Structural Basis for Preferential Recognition of Diaminopimelic Acid-type Peptidoglycan by a Subset of Peptidoglycan Recognition Proteins*

Jae-Hong Lim†, Min-Sung Kim*, Han-Eol Kim†, Tamaki Yano†, Yoshiteru Oshima†, Kamna Aggarwal†, William E. Goldman†, Neal Silverman‡, Shoichiro Kurata§, and Byung-Ha Oh†,‡

From the †Center for Biomolecular Recognition and Division of Molecular and Life Science, Department of Life Science, Pohang University of Science and Technology, Pohang, Kyungbuk 790-784, Korea, §Graduate School of Pharmaceutical Sciences, Tohoku University, Aramaki, Aoba-ku, Sendai 980-8578, Japan, ¶Division of Infectious Disease, Department of Medicine, University of Massachusetts Medical School, Worcester, Massachusetts 01605, and the †Department of Microbiology, Washington University School of Medicine, St. Louis, Missouri 63110

Drosophila peptidoglycan recognition protein (PGRP)-LCx and -LCa are receptors that preferentially recognize meso-diaminopimelic acid (DAP)-type peptidoglycan (PGN) present in Gram-negative bacteria over lysine-type PGN of Gram-positive bacteria and initiate the IMD signaling pathway, whereas PGRP-LE plays a synergistic role in this process of innate immune defense. How these receptors can distinguish the two types of PGN remains unclear. Here the structure of the PGRP domain of Drosophila PGRP-LE in complex with tracheal cytotoxin (TCT), the monomeric DAP-type PGN, reveals a buried ionic interaction between the unique carboxyl group of DAP and a previously unrecognized arginine residue. This arginine is conserved in the known DAP-type PGN-interacting PGRPs and contributes significantly to the affinity of the protein for the ligand. Unexpectedly, TCT induces infinite head-to-tail dimerization of PGRP-LE, in which the disaccharide moiety, but not the peptide stem, of TCT is positioned at the dimer interface. A sequence comparison suggests that TCT induces heterodimerization of the ectodomains of PGRP-LCx and -LCa in a closely analogous manner to prime the IMD signaling pathway, except that the heterodimer formation is nonperpetuating.

Innate immune defenses against pathogens are initiated by pattern recognition receptors that bind conserved stereotypical, rather than particular, molecular structures present in a wide spectrum of microorganisms but absent in the host (1). A representative example of such structures is peptidoglycan (PGN), the major constituent of the cell wall of both Gram-positive and -negative bacteria. The peptidoglycan recognition protein (PGRP) family is a class of pattern recognition receptors that bind, and sometimes cleave, PGN. A total of 13 and four PGRP family members have been identified in Drosophila and humans, respectively (2–5). PGRPs are often characterized based on their polypeptide length. Short form PGRPs, such as PGRP-SA and -SD, contain a single PGRP domain (~180 amino acids) and, in most cases, a signal sequence, leading to the secretion of the proteins. Long form PGRPs, such as PGRP-LC and -LE, contain other domain(s) in addition to the PGRP domain, often including a transmembrane region (6). The PGRP domain is similar in structure to N-acetylmuramoyl-L-alanine amidases, such as T7 lysozyme, and some PGRPs are similarly catalytic, whereas others lack a critical cysteine residue in the catalytic triad (7). Thus, these PGRPs lack catalytic activity but function instead as pattern recognition receptors and/or as antimicrobials. For example, murine PGRP-S is directly antimicrobial and contributes to the neutrophil-mediated killing of bacteria (8, 9), whereas PGRP-SA, -SD, -LC, and -LE are key pattern recognition receptors involved in activation of the Drosophila immune response through the Toll or IMD (immune deficiency) signaling pathway (10–16).

In particular, genetic studies revealed that circulating PGRP-SA and PGRP-SD detect Gram-positive bacteria and activate the Toll pathway (17–19), whereas PGRP-LC and -LE are receptors for the IMD pathway that recognize Gram-negative bacteria (20–23). Analysis of mutant flies in response to bacterial infection and a knockdown experiment using RNA interference showed that PGRP-LC is the major receptor in the IMD pathway, and PGRP-LE acts in parallel and synergistically with PGRP-LC (24, 25). PGRP-LC, composed of the N-terminal cytoplasmic, single pass transmembrane and extracellular PGRP domains (2), is the bona fide cell surface receptor for the IMD pathway (20–22, 26, 27). PGRP-LC exists in three different isoforms as a result of alternative mRNA splicing: PGRP-LCa, -LCx, and -LCy (3). The cytoplasmic domains of the three isoforms are identical and physically interact with the death domain protein IMD to activate downstream signaling (26), whereas their ectodomains are unique, sharing less than 40% sequence identity with each other (3).

