Peptidoglycan recognition protein (PGRP) specifically binds to peptidoglycan and is considered to be one of the pattern recognition proteins in the innate immunity of insect. The PGRP is an essential component for peptidoglycan to trigger the prophenoloxidase cascade that is now recognized to be an important insect defense mechanism. We cloned cDNA encoding PGRP from the silkworm fat body cDNA library. Northern blot analysis showed that the PGRP gene is constitutively expressed in the fat body, epithelial cell, and hemocytes of naive silkworms. Furthermore, a bacterial challenge intensified the gene expression, with the maximal period being from 6 to 36 h after infection. The upstream sequence of the cloned PGRP gene was shown to contain putative cis-regulatory elements similar to the NF-κB-like element, interferon-response half-element, and GATA motif element, which have been found in the promoters of the acute phase protein genes of mammals and insects. A homology search revealed that the homologs of silkworm PGRP are present in mice, nematodes, and bacteria. This suggests that the recognition of peptidoglycan as foreign is effected in both vertebrates and invertebrates by PGRP homologs with an evolutionally common origin.

Innate immunity plays vital roles in primary defense mechanisms against invading pathogens in both vertebrates and invertebrates. The common structures among pathogens are recognized to be non-selves in the immunity. This type of recognition is termed pattern recognition, as opposed to clonal recognition in which clonally selected immunoglobulins are employed (1). Lipopolysaccharide (LPS) and peptidoglycan (PG) are the cell wall components of bacteria and β-1,3-glucan (βG) is that of fungi. They are recognized by pattern recognition proteins that are present in plasma as free-floating molecules or on the cell surface as receptors.

Among the pattern recognition proteins in mammals, a humoral LPS-binding protein and the cellular receptor CD14 have been well characterized (2–5). Their roles in stimulating macrophages to produce cytokines are well understood. In contrast, pattern recognition proteins for PG and βG have not been characterized as thoroughly as those for LPS, although CD14 is also implicated for the recognition of PG as non-self (6, 7).

In insects, a number of proteins that could be pattern recognition proteins have been described. They are lectins (8–11), hemolin (12, 13), lipopolysaccharide-binding protein (14, 15), Gram-negative bacteria-binding protein (16), peptidoglycan recognition protein (PGRP) (17), and β-1,3-glucan recognition protein (βGRP) (18, 19). The latter two proteins are members of the peprophenoloxidase cascade, of which other members are known to be serine protease zymogens and prophenoloxidase (20). They have been shown to have specific affinity to PG or βG and to work at initiation points of the cascade. The peprophenoloxidase cascade is now recognized to be one of major insect defense mechanisms and possibly play vital roles in interfering the mechanisms and in recognizing of microbes as foreign. Thus, it is probable that PGRP and βGRP of the peprophenoloxidase cascade are employed generally in insects as pattern recognition molecules for PG and βG. We previously reported the purification and the characterization of silkworm PGRP and βGRP (17, 18). Pattern recognition proteins other than PGRP and βGRP are not known to be distributed among insects as widely as PGRP and βGRP.

It is not clear whether pattern recognition proteins with the same specificity in mammals and in insects have any common structural similarity. However, parallels between cellular signaling pathways for the synthesis of mammalian acute phase proteins and insect immune proteins after a bacterial challenge have been revealed (21). This indicates a common origin of the pathways in innate immunity in mammals and insects. Here, we report the cloning of PGRP cDNA from the fat body cDNA library and PGRP gene from the genomic library of the silkworm Bombyx mori. The homology search showed that PGRP is a homologous protein to bacteriophage T7 lysozyme, although it does not have lysozyme activity, and that proteins homologous to PGRP are expressed in mammals. This experimental evidence suggests that recognition proteins for the initial extracellular non-self recognition in innate immunity of vertebrates and invertebrates have developed from a common origin. Our experimental results also show that PGRP synthesis is induced by a bacterial challenge and suggest that expression of the PGRP gene is regulated by the Rel family of transcription factors.

EXPERIMENTAL PROCEDURES

Insect and Bacterial Challenge—Silkworms, B. mori (strain Kinshu × Showa), were reared on an artificial diet as described previously (22). The larvae on day 5 of the fifth instar were injected with 10 μl of late logarithmic phase Enterobacter cloacae (JCM1232) suspension (ΔOD450 = 0.1) in the physiological saline (10 mM bis-Tris propane buffer, pH 6.5, containing 150 mM NaCl) or with 10 μl of the saline as the control experiment.
purification of amino acid residues.—Underlined amino acid residues were Conse-terminus of the mature protein, and the cDNA encoding PGRP from ments which obtained after proteolysis of the mature protein and the peptide frag-
firmed by sequencing of the N terminus of right numbered at 1,556, beginning at the N terminus of the mature protein and the peptide fragments which obtained after proteolysis of the S-pyridylethylated PGRP by trypsin.

