



Infection-induced molecular pattern recognition proteins in larvae of *Protaetia brevitarsis seulensis* (Coleoptera: Cetoniidae)

KYEONGRIN BANG¹, JIAE LEE¹, SEJUNG HWANG¹, YOUNGWOONG CHO^{1,2}, JANGWOO PARK³ and SAEYOULL CHO^{1,2,*}

¹ Department of Applied Biology, College of Agriculture and Life Science, Kangwon National University, Chuncheon, Kangwon-do 200-701, South Korea; e-mails: bb1980@korea.kr, zloh1211@korea.kr, tpwj321@korea.kr, 5120057@kangwon.ac.kr, saeyoullcho@kangwon.ac.kr

² Department of Interdisciplinary Program in Smart Agriculture, College of Agriculture and Life Science, Kangwon National University, Chuncheon, South Korea

³ Osang Kinsect Co., Guri, South Korea; e-mail: pjw@k-insect.com

Key words. PGRPs, infection, gene expression, immunity

Abstract. We cloned and sequenced full-length peptidoglycan recognition protein (PGRP)-like cDNAs, named *PS PGRP-SA(a)*-like, *PS PGRP-SA(b)*-like, *PS PGRP-SB1*-like and *PS PGRP-SC*-like, from *Protaetia brevitarsis seulensis*. The amino acid sequences of PS PGRPs share 32.03–47.93% homology with those of PGRP family members in insects and mammals, including humans. We identified a conserved consensus sequence for amidase activity (His; H-Tyr; Y-His; H-Thr; T-Cys; C) and residues for binding peptidoglycan (PGN), one of the major bacterial cell wall components, including Asp (D) and Phe (F) for Lys-type PGN; and Gly (G), Trp (W) and Arg (R) for DAP-type PGN. The topological structures of PS PGRP-SA(a)-like, PS PGRP-SA(b)-like and PS PGRP-SC-like proteins are structurally similar to those of *Drosophila melanogaster* PGRP-SA, which has three α -helices and six β -strands. The β -strands are located in a central region and helix α 1 on the back and peripheral α 2 and α 3 helices are on the front. The three α -helices and six β -strands are also present in PS PGRP-SB1-like, but the topological structure differs from that of typical PGRP. Significantly increased levels of *PS PGRP-SA (a)*-like and *PS PGRP-SA (b)*-like mRNA were recorded when Gram-positive bacteria or yeast cells were injected into larvae. *PS PGRP-SB1*-like mRNA levels were up-regulated by infection by all three pathogens; however, expression of *PS PGRP-SC*-like mRNA was increased 20- or 30-fold only shortly after injection with Gram-negative bacteria.

INTRODUCTION

Micro-associated molecular patterns (MAMPs) and pathogen-associated molecular patterns (PAMPs), such as peptidoglycan (PGN), lipopolysaccharide (LPS), β -glucans, lipoproteins, CpG dinucleotides and flagellin, are molecular markers recognized by the insect innate immune system. There are various cellular immune responses of insect blood cells (haemocytes) and humoral immune responses mediated by various effector molecules, including antimicrobial peptides (AMPs) and the phenol oxidase (PO) cascade is part of the insect immune system (Janeway et al., 2002; Hoffmann, 2003; Cho & Cho, 2019). Humoral immune responses involving Toll and immune deficiency (IMD) pathways are mainly activated by insect pattern recognition receptors (PRRs) (Wang et al., 2019). Therefore, humoral immune responses are activated when MAMPs are recognized by insect PRRs. Peptidoglycan recognition proteins (PGRPs), C-type lectin receptors (CLRs), fibrinogen-related proteins (FREPs), thioester-containing pro-

teins (TEPs), macrophage scavenger receptors (SRs) and β -(1 \rightarrow 3)-glucan recognition protein (β GRP) are all insect PRRs, and activate complicated signalling pathways such as Toll, prophenol oxidase cascade and IMD pathways, thereby stimulating the production of AMPs and other defence molecules in host organisms (Yu et al., 2002; Lemaitre & Hoffmann, 2007; Ragan et al., 2009; Bang et al., 2013).

Among the PRRs, insect PGRP family proteins are found in various animals from insects to mammals (Liu et al., 2021). In humans, there is a family of four PGRPs: PGRP-L (Long PGRP, 64 kDa), PGRP-Ia (Intermediate PGRP, 38 kDa), PGRP-I β (46 kDa) and PGRP-S (Small PGRP, 24 kDa), where 'S' means short, 'L' long and 'I' intermediate, based on the size of their mRNA transcript, which are similar to insect PGRPs (Chaput & Boneca, 2017; Hou et al., 2020). Various types of PGRPs are reported in different insects. The first PGRP was identified in the haemolymph of the silkworm *Bombyx mori* and has a molecular weight

* Corresponding author; e-mail: saeyoullcho@kangwon.ac.kr

of ~19 kDa (Yoshida et al., 1996; Wang et al., 2019). Thirteen PGRP genes are isolated of the 19 in *Drosophila melanogaster* (Royet et al., 2011). There are 12 PGRPs in *B. mori*, seven in *Anopheles gambiae* and four in *Apis mellifera* (Christophides et al., 2002; Evans et al., 2006; Tanaka et al., 2008). Like mammals, PGRPs in insects are also classified into short (S) and long (L) types according to their transcript length. For example, *D. melanogaster* has PGRP-SA, SB1, SB2, SC1B, SC2, SD and PGRP-LA, LB, LC, LD, LE and LF [9,13]. PGRP-S1 to S6 are short types and PGRP-L1 to L6 are long types in the silkworm (Tanaka et al., 2008; Wang et al., 2019).

Among the PGRPs in *D. melanogaster*, PGN from most Gram-positive bacteria (PGN containing lysine, Lys-type PGN) is recognized by PGRP-SA or PGRP-SD, and PGN of most Gram-negative bacteria (PGN containing diaminopimelic acid, DAP-type PGN) by the transmembrane receptor protein PGRP-LC (Liu et al., 2021). Therefore, a highly conserved PGN-binding bacterial cell wall type 2 amidase domain is present in most PGRPs (PGN-binding region; ~165 amino acid residues), which consists of three α -helices and five β -folds, four of which are parallel, and the last is antiparallel, a structure that is homologous to bacteriophage T7 lysosome and bacterial type 2 amidases (Dziarski & Gupta 2006; Hou et al., 2020). Enzymatic and bactericidal PGRPs, such as PGRP-LB, PGRP-SC and PGRP-SB, in *D. melanogaster* directly cleave PGN into fragments by removing peptides from glycan chains (Iatsenko et al., 2016). PGRP-LC, PGRP-LE and PGRP-SA have no catalytic activity and only have a conserved domain that binds PGN and activates downstream signaling pathways upon pathogen ligand binding (Royet et al., 2011; Iatsenko et al., 2016). In mammals, the proteins PGRP-L, PGRP-I α , PGRP-I β are secreted and can act as bactericides.