On the other hand, PGRP-LE is a two-domain protein composed of the N-terminal acidic domain with unknown function and the C-terminal PGRP domain, which preferentially binds Gram-negative bacterial PGN (23). PGRP-LE is devoid of a transmembrane domain and probably functions as an intracellular receptor for TCT as well as an extracellular factor capable of recognizing PGN (25). PGRP-LE is also required for activation of the prophenoloxidase cascade, the most immediate response to microbial infection or septic injury (25). The cascade leads to the production of melanin, which immobilizes the microorganism and contains the infection (6, 28).

---

* Supported by the Brain Korea 21 Project.
† To whom correspondence should be addressed. Tel.: 82-54-279-2289; Fax: 82-54-279-2199; E-mail: bhoh@postech.ac.kr.
‡ Supported by the Brain Korea 21 Project.
§ The abbreviations used are: PGN, peptidoglycan; DAP, meso-diaminopimelic acid; IMD, immune deficiency; ITC, isothermal titration calorimetry; MurNAc, N-acetylmuramic acid; MurNAc(anh), anhydro form of N-acetylmuramic acid; PGRP, peptidoglycan recognition protein; TCT, tracheal cytotoxin.

8 This work was supported by Creative Research Initiatives of the Korean Ministry of Science and Technology. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

9 The atomic coordinates and structure factors (code 2CB3) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

10 Supported by the Brain Korea 21 Project.
PGN is a polymer consisting of glycan chains of alternating N-acetylglucosamine (GlcNAc) and N-acetylmuramic acid (MurNAc) residues attached to peptide stems. To a varying degree, these stem peptides are also cross-linked to each other via a variety of peptide linkages. A critical distinction between Gram-negative bacterial or *Bacillus* PGNs and other Gram-positive bacterial PGNs is the presence of meso-diaminopimelic acid (DAP) in the third position of the peptide stem in the former instead of lysine found in the latter (29). DAP-type PGN is a potent activator of the IMD pathway, whereas Lys-type PGN is much less effective (24, 30, 31). Intriguingly, a monomeric DAP-type PGN, GlcNAc-MurNAc-anh-L-Ala-D-isoGlu-meso-DAP-d-Ala, also activates the IMD pathway potently as polymeric DAP-type PGN (24, 32). The monomeric PGN, known as tracheal cytotoxin (TCT), is continuously released from some Gram-negative bacteria, including *Bordetella pertussis*, in the course of turnover and recycling of the cell wall PGN (33). Recognition of the diffusible monomeric PGN should be advantageous in detecting Gram-negative pathogens, since their PGN layer is buried underneath the outer cell membrane in Gram-negative bacteria, whereas Gram-positive bacterial PGN is largely exposed. The selective activation of the either Toll or IMD immune signaling pathway relies on the ability of *Drosophila* to recognize specific forms of PGN (30). Recent studies demonstrated that PGRP-SA preferentially recognizes PGN from Gram-positive bacteria over DAP-type PGN (27), and the activation of the IMD pathway by TCT through PGRP-LE depends strongly on the presence of DAP in the peptide stem and weakly on the (1,6)-anhydro form of MurNAc (MurNAc(anh)) of the ligand (24, 32).

To date, the crystal structures of four different PGRPs have been reported: PGRP-LB (7), PGRP-SA (34, 35), the ectodomain of PGRP-LC (36), and the PGRP domain of human PGRP-La either in the unliganded form (37) or in complex with MurNAc-L-Ala-d-isoGlu-L-Lys (38). They provided important insights into the structural bases for the zinc-dependent amidase activity or the LD-carboxypeptidase activity contained in some members of the PGRP family and for the recognition of PGN in the binding groove. However, the mechanisms of discrimination for DAP-*versus* lysine-type PGN remain unknown, and the molecular mechanisms whereby PGN recognition by PGRPs leads signal transduction are unclear.

We determined the structure of the PGRP domain of *Drosophila* PGRP-LE in complex with TCT to 2.4 Å resolution. The presented structure reveals the first full picture of the binding of a complete monomeric PGN to a PGRP and provides a structural basis for the discrimination of PGN types by different PGRPs. The structure also suggests a mechanism of the IMD pathway activation by monomeric as well as polymeric DAP-type PGN through PGRP-LC.

