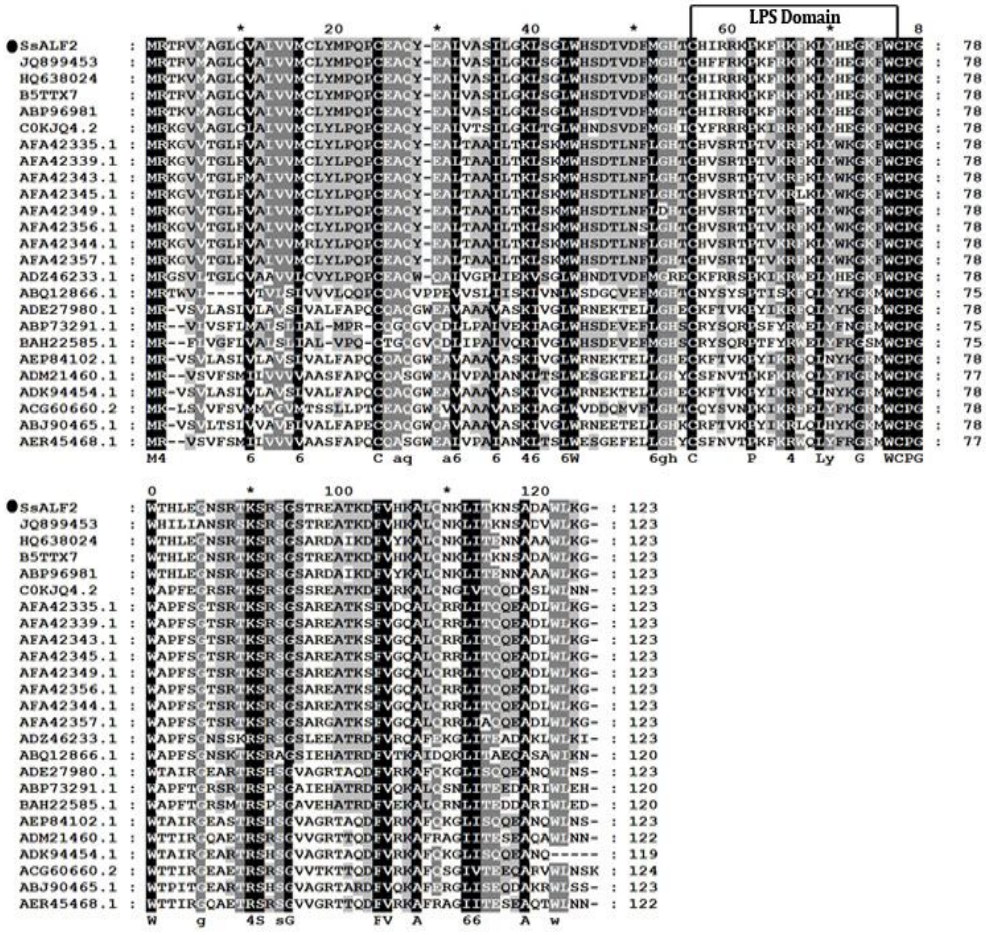


Fig.2.14 Multiple alignment of nucleotide sequence of SsALF2 with other crustacean ALFs obtained using GeneDoc program. Black and grey indicates conserved sequences.

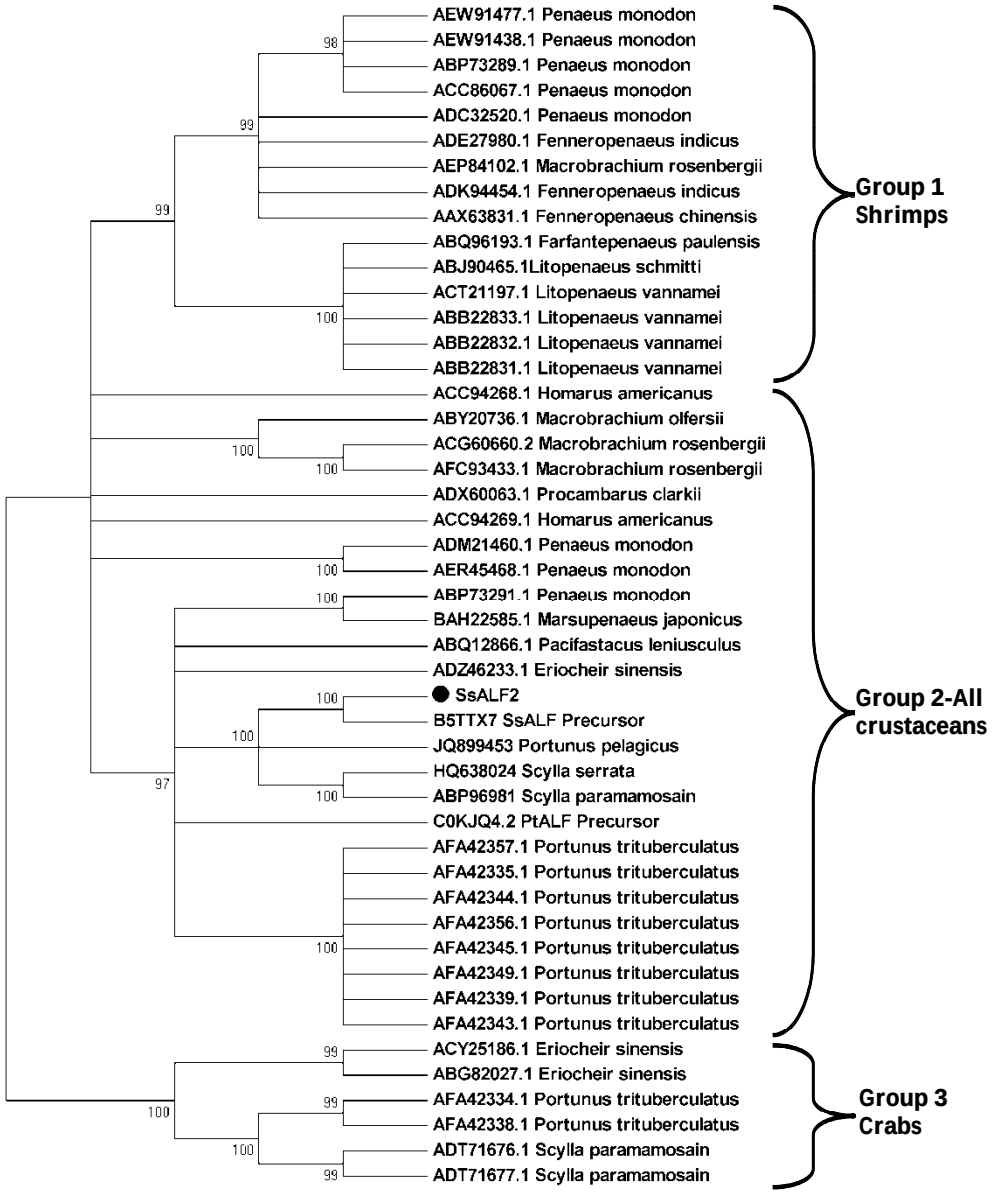


Fig.2.15 A bootstrapped neighbour-joining tree obtained using MEGA version 5.0 illustrating relationships between the deduced amino acid sequence of SsALF2 with other crustacean ALFs. Values at the node indicate the percentage of times that the particular node occurred in 1000 trees generated by bootstrapping the original deduced protein sequences.

```
gttgttatcatctgttgcccttgctgcttccctccaccttcttccctcatgatgaaaaaa
                                                  M  K  K
ggagtggccgccctgctgcggctgctgctggtgatgaggctgctgctgccgccccttct
 G  V  A  A  L  L  R  L  L  L  V  M  R  L  L  L  P  P  P  S
tgcgatcctcatgacgatcctctggcttcttctcttcttgaaatcctcgctgggcggtgc
 C  D  P  H  D  D  P  L  A  S  S  L  L  E  I  L  A  G  R  C
cacaactcctctgatgacttggtgggcccctgctgcttcctccgccgcccccctatcata
 H  N  S  S  D  D  L  V  G  P  C  C  F  L  R  R  P  P  I  I
agaatatttctgctgcaccacgagggctttttttggtgtgctgggtggccgcctttcgat
 R  I  F  L  L  H  H  E  G  F  F  W  C  A  G  W  P  P  F  D
gggtcgacgaggacaaacaggacggcggcgtcctcggaggacgccaccgacgacttgcgg
 G  S  T  R  T  N  R  T  A  A  S  S  E  D  A  T  D  D  L  R
cgcgctttttacacaaactactcatcccacaggatgatgctccgcggcggatgaacaac
 R  A  F  L  H  K  L  L  I  P  Q  D  D  A  P  R  R  M  N  N
taa
 *
```

Fig. 2.16 Nucleic acid (above) and deduced amino acid (below) sequence of PpALF1. The signal peptide sequence (green), mature peptide region (bright blue & light blue) and putative LPS domain (ight blue) have been highlighted.

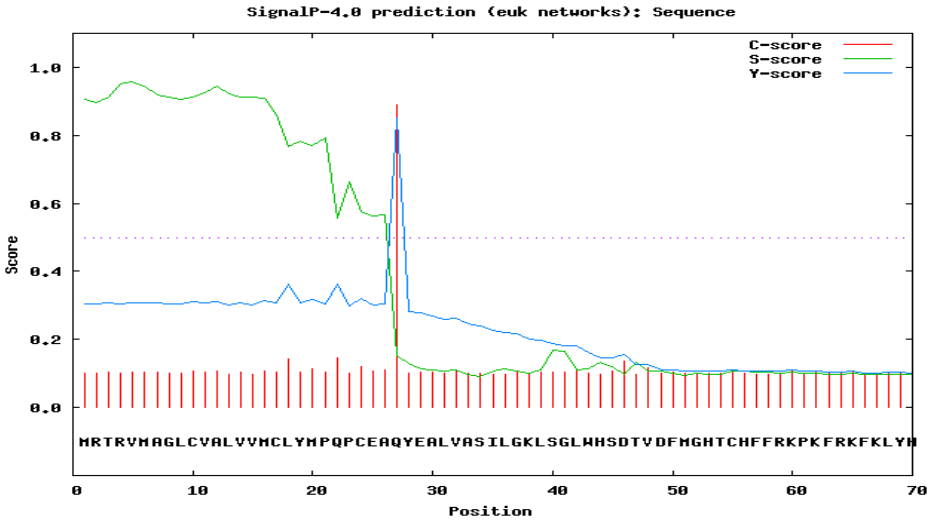


Fig.2.17 Signal peptide analysis of PpALF1 as predicted by the SignalP 3.0 server.

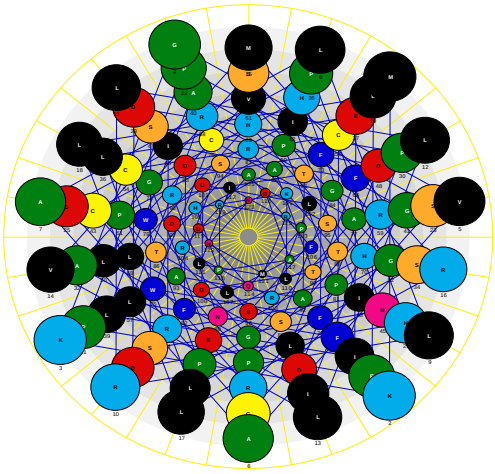


Fig.2.18 Schiffer-Edmundson Helical wheel representation of PpALF1 generated using DNASTAR software. Hydrophobic and hydrophilic amino acids are seen occupying opposite positions.

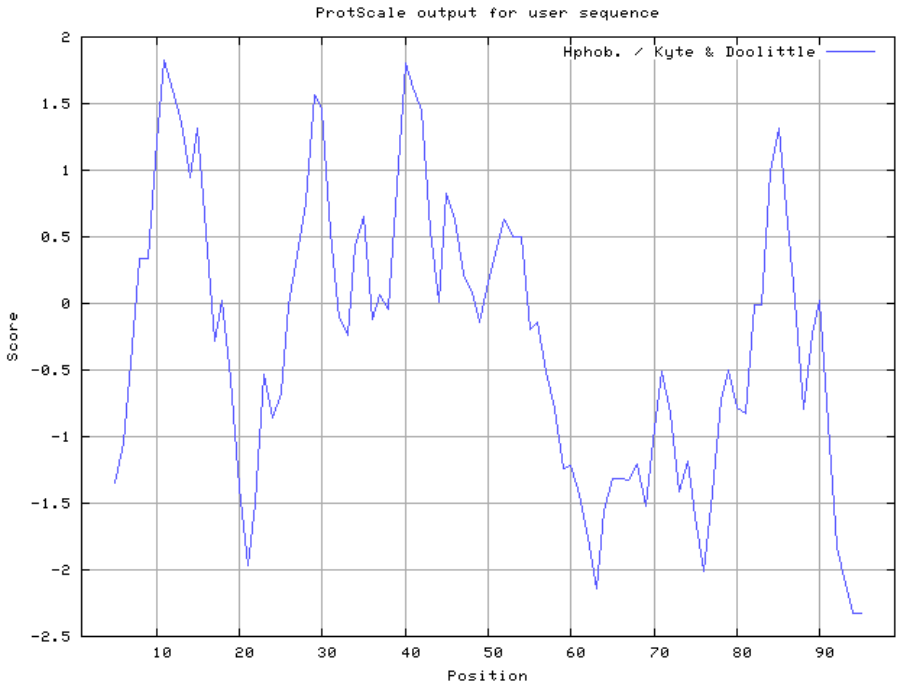


Fig. 2.19 Kyte-Doolittle plot showing hydrophobicity of PpALF1. The peaks above the mean (0.0) indicate the hydrophobic nature of the predicted protein.

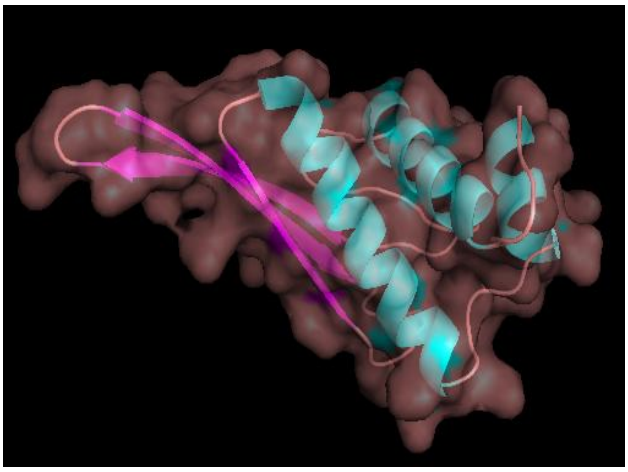


Fig.2.20 Secondary structure of PpALF1 created using PyMOL showing α-helix, β-sheet and extended structures

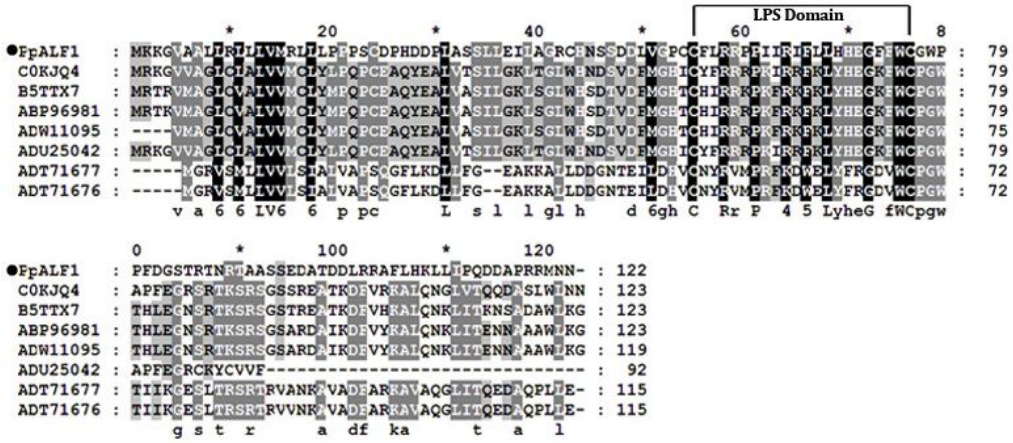


Fig.2.21 Multiple alignment of nucleotide sequence of PpALF1 with other crustacean ALFs obtained using GeneDoc program. Black and grey indicates conserved sequences.

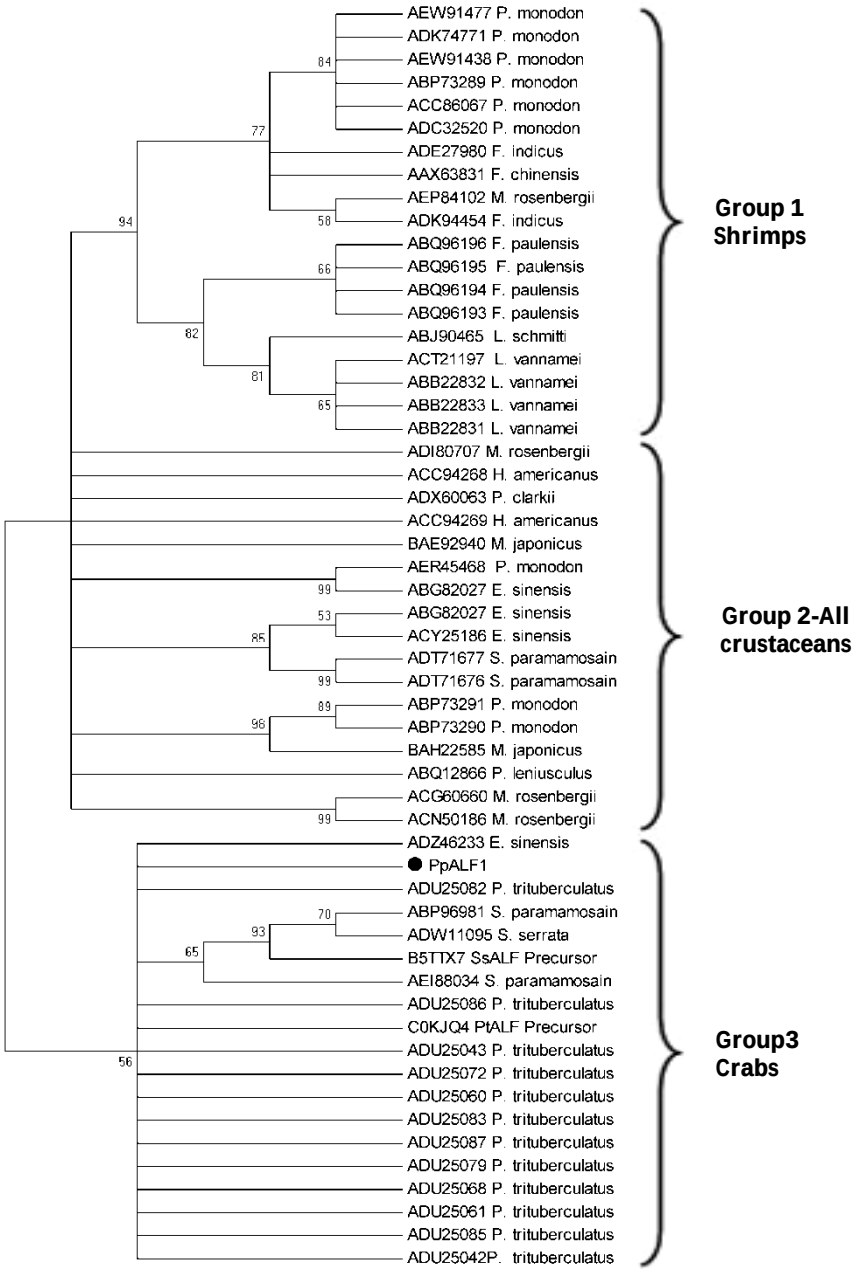


Fig.2.22 A bootstrapped neighbour-joining tree obtained using MEGA version 5.0 illustrating relationships between the deduced amino acid sequence of PpALF1 with other crustacean ALFs. Values at the node indicate the percentage of times that the particular node occurred in 1000 trees generated by bootstrapping the original deduced protein sequences.

Fig. 2.23 Nucleic acid (above) and deduced amino acid (below) sequence of PpALF2. The signal peptide sequence (green), mature peptide region (bright blue & light blue) and putative LPS domain (ight blue) have been highlighted.

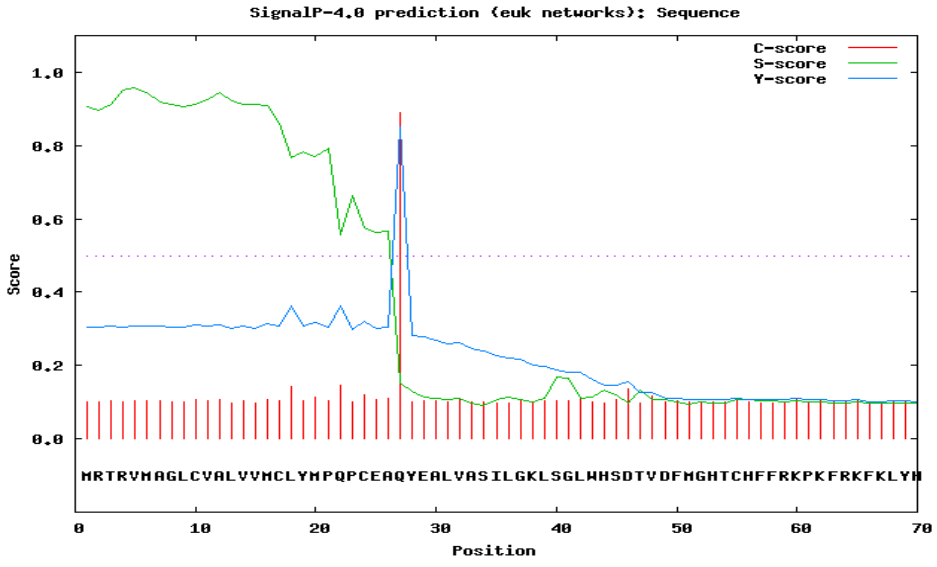


Fig.2.24 Signal peptide analysis of PpALF2 as predicted by the SignalP 3.0 server.

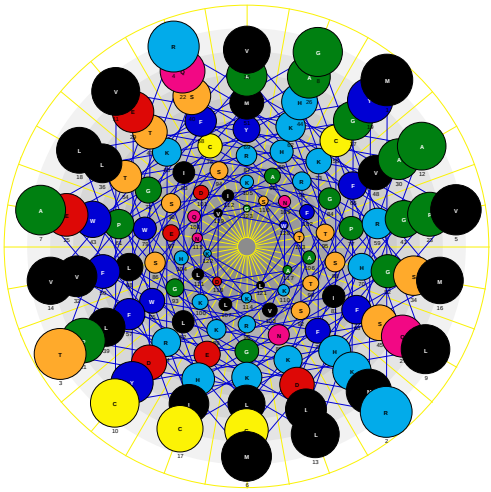


Fig.2.25 Schiffer-Edmundson Helical wheel representation of PpALF2 generated using DNASTAR software. Hydrophobic and hydrophilic amino acids are seen occupying opposite positions.

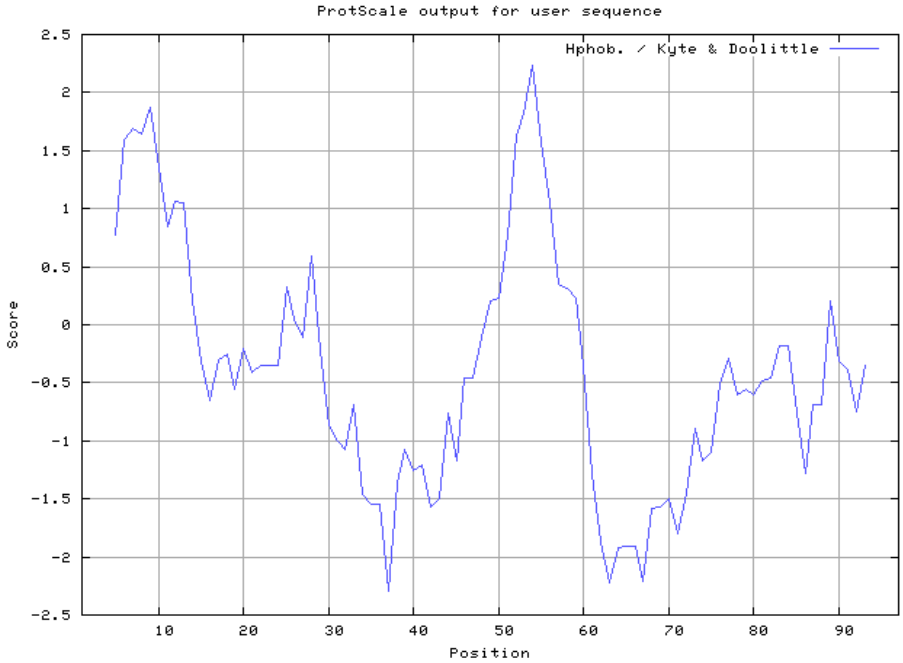


Fig. 2.26 Kyte-Doolittle plot showing hydrophobicity of PpALF2. The peaks above the mean (0.0) indicate the hydrophobic nature of the predicted protein.

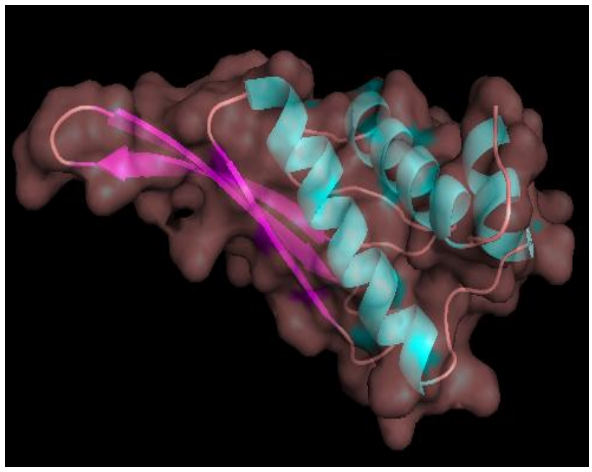


Fig.2.27 Secondary structure of PpALF2 created using PyMOL showing α-helix, β-sheet and extended structures

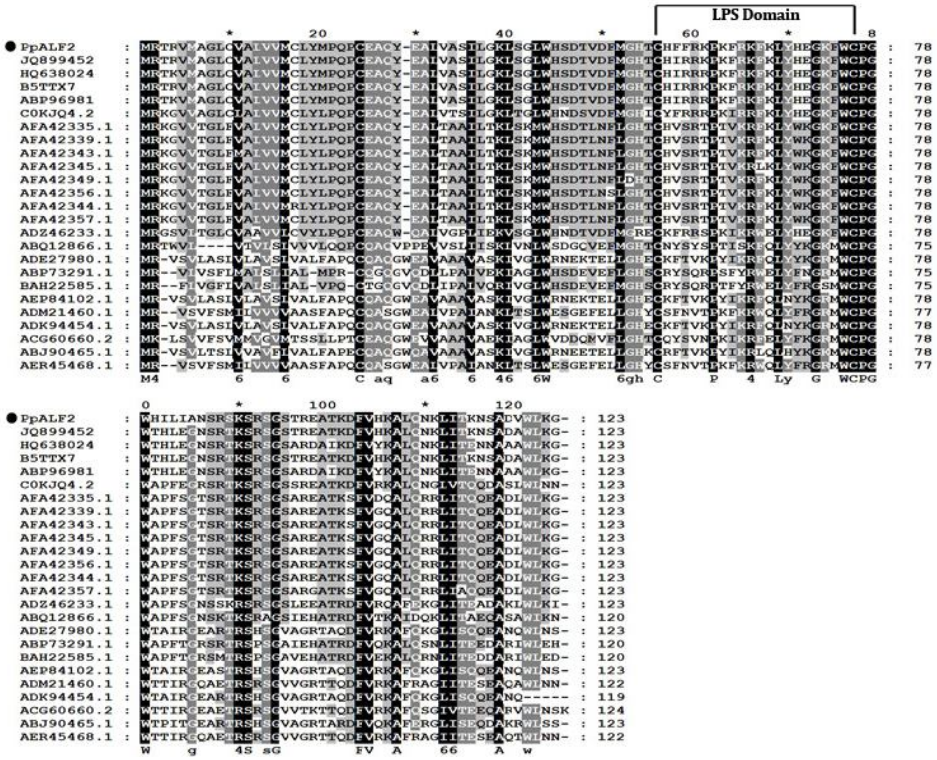


Fig.2.28 Multiple alignment of nucleotide sequence of PpALF2 with other crustacean ALFs obtained using GeneDoc program. Black and grey indicates conserved sequences.

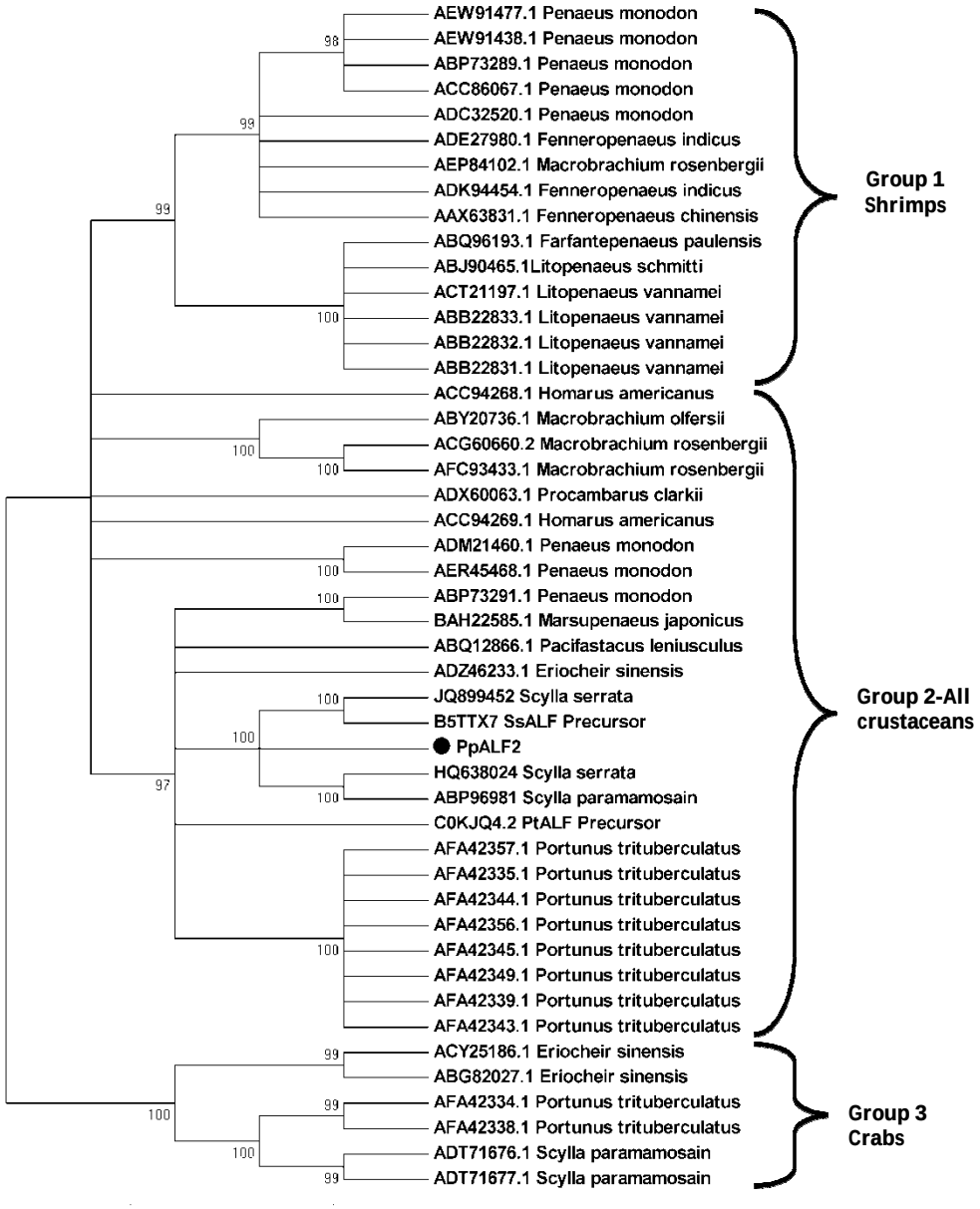


Fig.2.29 A bootstrapped neighbour-joining tree obtained using MEGA version 5.0 illustrating relationships between the deduced amino acid sequence of PpALF2 with other crustacean ALFs. Values at the node indicate the percentage of times that the particular node occurred in 1000 trees generated by bootstrapping the original deduced protein sequences.

Table 2.1 Primers used for the study

Target gene	Sequence (5'-3')	Product Size (bp)	Annealing Temp. (°C)	$MgCl_2$ Conc. (mM)
ALF-1	F: ggacagaagaaacattgaggacgacgca R: ggaaatcaaaaacatccattacaggtca	520	60	1.5
ALF-2	F: agggagtgggtgatgagcta R: tacggctattacgatccaaca	900	62	1.5
β-actin	F: cttgtggttgacaatggctccg R: tggtgaaggagtagccacgctc	520	60	1.5
T7	tgtaatacgactcactataggg	--	57	1.5
SP6	gatttaggtgacactatag	--	57	1.5

Table 2.2 List of ALF isoforms identified and characterized in the present study

No:	Name of AMP	Source Organism	ORF	GenBank Accession
1	SsALF1	*Scylla serrata*	Complete	**HQ638024**
2	SsALF2	*Scylla serrata*	Complete	**JQ899452**
3	PpALF1	*Portunus pelagicus*	Complete	**JQ745295**
4	PpALF2	*Portunus pelagicus*	Complete	**JQ899453**

Table 2.3 Physico-chemical properties of SsALF1

Parameters Studied	Results
Open reading frame (GeneTool)	123 aa
Length of mature peptide (No. of amino acids) (GeneTool)	97 aa
Signal Peptide (SignalP)	1-26 aa
Molecular weight (PROTPARAM, ExPASy)	11.17 kDa
Isoelectric point (PROTPARAM, ExPASy)	9.95
Charge(PROTPARAM, ExPASy)	+9
Hydrophobicity (PROTPARAM, ExPASy)	37 %
Structure (PyMol & DNASTAR)	α Helix & β Sheet

Table2.4 Result of BLASTp analysis of SsALF1

GenBank Accession	Description	Query Coverage	Identity to SsALF1
EF207786	*Scylla paramamosain* anti-lipopolysaccharide factor mRNA, complete cds	100 %	100 %
FJ013272	*Scylla serrata* anti-lipopolysaccharide factor precursor, mRNA, complete cds	100 %	93 %
GQ165621	*Portunus trituberculatus* anti-lipopolysaccharide factor isoform 3, mRNA, complete cds, alternatively spliced	98 %	76 %
EF523760	*Pacifastacus leniusculus* anti-lipopolysaccharide factor mRNA, complete cds	85 %	52 %
GU727863	*Fenneropenaeus indicus* anti-lipopolysaccharide factor mRNA, complete cds	98 %	41 %

Table2.5 Physico-chemical properties of SsALF2

Parameters Studied	Results
Open reading frame (GeneTool)	123 aa
Length of mature peptide (No. of amino acids) (GeneTool)	97 aa
Signal Peptide (SignalP)	1-26 aa
Molecular weight (PROTPARAM, ExPASy)	11.2 kDa
Isoelectric point (PROTPARAM, ExPASy)	10
Charge(PROTPARAM, ExPASy)	+10
Hydrophobicity (PROTPARAM, ExPASy)	34 %
Structure (PyMol & DNASTAR)	α Helix & β sheet

Table2.6 Result of BLASTp analysis of SsALF2

GenBank Accession No.	Description of the AMP	Query coverage	Identity to SsALF2
ABP96981	Anti-lipopolysaccharide factor [*Scylla paramamosain*]	100 %	93 %
ADW11095	Anti-lipopolysaccharide factor [*Scylla serrata*]	100 %	93 %
ACM89169	Anti-lipopolysaccharide factor [*Portunus trituberculatus*]	99 %	76 %
ADZ46233	Anti-lipopolysaccharide factor 3 [*Eriocheir sinensis*]	99 %	61 %
ABP73291	Anti-lipopolysaccharide factor isoform 2 [*Penaeus monodon*]	96 %	45 %
BAH22585	Anti-lipopolysaccharide factor 2 [*Marsupenaeus japonicus*]	96 %	44 %
ABQ12866	Anti-lipopolysaccharide factor [*Pacifastacus leniusculus*]	82 %	53 %

Table2.7 Physico-chemical properties of PpALF1

Parameters Studied	Results
Open reading frame (GeneTool)	123 aa
Length of mature peptide (No. of amino acids) (GeneTool)	99 aa
Signal Peptide (SignalP)	1-24 aa
Molecular weight (PROTPARAM, ExPASy)	13.86 kDa
Isoelectric point (PROTPARAM, ExPASy)	8.49
Charge(PROTPARAM, ExPASy)	+2
Hydrophobicity (PROTPARAM, ExPASy)	39 %
Structure (PyMol & DNASTAR)	α Helix & β sheet

Table 2.8 Result of BLASTp analysis of PpALF1

GenBank Accession No.	Description of the AMP	Query coverage	Identity to PpALF1
ADU25043	Anti-lipopolysaccharide factor isoform 2 [*Portunus trituberculatus*]	88 %	46 %
ADZ46233	Anti-lipopolysaccharide factor 3 [*Eriocheir sinensis*]	85 %	39 %
ABP96981	Antilipopolysaccharide factor [*Scylla paramamosain*]	94 %	38 %
ADW11095	Anti-lipopolysaccharide factor [*Scylla serrata*]	92 %	38 %

Table 2.9 Physicochemical properties of PpALF2

Parameters Studied	Results
Open reading frame (GeneTool)	123 aa
Length of mature peptide (No. of amino acids) (GeneTool)	97 aa
Signal Peptide (SignalP)	1-26 aa
Molecular weight (PROTPARAM, ExPASy)	10 kDa
Isoelectric point (PROTPARAM, ExPASy)	11.2
Charge(PROTPARAM, ExPASy)	+10
Hydrophobicity (PROTPARAM, ExPASy)	37 %
Structure (PyMol & DNASTAR)	α Helix & β sheet

Table2.10 Result of BLASTp analysis of PpALF2

GenBank Accession No.	Description of the AMP	Query coverage	Identity to PpALF2
ABP96981	Anti-lipopolysaccharide factor [*Scylla paramamosain*]	100 %	87 %
ADW11095	Anti-lipopolysaccharide factor [*Scylla serrata*]	100 %	87 %
ACM89169	Anti-lipopolysaccharide factor [*Portunus trituberculatus*]	99 %	75 %
ADZ46233	Anti-lipopolysaccharide factor 3 [*Eriocheir sinensis*]	99 %	61 %
ABP73291	Anti-lipopolysaccharide factor isoform 2 [*Penaeus monodon*]	96 %	44 %
BAH22585	Anti-lipopolysaccharide factor 2 [*Marsupenaeus japonicus*]	96 %	43 %
ABQ12866	Anti-lipopolysaccharide factor [*Pacifastacus leniusculus*]	82 %	51 %

3

Molecular Characterization and Phylogenetic Analysis of Crustins in Crustaceans

3.1 Introduction

Crustaceans are the most abundant and diverse group of animals inhabiting the marine environment. They live in a habitat overloaded with infectious organisms, such as viruses, bacteria, fungi and other parasites. To survive in such a potentially hostile and microbe-enriched environment they have efficient defense mechanisms. Crustaceans rely mostly on their innate immune response comprised of cellular and humoral responses, primarily linked to their haemolymph. Antimicrobial peptides (AMPs) are one of the major components of humoral part of the innate immune defense and are in fact, multifunctional molecules having rapid and efficient antimicrobial and antitumor effects, mitogenic activity and immunoregulatory activities. Different families of AMPs have been identified in crustaceans of which crustins are an important group of AMPs occurring more widely across the decapod crustaceans.

So far over 60 crustin sequences have been reported from a variety of decapods, including crabs, lobsters, shrimps, prawns and crayfishes. Crustins are generally defined as multi-domain cationic AMPs of size range 7-14 kDa and a *p*I of 7.0-8.7 possessing a whey acidic protein (WAP)

domain at the C-terminus (Smith et al., 2008). The term 'WAP' is derived from the name given to a family of proteins, originally discovered in the whey fraction of mammalian milk. Though all these milk proteins are characterized by the possession of two WAP domains, each comprising 50 amino acids (Ranganathan et al., 1999), numerous other non-milk WAPs are also known possessing one or more four disulfide cores (4DSCs). Amongst these non-milk WAPs are small secretory proteins with protease inhibitory properties or regulatory functions in growth, tissue differentiation or regulation and may sometimes be expressed in certain cancerous conditions (Schalkwijk et al., 1999; Bouchard et al., 2006). Analysis of numerous WAPs from vertebrates reveals a high degree of similarity between the WAP domain structures. Ranganathan et al. (1999) have proposed that the PROSITE definition of the domain structure be modified to the following identifying motifs whose positions of the eight conserved cysteines (C_1-C_8) are as follows: C_1-(Xn)-C_2-(Xn)-$\underline{C_3}$-$\underline{(X5)}$-$\underline{C_4}$-$\underline{(X5)}$-$\underline{C_5}$-$\underline{C_6}$-(X3,X5)-C_7-(X3,X4)-C_8, where, X indicates any residue and Xn is a stretch of n residues. The signature motif of the central four cysteines that form the basis of the 4DSC is underlined. The definition of a crustin is in agreement with this format and therefore excludes several other cysteine-rich crustacean AMPs that lack a WAP domain, such as defensins or penaeidins. Importantly, it embraces some crustacean molecules that have not so far been specifically designated as crustins or for which an antibacterial function is yet to be established. Moreover, it sets crustins within their own category of WAP-containing proteins (Schalkwijk et al., 1999; Talas-Ogras, 2004; Andrenacci et al., 2006; Nile et al., 2006) and distinguishes them from those having a WAP domain at the N-terminus or more than one at the C-terminus, even though there may be some

functional equivalence between these proteins in terms of their role in inflammation and/or antimicrobial properties (Hiemestra et al., 1996; Hiemstra, 2002; Hagiwara et al., 2003; Jimenez-Vega et al., 2007). The WAP domain, in contrast to the signal sequence, is highly conserved between species and in several crustins, especially those from shrimp, aspartic acid and lysine residues are positioned as follows: -C-XX-D-XX-C-XXXD-K-CC-X-D-. This arrangement, however, does not hold true in every case (Jiravanichpaisal et al., 2007). Other differences in the WAP domain of crustins could well come to light as more crustin-like sequences are determined for a wider range of crustacean species. Despite small differences in the sequences, the domain, itself, seems to form a tightly coiled structure enclosing two β-sheets and α-helical segment. Structural models of the WAP domain of three different crustins (Smith et al., 2008) suggest that the tertiary structure of this part of the molecule is well conserved between decapod species.

Because of their potential antimicrobial activity, crustins are considered as potential antimicrobial agents. The first crustin member to be identified was an 11.5 kDa protein purified from the granular haemocytes of the shore crab *C. maenas* that exhibits specific activity towards Gram-positive marine or salt-tolerant bacteria (Relf et al., 1999). Thereafter over 60 crustins and crustin-like sequences were identified from various decapod crustacean species such as crayfishes, shrimps, freshwater prawns, crabs, lobsters as well as non-decapod crustacean species such as amphipods, through EST-based approaches (Bartlett et al., 2002; Hauton et al., 2006; Zhang et al., 2007; Smith et al., 2008; Supungul et al., 2008). It was Bartlett and co-workers (2002) who proposed the name crustin, to describe transcripts found in two penaeid shrimp species (*L. vannamei*

and *L. setiferus*) with high sequence similarity to the crab 11.5 kDa protein, later designated as carcinin (Brockton et al., 2007).

The structure of crustin consists of a signal peptide sequence at the N-terminus followed by a C-terminal region containing a WAP domain. The putative signal sequence at the N-terminus comprises ~16–24 amino acids, which in many species shows strong representation of valine residues. The cleavage site, which marks the end of the signal sequence, at least as predicted by software programs, is usually between alanine and glycine, although in some crustins it lies between glycine and glutamine, alanine or threonine. WAP domain consists of 50 amino acid residues with eight cysteine residues at defined positions involved in the formation of a single four disulfide core (4DSC) bonds creating a tightly packed structure. Based on their structural features, Smith et al. (2008) categorized crustins into three main types (Type I, II and III). Type I crustins encompass the members most related to carcinin and have a cysteine-rich region of variable length between the signal peptide sequence and the WAP domain. This type of crustin is predominantly found in crabs (Relf et al., 1999; Sperstad et al., 2009a; Imjongjirak et al., 2009; Mu et al., 2010; Yue et al., 2010), lobsters (Stoss et al., 2004; Hauton et al., 2006; Christie et al., 2007; Pisuttharachai et al., 2009), crayfishes (Jiravanichpaisal et al., 2007), shrimps (Sun et al., 2010) and freshwater prawn (Dai et al., 2009). Whereas, Type II crustins possess a signal sequence followed by a long glycine-rich region, a cysteine-rich region and the WAP domain. The number of glycines varies between species but is usually between 20 and 50, and in shrimp, it is often arranged as repeat VGGGLG motifs that vary in number from 5 to 8 (Bartlett et al., 2002; Supungul et al., 2004; Vargas-Albores et al., 2004; Rosa et al., 2007; Zhang

et al., 2007; Amparyup et al., 2008a, 2008b; Supungul et al., 2008). This group is mainly found in penaeid shrimps (Bartlett et al., 2002; Rattanachai et al., 2004; Supungul et al., 2004; de Lorgeril et al., 2005; Rosa et al., 2007; Zhang et al., 2007; Antony et al., 2010) and crayfishes (Jiravanichpaisal et al., 2007). Type III crustins possess a short proline-rich peptide (PRP) region between the signal peptide region and the single WAP domain. However, these do not contain neither the characteristic cysteine-rich domain present in both Type I and II crustins nor the glycine region motif. Only a few of these are known and so far reports of them are confined to the shrimps, *P. monodon* (Supungul et al., 2004), *L. vannamei* (Jimenez-Vega et al., 2004), *M. japonicus* (Rojtinnakorn et al., 2002) and *F. chinensis*. In the literature these types of proteins have been termed not as crustins, but as single-whey domain (SWD) proteins (Jimenez-Vega et al., 2004), chelonianin-like proteins (Chen et al., 2005) or anti-leuko proteinase-like proteins (Rojtinnakorn et al., 2002).

The genomic organization of crustins, was also found to be distinct among different groups. The genes of two Type I crustins are encoded by four exons interrupted by three introns (Brockton et al., 2007; Imjongjirak et al., 2009). Whereas in case of Type III crustins there exists only three exons and two introns (Amparyup et al., 2008a). Furthermore, two crustin variants from the same shrimp species, *P. monodon,* was found to be encoded by different genes; crustinPm5 comprised of four exons separated by three introns, and crustin-likePm comprising only of two exons and one intron (Amparyup et al., 2008b; Vatanavicharn et al., 2009). Analysis of genomic and recombinantly expressed carcinin has revealed that the signal and mature sections are probably not encoded by separate exons (Brockton et al., 2007). It is unclear if the signal sequence

is directing trans-membrane transportation of the peptide, as in the case of insects and mammals, or if the mature protein is released from the haemocytes through regulated exocytosis, as in the case of penaeidins (Destoumieux et al., 2000a, 2000b; Munoz et al., 2002).

Crustin-encoding genes are mainly expressed in the circulating haemocytes, although some reports point to the possibility of expression in other tissues such as heart, ovary and intestines (Smith et al., 2008). However, as these tissues are highly vascularised, it is not clear if the signal from these organs is primarily due to the haemocytes. Interestingly, transcripts of different crustin types, have also been essentially identified in the epipodite, hematopoietic tissue, ovary and olfactory organ, respectively employing RT-PCR (Stoss et al., 2003; Jiravanichpaisal et al., 2007; Vatanavicharn et al., 2009; Sun et al., 2010). Patterns of expression during development are also very poorly understood, as most studies of crustins have been performed only on late-stage postlarvae or adults. Larvae of the shrimp, *P. monodon*, have been reported to express a crustin transcript at high levels at all stages of development from nauplii stage IV to juveniles (Jiravanichpaisal et al., 2007).

Molecular studies have revealed that some decapod crustaceans express more than one type of crustin that differ from one another only by a few amino acids, usually 1–4. This has been reported in *P. monodon* (Supungul et al., 2004, 2008), *L. setiferus* (Bartlett et al., 2002) and *M. japonicas* (Rattanachai et al., 2004). The potential number of variants of the encoded protein can thus, at least theoretically, be quite large. An interesting situation exists in *L. vannamei* for which Bartlett et al. (2002) have reported the occurrence of six possible crustin variants while Vargas-Albores et al. (2004) have discovered two. Depending on the

amino acids involved and their location within the protein, isoforms may vary slightly in biological activity and undergo positive selection if any confer survival advantages to their host under local conditions.

Major work on different aspects of crustins have been carried out in *L. vannamei* (Gross et al., 2001; Bartlett et al., 2002; Vargas-Albores et al., 2004; Jimenez-Vega et al., 2004), *L. setiferus* (Gross et al., 2001; Bartlett et al., 2002; Rosa et al., 2007), *P. monodon* (Jimenez-Vega et al., 2004; Supungul et al., 2004, 2008; Chen et al., 2004a, 2005; Jiravanichpaisal et al 2007; Zhang et al., 2007; Amparyup et al., 2008b), *M. japonicus* (Rattanachai et al., 2004; Zhang et al., 2007), *L. schmitti* (Rosa et al., 2007) *F. chinensis* (Zhang et al., 2007; Amparyup et al., 2008a), *F. brasiliensis* (Rosa et al., 2007), *F. paulensis* (Rosa et al., 2007), *F. subtilis* (Rosa et al., 2007), *S. paramamosain* (Imjongirak et al., 2009), *E. sinensis* (Mu et al., 2010, 2011), *H. araneus* (Sperstad et al., 2009), *P. trituberculatus* (Xu and Liu., 2011) and *C. maenas* (Brockton et al., 2007).

Crustins are widely regarded as antimicrobial effectors, hence, there have been several studies demonstrating their antimicrobial properties *in vitro*. Evidence of the role of crustins as direct antimicrobial defence effectors has been sought from studies of their expression in different tissues and following experimental infection. Most crustins seem to be constitutively expressed by the granular haemocytes often at very high levels and are found to be active against bacteria, fungi and virus (Relf et al., 1999; Jimenez-Vega et al., 2004; Zhang et al., 2007; Supungul et al., 2008; Imjongjirak et al., 2009; Antony et al., 2011; Krusong et al., 2012). Previous researches regarding the biological activities of crustins have revealed that all crustin groups are mainly active against Gram-positive bacteria with an average MIC of < 8 μM. Crustins have proved to possess

antibacterial activity against Gram-positive strains of the genera *Micrococcus, Aerococcus, Planococcus, Staphylococcus, Streptococcus, Corynebacterium* and *Bacillus* (Relf et al., 1999; Zhang et al., 2007; Supungul et al., 2008; Imjongjirak et al., 2009; Sperstad et al., 2009). On the contrary, a Type II crustin from *P. monodon* (crustin-likePm) showed strong antibacterial activity against both Gram-positive and Gram-negative bacteria, including the crustacean opportunist/pathogen *Vibrio harveyi,* with an average MIC of < 5 μM (Amparyup et al., 2008b). Biological activities of crustins such as protease inhibitory and antibacterial activities seem to be related to a particular structure of the WAP domain. Protease inhibitory activity is generally associated with the presence of a methionine residue adjacent to the second cysteine in the 4DSC. This residue is substituted by cationic or hydrophobic amino acids in WAP-containing proteins with antibacterial activity (Hagiwara et al., 2003). Despite the presence of the WAP domain in all crustin groups, anti-protease activities have only been reported for Type III crustins (Amparyup et al., 2008a; Jia et al., 2008). This anti-protease capacity could be important to inactivate microbial proteases during infection and/or regulate endogenous protease cascades, such as the proPO system that leads to melanization. Thus, type III crustins seem to be multifunctional immune proteins that combine both antimicrobial and anti-protease properties (Amparyup et al., 2008a; Jia et al., 2008).

Carcinin, which was purified from the haemocytes of *C. maenas* on the basis of its ability to inhibit the growth of bacteria, appears to be active against Gram-positive alone (Relf et al., 1999). The crustins purified from the spider crab, *H. araneus* and shrimp *F. chinensis*, also seem to kill Gram-positive bacteria (Haug et al., 2006; Zhang et al., 2007), but no detectable

activity against Gram-negative bacteria or fungi have been so far reported (Zhang et al., 2007). A crustin type II from *P. monodon* has been found to inhibit the growth of Gram-negative bacteria (Amparyup et al., 2008b). Thus, whilst several crustins clearly have antibacterial effects primarily against Gram-positive bacteria, broader specificity to deal with a wider range of pathogens can occur and seems to be achieved through sequence variations. The extent to which single nucleotide substitutions, as occurs in isoforms, alter the function remains largely untested (Smith et al., 2008).

Unfortunately, the exact mechanism of action of crustins still remains enigmatic. As all crustins share a common WAP domain, it is believed that, this part of the molecule must play a key role in its antimicrobial activities. The antimicrobial activity of crustins appears to be related to the tertiary structure of the 4DSC (tightly constrained by three disulfide bonds and containing a small α-helix). Zhang et al. (2007) revealed that the crustin-like CshFc from *F. chinensis,* which lacks the authentic WAP domain, did not affect any bacteria tested, in contrast to the WAP-containing Type II crustin CruFc, which inhibited the growth of Gram-positive bacteria *in vitro.*

The economic role of marine and freshwater crustaceans as a food source in the export market demand the need to augment fishery resources through the adoption of intensive culture practices. This has led to physiological stress on cultured organisms, often increasing the incidence of diseases (Lorenzon et al., 1999). Despite the development of safe and potent antibiotics, bacterial diseases remain a worldwide health crisis due to the emergence of multiple drug resistant pathogens (Smith et al., 2003). The use of AMPs as a therapeutic tool has been among the most

promising avenues investigated to date, for addressing antibiotic resistance.

The wide distribution of crustins in crustaceans indicates the importance of these AMPs in the innate immunity and adetailed research is binding to reveal the enigmatic character of crustins. However, so far crustin has not been reported from the mud crab, *S. serrata* and the blue swimmer crab, *P. pelagicus*. Both the crab species are exploited heavily from estuaries and inshore waters and are also widely cultured in India due to their high market value as export commodity. The present study is focused at molecular identification and characterization of crustin AMPs from the haemocytes of *S. serrata* and *P. pelagicus*. The study of crab immune system, especially the AMPs, is essential to better comprehend their immune potential and open up new perspectives for developing novel strategies to prevent and control infections in crab culture systems.

3.2 Materials and Methods

3.2.1 Experimental animals

The experimental organisms used in the present study include the mud crab, *S. serrata* and the blue swimmer crab, *P. pelagicus* belonging to decapod crustaceans of the brachyuran family, Portunidae. Details of the experimental organism are given in section 2.2.1 of Chapter 2.

3.2.2 RNase control

Basic precautions were taken when working with RNA as described in section 2.2.2 of Chapter 2.