The crystal structures of *D. melanogaster* proteins PGRP-SA, PGRP-SD, PGRP-LCa, PGRP-LE, PGRP-LF and the PGRP-LCx-LCa complex are known (Liu et al., 2019). All PGRPs have a conserved PGN-binding cleft with a bound zinc ion and a similar overall fold, but they differ in their N-terminal regions (Guan et al., 2005). Therefore, PGN, a component of bacterial and fungal cell walls, is recognized by a conserved PGN-binding cleft with a bound zinc ion within PGRPs and these complexes play a key role in triggering insect immune responses to pathogens (Liu et al., 2019).

This research is mainly focused on the PGRP genes in white-spotted flower chafers, *Protaetia brevitarsis seulensis* (Kolbe). This beetle is easy to maintain under laboratory conditions, inject with pathogens and for obtaining tissues including fat bodies, guts, and haemocytes. The cellular and humoral immune system of this beetle is especially well developed (Kwon et al., 2014; Bang et al., 2015; Lee et al., 2016). Using these beetles, we describe the isolation, characterization and immune-related functions of PGRP cDNAs, including the complete sequences of cDNA clones and analysis of their sequences in a phylogenetic context. In addition, to characterizing the pos-

sible anti-pathogen functions of the PGRP proteins, larvae were challenged with Gram-negative bacteria (*Escherichia coli* K12), Gram-positive bacteria (*B. cereus*) and yeast (*Saccharomyces cerevisiae*). We demonstrate that PGRP mRNA levels were up-regulated following pathogen infection and may be associated with innate immune functions.

MATERIAL AND METHODS

Insects

A method for rearing white-spotted flower chafers, *P. brevitarsis seulensis* (Kolbe) (Coleoptera: Cetoniidae) is described by Kwon (2014). Twenty larvae were reared in a clear plastic box (30 cm \times 20 cm \times 20 cm) in a constant environment (CE) incubator (MIR-553; Sanyo Electric Biomedical, Japan) at 25 \pm 1°C, 40% relative humidity and long photoperiod of 16L:8D under aseptic conditions in well-fermented sterile oakwood sawdust.

RNA extractions, cDNA preparation, sequence characterization and prediction of protein structures

We extracted total RNA from the fat bodies of each of the challenged and control larvae using the SV Total RNA Isolation System Kit (Promega Corp., Madison, WI, USA). Then, total RNA was dissolved in RNase free sterile water and RNA integrity was checked using an Agilent 2100 BioAnalyzer (Agilent Technologies, Englewood, CO, USA).

To obtain partial PGRP cDNA sequences, we searched for PGRP like sequences in high-throughput RNA sequencing (RNA-seq) expression profiling data using known PGRPs from *D. melanogaster*, *Tenebrio molitor* and *B. mori*. This analysis revealed four candidate partial cDNA sequences of PGRPs. To clone the full length of the PGRP cDNAs, we used 5' and 3' Rapid Amplification of cDNA Ends (RACE; Clontech, Palo Alto, CA) polymerase chain reaction (PCR) and nested PCR with gene specific primer sets (Table 1). PCR was done using the following conditions: 30 cycles of 30 s at 94°C, 30 s at 55°C and 1

Table 1. List of primers used in this study.

| Purpose | Primers | Sequences |
|------------|-------------------|----------------------------------|
| RACE -PCR | PGRP-SA(a)RACE F1 | 5'-tgggcaacaattgatgggctagtga-3' |
| | PGRP-SA(a)RACE R1 | 5'-gtgtattattgacgtgtcaggtagcg-3' |
| | PGRP-SA(b)RACE F1 | 5'-gtgggcagatgatcgatcgtagcaa-3' |
| | PGRP-SA(b)RACE R1 | 5'-gtgcgttacctagtcgatgcttagt-3' |
| | PGRP-SB1RACEF1 | 5'-aatttcgatggtttacgaagcg-3' |
| | PGRP-SB1RACER1 | 5'-ggcctgattagctaattgtag-3' |
| | PGRP-SCRACEF1 | 5'-tagcatgcaaatggcccgtag-3' |
| | PGRP-SCRACER1 | 5'-cgtaggtagctattgatagct-3' |
| Nested-PCR | PGRP-SA(a) F2 | 5'-gtacgtttaaagctggcactgca-3' |
| | PGRP-SA(a) R2 | 5'-ttagctaggaaaatgcacggac-3' |
| | PGRP-SA(b) F2 | 5'-gtacgataaagctgatgcattag-3' |
| | PGRP-SA(b) R2 | 5'-gtgactgaaatgctgcccgatg-3' |
| | PGRP-SB1RACEF2 | 5'-tatttcggggcccatgctagaa-3' |
| | PGRP-SB1RACER2 | 5'-gttcattccatatataggcat-3' |
| | PGRP-SCRACEF2 | 5'-atcatcgtagccctgggatgct-3' |
| | PGRP-SCRACER2 | 5'-gtggctaattgctgaatgcat-3' |
| RT-PCR | PGRP-SA(a)RTF1 | 5'-gtgacttgattagctaaggcc-3' |
| | PGRP-SA(a)RTR1 | 5'-aattgcttgatgcccgaattga-3' |
| | PGRP-SA(b)RTF1 | 5'-gtggaacgcgcttgattgc-3' |
| | PGRP-SA(b)RTR1 | 5'-ggccctgttaaatgctgatca-3' |
| | PGRP-SB1RACERTF1 | 5'-aatggcgcgtttgactggatg-3' |
| | PGRP-SB1RACERTR1 | 5'-gaattgcgtagctgattagc-3' |
| | PGRP-SCRACERTF1 | 5'-atggcctgattgctgattagc-3' |
| | PGRP-SCRTR1 | 5'-atatgcgatgctgatggctta-3' |
| | PsactinF1 | 5'-ggttggatgggtacgaagga-3' |
| | PsactinR1 | 5'-gcgggtgggtgaagagta-3' |

min at 72°C, followed by terminal extension at 72°C for 2 min and cooling at 30°C for 1 min. Ten microliters of PCR products were electrophoresed on 1.5% agarose gel and after staining with ethidium bromide was recorded using a direct screen camera. The pGEM-T Easy vector was used for cloning all PCR products (Promega Corp., Madison, WI, USA) and cloning products were analysed by DNA sequencer (Applied Biosystems, Foster City, CA). Two primers were used for sequencing: T7 (5'-taatacgactcactatagg-3') and SP6 (5'-attagtgacactatag-3'). To compare nucleotides or proteins, we used GenBank and the Basic Local Alignment Search Tool (BLAST: Blastn for nucleotides or Blastp for proteins) in the National Centre for Biotechnology Information server (<http://www.ncbi.nlm.nih.gov/>). The amino acid sequences of various PGRPs from insects and human were aligned using the Cluster algorithm, ClustalW ver 1.8 (Thompson et al., 1994) and phylogenetic trees constructed based on alignment by MEGA 7: Molecular Evolutionary Genetics Analysis Version 7.0 (Tamura et al., 2007). We used the Phyre² server to construct predicted tertiary protein structures (<http://www.sbg.bio.ic.ac.uk/phyre2>).

