

A peptidoglycan recognition protein in innate immunity conserved from insects to humans

DAIWU KANG*, GANG LIU, ANNIKA LUNDSTRÖM, EVA GELIUS, AND HÅKAN STEINER†

Department of Microbiology, Stockholm University, S-106 91 Stockholm, Sweden

Communicated by John H. Law, University of Arizona, Tucson, AZ, June 23, 1998 (received for review April 30, 1998)

ABSTRACT Innate nonself recognition must rely on common structures of invading microbes. In a differential display screen for up-regulated immune genes in the moth *Trichoplusia ni* we have found mechanisms for recognition of bacterial cell wall fragments. One bacteria-induced gene encodes a protein that, after expression in the baculovirus system, was shown to be a peptidoglycan recognition protein (PGRP). It binds strongly to Gram-positive bacteria. We have also cloned the corresponding cDNA from mouse and human and shown this gene to be expressed in a variety of organs, notably organs of the immune system—i.e., bone marrow and spleen. In addition, purified recombinant murine PGRP was shown to possess peptidoglycan affinity. From our results and the sequence homology, we conclude that PGRP is a ubiquitous protein involved in innate immunity, conserved from insects to humans.

It has been postulated that primary immune recognition is based on structures common among invading pathogens. Jane-way termed this pattern-recognition (1). Bacterial surface molecules, such as lipopolysaccharide (LPS) and peptidoglycan (PG), are known to elicit immune reactions ranging from cytokine release to fever. The molecular mechanism behind LPS activation is rather well characterized and includes a humoral LPS-binding protein (2) and the cellular receptor CD14 (3). In contrast, less is known about peptidoglycan-mediated immune reactions, although CD14 has been implied to be part of this pathway as well (4, 5). In the innate immune system of insects, bacterial cell envelope substances activate both the phenol oxidase cascade and the genes for humoral antibacterial effector proteins. The latter activation is through pathways ending in Rel-mediated gene activation (6). However, the initial extracellular nonself recognition is poorly understood.

To find new components of the immune system in insects, we have used the differential display technique (7). Genes from the moth *Trichoplusia ni* that are induced by a bacterial challenge have been cloned (8, 9). Here we describe a peptidoglycan-binding protein that we have not only found in insects but also cloned from mouse and human. It is a strong link between vertebrate and invertebrate immunity, pointing at their common origin.

MATERIALS AND METHODS

Insects and Infections. *T. ni* was reared (10) and challenged with *Enterobacter cloacae* (11).

Differential Display and cDNA Cloning. Primer T₁₂VC was used for reverse transcription of total RNA from normal and challenged *T. ni* larvae. First-strand cDNA was PCR amplified with primers T₁₂VC and GATCACGTAC. PCR products were

separated on a 6% sequencing gel. One of the differentially expressed bands (VC16.1) was cloned, sequenced, and used as probe to screen a cDNA library from challenged *T. ni* larvae, as described (8). Mouse peptidoglycan recognition protein (PGRP) was reverse transcription-PCR cloned from mouse spleen total RNA by using primers TACACAGCCCT-GCGTCCT and GGTGCTGAGTCTTTATTG. Human spleen total RNA (CLONTECH) was reverse transcribed by using primer TCGAATTCCTCGAGAAGCT₁₈. The cDNA was amplified by using primers AAYCTNCARCAITAYCA and CCNATRAANGTNATNCC, cloned, and sequenced. From this sequence information primers ACATGGGGTTC-CATAAGT and AACTGGGGCTGGTGCGAC were made and used in 5'- and 3'-rapid amplification of cDNA ends (RACE) PCR, respectively, employing the human bone marrow Marathon-Ready cDNA kit (CLONTECH). The complete sequence information was obtained by using primers CCGGGCGCTCCTAGCGGTC and ACAGTGGGGTTTT-TGGCC in RACE PCR with the same cDNA kit.

Northern Blotting and Localization. Total RNA from different tissues of *T. ni* last-instar larvae and at different times after infection was separated on an agarose gel and blotted to a Hybond-N membrane (Amersham) according to standard protocols. Mouse multiple tissue Northern blots and human multiple tissue dot blots were purchased from CLONTECH.

