www.trjfas.org
ISSN 1303-2712
Turkish Journal of Fisheries and Aquatic Sciences 14: 379-389 (2014) DOI: 10.4194/1303-2712-v14_2_08

Molecular Characterization of a Novel Cathepsin B from Striped Murrel *Channa striatus*: Bioinformatics Analysis, Gene Expression, Synthesis of Peptide and Antimicrobial Property

Jesu Arockiaraj[1,*], Venkatesh Kumaresan[1], Mukesh Kumar Chaurasia[1], Prasanth Bhatt[1], Rajesh Palanisamy[1], Mukesh Pasupuleti[2], Annie J. Gnanam[3], Marimuthu Kasi[4]

[1] SRM University, Faculty of Science and Humanities, Department of Biotechnology, Division of Fisheries Biotechnology and Molecular Biology, Kattankulathur 603 203, Chennai, Tamil Nadu, India.
[2] CSIR-Central Drug Research Institute, Lab PCN 206, Microbiology Division, B.S. 10/1, Sector 10, Jankipuram Extension, Sitapur Road, Lucknow-226031, Uttar Pradesh, India
[3] The University of Texas at Austin, Institute for Cellular and Molecular Biology, 1 University Station A4800, Austin, Texas, 78712, USA
[4] AIMST University, Faculty of Applied Sciences, Department of Biotechnology, Semeling, 08100 Bedong, Kedah, Malaysia.

* Corresponding Author: Tel.: +91.44 27452270; Fax: +91.44 27453903;
E-mail: jesuaraj@hotmail.com

Received 07 January 2014
Accepted 13 April 2014

Abstract

In this study, we have reported a full length cDNA of cathepsin B identified from the constructed cDNA library of snakehead murrel *Channa striatus* by genome sequence FLX technology. The identified full length *C. striatus* cathepsin B (*Cs*Cath B) is 1486 base pairs (bp) long which contains 990 bp open reading frame (ORF). The ORF region encodes 330 amino acids with a molecular mass of 36 k Da. This amino acid sequence contains three thiol protease motifs at 101-112, 275-285 and 292-311 with their respective active sites *viz.,* Cys^{107}, His^{277} and Asp^{297}. *Cs*Cath B exhibited the maximum similarity (87%) with Cath B from mangrove red snapper, *Lutjanus argentimaculatus*. Phylogenetically, *Cs*Cath B is clustered together with the fish groups belonging to perciformes. A predicted 3D model of *Cs*Cath B revealed 11 α-helix and 10 β-strands. *Cs*Cath B contains higher percentage (10%) of coils due to the presence of many glycine residues (36 residues). The highest gene expression ($P<0.05$) was noticed in liver. Further, the expression was induced with fungal (*Aphanomyces invadans*) and bacterial (*Aeromonas hydrophila*) infections. The predicted antimicrobial region of *Cs*Cath B was synthesized to study its antimicrobial property. The peptide exhibited the antimicrobial activity towards Gram negative and Gram positive bacteria. The overall results indicate that *Cs*Cath B is a potential molecule for further studies on murrel defense mechanism.

Keywords: Cathepsin B, murrel, fungus, bacteria, antimicrobial peptide.

Introduction

Cathepsins are lysosomal proteases secreted in response to external stimuli (Holt *et al*., 2006). Cathepsins exist in different forms of isomers that have either unique or overlapping functions. Turk *et al.* (2001) reported that most of the cathepsins are cysteine proteases (for e.g. B, C, F, H, K and L), whereas cathepsin A and G are serine proteases and cathepsin D and E are aspartic proteases. Cathepsins possess both exopeptidase as well as endopeptidase activity and they cleave their substrates non-specifically (Chwieralski *et al*., 2006). Cathepsin molecules are associated with cell death regulation and/or apoptosis-like caspases (Turk *et al*., 2000; Foghsgaard *et al*., 2001; Guicciardi *et al*., 2001; Salvesen, 2001). It is involved in class II major histocompatibility complex (MHC) maturation, keratinocyte differentiation, tumor progression as well as metastasis, remodeling of bone, osteoarthritis and rheumatoid arthritis (Friedrichs *et al*., 2003; Vasiljeva *et al*., 2007).

In humans, cathepsin B which has antiamyloidogenic and neuroprotective function plays an important role in removing the amyloid plaques in case of Alzheimer disease, by reducing the levels of amyloid-β peptides via proteolytic cleavage (Mueller-Steiner *et al*., 2006). Cathepsin B, L (cysteine protease) and cathepsin D (aspartic protease) along with other molecules like plasminogen activator and matrix metalloproteinase (MMP) degrade basement membrane and extracellular matrix (Adeni *et al*., 1995). Thomssen *et al.* (1995) revealed that malignant tumors have elevated levels of cathepsin B and cathepsin L when compared to nonmalignant tumors and also determined that their elevated levels lead to relapse, thus giving an insight into the prognosis.

Cathepsins are responsible for apoptosis of inflammatory cells such as eosinophils, neutrophils and basophils, which are further phagocytosed by macrophages. Among the cathepsins reported, cathepsin B is involved in inflammatory disorders, cancer and Alzheimer's disease (Aoki and Ueno, 1997). Release of cathepsin B into the cytosol induces cleavage of Bid [BH3 interacting domain] which is a member of Bcl-2 family. It activates the caspases and subsequent apoptosis of cells (Blomgran *et al*., 2007) thus proving that cathepsin B is an important molecule for the subsequent innate immune response. Many researchers (Itami *et al*., 1987; Takahashi *et al*., 1987; Aranishi, 1999) reported that the cysteine proteinases including cathepsin B and L extracted from fish skin mucus is a potential bacteriolysin which is involved in nonspecific immunity of fish. This phenomenon was already established in a few species (Aranishi, 1999).

Striped murrel *Channa striatus* is a tropical, freshwater, carnivorous and air breathing fish found in Indian subcontinent, China and Southeast Asian countries. The medicinal properties of this species have been extensively utilized in these countries. Jais *et al.* (1994) reported that it is used to treat skin diseases due to the presence of docosohexanoic acid (DHA), an essential fatty acid with neutraceutical values. The overall wound healing property of this species is attributed to its already existing antimicrobial, antifungal and platelet aggregation properties (Jais *et al*., 1994). Vitamin A, an important molecule required for wound healing is also present in very high concentrations in *C. striatus* (Aoki *et al*., 1997). Extracts of this fish have also been used in aggregation of platelets thus helping in blood clotting during haemorrhagic conditions. Murrel fillet extract has antinociceptive property thus leading to healing, which is due to the presence of hydromethanolic portion in the extract (Jais *et al*., 1997).

Infectious diseases, especially epizootic ulcerative syndrome (EUS) creates a serious problem for this species resulting in heavy economic losses (Lilley and Roberts, 1997). EUS is one of the most destructive diseases among murrels in the Asian Pacific region. It is very common in both northern and southern India and has spread over rivers, reservoirs and paddy fields to various states, causing substantial loss to farmers (Dhanaraj *et al*., 2008). Dhanaraj *et al.* (2008) reported that a fungus *Aphanomyces invadans* is the primary causative agent of EUS. The secondary infections are caused by various bacterial species especially *Aeromonas hydrophila*. Therefore, research on snakehead murrel immune system is necessary to establish a disease control method particularly against EUS. Though the information on cathepsin B from fish (Liu *et al*., 2008; Zhang *et al*., 2008; Whang *et al*., 2011), crustacean (Stephens *et al*., 2012) and mollusk (Wang *et al*., 2008) are available, there has been no such report of *C. striatus* cathepsin B (designated as *Cs*Cath B). To gain insight into the characterization of *Cs*Cath B and its role in *C. striatus*, a full length cDNA of *Cs*Cath B was identified from the *C. striatus* cDNA library constructed by Genome Sequencing FLX (GS FLX) technology. The transcriptional differentiation of *Cs*Cath B mRNA has been analyzed after challenging with *A. invadans* and *A. hydrophila*. Moreover, we predicted an antimicrobial region from *Cs*Cath B based on the earlier studies reported elsewhere and it was synthesized as a short peptide to study its antimicrobial property.