3.2.3 Haemolymph collection

Haemolymph was collected from the experimental organisms and was suspended in TRI reagent for total RNA isolation as given in section 2.2.3 of Chapter 2.

3.2.4 RNA isolation

Total RNA was extracted from the haemocytes of the experimental organisms using TRI™ reagent (Sigma) following manufacturer's protocol as described in section 2.2.4 of Chapter 2. RNA pellets were dried for 5-10 min and dissolved in RNase free water by repeated pipetting with a micropipette at 55-60 ^{0}C for 10-15 min.

3.2.5 Determination of the quantity and quality of RNA

RNA was quantified and qualified by measuring spectrophotometry and agarose gel electrophoresis as given in section 2.2.5 of Chapter 2.

3.2.6 cDNA synthesis

cDNA was synthesized using oligo d(T_{20}) primers as described previously in section 2.2.6 of Chapter 2.

3.2.7 PCR amplification

PCR amplification of cDNA were carried out using gene specific primers, designed using GeneTool software based on consensus sequences of crustins retrieved from GenBank. β-actin was used as the internal control to verify reverse transcription reaction. The sequences of the primers used to amplify β-actin and crustin genes are given in Table. 3.1.

PCR amplification of 1µl cDNA was performed in a 25 µl reaction volume using gene specific primers as explained in section 2.2.7 of Chapter 2.

Annealing temperature varied for the different genes as given in Table 3.1. The PCR products were visualized by electrophoresis in 1.5 % agarose gel.

3.2.8 Agarose gel electrophoresis

Electrophoresis was performed in 1.5 % agarose gel and documented as described in section 2.2.8 of Chapter 2.

3.2.9 Cloning of the PCR product

The PCR products were cloned onto the pGEM-T Easy vector as explained in section 2.2.9 of Chapter 2. *E. coli* DH5α competent cells were used for transforming the ligated cloning vector. To confirm the presence of the insert DNA in the transformed cells, colony PCR was performed for all the individually streaked colonies using vector specific primers viz. T7 and SP6 as well as gene specific primers (Table 3.1). Plasmid with the insert was then extracted and purified using GenElute HP' plasmid MiniPrep kit (Sigma) and was sent for sequencing at SciGenom, Kochi, India.

3.2.10 Sequencing and sequence analysis

Bioinformatics analysis of the sequences were performed using various softwares / programs as described in section 2.2.10 of Chapter 2.

3.3 Results

In the present study, three different isoforms of crustins could be identified, cloned and characterized from *S. serrata* (SsCrustin1 and SsCrustin2) and *P. pelagicus* (PpCrustin1) (Table 3.2).

3.1.1 Molecular characterization of crustins in *S. Serrata*

From the haemocytes of the mud crab, *S. serrata*, two new isoforms of AMPs belonging to crustin family could be identified and characterized, herein after referred to as SsCrustin1 and SsCrustin2.

3.3.1.1 SsCrustin1

3.3.1.1.1. Cloning and sequencing of SsCrustin1 cDNA

An initial RT-PCR reaction with *S. serrata* haemocyte RNA and gene specific primer crustin-1, a product of ~ 433 bp could be amplified which was subsequently cloned and sequenced. The SsCrustin1 cDNA encoded a polypeptide of 111 amino acids in the ORF with a putative signal peptide of 21 amino acid residues followed by a mature protein of 90 amino acids (Fig. 3.1 and 3.2). The full-length sequence was deposited in the NCBI GenBank under accession number **HQ638025.**

3.3.1.1.2 Analysis of SsCrustin1 sequence and peptide characteristics

The deduced amino acid sequence of SsCrustin1 was found to be rich in amino acid residues cysteine (13.3 %) and proline (11.1 %). A WAP domain could be detected in the C-terminus of SsCrustin1. A conserved eight cysteine residue region responsible for the formation of 4 disulfide core (4-DSC), could also be detected in the C_5 to C_{12} position (Fig. 3.1). The physico-chemical properties of SsCrustin1 including its molecular weight, net charge, isoelectric point and hydrophobicity predicted using PROTPARAM, ExPASy, APD and HELIQUEST programs are described in Table 3.3. The calculated molecular mass of the mature protein was 10.24 kDa. The isoelectric point (*pI*) was estimated to be 8.76 as predicted by the PROTPARAM software. The SsCrustin1 was found to be highly cationic

and possessed a net charge of +6 and hydrophobicity of 33 %. Peptide model of SsCrustin1 created using SWISS-MODEL server and PyMOL indicated a random coiled structure that is with two possible β-sheets but no helices (Fig. 3.3). The 12 cysteines in SsCrustin1 are considered to be important for maintaining the tertiary structure of the peptide.

3.3.1.1.3 Sequence alignment and phylogenetic analysis of SsCrustin1

Sequence comparison using BLAST algorithm showed that the deduced amino acid sequence of SsCrustin1 possessed an overall similarity of 81 %, 62 %, 73 %, 56 % and 39 % to the crustins of *S. paramamosain*, *P. trituberculatus*, *H. araneus*, *C. maenas* and *F. chinensis* respectively (Table 3.4). The deduced amino acid sequence of SsCrustin1 was found to be rich in amino acid residues, cysteine (13.3 %) and proline (11.1 %).

Multiple alignment performed for SsCrustin1 with other crustins of decapods revealed the presence of conserved regions within the sequence (Fig. 3.4).

The phylogenetic relationship between SsCrustin1 and other crustins of decapod crustaceans were analyzed using neighbor-joining (NJ) method. Molecular phylogenetic tree based on amino acid sequences suggests that all the crustin members possess the same ancestral origin, which has subsequently diverged at different phases of evolution. The tree could be broadly divided into four major groups; Group 1 included crustins from crabs, Group 2 included crustins of crayfishes, Group 3 included crustins of lobsters and Group 4 included crustins of shrimps. SsCrustin1 was found to cluster with Group 1 comprising the crab crustins. The bootstrap

distance tree calculated for the SsCrustin1 clearly indicated that the SsCrustin1 possessed great similarity to crustins of other crabs (Fig. 3.5).

3.3.1.2 SsCrustin2

3.3.1.2.1 Cloning and sequencing of SsCrustin2 cDNA

A 463 bp nucleotide sequence representing the complete cDNA sequence of SsCrustin2 with an ORF of 336bp was obtained by RT-PCR (Fig. 3.6). The analysis with SignalP software revealed the presence of a signal peptide with 21 amino acid residues at the N-terminus of SsCrustin2, with a predicted cleavage site between Ala_{21} and Ser_{22}, resulting in a 10 kDa (90 residue) mature protein (Fig. 3.7). The ORF of SsCrustin2 encoded a polypeptide of 111 amino acids with a molecular weight of 10 kDa and a *p*I of 8 as predicted by the PROTPARAM software. The sequence of SsCrustin2 was submitted in NCBI GenBank under the accession number **JQ753312**.

3.3.1.2.2 Analysis of SsCrustin2 sequence and peptide characteristics

The deduced amino acid sequence of SsCrustin2 was found to be rich in amino acid residues cysteine (13.3 %) and proline (11-12 %). A WAP domain could be detected in the C-terminus of SsCrustin2. The WAP domain was detected in the C_5–C_{12} position of SsCrustin2 and contained eight-cysteine residues in a conserved arrangement. SsCrustin2 possessed 5 residues between C_7 and C_8 (C_7 X_5 C_8). Several other consensus sequences also appear in the 4-DSC domain viz. (1) a conserved aspartate (D) residue between C_3 and C_4 and (2) KCC with C_5 and C_6.

SsCrustin2 lacked the glycine-rich repeats which are usually found at the N-terminal of the mature peptide of shrimp crustins. However, the C-

terminal segment included a high proportion of cysteine-rich region (13.3 %), which contained 12 cysteine residues that participate in the formation of disulphide bonds. SsCrustin2 lacked a methionine residue adjacent to the second cysteine within the 4-DSC. However, the WAP domain of SsCrustin2 appear to form part of a larger pattern of 12 cysteine residues within the crustin sequence which is conserved across the various crustin isoforms from several crustacean species. The physico-chemical properties of SsCrustin2 are described in Table 3.3. The mature protein of SsCrustin2 possessed a predicted molecular mass of 10 kDa and an isoelectric point (*pI*) of 8. The SsCrustin2 was found to be a cationic peptide with a net charge of +3 and its hydrophobicity was calculated as 33 %. Peptide models of SsCrustin2 created using SWISS MODEL and PyMOL software also indicated a tightly coiled structure enclosing two β-sheets but no helices or possibly a loop protein (Fig. 3.8).

3.3.1.2.3 Sequence alignment and phylogenetic analysis of SsCrustin2

Sequence comparison using BLAST algorithm showed that the deduced amino acid sequence of SsCrustin2 shared significant identity with crustins of portunid crabs (68–95 %) and other crabs (60-73 %) (Table 3.6). SsCrustin2 sequences were highly similar to the crustins of *S. serrata* and *S. paramamosain* and shared 88 % identity among each other. However, SsCrustin2 was found to share only a low homology with the crustins of crayfishes, lobsters and shrimps (less than 40 %).

Multiple alignment of SsCrustin2 with other crustins of decapods revealed the presence of conserved regions within the sequence and also revealed maximum identity to that of *S. serrata* and *S. paramamosain* (Fig. 3.9).

To evaluate the molecular evolutional relationships of SsCrustin2 against other crustin family members, a phylogenetic tree was constructed based on the amino acid sequences of crustins with WAP domain by the neighbor-joining method (Fig. 3.10). The tree topologies revealed the relationships of SsCrustin2 with other crustin-like peptides of decapod crustaceans. Phylogenetic analysis of crustins further revealed that crustin sequences were clustered according to species. Phylogenetic trees also showed that all crustins including SsCrustin2, possess the same ancestral origin and have a similar evolutionary status, which has subsequently diverged at different phases of evolution and also that they were phylogenetically ancient immune effector molecules which may play an essential role in the host defense mechanism. SsCrustin2 was closely related to crab crustins rather than to the other groups viz. shrimps, prawns, lobsters and crayfishes.

The phylogenetic tree could be broadly divided into two major groups, Group 1 included crustins from shrimps and prawns; and Group 2 included crustins from crayfishes, lobsters and crabs respectively. Within Group 2, there were three distinct subgroups: first subgroup of lobster and crayfish crustins, second subgroup of crab crustins to which the SsCrustin2 belonged to and third subgroup containing crustins from *P. trituberculatus* only. The bootstrap distance tree calculated for the SsCrustin2 clearly indicated that this possessed great similarity to another crustin isolated from *S. serrata* (SsCrustin1). Great variability could also be noticed in the crustin sequences of various decapods. Phylogenetic tree also showed that crustins from crabs including SsCrustin2, PpCrustin1, SsCrustin1 and crustins from *C. maenas, H. araneus, S. paramamosain* and *P. trituberculatus* clustered together, and

had a closer relationship with crayfish *P. clarkia* and lobster (*N. norvegicus* and *H. americanus*) crustins than to the shrimp and prawn crustins.

3.3.2 Molecular characterization of crustins in *P. pelagicus*

A new isoform of AMP belonging to crustin family could be identified, cloned and characterized from the haemocytes of *P. pelagicus*, herein after referred to as PpCrustin1. This is the first crustin to be reported from this crab species.

3.3.2.1 PpCrustin1

3.3.2.1.1 Cloning and sequencing of PpCrustin1 cDNA

The complete cDNA sequence of PpCrustin1 obtained was comprised of a 457 bp nucleotide sequence with an ORF of 336 bp encoding a polypeptide of 111 amino acids (Fig. 3.11). The analysis with SignalP software revealed the presence of a signal peptide with 21 amino acid residues at the N-terminus of PpCrustin1, with a predicted cleavage site between Ala_{21} and Ser_{22}, resulting in a 10 kDa (90 residue) mature protein (Fig. 3.12). The sequences of PpCrustin1 were submitted in NCBI GenBank under the accession number **JQ965930**.

3.3.2.1.2 Analysis of PpCrustin1 sequence and peptide characteristics

The deduced amino acid sequence of PpCrustin1 was found to be rich in amino acid residues cysteine (13.3 %) and proline (11-12 %). A WAP domain could be detected in the C-terminus of PpCrustin1. The WAP domain was the only domain found in PpCrustin1 sequence obtained in the present study. This domain forms a tightly packed structure described

on PROSITE as a four-disulphide core (4DSC). The WAP domain was detected in the C_5–C_{12} position of PpCrustin1 and contained eight-cysteine residues in a conserved arrangement. PpCrustin1 possessed 5 residues between C_7 and C_8 (C_7 X_5 C_8). Several other consensus sequences also appear in the 4-DSC domain viz. (1) a conserved aspartate (D) residue between C_3 and C_4 and (2) KCC with C_5 and C_6. PpCrustin1 lacked the glycine-rich repeats which are usually found at the N-terminal of the mature peptide of shrimp crustins. The C-terminal segment included a high proportion of cysteine-rich region (13.3 %), which contained 12 cysteine residues that participate in the formation of disulphide bonds. PpCrustin1 lacked a methionine residue adjacent to the second cysteine within the 4-DSC. The molecular weight, isolelectric point, charge and hydrophobicity of PpCrustin1 predicted using softwares are given in Table 3.7. The molecular weight of mature peptide of PpCrustin1 was calculated as 10 kDa and a p*I* of 8 was predicted by the PROTPARAM software. PpCrustin1 showed a high cationicity with a charge of +6 and its hydrophobicity was 34 %. Peptide models of PpCrustin1 created using SWISS MODEL and PyMOL software also indicated a tightly coiled structure enclosing two β-sheets but no helices or possibly a loop protein (Fig. 3.13).

3.3.1.1.3 Sequence alignment and phylogenetic analysis of PpCrustin1

Sequence comparison using BLAST algorithm showed that the deduced amino acid sequence of PpCrustin1 shared significant identity with crustins of portunid crabs (68–95 %) and other crabs (60-73 %) as given in Table 3.8. PpCrustin1 sequences were highly similar to the crustins of *S. serrata* and *S. paramamosain.* However, PpCrustin1 was found to share

only a low homology with the crustins of crayfishes, lobsters and shrimps (less than 40 %). Multiple alignment of PpCrustin1 with other crustins of decapods revealed the presence of conserved regions within the sequence and also revealed maximum identity to that of *S. serrata* and *S. paramamosain* (Fig. 3.14). PpCrustin1 showed variations at six amino acid positions in their signal peptide region. In PpCrustin1, the aspartic acid and lysine residues were found to be positioned as follows:-C-XX-D-XX-C-XXXD-K-CC-X-D-. To evaluate the molecular evolutional relationships of PpCrustin1 against other crustin family members, a phylogenetic tree was constructed based on the amino acid sequences of crustins with WAP domain by the neighbor-joining method (Fig. 3.15). The tree topologies revealed the relationships of PpCrustin1 with other crustin-like peptides of decapod crustaceans. Phylogenetic analysis of crustins further revealed that crustin sequences were clustered according to species. Phylogenetic trees also showed that all crustins including PpCrustin1, possess the same ancestral origin and have a similar evolutionary status, which has subsequently diverged at different phases of evolution. PpCrustin1 was closely related to crab crustins rather than to the other groups viz. shrimps, prawns, lobsters and crayfishes.

The phylogenetic tree could be broadly divided into two major groups, Group 1 included crustins from shrimps and prawns; and Group 2 included crustins from crayfishes, lobsters and crabs respectively. Within Group 2, there were three distinct subgroups: first subgroup of lobster and crayfish crustins, second subgroup of crab crustins to which PpCrustin1 belonged to and third subgroup containing crustins from *P. trituberculatus* alone. The bootstrap distance tree calculated for PpCrustin1 clearly indicated that this possessed great similarity to crustin

isolated from *S. serrata* (SsCrustin1). Phylogenetic tree also showed that crustins from crabs including PpCrustin1, SsCrustin1, SsCrustin2 and crustins from *C. maenas, H. araneus, S. paramamosain* and *P. trituberculatus* clustered together, and had a closer relationship with crayfish *P. clarkia* and lobster (*N. norvegicus* and *H. americanus*) crustins than to the shrimp and prawn crustins.

3.4 Discussion

Crustins are cationic AMPs of ~ 7–14 kDa with a characteristic four-disulphide core containing WAP domain, present in the haemocytes of crustaceans. The present study reports the first crustin sequences from two portunid crabs, viz. the mud crab *S. serrata* (SsCrustin1 and SsCrustin2) and the blue swimmer crab *P. pelagicus* (PpCrustin1). All the three crustin isoforms possessed the complete ORF of 336 bp encoding 111 aminoacids with a predicted molecular weight of ~ 10-11 kDa and a *p*Iof 8-9. All the three isoforms possessed a signal sequence containing 21 aminoacids residues, which was followed by a mature peptide with a WAP domain at the C-terminus. The presence of conserved WAP domain and its characteristic features and high similarity with already known crustins proved the identity of these new isoforms belonging to the crustin family.

Signal peptide of both SsCrustin2 and PpCrustin1 possessed a predicted cleavage site between Ala_{21} and Ser_{22}, resulting in a 10 kDa (90 residue) mature protein. Whereas in case of SsCrustin1 it was found between Ser and Arg at the same position i.e. between 21 and 22. All the three crustin isoforms were found to be cationic. Considering the net charge of these isoforms, SsCrustin2 possessed the lowest net charge of +3 followed by +6 of SsCrustin1 and PpCrustin1 respectively. However, not much variation

in the hydrophobicity could be noticed and the hydrophobicity was found to range between 33-34 %. All the above features are in agreement with the required features of crustins. The above mentioned features of the newly identified crustins are in agreement with Smith and co-workers (2008), who defined crustins as cationic cysteine-rich antibacterial polypeptides of ~ 7–14 kDa, with an isoelectric point usually in the range of 7.0–8.7, present in crustaceans that contain one WAP domain at the C-terminus.

The deduced amino acid sequence of the newly identified crustins were found to be rich in amino acid residues cysteine (13.3 %) and proline (11-12 %). They lacked the glycine-rich repeats which are usually found at the N-terminal of the mature peptide of shrimp crustins (Antony et al., 2010). However, the C-terminal segment included a high proportion of cysteine-rich region, which contained 12 cysteine residues that participate in the formation of disulphide bonds. They also lacked a methionine residue adjacent to the second cysteine within the 4-DSC, indicating that this WAP domain has antibacterial function (Hauton et al., 2006).

The WAP domain was the only domain found in the sesequences obtained in the present study. This domain forms a tightly packed structure described on PROSITE as a four-disulphide core (4DSC). The WAP domain was detected in the C_5–C_{12} position and contained eight cysteine residues in a conserved arrangement which is believed to be responsible for the formation of 4-DSC as described by Imjongjirak et al. (2009). PROSITE database analysis revealed the existence of WAP-type 4-DSC domain signature, C_1-(X_n)-C_2-(X_n)-C_3-(X_5)-C_4-(X_5)-C_5-C_6-(X_{3-5})-C_7-(X_{3-4})-C_8 (X indicates any residue and X_n is a stretch of n residues.). However these crustins possessed five residues between C_7 and C_8 (C_7 X_5 C_8). Several

other consensus sequences also appear in the 4-DSC domain viz. a conserved aspartate (D) residue between C_3 and C_4 and KCC with C_5 and C_6. According to the previous reports on the crustin-like peptides, the 4-DSC domain has proved to play important roles in the biological function of crustins (Zhang et al., 2007).

However, the WAP domain of SsCrustin2 and PpCrustin1 appears to form part of a larger pattern of 12 cysteine residues within the crustin sequence which is conserved across the various crustin isoforms from several crustacean species. This pattern of 12 cysteine residues is also found in other crustin/carcinin-like transcripts from *S. serrata, C. maenas*, *M. japonicus*, *P. argus*, *H. americanus*, *C. sapidus* and *P. monodon*. These 12 cysteines are considered to be important for maintaining the tertiary structure of the peptide just as that reported in the case of other crustins (Gross et al., 2001; Zhang et al., 2007). As per Brockton et al. (2007), if all the 12 cysteine residues remain conserved in these sequences and if all six disulphide bonds are formed in the tertiary structure, then it would be the first family of six disulphide-bonded AMPs. These 12 cysteine residues present in crustin/carcinin-like peptides represents a new structural family that appear to be unique to the crustaceans. Peptide models of the three crustin isoforms created using PyMOL software indicated a tightly coiled structure enclosing two β-sheets but no helices or possibly a loop protein.

All the three newly identified crustins shared significant similarities with crustins of portunid crabs (68–95 %) and other crabs (60–73 %). However, they were found to be the first isoform isolated from its own species. SsCrustin2 and PpCrustin1 sequences were highly similar to the crustins of *S. serrata* and *S. paramamosain* and shared 88 % identity

among each other. However, both SsCrustin2 and PpCrustin1 were found to share only a low homology with the crustins of crayfishes, lobsters and shrimps (less than 40 %).

Multiple alignment performed for these crustin isoforms with other crustins of decapods revealed the presence of conserved regions within the sequence and also revealed maximum identity to that of *S. serrata* and *S. paramamosain*. The signal peptide region of crustins was not found to be highly conserved as in the case of other AMP families. SsCrustin2 and PpCrustin1 showed variations at six amino acid positions in their signal peptide region. Whereas, the mature peptide region of SsCrustin2 and PpCrustin1 shared greater similarities among each other, except at the 2^{nd}, 7^{th}, 8^{th}, 20^{th}, 35^{th}, 38^{th} and 56^{th} amino acid positions. However, these amino acid positions were also found to vary in various crustin isoforms from other portunid crabs. The WAP domain, in contrast to the signal sequence, is highly conserved between species. In both SsCrustin2 and PpCrustin1, the aspartic acid and lysine residues were found to be positioned as follows: -C-XX-D-XX-C-XXXD-K-CC-X-D-, which is in agreement with Ranganathan et al. (1999). As per Smith et al. (2008), both SsCrustin2 and PpCrustin1 could be grouped under type-I crustins; being the most commonly found crustin type in crabs, lobsters and crayfishes. Type I crustin is believed to follow a similar domain arrangement as carcinins, the first crustacean WAP domain-containing AMP discovered (Relf et al., 1999).

Phylogenetic analysis showed that these isoforms possess the same ancestral origin and have a similar evolutionary status like other crustins, which has subsequently diverged at different phases of evolution. The tree topologies revealed the relationships of these potunid crustins with

other crustin-like peptides of decapod crustaceans. Phylogenetic analysis of crustins further revealed that crustin sequences were clustered according to species. Phylogenetic trees also showed that all crustins including SsCrustin1, SsCrustin2 and PpCrustin1, possess the same ancestral origin and have a similar evolutionary status, which has subsequently diverged at different phases of evolution and also that they were phylogenetically ancient immune effector molecules which may play an essential role in the host defense mechanism. These new portunid crustin isoforms were found to be closely related to crab crustins rather than to the other groups viz. shrimps, prawns, lobsters and crayfishes.

Great variability could also be noticed in the crustin sequences of various decapods. Phylogenetic tree also showed that crustins from crabs including SsCrustin1, SsCrustin2 and PpCrustin1 and crustins from *C. maenas, H. araneus, S. paramamosain* and *P. trituberculatus* clustered together, and had a closer relationship with crayfish *P. clarkia* and lobster (*N. norvegicus* and *H. americanus*) crustins than to the shrimp and prawn crustins. This might imply the complex evolution of crustins in these groups and potential similarity in biological functions.

The wide distribution of crustins in crustaceans indicates the importance of these AMPs in the innate immunity. From the characteristics of these crustin isoforms itself it is possible to predict that these peptides could function as a broad spectrum antibacterial peptide. Further research on the expression profile of these molecules in response to various environmental conditions and microbial infection would reveal their role in the protection of the animals from the onslaught of diseases. It could be stipulated that discovery of novel crustins might pave way to the

discovery of promising therapeutic / prophylactic agents in health management and disease control in crustacean aquaculture.

Fig. 3.1 Nucleic acid (above) and deduced amino acid (below) sequence of SsCrustin1. The signal peptide sequence (green), mature peptide region bright blue and light blue) and putative WAP domain (light blue) have been highlighted.

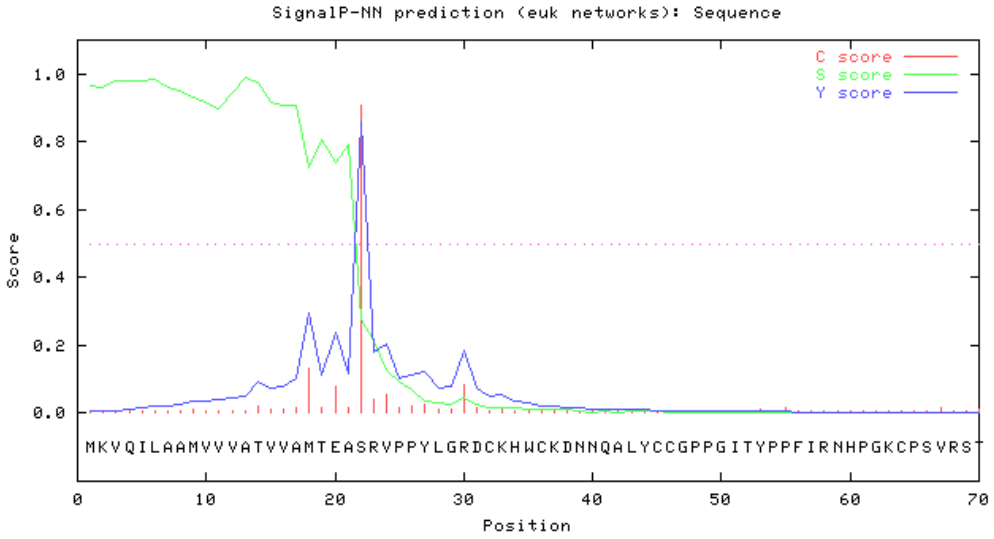


Fig 3.2 Signal peptide analysis of SsCrustin1 as predicted by the SignalP 3.0 server.

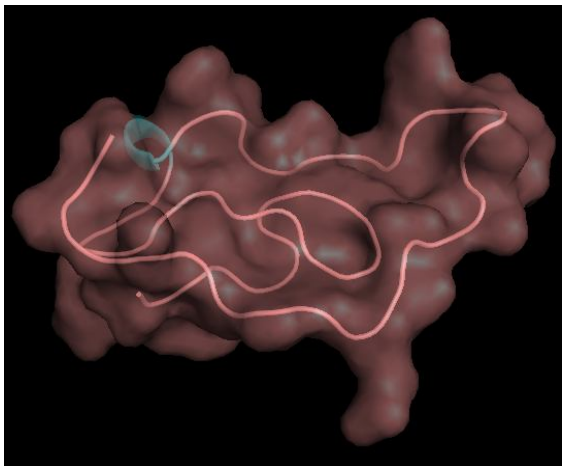


Fig.3.3 Secondary structure of SsCrustin1 created using PyMOL showing β-sheet and extended structures

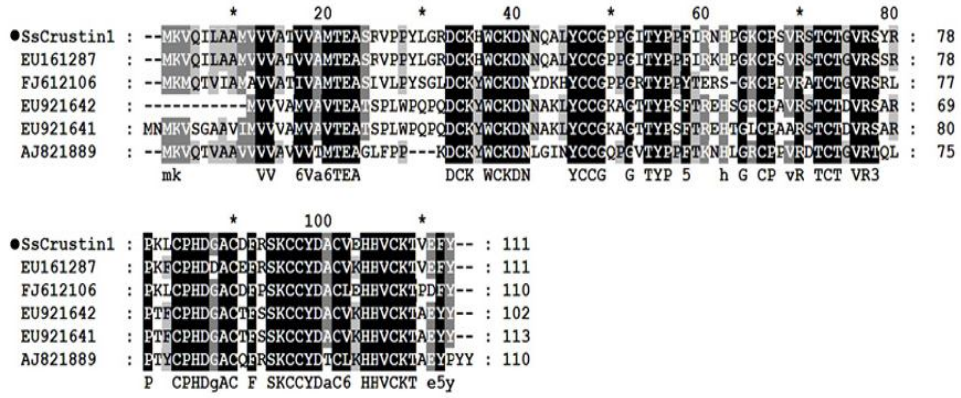


Fig.3.4 Multiple alignment of nucleotide sequence of SsCrustin1 with other crustins obtained using GeneDoc program. Black and grey indicates conserved sequences.

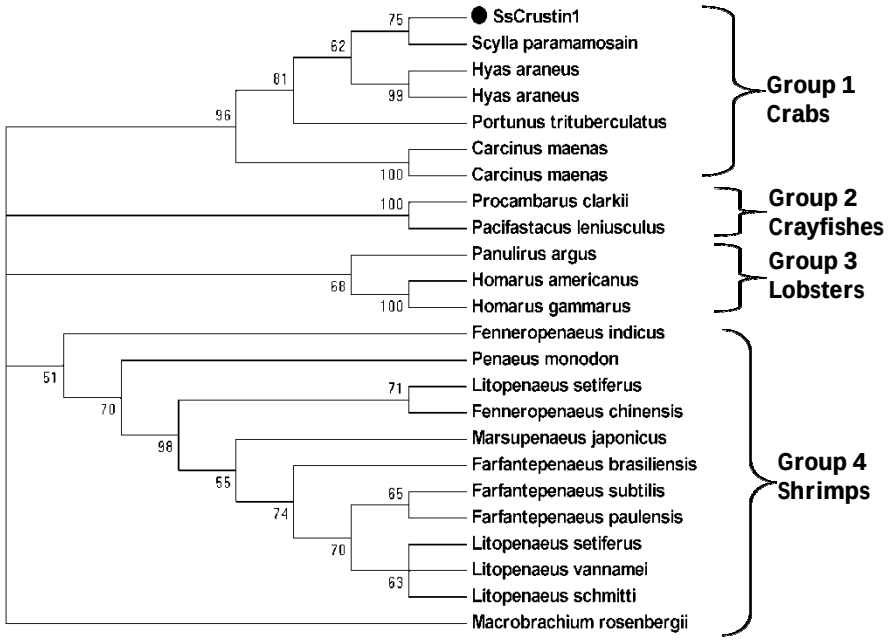


Fig.3.5 A bootstrapped neighbour-joining tree obtained using MEGA version 5.0 illustrating relationships between the deduced amino acid sequence SsCrustin1 with other crustins. Values at the node indicate the percentage of times that the particular node occurred in 1000 trees generated by bootstrapping the original deduced protein sequences.

caacggcttcttcagacctattgaaatatgaaggtgaaaattttagcagccatggtggtt

M K V K I L A A M V V

gtggctaccgttgtggccatgacagaagcatccctagtacctccatatccaggtcgggat

V A T V V A M T E A S L V P P Y P G R D

tgtaagcactggtgcaaagacaacaatgaagcactctactgctgcggccctccaggaatt

C K H W C K D N N E A L Y C C G P P G I

acctatccacctttaattagagagcaccctggtaaatgtccttcagtccgctctacatgt

T Y P P L I R E H P G K C P S V R S T C

actggtgtgaggtcatctagaccaaagctctgccccacgatggtgcttgtgactttaga

T G V R S S R P K L C P H D G A C D F R

agcaagtgttgctatgacgcctgtgtggagcaccacgtatgcaagactgttgaattctac

S K C C Y D A C V E H H V C K T V E F Y

taaacaacaattacaccagacctgagctgaaaatctttcacaagtaaaaaaatttcaaca

*

aactgtattcactggatttttttttttaaaccctgaagtagggt

Fig.3.6 Nucleic acid (above) and deduced amino acid (below) sequence of SsCrustin2. The signal peptide sequence (green), mature peptide region bright blue and light blue) and putative WAP domain (light blue) have been highlighted.

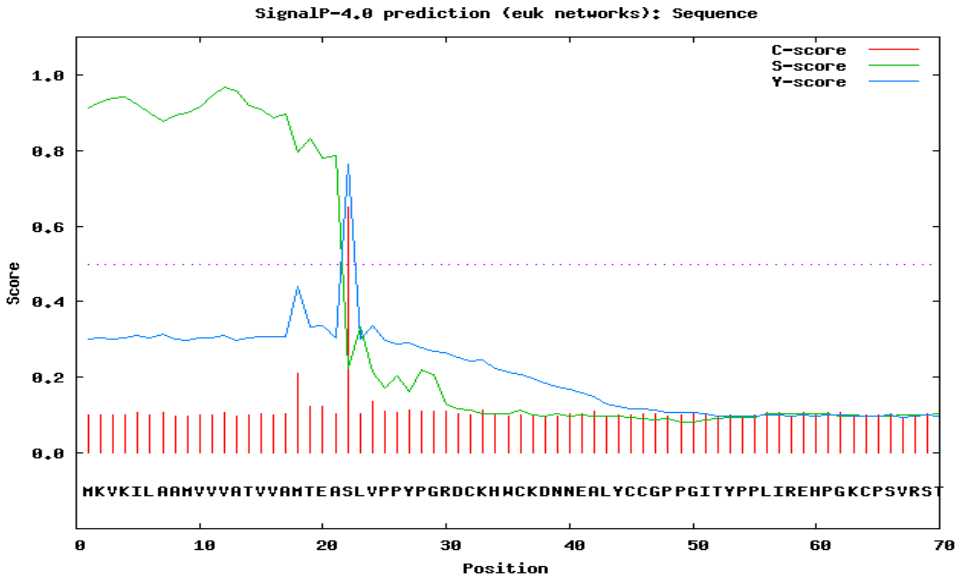


Fig.3.7 Signal peptide analysis of SsCrustin2 as predicted by the SignalP 3.0 server

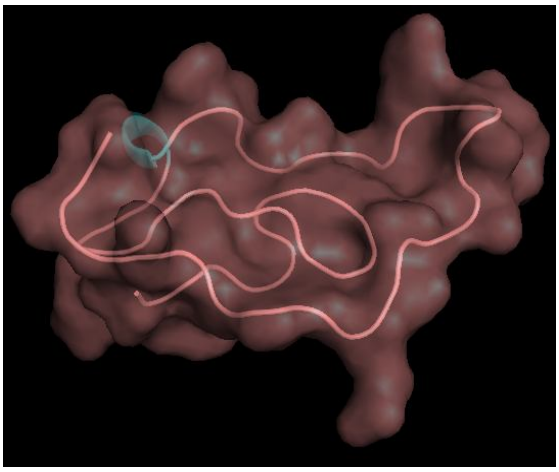


Fig.3.8 Secondary structure of SsCrustin2 created using PyMOL showing β-sheet and extended structures

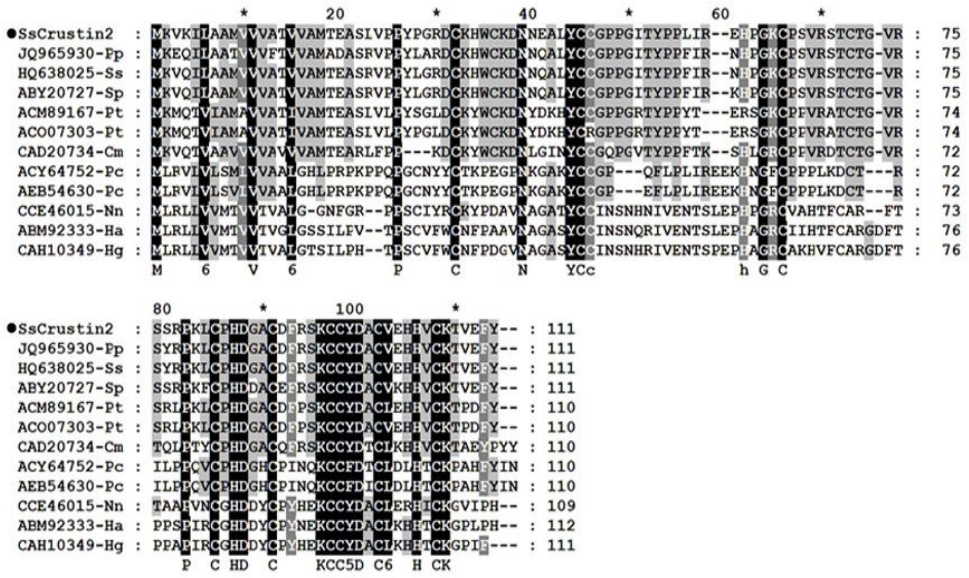


Fig.3.9 Multiple alignment of nucleotide sequence of SsCrustin2 with other crustins obtained using GeneDoc program. Black and grey indicates conserved sequences.

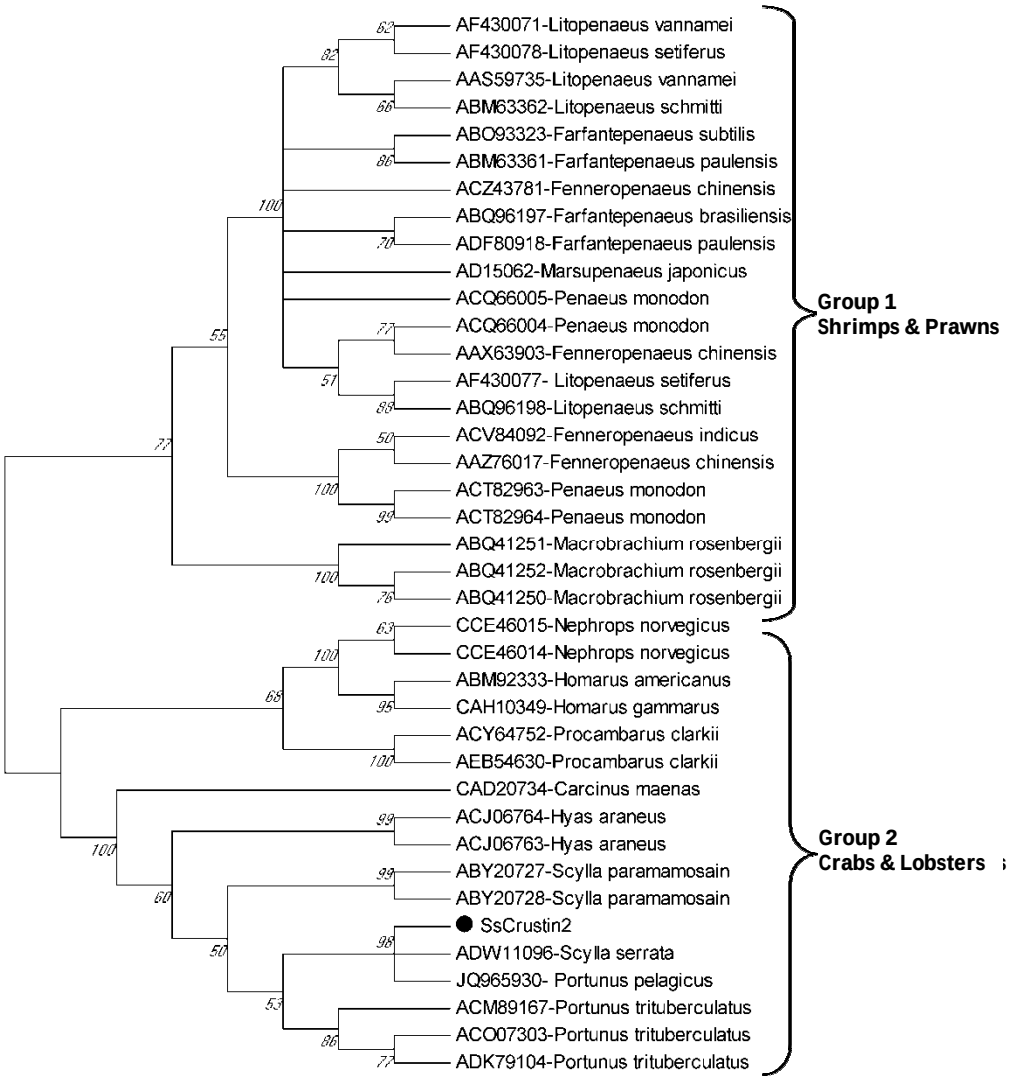


Fig.3.10 A bootstrapped neighbour-joining tree obtained using MEGA version 5.0 illustrating relationships between the deduced amino acid sequence SsCrustin2 with other crustins. Values at the node indicate the percentage of times that the particular node occurred in 1000 trees generated by bootstrapping the original deduced protein sequences.

Fig.3.11 Nucleic acid (above) and deduced amino acid (below) sequence of PpCrustin1. The signal peptide sequence (green), mature peptide region (bright blue and light blue) and putative WAP domain (light blue) have been highlighted.

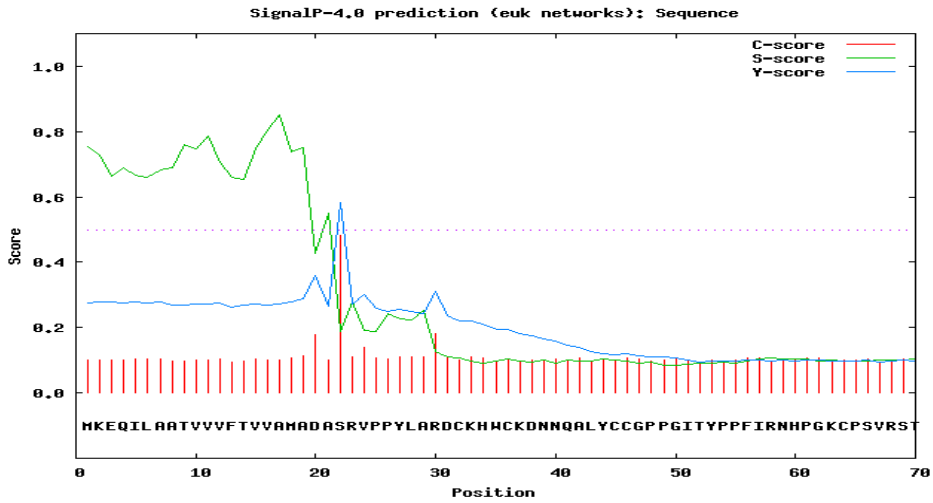


Fig.3.12 Signal peptide analysis of PpCrustin1 as predicted by the SignalP 3.0 server.

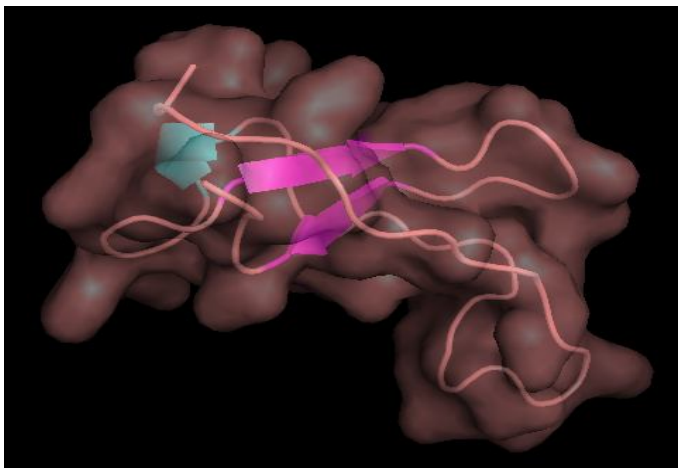


Fig.3.13 Secondary structure of PpCrustin1 created using PyMOL showing β-sheet and extended structures

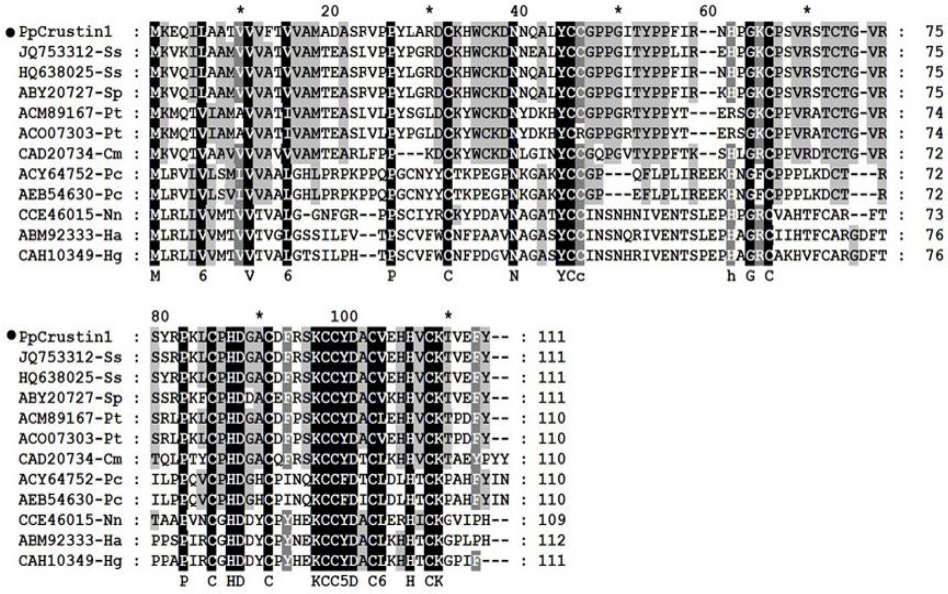


Fig.3.14 Multiple alignment of nucleotide sequence of PpCrustin1 with other crustins obtained using GeneDoc program. Black and grey indicates conserved sequences.

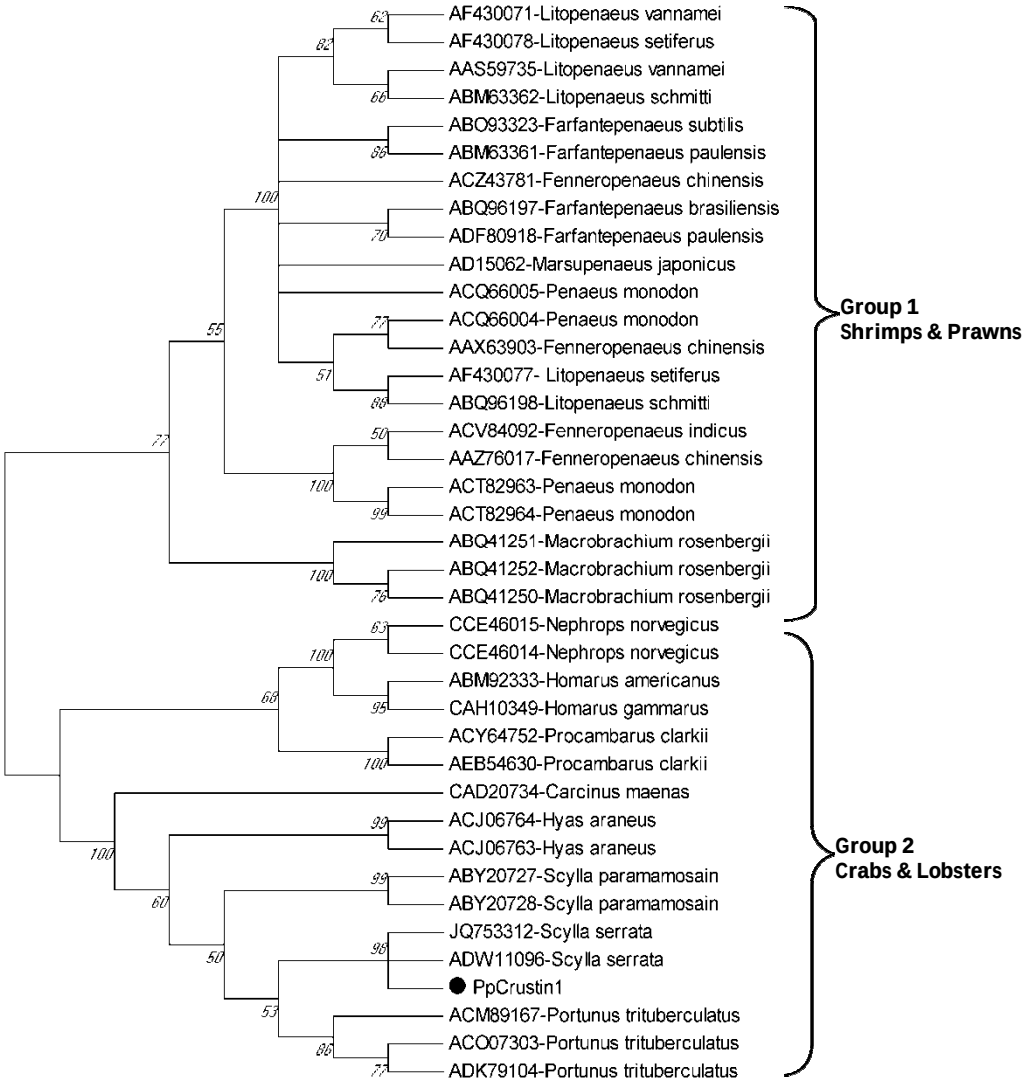


Fig.3.15 A bootstrapped neighbour-joining tree obtained using MEGA version 5.0 illustrating relationships between the deduced amino acid sequence PpCrustin1 with other crustins. Values at the node indicate the percentage of times that the particular node occurred in 1000 trees generated by bootstrapping the original deduced protein sequences.

Table 3.1 Primers used for the study

Primer	Primer Sequence (5'-3')	Product size (bp)	Annealing Temp. (°C)	$MgCl_2$ Conc. (mM)
Crustin-1	F: gagagcagaattagacactgt R: atatagtataacataaccatacttc	433	60	1.5
Crustin-2	F: agggagtgggtgatgagcta R: tacggctattacgatccaaca	463	60	1.5
β-actin	F: cttgtggttgacaatggctccg R: tggtgaaggagtagccacgctc	520	60	1.5
T7	tgtaatacgactcactataggg	--	57	1.5
SP6	gatttaggtgacactatag	--	57	1.5

Table 3.2 List of Crustins Identified in the present study

No:	Name of AMP	Source Organism	ORF	GenBank Accession
1	SsCrustin1	*Scylla serrata*	Complete	**HQ638025**
2	SsCrustin2	*Scylla serrata*	Complete	**JQ753312**
3	PpCrustin1	*Portunus pelagicus*	Complete	**JQ965930**

Table 3.3 Physicochemical properties of SsCrustin1

Parameters Studied	Results
Open reading frame (GeneTool)	111 aa
Length of mature peptide (No. of amino acids) (GeneTool)	90 aa
Signal peptide (SignalP)	1-21 aa
Molecular weight (PROTPARAM, ExPASy)	10.24 kDa
Isoelectric point (PROTPARAM, ExPASy)	8.76
Charge (PROTPARAM, ExPASy)	+6
Hydrophobicity (PROTPARAM, ExPASy)	33 %
Structure (PyMol & DNASTAR)	β-sheet

Table3.4 Result of BLASTp analysis of SsCrustin1

GenBank Accession No.	Description of the AMP	Query coverage	Identity to SsCrustin1
EU161287	*Scylla paramamosain* crustin antimicrobial peptide mRNA, complete cds	99 %	98 %
FJ467931	*Portunus trituberculatus* crustin mRNA, complete cds	90 %	82 %
EU161288	*Scylla paramamosain* crustin antimicrobial peptide gene, complete cds	83 %	99 %
GU373914	*Portunus trituberculatus* type III crustin mRNA, complete cds	54 %	82 %

Table3.5 Physicochemical properties of SsCrustin2

Parameters Studied	Results
Open reading frame (GeneTool)	111 aa
Length of mature peptide (No. of amino acids) (GeneTool)	90 aa
SIGNAL Peptide (SignalP)	1-21 aa
Molecular weight (PROTPARAM, ExPASy)	10.0 kDa
Isoelectric point (PROTPARAM, ExPASy)	8.0
Charge (PROTPARAM, ExPASy)	+3
Hydrophobicity (PROTPARAM, ExPASy)	33 %
Structure (PyMol & DNASTAR)	β-sheet

Table 3.6 Result of BLASTp analysis of SsCrustin2

GenBank Accession No.	Description of the AMP	Query coverage	Identity to SsCrustin2
ADW11096	Crustin [*Scylla serrata*]	100 %	94 %
ABY20727	Crustin [*Scylla paramamosain*]	100 %	91 %
ACM89167	Crustin [*Portunus trituberculatus*]	100 %	74 %
CAH25399	Carcinin-like protein [*Carcinus maenas*]	99 %	65 %

Table 3.7 Physicochemical properties of PpCrustin1

Parameters Studied	Results
Open reading frame (GeneTool)	111 aa
Length of mature peptide (No. of amino acids) (GeneTool)	90 aa
SIGNAL Peptide (SignalP)	1-21 aa
Molecular weight (PROTPARAM, ExPASy)	10.0 kDa
Isoelectric point (PROTPARAM, ExPASy)	8.0
Charge (PROTPARAM, ExPASy)	+6
Hydrophobicity (PROTPARAM, ExPASy)	34 %
Structure (PyMol & DNASTAR)	β-sheet

Table3.8 Result of BLASTp analysis of PpCrustin1

GenBank Accession No.	Description of the AMP	Query coverage	Identity to PpCrustin1
ADW11096	Crustin [*Scylla serrata*]	100 %	95 %
ABY20727	Crustin [*Scylla paramamosain*]	100 %	89 %
ACM89167	Crustin [*Portunus trituberculatus*]	100 %	68 %
CAH25399	Carcinin-like protein [*Carcinus maenas*]	99 %	63 %

4

Molecular Characterization and Phylogenetic Analysis of Penaeidins in Crustaceans

4.1 Introduction

Penaeid shrimps are among the most widely cultured marine organisms and hence possess high commercial importance. Infectious diseases caused by various viruses, bacteria and fungi have evolved as a serious threat to the shrimp aquaculture industry over the years. In shrimps, the innate immune responses play a major role in combating invading pathogens and prevent them against microbial attack. Antimicrobial peptides (AMPs) were known to play a critical role in shrimp innate immune defense mechanism. Penaeidins are a unique family of AMPs found to be confined to the Dendrobranchiata, a suborder of decapods crustaceans comprising of shrimps. They are cationic AMPs possessing a molecular weight of 5–7 kDa, characterized by an unstructured N-terminal proline-rich domain and a C-terminal region containing six cysteine residues that are engaged in three intramolecular disulfide bridges (Destoumieux et al., 1997; Yang et al., 2003; Cuthbertson et al., 2005).

Penaeidins were first isolated from the haemolymph of the Pacific white shrimp, *Litopenaeus vannamei* and showed antimicrobial activity mainly

against Gram-positive bacteria and fungi (Destoumieux et al., 1997). More studies based on genomic approaches have revealed the presence of penaeidins in different shrimp species. Several penaeidin isoforms have also been isolated and characterized from penaeid shrimps such as *L. vannamei* (Bartlett et al., 2002; Cuthbertson et al., 2002); *L. setiferus* (Cuthbertson et al., 2002; Bartlett et al., 2002); *L. stylirostris* (Munoz et al., 2004); *M. japonicus* (Rojtinnakorn et al., 2002); *P. monodon* (Supungul et al., 2002; Chen et al., 2004b; Chiou et al., 2005, 2006; Hu et al., 2006); *F. chinensis* (Kang et al., 2004); *F. paulensis* (Barracco et al., 2005) and *L. schmitti* (Barracco et al., 2005).

Penaeidins are composed of an N-terminal domain rich in proline residues and a C-terminal domain containing six cysteines that form three disulphide bridges. These features are usually associated with two distinct groups of AMPs found in insects, such as the proline rich peptides, active against Gram-negative bacteria, and defensins active against Gram-positive bacteria. Besides this chimera-like overall structure, penaeidins undergo post-translational modifications, such as C-terminal amidation (Destoumieux et al., 1997) which has also been observed in other marine invertebrate antimicrobial peptides such as the tachyplesins from *Tachypleus tridentatus* (Nakamura et al., 1988) and in insect cecropins (Li et al., 1988) where it is shown to be functionally important by increasing antimicrobial activity compared with the same peptides, which have a free carboxyl group.

