



## Molecular cloning and mRNA expression of the peptidoglycan recognition protein gene HcPGRP1 and its isoform HcPGRP1a from the freshwater mussel *Hyriopsis cumingi*

Ye Tao<sup>1,2</sup>, Zi-yan Yang<sup>1,3</sup>, Xin Zhang<sup>4</sup> and Hong-juan Wu<sup>1</sup>

<sup>1</sup>School of Environmental Science and Engineering, Huazhong University of Science and Technology, Wuhan, China.

<sup>2</sup>Department of Environmental Protection of Henan Province, Zhengzhou, China.

<sup>3</sup>School of Environment and Municipal Engineering, North China University of Water Resources and Electric Power, Zhengzhou, China.

<sup>4</sup>Henan Rural Energy & Environment Agency, Henan Provincial Department of Agriculture, Zhengzhou, China.

### Abstract

Peptidoglycan recognition proteins (PGRPs) are innate immune molecules that have been structurally conserved throughout evolution in invertebrates and vertebrates. In this study, peptidoglycan recognition protein HcPGRP1 and its isoform HcPGRP1a were identified in the freshwater mussel *Hyriopsis cumingi*. The full-length cDNAs of HcPGRP1 (973 bp) and HcPGRP1a (537 bp) encoded polypeptides with 218 and 151 amino acids, respectively. Sequence analysis showed that HcPGRP1 had one C-terminal PGRP domain that was conserved throughout evolution. Phylogenetic analysis showed that HcPGRP1 clustered closely with EsPGRP4 of *Euprymna scolopes*. Real-time PCR showed that the mRNA transcripts of HcPGRP1 and HcPGRP1a were constitutively expressed in various tissues, with the highest level in hepatopancreas. Stimulation with lipopolysaccharide (LPS) and peptidoglycan (PGN) significantly up-regulated HcPGRP1 mRNA expression in hepatopancreas and foot, but not in gill, whereas HcPGRP1a expression was significantly up-regulated in all three tissues. Our results indicate that HcPGRP1 is both a constitutive and inducible protein that may be involved in immune responses (recognition and defense) against invaders.

**Keywords:** expression pattern, gene cloning, *Hyriopsis cumingi*, innate immunity, peptidoglycan recognition protein.

Received: November 21, 2013; Accepted: April 10, 2014.

### Introduction

The innate immune system is the first line of defense against invading microorganisms in vertebrates and the only line of defense in invertebrates and plants (Medzhitov and Janeway, 2002; Hoffmann, 2003). This system recognizes microorganisms through a series of pattern recognition receptors (PRRs) that have been highly conserved throughout evolution (Hoffmann *et al.*, 1999; Janeway and Medzhitov, 2002). Peptidoglycan recognition proteins (PGRPs) are a type of PRR and play a very important role in various activities associated with the innate immune response of invertebrates and vertebrates, including pathogen recognition, degradation of the amidase activity of peptidoglycan (PGN), induction of phagocytosis, activation of Toll

or immune deficiency (IMD) signal pathways and the prophenoloxidase cascade (Yoshida *et al.*, 1996; Leulier *et al.*, 2003; Mellroth *et al.*, 2003; Bischoff *et al.*, 2004; Chang *et al.*, 2004; Takehana *et al.*, 2004).

The first PGRP, a 19-kDa protein, was discovered in the hemolymph and cuticle of a silkworm (*Bombyx mori*) (Yoshida *et al.*, 1996). Subsequently, highly diversified PGRP homologs were identified in insects and mammals (Werner *et al.*, 2000; Liu *et al.*, 2001; Christophides *et al.*, 2002), as well as in mollusks (Su *et al.*, 2007). Currently, more than 20 PGRP genes have been identified from mollusks. Based on the protein primary structure and length, mollusk PGRPs have been divided into two classes: short PGRPs (PGRP-S) and long PGRPs (PGRP-L). Short PGRPs are small extracellular proteins (19-20 kDa) that usually contain a signal peptide sequence and a PGRP domain. In contrast, long PGRPs have long transcripts and are either intracellular or membrane-spanning proteins (30-90 kDa). In addition to the PGRP domain in the C-ter-

minal, PGRP-Ls also contain a large N-terminal domain of unknown function (Guan and Mariuzza, 2007). Long PGRPs often have multiple splice forms, such as BgPGRP-LA from *Biomphalaria glabrata* that has three isoforms (-LA, LA1 and -LA2) (Zhang *et al.*, 2007).

All PGRP genes contain a highly conserved PGRP domain that can recognize and bind PGN; the structure of this region is similar to that of bacteriophage T7 lysozyme that can hydrolyze the bond between N-acetylmuramic acid and the peptide of bacterial PGN (Kim *et al.*, 2003; Dziarski, 2004). Many PGRPs have amidase activity and are characterized by a Zn<sup>2+</sup> binding site. The active site consists of two histidines, one tyrosine and one cysteine residue, with the cysteine residue being crucial for amidase activity. In *Drosophila*, the mutation of cysteine (Cys80) in PGRP-SA to tyrosine (Tyr80) abolishes Toll activation by Gram-positive bacteria (Michel *et al.*, 2001). In addition, the mutation of cysteine (Cys419) to alanine (Ala419) in human PGLYRP-2 results in a loss of amidase activity, indicating that Cys419 is important for this activity (Wang *et al.*, 2003).

In view of the significant roles of PGRPs in innate immune defense, these proteins have also been studied in mollusks. AiPGRP from *Argopecten irradians* is induced primarily by PGN and may be involved in the scallop immune response to infection by Gram-positive bacteria (Ni *et al.*, 2007). In addition to binding PGN, some PGRPs may also bind other pathogen-associated molecular pattern (PAMP) molecules. For example, CfPGRP-S1 mRNA expression was up-regulated after stimulation by PGN and LPS, indicating a possible role in the immune defense against Gram-positive and Gram-negative bacteria (Yang *et al.*, 2012).

*Hyriopsis cumingii*, an economically important freshwater pearl mussel species used for monitoring aquatic environments, is widely cultured in southern China. In this study, HcPGRP1 and its isoform HcPGRP1a were cloned from *H. cumingii*. In addition, the tissue distribution and inducibility of transcripts in various tissues challenged with LPS or PGN was investigated. The findings reported here should be useful in understanding the immune significance of PGRP in mollusks.