Baculovirus Expression, Protein Expression and Purification, and Raising of Antiserum. To add six histidine residues to the C terminus of *T. ni* PGRP for baculovirus expression, cDNA clone pVC16.1.6.1b was amplified by PCR with primers CCGGATAACAATTTCACACAGGAA and CCGCTCGA-GTTAATGATGATGATGATGATGACCGTTATCTAGG-AAATGGTC. The PCR product was cleaved with *Eco*RI and *Xho*I and cloned into pBacPAK9 (CLONTECH). The construct was confirmed by DNA sequencing and recombinant baculovirus vBac. VC16His was obtained by using standard methods (12). Sf-9 cells were collected 48 h after infection, resuspended in His-Bind buffer (Novagen), sonicated three times for 20 s, and centrifuged at 15,000 × *g* for 20 min. The supernatant was recovered for protein purification. The recombinant PGRP with a C-terminal His tag was purified batchwise with a His-Bind resin, according to the manufacturer's recommendations (Novagen). The purified protein was separated by electrophoresis on an SDS/15% polyacrylamide gel. After staining, a band corresponding to molecular mass of 20 kDa was excised and homogenized in PBS. Immunization of rabbits was with three injections of 50 μg of protein each time. A 10,000-fold dilution of this antiserum was used for Western blotting experiments. Detection was with a goat anti-rabbit IgG conjugated to alkaline phosphatase (Sigma).

Abbreviation: PGRP, peptidoglycan recognition protein.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. AF076481, AF076482, and AF076483).

*Present address: Cell and Molecular Biology Department, Astra Draco AB, S-221 00 Lund, Sweden.

†To whom reprint requests should be addressed. e-mail: Hakan.Steiner@mibi.su.se.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

© 1998 by The National Academy of Sciences 0027-8424/98/9510078-5\$2.00/0
PNAS is available online at www.pnas.org.

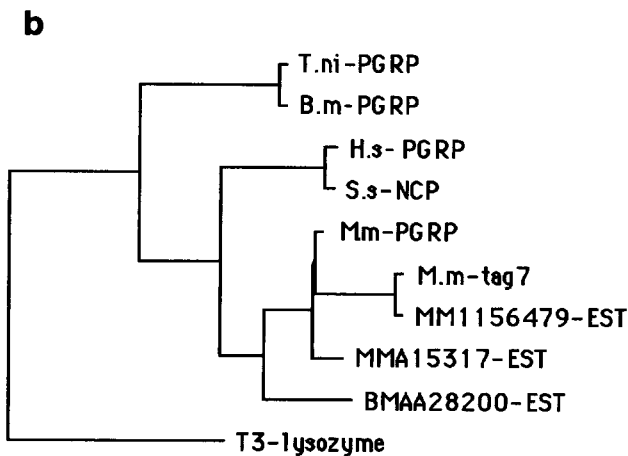
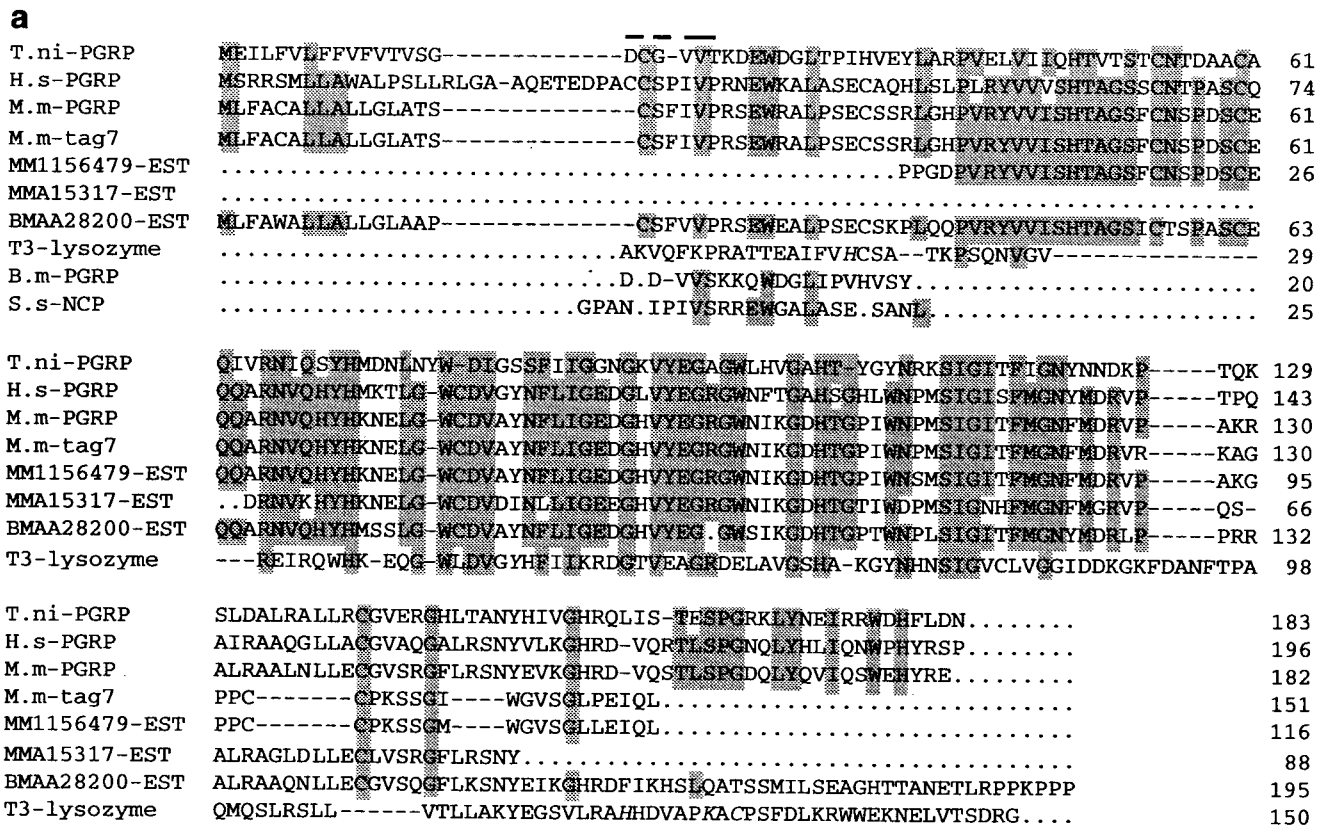


