Fish & Shellfish Immunology 34 (2013) 1505–1513

Contents lists available at SciVerse ScienceDirect

Fish & Shellfish Immunology

journal homepage: www.elsevier.com/locate/fsi

Expression profile of cystatin B ortholog from Manila clam (*Ruditapes philippinarum*) in host pathology with respect to its structural and functional properties

H.K.A. Premachandra [a], Don Anushka Sandaruwan Elvitigala [a], Ilson Whang [a,*], Eunmi Kim [b], Mahanama De Zoysa [c], Bong-Soo Lim [d], Sang-Yeob Yeo [e], Seokryel Kim [f], Myoung-Ae Park [f], Hae-Chul Park [b], Jehee Lee [a,d,*]

[a] *Department of Marine Life Sciences, School of Marine Biomedical Sciences, Jeju National University, Jeju Special Self-Governing Province 690-756, Republic of Korea*
[b] *Graduate School of Medicine, Korea University, Ansan, Gyeonggido 425-707, Republic of Korea*
[c] *College of Veterinary Medicine, Chungnam National University, Yuseong-gu, Daejeon 305-764, Republic of Korea*
[d] *Marine and Environmental Institute, Jeju National University, Jeju Special Self-Governing Province 690-814, Republic of Korea*
[e] *Department of Biotechnology, Division of Applied Chemistry & Biotechnology, Hanbat National University, Daejeon 305-719, Republic of Korea*
[f] *Pathology Division, National Fisheries and Research & Developmental Institute, Busan 619-705, Republic of Korea*

ARTICLE INFO

Article history:
Received 28 June 2012
Received in revised form
18 February 2013
Accepted 11 March 2013
Available online 22 March 2013

Keywords:
Cystatin B
Manila clam
Structural characterization
Papain inhibitory activity
Transcriptional analysis

ABSTRACT

Cystatins are a well-characterized group of cysteine protease inhibitors, which play crucial roles in physiology and immunity. In the present study, an invertebrate ortholog of cystatin B was identified in Manila clam (*Ruditapes philippinarum*) (RpCytB) and characterized at the molecular level, demonstrating its inhibitory activity against the well-known cysteine protease, papain. The complete coding sequence of RpCytB (297 bp in length) encodes a 99 amino acid peptide with a calculated molecular mass of 11 kDa and a theoretical isoelectric point of 5.9. The derived peptide was found to harbor typical features of cystatin proteins, including the 'Q-X-V-X-G' motif, which was identified as QLVAG in RpCytB. Phylogenetic analysis of RpCytB revealed close evolutionary relationships with its invertebrate counterparts, especially those from mollusks. Recombinant RpCytB (rRpCytB) was overexpressed as a fusion with maltose binding protein (MBP) in *Escherichia coli* BL21 (DE3) cells. Purified rRpCytB fusion protein exhibited a detectable inhibitory activity against papain, while the control MBP showed an almost constant negligible activity. While quantitative RT-PCR detected ubiquitous RpCytB expression in all tissues examined, the expressions in hemocytes and gills were relatively higher. Upon *in vivo* immune challenge with lipopolysaccharide (LPS), the expression of RpCytB in gills and hemocytes was down-regulated. Similar challenges with poly I:C and intact *Vibrio tapetis* bacteria revealed a complicated transcriptional regulation, wherein mRNA expression levels fluctuated over time of exposure. Moreover, a precise induction of RpCytB expression after bacterial infection was detected in gills by *in situ* hybridization. Collectively, our findings in this study indicate that RpCytB expression is sensitive to host pathological conditions and may contribute cysteine protease inhibitory activity to modulate the immune response.

1. Introduction

Proteolysis is a universal mechanism mediated by proteinases that act to catalyze the hydrolysis of amide bonds, thereby regulating the function and fate of myriad proteins in physiologic and pathogenic conditions [1]. Proteinases are distributed among almost all the taxon levels in the biosphere, from viruses to the complicated vertebral organisms [2]. The cysteine proteinases represent one of the five major types of proteinases, and they have been classified according to the nature of their catalytic sites. The

* Corresponding authors. Marine Molecular Genetics Lab, Department of Marine Life Sciences, College of Ocean Science, Jeju National University, 66 Jejudaehakno, Ara-Dong, Jeju 690-756, Republic of Korea. Tel.: +82 64 754 3472; fax: +82 64 756 3493.
E-mail addresses: ilsonwhang@hanmail.net (I. Whang), jehee@jejunu.ac.kr, jeheedaum@hanmail.net (J. Lee).

1050-4648/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.fsi.2013.03.349

most well studied proteinase is the thiol (-SH)-dependent group, which utilizes cysteine residues as nucleophiles to attack the amide bonds of a target protein [2,3]. Cysteine proteinases are involved in crucial events in physiology of an organism, including lysosome degeneration, antigen presentation, proteolytic processing of pro-enzymes and pro-hormones, fertilization, proliferation, differentiation, and apoptosis [3,4]. These enzymes are relatively less abundant in prokaryotes, but show closely related phylogenetic relationship among them [5]. In Gram-positive and Gram-negative bacteria, the cysteine proteinases are mainly involved in peptidoglycan turnover, but also function in various housekeeping processes or as virulence factors that may target specific components of the host antimicrobial defense system [6]. Viruses also produce cysteine proteinases as crucial elements of their pathogenesis [7]. However imbalanced activity of endogenous cysteine proteinases may lead to numerous pathological conditions, such as tumorigenesis and multiple sclerosis [8]. Therefore, certainly the eukaryotic systems have evolved precise regulatory mechanisms for cysteine proteinases, modulating their expression, secretion, and maturation through specific degradation of mature enzymes and targeted inhibition [9].

Cystatins are considered as the largest and the best described group of natural exosite binding cysteine proteinase inhibitors, which obstruct the access of substrate without interacting with the enzyme catalytic center [10]. Cystatins predominantly interact with cysteine proteases, such as the plant-derived papain and human cathepsins B, H, and L [11]. The cystatins are now considered as cystatin superfamily, which is mainly divided into three families, including the stafins (family 1: stefins A and B, also known as cystatins A and B), cystatins (family 2), and kininogens (family 3). However, recent studies have identified a novel cystatin family in invertebrates, designated as family 4 [12,13].

The stafins are unglycosylated inhibitors (~11 kDa in size), which lack a signal sequence and disulfide bonds, and are generally expressed intracellularly [14]. Cystatin A and B proteins show extensive sequence similarity in their primary structure [14]. Cystatin B is unique among the cysteine proteinase inhibitor superfamily, in that it contains a free cysteine in the N-terminal segment of the proteinase-binding region that facilitates particularly tight binding of target proteinases. Cystatin B reversibly inhibits the papain-like cysteine proteinases, and was found to play a protective role in organisms due to its omnipresent distribution in cells and tissues [15,16]. It most likely prevents inappropriate proteolysis caused by the action of lysozymal cysteine proteinases, primarily the cathepsins B, H, K, L, and S [17]. In addition, cystatin B has been implicated in innate immune responses to bacterial infections in vertebrate and invertebrate lineages [18–20], as well as in responses against viral infections [21]. These observations may reflect the potential inhibitory action of cystatin B toward pathogen-derived cysteine proteinases, which are secreted as virulent factors or as suppressors of host cysteine proteinases to promote survival of the pathogenic microorganisms.

To date, cystatin B orthologs have been identified in several vertebrates and some invertebrates [13,19,20,22], but only a single cystatin B was characterized from invertebrate species, the duck leech *Theromyzon tessulatum* [20], except to our recent study on cystatin B from disk abalone *Haliotis discus discus* [22]. However, another member from family 1 cystatins has been characterized in the Chinese mitten crab *Eriocheir sinensis* [13]. Considering the broad phylogenetic distribution of cysteine proteinases, it is likely that functional cystatin B variants exist in invertebrate lineages.