Cloning and Sequencing of PGRP cDNA.—A silkworm fat body cDNA library was constructed in the vector λZAP (Stratagene). Using internal peptide sequences, degenerate oligonucleotides corresponding to KQWDG and WPWLE were synthesized. Their sequences in 5′-AAGAATTCC(A/G)AA(C/G)AA(C/G)AATGGGA(A/T)GG-3′ (sense primer) and 5′-AAGAATTCTCC(A/G)AA(C/G)ACACATCTGA(C/G)TGG-CCTGGGA(A/T)GG-3′ (antisense primer), respectively. These primers were used for the polymerase chain reaction. The reaction was performed using the silkworm fat body cDNA library as a template under the following conditions: 35 cycles comprising 94 °C for 1 min, 60 °C for 2 min, and 72 °C for 3 min. A 480-bp fragment was amplified, subcloned into a plasmid vector pBluescript (Stratagene), and sequenced using an automatic DNA sequencer (PE Applied Biosystems, model 377). The cloned polymerase chain reaction product labeled with [α-32P]dCTP was used to probe the λZAP fat body cDNA library. Hybridization was carried out at 42 °C for 16 h in 2 × PIPES buffer (0.8 M NaCl, 40 mM PIPES, pH 6.5), 50% formamide, 0.5% SDS, and 100 μg/ml denatured salmon sperm DNA. The membrane was washed twice in 0.1 × SSC (15 mM NaCl, 1.5 mM sodium citrate, pH 7.2) at 53 °C for 15 min and subjected to autoradiography. Three positive clones were obtained and sequenced. The clone that had the longest insert and a complete open reading frame was used as PGRP cDNA clone.

Northern Blot Analysis.—In the detection of PGRP transcript in naive silkworms, poly(A) RNA preparations from hemocytes, fat body, epidermal cells, midgut, silk gland, and malpighian tubules were prepared by chromatography on oligo(dT)-cellulose (Amersham Pharmacal Biotech). 5 μg of poly(A)+ RNA of each preparation were separated by electrophoresis in 1% agarose gels with 10 mM sodium phosphate, pH 7.0, transferred to a Hybond-N+ membrane, and hybridized with a 32P-labeled PGRP cDNA probe. The hybridization was performed for 16 h at 42 °C in 50% formamide, 4 × SSPE (1.2 M NaCl, 40 mM sodium phosphate, pH 7.4, and 4 mM EDTA), 5 × Denhardt’s (0.1% polyvinylpyrolidone, 0.1% bovine serum albumin, and 0.1% Ficoll), 0.1% SDS, and 100 μg/ml denatured salmon sperm DNA. The membrane was washed once in 2 × SSPE containing 0.1% SDS and subsequently twice in 1 × SSPE containing 0.1% SDS at 56 °C for 10 min. In the experiments to demonstrate the induction of PGRP gene in the fat body, total RNA from the fat body of silkworms was collected at intervals after the bacterial challenge. 20-μg aliquots of the RNA preparations were subjected to Northern blot analyses as above except that silkworm cecropin B (0.4 kb) (23) and α-tubulin (1.3 kb) probes were used in addition to the PGRP probe. The silkworm cecropin B (0.4 kb) and α-tubulin (1.3 kb) probes were obtained as polymerase chain reaction products, and the identities of their sequence to those having been reported were confirmed by referring to GenBank accession numbers S60579 and X83429, respectively.

RESULTS

Nucleotide Sequence of PGRP cDNA and the Deduced Amino Acid Sequence.—PGRP cDNA clones were obtained by screening a fat body cDNA library of the silkworm B. mori. The nucleotide sequence and the deduced amino acid sequence are shown in Fig. 1. The open reading frame composed of nucleotides 31–618 encodes 196 amino acid residues. The first 23 amino acid residues seemed to be a signal peptide because the deduced amino acid sequence beginning from the 24th residue was identical to that observed previously at the N terminus of PGRP. Thus, the mature protein consists of 173 residues with a calculated molecular mass of 19290 Da. 49% of the deduced mature protein sequence, including the N-terminal, was confirmed to be identical to that determined by direct sequencing of peptides obtained after trypsin digestion of PGRP (Fig. 1).
The PGRP gene was revealed to contain four exons (exon I, nucleotides 1–100; exon II, 101–299; exon III, 300–418; and exon IV, 419–1185). This observation supports further that our cloned PGRP gene encodes a protein with 196 amino acid residues. The deduced amino acid sequence of PGRP is consistent with our previous observation that purified PGRP did not manifest reactivity to eight commercially available lectins. We have produced recombinant PGRP in a baculovirus expression system and confirmed that the protein possesses the ability to specifically bind to peptidoglycan.

