

Molecular Characterization of a Newly Identified Subfamily Member of Penaeidin from two Penaeid Shrimps, *Fenneropenaeus indicus* and *Metapenaeus monoceros*

V. V. Afsal¹ · Swapna P. Antony^{1,2} · Rosamma Philip¹ · I. S. Bright Singh²

Published online: 25 November 2015

© Springer Science+Business Media New York 2015

Abstract Penaeidins are a major group of antimicrobial peptides found in penaeid shrimps. This study reports a new isoform of penaeidin from the hemocytes of Indian white shrimp, *Fenneropenaeus indicus* (Fi-PEN, **JX657680**), and the pink shrimp, *Metapenaeus monoceros* (Mm-PEN, **KF275674**). Mm-PEN is also the first antimicrobial peptide to be identified from *M. monoceros*. The complete coding sequences of the newly identified Fi-PEN and Mm-PEN consisted of an ORF of 338 bp encoding 71 amino acids with a predicted molecular weight of 5.66 kDa and a pI of 9.38. The penaeidins had its characteristic signal peptide region (19 amino acids), which was followed by a mature peptide with a proline-rich domain (24 amino acids) at the N-terminal region and a cysteine-rich domain (28 amino acids) at the C-terminal region, designating it to penaeidin-3 subgroup. Structural analysis revealed an alpha-helix in its secondary structure and an extended structure at the proline-rich domain. The newly identified penaeidin isoform showed maximum similarity of 63 % to a penaeidin-3 isoform of *P. monodon*, which further proves it to be a new isoform. Phylogenetic analysis showed that it possessed similar evolutionary status like other penaeidins, which has subsequently diverged at different phases of evolution. The wide distribution of penaeidins in penaeid

shrimps indicates the importance of these AMPs in the innate immunity.

Keywords Antimicrobial peptide · Penaeidin · Innate immunity · *Fenneropenaeus indicus* · *Metapenaeus monoceros*

Introduction

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.

Penaeidins were first isolated from the hemolymph of the Pacific white shrimp, *Litopenaeus vannamei*, and showed antimicrobial activity mainly against Gram-positive bacteria and fungi [1]. More studies based on genomic approaches have revealed the presence of penaeidins in different shrimp species. To date, penaeidins have been characterized from *L. vannamei* [2], *L. setiferus* [3], *L. stylirostris* [4], *L. schmitti* [5], *Farfantepenaeus paulensis* [5], *Penaeus monodon* [6], *Marsupenaeus japonicas* [7], *Fenneropenaeus chinensis* [8] and *F. indicus* [9, 10].

Penaeidins are composed of an N-terminal domain rich in proline residues and a C-terminal domain containing six cysteines that form three disulfide 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

✉ Rosamma Philip
rosammap@gmail.com; rose@cusat.ac.in

¹ Department of Marine Biology, Microbiology and Biochemistry, School of Marine Sciences, Cochin University of Science and Technology, Fine Arts Avenue, Kochi, Kerala 682 016, India

² National Centre for Aquatic Animal Health, Cochin University of Science and Technology, Fine Arts Avenue, Kochi, Kerala 682 016, India

overall structure, penaeidins undergo posttranslational modifications, such as C-terminal amidation [1]. Penaeidins are highly variable with respect to genetic composition which is clearly exhibited by the existence of different isoforms in each subgroup. There are four subgroups of penaeidins in penaeid shrimps based on their primary amino acid sequence diversity, namely penaeidin-2, penaeidin-3, penaeidin-4 and penaeidin-5, and each subgroup has several isoforms [1, 2]. Moreover, each class has been defined by the conserved key residues of eight specific amino acids located in precise positions [11]. Penaeidins are found to display activity against Gram-positive bacteria, filamentous fungi, viruses and protozoans and are also found to possess chitin-binding properties [1]. Different classes exhibit variations in their potency and target specificity against various strains of microorganisms.

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. After cleavage of signal peptide, mature peptides can be posttranslationally 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 [1]. The amino acid sequences deduced from the cDNA revealed that the penaeidins isolated from hemocytes are synthesized as precursor molecules consisting of a signal peptide (19–21 amino acids) immediately preceding the bioactive molecule [1]. The role of posttranslational modifications observed in native penaeidin, such as 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, but they increased the stability of penaeidin, to proteolysis [1]. The family appears to be characterized by (1) a highly conserved signal peptide, (2) an N-terminal proline-rich domain with the following signature (Y,F)T(R,G)P(X)2(R,K)P and (3) a C-terminal cysteine-rich domain with the following signature C(X)2-3C(X)7RXC C(X)5CC.