Over the past decade, marine food sources have overtaken a large portion of the world food market, representing a promising supply of nutritional resources to compensate for the decrease in terrestrial sources that has resulted from overconsumption. This trend has promoted mariculture to the status of an economically important component of the global economy. Manila clam (*Ruditapes philippinarum*) is an edible marine bivalve mollusk species, which is harvested in large scale by the mariculture industry. As such, there is significant interest in gaining a detailed understanding of the host–pathogen interactions and immune modulations in this mariculture creature to prevent drastic losses of its harvest due to different pathogenic infections.

Here, we describe the identification and characterization of a novel cystatin B ortholog from Manila clam (RpCytB) from a structural and functional standpoint. Investigations into its potential role in immune responses against *Vibrio tapetis*, lipopolysaccharides (LPS), and polyinosinic:polycytidylic acid (poly I:C) were carried out as well.

2. Materials and methods

2.1. Identification of the cystatin B subunit

The partial cDNA sequence encoding RpCytB was identified using the basic local alignment tool (BLAST; http://blast.ncbi.nlm.nih.gov/Blast.cgi) from our previously established Manila clam cDNA sequence database [23].

2.2. In silico analysis of DNA and protein sequences

The amino acid sequence corresponding to RpCytB coding sequence was derived using DNAssist 2.2 (version 3.0). The orthologous sequences of RpCytB were compared by the BLAST search program. Pairwise sequence alignment (http://www.ebi.ac.uk/Tools/psa) and multiple sequence alignment were performed by the ClustalW2 program [24]. Phylogenetic analysis was carried out by the Neighbor-Joining method with bootstrapping values taken from 1000 replicates using Molecular Evolutionary Genetics Analysis (MEGA) software, version 5 [25]. Prediction of characteristic protein domains was carried out by the ExPASy-prosite database (http://prosite.expasy.org) and MotifScan scanning algorithm (http://myhits.isb-sib.ch/cgi-bin/motif_scan). Some of the physicochemical properties of RpCytB were determined by the ExPASy ProtParam tool (http://web.expasy.org/protparam).

2.3. Cloning of the RpCytB open reading frame (ORF)

To generate a recombinant RpCytB protein fused with maltose binding protein (MBP), designated as rRpCytB, the coding sequence of RpCytB was first PCR amplified using gene-specific cloning primers F2 and R2 (Table 1) designed to introduce compatible restriction sites (*Eco*RI and *Pst*1) for cloning into the pMAL-c2X expression vector. The PCR amplification was carried out in a TaKaRa thermal cycler (TaKaRa Korea Biomedical Inc, Korea) using a 50 μL reaction mixture composed of 5 U of Ex*Taq* polymerase (TaKaRa), 5 μL of 10x Ex*Taq* buffer, 8 μL of 2.5 mM dNTPs, 80 ng of

Table 1
Primers used in this study.

ID	Sequence, 5′–3′	Purpose
F1	GGAGGTGCAGGTGATGTTATG	RpCytB qRT-PCR, forward primer
R1	CCACTGAATCTTTTCCTGCTTTC	RpCytB qRT-PCR, reverse primer
F2	GAGAGAGAATTCATGTGTGGAG GTGCAGGT	Cloning, introduction of *Eco*RI, forward primer
R2	GAGAGACTGCAGCTAGAAATATT CAACAGCATCTTCTACTTTC	Cloning, introduction of *Pst*I, reverse primer
F3	CTCCCTTGAGAAGAGCTACGA	β-actin qRT-PCR, forward primer
R3	GATACCAGCAGATTCCATACCC	β-actin qRT-PCR, reverse primer

template, and 20 pmol of each primer. The thermal cycling conditions were: initial incubation at 94 °C for 3 min, followed by 35 cycles of 94 °C for 30 s, 55 °C for 30 s and 72 °C for 30 s, and a final extension at 72 °C for 2 min. The amplified products and pMAL-c2X vector were digested with the appropriate restriction enzymes, and confirmed by agarose gel electrophoresis. Appropriate sized fragments were excised and purified using the Accuprep Gel Purification Kit (Bioneer Co., Daejeon, Korea) and ligated into the linearized pMAL-c2X by incubation with Mighty Mix (TaKaRa) overnight at 4 °C. Subsequently, the pMAL-c2X/RpCytB construct was transformed into *Escherichia coli* (*E. coli*) DH5α cells. The size of the construct and sequence of the ORF were confirmed by electrophoresis and sequencing, respectively.

2.4. Overexpression and purification of rRpCytB fusion protein

The recombinant fusion construct, rRpCytB-MBP, was transformed into BL21 (DE3) competent cells and grown on agar plates. A single colony was selected, inoculated in 5 mL of Luria–Bertani (LB) broth supplemented with 100 μg/mL of ampicillin, and grown overnight at 37 °C. Then, the inoculum was transferred into 500 mL of LB broth containing ampicillin, and grown at 37 °C in a shaking incubator (200 rpm). When the optical density (OD) reached 0.6 (at 600 nm), isopropy-β-thiogalactopyranoside (IPTG) was added at a final concentration of 0.5 mM and the mixture was incubated for 8 h at 20 °C to induce recombinant protein expression. Thereafter, cells were cooled on ice for 30 min and harvested by centrifugation (3500 × *g* for 30 min at 4 °C). The pellet was re-suspended in column buffer (20 mM Tris–HCL, pH 7.4 and 200 mM NaCl) and stored at −20 °C for overnight. After thawing, the cells were ruptured by cold sonication in the presence of lysozyme (1 mg/mL), and the resultant slurry was separated by centrifugation (9000 × *g* for 30 min at 4 °C). The supernatant was defined as crude extract and the recombinant protein was purified using a pMAL protein purification technique [26]. Following the purification, the concentration of the purified fusion product was determined using the Bradford method [27], with bovine serum albumin as the reference. Protein purity was analyzed by 12% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), under reduced conditions.

2.5. Cysteine protease inhibitory activity assay of rRpCytB

Recombinant protein activity was assayed by the previously published procedure with some modifications [19]. Briefly, papain (P3375; Sigma-Aldrich, St. Louis, MO, USA) was dissolved in potassium phosphate buffer (PPB; pH 7.4) to give a final concentration of 1 mg/mL. Subsequently, different volumes of the rRpCytB and papain were mixed together to generate the following rRpCytB/papain concentration ratios in 125 μL final volumes: 0/1 (control), 1/16, 1/8, 1/4, 1/2, 1/1, 2/1. The mixtures were then incubated at 25 °C for 10 min, after which 125 μL of azo-casein (0.5% in PPB) was added to the mixture and was allowed to react by incubating at 37 °C for 30 min. The reaction was terminated by adding an equal volume of 10% trichloroacetic acid. Same procedure was followed simultaneously for the MBP as the control. Finally, the mixture was centrifuged and the supernatant was subjected to measure the OD_{440}. The relative activity was calculated as: $100 \times [1 - (OD_{440}$ of test sample/OD_{440} of control)] [19]. All the samples were tested in three replicates at each point.

2.6. Experimental animals and tissue collection

Healthy clams (average size: 35 ± 5 mm) were collected from the east costal region of Jeju Island (Korea). Clams were maintained at 21 ± 1 °C in 250 L tanks with aerated, sand-filtered seawater having salinity of 32 ± 1‰. Animals were acclimatized to laboratory conditions for one week prior to the experiment. Hemolymph (1–2 mL/clam) was collected from the posterior adductor muscle sinus of healthy, unchallenged clams using a 26G syringe; samples were immediately centrifuged (3000 × *g* at 4 °C for 10 min) to harvest the hemocytes. Tissues from adductor muscle, mantel, siphon, gill, and foot were collected from three animals, snap-frozen in liquid nitrogen, and stored at −80 °C until use for RNA extraction.