**Structure of PGRP Gene**—We isolated three positive clones from the silkworm genomic library. The restriction maps of the three clones indicated that a 2.1-kbp SacI fragment overlaps both the 2.5- and 4.3-kbp BamHI fragments (Fig. 3B). Both of the BamHI fragments and the 2.1-kbp SacI fragment were subcloned separately into pBluescript. The sequences of these fragments were examined after subcloning the various restriction fragments. By comparison of the nucleotide sequence of the PGRP gene with that of the PGRP cDNA clone, the PGRP gene was revealed to contain four exons (exon I, nucleotides 1–100; exon II, 101–299; exon III, 300–418; and exon IV, 419–739 by nucleotide numbering of the PGRP cDNA clone in Fig. 1) interspersed with three introns. The 5’ end of exon I is putative because the starting point of transcription of the PGRP gene has not been studied. The 2.5-kbp BamHI fragment contained 1966 bp of 5’ upstream sequence of the initiating Met codon (Fig. 3C). The TATA-box (TATATA) is located from nucleotide −73 to −68. Several putative cis-regulatory sequences were observed within −210 bp from the TATA-box. They were cAMP response motif (TGACGTG), NF-kB-like motif (AGGGATTTC), GATA motif (TGATTA), and five interferon response half-motifs (GAAANN) that have been found in promotor regions of many acute phase protein genes.

Southern blot analyses of genomic DNA digested with EcoRI, BamHI, or SacI were performed by using the PGRP cDNA probe. As is shown in Fig. 3A, only one hybridized band was observed with the digest by EcoRI or BamHI, and the digest with SacI gave two hybridized bands (2.1 and 6.0 kbp). Because the cloned PGRP gene had one restriction site for each of BamHI and SacI, the digest with BamHI should have given two hybridization bands (2.1 kbp in addition to the observed 6.0 kbp). The reason we could not observe the 2.1-kbp fragment seems to be that the fragment contains only a short exon and the PGRP cDNA probe barely hybridized with the fragment. Thus, the results of the Southern blot analyses indicate the presence of a single copy of the PGRP gene in the silkworm genome.

**Search for the Homologs of PGRP**—We searched for the existence of silkworm PGRP homologs in the database by using the PGRP cDNA as a probe detected a 0.8-kbp transcript of PGRP in hemocytes, fat body, and epidermal cells but not in malpighian tubules, silk gland, and midgut of naive silkworms. This result indicates that PGRP mRNA is constitutively expressed in the fat body, epidermal cells, and hemocytes. Major sites for the synthesis of PGRP seemed to be the fat body and epidermal cells (Fig. 2A). The expression of PGRP gene and cecropin B gene in fat body was induced by injection of bacteria (E. cloacae) to silkworm larvae, as shown in Fig. 2B. The induction kinetics was similar in both genes: the induction was detected at 6 h and reached its maximum at 24 h after injection. When saline was injected instead of bacteria in the above experiment, the induction kinetics were similar, but none of the transcripts of PGRP and cecropin B genes increased in the fat body for 36 h after the injection (data not shown).

**Northern blot analysis using PGRP cDNA as a probe detected a 0.8-kbp transcript of PGRP** in hemocytes, fat body, and epidermal cells but not in malpighian tubules, silk gland, and midgut of naive silkworms.

**Expression of PGRP mRNA—Northern blot analysis using PGRP cDNA as a probe detected a 0.8-kbp transcript of PGRP in**...
Considering the ability of PGRP to bind to peptidoglycan, the homology is understandable. At the same time, the present results concerning the PGRP sequence did not offer any clues to answering our question: What kind of activity does the complex of PGRP and peptidoglycan have? Because binding of PGRP to peptidoglycan is known to lead to the activation of protease zymogens of the prophenoloxidase cascade, we had speculated that PGRP itself is a protease zymogen of which activation is caused by its binding to peptidoglycan. However, the deduced PGRP sequence seemed to deny this possibility. We now speculate that the complex of peptidoglycan and PGRP activates a protease zymogen that is a member of the prophenoloxidase cascade. In the blood coagulation system of horseshoe crab, the factors that directly interact with elicitors such as lipopolysaccharide and \(-1,3\)-glucan are zymogen of proteases (26). In this respect, the component located at the initiation point of the prophenoloxidase cascade is different from that of the blood coagulation cascade.