## Materials and Methods

### Experimental animals

Pearl mussels (*H. cumingii*) ~2 years old (shell length: 55 ± 5 mm) obtained from an aquaculture center in Yueshan Village, Ezhou city, China, were maintained in aerated fresh lake water at 20 ± 1 °C before the experiments. For the experiments, mussels were transferred to 10 L of lake water containing *Microcystis aeruginosa* 905 (purchased from the Institute of Hydrobiology, Chinese Academy of Sciences), with a microcystin-LR (MC-LR) concentration of 40–60 µg/L each day. After five days, the hepatopancreas was excised from the mussels and homoge-

nized in Trizol reagent to obtain the RNA used to construct a suppression subtractive hybridization (SSH) cDNA library and clone the full-length cDNA of PGRP.

### SSH cDNA library construction and EST analysis

The differential expression of hepatopancreas genes in mussels exposed to MC-LR was assessed with an SSH cDNA library constructed using a PCR-Select cDNA subtraction kit (Clontech, USA). About 400 positive clones were sequenced and 98 high quality sequences were identified; these sequences corresponded to genes involved in apoptosis, signal transduction, cytoskeletal remodel, innate immunity, material and energy metabolism, translation and transcription (Yang *et al.*, 2012). BLAST analysis of all the EST sequences revealed that one EST (EST No.j-7; 406 bp) was homologous to PGRP 4 in *E. scolopes* (score 150 bits, expect 2e-42, identities 63%, positives 80%, GenBank No. AAY27976); this EST was used to design primers for cloning the HcPGRP1 gene.

### Cloning of the full-length cDNA of HcPGRP1

Total RNA was extracted from the hepatopancreas of MC-LR-challenged mussels using Trizol reagent according to the manufacturer's instructions (Invitrogen, USA). cDNA was then synthesized and amplified using a RevertAid first strand cDNA synthesis kit (MBI Fermentas, Germany) according to the manufacturer's instructions. Two groups of specific primers (5' GSP1, 5' GSP2 and 3' GSP1, 3' GSP2) were designed (Table 1) from the partial PGRP sequence identified in the SSH cDNA library and were used along with adaptor primers (UPM) for 5' and 3' rapid amplification of cDNA ends (RACE). The RACE reactions were done using a SMART RACE cDNA amplification kit (Clontech) according to the manufacturer's instructions. Two rounds of PCR amplification were done using the following conditions: 94 °C for 3 min, followed by seven cycles of 94 °C for 30 s, 66 °C (first round) or 68 °C (second round) for 30 s, 72 °C for 90 s and 28 cycles of 94 °C for 30 s, 63 °C (first round) or 65 °C (second round) for 30 s, 72 °C for 90 s, and a final extension at 72 °C for 10 min. The PCR products were separated by electrophoresis on agarose gels and the desired fragments were recovered with an E.Z.N.A.<sup>®</sup> Gel Extraction kit (Omega), cloned into the pMD-18 T vector (TaKaRa, Japan) and transformed into competent *Escherichia coli* TOP10. The positive recombinants were identified by blue-white color selection in ampicillin-containing LB plates and then sequenced at Shanghai Sangon Biological Engineering Technology & Services Co. Ltd. (China).

### Cloning of HcPGRP1 isoform cDNA

To obtain a possible splicing variant of HcPGRP1 in *H. cumingii*, the primer pair isoPGRP1F/isoPGRP1R (Table 1) was designed based on the full cDNA sequence of HcPGRP1. A mixture of cDNA templates from different

**Table 1** - Primer sequences used in this study.

Primer	Sequence (5' to 3')	Feature
Oligo(dT)-adaptor	GGCCACGCGTCGACTAGTACT17	Universal primers mix
UPM-long	TGGTATCAACGCAGAGTCTAATACGACTCACTATAGGGC	Adaptor primer
UPM-short	TTATGTCTGCGATAGCGACCTCTCA	
5'-GSP1	GTGTGATTTTGCCCTTCAGGATGC	5'-RACE
5'-GSP2	GCAGTTAGAGCAGCAGAGTTAGGCAGT	
3'-GSP1	ACCAATAAACTGCCTAACTCTGCTGC	3'-RACE
3'-GSP2	TGAAGGGCAAAATCACACCAGACTAC	
Iso- PGRP1F	AAGGTCGCAGAACCCTGCTCGT	Isoform cloning
Iso- PGRP1R	CTGTATGTGACAACGTTGTGTAC	
RT-PGRP1F	CACAGTGGAGATGCAGAAAATAC	Realtime PCR
RT-PGRP1R	CAGCAGAGTTAGGCAGTTTATTG	
RT-PGRP1aF	TTCTTGTGAAGGTTGGGATGACAT	
RT-PGRP1aR	TTTGCCTTCAGGATGCCACAGT	
ActinF	GAGACAACCTACAACAGCATCAT	
actinR	GAAGCCAAAATGGGACCACCGAT	

F – forward, R – reverse.

organs of grass carp was used to clone the splicing isoform of HcPGRP1a.

### Sequence analysis

Sequence similarity with other known proteins was analyzed using the BLAST program. The alignment of multiple sequences was done with Clustal W (version 1.83). The theoretical signal peptide was predicted using the SignalP 3.0 Server. The transmembrane domain was predicted with the TMHMM program. Possible N-glycosylation sites were identified using NetNGlyc 1.0 and a Pfam protein family search was done with the Pfam HMM search program. Physico-chemical parameters were analyzed using the ProtParam program. A phylogenetic tree was constructed based on the deduced amino acid sequences using the neighbor-joining (NJ) method with 1000 bootstrap replicates and MEGA v4.0 software.

### Gene expression pattern of HcPGRP1 and HcPGRP1a using real-time quantitative PCR

To investigate the tissue distribution of HcPGRP1 and HcPGRP1a transcripts, total RNA was extracted from various tissues (hemocytes, hepatopancreas, gonad, kidney, intestinal, gill, mantle, adductor muscle and foot) of healthy mussels. Total RNA was also extracted from hepatopancreas, gonad, kidney, gill and foot tissues of *H. cumingii* challenged with LPS or PGN for various intervals.

One hundred and twenty mussels were used for the stimulation experiment. The mussels were randomly divided into three groups (n = 40 each). The sample groups were injected with 100 µL of LPS (1 mg/mL in PBS) from *E. coli* 055:B5 (Sigma-Aldrich) or 100 µL of PGN

(1 mg/mL in PBS) from *Staphylococcus aureus* (Sigma-Aldrich), while the control group was injected with 100 µL of PBS. Tissue samples were obtained 3, 6, 12, 18, 24 and 36 h after challenge with PBS, LPS or PGN. Total RNA was extracted using Trizol reagent (Invitrogen) as described by the manufacturer. After treatment with RNase-free DNase, 2 µg of RNA from each sample was reverse-transcribed with a RevertAid™ First Strand cDNA synthesis kit (Fermentas) at 42 °C using an oligo (dT) 18 primer.