In shrimps, hemocytes are the main source of penaeidin production. About 30–40 % of the circulating hemocytes express penaeidins [4, 8]. Penaeidins are constitutively expressed in their mature and active form in granular hemocytes of naive shrimps and are stored within cytoplasmic granules of granular hemocyte populations [4]. The penaeidins could be secreted or released from hemocytes by degranulation into the blood upon immune response stimulation. The population of hyaline cells is devoid of penaeidins. Distribution of penaeidin transcripts and proteins is restricted to hemocytes that are present. In most of the cases, the various penaeidin subgroups have been reported to be expressed in one single individual,

suggesting that the various penaeidins may have different biological functions in shrimp.

The present study targets the molecular identification, characterization and phylogenetic analysis of a new penaeidin isoform from the hemocytes of two penaeid shrimps, *F. indicus* and *M. monoceros*. The study also reports the first AMP to be identified from *M. monoceros*. Identifying AMPs in shrimps will not only help to progress basic knowledge about shrimp immunity but also offer possible applications for disease management in aquaculture.

Materials and Methods

Total RNA was extracted from the hemocytes of the experimental organisms using TRI reagent (Sigma) following manufacturer's protocol. First-strand cDNA was generated in a 20- μ l reaction volume containing 5 μ g total RNA, 1 \times RT buffer, 2 mM dNTP, 2 μ M oligo d(T₂₀), 20 U of RNase inhibitor and 100 U of M-MLV reverse transcriptase (Fermentas, Inc., USA). 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 μ l of cDNA was performed in a 25- μ l reaction volume containing 1 \times standard Taq buffer (10 mM Tris-HCl, 50 mM KCl, pH 8.3), 2.5 mM MgCl₂, 200 μ M dNTPs, 0.4 μ M each primer and 1U Taq DNA polymerase (Fermentas Inc., USA). PCR amplifications were performed using the forward primer (5'-acctgaccctcacctgcagaggcc-3') and reverse primer (5'-acctacatccttccacaag-3'), designed using GeneTool software based on consensus sequences of penaeidins already deposited in the GenBank. The thermal profile used for the PCR amplification was 94 °C for 2 min followed by 35 cycles of 94 °C for 15 s, 65 °C for 30 s and 68 °C for 30 s and a final extension at 68 °C for 10 min. The PCR products were purified and sequenced with ABI Big Dye Terminator Cycle Sequencing Kit and analyzed in the ABI prism 377 Automated DNA sequencer at SciGenom, India.

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). Gene translation and prediction of deduced proteins were performed using ExPASy. The signal peptide was predicted by SignalP program. Multiple sequence alignments were performed with amino acid sequences of known penaeidins of shrimps using ClustalW and GeneDoc programs. Amino acid sequences of all known penaeidins were retrieved from GenBank, and phylogenetic and molecular evolutionary analyses were conducted by the neighbor-joining (NJ) method using MEGA version 5. The structural modeling of the new penaeidin isoform was performed using the

software ViewerLite 4.2 (Accelrys Inc., USA), with the PDB data generated by SWISS-MODEL prediction algorithm based on the template IueoA (99.9 Å). Helical wheel modeling was also performed using helical wheel projection program, and the formation of disulfide bridges was predicted by DiANNA 1.1 web server. The peptide was analyzed for its antimicrobial activity using antimicrobial peptide predictor program. The nucleotide sequences and deduced amino acid sequence of the penaeidin AMP were submitted to GenBank.

Results and Discussion

From the present study, a new isoform of penaeidin, belonging to subgroup III, was identified from the hemocytes of Indian white shrimp, *F. indicus*, and also from the pink shrimp, *M. monoceros*, hereinafter referred to as Fi-PEN and Mm-PEN, respectively. 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* hemocytes by RT-PCR (Fig. 1a). 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. 1b). The nucleotide sequence and deduced amino acid sequence were submitted to NCBI GenBank under the accession numbers **JX657680** (Fi-PEN) and **KF275674** (Mm-PEN).

The molecular weight (MW) of the mature peptide was found to be 5.66 kDa and theoretical isoelectric point (pI), 9.38. 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 extinction coefficient was computed to be 8855, assuming that all pairs of Cys residues form cystines, and 8480 when all Cys residues are reduced. 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 %). The new penaeidin isoform was predicted to possess a protein-binding potential of 1.72 kcal/mol, and the total net charge was calculated to be +6.