**EXPERIMENTAL PROCEDURES**

*Plasmid Construction and Preparation of PGRP-LE—* The gene segment coding for the PGRP domain (residues 173–345) of *Drosophila* PGRP-LE was cloned into the pPROEx HTa vector (Invitrogen) by a standard polymerase chain reaction method, and the resulting recombinant protein was expressed in *Escherichia coli* BL21 (DE3) cells. Bacterial lysate was prepared by sonication in buffer A consisting of 20 mM Tris-HCl (pH 8.0), 100 mM NaCl, and 5 mM β-mercaptoethanol, and the supernatant was applied to a Ni²⁺-nitrilotriacetic acid column (Qiagen). The column was washed with buffer A containing 20 mM imidazole and subsequently washed with a 1:1 mixture of buffer A and 6 M guanidine hydrochloride (pH 8.0). This step for mild protein denaturation was essential to yield a homogeneous form of the protein in subsequent purification steps. The column was washed with buffer A for on-resin protein renaturation. The bound protein was eluted with buffer A containing 300 mM imidazole and then subjected to TEV protease cleavage to remove the N-terminal His₆ tag. The protein was further purified with a linear gradient of 0.1–1 M NaCl on a Hitrap Q column (Amersham Biosciences) and finally by gel filtration on a Superdex 75 column (Amersham Biosciences) equilibrated with 20 mM Tris-HCl (pH 7.0), 100 mM NaCl, and 3 mM dithiothreitol. All of the PGRP-LE mutants used in this study were created using the QuikChange kit (Strategene) and purified as described above except for the denaturation-renaturation step, which was dispensable due to their weaker affinity for PGN fragments or TCT.

**Crystallography and Structure Determination—* The PGRP domain of PGRP-LE was crystallized by the hanging drop vapor diffusion method by mixing and equilibrating 1.5 l of TCT dissolved in the reservoir solution at 50 mM with each of the protein solution (30–40 mg/ml) and a precipitant solution containing 10% polyethylene glycol 8000 and 100 mM HEPES (pH 7.0–7.6). Cubic crystals grew at 23 °C in 1–2 days. To obtain the crystals of the protein in complex with TCT, 0.2 μl of TCT dissolved in the reservoir solution at 50 mM was added to the hanging drops. Before data collection, the crystals were immersed briefly in a cryoprotectant solution, which was the reservoir solution plus 21% glycerol. A diffraction data set was collected on beamline 4A at the Pohang Accelerator Laboratory and processed using the programs DENZO and SCALEPACK (39). The structure was determined by the molecular replacement method with the CCP4 version of MolRep (40) using the structure of PGRP-LB (Protein Data Bank code 1OHT) as a search model. The solution consisted of four molecules in the asymmetric unit. The resulting model was refined in conjunction with model rebuilding using the programs CNS (41) and O (42), respectively. Crystallographic data statistics are summarized in Table 1.

**Preparation of TCT—* E. coli PGN was prepared essentially as described previously (30). In brief, *E. coli* cells were grown overnight at 37 °C in LB medium. The culture was rapidly chilled to 0–4 °C. Cells were collected and washed with a cold 0.85% NaCl solution and centrifuged again. The cells were rapidly suspended with vigorous stirring in 200 ml of hot (95–100 °C) 4% SDS solution for 30 min. After standing overnight at room temperature, the suspension was centrifuged for 30 min at 200,000 × g. The pellet was washed several times and resuspended with distilled water. For production of TCT, sodium acetate buffer (pH 4.5) was added to the resuspended solution, which contained polymeric PGN, to a final concentration of 0.3 M. After sonification, the preparation was treated with a transglycosylase SlY, purified in house as reported (43). After incubation for over 2 days at 37 °C with shaking, the reaction was stopped by adding 0.6% phosphoric acid. After centrifuga-

### TABLE 1

| Data collection and structure refinement |
|-----------------------------------------|
| **Refinement statistics**                |
| Space group                             | *P*2₁ |
| Unit cell parameters                     | *a = b = c = 217.53 Å |
| Resolution (Å)                          | 3.00–2.4 |
| Completeness                            | 92.9 (90.4) |
| Rmerge (%)                              | 8.9 (20.2) |
| # of refined atoms                      | 5483/210/268 |
| protein/water/TCT                       | 19.9/21.5 |
| Root mean square deviation bond length (Å) | 0.0060 |
| Root mean square deviation bond angle (degrees) | 1.3046 |
| **Ramachandran plot (%)**               |
| Most favored region                     | 89.9 |
| Additionally allowed region             | 9.8 |
| Generously allowed region               | 0.3 |

*The numbers in parenthesis are statistics from the highest resolution shell.*
tion, the supernatant was loaded on a Nucleosil C18 5 μm reverse-phase high pressure liquid chromatography column (4.6 × 250 mm; Alltech, France) and eluted with a linear gradient of 0–30% methanol in 50 mM sodium phosphate buffer (pH 4.5). The TCT-containing fractions were desalted with a Hi-Trap desalting column (Amersham Biosciences), and the identity of TCT was confirmed by mass spectrometric analysis.

Isothermal Titration Calorimetry—All measurements were carried out at 25 °C on a MicroCalorimetry System (Microcal Inc.). The proteins in this study were dialyzed against a buffer solution containing 20 mM Tris-HCl (pH 8.0), 100 mM NaCl, and 5 mM β-mercaptoethanol. TCT was dissolved in the same buffer. The samples were degassed for 20 min and centrifuged to remove any residuals prior to the measurements. Dilution enthalpies were determined in separate experiments (titrant into buffer) and subtracted from the enthalpies of the binding between the proteins and TCT. Data were analyzed using Origin software (Microcal).