Expression of PS PGRPs mRNA and real-time quantitative RT-PCR (qPCR)

To induce PS PGRPs mRNA, the last larval instar was injected with ultraviolet (UV)-killed Gram negative bacteria 4×10^4 *E. coli* K12 or Gram positive 3×10^5 *B. cereus* or 5×10^3 yeast *Saccharomyces cerevisiae* and control larvae with LB media. One hundred last instar larvae were divided into four injected groups (25 larvae in each group), into which 10 µl of gram positive bacteria, gram negative bacteria, yeast or sterile water were injected, respectively. Then, total RNA was isolated from fat bodies of pooled samples (five individuals) of each group at 4 h, 8 h, 12 h, 24 h and 48 h post infection (Promega Corp., Madison, WI, USA). Approximately 1 µg total RNA from each group was pooled. Total RNA from fat body were extracted using the SV Total RNA Isolation system kit (Promega Corp., Madison, WI, USA) and then, cDNAs were constructed using GoScriptTM Reverse Transcriptase according to the manufacturer's procedures (Invitrogen, Karlsruhe, Germany). Briefly, we converted ~1 µg of total RNA into single-stranded (ss) DNA using control primer one (5 µM dT-ACP1), two (dT-ACP2) which prime form polyA tail, dNTP (1 mM each), 2.5 µL of 25 mM MgCl₂, 2 µL of RNase inhibitor (40 U/µL; Promega, Madison, WI, USA), and 1 µL of reverse transcriptase (200 U/µL; Promega Corp., Madison, WI, USA). The total ssDNA reaction mixture was 20 µL and subjected to 60 min at 42°C, then 15 min at 70°C. Then it was subjected to the following conditions in a Rotor-Gene-Q (Qiagen, USA) RT-PCR machine: 35 cycles of 94°C for 30 s, 55°C for 20 s and 72°C for 10 s with a single 470 nm green colour fluorescence measurement in 5 µl SYBR GreenI master mix (Elpis-biotech, Korea), 0.2 µl of cDNA, and 0.4 µM each of gene-specific primers (Table 1). The *β-actin* with the following primers, *PsactinF1* and *PsactinR1* were used to normalize the qPCR reactions and Prism 5.0 software was used for the statistical analyses of the qPCR data (GraphPad Software, San Diego, CA, USA). For statistical tests, we used Student's two-tailed t-test or one-way ANOVA at a probability (*p*) value of less than 5%.

RESULTS

Identification, sequence analysis and phylogenetic analysis of PGRP-like cDNAs in larvae of *P. brevitarsis seulensis*

To identify *PGRP*-like cDNAs, we searched *PGRP*-like cDNA sequences from high-throughput RNA sequencing

(RNA-seq) expression data generated by our group (Bang et al., 2015). We searched over 35,000 sequences and isolated four *PGRP*-like sequences of 155 bp, 243 bp, 222 bp and 242 bp. To clone whole sequences of *PGRP*-like cDNAs, we performed 5' and 3' rapid amplification of cDNA ends (*RACE*) PCR with standard nested PCR primer sets designed based on the partial sequences. Finally, we cloned and sequenced full-length *PGRP*-like cDNAs named *PS PGRP-SA(a)*-like (OK413871), *PS PGRP-SA(b)*-like (OK413872), *PS PGRP-SB1*-like (OK413873) and *PS PGRP-SC*-like (OK413874) (Fig. 1). The full-length *PS PGRP-SA(a)*-like transcript includes 594 bp encoding a polypeptide of 198 amino acids (aa) with a calculated molecular mass of 22.1 kDa and predicted isoelectric point (pI) of 7.68 (Fig. 1A). The *PS PGRP-SA(b)*-like and *PS PGRP-SB1*-like transcripts are 564 and 594 nucleotides encoding 188 and 198 aa polypeptides, with calculated molecular masses of 20.8 and 20.9 kDa and predicted pI values of 8.33 and 8.97, respectively. The *PS PGRP-SC*-like transcript is 570 nucleotides encoding a 190 aa polypeptide with a calculated molecular mass of 19.7 kDa and a predicted pI of 9.82. The first 23, 19, 19 and 20 aa of the N-terminal of *PS PGRP-SA(a)*-like, *PS PGRP-SA(b)*-like, *PS PGRP-SB1*-like and *PS PGRP-SC*-like proteins comprise predicted secretion signal peptides (Fig. 1, red letters). We searched *PS PGRP*-like protein domains against the Pfam database (<http://pfam.sanger.ac.uk>) and found that *PS PGRP-SA(a)*-like, *PS PGRP-SA(b)*-like, *PS PGRP-SB1*-like and *PS PGRP-SC*-like aa residues 53–180, 41–171, 41–171 and 43–173, respectively, which correspond to conserved *PGRP* domains (Fig. 1, blue letters). We confirmed that these proteins are members of the *PGRP* family in insects. As described above, these *PGRP*-like proteins share sequence homology and have similar molecular weights, pI values and *PGRP* protein domains.

Most *PGRPs* contain a conserved consensus sequence (His; H-Tyr; Y-His; H-Thr; T-Cys; C) for amidase activity and residues for binding to PGN, one of the major bacterial cell wall components: Asp (D) and Phe (F) for Lys-type PGN; and Gly(G), Trp (W) and Arg (R) for DAP-type PGN. As shown in Fig. 1B, we identified H-T-H-T-C residues in *PS PGRP-SC*-like protein sequences, which are also conserved in *Aedes aegypti* (*PGRP-LB* and *PGRP-S3*), *Bombyx mori* (*PGRP-LP1*), *Drosophila melanogaster* (*PGRP-LB(a)* and *PGRP-SCs*) and *Samia cynthia ricini* (*PGRP-B*, C, and D). In addition, the *PS PGRP-SC*-like protein contains conserved Gly(G), Trp (W) and Arg (R) residues for binding DAP-type PGN (Fig. 1B), while *PS PGRP-SA(b)*-like contains conserved Asp (D) and Phe (F) residues for binding Lys-type PGN (Fig. 1B). These DAP-type PGN-binding residues (G, W and R) are only found in *A. aegypti* (*PGRP-S3*). The Lys-type PGN-binding sequence (D and F) in *PGRP-SA(b)* is also found in *D. melanogaster PGRP-SA*.