The analysis with SignalP software revealed the presence of a signal peptide with 19 amino acid residues at the

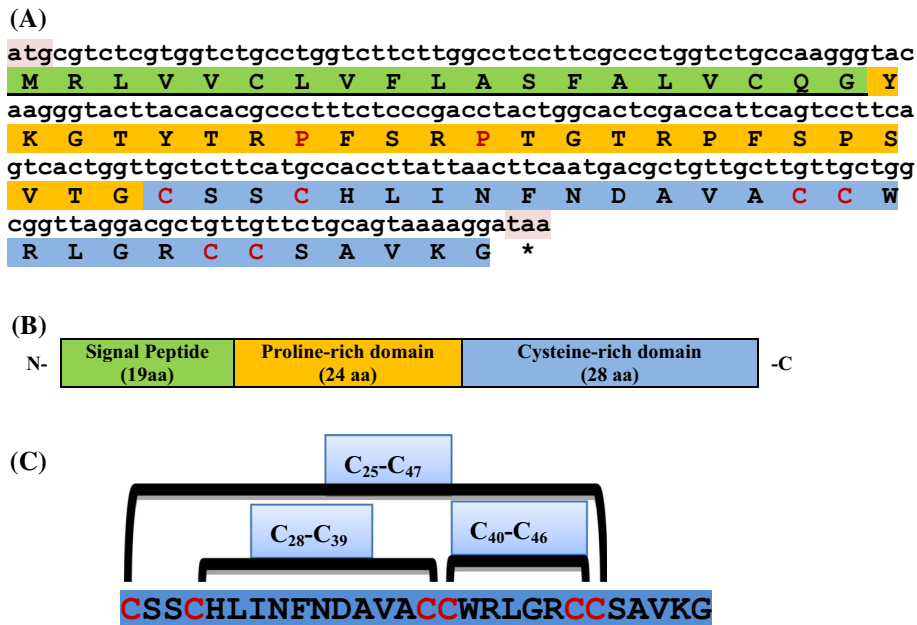


Fig. 1 a Nucleotide and amino acid sequences corresponding to the open reading frame of Fi-PEN (**JX657680**) and Mm-PEN (**KF275674**). **b** Schematic representation of the domain organization of Fi-PEN and Mm-PEN. The underlined amino acid residues highlighted in *green* color indicate a putative signal sequence; region highlighted in *yellow* indicates the proline-rich domain; and the region highlighted in *blue* indicates the cysteine-rich domain. The two

conserved proline residues of the proline-rich domain and the six cysteine residues of the cysteine-rich domain engaged in the formation of disulfide bridges are highlighted in *red* color. An *asterisk* is the stop codon. **c** Formation of disulfide bridges (C₂₅–C₄₇, C₂₈–C₃₉ and C₄₀–C₄₆) by the six cysteine residues of the cysteine-rich domain as predicted by DiANNA 1.1 web server (Color figure online)

N-terminus, with a predicted cleavage site between positions 19 and 20, i.e., CQG-YK (Fig. 1a). 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 consisting of 28 amino acid residues (Fig. 1a, b, c). Generally, the N-terminal proline-rich domain is found to be longer than the cysteine-rich domain; cysteine-rich domain is stabilized by three disulfide bonds and found to be more conserved across classes [11, 12]. However, in case of the newly identified Fi-PEN and Mm-PEN, 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. 2a). Structural analysis revealed the presence of an alpha-helix in the cysteine-rich domain which was stabilized by disulfide bonds (Fig. 2b). However, the proline-rich domain formed an extended structure. This is in agreement with the solution structure described by Yang et al. [2] and Cuthbertson et al. [12].

Sequence comparison using BLAST algorithm showed that the deduced amino acid sequence of Fi-PEN and Mm-PEN 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 *Farfantepenaeus 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 %). Fi-PEN was found to be a different isoform from the already reported penaeidins of *F. indicus* viz. Fi-penaeidin [9, 10], and Mm-PEN is the first AMP to be reported from *M. monoceros*.