Penaeidins are highly variable with respect to genetic composition which is clearly exhibited by the existence of different isoforms in each sub group. Based on the amino acid sequence and the position of some precise residues in both the N- and C-terminal regions, penaeidins are classified

into four distinct subgroups: Penaeidin 2 (Pen2), Penaeidin 3 (Pen3), Penaeidin 4 (Pen4) and Penaeidin 5 (Pen5) which have been identified in *L. vannamei* (Destoumieux et al., 1999; Cuthbertson et al., 2002), *F. chinensis* (Cuthbertson et al., 2004; Kang et al., 2007), *L. setiferus* (Cuthbertson et al., 2004) and *P. monodon* (Chen et al., 2004b).

All penaeidin precursors comprise a highly conserved signal peptide followed by a cationic mature peptide (5.48 - 6.62 kDa) with a calculated isoelectric point above 9 (Bachère et al., 2004). After cleavage of signal peptide, mature peptides can be post-translationally processed by the formation of a pyroglutamic acid in the N-terminus and/or by a C-terminal amidation involving the elimination of a glycine residue (Destoumieux et al., 1997, 2000a). The amino acid sequences deduced from the cDNA revealed that the penaeidins isolated from haemocytes are synthesized as precursor molecules consisting of a signal peptide (19–21 amino acids) immediately preceding the bioactive molecule (Destoumieux et al., 1997). The role of post-translational modifications observed in native penaeidin, like C-terminal amidation and N-terminal pyroglutamic acid has been studied with a set of penaeidin variants. The results showed that these modifications had little effect on penaeidin AMPs (Destoumieux et al., 1999), but they increased the stability of penaeidin, to proteolysis (Destoumieux et al., 1997). The family appears to be characterized by (i) a highly conserved signal peptide (ii) an N-terminal proline-rich domain with the following signature (Y,F)T(R,G)P(X)2(R,K)P, and (iii) a C-terminal cysteine-rich domain with the following signature C(X)2-3C(X)7RXC C(X)5CC.

One of the most striking characteristics of the penaeidins is the presence of a 5' UTR and leader sequence that is nearly identical within and

between species. This contrasts with the proline rich domain (PRD) and cysteine rich domain (CRD) which are only about 40 % similar across penaeidin classes regardless of species (O'Leary and Gross, 2006). Adding to the structural complexity of the penaeidin family, each of these classes display a significant degree of isoform diversity (Bachere et al., 2000a; Gueguen et al., 2006). Multiple isoforms, generated by substitutions and deletions within the proline and cysteine-rich domains, have been reported at the mRNA levels for all the classes of penaeidins suggesting that this is a highly diverse gene family. O'Leary and Gross (2006) examined the genomic sources for the penaeidins and has proved that penaeidin class is encoded by a unique gene and that isoform diversity is generated by polymorphism within each penaeidin gene locus.

Most groups working on shrimps from different parts of the world have now isolated numerous penaeidin sequences mostly by genomic approaches, and this family appears to be ubiquitous among penaeid shrimp species. Currently more than 200 sequences can be found in the EMB/GenBank/DDBJ databases, some of which are yet to be published. Expressed sequence tag (EST) analysis from haemocyte cDNA libraries has shown that penaeidin transcripts are abundant. In fact, penaeidins appear to represent 10.7 and 20 % of all the sequences isolated from haemocytes of *L. vannamei* and *L. setiferus*, respectively (Gross et al., 2001). Between the penaeidin subgroups, PEN3 is the most abundant at both levels of peptide and mRNA in *L. vannamei* haemocytes (Destoumieux et al., 1997, 2000b). It represents more than 90 % of all the penaeidin mRNA sequences detected in both shrimp species (Cuthbertson et al., 2002). Gueguen et al. (2006) have developed a penaeidin database, named as 'PenBase' (http://www.penbase.immunaqua.com) to provide

comprehensive information about penaeidin properties, diversity and nomenclature.

In shrimps, haemocytes are the main source of penaeidin production (Destoumieux et al., 2000a, 2000b; Munoz et al., 2002, 2003, 2004; Kang et al., 2004). Penaeidins are constitutively expressed in their mature and active form in granular haemocytes of naive shrimps and were stored within cytoplasmic granules of granular haemocyte populations (Munoz et al., 2002, 2003). The penaeidins could be secreted or released from haemocytes by degranulation into the blood upon immune response stimulation (Bachere et al., 2000). About 30–40 % of the circulating haemocytes express penaeidins (Meister et al., 1997). The population of hyaline cells is devoid of penaeidins. Distribution of penaeidin transcripts and proteins is restricted to haemocytes that are present strikingly in almost all the shrimp tissues, both circulating in blood vessels irrigating the tissues or infiltrating tissues such as the brain, subcuticular epithelia, midgut cecum, or muscle. In most of the cases, the various penaeidin subgroups have been shown to be expressed in one single individual (Gross et al., 2001), suggesting that the various penaeidins may have different biological functions in shrimp.

Penaeidins are the well-characterized AMP family at the level of gene expression and biological activities. The spectrum of the antimicrobial activity of penaeidins has been studied in detail using native peptides purified from shrimp haemolymph, synthetic peptides and recombinant variants produced in heterologous expression systems (Destoumieux et al., 1997, 1999; Cuthbertsonet al., 2004, 2005, 2006; Li et al., 2005; Kang et al., 2007). Penaeidin anti-bacterial activity is predominantly directed against Gram-positive bacteria via a strain-specific inhibition mechanism

and through multiple modes of action. Penaeidins have been shown to be particularly effective against Gram-positive bacteria such as *Micrococcus, Bacillus, Staphylococcus* (MIC of about 0.3-2.5 µM). But penaeidins display very weak anti-bacterial activity *in vitro* against Gram-negative strains including *Vibrio* species. However Fenchi PEN5, appears to be active against Gram-negative *Klebsiella pneumonia*.

Penaeidins combine two domains in their overall structure; one proline-rich and the other cysteine-rich, usually observed in distinct groups of antimicrobial peptides (Destoumieux et al., 2000a). The overall biological activity of penaeidin may be associated with distinct properties of their two characteristic regions. The penaeidin N-terminal proline-rich region shares sequence similarities with other proline-rich AMPs (Gennaro et al., 1989; Schnapp et al., 1996). A synthetic peptide corresponding to the N-terminal proline-rich domain of penaeidin (residues 1–20) was produced and found to be inactive against both bacterial and fungal strains. Hence, it is assumed that the penaeidin proline rich region is similarly involved more in target cell interaction than in direct antimicrobial activity. This is also consistent with the spectrum of penaeidin antimicrobial activity, which is directed mainly against fungi and Gram-positive bacteria (Destoumieux et al., 1999), in contrast to most of the known proline-rich peptides, which have essentially anti-Gram-negative properties (Bulet et al., 1999).

The penaeidin C-terminal domain was shown to confer to the whole molecule an ability to attach chitin tightly (Destoumieux et al., 2000b). The C-terminal domain of penaeidins presents some similarities and partial conservation of a primary sequence motif common to several chitin-binding proteins isolated from plants (Raikhel et al., 1993). Chitin-

binding ability is most often related to an anti-fungal activity (Cuthbertson et al., 2006). Penaeidins possess broad-spectrum fungicidal activity against filamentous fungi such as *Fusarium, Nectria, Alternaria, Neurospora, Botritys* and *Penicillium* (MIC of about 1.25-2.5 µM). However, they are found to be inactive against yeast such as *Saccharomyces cerevisiae* or *Candida albicans* (Destoumieux et al., 1999). Interestingly, penaeidins are active against the shrimp pathogen *Fusarium oxysporum*, which is responsible for infections in different penaeid shrimps (Rhoobunjongde et al., 1991). Many phytopathogenic fungal strains such as *Nectria haematococca*, *Alternaria brassicola*, *Neurospora crassa* and *Botritys cinerea* were shown to be sensitive to the peptides, indicating that penaeidins could have applications in agronomy as therapeutic agents. Penaeidin fungicidal activity against the different strains was shown to be associated with inhibition of spore germination. At lower concentrations, the peptides cause growth inhibition of fungal hyphae, resulting in morphological abnormalities.

Besides their antimicrobial properties, penaeidins are also able to bind to chitin. A conserved chitin-binding motif is recognized in the cysteine-rich domain, whereas the PRP-rich domain is preferentially involved with the antimicrobial activities (Destoumieux et al., 2000a; Cuthbertson et al., 2004). According to Destoumieux et al. (2000b), the chitin-binding domain of penaeidins could have supplementary functions, such as the capability to bind to shrimp carapace upon injury and thus have a contribution in wound healing and/or molting processes. Therefore, it is suggested that they can be involved in chitin assembly or wound healing and may be essential in shrimp protection during the moulting cycles (Destoumieux et al., 2000b).

One can assume that the peptides, bound to the shrimp cuticle tissues through chitin-binding property carried by the C-terminal domain, exert antimicrobial activity and initiate opsono-phagocytosis via the free N-terminal end. This domain can adopt a conformation upon interaction with the bacterial membrane, which displays antimicrobial activity. From all hypotheses, the two domains are likely to have complementary activities, and their presence in one single peptide would be necessary for its full activity. Studies are being undertaken to address this question. Penaeidins may participate in protein–protein interactions and thus can display various functions. Chitin binding ability of penaeidin could participate in antimicrobial activity and wound healing and chitin assembly. The peptides may play a role in the protection of the animals during molting cycle, when the animals are particularly exposed to potential infections. This dual function of penaeidins is likely determinant for the survival of the animals. To address this question, it is now under investigation whether the penaeidins conserve their antimicrobial activity when bound to the shrimp cuticle (Bachere et al., 2004).

Post-translational modifications were found to have little effect on antimicrobial properties of penaeidin (Destoumieux et al., 1999), but possibly they increase the stability of penaeidins, which are highly resistant to proteolysis (Destoumieux et al., 1997). N-terminal blocking has no effect on penaeidin antimicrobial properties (Destoumieux et al., 2000a). Moreover, when C-terminal amidation is replaced by an extra glycine residue, penaeidin antifungal properties are unaltered, and antibacterial activity is decreased only two-fold. The partial loss of antibacterial activity that occurs in non-amidated penaeidins may be due

to the loss of a positive charge at the C-terminus, and consequently to less efficient interaction of the peptide with bacterial membranes.

Penaeidins, are the most intensively studied AMP in terms of biological properties, antimicrobial activities, gene expression and localization in response to infection (Bachere et al., 2004). In fact, increased resistance of bacteria towards antibiotic drugs has stimulated intensive effort for discovery and characterization of AMPs as sources or templates for the design of new therapeutic antibiotics. Most of the cultured penaeid shrimps are prone to infection by microbes at various stages of its growth; the losses due to disease can be enormous. As a result, a good amount of research has been focused in isolation and characterization of proteins playing a role in immune regulation. With strong and broad spectrum antimicrobial activity against Gram-positive bacteria and fungi, the penaeidins are becoming potential therapeutic agents for prophylactic treatment of fungal and bacterial infectious diseases.

Hence, the present study was focused on the discovery of new penaeidin isoforms from shrimps through molecular approach and to investigate their characteristics and phylogeny so as to understand the role of penaeidins in the innate immune defense of crustaceans.

4.2 Materials and Methods

4.2.1 Experimental animals

The experimental organisms used in the present study include the Indian White Shrimp, *Fenneropenaeus indicus* and the Pink Shrimp, *Metapenaeus monoceros* belonging to decapod crustaceans of the family, Penaeidae. Live specimens of *F. indicus* and *M. monoceros* were collected from Cochin

barmouth region, in Vypeen, Kochi and were brought to the laboratory (Fig. 4.1 a and b).

4.2.2 RNase control

Basic precautions were taken when working with RNA as described in section 2.2.2 of Chapter 2.

4.2.3 Haemolymph collection

Haemolymph was collected from the rostral sinus of the experimental organisms and was suspended in TRI reagent for total RNA isolation as given in section 2.2.3 of Chapter 2.

4.2.4 RNA isolation

Total RNA was extracted from the haemocytes of the experimental organisms using TRI™ reagent (Sigma) following manufacturer's protocol as described in section 2.2.4 of Chapter 2. RNA pellets were dried for 5-10 min and dissolved in RNase free water by repeated pipetting with a micropipette at 55-60 ^{0}C for 10-15 min.

4.2.5 Determination of the quantity and quality of RNA

RNA was quantified and qualified by spectrophotometry and agarose gel electrophoresis as given in section 2.2.5 of Chapter 2.

4.2.6 cDNA synthesis

cDNA was synthesized using oligo d(T_{20}) primers as described previously in section 2.2.6 of Chapter 2.

4.2.7 PCR amplification

PCR amplification of cDNA were carried out using gene specific primers, designed using GeneTool software based on consensus sequences of

penaeidins retrieved from GenBank. β-actin was used as the internal control to verify reverse transcription reaction. The sequences of the primers used to amplify β-actin and penaeidin genes are given in Table. 4.1. PCR amplification of 1µl cDNA was performed in a 25 µl reaction volume using gene specific primers as explained in section 2.2.7 of Chapter 2. Annealing temperature varied for the different genes as given in Table 4.1. The PCR products were visualized by electrophoresis in 1.5 % agarose gel.

4.2.8 Agarose gel electrophoresis

Electrophoresis was performed in 1.5 % agarose gel and documented as described in section 2.2.8 of Chapter 2.

42.9 Cloning of the PCR product

The PCR products were cloned onto the pGEM-T Easy vector as explained in section 2.2.9 of Chapter 2. *E. coli* DH5α competent cells were used for transforming the ligated cloning vector. To confirm the presence of the insert DNA in the transformed cells, colony PCR was performed for all the individually streaked colonies using vector specific primers viz. T7 and SP6 as well as gene specific primers (Table 4.1). Plasmid with the penaeidin insert was then extracted and purified using GenElute HP' plasmid MiniPrep kit (Sigma)and was sent for sequencing at SciGenom, Kochi, India.

4.2.10 Sequencing and sequence analysis

Bioinformatics analysis of the sequences were performed using various softwares / programs as described in section 2.2.10 of Chapter 2.

4.3 Results

In the present study two penaeidin isoforms could be identified, cloned and characterized, each from *F. indicus* and *M. monoceros* respectively (Table 4.2). As these isoforms were found to be similar both have been analysed and discussed together in the following sections. The new penaeidin isoform was found to belong to subgroup-3, and is herein after referred to as FiPEN and MmPEN respectively.

4.3.1 Molecular characterization of penaeidin in *F. indicus* and *M. monoceros*

4.3.1.1 FiPEN and MmPEN

4.3.1.1.1 Cloning and sequencing of FiPEN and MmPEN cDNA

A 338 bp fragment cDNA with an open reading frame (ORF) of 216 bp was obtained from the mRNA of *F. indicus* and *M. monoceros* haemocytes by RT-PCR. This fragment could be successfully cloned and sequenced in the present study. The ORF encoded 71 amino acid residues and consisted of a signal peptide region followed by a proline-rich domain (PRD) and a cysteine-rich domain (CRD), characteristic of the penaeidins (Fig. 4.2). The penaeidins had its characteristic signal peptide region possessing 19 amino acids followed by a cleavage site, then the mature peptide region comprising of a proline-rich domain with 24 amino acids at the N-terminal region as well as a cysteine-rich domain with 28 amino acids at the C- terminal region (Fig. 4.3). These features designated this penaeidin isoform to penaeidin-3 subgroup. The nucleotide sequence and deduced amino acid sequence were submitted to NCBI GenBank under the accession numbers, **JX657680** (FiPEN) and **KF275674** (MmPEN).

4.3.1.1.2 Analysis of sequence and peptide characteristics of FiPEN and MmPEN

The mature peptide region of FiPEN and MmPEN possessed a predicted molecular weight of 5.66 kDa and an isoelectric point, p*I* of 9.38, as predicted by the PROTPARAM software. Both FiPEN and MmPEN are proved to be highly cationic peptides with a charge of +6 and they possess a hydrophobicity of 42 %. The total number of negatively charged residues (Asp + Glu) was one, while the total number of positively charged residues (Arg + Lys) was found to be seven. The estimated half-life was predicted to be 2.8 h (in mammalian reticulocytes, *in vitro*), 10 min (in yeast, *in vivo*) and 2 min (in *E.coli*, *in vivo*) respectively. The aliphatic index and the grand average of hydropathicity (GRAVY) were found to be 45 and -0.206, respectively. The deduced amino acid sequence was found to be rich in amino acid residues cysteine and serine (11.5 %) followed by arginine, glycine and threonine (9.6 %) as calculated by the PROTPARAM tool of ExPASy program. The new penaeidin isoforms were analyzed for their antimicrobial activity using Antimicrobial Peptide Predictor Program which revealed that they possess a protein binding potential of 1.72 kcal/mol and the total net charge was calculated to be +6. The physico-chemical properties of FiPEN and MmPEN are given in Table 4.3.

Helical wheel modeling was also performed using helical wheel projection program, which revealed clustering of hydrophobic and hydrophilic residues of the peptide (Fig. 4.4). Structural analysis revealed the presence of an α-helix in the cysteine-rich domain which was stabilized by disulfide bonds (Fig. 4.5). The proline-rich domain was found to form an extended structure.

4.3.1.1.3 Sequence alignment and phylogenetic analysis of FiPEN and MmPEN

The new penaeidin isoform shared a maximum identity of 63 % with a penaeidin-3 isoform of *P. monodon* and penaeidin-2 of *Farfantepenaeus paulensis*. This was followed by identity with penaeidin-3 of *L. schmitti* (62 %) and penaeidin-5 of *F. chinensis* (61 %) (Table 4.4).

Multiple alignment performed with other known penaeidins revealed conserved regions within the peptide (Fig. 4.6). Like other penaeidins, the signal peptide region of this new penaeidin isoform was found to be highly conserved (MRLVVCLVFLASFALVCQG). The mature peptide showed the presence of a proline-rich domain at N-terminal region and a cysteine-rich domain at C-terminal region which is the characteristic feature of penaeidins. FiPEN and MmPEN were also characterized by highly conserved amino acid sequences in the mature peptide viz. a threonine and two proline residues conserved in the N-terminal domain; and the conserved cysteine array of the C-terminal domain. The sequence showed six highly conserved cysteine residues (Cys_{25}, Cys_{28}, Cys_{39}, Cys_{40}, Cys_{46}, Cys_{47}) which are engaged in the formation of three disulphide bridges. DiANNA 1.1 web server predicted the formation of three disulphide bridges between C_{25}-C_{47}, C_{28}-C_{39} and C_{40}-C_{46}.

Phylogenetic tree (Fig. 4.7) constructed to study the evolutional relationships of FiPEN and MmPEN against other penaeidins revealed that it is related to penaeidins of *F. paulensis, F. brasiliensis, P. monodon* and *F. chinensis*. The analysis further revealed that the penaeidin sequences were clustered not according to species, but also according to the subgroup. The tree could be broadly divided into two major groups.

Group 1 consisted of penaeidins from *Litopenaeus* sp., *Fenneropenaeus* sp., *Farfantepenaeus* sp. and *Penaeus* sp. Group 1 could again be divided into four subgroups: subgroup 1 consisting of penaeidin-3 of *Litopenaeus* sp. and *Fenneropenaeus* sp.; subgroup 2 consisting of penaeidin-4 of *Litopenaeus* sp.; subgroup 3 consisting of penaeidin-2, 3 and 5 of *Litopenaeus* sp., *Fenneropenaeus* sp., and *Penaeus* sp.; in which the penaeidin-2,3 and 5 formed clear distinct branches; and subgroup 4 consisting of penaeidin-3 of *L. vannamei* alone.

Group 2 could be divided into four subgroups. Subgroup 1 consisted of penaeidin-2 of *F. paulensis*; The FiPEN and MmPEN formed the subgroup 2, which was the only member of that subgroup; subgroup 3 consisted of penaeidins from *Farfantepenaeus* sp. and *Fenneropenaeus* sp.; and subgroup 4 consisted of penaeidins of *Penaeus* sp. and *Fenneropenaeus* sp. The tree topologies revealed that all penaeidins possess the same ancestral origin and have a similar evolutionary status, and that they were phylogenetically ancient immune effector molecules which may play an essential role in the host defense mechanism.

4.4 Discussion

From the present study a new isoform of penaeidin, belonging to subgroup-3, could be identified from the haemocytes of Indian white shrimp *F. indicus* (FiPEN) and also from the pink shrimp, *M. monoceros* (MmPEN). FiPEN was the first isoform of penaeidin-3 to be reported from *F. indicus*. At the same time MmPEN is the first AMP to be reported from *M. monoceros*. An ORF possessing 216 bp encoding 71 amino acid could be obtained from the hameocytes of penaeid shrimps in the present study by RT-PCR. Both isoforms were found to be 100 % similar (Fig. 4.2). The

sequence possessed a signal peptide region followed by a proline-rich domain (PRD) and a cysteine-rich domain (CRD), characteristic of the penaeidins (Fig. 4.3). The physico-chemical parameters including molecular weight (5.66 kDa), *p*I (9.38) were in agreement with the general characteristic features of penaeidins. This penaeidin isoform was found to be highly cationic and possessed hydrophobicity of 42 %. The deduced amino acid sequence was found to be rich in amino acid residues cysteine and serine (11.5 %) followed by arginine, glycine and threonine (9.6 %) as calculated by the PROTPARAM tool of ExPASy program. The new penaeidin isoform was analyzed for its antimicrobial activity using Antimicrobial Peptide Predictor program which revealed that it possess a protein binding potential of 1.72 kcal/mol and the total net charge was calculated to be +6. All the above features are in agreement with the required features of penaeidins (Destoumieux et al., 1999).

The analysis with SignalP software revealed the presence of a signal peptide with 19 amino acid residues at the N-terminus, with a predicted cleavage site between positions 19 and 20, i.e. CQG-YK (Fig. 4.3). The signal peptide was followed by a proline-rich domain (PRD) consisting of 24 amino acid residues at the N-terminal region and a C-terminal cysteine-rich domain (CRD) with six cysteine residues and consisting of 28 amino acid residues. Generally the N-terminal proline-rich domain is found to be longer than the cysteine-rich domain; cysteine-rich domain being stabilized by three disulfide bonds and found to be more conserved across classes (Yang et al., 2003; Gueguen et al., 2006). However, in case of the newly identified FiPEN and MmPEN, cysteine-rich domain (28 amino acid residues) was found to be longer than the proline-rich domain (24 amino acid residues).

Helical wheel modeling was also performed using helical wheel projection program, which revealed clustering of hydrophobic and hydrophilic residues of the peptide (Fig. 4.4). Structural analysis revealed the presence of an α-helix in the cysteine-rich domain which was stabilized by disulfide bonds (Fig. 4.5). However, the proline-rich domain formed an extended structure. This is in agreement with the solution structure described by Yang et al. (2003) and Cuthbertson et al. (2002).

Sequence comparison using BLAST algorithm showed that the deduced amino acid sequence of FiPEN and MmPEN shared identity with other penaeidins. The new penaeidin isoform shared a maximum identity of 63 % with a penaeidin-3 isoform of *P. monodon* and penaeidin-2 of *F. paulensis*, which proves it to be a new isoform. This was followed by identity with penaeidin-3 of *L. schmitti* (62 %) and penaeidin-5 of *F. chinensis* (61 %). FiPEN was found to be a different isoform from the already reported penaeidins of *F. indicus* viz. Fi-Penaeidin (Antony et al., 2011; Vaseeharan et al., 2012) and MmPEN is the first AMP to be reported from *M. monoceros*.

Multiple alignment performed with other known penaeidins revealed conserved regions within the peptide (Fig. 4.6). Like other penaeidins, the signal peptide region of this new penaeidin isoform was found to be highly conserved (MRLVVCLVFLASFALVCQG). The mature peptide showed the presence of a proline-rich domain at N-Terminal region and a cysteine-rich domain at C-terminal region which is the characteristic feature of penaeidins. FiPEN and MmPEN were also characterized by highly conserved amino acid sequences in the mature peptide viz. a threonine and two proline residues conserved in the N-terminal domain; and the conserved cysteine array of the C-terminal domain which is in

agreement with the penaeidin signature assigned by Gueguen et al. (2006). As per Destoumieux et al. (1997) and Gueguen et al. (2006) this overall structure of penaeidins is quite unique among the AMP families. The sequence showed six highly conserved cysteine residues (Cys_{25}, Cys_{28}, Cys_{39}, Cys_{40}, Cys_{46}, Cys_{47}) which are engaged in the formation of three disulphide bridges. DiANNA 1.1 web server predicted the formation of three disulphide bridges between C_{25}-C_{47}, C_{28}-C_{39} and C_{40}-C_{46}.

Phylogenetic tree (Fig. 4.7) to study the evolutional relationships of FiPEN and MmPEN against other penaeidins revealed that it is related to penaeidins of *F. paulensis, F. brasiliensis, P. monodon* and *F. chinensis*. The analysis further revealed that the penaeidin sequences were clustered not according to species, but according to the subgroup. The tree could be broadly divided into two major groups. Group 1 consisted of penaeidins from *Litopenaeus* sp., *Fenneropenaeus* sp., *Farfantepenaeus* sp. and *Penaeus* sp. Group 1 could again be divided into four subgroups: subgroup 1 consisting of penaeidin-3 of *Litopenaeus* sp. and *Fenneropenaeus* sp.; subgroup 2 consisting of penaeidin-4 of *Litopenaeus* sp.; subgroup 3 consisting of penaeidin-2, -3 and -5 of *Litopenaeus* sp., *Fenneropenaeus* sp., and *Penaeus* sp.; in which the penaeidin-2, -3 and -5 formed clear distinct branches; and subgroup 4 consisting of penaeidin-3 of *L. vannamei* alone.

However, penaeidins of Group 2 was found to form a diverse branch and hence might be distantly related to group 1, as evident from the phylogenetic tree. Group 2 could be divided into four subgroups. Subgroup 1 consisted of penaeidin-2 of *F. paulensis*; The FiPEN and MmPEN formed the subgroup 2, which was the only member of that subgroup; subgroup 3 consisted of penaeidins from *Farfantepenaeus* sp. And *Fenneropenaeus* sp.; and subgroup 4 consisted of penaeidins of

Penaeus sp. and *Fenneropenaeus* sp. The tree topologies revealed that all penaeidins possess the same ancestral origin and have a similar evolutionary status, and that they were phylogenetically ancient immune effector molecules which may play an essential role in the host defense mechanism.

From the tree topology and BLAST algorithm, it would be difficult to designate the subgroup of FiPEN and MmPEN to which they belong to. However, the characteristic features of various penaeidin subgroups as evident from the Penbase database revealed that FiPEN and MmPEN belonged to penaeidin-3 subgroup. According to Penbase, penaeidins possessing a signal peptide of 19 amino acid residues, a proline-rich domain of 21-31 amino acid residues and a cysteine-rich domain of 25-35 amino acid residues could be designated to subgroup 3. Analysis of the amino acid sequence of FiPEN and MmPEN revealed the presence of 19 amino acid residues in the signal peptide region; 24 amino acid residues in the proline-rich region and 28 amino acid residues in the cysteine –rich region; fulfilling all the conditions for designating it to subgroup 3 of penaeidins.

All penaeidins characterized so far are generally inactive against Gram-negative bacteria. However, penaeidins do exhibit activity against Gram-positive bacteria and fungi, which may or may not be specific. Penaeidins are also believed to possess chitin binding property because of its characteristic cysteine-rich domain. Microbial target specificity of penaeidins are found to be related to structural characteristics that cannot be deduced directly from primary amino acid sequence comparisons between penaeidin isoforms (Cuthbertson et al., 2002). A detailed research is binding to reveal the possible antimicrobial

properties of the newly identified penaeidin isoform. Possibly many other isoforms exist in penaeid shrimps which belong to various subgroups and which might vary in their biological properties, and are yet to be discovered. Overall, the presence of different penaeidin subgroups and isoforms within the various penaeid species indicate that shrimp penaeidin make up a large and diverse family. The wide distribution of penaeidins in penaeid shrimps indicates the importance of these AMPs in the innate immunity. Discovery of novel penaeidins might pave way to the discovery of promising therapeutic and/or prophylactic agents in health management and disease control in shrimp aquaculture.

(a) **(b)**

Fig.4.1 Experimental organism used for the study: (a) *Fenneropenaeus indicus* (b) *Metapenaeus monoceros*

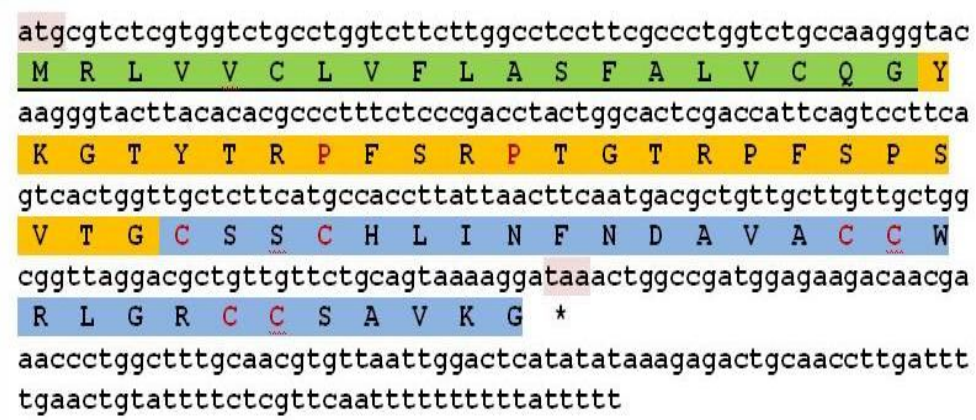


Fig. 4.2 Nucleic acid (above) and deduced amino acid (below) sequence of FiPEN and MmPEN. The signal peptide sequence (green), proline-rich domain (yellow) and cysteine-rich domain (blue) have been highlighted.

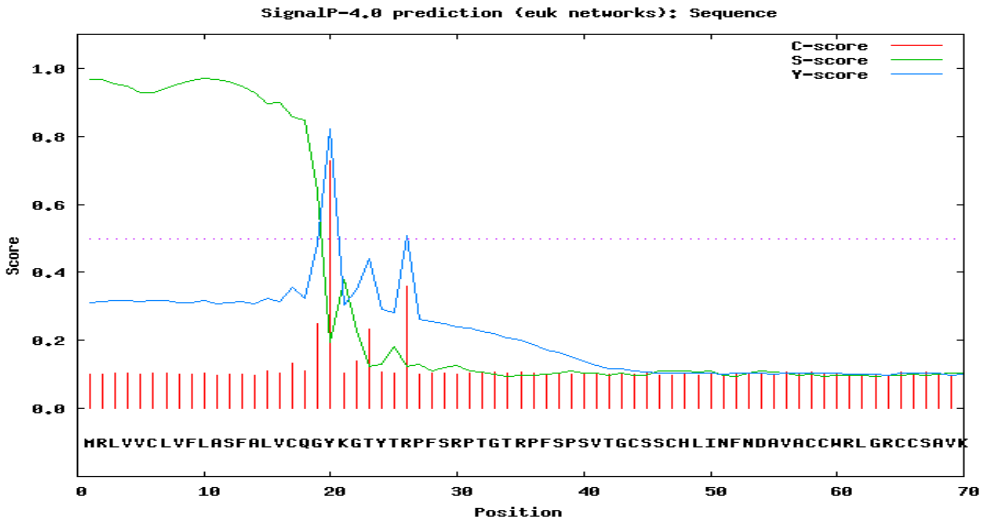


Fig.4.3 Signal peptide analysis of FiPEN and MmPEN as predicted by the SignalP 3.0 server.

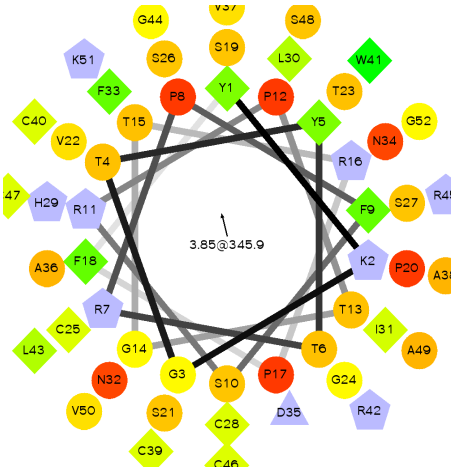


Fig.4.4 Schiffer-Edmundson Helical wheel representation of FiPENand MmPEN generated using DNASTAR software. Hydrophobic and hydrophilic amino acids are seen occupying opposite positions.

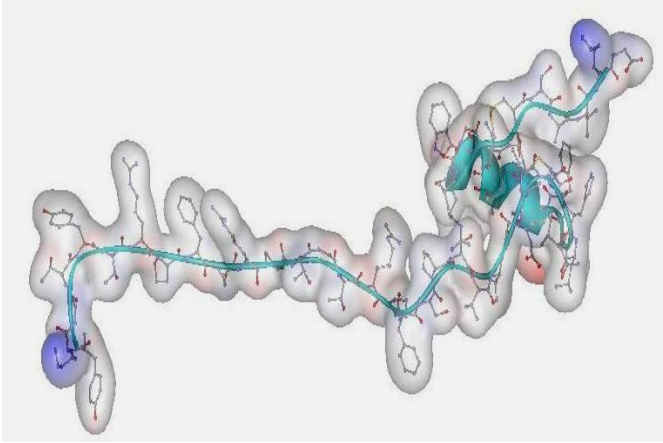


Fig.4.5 Secondary structure of FiPEN and MmPEN created using PyMOL showing α-helix and extended structures

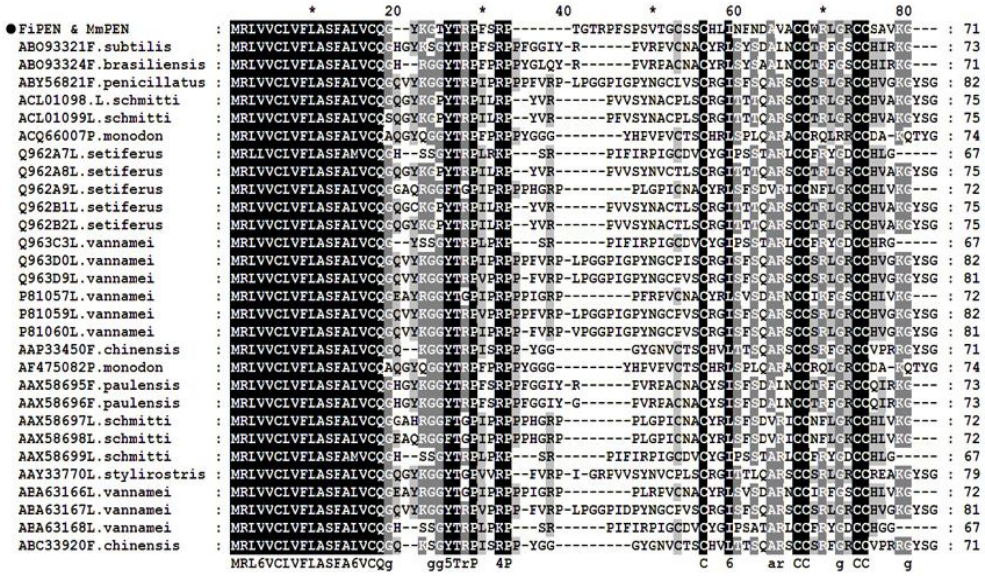


Fig.4.6 Multiple alignment of nucleotide sequence of FiPEN and MmPEN with other penaeidins obtained using GeneDoc program. Black and grey indicates conserved sequences.

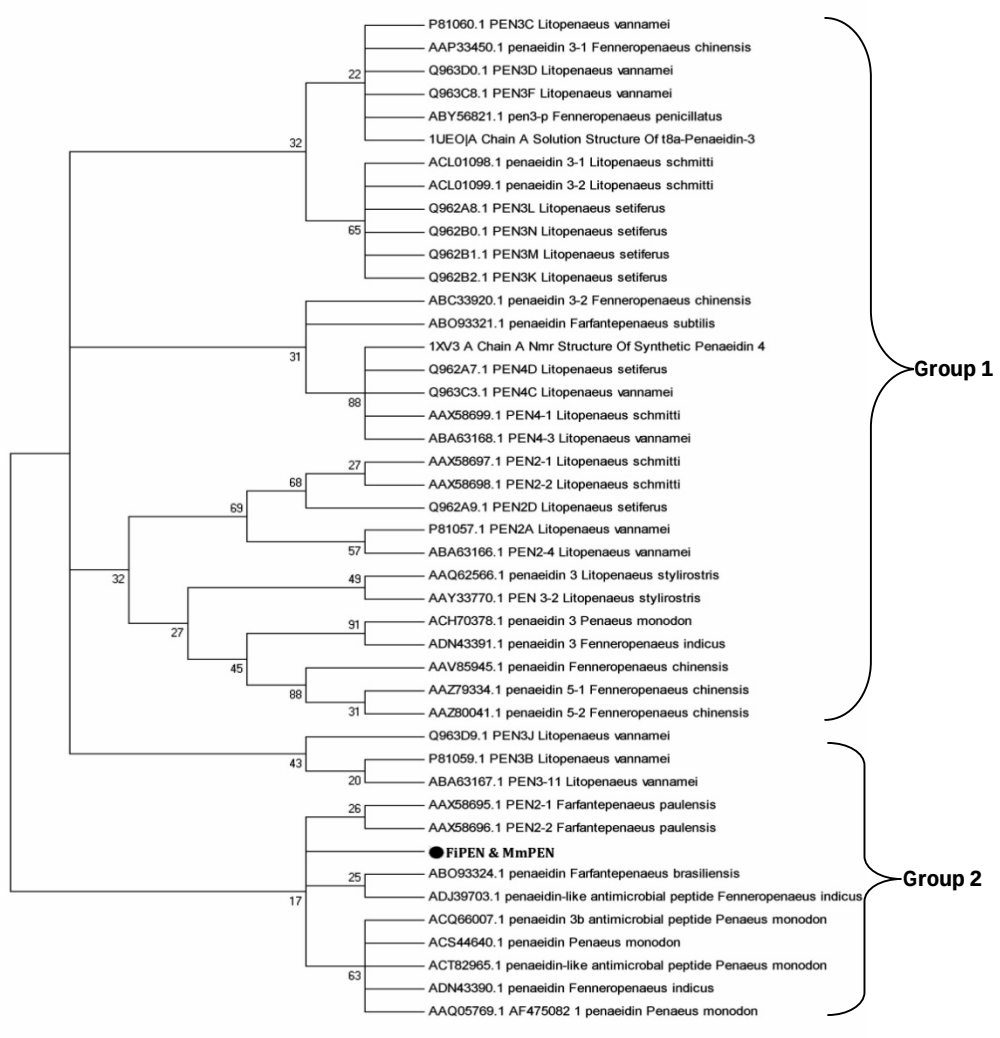


Fig.4.7 A bootstrapped neighbour-joining tree obtained using MEGA version 5.0 illustrating relationships between the deduced amino acid sequence of FiPEN and MmPEN with other penaeidins. Values at the node indicate the percentage of times that the particular node occurred in 1000 trees generated by bootstrapping the original deduced protein sequences.

Table 4.1 Primers used for the study

Primer	Primer Sequence (5'-3')	Product Size (bp)	Annealing Temp. (°C)	$MgCl_2$ Conc. (mM)
Penaeidin	F:acctgaccctcacctgcagaggcc R: acctacatcctttccacaag	338	65	1.5
β-actin	F:cttgtggttgacaatggctccg R: tggtgaaggagtagccacgctc	520	60	1.5
T7	tgtaatacgactcactataggg	--	57	1.5
SP6	gatttaggtgacactatag	--	57	1.5

Table 4.2 List of penaeidins identified

No:	Name of AMP	Source Organism	ORF	GenBank Accession
1	FiPEN	*Fenneropenaeus indicus*	Complete	**JX657680**
2	MmPEN	*Metapenaeus monocerus*	Complete	**KF275674**

Table 4.3 Physicochemical properties of FiPEN and MmPEN

Parameters Studied	Results
Open reading frame (GeneTool)	71 aa
Length of mature peptide (No. of amino acids) (GeneTool)	52 aa
Signal peptide (SignalP)	1-19 aa
Molecular weight (PROTPARAM, ExPASy)	5.66 kDa
Isoelectric point (PROTPARAM, ExPASy)	9.38
Charge (PROTPARAM, ExPASy)	+8
Hydrophobicity (PROTPARAM, ExPASy)	42 %
Structure (PyMol and DNASTAR)	α-helix

Table 4.4 Result of BLASTp analysis of FiPEN and MmPEN

GenBank Accession No.	Description of the AMP	Query coverage	Identity to FiPEN and MmPEN
AFN37209	Penaeidin 3 [*Fenneropenaeus indicus*]	100 %	76 %
ACH70378	Penaeidin 3 [*Penaeus monodon*]	100 %	63 %
AAX58696	Antimicrobial peptide PEN2 [*Farfantepenaeus paulensis*]	100 %	63 %
AAV85945	Penaeidin [*Fenneropenaeus chinensis*]	100 %	61 %
ABO93324	Penaeidin [*Farfantepenaeus brasiliensis*]	100 %	54 %

5

Recombinant Expression of Anti-lipopolysaccharide factor and its Functional Characterization

5.1 Introduction

AMPs are potent, not easy to generate drug-resistance molecules and therefore possess great potential to be novel therapeutic agents (Hoffmann, 1995; Boman, 1995; Hancock and Chapple, 1999; Zasloff, 2002; Ganz, 2003). In order to better characterize a novel AMP, to better understand its mechanisms of action and structure–function relationships, as well as to develop new antibiotics for widespread clinical use, huge amounts of that compound with high purity, quality as well as activity should be readily available in a cost-effective manner. Due to the small amount of pure peptides directly recovered from its host organism, further studies always depend on a strategy to recover more compound of interest. In general there exists three different approaches that can be employed; 1) direct isolation of the target AMP from its natural host, 2) chemical synthesis of the target AMP and 3) heterologous expression of the target AMP.

AMPs are naturally present in very low concentrations in most organisms and the concentrations are affected by unknown environmental factors and / or pathogenic microbes (Li et al., 2010). Hence, direct recovery from

the host organism is neither economically nor practically feasible. Along with that, the process of extraction is time consuming, environmental unfriendly as well as expensive and therefore not recommended (Li et al., 2008).

Solid phase peptide synthesis is a widely used method to produce antibacterial peptides, but the cost is very high and the synthetic process is usually reported to cause environmental pollution problems (Andersson et al., 2000). Moreover, the production costs are extremely high, especially for those possessing one or more disulfide bonds, which hamper the production of such molecules (Li et al., 2010). For these reasons, chemical synthesis of AMPs is usually considered suitable for producing smaller AMPs with low cysteine content and those lacking post-translational modifications. Hence, recombinant expression of AMPs appears to offer a cost effective method for large-scale production (Rao et al., 2005).

The expression host and vector system are selected on the basis of AMP features such as size, its intracellular localization or secretion, folding as well as glycosylation pattern. Almost 97.4 % of the hosts so far used for heterologous AMP production include bacteria and yeasts (Li et al., 2008). There are many eukaryotic host systems available for AMP production, such as yeasts *Pichia pastoris*, *Saccharomyces cerevisiae* and *Yarrowia lipolytica*; the baculovirus expression system in insect cells; plant and human cells. Although eukaryotic hosts usually improve the production of AMPs, especially that require post-translational modifications, these strategies are time-consuming and expensive. However, the prokaryotic expression systems, especially the bacterium *Escherichia coli* on the other hand is one of the most commonly used hosts for AMP production. Other

than *E. coli*, the Gram-positive *Bacillus subtilis* and other probiotic microorganisms have also been used in AMP expression (Feng et al., 2012; Mandal et al., 2014).

E. coli is the most widely used prokaryotic expression system, as they are easy to culture, grow fast, easily inducible, available commercially, produce high yields of recombinant protein as well as cost-effective (Ingham and Moore, 2007). The most frequently used *E. coli* expression strains include BL21 (DE3), pLysS Origami, Rosetta and Rosettagami (DE3). Strains such as Origami, Origami B and Rosetta-gami™ also enhance the formation of disulfide bonds (Prinz et al., 1997; Aslund et al., 1999). The most commonly used expression vectors for AMP production in bacterial host is the pET series. Presence of fusion tags facilitates detection and purification of the target protein. Some of the widely used fusion tags include S•Tag™, T7•Tag®, GST•Tag™, His•Tag®, HSV•Tag® and Nus•Tag™. Tags such as His•Tag, GST•Tag, S•Tag, and T7•Tag sequences can be used for affinity purification using the corresponding resin and buffer kits. Among these His•Tag sequence is very useful as a fusion partner for purification of proteins expressed as inclusion bodies. These tags can be removed using protease cleavage sites using enzymes such as thrombin, Factor Xa and enterokinase between the 5' tag and the target sequence followed by purification. Expression in *E. coli* has advantages such as fast growth, low cost and high expression. Nevertheless, AMPs are difficult subjects for heterologous expression in microbial hosts due to their intrinsic antibacterial activity (Li et al., 2009; 2014).

Among yeast expression systems, yeasts such as *S. cerevisiae* and *P. pastoris* are largely used. Yeasts have some advantages over bacterial

systems with regard to post-transcriptional and post-translational modifications, such as glycosylation. Yeast expression has the ability to direct the secretion of peptides with disulfide bridge formation, and the expressed proteins can be easily separated from the fermentation broth (Porro et al., 2005). In the case of AMP production, *P. pastoris* is the widely used host compared to *S. cerevisiae.* However, it requires the establishment of a more suitable promoter in order to achieve the desired yield of recombinant AMPs. Also, in case of AMPs, the expression level is usually found to be low in yeast expression systems, as well as unfavourable degradation has been found to occur in case of AMPs contacting multiple basic residues (Porro et al., 2005). Baculovirus expression system provides a rapid and efficient method to generate recombinant baculoviruses (Ciccarone et al., 1997).

Plants such as soy, maize, cotton, canola, and rice have also been used as peptide expression platforms. Genetically modified plants are mainly aimed at crop improvement and also the nutritional quality of plants. Hence, there are no extensive efforts in quantifying the amount of recombinant AMP produced. However, it has got an advantage that transgenic plants can be directly used for microbial control by simply expressing the peptide in the desired crop without need for purifying the peptide (Desai et al., 2010).

ALFs are basic peptides, which have the ability to bind and neutralize LPS (Morita et al., 1985). ALFs have been so far isolated and characterized from Chelicerates such as *Tachypleus* and *Limulus* (Tanaka et al., 1982; Vallespi et al., 2003) and from crustaceans such as shrimps viz. *P. monodon* (Somboonwiwat et al., 2005), *L. vannamei* (de la Vega et al., 2008), *L. schmitti* (Rosa et al., 2008), *F. chinensis* (Liu et al., 2005); prawns

viz. *M. rosenbergii* (Ren et al., 2012); crabs viz. *S. paramamosain* (Imjongjirak et al., 2007), *E. sinensis* (Li et al., 2008), *S. serrata* (Yedery and Reddy, 2009), *P. trituberculatus* (Yue et al., 2010; Liu et al., 2012) and crayfishes viz. *P. clarki* (Sun et al., 2011).

Recombinant expression of ALFs have been carried out in shrimps, *P. monodon* (Somboonwiwat et al., 2005), crayfish *P. clarki* (Sun et al., 2011) and crabs such as *S. serrata* (Yedery and Reddy, 2009), *E. sinensis* (Zhang et al., 2010) and *S. paramamosain* (Liu et al., 2012; Zhu et al., 2014; Sun et al., 2015). Among these, ALFs from *P. monodon* and *S. paramamosain* were expressed in yeast *P. pastoris* (Somboonwiwat et al., 2005; Liu et al., 2012) and the rest of the ALFs in *E. coli* expression systems (Yedery and Reddy, 2009; Zhang et al., 2010; Sun et al., 2011, 2015; Zhu et al., 2014).

Pioneer studies reporting recombinant expression of an ALF isoform from crustacean was by Somboonwiwat and co-workers in 2005. They performed molecular and functional characterization of the most abundant isoform of ALF, found in *P. monodon*, named as ALFPm3 by expressing it heterologously in the yeast *P. pastoris.* ALF was produced in large-scale in a fermenter with a yield of 262 mg/l and was further purified to homogeneity by single chromatography step on expanded-bed streamline SP6XL. The rALFPm3 was further characterized by employing N-terminal sequencing and mass spectrometry. Antimicrobial assays demonstrated that rALFPm3 possess broad spectrum anti-bacterial activities against both Gram-positive and Gram-negative bacteria, associated with bactericidal effect as well as anti-fungal properties against filamentous fungi. The rALFPm3 was found to be highly efficient against various pathogenic *Vibrio* strains of shrimps (Somboonwiwat et al., 2005).

Pioneer works on heterologous expression of crab ALFs was performed by Yedery and Reddy in 2009. They identified and characterized an ALF isoform from *S. serrata* (*Ss*ALF). The recombinant form of SsALF (rSsALF) was expressed with a His-tag, in *E. coli* Rosetta B(DE3) pLacI, using the pTriEx-4 Ek/LIC vector. ALF was purified employing BugBuster Master Mix, Ni-NTA column, dialysis and enterokinase treatment. The purified rSsALF protein was then subjected to functional characterization. rSsALF was found to possess antimicrobial activity against both Gram-positive (*S. aureus* and *S. pyogenes*) and Gram-negative bacteria (*E. coli* and *P. aeruginosa*) tested. MIC was found to be higher in case of Gram-positive bacteria (100-200 μg/ml) than Gram-negative bacteria (25-100 μg/ml). The recombinant protein was also able to neutralize LPS-induced expression on SsALF *in vivo* as demonstrated by real-time PCR. rSsALF was also able to permeabilize artificial phospholipid membranes as demonstrated by calcein enclosed liposome model.

From *E. sinensis*, Zhang et al. (2010) identified, and performed recombinant expression of an ALF isoform (EsALF-2). The recombinant EsALF-2 was expressed with a His-tag, in *E. coli* BL21(DE3)-pLysS, using the pET-32a+ vector. The recombinant EsALF-2 protein (rEsALF-2) was purified by nickel affinity chromatography Mag Extractor His-Tag NPK-700. Purified protein was refolded before subjecting to antimicrobial assay. The recombinant protein of EsALF-2 showed antimicrobial activity against tested microbes such as *L. anguillarum* (75 μg/ml) and *P. pastoris* (18.75 μg/ml).

Recombinant expression of ALF from *P. clarki* (*Pc*ALF1) was done by Sun et al. (2011). Recombinant *Pc*ALF1 was produced in *E. coli* BL21-DE3 cells using pET30a+ vector. PcALF1 was expressed as inclusion bodies and was

further purified employing a combination of high-affinity Ni-IDA resin, dialysis and Amicon ultra-15 centrifugal filter units of 5 kDa cut off before subjecting to functional characterization. 20 µg of the purified recombinant protein of PcALF1 showed antibacterial activity against Gram-positive S. *aureus, B. subtilis, B. cereus, B. thuringiensis* and *B. megaterium* and against Gram-negative *E. coli* and *P. aeruginosa*. The rPcALF1 also exhibited clearance activity on *V. anguillarum* in a dose-dependent manner *in vivo.* It also showed activity against *Candida albicans*. However, the MICs for various tested microbes were not calculated in this experiment.

Recombinant expression of ALF from the mud crab, *S. paramamosain* (Sp-ALF1 and Sp-ALF2) was performed by Liu et al. (2012). *P. pastoris* was used as the expression host and pPIC9K, the vector. Demonstration of antibacterial activity revealed that recombinant SpALF possessed activity against both Gram-positive (MIC ranging from 1-12 µM) and Gram-negative (1 to >25µM) bacteria tested. Sp-ALF also demonstrated strong anti-WSSV activity.

Liu et al. (2012) performed recombinant expression of a new isoform of ALF from the swimming crab *P. trituberculatus* in *E. coli* BL21(DE3)-pLysS and pET32a+ vector. Analysis of the results revealed that recombinant PtALF possessed antimicrobial activity against Gram-negative bacteria *V. alginolyticus* and *P. aeruginosa*. PtALF also possessed minor activity against *E. tarda*. MIC values of antibacterial activity were found to range between 3.89 – 31.08 µM. However, it did not inhibit the growth of tested Gram-positive bacteria such as *M. luteus* and *S. aureus* as well as fungus *P. pastoris*.

Another report on recombinant expression of ALFs of *S. paramamosain* was by Zhu et al. (2014). Recombinant expression was attained by cloning to pGEX4T1 expression vector and the expression host used was *E. coli* Rosetta (DE3) cells. Cultures were induced with 0.1 mM IPTG and the bacterial cells were subjected to sonication lysis. Glutathoine Sepharose 4B chromatography was employed to purify the recombinant SpALF4 protein and was subjected to antimicrobial assays. Results showed that rSpALF4 could inhibit the growth of Gram-negative bacteria such as *V. harveyi*, *V. anguillarum*, *V. alginolyticus*, *A. hydrophila*, *P. putida;* Gram-positive bacteria such as *S. aureus* and *B. megaterium* as well as a fungus *C. albicans* to varying degrees. Bacterial binding disclosed that it could also bind to all the aforementioned microorganisms except *S. aureus.*

Rosa et al. (2013) studied the functional divergence of various shrimp ALFs. ALF from *L. stylirostris* and *P. monodon* were recombinantly expressed in *E. coli*. Antibacterial and antifungal activities of the recombinant ALFs were demonstrated with a wide range of microorganisms. Analysis of the results showed that both ALFs showed greater variation in their activity spectrum. ALFs from *L. stylirostris* was found to possess impaired LPS binding activities and only limited antimicrobial activity when compared to that of *P. monodon* ALF. MIC of *L. stylirostris* ALF ranged from 2.5 to >10 μM; whereas that of *P. monodon* ALF was found to be ranging from 0.15 to >10 μM.

Sun et al., 2015 reported recombinant expression of ALF from *S. paramamosain.* Recombinant pET32a+ with SpALF5 was transformed to *E.coli* Rosetta-gami TM2 (DE3) plysS competent cells. Recombinant protein was purified by Ni-NTA His bind resin and was further subjected to antimicrobial assays. The recombinant SpALF5 protein showed a

varying degree of binding activity towards bacteria and fungus including Gram-negative bacteria (*V. parahemolyticus, V. alginolyticus, E. coli, A. hydrophila*) and fungus (*S. cerevisiae*), but could only inhibit the growth of some Gram-positive bacteria like *S. aureus.*

In the present study, in order to further characterize the biological activity of *S. serrata* ALF (SsALF2), a recombinant expression of the protein was carried out in *E. coli* and its antimicrobial activity was tested against various microorganisms. Along with that the cytotoxic effects of the expressed protein were also performed.

5.2 Materials and Methods

5.2.1 Target gene for recombinant expression

A novel ALF isoform identified from *S. serrata* viz., SsALF2 was selected for recombinant expression. The mature peptide region of SsALF2 excluding the signal peptide region (named as SSA) possessing 294 bp was selected for recombinant expression. The target gene possessed 91 amino acid residues in its mature peptide region with a predicted molecular weight of 11.2 kDa. The physico-chemical parameters and structural features of SsALF2 have already been discussed in detail in Chapter 2. The nucleotide and amino acid sequence of the target gene is given in Fig. 5.1.