The amino acid sequences of deduced *PS PGRP-SA(a)*-like, *PS PGRP-SA(b)*-like, *PS PGRP-SB1*-like and *PS PGRP-SC*-like protein were compared with other *PGRP* orthologs (as a phylogenetic outgroup) from vertebrates

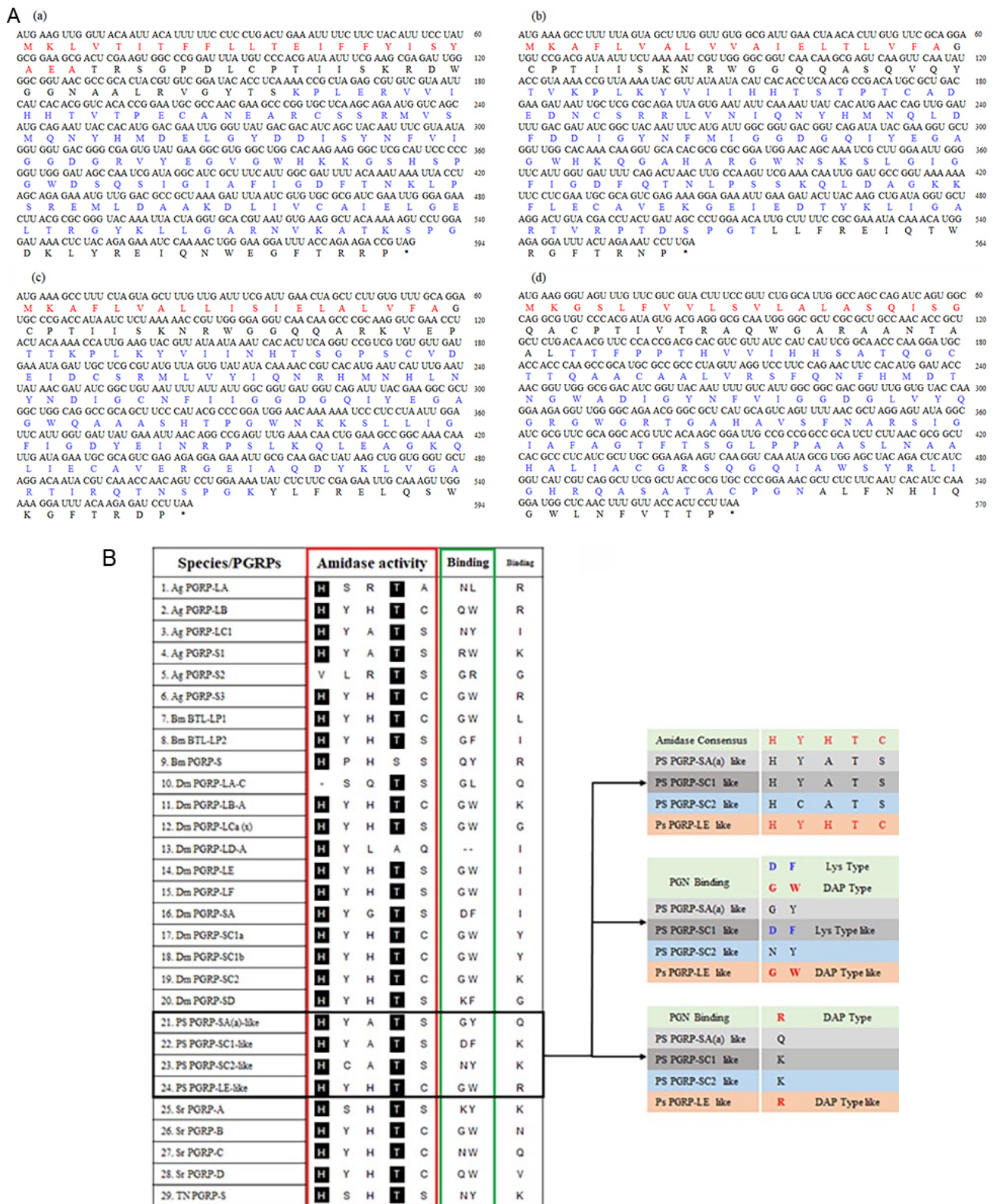


Fig. 1. (A) Nucleic acid and amino acid sequences of PS PGRP-SA(a)-like, PS PGRP-SA(b)-like, PS PGRP-SB1-like and PS PGRP-SC. (a) PS PGRP-SA(a)-like: the 594-nucleotide gene encodes a 198 amino acid polypeptide. (b) PS PGRP-SA(b)-like: the 564-nucleotide gene encodes a 188 amino acid polypeptide. (c) PS PGRP-SB1-like: the 594 bp gene encodes a 198 amino acid polypeptide. (d) PS PGRP-SC-like: the 570-nucleotide gene encodes a 190 amino acid polypeptide. The first 23, 19, 19 and 20 amino acids of the N-terminal of PS PGRP-SA(a)-like, PS PGRP-SA(b)-like, PS PGRP-SB1-like and PS PGRP-SC-like comprise predicted secretion signal peptides (indicated by red letters). PS PGRP-SA(a)-like, PS PGRP-SA(b)-like, PS PGRP-SB1-like and PS PGRP-SC-like residues 53–180, 41–171, 41–171 and 43–173, respectively, correspond to conserved PGRP domains (indicated by blue letters). (B) PS PGRPs contain a conserved consensus sequence for amidase activity (His; H-Tyr; Y-His; H-Thr; T-Cys; C) and bacterial cell wall component binding residues Asp (D) and Phe (F) for Lys-type PGN, and Gly(G), Trp (W) and Arg (R) for DAP-type PGN. PS PGRP-SC-like protein sequences contain H-T-H-T-C residues and conserved Gly(G), Trp (W) and Arg (R) residues for DAP-type PGN. The Lys-type PGN binding sequence is present in PGRP-SA(b).