2.7. Immune-challenge experiment

To determine the immune-related responses of RpCytB expression, mRNA levels were measured in clams following the challenge with *V. tapetis* (a Gram-negative bacterial pathogen), LPS (an endotoxin of Gram-negative bacterial cell walls), or poly I:C (a dsRNA viral mimic). As described previously, 100 μL of *V. tapetis* (1.9×10^6 cells/mL in 0.9% saline), LPS (100 μg/clam in 0.9% saline), or poly I:C (100 μg/clam in 0.9% saline) was injected directly into adductor muscle [28]. A group of un-injected clams served as negative controls. Gills and hemocytes tissues were collected at 3, 6, 12, 24 and 48 h post-injection (p.i.) from control and challenged groups, including five clams from each group. Total RNA was extracted from various tissues of at least five clams and used for cDNA synthesis.

2.8. Total RNA extraction and cDNA synthesis

Total RNA was isolated from the collected tissues (Section 2.6 and 2.7) using the Tri Reagent (Sigma–Aldrich). The concentration was quantified by measuring the OD_{260} in a UV-spectrophotometer (Bio-Rad, Hercules, CA, USA). Purified RNA samples were diluted up to 1 μg/μL and pooled to perform multi-tissue cDNA synthesis. The PrimeScript™ First-Strand cDNA Synthesis Kit (TaKaRa) was used to synthesize the first-stranded cDNA from 2.5 μg of pooled RNA. The resultant cDNA was then diluted 40-fold (total 800 μL) before storage at −20 °C.

2.9. RpCytB mRNA expression analysis by quantitative real time-polymerase chain reaction (qRT-PCR)

qRT-PCR was used to determine the expression levels of RpCytB in collected tissues (Section 2.6) and immune-challenged hemocytes and gills. Total RNA was extracted at different time points following immune induction, and the first-stranded cDNA synthesis was carried out as described in Section 2.8. qRT-PCR was carried out as described previously [29], using the gene-specific primer pair (F1 and R1) along with β-actin primers (F3 and R3) (Table 1). The expression level of adductor muscle mRNA was considered as the basal level, by which expression in all other tissues was compared. To determine the relative expressions, the observed expression after challenge was first normalized to saline-injected controls and then compared with expression in the unchallenged group.

2.10. In situ hybridization

Both healthy and *V. tapetis* infected (48 h p.i.) clams were dissected and gill tissues were subsequently fixed in 4% paraformaldehyde in PBS at 4 °C for overnight. Thereafter, tissues were mounted in 1.5% agarose, 2.5% sucrose solution followed by immersion in 30% sucrose at 4 °C for overnight. Then the cryomold embedded tissues floating on methyl butane were frozen in liquid nitrogen followed by sectioning at 10–15 μm micro-sections and air dried for 3 h. Afterward, micro-sections were subjected to dehydration in 100% methanol for 10 min at room temperature, followed by rehydration with series of MeOH in 1X PBS (75%, 50%, and 25%) for 10 min at room temperature. Then the tissue sections were

```
-85    GGTGAAAGTGTAAAGGTTTCGTCTTGCTTTGAACAGTTTGTGATAATTTCCGTTTCTA
-27  CTGTAGATTCATCTTAATATGGCTCAAATGTGTGGAGGTGCAGGTGATGTTATGCCAGCT
  1                                M  C  G  G  A  G  D  V  M  P  A
 34  GATGAGGAAGTGAAAGGTTATTGCAATGAGGTGAAAGCTGACATACTGAAGAAAGCAGGA
 12   D  E  E  V  K  G  Y  C  N  E  V  K  A  D  I  L  K  K  A  G
 94  AAAGATTCAGTGGAAATATTTGAACCTGTCCACTACAGAAGCCAGCTTGTAGCTGGAGTC
 32   K  D  S  V  E  I  F  E  P  V  H  Y  R  S  Q  L  V  A  G  V
154  AATTACTTTGTAAAGATCAGAATTGGCTCTGGTGGAGAATGTTTACATGCTAGAATTTTC
 52   N  Y  F  V  K  I  R  I  G  S  G  G  E  C  L  H  A  R  I  F
214  AAGGGTCTTCCTCATACAGGAGGGAACTTAGAAGTGAGCAGCGTACAAACAAATAAGAAA
 72   K  G  L  P  H  T  G  G  N  L  E  V  S  S  V  Q  T  N  K  K
274  GTAGAAGATGCTGTTGAATATTTCTAGTTGTTGAAAGGCAGAGTTAGCAGGGACACTTGA
 92   V  E  D  A  V  E  Y  F  *
334  TTGCACAACAGTGGTCAATAATTAGCCAAGCTTGATTTAAGGGCTTTATTTTACCTTTAA
394  ATATTATAACAGTAAATACTCAACTTAAAAA
```

Fig. 1. Nucleotide and deduced amino acid sequences of RpCytB. The start codon (ATG) and stop codon (TAG) are in bold and underlined. The polyadenylation signal sequence (AGTAAA) is indicated by bolded letters. The QXVXG motif is boxed and the PH motif is double underlined. Characteristic and conserved glycine (G) residue is in italics. Cysteine protease inhibitory signature is bolded.

washed 3 times with PBS for 3 min each and treated with proteinase K (1 μg/mL) for 5 min at room temperature. A second fixation was done with 4% paraformaldehyde in PBS for 5 min followed by 3 more washes with PBS as described above. RNA probes were prepared using digoxigenin (DIG)-RNA labeling and detection kit from Boehringer Mannheim Biochemicals, following the manufacture's protocol. Hybridization and visualizing of RNA probes were carried out as described by Schulte-Merker et al. [30].

2.11. Statistical analysis

To determine the statistical significance ($P < 0.05$) between the experimental and control groups, a two-tailed, paired *t*-test was carried out with SPSS version 16.0 software (Chicago, IL, USA). All the data are presented as mean ± standard deviation.

3. Results and discussion

3.1. Sequence characterization and comparison

The cDNA of RpCytB was determined to consist of 509 nucleotides, comprising a 297 bp open reading frame (ORF) that encodes 99 amino acids, an 85 bp 5′-untranslated region (UTR), and a 127 bp 3′-UTR. The 3′-UTR contains a polyadenylation signal (^{404}AGTAAA^{409}) and a single RNA instability motif (^{368}ATTTA^{372}) (Fig. 1). Moreover, the predicted molecular mass was 11 kDa, which showed a consistency with the average molecular mass of family 1 cystatins [14], and the theoretical isoelectric point was 5.9. The cDNA sequence and amino acid sequence was deposited in the NCBI GenBank database under the accession number JQ972711.

Cystatins are generally comprised of three conserved motifs, which are important for forming a wedge-shaped structure and are involved in obstruction of the cysteine protease active site [11]. Specifically, these motifs are an N-terminal glycine residue, a glutamine–valine–glycine (Q–X–V–X–G) loop, and a C-terminal proline–tryptophan (PW) motif that forms a hairpin loop in the tertiary structure. However, in cystatin B, the tryptophan in the PW motif is replaced by a histidine residue to form a PH motif. As shown in Figs. 1 and 2, all three motifs were identified in RpCytB, highlighting the characteristic structural features of a typical cystatin B. Moreover, the cysteine protease inhibitory signature was detected in the RpCytB amino acid sequence.