5.2.2 Designing primers with restriction sites and PCR amplification

5.2.2.1 Primer designing

Primers were designed by adding restriction site sequence of Nco1 (CCATGG) to the 5' end of forward primer and Hind III (AAGCTT) to the reverse primer respectively. 6x Histidine tag (a stretch of 6 His amino

acids) was also added to 5' end of forward primer. The details of the primers are given in Table 5.1.

5.2.2.2 PCR amplification

cDNA synthesized using total RNA extracted from the haemocytes of the mud crab, *S. serrata* was used as template for PCR amplification. The details of the ALF isoform SsALF2 is explained in detail in Chapter 2. The mature peptide region of SsALF2 was amplified using restriction primers mentioned in section 5.2.2.1. PCR amplification of 1 µl of cDNA was performed in a 25 µl reaction volume containing 1x standard Taq buffer (10 mM Tris-HCl, 50 mM KCl, pH 8.3), 2.5 mM $MgCl_2$, 200 µM dNTPs, 0.4 µM each primer and 1U Taq DNA polymerase). The thermal profile used for the PCR amplification was 94 ^{0}C for 2 min followed by 35 cycles of 94 ^{0}C for 15 sec, 60 ^{0}C for 30 sec and 68 ^{0}C for 30 sec and a final extension at 68 ^{0}C for 10 min. PCR products were analyzed by electrophoresis in 1.5 % agarose gel in TBE buffer, stained with ethidium bromide and visualized under UV light. The PCR products were cloned and confirmed by sequencing before proceeding to further experiments.

5.2.3 Cloning the target gene to the cloning vector, pGEM-T Easy vector

The PCR products were cloned onto the pGEM-T Easy vector as explained in section 2.2.9 of Chapter 2. *E. coli* DH5α competent cells were used for transforming the ligated cloning vector. To confirm the presence of the insert DNA in the transformed cells, colony PCR was performed for all the individually streaked colonies using vector specific primers viz. T7 and SP6 as well as gene specific primers (Table 5.1). Plasmid with the insert

was then extracted and purified using GenElute HP plasmid MiniPrep kit (Sigma) and was sent for sequencing at SciGenom, Kochi, India.

5.2.4 Restriction digestion

Plasmids after confirming the presence of the insert was subjected to restriction digestion using the restriction enzymes *Nco*1 and *Hind*III (FastDigest restriction enzymes, Thermo), thereby releasing the 6x His tagged target gene. For performing restriction digestion 100 µl of purified plasmid was incubated with 10 µl of reaction buffer and 1 unit each of *Nco*1 and *Hind* III for 1 h at 37 °C followed by an inactivation at 65 °C for 20 min. Restriction digestion was confirmed by agarose gel electrophoresis by checking the presence of released insert.

5.2.5 Construction of expression vector, pET32a+

5.2.5.1 Restriction digestion of the expression vector

The expression vector, pET32a+ translation vector was also subjected to restriction digestion with *Nco*1 and *Hind* III enzymes as explained in previous section 5.2.5.

5.2.5.2 Gel elution of restriction digested insert and expression vector

The restriction digested target insert gene and the expression vector, pET32a+ were gel purified using GenElute Gel extraction kit (Sigma, USA) following manufacturer's instructions. Briefly, the agarose gel containing DNA fragment of appropriate sizes were excised and transferred to 2 ml pre-weighed vials. The excised gel was then incubated with gel solubilization solution at 60 ^{0}C for 10 min with repeated vortexing. Following incubation, 1x gel volume of 100 % isopropanol was added and

was gently mixed until it became homogenous. The solubilized gel solution was then loaded onto the binding column and was subjected to centrifugation at 12000 xg for 1 min. This was followed by two washing steps using 700 µl of the provided wash solution and further centrifugation at 12000 xg for 1 min. Finally about 50 µl of elution buffer (10 mM Tris-HCl, pH 9.0) was added to the column and was centrifuged at 12000 x g for 1 min and stored at -20 ^{0}C. The products were confirmed by agarose gel electrophoresis and the concentration was measured spectrophotometrically at 260/280 nm in a UV-VIS spectrophotometer (U2800, Hitachi, Japan).

5.2.5.3 Ligation and transformation of pET32a+ to *E. coli* DH5α competent cells

The restriction digested and purified target gene was ligated to digested pET32a+ vector (Novagen, UK) by following the manufacturer's instructions. Briefly, 10 µl ligation mixture containing 1 µl pET32a+ vector (50 ng µl^{-1}), 4 µl target gene, 1 µl ligation buffer (10x), 1 µl T4 DNA ligase (1U µl $^{-1}$) and 3 µl MilliQ were incubated at 4 °C overnight. The ligated products were then transformed to *E. coli* DH5α competent cells as discussed in section 5.2.3.3. The clones were selected and patched on LB/ampicillin (100 µg µl^{-1}) plates to further confirm the transformation process. All individually streaked colonies were subjected to colony PCR using vector specific primers viz. T7-forward and T7-reverse primers as well as gene specific primers (Table 5.1). White colonies (template) picked from the transformed plate were dispensed into the PCR reaction mix (25 µl) containing 1x standard Taq buffer (10 Mm Tris-HCl, 50 mM KCl, pH 8.3), 200 µM dNTPs, 0.4 µM each primer (T7 forward and T7 reverse) and 1U Taq DNA polymerase. The thermal profile used was an

initial denaturation at 95 °C for 5 min followed by 35 cycles of denaturation at 94 °C for 15 sec, annealing at 52 °C for 20 sec and extension at 72 °C for 1 min and a final extension at 72 °C for 10 min. PCR products were then visualized by electrophoresis and documented using Gel documentation system (Gel Doc™ XR+ imaging system, Bio-Rad, USA).

5.2.5.4 Plasmid extraction and sequencing

After confirmation, the transformed *E. coli* DH5α containing recombinant expression vector construct was propagated in 10 ml LB ampicillin (100 μg μl^{-1}) medium at 37 °C, with shaking at 250 rpm. Plasmid extraction was carried out as discussed in section 5.2.4. The cloned plasmids were sequenced at SciGenom, Kochi, India and the sequences were analyzed using GeneTool software for the in-frame confirmation of the genes with the ATG of the pET32a+ system.

5.2.6 Transformation into expression host, *E. coli* Rosettagami B (DE3) pLysS

5.2.6.1 Selection of expression host

E. coli Rosetta-gami™ B (DE3) pLysS was selected as the expression host for heterologous production of SSA. *E. coli* Rosetta-gami™ B (DE3) pLysS belongs to Rosetta-gami B host strains. Rosetta-gami B combines the key features of BL21, Origami, and Rosetta to enhance both the expression of eukaryotic proteins and the formation of target protein disulfide bonds in the bacterial cytoplasm. Rosetta-gami B strains are derived from Origami B cells, a kanamycin-sensitive K-12 strain carrying the *trxB* and *gor* mutations for disulfide bonds formation in the cytoplasm. The Rosetta strains supply tRNAs for the codons AUA, AGG, AGA, CUA, CCC, and GGA on a compatible chloramphenicol-resistant plasmid, pRARE. The codons

are under the control of their native promoter. By supplying these rare codons, the Rosetta strains facilitate universal translation, which would otherwise be limited by the *E. coli* codon. BL21 *lacZY* deletion mutant allows precise control with IPTG.

5.2.6.2 Transformation to the expression host

E. coli Rosetta-gami™ B (DE3) pLysS cells were made chemically competent using $CaCl_2$ by following the protocol explained in section 5.2.3.2. The competent cells formed are stored in -80 °C with addition of 10-15 % glycerol. About 5 µl of the ligated pET32a+ plasmid were transformed to the competent *E.coli* Rosetta-gami™ B (DE3) pLysS cells employing heat shock method as explained in section 5.2.3.3. The transformation mixture (200 µl) was spread onto LB agar plates supplemented with ampicillin (100 µg/ml). Plated 100 µl of each transformation culture onto duplicate LB plates supplemented with ampicillin. Plates were then incubated overnight at 37 °C. Colonies obtained were patched onto LB plates supplemented with ampicillin.

5.2.7 Recombinant expression of fusion protein

Single colonies of recombinant pET32a+ SSA constructs in *E. coli* Rosetta-gami™ B (DE3) pLysS were inoculated into 5 ml LB broth media supplemented with ampicillin (100 µg μl^{-1}) and kanamycin (50 µg μl^{-1}) and was incubated at 37 °C at 250 rpm overnight. One ml of this overnight culture was then propagated in 100 ml antibiotic supplemented LB broth by incubating at 37 °C for 3-4 h until OD_{600} reaches 0.6. Cells were induced with 0.1 mM IPTG when OD_{600} reaches 0.6. Uninduced culture, collected prior to IPTG induction, served as the negative control. Both the induced and uninducedcultures were incubated further for 3 h at 37 °C with

shaking at 250 rpm. On a time-course basis 2 ml aliquot of the induced culture was removed each hour upto six hours and was subjected to SDS-PAGE for checking the level of expression of recombinant SSA (rSSA). The cells were harvested by centrifugation at 12000 xg for 2 min at room temperature and were stored at -20 °C until further processing.

5.2.8 Purification of the recombinant AMP

5.2.8.1 SDS-PAGE analysis

Level of expression of recombinant SSA was analyzed by performing Tricine – sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE). Prior to loading, the cell pellets were subjected to boiling for 15 min along with 10 µl of sample buffer containing 150 mM Tris-Cl, pH 7, 12 % SDS, 30 % glycerol, 6 % mercaptoethanol and 5 % Coomassie brilliant blue R-250. The samples were then centrifuged and the supernatant was loaded onto 16 % Tricine – SDS-PAGE gel as per Schägger (2006). The stacking and resolving / running gels were prepared at a concentration of 4 % and 16 % respectively. Electrophoresis was performed using 1 x Tris (pH 8.9) as the anode buffer and 1 x Tris-tricine (pH 8.3) as the cathode buffer at a voltage of 50 V in stacking gel and 120 V in resolving gel respectively (4-gel Mini-PROTEAN® Tetra cell protein electrophoresis unit, BioRad, USA). After electrophoretic separation, gels were stained in staining solution containing coomasie brilliant blue stain R-250 (0.5 % coomasie brilliant blue R-250, 40 % methanol and 10 % acetic acid in distilled water) and de-stained using de-staining solution (10 % methanol and 10 % acetic acid in distilled water) and was documented using Gel-DOC™ XR+ imaging system (BioRad, USA). Recombinant protein expression was

confirmed by the presence of band at the expected size range. BioLit low molecular weight protein ladder (3 – 40 kDa) was used for determining the molecular weight.

5.2.8.2 Ni-NTA column purification

Ni-NTA column purification was employed for purifying the recombinant SSA possessing a 6x His tag. Qiagen Ni-NTA superflow cartridges pre-filled with 1ml Ni-NTA superflow and ready to use for purification of 6x His-tagged proteins was used for purification. The peptides were purified under denaturing condition following manufacturer's protocol. Briefly, cell lysate was prepared by mixing the thawed cell pellets with 700 µl of buffer B (7 M Urea, 100 mM NaH_2PO_4, pH 8.0) and was vortexed continuously for 15 min. followed by centrifugation at 6000 xg for 30 min. The supernatant was subjected to syringe filtration containing a cartridge equilibrated with buffer B at a flow rate of 1 ml min^{-1}. Further washing steps were performed with buffer C (wash buffer-8 M Urea, 100 mM NaH_2PO_4, pH 6.3) at the same flow rate. Finally the recombinant protein was eluted using buffer E (elution buffer- 8 M urea, 100 mM NaH_2PO_4, pH 4.5) and the elute was subjected to Tricine-SDS-PAGE for confirming the presence of recombinant protein and was stored at -20 ^oC for further use.

5.2.9 Enterokinase treatment

Enterokinase was used to cleave the fusion proteins and to release the recombinant SSA. Briefly, 1 ml of refolded proteins was mixed with 2 units of enterokinase (Sigma; E0885-40UN). The reaction was incubated at 37 ^oC for 14-16 h for digestion of the recombinant protein. Enterokinase treated sample was analyzed on Tricine SDS-PAGE to

confirm digestion. Refolded protein without enterokinase treatment was loaded as the control.

5.2.10 Concentration of recombinant protein using Amicon Cut off filtration

The eluted samples were concentrated to one tenth of original volume using Millipore's Amicon Ultra-Centrifugal 3 kDa cut-off membrane by centrifuging at 5000 x g for 30 min. The concentrated samples were reconstituted to the original volume using the refolding buffer (Tris-Cl pH- 8, 50 mM, EDTA- 0.1 mM, NaCl- 0.15M) and centrifuged at 5000x g for 30 min. The concentrated sample was given 10-12 washes with the refolding buffer to remove the denaturing salts and thereby refolding the protein to its native form.

5.2.11 Quantification of recombinant AMP

Recombinant SSA was quantified using Quant-iT™ protein assay kit in a Qubitfluorometer (Invitrogen) following manufacturer's protocol. The assay is accurate for initial sample concentrations from 12.5 µg ml^{-1} to 5 mg ml^{-1} and exhibits low protein-to-protein variation. The assays were performed at room temperature, and the signal was stable for 3 h. Briefly, the Quant-iT working solution was made by diluting the Quant-iT protein reagent in Quant-iT protein buffer (1:200). The Quant-iT working solution was mixed well without any air bubbles and 190 µl was aliquoted to 0.5 ml tubes. To this, 10 µl of the recombinant peptide, SSA was added and mixed by mild vortexing without any air bubbles, as air bubbles cause error in the readings. The tubes were incubated for 15 min at room temperature. The samples were then measured using the Quant-iT protein program. The readings were recordedin µg ml^{-1}. This value

corresponded to the concentration after the sample was diluted into the assay tube. The sample concentration was calculated using the following equation:

Concentration of sample = QF value x (200/X)

Where, QF value = the value given by the Qubit Fluorometer,

X = the number of µl of sample added to the assay tube.

5.2.12 Antimicrobial assay

Antimicrobial activity of the recombinant SSA was checked against both Gram-positive and Gram-negative bacteria employing liquid growth inhibition assay as described by Huang et al. (2006) with modifications. Six dilutions (ranging from 0.1 to 10 µM) of the recombinant peptide in triplicates were subjected to assay.

5.2.12.1 Microorganisms used

The microbial strains used to determine antimicrobial activity included two Gram-positive bacteria viz. *B. cereus* (MCCB 101) and *S. aureus* (MTCC 3061) and five Gram-negative bacteria viz. *E. tarda* (MTCC 2400), *P. aeruginosa* (MCCB 119), *A. hydrophila* (MCCB 113), *V. cholerae* (MCCB 129) and *V. parahaemolyticus* (MCCB 133). The cell number of the bacterial strains corresponding to the optical density was found out employing spread plate method. During assay, bacterial strains were tested for purity by repeated streaking on sterile nutrient agar plates. The isolated colonies obtained were picked and streaked onto sterile nutrient agar slants, which were then incubated at 37 ^{o}C for 24 h. To this, sufficient quantity of HEPES buffer (4-[2-hydroxyethyl]-1-piperazine ethane sulfonic acid: 50 mM HEPES buffer was prepared by dissolving 26.03 g of

HEPES in 400 ml of MilliQ water. pH was adjusted to 7 using 1N NaOH. This was made up to a final concentration of 500 ml using MilliQ water) was added. The bacterial cells were scraped off using sterile inoculation loop and mixed well for uniformity. The Optical Density (OD) was measured at 600 nm. OD was adjusted so that every 10µl of bacterial suspension contained approximately 10^3 colony forming units (cfu).

5.2.12.2 Liquid growth inhibition assay

Antimicrobial activity of the recombinant SSA was tested employing liquid growth inhibition assay as described by Huang et al. (2006) with modifications. For performing antimicrobial assay, six dilutions of the recombinant SSA ranging from 0.1 to 10 µM along with positive and negative controls in triplicates were subjected to test. Briefly, 10 µl aliquots of each test concentrations of the recombinant SSA were mixed with 10 µl of the bacterial suspensions containing ~10^3 cells/ml in microtiter wells. Negative control constituted of 10 µl of the bacterial suspension mixed with 10 µl of 50 mM Tris-HCl instead of the peptide sample and the positive control constituted of 10 µl of the bacterial suspension mixed with 10 µl of ampicillin solution (1 mg/ml). After an incubation of 2 h at room temperature, 80 µl of nutrient broth was added to each well, followed by incubation of approximately 5 to 6 hrs at 37 ˚C. The incubation was succeeded by the addition of 25 µl of MTT (5 µg/ml solution of MTT (3-[4,5-dimethylthiazol-2-yl]-2,5-diphenyl tetrazolium bromide) prepared in phosphate buffer saline). It was then incubated at room temperature for 30 min. Then the reaction was stopped by adding 125 µl of acidified isopropanol (0.1 N HCl in isopropanol) and was incubated on a shaker at room temperature for 30 min. Bacterial growth was estimated colorimetrically by measuring absorbance at a wavelength

of 570 nm and reference wavelength of 630 nm using a microplate reader (Tecan, USA). Percentage of microbial growth inhibition was calculated using the following formula:

Growth % = (OD of microtitre wells containing bacteria + peptide sample/OD of positive control) x 100

Inhibition % = 100 – Growth %

5.2.12.3 Minimum inhibitory concentration

Minimum inhibitory concentration (MIC) of the recombinant SSA was determined by the above mentioned standard micro dilution method using 96-well microtiter plates. The MIC was expressed as the lowest concentration of the peptide that causes 100 % bacterial growth inhibition. MIC was calculated as the average value from three independent experiments, each performed in triplicate. MIC was again confirmed using PriProbit computer based program.

5.2.13 Cytotoxicity / Methyl thiazol tetrazolium (MTT) assay with NCI-H460 cells

Cytotoxicity assay was performed for recombinant SSA employing MTT assay on NCI-H460 cells. Briefly, 1×10^6 NCI-H460 cells were inoculated into each well of a 96 well tissue culture plate containing minimal essential medium (MEM). The tissue culture plate was then incubated for 12 h at 37°C. Following incubation, the wells were washed with phosphate buffered saline (PBS) and the medium was exchanged with MEM containing a series of two fold dilutions of rSSA ranging from 0.39 to 50 μM. MEM without synthetic peptide served as the negative control. MEM incubated with two commercially important potent synthetic AMPs

viz. melittin and magainin served as the positive controls. Each sample was treated in triplicates. After 24 h incubation at 37 °C, cells were observed for morphological changes under inverted phase contrast microscope (Leica, Germany). 50 µl MTT reagent (5 mg/ml) was then added to each well and was incubated for 4 h at 37 °C. Formazan crystals were dissolved by addition of 20 µl DMSO solution. The absorbance of each well was determined using the microplate reader (Tecan Microplate Reader) at 570 nm. Cell viability of NCI-H460 cells was calculated as a percentage of viable cells in the peptide treated group versus control group.

5.3 Results

5.3.1 PCR amplification of the target gene using primers with restriction sites

SsALF2, a novel ALF isoform, identified and characterized from *S. serrata* was selected for recombinant expression and for analyzing its biological activity. The mature peptide region of SsALF2, excluding the signal peptide region was subjected to recombinant expression and further functional characterization. The selected region, designated as SSA, possessed the active region i.e. the LPS binding domain and consisted of the sequence:

'QYEALVASILGKLSGLWHSDTVDFMGHTCHFFRKPKFRKFKLYHEGKFWCPG WHILIGNSRSKSRSGSTREATKDFVHKALQNKLITKNSADVWLKG'.

The physico-chemical properties of the mature peptide region have been explained in detail in Chapter 2. The target gene of SSA possessed a length of 97 amino acids and a molecular weight of 11.2 kDa. The SSA was found to be highly cationic with a net charge of +10.5. SSA possessed an

isoelectric point (p*I*) of 10.02 and hydrophobicity of 37 % as predicted by the PROTPARAM and HELIQUEST programs. PCR amplification of the target gene, SSA, was performed using primers designed by adding restriction site sequence of Nco1 (CCATGG) and 6x Histidine tag to the 5' end of forward primer and Hind III (AAGCTT) to the 3'end of the reverse primer (Table 5.1). A 294 bp target geneamplicon was obtained by PCR amplification of *S. serrata* cDNA using these primers (Fig. 5.2).

5.3.2 Cloning the amplified product to a cloning vector, pGEM®-T Easy vector

The PCR amplicons was cloned onto pGEM®-T Easy cloning vector and transformed to *E. coli* DH5α competent cells. Transformed cells were plated on to LB plates supplemented with antibiotics and positive colonies were identified based on blue white screening. The white coloured colonies obtained were patched onto LB plates and the presence of plasmid with insert was confirmed by colony PCR using vector specific (T7 and SP6) and gene specific primers (SSA-F and SSA-R) (Table 5.1). PCR amplicons using vector specific primers possessed a size of 456 bp and that of gene specific primer 330 bp (including His tag and restriction sites) respectively (Fig. 5.3a and b). From the positive colonies, plasmids with inserts were extracted using GenElute MiniPrep plasmid isolation kit, Sigma and was further confirmed by PCR using vector specific and gene specific primers using plasmid as the template (Fig. 5.4a and b). Plasmids were sequenced and sequences were checked for the presence and orientation of the restriction sites, 6x His tag and the target gene.

5.3.3 Restriction digestion

Plasmids after confirmation of the presence of the restriction site, 6x His tag and the orientation of the target gene were subjected to restriction digestion using Nco1 and Hind III restriction enzymes. Agarose gel analysis of the restriction digested products confirmed the release of inserts possessing a size of 330 bp (Fig. 5.5a). pET32a+ expression vectors were also restriction digested subsequently (Fig. 5.5b).

5.3.4 Cloning of the target gene to an expression vector, pET32a+

Restriction digested target gene and pET32a+ vector were subjected to ligation and were transformed to *E. coli* DH5α competent cells. Amplicons of size 1080 bp (330 + 750 = 1080 bp) was obtained by PCR amplification of the colonies using vector specific primers (T7-F and T7-R) and gene specific primers (SSA-F and SSA-R) respectively. Plasmids were extracted from positive colonies and were confirmed for the presence of insert in correct frame by sequencing using vector specific primers. After confirmation of the correct frame of the 6x His tag and the target gene in pET32a+, the vector was transformed to the expression host, *E.coli* Rosetta-gami™ B (DE3) pLysS competent cells. Recombinant clones possessing the target gene, SSA was obtained on LB agar plates with ampicillin.

5.3.5 Recombinant expression of the fusion protein

Single colonies of recombinant pET32a+ SSA constructs in *E. coli* Rosetta-gami™ B (DE3) pLysS were propagated in LB broth media supplemented with antibiotics overnight. One ml of this overnight culture was further propagated in 100 ml antibiotic supplemented LB broth until OD_{600} reaches 0.6. The expression of SSA was induced with 0.1 mM IPTG when

OD_{600} reaches 0.6. Uninduced culture, collected prior to IPTG induction, served as the negative control. Both the induced and uninduced cultures were incubated further for 3 h at 37 °C with shaking at 250 rpm. On a time-course basis 2 ml aliquot of the induced culture was removed each hour up to six hours and was subjected to SDS-PAGE for checking the level of expression of recombinant SSA. The cells were harvested by centrifugation and an initial Tricine-SDS-PAGE analysis of the total bacterial extracts of induced bacteria were compared at various time points with that of the uninduced sample. The induced lysates clearly demonstrated the presence of a strong band at 29.2 kDa which is consistent with the expected molecular weight of rSSA protein with an N-terminal His tag as well as the fusion protein.

Recombinant protein expression could be confirmed by the presence of band at the expected size range. Level of expression was not found to be varying with time (Fig. 5.6). Bulk production of rSSA in 1000 ml LB medium was carried out and cells were harvested after 3 h post IPTG induction. The recombinant expression was further confirmed using SDS-PAGE.

5.3.6 Purification, concentration and quantification of the recombinant protein

Purification and concentration of the rSSA was performed employing a combination of Ni-NTA column purification, enterokinase treatment, Amicon cut off filtration and refolding processes. Fusion protein of size 29.2 kDa (11.2 kDa of SSA + 18 kDa of vector proteins) could be obtained using Ni-NTA column purification of rSSA (Fig. 5.7a). The fusion protein was subjected to enterokinase treatment thereby releasing the rSSA of

11.2 kDa (Fig. 5.7b). Tricine SDS- PAGE profiles of the purified fraction clearly demonstrated an efficient purification of ~ 11.2 kDa rSSA in the eluted fraction. The released product was then concentrated using Amicon cut off filtration and was refolded using refolding buffer. The concentrated protein was quantified by Quant-iT™ protein assay kit and the concentration of the rSSA was found to be 230 μg/ml (20.5 μM).

5.3.7 Antimicrobial assay

To determine the antimicrobial activity of rSSA, the recombinant protein was prepared and tested for the ability to inhibit proliferation of microorganisms using MIC method. The assay was performed with rSSA at concentrations ranging from 0.1 to 10 μM against a group of Gram-positive and Gram-negative bacterial strains employing liquid growth inhibition assay. The antibacterial activity of rSSA, expressed as minimum inhibitory concentrations (MIC) against two Gram-positive bacteria and five Gram-negative bacteria tested is listed in Table 5.2. Analysis of the results showed that, SSA demonstrated a higher activity against Gram-negative bacteria. While Gram-positive *S. aureus* and *B. cereus* were not found to be inhibited at the tested concentrations. Also, among the five Gram-negative bacteria tested, rSSA showed activity against only three such as *V. cholerae* (MIC=11.5 μM), *E. tarda* (MIC=16.9 μM) and *P. aeruginosa* (MIC=77.1 μM). However, gram-negative bacteria such as *A. hydrophila* and *V. parahaemolyticus* were not found to be affected by the rSSA. Recombinant SSA was found to be highly active against *V. cholerae* and *E. tarda* at all tested concentrations (O.1-10 μM). The antibacterial activity of rSSA against the tested bacteria at various concentrations is illustrated in Fig. 5.8 (a-c).

5.3.8 Cytotoxicity / Methyl thiazol tetrazolium (MTT) assay with NCI-H460 cells

The cytotoxicity assay of rSSA was performed against NCI-H460 cells employing MTT assay. This assay measures the reducing potential of the cell using a colorimetric reaction. Viable cells will reduce the MTS reagent to a coloured formazan product. Analysis of the results proved that the rSSA is non-cytotoxic to the tested cells when compared to that of the two control AMPs used. The control melittin, was found to be highly cytotoxic even at smaller concentrations. A reduction of < 95 % was observed for NCI-H460 cells even at a smaller tested concentration of 6.25 μM in case of melittin. The IC_{50} of melittin against NCI-H460 cells was estimated to be ~ 3.3 μM. Whereas, in the case of rSSA, even at the highest concentration tested i.e. 50 μM, the peptide was found to inhibit only 35 % of growth in NCI-H460 cells. The result of cytotoxic assay against NCI-H460 cells is represented in Fig. 5.9.

5.4 Discussion

Two attractive properties of AMPs make them promising candidates as therapeutic agents; (1) they demonstrate a broad range of antimicrobial activities and (2) they mainly target microbial membranes, impeding the ability of microbes to develop resistance against them (Hancock and Sahl, 2006). However, a number of fundamental aspects such as its antimicrobial spectrum, mechanism of action, efficacy and safety of these molecules must be addressed before it can be brought to clinical trials. Extensive functional and structural studies need to be conducted for answering these questions; for which enough AMPs in higher concentration should be available in pure and active form. Apart from

that, peptide availability is one of the major factors that determine the feasibility of their widespread application as therapeutic agent. Therefore, for performing both basic researches as well as for clinical applications, high quality peptides should be readily available in a cost-effective manner.

Among the different approaches that can be employed for AMP production, heterologous expression of AMPs employing recombinant DNA technology provides an economical means for AMP production. Indeed, many AMPs have been successfully obtained through recombinant production in various heterologous hosts (Ingham and Moore, 2007). Among the systems available for recombinant protein production, *E. coli* has been the most widely used host (Li and Chen, 2008). Recombinant expression of these peptides in *E. coli* appears to offer a cost effective method for the production of these peptides as well as has got advantages such as fast growth, low cost and high expression. However, AMPs are difficult subjects for heterologous expression in microbial hosts due to their intrinsic antibacterial activity (Li, 2009).

Expression of ALFs in *E. coli* faces two challenges, such as (1) the antibacterial nature of ALFs makes them potentially fatal to the producing host and (2) the small size and high cationic property of ALF makes them highly susceptible to proteolytic degradation. A strategy that effectively overcomes both of these obstacles is to fuse the peptide of interest to a carrier protein. The target peptide can be released from the fusion at a later stage by enzymatic or chemical cleavage at corresponding site around the carrier-peptide junction. Also, the carrier protein protects the peptide from cellular proteases and the host from the toxic effects of these peptides. For fusion expression of more than 25 % AMPs, thioredoxin

carrier proteins are being used (Li, 2009). A recent study has showed that, among 13 different carrier proteins tested, thioredoxin fusion gave the highest absolute yield of the fused peptide. Different from its use as a fusion carrier, free thioredoxin increases the solubility of the co-expressed proteins/peptides mainly through promoting correct disulfide bond formation. However, thioredoxin require additional affinity tags for initial purification of the fusion proteins. The most commonly used affinity tag for this purpose is the hexa-histidine tag (6x His-tag). His-tag has also been proved sufficient for masking the toxicity of the AMPs.

Insoluble expression of AMPs is believed to be more efficient than soluble fusion in masking the toxic effects of AMPs as well as protecting them from degradation. However, several AMPs have been successfully expressed as soluble fusions and it is not always necessary to force fusion proteins into inclusion bodies for protecting the host and peptides. Also, insoluble expression of AMPs allows quick purification of the fusion as they can be easily collected from the cell lysate by centrifugation and can be further purified by simple washing procedure. Presence of His-tag, enables immobilized metal-affinity chromatography (IMAC) purification of AMPs under denaturing conditions. All fusion proteins require exogenous chemicals or proteases to release the target peptides. Enterokinase is one of the most widely used enzymes for this purpose. His-tag may or may not be removed from the target peptide before proceeding with downstream applications. Refolding is required to restore activity of most of the recombinantly expressed AMPs, as they contain multiple disulfide bonds. The native AMP genes often contain codonsthat are rarely used in *E. coli* and the codon bias may result in poor expression of the heterologous gene. The codon bias problem can be

solved by using specially engineered host strains that supply tRNAs for rare codons. Further optimization of environmental factors such as IPTG concentration and growth temperature may be required in order to obtain soluble fusions. In general, induction with low concentration of IPTG and growth at reduced temperature tends to improve solubility of the fusion proteins (Schein and Noteborn, 1988; Sorensen and Mortensen, 2005). In case of ALFs, special host strains that are engineered to promote disulfide bond formation (e.g. Novagen's Origami cells) can greatly increase the chance of obtaining soluble fusions with proper folding (Richard et al., 2004).

ALFs are small basic peptides that can bind and neutralize LPS, mediating degranulation and activation of an intracellular coagulation cascade. ALFs are considered as predominant candidates for potential therapeutic agents of viral and bacterial infectious diseases (Liu et al., 2005; Somboonwiwat et al., 2005; Imjongjirak et al., 2007; de la Vega et al., 2008; Rosa et al., 2008). The goal of this study was to investigate and characterize the biological properties of an ALF isoform identified in the crab, *S. serrata* (SsALF2). For that, we have chosen to produce this immune effector in the recombinant system of *E. coli*, which have been shown to be appropriate for the production of large amount of active molecules.

Briefly, the mature peptide region of SsALF2 was recombinantly expressed in *E. coli* Rosetta gami DE3 pLYsS cells using the pET32a+ vector possessing thioredoxin carrier protein. At the N-terminus 6 x His-tag was added for affinity purification purpose. Enterokinase cleavage was used to remove the fusion tag and was further concentrated and refolded to obtain the active form before subjecting it to functional

characterization. Tricine–SDS–PAGE was used to characterize the recombinant peptide and also to confirm enzyme cleavage and purification. Size exclusion purification was also done to separate rSSA.

The deduced protein sequence of SsALF2 revealed a conserved structure comprising a signal peptide, an LPS-binding domain, and the two conserved cysteine residues at both ends of the LPS-binding domain. The two cysteine residues formed a disulfide loop, within which the positively charged amino acid residues were mainly clustered. The mature peptide region of SsALF2 excluding the signal peptide region possessing 294 bp was selected for recombinant expression protocols. The selected region, designated as SSA, possessed the active region i.e. the LPS binding domain and consisted of the sequence:

'QYEALVASILGKLSGLWHSDTVDFMGHTCHFFRKPKFRKFKLYHEGKFWCPG WHILIGNSRSKSRSGSTREATKDFVHKALQNKLITKNSADVWLKG'.

SSA possessed a length of 97 amino acids and a molecular weight of 11.2 kDa and was found to be highly cationic with a net charge of +10.5. SSA possessed an isoelectric point (p*I*) of 10.02 and hydrophobicity of 37 % as predicted by the PROTPARAM and HELIQUEST programs. All the above mentioned features along with the positively charged amino acids in the disulfide loop have been demonstrated to be important for biological activity of the ALF molecules (Ried et al., 1996; Somboonwiwat et al., 2008; Yang et al., 2009).

In the present study, this mature peptide region of SsALF2 was recombinantly expressed and functionally characterized for its antibacterial and cytotoxic activity. This study presents the second report of recombinant expression of an ALF isoform from *S. serrata*. A 294 bp

nucleotide sequence representing the mature peptide region of SsALF2 was obtained by PCR amplification of *S. serrata* cDNA using specific primers designed. PCR amplicons were first cloned to pGEMT Easy vector and after confirming the sequence, it was ligated to pET32a+ vector. The plasmid was again sequenced to confirm the frame of insert and further transformed to expression vector, *E. coli* Rosetta gami (DE3) pLysS strain for recombinant peptide production. The recombinant expression was found to be maximum after 1 h of IPTG induction at a concentration of 0.1 mM and incubation at 37 ^{o}C and 250 rpm. It has been reported that when protein production is induced at an earlier stage of growth, lysis of bacterial cells would be easier (Moon et al., 2007). The cell pellets were then lysed under denaturing condition and the recombinant SSA were purified on Ni-NTA column, to which proteins with His-tag alone would bind. The purified products were confirmed by Tricine-SDS-PAGE. The step-wise reduction of urea with refolding buffer provided an extensive washing as well as the refolding of the recombinant peptides. The refolding buffer was then subsequently replaced with 50 mM Tris-Cl solution so as to remove EDTA which would otherwise hinder enterokinase treatment as well as the antimicrobial assay. The denatured rSSA were successfully refolded as judged by the clearness of the solution (misfolded proteins would precipitate in short period of time) and cleavage by enterokinase enzyme. The recombinant protein was purified to homogeneity and its antimicrobial activity determined with a MIC assay.

Antibacterial activity of this purified recombinant peptide, rSSA was assessed by employing liquid growth inhibition assay against seven bacterial strains such as Gram-negative *V. cholerae*, *E. tarda*, *P. aeruginosa*,

A. hydrophila and *V. parahaemolyticus* as well as Gram-positive *B. cereus* and *S. aureus*. The MIC results clearly indicated a strong inhibition by rSSA to the growth of Gram-negative bacterial cells than observed in case of Gram-positive bacteria.

Analysis of the results showed that rSSA to be strongly active against gram-negative bacteria such as *V. cholerae* (11.5 μM), *E. tarda* (16.9 μM) and *P. aeruginosa* (77.1 μM). At the same time, rSSA was not found to be inhibiting the growth of any Gram-positive bacterial strains tested such as *B. cereus* and *S. aureus*. Also, not all Gram-negative bacteria tested were found to susceptible to rSSA as in the case of *A. hydrophila* and *V. parahaemolyticus*.

The results are in agreement with the observation of another ALF isoform identified and characterized from *L. polyphemus* (LALF), that exhibited anti-bacterial activity against Gram-negative bacteria alone, and not against the tested Gram-positive bacteria (Morita et al., 1985). Reports on ALF of *S. serrata* by Yeddey and Reddy (2009) also reported that the recombinant ALF from *S. serrata* was found to be active against Gram-negative bacteria alone. No activity could be seen against Gram-positive bacterial strains. Similar cases have been reported for ALF from *P. trituberculatus* also. Where, the recombinant peptide showed activity against Gram-negative bacteria alone (Liu et al., 2012). Such findings could be attributed to the presence of LPS in the Gram-negative bacterial cell wall, which has been demonstrated to be the binding target of ALF, as reported earlier (Weiss et al., 2000). Along with that, it is a well-established fact that the ability of AMPs to interact with and insert into negatively charged cytoplasmic membranes of their target bacterial cells are attributed to the presence of spatially segregated hydrophobic and

cationic domains and several membrane-permeabilization mechanisms employing liposome models have been proposed (Shai, 2002).

At the same time, some ALFs have been reported to possess broad spectrum antimicrobial activity including Gram-negative, Gram-positive bacteria, fungi and viruses (Somboonwiwat et al., 2005; Imjongjirak et al., 2007; Sun et al., 2011; Liu et al., 2012). However, in case of shrimp ALFs, as in case of *P. monodon*, broad spectrum antimicrobial activity including a spectrum of anti-fungal activity and anti-bacterial activities against both Gram-positive and Gram-negative bacteria have been demonstrated (Somboonwiwat et al., 2005). Another ALF from *E. sinensis* have also been demonstrated similar bactericidal activity against both Gram-positive and Gram-negative bacteria (Li et al., 2008). Whereas, yet another isoform from the same species i.e. *E. sinensis*, indicated that the purified form displayed anti-bacterial activity against a Gram-negative bacterium and anti-fungal activity, but not against Gram-positive bacteria (Zhang et al., 2010). ALFs present in crayfish *P. clarkii*, rPcALF1, when tested with a series of microorganisms, displayed an extensive binding activity with all the Gram-positive and Gram-negative bacteria and fungi (Sun et al., 2011).

As suggested by Brogden (2005) and Imjongjirak et al. (2011), it may be presumed that the differences of the charged amino acids such as arginine and lysine could account for the different antimicrobial activities of these ALF molecules. The antimicrobial activity of recombinant ALF from *S. paramamosain* also showed activity against both Gram-positive and Gram-negative bacteria. It has been demonstrated by ELISA-based binding assay, that the antimicrobial and bactericidal activities of ALFs rely on the binding activity with cell wall components of the pathogenic organisms (Sun et al., 2011).

The MIC values obtained are also in agreement with that of previous results. rSSA showed high activities against Gram-negative bacteria. The rSSA was found to strongly inhibit the growth of *V. cholerae* (MIC=11.5 μM), *E. tarda* (MIC=16.9 μM) and *P. aeruginosa* (MIC=77.1 μM), the results are in conformity with that of the previous report on the biological activity of another isoform of ALF from *S. serrata*. In that case also recombinantly produced ALF isoform was found to possess higher activity against Gram-negative bacteria than to Gram-positive bacteria. The MIC was found to be higher for Gram-positive bacteria viz. *S. aureus* and *S. pyogenes* ranging between 9.1 to >18.2 μM. Whereas in case of Gram-negative such as *E. coli* and *P. aeuriginosa*, the MIC values were found to be 4.5 and 9.1 μg/ml respectively (Yedery and Reddy, 2009).

A shrimp ALF from the black tiger shrimp *P. monodon* have been produced by recombinant expression in *P. pastoris* system by Somboonwiwat and colleagues (Somboonwiwat et al., 2005). The rALFPm3 demonstrated a broad spectrum of anti-fungal properties against filamentous fungi, and anti-bacterial activities against both Gram-positive and Gram-negative bacteria. The MIC was found to vary between 0.19 to 6.25 μM in case of Gram–negative bacteria. It may also be noted that rALFPm3 showed an anti-microbial activity against all the Gram-positive bacteria, with the exception of *S. aureus* which remained unaffected even in the concentration range of 50–100 μM. Comparatively, in the present study, we have checked activity against two Gram-positive bacteria viz. *B. cereus* and *S. aureus*. The maximum concentration which we could apply for checking antibacterial properties was 10 μM. May be at a higher concentration of rSSA of 100 μM, it might be active against Gram-positive. The MIC of rALFPm3 against Gram-negative bacteria was found

to be 0.1-6.25 μM with an exception of *V. penaeidicida* where the MIC was found to be 50 μM. MIC against other vibrios such as *V. harveyi*, *V. anguillarum* and *V. alginolyticus* was within the range from 0.4 to 1.56 μM. Hence it should be stipulated that there does exists variation in the MIC required for specific pathogens even if they belong to the same genus.

Comparatively, as expected, the antibacterial activity of rSSA was found to be high at higher concentration of the peptide and was found to reduce gradually with reduction in concentration. To date, very few AMPs have been reported to be highly active against *Vibrio* strains apart from the mytilins, which is also isolated from a marine invertebrate, the mussel *Mytilus galloprovincialis* (Mitta et al., 2000), and which display MIC values similar to those observed for rSSA.

Antibacterial activity of rSSA against Gram-negative bacteria belonging to genus *Edwardsiella*, *Vibrio* and *Pseudomonas* suggested that rSSA could play an essential role in defence mechanism in *S. serrata*. Apart for a direct antimicrobial activity against potentially harmful bacteria, rSSA would be a major effector of the shrimp immune system, which could be involved due to its LPS-binding property to various reactions such as inhibition of inflammatory reaction or anticoagulant activity. Analysis of the mature SSA sequence using the HELIQUEST program clearly indicated the significant presence of hydrophobic amino acids in the protein. These hydrophobic amino acid residues might be part of α-helix and might be able to span across a lipid bilayer or membrane. The presence of α-helical regions, comprising greater portion of the peptide structure, was also confirmed by the SWISS-MODEL and PyMOL structures. As revealed by the AMP Database programthe presence of putative LPS binding region ofthe recombinant ALF, rSSA situated between the two conserved

cysteine residues exhibited a strong positive charge of +10.5 with 37 % hydrophobicity. This explains the possibility of the LPS binding domain to bind to the highly negative charged LPS layer. Overall the protein demonstrated the presence of amphipathic regions, a characteristic feature of well-studied AMPs.

Some AMPs are known to cause harmful side effects, such as toxicity to cells at peptide concentrations demonstrating inhibition of bacterial growth (Reddy et al., 2004). The cytotoxicity assay performed during the present study had showed that the recombinant SSA does not cause any cytotoxicity to human NCI-H460 cells even at a concentration of 50 μM. Among the two important AMPs used as controls for the study, Melittin was found to be highly cytotoxic even at very low conc of < 1 μM whereas magainin was found to be non-cytotoxic. According to literature, the cytotoxic potential depends on the type of peptide and its concentration. The result reveals the possibility of rSSA as a safe therapeutic agent. Despite the results obtained from the cytotoxicity study outlined further *in vivo*, work involving human and/or animal studies would be required to correlate the true antimicrobial and toxic potential of the ALF, rSSA clinically.

In conclusion, in the present study, an ALF isoform identified from *S. serrata* was cloned and expressed in *E. coli* Rosetta-gami™ B (DE3) pLysS expression system. The recombinant ALF, rSSA was purified and assayed for their antibacterial property. A recombinant ALF exhibiting a bactericidal activity specifically towards Gram-negative bacterial species could be demonstrated in this study. The rSSA was found to exhibit strong antibacterial activity against *V. cholerae, E. tarda* and *P. aeuriginosa*. Results of the study indicate the possible role of ALFs in the immune

responses of crabs. The increasing number of ALF being identified and characterized implies there are abundant and diverse ALFs in *S. serrata* with various biological functions. Recombinant ALF, rSSA has the potential to be developed as antibacterial agents against *Vibrio* strains as well as against *E. tarda,* which are pathogenic to humans and fishes. The recombinant ALF, rSSA could also be used to develop as a potential therapeutic agent for preventing LPS-induced sepsis in humans / aquaculture. It will be of a prime importance to further analyze the respective kinetics of expression and localization of SSA effector in response to *Vibrio* and *Edwardsiella* infection. Advances in characterization of AMPs may lead to a better understanding of the crab immune defense mechanisms especially that of *S. serrata* and will surely give new insights into the health management and disease control in crab aquaculture.

At present, increased resistance of bacteria against antibiotic drugs has attracted intensive effort to screen antimicrobial peptides as sources or templates for designing novel therapeutic antibiotics. Regarding its antimicrobial properties, the crab SsALF2 appear to be good candidate for further study about their potential use in aquaculture farming and, importantly, as a promising alternative to conventional antibiotics. Finally, due to its antimicrobial properties, the crab ALF appears to be a good candidate for further investigation about its potential use in larviculture as therapeutic agent and as alternative to conventional antibiotics. The use of recombinant DNA technology is a promising method to produce bulk quantity for AMPs for clinical research. Recent advances in technology have led to the development of new cloning methods for overexpression and purification of recombinant AMPs.

Further improvement in these technologies is expected as clinical interest shifts from antibiotics to AMPs.

```
cagtatgaagctctggtagcttccattcttggaaagctgtcgggactgtggcacagcgac
 Q  Y  E  A  L  V  A  S  I  L  G  K  L  S  G  L  W  H  S  D
acagtggacttcatgggacacacctgccatttttccgcaagccgaagttcaggaaattt
 T  V  D  F  M  G  H  T  C  H  F  F  R  K  P  K  F  R  K  F
aagctttaccacgagggcaagttttggtgtccgggttggcatatcttgattggcaattcg
 K  L  Y  H  E  G  K  F  W  C  P  G  W  H  I  L  I  G  N  S
aggtccaagagcaggtcggggtcaaccagggaagccaccaaggacttcgtgcacaaagct
 R  S  K  S  R  S  G  S  T  R  E  A  T  K  D  F  V  H  K  A
ttacaaaacaaactcatcacgaagaatagcgcggacgtctggctgaaggggtga
 L  Q  N  K  L  I  T  K  N  S  A  D  V  W  L  K  G  *
```

Fig. 5.1 Nucleotide (above) and deduced amino acid (below) sequence of the mature peptide region of SsALF2 selected for recombinant expression. The stop codon is indicated by asterisk

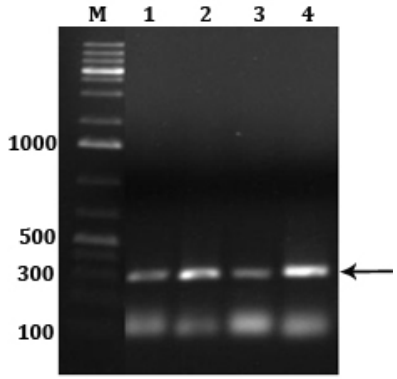
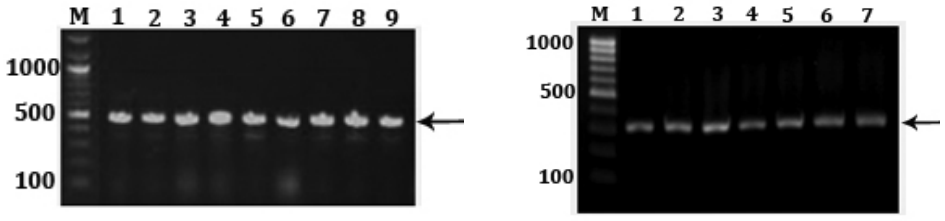
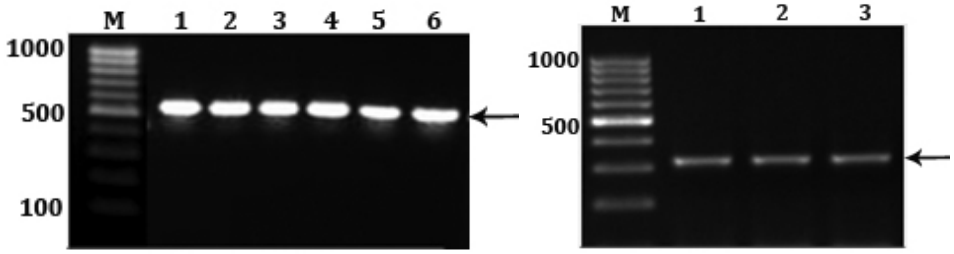


Fig. 5.2 Agarose gel electrophoretogram of the PCR amplified mature peptide region of SsALF2. Lane M: 100 bp ladder; Lane 1-4: PCR amplified product (330 bp).



(a) **(b)**

Fig. 5.3. Agarose gel electrophoretogram of the colony PCR performed using **(a)** vector specific primers, T7-F and SP6. Lane M: 100 bp ladder; Lane 1-9: PCR amplified product (456 bp). **(b)** gene specific primers, SSA-F and SSA-R. Lane M: 100 bp ladder; Lane 1-7: PCR amplified product (330 bp).



(a) **(b)**

Fig. 5.4. Agarose gel electrophoretogram of the PCR amplification of plasmid performed using **(a)** vector specific primers, T7-F and SP6. Lane M: 100 bp ladder; Lane 1-6: PCR amplified product (456 bp). **(b)** gene specific primers, SSA-F and SSA-R. Lane M: 100 bp ladder; Lane 1-3: PCR amplified product (330 bp).

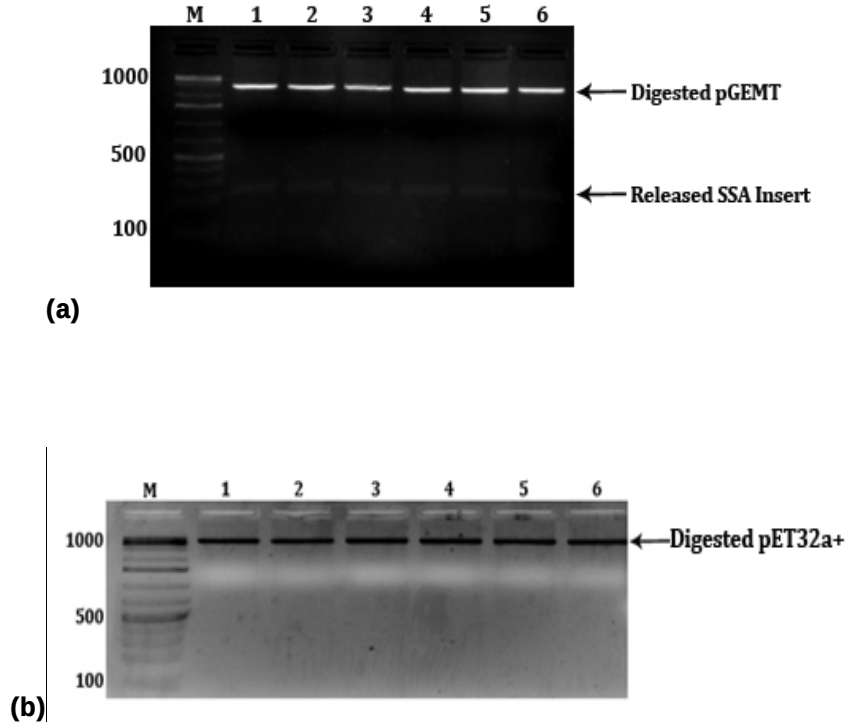


Fig. 5.5 Agarose gel electrophoretogram of the plasmids digested with NCo1 and HindIII restriction enzymes. **(a)** digested pGEMT–SSA plasmid **(b)** digested pET32a+ expression vector. Lane M: 100 bp ladder; Lane 1-6: restriction digested plasmids.

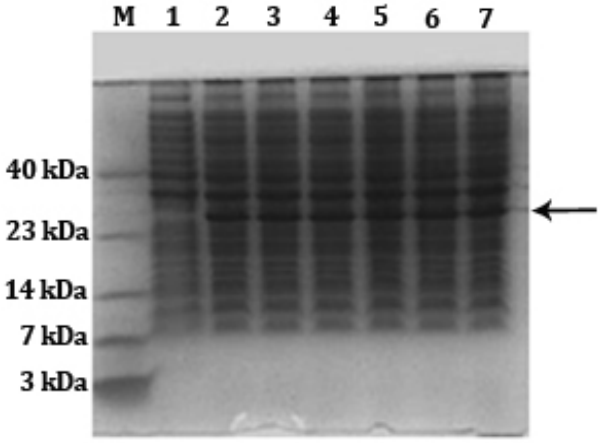
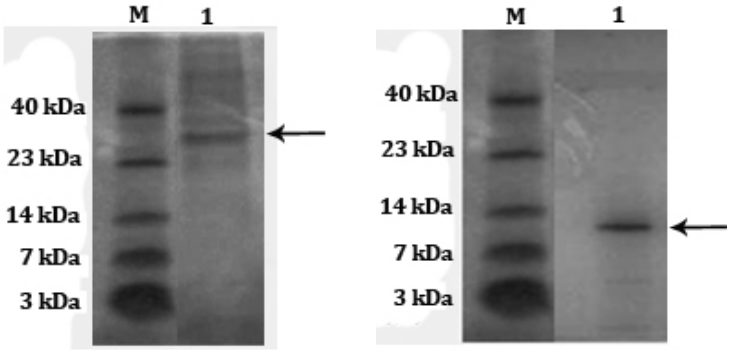


Fig. 5.6 Tricine SDS-PAGE analysis of the cells containing recombinantly expressed ALF, rSSA, before and after IPTG induction on a time-course basis. Lane M: Low molecular weight protein ladder; Lane 1: uninduced control (before IPTG induction); Lane 2-7: IPTG induced cells after 1-6 hours of induction.



(a) **(b)**

Fig. 5.7 Tricine SDS-PAGE analysis of recombinantly expressed ALF, rSSA at different steps of purification. **(a)** Ni-NTA purified recombinant peptide before restriction digestion (29.2 kDa) **(b)** Ni-NTA purified recombinant peptide after restriction digestion (11.2 kDa). Lane M: Low molecular weight protein marker; Lane 1: purified recombinant SSA.