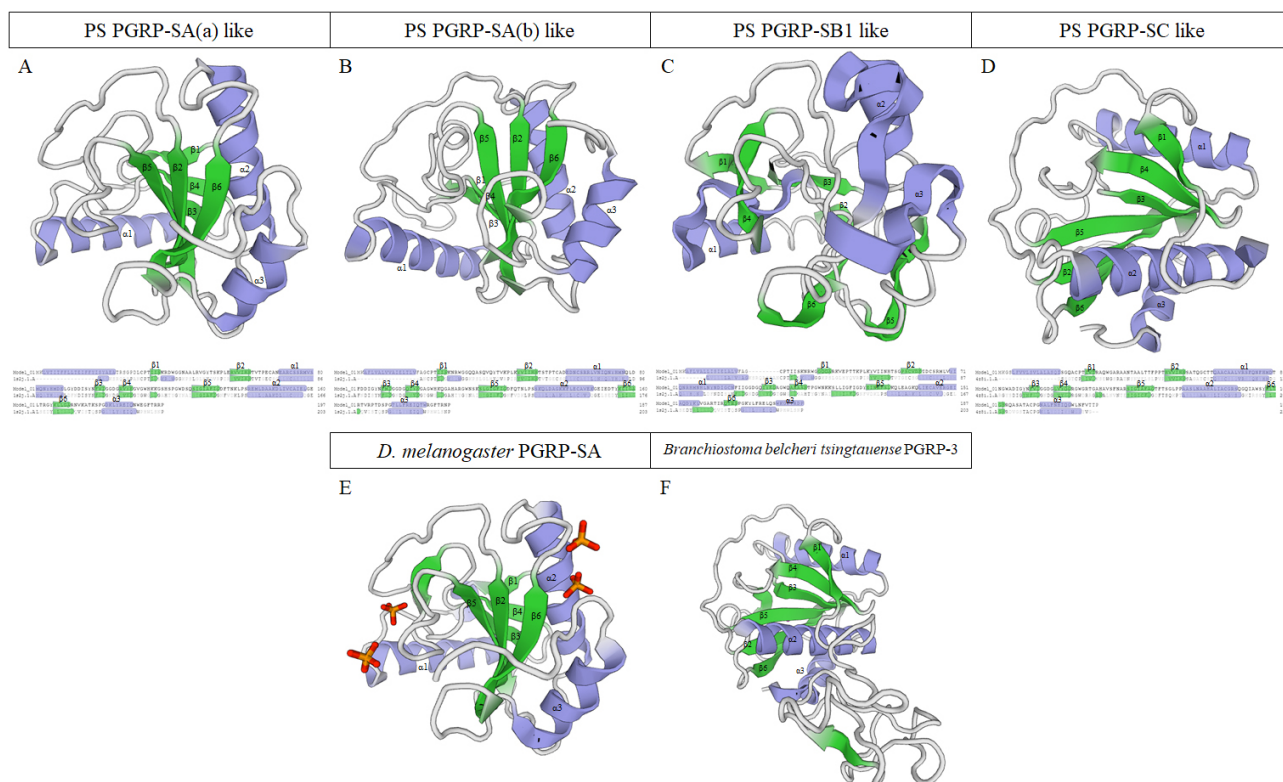


Fig. 2. Predicted structures of PS PGRPs. (A) PS PGRP-SA(a)-like, (B) PS PGRP-SA(b)-like, (C) PS PGRP-SB1-like, (D) PS PGRP-SC, (E) *D. melanogaster* PGRP-SA and (F) *B. belcheri tsingtauense* PGRP-3. The topological structures of PS PGRP-SA(a)-like and PS PGRP-SA(b)-like proteins are similar to the *D. melanogaster* PGRP-SA crystal structure (E). The β -sheets of PS PGRP-SA(a)-like and PS PGRP-SA(b)-like proteins, consisting of strands β 1– β 6, are located in a central pocket, and the peripheral helix α 1 is located on the front, while peripheral α 2 and α 3 helices are on the back of the pocket. PS PGRP-SB1-like also has six β -strands and three α -helices, but the topological structure differs from that of typical structures of PGRPs. The PS PGRP-SC-like protein is structurally similar to *B. belcheri tsingtauense* PGRP-3, and also consists of six β -strands and three α -helices (34% α -helices and 15% β -strands), with β -strands located in a central cavity: helix α 1 located on the back and peripheral α 2 and α 3 helices located on the front of the central cavity.

and invertebrates. The results showed that PS PGRP-SA(a)-like and PS PGRP-SA(b)-like are clustered with *D. melanogaster* PGRP-SA. The PS PGRP-SB1-like is clustered with *D. melanogaster* PGRP-SB and PS PGRP-SC-like formed an independent cluster with *D. melanogaster* PGRP-SC (data not shown). The aa sequence identity between PS PGRP-SA(a)-like and PGRP-SA from *D. melanogaster* is 45.74%, compared with 47.93% between PS PGRP-SA(b) and PGRP-SA from *D. melanogaster*. The aa sequence of the PS PGRP-SB1-like protein is 39.77% homologous with that of PGRP-SB1 of *D. melanogaster*. PS PGRP-SC-like and PGRP-SC1 from *D. melanogaster* share 44.26% aa sequence homology. In addition, the aa sequence of PS PGRP-SC shares 44.51% homology with human PGRP-1. Similarly, the aa sequence of PS-PGRP family members in this insect share 39.77–47.93% homology with PGRP family members in other species, including human.

Structural features of PS PGRP-like proteins

To better understand the relationships between PS PGRP families, the SWISS-MODEL server (<https://swissmodel.expasy.org/interactive>) was used to build three-dimensional models of the corresponding PGRP domains. The topological structures of PS PGRP-SA(a)-like and PS PGRP-SA(b)-like proteins are homologous with the *D.*

melanogaster PGRP-SA structure determined by X-ray diffraction (Fig. 2A, B, and E). The structures of PS PGRP-SA(a)-like and PS PGRP-SA(b)-like proteins superposed well with *D. melanogaster* PGRP-SA (Fig. 2A, B and E). As shown in Fig. 2E, PS PGRP-SA(a)-like and PS PGRP-SA(b)-like proteins also consist of six β -strands and three α -helices. According to Phyre2 JSmol, PS PGRP-SA(a) has 35% α -helices, 17% β -strands, 8% transmembrane helices and 12% random coil, compared with 38% α -helices, 16% β -strands and 10% random coil for PS PGRP-SA(b)-like protein (Fig. 2A and B). The β -sheets of PS PGRP-SA(a)-like and PS PGRP-SA(b)-like, consisting of β 1– β 6 strands, are located in a central pocket, the peripheral α 1 helix is located on the front face and the peripheral α 2 and α 3 helices are on the back of the pocket. The β 1, β 3 and β 4 strands are paired in the same direction, while β 2, β 5 and β 6 strands are also paired. The paired β 1, β 3 and β 4 strands face the paired β 2, β 5 and β 6 strands (Fig. 2A and B). The topological structures of PS PGRP-SC-like proteins also share homology with PS PGRP-SA(a) and PS PGRP-SA(b), as well as *D. melanogaster* PGRP-SA. Likewise, PS PGRP-SC-like shares homology with *Branchiostoma belcheri tsingtauense* PGRP-3, which also consists of six β -strands and three α -helices (34% α -helices and 15% β -strands) and its β -strands are located in a central cavity: helix α 1 is lo-