Materials and Methods

Fish

Healthy *C. striatus* (average body weight of 50 g) were obtained from Center for Aquaculture Research and Extension (CARE), St. Xavier's College (Autonomous), Palayamkottai, Tamil Nadu, India. Fishes were maintained in flat-bottomed plastic tanks (250 L) with aerated and filtered freshwater (water quality: dissolved oxygen, 5.8 ± 0.2 mg/L ; water temperature, 28 ± 1 °C and pH, 7.2 ± 0.2). The fishes were acclimatized for 1 week before being challenged to *A. invadans* and *A. hydrophila*. A maximum of 20 fishes per tank were maintained during the experiment.

cDNA Library Construction, Identification and Bioinformatics Analysis of *Cs*Cath B

A full length cathepsin B was identified from the constructed *C. striatus* cDNA library by the genome sequence FLX^{TM} technology. The detailed procedure on *C. striatus* cDNA library construction was described in our earlier studies (Arockiaraj *et al*., 2013a; Abirami *et al*., 2013). From the established cDNA library of *C. striatus* sequence database, we identified a full length cathepsin B gene, which we designated as *Cs*Cath B. The full-length *Cs*Cath B sequence was compared with other sequences available in NCBI database (http://blast.ncbi.nlm.nih.gov/Blast) and the similarities were analyzed (Arockiaraj *et al*., 2012a). The open reading frame (ORF) and amino acid sequence of *Cs*Cath B was obtained by using DNAssist (ver. 2.2.). Characteristic domains or motifs were identified using the PROSITE profile database (http://prosite.expasy.org/scanprosite/). The N-terminal transmembrane sequence was determined by DAS transmembrane prediction program (http://www.sbc.su.se/~miklos/DAS). Signal peptide analysis was done using the SignalP (http://www.cbs.dtu.dk). Multiple sequence alignment was carried out on ClustalW (ver. 2) (http://www.ebi.ac.uk/Tools/msa/clustalw2/) program to find out the evolutionarily conserved residues among the different organisms. The evolutionary history of *Cs*Cath B was inferred using the Neighbor-Joining method on MEGA 5. The evolutionary distances were computed using the Poisson correction method (Uinuk-Ool *et al*., 2003). The 3D structure of the *Cs*Cath B protein was predicted by utilizing the I-Tasser server (http://zhanglab.ccmb.med.umich.edu/I-TASSER).
The obtained model was validated using Ramachandran plot analysis (http://mordred.bioc.cam.ac.uk/~rapper/rampage.php). The antimicrobial region was predicted using the AMPA web server (Torrent *et al*., 2012). The window size and the threshold values were set as default.

Immune Challenge Experiment

For fungus induced mRNA expression analysis, the fish were injected with *A. invadans* (10^2 spores). In our earlier report (Bhatt *et al*., 2013), we have clearly explained the isolation of *A. invadans* from the infected *C. striatus* muscle, culture in the laboratory, identification and injection to the fishes. For bacterial challenge, the fish were injected intraperitonealy with *A. hydrophila* (5×10^6 CFU/ml) suspended in 1X phosphate buffer saline (100 µl/fish). *A. hydrophila* was also isolated and identified from the muscle sample of EUS infected *C. striatus* as described by Dhanaraj *et al.* (2008). Samples were collected before (0 h), and after injection (3, 6, 12, 24 and 48 h) and were immediately snap-frozen in liquid nitrogen and stored at -80°C until total RNA was isolated. Using a sterilized syringe, the blood (0.5-1.0 ml per fish) was collected from the fish caudal fin and immediately centrifuged at 4000 X g for 10 min at 4°C to allow blood cell collection for total RNA extraction. PBS (1X) were prepared and served as control (100 µl/fish). Five fishes were collected from each time schedule in *A. invadans, A. hydrophila* and PBS induced groups.

RNA Isolation and cDNA Conversion

Total RNA from the control and infected fish were isolated using Tri Reagent™ (Life Technologies), according to the manufacture's protocol with slight modifications (Arockiaraj *et al*., 2011a, 2011b). Using 2.5 µg of RNA, first strand cDNA synthesis was carried out using a SuperScript® VILO™ cDNA Synthesis Kit (Life technologies) as suggested by the manufacturer with slight modifications (Arockiaraj *et al*., 2013b, 2013c). The resulting cDNA solution was stored at -20°C for further analysis.

Gene Expression Studies

The relative expression of *Cs*Cath B in blood, gills, liver, heart, spleen, intestine, head, kidney, skin, muscle and brain were measured by quantitative real time polymerase chain reaction (qRT-PCR) (Arockiaraj *et al*., 2012b and 2012c). qRT-PCR was carried out using a BIO-RAD CFX384 Touch Real-Time PCR Detection System in 20µl reaction volume containing 4µl of cDNA from each tissue, 10µl of Fast SYBR® Green Master Mix, 0.5µl of each primer (20 pmol/µl) and 5µl dH_2O. The qRT-PCR cycle profile was 1 cycle of 95°C for 10s, followed by 35 cycle of 95°C for 5s, 58°C for 10s and 72°C for 20s and finally 1 cycle of 95°C for 15s, 60°C for 30s and 95°C for 15s. The same qRT-PCR cycle profile was used for the internal control gene, β-actin. The internal control primers were designed from the β-actin of *C. striatus* (GenBank Accession Number EU570219). The primer details of gene specific primer (*Cs*Cath B) and internal control (β-actin) are as follows: *Cs*Cath B F1: CACACCCAAGTGCGTCTATAA and *Cs*Cath B R2: GAATCTCCT CCTCGGTTGAAAG; β-actin F3: TCTTCCAGCCTTCCTTCCTTGGTA and β-actin R4: GACGT CGCACTTCATGATGCTGTT. After the PCR program, data were analyzed with BIO-RAD software. To maintain consistency, the baseline was set automatically by the software. The comparative CT method ($2^{-\delta\delta CT}$ method) was used to analyze the expression level of *Cs*Cath B (Livak and Schmittgenm, 2001). The computed gene expression of *Cs*Cath B was compared with the corresponding expression level of brain for the tissue specific gene expression. For examination of the relative fold change after being challenged with *A. invadans* and *A. hydrophila*, the relative gene expression at each time point of infected fish was compared to the corresponding PBS injected control.

***C. striatus Cs*Cath B Peptide Synthesis**

The predicted antimicrobial region ([55]QKLCGTKLNGPK[66]) was synthesized by solid-phase peptide synthesis method (Sigma-Aldrich). Then, the peptide was purified using reverse phase high pressure liquid chromatography (HPLC) by performing a Hitachi HPLC (Chromaster). The purity of the synthesized peptide was analyzed using HPLC analytical column. The integrity of the purified peptide was subjected to MS/MS analysis to determine the quality of the peptide as described by Mann and Aebersold (2003). The purified peptide was dissolved in sterilized water in order to minimize the risk of contamination as suggested by Somboonwiwat *et al.* (2005).