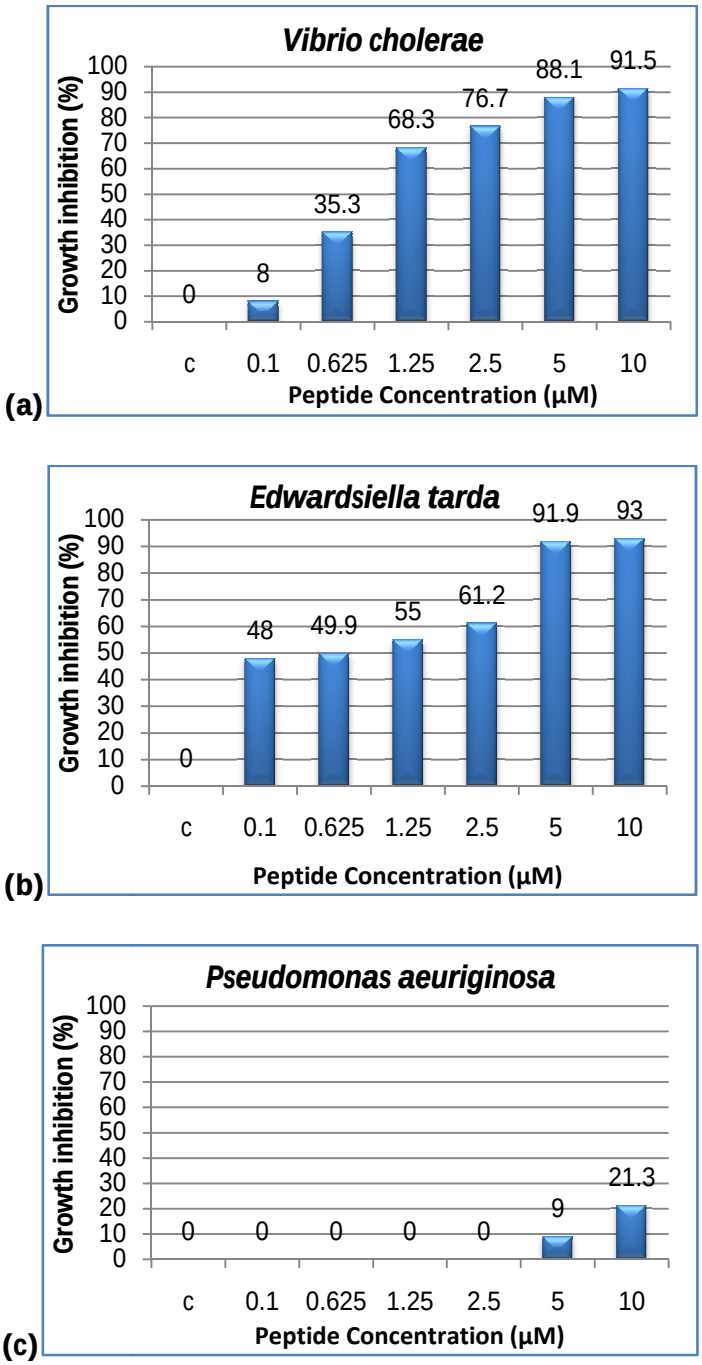


Fig. 5.8. (a-c) Antimicrobial activity of the recombinant ALF, rSSA measured by liquid growth inhibition assay at various concentrations against **(a)** *V. cholerae* **(b)** *E. tarda* **(c)** *P. aeruginosa.*

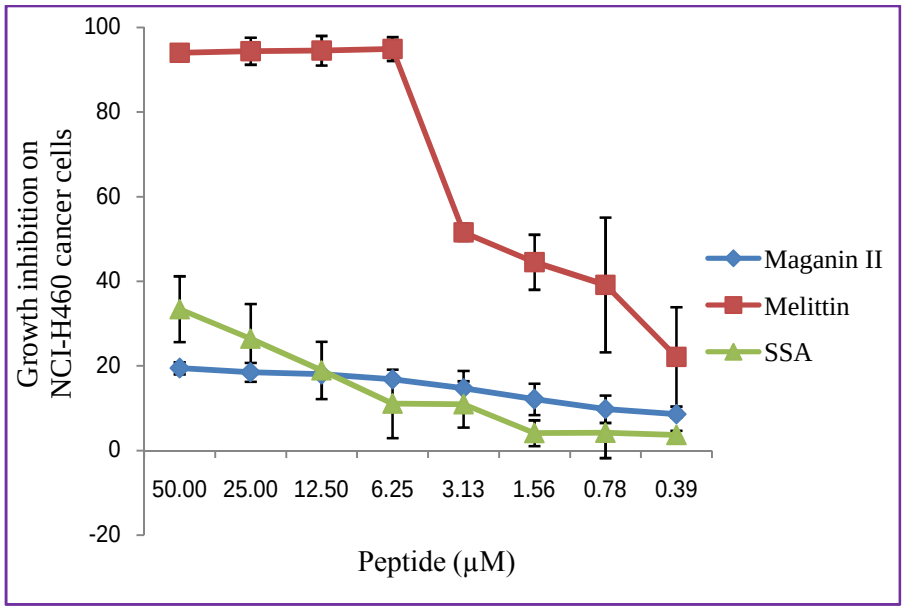


Fig. 5.9 Cytotoxicity Assay of the recombinant ALF, rSSA against NCI-H460 cells at various concentrations

Table 5.1 Details of primers used for the study

Primer	Sequence (5'-3')
SSA-F	catgccatgggccatcatcatcatcatcagtatgaagctctggtagc
SSA-R	aagctttcacccttcagccagg
T7-F	taatacgactcactataggg
SP6	gatttaggtgacactatag
T7-R	ctagttattgctcagcggtg

Table 5.2 Antimicrobial activity and MIC of recombinant ALF, rSSA against Gram-positive and Gram-negative bacteria as inferred from liquid growth inhibition assay and Priprobit software. MIC is expressed as the lowest concentration of the peptide that causes 100 % growth inhibition. NA: not active up to 10 μM.

Microorganisms used	MIC (μM)
Gram-positive bacteria	
Bacillus cereus	NA
Staphylococcus aureus	NA
Gram-negative bacteria	
Edwardsiella tarda	16.9
Vibrio cholerae	11.5
Vibrio parahemolyticus	NA
Pseudomonas aeuruginosa	77.1
Aeromoas hydrophila	NA

6

Structural and Functional Characterization of a Synthetic Anti-lipopolysaccharide Factor

6.1 Introduction

Developing new lead structures as drugs against multi-resistant bacteria possessing little toxicity towards the host organism is an urgent need of modern medicine. Recently, antimicrobial peptides (AMPs) have come to the forefront as potential antibiotic alternates due to their robust killing activity against a wide range of microbes, including drug-resistant strains. Due to the non-specific membrane-disturbing mode of action, high specificity for prokaryotes and low toxicity towards eukaryotic cells; cationic AMPs have become highly attractive as promising therapeutics to overcome bacterial resistance (Bax et al. 2000; Mor 2000; Hancock and Rozek 2002). It has been demonstrated that most of the potential AMPs contain cationic and hydrophobic residues in elevated proportion and is thus capable of interacting with microbial membranes through non-specific interactions with the membrane lipids (Arouri et al., 2009; Brandenburg et al., 2012; Seo et al., 2012). The short time-frame of interaction promotes the rapid death of microbes and also decreases the probability of resistance development by the microorganism (Fernebro, 2011).

Current research on AMPs is focused on exemplifying the biological and biomedical importance of these molecules. Important target organisms for demonstrating the biological activities of these peptide molecules including Gram-positive and Gram-negative bacteria along with many other different targets such as fungi, protozoa and enveloped viruses (Hancock and Diamond, 2000; Giuliani et al., 2007). There are also reports describing immunomodulatory activities for AMPs (Jerala and Porro, 2004; Bowdish et al., 2005; McPhee et al., 2005). In addition to these targets, there are studies that report a broad spectrum of anticancer activities by these peptides (Moore et al., 1994; Mader and Hoskin, 2006; Hoskin and Ramamoorthy, 2008; Berge et al., 2010).

Although AMPs have been essentially explored and developed as potential alternatives for fighting infectious diseases, their probable use as anticancer peptides in cancer therapy either alone or in combination with other conventional drugs has been regarded as a therapeutic strategy to explore. Cancer remains a major cause of death affecting millions of people and is triggered by the growth and spreading of abnormal cells in an uncontrolled manner. The worldwide statistics reveal that the most commonly diagnosed cancers include lung, breast and colorectal cancers (Parkin et al., 2005; Ferlay et al., 2010). Several research activities have been devoted in creating new therapeutics that are at the same time more selective and less harmful for cancer patients. In fact, as chemical agents are designed to attack the rapidly dividing cancer cells, they are expected to induce side effects on normal cells that divide at the same rate. In the current scenario of increasing cancer patients and where conventional therapies suffer from deficits and drawbacks, new treatment options are a mandate. In this context, AMPs, especially the anticancer peptides have

been proved to be a resourceful strategy for drug discovery and for developing molecular targets for cancer. Even though anticancer peptides are expected to be selective towards tumour cells without impairing the normal body cells and its functions, the development of a selective anticancer peptide has been a challenge. Also, it is not yet possible to predict antitumor activity based on its structures.

Despite the specificity and therapeutic advantages of AMPs over conventional antibiotics, there exist some constraints in producing and characterizing AMPs. Even though various multi-cellular organisms produce AMPs naturally, extraction procedures are time consuming and typically exhibit low yield. Besides, AMPs can be synthesized from individual amino acids; however, the production and purification costs for these compounds are quite high (Rotem and Mor, 2009). In order to better understand the mechanisms of action of AMPs, to reveal their structure–function relationships as well as to perform functional characterization leading to development of new therapeutics for widespread clinical use, huge amounts of AMPs with high purity and quality should be readily available in a cost-effective manner.

Recently, solid phase peptide synthesis is the widely used method to produce huge amounts of AMPs with high purity and quality. The major advantage of chemical synthesis is that it produces highly pure protein. Also, chemical synthesis is extremely useful in producing toxic peptides that are difficult to produce by biological expression systems. Another important application of chemical synthesis is that proteins can be labelled at specific sites depending on its application and also allows modification of the peptides with natural or unnatural amino acids. However, there are some disadvantages for chemical synthesis such as,

(1) it works well only for small proteins and peptides; (2) yield is often quite low with chemical synthesis and (3) the method is very expensive for longer polypeptides.

Anti-lipopolysaccharide factors (ALFs) are cationic AMPs which constitute one of the key effector molecules in innate immune system of crustaceans and are capable of binding and neutralizing lipopolysaccharides (LPS). ALFs are believed to possess broad spectrum of antimicrobial activities. A growing number of studies on crustacean ALFs have become available in recent years, describing primarily the molecular cloning, sequencing and expression analysis of ALF in various crustaceans (Supungul et al., 2004; Liu et al., 2005; Nagoshi et al., 2006; Somboonwiwat et al., 2006). ALFs, possess a 24 residue LPS binding domain in their mature peptide region that could recognize LPS, a major component of the Gram-negative bacterial cell wall. This LPS binding domain consisted of conserved amino acid residues exposed in the four-stranded β-sheet.

Though ALFs have been identified and characterized from Chelicerates such as *Limulus* and *Tachypleus*, and from decapod crustaceans including shrimps, prawns, crabs and lobsters (Tanaka et al., 1982; Gross et al., 2001; Supungul et al., 2002; Liu et al., 2006; Imjongjirak et al., 2007; Li et al., 2008; Beale et al., 2008; Rosa et al., 2008; Yedery and Reddy, 2009a; Tassanakajon et al., 2010; Yue et al., 2010); studies demonstrating the biological activity of synthetic ALF is very limited.

Studies on chemically synthesized ALF started with earlier works of Alpert and co-workers (1992) on ALFs present in *Limulus*. His work mainly focused on *in vivo* neutralization effect of the 22 amino acid residue (CHYRIKPTFRRLKWKYKGKFWC) present in *Limulus* ALF (*L*ALF).

Later in 1993, this 22 amino acid residue of *L*ALF was described as the potential binding site for LPS. Employing chemical synthesis, the immunological activity of this 22 amino acid long potential LPS binding site was further characterized by Vallespi and co-workers (2000). They also proposed that *L*ALF modifies LPS-induced response and possessed *in vitro* antiviral effects on Hep-2 cell lines and tumor necrosis factor. For the first time, studies by Vallespi et al. (2000) also demonstrated the anti-inflammatory properties of the ALF-derived peptides. The effect of this synthetic peptide on the cytokine gene expression and bacterial acute infection in mice was also studied (Vallespi et al., 2003). The results showed that the synthetic 22 residue peptide favoured host resistance against Gram-negative bacteria acute infection and increased the survival of challenged mice.

Another ALF isoform isolated from a Chelicerate, *Tachypleus,* named tachyplesin, was synthesized by solid phase process and its specific interactions with LPS has been carried out by Hirakura et al. (2002) and Hong et al. (2012). In 2007, Imura et al. studied the mechanism of action of synthetic tachyplesin synthesized by Fmoc-based solid phase process and also the effectsof PEGylation on its biological activity. It was found that both the peptides (with and without PEGylation) induced lipid flip-flop coupled to leakage and was translocated into the inner leaflet of the bilayer, indicating that tachyplesin forms a toroidal pore. The study also proved that PEGylation did not alter the basic mechanism of membrane permeabilization of the parent peptide. However, it was found that the peptides showed great differences in their biological activities. It was found that the process of PEGylation, though, reduced the antimicrobial

activity of tachyplesin, the process also significantly lowered its cytotoxic effects.

Studies on synthetic shrimp ALFs (*S*ALF) were first initiated by Pan and co-workers (2007). They synthesized 22 amino acid long ($SALF_{55-76}$ = ECKFTVKPYLKRFQVYYKGRMWCP) peptide in linear and cyclic forms by solid-phase procedure of Fmoc/tBu chemistry and characterized its biological activity. The analysis of the results showed that only cyclic peptide exhibited antibacterial activity against the tested organisms; whereas linear peptide did not show any activity against any of the bacteria tested except for *Listeria* and *Vibrio* with an MIC ranging from 25-100 μM. However, the cyclic synthetic peptide showed broad spectrum of antibacterial activity against almost all the bacteria tested with an MIC ranging from 12.5-100 μM and was also found to protect mice against challenge with *P. aeruginosa*. The study also reported tissue-specific up-regulation of interleukins, interferons, Toll-like receptors and tumor necrosis factors. This synthetic peptide also possessed broad cytotoxic activity against HeLa, HT-1080, and MCF-7 cells. Because of its multifunctional properties, Pan et al. (2007) suggested this 22 residue long peptide as a prophylactic agent for bacterial infectious diseases, as well as for septic shock.

In another work by Pan et al. (2010), *in vitro* activities of two synthetic ALF isoforms (linear and cyclic forms of $SALF_{55–76}$ = ECKFTVKPYLKRFQVYYKGRMWCP) of *P. monodon* against bacterial pathogens of duck were performed. In this work, the LPS binding domain was synthesized using a solid-phase procedure of Fmoc/tBu chemistry. Both linear and cyclic synthetic ALFs showed activity against the pathogens tested and the MIC values ranged between 6-25 μM.

Functional characterization of synthetic ALF (22 residue) from the freshwater prawn *M. rosenbergii* (*Mr*ALF$_{54-77}$ = TCQYSVNPKIKRFELYFKGRMWCP) was carried out by Arockiaraj et al. (2014). The peptide showed antimicrobial activity against both Gram-negative and Gram-positive bacteria. The bactericidal assay showed that the peptide recognized the LPS of bacterial cell walls and further binding to this layer helped in distinguishing the pathogens.

Li et al. (2014) synthesized another shrimp ALF peptide corresponding to the LPS-binding domain (CSFNVTPKFKRWQLYFRGRMWC) of *Fc*ALF (ALF from *F. chinensis*). The peptide contained modifications such as one flanking amino acid residue at the N and C terminal ends, inclusion of a disulfide bond between the two cysteine residues, and in case of basic ALF peptides, some amino acids were replaced with lysine or arginine residues. All peptides were acetylated at the N terminal and amidated in the C terminal. Results showed that peptides with modifications possessed strong antibacterial activity against Gram-positive bacteria *Micrococcus luteus* and *M. lysodeikticus* with MIC ranges of 2–4 μM and 1–2 μM respectively and also possessed significant inhibition activity against white spot syndrome virus (WSSV). Also, the biological activity of the synthetic peptide was found to be enhanced with modifications.

Another important work on *F. chinensis* ALF was by Guo et al. (2014). They designed and synthesized a peptide corresponding to the LPS-binding domain of ALF from the Chinese shrimp (*Fc*ALF). Modifications such as a flanking amino acid residue in both N and C terminal end, a disulfide bond between two cysteine residues, methylation of thiol group of cysteine residues (so as to prevent disulphide loop formation), replacement of certain amino acids by neutral amino acids, N-terminal

acetylation and C-terminal amidation were also performed. Results showed that modified peptides exhibited increased antibacterial activities against Gram-negative bacteria such as *E. coli* and *V. anguillarum* and also Gram-positive bacteria such as *M. luteus* and *M. lysodeikticus* with MIC ranging from 32–64, 2–4, 1–2, and 32–64 μM, respectively. The disulfide loop and the basic amino acids in the LPS-binding domain of ALF were found to play key roles in its antibacterial activities. In addition, the modified peptide could reduce the propagation of white spot syndrome virus (WSSV) *in vivo*. The study also proved that lysine residue is indispensable for its antiviral property.

Pioneer studies on synthetic crab ALFs were performed by Imongjirak and co-workers in 2007. They synthesized putative LPS binding domain of an ALF isoform present in *S. paramamosain* from positions 54 to 77 (TCHIRRRPKFRKFKLYHEGKFWCP) with a disulfide bond between the 2nd and 23rd cysteine residues. Modifications such as N-terminal residue acetylation and the C-terminal amidation were also included in the peptide. Synthesis of LPS-binding domain of *Sp*ALF was done by Fmoc (N-(9-fluorenyl) methoxycarbonyl) chemistry by the peptide synthesizer. The synthetic peptide revealed strong antimicrobial activity against several bacteria such as the Gram-positive *M. luteus* and Gram-negative *Vibrio harveyi* and the MIC values ranged from 0.1 to 50 μM. However, the peptide was not found to be active against certain Gram-negative bacteria tested such as *S. typhimurium*, *E. cloacae* and *K. pneumoniae* even at 50 μM concentrations.

Sharma et al. (2011) designed and synthesized another 24 amino acid – long linear and cyclic peptides based on the putative LPS binding domain of ALF from the mud crab *S. serrata* and has demonstrated its ability to

bind to LPS. The antibacterial activities of this synthetic peptide against vaginal pathogens were demonstrated by MIC, CFU and phagocytosis assays. The cyclic synthetic peptide was not found to be cytotoxic to human vaginal epithelial cells (HeLa-S3), macrophages and rabbit erythrocytes even at high concentration (64.64 μM).

As there exists only two reports regarding the demonstration of biological activity ALF present in crabs, the present chapter aims at structural and functional characterization of a synthetic ALF molecule present in crab. As explained in Chapter-2, four ALFs viz. SsALF1, SsALF2, PpALF1 and PpALF2 have been identified and characterized from crabs. PpALF1 being the most novel isoform among the four was selected for chemical synthesis and further demonstration of its biological activities. The objective of this chapter was to synthesize the LPS binding domain present in the mature peptide region of the AMP, PpALF1 and to functionally characterize the synthetic peptide with respect to its antibacterial activity against Gram-positive and Gram-negative bacteria. As detailed research on several anticancer peptides is crucial for optimizing drug development, in this chapter we have also tried to analyse the anticancer activity of the synthetic ALF, along with its cytotoxic effects.

6.2 Materials and Methods

6.2.1 Design, sequence analysis and synthesis of the target peptide for synthesis

Out of the nine AMPs characterized in the present study, the novel ALF isoform from *P. pelagicus* (PpALF1) was selected for synthesis and further functional characterization. The presence of LPS binding domain in

PpALF1 was identified from the multiple alignment results of ALF sequences from crustaceans using ClustalW program (http://www.ebi.ac.uk/Tools/ msa/clustalw2/) and was commercially synthesized at VCPBIO Ltd., China by standard solid phase synthesis method. The linear synthetic peptide, hereinafter denoted as PPA, was synthesized without any further modification and was subjected to structural and functional characterization.

6.2.2 Analysis of peptide characteristics

The physico-chemical properties and amino acid composition of the synthetic peptide were predicted and confirmed using PROTPARAM, ExPASy and APD programs (http://www.au.expasy.org/; http://aps.unmc.edu/ AP/main.php). The hydrophobicity (<H>) and hydrophobic moment (<μH>) were calculated by using HELIQUEST program (http:// heliquest.ipmc.cnrs.fr). To find out the amphipathicity of the synthetic peptide, PPA, a helical wheel representation of the peptide was also generated using the same program. The structural model of the synthetic peptide was also created using SWISS-MODEL server.

6.2.3 Determination of purity and mass

Purity of the synthetic peptide was confirmed by performing reverse-phase HPLC (RP-HPLC) on a C18 column (4.6 mm x 250 mm, Inertsil-ODS-SP), eluted at 1.0 ml/min with a 30–100 % step gradient of Solvent A (0.1 % trifluoroacetic in 100 % acetonitrile) and solvent B (0.1 % trifluoroacetic in 100 % H_2O) over 30 min. Absorbance was monitored at 220 nm. Also, the homogeneity, identity and mass of the synthetic peptide were assessed by MALDI-TOF mass spectrometry using a Thermo Finnigan LCQ Duo mass spectrometer with an electrospray source and

Xcaliber software. Sample was dissolved in 50 % (v/v) acetonitrile and 50 % of 0.1 % trifluoracetic acid (v/v) and analyzed by Electrospray Ionization (ESI). The spectra were obtained in the continuous acquisition mode, scanning from m/z 400 to 2000 at a scan time of 5 sec. Both RP-HPLC and MALDI-TOF MS analysis were performed by the peptide synthesizers.

6.2.4 Antimicrobial assay

Antimicrobial activity of the synthetic ALF, PPA was checked against both Gram-positive and Gram-negative bacteria employing liquid growth inhibition assay as described by Huang et al. (2006) with modifications. Six dilutions (ranging from 3 to 120 μM) of the recombinant peptides in triplicates were subjected to assay.

6.2.4.1 Microorganisms used

The microbial strains used to determine antimicrobial activity included two Gram-positive bacteria viz. *B. cereus* (MCCB 101) and *S. aureus* (MTCC 3061) and five Gram-negative bacteria viz. *E. tarda* (MTCC 2400), *P. aeruginosa* (MCCB 119), *A. hydrophila* (MCCB 113), *V. cholerae* (MCCB 129) and *V. parahaemolyticus* (MCCB 133). The details are as given in section 5.2.13.1 of Chapter 5.

6.2.4.2 Liquid growth inhibition assay

Antimicrobial activity was tested employing liquid growth inhibition assayas described in section 5.2.13.2 of Chapter 5. Six dilutions of the synthetic PPA ranging from 3-120 μM of the synthetic peptide PPA along with positive and negative controls in triplicates were subjected to antimicrobial assay.

6.2.4.3 Minimum inhibitory concentration

Minimum inhibitory concentration (MIC) of the synthetic PPA was determined as described in section 5.2.13.3 of Chapter 5.

6.2.5 Anticancer assay

The synthetic PPA was tested for its possible anticancer effect against 60 human cell lines at the National Cancer Institute (NCI), Maryland, US. Anticancer assay was performed as part of the *in vitro* Cell Line Screening Project (IVCLSP) of NCI, which is a dedicated service providing direct support to the anticancer drug discovery program implemented in 1990. The operation (NCI60) of this screen utilizes 60 different human tumor cell lines, representing leukemia, melanoma and cancers of the lung, colon, brain, ovary, breast, prostate, and kidney. The details of the 60 human cancer cell lines used for the assay are as given in Figure 6.9. As per the NCI instruction, the compound needs to be subjected to an initial virtual screening by *in silico* methods. The structure of the compound in *.mol* format will be subjected to virtual screening. Only compounds that give satisfactory results during *in silico* analysis will be accepted for the NCI60 programme. Following that the anticancer activity will be screened in a two-stage process. First the activity of 10 μM (single dose) of sample is evaluated against the 60 cell lines and then the samples which exhibit significant growth inhibition are further evaluated at five concentration levels against the 60 cell lines.

Following the NCI instructions, the structure of PPA in *.mol* format was send to NCI for *in silico* analysis (Fig. 6.1). Subsequently, obtaining satisfactory results for synthetic PPA (results not provided by NCI), 10-12 mg of the synthetic PPA was sent to NCI for conducting single dose assay

at 10 μM conc. Briefly, the cell lines were grown in RPMI 1640 medium containing 5 % fetal bovine serum and 2 mM L-glutamine. The cells were then inoculated into 96 well microtiter plates in 100 μL at plating densities ranging from 5,000 to 40,000 cells/well depending on the doubling time of individual cell lines. After inoculation, the microtiter plates were incubated at 37 °C, 5 % CO_2, 95 % air and 100 % relative humidity for 24 h. After 24 h incubation, two plates of each cell line were fixed *in situ* with trichloro acetic acid (TCA) so as to keep a record of the cell population for each cell line at the time of adding the sample (Tz). PPA was solubilized in dimethyl sulfoxide (DMSO) and stored frozen prior to use. At the time of drug addition, an aliquot of frozen concentrate was thawed and diluted to a concentration of 10 μM with complete medium containing 50 μg/ml gentamicin. Following addition of PPA, the plates were incubated for an additional 48 h at 37 °C, 5 % CO_2, 95 % air, and 100 % relative humidity. For adherent cells, the assay was terminated by the addition of cold TCA. Cells were then fixed *in situ* by the gentle addition of 50 μl of cold 50 % (w/v) TCA (final concentration, 10 % TCA) and incubated for 60 min at 4 °C. The supernatant was then discarded, and the plates were washed five times with tap water and air dried. Sulforhodamine B (SRB) solution (100 μl) at 0.4 % (w/v) in 1 % acetic acid was added to each well, and plates were incubated for 10 min at room temperature. After staining, unbound dye was removed by washing five times with 1 % acetic acid and the plates were air dried. Bound stain was subsequently solubilized with 10 mM trizma base, and the absorbance was read on an automated plate reader at a wavelength of 515 nm. The methodology followed for suspension cells was almost same, except that the assay is terminated by fixing settled cells at the bottom of

the wells by gently adding 50 µl of 80 % TCA (final concentration, 16 % TCA). Using the absorbance measurements [time zero, (Tz), control growth, (C), and test growth in the presence of sample (Ti)], the percentage growth was calculated as:

Growth % = [(Ti-Tz) / (C-Tz)] x 100

6.2.6 Cytotoxicity / Methyl thiazol tetrazolium (MTT) assay with NCI-H460 cells

Cytotoxicity assay was performed for synthetic PPA employing MTT assay on NCI-H460 cells as explained in the section 5.2.14. of Chapter 5.

6.3 Results

6.3.1 Design, sequence analysis and synthesis of the target peptide for synthesis

PpALF1, a novel ALF isoform, identified and characterized from *P. pelagicus* was selected for synthesis and for analyzing its biological activity. Amino acid sequences of known crustacean ALFs were aligned using the ClustalW program to identify the LPS binding region located between the conserved cysteine residues (Fig. 6.2). The mature peptide region of PpALF1, containing the active region i.e. the LPS binding domain corresponding to amino acids 54 to 77 of the mature peptide consisting of the sequence CCFLRRPPIIRIFLLHHEGFFWCA was commercially synthesized at VCPBIO Ltd., China by standard solid phase synthesis method. This synthetic LPS binding domain of PpALF1 is designated as PPA. The amino acid sequence comparison of the full length PpALF1 and the synthesized active LPS region (PPA) is as given in Fig. 6.3. The linear synthetic PPA was synthesized with a purity of 71.61 % without any

further modification and was subjected to structural and functional characterization. About 12.5 mg of the lyophilized synthetic peptide was supplied by the manufacturers.

6.3.2 Analysis of peptide characteristics

The physico-chemical properties of the synthetic peptide including its molecular weight, net charge, isoelectric point and hydrophobicity predicted using PROTPARAM, ExPASy, APD and HELIQUEST programs are described in Table 6.1. The synthetic peptide PPA possessed a length of 24 amino acids and a molecular weight of 2.975. The PPA was found to be highly cationic with a net charge of +2. The synthetic PPA possessed an isoelectric point (p*I*) of 8.74 and hydrophobicity of 62 % as predicted by the PROTPARAM and HELIQUEST programs. The synthetic PPA molecule also showed the conservation of two cysteine residues at positions Cys_2 and Cys_{23}, important for one disulfide bond (loop) formation in the peptide. The percentage of each amino acid of the synthetic PPA is given in Table 6.2. PPA was found to be rich in Phenyl alanine (16 %) which was followed by equal concentration of amino acids such as arginine, cysteine, leucine and isoleucine (12 % each).

The antimicrobial property of the synthetic ALF, PPA was predicted using APD and HELIQUEST programs. The results clearly indicated that the synthetic peptide taken for analysis possesses antimicrobial property. Among the residues taken for analysis, 62 % of amino acids of synthetic ALF, PPA were hydrophobic residues and its total net charge was determined to be +2, thereby confirming the antimicrobial nature of the peptide. Based on the prediction of APD program, PPA may form α-helices and possess at least five hydrophobic residues on the same hydrophobic

surface. The peptide was predicted to interact with membranes and possess a chance to be a potential AMP (result based on APD program http://aps.unmc.edu/AP/main.php). The α-helix wheel projections for PPA constructed using HELIQUEST program is given in Fig. 6.4. The analysis confirmed the peptide to be amphipathic in nature due to the presence of hydrophobic face. Analysis of results obtained from SWISS-MODEL server showed that PPA exists as anti parallel β-sheets (Fig. 6.5).

6.3.3 Determination of purity and mass

The purity and molecular weight of the synthetic peptide was confirmed by RP-HPLC and ESI-mass spectrometry. Purity of the synthetic peptide was checked by RP-HPLC using C18 column. The HPLC chromatogram of synthetic ALF, PPA is as shown in Fig. 6.6. The retention time of the synthetic PPA was found to be 12.86 min. The molecular weight of synthetic PPA (2.976 kDa) was verified by ESI mass spectroscopy. The ESI mass spectrum of the synthetic PPA is as shown in Fig. 6.7. The mass spectrum shows the mass to charge ratio (m/z) from 400 to 2000 of all the ionized molecules present in the sample. The most abundant ion in the spectrum is seen at a mass to charge ratio of 744.95. This corresponds to PPA ionized to +4 (rounded off MW = 2976 + 4H+ = 2980). The mass to charge ratio is thus 2980/4 = 745. Two other relatively abundant ions attributed to PPA were m/z of 596.20 ionized to +5 (MW = 2976 Da + 5H+ = 2981; 2981/5 = 596.2) and 992.5 (MW = 2976 + 3 = 2979; 2979/3 = 993).

6.3.4 Antimicrobial assay

PPA was tested at concentrations ranging from 3 to 120 μM concentrations against a group of Gram-positive and Gram-negative

bacterial strains employing liquid growth inhibition assay. The antibacterial activity of PPA, expressed as minimum inhibitory concentrations (MIC) against two Gram-positive bacteria and five Gram-negative bacteria tested is listed in Table 6.3. Analysis of the results showed that, PPA possessed activity against both Gram-positive and Gram-negative bacteria. However, all bacterial strains tested did not show susceptibility to the antibacterial action of PPA. In particular, the peptide was found to be highly active against Gram-positive *S. aureus* (MIC=124.4 μM), followed by Gram-negative *V. cholerae* (MIC=147.4 μM) and *V. parahaemolyticus* (MIC=237.2 μM), but less active against *P. aeruginosa* (MIC=1.04 mM), *E. tarda* (MIC=2.09 mM) and *A. hydrophila* (MIC_{50}=31.5 mM). PPA was not found to possess any activity against the Gram-positive bacteria *B. cereus*. The antibacterial activity of synthetic ALF, PPA against the tested bacteria at various concentrations is illustrated in Fig. 6.8 (a-f).

6.3.5 Anticancer assay

Screening for anticancer activity was performed at the National Cancer Institute (NCI), Maryland, US against 60 human cancer cell lines. The tested cancer cell lines included seven colon cancer cell lines, seven renal cancer cell lines, eight non-small cell lung cancer cell lines, six CNS cancer cell lines, nine melanoma cell lines, six leukemia cell lines, seven ovarian cancer cell lines, six breast cancer cell lines and two prostate cancer cell lines. The growth % of cancer cell lines against 10 μM of synthetic peptide, PPA was determined. Analysis of the results received from NCI revealed that 10 μM concentration of synthetic PPA could cause 12 % growth reduction against a colon cancer cell line, HCT-116. Other cancer cell lines inhibited by 10 μM PPA at a minor level includes 8.24 % growth reduction against non-small cell lung cancer cell line, HOP-62; 7.28 % growth

reduction against non-small cell lung cancer cell line, A549; 5.05 % growth reduction against CNS cancer cell line, U251; 5.25 % growth reduction against ovarian cancer cell line, IGROV1 and 4.95 % growth reduction against breast cancer cell line, MDA-MB-231. The detailed result of screening of PPA against 60 human cancer cell lines as received from NCI is given in Fig. 6.9.

6.3.6 Cytotoxicity / Methyl thiazol tetrazolium (MTT) assay with NCI-H460 cells

The cytotoxicity assay of PPA was performed against NCI-H460 cells employing MTT assay. This assay measures the reducing potential of the cell using a colorimetric reaction. Viable cells will reduce the MTT reagent to a coloured formazan product. Analysis of the results proved that the synthetic PPA is non-cytotoxic to the tested cells when compared to that of the two control AMPs used. The control melittin, was found to be highly cytotoxic even at smaller concentrations. A reduction of < 95 % was observed for NCI-H460 cells even at a smaller tested concentration of 6.25 μM in case of melittin. The IC_{50} of melittin against NCI-H460 cells was estimated to be ~ 3.3 μM. Whereas, in the case of PPA, even at higher concentrations,the peptide did not exhibit any inhibition of growth of NCI-H460 cells. The result of cytotoxic assay against NCI-H460 cells is represented in Fig. 6.10.

6.4 Discussion

The extensive clinical use of classical antibiotics has led to the growing emergence of many medically relevant resistant pathogen strains. Antibiotic resistance was known even before the widespread manufacturing of antibiotics, starting with penicillin, but the pace of

antibiotic resistance seems to have hastened. Therefore, the development of a new class of antimicrobials with a different mechanism of action other than the present conventional antibiotics has become critical. The cationic AMPs represent such a new class of antimicrobials (Hancock 1997; Andreu and Rivas 1998; Sitaram and Nagara, 1999). The development of resistance to membrane active peptides whose sole target is the cytoplasmic membrane is not expected as this requires substantial changes in the lipid composition of the microbial cell membranes (Hancock 2001). ALFs have been described as highly cationic peptides with a broad spectrum of antimicrobial activities. So far, ALFs have been reported and characterized from only two groups of organisms viz. marine chelicerates and crustaceans.

Out of the nine AMPs identified and characterized in the present study, an AMP belonging to ALF family identified and characterized from the crab, *P. pelagicus* was selected for chemical synthesis and for demonstrating its biological activity. The active regions of ALF was selected because of the following reasons: (a) Structural and functional characterization of synthetic crustins and penaeidins have been studied in detail by several researchers (Destoumieux et al., 1997, 1999; Cuthbertson et al., 2004, 2005, 2006; Li et al., 2005; Kang et al., 2007). (b) In case of penaeidins, the active region to be synthesized is a 52 amino acid long region; and the active region of crustin is comprised of 45 amino acid long WAP domain. Hence, it was not economically feasible to synthesize longer peptides corresponding to the active regions of crustins and penaeidins.

AMP belonging to ALF family identified and characterized in the present study (PpALF1) was selected for chemical synthesis and further characterization, because of the fact that there are hardly any studies on

synthetic ALFs especially from crabs. The only publication on chemical synthesis and functional characterization of a crab ALF is by Imjongjirak and co-workers (2007) of an ALF isoform from *S. paramamosain* and by Sharma et al. (2011) in *S. serrata*.

ALFs, possess a 24 residue LPS binding domain in their mature peptide region that could recognize LPS, a major component of the Gram-negative bacterial cell wall. This LPS-binding domain has been demonstrated as pivotal for the antibacterial activity of ALF molecules (Supungul et al., 2004; Nagoshi et al., 2006; Imjongjirak et al., 2007; de la Vega et al., 2008; Rosa et al., 2013). LPS binding domain consists of conserved amino acid residues exposed in the four-stranded β-sheet. An *in vitro* synthesized peptide of LPS-binding domain has been demonstrated to interact with LPS (Nagoshi et al., 2006).

The present study aims at designing, synthesis and evaluation of the LPS binding domain of ALF family of AMPs for potential use in developing new class therapeutics. The novel ALF isoform from *P. pelagicus* reported in the present study, which showed only 46 % similarity to already reported ALF isoforms, viz. PpALF1 was selected for chemical synthesis and its further structural and functional characterization. PpALF1 is a 123 amino acid long ALF isoform possessing a molecular weight of 13.86 kDa. In the present study, the efficacy of amino acids corresponding to amino acids 54 to 77 of the mature peptide region, corresponding to the LPS binding domain of PpALF1 was investigated. This 24 amino acid residue long LPS binding domain, designated as PPA, was synthesized at VCPBIO Ltd., China.

The synthetic peptide possessed the sequence: CCFLRRPPIIRIFLLHHEGFFWCA. Synthetic ALF, PPA was found to be a cationic peptide mainly enriched by residues such as phenyl alanine, arginine, cysteine, leucine, isoleucine, with a p*I* of 8.74 and a net charge of +2 at pH 7.0. PPA was synthesized as a linear peptide without any further modification. About 12.5 mg of PPA was synthesized with a purity of 71.61 %, so that enough concentration of the peptide will be available for the study and is also economically feasible. The purity and MW was confirmed by HPLC and MS.

Its physicochemical parameters were in agreement with the required properties of an AMP. The synthetic PPA possessed a molecular weight of 2.975 kDa and was found to be highly cationic with a net charge of +2 and 62 % hydrophobicity. The peptide structure consisted of two β-strands which are in turn linked by a disulfide bond to form an amphipathic loop rich in cationic amino acid side chains. The disulfide bond (loop) formation in the peptide is effected by two conserved cysteine residues at positions Cys_2 and Cys_{23}. Some earlier researchers have synthesized cyclic peptides by introducing a disulphide loop between these cysteine residues. According to analysis of the results the cyclic peptides are found to be more active than the linear peptides synthesized. The antimicrobial property of PPA predicted using APD and HELIQUEST programs also clearly indicated the possible potential antimicrobial property of the linear synthetic PPA.

Functional characterization including the demonstration of the antimicrobial and anticancer activity as well as the cytotoxicity of the synthetic PPA were carried out in the present study. Analysis of the results showed that the synthetic ALF, PPA, exhibited both antibacterial

activity as well as anti-cancer activity to some extent, without causing any cytotoxicity to human cells.

The antimicrobial activity of PPA was tested by liquid growth inhibition assay (MIC) against seven bacteria including two Gram-positive and five Gram-negative bacteria. The antimicrobial activity of PPA was tested at various concentrations ranging from 3 to 120 μM. The peptide was found to be equally active against Gram-positive and Gram-negative bacteria with the exception of Gram-positive *B. cereus*, which was found to be resistant to PPA. The highest activity was shown against important pathogens such as Gram-positive *S. aureus* and the Gram-negative *V. cholerae* with an MIC ranging from 124.4-147.4 μM. This was followed by Gram-negative *V. parahaemolyticus*, *P. aeruginosa, E. tarda* and *A. hydrophila*. The antimicrobial activity showed by the ALF, PPA against *S. aureus* and *V. cholerae* strains is very interesting considering the fact that both are clinically important pathogens. *S. aureus* is one of the most common intestinal tract pathogenic bacteria in both animals and humans. Moreover, *S. aureus* is also involved in food spoilage and an application of PPA in the food processing may also be possible, along with application in clinical or veterinary health. Previous studies on *S. aureus* have shown that it is very resistant to AMPs found in human beings, such as defensins (Peschel et al., 1999). Hence, new antimicrobial agents that are effective against s*taphylococci* and vibrios to treat infection are needed in the present scenario.

Previous work on synthetic crab ALF from *S. paramamosain* (*Sp*ALF) by Imjongjirak et al. (2007) has also showed high antimicrobial activity against a wide range of Gram-positive and Gram-negative bacteria. In his work he has used five gram-positive bacteria such as *S. aureus, S.*

haemolyticus, *A. viridans*, *B. megaterium* and *M. luteus* and five Gram-negative bacteria such as *S. typhimurium*, *E. coli 363*, *Enterobacter cloacae*, *V. harveyi* and *K. pneumonia*. Analysis of the results has showed that the synthetic LPS region of *Sp*ALF was found to be more active against Gram-positive bacteria especially *S. haemolyticus*, *A. viridans*, *B. megaterium* and *M. luteus* with an MIC ranging from 1-12 μM and was also found to be active against *S. aureus* possessing an MIC of 50 μM. In the present study the MIC calculated for *S. aureus* was found to be 124.4 μM.

Antibacterial activity of the synthetic LPS domain of *S. serrata* (*Ss*ALF) has been studied by Sharma et al. (2011). In his study the antibacterial activity of the synthetic LPS region of *Ss*ALF was tested against two Gram-positive such as *S. aureus* and *S. pyogenes* and two Gram-negative bacteria such as *E. coli* and *P. aeuriginosa*. Analysis of the MICs showed that the synthetic peptide was more active against Gram-negative bacteria (MIC= 32-64 μM) than against Gram-positive bacteria (MIC > 65 μM). This is in agreement with results of the present study as in this study also the synthetic PPA was found be more active against Gram-negative when compared to that of Gram-positive bacteria. The MIC of *S. aureus* is also in agreement with the results obtained for that of synthetic *Ss*ALF (Sharma et al., 2011).

Generally AMPs are known to exhibit antibacterial properties via their interaction with the bacterial cytoplasmic membranes and this property is attributed to the presence of hydrophobic and cationic residues which are able to bind to negatively charged phospholipids such as phosphatidylcholine and phosphatidylinositol, highly present in the bacterial membranes (Hancock and Chapple, 1999). These interactions are known to cause destabilization of microbial membranes, including

pore formation, thereby resulting in the death of the target cells (Shai, 2002). This phenomenon has been extensively studied and proven using artificial membrane models / liposomes as well as in the bacteria and yeast cells (Lindgren et al., 2000).

According to previous literature, both Gram-positive and Gram-negative bacteria have their own unique response system to AMPs (Gunn and Miller, 1996; Moskowitz et al., 2004; Li et al., 2007). The majority of AMPs act mainly on the membrane of pathogenic bacteria, by increasing their permeability to ions and other metabolites to potentially lethal levels through a detergent-like effect or pore formation. *Limulus* ALF has shown to bind to LPS layer of bacteria and possess a strong antibacterial activity, particularly on the growth of Gram-negative bacteria (Aketagawa et al., 1986; Muta et al., 1987) and has also been found to be effective in protecting animals from an LPS challenge and/or Gram-negative bacterial infection (Warren et al., 1992).

The positively charged amino acid, arginine, with its cationic nature, is believed to provide means for electrostatic interaction with target membranes and has the ability to form hydrogen bonds with negatively charged parts of the bacterial membrane such as LPS, teichoic acid, or phosphatidyl group. Also, when tryptophan is also present, arginine will interact with tryptophan, thereby allowing an easier penetration into the lipid bilayer (Chan et al., 2006). The ALF, PPA possesses high concentration of arginine (12.5 %) which could be attributed to its antimicrobial activity. Along with that the presence of 4.2 % tryptophan adds up to its activity. Chan et al. (2006) have also proposed many interactions between cationic amino acids and aromatic residues, which was also present in high concentration in PPA i.e. phenyl alanine (16.7 %),

which further improves the activity of PPA. The structure of PPA with an extended β-conformation pointing in opposite directions also helps to form a cationic amphipathic loop which adds up to its activity. A similar structural conformation was observed for the *Ss*ALF peptide when analyzed on the AMP database (Sharma et al., 2011). Also, studies have suggested that the interaction between LPS and the amphipathic loop of cyclic ALF peptides are of an electrostatic nature whereby the positive charges of the peptides are assumed to bind to the negatively charged groups of LPS (phosphates and carboxylates) (Andraa et al., 2004).

The antibacterial properties of PPA could be attributed to their high net positive charge of +2 and a hydrophobic ratio of 62 %. The cation–π interactions that are possible between all aromatic amino acids and arginine/ lysine and combination of electrostatic and hydrophobic interactions with membrane components, may be the reason behind PPA's exquisite ability to select between bacterial-like and eukaryotic-like membranes. The antibacterial activity of PPA can also be attributed to the presence of alternating series of positively charged and hydrophobic residues.

Though the exact mode of action of PPA is not demonstrated in the present study, it could be speculated that, as for other cationic non-amphipathic AMPs, the positive charges carried by the peptide are essential for membrane binding through electrostatic interaction between residues with anionic phospholipids. The peptide-membrane interaction could be responsible for membrane aggregation by peptides and further disruption of the membrane, resulting in the death of the microbes (Lamaziere et al., 2007).

There have been reports regarding the effects of specific amino acids and its relation with the antibacterial activity of synthetic peptides (Guo et al., 2014). The basic amino acids along with the disulphide loop have been found essential for the antibacterial activities of the LPS-binding domain. However, for demonstrating anti-WSSV property, only the lysine residue is indispensable.

The first attempt to testify the effects of the sequence features of the LPS-binding domain on its antimicrobial activities was done by Guo et al., 2014. They have demonstrated anti-WSSV property of shrimp ALF (*Fc*ALF) with modified synthetic peptide. According to Guo et al. (2014), replacement specific amino acids in the LPS domain with lysine residues is indispensable for its antiviral property, which doesn't showed the antviral activity otherwise. Addition of modifications such as disulfide loop between 2nd and 23rd cysteine residues, N-terminal acetylation, C-terminal amidation, replacement of amino acids in the LPS binding domain has proved to show key roles in its antibacterial and / or other biological activities (Guo et al., 2014). The synthetic LPS domain peptide of *Pm*ALF has showed protection to Hpt cell cultures from WSSV infection (Tharntada et al., 2009). In a previous work by Guo et al., 2014, synthetic LPS of *F. chinensis* (*Fc*ALF) were found to reduce its activity potential against Gram-positive and Gram-negative bacteria when the disulfide bond was broken up. Similar results have been reported by Pan et al. (2007), where linear peptide (*Pm*ALF) lost its antibacterial activity for most tested bacteria as compared to the cyclic peptide. As a result, the synthetic *Sp*ALF peptide, which consisted of 24 amino acid residues in the predicted LPS binding region exhibited potent antimicrobial activity

against Gram-positive and Gram-negative bacteria. The result suggested the antimicrobial activity of *Sp*ALF-derived peptide.

Screening for anticancer activity was performed at the National Cancer Institute (NCI), Maryland, US against 60 human cancer cell lines. Analysis of the results received from NCI revealed that even at 10 μM concentration, synthetic ALF, PPA could cause 12 % growth reduction against a colon cancer cell line, HCT-116 followed by 8.24 % growth reduction against non-small cell lung cancer cell line, HOP-62; 7.28 % growth reduction against non-small cell lung cancer cell line, A549; 5.05 % growth reduction against CNS cancer cell line, U251; 5.25 % growth reduction against ovarian cancer cell line, IGROV1 and 4.95 % growth reduction against breast cancer cell line, MDA-MB-231.

The outer layer of mammalian cells is mainly composed of the zwitterionic phospholipids viz. phosphatidyl choline and sphingomyelin along with cholesterol (Turner and Rouser 1970; Verkleij et al. 1973). Whereas, the bacterial membranes contain mainly anionic phospholipids and lack cholesterol (Brock 1974). Also, the outer surfaces of Gram-negative bacteria contain lipopolysaccharides, while those of Gram-positive bacteria contain teichoic acid, which in both cases add to the negative charge of the bacterial surface (Brock 1974). The cationic nature of the synthetic ALF, PPA clearly contributes to their preferential recognition by the negatively charged outer surfaces of bacterial membranes (Oren and Shai 1998; Shai 1999). Since AMPs are generally cationic and amphipathic in nature it can bind to anionic constituents of cells and cause membrane permiabilization. Because of these properties some AMPs are known to cause harmful side effects, such as toxicity to cells at peptide concentrations demonstrating inhibition of bacterial

growth (Reddy et al., 2004). Hence a cytotoxicity assay on NCI-H460 cells was performed employing MTT assay. The assay also two controls viz. melittin and magainin which are also AMPs possessing broad spectrum of activity. Melittin is an AMP isolated from honey bee and magainin from frog. Analysis of the results showed that even at a high concentration of 50 μM PPA was not found to be cytotoxic to NCI-H460 cells. At the same time, melittin was found to be highly cytotoxic even at very low conc of < 1 μM. Magainin was found tobe less cytotoxic. However exhibited killing properties of 20 % of cells > 6 μM range. This result is in agreement with the cytotoxic effect of synthetic LPS of *S. serrata* (Sharma et al., 2011).

The cytotoxicity assay of PPA was performed against NCI-H460 cells employing MTT assay. Analysis of the results proved that the synthetic ALF, PPA is non-cytotoxic to the tested cells even at higher concentrations. Even at higher concentrations PPA was not inhibiting the growth of NCI-H460 cells.

This higher range of activity of the synthetic peptides of previous works might be because of the peptide modifications they have performed such as introducing disulfide bond between the 2nd and 23rd cysteine residues, N-terminal acetylation, C-terminal amidation and also the purity of the synthetic peptide tested in previous cases was ~ 95 %. Comparison of our linear peptide without modification at 76 % purity with the other synthetic peptides itself explains the reason for reduced activity of PPA. Also, cyclization of PPA by introducing the disulphide loop will surely increase the antibacterial potential of PPA against wide range of bacteria. It is interesting and quite promising to note that even without any of the modifications, the 24 amino acid long synthetic PPA showed activity

against both Gram-positive and Gram-negative bacteria and also inhibition to growth of cancer cell lines at least to some level.

The synthetic PPA containing putative LPS binding domain revealed antimicrobial activity against both Gram-positive and Gram-negative bacteria especially on the growth of *S. aureus* and *V. cholerae*, which suggested that PpALF1 could play an essential role in defence mechanism in *P. pelagicus*.

Finally, the tested synthetic peptide is a good starting point to design new synthetic derivatives with modified physico-chemical properties, so as to enhance its antimicrobial activity against pathogens or against cancerous cell lines and thereby improving its pharmaceutical potential (Huang et al. 2010; Brogden and Brogden 2011).The present study also proves the importance of disulphide bond in the biological activity of ALFs.

In conclusion, PPA characterized in this study, fulfill important requirements for its use as therapeutic agent to prevent or treat bacterial infections and cancer to a lesser level. The synthetic PPA exhibiting a bactericidal activity towards six bacterial species is reported in this study. With some modifications in the peptide such as introduction of disulphide bond, N-terminal acetylation, C-terminal amidation, cyclization, antimicrobial activities and anticancer activity of PPA could be enhanced for potential applications in aquaculture / medicine.

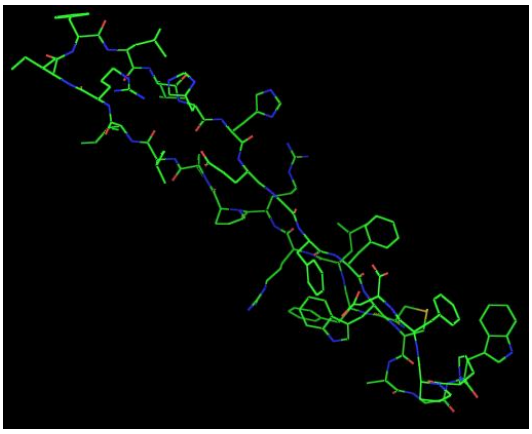


Fig. 6.1 The 2-D structure of synthetic ALF, PPA in .mol format sent to NCI for *in silico* analysis of possible anticancer activity

Fig. 6.2 Multiple alignment of amino acid sequences of PpALF1 with other crustacean ALFs performed to identify the LPS binding region. LPS binding domain between the conserved cysteine residues has been indicated. Black and grey indicates conserved sequences.

- **AMINO ACID SEQUENCE of full length PpALF1**

  MKKGVAALLRLLLVMRLLLPPPSCDPHDDPLASSLLEILAGRCHNSSDDLVGP CCFLRRPPIIRIFLLHHEGFFWCAGWPPFDGSTRTNRTAASSEDATDDLRRAF LHKLLIPQDDAPRRMNN*

- **AMINO ACID SEQUENCE of synthetic peptide, PPA**

  CCFLRRPPIIRIFLLHHEGFFWCA

Fig.6.3 Amino acid sequence comparison of full length PpALF1 and the synthetic ALF, PPA. The amino acid sequence (54 to 77) corresponding to the LPS binding domain has been underlined.

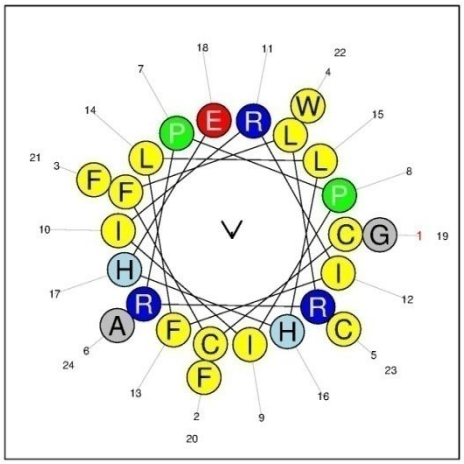


Fig 6.4 The helical wheel diagram of PPA predicted using HELIQUEST tool. The structure was constructed to find out the amphipathicity of the synthetic ALF, PPA. The amino and carboxy terminal ends are mentioned as N and C, respectively. Residues are numbered consecutively from the N terminus to the C terminus, with hydrophobic residues shown in yellow.

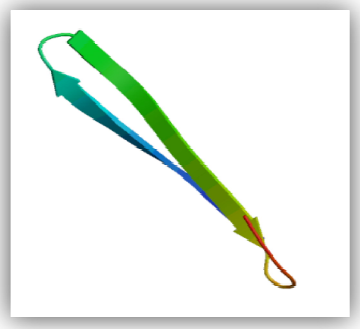


Fig. 6.5 Structural model of synthetic ALF, PPA created using SWISS-Model server depicting anti parallel β-sheets

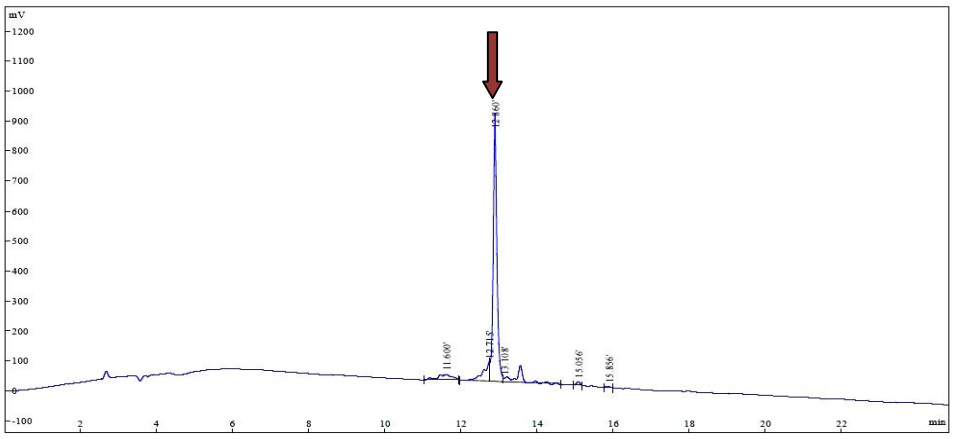


Fig. 6.6 RP-HPLC chromatogram of synthetic ALF, PPA. The retention time of the synthetic PPA, 12.86 min, has been marked with an arrow.

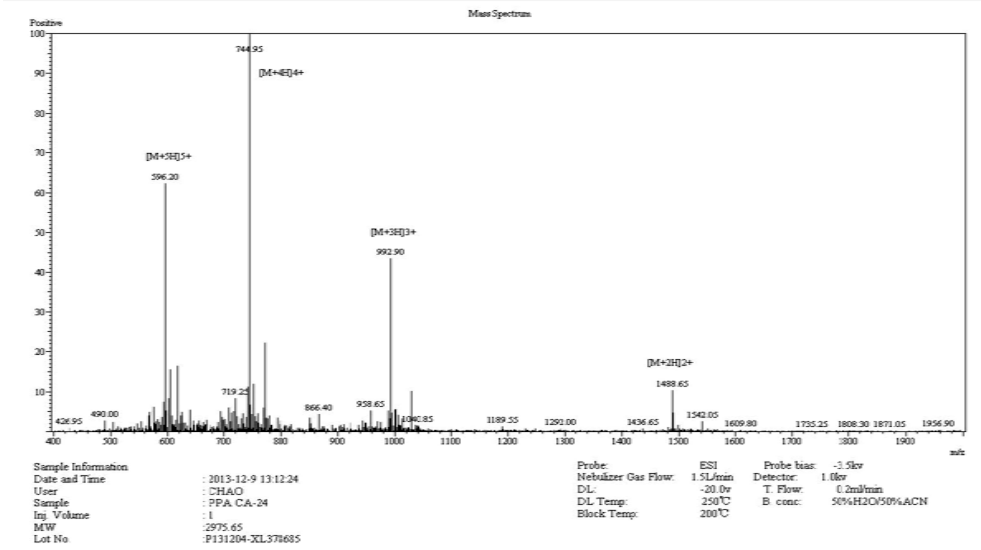


Fig. 6.7 ESI Mass spectrum of synthetic ALF, PPA demonstrating its molecular weight. Most abundant ion in spectrum is seen at m/z of 744.95 [M+4H]4+ followed by 596.20 [M+5H]5+ and 992.50 [M+3H]3+.