Pairwise sequence alignment with known cystatin B orthologs revealed that RpCytB has relatively low percentage identity and similarity. The highest identity and similarity were with cystatin B from Pacific oyster *Crassostrea gigas*, and the lowest identity and

```
                          #                                    ♦♦♦♦♦
Gallus gallus             --MLCGGVSKARPATSETQQIADEVKPQLEEKEGK-TFDVFTAVEFKTQLVAGTNYFIKV 57
Taeniopygia guttata       --MLCGGASAARPATEETQRIAEQVKAQLEEKEGK-TFDVFTAVEFKTQVVAGTNYFIKV 57
Rattus norvegicus         --MMCGAPSATMPATTETQEIADKVKSQLEEKANQ-KFDVFKAISFRRQVVAGTNFFIKV 57
Homo sapiens              --MMCGAPSATQPATAETQHIADQVRSQLEEKENK-KFPVFKAVSFKSQVVAGTNYFIKV 57
Ruditapes philippinarum   ---MCGGAGDVMPADEEVKGYCNEVKADILKKAGKDSVEIFEPVHYRSQLVAGVNYFVKI 57
Crassostrea gigas         --MMCGGTSDAKPADAEIQAIINEVKSAVEDKVGK-KLDTYTPVSYKTQVVAGTNYFVKV 57
Danio rerio               MPLVCGGTSEAKDANEEVQKICDQMKADAEEKAGR-KFDVFTAKSFKTQLVAGTNYFVKV 59
Anoplopoma fimbria        MAMMCGGAGPVQDATKEIQELCDSVKVHAEKKAGK-TYDVFIAKSYMSQTVAGTNYFIKV 59
Perinereis aibuhitensis   --MMSGGLGPTLPADPETQVICDEVKNQLEGEVGR-NFAEYNAILFRSQVVAGTVLFIKV 57
Theromyzon tessulatum     -MPLCGGTSDVKQADGKTQEIVDKVKHHVESKINK-SFKEFKAIQFSSQVVAGTNYFVKV 58
                             :.*. . .  *  : :   :.::    : .: .   : .  :  * ***.  *:*:
                                        ▼▼
Gallus gallus             HVG-NDEFMHLRVFRSLPHEDKPLSLHGYQSSKTKHDELTYF--- 98
Taeniopygia guttata       HVG-NEEFLHLRVFKSLPHENEQLSLHSYQGSKTKHDELAYF--- 98
Rattus norvegicus         DVG-EEKCVHLRVFEPLPHENKPLTLSSYQTDKEKHDELTYF--- 98
Homo sapiens              HVG-DEDFVHLRVFQSLPHENKPLTLSNYQTNKAKHDELTYF--- 98
Ruditapes philippinarum   RIGSGGECLHARIFKGLPHTGGNLEVSSVQTNKKVEDAVEYF--- 99
Crassostrea gigas         NAG--DEHLHLRIFAPLPHTGNPKELADLQRGKTAEDDLSYF--- 97
Danio rerio               HVG-DEDYVHLRVYKTLPHAGETLQLTSIQTAKAHHDAIEYF--- 100
Anoplopoma fimbria        HVG-GDDHVHLRVWKKLSCHGGELELSNMQQSKTQHDPVEYFS-- 101
Perinereis aibuhitensis   HVG-HADYIHIRIFIPLPGPG-KTSLGGYQLHKTKDDPIKYFNNN 100
Theromyzon tessulatum     HVG-EEEYLHLRVFAPLPSSGGQPELHGIEYPKTKHDELKYFNPN 102
                            *   . :* *::  *.  .    : . :  *  .* : **
```

Fig. 2. Multiple sequence alignment of vertebrate and invertebrate cystatin B proteins. Sequence alignments were generated by the ClustalW algorithm. Conserved residues are highlighted in gray. Characteristic and conserved glycine (G) residues are denoted by the "#" symbol. Characteristic and conserved residues of the QXVXG motif and the PH motif are indicated by the "◆" and "▼" symbols, respectively.

Table 2
Similarity and identity of RpCytB with cystatin B orthologs of other species.

Species	Common name	Accession no.	Identity, %	Similarity, %
Homo sapiens	Human	NP_000091	36.0	58.0
Gallus gallus	Chicken	NP_001185577	37.0	59.0
Rattus norvegicus	Rat	NP_036970	37.0	62.0
Crassostrea gigas	Pacific oyster	ADI33157	47.0	63.0
Anoplopoma fimbria	Sable fish	ACQ58093	44.7	62.1
Theromyzon tessulatum	Duck leech	AAN28679	37.5	58.0
Perinereis aibuhitensis	Clam worm	ACL12062	34.0	54.4

similarity were with that from clam worm *Perinereis aibuhitensis* (Table 2). Furthermore, the multiple sequence alignment further confirmed that the N-terminal glycine (G) residue, the Q–X–V–X–G motif (here QLVAG), and the C-terminal PH motif were well conserved in RpCytB (Fig. 2). However, the PH motif in cystatin B from *Anoplopoma fimbria* was completely replaced by serine (S) and cysteine (C) residues, and in *P. aibuhitensis* and *T. tessulatum* was replaced by glycine (G) and serine (S) residues.

3.2. Phylogenetic analysis of RpCytB

The evolutionary tree generated by phylogenetic analysis of RpCytB revealed four main clusters that corresponded to the four different families of cystatins. RpCytB clustered with family 1, as expected, and was placed in the sub-cluster with cystatin B from mollusks, further affirming its orthologous nature toward cystatin B variants (Fig. 3). Moreover, this close evolutionary relationship provides evidence to support the notion that RpCytB originated from a common invertebrate ancestor.

3.3. Overexpression and purification of RpCytB

In order to characterize the biochemical properties of RpCytB, the coding sequence was cloned into the pMAL-c2X expression vector and expressed as a fusion protein with MBP (Fig. 4). SDS-PAGE resolved a single purified protein band, confirming successful purification of the fusion protein and indicating a molecular mass of ~55 kDa. The size of the band was consistent with the predicted molecular mass of RpCytB (11 kDa), since the molecular mass of MBP is ~42.5 kDa.

3.4. Papain inhibitory activity of rRpCytB

To investigate the inhibitory activity of rRpCytB on cysteine proteases, its ability to suppress the hydrolysis activity of papain against azo-casein was measured. The percent inhibitory activity was plotted against the recombinant protein:papain concentration ratio (Fig. 5). rRpCytB exhibited a dose-dependent inhibition of