Antimicrobial Property of Synthesized *Cs*Cath B Peptide

The antimicrobial property of the peptide was analyzed as explained in our earlier study (Arockiaraj *et al*., 2013d). The study was performed using various Gram negative (*A. hydrophila, E. coli, Edwardsiella tarda, Vibrio parahaemolyticus, V. alginolyticus* and *V. harveyi*) and Gram positive (*Bacillus subtilis, Streptococcus iniae, Staphylococcus aureus, Enterococcus faecium* and *Lactococcus lactis*) bacteria. Ampicillin (100 µg) and the same volume of DEPC treated nuclease-free de-ionized water was used as positive and negative controls respectively. The 1.5% broth agar containing bacteria (OD_{600} = 0.1) was poured onto petriplates. The peptide at different concentrations (0, 25, 50 and 100 µg), positive control and negative control were added into the individual wells in the agar plates and incubated at 35 °C for 20 h. The diameter of the inhibition zone (millimeter) was determined. The assays were conducted in three duplications and the values are presented here as average ± standard deviation.

Statistical Analysis

All the statistical analysis was performed in SPSS (ver. 11.5). The data were subjected to one-way ANOVA and the mean comparisons were performed by Tukey's Multiple Range Test and the significance was determined at $P<0.05$ level.

Results

Bioinformatics Characterization of *Cs*Cath B

*Cs*Cath B cDNA was obtained from the constructed cDNA library of *C. striatus* using Genome Sequencing FLX technology. This sequence was submitted to NCBI database under the accession number JX469845. *Cs*Cath B was subjected to analysis for determining the physico-chemical properties using DNAssist software. The data revealed that, *Cs*Cath B is 1486 base pairs (bp) long, with 124 bp 5' untranslated region (UTR), 990 bp open reading frame (ORF) and 372 bp 3' UTR. The ORF region encodes 330 amino acids with a theoretical molecular weight of 36 kDa and isoelectric point 6. This amino acid contains three thiol protease motifs at ^{101}QGSCGSCWAFGA^{112} (thiol cysteine protease), ^{275}GGHAIKVLGWG^{285} (thiol histidine protease) and ^{292}YWLCANSWNTDWGDNGFFKF^{311} (thiol asparagine protease) with their respective active sites *viz*., Cys^{107}, His^{277} and Asp^{297} (Figure 1). An occluding loop is present in the amino acid region of *Cs*Cath B at ^{179}PYTIAPCEHHVNGSRPPCTGE^{199}. Other than these gene specific hits, another 23 high probability hits were also observed from *Cs*Cath B and is presented in Table 1.

The protein sequence of *Cs*Cath B was compared with five other homologous sequences of cathepsin B from different species including *Lutjanus argentimaculatus, Oplegnathus fasciatus, Xenopus laevis, Gallus gallus* and *Homo sapiens* using

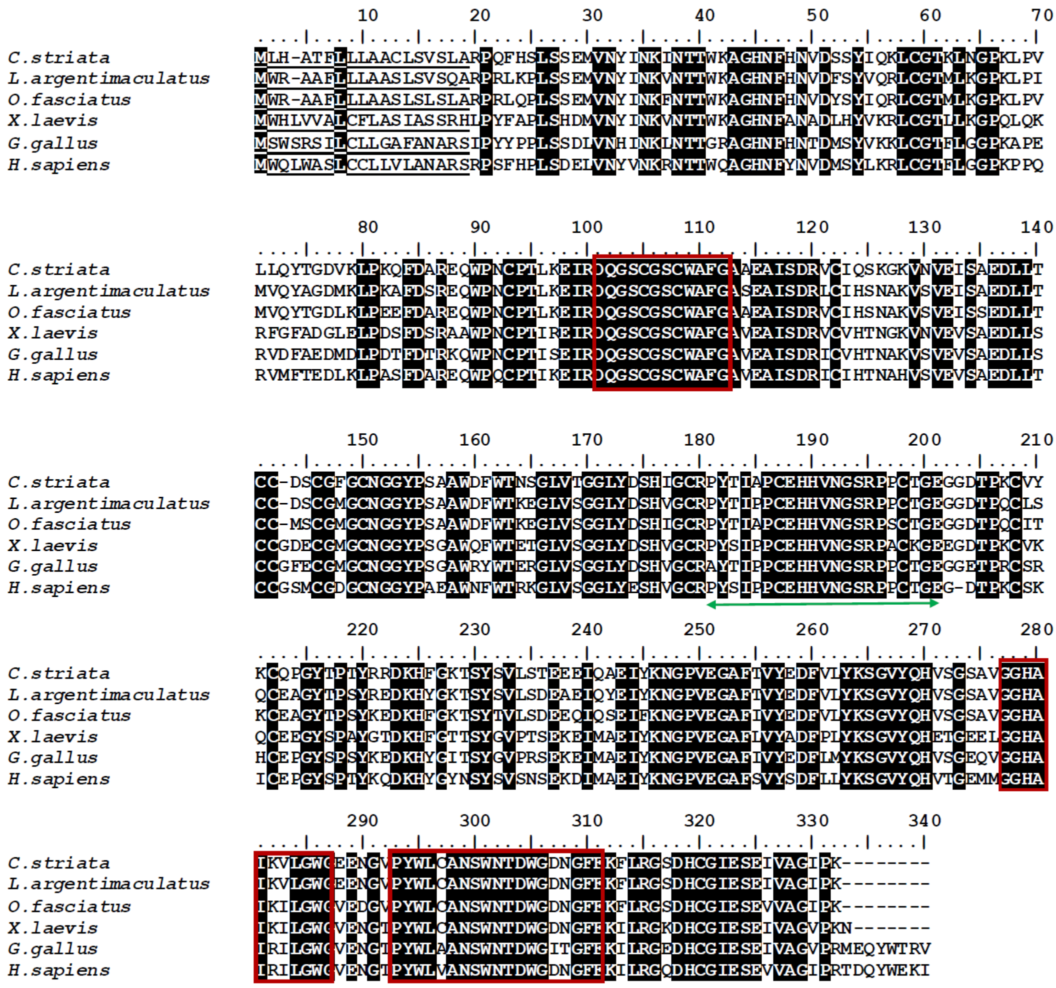

Figure 1. Multiple sequence alignment of *Cs*Cath B. This analysis was performed on ClustalW (ver. 2.0) using sequences of cathepsin B belonging to different phyla [*Lutjanus argentimaculatus* (ACO82382), *Oplegnathus fasciatus* (AEA48884), *Xenopus laevis* (NP_001079570), *Gallus gallus* (NP_990702) and *Homo sapiens* (NP_001899)]. The Signal sequence is underlined. The gene specific motifs thiol (cysteine, histidine and asparagine) protease is boxed in red color. The occluding loop is marked in green arrow. The identical residues are shaded in black color. The numbers represent the position of the amino acid residue. The dashes (-) represent the gap.

ClustalW. The results of multiple sequence alignment reveals that *Cs*Cath B showed 87%, 86%, 71%, 69% and 72% similarity with *L. argentimaculatus, O. fasciatus, X. laevis, G. gallus* and *H. sapiens* respectively (data not shown). Moreover, this result shows that *Cs*Cath B possesses mature form light chain between 80 and 125 and this light chain carries a thiol cysteine protease motif which is conserved in all the sequence that has been taken for multiple sequence analysis (Figure 1). Another mature form heavy chain is present between 126 and 327 with this heavy chain carrying a thiol histidine protease motif, thiol asparagine protease and a putative occluding loop, which were all conserved in the sequences taken for multiple sequence analysis (Figure 1). *Cs*Cath B shares a structural similarity with human cathepsin B, both containing a pre-region (1-19 residues) and a pro-region (20-77) before the mature form light chain.

The phylogenetic tree showed five different clades which includes higher vertebrates (mammals and birds), lower vertebrates (amphibians and fishes) and invertebrates (arthropods). *Cs*Cath B clustered together with the fish groups (*O. fasciatus* and *L. argentimaculatus*) belonging to perciformes (Figure 2). *Cs*Cath B has 81% bootstraps identity with its fish groups. Cathepsin B from an insect *Triatoma sordid* was set as an out group.