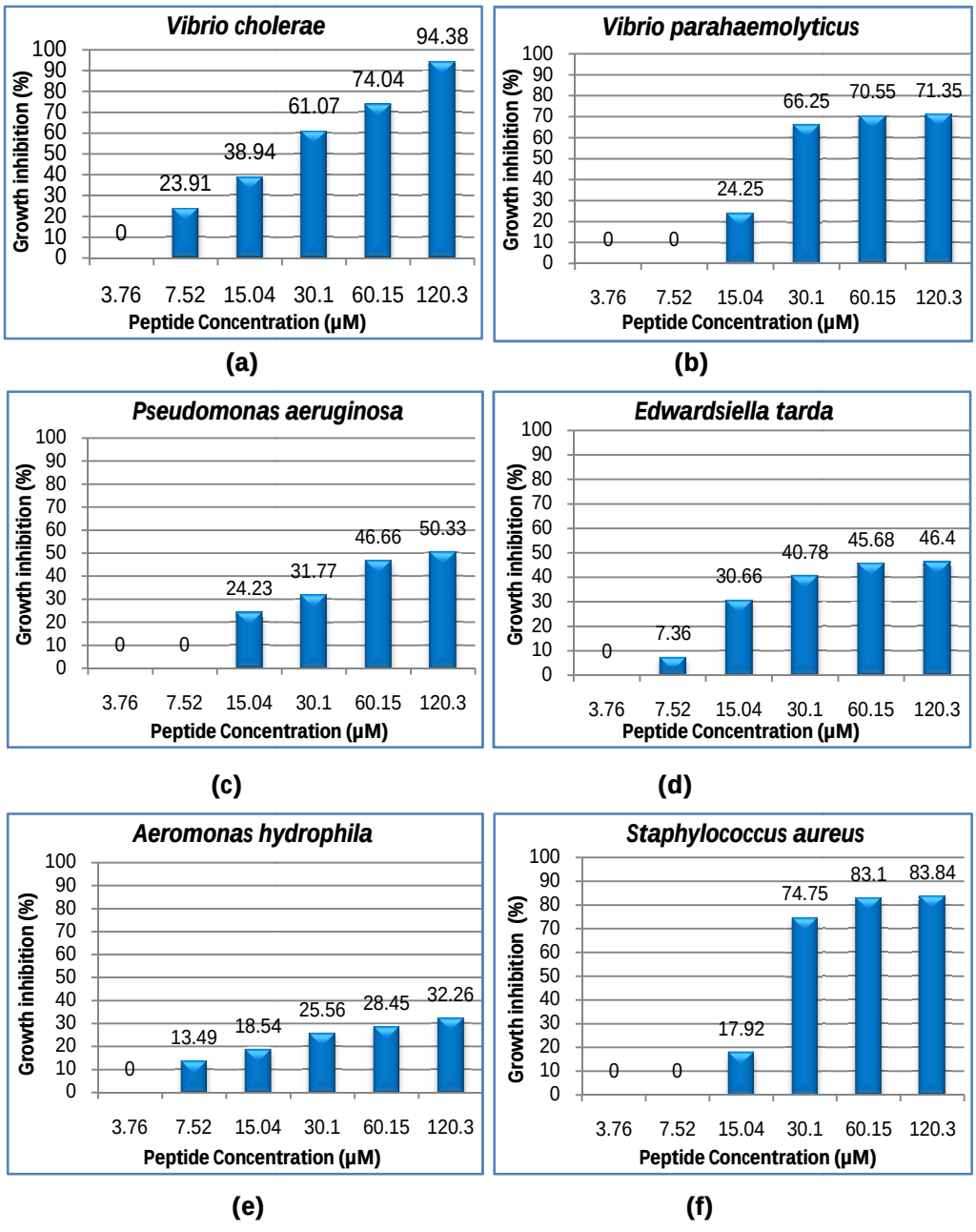


Fig. 6.8.(a-f) Antibacterial activity of synthetic ALF, PPA measured by liquid growth inhibition assay at various concentrations against **(a)** *V. cholerae* **(b)** *V. parahaemolyticus* **(c)** *P. aeruginosa* **(d)** *E. tarda* **(e)** *A. hydrophila* **(f)** *S. aureus.*

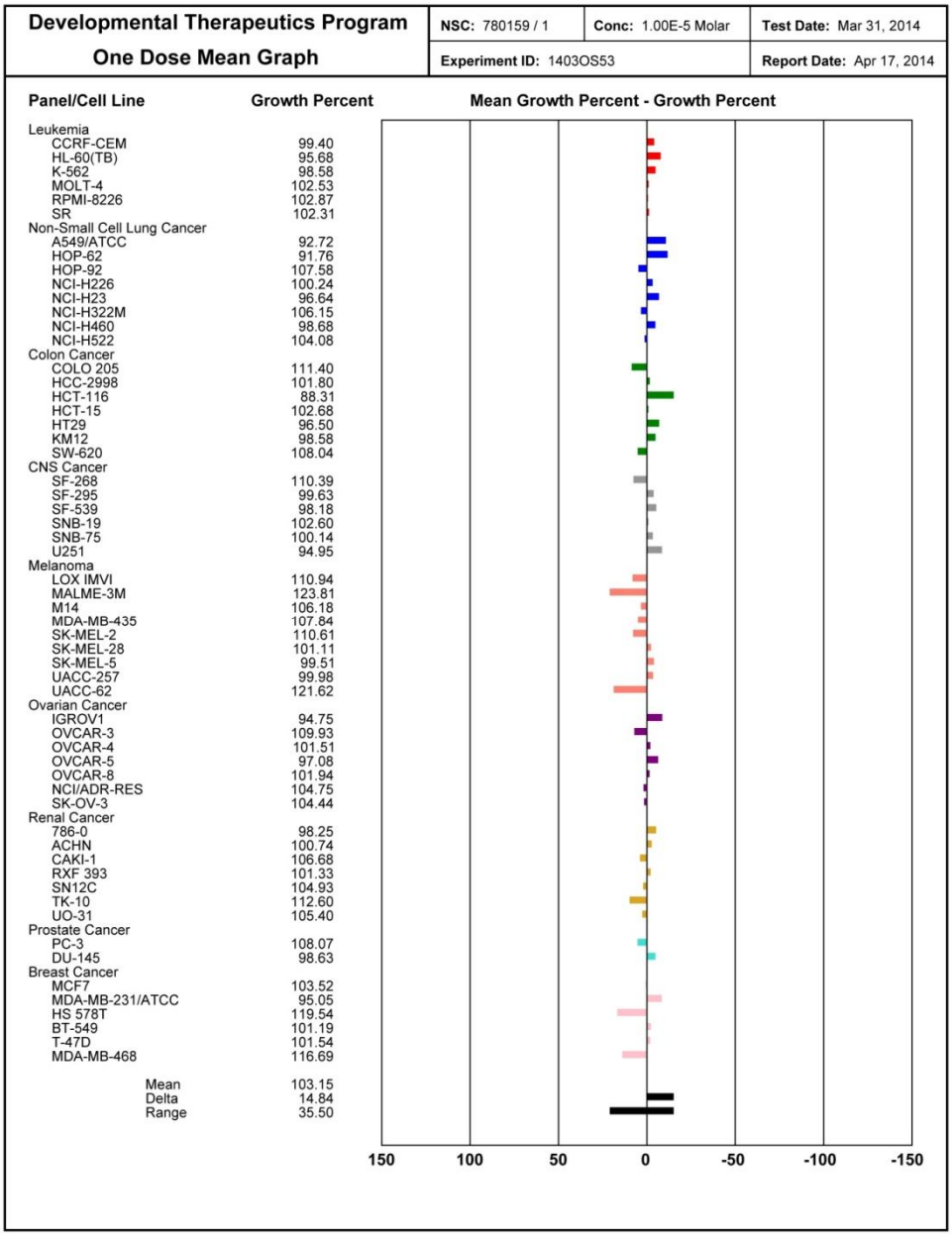


Fig. 6.9 Results of anticancer assay of the synthetic ALF, PPA, performed against 60 human cancer cell lines as provided by NCI, USA. Growth % of tested 60 human cancer cell lines against synthetic ALF, PPA at a concentration of 10 μM.

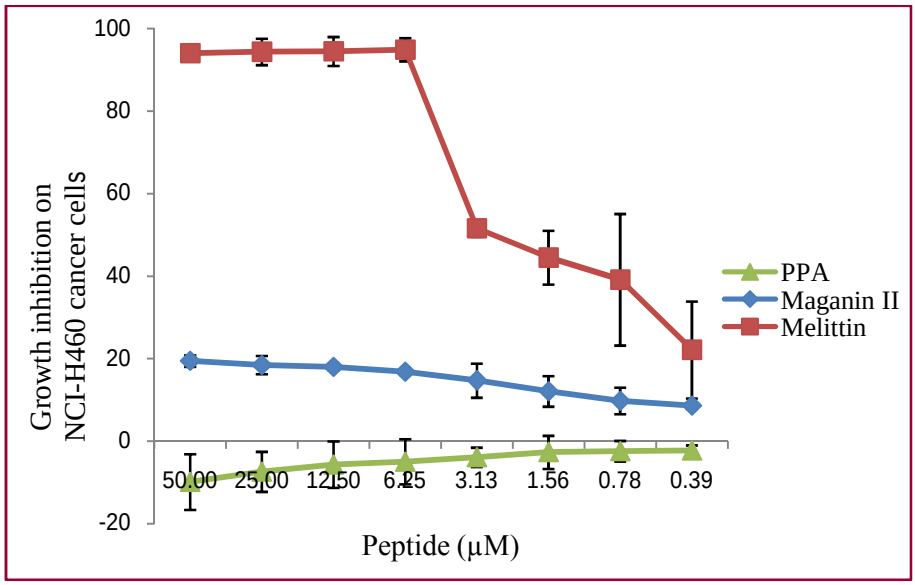


Fig. 6.10 Cytotoxicity assay of the synthetic ALF, PPA against NCI-H460 cells at various tested concentrations

Table 6.1 Physico chemical parameters of the synthetic ALF, PPA used in the study.

Parameters	Results
Sequence	CCFLRRPPIIRIFLLHHEGFFWCA
Length (No. of amino acids)	24 aa
Net Charge	+2
Molecular weight	2.975 kDa
Isoelectric point (*p*I)	8.74
Hydrophobicity	62 %
Molecular formulae	C142H209N38O27S3

Table 6.2 Percentage of each amino acid present in synthetic ALF, PPA

The percentage of each amino acid:
The Ile ratio = 12 %
The Val ratio = 0 %
The Leu ratio = 12 %
The Phe ratio = 16 %
The Cys ratio = 12 %
The Met ratio = 0 %
The Ala ratio = 4 %
The Trp ratio = 4 %
The Gly ratio = 4 %
The Pro ratio = 8 %
The Thr ratio = 0 %
The Ser ratio = 0 %
The Tyr ratio = 0 %
The Gln ratio = 0 %
The Asn ratio = 0 %
The Glu ratio = 4 %
The Asp ratio = 0 %
The His ratio = 8 %
The Lys ratio = 0 %
The Arg ratio = 12 %

Table 6.3 Antimicrobial activity and MIC of the synthetic ALF, PPA against Gram-positive and Gram-negative bacteria as inferred from liquid growth inhibition assay and Priprobit software. MIC is expressed as the lowest concentration of the peptide that causes 100 % growth inhibition. NA: not active up to 120 μM.

Microorganisms used	MIC (μM)
Gram-positive bacteria	
Bacillus cereus	NA
Staphylococcus aureus	124.4 μM
Gram-negative bacteria	
Edwardsiella tarda	2.09 mM
Pseudomonas aeruginosa	1.04 mM
Aeromonas hydrophila	31.5 mM
Vibrio cholerae	147.4 μM
Vibrio parahaemolyticus	237.2 μM

7

Summary and Conclusion

The emergence of pathogenic bacteria which are resistant to conventional antibiotics calls for an increased focus on the identification and characterization of antimicrobial compounds with novel and unique mechanisms of action. Antimicrobial peptides (AMPs) have clear and obvious therapeutic potential because of their broad spectrum of activity, rapid killing property, lesser chances of development of resistance and being active in animal models. It is generally accepted that these peptides have the tendency to selectively interact with phospholipid bilayers in negatively charged bacterial surfaces and membranes. AMPs possess extremely broad spectrum of activity against Gram-negative and Gram-positive bacteria, viruses, fungi, and protozoa. AMPs are active even at low concentrations, i.e. in micromolar concentrations or below and possess killing effect within minutes *in vitro*. In the aquatic environment, marine invertebrates are exposed to millions of potential pathogens, through contact, ingestion, and inhalation. As invertebrates rely purely on innate immunity, to combat infectious agents, it is believed that the immune effectors present in these animals are efficient and rapid inhibitors of microbial growth. However, marine invertebrates, especially

marine crustaceans, are still relatively less explored in terms of potential bioactive molecules.

The present study was aimed at identification and characterization of antimicrobial peptides (AMPs) present in marine crustaceans. Recombinant production of a potent AMP in a prokaryotic expression system, *E. coli* was also done. The recombinant and synthetic AMPs were subjected to functional characterization in terms of antibacterial, anticancer and cytotoxic activities.

Salient findings

- ✓ From crustaceans totally nine AMPs belonging to ALF, crustin and penaeidin family could be identified and characterized
- ✓ ALF genes with complete ORF could be identified from the haemocytes of the mud crab, *Scylla serrata* (SsALF1, SsALF2) and the blue swimmer crab, *Portunus pelagicus* (PpALF1, PpALF2).
- ✓ PpALF1, PpALF2 were the first ALF isoforms to be reported from *P. pelagicus*. SsALF1 and SsALF2 were found to be new ALF isoforms from *S. serrata*.
- ✓ Three crustin genes with complete ORF could be characterized from the haemocytes of the mud crab, *S. serrata* (SsCrustin1, SsCrustin2) and the blue swimmer crab, *P. pelagicus* (PpCrustin1).
- ✓ PpCrustin1 was the first crustin sequence to be reported from *P. pelagicus* and SsCrustin1, SsCrustin2 were found to be new crustin isoforms from *S. serrata*.

- ✓ From penaeid shrimps two penaeidin genes with complete ORF could be characterized from *F. indicus* (FiPEN) and *M. monoceros* (MmPEN).
- ✓ MmPEN was the first penaeidin to be reported from *M. monoceros* and FiPEN was found to be a new penaeidin isoform present in *F. indicus.*
- ✓ The phylogenetic analysis performed for these AMP families revealed that the AMPs identified in the present study were evolutionarily closely related to AMPs of other decapod species. The molecular phylogenetic tree implied that each AMP family possesses a same ancestral origin, which has subsequently diverged at different phases of evolution.
- ✓ All the nucleotide sequences and the deduced amino acid sequences were submitted to GenBank under the following accession numbers:
  - ✓ ALF
    - ✓ SsALF1 - **HQ638024**
    - ✓ SsALF2 - **JQ899453**
    - ✓ PpALF1 - **JQ745295**
    - ✓ PpALF2 - **JQ899452**
  - ✓ Crustin
    - ✓ SsCrustin1- **HQ638025**
    - ✓ SsCrustin2 - **JQ753312**
    - ✓ PpCrustin1- **JQ965930**

- ✓ Penaeidin
  - ✓ FiPEN - **JX657680**
  - ✓ MmPEN- **KF275674**

- ✓ Heterologous expression of mature peptide region of the anti-lipopolysaccharide factor, SsALF2 designated as SSA was performed using pET32a+ expression vector system and *E.coli* Rosetta-gami™ B (DE3) pLysS expression host system.
- ✓ Functional characterization of recombinant ALF, rSSA in terms of antimicrobial activity and cytotoxicity were analyzed.
- ✓ Recombinant ALF, SSA was found to possess significant antibacterial activity against Gram-negative bacteria but not against the tested Gram-positive bacteria. Recombinant SSA was found to be active against *V. cholerae* (MIC=11.5 μM), *E. tarda* (MIC= 16.9 μM), *P. aeruginosa* (MIC= 77.1 μM) and showed no activity against Gram-positive *B. cereus*, *S. aureus* and the Gram-negative *V. parahemolyticus* and *A. hydrophila*
- ✓ Recombinant SSA was not found to be cytotoxic to NCI-H460 cells
- ✓ LPS binding domain of PpALF1, designated as PPA, was synthesized and was subjected to functional characterization in terms of antimicrobial, anticancer and cytotoxicity.
- ✓ Synthetic ALF, PPA possessed antibacterial activity against both Gram-positive and Gram-negative bacteria. Synthetic PPA was found to be active against *S. aureus* (MIC=124.4 μM), *V. cholerae* (MIC=147.4 μM) and *V. parahaemolyticus* (MIC=237.2 μM) and less active against *P. aeruginosa* (MIC=1.03 mM), *E. tarda*

(MIC=2.09 mM), *A. hydrophila* (MIC=31.5 mM). *B. cereus* was found to be resistant to synthetic PPA.

- ✓ Synthetic ALF, PPA showed 12 % growth reduction against colon cancer cell line, HCT-116 at a concentration of 10 μM.
- ✓ Synthetic PPA was found to be non cytotoxic to HEp 2 cells.

Conclusion

During the last two decades, much progress has been made in the understanding of the invertebrate innate immune system, especially AMPs. Knowledge of AMPs and its functions will be particularly important for further establishment of disease control in aquaculture systems. The present study provides new clues towards understanding the fundamentals of crustacean immunity as well as its biological activity to some extent. However, they need to be thoroughly investigated in terms of their functional properties as well as its mechanism of action. Such knowledge would form a baseline information for the use of these peptides as prototypes of innovative drugs. Furthermore, it is likely that new peptide antibiotics could be developed based on the known structures of these molecules. The challenge in the near future will be to produce them, at a reasonable cost and develop novel therapeutics with powerful antimicrobial activity. Modification such as cyclization and / or amino acid substitution might improve the activity of the synthetic peptide also. New analogues of this molecule with improved antibacterial activity and low toxicity could also be designed for applications in aquaculture/medicine.

Bibliography

Aketagawa, J., Miyata, T., Ohtsubo, S., Nakamura, T., Morita, T., Hayashida, H., Iwanaga, S., Takao, T., Shimonishi, Y. 1986. Primary structure of Limulus anticoagulant anti-lipopolysaccharide factor. J. Biol. Chem. 261, 7357–7365.

Alberola, J., Rodriguez, A., Francino, O., Roura, X., Rivas, L., Andreu, D. 2004. Safety and efficacy of antimicrobial peptides against naturally acquired Leishmaniasis. Antimicrob. Agents Ch. 48, 641–643.

Alpert, G., Baldwin, G., Thompson, C., Wainwright, N., Novitsky, T.J., Gillis, Z. 1992. Limulus anti-lipopolysaccharide factor protects rabbits from meningococcal endotoxin shock. J. Inf. Dis. 165, 494–500.

Altman, H., Steinberg, D., Porat, Y., Mor, A., Fridman, D., Friedman, M., Bachrach, G. 2006. *In vitro* assessment of antimicrobial peptides as potential agents against several oral bacteria. J. Antimicrob. Chemother. 58, 198–201.

Amparyup, P., Donpudsa, S., Tassanakajon, A. 2008a. Shrimp single WAP domain (SWD)-containing protein exhibits proteinase inhibitory and antimicrobial activities. Dev. Comp. Immunol. 32, 1497-1509.

Amparyup, P., Kondo, H., Hirono, I., Aoki, T. and Tassanakajon, A. 2008b. Molecular cloning, genomic organization and recombinant expression of a crustin-like antimicrobial peptide from black tiger shrimp *Penaeus monodon*. Mol. Immunol. 45, 1085-1093.

Amparyup, P., Kondo, H., Hirono, I., Aoki, T., Tassanakajon, A. 2008. Molecular cloning, genomic organization and recombinant expression of a crustin-like antimicrobial peptide from black tiger shrimp *Penaeus monodon.* Mol. Immunol. 45, 1085–1093.

Andersson, L., Blomberg, L., Flegel, M., Lepsa, L., Nilsson, B., Verlander, M. 2000. Large-scale synthesis of peptides. Biopolymers. 55, 227-250.

Andra, J., Lamata, M., de Martinez Tejada, G., Bartels, R., Koch, M.H., Brandenburg, K. 2004. Cyclic antimicrobial peptides based on Limulus anti-lipopolysaccharide factor for neutralization of lipopolysaccharide. Biochem. Pharmacol. 68, 1297–1307.

Andrenacci, D., Grimaldi, M.R., Panetta, V., Riano, E., Rugarli, E.I., Graziani, F. 2006. Functional dissection of the Drosophila Kallmann's syndrome protein DmKal-1. BMC. Genet. 7, Pp. 47.

Andreu, D., Rivas, L. 1999. Animal antimicrobial peptides: an overview. Biopolymers. 47, 415-433.

Antony, S.P., Philip, R., Joseph, V., Singh, I.S.B. 2011. Anti-lipopolysaccharide factor and crustin-III, the anti-white spot virus peptides in *Penaeus monodon*: Control of viral infection by up-regulation. Aquaculture, 319, 11–17.

Antony, S.P., Philip, R., Sajeevan, T.P., Joseph, S.P., Singh, I.S.B. 2012. Differential expression of antimicrobial peptides in *Penaeus monodon* in response to the administration of marine yeasts / β-glucans and white spot virus challenge. Blue Biotechnology Journal. 1, 1-22.

Antony, S.P., Singh I.S.B., Philip, R. 2010. Molecular characterization of a crustin-like, putative antimicrobial peptide, Fi-crustin, from the Indian White Shrimp, *Fenneropenaeus indicus*. Fish Shellfish Immunol. 28, 216-220.

Arnold, K., Bordoli, L., Kopp, J., Schwede, T. 2006. The SWISS-MODEL Workspace: A web-based environment for protein structure homology modeling. Bioinformatics. 22, 195-201.

Arockiaraj, J., Kumaresan, V., Bhatt, P., Palanisamy, R., Gnanam, A.J., Pasupuleti, M., Chaurasia, M.K. 2014. A novel single-domain peptide, anti-LPS factor from prawn: Synthesis of peptide, antimicrobial properties and complete molecular characterization. Peptides. 53, 79-88.

Arouri, A., Dathe, M., Blume, A. 2009. Peptide induced demixing in PG/PE lipid mixtures: a mechanism for the specificity of antimicrobial peptides towards bacterial membranes? Biochim. Biophys. Acta. 1788, 650–659.

Arrighi, R.B.G., Nakamura, C., Miyake, J., Hurd, H., Burgess, J.G. 2002. Design and activity of antimicrobial peptides against sporogonic-stage parasites causing murine malarias. Antimicrob. Agents Ch. 46, 2104-2110.

Aslund, F., Bessette, P.H., Georgiou, G., Beckwith, J. 1999. Efficient folding of proteins with multiple disulfide bonds in the *Escherichia coli* cytoplasm. Innovations. 10, 11–12.

Bachere, E. 2000a. Penaeidins, antimicrobial peptides with chitin-binding activity, are produced and stored in shrimp granulocytes and released after microbial challenge. J. Cell Sci. 113, 461-469.

Bachere, E., Gueguen, Y., Gonzalez, M., de Lorgeril, J., Garnier, J., Romestand, B. 2004. Insights into the anti-microbial defense of marine invertebrates: the penaeid shrimps and the oyster *Crassostrea gigas*. Immunol. Rev. 198, 149–168.

Baker, M.A., Maloy, W.L., Zasloff, M., Jacob, L.S. 1993. Anticancer efficacy of magainin 2 and analogue peptides. Cancer Res. 53, 3052-3057.

Bals, R. 2000. Epithelial antimicrobial peptides in host defense against Infection. Respir. Res. 1, 141–150.

Barbault, F., Landon, C., Guenneugues, M., Meyer, J.P., Schott, V., Dimarcq, J.L., Vovelle, F. 2003. Solution structure of Alo-3: a new knottintype antifungal peptide from the insect *Acrocinus longimanus*. Biochemistry. 42, 14434–14442.

Barracco, M.A., de Lorgeril, J., Gueguen, Y., Bachere, E. 2005. Molecular characterization of Penaeidins from two Atlantic Brazilian shrimp species, *Farfantepenaeus paulensis* and *Litopenaeus schmitti*. FEMS Microbiol. Lett. 250, 117-120.

Bartlett, T.C., Cuthbertson, B.J., Shepard, E.F., Chapman, R.W., Gross, P.S., Warr, G.W. 2002. Crustins, homologues of an 11.5 kDa antibacterial peptide, from two species of penaeid shrimp, *Litopenaeus vannamei* and *Litopenaeus setiferus*. Mar. Biotechnol. 4, 278-293.

Battison, A.L., Summerfield, R., Patrzykat, A. 2008. Isolation and characterisation of two antimicrobial peptides from hemocytes of the American lobster *Homarus americanus*. Fish Shellfish Immunol. 25, 181–187.

Bax, R., Mullan, N., Verhoef, J. 2000. The millennium bugs — the need for and development of new antibacterials. Int. J. Antimicrob. Agents. 16, 51–59.

Beale, K.M., Towle, D.W., Jayasundara, N., Smith, C.M., Shields, J.D., Small, H.J. 2008. Anti-lipopolysaccharide factors in the American lobster *Homarus americanus*: Molecular characterization and transcriptional response to *Vibrio fluvialis* challenge. Comp. Biochem. Physiol. Part D. 3, 263–269.

Berge, G., Eliassen, L.T., Camilio, K.A., Bartnes, K., Sveinbjornsson, B., Rekdal, O. 2010. Therapeutic vaccination against a murine lymphoma by intratumoral injection of a cationic anticancer peptide. Cancer Immunol. Immunother. 59, 1285–1294.

Bogdan, C., Röllinghoff, M., Diefenbach, A. 2000. Reactive oxygen and reactive nitrogen intermediates in innate and specific immunity. Curr. Opin. Immunol. 12, 64-76.

Boman, H.G. 1995. Peptide antibiotics and their role in innate immunity. Annu. Rev. Immunol. 13, 61-92.

Boman, H.G. 2003. Antibacterial peptides: basic facts and emerging concepts. J. Int. Med. 254, 197-215.

Boman, H.G., Agerberth, B., Boman, A. 1993. Mechanisms of action on *Escherichia coli* of cecropin P1 and PR-39, 2 antibacterial peptides from pig intestine. Infect. Immun. 617, 2978-2984.

Boman, H.G., Faye, I., Gudmundsson, G.H. 1991. Cell-free immunity in cecropia: a model system for antibacterial proteins. Eur. J. Biochem. 201, 23-31.

Boman, H.G., Faye, I., Hofsten, P.V., Kockum, K., Lee, J.Y., Xanthopoulos, K.G. 1985. On the primary structure of lysozyme, cecropins and attacins from *Hyalophora cecropia*. Dev. Comp. Immunol. 9, 551-558.

Bonomo, R.A. 2000. Multiple antibiotic-resistant bacteria in long-term-care facilities: an emerging problem in the practice of infectious diseases. Clin. Infect. Dis. 31, 1414-1422.

Bouchard, D., Morisset, D., Bourbonnais, Y., Tremblay, G.M. 2006. Proteins with whey-acidic-protein motifs and cancer. Lancet Oncol. 7, 167–74.

Bowdish, D.M., Davidson, D.J., Hancock, R.E.W. 2005. A re-evaluation of the role of host defense peptides in mammalian immunity. Curr. Prot. Pept. Sci. 6, 35–51.

Bowman, H.G. 2003. Antibacterial peptides: basic facts and emerging concepts. J. Intern. Med. 254, 197–215.

Brandenburg, L., Merres, J., Albrecht, L., Varoga, D., Pufe, T. 2012. Antimicrobial peptides: Multifunctional drugs for different applications. 4, 539–560.

Brewer, D., Hunter, H., Lajoie, G. 1998. NMR studies of the antimicrobial salivary peptides histatin 3 and histatin 5 in aqueous and nonaqueous solutions. Biochem. Cell Biol. 76, 247-256.

Brock, T.D. 1974. Biology of Microorganisms. Prentice-Hall, Inc, Englewood Cliffs N.J.

Brockton, V., Hammond, J.A., Smith, V.J. 2007. Gene characterisation, isoforms and recombinant expression of carcinin, an antibacterial protein from the shore crab, *Carcinus maenas*. Mol. Immunol. 44, 943–949.

Brogden, K.A. 2005. Antimicrobial peptides: Pore formers or metabolic inhibitors in bacteria? Nat. Rev. Microbiol. 3, 238-250.

Brogden, K.A., Ackermann, M., McCray Jr., P.B., Tack, B.F. 2003. Antimicrobial peptides in animals and their role in host defenses. Int. J. Antimicrob. Agents. 22,465-478.

Brogden, N.K., Brogden, K.A. 2011. Will new generations of modified antimicrobial peptides improve their potential as pharmaceuticals? Int. J. Antimicrob. Agents. 38, 217-225.

Brotz, H., Bierbaum, G., Leopold, K., Reynolds, P.E., Sahl, H.G. 1998. The lantibiotic mersacidin inhibits peptidoglycan synthesis by targeting lipid II. Antimicrob. Agents Ch. 42, 154–160.

Brown, K.L., Hancock, R.E. 2006. Cationic host defense (antimicrobial) peptides. Curr. Opin. Immunol. 18, 24-30.

Brumfitt, W., Salton, M., Hamilton-Miller, J. 2002. Nisin, alone and combined with peptidoglycan-modulating antibiotics: activity against methicillin-resistant *Staphylococcus aureus* and vancomycin-resistant *Enterococci*. J. Antimicrob. Chemother. 50, 731–734.

Bu, X., Wu, X., Xie, G., Guo, Z. 2002. Synthesis of tyrocidine A and its analogues by spontaneous cyclization in aqueous solution. Org. Lett. 4, 2893–2895.

Bulet, P., Hetru, C., Dimarq, J.L., Hoffman, D. 1999. Antimicrobial peptides in insects; structure and function. Dev. Comp. Immunol. 23, 329-344.

Bulet, P., Stocklin, R., Menin, L. 2004. Antimicrobial peptides: from invertebrate to vertebrate. Immunol. Rev. 198, 169-184.

Bulet, P., Urge, L., Ohresser, S., Hetru, C., Otvos, L. 1996. Enlarged scale chemical synthesis and range of activity of drosocin, an O-glysosylated antibacterial peptide of *Drosophila*. Eur. J. Biochem. 238, 64-69.

Burge, E.J., Madigan, D.J., Burnett, L.E., Burnett, K.G. 2007. Lysozyme gene expression by haemocytes of Pacific White Shrimp, *Litopenaeus vannamei*, after injection with *Vibrio*. Fish Shellfish Immunol. 22, 327-39.

Cabiaux, V., Agerberth, B., Johansson, J., Homblé, F., Goormaghtigh, E., Ruysschaert, J.M. 1994. Secondary structure and membrane interaction of PR-39, a Pro+Arg-rich antibacterial peptide. Eur. J. Biochem. 224, 1019-1027.

Carriel-Gomes, M.C., Kratz, J.M., Barracco, M.A., Bachere, E., Barardi, C.R., Simoes, C.M. 2007. *In vitro* antiviral activity of antimicrobial peptides against herpes simplex virus 1, adenovirus, and rotavirus. Mem. Inst. Oswaldo Cruz. 102, 469–472.

Casteels, P., Ampe, C., Jacobs, F., Tempst, P. 1993. Functional and chemical characterization of hymenoptaecin, an antibacterial polypeptide that is infection-inducible in the honeybee *Apis mellifera*. J. Biol. Chem. 268, 7044–7054.

Casteels, P., Tempst, P. 1994. Apidaecin-type peptide antibiotics function through a nonporeforming mechanism involving stereospecificity. Biochem. Biophys. Res. Commun. 199, 339-345.

Chan, D.I., Prenner, E.J., Vogel, H.J. 2006. Tryptophan- and arginine-rich antimicrobial peptides: Structures and mechanisms of action. Biochim. Biophys. Acta. 1758, 1184–1202.

Chen, J.Y., Chuang, H., Pan, C.Y., Kuo, C.M. 2005. cDNA sequence encoding an antimicrobial peptide of chelonianin from the tiger shrimp *Penaeus monodon*. Fish Shellfish Immunol. 18, 179-183.

Chen, J.Y., Pan, C.Y., Kuo, C.M. 2004a. cDNA sequence encoding an 11.5-kDa antibacterial peptide of the shrimp *Penaeus monodon*. Fish Shellfish Immunol. 16, 659–664.

Chen, J.Y., Pan, C.Y., Kuo, C.M. 2004b. Molecular cloning and sequencing of shrimp *Penaeus monodon* penaeidin-5 cDNA. Fish Shellfish Immunol. 16, 665–670.

Chiou, T.T., Lu, J.K., Wu, J.L., Chen, T.T., Ko, C.F., Chen, J.C. 2006. Expression and characterization of tiger shrimp, *Penaeus monodon* penaeidin [mo-penaeidin] in various tissues, during early embryonic development and moulting stages. Dev. Comp. Immunol. 31, 132-142.

Chiou, T.T., Wu, J.L., Chen, T.T., Lu, J.K. 2005. Molecular cloning and characterization of cDNA of penaeidin-like antimicrobial peptide from tiger shrimp *Penaeus monodon*. Mar. Biotechnol. 7, 119-127.

Christie, A.E., Rus, S., Goiney, C.C., Smith, C.M., Towle, D.W., Dickinson, P.S. 2007. Identification and characterization of a cDNA encoding a crustin-like, putative antibacterial protein from the American lobster *Homarus americanus*. Mol. Immunol. 4413, 3333-3337.

Ciccarone, V.C., Polayes, D., Luckow, V.A. 1997. In: Reischt, U. (Ed.), Generation of recombinant baculovirus DNA in *E. coli* using baculovirus shuttle vector. Vol. 13. Humana Press, Totowa, New Jercy.

Cleveland, J., Montville, T.J., Nes, I.F., Chikindas, M.L. 2001. Bacteriocins: safe, natural antimicrobials for food preservation. J. Food Microbiol. 71, 1-20.

Cruciani, R.A., Barker, J.L., Zasloff, M., Chen, H.C., Colamonici, O. 1991. Antibiotic magainins exert cytolytic activity against transformed cell lines through channel formation. Proc. Natl. Acad. Sci. USA. 88, 3792-3796.

Cuthbertson, B.J., Bullesbach, E.E., Fievet, J., Bachere, E., Gross, P.S. 2004. A new class penaeidin class 4 of antimicrobial peptides from the Atlantic white shrimp *Litopenaeus setiferus* exhibits target specificity and an independent proline-rich-domain function. Biochem. J. 381, 79–86.

Cuthbertson, B.J., Bullesbach, E.E., Gross, P.S. 2006. Discovery of synthetic penaeidin activity against antibiotic-resistant fungi. Chem. Biol. Drug Res. 68, 120–127.

Cuthbertson, B.J., Shepard, E.F., Chapman, R.W., Gross, P.S. 2002. Diversity of the penaeidin antimicrobial peptides in two shrimp species. Immunogenetics. 54, 442–445.

Cuthbertson, B.J., Yang, Y., Bachere, E., Bullesbach, E.E., Gross, P.S., Aumelas, A. 2005. Solution structure of synthetic penaeidin-4 with structural and functional comparisons with Penaeidin-3. J. Biol. Chem. 280, 16009–16018.

Daher, K.A., Selsted, M.E., Lehrer, R.I. 1986. Direct inactivation of viruses by human granulocyte defensins. J. Virol. 60, 1068-1074.

Dai, Z.M., Zhu, X.J., Yang, W.J. 2009. Full-length normalization subtractive hybridization: a novel method for generating differentially expressed cDNAs. Mol. Biotechnol. 43, 257-263.

Dawson, R.M., Liu, C.Q. 2008. Properties and applications of antimicrobial peptides in biodefense against biological warfare threat agents. Crit. Rev. Microbiol. 34, 89–107.

de la Vega, E., O'Leary, N.A., Shockey, J.E., Robalino, J., Payne, C., Browdy, C. L., Warr, G.W., Gross, P.S. 2008. Anti-lipopolysaccharide factor in *Litopenaeus vannamei Lv*ALF: A broad spectrum antimicrobial peptide essential for shrimp immunity against bacterial and fungal infection. Mol. Immunol. 45, 1916–1925.

de Lucca, A.J., Walsh, T.J. 1999. Antifungal peptides: Novel therapeutic compounds against emerging pathogens. Antimicrob. Agents Ch. 43, 1-11.

Delatorre, P., Olivieri, J.R., Neto, R.J., Lorenzi, C.C.B., Canduri, F., Fadel, V., Konno, K., Palma, M.S., Yamane, T., de Azevedo, W.F. 2001. Preliminary cryocrystallography analysis of an eumenine mastoparan toxin isolated from the venom of the wasp *Anterhynchium flavomarginatum* micado. Biochim. Biophys. Acta. 1545, 372-376.

Delves-Broughton, J. 1990. Nisin and its uses as a food preservative. Food Technol. 44, 100-117.

Dennison, S.R., Harris, F., Phoenix, D.A. 2003. Antimicrobial activity of aurein 2.5 against yeasts. FEMS Microbiol Lett. 2013, 346, 140-145.

Desai, P.N., Shrivastava, N., and Padh, H. 2010. Production of heterologous proteins in plants: strategies for optimal expression. Biotech. Adv. 28, 427–435.

Destoumieux, D., Bulet, P., Loew, D., Van Dorsselaer, A., Rodriguez, J. and Bachere E. 1997. Penaeidins: a new family of antimicrobial peptides in the shrimp *Penaeus vannamei* (Decapoda). J. Biol. Chem. 272, 28398–28406.

Destoumieux, D., Bulet, P., Strub, J.M., van Dorsselaer, A., Bachere, E. 1999. Recombinant expression and range of activity of penaeidins, antimicrobial peptides from penaeid shrimp. Eur. J. Biochem. 266, 335-346.

Destoumieux, D., Munoz, M., Bulet, P., Bachere, E. 2000a. Penaeidins, a family of antimicrobial peptides from penaeid shrimp Crustacea, Decapoda. Cell Mol. Life. Sci. 57, 1260–1271.

Destoumieux, D., Munoz, M., Cosseau, C., Rodriguez, J., Bulet, P., Comps, M., Bachere, E. 2000b. Penaeidins, antimicrobial peptides with chitin-binding activity, are produced and stored in shrimp granulocytes and released after microbial challenge. J. Cell Sci. 113, 461-469.

Destoumieux-Garzon, D., Saulnier, D., Garnier, J., Jouffrey, C., Bulet, P., Bachère, E. 2001. Crustacean immunity. Antifungal peptides are generated from the C terminus of shrimp hemocyanin in response to microbial challenge. J. Biol. Chem. 276, 47070-47077.

Dubos, R.J. 1939. Studies on a bactericidal agent extracted from a soil *Bacillus*: I. Preparation of the agent. Its activity *in vitro*. J Exp Med 70, 1–10.

Duclohier, H., Wroblewski, H. 2001. Voltage-dependent pore formation and antimicrobial activity by alamethicin and analogues. J. Membr. Biol. 184, 1-12.

Fehlbaum, P., Bulet, P., Chernysh, S., Briand, J.P., Roussel, J.P., Leitellier, L., Hetru, C., Hoffmann, J.A. 1996. Structure-activity analysis of thanatin, a 21-residue inducible insect defense peptide with sequence homology to frog skin antimicrobial peptides. Proc. Natl. Acad. Sci. USA. 93, 1221-1225.

Feng, X., Liu, C., Guo, J., Song, X., Li, J., Xu, W., Li, Z. 2012. Recombinant expression, purification, and antimicrobial activity of a novel hybrid antimicrobial peptide LFT33. Appl. Microbiol. Biotechnol. 95, 1191–1198.

Ferlay, J., Shin, H.R., Bray, F., Forman, D., Mathers, C., Parkin, D.M. 2010. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. Int J Cancer. 127, 2893-917.

Fernandes, J.M, Kemp, G.D, Molle, M.G, Smith, V.J. 2002. Antimicrobial properties of histone H2A from skin secretions of rainbow trout, *Oncorhynchus mykiss*. Biochem. J. 368, 611-620.

Fernebro, J. 2011. Fighting bacterial infections-future treatment options. Drug Resist. Updat. 14, 125-139.

Fjell, C.D., Hiss, J.A., Hancock, R.E., Schneider, G. 2012. Designing antimicrobial peptides: Form follows function. Nat. Rev. Drug Discov. 11, 37–51.

Fontana, R., Mendes, M.A., De Souza, B.M., Konno, K., Cesar, L.M., Malaspina, O., Palma, M.S. 2004. Jelleines: a family of antimicrobial peptides from the Royal Jelly of honeybees (*Apis mellifera*). Peptides. 25, 919-928.

Fujimura, M., Ideguchi, M., Minami, Y., Watanabe, K., Tadera, K. 2004. Purification, characterization, and sequencing of novel antimicrobial peptides, Tu-AMP 1 and Tu-AMP 2, from bulbs of tulip *Tulipa gesneriana* L. Biosci. Biotechnol. Biochem. 68, 571–577.

Gallo, R.L., Murakami, M., Ohtake, T., Zaiou, M. 2002. Biology and clinical relevance of naturally occurring antimicrobial peptides. J. Allergy Clin. Immunol. 110, 823-831.

Ganz, T. 2003. Defensins: antimicrobial peptides of innate immunity. Nat. Rev. Immunol. 3, 710-720.

Ganz, T., Lehrer, R.I. 1995. Defensins. Pharmacol. Ther. 66, 191-205.

Gennaro, R., Skerlavaj, B., Romeo, D. 1989. Purification, composition and activity of two bactenecins, antibacterial peptides of bovine neutrophils. Infect. Immun. 57, 3142–3146.

Giangaspero, A., Sandri, L., Tossi, A. 2001. Amphipathic α-helical antimicrobial peptides. Eur. J. Biochem. 268, 5589-5600.

Giuliani, A., Pirri, G., Nicoletto, S.F. 2007. Antimicrobial peptides: an overview of a promising class of therapeutics. Cent. Eur. J. Biol. 2, 1-33.

Goumon, Y., Strub, J.M., Moniatte, M., Nullans, G., Poteur, L., Hubert, P., van Dorsselaer, A., Aunis, D., Metz-Boutigue, M.H. 1996. The C-terminal bisphosphorylated proenkephalin-a-(209–237)-peptide from adrenal medullary chromaffin granules possesses antibacterial activity. Eur. J. Biochem. 235, 516-525.

Gross, P.S., Bartlett, T.C., Browdy, C.L., Chapman, R.W., Warr, G.W. 2001. Immune gene discovery by expressed sequence tag analysis of hemocytes and hepatopancreas in the Pacific White Shrimp, *Litopenaeus vannamei*, and the Atlantic White Shrimp, *L. setiferus*. Dev. Comp. Immunol. 25, 565–577.

Guaní-Guerra, E., Santos-Mendoza, T., Lugo-Reyes, S.O. 2010. Antimicrobial peptides: General overview and clinical implications in human health and disease. Clin. Immunol. 135, 1-11.

Gudmundsson, G.H., Agerberth, B., Odeberg, J., Bergman, T., Olsson, B., Salcedo, R. 1996. The human gene FALL39 and processing of the cathelin precursor to the antibacterial peptide LL-37 in granulocytes. Eur. J. Biochem. 238, 325-332.

Gueguen, Y., Garnier, J., Robert, L., Lefranc, M.P., Mougenot, I., de Lorgeril, J., Janech, M., Gross, P.S., Warr, G.W., Cuthbertson, B., Barracco, M.A., Bulet, P., Aumelas, A., Yang, Y., Bo, D., Xiang, J., Tassanakajon, A., Piquemal, D., Bachere, E. 2006. PenBase, the shrimp antimicrobial peptide penaeidin database: Sequence-based classification and recommended nomenclature. Dev. Comp. Immunol. 30, 283–288.

Guex, N., Peitsch, M.C. 1997. SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modelling. Electrophoresis. 18, 2714-2723.

Guo, S., Li, S., Li, F., Zhang, X., Xiang, J. 2014. Modification of a synthetic LPS-binding domain of anti-lipopolysaccharide factor from shrimp reveals strong structure-activity relationship in their antimicrobial characteristics. Dev. Comp. Immunol. 45, 227–232.

Hagiwara, K., Kikuchi, T., Endo, Y., Huqun, U.K., Takahashi, M., Shibata, N. 2003. Mouse SWAM1 and SWAM2 are antibacterial proteins composed of a single whey acidic protein motif. J. Immunol. 170, 1973–1979.

Hale, J.D., Hancock, R.E.W. 2007. Alternative mechanisms of action of cationic antimicrobial peptides on bacteria. Expert Rev. Anti Infect. Ther. 5, 951–959.

Hancock, R.E. Diamond, G. 2000. The role of cationic antimicrobial peptides in innate host defences. Trends Microbiol. 8, 402–410.

Hancock, R.E., Chapple, D.S. 1999. Peptide antibiotics. Antimicrob. Agents Chemother. 43, 1317-1323.

Hancock, R.E., Diamond, G. 2000. The role of cationic antimicrobial peptides in innate host defences. Trends Microbiol. 8, 402-410.

Hancock, R.E., Lehrer, R. 1998. Cationic peptides: a new source of antibiotics. Trends Biotechnol. 16, 82-88.

Hancock, R.E., Rozek, A. 2002. Role of membranes in the activities of antimicrobial cationic peptides. FEMS Microbiol. Lett. 206,143-149.

Hancock, R.E.W. 1997. Peptide antibiotics. Lancet. 349, 418-422.

Hancock, R.E.W. 2001. Cationic peptides: effectors in innate immunity and novel antimicrobials. Infectious Diseases. Lancet. 1, 156-164.

Harwig, S.S., Swiderek, K.M., Lee, T.D., Lehrer, R.I. 1995. Determination of disulphide bridges in PG-2, an antimicrobial peptide from porcine leukocytes. J. Pept. Sci. 1, 207-215.

Haug, T., Stensvag, K., Vasskog, T., Sandsdalen., Styrvold, O. 2006. Antibacterial peptides isolated from the haemocytes of the small spider crab, *Hyas araneus*. In: Abstract of the 10th international congress of ISDCI, Charleston, SC, USA.

Hauton, C., Brokton, V., Smith, V.J. 2006. Cloning of a crustin-like single whey-acidic-domain, antibacterial peptide from the haemocytes of the European lobster, *Homarus gammarus*, and its response to infection with bacteria. Mol. Immunol. 43, 1490-1496.

Helmerhorst, E.J., Breeuwer, P., van't Hof, W., Walgreen-Weterings, E., Oomen, L.C., Veerman, E.C., Amerongen, A.V., Abee, T. 1999. The cellular target of histatin 5 on *Candida albicans* is the energized mitochondrion. J. Biol. Chem. 274, 7286–7291.

Herbinière, J., Braquart-Varnier, C., Greve, P., Strub, J.M., Frere, J., Van Dorsselaer, A., Martin, G. 2004. Armadillidin: a novel glycine-rich antibacterial peptide directed against gram-positive bacteria in the woodlouse *Armadillidium vulgare* (Terrestrial Isopod, Crustacean). Dev. Comp. Immunol. 29, 489-499.

Hiemestra, P.S., Massen, R.J., Stolk, J., Heinzel-Wieland, R., Staffens, G.J., Dijkman, J.H. 1996. Antibacterial activity of anti-leukoprotease. Infect. Immunol. 64, 4520–4522.

Hiemstra, P.S. 2002. Novel roles of protease inhibitors in infection and inflammation. Biochem. Soc. Trans. 30, 116–120.

Hirakura, Y., Kobayashi, S., Matsuzaki, M. 2002. Specific interactions of the antimicrobial peptide cyclic β-sheet tachyplesin I with lipopolysaccharides. Biochim. Biophys. Acta. 1562, 32–36.

Hoebe, K., Jansen, E., Beutler, B. 2004. The interface between innate and adaptive immunity. Nat. Immunol. 5, 971-974.

Hoess, A., Watson, S., Siber, G.R., Liddington, R. 1993. Crystal structure of an endotoxin-neutralizing protein from the horseshoe crab, Limulus anti-LPS factor, at 1.5 A resolution. EMBO J. 12, 3351- 3356.

Hoffmann, J.A. 1995. Innate immunity of insects. Immunology. 7, 4–10.

Hoffmann, J.A., Kafatos, F.C., Janeway, C.A., Ezekowitz, R.A. 1999. Phylogenetic perspectives in innate immunity. Science. 284, 1313-1318.

Hong, J., Dai, J.G., Guan, W.T., Jin, G., Huang, Z.L., Zhang, L.J., Dang, J., Zhang, Y. 2012. Tachyplesin I induce drug resistance in bacteria *in vitro*. J. Anim. Vet. Adv. 11, 939–945.

Hoskin, D.W., Ramamoorthy, A. 2008. Studies on anticancer activities of antimicrobial peptides. Biochim. Biophys. Acta 1778, 357-375.

Hsu, C.H., Chen, C., Jou, M.L., Lee, A.Y., Lin, Y.C., Yu, Y.P., Huang, W.T., Wu, S.H. 2005. Structural and DNA-binding studies on the bovine antimicrobial peptide, indolicidin: evidence for multiple conformations involved in binding to membranes and DNA. Nucleic Acids Res. 33, 4053–4064.

Hu, S.Y., Huang, J.H., Huang, W.T., Yeh, Y.H., Chen, M.H.C., Gong, H.Y., Chiou, T.T., Yang, T.H., Chen, T.T., Lu, J.K., Wu, J.L. 2006. Structure and function of antimicrobial peptide penaeidin-5 from the black tiger shrimp *Penaeus monodon*. Aquaculture. 260, 61-68.

Huang, H.W. 2000. Action of antimicrobial peptides: two-state model. Biochemistry. 39, 8347-8352.

Huang, W.S., Wang, K.J., Yang, M., Cai, J.J., Li, S.J., Wang, G.Z. 2006. Purification and part characterization of a novel antibacterial protein Scygonadin, isolated from the seminal plasma of mud crab, *Scylla serrata* (Forskal, 1775). J. Exp. Mar. Biol. Ecol. 339, 37-42.

Huang, Y., Huang, J., Chen, Y. 2010. Alpha-helical cationic antimicrobial peptides: relationships of structure and function. Protein Cell. 1, 143-152.

Imjongjirak, C., Amparyup, P., Tassanakajon, A. 2011. Molecular cloning, genomic organization and antibacterial activity of a second isoform of antilipopolysaccharide factor (ALF) from the mud crab, *Scylla paramamosain*. Fish Shellfish Immunol. 30, 58–66.

Imjongjirak, C., Amparyup, P., Tassanakajon, A., Sittipraneed, S. 2009. Molecular cloning and characterization of crustin from mud crab *Scylla paramamosain*. Mol. Biol. Rep. 36, 841–850.

Imjongjirak, C., Amparyup, P., Tassanakajon, A., Sittipraneed, S. 2007. Antilipopolysaccharide factor ALF of mud crab *Scylla paramamosain*: Molecular cloning, genomic organization and the antimicrobial activity of its synthetic LPS binding domain. Mol. Immunol. 44, 3195–3203.

Imura, Y., Nishida, M., Ogawa, Y., Takakura, Y., Matsuzaki, K. 2007. Action mechanism of tachyplesin I and effects of PEGylation. Biochim. Biophys. Acta. 1768, 1160–1169.

Ingham, A.B., Moore, R.J. 2007. Recombinant production of antimicrobial peptides in heterologous microbial systems. Biotechnol. Appl. Biochem. 47, 1–9.

Iwanaga, S., Lee, B.L. 2005. Recent advances in the innate immunity of invertebrate animals. J. Biochem. Mol. Biol. 38, 128–150.

Iwanaga, S., Morita, T., Harada, T., Nakamura, S., Niwa, M., Takada, K., Kimura, T., Sakakibara, S. 1978. Chromogenic substrates for horseshoe crab clotting

enzyme. Its application for the assay of bacterial endotoxins. Haemostasis. 7, 183-188.

Jenssen, H., Hamill, P., Hancock, R.E.W. 2006a. Peptide antimicrobial agents. Clin. Microbiol. Rev. 19, 491–511,

Jenssen, H., Gutteberg, T.J., Lejon, T. 2006b. Modelling the anti-herpes simplex virus activity of small cationic peptides using amino acid descriptors. J. Pept. Res. Suppl. 1, 48–56.

Jerala, R., Porro, M. 2004. Endotoxin neutralizing peptides. Curr. Top. Med. Chem. 4, 1173–1184.

Jia, Y.P., Sun, Y.D., Wang, Z.H., Wang, Q., Wang, Z.W., Zhao, Z.F., Wang, J.X. 2008. A single whey acidic protein domain (SWD)-containing peptide from fleshy prawn with antimicrobial and proteinase inhibitory activities. Aquaculture. 284, 246 – 259.

Jiang, Z., Vasil, A.I., Hale, J.D., Hancock, R.E.W., Vasil, M.L., Hodges, R.S. 2008. Effects of net charge and the number of positively charged residues on the biological activity of amphipathic alpha-helical cationic antimicrobial peptides. Biopolymers. 90, 369–383.

Jiménez-Vega, F., Vargas-Albores, F. 2007. A secretory leukocyte proteinase inhibitor (SLPI)-like protein from *Litopenaeus vannamei* haemocytes. Fish Shellfish Immunol. 23, 1119-1126.

Jimenez-Vega, F., Yepiz-Plascencia, G., Soderhall, K., Vargas- Albores, F. 2004. A single WAP domain-containing protein from *Litopenaeus vannamei* hemocytes. Biochem. Biophys. Res. Commun. 314, 681–687.

Jimenez-Vega, F., Yepiz-Plascencia, G., Soderhall, K., Vargas- Albores, F. 2004. A single WAP domain-containing protein from *Litopenaeus vannamei* hemocytes. Biochem. Biophys. Res. Commun. 314, 681–687.

Jiravanichpaisal, P., Lee, B.L., Söderhäll, K. 2006. Cell-mediated immunity in arthropods: hematopoiesis, coagulation, melanization and opsonization. Immunobiology. 211, 213-236.

Jiravanichpaisal, P., Puanglarp, N., Petkon, S., Donnuea, S., Soderhall, I., Soderhall, K. 2007. Expression of immune-related genes in larval stages of the giant tiger shrimp, *Penaeus monodon*. Fish Shellfish Immunol. 23, 815-824.

Jones, T.J., Overstreet, R.M., Lotz, J.M., Frelier, P.F. 1994. *Paraophioidina scolecoides* n. sp., a new aseptate gregarine form cultured Pacific white shrimp *Penaeus vannamei.* Dis. Aquat. Organ. 19, 67-75.