FIG. 1. (a) Deduced amino acid sequence of *T. ni* PGRP, murine *M.m*-PGRP, and human *H.s*-PGRP cDNA (GenBank accession nos. AF076481, AF076482, and AF076483). Homologs found by BLAST and FASTA database searches (29) were *M.m*-tag7 (X86374) from mouse; MM1156479 and MMA15317, both expressed sequence tags (ESTs) from mouse; BMAA28200 from the nematode *Brugia malayi*; T3 lysozyme (P20331) from phage T3; *B.m*-PGRP (1839611) from *Bombyx mori*; and *S.s*-NCP (P80552) from pig neutrophils. Determined or predicted (15) signal peptides are separated from the mature proteins by a gap (-). Missing sequence data are indicated by periods. Conserved residues in all but one of the given sequences (excluding that of T3 lysozyme) are printed on gray background. Residues obtained by amino acid sequencing of the baculovirus-expressed *T. ni* PGRP are overlined. The T3 lysozyme active site residues (His-17, Tyr-46, His-122, Lys-128, and Cys-130) are shown in italics. (b) Phylogenetic tree derived from the amino acid sequences by using the branch-and-bound algorithm of the PAUP program (20). Branch lengths are shown proportional to the minimal number of changes.

Murine histidine-tagged PGRP was produced similarly except that its cDNA was PCR amplified with the primers TACACAGCCCTGCTGCGTCCT and CCGCTCGAGTTAATGATGATGATGACCCTCTCGGTAGTGTTCCAG.

Protein Sequencing. Affinity-purified recombinant PGRP was also used for protein sequencing. After SDS/PAGE, the protein was blotted to a Immobilon P membrane (Millipore) and stained with Coomassie brilliant blue. The band corresponding to 20 kDa was cut out and sequenced on an Applied

Biosystems 476A, essentially according to the manufacturer's instructions.

Peptidoglycan and Bacteria Binding Assay. Insoluble peptidoglycan was prepared from *Micrococcus luteus* as described (13). The peptidoglycan binding assay was performed as described by Yoshida *et al.* (14), incubating 0.32 mg of peptidoglycan in 280 μ l of 10 mM maleate buffer, pH 6.5/0.15 M NaCl with 3-6 μ g of PGRP in 40 μ l of 1 M imidazole/0.5 M NaCl/20 mM Tris-HCl, pH 7.9, for 30 min at 4°C. One-sixteenth of the supernatant and of the pellet was removed for

analysis. The protein was separated from the peptidoglycan by boiling in 2× SDS/PAGE loading buffer, subjected to electrophoresis on an SDS/15% polyacrylamide gel, and stained with Coomassie brilliant blue.

