Peptidoglycan Recognition by Pal, an Outer Membrane Lipoprotein[†,‡]

Lisa M. Parsons, Florence Lin, and John Orban*

Center for Advanced Research in Biotechnology, University of Maryland Biotechnology Institute, 9600 Gudelsky Drive, Rockville Maryland 20850

Received November 1, 2005; Revised Manuscript Received December 23, 2005

ABSTRACT: Peptidoglycan-associated lipoprotein (Pal) is a potential vaccine candidate from *Haemophilus influenzae* that is highly conserved in Gram-negative bacteria and anchored to the outer membrane through an N-terminal lipid attachment. Pal stabilizes the outer membrane by providing a noncovalent link to the peptidoglycan (PG) layer through a periplasmic domain. Using NMR spectroscopy, we determined the three-dimensional structure of a complex between the periplasmic domain of Pal and a biosynthetic peptidoglycan precursor (PG-P), UDP-*N*-acetylmuramyl-L-Ala-α-D-Glu-*m*-Dap-D-Ala-D-Ala (*m*-Dap is *meso*-diaminopimelate). Pal has a binding pocket lined with conserved surface residues that interacts exclusively with the peptide portion of the ligand. The *m*-Dap residue, which is mainly found in the cell walls of Gram-negative bacteria, is sequestered in this pocket and plays an important role by forming hydrogen bond and hydrophobic contacts to Pal. The structure provides insight into the mode of cell wall recognition for a broad class of Gram-negative membrane proteins, including OmpA and MotB, which have peptidoglycan-binding domains homologous to that of Pal.

Gram-negative bacteria have a thin structural peptidoglycan (PG)[1] layer located in the periplasm between a lipopolysaccharide-containing outer membrane and an inner cytoplasmic membrane. PG is a polymer composed of *N*-acetylglucosamine (NAG) and *N*-acetylmuramic acid (NAM) connected in an alternating fashion through β-1,4-glycosidic bonds. Short peptides are attached to the NAM groups, and these can cross-link to form the cell wall matrix (*1*). A number of proteins interact noncovalently with the cell wall through a periplasmic PG-binding domain that is widespread in Gram-negative bacteria (Figure 1a). Members of this sequence family include the outer membrane proteins Pal and OmpA, the inner membrane flagellar motor protein MotB, and a number of other proteins with as yet unknown function. Pal is linked to the outer membrane through an N-terminal lipid group, while both OmpA and MotB have additional integral membrane domains.

Pal is part of a larger network of proteins encoded by the Pal/Tol operon, including TolA, TolB, TolQ, and TolR. TolA is anchored to the periplasmic side of the inner membrane by its N-terminal domain and forms a complex with the inner membrane-spanning proteins TolQ and TolR (*2*). TolB is a periplasmic protein that can bind both the C-terminal domain of TolA (*3*) and Pal (*4*), thereby forming a possible structural link between the inner and outer membranes. While this system has been studied extensively, the exact function of the Pal/Tol network remains unclear. Mutations in Pal lead to vesicle formation and a loss of outer membrane integrity (*5*, *6*). Further, mutations in any of the Tol/Pal genes affect the uptake of some organic compounds (*7*). Earlier work showed that the Tol proteins are involved in protein import as they have been parasitized to facilitate transport of bacteriophage and colicins (toxins excreted by competing bacteria) into the cell (*2*). More recently, it was demonstrated that Pal and TolA are required for proper surface expression of the cell surface O-antigen lipopolysaccharide in *Escherichia coli* (*8*).

Pal is one of the components released by *E. coli* during sepsis and may contribute to sepsis-induced inflammation (*9*). In addition, Pal plays an important role in the virulence of at least one species; it was shown that infecting human subjects with bacteria containing a mutant Pal in a human model of *Haemophilus ducreyi* infection affected the ability of the bacteria to progress to the pustular stage of disease (*10*). Humans can mount an immune response against Pal from *Haemophilus influenzae* (also known as P6), and the antibodies produced are bactericidal (*11*). Thus, a number of efforts have been made and are currently underway to utilize Pal to develop a vaccine against nontypeable *H. influenzae* infections (*12*−*14*).