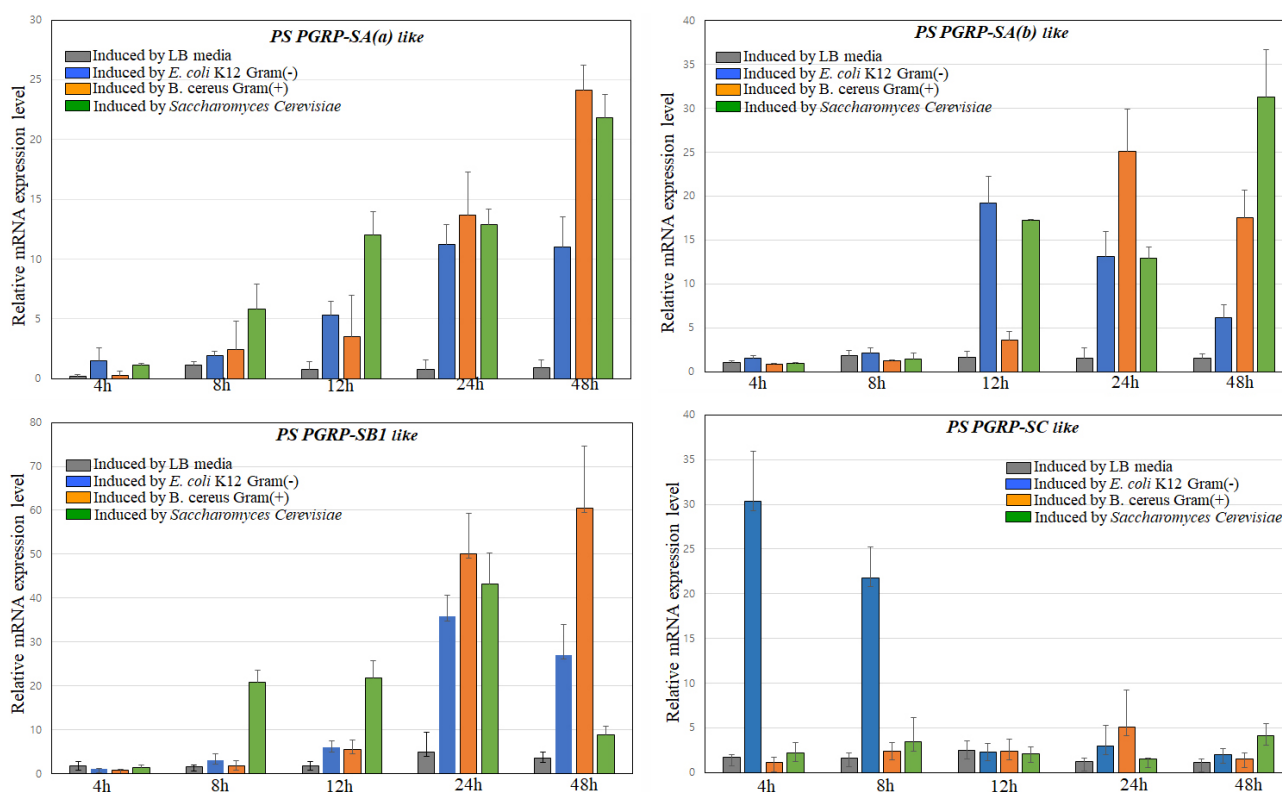


Fig. 3. Expression analysis of PGRP-like mRNAs induced by Gram-negative bacteria (*E. coli* K12), Gram-positive bacteria (*B. cereus*) or yeast (*Saccharomyces cerevisiae*). Control larvae were injected with LB media. From 4 to 48 h after injection, *PS PGRP-SA(a)*-like, *PS PGRP-SA(b)*-like, *PGRP-SB1*-like and *PS PGRP-SC*-like transcript levels in the fat body were measured by quantitative real-time PCR (qRT-PCR), with β -actin mRNA as an internal control. *PS PGRP-SA(a)*-like mRNA levels were increased 10- to 20-fold at 12 and 48 h relative to control larvae. *PS PGRP-SA(b)*-like mRNA levels were up-regulated at 24 h post-infection with Gram-positive bacteria, but they decreased thereafter and levels were highly up-regulated following infection with yeast at 12, 24 and 48 h. *PS PGRP-SB1*-like mRNA levels were significantly up-regulated in response to infection by all three pathogens (over 30-fold at 24 h in all post-challenged larvae). *PS PGRP-SC*-like mRNA transcript levels were immediately increased by more than 30-fold within 4 h and then slightly decreased to over 20-fold at 8 h, but expression levels decreased at 12, 24 and 48 h post-challenge. RT-PCR results are expressed as the mean of three replicates \pm SEM, each containing five larvae (fifteen larvae per condition). * $P < 0.05$ (t-test).

cated on the back face and the peripheral $\alpha 2$ and $\alpha 3$ helices are on the front face (Fig. 2D and 2F). *PS PGRP-SB1*-like also has six β -strands and three α -helices (35% α -helices and 15% β -strands), but the topological structure differs from that of typical structures of PGRPs (Fig. 2C).

Expression of PGRP-like mRNAs induced by bacteria and fungi

To investigate the involvement of *PS PGRP-SA(a)*-like, *PS PGRP-SA(b)*-like, *PGRP-SB1*-like and *PS PGRP-SC*-like mRNAs in rapidly detecting various infectious pathogens, larvae of *P. brevitarsis seulensis* were injected with ultraviolet (UV)-killed Gram-negative *E. coli* K12, Gram-positive *Bacillus cereus* or yeast *Saccharomyces cerevisiae* cells. Control larvae were injected with LB media. From 4 to 48 h after injection, the levels of *PS PGRP-SA(a)*-like, *PS PGRP-SA(b)*-like, *PGRP-SB1*-like and *PS PGRP-SC*-like transcripts in the fat body were measured by quantitative real-time PCR (qRT-PCR), with β -actin mRNA as an internal control (Fig. 3A–D). Expression of *PS PGRP-SA(a)*-like mRNA in *E. coli* K12-challenged larvae was ~5-fold greater between 4 and 48 h compared with the controls (Fig. 3A). In the case of *B. cereus*- and yeast *S. cerevisiae*-challenged larvae, expression of *PS PGRP-SA(a)*-like

mRNA was over 10- to 20-fold greater between 12 and 48 h (Fig. 3A). We also observed a significant increase in the levels of *PS PGRP-SA(b)*-like mRNA when challenged with *B. cereus* or *S. cerevisiae* (Fig. 3B). As shown in Fig. 3B, *PS PGRP-SA(b)*-like mRNA levels were greater 24 h post-infection with *B. cereus* and decreased thereafter. We also observed that *PS PGRP-SA(b)*-like mRNA expression was highly up-regulated in larvae challenged with *S. cerevisiae* at 12, 24 and 48 h post-infection (Fig. 3B).

In the case of *PS PGRP-SB1*-like mRNA, we observed a significant up-regulation of mRNA levels in response to infection of all three pathogens (over 30-fold at 24 h post-challenge; Fig. 3C). However, *PS PGRP-SB1*-like mRNA expression then decreased slightly at 48 h in *E. coli* K12- and *S. cerevisiae*-challenged larvae, but its expression remained up-regulated in response to infection with *B. cereus* (over 60-fold; Fig. 3C).