Escherichia coli D22, *Bacillus megaterium*, and *Micrococcus luteus* were grown in Luria–Bertani (LB) medium. At OD₆₀₀ = 0.4, the bacteria (4 ml) were harvested, washed in PBS, and resuspended in 40 μl of PGRP (0.8 μg). After 30-min incubation at room temperature, the bacterial suspensions were centrifuged and the bacteria were washed and resuspended in 40 μl of PBS. Samples (1/4 of bacteria and supernatants) were subjected to SDS/PAGE and Western blot analysis using the polyclonal antiserum against PGRP as described. PGRP incubated in the absence of bacteria was used as a control throughout the experiment.

RESULTS AND DISCUSSION

From the differential display analysis of bacteria-induced genes in last instar larvae, one of the differentially amplified bands, VC16.1, was used to clone a full-length cDNA of 660 nucleotides (nt) from a library made from immunized *T. ni* larvae. The deduced amino acid sequence of the corresponding protein is given in Fig. 1*a*. In view of its activity (see below), we named it the peptidoglycan recognition protein (PGRP).

Excluding a putative signal sequence (15), it is a 19-kDa protein with a pI of 7.3, containing four cysteine residues.

The inducibility of PGRP was confirmed with Northern blot analysis of RNA samples from carefully matched immunized and naive larvae (Fig. 2*a*). A major 0.80-kb transcript and a minor 1.1-kb transcript are both strongly induced upon challenge from levels below detection in naive larvae. The PGRP expression kinetics is similar to that of other immune-induced genes in this species (9), with maximal expression at between 20 and 40 h after infection.

The tissue distribution of PGRP mRNA expression was also analyzed by using Northern blotting (Fig. 2*b*). Strong expression was observed in the fat body (an organ functionally equivalent to the mammalian liver) of induced insects. In addition, a weak signal was evident in hemocyte samples from induced insects. Neither the induced gut nor the tested organs of naive larvae contained detectable signal. Thus, the tissue distribution of PGRP mRNA expression is typical of an immune gene in insects (16).

To address the function of PGRP, we expressed a construct in Sf-9 insect cells by using a baculovirus vector. The construct was a fusion of the *T. ni* PGRP coding region and a polyhistidine coding tail (see *Materials and Methods*). After affinity purification of the recombinant protein, the suggested N-terminus was confirmed by five cycles of amino acid sequenc-