While crystal structures have been reported for a truncated Pal from *E. coli* (PDB accession code 1oap) and for the OmpA-like motif of RmpM from *Neisseria meningitidis* (*15*), the mechanism by which these PG-binding domains recognize the cell wall has remained unknown. We describe here the solution NMR structure of the periplasmic domain of Pal from *H. influenzae* bound to the peptidoglycan precursor (PG-P), UDP-*N*-acetylmuramyl-L-Ala-α-D-Glu-*m*-Dap-D-

[†] Supported by NIH Grants GM57890 and 1S10RR15744 and the W. M. Keck Foundation. L.M.P. is supported by a PRAT Fellowship from NIGMS.

[‡] Structure coordinates for the Pal/PG-P complex are deposited in the PDB (accession code 2aiz). NMR assignments for Pal, free PG-P, and bound PG-P are deposited in BioMagResBank under the BMRB accession codes 6465, 6856, and 6858, respectively.

* Corresponding author. Phone, 240-314-6221; fax, 240-314-6255; e-mail, orban@umbi.umd.edu.

[1] Abbreviations: Pal, peptidoglycan-associated lipoprotein; PG, peptidoglycan; PG-P, peptidoglycan precursor; NAM, *N*-acetylmuramic acid; NAG, *N*-acetylglucosamine; OmpA, outer membrane protein A; *m*-Dap, *meso*-diaminopimelate; H-bond, hydrogen bond.