Upon infection with *E. coli* K12, levels of *PS PGRP-SC*-like mRNA transcripts immediately increased by more than 30-fold within 4 h and then slightly decreased to over 20-fold at 8 h, but up-regulation did not persist at 12, 24 and 48 h post-challenge (Fig. 3D). By contrast, challenge with *B. cereus* or *S. cerevisiae* did not result in an increase in *PS PGRP-SC*-like mRNA levels at 4 to 48 h (Fig. 3D). Ex-

pression of *PS PGRP-SA(a)*-like and *PS PGRP-SA(b)*-like mRNAs in the immune response to infection with Gram-positive bacteria or yeast was evident through similar gene expression patterns. *PS PGRP-SB1*-like mRNA was up-regulated by infection with Gram-positive bacteria, Gram-negative bacteria and yeast, but expression of *PS PGRP-SC*-like mRNA was only increased by 20- or 30-fold at 4h following challenge with Gram-negative bacteria.

DISCUSSION

In this study, we cloned four small-type PGRPs from *P. brevitarsis seulensis*, predicted their protein structures and expression analysis using quantitative real-time PCR in order to provide clues to understanding their functions. PGRPs are found in all animals, but not in lower metazoans or plants, and their domains are roughly ~42% identical and ~55% similar to each other (Guan et al., 2005; Dziarski & Gupta, 2006). The PGRP domain is ~165 aa long, it is structurally homologous to bacteriophage and bacterial type 2 amidases, and is found in all PGRPs in vertebrates and invertebrates (Royet et al., 2011). Likewise, the aa sequences of PS PGRPs share 32.03–47.93% homology with those of PGRP family members in insects and mammals, including humans. In addition, models based on crystal structures of PGRPs in *D. melanogaster* and humans revealed homology with T7 lysozymes and type 2 bacteriophage amidases, which are composed of three peripheral α -helices and several central β -sheet strands (Dziarski & Gupta, 2006). Consistent with the crystal structure of PGRP-SA from *D. melanogaster* (Fig. 3E), PS PGRP-SA(a)-like, PS PGRP-SA(b)-like and PS PGRP-SC-like proteins share a typical PGRP structure. All β -sheets, consisting of strands β 1 to β 6, are located in a central pocket, and the peripheral α 1 helix is located on the front, while peripheral α 2 and α 3 helices are on the back. Based on their structural similarities, we can assume that they function as PS PGRPs. In addition, our expression data support the predicted structural analysis.

As shown in Fig. 1A and 2, the four cloned PS PGRPs are 188–198 aa in length, similar to insect and human short-type PGRPs. Numerous PGRPs are reported as long (L) or short (S) types based on their transcript size, including proteins in flies (*D. melanogaster*), silkworms (*B. mori*) and yellow fever mosquitoes (*A. aegypti*). We did not observe transmembrane domain sequences in any of the PS PGRPs, and all produced relatively short transcripts with signal sequences. Therefore, we assumed that all four PS PGRPs are short-type PGRPs-like proteins that are secreted as extracellular proteins. In *D. melanogaster*, short-type PGRPs are structurally similar to each other and are exported into different tissues of the body as inducible proteins (Werner et al., 2000). For example, *D. melanogaster* PGRP-SA, which is homologous to PS PGRPs-SA(a) and PS PGRP-SA(b), is also secreted and induced by an infection of the epidermis. Thus, PGRP-SA proteins perform specific roles in the epidermis, which acts as a barrier to infection and has its own inducible antibacterial peptides (Werner et al., 2000). As shown in Fig. 3, expression of

PS PGRP-SA(a)- and *-SA(b)*-like mRNAs in larvae challenged with Gram-negative bacteria (*E. coli* K12) was low compared with expression in response to Gram-positive bacteria (*B. cereus*) or yeast (*S. cerevisiae*). These results suggest that PS PGRP-SA(a) and -SA(b) proteins might be involved in the immunological functions of the epidermis associated with fighting infection by Gram-positive bacteria and yeast. However, in another study, PGRP-SA from *D. melanogaster* is also expressed in the fat body, where it acts as an activator of the Toll pathway following infection by Gram-positive bacterial (Michel et al., 2001). Further expression studies in various tissues in response to infection are clearly required.

D. melanogaster proteins PGRP-SB1 and PGRP-SC are secreted from the fat body, possess bactericidal amidase activity and regulate AMPs through the IMD pathway. Herein, we observed that *PS PGRP-SB1*-like mRNA was strongly induced by Gram-negative bacteria, Gram-positive bacteria and yeast (Fig. 3). Therefore, PS PGRP-SB1-like appears to be distinct from PGRP-SB1 and PGRP-SC in *D. melanogaster*. As shown in the sequence analysis in Fig. 1B, the consensus sequences for bacterial amidase activity and bacterial cell wall binding-type (diaminopimelic acid, DAP-type; PGN containing lysine, Lys-type) were not identified in the *PS PGRP-SB1* DNA sequence. Thus, PS PGRP-SB1 might only function as an immune effector to activate IMD and Toll pathways, since we observed that *PS PGRP-SB1*-like mRNA levels were up-regulated following infection by all three pathogens.

Insect immune responses are triggered by microbial PGN, LPS, β -glucans, lipoproteins, CpG dinucleotides or flagellin, which are all products of microbial metabolism. Indeed, the cell walls of almost all bacteria are composed of PGN. Many Gram-negative bacterial cell walls are made of cross-linked PGN containing DAP (DAP-type), while PGN containing lysine (Lys-type) is found in many Gram-positive bacterial cell walls. Dap-type and Lys-type PGN are the primary ligands of PGRP molecules. The H-Y-H-T-C sequence is highly conserved for the recognition and hydrolysis of PGN (Mellroth et al., 2003; Liu et al., 2021). PS PGRP-SC-like protein has this consensus sequence for amidase activity, as well as G-W-R corresponding to DAP-type PGN-binding sites (Fig. 1B). In addition, we confirmed that *PS PGRP-SC*-like mRNA transcript levels were immediately increased 20–30-fold shortly after (4–8 h) infection with Gram-negative bacteria (*E. coli* K12; Fig. 3). By contrast, larvae challenged with Gram-positive bacteria or yeast cells did not display elevated *PS PGRP-SC*-like mRNA levels. In *D. melanogaster*, PGRP-SB2, PGRP-SC1a, PGRP-SC1b and PGRP-LB possess bactericidal activity and zinc-dependent amidase activity that removes peptides from glycan chains, thereby reducing or eliminating the biological activity of PGN (Zaidman-Rémy et al., 2006). Therefore, we assumed that PGRP-SC-like in *P. brevitarsis seulensis* encodes a secreted PGRP with amidase activity against DAP-type PGN and may also play a role as an immune effector to regulate the IMD pathway.