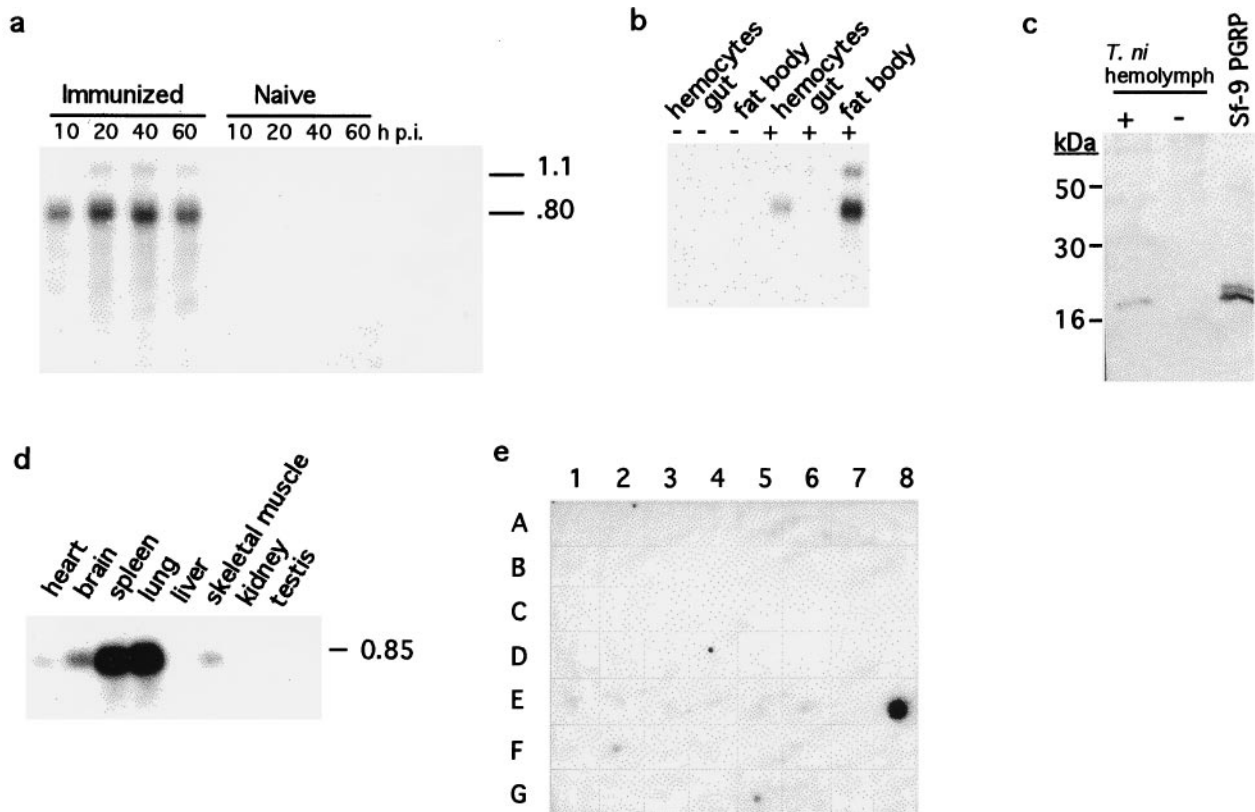


FIG. 2. Inducibility and tissue distribution of PGRP mRNA and protein. Plus and minus signs indicate samples from animals with and without bacterial challenge, respectively. For hybridization, radiolabeled PGRP-specific cDNA probes from the tested species were used. (a) Northern blot analysis of RNA from *T. ni* larvae. RNA was extracted at the indicated times after infection. Molecular sizes (in kb) of the two transcripts are shown on the right. (b) Northern blot analysis of RNA from fat body, gut, and hemocytes of last-instar larvae. (c) Western blot of *T. ni* hemolymph proteins from immunized and naive last-instar larvae. Purified recombinant *T. ni* PGRP from Sf-9 cells, which was used to raise rabbit antibodies, is included as reference. (d) Northern blot analysis of PGRP mRNA expression in various mouse tissues (CLONTECH). Molecular size (in kb) of the hybridizing material is shown on the right. (e) Dot-blot analysis of PGRP mRNA expression in human tissues (CLONTECH). Bullets indicate relative level of expression. A: 1, whole brain; 2, amygdala; 3, caudate nucleus; 4, cerebellum; 5, cerebral cortex; 6, frontal lobe; 7, hippocampus; and 8, medulla oblongata. B: 1, occipital lobe; 2, putamen; 3, substantia nigra; 4, temporal lobe; 5, thalamus; 6, subthalamic nucleus; and 7, spinal cord. C: 1, heart; 2, aorta; 3, skeletal muscle; 4, colon; 5, bladder; 6, uterus; 7, prostate; and 8, stomach. D: 1, testis; 2, ovary; 3, pancreas; 4, pituitary gland; 5, adrenal gland; 6, thyroid gland; 7, salivary gland; and 8, mammary gland. E: 1, kidney ●; 2, liver ●; 3, small intestine ●; 4, spleen ●; 5, thymus ●; 6, peripheral leukocyte ●; 7, lymph node; and 8, bone marrow ●●●. F: 1, appendix; 2, lung ●; 3, trachea; and 4, placenta. G: 1, fetal brain; 2, fetal heart; 3, fetal kidney; 4, fetal liver; 5, fetal spleen ●; 6, fetal thymus; and 7, fetal lung.

ing (residues overlined in Fig. 1*a*). For Western blots (Fig. 2*c*), a polyclonal antibody raised against this fusion protein was used. The strong gene activation observed upon immune stimulation thus results in PGRP protein synthesis, signal peptide cleavage, and secretion into the hemolymph.

In *Bombyx mori* a peptidoglycan-binding protein has been shown to be a mediator of peptidoglycan activation of the pro-phenol oxidase system (14). A short N-terminal sequence of this protein shares 60–65% sequence identity with the corresponding positions in mature *T. ni* PGRP. The band pattern in Fig. 3*a* shows that purified recombinant *T. ni* PGRP also binds to insoluble peptidoglycan from *Micrococcus luteus* *in vitro*. This finding led us to name the protein peptidoglycan recognition protein.