We predicted five different 3D models of *Cs*Cath B protein and the quality of the models were evaluated using Ramachandran Plot analysis. The analysis indicated that among the 330 amino acids residues of *Cs*Cath B, model-1 shares 290 amino acid residues in favored region (88.40%), 20 residues in allowed region (6.01%) and 17 residues in outlier (or disallowed) region (5.50%) (data not shown). Hence, model-1 was selected as the best model for further analysis and is given in Figure 3. The predicted 3D model of *Cs*Cath B (Figure 3) shows 11 α-helix (23%) at 3-16, 27-36, 51-57, 85-88, 94-96, 107-123, 135-139, 153-161, 218-221, 235-244 and 257-259, 10

Table 1. Details of high probability hits from *Cs*Cath B amino acid

Hits (Nos.)	Position of amino acid
N-glycosylation site (2)	37-40 & 190-193
Protein kinase C phosphorylation site (5)	39-41, 94-96, 117-119, 203-205 & 217-219
N-myristoylation site (10)	59-64, 102-107,105-110, 111-116, 147-152, 150-155, 164-169, 169-174, 271-276 & 314-319
Casein kinase II phosphorylation site (6)	94-97, 133-136, 139-142, 233-236, 234-237 & 253-256

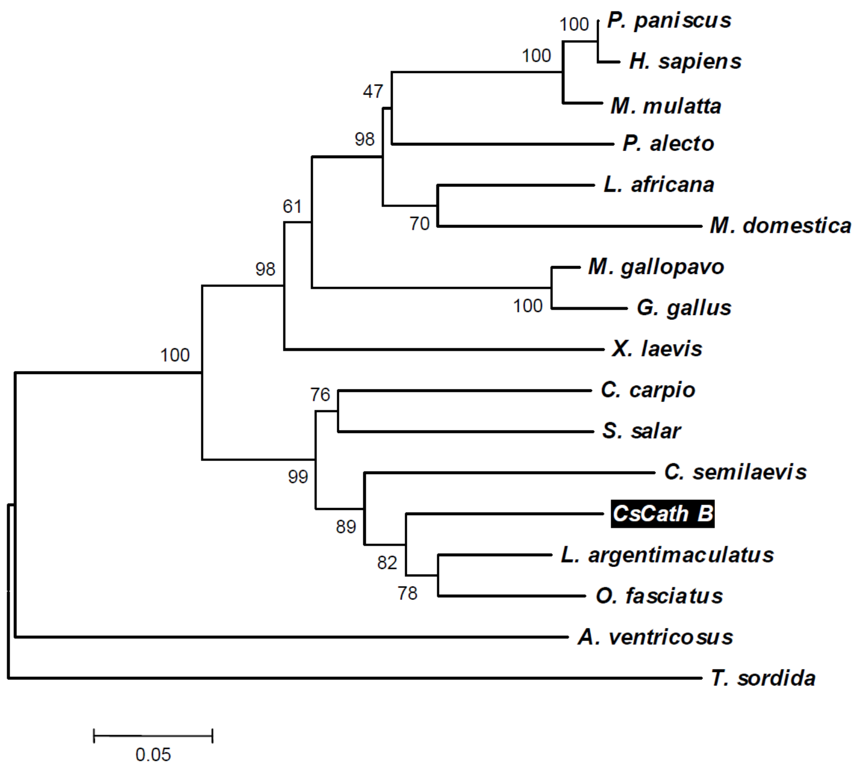

Figure 2. Phylogenetic analysis of *Cs*Cath B with other homologous constructed on MEGA (ver. 5.1) using Neighbour-Joining Method. The numbers mentioned at nodes indicate bootstraps in percentage after 1000 replications. For GenBank accession number and complete species details please refer Table 2.

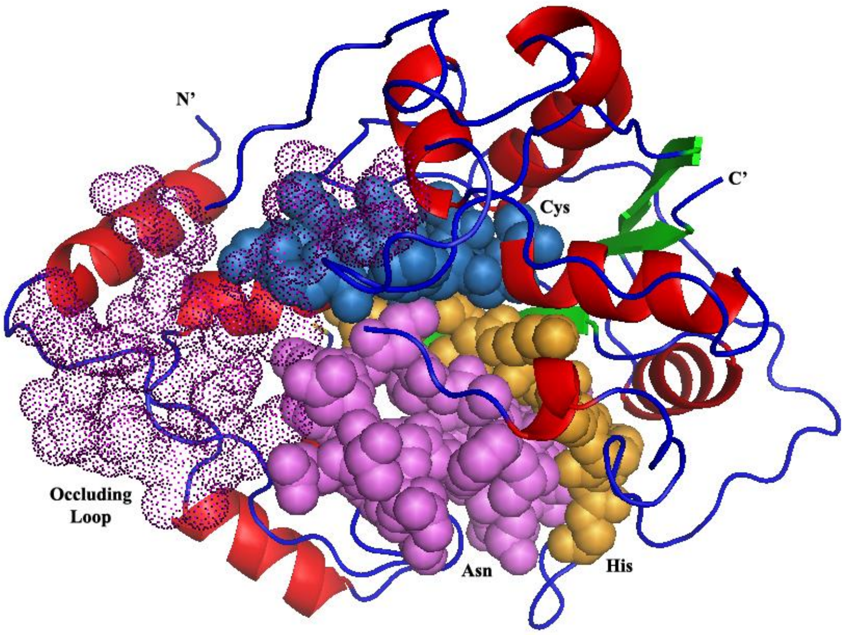

Figure 3. The predicted 3D structure of *Cs*Cath B constructed on I-TASSER program. Methionine (1) and lysine (330) represent the N-terminal and C-terminal regions respectively. The α-helices, β-sheets and coils are presented in red, green and blue colors respectively. The active sites of *Cs*Cath B were highlighted in different color balls *viz.*, thiol cysteine proteases in blue color, thiol histidine proteases in brown color and thiol asparagine proteases in pink color. The occluding loop region is highlighted as dots.

β-strands (12%) at 41-43, 131-133, 166-168, 229-231, 248-254, 260-262, 275-282, 291-297, 307-313 and 324-326 and 65% random coils. Higher percentage of coils is present in *Cs*Cath B, which is due to the presence of enormous amount of glycine (36 residues = 10%) in the sequence. This polypeptide has small side chains, hence, it is unable to contribute the formation of α-helix and β-sheet.

Gene Expression of CsCath B

To study the tissue distribution of *Cs*Cath B transcripts, total RNA was isolated from various tissues including blood, heart, liver, spleen, intestine, kidney, head kidney, gills, skin, muscle and brain. The isolated RNA were converted into cDNA and subjected to gene expression analysis using quantitative real time PCR. The tissue distribution result shows that the *Cs*Cath B transcript was expressed in all the examined tissues. Significantly ($P<0.05$) highest gene expression was noticed in liver and lowest expression in brain (Figure 4A). Based on the results of tissue distribution, gene expression in liver tissue was studied after being infected with *A. invadans* and *A. hydrophila*. In *A. invadans* infected *C. striatus*, Cath B mRNA expression almost remains in the basal level until 3 h post injection (p.i.) and then the expression started increasing and finally it reached significantly ($P<0.05$) higher expression at 48 h p.i. (Figure 4B). In *A. hydrophila* infected *Cs*Cath B, mRNA expression was significantly ($P<0.05$) higher at 24 h p.i. compared to PBS injected control (14 fold) (Figure 4C) and then the expression level started decreasing.