ACKNOWLEDGEMENTS. This work was supported by Korea Institute of Planning and Evaluation for Technology in Food, Agriculture, Forestry and Fisheries (IPET) through Industrialization technology development to respond to crop viruses and pests Project, funded by Ministry of Agriculture, Food and Rural Affairs (MAFRA) (No. 121052-2)

REFERENCES

- BANG K., PARK S. & CHO S. 2013: Characterization of a β -1, 3-glucan recognition protein from the beet armyworm, *Spodoptera exigua* (Insecta: Lepidoptera: Noctuidae). — *Insect Sci.* **20**: 575–584.
- BANG K., HWANG S., LEE J. & CHO S. 2015: Identification of immunity-related genes in the larvae of *Protaetia brevitarsis seulensis* (Coleoptera: Cetoniidae) by a next-generation sequencing-based transcriptome analysis. — *J. Insect Sci.* **15**: 142, 6 pp.
- CHAPUT C. & BONECA I.G. 2017: Peptidoglycan detection by mammals and flies. — *Microb. Infect.* **9**: 637–647.
- CHO Y. & CHO S. 2019: Hemocyte-hemocyte adhesion by granulocytes is associated with cellular immunity in the cricket, *Gryllus bimaculatus*. — *Sci. Rep.* **9**: 1–12.
- CHRISTOPHIDES G.K., ZDOBNOV E., BARILLAS-MURY C., BIRNEY E., BLANDIN S., BLASS C., BREY P.T., COLLINS F.H., DANIELLI A. & DIMOPOULOS G. 2002: Immunity-related genes and gene families in *Anopheles gambiae*. — *Science* **298**: 159–165.
- DZIARSKI R. & GUPTA D. 2006: The peptidoglycan recognition proteins (PGRPs). — *Genome Biol.* **7**: 1–13.
- EVANS J.D., ARONSTEIN K., CHEN Y.P., HETRU C., IMLER J., JIANG H., KANOST M., THOMPSON G.J., ZOU Z. & HULTMARK D. 2006: Immune pathways and defence mechanisms in honey bees *Apis mellifera*. — *Insect Mol. Biol.* **15**: 645–656.
- GUAN R., WANG Q., SUNDBERG E.J. & MARIUZZA R.A. 2005: Crystal structure of human peptidoglycan recognition protein S (PGRP-S) at 1.70 Å resolution. — *J. Mol. Biol.* **347**: 683–691.
- HOFFMANN J.A. 2003: The immune response of *Drosophila*. — *Nature* **426**: 33–38.
- HOU H., GUO M., GENG J., WEI X., HUANG D. & XIAO J. 2020: Genome-wide analysis of peptidoglycan recognition protein genes in fig wasps (Hymenoptera, Chalcidoidea). — *Insects* **11**: 597, 15 pp.
- IATSENKO I., KONDO S., MENGIN-LECREULX D. & LEMAITRE B. 2016: PGRP-SD, an extracellular pattern-recognition receptor, enhances peptidoglycan-mediated activation of the *Drosophila* Imd pathway. — *Immunity* **45**: 1013–1023.
- JANEWAY JR C.A. & MEDZHITOV R. 2002: Innate immune recognition. — *Annu. Rev. Immunol.* **20**: 197–216.
- KWON H., BANG K. & CHO S. 2014: Characterization of the hemocytes in larvae of *Protaetia brevitarsis seulensis*: involvement of granulocyte-mediated phagocytosis. — *PLoS ONE* **9**: e103620, 12 pp.
- LEE J., HWANG S. & CHO S. 2016: Immune tolerance to an intestine-adapted bacteria, *Chryseobacterium* sp., injected into the hemocoel of *Protaetia brevitarsis seulensis*. — *Sci. Rep.* **6**: 1–14.
- LEMAITRE B. & HOFFMANN J. 2007: The host defense of *Drosophila melanogaster*. — *Annu. Rev. Immunol.* **25**: 697–743.
- LIU Y., ZHAO X., HUANG J., CHEN M. & AN J. 2019: Structural insights into the preferential binding of PGRP-SAs from bumblebees and honeybees to dap-type peptidoglycans rather than lys-type peptidoglycans. — *J. Immunol.* **202**: 249–259.
- LIU F., LI H., YANG P. & RAO X. 2021: Structure-function analysis of PGRP-S1 from the oriental armyworm, *Mythimna separata*. — *Arch. Insect Biochem. Physiol.* **106**: e21763, 18 pp.
- MELLROTH P., KARLSSON J. & STEINER H. 2003: A scavenger function for a *Drosophila* peptidoglycan recognition protein. — *J. Biol. Chem.* **278**: 7059–7064.
- MICHEL T., REICHHART J., HOFFMANN J.A. & ROYET J. 2001: *Drosophila* Toll is activated by gram-positive bacteria through a circulating peptidoglycan recognition protein. — *Nature* **414**: 756–759.
- RAGAN E., AN J.C., JIANG H. & KANOST M.R. 2009: Roles of haemolymph proteins in antimicrobial defences of *Manduca sexta*. In Rolff J. & Reynolds S. (eds): *Insect Infection and Immunity. Evolution, Ecology, and Mechanisms*. Oxford University Press, pp. 34–48.
- ROYET J., GUPTA D. & DZIARSKI R. 2011: Peptidoglycan recognition proteins: modulators of the microbiome and inflammation. — *Nature Reviews Immunol.* **11**: 837–851.
- TAMURA K., DUDLEY J., NEI M. & KUMAR S. 2007: MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. — *Mol. Biol. Evol.* **24**: 1596–1599.
- TANAKA H., ISHIBASHI J., FUJITA K., NAKAJIMA Y., SAGISAKA A., TOMIMOTO K., SUZUKI N., YOSHIYAMA M., KANEKO Y. & IWASAKI T.A. 2008: Genome-wide analysis of genes and gene families involved in innate immunity of *Bombyx mori*. — *Insect Biochem. Mol. Biol.* **38**: 1087–1110.
- THOMPSON J.D., HIGGINS D.G. & GIBSON T.J. 1994: CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. — *Nucleic Acids Res.* **22**: 4673–4680.
- WANG Q., REN M., LIU X., XIA H. & CHEN K. 2019: Peptidoglycan recognition proteins in insect immunity. — *Mol. Immunol.* **106**: 69–76.
- WERNER T., LIU G., KANG D., EKENGREN S., STEINER H. & HULTMARK D.A. 2000: Family of peptidoglycan recognition proteins in the fruit fly *Drosophila melanogaster*. — *Proc. Natn. Acad. Sci.* **97**: 13772–13777.
- YOSHIDA H., KINOSHITA K. & ASHIDA M. 1996: Purification of a peptidoglycan recognition protein from hemolymph of the silkworm, *Bombyx mori*. — *J. Biol. Chem.* **271**: 13854–13860.
- YU X., ZHU Y., MA C., FABRICK J.A. & KANOST M.R. 2002: Pattern recognition proteins in *Manduca sexta* plasma. — *Insect Biochem. Mol. Biol.* **32**: 1287–1293.
- ZAIDMAN-RÉMY A., HERVÉ M., POIDEVIN M., PILI-FOURRY S., KIM M., BLANOT D., OH B., UEDA R., MENGIN-LECREULX D. & LEMAITRE B. 2006: The *Drosophila* amidase PGRP-LB modulates the immune response to bacterial infection. — *Immunity* **24**: 463–473.

Received August 31, 2021; revised and accepted December 17, 2021
Published online January 28, 2022