Our search for proteins with the sequence similarity to PGRP indicated that homologs to PGRP are present in mice and nematodes. The mouse Tag7 protein has been reported to be a cytokine (27). It remains to be determined whether the silkworm PGRP has a similar biological activity. The ancestral protein of the homologs to PGRP seems to have appeared before protostomia and deuterostomia diverged in evolution and to be present widely in animals. This wide distribution of PGRP suggests the possibility that insect PGRP and its homologs may generally play a role in the recognition of peptidoglycan in innate immunity. The function of PGRP in triggering the prophenoloxidase cascade may be peculiar to insects. Peptidoglycan manifests various biological activities as have been mentioned in the introduction of our previous paper (17) and of papers by Dziarski and co-workers (6, 7). Despite the myriad activities, the mechanism for the recognition of peptidoglycan as foreign has been poorly understood. Recently, Dziarski and co-workers showed that CD14 on murine macrophages has affinity to peptidoglycan and is involved in activation of the transcription factor NF-kB by peptidoglycan (7). It has been established that CD14 is a receptor for LPS in mammals (3). In this case, however, another protein, the LPS-binding protein, in plasma is assumed to form a complex with LPS, and \(K_D\) of CD14 for the complexed LPS becomes very low in comparison.
with that for noncomplexed LPS. Thus, the LPS-binding protein increases the sensitivity of cells with CD14 to LPS. It would be worth testing whether PGRP works in the same way as the LPS-binding protein.

A single copy of the PGRP gene was indicated to be present in the silkworm genome (Fig. 3A). We found a CAMP response element, NF-κB-like element, and five interferon response half-elements in the promoter region of the gene. The transcripts of the gene were detected in the fifth instar larvae before challenging them with Gram-negative bacteria or peptidoglycan. However, the transcription was up-regulated after challenging them with the microbes or the cell wall component (Fig. 2). The presence of cis-regulatory elements like an NF-κB element and the enhancement of transcription by cell wall components of microbes are common features of acute phase proteins of insects and mammals (28, 29). In this respect, PGRP can be classified to be an acute phase immune protein. Whether up-regulated expression of the PGRP gene upon bacterial challenge has merits for silkworms and, if so, what those merits are remain to be studied.

The intracellular events leading to the induction of acute phase immune protein synthesis are now being intensely studied in both insects and mammals. Proteins belonging to the Rel family are commonly employed in the signaling pathway for the induction in different organisms (21, 30, 31). Toll has been shown to be a receptor for a protein named spätzle and to be involved in human innate immunity as well as adaptive immunity (33). All these observations point to a common origin of the intracellular signaling pathway. Our present results indicated the presence of homologous pattern recognition proteins for peptidoglycan in both insects and mammals. The results seem to imply that the mechanism for extracellular recognition of microbes as foreign is conserved between insects and mammals, indicating a yet closer relation between innate immune systems in such phylogenetically remote organisms as insects and mammals.

Addendum—After the submission of the manuscript of the present paper, a paper reporting a cDNA for PGRP from a moth Trichoplusia ni and its mouse and human homologs was published (34).