Sequence Alignment—Sequences of the PGRP domains of 18 different PGRPs (Fig. 4c) were aligned using the program ClustalW (44). The accession numbers in the Swiss-Prot data base are as follows: D-LA (Q9VSV8), D-LB (Q9VGN3), D-LC (Q9GNK5), D-LD (Q9GN97), D-LE (Q9VXN9), D-LF (Q9VSW0), D-SA (Q9VYX7), D-SB1 (Q9VV97), D-SB2 (Q9VV96), D-SC1B (Q95SQ9), D-SC2 (Q9V4X2), D-SD (Q9VV97), H-L (Q96PD5), H-Iα (Q96LB9), H-Iβ (Q96LB8), and H-S (O75594), where “D” and “H” represent Drosophila melanogaster and human, respectively.

RESULTS
Structure Determination and Overall Structure—The PGRP domain (residues 173–345) of Drosophila PGRP-LE (hereafter referred to as PGRP-LE) was produced in E. coli. The monomeric form of PGRP-LE could be obtained, but nondiscrete oligomeric forms were overwhelming, which eluted as a broad peak from a gel filtration column. We suspected that the affinity of PGRP-LE for PGN is quite high, such that heterogeneously sized PGN fragments were co-purifying with the protein. In retrospect, TCT binding to the protein probably caused the oligomerization of the protein (see below). Partial denaturation fol-
lowed by renaturation of the protein as described under “Experimental Procedures” resulted in a dramatic increase in the monomeric form that migrated as a single band on a native gel and exhibited a typical circular dichroism spectrum for the α/β fold. Large crystals were obtained with this protein sample, but they diffracted only up to 6 Å. To obtain the crystals of the protein in complex with TCT, a 3-fold molar excess of TCT was soaked into the crystals. The TCT-soaking improved the x-ray diffraction from 6 to 2.4 Å. The crystal structure of the PGRP-LE/TCT complex was solved with molecular replacement method by using the structure of PGRP-LB (7) as the search model. The solvent content of the crystals was exceptionally high (78%), and the asymmetric unit contained four PGRP-LE molecules, all of which exhibited a nearly identical local environment. The structure of PGRP-LE consists of three α-helices and one central β-sheet composed of six β-strands.

Clear and strong electron density for the bound TCT molecule (Fig. 1, a and b), fully occupying the entire PGN-binding cleft, was observed at an initial stage of the structure refinement. The tetrapeptide stem of TCT lies in the cleft in an extended conformation (Fig. 1c). The Cα atom of meso-DAP is the point at which its long side chain and the fourth residue, Δ-Ala, bifurcate to span together the shallow and wide end of the PGN binding groove. The plane of MurNAc(anh) is perpendicular to the peptide stem and leans tightly against the deep binding pocket. The plane of GlcNAc is perpendicular to MurNAc(anh) and therefore is parallel with the peptide stem. The binding of TCT is unlikely to accompany a change in the polypeptide backbone conformation, because the shape of the ligand binding groove is essentially the same as that of isolated PGRP-LB (7).