Multiple alignment performed with other known penaeidins revealed conserved regions within the peptide (Fig. 3a). 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. Fi-PEN and Mm-PEN 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. [11]. As per Destoumieux et al. [1] and Gueguen et al. [11], this overall structure of penaeidins is quite unique among the AMP families. The sequence showed six highly conserved cysteine residues (Cys₂₅, Cys₂₈, Cys₃₉, Cys₄₀, Cys₄₆ and

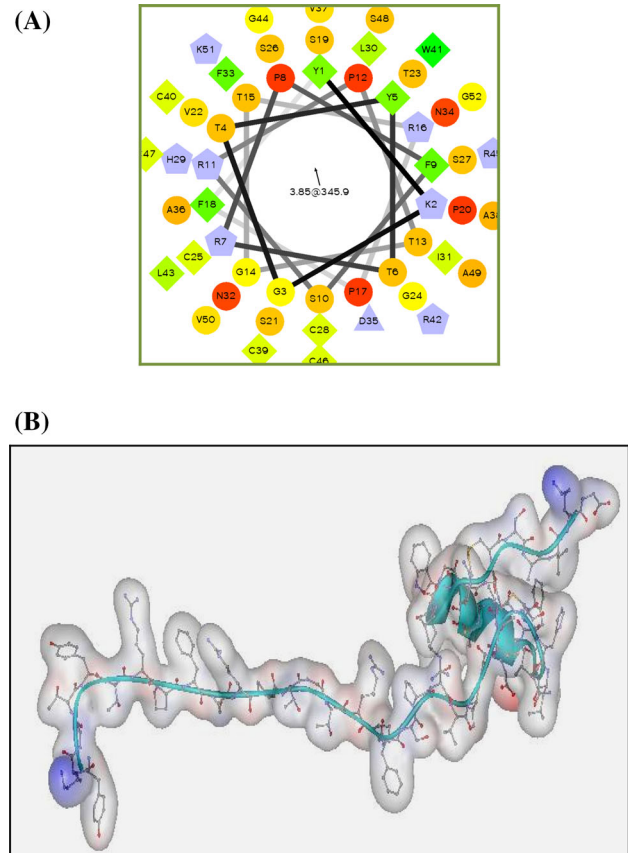
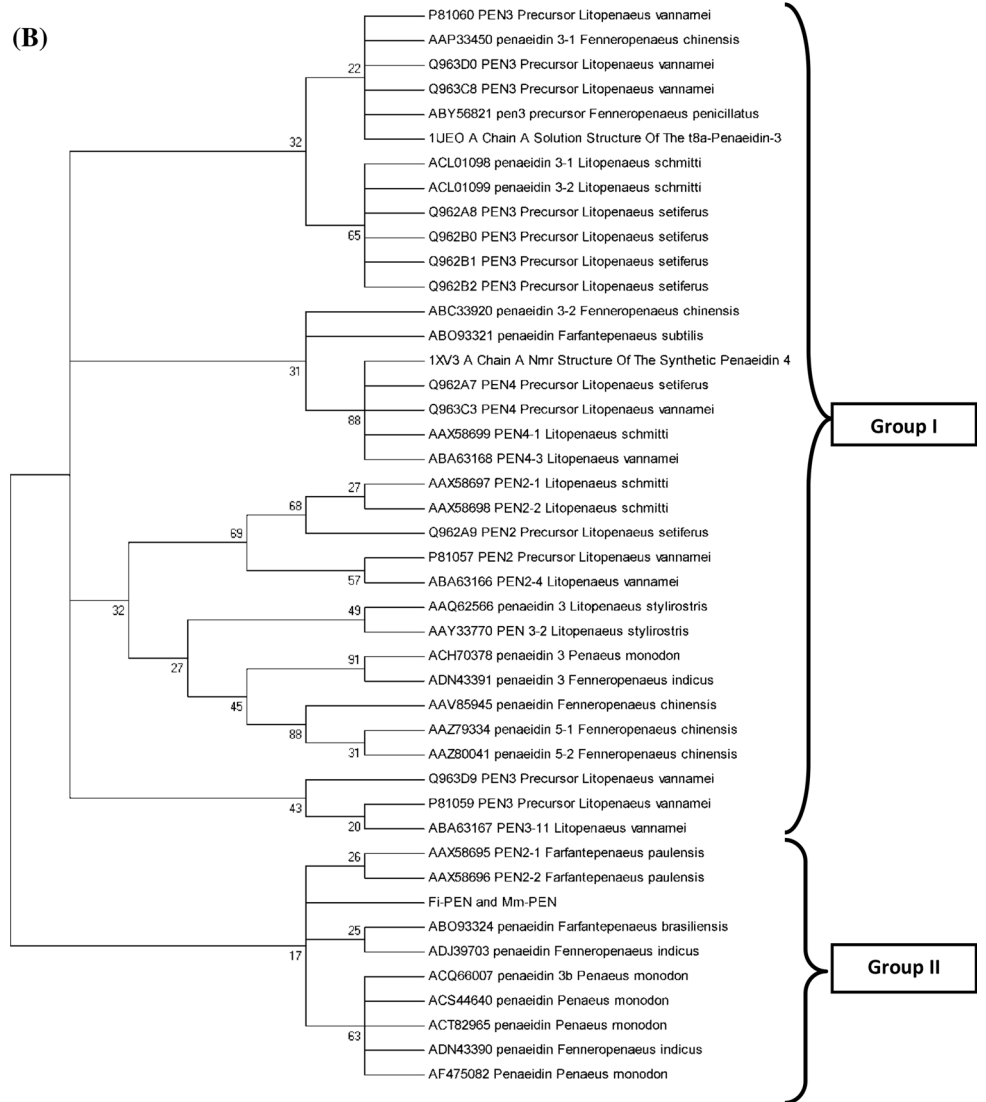
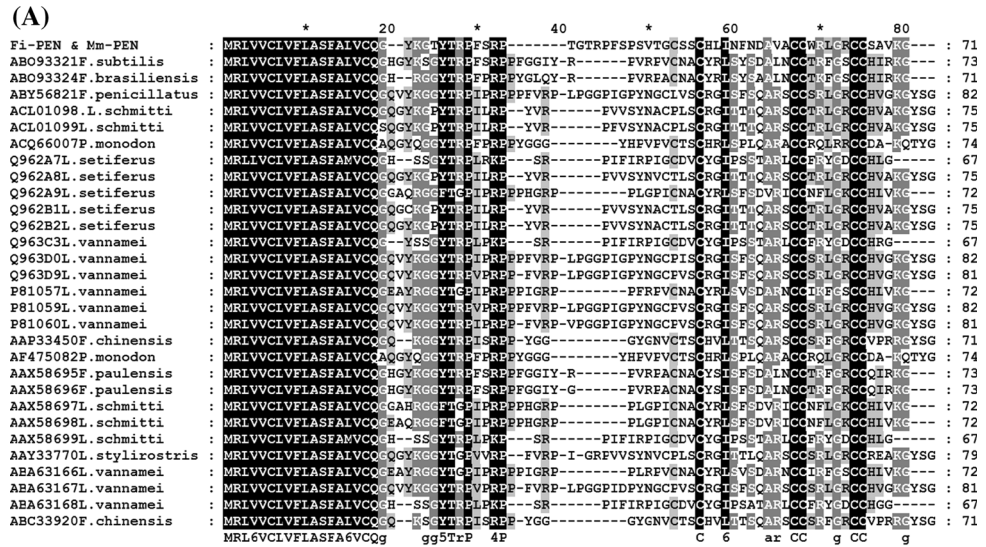