Prediction of Antimicrobial Region in CsCath B

The antimicrobial region of *Cs*Cath B (^{55}QKLCGTKLNGPK^{66}) was predicted through AMPA web server program. The sequences contain proline hinge and lysine rich regions. The results clearly indicate the antimicrobial property of the peptide. Protparam analysis showed the instability index of the *Cs*Cath B antimicrobial peptide to be -24.58 which makes the peptide highly stable. The hydrophobic index of the *Cs*Cath B peptide is 25% and its total net charge was determined to be +2, thereby confirming the antimicrobial nature of the peptide. BLAST alignment showed that the peptide has 42% identity with gramicidin S from *Bacillus brevis*, an antimicrobial peptide in which the lysine and proline residues remained conserved.

Synthesis of *Cs*Cath B Peptide and its Antimicrobial Activity

The predicted antimicrobial region of *Cs*Cath B peptide was synthesized and its integrity was measured to be 87.5%. The *Cs*Cath B derived peptide was used to examine its antimicrobial capacity. The antimicrobial activity of the *Cs*Cath B peptide at various concentrations against different Gram negative and Gram positive bacteria was tested using agar well diffusion method. The peptide exhibited

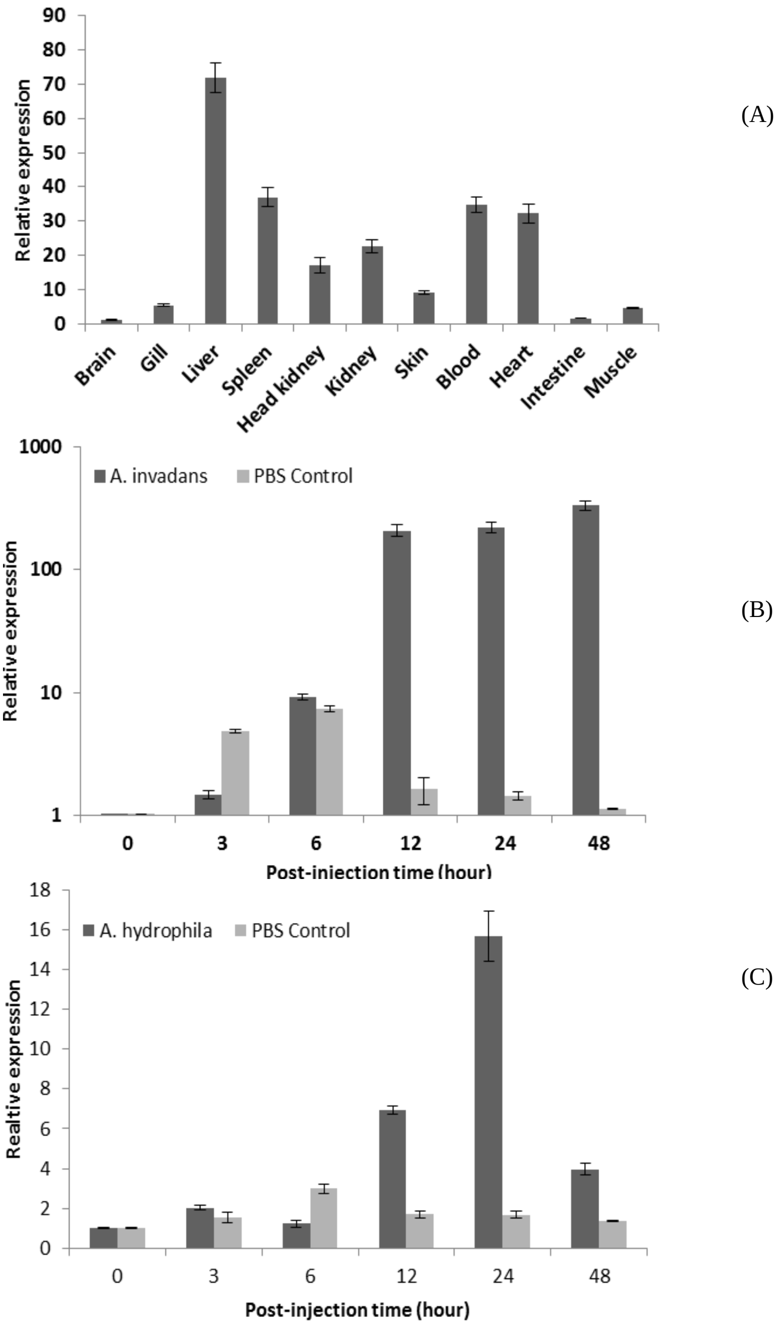

Figure 4. Gene transcript patterns of *Cs*Cath B by real time PCR. 4A: Tissue distribution of *Cs*Cath B in different tissues of *C. striatus*. Data are expressed as a ratio to *Cs*Cath B mRNA transcription in brain. 4B and 4C: The time course of *Cs*Cath B mRNA transcription in liver at 0, 3, 6, 12, 24, and 48 h post injection with *A. invadans* and *A. hydrophila* respectively. Data are expressed as a ratio to *Cs*Cath B mRNA in sample from PBS injected control group.

antimicrobial property against both Gram negative and Gram positive bacteria. Significantly ($P < 0.05$) highest inhibition zone was noticed in Gram negative bacteria *A. hydrophila* followed by *E. coli, V. harveyi, E. tarda, V. alginolyticus* and *V. parahaemolyticus* and Gram positive bacteria *S. iniae, S. aureus, L. lactis, E. faecium* and *B. subtilis.* The positive control yielded the highest activity and the negative control showed no activity (Figure 5). The results indicated that the antimicrobial activity of the synthesized peptide is concentration dependent.

Discussion

The identified *Cs*Cath B gene encodes a protein having structural features that are distinct to vertebrate cathepsin family due to the presence of thiol protease with conserved cysteine, histidine and asparagine active sites at Cys107, His277 and Asn297, thus maintaining the functional aspects of cathepsin B (Whang *et al*., 2011). Lacaille *et al.* (2002) reported that these active sites play crucial roles in the formation and stabilization of the catalytic

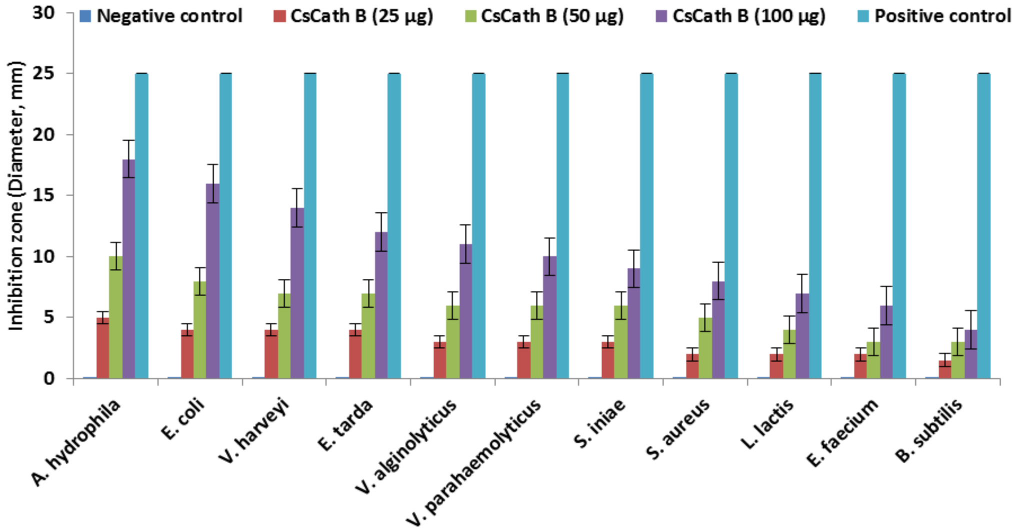

Figure 5. Antimicrobial capacity of recombinant *Cs*Cath B protein against the Gram negative and Gram positive bacteria. The diameters of the inhibition zone showing the expression of the recombinant *Cs*Cath B protein at various concentrations along with its positive and negative controls.