We could also demonstrate that PGRP binds to live bacteria. In Fig. 3*b*, PGRP is shown to bind strongly to *Micrococcus luteus*, less so to *Bacillus megaterium*, and not to *Escherichia*

coli. This pattern correlates well with the external exposure of peptidoglycan on these bacterial species.

T. ni PGRP also shares 28% sequence identity and 50% similarity with bacteriophage T3 lysozyme (an *N*-acetylmuramoyl-L-alanine amidase). Consequently, we tested the recombinant PGRP for amidase activity on *E. coli* cell walls (17), but we found none. We conclude that we have found the gene for a protein that can bind peptidoglycan without cleaving it. It is interesting that the peptidoglycan binding must also be the selection force behind this old connection to a bacteriophage lysozyme. However, the *T. ni* PGRP has lost all of the five suggested active-site residues present in the phage enzyme (17), in consonance with the observed lack of amidase activity.

The insect immune system has been regarded as a model for innate immunity. Peptidoglycan is a fundamental component of the bacterial cell wall and is thus a candidate for a pattern to be recognized by any immune system. A PGRP-like protein should therefore be expected to be ubiquitous. To extend our observations to vertebrates, we PCR-cloned PGRP homologs from both murine and human spleen cDNA as described in *Materials and Methods*. The deduced respective protein sequences (M.m-PGRP and H.s-PGRP in Fig. 1*a*) both share 43% sequence identity with *T. ni* PGRP, making them likely orthologs of this protein. Recently, a number of database sequence entries have corroborated the ubiquitous nature of PGRP. In addition to a mouse cDNA clone (M.m-tag7 in Fig. 1*a*) obtained from mammary gland adenocarcinoma cells (18), a number of expressed sequence tags (ESTs) from mouse and one from the human filarial nematode *Brugia malayi* (BMAA28200), as well as a short N-terminal protein sequence from a pig neutrophil protein (19), have been reported. All sequences show definite homology to PGRP.

To illustrate the relationship between these putative peptidoglycan-binding proteins, the most parsimonious phylogenetic tree (20) was constructed, using the entire sequences, with T3 lysozyme as an outgroup (Fig. 1*b*). Branch lengths are drawn proportional to the number of changes despite the different lengths of the sequences. The large divergence among the mouse PGRP sequences is probably artifactual and comes from uncertainties in the EST sequences. We have also noticed that by introducing shifts in reading frames of the M.m-tag7 sequence, it might be identical to, or an allelic variant of, the M.m-PGRP. The presented tree is in full agreement with established phylogeny, except for the *B. malayi* EST sequence, which forms a clade with the murine sequences. Though nematodes have been shown to possess an innate immune system (21), it is likely that this clone is actually derived from the rodent host (a jird) used to cultivate the nematode. These nematodes have mechanisms to acquire host proteins (22) and covalently incorporate them into microfilariae (23), but there are few examples of host–parasite horizontal gene transfer (24). In the present case, a transferred gene should have been more like the human natural host gene than a PGRP gene from the experimental rodent system.

The tissue distribution of human and mouse PGRP expression was analyzed (Fig. 2*d* and *e*). In the tested mouse organs, strong expression is seen in spleen and lung. In humans, bone marrow is the major site of expression, but weak expression is also seen in kidney, liver, small intestine, spleen, thymus, peripheral leukocyte, lung, and fetal spleen. The expression profile is consistent with an immunity function of PGRP also in mammals.

To demonstrate that mammalian PGRP actually has a function similar to that of insect PGRP, recombinant murine PGRP was produced in the baculovirus expression system as described for *T. ni* PGRP. Fig. 3*c* shows that the mammalian protein possesses a peptidoglycan affinity similar to that of insect PGRP, making them likely orthologs.