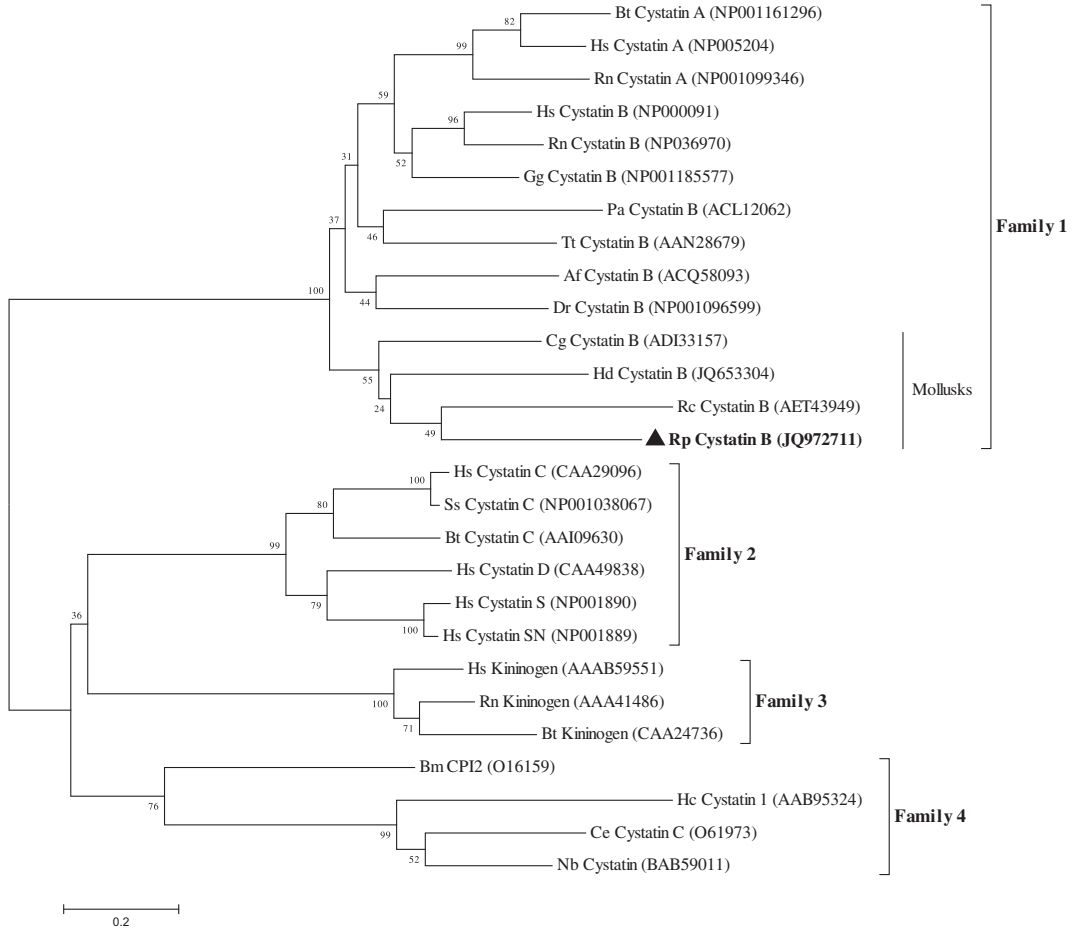

Fig. 3. Phylogenetic tree for cystatin superfamily proteins. The tree was constructed by ClustalW alignment of deduced amino acid sequences of cystatin variants from different animals, and estimated by the Neighbor-Joining method in MEGA ver. 5.0. Bootstrapping values are shown for the corresponding lineages. The cystatins belonging to the four known families clustered appropriately. Abbreviations: Bt, *Bos taurus*; Hs, *Homo sapiens*; Rn, *Rattus norvegicus*; Gg, *Gallus gallus*; Pa, *Perinereis aibuhitensis*; Tt, *Theromyzon tessulatum*; Af, *Anoplopoma fimbria*; Dr, *Danio rerio*; Cg, *Crassostrea gigas*; Hd, *Haliotis discus discus*; Rc, *Reishia clavigera*; Rp, *Ruditapes philippinarum*; Ss, *Sus scrofa*; Bm, *Brugia malayi*; Hc, *Haemonchus contortus*; Ce, *Caenorhabditis elegans*; Nb, *Nippostrongylus brasiliensis*.

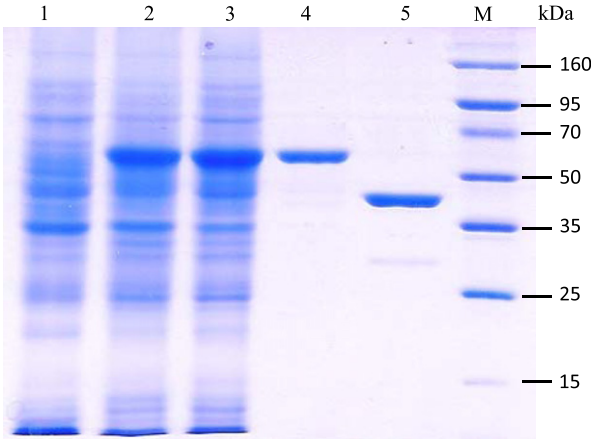

Fig. 4. SDS-PAGE analysis of overexpressed and purified recombinant RpCytB fusion protein. Lanes: 1, total cellular extract from *E. coli* BL21 (DE3) prior to IPTG induction; 2, crude extract of cell lysate after IPTG induction; 3, supernatant of IPTG-induced cell lysate; 4, purified recombinant fusion protein (rRpCytB-MBP); 5, purified recombinant MBP; M, low molecular weight protein marker (Enzynomics).

papain activity. The 1:1 concentration ratio exhibited the highest inhibitory activity (82%) against papain. Further increases in the rRpCytB concentration did not significantly increase the inhibitory activity ($P < 0.05$), as evidenced by a plateau shape on the graph and suggesting that the 1:1 ratio might provide the optimal conditions for the reaction elevation of its activity at provided conditions. These results agree with several previously reported characterizations of cystatins from invertebrates. For example, when the concentration ratio of cystatin from jellyfish *Cyanea capillata* to papain reached 1:1, the activity of papain was completely inhibited [31]. A parallel study from disk abalone *H. discus discus* also observed almost similar level of inhibition of papain activity toward the azo-casein [22]. Furthermore, in Chinese mitten crab *E. sinensis*, Horseshoe crab *Tachypleus tridentatus*, and a tick species *Haemaphysalis longicornis*, the respective cystatins produced detectable inhibition of papain activity [13,32,33]. Since the papain-like C1 family cysteine proteases are crucial for survival of microorganisms, especially bacteria and viruses [34], it is reasonable that host systems have evolved antimicrobial defenses to inhibit such pathogen-derived proteinases. Thus, the detected papain inhibitory activity of rRpCytB may reflect a potential role in Manila clam antimicrobial defense.

As a control, the papain inhibition experiment was conducted with the MBP fusion protein alone, in place of the rRpCytB-MBP recombinant fusion protein product. The results demonstrated a negligible papain inhibitory activity, compared to the results obtained with the recombinant fusion protein (Fig. 5). Therefore, the MBP protein was considered as an inert component of the rRpCytB fusion protein.

3.5. Tissue-specific transcriptional profile of RpCytB

RpCytB mRNA expression in various tissues collected from healthy clams was detected by qRT-PCR, standardized against β-actin, and normalized to the expression level in adductor muscle (least expression). A ubiquitous RpCytB transcriptional profile was observed, supporting the fact that family 1 cystatins have a broad distribution range in cells and tissues [35,36]. However, the strongest expression was detected in hemocytes (~170-fold higher than adductor muscle), and a moderately high expression level was identified in gill (Fig. 6). All of the other tissues showed relatively low RpCytB transcriptional profiles, compared to adductor muscle. Similar results for tissue specific expression from disk abalone also detected as the highest expression in hemocytes, followed by gills [22]. The pattern of differential expression among the different tissues may indicate its functional roles in the Manila clam system. The highest expressing tissues, hemolymph and gills, represent important components of the innate immune responses in invertebrates, possibly indicating a role for RpCytB in host defense. Evolutionary simple organisms, like mollusks are solely depending on the innate immunity to combat infectious pathogens. In this regards, hemolymph particularly hemocytes play an indispensable role in cellular responses of first line innate immunity in mollusks which can kill microbes by phagocytosis and cytotoxic reactions [37]. Similarly, gill tissue in mollusks is a key respiratory organ whose surface is exposed to the external environment and frequently in contact with pathogenic microbes, and it was believed to be involved with the mollusks immune system [38]. Therefore, we suggest that RpCytB in clam might play a critical role associated with immune regulation in these two tissues.