REFERENCES

1. Janeway, C. A., Jr. (1989) Cold Spring Harbor Symp. Quant. Biol. 54, 1–13
2. Schumann, R. R., Leung, S. R., Flaggs, G. W., Gray, P. W., Wright, S. D., Mathison, J. C., Tobias, P. S., and Ulevitch, R. J. (1990) Science 249, 1429–1431
3. Wright, S. D., Ramos, R. A., Tobias, P. S., Ulevitch, R. J., and Mathison, J. C. (1990) Science 249, 1431–1433
4. Gegner, J. A., Ulevitch, R. J., and Tobias, P. S. (1995) J. Biol. Chem. 270, 5320–5325
5. Viriyakosol, S., and Kirkland, T. N. (1995) J. Biol. Chem. 270, 361–368
6. Gupta, D., Kirkland, T. N., Viriyakosol, S., and Dziarski, R. (1996) J. Biol. Chem. 271, 23310–23316
7. Dziarski, R., Tapping, R. I., and Tobias, P. S. (1998) J. Biol. Chem. 273, 8680–8690
8. Komano, H., and Natori, S. (1985) Dev. Comp. Immunol. 9, 31–40
9. Pendland, J. C., and Boucias, D. G. (1986) Dev. Comp. Immunol. 10, 477–487
10. Amanai, K., Sakurai, S., and Ohtaki, T. (1990) Comp. Biochem. Physiol. 97B, 471–476
11. Kotani, E., Yamakawa, M., Iwamoto, S., Tashiro, M., Sumida, M., Matsubara, F., Tanai, K., Okuda, K. K., Kato, Y., and Mori, H. (1995) Biochim. Biophys. Acta 1269, 245–258
12. Sun, S.-C., Lindström, I., Roman, H. G., Faye, I., and Schmidt, O. (1990) Science 250, 1729–1732
13. Schmidt, O., Faye, I., Lindström-Dinnetz, I., and Sun, S.-C. (1993) Dev. Comp. Immunol. 17, 195–200
14. Komori, T., and Natori, S. (1991) J. Biol. Chem. 266, 13318–13323
15. Koizumi, N., Morozumi, A., Imamura, M., Tanaka, E., Iwahana, H., and Sato, R. (1997) Eur. J. Biochem. 248, 217–224
16. Lee, W.-J., Lee, J. D., Kravchenko, V. V., Ulevitch, R. J., and Brey, P. T. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 7888–7893
17. Yoshida, H., Kinoshita, K., and Ashida, M. (1996) J. Biol. Chem. 271, 13854–13860
18. Ochani, M., and Ashida, M. (1988) J. Biol. Chem. 263, 12056–12062
19. Söderhall, K., Røgn, W., Söderhäll, I., Newton, R. P., and Ratcliffe, N. A. (1988) Insect Biochem. 18, 323–330
20. Ashida, M., and Brey, P. T. (1997) in Molecular Mechanisms of Immune Responses in Insects (Brey, P. T., and Hultmark, D., eds) pp. 135–172, Chapman and Hall, London
21. Hultmark, D. (1994) Nature 367, 116–117
22. Yoshida, H., and Ashida, M. (1996) Insect Biochem. 26, 539–545
23. Kato, Y., Tanai, K., Hirochika, H., and Yamakawa, M. (1993) Insect Biochem. Mol. Biol. 23, 285–290
24. Engstrem, Y. (1997) in Molecular Mechanisms of Immune Responses in Insects (Brey, P. T., and Hultmark, D., eds) pp. 211–244, Chapman and Hall, London
25. Cheng, X., Zhang, X., Pflugrath, J. W., and Studier, F. W. (1994) Proc. Natl. Acad. Sci. U. S. A. 91, 4034–4038
26. Muta, T., and Iwagami, S. (1996) in Progress in Molecular and Subcellular Biology: Invertebrate Immunology (Rinkevich, B., and Muller, W. E. G., eds) Vol. 15, pp. 154–189, Springer-Verlag, Berlin
27. Kiselev, S. L., Kustikova, O. S., Korobko, E. V., Prokhortchouk, E. B., Kabish, A. A., Lukandin, E. M., and Georgiev, G. P. (1998) J. Biol. Chem. 273, 18633–18639
28. Kappler, C., Meister, M., Lagueux, M., Gateff, E., Hoffmann, J. A., and Reichhart, J. M. (1993) EMBO J. 12, 1561–1568
29. Engström, Y., Kadalayil, L., Sun, S. C., Samakovlis, C., Hultmark, D., and Faye, I. (1995) J. Mol. Biol. 232, 227–333
30. Baeuerle, P. A., and Henkel, T. (1994) Annu. Rev. Immunol. 12, 141–179
31. Siebenlist, U., Fränzle, G., and Brown, K. (1994) Annu. Rev. Cell Biol. 10, 405–455
32. Lemaitre, B., Nicolas, E., Michaut, L., Reichhart, J.-M., and Hoffmann, J. A. (1996) Cell 86, 973–983
33. Medahtive, R., Preston-Hurburt, P., and Janeway, C. A., Jr. (1997) Nature 388, 394–397
34. Kang, D., Liu, G., Lundström, A., Gelius, E., and Steiner, H. (1998) Proc. Natl. Acad. Sci. U. S. A. 95, 10078–10082