In vertebrates, peptidoglycan has been shown to bind directly to CD14 (4, 5), though a conflicting result has been

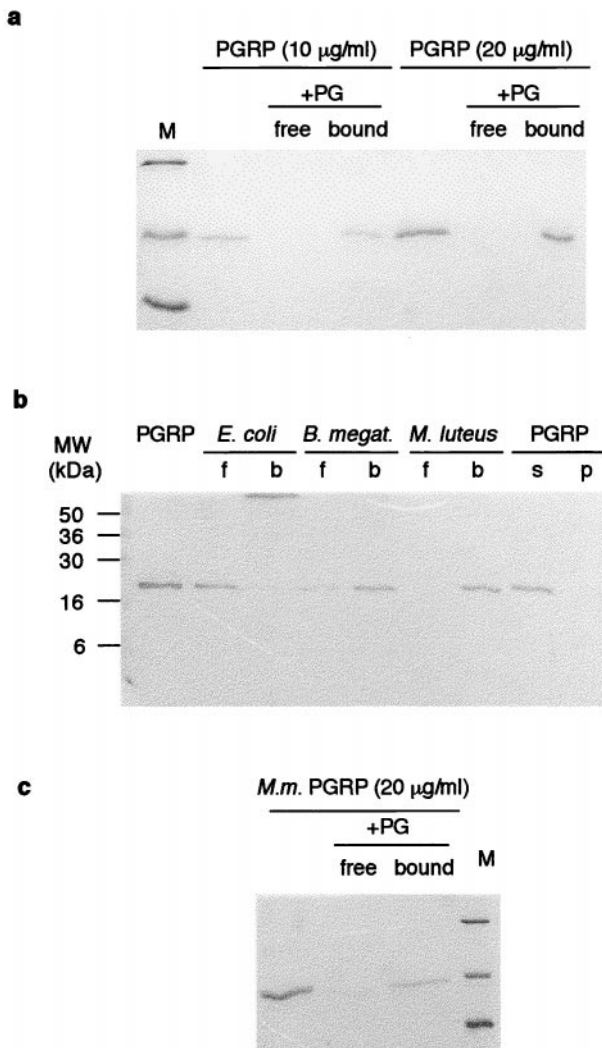


FIG. 3. Binding of PGRP to *Micrococcus luteus* peptidoglycan and to live bacteria. (a) Purified *T. ni* PGRP was incubated with peptidoglycan (PG) and free protein was isolated from bound protein as described in the text. The marker proteins have molecular masses of 31.0, 21.5, and 14.4 kDa. (b) Live *E. coli*, *B. megaterium*, and *M. luteus* bacteria were incubated with *T. ni* PGRP. Free (f) PGRP was separated from bound (b) by centrifugation. As a control the supernatant (s) and the “pellet” (p) of PGRP without added bacteria were analyzed similarly. Samples were run on SDS/PAGE and PGRP was visualized by Western blotting using anti-PGRP antibodies. (c) Binding of purified recombinant murine PGRP to peptidoglycan was tested as described for a.

reported (25). These results do not exclude the necessity of a humoral peptidoglycan adapter of the kind we report here; the need for serum (25) may indicate its presence. The intracellular events leading to immune gene activation in insects is under intense study (6, 26, 27). The peptidoglycan receptor role of CD14 in mammalian immunity could in *Drosophila* be played by Toll, as both contain leucine-rich extracellular domains. Recently, a human Toll homolog was shown to participate in activation of innate as well as the adaptive immunity (28), indicating that the entire signaling pathway is conserved between insects and mammals.

We are now in a position to test PGRP function by mutational analysis, as we have mapped the PGRP gene in *Drosophila* (unpublished work). The role of PGRP for the diverse effects of peptidoglycan in vertebrate immunity is also open for testing. An intriguing observation is that two of the PGRP homologs have been isolated after experimental nematode infections: in the jird with *Brugia malayi* and in the pig with *Ascaris suum* (19) to evoke allergic reactions (BMAA28200-EST and S.s-NCP in Fig. 1).

The peptidoglycan-binding protein described in this work is yet another strong link between invertebrate and vertebrate immunity, pointing at a common origin, comparable to lysozyme, Toll, and Rel proteins. PGRP not only is a strong link between vertebrate and invertebrate immunity but also could serve as a link between the pro-phenol oxidase cascade and humoral immunity in insects. Ashida and co-workers (14) have shown clearly that a PGRP protein mediates activation of the pro-phenol oxidase cascade in *Bombyx mori*. Here we have shown that a similar peptidoglycan-binding protein is synthesized from an "insect immune gene"—i.e., a gene up-regulated by a bacterial infection in *T. ni*. We have also demonstrated that a corresponding gene is part of the vertebrate immune system, a system devoid of phenol-oxidase-mediated immune mechanisms.