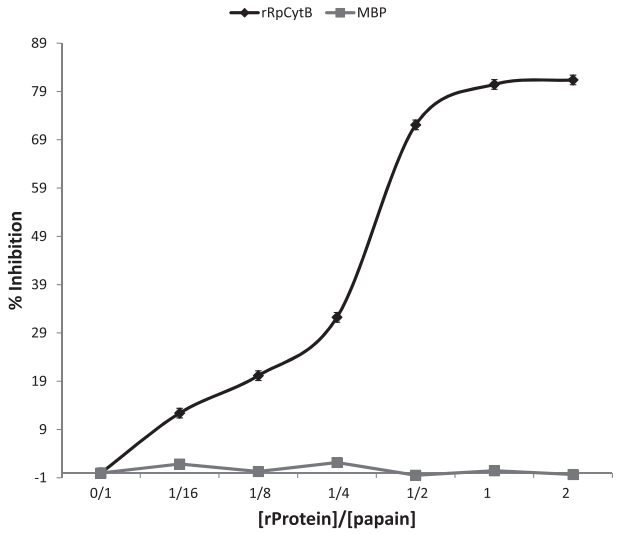

Fig. 5. Papain inhibitory activity of rRpCytB recombinant fusion protein. Azo-casine was used as the substrate, and the experiment was conducted with different ratios of protein:papain. Maltose binding protein (MBP) was used as a control protein. Error bars represent the standard deviation ($n = 3$).

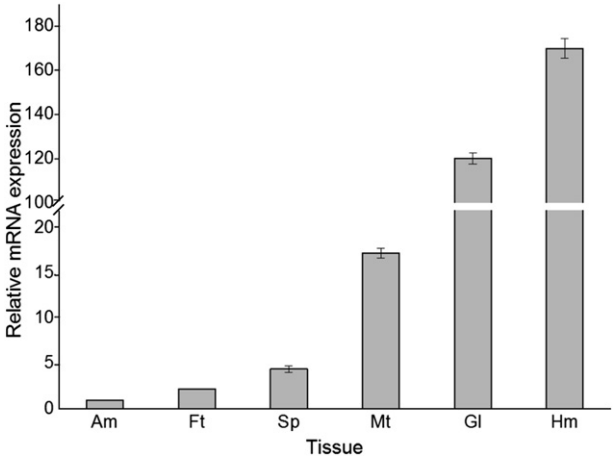

Fig. 6. Tissue-specific expression profile of RpCytB mRNA determined by qRT-PCR. Abbreviations: Hm, hemocytes; Am, adductor muscle; Mt, mantle; Sp, siphon; Gl, gill; Ft, foot. Error bars represent the standard deviation ($n = 3$).

Although gill and hemolymph showed relatively high transcript levels compared to the expression level of heart in family 1 cystatin from Chinese mitten crab, tissue specific expression profile exhibited relatively lower mRNA expression level in every tissue, where highest expression was reported in muscle tissue (~20-fold compared to heart) [13]. The cystatin B from duck leech *T. tessulatum* was shown to be expressed only in large coelomic cells, where it contributes to host immune defense [20]. The L-cystatin from horseshoe crab *T. tridentatus* was found to be primarily expressed in hemolymph and heart, with the mRNA level being negligible in hepatopancreas, intestine, stomach, and muscle [33]. Studies of vertebrate lineages have revealed distinctive tissue distribution patterns, as well. Mouse cystatin B was shown to be widely expressed in heart, liver, kidney, spleen, testis, brain, lung, and skeletal muscle, and to be involved in cerebellar apoptosis [39]. Furthermore, the cystatin B transcription profile of *Scophthalmus maximus* turbot showed constitutive distribution in every tissue examined, but with more pronounced expression in muscle, brain, and heart [19]. Hence, the distribution pattern of the cystatin B orthologs showed species specific behavior.

3.6. Immune-regulated transcriptional profile of RpCytB

Cathepsins are well-characterized cysteine proteases involved in innate immune responses, with principal signaling activities through the toll-like receptors (TLRs) [40]. As cystatins, including cystatin B, are known to play an inhibitory role on cysteine proteases, such as cathepsin B, H and L [11], a considerable down-regulation of cystatin gene expression is expected to occur during an immune response to an invasive pathogen. Furthermore, in human placental macrophages, down-regulated expression of cystatin B was reported to facilitate the inhibition of human immunodeficiency virus replication via a STAT-1 interaction [21]. On the other hand, cystatin expression can be elevated under some other pathological conditions, reflecting the stimulatory role of these molecules to induce production of different cytokines, such as IL-10 and TNFα, in mammalian macrophages [41,42].

Significant transcriptional modulation of RpCytB was observed in gill tissue and hemocytes, both of which are involved in immunity. Systemic challenge with a well-characterized bacterial cell wall component, LPS, remarkably influenced RpCytB mRNA expression in hemocytes, with significant down-regulation being detected at every time point after the endotoxin injection ($P < 0.05$; Fig. 7A). Moreover, except at 3 h p.i., gill tissue also exhibited a generally under-expressed transcriptional profile upon LPS exposure. These observations strongly indicate a potential active involvement of cathepsin-like proteinases in Manila clam tissues in response to Gram-negative bacterial infections, whereby cystatins would be made less abundant in the affected tissues.

Upon infection with the live bacterial pathogen, *V. tapetis*, a complex transcriptional response of RpCytB was observed. In hemocytes, significant elevations of RpCytB were detected at 3 h and 48 h p.i. ($P < 0.05$), whereas in gill, RpCytB mRNA level was induced only at 48 h p.i. (4-fold higher relative mRNA expression level than the basal expression; Fig. 7B). In contrast, RpCytB expression was significantly down-regulated at 6 h, 12 h, and 24 h p.i. in gill, but only under-expressed at 12 h p.i. ($P < 0.05$) in hemocytes. The overall transcriptional modulation of cystatin B that was observed in Manila clam upon *V. tapetis* stimulation suggests a complicated regulatory mechanism for RpCytB in host pathology. Similar results were reported after the bacterial infections in disk abalone cystatin B, which showed a late phase induction in gills compared to the hemocytes and significant down-regulation toward the late phase of the experiment in hemocytes [22]. Correspondingly, the expression of invertebrate linage cystatins from duck leech and Chinese mitten crab were reported to be augmented at the transcriptional level upon bacterial induction [13,20]. In duck leech [20], the transcript level of its cystatin B homolog was elevated in their large coelomocytes at 2 h p.i. upon *E. coli* stimulation, at 3 h and 24 h p.i. upon *Micrococcus luteus* stimulation, and at 2 h p.i. upon induction with a mixture of the two bacteria. Likewise, in Chinese mitten crab, the mRNA level of a family 1 cystatin member was enhanced in hemolymph upon *Listonella anguillarum* infection at 24 h p.i., while *Pichia pastoris* infection enhanced transcription at 3 h (peak), 6 h, 12 h, and 48 h p.i. However, when the Chinese mitten crab was challenged with *L. anguillarum*, family 1 cystatin expression was down-regulated at 3 h p.i, which agreed with down-regulated transcriptional profile observed for RpCytB in response to *V. tapetis* [13].