Structural Basis for TCT-induced Formation of an Infinite Array of PGRP-LE Dimers—The addition of TCT caused PGRP-LE to form oligomers, whose average size increased proportionally with the concentration of TCT (Fig. 2a). Surprisingly, TCT is bound by two molecules of PGRP-LE in the crystal. Whereas one molecule of PGRP-LE (referred to as the first molecule) interacts with TCT at the binding groove as described above, another molecule (referred to as second molecule) interacts with the GlcNAc-MurNAc(anh) moieties of the ligand using one face of its α1 (Fig. 2b), whichotherwise is solvent-exposed. Therefore, each PGRP-LE molecule interacts with one TCT molecule in the PGN-binding groove and with the disaccharide unit of the other TCT molecule at the interface between two protein molecules, resulting in an infinite array of the dimer in a head-to-tail fashion in the crystal (Fig. 2b). In a macroscopic view, the PGRP-LE/TCT complex forms infinitely long oligomers (Fig. 2b). At the dimer interface, the protein–TCT interactions are mediated exclusively by three residues on α1 of the second molecule, including Glu231, which makes two direct hydrogen bonds with the GlcNAc moiety (Fig. 3a). Only the disaccharide of TCT, but not the peptide stem, interacts with the second molecule of PGRP-LE, which highlights the crucial role of the sugar moieties in the dimerization of the protein. The protein–protein interactions at the dimer interface, which are hydrophobic as well as hydrophilic, are mediated by four residues on α1 of the second molecule (Ala216, Val219, Arg223, Asp224) and three residues in another part of the second molecule (Fig. 3b). Of these, the side chain of Ser225 of the first molecule is involved in a buried hydrogen bond with the backbone carbonyl of Ala216 on α1 of the second molecule. This interaction mediates a tight local packing of the two subunits, together with the nearby hydrophobic interactions shown in Fig. 3b. As such, α1 makes a predominant contribution to the interactions at the dimer interface and thus appears as if it is a clamp that holds down the disaccharide unit of TCT to the docking groove of the first PGRP-LE molecule (Fig. 2b). Since the perpetuation of the dimer in the crystal is consistent with the TCT-induced aggregation of PGRP-LE in solution, we tested whether perturbation of the protein–TCT or protein–protein interactions at the dimer interface in the crystal affects the aggregation of PGRP-LE in solution. First, we generated a PGRP-LE mutant, containing leucine in place of Glu231, which is the key residue interacting with the sugar moiety of TCT (Fig. 3a). The mutant did not oligomerize in the presence of TCT (Fig. 3c). Second, we substituted Ser225 located at the dimer interface with glutamate. At this position, glutamate could not be modeled without a severe steric clash with the adjacent molecule in the presented structure and was expected to prevent the dimer formation. Consistently, the PGRP-LE(S232E) mutant remained as a monomer in solution in the presence of TCT (Fig. 3c). In this oligomerization-deficient mutant, the binding of TCT to the PGN-binding groove should be unaffected, since the serine residue on α1 is not a part of the groove. The property of the two mutants clearly indicates that the TCT-induced oligomerization of PGRP-LE in solution is through the head-to-tail dimer formation seen in the crystal structure of the protein in complex with TCT.

Drastically Enhanced Binding Affinity of PGRP-LE for TCT by Dimerization—The crystal packing interactions of TCT-unbound and TCT-bound PGRP-LE were virtually identical, since the space group and the cell parameters of the PGRP-LE crystals did not change by the TCT soaking. This observation indicated that PGRP-LE has a tendency to form the infinite array of the dimeric unit in the absence of TCT, but the intermolecular interactions are too weak, requiring an extremely high concentration of the protein for the oligomerization, e.g. the protein concentration employed for the crystallization (40 mg/ml). In order to learn whether TCT facilitates intersubunit interactions of PGRP-LE and thereby enhances its affinity for the protein, we analyzed the interactions of the wild-type and the PGRP-LE(S232E) mutant with TCT by isothermal titration calorimetry (ITC). The apparent dissociation constant (K_D) of the wild-type protein for TCT was 27 nM, whereas that of the oligomerization-deficient mutant was 820 nM (Fig. 3d). In the PGRP-LE(S232E) mutant, the interaction between TCT and the docking groove should be solely responsible for the binding between the two. Therefore, TCT binds to the docking groove with a low affinity, but the binding enhances its affinity for PGRP-LE drastically via inducing the large aggregates of the protein. In the crystal, the first molecule of PGRP-LE is in contact with another protein molecule in addition to the second molecule. The second molecule buries 650 Å² of accessible surface area of the first molecule at the dimer interface, whereas the third molecule buries 121 Å² of the area with poor electrostatic and shape complementarity (data not shown). Therefore, predominant intermolecular interactions in the crystalline oligomer are between the first and second molecules of PGRP-LE and TCT at the dimer interface. In solution, similar intermolecular interactions are likely to be facilitated by TCT. The drastic increase of the TCT binding affinity through the dimer formation (30-fold decrease in the K_D value) is ascribed to the α1 clamp providing not only the ligand–protein and protein–protein interactions but also an environment shielded from the bulk solvent in which the strength of hydrophilic interactions should increase in principle. By the α1 clamp, the MurNAc moiety of TCT is almost completely buried, indicating that polymeric DAP-type PGN cannot induce a similar dimer formation of PGRP-LE due to the presence of the β(1-4)-glycosidic bond between MurNAc and the next sugar residue, sterically crashing with the second molecule in such a complex. Instead, PGRP-LE would bind as a monomer to the repeating unit of the PGN polymer through the interactions between the docking groove and the ligand.