site of the activated enzyme. Bioinformatics analysis suggested that CsCath B the *Cs*Cath B is a typical cathepsin B cysteine protease with a typical signal peptide sequence between Met[1] and Ala[19], a pro-domain between Arg[20] and ASP[77] and a mature domain between Leu[80] and Val[327] (Stephens *et al.*, 2012). Illy *et al.* (1997) reported that cathepsin B has a unique carboxy- peptidyl activity, which is attributed to the presence of occluding loop. In *Cs*Cath B, we also noticed a putative occluding loop at 179-199. Illy *et al.* (1997) reported that occluding loop is interrupts in binding of extended peptides as well as binding of protein protease inhibitors. Moreover, Illy *et al.* (1997) stated that this occluding loop with two histidine residues at His[187] and His[188] has the ability to accept negatively charged carboxylate ion and this ion provides basis for peptidase activity of cathepsin B. Musil *et al.* (1991) reported that all the cathepsin B carries His[187], His[188], Glu[249] and Glu[323] as conserved residues, which are responsible for exopeptidase and endopeptidase activity. The same pattern was noticed in *Cs*Cath B also. Moreover, the sequence analysis showed that *Cs*Cath B contains twelve conserved cysteine residues at 59, 93, 105, 108, 122, 141, 142, 146, 150, 179, 187 and 198, which suggest the potential formation of six disulphide bridges. However, two additional conserved cysteines (208 and 212) were noticed in *Cs*Cath B at the C-terminal region similar to cath B from *Litopenaeus vanameii.* Two potential N-glycosylation sites were identified in *Cs*Cath B at 37-40 and 190-193, which are necessary for the intracellular transport mechanism as reported by Ma *et al.* (2010).

The multiple sequence alignment results indicate that *Cs*Cath B amino acid sequence was homologous to other known cathepsin B due to the presence of pre-region, pro-region, mature form light chain and heavy chain. The phylogenetic analysis of *Cs*CathB also provide evidence to prove that the identified sequence from *C. striatus* shares homology with cathepsin B of most of the known species due to the gene specific domains which are necessary for their function.

The 3D model analysis of *Cs*Cath B revealed that the gene specific motifs, thiol cysteine protease is present in the α-helical region between 101 and 112 and thiol histidine protease and thiol asparagine protease are present in the β-strands at 275-285 and 292-311 respectively. The thiol protease motifs with active sites at Cys[107], His[277] and Asn[297] are conserved, thus maintaining functional aspects of cathepsin B (Whang *et al*., 2011).

*Cs*Cath B was detected in all the tissues taken for analysis and the highest expression was observed at liver. Zhang *et al.* (2008), Whang *et al.* (2011), Akoi *et al.* (2003) and Feng *et al.* (2011) also observed the gene expression of cathepsin B in many tissues and the highest at the haemopoietic organs like liver, heart and kidney in various fishes. Moreover, the fungal and bacterial infection induced significant induction of *Cs*Cath B expression in liver. This observation indicates an involvement of *Cs*Cath B in host immune response against fungal and bacterial infection. These pathogen induced gene expression is related to inflammation, cytokine activity, antigen presentation and binding activity as reported by Trent *et al.* (2006). The variation in the gene expression of *Cs*Cath B in different time point is due to varied pathogenicity levels during the infection. As reported in the literature (Darawiroj *et al*., 2008), it has been explained that the feasibility of difference in gene expression is due to immune induction in fish tissues.

Cathepsin B has been reported as antimicrobial protein in many organisms including Japanese eel (Aranishi, 1999). The antimicrobial prediction program of *Cs*Cath B peptide showed that the proline hinge region and lysine rich region increases the antimicrobial capacity in *Cs*Cath B peptide. Hiromi and Jimmy (2008) reported that the lysine rich antimicrobial peptides are highly bioactive. Moreover, Markossian *et al.* (2004) and Sitaram (2006) stated that the proline hinge in antimicrobial peptide plays a major role in peptide's membrane translocation, membrane permeabilization and antimicrobial activity. Otvos (2002) also reported that the proline hinge region is predominant in many antimicrobial peptides (AMPs), for example proline-rich antimicrobial peptides. In addition, Thennarasu and Nagaraj (1996) and Suh *et al.* (1999) observed that mutating proline residues generate remarkable changes in the properties. Nguyen *et al.* (1990) found that the protease enzymes especially cathepsin B and L extracted from the fish mucus are involved in degrading the proteoglycans which are the major components of bacterial cell walls. Similarly, Aranishi (1999) also observed the antimicrobial nature of cysteine proteases in Japanese eel. Therefore, we studied the antimicrobial nature of *Cs*Cath B derived peptide. Hence, we predicted an antimicrobial region through AMPA web server program, which determines IC_{50} (half maximal inhibitory concentration) of AMP. The results showed that the peptide have the ability to inhibit the activity of both Gram negative and Gram positive bacteria. Itami *et al.* (1987) and Takahashi *et al.* (1987) also reported the bacteriolysis nature of cysteine protease. Overall, the observation indicated that cathepsin B may be a novel bacteriolysin which is involved in the nonspecific defense system of murrel, thus showing it antimicrobial property.

Acknowledgements

This research is supported by DBT's Prestigious Ramalingaswami Re-entry Fellowship (D.O.NO.BT/HRD/35/02/2006), Department of Biotechnology, Ministry of Science and Technology, Government of India, New Delhi.

References

Abirami, A., Venkatesh, K., Akila, S., Rajesh, P., Nagaram, P., Prasanth, B., Arpita, R., Thirumalai, M.K., Gnanam, A.J., Mukesh, P., Kasi, M. and Arockiaraj, J. 2013. Fish lily type lectin-1 contains β-prism 2 architecture: Immunological characterization. Molecular Immunology, 56: 497-506. doi: 10.1016/j.molimm.2013.06.020.

Adeni, A., Huet, G., Zerimech, F., Hecquet, B., Balduyck, M. and Peyrata, J.P. 1995. Cathepsin B, L, and D activities in colorectal carcinomas: relationship with clinico-pathological parameters. Cancer Letters, 96: 267-275. doi: 10.1016/0304-3835(95)03930-U.

Aoki, H., Ahsan, M.N. and Watabe, S. 2003. Molecular cloning and characterization of cathepsin B from the hepatopancreas of northern shrimp *Pandalus borealis*. Comparitive Biochemistry Physiology, Part B, 134: 681-694. doi: 10.1016/S1096-4959(03)00023-X.

Aoki, T., Nakano, T. and Ueno, R. 1997. Purification and some properties of a latent form cathepsin L from mackerel whitemuscle. Fish Sciences, 63: 824-829.

Aoki, T. and Ueno, R. 1997. Involvement of cathepsins B and L in the post-mortem autolysis of mackerel muscle. Food Research International, 30: 585-591. doi: 10.1016/S0963-9969(98)00014-3.

Aranishi, F. 1999. Lysis of pathogenic bacteria by epidermal cathepsins L and B in the Japanese eel. Fish Physiology and Biochemistry, 20: 37-41. doi: 10.1023/A:1007763711158.

Arockiaraj, J., Annie, J.G., Dhanaraj, M., Ranganath, G., James, M., Arun, S., Saravanan, M., Kasi, M. and Subha, B. 2013b. Crustin, a WAP domain containing antimicrobial peptide from freshwater prawn *M. rosenbergii*: Immune characterization. Fish and Shellfish Immunology, 34: 109-118. doi: 10.1016/j.fsi.2012.10.009.

Arockiaraj, J., Annie, J.G., Dhanaraj, M., Thirumalai, M.K., Mukesh, P., Milton, J. and Kasi, M. 2013c. *Macrobrachium rosenbergii* cathepsin L: Molecular characterization and gene expression in response to viral and bacterial infections. Microbiology Research, 168: 569-579. doi: 10.1016/j.micres.2013.04.007.