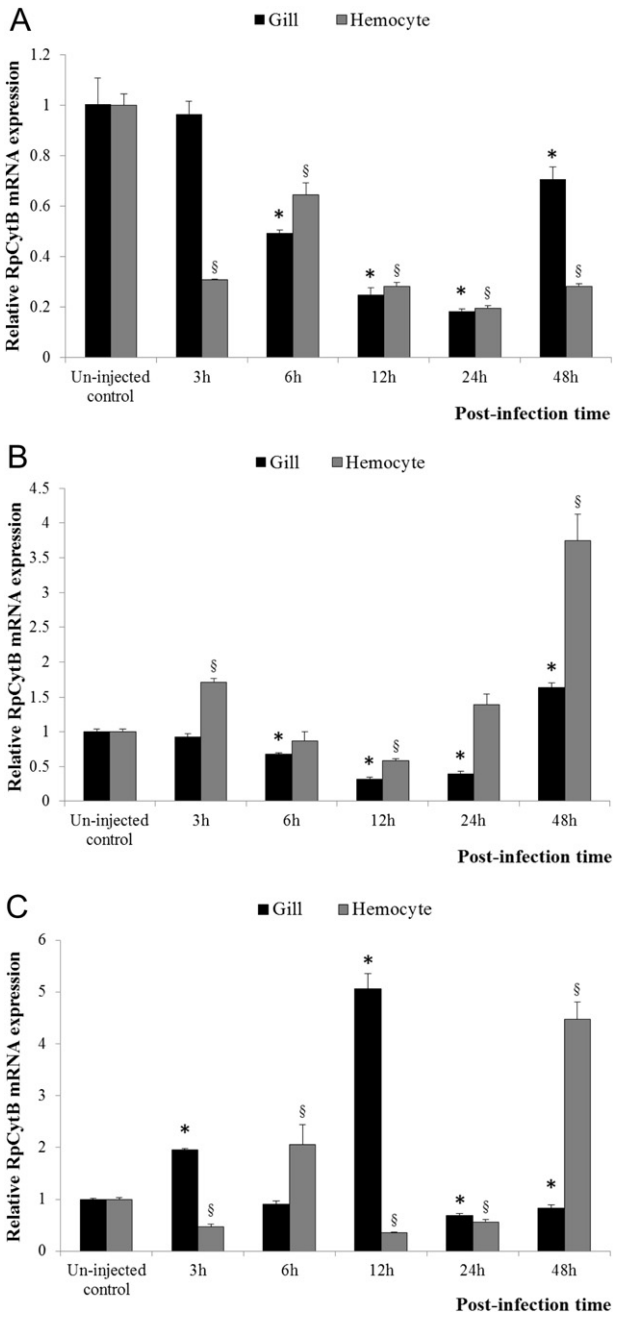

Fig. 7. Expression profile of RpCytB in gill and hemocytes upon (A) LPS induction, (B) *V. tapetis* induction, and (C) poly I:C induction. Error bars represent the standard deviation ($n = 3$); ($P < 0.05$).

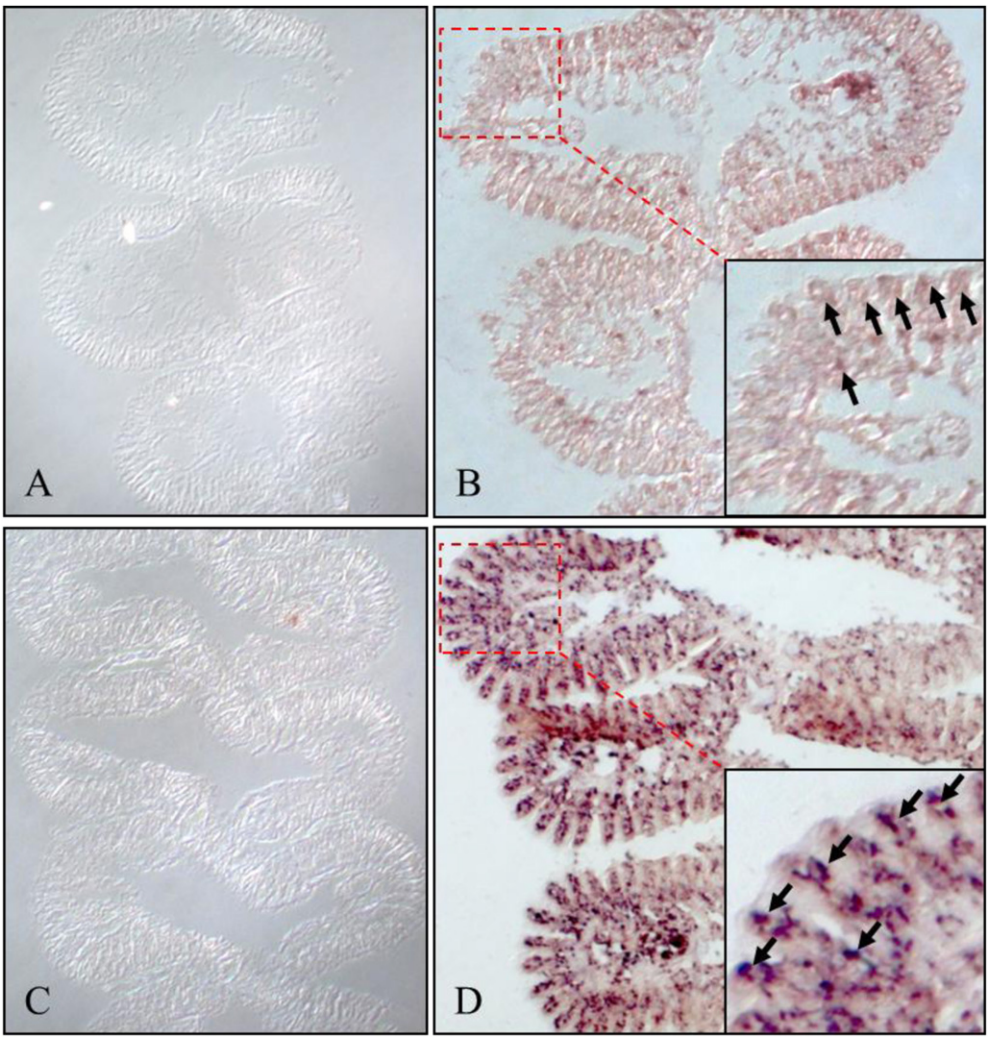

Fig. 8. RpCytB mRNA localization in the gill tissue by *in situ* hybridization in healthy (A and B) and *V. tapetis* infected (C and D) clams. Tissue sections were hybridized with RpCytB sense (A and C) and antisense (B and D) digoxigenin-labeled RNA probes. Arrow heads indicate the RpCytB expressing cells.

In our study, transcriptional response of an invertebrate cystatin B upon challenge with the viral dsRNA mimic immune stimulant, poly I:C, was investigated. Similar to the complex regulatory mechanism related to *V. tapetis* induction, poly I:C-stimulated clams showed a general down-regulated transcriptional profile of RpCytB, along with two early and late positively regulated responses (gill: at 3 h p.i. and 12 h p.i., hemocytes: at 6 h and 48 h p.i.; Fig. 7C). The early response elicited in gill, compared to that observed in hemocytes, may be attributed to the frequent exposure of gill tissues to various pathogens, through its continuous interaction with the outer environment [41,42].

Moreover, in order to understand the RpCytB expression in gill tissues independent to the hemocytes in gill hemolymph, *in situ* hybridization test was performed in both healthy and *V. tapetis* infected clams. RpCytB expressions were precisely detected in both healthy and bacteria infected clams gill filament cells, with a remarkable induction of the mRNA expression after 48 h of bacterial infection (Fig. 8) affirming the qRT-PCR data.

According to the overall transcriptional behavior upon immune challenges and its structural and functional properties, RpCytB appears to play roles in complicated immune modulatory events under pathological conditions of Manila clam. This cystatin B homolog may act directly on pathogen-derived or host immune-related factors, possibly leading to inhibition of cysteine proteinases.

Acknowledgments

This research was supported by National Fisheries Research and Development Institute (RP-2012-AQ-056).