Signature Interactions of DAP-type PGN—TCT (molecular mass of 921 Da) interacts with numerous residues at the ligand-binding groove (Fig. 4a). Remarkably, the carboxylate end of meso-DAP,
which is the signature motif of Gram-negative PGN, is oriented inward facing the bottom of the docking groove and forms a bifurcate salt bridge with Arg254 buried almost entirely inside the protein (Fig. 4b). The importance of the “signature interaction” in the recognition of DAP-type PGN is clear from the sequence alignment (Fig. 4c) showing that the Arg254 residue is conserved in the PGRP molecules that are known to interact preferentially with Gram-negative bacterial PGN: PGRP-LCx (27), PGRP-LB (7), and human...
PGRP-S (45). In contrast, PGRP-SA, which shows specificity for Gram-positive bacterial PGN (27), contains threonine at the corresponding position (Fig. 4b). In order to assess the contribution of the signature interaction to the binding affinity for TCT, we generated a mutant, PGRP-LE(R254T), containing the substitution of Arg254 with threonine. The mutant interacted with TCT with a $K_D$ of 347 nM (Fig. 4d), which is a 13-fold increase from that of the wild-type PGRP-LE. This great reduction in the binding affinity confirms the critical role of the selectively conserved arginine residue in the binding/recognition of the DAP-type PGN. In order to evaluate the effect of the R254T mutation only on the TCT-binding to the docking groove, which is relevant to the binding of the protein to polymeric DAP-type PGN, the R254T mutation was introduced into the oligomerization-deficient PGRP-LE(S232E) mutant, and the binding between the double mutant and TCT was analyzed by ITC. However, the interaction between the two was too weak to produce a meaningful titration curve under the experimental conditions. An alternative, qualitative assay showed that the binding affinity of the double mutant, PGRP-LE(S232E/R254T), for E. coli polymeric PGN was markedly reduced compared with the PGRP-LE(S232E) mutant (Fig. 4d), underscoring the importance of Arg254 in the recognition of polymeric as well as monomeric DAP-type PGN.

The other signature motif of Gram-negative PGN is the internal (1,6)-anhydro bond of MurNAc. However, in contrast with the DAP residue, this anhydro form of MurNAc is present only in 5% of the glycan chains of natural Gram-negative bacterial PGN (46), whereas this sugar in the rest of the chains is linked to GlcNAc by the β-(1,4)-glycosidic bond. At the docking groove of the first molecule, O6, the bridging oxygen of the anhydro bond is within hydrogen-bonding distance of the guanidine group of Arg233, whereas the ether oxygen O1 of MurNAc makes a polar interaction with the plane of the guanidine group of the second molecule (Fig. 4b). Besides these favorable interactions, the (1,6)-anhydro bond, which restricts the conformational freedom of the MurNAc sugar ring, should provide a thermodynamic advantage in the binding of TCT to the protein. In order to see if the anhydro form may provide better protein-ligand interactions than the hydro form, we modeled the hydro form of MurNAc with C1 in the β-configuration. In this hydro form, Arg233 does not interact with the ether oxygen O1 or C6 hydroxyl group (data not shown), suggesting that the conservation of this arginine residue is advantageous for binding TCT.
FIGURE 4. TCT-binding groove of PGRP-LE and signature interactions. a, stereo view of the binding groove. The residues interacting with TCT (interatomic distance less than 3.9 Å) are shown in ball-and-stick representations. The dotted lines indicate hydrogen or ionic bonds. The labels in parentheses indicate the corresponding residues of PGRP-LCx. Only those different in the two proteins are shown. b, close-up views of two signature interactions. The top panel shows the bifurcate salt bridge between the carboxylate group of DAP and Arg254. The main chain -NH group of Asn236 also interacts with DAP. The bottom panel shows the polar interactions of the MurNAc(anh) with Arg233 of the first molecule and Arg223 of the second molecule. c, conservation of Arg254. Sequences of all of the known Drosophila and human PGRPs were aligned, and segments of 50 residues containing Arg254 of LE crystal structure were aligned. Only those different in the two proteins are shown. d, functional studies of the Arg254 mutants. Titrations with TCT indicated that the dissociation constant $K_D$ for R254T is $347 \pm 29.7 \text{ nM}$.
but would give no advantage or only a marginal advantage in the binding of polymeric DAP-type PGN. Of note, Arg223 of PGRP-LC is conserved in only a few PGRP molecules, including PGRP-LB, unlike Arg254, which is conserved in many PGRP molecules (Fig. 4c), the majority of which are expected to interact with DAP-type PGN preferentially.

A Model for the Activation of the IMD Pathway by DAP-type PGN—The activation of the IMD pathway by polymeric Gram-negative bacterial PGN requires PGRP-LCx only, whereas both PGRP-LCx and -LCa isoforms are required for the recognition of TCT (24, 27). The amino acid sequences of the PGRP domains of PGRP-LC and LCx are 47% identical, indicating that the structures of the two are highly similar to each other. Compared with the overall homology, the residues of PGRP-LC interacting with TCT using their side chains in the docking groove are even more conserved in PGRP-LCx (65% identity), including the key residue Arg254 involved in the interaction.

The residues of PGRP-LE interacting with TCT using their side chains in the docking groove are even more conserved in PGRP-LCx (65% identity), including the key residue Arg254 involved in the interaction. It seems to be crucial for binding TCT present at a low concentration but would give no advantage or only a marginal advantage in the binding of polymeric DAP-type PGN.