HcPGRP1 expression was examined by quantitative real-time PCR (qRT-PCR) with a SYBR Green supermix (Bio-Rad, USA) in a CFX96 C1000 thermal cycler (Bio-Rad, USA). Two pairs of gene-specific primers (RT-PGRP1F and RT-PGRP1R; RT-PGRP1aF and RT-PGRP1aR) (Table 1) designed by Primer premier 5.0 for HcPGRP1 and HcPGRP1a were used to amplify PCR products of 209 bp and 223 bp, respectively. Two β-actin primers (actin-F and actin-R; Table 1) were used to amplify a 218-bp gene fragment as an internal control for qRT-PCR. The amplification efficiency of all primers was determined using standard curves and primers with an efficiency of 0.95-1.0 chosen for this study. The qRT-PCR amplifications were done in triplicate along with the internal control gene in 96-well plates. All analyses were based on the CT values of the PCR products.

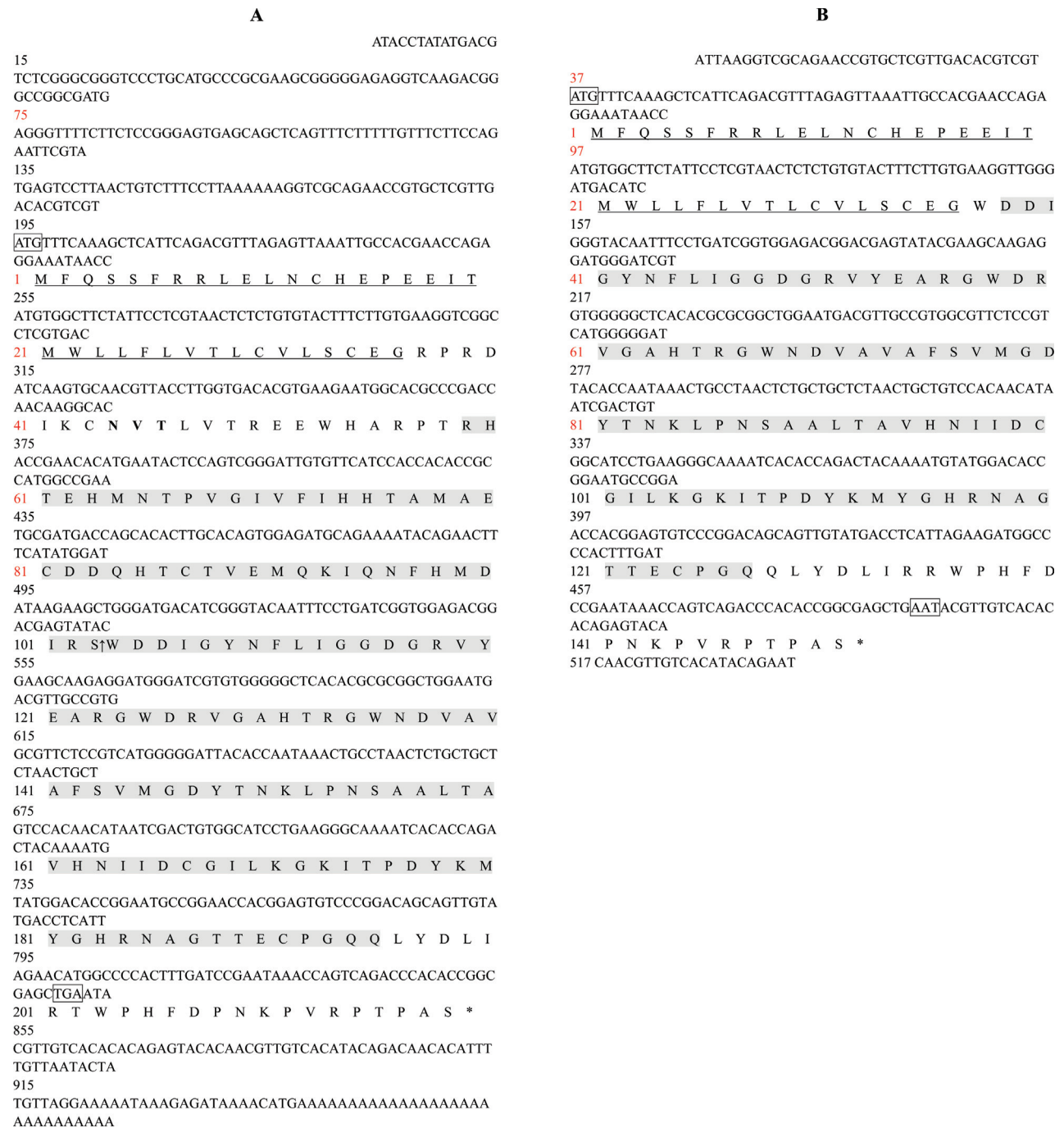
The  $2^{-\Delta\Delta CT}$  method was used to analyze the level of HcPGRP1 and HcPGRP1a expression (Schmittgen, 2001). The relative expression of the target gene was normalized to the expression of β-actin and then expressed as a fold change relative to the corresponding control group. All data showing the relative mRNA expression are the mean of at least three independent experiments. Statistical compari-

sons were done using Students *t*-test, with  $p < 0.05$  indicating significance.

## Results

### Molecular features of HcPGRP1 and HcPGRP1a cDNA

The cDNA sequence of *H. cumingii* PGRP was confirmed by blastn analysis on NCBI and was designated as



**Figure 1** - Complementary DNA and predicted amino acid sequences of HcPGRP1 (A) and its isoform HcPGRP1a (B) from *H. cumingii*. The signal peptide (1-36) is underlined, the PGRP domain profile is shaded in dark gray, the start codon and stop codons are boxed and the N-glycosylation site is in bold. The arrow indicates the cleavage site and the asterisk indicates termination.

signal peptide with a putative cleavage site located after position 36 (MFQ-EG) (Figure 1A). An N-glycosylation site was predicted at position 44 (NVT) and an amidase\_II domain (also known as a PGRP domain or T phage lysozyme homology domain) was predicted at the C-terminal end of HcPGRP1.