Fig. 2 **a** Helical wheel diagram demonstrating the hydrophilic and hydrophobic residues of Fi-PEN and Mm-PEN. Hydrophilic residues are represented as *circles*, hydrophobic residues as *diamonds*, potentially negatively charged as *triangles* and potentially positively charged as *pentagons*. Hydrophobicity is color-coded: The most hydrophobic residue is *green*, and the amount of *green* is decreasing proportionally to the hydrophobicity, with zero hydrophobicity coded as *yellow*. Hydrophilic residues are coded *red* with *pure red* being the most hydrophilic (uncharged) residue, and the amount of *red* is decreasing proportionally to the hydrophilicity. The potentially charged residues are *light blue*. **b** Predicted three-dimensional structure of Fi-PEN and Mm-PEN generated using ViewerLite 4.2 software (Color figure online)

Cys₄₇) which are engaged in the formation of three disulfide bridges. DiANNA 1.1 web server predicted the formation of three disulfide bridges between C₂₅–C₄₇, C₂₈–C₃₉ and C₄₀–C₄₆.

Phylogenetic tree (Fig. 3b) constructed to study the evolutionary relationships of Fi-PEN and Mm-PEN 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 only according to species, but also according to the subgroup. The tree could be broadly divided into two major groups. Group I consisted of penaeidins from *Litopenaeus* sp., *Fenneropenaeus* sp., *Farfantepenaeus* sp. and *Penaeus* sp. Group I could again

Fig. 3 a Multiple alignment of nucleotide sequence of Fi-PEN (**JX657680**) and Mm-PEN (**KF275674**) with other penaeidins obtained using GenDoc program version 2.7.0. *Black* and *gray* indicate conserved sequences. **b** Bootstrapped neighbor-joining tree obtained using MEGA version 5.0 illustrating relationships between the deduced amino acid sequences of the Fi-PEN (**JX657680**) and Mm-PEN (**KF275674**) with other penaeidins



be divided into four subgroups: subgroup I consisting of penaeidin-3 of *Litopenaeus* sp. and *Fenneropenaeus* sp.; subgroup II consisting of penaeidin-4 of *Litopenaeus* sp.; subgroup III 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 IV consisting of penaeidin-3 of *L. vannamei* alone.

However, penaeidins of group II were found to form a diverse branch and hence might be distantly related to group I, as evident from the phylogenetic tree. Group II could be divided into four subgroups. Subgroup I consisted of penaeidin-2 of *F. paulensis*; subgroup II consisted of Fi-PEN and Mm-PEN; subgroup III consisted of penaeidins from *Farfantepenaeus* sp. and *Fenneropenaeus* sp.; and subgroup IV 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 Fi-PEN to which it belongs. However, the characteristic features of various penaeidins subgroups revealed that Fi-PEN 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 III. Analysis of the amino acid sequence of Fi-PEN 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, which fulfilled all the conditions for designating it to subgroup III 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 is found to be related to structural characteristics that cannot be deduced directly from primary amino acid sequence comparisons between penaeidin isoforms [2]. 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 indicates that shrimp penaeidin makes up a

large and diverse family. The wide distribution of penaeidins in penaeid shrimps indicates the importance of these AMPs in the innate immunity.

Acknowledgments The authors are grateful to the Ministry of Earth Sciences (MoES), Government of India, for the research Grant (MoES/10-MLR/01/2012) with which the work was carried out. Authors also thank The Co-ordinator, National Centre for Aquatic and Animal Health (NCAAH), CUSAT, for scientific support.

Funding The study was funded by Ministry of Earth Sciences (MoES), Government of India (MoES/10-MLR/01/2012).

Compliance with Ethical Standards

Conflict of interest Afsal V.V., Swapna P. Antony, Rosamma Philip and I.S. Bright Singh declare that they have no conflict of interest.

Human and Animal Rights This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Destoumieux D, Bulet P, Loew D, Dorsselaer AV, Rodriguez J, Bachere E (1997) Penaeidins, a new family of antimicrobial peptides isolated from the shrimp *Penaeus vannamei* (Decapoda). *J Biol Chem* 272:28398–28406
2. Cuthbertson BJ, Shepard EF, Chapman RW, Gross PS (2002) Diversity of the penaeidin antimicrobial peptides in two shrimp species. *Immunogenetics* 54:442–445
3. Bartlett TC, Cuthbertson BJ, Shepard EF, Chapman RW, Gross PS, Warr GW (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
4. Munoz M, Vandenbulcke F, Garnier J, Gueguen Y, Bulet P, Saulnier D, Bachère E (2004) Involvement of penaeidins in defense reactions of the shrimp *Litopenaeus stylirostris* to a pathogenic vibrio. *Cell Mol Life Sci* 61:961–972
5. Barracco MA, 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
6. Supungul P, Klingbunga 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
7. Rojtinnakorn J, Hirano I, Itami T, Takahashi Y, Aoki T (2002) Gene expression in hemocytes of kuruma prawn, *Penaeus japonicus*, in response to infection with WSSV by EST approach. *Fish Shellfish Immunol* 3:69–83
8. Kang CJ, Wang JX, Zhao XF, Yang XM, Shao HL, Xiang JH (2004) Molecular cloning and expression analysis of Ch-penaeidin, an antimicrobial peptide from Chinese shrimp, *Fenneropenaeus chinensis*. *Fish Shellfish Immunol* 6:513–525
9. Antony SP, Singh ISB, Philip R (2011) Molecular characterization and phylogenetic analysis of a penaeidin-like antimicrobial peptide, Fi-penaeidin from *Fenneropenaeus indicus*. *Aquaculture* 319:298–301
10. Vaseeharan B, Shanthi S, Chen JC, 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
11. Gueguen Y, Garnier J, Robert L, Lefranc MP, Mougnot I, de Lorgeril J, Janech M, Gross PS, Warr GW, Cuthbertson B, Barracco MA, Bulet P, Aumelas A, Yang Y, Bo D, Xiang J, Tasanakajon 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
 12. Yang Y, Poncet J, 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