References

[1] Neurath H, Walsh KA. Role of proteolytic enzymes in biological regulation (a review). Proc Natl Acad Sci U S A 1976;73:3825–32.
[2] Potempa J, Golonka E, Filipek R, Shaw LN. Fighting an enemy within: cytoplasmic inhibitors of bacterial cysteine proteases. Mol Microbiol 2005;57:605–10.
[3] Chapman HA, Riese RJ, Shi GP. Emerging roles for cysteine proteases in human biology. Annu Rev Physiol 1997;59:63–88.
[4] Grzelakowska-Sztabert B. Molecular mechanisms of apoptosis induced by activation of membrane receptors from the TNF-R superfamily. Postepy Biochem 1998;44:8–21.
[5] Potempa J, Pike RN. Bacterial peptidases. In: Russell W, Herwald H, editors. Concepts in bacterial virulence. Contrib Microbiol Basel. Karger; 2005. p. 132–80.
[6] American College of Obstetricians and Gynecologists. ACOG educational and practice bulletins: list of titles-August 2005. Obstet Gynecol 2005;106:427–8.
[7] Rudenskaya GN, Pupov DV. Cysteine proteinases of microorganisms and viruses. Biochemistry (Mosc) 2008;73:1–13.
[8] Berdowska I, Siewinski M. The role of cysteine cathepsins and their inhibitors in physiological and neoplastic processes. Postepy Biochem 2000;46:73–84.
[9] Rzychon M, Chmiel D, Stec-Niemczyk J. Modes of inhibition of cysteine proteases. Acta Biochim Pol 2004;51:861–73.

[10] Bode W, Huber R. Structural basis of the endoproteinase-protein inhibitor interaction. Biochim biophys acta 2000;1477:241–52.
[11] Turk V, Bode W. The cystatins: protein inhibitors of cysteine proteinases. FEBS Lett 1991;285:213–9.
[12] Khaznadji E, Collins P, Dalton JP, Bigot Y, Moire N. A new multi-domain member of the cystatin superfamily expressed by *Fasciola hepatica*. Int J Parasitol 2005;35:1115–25.
[13] Li F, Gai X, Wang L, Song L, Zhang H, Qiu L, et al. Identification and characterization of a cystatin gene from Chinese mitten crab *Eriocheir sinensis*. Fish Shellfish Immunol 2010;29:521–9.
[14] Ochieng J, Chaudhuri G. Cystatin superfamily. J Health Care Poor Underserved 2010;21:51–70.
[15] Rahman KM, Khan HM, Haq JA. Incidence of cutaneous diphtheria in Bangladesh. Bangladesh Med Res Counc Bull 1983;9:49–53.
[16] Abrahamson M. Cystatins. Meth Enzymol 1994;244:685–700.
[17] Turk B, Turk V, Turk D. Structural and functional aspects of papain-like cysteine proteinases and their protein inhibitors. Biol Chem 1997;378:141–50.
[18] Suzuki T, Hashimoto S, Toyoda N, Nagai S, Yamazaki N, Dong HY, et al. Comprehensive gene expression profile of LPS-stimulated human monocytes by SAGE. Blood 2000;96:2584–91.
[19] Xiao PP, Hu YH, Sun L. *Scophthalmus maximus* cystatin B enhances head kidney macrophage-mediated bacterial killing. Dev Comp Immunol 2010;34:1237–41.
[20] Lefebvre C, Cocquerelle C, Vandenbulcke F, Hot D, Huot L, Lemoine Y, et al. Transcriptomic analysis in the leech *Theromyzon tessulatum*: involvement of cystatin B in innate immunity. Biochem J 2004;380:617–25.
[21] Luciano-Montalvo C, Melendez LM. Cystatin B associates with signal transducer and activator of transcription 1 in monocyte-derived and placental macrophages. Placenta 2009;30:464–7.
[22] Premachandra HKA, Wan Q, Elvitigala DA, De Zoysa M, Choi CY, Whang I, et al. Genomic characterization and expression profiles upon bacterial infection of a novel cystatin B homologue from disk abalone (*Haliotis discus discus*). Dev Comp Immunol 2012;38:495–504.
[23] Lee Y, Whang I, Umasuthan N, De Zoysa M, Oh C, Kang DH, et al. Characterization of a novel molluscan MyD88 family protein from manila clam, *Ruditapes philippinarum*. Fish Shellfish Immunol 2011;31:887–93.
[24] Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994;22:4673–80.
[25] Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 2011;28:2731–9.
[26] Maina CV, Riggs PD, Grandea 3rd AG, Slatko BE, Moran LS, Tagliamonte JA, et al. An *Escherichia coli* vector to express and purify foreign proteins by fusion to and separation from maltose-binding protein. Gene 1988;74:365–73.
[27] Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 1976;72:248–54.
[28] Umasuthan N, Revathy KS, Lee Y, Whang I, Choi CY, Lee J. A novel molluscan sigma-like glutathione S-transferase from Manila clam, *Ruditapes philippinarum*: cloning, characterization and transcriptional profiling. Comp Biochem Physiol C Toxicol Pharmacol 2012;155:539–50.
[29] Umasuthan N, Whang I, Lee Y, Lee S, Kim Y, Kim H, et al. Heparin cofactor II (RbHCII) from rock bream (*Oplegnathus fasciatus*): molecular characterization, cloning and expression analysis. Fish Shellfish Immunol 2011;30:194–208.
[30] Schulte-Merker S, Ho RK, Herrmann BG, Nüsslein-Volhard C. The protein product of the zebrafish homologue of the mouse T gene is expressed in nuclei of the germ ring and the notochord of the early embryo. Development 1992;116:1021–32.
[31] Yang Y, Cun S, Peng L, Xie X, Wei J, Yang W, et al. cDNA cloning, identification and characterization of a novel cystatin from the tentacle of *Cyanea capillata*. Biochimie 2003;85:1033–9.
[32] Zhou J, Liao M, Ueda M, Gong H, Xuan X, Fujisaki K. Characterization of an intracellular cystatin homolog from the tick *Haemaphysalis longicornis*. Vet Parasitol 2009;160:180–3.
[33] Agarwala KL, Kawabata S, Hirata M, Miyagi M, Tsunasawa S, Iwanaga S. A cysteine protease inhibitor stored in the large granules of horseshoe crab hemocytes: purification, characterization, cDNA cloning and tissue localization. J Biochem 1996;119:85–94.
[34] Katunuma N, Shiota H, Le QT. Medical significance of cysteine protease inhibitors in mammalian secretory fluids. J Med Invest 2003;50:154–61.
[35] Cimerman N, Mesko Brguljan P, Krasovec M, Suskovic S, Kos J. Serum concentration and circadian profiles of cathepsins B, H and L, and their inhibitors, stefins A and B, in asthma. Clin Chim Acta 2001;310:113–22.
[36] Kos J, Lah TT. Cysteine proteinases and their endogenous inhibitors: target proteins for prognosis, diagnosis and therapy in cancer (review). Oncol Rep 1998;5:1349–61.
[37] Pruzzo C, Gallo G, Canesi L. Persistence of vibrios in marine bivalves: the role of interactions with haemolymph components. Environ Microbiol 2005;7:761–72.
[38] Chakraborty S, Ray M, Ray S. Toxicity of sodium arsenite in the gill of an economically important mollusc of India. Fish Shellfish Immunol 2010;29:136–48.
[39] Pennacchio LA, Bouley DM, Higgins KM, Scott MP, Noebels JL, Myers RM. Progressive ataxia, myoclonic epilepsy and cerebellar apoptosis in cystatin B-deficient mice. Nat Genet 1998;20:251–8.
[40] Creasy BM, McCoy KL. Cytokines regulate cysteine cathepsins during TLR responses. Cell Immunol 2011;267:56–66.
[41] Vray B, Hartmann S, Hoebeke J. Immunomodulatory properties of cystatins. Cell Mol Life Sci 2002;59:1503–12.
[42] Verdot L, Lalmanach G, Vercruysse V, Hoebeke J, Gauthier F, Vray B. Chicken cystatin stimulates nitric oxide release from interferon-gamma-activated mouse peritoneal macrophages via cytokine synthesis. Eur J Biochem 1999;266:1111–7.