Based on the full-length cDNA sequence, a pair of primers was designed to obtain an alternative splicing isoform of HcPGRP1, which was named HcPGRP1a (GenBank acc. no. KF479261) (Figure 1B). The complete cDNA sequence of HcPGRP1a was 537 bp, with an open reading frame (ORF) of 458 bp encoding a polypeptide of 151 amino acids. Alignment of the cDNA sequences of HcPGRP1 and HcPGRP1a showed that alternative splicing in HcPGRP1 is mainly generated by deletion of the sequence between R37-S103 (RPDIDKCNVTLVTRREEWHARPTRHTEHMNTPVGI VFIHHTAMAECDDQHTCTVEMQKIQNFHMDIRS) while it retained the entire signal peptide sequence and part of the PGRP domain. Table 2 shows the most important biochemical parameters of HcPGRP1 and HcPGRP1a.

### Sequence analysis and phylogenetic characterization of HcPGRP1

BLAST analysis showed that the sequence identity of HcPGRP1 with other homologs ranged from 32% to 56%. The highest percentages of identity and similarity (56% and 70%, respectively) were with the EsPGRP4 homolog (GenBank accession no. AAY27976) from the Pacific oyster *E. scolopes*. Based on the multiple sequence alignments between HcPGRP1 and other animal PGRPs (Figure 2), four catalytic residues (His74, Tyr109, His183 and Cys191) for the T7 lysozyme Zn<sup>2+</sup> binding were found to be highly conserved in HcPGRP1.

A phylogenetic tree was constructed using the neighbor-joining (NJ) method with 1000 bootstraps based on the multiple alignments (Figure 3). Based on the overall amino acid sequences of 25 PGRPs from other aquatic mollusks, the phylogenetic tree showed that HcPGRP1 and EsPGRP4 from *E. scolopes* were grouped together in a minimum cluster and then clustered with the other PGRP subfamily from *E. scolopes*, suggesting there was a close evolutionary relationship between these proteins.

### Tissue distribution of HcPGRP1 and HcPGRP1a mRNA

SYBR Green real-time PCR analysis was used to study the distribution of mRNA expression of the two PGRPs, with the housekeeping gene  $\beta$ -actin used as an internal standard. HcPGRP1 and HcPGRP1a mRNA was constitutively expressed at different levels in a variety of tissues (Figure 4). The level of HcPGRP1 mRNA expression was highest in hepatopancreas followed by intestine, and lowest in hemocytes. The expression of HcPGRPS1a was also highest in hepatopancreas followed by intestine, but was lowest in adductor muscle.

### Expression pattern of HcPGRP1 and HcPGRP1a in mussels challenged with LPS or PGN

Time-course experiments were used to investigate the temporal variation in HcPGRPS1 and HcPGRP1a transcription *in vivo* for up to 36 h after a challenge with LPS or PGN. Based on the initial expression patterns determined by RT-PCR, three *H. cumingi* tissues (hepatopancreas, gill and foot) were selected to evaluate the patterns of HcPGRP1 and HcPGRP1a gene expression. The  $\beta$ -actin gene was used as an internal control.

Figure 5A-C shows that stimulation with LPS significantly up-regulated HcPGRP1 gene expression in hepatopancreas and foot, with the highest expression occurring 18 h and 3 h post-stimulation, respectively ( $p < 0.05$ ); there was no significant change in HcPGRP1 expression in gill. In contrast, HcPGRP1a expression was significantly up-regulated in hepatopancreas, gill and foot, with the greatest increase occurring 12 h, 12 h and 6 h post-stimulation, respectively ( $p < 0.05$ ). In mussels stimulated with PGN (Figure 5D-F), the expression of both HcPGRP1 and HcPGRP1a was significantly up-regulated in hepatopancreas, gill and foot. HcPGRP1 mRNA expression was highest in hepatopancreas, gill and foot at 12 h, 12 h and 24 h post-stimulation, respectively ( $p < 0.05$ ), whereas that of HcPGRP1a mRNA was highest at 12 h, 6 h and 18 h, respectively ( $p < 0.05$ ).

### Discussion

In this study, a PGRP gene (designated as HcPGRP1) and a splice variation (HcPGRP1a) were cloned from *H. cumingi*. Both of these genes and the previously cloned HcPGRPS1 (Yang *et al.*, 2013) belonged to the PGRP gene family of *H. cumingi*. HcPGRP1 had a predicted molecular

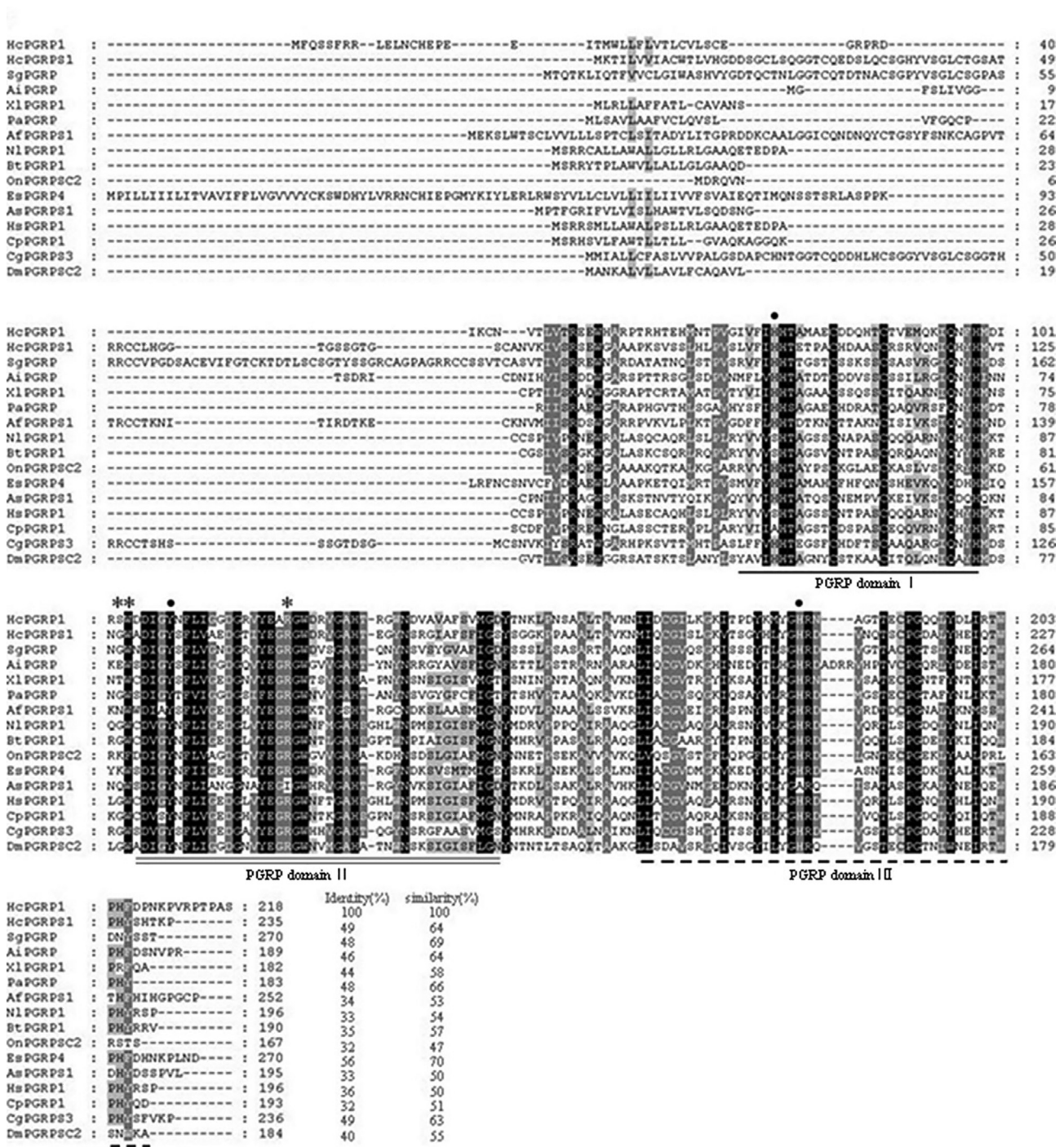
**Table 2** - The biochemical parameters of *H. cumingi* HcPGRP1 and its isoform HcPGRP1a.