Arockiaraj, J., Puganeshwaran, V., Sarasvathi, E., Arun, S., Rofina, Y.O. and Subha, B. 2012c. Molecular functions of chaperonin gene, containing tailless complex polypeptide 1 from *Macrobrachium rosenbergii*. Gene, 508: 241-249. doi: 10.1016/j.gene.2012.07.050.

Arockiaraj, J., Puganeshwaran, V., Sarasvathi, E., Arun, S., Tahereh, A., Rofina, Y.O. and Subha, B. 2011b. Gene profiling and characterization of arginine kinase-1 (*Mr*AK-1) from freshwater giant prawn (*Macrobrachium rosenbergii*). Fish and Shellfish Immunology, 31: 81-89. doi: 10.1016/j.fsi.2011.04.004.

Arockiaraj, J., Sarasvathi, E., Puganeshwaran, V., Arunsingh, S.V., Othman, R.Y. and Subha, B. 2012b. Immunological role of thiol-dependent peroxiredoxin gene in *Macrobrachium rosenbergii*. Fish and Shellfish Immunology, 33: 121-129. doi: 10.1016/j.fsi.2012.04.010.

Arockiaraj, J., Annie, J.G., Dhanaraj, M., Mukesh, P., Milton, J. and Arun, S. 2013a. An upstream initiator caspase 10 of snakehead murrel *Channa striatus*, containing DED, p20 and p10 subunits: Molecular cloning, gene expression and proteolytic activity. Fish and Shellfish Immunology, 34: 505-513. doi: 10.1016/j.fsi.2012.11.040.

Arockiaraj, J., Annie, J.G., Venkatesh, K., Rajesh, P., Prasanth, B., Thirumalai, M.K., Arpita, R., Mukesh, P. and Kasi, M. 2013d. An unconventional antimicrobial protein histone from freshwater prawn *Macrobrachium rosenbergii*: Analysis of immune properties. Fish and Shellfish Immunology, 35: 1511-1522. doi: 10.1016/j.fsi.2013.08.018.

Arockiaraj, J., Sarasvathi, E., Puganeshwaran, V., Arun, S., Rofina, Y.O. and Subha, B. 2011a. Effect of infectious hypodermal and hematopoietic necrosis virus (IHHNV) infection on caspase 3c expression and activity in freshwater prawn *Macrobrachium rosenbergii*. Fish and Shellfish Immunology, 32: 161-169. doi: 10.1016/j.fsi.2011.11.006.

Arockiaraj, J., Sarasvathi, E., Puganeshwaran, V., Arun, S., Rofina, Y.O. and Subha, B. 2012a. Molecular cloning, characterization and gene expression of an antioxidant enzyme catalase (*Mr*Cat) from *Macrobrachium rosenbergii*. Fish and Shellfish Immunology, 32: 670-682. doi: 10.1016/j.fsi.2012.01.013.

Bhatt, P., Venkatesh, K., Rajesh, P., Mukesh, K.C., Annie, J.G., Mukesh, P. and Arockiaraj, J. 2013. Immunological role of C4 CC chemokine-1 from snakehead murrel *Channa striatus.* Molecular Immunology, 57: 292-293. doi: 10.1016/j.molimm.2013.10.012.

Blomgran, R., Zheng, L. and Stendahl, O. 2007. Cathepsin-cleaved Bid promotes apoptosis in human neutrophils via oxidative stress-induced lysosomal membrane permeabilization. Journal of Leucine Biology, 81: 1213-1223. doi:10.1189/jlb.0506359.

Chwieralski, C.E., Welte, T. and Bühling, F. 2006. Cathepsin-regulated apoptosis. Apoptosis, 11: 143-149. doi: 10.1007/s10495-006-3486-y.

Darawiroj, D., Kondo, H., Hirono, I. and Aoki, T. 2008. Immune-related gene expression profiling of yellowtail (*Seriola quinqueradiata*) kidney cells stimulated with ConA and LPS using microarray analysis. Fish and Shellfish Immunology, 24: 260-266. doi: 10.1016/j.fsi.2007.07.011.

Dhanaraj, M., Haniffa, M.A., Ramakrishnan, C.M. and Arunsingh, S.V. 2008. Microbial flora from the epizootic ulcerative syndrome (EUS) infected murrel *Channa striatus* (Bloch, 1797) in Tirunelveli region. Turkish Journal of Veterinary and Animal Sciences, 32: 221-224.

Feng, T., Zhang, H., Liu, H., Zhou, Z., Niu, D., Wong, L., Kucuktas, H., Liu, X., Peatman, E. and Liu, Z. 2011. Molecular characterization and expression analysis of the channel catfish cathepsin D genes. Fish and Shellfish Immunology, 31: 164-169. doi: 10.1016/j.fsi.2011.04.006.

Foghsgaard, L., Wissing, D., Mauch, D., Lademann, U., Bastholm, L., Boes, M., Elling, F., Leist, M. and Jaattela, M. 2001. Cathepsin B acts as a dominant execution protease in tumor cell apoptosis induced by tumor necrosis factor. Journal of Cell Biology, 153: 999-1010. doi: 10.1083/jcb.153.5.999.

Friedrichs, B., Tepel, C., Reinheckel, T., Deussing, J., von Figura, K., Herzog, V., Peters, C., Saftig, P. and Brix, K. 2003. Thyroid functions of mouse cathepsins B, K, and L. Journal of Clinical Investigation, 111: 1733-1745. *doi:* 10.1172/JCI200315990.

Guicciardi, M.E., Miyoshi, H., Bronk, S.F. and Gores, G.J. 2001. Cathepsin B knockout mice are resistant to tumor necrosis factor-α mediated hepatocyte apoptosis and liver injury: implications for therapeutic applications. American Journal of Pathology, 159: 2045-2054. doi: 10.1016/S0002-9440(10)63056-8.

Hiromi, S. and Jimmy, B.F. 2008. Lysine-Enriched Cecropin-Mellitin Antimicrobial Peptides with Enhanced Selectivity. Antimicrobial Agents Chemotherapy, 52: 4463-4465. doi: 10.1128/AAC.00810-08.

Holt, O.J., Gallo, F. and Griffiths, G.M. 2006. Regulating secretory lysosomes. J. Biochem., 140: 7-12. doi: 10.1093/jb/mvj126.

Illy, C., Quraishi, O., Wang, J., Purisima, E., Vernet, T. and Mort, JS. 1997. Role of the occluding loop in cathepsin B activity. Journal of Biological Chemistry, 272: 1197-1202. doi: 10.1074/jbc.272.2.1197.

Itami, T., Takahashi, Y. and Sato, K. 1987. Bacteriolytic activities in the skin mucus of ayu. Nippon Suisan Gakkaishi, 53: 401-406.

Jais, A.M.M., Dambisya, Y.M. and Lee, T.L. 1997. Antinociceptive activity of *Channa striatus* extracts in mice. Journal of Ethnopharmacology, 57: 125-130. doi: 10.1016/S0378-8741(97)00057-3.

Jais, A.M.M., Mc Culloch, R. and Croft, K.D. 1994. Fatty acid and amino acid composition of Haruan- Potential for Wound healing. General Pharmacology, 25: 947-950. doi: 10.1016/0306-3623(94)90101-5.

Lacaille, F., Kaleta, J. and Brömme, D. 2002. Human and parasitic papain-like cysteine proteases: their role in physiology and pathology and recent developments in inhibitor design. Chemical Reviews, 102:4459-88. doi: 10.1021/cr0101656.