Identification of the bacterial species would be crucial for mounting the proper host immune responses. By combined structural, biochemical, and biophysical studies on the interaction between PGRP-LC and TCT, we addressed the important question of how PGRPs can discriminate Gram-positive versus Gram-negative bacterial PGN. The key chemical difference between the two types of PGN lies in the presence of DAP or lysine at the third residue position of the peptide stem. These two amino acids differ only by the presence or absence of a carboxyl group on their side chain. The subtle difference is recognized by an arginine residue, which is now found to be conserved in all of the known DAP-type PGN-interacting PGRPs. This residue, Arg254 in PGRP-LC, provides the guanidine group to charge balance the carboxylate group of the DAP residue at the bottom of the PGN-binding groove. When this prominent ionic interaction between the two was disrupted by a change of the arginine residue into threonine, the affinity for the protein–ligand interaction at the binding groove was severely affected. Therefore, Arg254 is the key determinant that allows the preferential recognition of DAP-type PGN over lysine-type PGN by PGRP-LC. Conservation of this arginine residue may be utilized as a convincing criterion, albeit inconclusive, for identifying PGRP family members that preferentially interact with DAP-type PGN. According to this criterion, human PGRP-L and PGRP-S are expected to interact with DAP-type PGN preferentially. Recently, a DAP-containing muramylpentapeptide was demonstrated to interact with human PGRP-S much more tightly than lysine-containing muramylpentapeptide, with a 103-fold difference in the KD values (45). However, the conservation of this arginine is not an absolute criterion, because PGRP-SD, which is known to respond to Gram-positive bacteria, has arginine at the corresponding position (Fig. 4c).

DISCUSSION

PGN is a primary pathogen-associated molecular pattern present in virtually all bacteria. With a repertoire of PGRP proteins, flies can detect and identify invading bacterial species based on selective recognition of the differences in the chemical composition of this cell wall component. Identification of the bacterial species would be crucial for mounting the proper host immune responses. By combined structural, biochemical, and biophysical studies on the interaction between PGRP-LC and TCT, we addressed the important question of how PGRPs can discriminate Gram-positive versus Gram-negative bacterial PGN. The key chemical difference between the two types of PGN lies in the presence of DAP or lysine at the third residue position of the peptide stem. These two amino acids differ only by the presence or absence of a carboxyl group on their side chain. The subtle difference is recognized by an arginine residue, which is now found to be conserved in all of the known DAP-type PGN-interacting PGRPs. This residue, Arg254 in PGRP-LC, provides the guanidine group to charge balance the carboxylate group of the DAP residue at the bottom of the PGN-binding groove. When this prominent ionic interaction between the two was disrupted by a change of the arginine residue into threonine, the affinity for the protein–ligand interaction at the binding groove was severely affected. Therefore, Arg254 is the key determinant that allows the preferential recognition of DAP-type PGN over lysine-type PGN by PGRP-LC. Conservation of this arginine residue may be utilized as a convincing criterion, albeit inconclusive, for identifying PGRP family members that preferentially interact with DAP-type PGN. According to this criterion, human PGRP-L and PGRP-S are expected to interact with DAP-type PGN preferentially. Recently, a DAP-containing muramylpentapeptide was demonstrated to interact with human PGRP-S much more tightly than lysine-containing muramylpentapeptide, with a 103-fold difference in the KD values (45). However, the conservation of this arginine is not an absolute criterion, because PGRP-SD, which is known to respond to Gram-positive bacteria, has arginine at the corresponding position (Fig. 4c).
ably, PGRP-SD may bind DAP-type PGN with higher affinity than it binds lysine-type PGN, but the subsequent signaling event may not take place. At least, the arginine residue, presumed to be slightly exposed at the bottom of the docking groove of the protein, would not interfere with the binding of lysine-type PGN. In Drosophila PGRP-SA, -LD, and -SB2 and human PGRP-1α and -β, the equivalent of Arg254 of PGRP-LE is substituted with a nonhomologous amino acid. These substitutions would not necessarily indicate that those PGRPs bind lysine-type PGN preferentially. Instead, they would mean that the unique carboxylate group of DAP will not contribute to the binding affinity of these proteins for DAP-type PGN. Consistently, human PGRP-1α, which contains valine at the position corresponding to Arg254 of -LE, exhibited virtually the same affinity for the two muramylpeptides that differ only by containing either lysine or DAP at the third position (45). This observation can be rationalized by the crystal structure of human PGRP-1α in complex with a lysine-containing muramyltripeptide, showing that the valine residue is buried inside the protein, and the side chain amino group of the lysine residue is not involved in an interaction with the protein (38).