Molecular name	Amino acids	Molecular mass (kDa)	Theoretical pI	Instability index	Aliphatic index	GRAVY
HcPGRP1	218	24.98	6.29	41.16	72.43	-0.416
HcPGRP1a	151	17.10	6.18	52.88	77.48	-0.302

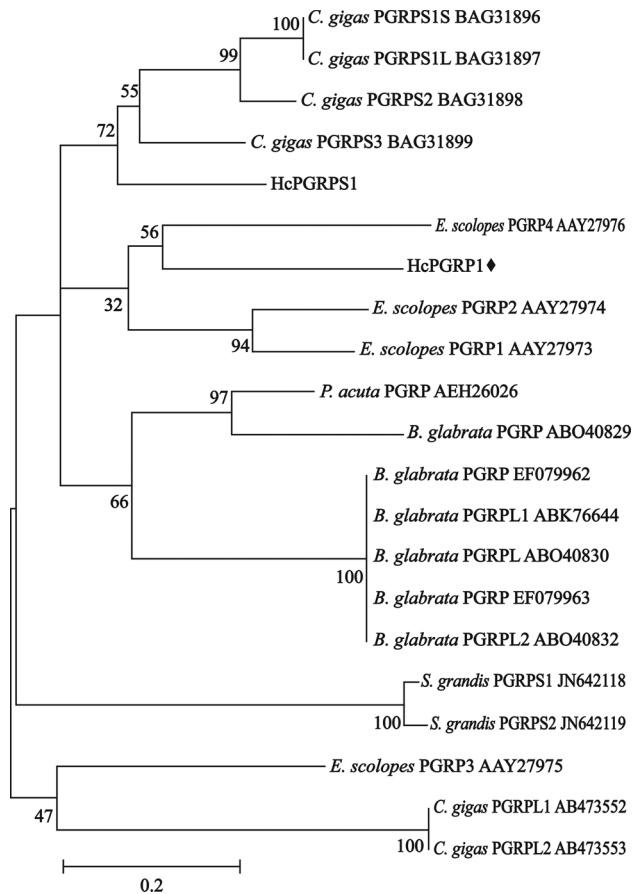
GRAVY – grand average of hydropathicity.

mass of 24.98 kDa with a signal peptide of 36 amino acids but no transmembrane domain, indicating that it was most likely to be a secretory protein. BLAST and phylogenetic

analyses showed that HcPGRP1 was homologous to EsPGRP4 from the mollusk *E. scolopes* (Goodson et al., 2005) and clustered closely with other PGRPs from *E.*



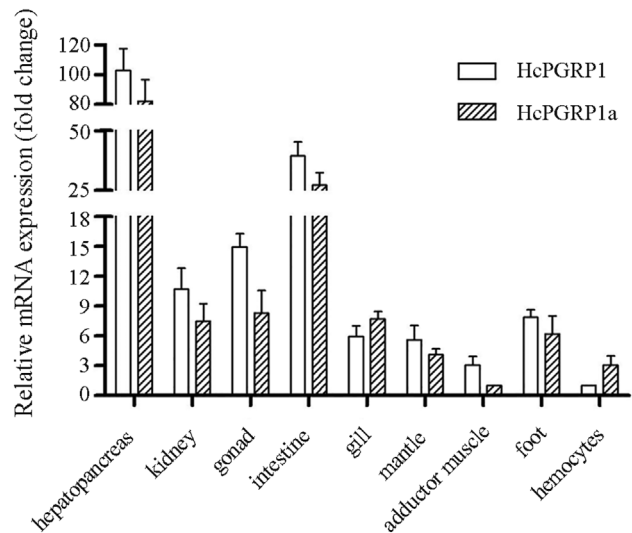
**Figure 2** - Multiple alignments of HcPGRP1 with other known PGRPs in aquatic mollusks. PGRP domains I, II and III are underlined. The dots indicate the amino acids required for T7 lysozyme Zn<sup>2+</sup>-binding and amidase activity. The Trp (W) and Arg (R) that are believed to determine preferential DAP specificity are indicated by an asterisk. The sequences used for the analysis were PGRPs from *Argopecten irradians* (AiPGRP; AAR92030), *Armigeres subalbatus* (AsPGRPS1; AEX31477), *Azumapecten farreri* (AfPGRPS1; AAY53765), *Bos taurus* (BtPGRP1; NP\_776998), *Cavia porcellus* (CpPGRP1; XP\_003464642), *Crassostrea gigas* (CgPGRPS3; BAG31899), *Drosophila melanogaster* (DmPGRPS2; CAD89179), *Euprymna scolopes* (EsPGRP4; AAY27976), *Homo sapiens* (HsPGRP1; O75594), *Hyriopsis cumingi* (HcPGRPS1; KC905181), *Nomascus leucogenys* (NlPGRP1; XP\_003277649), *Oreochromis niloticus* (OnPGRPS2; XP\_003441739), *Physella acuta* (PaPGRP; AEH26026), *Solen grandis* (SgPGRP; AEW43446) and *Xenopus laevis* (XlPGRP1; NP\_001088771).