We thank K. Berzins for raising the rabbit antiserum and D. Hultmark for critically reading the manuscript. This work was supported by grants from the Swedish Natural Science Research Council and from the Carl Trygger Foundation.

- Janeway, C. A., Jr. (1989) *Cold Spring Harbor Symp. Quant. Biol.* **54**, 1–13.
- Schumann, R. R., Leong, S. R., Flaggs, G. W., Gray, P. W., Wright, S. D., Mathison, J. C., Tobias, P. S. & Ulevitch, R. J. (1990) *Science* **249**, 1429–1431.
- Wright, S. D., Ramos, R. A., Tobias, P. S., Ulevitch, R. J. & Mathison, J. C. (1990) *Science* **249**, 1431–1433.
- Gupta, D., Kirkland, T. N., Viriyakosol, S. & Dziarski, R. (1996) *J. Biol. Chem.* **271**, 23310–23316.
- Weidemann, B., Schletter, J., Dziarski, R., Kusumoto, S., Stelter, F., Rietschel, E. T., Flad, H. D. & Ulmer, A. J. (1997) *Infect. Immun.* **65**, 858–864.
- Hultmark, D. (1994) *Nature (London)* **367**, 116–117.
- Liang, P. & Pardee, A. B. (1992) *Science* **257**, 967–971.
- Kang, D. W., Liu, G., Gunne, H. & Steiner, H. (1996) *Insect Biochem. Mol. Biol.* **26**, 177–184.
- Kang, D. W., Lundström, A. & Steiner, H. (1996) *Gene* **174**, 245–249.
- Shorey, H. H. & Hale, R. L. (1965) *J. Econ. Entomol.* **58**, 522–524.
- Andersons, D., Gunne, H., Hellers, M., Johansson, H. & Steiner, H. (1990) *Insect Biochem.* **20**, 537–543.
- O'Reilly, D. R., Miller, L. K. & Luckow, V. A. (1992) *Baculovirus Expression Vectors* (Freeman, New York).
- Araki, Y., Nakatani, T., Nakayama, K. & Ito, E. (1972) *J. Biol. Chem.* **247**, 6312–6322.
- Yoshida, H., Kinoshita, K. & Ashida, M. (1996) *J. Biol. Chem.* **271**, 13854–13860.
- von Heijne, G. (1986) *Nucleic Acids Res.* **14**, 4683–4690.
- Samakovlis, C., Kimbrell, D. A., Kylvsten, P., Engström, Å. & Hultmark, D. (1990) *EMBO J.* **9**, 2969–2976.
- Cheng, X., Zhang, X., Pflugrath, J. W. & Studier, F. W. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 4034–4038.
- Kustikova, O. S., Kiselev, S. L., Borodulina, O. R., Senin, V. M., Afanas'eva, A. V. & Kabishev, A. A. (1996) *Genetika* **32**, 621–628.
- Fornhem, C., Peterson, C. G. B. & Alving, K. (1996) *Int. Arch. Allergy Immunol.* **110**, 132–142.
- Swafford, D. L. (1991) PAUP, Phylogenetic Analysis Using Parsimony, Version 3.0 (Illinois Natural History Survey, Champaign, IL).
- Kato, Y. & Komatsu, S. (1996) *J. Biol. Chem.* **271**, 30493–30498.
- Maizels, R. M., Denham, D. A. & Sutanto, I. (1985) *Mol. Biochem. Parasitol.* **17**, 277–288.
- Mehta, K., Chandrashekar, R. & Rao, U. R. (1996) *Mol. Biochem. Parasitol.* **76**, 105–114.
- Houck, M. A., Clark, J. B., Peterson, K. R. & Kidwell, M. G. (1991) *Science* **253**, 1125–1128.
- Heumann, D., Barras, C., Severin, A., Glauser, M. P. & Tomasz, A. (1994) *Infect. Immun.* **62**, 2715–2721.
- Dushay, M. S., Åsling, B. & Hultmark, D. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 10343–10347.
- Lemaitre, B., Nicolas, E., Michaut, L., Reichhart, J. M. & Hoffmann, J. A. (1996) *Cell* **86**, 973–983.
- Medzhitov, R., Preston-Hurlbert, P. & Janeway, C. A., Jr. (1997) *Nature (London)* **388**, 394–397.
- Devereux, J., Haeberli, P. & Smithies, O. (1984) *Nucleic Acids Res.* **12**, 387–395.