Lilley, J.H. and Roberts, R.J. 1997. Pathogenicity and culture studies comparing the *Aphanomyces* involved in epizootic ulcerative syndrome (EUS) with other similar fungi. Journal of Fish Diseases, 20: 135-144. doi: 10.1046/j.1365-2761.1997.d01-116.x.

Liu, H., Yin, L., Zhang, N., Li, S. and Ma, C. 2008. Isolation of cathepsin B from the muscle of silver carp (*Hypophthalmichthys molitrix*) and comparison of cathepsins B and L actions on surimi gel softening. Food Chemistry, 110: 310-318. doi:10.1016/j.foodchem.2008.01.068.

Livak, K.J. and Schmittgenm, T.D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. Methods, 25: 402-408. doi:10.1006/meth.2001.1262.

Ma, J., Zhang, D., Jiang, J., Cui, S., Pu, H. and Jiang, S. 2010. Molecular characterization and expression analysis of cathepsin L1 cysteine protease from pearl oyster *Pinctada fucata.* Fish and Shellfish Immunology, 29: 501-507. doi: 10.1016/j.fsi.2010.05.006.

Mann, M. and Aebersold, R. 2003. Mass spectrometry-based proteomics. Nature, 422: 198-207. doi:10.1038/nature01511.

Markossian, K.A., Zamyatnin, A.A. and Kurganov, B.I. 2004. Antibacterial proline-rich oligopeptides and their target proteins. Biochemistry (Mosc), 69: 1082-1091. doi: 10.1023/B:BIRY.0000046881.29486.51.

Mueller-Steiner, S., Zhou, Y., Arai, H., Roberson, E.D., Sun, B., Chen, J., Wang, X., Yu, G., Esposito, L., Mucke, L. and Gan, L. 2006. Antiamyloidogenic and neuroprotective functions of cathepsin B: Implications for Alzheimer's disease. Neuron, 51: 703-714. doi:10.1016/j.neuron.2006.07.027.

Musil, D., Zucic, D., Turk, D., Engh, R.A., Mayr, I., Huber, R., Popovic, T., Turk, V., Towatari, T. and Katunuma, N. 1991. The refined 2.15 Å X-ray crystal structure of human liver cathepsin B: the structural basis for its specificity. EMBO Journal, 10: 2321-2330.

Nguyen, Q., Mort, J.S. and Roughley, P.J. 1990. Cartilage proteoglycan aggregate is degraded more extensively by cathepsin L than by cathepsin B. Biochemistry Journal., 266: 569-573.

Otvos, L Jr. 2002. The short proline-rich antibacterial peptide family. Cellular and Molecular Life Sciences, 59: 1138-1150. doi: 10.1007/s00018-002-8493-8.

Salvesen, G.S. 2001. A lysosomal protease enters the death scene. Journal of Clinical Investigation, 107: 21-23. doi: 10.1172/JCI11829.

Sitaram, N. 2006. Antimicrobial peptides with unusual amino acid compositions and unusual structures. Current Medicinal Chemistry, 13: 679-696. doi: 10.2174/092986706776055689.

Somboonwiwat, K., Marcos, M., Tassanakajon, A., Klinbunga, S., Aumelas, A., Romestand, B., Gueguen, Y., Boze, H., Moulin, G. and Bachère, E. 2005. Recombinant expression and anti-microbial activity of anti-lipopolysaccharide factor (ALF) from the black tiger shrimp Penaeus monodon. Developmental and Comparative Immunology, 29: 841-851. doi: 10.1016/j.dci.2005.02.004.

Stephens, A., Rojo, L., Araujo-Bernal, S., Garcia-Carreño, F. and Muhlia-Almazan, A. 2012. Cathepsin B from the white shrimp *Litopenaeus vannamei*: cDNA sequence analysis, tissues-specific expression and biological activity. Comparative Biochemistry and Physiology Part B, 161:32-40. doi: 10.1016/j.cbpb.2011.09.004.

Suh, J.Y., Lee, Y.T., Park, C.B,, Lee, K.H,, Kim, S.C. and Choi, B.S. 1999. Structural and functional implications of a proline residue in the antimicrobial peptide gaegurin. Europena Journal of Biochemistry, 266: 665-674. doi: 10.1046/j.1432-1327.1999.00917.x.

Takahashi, Y., Kajiwaki, T., Itami, T. and Konegawa, K. 1987. Enzymatic properties of the bacteriolytic substances in the skin mucus of yellowtail. Nippon Suisan Gakkaishi, 53: 425-431.

Thennarasu, S. and Nagaraj, R. 1996. Specific antimicrobial and hemolytic activities of 18-residue peptides derived from the amino terminal region of the toxin pardaxin. Protein Engineering 9: 1219-1224. doi: 10.1093/protein/9.12.1219.

Thomssen, C., Schmitt, M., Goretzki, L., Oppelt, P., Pache, L., Dettmar, P., Janicke, F. and Graeff, H. 1995. Prognostic value of the cysteine proteases Cathepsin B and Cathepsin L in human breast cancer. Clinical Cancer Research, 1: 741-746.

Torrent, M., Tommaso, P.D., Pulido, D., Nogués, M.V., Notredame, C., Boix, E. and Andreu, D. 2012. AMPA: an automated web server for prediction of protein antimicrobial regions. Bioinformatics, 28: 130-131. doi: 10.1093/bioinformatics/btr604.

Trent, M.S., Stead, C.M., Tran, A.X. and Hankins, J.V. 2006. Diversity of endotoxin and its impact on pathogenesis. Journal of Endotoxin Research, 12: 200-223. doi: 10.1177/09680519060120040201.

Turk, B., Turk, D. and Turk, V. 2000. Lysosomal cysteine proteases: more than scavengers. Biochemica et Biophysica Acta, 1477: 98-111. doi: 10.1016/S0167-4838(99)00263-0.

Turk, V., Turk, B. and Turk, D. 2001. Lysosomal cysteine proteases: facts and opportunities EMBO Journal, 20: 4629-4633. doi: 10.1093/emboj/20.17.4629

Uinuk-Ool, T.S., Takezaki, N., Kuroda, N., Figueroa, F., Sato, A., Samonte, I.E., Mayer, W.E. and Klein, J. 2003. Phylogeny of antigen-processing enzymes: cathepsins of a cephalochordate, an agnathan and a bony fish. Scandinavian Journal of Immunology, 58: 436-448. doi: 10.1046/j.1365-3083.2003.01322.x.

Vasiljeva, O., Reinheckel, T., Peters, C., Turk, D., Turk, V. and Turk, B. 2007. Emerging roles of cysteine cathepsins in disease and their potential as drug targets. Current Pharmaceutical Design, 13: 387-403. **doi:** 10.2174/138161207780162962.

Wang, X., Liu, B., Wang, G., Tang, B. and Xiang, J. 2008. Molecular cloning and functional analysis of cathepsin B in nutrient metabolism during larval development in *Meretrix meretrix*. Aquaculture, 282: 41-46. doi: 10.1016/j.aquaculture.2008.06.014.

Whang, I., De Zoysa, M., Nikapitiya, C., Lee, Y., Kim, Y., Lee, S., et al. 2011. Molecular characterization and expression analysis of Cathepsin B and L cysteine proteases from rock bream (*Oplegnathus fasciatus*). Fish and Shellfish Immunology, 30: 763-772. doi: 10.1016/j.fsi.2010.12.022.

Zhang, F.T., Zhang, Y.B., Chen, Y.D., Zhu, R., Dong, C.W., Li, Y.Y., Zhang, Q.Y. and Gui, J.F. 2008. Expressional induction of *Paralichthys olivaceus* cathepsin B gene in response to virus, poly I:C and lipopolysaccharide. Fish and Shellfish Immunology, 25: 542-549. doi: 10.1016/j.fsi.2008.07.018.