**Figure 3** - Phylogenetic analysis of HcPGRP1 from *H. cumingii* with other known mollusks PGRPs orthologs. The tree was constructed by the neighbor-joining (NJ) algorithm using the Mega 4.0 program based on the multiple sequence alignment provided by ClustalW. Bootstrap values of 1000 replicates (%) are indicated for the branches.

*scolopes*, suggesting that there was a close evolutionary relationship among these proteins. BLAST analysis also revealed that the HcPGRP1 gene shared 50% identity with HcPGRPS1 and < 56% identity with EsPGRP4, while the phylogenetic analysis placed HcPGRP1 and HcPGRPS1 on different evolutionary branches. This result suggests that these two proteins from the same organism belong to different subfamilies.

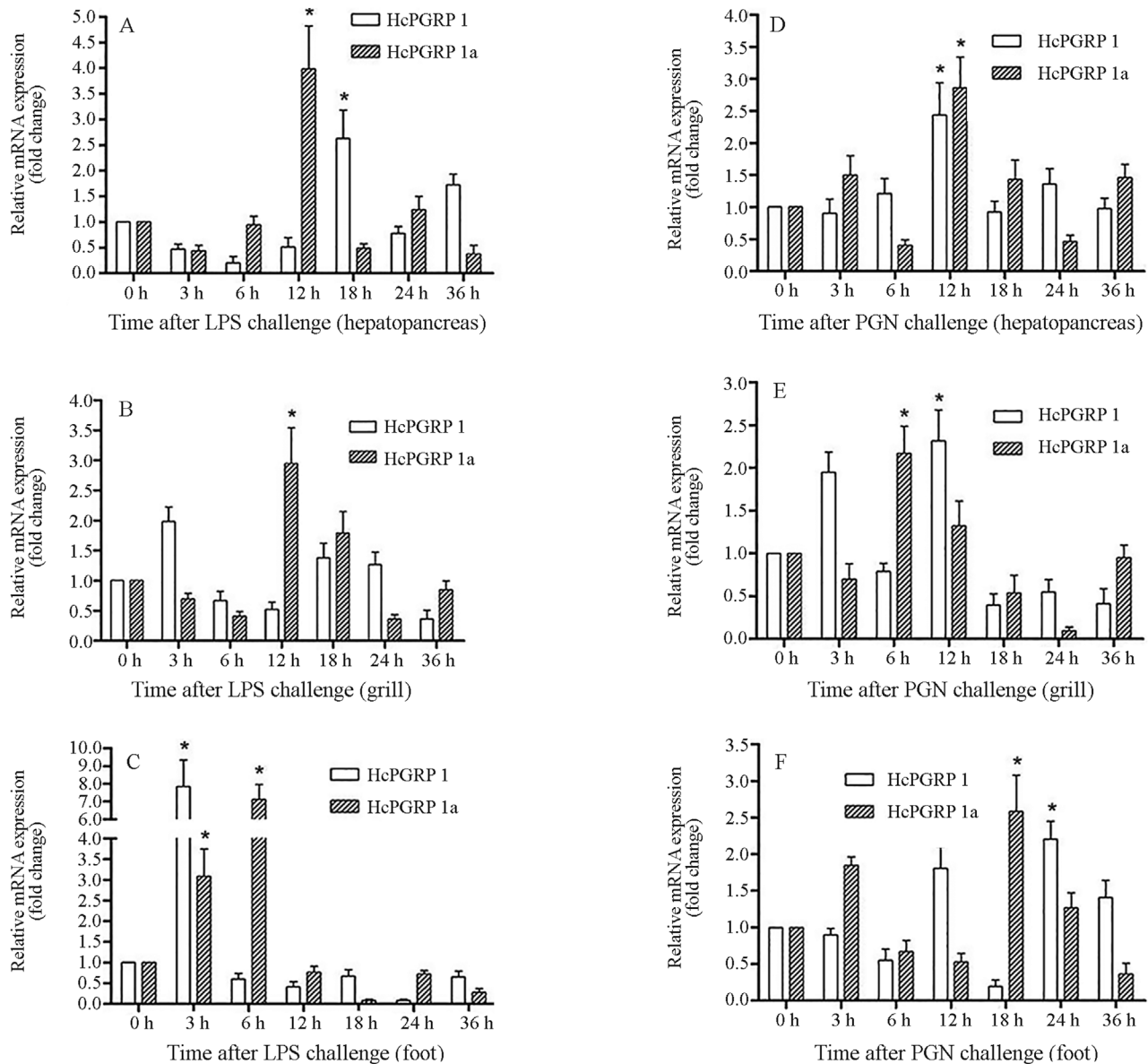
Multiple alignments with other short PGRPs showed that the four amino acids (His17, Tyr46, His122 and Cys130) that are required for the T7 lysozyme Zn<sup>2+</sup> binding and amidase activity were conserved in HcPGRP1 (His74, Tyr109, His183 and Cys191) (Cheng *et al.*, 1994); these sites have also been found in HcPGRPS1 from *H. cumingii* (Yang *et al.*, 2013), CgPGRPS3 from *Crassostrea gigas* (Itoh and Takahashi, 2008), AiPGRP from *A. irradians* (Ni *et al.*, 2007) and DmPGRPSC2 from *Drosophila* (Bischoff *et al.*, 2006). This similarity suggests that these molecules might serve as amidases involved in the elimination of PGN during the immune response against bacteria.



**Figure 4** - Distribution of HcPGRP1 and HcPGRP1a transcripts in *H. cumingii* tissues. The tissues screened by qRT-PCR included hemocytes, hepatopancreas, gonad, kidney, intestine, gill, mantle, adductor muscle and foot.  $\beta$ -Actin was used as an internal control. Vertical bars indicate the mean  $\pm$  SD (n = 3).

Mollusk PGRPs may have developed special functions during their evolution. For example, EsPGRPs 1, 2, 3 and 4 from *E. scolopes* may serve as signal transducers to trigger the Toll/NF- $\kappa$ B phosphorylation cascade. However, the replacement of Cys160 by Ser160 in EsPGRP4 results in a loss of catalytic activity and limited PGN recognition or signal transduction functions (Goodson *et al.*, 2005). Some PGRPs specifically or preferentially recognize Dap-type or Lys-type PGNs, with this specificity being determined by the three amino acids in the PGN binding groove. For example, in mollusks, rCfPGRPS1 shows high affinity for Lys-type PGN (Yang *et al.*, 2010) and HcPGRPS1 displays PGN-binding activity towards DAP-type and Lys-type PGN (Yang *et al.*, 2013), whereas human PGLYRP-1 (Gly68, Trp69 and Arg88) and *Drosophila* PGRP-LE (Gly234, Trp235 and Arg254) bind Dap-type PGN (Swaminathan *et al.*, 2000; Kumar *et al.*, 2005; Chang *et al.*, 2006). As shown here, HcPGRP1 retained only two of the four amino acids (Trp104 and Arg123) that are normally conserved, which suggested that the binding specificity of this protein was for DAP-type PGN.