Kagan, B.L., Selsted, M.E., Ganz, T., Lehrer, R.I. 1990. Antimicrobial defensin peptides form voltage-dependent ion-permeable channels in planar lipid bilayer membranes. Proc. Natl. Acad. Sci. USA. 87, 210-214.

Kalfa, V.C., Jia, H.V.P., Kunkle, R.A., McCray, P.B., Tack, B.F., Brogden, K.A. 2001. Congeners of SMAP29 kill ovine pathogens and induce ultrastructural damage in bacterial cells. Antimicrob. Agents Ch. 45, 3256–3261.

Kang, C.J., Wang, J.X., Zhao, X.F., Yang, X.M., Shao, H.L., Xiang, J.H. 2004. Molecular cloning and expression analysis of Ch-penaeidin, an antimicrobial peptide from Chinese shrimp, *Fenneropenaeus chinensis*. Fish Shellfish Immunol. 16, 513-525.

Kang, C.J., Xue, J.F., Liu, N., Zhao, X.F., Wang, J.X. 2007. Characterization and expression of a new subfamily member of penaeidin antimicrobial peptides (penaeidin 5) from *Fenneropenaeus chinensis*. Mol. Immunol. 44, 1535-1543.

Khoo, L., Robinette, D.W., Noga, E.J. 1999. Callinectin, an antibacterial peptide from blue crab, *Callinectes sapidus*, hemocytes. Mar. Biotechnol. 1, 44-51.

Kieffer, A.E., Goumon, Y., Ruh, O., Chasserot-Golaz, S., Nullans, G., Gasnier, C., Aunis, D., Metz-Boutigue, M.H. 2003. The N- and C-terminal fragments of ubiquitin are important for the antimicrobial activities. FASEB J. 17, 776-778.

Kikukawa, T., Araiso, T. 2002. Changes in lipid mobility associated with *alamethicin* incorporation into membranes. Arch. Biochem. Biophys. 405, 214-222.

Kragol, G., Lovas, S., Varadi, G., Condie, B.A., Hoffmann, R., Otvos Jr. L. 2001. The antibacterial peptide pyrrhocoricin inhibits the ATPase actions of DnaK and prevents chaperone-assisted protein folding. Biochemistry. 40, 3016–3026.

Krusong, K., Poolpipat, P., Supungul, P., Tassanakajon, A. 2012. A comparative study of antimicrobial properties of crustinPm1 and crustinPm7 from the black tiger shrimp *Penaeus monodon*. Dev. Comp. Immunol. 36, 208–215.

Krutzik, S.R., Sieling, P.A., Modlin, R.L. 2001. The role of Toll-like receptors in host defense against microbial infection. Curr. Opin. Immunol. 13, 104-108.

Lai, Y., Gallo, R.L. 2009. AMPed up immunity: how antimicrobial peptides have multiple roles in immune defense. Trends Immunol. 30, 131-141

Lamaziere, A., Burlina, F., Wolf, C., Chassaing, G., Trugnan, G., Ayala-Sanmartin, J. 2007. Non-metabolic membrane tubulation and permeability induced by bioactive peptides. Plos One. 2(2), e201.

Lamb, H.M., Wiseman, L.R. 1998. Pexiganan acetate. Drugs. 56, 1047– 1054.

Lavine, M.D., Strand, M.R. 2002. Insect hemocytes and their role in immunity. Insect Biochem. Mol. Biol. 3210, 1295-1309.

Lawyer, C., Pai, S., Watabe, M., Borgia, P., Mashimo, T., Eagleton, L., Watabe, K. 1996. Antimicrobial activity of a 13 amino acid tryptophan-rich peptide derived from a putative porcine precursor protein of a novel family of antibacterial peptides. FEBS Lett. 390, 95-98.

Lee, S.Y., Lee, B.L., Soderhall, K. 2003. Processing of an antibacterial peptide from hemocyanin of the freshwater crayfish *Pacifastacus leniusculus*. J. Biol. Chem. 27810, 7927-7933.

Lee, S.Y., Lee, B.L., Soderhall, K. 2004. Processing of crayfish hemocyanin subunits into phenoloxidase. Biochem. Biophys. Res. Commun. 3222, 490-496.

Lehrer, R.I., Barton, A., Daher, K.A., Harwig, S.S., Ganz, T., Selsted, M.E. 1989. Interaction of human defensins with *Escherichia coli*. Mechanism of bactericidal activity. J. Clin. Invest. 84, 553-561.

Lehrer, R.I., Szklarek, D., Ganz, T., Selsted, M.E. 1985. Correlation of binding of rabbit granulocyte peptides to *Candida albicans* with candidacidal activity. Infect. Immun. 49, 207–211.

Leippe, M., Ebel, S., Schoenberger, O.L., Horstmann, R.D., Muller-Eberhard, H.J. 1991. Proc. Natl. Acad. Sci. USA. 88, 7659-7663.

Lemaitre, B., Nicolas, E., Michaut, L., Reichhart, J.M., Hoffmann, J.A. 1996. The dorsoventral regulatory gene cassette spaetzle/Toll cactus controls the potent antifungal response in *Drosophila* adults. Cell. 86, 1-20.

Li, C., Blencke, H. M., Paulsen, V., Haug, T., Stensvåg, K. 2010. Powerful work horses for antimicrobial peptide expression. Bioeng. Bugs. 1, 217-220.

Li, C., Song, L., Zhao, J., Zhu, L., Zou, H. 2007. Preliminary study on a potential antibacterial peptide derived from histone H2A in hemocytes of scallop *Chlamys farreri.* Fish Shellfish Immunol. 22, 663-672.

Li, H., Lu, X., Li, C., Chen, J. 2014. Molecular characterization and functional analysis of two distinct liver-expressed antimicrobial peptide 2 (LEAP-2) genes in large yellow croaker (*Larimichthys crocea*). Fish Shellfish Immunol. 38, 330-339.

Li, L., Wang, J., Zhao, X., Kanga, C., Liua, N., Xiang, J., Sueda, S., Kondo, H. 2005. High level expression, purification, and characterization of the shrimp antimicrobial peptide, Ch-penaeidin, in *Pichia pastoris*. Prot. Exp. Pur. 39, 144–151.

Li, S., Zhang, X., Sun, Z., Li, F., Xiang, J. 2013. Transcriptome Analysis on Chinese Shrimp *Fenneropenaeus chinensis* during WSSV acute infection. PLoS ONE. 8, e58627.

Li, Y. 2009. Carrier proteins for fusion expression of antimicrobial peptides in *Escherichia coli*. Biotechnol. Appl. Biochem. 54, 1-9.

Li, Y.F., Chen, Z.X. 2008. RAPD: a database of recombinantly-produced antimicrobial peptides. FEMS Microbiol. Lett. 289, 126–129.

Li, Z.Q., Merrifield, R.B., Boman, A., Boman, H.G. 1988. Effects on electrophoretic mobility and antibacterial spectrum of removal of two residues from synthetic sarcotoxin IA and addition of the same residues to cecropin B. FEBS Lett. 231, 299-302.

Lin, M.C., Pan, C.Y., Hui, C.F., Chen, J.Y., Wu, J.L. 2013. Shrimp anti-lipopolysaccharide factor (SALF), an antimicrobial peptide, inhibits proinflammatory cytokine expressions through the MAPK and NF-κB pathways in LPS-induced HeLa cells. Peptides. 40, 42–48.

Linde, C.M.A., Hoffner, S.E., Refai, E., Andersson, M. 2001. *In vitro* activity of PR-39, a proline-arginine-rich peptide, against susceptible and multi-drug-resistant *Mycobacterium tuberculosis*. J. Antimicrob. Chemother. 47, 575-580.

Lindgren, M., Hällbrink, M., Prochiantz, A., Langel, U. 2000. Cell-penetrating peptides. Trends Pharmacol. Sci. 21, 99-103.

Lindholm, P., Goransson, U., Johansson, S., Claeson, P., Gullbo, J., Larsson, R., Bohlin, L., Backlund, A. 2002. Cyclotides: a novel type of cytotoxic agents. Mol. Cancer Ther. 1, 365-369.

Liu, F., Liu, Y., Li, F., Dong, B., Xiang, J. 2005. Molecular cloning and expression profile of putative antilipopolysaccharide factor in Chinese shrimp *Fenneropenaeus chinensis*. Mar. Biotechnol. 7, 600–608.

Liu, H., Chen, R., Zhang, Q., Wang, Q., Li Q., Peng C., Cai, L., Zheng, C., Wang, K. 2012. Characterization of two isoforms of antiliopolysacchride factors (Sp-ALFs) from the mud crab *Scylla paramamosain*. Fish Shellfish Immunol. 33, 1–10.

Liu, H., Jiravanichpaisal, P., Soderhall, I., Cerenius, L., Soderhall, K. 2006. Antilipopolysaccharide factor interferes with white spot syndrome virus replication in vitro and in vivo in the crayfish *Pacifastacus leniusculus*. J. Virol. 80, 10365–10371.

Liu, T., Gao, Y., Wang, Y., Xu, T. 2014. Characterization, evolution and functional analysis of the liver-expressed antimicrobial peptide 2 (LEAP-2) gene in miiuy croaker. Fish Shellfish Immunol. 41, 191-199.

Liu, Y., Cui, Z., Li, X., Song, C., Li, Q., Wang, S. 2012. Molecular cloning, expression pattern and antimicrobial activity of a new isoform of anti-lipopolysaccharide factor from the swimming crab *Portunus trituberculatus*. Fish Shellfish Immunol. 33, 85–91.

Liu, Y., Cui, Z., Li, X., Song, C., Shi, G. 2013. A newly identified anti-lipopolysaccharide factor from the swimming crab *Portunus trituberculatus* with broad spectrum antimicrobial activity. Fish Shellfish Immunol. 34, 463-464.

Lorenzon, S., de Guarrini, S., Smith, V.J., Ferrero, E.A. 1999. Effects of LPS injection on circulating haemocytes in crustaceans *in vivo.* Fish Shellfish Immunol. 9, 31–50.

Lorgeril, J.D., Saulnier, D., Janech, M.G., Gueguen, Y., Bachere, E. 2005. Identification of genes that are differentially expressed in haemocytes of the Pacific blue shrimp *Litopenaeus stylirostris* surviving an infection with *Vibrio penaeicida*. Physiol. Genomics. 21, 174-183.

Mader, J.S., Hoskin, D.W. 2006. Cationic antimicrobial peptides as novel cytotoxic agents for cancer treatment. Expert Opin Investig Drugs. 15, 933-946.

Maloy, W.L., Kari, U.P. 1995. Structure-activity studies on magainins and other host defense peptides. Biopolymers. 37, 105-122.

Mandal, S.M., Silva, O.N., Franco, O.L. 2014. Recombinant probiotics with antimicrobial peptides: a dual strategy to improve immune response in immunocompromised patients. Drug Discov. Today, 19, 1045–1050.

Mangoni, M.L., Rinaldi, A.C., Di Giulio, A., Mignogna, G., Bozzi, A., Barra, D., Simmaco, M. 2000. Structure-function relationships of temporins, small antimicrobial peptides from amphibian skin. Eur. J. Biochem. 267, 1447-1454.

Marr, A.K., Gooderham, W.J., Hancock, R.E. 2006. Antibacterial peptides for therapeutic use: Obstacles and realistic outlook. Curr. Opin. Pharmacol. 6, 468–472.

Matsuzaki, K. 1999. Why and how are peptide-lipid interactions utilized for self-defense? Magainins and tachyplesins as archetypes. Biochim. Biophys. Acta 1462, 1-10.

Matsuzaki, K., Sugishita, K., Fujii, N., Miyajima, K. 1995. Molecular basis for membrane selectivity of an antimicrobial peptide, magainin 2. Biochemistry. 34, 3423-3429.

Matsuzaki, K., Yoneyama, S., Fujii, N., Miyajima, K., Yamada, K., Kirino, Y., Anzai, K. 1997. Membrane permeabilization mechanisms of a cyclic antimicrobial peptide, tachyplesin I, and its linear analog. Biochemistry. 36, 9799-9806.

McManus, A.M., Nielsen, K.J., Marcus, J.P., Harrison, S.J., Green, J.L., Manners, J.M., Craik, D.J. 1999. MiAMP1, a novel rotein from Macadamia integrifolia adopts a Greek key beta-barrel fold unique amongst plant antimicrobial proteins. J. Mol. Biol. 293, 629–638.

McPhee, J.B., Scott, M.G., Hancock, R.E. 2005. Design of host defence peptides for antimicrobial and immunity enhancing activities. Comb. Chem. High Throughput Screen. 8, 257–272.

Meister, M., Lemaitre, B., Hoffmann, J.A. 1997. Antimicrobial peptide defense in *Drosophila*. Bioessays. 19, 1019–1026.

Mekata, T., Sudhakaran, R., Okugawa, S., Inada, M., Kono, T., Sakai, M., Itami, T. 2010. A novel gene of tumor necrosis factor ligand superfamily from kuruma shrimp, *Marsupenaeus japonicus*. Fish Shellfish Immunol. 28, 571–578.

Melo, M.N., Ferre, R., Castanho, M.A. 2009. Antimicrobial peptides: linking partition, activity and high membrane-bound concentrations. Nat. Rev. Microbiol. 7, 245–250.

Mitta, G., Hubert, F., Dyrynda, E.A., Boudry, P., Roch, P. 2000. Mytilin B and MGD2, two antimicrobial peptides of marine mussels: Gene structure and expression analysis. Dev. Comp. Immunol. 24, 381-393.

Miyasaki, K.T., Lehrer, R.I. 1998. β-Sheet antibiotic peptides as potential dental therapeutics. Int. J. Antimicrob. Agents. 9, 269-280.

Moerman, L., Bosteels, S., Noppe, W., Willems, J., Clynen, E., Schoofs, L., Thevissen, K., Tytgat, J., van Eldere, J., van der Walt, J., Verdonck, F. 2002. Antibacterial and antifungal properties of α-helical, cationic peptides in the venom of scorpions from Southern Africa. Eur. J. Biochem. 269, 4799–4810.

Montero, T., Vallespi, M.G., Garay, H., Reyes, O., Aran, M.J. 2003. A Limulus anti-LPS factor-derived peptide modulates cytokine gene expression and promotes resolution of bacterial acute infection in mice. Int. Immunopharmacol. 3, 247-256.

Moon, W.J., Hwang, D.K., Park, E.J., Kimb, Y.M., Chae, Y.K. 2007. Recombinant expression, isotope labeling, refolding and purification of an antimicrobial peptide, piscidin. Prot. Exp. Pur. 51, 141–146.

Moore, A.J., Devine, D.A., Bibby, M.C. 1994. Preliminary experimental anticancer activity of cecropins. Pept. Res. 7, 265-269.

Morita, T., Ohtsubo, S., Nakamura, T., Tanaka, S., Iwanaga, S., Ohashi, K., Niwa, M. 1985. Isolation and biological activities of Limulus anticoagulant (anti-LPS factor) which interacts with lipopolysaccharide (LPS). J. Biochem. 97, 1611–1620.

Moskowitz, C.H., Kewalramani, T., Nimer, S.D.,Gonzalez, M., Zelenetz, A.D., Yahalom, J. 2004. Effectiveness of high dose chemoradio-therapy and autologous stem cell transplantation for patients with biopsy-proven primary refractory Hodgkin's disease. Br. J. Haematol. 124, 645–652.

Mu, C., Zheng, P., Zhao, J., Wang, L., Qiu, L., Zhang, H., Gai, Y., Song, L. 2011. A novel type III crustin (CrusEs2) identified from Chinese mitten crab *Eriocheir sinensis*. Fish Shellfish Immunol. 31, 142–147.

Mu, C., Zheng, P., Zhao, J., Wang, L., Zhang, H., Gai, Y., Song, L. 2010. Molecular characterization and expression of a crustin-like gene from Chinese mitten crab, *Eriocheir sinensis*. Dev. Comp. Immunol. 34, 734–740.

Munoz, M., Vandenbulcke, F., Garnier, J., Gueguen, Y., Bulet, P., Saulnier, D., Bachere, E. 2004. Involvement of penaeidins in defense reactions of the shrimp *Litopenaeus stylirostris* to a pathogenic *Vibrio.* Cell Mol. Life Sci. 61, 961-972.

Munoz, M., Vandenbulcke, F., Gueguen, Y., Bachere, E. 2003. Expression of penaeidin antimicrobial peptides in early larval stages of the shrimp *Penaeus vannamei*. Dev. Comp. Immunol. 27, 283-289.

Munoz, M., Vandenbulcke, F., Saulnier, D., Bachere, E. 2002. Expression and distribution of penaeidin antimicrobial peptides are regulated by haemocyte reactions in microbial challenged shrimp. Eur. J. Biochem. 269, 2678–2689.

Muta, T., Miyata, T., Tokunaga, F., Nakamura, T., Iwanaga, S. 1987. Primary structure of anti-lipopolysaccharide factor from American horseshoe crab, *Limulus polyphemus*. J. Biochem. 1016, 1321-1330.

Mygind, P.H., Fischer, R.L., Schnorr, K.M., Hansen, M.T., Sonksen, C.P., Ludvigsen, S., Raventos, D., Buskov, S., Christensen, B., De Maria, L., Taboureau, O., Yaver, D., Elvig-Jørgensen, S.G., Sørensen, M.V., Christensen, B.E., Kjaerulff, S., Frimodt-Moller, N., Lehrer, R.I., Zasloff, M., Kristensen, H.H. 2005. Plectasin is a peptide antibiotic with therapeutic potential from a saprophytic fungus. Nature. 437, 975–980.

Nagoshi, H., Inagawa, H., Morii, K., Harada, H., Kohchi, C., Nishizawa, T., Taniguchi, Y., Uenobe, M., Honda, T., Kondoh, M., Takahashi, Y., Soma, G. 2006. Cloning and characterization of a LPS regulatory gene having an LPS binding domain in kuruma prawn *Marsupenaeus japonicus*. Mol. Immunol. 43, 2061-2069.

Nakamura, T., Furunaka, H., Miyata, T., Tokunaga, F., Muta, T., Iwanaga, S. 1988. Tachyplesin, a class of antimicrobial peptide from the hemocytes of the

horseshoe crab (*Tachypleus tridentatus*). J. Biol. Chem. 263, 16709–16713.

Nguyen, L.T., Haney, E.F., Vogel, H.J. 2011. The expanding scope of antimicrobial peptide structures and their modes of action. Trends Biotechnol. 29, 464–472.

Nile, C.J., Townes, C.L., Hirst, B.H., Hall, J. 2006. The novel avian protein, AWAK, contains multiple domains with homology to protease inhibitory modules. Mol. Immunol. 43, 388–394.

Nissen-Meyer, J., Nes, I.F. 1997. Ribosomally synthesized antimicrobial peptides: their function, structure, biogenesis, and mechanism of action. Arch. Biochem. Biophys. 167, 67-77.

O'Leary, N.A., Gross, P.S. 2006. Genomic structure and transcriptional regulation of the penaeidin gene family from *Litopenaeus vannamei*. Gene. 3711, 75-83.

Oren, Z., Lerman, J.C., Gudmundsson, G.H., Agerberth, B., Shai, Y. 1999. Structure and organization of the human antimicrobial peptide LL-37 in phospholipid membranes: relevance to the molecular basis for its non-cell-selective activity. Biochem. J. 341, 501–513.

Oren, Z., Shai, Y. 1998. Mode of action of linear amphipathic alpha- helical antimicrobial peptides. Biopolymers. 47, 451–463.

Otvos Jr. L. 2000. Antibacterial peptides isolated from insects. J. Pept. Sci. 6, 497-511.

Pan, C., Chao, T., Chen, J., Chen, J., Liu, W., Lin, C., Kuo, C. 2007. Shrimp (*Penaeus monodon*) anti-lipopolysaccharide factor reduces the lethality of *Pseudomonas aeruginosa* sepsis in mice. Int. Immunopharmacol. 7, 687–700.

Pan, C., Chen, J., Lin, T., Lin, C. 2009. *In vitro* activities of three synthetic peptides derived from epinecidin-1 and an anti-lipopolysaccharide factor against *Propioni bacteriumacnes*, *Candida albicans* and *Trichomonas vaginalis*. Peptides. 30, 1058–1068.

Pan, C., Chow, T., Yu, C., Yu, C., Chen, J., Chen, J. 2010. Peptides Antimicrobial peptides of an anti-lipopolysaccharide factor, epinecidin-1 and hepcidin reduce the lethality of *Riemerella anatipestifer* sepsis in ducks. Peptides, 31, 806–815.

Pan, C.Y., Chao, T.T., Chen, J.C., Chen, J.Y., Liu, W.C., Lin, C.H., Kuo, C.M. 2007. Shrimp *Penaeus monodon* anti-lipopolysaccharide factor reduces the lethality of *Pseudomonas aeruginosa* sepsis in mice. Int. Immunopharmacol. 75, 687-700.

Park, C.B., Yi, K.S., Matsuzaki, K., Kim, M.S., Kim, S.C. 2000. Structure-activity analysis of buforin II, a histone H2A-derived antimicrobial peptide: the proline hinge is responsible for the cell-penetrating ability of buforin II. Proc. Nat. Acad. Sci. USA. 97, 8245-8250.

Park, I.Y., Park, C.B., Kim, M.S., Kim, S.C. 1998. Parasin I, an antimicrobial peptide derived from histone H2A in the catfish, *Parasilurus asotus*. FEBS Lett. 437, 258-262.

Park, Y., Lee, D.G., Hahm, K.S. 2004. HP2–9-magainin 21–12, a synthetic hybrid peptide, exerts its antifungal effect on *Candida albicans* by damaging the plasma membrane. J. Pept. Sci. 10, 204–209.

Parkin, D.M., Bray, F., Ferlay, J., Pisani, P. 2005. Global cancer statistics, 2002. Cancer J Clin. 55, 74-108.

Pasupuleti, M., Walse, B., Svensson, B., Malmsten, M., Schmidtchen, A. 2008. Rational design of antimicrobial C3a analogues with enhanced effects against *Staphylococci* using an integrated structure and function-based approach. Biochemistry. 47, 9057-9070.

Patat, S., Carnegie, R.B., Kingsbury, C., Gross, P.S., Chapman, R. 2004. Antimicrobial activity of histones from hemocytes of the Pacific white shrimp. Eur. J. Biochem. 271, 4825-4833.

Patrzykat, A., Friedrich, C.L., Zhang, L., Mendoza, V., Hancock, R.E. 2002. Sublethal concentrations of pleurocidin derived antimicrobial peptides inhibit

macromolecular synthesis in *Escherichia coli*. Antimicrob. Agents Ch. 46, 605–614.

Perez-Paya, E., Houghten, R.A., Blondelle, S.E. 1994. Determination of the secondary structure of selected melittin analogs with different hemolytic activities. Biochem. J. 299, 587-591.

Peschel, A., Otto, M., Jack, R. W., Kalbacher, H., Jung, G., Gotz, F. 1999. Inactivation of the dlt operon in *Staphylococcus aureus* confers sensitivity to defensins, protegrins, and other antimicrobial peptides. J. Biol. Chem. 274, 8405- 8410.

Pisuttharachai, D., Yasuike, M., Aono, H., Yano, Y., Murakami, K., Kondo, H., Aoki, T., Hirono, I. 2009. Characterization of two isoforms of Japanese spiny lobster *Panulirus japonicas* defensin cDNA. Dev. Comp. Immunol. 33, 434-438.

Ponprateep, S., Tharntada, S., Somboonwiwat, K., Tassanakajon, A. 2012. Gene silencing reveals a crucial role for anti-lipopolysaccharide factors from *Penaeus monodon* in the protection against microbial infections. Fish Shellfish. Immunol. 32, 26-34.

Porro, D., Sauer, M., Branduardi, P., Mattanovich, D. 2005. Recombinant protein production in yeasts. Mol. Biotechnol. 31, 245–259.

Powers, J.P., Hancock, R.E.W. 2003. The relationship between peptide structure and antibacterial activity. Peptides. 24, 1681-1691.

Prenner, E.J., Lewis, R.N.A.H., Kondejewski, L.H., Hodges, R.S., McElhaney, R.N. 1999. Differential scanning salorimetric study on the effect of the antimicrobial peptide Gramicidin S on the thermotropic phase behavior of phosphatidyl choline, phosphatidyl ethanolamine and phosphatidyl glycerol lipid bilayer membranes. Biochim. Biophys. Acta. 1417, 211-223.

Prinz, W.A., Aslund, F., Holmgren, A., Beckwith, J. 1997. The role of the thioredoxin and glutaredoxin pathways in reducing protein disulfide bonds in the *Escherichia coli* cytoplasm. J. Biol. Chem. 272, 15661-15667.

Pushpanathan, M., Gunasekaran, P., Rajendhran, J. 2013. Antimicrobial Peptides: Versatile Biological Properties . Int. J. Pep. 675391, 1-15.

Pushpanathan, M., Rajendhran, J., Jayashree, S., Sundarakrishnan, B., Jayachandran, S., Gunasekaran, P. 2012. Direct cell penetration of the antifungal peptide, MMGP1, in *Canadida albicans*. J. Pept. Sci. 18, 657–660.

Raikhel, N.V., Lee, H.I., Broekaert, W.F. 1993. Structure and functions of chitin-binding proteins. Ann. Rev. Plant Physiol. Plant Mol. Biol. 44, 591–615.

Ranganathan, S., Simpson, K.J., Shaw, D.C., Nicholas, K. 1999. The whey acidic protein family: a new signature motif and three dimensional structure by comparative modeling. J. Mol. Graph Model. 17, 106–113.

Rao, A. 1999. Conformation and antimicrobial activity of linear derivatives of tachyplesin lacking disulfide bonds. Arch. Biochem. Biophys. 361, 127–134.

Rao, X., Hu, J., Li, S., Jin, X., Zhang, C., Cong, Y., Hu, X., Tan, Y., Huang, J., Chen, Z., Zhu, J., Hu, F. 2005. Design and expression of peptide antibiotic hPAB-beta as tandem multimers in *Escherichia coli*. Peptides. 26, 721–729.

Rattanachai, A., Hirono, I., Ohira, T., Takahashi, Y., Aoki, T. 2004. Cloning of kuruma prawn *Marsupenaeus japonicas* crustin-like peptide cDNA and analysis of its expression. Fish. Sci. 70, 765–771.

Reddy, K.V., Yedery, R.D., Aranha, C. 2004. Antimicrobial peptides: Premises and promises. Int. J. Antimicrob. Agents, 24, 536-547.

Relf, J.M., Chisholm, J.R.S., Kemp, G.D., Smith, V.J. 1999. Purification and characterization of a cysteine-rich 11.5 kDa antibacterial protein from the granular hemocytes of the shore crab, *Carcinus maenas*. Eur. J. Biochem. 264, 350–357.

Ren, Q., Du, Z.Q., Li, M., Zhang, C.Y., Chen, K.P. 2012. Cloning and expression analysis of an anti-lipopolysaccharide factor from giant freshwater prawn, *Macrobrachium rosenbergii.* Mol. Biol. Rep. 39, 7673-7680.

Rhoobunjongde, W., Hatai, K., Wadas, S., Kubota, S. 1991. Fusarium-moniliforme (sheldon) isolated from gills of kuruma prawn *Penaeus japonicus* (bate) with black gill disease. Nippon Suisan Gakkaishi. 57, 629-635.

Richard, C., Drider, D., Elmorjani, K., Marion, D., Prevost, H. 2004. Heterologous expression and purification of active divercin V41, a class IIa bacteriocin encoded by a synthetic gene in *Escherichia coli*. J. Bacteriol. 186, 4276–4284.

Ried, C., Wahl, C., Miethke, T., Wellnhofer, G., Landgraf, C., Schneider-Mergener, J., Hoess, A. 1996. High affinity endotoxin-binding and neutralizing peptides based on the crystal structure of recombinant Limulus antilipopolysaccharide factor. J. Biol. Chem. 271, 28120–28127.

Rinaldi, A.C., Di Giulio, A., Liberi, M., Gualtieri, G., Simmaco, M., Barra, D., Bozzi, A. 2001. Effects of temporins on molecular dynamics and membrane permeabilization in lipid vesicles. J. Pept. Res. 58, 213-220.

Rinaldi, A.C., Mangoni, M.I., Rufo, A., Luzi, C., Simmaco, M., Barra, D., Zhao, H., Kinnunen, P.K.J., Bozzi, A., Di Giulio, A. 2002. Temporin L: antimicrobial, cytotoxic activities and effects on membrane permeabilization in lipid vesicles. Biochem. J. 368, 91-100.

Rojtinnakorn, J., Hirono, I., Itami, T., Takahashi, Y., Aoki, T. 2002.Gene expression in haemocytes of kuruma prawn, *Penaeus japonicus*, in response to infection with WSSV by EST approach. Fish Shellfish Immunol. 13, 69–83.

Rolland, J.L., Abdelouahab, M., Dupont, J., Lefevre, F., Bachère, E., Romestand, B. 2010. Stylicins, a new family of antimicrobial peptides from the Pacific blue shrimp *Litopenaeus stylirostris*. Mol. Immunol. 47, 1269–1277.

Rosa, R.D., Bandeira, P.T., Barracco, M.A. 2007. Molecular cloning of crustins from the hemocytes of Brazilian penaeid shrimps. FEMS Microbiol. Lett. 274, 287–290.

Rosa, R.D., Barracco, M.A. 2010. Antimicrobial peptides in crustaceans. Inver. Surv. J. *7*, 262–284.

Rosa, R.D., Stoco, P.H., Barracco, M.A. 2008. Cloning and characterisation of cDNA sequences encoding for anti-lipopolysaccharide factors ALFs in Brazilian palaemonid and penaeid shrimps. Fish Shellfish Immunol. 25, 693–696.

Rosa, R.D., Vergnes, A., de Lorgeril, J., Goncalves, P., Perazzolo, L.M., Sauné, L., Romestand, B., Fievet, J., Gueguen, Y., Bachère,E., Destoumieux, D. 2013. Functional divergence in shrimp anti-lipopolysaccharide factors (ALFs): from recognition of cell wall components to antimicrobial activity. PLoS ONE 8(7), e67937.

Rotem, S., Mor, A. 2009. Antimicrobial peptide mimics for improved therapeutic properties. Biochim. Biophys. Acta. 1788, 1582-1592.

Rozek, A., Friedrich, C.L., Hancock, R.E.W. 2000. Structure of the bovine antimicrobial peptide indolicidin bound to dodecylphosphocholine and sodium dodecl sulfate micelles. Biochemistry. 39, 15765–15774.

Salzet, M. 2001. Vertebrate innate immunity resembles a mosaic of invertebrate immune responses. Trends Immunol. 22, 285-288.

Samakovlis, C., Kylsten, P., Kimbrell, D.A., Engstrom, A., Hultmark, D. 1991. The andropin gene and its product, a male-specific antibacterial peptide in *Drosophila melanogaste*r. EMBO J. 10, 163-169.

Schalkwijk, J., Wiedow, O., Hirose, S. 1999. The trappin gene family: proteins defined by an N-terminal transglutaminase substrate domain and a C-terminal four-disulphide core. Biochem. J. 340, 569–77.

Schein, C.H., Noteborn, M.H.M. 1988. Formation of soluble recombinant proteins in *Escherichia coli* is favored by lower growth temperature. Nat. Biotechnol. 6, 291-294.

Schnapp, D., Kemp, G.D., Smith, V.J. 1996. Purification and characterization of a proline-rich antibacterial peptide, with sequence similarity to bactenecin-7, from the haemocytes of the shore crab, *Carcinus maenas*. Eur. J. Biochem. 2403, 532-539.

Schwede, T., Kopp, J., Guex, N., Peitsch, M.C. 2003. SWISS-MODEL: an automated protein homology-modeling server. Nucleic Acids Res. 31, 3381-3385.

Scocchi, M., Bontempo, D., Boscolo, S., Tomasinsig, L., Giulotto, E., Zanetti, M. 1999. Novel cathelicidins in horse leukocytes. FEBS Lett. 457, 459-464.

Selsted, M.E., Novotny, M.J., Morris, W.L., Tang, Y.Q., Smith, W., Cullen, J.S. 1992. Indolicidin, a novel bactericidal tridecapeptide amide from neutrophils. J. Biol. Chem. 267, 4292-4295.

Seo, P.J., Park, M.J., Lim, M.H., Kim, S.G., Lee, M., Baldwin, I.T., Park, C.M. 2012. A self-regulatory circuit of circadian clock-associated underlies the circadian clock regulation of temperature responses in *Arabidopsis*. Plant Cell. 24, 2427–2442.

Shagger, H. 2006. Tricine-SDS- PAGE. Nat. Prot. 1, 16-23.

Shai, Y. 1999. Mechanism of the binding, insertion and destabilization of phospholipid bilayer membranes by alphahelical antimicrobial and cell non-selective membrane-lytic peptides. Biochim. Biophys. Acta. 1462, 55–70.

Shai, Y. 2002. Mode of action of membrane active antimicrobial peptides. Biopolymers. 66, 236–248.

Sharma, S., Yedery, R.D., Patgaonkar, M.S., Selvaakumar, C., Reddy, K.V., 2011. Antibacterial activity of a synthetic peptide that mimics the LPS binding domain of Indian mud crab, *Scylla serrata* anti-lipopolysaccharide factor (SsALF) also involved in the modulation of vaginal immune functions through NF-kB signaling. Microb. Pathog. 50, 179–191.

Shi, J., Ross, C.R., Chengappa, M.M., Sylte, M.J., McVey, D.S., Blecha, F. 1996. Antibacterial activity of a synthetic peptide PR-26 derived from PR-39, a proline–arginine-rich neutrophil antimicrobial peptide. Antimicrob. Agents Ch. 40, 115–121.

Sitaram, N., Nagaraj, R. 1999. Interaction of antimicrobial peptides with biological and model membranes: structural and charge requirements for activity. Biochim. Biophys. Acta. 1462, 29–54.

Smith, V.J., Brown, J.H., Hauton, C. 2003. Immunostimulations in crustaceans: does it really protect against infection? Fish Shellfish Immunol. 15, 71-90.

Smith, V.J., Desbois, A.P., Dyrynda, E.A. 2010. Conventional and unconventional antimicrobials from fish, marine invertebrates and micro-algae. Mar. Drugs. 8, 1213–1262.

Smith, V.J., Fernandes, J.M., Kemp, G.D., Hauton, C. 2008. Crustins: Enigmatic WAP domain-containing antibacterial proteins from crustaceans. Dev. Comp. Immunol. 32, 758-772.

Somboonwiwat, K., Bachere, E., Rimphanitchayakit, V., Tassanakajon, A. 2008. Localization of anti-lipopolysaccharide factor ALF*Pm*3 in tissues of the black tiger shrimp, *Penaeus monodon*, and characterization of its binding properties. Dev. Comp. Immunol. 3210, 1170-1176.

Somboonwiwat, K., Marcos, M., Tassanakajon, A., Romestand, B., Gueguen, Y. 2005. Recombinant expression and anti-microbial activity of anti-lipopolysaccharide factor (ALF) from the black tiger shrimp *Penaeus monodon*, 29, 841–851.

Somboonwiwat, K., Supungul, P., Rimphanitchayakit, V., Aoki, T., Hirono, I., Tassanakajon, A. 2006. Differentially expressed genes in hemocytes of *Vibrio harveyi*-challenged shrimp *Penaeus monodon*. J. Biochem. Mol. Biol. 391, 26-36.

Sorensen, H.P., Mortensen, K.K. 2005. Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. J. Biotech. 115, 113–128.

Sperstad, S.V., Haug, T., Vasskog, T., Stensvåg, K. 2009. Hyastatin, a glycine-rich multi-domain antimicrobial peptide isolated from the spider crab (*Hyas araneus*) hemocyes. Mol. Immunol. 46, 2604–2612.

Splith, K., Neundorf, I. 2011. Antimicrobial peptides with cell-penetrating peptide properties and vice versa. Eur. Biophys. J. 40, 387–397.

Stefano, G.B., Salzet, B., Fricchione, G.L. 1998. Enkelytin and opioid peptide association in invertebrates and vertebrates: immune activation and pain. Immunol. Today. 19, 265–268.

Stensvag, K., Haug, T., Sperstad, S.V., Rekdal, O., Indrevoll, B., Styrvold, O,B. 2008. Arasin 1, a proline-arginine-rich antimicrobial peptide isolated from the spider crab, *Hyas araneus*. Dev. Comp. Immunol. 32, 275–285.

Stoss, T.D., Nickell, M.D., Hardin, D., Derby, C., Mc Clintock, T.S. 2003. Inducible transcript expressed by reactive epithelial cells at sites of olfactory sensory neuron proliferation. J. Neurobiol. 58, 355-368.

Stoss, T.D., Nickell, M.D., Hardin, D., Derby, C.D., McClintock, T.S. 2004. Inducible transcript expressed by reactive epithelial cells at sites of olfactory sensory neuron proliferation. J. Neurobiol. 58, 355–368.

Subbalakshmi, C., Sitaram, N. 1998. Mechanism of antimicrobial action of indolicidin. FEMS Microbiol. Lett. 160, 91-96.

Sugumaran, M. 2002. Comparative biochemistry of eumelanogenesis and the protective roles of phenoloxidase and *melanin* in insects. Pigment Cell Res. 15, 2-9.

Sun, C., Du, X.J., Xu, W.T., Zhang, H.W., Zhao, X.F., Wang, J.X. 2010. Molecular cloning and characterization of three crustins from the Chinese white shrimp, *Fenneropenaeus chinensis*. Fish Shellfish Immunol. 28, 517-524.

Sun, C., Xu, W.-T., Zhang, H.-W., Dong, L.-P., Zhang, T., Zhao, X.-F., Wang, J.X. 2011. An anti-lipopolysaccharide factor from red swamp crayfish, *Procambarus clarkii*, exhibited antimicrobial activities *in vitro* and *in vivo*. Fish Shellfish Immunol. 30, 295–303.

Sun, W., Wan, W., Zhu, S., Wang, S., Wang, S., Wen, X., Li, S. 2015. Characterization of a novel anti-lipopolysaccharide factor isoform (SpALF5) in mud crab, *Scylla paramamosain*. Mol. Immunol. 64, 262–275.

Supungul, P., Klinbunga, S., Pichyangkura, R., Hirono, I., Aoki, T., Tassanakajon, A. 2004. Antimicrobial peptides discovered in the black tiger shrimp *Penaeus monodon* using the EST approach. Dis. Aquat. Organ. 6, 123-135.

Supungul, P., Klinbunga, S., Pichyangkura, R., Jitrapakdee, S., Hirono, I., Aoki, T., Tassanakajon, A. 2002. Identification of immune-related genes in

hemocytes of black tiger shrimp *Penaeus monodon*. Mar. Biotechnol. 4, 487-494.

Supungul, P., Tang, S., Maneeruttanarungroj, C., Rimphanitchayakit, V., Hirono, I., Aoki, T., Tassanakajon, A. 2008. Cloning, expression and antimicrobial activity of crustin*Pm*1, a major isoform of crustin, from the black tiger shrimp *Penaeus monodon*. Dev. Comp. Immunol. 32, 61–70.

Talas-Ogras, T. 2004. Screening antimicrobial activities of basic protein fractions from dry and germinated wheat seeds. Biol. Plant. 48, 583–588.

Tamamura, H., Ishihara, T., Otaka, A., Murakami, T., Ibuka, T., Waki, M., Matsumoto, A., Yamamoto, N., Fujii, N. 1996. Analysis of the interaction of an anti-HIV peptide, T22 [Tyr5, 12, Lys7]-polyphemusin II, with gp120 and CD4 by surface plasmon resonance. Biochim. Biophys. Acta. 1298, 37-44.

Tamamura, H., Waki, M., Imai, M., Otaka, A., Ibuka, T., Waki, K., Miyamoto, K., Matsumoto, A., Murakami, T., Nakashima, H., Yamamoto, N., Fujii, N. 1998. Downsizing of an HIV-cell fusion inhibitor, T22 [Tyr5,12, Lys7]-polyphemusin II, with the maintenance of anti-HIV activity and solution structure. Bio-org. Med. Chem. 6, 473-479.

Tamura, K., Dudley, J., Nei, M., Kumar, S. 2007. MEGA 4: Molecular Evolutionary Genetics Analysis MEGA software version 4.0. Mol. Biol. Evol. 24, 1596-1599.

Tanaka, S., Nakamura, T., Morita, T., Iwanaga, S. 1982. Limulus anti-LPS factor: an anticoagulant which inhibits the endotoxin mediated activation of limulus coagulation system. Biochem. Biophys. Res. Commun. 105, 717–723.

Tang, T., Li, L., Sun, L., Bu, J., Xie, S., Liu, F. 2014. Functional analysis of *Fenneropenaeus chinensis* anti-lipopolysaccharide factor promoter regulated by lipopolysaccharide and (1,3)-β-D-glucan. Fish Shellfish Immunol. 38, 348–353.

Tassanakajon, A., Amparyup, P., Somboonwiwat, K., Supungul, P. 2010. Cationic antimicrobial peptides in penaeid shrimp. Mar. Biotechnol. 12, 487-505.

Tenenholz, T.C., Klenk, K.C., Matteson, D.R., Blaustein, M.P., Weber, D.J. 2000. Structural determinants of scorpion toxin affinity: the charybdotoxin α-KTX family of K+-channel blocking peptides. Rev. Physiol. Biochem. Pharmacol. 140, 135-185.

Terras, F.R.G., Schoofs, H.M.E., De Bolle, M.F.C., Van Leuven, F., Rees, S.B., Vanderleyden, J., Cammue, B.P.A., Broekaert, W.F. 1992. Analysis of two novel classes of antifungal proteins from the radish *Raphanus satilus* L. seeds. J. Biol. Chem. 267, 15301–15309.

Tharntada, S., Ponprateep, S., Somboonwiwat, K., Liu, H., Soderhall, I., Soderhall, K., Tassanakajon, A. 2009. Role of anti-lipopolysaccharide factor from the black tiger shrimp, *Penaeus monodon*, in protection from white spot syndrome virus infection. J. Gen. Virol. 90, 1491-1498.

Torrent, M., Andreu, D., Nogués, V.M., Boix, E. 2011. Connecting peptide physicochemical and antimicrobial properties by a rational prediction model. PloS One. 6, e16968, 201.

Tossi, A., Sandri, L. 2002. Molecular diversity in gene-encoded, cationic antimicrobial polypeptides. Curr. Pharm. Des. 8, 743–761.

Towle, D.W., Smith, C.M. 2006. Gene discovery in *Carcinus maenas* and *Homarus americanus* via expressed sequence tags. Integr. Comp. Biol. 46, 912-918.

Tsai, H., Bobek, L.A. 1998. Human salivary histatins: promising anti-fungal therapeutic agents. Crit. Rev. Oral Biol. Med. 9, 480-497.

Turner, J.D., Rouser, G. 1970. Removal of lipid from intact erythrocytes and ghosts by aqueous solutions. Anal. Biochem. 38, 437-445.

Tytler, E.M., Anantharamaiah, G.M., Walker, D.E., Mishra, V.K., Palgunachari, M.N., Segrest, J.P. 1995. Molecular basis for prokaryotic specificity of magainin-induced lysis. Biochemistry. 34, 4393-4401.

Underhill, D.M., Orinsky, A. 2002. Toll-like receptors: key mediators of microbe detection. Curr. Opin. Immunol. 14, 103-110.

Uteng, M., Hauge, H.H., Markwick, P.R., Fimland, G., Mantzilas, D., Nissen-Meyer, J., Muhle-Goll, C. 2003. Three-dimensional structure in lipid micelles of the pediocin-like antimicrobial peptide sakacin P and a sakacin P variant that is structurally stabilized by an inserted C-terminal disulfide bridge. Biochemistry. 42, 11417–11426.

Vallespi, M.G., Alvarez-Obregon, J.C., Rodriguez-Alonso, I., Montero, T., Garay, H. 2003. A Limulus anti-LPS factor-derived peptide modulates cytokine gene expression and promotes resolution of bacterial acute infection in mice. Int. Immunopharmacol. 3, 247–256.

Vallespi, M.G., Glaria, L.A., Reyes, O., Garay, H.E., Ferrero, J., Arana, M.J. 2000. A Limulus anti-lipopolysaccharide factor derived peptide exhibits a new immunological activity with potential applicability in infectious diseases. Clin. Diagn. Lab. Immunol. 7, 669–675.

Van der Weerden, N.L., Hancock, R.E., Anderson, M.A. 2010. Permeabilization of fungal hyphae by the plant defensin NaD1 occurs through a cell wall-dependent process. J. Biol. Chem. 285, 37513–37520.

van't Hof, W., Veerman, E.C., Helmerhorst, E.J., Amerongen, A.V. 2001. Antimicrobial peptides: Properties and applicability. Biol. Chem. 382, 597–619.

Vandenbergh, P.A., Pucci, M.J., Kunka, B.S., Vedamuthu, E.R. 1989. Method for inhibiting *Listeria monocytogenes* using a bacteriocin. European Patent Application 89101125.6.

Vargas-Albores, F., Yepiz-Plascencia, G., Jimenez-Vega, F., Avila-Villa, A. 2004. Structural and functional differences of *Litopenaeus vannamei* crustins. Comp. Biochem. Physiol. B-Biochem. Mol. Biol. 138, 415-422.

Vaseeharan, B., Shanthi S., Chen J.C., Espineira M. 2012. Molecular cloning, sequence analysis and expression of Fein-Penaeidin from the haemocytes of Indian white shrimp *Fenneropenaeus indicus*. Results Immunol. 2, 35–43.

Vatanavicharn, T., Supungul, P., Puanglarp, N., Yingvilasprasert, W., Tassanakajon, A. 2009. Genomic structure, expression pattern and functional characterization of crustinPm5, a unique isoform of crustin from *Penaeus monodon*. Comp Biochem Physiol B. Biochem. Mol. Biol. 153, 244–252.

Verkleij, A.J., Zwaal, R.F.A., Roelofsen, B., Comfurius, P., Kastelijn, D., Van Deenen, L.L.M. 1973. The asymmetric distribution of phospholipids in the human red cell membrane. A combined study using phospholipases and freeze-etch electron microscopy. Biochim. Biophys. Acta. 323, 178–193

Vogel, H.J., Schibli, D.J., Jing, W., Lohmeier-Vogel, E.M., Epand, R.F., Epand, R.M. 2002. Towards a structure-function analysis of bovine lactoferricin and related tryptophan and arginine-containing peptides. Biochem. Cell Biol. 80, 49-63.

Wachinger, M., Kleinschmidt, A., Winder, D., Von Pechmann, N., Ludvigsen, A., Neumann, M., Holle, R., Salmons, B., Erfle, V., Brack-Werner, R. 1998. Antimicrobial peptides melittin and cecropin inhibit replication of human immunodeficiency virus 1 by suppressing viral gene expression. J. Gen. Virol. 79, 731-740.

Wang, H.C., Wang, H.C., Leu, J.H., Kou, G.H., Wang, A.H.J., Lo, C.F. 2007. Protein expression profiling of the shrimp cellular response to white spot syndrome virus infection. Dev. Comp. Immunol. 31, 672-686.

Wang, Q., Zhu, F., Xin, Y., Liu, J., Luo, L., Yin, Z. 2011. Expression and purification of antimicrobial peptide buforin IIb in *Escherichia coli*. Biotechnol. Lett. 33, 2121-2126.

Wang, Y.G., Lee, K.L., Najiah, M., Shariff, M., Hassan, M.D. 2002. A new bacterial white spot syndrome BWSS in cultured tiger shrimp *Penaeus monodon* and its comparison with white spot syndrome WSS caused by virus. Dis. Aquat. Organ. 411, 9–18.

Warren, H.S., Glennon, M.L., Wainwright, N., Amato, S.F., Black, K.M., Kirsch, S.J., Riveau, G.R., Whyte, R.I., Zapol, W.M., Novitsky, T.J. 1992. Binding and

neutralization of endotoxin by Limulus anti-lipopolysaccharide factor. Infect. Immun. 60, 2506–2513.

Wasiluk, K.R., Leslie, D.B., Vietzen, P.S., Mayo, K.H., Dunn, D.L. 2004. Structure/function studies of an endotoxin-neutralizing peptide derived from bactericidal/permeability-increasing protein. Surgery. 136, 253-260.

Weiss, C.A., Wasiluk, K.R., Kellogg, T.A., Dunn, D.L. 2000. Bactericidal and endotoxin neutralizing activity of a peptide derived from Limulus anti-lipopolysaccharide factor. Surgery 128, 339–344.

Wiesner, J., Vilcinskas, A. 2010. Antimicrobial Peptides: The Ancient Arm of the Human Immune System. Virulence. 5, 440-464.

Xu, Q.H., Liu, Y. 2011. Gene expression profiles of the swimming crab *Portunus trituberculatus* exposed to salinity stress. Mar. Biol. 158, 2161-2172.

Yang, Y., Boze, H., Chemardin, P., Padilla, A., Moulin, G., Tassanakajon, A. 2009. NMR structure of rALF-Pm3, an anti-lipopolysaccharide factor from shrimp: model of the possible lipid A-binding site. Biopolymers. 91, 207-220.

Yang, Y., Poncet, J.L., Garnier, J., Zatylny, C., Bachere, E., Aumelas, A. 2003. Solution structure of the recombinant Penaeidin-3, a shrimp antimicrobial peptide. J. Biol. Chem. 278, 36859–36867.

Yang, Y.S., Mitta, G., Chavanieu, A., Calas, B., Sanchez, J.F., Roch, P., Aumelas, A., 2000. Solution structure and activity of the synthetic four-disulfide bond Mediterranean mussel defensin (MGD-1). Biochemistry. 39, 14436–1444.

Yeaman, M.R., Yount, N.Y. 2003. Mechanisms of antimicrobial peptide action and resistance. Pharmacol. Rev. 55, 27-55.

Yedery, R.D., Reddy, K.V.R. 2009. Identification, cloning, characterization and recombinant expression of an anti-lipopolysaccharide factor from the hemocytes of Indian mud crab, *Scylla serrata*. Fish Shellfish Immunol. 27, 275–84.

Yokoyama, S., Iida, Y., Kawasaki, Y., Minami, Y., Watanabe, K., Yagi, F. 2009 The chitin-binding capability of Cy-AMP1 from cycad is essential to antifungal activity. J. Pept. Sci. 15, 492–497.

Yount, N.Y., Bayer, A.S., Xiong, Y.Q., Yeaman, M.R. 2006. Advances in antimicrobial peptide immunobiology. Biopolymers. 84, 435-458.

Yount, N.Y., Bayer, A.S., Xiong, Y.Q., Yeaman, M.R. 2006. Advances in antimicrobial peptide immunobiology. Biopolymers. 84, 435-458.

Yount, N.Y., Yeaman, M.R. 2005. Immunocontinuum: perspectives in antimicrobial peptide mechanisms of action and resistance. Prot. Pep. Lett. 12, 49-67.

Yue, F., Pan, L., Miao, J., Zhang, L., Li, J. 2010. Molecular cloning, characterization and mRNA expression of two antibacterial peptides, crustin and anti-lipopolysaccharide factor in swimming crab *Portunus trituberculatus*. Comp. Biochem. Physiol. 156B, 77–85.

Zaltash, S., Palmblad, M., Curstedt, T., Johansson, J., Persson, B. 2000. Pulmonary surfactant protein B: a structural model and a functional analogue. Biochim. Biophys. Acta. 1466, 179-186.

Zasloff, M. 1987. Magainins, a class of antimicrobial peptides from *Xenopus* skin: isolation, characterization of two active forms, and partial cDNA sequence of a precursor. Proc. Natl. Acad. Sci. USA. 84, 5449-5453.

Zasloff, M. 2002. Antimicrobial peptides of multicellular organisms. Nature. 415, 389-395

Zhang, J., Li, F., Wang, Z., Xiang, J. 2007. Cloning and recombinant expression of a crustin-like gene from Chinese shrimp, *Fenneropenaeus chinensis*. J. Biotech. 127, 605-614.

Zhang, L., Falla, T.J. 2010. Potential therapeutic application of host defense peptides. Methods Mol. Biol. 618, 303-327.

Zhang, L., Rozek, A., Hancock, R.E.W. 2001. Interaction of cationic antimicrobial peptides with model membranes. J. Biol. Chem. 276, 35714– 35722.

Zhao, C., Ganz, T., Lehrer, R.I. 1995. Structures of genes for two cathelin-associated antimicrobial peptides: prophenin-2 and PR-39. FEBS Lett. 376, 130-134.

Zhu, L., Lan, J-F., Huang, Y-Q., Zhang, C., Zhou, J-F., Fang, W-H., Yao, X-J., Wang, H., Li, X-C. 2014. SpALF4: A newly identified anti-lipopolysaccharide factor from the mud crab *Scylla paramamosain* with broad spectrum antimicrobial activity, Fish Shellfish Immunol. 36, 172-180.

GenBank Submissions

1. GenBank accession number **HQ638024.** 2011. **Afsal,V.V.**; Swapna P. Antony; Rosamma Philip. *Scylla serrata* anti-lipopolysaccharide factor mRNA, partial cds.
2. GenBank accession number **JQ745295.** 2012. **Afsal,V.V.**, Swapna P. Antony, Bright Singh, I. S. and Philip, R. *Portunus pelagicus* anti-lipopolysaccharide factor (ALF) mRNA, complete cds.
3. GenBank accession number **JQ899452.** 2012. **Afsal,V.V.**, Swapna P. Antony and Philip, R. *Portunus pelagicus* anti-lipopolysaccharide factor (ALF) precursor, mRNA, complete cds.
4. GenBank accession number **JQ899453.** 2012. **Afsal,V.V.**, Swapna P. Antony and Philip, R. *Scylla sp.*anti-lipopolysaccharide factor (ALF) precursor, mRNA, complete cds.
5. GenBank accession number **HQ638025.** 2011. **Afsal,V.V.**; Swapna P. Antony; Rosamma Philip. *Scylla serrata* crustin mRNA, complete cds.
6. GenBank accession number **JQ753312.** 2012. **Afsal,V.V.**; Swapna P. Antony; Bright Singh, I. S. and Rosamma Philip. *Scylla sp.*crustin mRNA, complete cds.
7. GenBank accession number **JQ965930.** 2012. **Afsal,V.V.**; Swapna P. Antony and Rosamma Philip. *Portunus pelagicus* crustin, mRNA, complete cds.
8. GenBank accession number **JX657680.** 2013. **Afsal,V.V.**, Swapna P. Antony and Philip, R. *Fenneropenaeus indicus* penaeidin mRNA, complete cds.
9. GenBank accession number **KF275674.** 2013**. Afsal,V.V.**, Swapna P. Antony and Philip, R. *Metapenaeus monoceros* penaeidin mRNA, complete cds.

Publications

List of Publications

1. **Afsal V.V.**; Swapna P. Antony; Naveen Sathyan; I. S. Bright Singh; Rosamma Philip. 2011. Molecular characterization and phylogenetic analysis of two antimicrobial peptides: Anti-lipopolysaccharide factor and Crustin from the brown mud crab, *Scylla serrata*. **Results in Immunology**. 1, 6-10.