It may be assumed that the side chain amino group of the lysine residue in the third position of the peptide stem is specifically recognized by some PGRPs that preferentially recognize lysine-type PGN over DAP-type PGN. However, the side chain amino group of the lysine residue is the point for cross-linking of two peptide stems through an amide bond in polymeric lysine-type PGN, exactly as the DAP residue in DAP-type polymeric PGN. Therefore, the lysine residue is unlikely to serve as the determinant for the discrimination. Given the diverse cross-linking patterns observed for Gram-positive bacterial PGNs (29), we speculate that the species-specific interpeptide bridges of lysine-type PGN may be the crucial determinant for preferential recognition by some PGRPs.

Apparently, PGRP-LC isoforms have evolved to recognize both polymeric and monomeric DAP-type PGN. The mode of TCT-binding to the docking groove of PGRP-LE suggests that many molecules of PGRP-LCx can bind to the array of monomeric PGN units on the cell wall of Gram-negative bacteria. The clustering of PGRP-LCx on the molecular pattern would result in close contact of the cytoplasmic domains of the receptor molecules and drive signal transduction, ultimately leading to the activation of transcription factor Relish. This signaling event may require damage on the cell wall to expose the otherwise hidden PGN layer. At an early stage of infection, without damage on the cell wall, TCT could be released from invading bacteria into the hemolymph as a result of metabolic recycling of the cell wall component. However, TCT would be present at an extremely low concentration, and furthermore it is devoid of a repetitive structure. We suggest that PGRP-LCx copes with the low concentration and the monomeric nature of TCT by recruiting PGRP-LCa that clamps down on the ligand and greatly enhances the ligand-binding affinity of the heterodimer compared with that of PGRP-LCx alone. Moreover, the induced ternary complex should juxtapose the identical cytoplasmic domains of these two receptor molecules and activate the cytoplasmic signaling.

In conclusion, we provided the complete view of the binding of the monomeric DAP-type PGN to the docking groove of PGRP-LE, in which two features are outstanding. First, the previously unrecognized arginine residue interacts with the unique carboxylate group of DAP, providing an explanation to the discrimination between DAP- and lysine-type PGN by PGRP-LE and -LC. Second, the four terminal groups of TCT, engaged in the glycosidic or peptide bond in the polymeric PGN, are sterically unobstructed and lie on the shallow ends of the docking groove. This binding mode provides a structural
Crystal Structure of PGRP-LE and Tracheal Cytotoxin Complex

basis for the clustering of PGRP-LCx by the polymeric PGN on the surface of invading bacteria. The unexpected TCT-induced head-to-tail dimer formation of PGRP-LE led us to suggest that the heterodimerization of PGRP-LCx and -LCa mediated by TCT would be closely similar to the dimerization of PGRP-LE/TCT complex, physically and thermodynamically.

Acknowledgments—We greatly appreciate the help of Dr. D. Mengin-Lecreulx in the preparation of TCT. This study made use of beamline 4A at the Pohang Accelerator Laboratory (Pohang, Korea).

REFERENCES

1. Medzhitov, R., and Janeway, C. A. J. (2002) Science 296, 298–300
2. Werner, T., Liu, G., Kang, D., Eken gren, S., Steiner, H., and Hultmark, D. (2000) Proc. Natl. Acad. Sci. U. S. A. 97, 13772–13777
3. Werner, T., Borge-Renberg, K., Mellroth, P., Steiner, H., and Hultmark, D. (2003) J. Biol. Chem. 278, 26319–26322
4. Liu, C., Xu, Z., Gupta, D., and Dziarski, R. (2004) J. Biol. Chem. 279, 34686–34694
5. Kang, D., Liu, G., Lundstrom, A., Gelius, E., and Steiner, H. (1998) Proc. Natl. Acad. Sci. U. S. A. 95, 10078–10082
6. Dziarski, R. (2004) Mol. Immunol. 40, 877–886
7. Kim, M. S., Byun, M., and Oh, B. H. (2003) J. Biol. Chem. 278, 689–697
8. Werner, T., Liu, G., Kang, D., Eken gren, S., Steiner, H., and Hultmark, D. (1999) Mol. Cell 4, 31873–31882
9. Guan, R., Malchiodi, E. L., Wang, Q., Schuck, P., and Mariuzza, R. A. (2004) J. Biol. Chem. 279, 31873–31882
10. Guan, R., Roychowdhury, A., Ember, B., Kumar, S., Boons, G. J., and Mariuzza, R. A. (2004) Proc. Natl. Acad. Sci. U. S. A. 101, 17168–17173
11. Otwinowski, Z., and Minor, W. (1997) Methods Enzymol. 276, 307–326
12. Vagin, A., and Teplyakov, A. (2000) Acta Crystallogr. Sect. D 56, 1622–1624
13. Brunger, A., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Pannu, N. S., Read, R. J., Rice, L. M., Simonov, T., and Warren, G. L. (1998) Acta Crystallogr. Sect. D 54, 905–921
14. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) Acta Crystallogr. Sect. A 47, 110–119