Alternative splicing, a process by which multiple different functional messenger RNAs can be synthesized from a single gene, plays a key role in the expansion of proteomic and regulatory complexity (Nilsen and Graveley, 2010). Some previous reports had shown that multiple alternative splicing isoforms exist in PGRP gene families, *e.g.*, *Drosophila* has 13 PGRP genes that encode 19 proteins, *Anopheles* has seven PGRP genes that encode nine proteins (Werner *et al.*, 2000; Christophides *et al.*, 2002). In addition, *B. glabrata* BgPGRP-LA has three isoforms (-LA, LA1 and-LA2) (Zhang *et al.*, 2007). As shown here,



**Figure 5** - Temporal expression of HcPGRP1 and HcPGRP1a mRNA in tissues (hepatopancreas, foot and gill) during the time course challenge with LPS (A-C) or PGN (D-F). Statistical comparisons of the control (untreated) and challenged (treated) groups were done using one-way analysis of variance (ANOVA) and SPSS 13.0 software. Vertical bars indicate the mean  $\pm$  SD ( $n = 3$ ). \* $p < 0.05$  compared to relative mRNA expression at 0 h.

we cloned a splice variation (HcPGRP1a) of HcPGRP1 that lacked part of the PGRP domain. Alternative splicing isoforms probably interact with normal forms to produce stimulatory and inhibitory effects (Rosenstiel *et al.*, 2006; Chang *et al.*, 2011).

Mollusk PGRPs show highly variable expression in various tissues. For example, bay scallop (*A. irradians*) AiPGRP is predominantly expressed in hemocytes (Ni *et al.*, 2007), whereas Pacific oyster (*C. gigas*) CgPGRP-S1L and CgPGRP-S3 are mainly expressed in mantle and digestive diverticula, respectively (Itoh and Takahashi, 2008). In *H. cumingi*, the pattern of HcPGRP1 and HcPGRP1a expression was similar to that of *H. cumingi* HcPGRPS1 (Yang *et al.*, 2013) and *Solen grandis* SgPGRP-S1 (Wei *et al.*,

2012); both genes were detected in all of the tissues examined, with the highest expression in hepatopancreas. The selective expression of PGRPs in different tissues suggests that they have different functions in the body. The responsiveness of PGRPs after exposure to microorganisms also varies considerably among host species. Stimulation with PGN, LPS and glucan markedly up-regulates the expression of CfPGRP-S1 in hemocytes of *Chlamys farreri*, indicating that this inducible protein is involved in the immune response to invading microbes (Yang *et al.*, 2010). Similarly, SgPGRP-S1 and SgPGRP-S2 from *S. grandis* and HcPGRPS1 from *H. cumingi* were significantly induced in response to stimulation by LPS or PGN (Wei *et al.*, 2012; Yang *et al.*, 2013). As shown here, microbial ligands (LPS



and PGN) also induced the expression of HcPGRP1 and HcPGRP1a in hepatopancreas, gill and foot, indicating the importance of HcPGRP1 as innate pattern recognition receptors in immune defense. Interestingly, the increase in HcPGRP1 and HcPGRP1a expression in hepatopancreas and foot tissue in response to LPS was greater than for PGN, which suggests that LPS may be the potential ligand for HcPGRP1.

## Acknowledgments

This study was supported by the National Natural Science Foundation of China (Grant no. 31072209).