2. **Afsal V.V**., Swapna P. Antony, V.N. Sanjeevan, P.R. Anil Kumar, I.S. Bright Singh, Rosamma Philip. 2012. A new isoform of anti-lipopolysaccharide factor identified from the blue swimmer crab, *Portunus pelagicus*: Molecular characteristics and phylogeny. **Aquaculture**. 356-357, 119–122.

Results in Immunology 1 (2011) 6–10

Contents lists available at ScienceDirect

Results in Immunology

journal homepage: www.elsevier.com/locate/rinim

Short Communication

Molecular characterization and phylogenetic analysis of two antimicrobial peptides: Anti-lipopolysaccharide factor and crustin from the brown mud crab, *Scylla serrata*

V.V. Afsal, Swapna P. Antony, Naveen Sathyan, Rosamma Philip *

Department of Marine Biology, Microbiology and Biochemistry, School of Marine Sciences, Cochin University of Science and Technology (CUSAT), Fine Arts Avenue, Kochi 682 016, Kerala, India

ARTICLE INFO

Article history:
Received 2 June 2011
Received in revised form
2 June 2011
Accepted 6 June 2011
Available online 8 July 2011

Keywords:
Antimicrobial peptide
Scylla serrata
Crustin
Anti-lipopolysaccharide factor
Innate immunity

ABSTRACT

AMPs are evolutional weapons, widely used by animals and plants in their innate immune system to fend off invading microbes. The present study reports characterization of a new ALF isoform (Sc-ALF; **HQ638024**) and the first crustin (Sc-crustin; **HQ638025**) from the mud crab, *Scylla serrata*. The full-length cDNA of Sc-ALF consisted of 477 bp with an ORF of 123 amino acids and a putative signal peptide of 26 amino acids. Sc-ALF had a predicted molecular weight (MW) of 11.17 kDa and theoretical isoelectric point (*pI*) of 9.95. Two highly conserved cysteine residues and putative LPS binding domain were observed in Sc ALF. Comparison of amino acid sequences with neighbor joining tree indicated that Sc-ALF shared maximum similarity with ALF of *S. paramamosain*. Peptide model of Sc-ALF created using SWISS-MODEL server was found to consist of two α-helices crowded against a four-strand β-sheet. The full-length cDNA of Sc-crustin consisted of 433 base pairs with an ORF of 111 amino acids and a putative signal peptide of 21 amino acids. Comparison of amino acid sequences with a neighbor-joining tree revealed that Sc-crustin shared high identity with other known crustins characterized from *S. paramamosain*, *P. trituberculatus*, *H. araneus*, *C. maenas* and *F. chinensis*. A whey-acidic-protein domain could be detected at the C-terminus with the characteristic four disulfide core. Sc-crustin had a predicted MW of 10.24 kDa and a *pI* of 8.76. Peptide model of Sc-crustin created using SWISS-MODEL server indicated a random coiled structure that is with two possible β-sheets but no helices.

1. Introduction

The economic role of marine and freshwater crustaceans as a food source in the export market demands the need to augment fishery resources through the adoption of intensive culture practices. This has led to physiological stress on cultured organisms, often increasing the incidence of diseases [1]. Despite the development of safe and potent antibiotics, bacterial diseases remain a worldwide health crisis due to the emergence of multiple drug resistant pathogens [2]. The use of antimicrobial peptides (AMPs) as a therapeutic tool has been among the most promising avenues investigated to date for addressing antibiotic resistance [3]. AMPs are found in a wide range of prokaryotic and eukaryotic organisms from plants to human beings [4–7]. In crabs, several AMPs have been isolated and characterized, viz. the 6.5 kDa AMP and a cysteine-rich 11.5 kDa AMP (carcinin) from the shore crab, *Carcinus maenas* [8,9], callinectin from the blue crab, *Callinectes sapidus* [10], scygonadin, an anionic antimicrobial peptide from the mud crab, *Scylla serrata* [11], anti-lipopolysaccharide factor (ALF) and crustin from the mud crab, *Scylla paramamosain* [12,13] and arasin and hyasin from the spider crab, *Hyas araneus* [14]. Hemocytes have been proved to be the site of production and storage of AMPs in invertebrates such as horseshoe crabs, mussels and decapod crustaceans [9,15–18].

The brown mud crab, *Scylla serrata*, is a decapod crustacean of the brachyuran family, Portunidae. They are tolerant to wide range of environmental parameters and there has been a huge interest in the aquaculture of this species due to its high demand/price, high flesh content and rapid growth rates in captivity. Mud crabs are a highly delicious seafood commodity and are therefore an important candidate species for aquaculture. *S. serrata* is a well known commercial species in India, Philippines and Vietnam. Like many other decapod crustaceans [8,19–21], the mud crab also possesses broad-spectrum antibacterial activity in its hemolymph that constitutes part of its nonspecific defenses. However, there are hardly any published works on major AMP families (ALF and crustins) present in this species. The present study aims at the

* Corresponding author. Tel.: +91 484 2368120; fax: +91 484 2381120.
E-mail addresses: rose@cusat.ac.in, rosammap@gmail.com (R. Philip).

2211-2839/$ - see front matter © 2011 Elsevier Ltd All rights reserved.
doi:10.1016/j.rinim.2011.06.001

molecular characterization and phylogenetic analysis of two major families of AMPs, viz. anti-lipopolysaccharide factor and crustin from *S. serrata*.

2. Materials and methods

Healthy adults of *S. serrata* (~300 g body weight) were collected from the backwater stream along Vypeen, Kochi, India. Hemolymph was collected from the base of abdominal appendages using specially designed capillary tubes (RNase-free) rinsed using pre-cooled anticoagulant solution (RNase-free 10% sodium citrate, pH 7.0). Total RNA was extracted from the hemocytes using TRI reagent (Sigma) following manufacture's protocol. RNA was quantified by spectrophotometry at 260 and 280 nm. Only RNAs with absorbance ratios (A_{260}:A_{280}) greater than 1.8 were used for the present work. First strand cDNA was generated in a 20 μl reaction volume containing 5 μg total RNA, 1*x* RT buffer, 2 mM dNTP, 2 μM oligo d(T_{20}), 20 U of RNase inhibitor and 100 U of M-MLV reverse transcriptase (Ferementas, Inc.). The reaction was conducted at 42 °C for 1 h followed by an inactivation step at 85 °C for 15 min. PCR amplification of 1 μl of cDNA was performed in a 25 μl reaction volume containing 1*x* standard Taq buffer (10 mM Tris–HCl, 50 mM KCl, pH 8.3), 3.5 mM $MgCl_2$, 200 μM dNTPs, 0.4 μM each primer and 1 U Taq DNA polymerase (Fermentas Inc.). PCR primers were designed using GeneTool software based on consensus sequence. Amplifications were performed using the primers (1) Sc-ALF-F (5′-ggacagaagaaacattgaggacgacgca-3′), Sc-ALF-R (5′-ggaaatcaaaaacatc-cattacaggtca-3′) and (2) Sc-Crus-F (5′-gagagcagaattagacactgt-3′), Sc-Crus-R (5′-atatagtataacataaccatacttc-3′). The thermal profile used

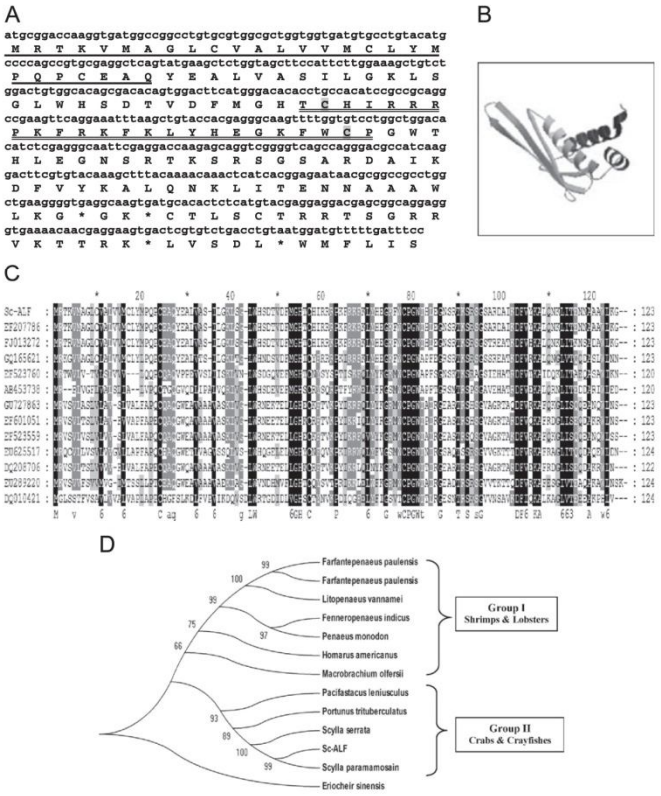


Fig. 1. (A) Nucleotide and amino acid sequences of Sc-ALF from the haemocyte of the mud crab, *Scylla serrata* (HQ638024). The underlined amino acid residues indicate a putative signal sequence. LPS binding domain characteristic of the ALF family is double underlined and the two conserved cysteine residues important for one disulfide bond (loop) formation are highlighted in gray. An asterisk is the stop codon. (B) Structural model of Sc-ALF (HQ638024) of *Scylla serrata* created using SWISS-MODEL server. (C) Multiple alignment of nucleotide sequence of the Sc-ALF (HQ638024) with other ALFs (*Scylla paramamosain* EF207786, *Scylla serrata* FJ013272, *Portunus trituberculatus* GQ165621, *Pacifastacus leniusculus* EF523760, *Marsupenaeus japonicas* AB453738, *Fenneropenaeus indicus* GU727863, *Farfantepenaeus paulensis* EF601051, *Penaeus monodon* EF523559, *Homarus americanus* EU625517, *Litopenaeus vannamei* DQ208706, *Macrobrachium olfersii* EU289220, *Litopenaeus stylirostris* DQ010421) obtained using GeneDoc programme Version 2.7.0. Black and gray indicate conserved sequences. (D) A bootstrapped neighbor-joining tree obtained using MEGA version 4.0 illustrating relationships between the deduced amino acid sequence of the Sc- ALF (HQ638024) with other ALFs of decapod crustaceans (*Scylla paramamosain* EF207786, *Scylla serrata* FJ013272, *Portunus trituberculatus* GQ165621, *Pacifastacus leniusculus* EF523760, *Fenneropenaeus indicus* GU727863, *Farfantepenaeus paulensis* EF601051, *Penaeus monodon* EF523559, *Homarus americanus* EU625517, *Litopenaeus vannamei* DQ208706, *Macrobrachium olfersii* EU289220, *Farfantepenaeus paulensis* EF601052. *Eriocheir sinensis* (GU014699) was used as an outgroup. Values at the node indicate the percentage of times that the particular node occurred in 1000 trees generated by bootstrapping the original deduced protein sequences.

Table 1
Result of BLAST analysis of Sc-ALF (**HQ638024**).

GenBank accession no.	Description	Query coverage (%)	*E*-value	Max identity (%)
EF207786	*Scylla paramamosain* anti-lipopolysaccharide factor mRNA, complete cds	100	2e−65	100
FJ013272	*Scylla serrata* anti-lipopolysaccharide factor precursor, mRNA, complete cds	100	4e−61	93
GQ165621	*Portunus trituberculatus* anti-lipopolysaccharide factor isoform 3 (ALF) mRNA, complete cds, alternatively spliced	98	6e−50	76
EF523760	*Pacifastacus leniusculus* anti-lipopolysaccharide factor mRNA, complete cds	85	1e−24	52
GU727863	*Fenneropenaeus indicus* anti-lipopolysaccharide factor (ALF) mRNA, complete cds	98	6e−23	41

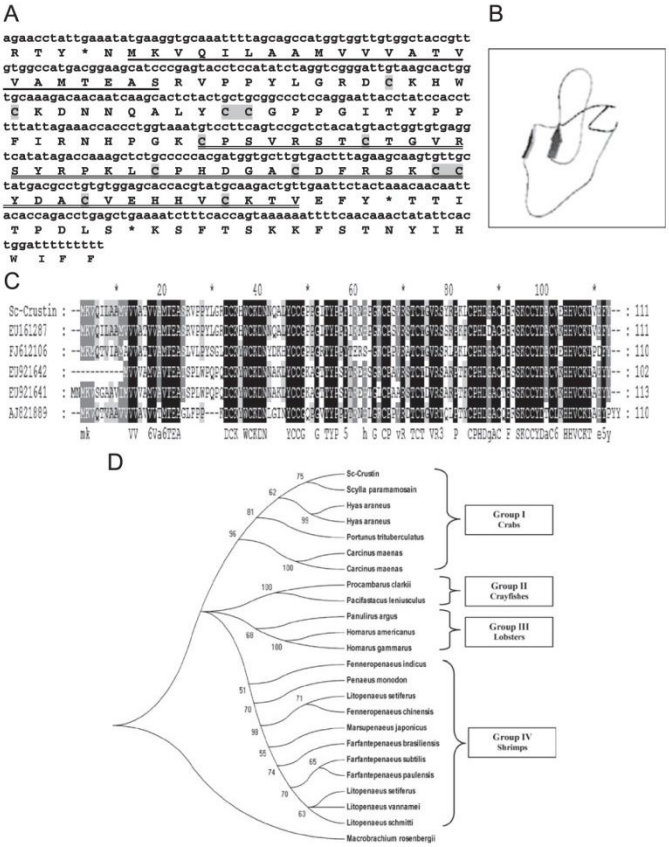


Fig. 2. (A) Nucleotide and amino acid sequences of Sc-crustin from the haemocyte of the mud crab, *Scylla serrata* (HQ638025). The underlined amino acid residues indicate a putative signal sequence. Cysteine residues that participate in the formation of intramolecular disulphide bonds are highlighted in gray and the WAP domain is double underlined. An asterisk is the stop codon. (B) Structural model of Sc-crustin (HQ638025) of *Scylla serrata* created using SWISS-MODEL server. (C) Multiple alignment of nucleotide sequence of the Sc-crustin (HQ638025) with other crustins (*Scylla paramamosain* EU161287, *Portunus trituberculatus* FJ612106, *Hyas araneus* EU921642, *Hyas araneus* EU921641, *Carcinus maenas* AJ821889) obtained using GeneDoc programme Version 2.7.0. Black and gray indicate conserved sequences. (D) A bootstrapped neighbor-joining tree obtained using MEGA version 4.0 illustrating relationships between the deduced amino acid sequence of the Sc-crustin (HQ638025) with other crustins of decapod crustaceans (*Scylla paramamosain* EU161287, *Portunus trituberculatus* FJ612106, *Hyas araneus* EU921642, *Hyas araneus* EU921641, *Carcinus maenas* AJ821889, *Carcinus maenas* AJ237947, *Procambarus clarkia* GQ301202, *Pacifastacus leniusculus* EF523613, *Marsupenaeus japonicus* AB121744, *Litopenaeus setiferus* AF430079, *Penaeus monodon* FJ380049, *Farfantepenaeus brasiliensis* EF601055, *Farfantepenaeus subtilis* EF450744, *Fenneropenaeus chinensis* FJ853147, *Litopenaeus vannamei* AY488497, *Farfantepenaeus paulensis* EF182747, *Litopenaeus setiferus* AF430078, *Fenneropenaeus indicus* GQ469987, *Litopenaeus schmitti* EF182748, *Homarus americanus* EF193003, *Homarus gammarus* AJ786653, *Panulirus argus* AY340636). *Macrobrachium rosenbergii* (EF364561) was used as an outgroup. Values at the node indicate the percentage of times that the particular node occurred in 1000 trees generated by bootstrapping the original deduced protein sequences.

Table 2
Result of BLAST analysis of Sc-crustin (**HQ638025**).

GenBank accession no.	Description	Query coverage (%)	*E*-value	Max identity (%)
EU161287	*Scylla paramamosain* crustin antimicrobial peptide mRNA, complete cds	100	3e−44	81
FJ612106	*Portunus trituberculatus* crustin antimicrobial peptide mRNA, complete cds	100	1e−30	62
EU921642	*Hyas araneus* crustin Ha2 mRNA, partial cds	73	3e−29	73
AJ821889	*Carcinus maenas* mRNA for carcinin-like protein, isolate IV	99	2e−27	56
FJ853148	*Fenneropenaeus chinensis* crustin 3 mRNA, complete cds	67	2e−08	39

was 94 °C for 2 min followed by 35 cycles of 94 °C for 15 s, 60 °C for 30 s and 68 °C for 30 s and a final extension at 68 °C for 10 min. PCR products were analyzed by electrophoresis in 1.5% agarose gels in TBE buffer, stained with ethidium bromide and visualized under UV light. Purified PCR products were sequenced at Scigenom, India.

The sequence homology and the deduced amino acid sequence comparisons were carried out using BLAST algorithm (tblastn) at the National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/blast). Gene translation and prediction of deduced protein were performed with ExPASy (http://www.au.expasy.org/). The signal peptide was predicted by SignalP program (http://www.au.expasy.org/). The multiple sequence alignments were performed with amino acid sequences of known crustins and ALFs from decapod crustaceans using CLUSTALW and GENDOC. Amino acid sequences of all known crustins and ALFs were retrieved from the NCBI GenBank and phylogenetic tree was constructed by the Neighbor-Joining (NJ) method and the Maximum Likelihood (ML) method using MEGA version 4.0 [22]. The structural models of the AMPs were created using SWISS-MODEL server. The nucleotide sequences and deduced amino acid sequences of the antimicrobial peptides were submitted to GenBank.

3. Results and discussion

In the present study two AMPs belonging to ALF and crustin families were characterized from the hemocytes of *S. serrata*, herein after referred to as Sc-ALF and Sc-crustin, respectively. The ORF of Sc-ALF consisted of 123 amino acids (Fig. 1A). BLAST analysis of the nucleotide sequences revealed the relation of Sc-ALF to other ALFs present in other decapod crustaceans. Sc-ALF was found to be 93% similar to an ALF isoform characterized from *S. serrata*. However, a 100% similar ALF molecule was found to be present in *S. paramamosain* [12]. Sc-ALF also shared similarity to ALFs of *Portunus trituberculatus* (76%), *Pacifastacus leniusculus* (52%) and *Fenneropenaeus indicus* (41%) (Table 1). Analysis with the SignalP software revealed the presence of a signal peptide with 26 amino acids at the N-terminal region of the Sc-ALF (Fig. 1A). The mature peptide consisted of 97 amino acid residues with a predicted molecular weight (MW) of 11.17 kDa. The Sc-ALF was highly cationic and the isoelectric point (*pI*) was estimated to be 9.95 as predicted by the PROTPARAM software. The sequence was deposited in the NCBI GenBank under accession number **HQ638024**. Sequence comparison of Sc-ALF amino acid revealed conserved amino acid residues in the region of LPS binding domain (Fig. 1A). The deduced amino acid sequence of Sc-ALF showed a 24 amino acid domain from residue 54 through 77, which was necessary for LPS binding and neutralization [12]. The Sc-ALF molecule also showed the conservation of two cysteine residues at positions Cys^{55} and Cys^{76}, important for one disulfide bond (loop) formation in the peptide (Fig. 1A). The deduced amino acid sequence of Sc-ALF was found to be rich in positively charged amino acid residues, arginine (10.3%) and lysine (7.2%) forming a cluster within the disulfide loop regarded as the functional domain of ALF as described by Imjongjirak et al. [12]. Peptide model of Sc-ALF created using SWISS-MODEL server consisted of two α-helices crowded against a four-strand β-sheet. Two of the β-strands are in turn linked by a disulfide bond to form an amphipathic loop rich in cationic amino acid side chains (Fig. 1B). Multiple alignment performed for Sc-ALF with other ALFs revealed the presence of conserved regions within the sequence (Fig. 1C). The phylogenetic relationship between Sc-ALF and other ALFs of decapod crustaceans was analyzed using the Neighbor-Joining (NJ) method (Fig. 1D). Molecular phylogenetic tree based on amino acid sequences suggests that all the ALF members possess the same ancestral origin, which has subsequently diverged at different phases of evolution. The tree could be broadly divided into two major groups, Group I included ALFs from shrimps and lobsters and Group II consisted of ALFs from crabs and crayfishes. The bootstrap distance tree calculated for the Sc-ALF clearly indicated that the Sc-ALF possessed great similarity to ALFs of other crabs (Fig. 1D).

The full-length cDNA of Sc-crustin was 433 bp in length, encoding 144 amino acids (Fig. 2A). The Sc-crustin cDNA encoded a polypeptide of 111 amino acids in the ORF with a putative signal peptide of 21 amino acid residues and a mature protein of 90 amino acids (Fig. 2A). The calculated molecular mass of the mature protein was 10.24 kDa. The isoelectric point (*pI*) was estimated to be 8.76 as predicted by the PROTPARAM software. The full-length sequence was deposited in the NCBI GenBank under accession number **HQ638025**. Sequence comparison using BLAST algorithm showed that the deduced amino acid sequence of Sc-crustin possessed an overall similarity of 81%, 62%, 73%, 56% and 39% to the crustins of *S. paramamosain*, *P. trituberculatus*, *H. araneus*, *C. maenas* and *F. chinensis*, respectively (Table 2). The deduced amino acid sequence of Sc-crustin was found to be rich in amino acid residues cysteine (13.3%) and proline (11.1%). A WAP domain could be detected in the C-terminus of Sc-crustin. As described by Imjongjirak et al. [13] a conserved eight-cysteine residue region responsible for the formation of 4 disulfide core (4-DSC) could also be detected in the C5–C12 position (Fig. 2A). The 12 cysteines in Sc-crustin (Fig. 2A) are considered to be important for maintaining the tertiary structure of the peptide just as that reported in the case of shrimp crustins [23,24]. Peptide model of Sc-crustin created using SWISS-MODEL server indicated a random coiled structure, that is, with two possible β-sheets but no helices (Fig. 2B). Multiple alignment performed for Sc-crustin with other crustins of decapods revealed the presence of conserved regions within the sequence (Fig. 2C).

In the present study, the BLAST homology search for the AMPs (Sc-ALF and Sc-crustin) from the brown mud crab, *S. serrata* showed maximum similarity to those from the green mud crab, *S. paramamosain*, a commercially important species in Queensland, Australia. The brown mud crab is a well known commercial species in India, Philippines and Vietnam. The brown mud crab differs from the green mud crab in having one less spine on the wrist and behind the fingers of the male. It also lacks the overall pale green mottling on the legs/rear paddles and is comparatively small in size. The study emphasizes the similarity in the defense peptides of these taxonomically related species.

This report presents the characterization and phylogenetic analysis of a new ALF isoform (Sc-ALF) and the first crustin

sequence (Sc-crustin) from *S. serrata*. Sc-ALF gave 93% similarity to an ALF isoform from *S. serrata*. Sc-crustin is the first report of a crustin sequence from *S. serrata*. Discovery of novel AMPs and its antimicrobial spectrum might pave way to unravel the obscurity of crustacean immunity. Further research on the expression profile of these molecules in response to various environmental conditions and microbial infection would reveal their role in the protection of the animals from the onslaught of diseases.

Acknowledgment

The authors are grateful to the Ministry of Earth Sciences (MoES), Government of India for the research grant (MoES/10-MLR/2/2007) with which the work was carried out.

References

[1] Lorenzon S, de Guarrini S, Smith VJ, Ferrero EA. Effects of LPS injection on circulating haemocytes in crustaceans *in vivo*. Fish Shellfish Immunol 1999;9:31–50.
[2] Smith V, Brown J, Hauton C. Immunostimulation in crustaceans: does it really protect against infection? Fish Shellfish Immunol 2003;15:71–90.
[3] Sheynis T, Sykora J, Benda A, Kolusheva S, Hof M, Jelinek R. Bilayer localization of membrane-active peptides studied in biomimetic vesicles by visible and fluorescence spectroscopies. Eur J Biochem 2003;270:4478–87.
[4] Bulet P, Hetru C, Dimarcq JL, Hoffmann D. Antimicrobial peptides in insects; structure and function. Dev Comp Immunol 1999;23:329–44.
[5] Zasloff M. Antimicrobial peptides of multicellular organisms. Nature 2002;415:389–95.
[6] Boman HG. Antibacterial peptides: basic facts and emerging concepts. J Int Med 2003;254:197–215.
[7] Tincu JA, Taylor SW. Antimicrobial peptides from marine invertebrates. Antimicrob Agents Chemother 2004;48:3645–54.
[8] Schnapp D, Kemp GD, Smith VJ. Purification and characterization of a proline-rich antibacterial peptide, with sequence similarity to bactenecin-7, from the haemocytes of the shore crab, *Carcinus maenas*. Eur J Biochem 1996;240:532–9.
[9] Relf JM, Chisholm JR, Kemp GD, Smith VJ. Purification and characterization of a cysteine-rich 11.5-kDa antibacterial protein from the granular haemocytes of the shore crab, *Carcinus maenas*. Eur J Biochem 1999;2642:350–7.
[10] Khoo L, Robinette DW, Callinectin Noga EJ. an antibacterial peptide from blue crab, *Callinectes sapidus*, haemocytes. Mar Biotechnol 1999;1:44–51.
[11] Wang KJ, Huang WS, Yang M, Chen HY, Bo J, Li SJ, et al. A male-specific expression gene, encodes a novel anionic antimicrobial peptide, scygonadin, in *S. serrata*. Mol Immunol 2007;44:1961–8.
[12] Imjongjirak C, Amparyup P, Tassanakajon A, Sittipraneed S. Antilipopolysaccharide factor ALF of mud crab *Scylla paramamosain*: molecular cloning, genomic organization and the antimicrobial activity of its synthetic LPS binding domain. Mol Immunol 2007;44:3195–203.
[13] Imjongjirak C, Amparyup P, Tassanakajon A, Sittipraneed S. Molecular cloning and characterization of crustin from mud crab *Scylla paramamosain*. Mol Biol Rep 2009;36:841–50.
[14] Stensvag K, Haug T, Sperstad SV, Rekdal O, Indrevoll B, Styrvold OB. Arasin 1, a proline-arginine-rich antimicrobial peptide isolated from the spider crab, *Hyas araneus*. Dev Comp Immunol 2008;32:275–85.
[15] Shigenaga T, Muta T, Toh Y, Tokunaga F, Iwanaga S. Antimicrob. tachyplesin peptide precursor. cDNA cloning and cellular localization in the horseshoe crab *Tachypleus tridentatus*. J Biol Chem 1990;265:21350–4.
[16] Destoumieux D, Munoz M, Bulet P, Penaeidins Bachere E. a family of antimicrobial peptides from penaeid shrimp Crustacea, Decapoda. Cell Mol Life Sci 2000;57:1260–71.
[17] Mitta G, Hubert F, Dyrynda EA, Boudry P, Roch P, Mytilin B. and MGD2, two antimicrobial peptides of marine mussels: Gene structure and expression analysis. Dev Comp Immunol 2000;24:381–93.
[18] Hauton C, Brokton V, Smith VJ. Cloning of a crustin-like single whey-acidic-domain, antibacterial peptide from the haemocytes of the European lobster, *Homarus gammarus*, and its response to infection with bacteria. Mol Immunol 2006;43:1490–6.
[19] Stewart JE, Zwicker BM. Induction and natural bactericidal activities of the lobster *Homarus americanus*: products of haemocyte–plasma interactions. Can J Microbiol 1972;18:1499–511.
[20] Chisholm JRS, Smith VJ. Antibacterial activity in the haemocytes of the shore crab, *Carcinus maenas* (L.). J Mar Biol Assoc UK 1992;72:529–42.
[21] Noga EJ, Arroll TA, Fan Z. Specificity and some physicochemical characteristics of the antibacterial activity from blue crab *Callinectes sapidus*. Fish Shellfish Immunol 1996;6:403–12.
[22] Tamura K, Dudley J, Nei M, Kumar S. MEGA 4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 2007;24:1596–9.
[23] Gross PS, Barlett TC, Browdy CL, Chapman RW, Warr GW. Immune gene discovery by expressed sequence tag analysis of haemocytes and hepatopancreas in the Pacific white shrimp, *Litopenaeus vannamei*, and Atlantic white shrimp, *Litopenaeus setiferus*. Dev Comp Immunol 2001;25:565–77.
[24] Zhang J, Li F, Wang Z, Xiang J. Cloning and recombinant expression of a crustin-like gene from Chinese shrimp, *Fenneropenaeus chinensis*. J Biotechnol 2007;127:605–14.

Aquaculture 356-357 (2012) 119–122

Contents lists available at SciVerse ScienceDirect

Aquaculture

journal homepage: www.elsevier.com/locate/aqua-online

A new isoform of anti-lipopolysaccharide factor identified from the blue swimmer crab, *Portunus pelagicus*: Molecular characteristics and phylogeny

V.V. Afsal [a], Swapna P. Antony [a], V.N. Sanjeevan [b], P.R. Anil Kumar [a], I.S. Bright Singh [c], Rosamma Philip [a,*]

[a] Department of Marine Biology, Microbiology and Biochemistry, School of Marine Sciences, CUSAT, Fine Arts Avenue, Kochi-16, Kerala, India
[b] Centre for Marine Living Resources and Ecology, 6th Floor, C-Block, Kendriya Bhavan, P.B. No. 5415, CSEZ P.O., Kochi-682 037, India
[c] National Center for Aquatic Animal Health (NCAAH), CUSAT, Fine Arts Avenue, Kochi-16, Kerala, India

ARTICLE INFO

Article history:
Received 11 May 2012
Accepted 21 May 2012
Available online 26 May 2012

Keywords:
Antimicrobial peptide
Portunus pelagicus
Anti-lipopolysaccharide factor
Crustacean immunity

ABSTRACT

Anti-lipopolysaccharide factors are small proteins that bind and neutralize lipopolysaccharide and exhibit potent antimicrobial activities. This study presents the molecular characterization and phylogenetic analysis of the first ALF isoform (Pp-ALF1; JQ745295) identified from the hemocytes of *Portunus pelagicus*. The full length cDNA of Pp-ALF1 consisted of 880 base pairs encoding 293 amino acids with an ORF of 123 amino acids and contains a putative signal peptide of 24 amino acids. Pp-ALF1 possessed a predicted molecular weight (MW) of 13.86 kDa and theoretical isoelectric point (*pI*) of 8.49. Two highly conserved cysteine residues and putative LPS binding domain were observed in Pp-ALF1. Peptide model of Pp-ALF1 consisted of two α-helices crowded against a four-strand β-sheet. Comparison of amino acid sequences and neighbor joining tree showed that Pp-ALF1 has a maximum similarity (46%) to ALF present in *Portunus trituberculatus* followed by 39% similarity to ALF of *Eriocheir sinensis* and 38% similarity to ALFs of *Scylla paramamosain* and *Scylla serrata*. Pp-ALF1 is found to be a new isoform of ALF family and its characteristic similarity with other known ALFs signifies its role in protection against invading pathogens.

1. Introduction

Fifteen divergent families of AMPs or single peptides sharing common molecular features with the known AMP families displaying antimicrobial activities against a number of specific microorganisms have been identified and characterized in a number of decapod crustaceans (Rosa and Barracco, 2010). Anti-lipopolysaccharide factors (anti-LPS factors or ALFs) are the potent antimicrobial peptides that can bind and neutralize lipopolysaccharides (LPS) forming a key effector molecule of the innate immune system in crustaceans. ALFs contain a signal peptide and a conserved LPS-binding domain consisting of a conserved cluster of positively charged residues within the disulfide loop (Hoess et al., 1993; Muta et al., 1987; Somboonwiwat et al., 2008; Yang et al., 2009). These peptides are part of a very well-characterized family of crustacean AMPs with different subgroups and variants that are either encoded by distinct genes or generated by alternative mRNA splicing (Tharntada et al., 2008). The high diversity of the variants (subgroups and isoforms) found in most crustacean AMP classes can presumably confer a broad spectrum of activity to a single AMP family. Limulus ALF has proved to possess strong antibacterial activity, especially on the growth of G-negative R type bacteria (Wang et al., 2002). Crustacean ALFs are also multifunctional proteins exhibiting a potent (MIC<6.25 μM) and broad spectrum antimicrobial activity against a large number of both Gram-positive and Gram-negative bacteria, including several opportunistic/pathogenic *Vibrio* species, fungi and human enveloped virus (Carriel-Gomes et al., 2007; Somboonwiwat et al., 2005; Yedery and Reddy, 2009). ALF has been reported to have a specific role in reducing white spot syndrome viral replication in crayfishes and shrimps (Antony et al., 2011; Liu et al., 2006). Although more than 200 penaeidin and crustin relative sequences have been registered in the GenBank to date, information on ALFs of crustaceans is comparatively less. ALFs were first purified from the hemolymph cells (amoebocytes) of two marine chelicerate arthropods, the horseshoe crabs, *Limulus polyphemus* and *Tachypleus tridentatus* (Tanaka et al., 1982). In crustaceans, homologues to horseshoe crab ALFs were first identified from the hemocytes of the penaeid shrimp species *Litopenaeus setiferus* (Gross et al., 2001) and *Penaeus monodon* (Supungul et al., 2002) using an EST approach. To date, the genes encoding ALFs have been identified only in decapod crustaceans, penaeid shrimps (Ponprateep et al., 2012; Tassanakajon et al., 2010), freshwater prawns (Lu et al., 2009; Ren et al., 2012; Rosa et al., 2008), crayfish (Liu et al., 2012), lobster (Beale et al., 2008) and crabs (Afsal et al., 2011; Imjongjirak et al., 2007; Li et al., 2008; Liu et al., 2012; Yedery and Reddy, 2009; Yue et al., 2010). However, there is hardly any published work on major AMP families present in the blue swimmer crab, *Portunus pelagicus*. This species is a large Portunid crab found in shallow waters between 10 and 50 m depth, including areas near reefs, mangroves, sea grass and algal beds. *P. pelagicus* is

* Corresponding author. Tel.: +91 4842368120; fax: +91 4842381120.
E-mail addresses: rosammap@gmail.com, rose@cusat.ac.in (R. Philip).

0044-8486/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.aquaculture.2012.05.028

similar to *Portunus trituberculatus* in its general appearance but, easily distinguished because it has four frontal teeth (three in case of *P. trituberculatus*) and possesses three spines in the merus of chelipeds (four in case of *P. trituberculatus*). *Portunus sanguinolentus* is easily distinguished by their three prominent maroon to red spots on posterior 1/3 of carapace. *P. pelagicus* is a highly prized commercially important crab and their characteristic features such as fast growth, ease of larviculture, high fecundity and relatively high tolerance to both nitrate and ammonia make this species ideal for aquaculture.

Studies on characterization and structure elucidation of ALFs in *P. pelagicus* may be useful in understanding the immune defense mechanisms in crabs giving new insights into disease control in crab aquaculture. This paper presents the molecular characterization and phylogenetic analysis of a new isoform of ALF identified from *P. pelagicus* and this could be the first report of an antimicrobial peptide from the species.

2. Materials and methods

Healthy adults of *P. pelagicus* (~250 g body weight) were collected from the backwater stream along Vypeen, Kochi, India. Hemolymph was collected from the base of abdominal appendages using specially designed capillary tubes (RNase-free) rinsed using pre-cooled anticoagulant solution (RNase free 10% sodium citrate, pH 7.0). Total RNA was extracted from the hemocytes using TRI Reagent (Sigma) following the manufacturer's protocol. RNA was quantified by spectrophotometry at 260 and 280 nm. First strand cDNA was generated in a 20 µl reaction volume containing 5 µg total RNA, 1× RT buffer, 2 mM dNTP, 2 µM oligo d(T_{20}), 20 U of RNase inhibitor and 100 U of M-MLV reverse transcriptase (Fermentas, Inc.). The reaction was conducted at 42 °C for 1 h followed by an inactivation step at 85 °C for 15 min. PCR amplification of 1 µl of cDNA was performed in a 25 µl reaction volume containing 1× standard Taq buffer (10 mM Tris–HCl, 50 mM KCl, pH 8.3), 3.5 mM $MgCl_2$, 200 µM dNTPs, 0.4 µM each primer and 1 U Taq DNA polymerase (Fermentas Inc.). PCR primers were designed using GeneTool software based on consensus sequences of already reported ALFs of decapods. Amplifications were performed using the primers Pp-ALF1-F (5′-agggagtgggtgatgagcta-3′) and Pp-ALF1-R (5′-tacggctattacgatccaaca-3′). The thermal profile used was 94 °C for 2 min followed by 35 cycles of 94 °C for 15 s, 60 °C for 30 s and 68 °C for 30 s and a final extension at 68 °C for 10 min. PCR products were analyzed by electrophoresis in 1.5% agarose gels in TBE buffer, stained with SYBR® Safe and visualized under UV light. Purified PCR products were sequenced at SciGenom, India.

```
gttgttatcatctgttgcccttgctgcttccctccaccttcttccctcatgatgaaaaaa
                                                  M  K  K
ggagtggccgccctgctgcggctgctgctggtgatgaggctgctgctgccgccccttct
 G  V  A  A  L  L  R  L  L  L  V  M  R  L  L  L  P  P  P  S
tgcgatcctcatgacgatcctctggcttcttctcttcttgaaatcctcgctgggcggtgc
 C  D  P  H  D  D  P  L  A  S  S  L  L  E  I  L  A  G  R  C
cacaactcctctgatgacttggtgggcccctgctgcttcctccgccgccccctatcata
 H  N  S  S  D  D  L  V  G  P  C  C  F  L  R  R  P  P  I  I
agaatatttctgctgcaccacgagggctttttttggtgtgctgggtggccgcctttcgat
 R  I  F  L  L  H  H  E  G  F  F  W  C  A  G  W  P  P  F  D
gggtcgacgaggacaaacaggacggcggcgtcctcggaggacgccaccgacgacttgcgg
 G  S  T  R  T  N  R  T  A  A  S  S  E  D  A  T  D  D  L  R
cgcgctttttacacaaactactcatcccacaggatgatgctccgcggcggatgaacaac
 R  A  F  L  H  K  L  L  I  P  Q  D  D  A  P  R  R  M  N  N
taaagaagaagagatgagctgcggaggacgacgaggagaataatgataacaacatagaaa
 *
ttgccggacgacttgttattatatatttttttcctcagactagttggtggatgctgcgt
tgttgatgcaatgatggaagtagagaagcgtgaaggaacgactatgatgtgaagataccac
ctcaaaaaatttttttttattatttttttatgaggtagcctcgtcatttttttcttctct
tttcctccaccatcgccgttgatgctagactggtggacgacaatgaagagataaagaagt
agatgatattattatttctaccactacatctttttatatgaggaacaaggagggccgcc
tcctcttttttgttgttgtggtagtgggacgactatctttttttttttgtgtgtgggagga
agaaagaaaagaaaaatatctttatttgtacttgagctgt
```

Fig. 1. Nucleotide and amino acid sequences of Pp-ALF1 from the hemocyte of the blue swimmer crab, *Portunus pelagicus* (JQ745295). The underlined amino acid residues indicate a putative signal sequence. LPS binding domain characteristic of the ALF family is double underlined and the two conserved cysteine residues important for one disulfide bond (loop) formation is highlighted in gray. An asterisk is the stop codon.

Table 1
Result of BLASTp analysis of Pp-ALF1 (JQ745295).

Accession no.	Description of the AMP	Query coverage	E value	Max identity
ADU25043	Anti-lipopolysaccharide factor isoform 2 [*Portunus trituberculatus*]	88%	3e −23	46%
ADU25072	Anti-lipopolysaccharide factor [*Portunus trituberculatus*]	77%	1e −20	46%
ADZ46233	Anti-lipopolysaccharide factor 3 [*Eriocheir sinensis*]	85%	7e −16	39%
ABP96981	Anti-lipopolysaccharide factor [*Scylla paramamosain*]	94%	5e −18	38%
ADW11095	Anti-lipopolysaccharide factor [*Scylla serrata*]	92%	8e −18	38%

The sequence homology and the deduced amino acid sequence comparisons were carried out using BLAST algorithm (BLASTn and BLASTp) at the National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/blast). Gene translation and prediction of deduced protein were performed with ExPASy (http://www.au.expasy.org/). The signal peptide was predicted by SignalP program (http://www.au.expasy.org/). The multiple sequence alignments were performed with amino acid sequences of known ALFs from decapod crustaceans using CLUSTALW and GENDOC. Amino acid sequences of all known ALFs were retrieved from the NCBI GenBank and phylogenetic tree was constructed by the Neighbor-Joining (NJ) method using MEGA version 4.0 (Tamura et al., 2007). The structural models of the AMPs were created using SWISS-MODEL server. The nucleotide and deduced amino acid sequences of the antimicrobial peptide were submitted to NCBI GenBank.

3. Results and discussion

In the present study, an AMP belonging to ALF family was identified and characterized from the hemocytes of *P. pelagicus*, herein after referred to as Pp-ALF1. An 880 bp fragment cDNA encoding an ORF of 123 amino acids was obtained from the mRNA of *P. pelagicus* hemocyte by RT-PCR (Fig. 1). BLAST analysis of the nucleotide sequences revealed the relation of Pp-ALF1 to other ALFs present in decapod crustaceans. Pp-ALF1 showed only 46% similarity to ALF isoforms present in *P. trituberculatus* followed by 39% similarity to ALF of *Eriocheir sinensis* and 38% similarity to ALFs of *Scylla paramamosain* and *Scylla serrata* (Table 1). Analysis with the SignalP software revealed the presence of a signal peptide with 24 amino acid residues at the N-terminal region of the Pp-ALF1 (Fig. 1). The mature peptide consisted of 123 amino acid residues with a predicted

Fig. 2. Structural model of Pp-ALF1 (JQ745295) of *Portunus pelagicus* created using SWISS-MODEL server.

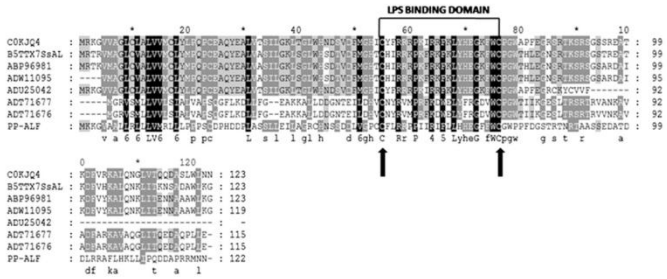


Fig. 3. Multiple alignment of nucleotide sequence of the Pp-ALF1 (JQ745295) with other ALFs obtained using GeneDoc program Version 2.7.0. The LPS-binding domains are enclosed with a bracket. The conserved cysteine residues are marked with arrowheads below the alignment. Black and gray indicate conserved sequences.

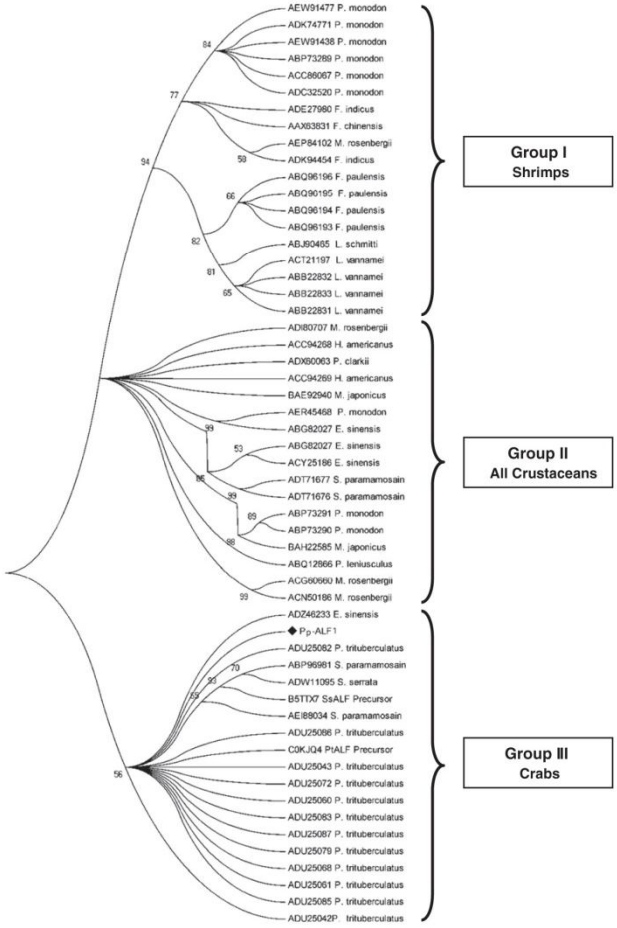


Fig. 4. A bootstrapped neighbor-joining tree obtained using MEGA version 4.0 illustrating relationships between the deduced amino acid sequences of the Pp-ALF1 (JQ745295) with other ALFs of decapod crustaceans. Values at the node indicate the percentage of times that the particular node occurred in 1000 trees generated by bootstrapping the original deduced protein sequences.

molecular weight (MW) of 13.86 kDa. The Pp-ALF1 was highly cationic and the isoelectric point (*pI*) was estimated to be 8.49 as predicted by the PROTPARAM software. The sequence was deposited in the NCBI GenBank under accession number JQ745295.

Sequence comparison of Pp-ALF1 peptide revealed that it possessed a signal peptide and conserved amino acid residues in the region of LPS binding domain (Fig. 1). The deduced amino acid sequence of Pp-ALF1 showed a 24 amino acid domain from residue 54 through 77 that was necessary for LPS binding and neutralization (Imjongjirak et al., 2007). The Pp-ALF1 molecule also showed the conservation of two cysteine residues at positions Cys^{55} and Cys^{76}, important for one disulfide bond (loop) formation in the peptide (Fig. 1). The deduced amino acid sequence of Pp-ALF1 was found to be rich in positively charged amino acid residues, arginine (10.6%) forming a cluster within the disulfide loop regarded as the functional domain of ALF as described by Imjongjirak et al. (2007). Peptide model of Pp-ALF1 created using SWISS-MODEL server consisted of two α-helices crowded against a four-strand β-sheet. Two of the β-strands are in turn linked by a disulfide bond to form an amphipathic loop rich in cationic amino acid side chains (Fig. 2). Multiple alignment performed for Pp-ALF1 with other ALFs revealed the presence of conserved regions within the sequence (Fig. 3).

The phylogenetic relationship between Pp-ALF1 and other ALFs of decapod crustaceans was analyzed using neighbor-joining (NJ) method (Fig. 4). The phylogenetic tree clearly showed that the ALF sequences clustered according to species. The crab and shrimp ALFs were found to cluster separately. The tree could be broadly divided into three major groups. Group I consisted of ALFs from shrimps; whereas Group II consisted of ALFs from all crustaceans including shrimps, crabs, crayfishes and lobsters. Group III was found to consist of ALFs found in crabs alone and Pp-ALF1 belonged to this group (Group III). The analysis showed that Pp-ALF1 is more closely related to crab ALFs rather than to the shrimp, crayfish and lobster ALFs. Also, within the crab cluster, ALFs of *Scylla* sp. grouped together and were found to be different from that of the *Portunus* and *Eriocher* sp. Also, Pp-ALF1 was found to be different from the other ALF isoforms of *P. trituberculatus*. The bootstrap distance tree calculated clearly indicated Pp-ALF1 to be a new isoform of the ALF family. The detection of ALF in the blue swimmer crab indicated Pp-ALF1 to be an immune effector molecule that may play an essential role in defense mechanism of these crabs. Also, molecular phylogenetic tree based on amino acid sequence suggests that all the ALF members are of the same ancestral origin, which has subsequently diverged at different phases of evolution (Fig. 4).

To conclude, this report presents the characterization and phylogenetic analysis of the first ALF isoform (Pp-ALF1) from the commercially important *P. pelagicus*. Discovery of novel AMPs and its antimicrobial spectrum might pave way to unravel the obscurity of crustacean immunity. Further research on the expression profile of these molecules in response to various environmental conditions and microbial infection would reveal their role in the protection of the animals from the onslaught of diseases.

Acknowledgment

The authors are grateful to the Ministry of Earth Sciences (MoES), Govt. of India for the research grant (MoES/10-MLR/2/2007) with which the work was carried out.

References

Afsal, V.V., Antony, S.P., Sathyan, N., Singh, I.S.B., Philip, R., 2011. Molecular characterization and phylogenetic analysis of two antimicrobial peptides, anti-lipopolysaccharide factor and crustin from the brown mud crab, *Scylla serrata*. Results Immunology 1, 6–10.

Antony, S.P., Philip, R., Joseph, V., Singh, I.S.B., 2011. Anti-lipopolysaccharide factor and crustin-III, the anti-white spot virus peptides in *Penaeus monodon*, control of viral infection by up-regulation. Aquaculture 319, 11–17.

Beale, K.M., Towle, D.W., Jayasundara, N., Smith, C.M., Shields, J.D., Small, H.J., et al., 2008. Anti-lipopolysaccharide factors in the American lobster *Homarus americanus*, molecular characterization and transcriptional response to *Vibrio fluvialis* challenge. Comparative Biochemistry and Physiology 3D, 263–269.

Carriel-Gomes, M.C., Kratz, J.M., Barracco, M.A., Bachere, E., Barardi, C.R., Simoes, C.M., 2007. In vitro antiviral activity of antimicrobial peptides against herpes simplex virus 1, adenovirus, and rotavirus. Memórias do Instituto Oswaldo Cruz 102, 469–472.

Gross, P.S., Barlett, T.C., Browdy, C.L., Chapman, R.W., Warr, G.W., 2001. Immune gene discovery by expressed sequence tag analysis of haemocytes and hepatopancreas in the Pacific white shrimp, *Litopenaeus vannamei*, and Atlantic white shrimp, *Litopenaeus setiferus*. Developmental and Comparative Immunology 25, 565–577.

Hoess, A., Watson, S., Siber, G.R., Liddington, R., 1993. Crystal structure of an endotoxin-neutralizing protein from the horseshoe crab, *Limulus* anti-LPS factor, at 1.5 A resolution. EMBO Journal 12, 3351–3356.

Imjongjirak, C., Amparyup, P., Tassanakajon, A., Sittipraneed, S., 2007. Anti-lipopolysaccharide factor ALF of mud crab *Scylla paramamosain*, Molecular cloning, genomic organization and the antimicrobial activity of its synthetic LPS binding domain. Molecular Immunology 44, 3195–3203.

Li, C., Zhao, J., Song, L., Mu, C., Zhang, H., Gai, Y., et al., 2008. Molecular cloning, genomic organization and functional analysis of an anti-lipopolysaccharide factor from Chinese mitten crab *Eriocheir sinensis*. Developmental and Comparative Immunology 32, 784–794.

Liu, H., Jiravanichpaisal, P., Söderhäll, I., Cerenius, L., Söderhäll, K., 2006. Anti-lipopolysaccharide factor interferes with white spot syndrome virus replication in vitro and in vivo in the crayfish *Pacifastacus leniusculus*. Journal of Virology 80, 10365–10371.

Liu, Y., Cui, Z., Li, X., Song, C., Li, Q., Wang, S., 2012. A new anti-lipopolysaccharide factor isoform (PtALF4) from the swimming crab *Portunus trituberculatus* exhibited structural and functional diversity of ALFs. Fish & Shellfish Immunology 32, 724–731.

Lu, K.Y., Sung, H.J., Liu, C.L., Sung, H.H., 2009. Differentially enhanced gene expression in hemocytes from *Macrobrachium rosenbergii* challenged in vivo with lipopolysaccharide. Journal of Invertebrate Pathology 100, 9–15.

Muta, T., Miyata, T., Tokunaga, F., Nakamura, T., Iwanaga, S., 1987. Primary structure of anti-lipopolysaccharide factor from American horseshoe crab, *Limulus polyphemus*. Journal of Biochemistry 101, 1321 1330.

Ponprateep, S., Tharntada, S., Somboonwiwat, K., Tassanakajon, A., 2012. Gene silencing reveals a crucial role for anti-lipopolysaccharide factors from *Penaeus monodon* in the protection against microbial infections. Fish & Shellfish Immunology 32, 26–34.

Ren, Q., Du, Z.Q., Li, M., Zhang, C.Y., Chen, K.P., 2012. Cloning and expression analysis of an anti-lipopolysaccharide factor from giant freshwater prawn, *Macrobrachium rosenbergii*. Molecular Biology Reports 7, 7673–7680.

Rosa, R.D., Barracco, M.A., 2010. Antimicrobial peptides in crustaceans. Invertebrate Survival Journal 7, 262–284.

Rosa, R.D., Stoco, P.H., Barracco, M.A., 2008. Cloning and characterization of cDNA sequences encoding for anti-lipopolysaccharide factors (ALFs) in Brazilian palaemonid and penaeid shrimps. Fish & Shellfish Immunology 25, 693–696.

Somboonwiwat, K., Marcos, M., Tassanakajon, A., Klinbunga, S., Aumelas, A., Romestand, B., et al., 2005. Recombinant expression and anti-microbial activity of antilipopolysaccharide factor ALF from the black tiger shrimp *Penaeus monodon*. Developmental and Comparative Immunology 2910, 841–851.

Somboonwiwat, K., Bachère, E., Rimphanitchayakit, V., Tassanakajon, A., 2008. Localization of anti-lipopolysaccharide factor (ALFPm3) in tissues of the black tiger shrimp, *Penaeus monodon*, and characterization of its binding properties. Developmental and Comparative Immunology 32, 1170–1176.

Supungul, P., Klinbunga, S., Pichyangkura, R., Jitrapakdee, S., Hirono, I., Aoki, T., Tassanakajon, A., 2002. Identification of immune-related genes in hemocytes of black tiger shrimp *Penaeus monodon*. Marine Biotechnology 4, 487–494.

Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA4, molecular evolutionary genetics analysis (MEGA) software version 4.0. Molecular Biology and Evolution 24, 1596–1599.

Tanaka, S., Nakamura, T., Morita, T., Iwanaga, S., 1982. Limulus anti-LPS factor, an anticoagulant which inhibits the endotoxin mediated activation of limulus coagulation system. Biochemical and Biophysical Research Communications 105, 717–723.

Tassanakajon, A., Amparyup, P., Somboonwiwat, K., Supungul, P., 2010. Cationic antimicrobial peptides in penaeid shrimp. Marine Biotechnology 12, 487–505.

Tharntada, S., Somboonwiwat, K., Rimphanitchayakit, V., Tassanakajon, A., 2008. Anti-lipopolysaccharide factors from the black tiger shrimp, *Penaeus monodon*, are encoded by two genomic loci. Fish & Shellfish Immunology 24, 46–54.

Wang, D.N., Liu, J.W., Yang, G.Z., Zhang, W.J., Wu, X.F., 2002. Cloning of anti-LPS factor cDNA from *Tachypleus tridentatus*, expression in *Bombyx mori* larvae and its biological activity in vitro. Molecular Biotechnology 21, 1–7.

Yang, Y., Boze, H., Chemardin, P., Padilla, A., Moulin, G., Tassanakajon, A., et al., 2009. NMR structure of rALF-Pm3, an anti-lipopolysaccharide factor from shrimp, model of the possible lipid A-binding site. Biopolymers 91, 207–220.

Yedery, R.D., Reddy, K.V., 2009. Identification, cloning, characterization and recombinant expression of an anti-lipopolysaccharide factor from the hemocytes of Indian mud crab, *Scylla serrata*. Fish & Shellfish Immunology 27, 275–284.

Yue, F., Pan, L., Miao, J., Zhang, L., Li, J., 2010. Molecular cloning, characterization and mRNA expression of two antibacterial peptides, crustin and anti-lipopolysaccharide factor in swimming crab *Portunus trituberculatus*. Comparative Biochemistry and Physiology 156B, 77–85.