## References

- Bischoff V, Vignal C, Boneca IG, Michel T, Hoffmann JA and Royet J (2004) Function of the *Drosophila* pattern-recognition receptor PGRP-SD in the detection of Gram-positive bacteria. *Nat Immunol* 5:1175-1180.
- Bischoff V, Vignal C, Duvic B, Boneca IG, Hoffmann JA and Royet J (2006) Downregulation of the *Drosophila* immune response by peptidoglycan-recognition proteins SC1 and SC2. *PLoS Pathog* 2:e14.
- Chang CI, Pili-Floury S, Hervé M, Parquet C, Chelliah Y, Lemaitre B, Mengin-Lecreux D and Deisenhofer J (2004) A *Drosophila* pattern recognition receptor contains a peptidoglycan docking groove and unusual L,D- carboxypeptidase activity. *PLoS Biol* 2:1293-1302.
- Chang CI, Chelliah Y, Borek D, Mengin-Lecreux D and Deisenhofer J (2006) Structure of tracheal cytotoxin in complex with a heterodimeric pattern-recognition receptor. *Science* 311:1761-1764.
- Chang MX, Wang TH, Nie P, Zou J and Secombes CJ (2011) Cloning of two rainbow trout nucleotide-binding oligomerization domain containing 2 (NOD2) splice variants and functional characterization of the NOD2 effector domains. *Fish Shellfish Immunol* 30:118-127.
- Cheng X, Zhang X, Pflugrath JW and Studier EW (1994) The structure of bacteriophage T7 lysozyme, a zinc amidase and an inhibitor of T7 RNA polymerase. *Proc Natl Acad Sci USA* 91:4034-4038.
- Christophides GK, Zdobnov E, Barillas-Mury C, Birney E, Blandin S, Blass C, Brey PT, Collins FH, Danielli A, Dimopoulos G, *et al.* (2002) Immunity related genes and gene families in *Anopheles gambiae*. *Science* 298:159-165.
- Dziarski R (2004) Peptidoglycan recognition proteins (PGRPs). *Mol Immunol* 40:877-886.
- Goodson MS, Kojadinovic M, Troll JV, Scheetz TE, Casavant TL, Soares MB and McFall-Ngai MJ (2005) Identifying components of the NF-kappaB pathway in the beneficial *Euprymna scolopes-Vibrio fischeri* light organ symbiosis. *Appl Environ Microbiol* 71:6934-6946.
- Guan R and Mariuzza RA (2007) Peptidoglycan recognition proteins of the innate immune system. *Trends Microbiol* 15:127-134.
- Hoffmann JA (2003) The immune response of *Drosophila*. *Nature* 426:33-38.
- Hoffmann JA, Kafatos FC, Janeway CA and Ezekowitz RAB (1999) Phylogenetic perspectives in innate immunity. *Science* 284:1313-1318.
- Naoki I and Keisuke GT (2008) Distribution of multiple peptidoglycan recognition proteins in the tissues of Pacific oyster, *Crassostrea gigas*. *Comp Biochem Physiol B* 150:409-417.
- Janeway C and Medzhitov R (2002) Innate immune recognition. *Annu Rev Immunol* 20:197-216.
- Kim MS, Byun M and Oh BH (2003) Crystal structure of peptidoglycan recognition protein LB from *Drosophila melanogaster*. *Nat Immunol* 4:787-793.
- Kumar S, Roychowdhury A, Ember B, Wang Q, Guan RJ, Mariuzza RA and Boons GJ (2005) Selective recognition of synthetic lysine and meso-diaminopimelic acid-type peptidoglycan fragments by human peptidoglycan recognition proteins I alpha and S. *J Biol Chem* 280:37005-37012.
- Leulier F, Parquet C, Pili-Floury S, Ryu JH, Caroff M, Lee WJ, Mengin-Lecreux D and Lemaitre B (2003) The *Drosophila* immune system detects bacteria through specific peptidoglycan recognition. *Nat Immunol* 4:478-484.
- Liu C, Xu ZJ, Gupta D and Dziarski R (2001) Peptidoglycan recognition proteins: A novel family of four human innate immunity pattern recognition molecules. *J Biol Chem* 276:34686-34694.
- Livak KJ and Schmittgen TD (2001) Analysis of relative gene expression data using realtime quantitative PCR and the 2(-Delta Delta C) method. *Methods* 25:402-408.
- Medzhitov R and Janeway CA (2002) Decoding the patterns of self and nonself by the innate immune system. *Science* 296:298-300.
- Mellroth P, Karlsson J and Steiner H (2003) A scavenger function for a *Drosophila* peptidoglycan recognition protein. *J Biol Chem* 278:7059-7064.
- Michel T, Reichhart JM, Hoffmann JA and Royet J (2001) *Drosophila* toll is activated by Gram-positive bacteria through a circulating peptidoglycan recognition protein. *Nature* 414:756-759.
- Ni D, Song L, Wu L, Wu LT, Chang YQ, Yu YD, Qiu LM and Wang LL (2007) Molecular cloning and mRNA expression of peptidoglycan recognition protein (PGRP) gene in bay scallop (*Argopecten irradians*, Lamarck 1819). *Dev Comp Immunol* 31:548-558.
- Nilsen TW and Graveley BR (2010) Expansion of the eukaryotic proteome by alternative splicing. *Nature* 28:457-463.
- Rosenstiel P, Huse K, Till A, Hampe J, Hellmig S, Sina C, Billmann S, von Kampen O, Waetzig GH, Platzer M, *et al.* (2006) A short isoform of NOD2/CARD15, NOD2-S, is an endogenous inhibitor of NOD2/receptor-interacting protein kinase 2-induced signaling pathways. *Proc Natl Acad Sci USA* 103:3280-3285.
- Su J, Ni D, Song LS, Zhao JM and Qiu LM (2007) Molecular cloning and characterization of a short type peptidoglycan recognition protein (CfPGRP-S1) cDNA from Zhikong scallop *Chlamys farreri*. *Fish Shellfish Immunol* 23:646-656.
- Swaminathan CP, Brown PH, Roychowdhury A, Wang Q, Guan R, Silverman N, Werner T, Liu G, Kang D, Ekengren S, *et al.* (2000) A family of peptidoglycan recognition proteins in the fruit fly *Drosophila melanogaster*. *Proc Natl Acad Sci USA* 97:13772-13777.
- Takehana A, Yano T, Mita S, Kotani A, Oshima Y and Kurata S (2004) Peptidoglycan recognition protein (PGRP)-LE and PGRP-LC act synergistically in *Drosophila* immunity. *EMBO J* 23:4690-4700.

- Wang ZM, Li X, Cocklin RR, Wang MH, Wang M, Fukase K, Inamura S, Kusumoto S, Gupta D and Dziarski R (2003) Human peptidoglycan recognition protein-L is an N-acetylmuramoyl-L-alanine amidase. *J Biol Chem* 278:49044-49052.
- Wei XM, Yang JM, Yang DL, Xu J, Liu XQ, Yang JL, Fang JH and Qiao HJ (2012) Molecular cloning and mRNA expression of two peptidoglycan recognition protein (PGRP) genes from mollusk *Solen grandis*. *Fish Shellfish Immunol* 32:178-185.
- Werner T, Liu G, Kang D, Ekengren S, Steiner H and Hultmark D (2000) A family of peptidoglycan recognition proteins in the fruit fly *Drosophila melanogaster*. *Proc Natl Acad Sci USA* 97:13772-13777.
- Yang JL, Wang W, Wei XM, Qiu LM, Wang LL, Zhang H and Song LS (2010) Peptidoglycan recognition protein of *Chlamys farreri* (CfPGRP-S1) mediates immune defenses against bacterial infection. *Dev Comp Immunol* 34:1300-1307.
- Yang ZY, Wu HJ and Li Y (2012) Toxic effect on tissues and differentially expressed genes in hepatopancreas identified by suppression subtractive hybridization of freshwater pearl mussel (*Hyriopsis cumingii*) following microcystin-LR challenge. *Environ Toxicol* 27:393-403.
- Yang ZY, Li JH, Wu HJ, Li Y and Wang XY (2013) Molecular cloning and functional characterization of a short peptidoglycan recognition protein (HcPGRPS1) from the freshwater mussel, *Hyriopsis cumingii*. *Mol Immunol* 56:729-738.
- Yoshida H, Kinoshita K and Ashida M (1996) Purification of peptidoglycan recognition protein from hemolymph of the silkworm, *Bombyx mori*. *J Biol Chem* 271:13854-13860.
- Zhang SM, Zeng Y and Loker ES (2007) Characterization of immune genes from the schistosome host snail *Biomphalaria glabrata* that encode peptidoglycan recognition proteins and Gram-negative bacteria binding protein. *Immunogenetics* 59:883-898.

## Internet Resources

- BLAST program; <http://www.ncbi.nlm.nih.gov/BLAST>.
- Theoretical signal peptide prediction SignalP 3.0 Server, <http://www.cbs.dtu.dk/services/SignalP/>.
- Transmembrane domain prediction TMHMM program, <http://www.cbs.dtu.dk/services/TMHMM-2.0/>.
- N-glycosylation sites prediction NetNGlyc 1.0 program <http://www.cbs.dtu.dk/services/NetNGlyc/>.
- Pfam HMM search program, <http://pfam.janelia.org>.
- Physico-chemical parameter analysis ProtParam program, <http://www.expasy.ch/tools/protparam.html>.

*Associate Editor: Célia Maria de Almeida Soares*

License information: This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.