10.1021/bi052227i CCC: $33.50
Published on Web 01/31/2006

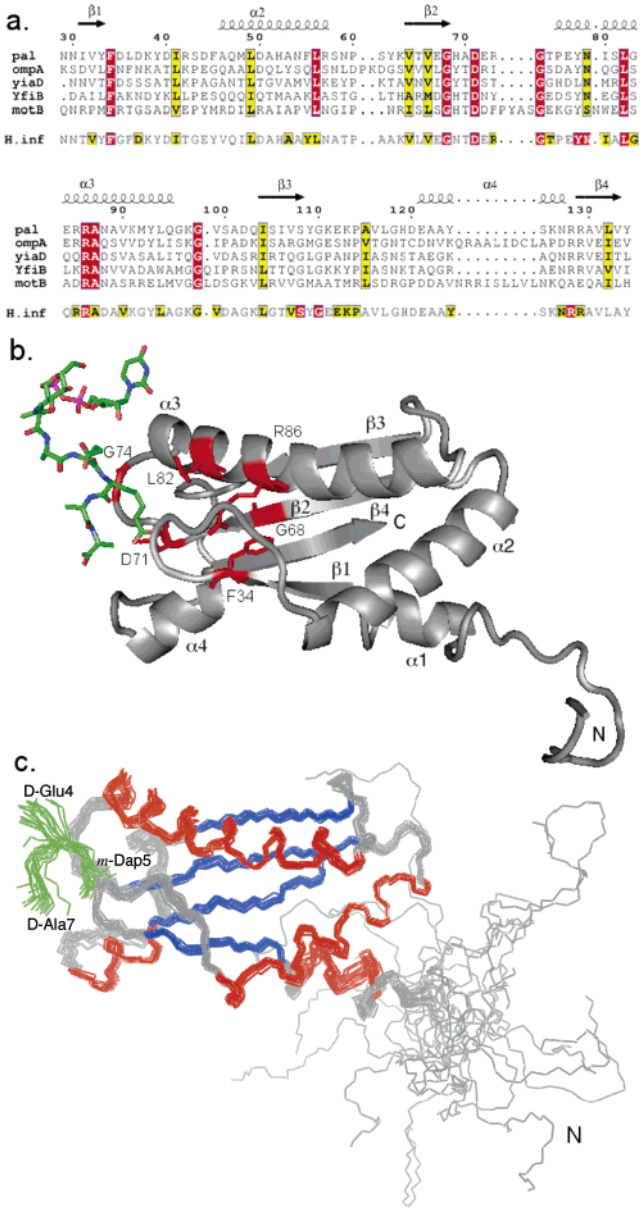

FIGURE 1: (a) Multiple alignment of Pal-like sequences. Top five sequences: protein sequence alignment of five proteins from *E. coli* containing the PG-binding domain motif. Residues in red are invariant, while residues in yellow are similar. Percent identity of each to Pal is OmpA 32%, YiaD 25%, YfiB 23%, and MotB 27%. Sixth sequence: sequence of *H. influenzae* Pal colored as above showing the conservation of Pal proteins aligned from 50 species. Gaps were added or moved to align this sequence with the *E. coli* protein sequences. The numbering and secondary structure is according to the mature *H. influenzae* protein. Alignment was done using ClustalW (*37*). (b) Solution structure of the Pal/PG-P NMR complex. Pal is shown in ribbon form (gray) with residues invariant in all six sequences labeled and highlighted (red). PG-P is shown in green. (c) Backbone superposition of the 20 lowest energy structures for the Pal/PG-P complex. β-Strands are highlighted in blue, α-helices are shown in red, loop or coil regions are gray, and residues 4−7 of PG-P are in green. The ensemble conformations for the *m*-Dap5 side chain are also shown in green. The orientation is the same as for panel b.

Ala-D-Ala (*m*-Dap is *meso*-diaminopimelate), a key intermediate of PG biosynthesis. This represents the first three-dimensional structure of such proteins complexed with a PG unit. The structure shows that Pal interacts exclusively with the peptide portion of the ligand. In particular, extensive H-bond and hydrophobic contacts to the *m*-Dap group, which is found in the cell walls of all Gram-negative bacteria, indicate that Pal specifically recognizes un-cross-linked Dap groups in the PG layer.

MATERIALS AND METHODS

Sample Preparation. The cloning and purification of the *H. influenzae* Pal/PG-P complex was carried out as follows. The Pal gene minus a 19 residue N-terminal signaling sequence was amplified from genomic DNA using PCR and placed into a pET-28a (Novagen) vector with a Tev-cleavable His_6-tagged GST fusion at the N-terminus. Protein was expressed in *E. coli* BL21 (DE3) cells grown in LB or labeled minimal media at 37 °C. Cells were harvested by centrifugation at 7000*g* for 10 min, resuspended in binding buffer (50 mM sodium phosphate, 300 mM sodium chloride, and 10 mM imidazole, pH 8.0), and lysed by sonication. The resulting lysate was centrifuged at 25 000*g* for 25 min. Pal/PG-P was purified by elution from a Ni−NTA column (Qiagen) using standard procedures. The His_6-tag was removed by incubating the purified Pal/PG-P sample (in 20 mM Tris, 100 mM NaCl, 1 mM DTT, and 0.5 mM EDTA, pH 8.0) with His_6-tagged Tev protease overnight at 30 °C. The GST tag and protease were then removed by passing the solution over a Ni−NTA column. Samples for NMR analysis were dialyzed into NMR buffer (50 mM sodium phosphate and 50 mM NaCl, pH 6.8) and concentrated to approximately 0.8 mM.

Samples of purified PG-P and peptidoglycan-free Pal were also used to assist in the NMR assignment process, and their preparation is described here. Supernatant containing $^{13}C/^{15}N$-labeled Pal with bound PG-P was loaded onto a Ni−NTA column following the manufacturer's guidelines for sample loading and washed with buffer (binding buffer with 20 mM imidazole). The column was washed with distilled water (6 bed volumes) to remove any salt, then placed in a 47 °C water bath, and equilibrated for 7−10 min. While still in the water bath, the column was washed with distilled water at 47 °C (6 bed volumes) to elute pure PG-P (ca. 0.5 mg/L of culture). Further washing of the column at room temperature with wash buffer (6 bed volumes) followed by elute buffer (binding buffer with 250 mM imidazole) gave peptidoglycan-free Pal. The His_6-tag was removed as described above. The identity of PG-P was confirmed by MALDI-TOF (Perseptive Biosystems) using a 2,5-dihydroxybenzolic acid (Aldrich) matrix. The mass spectrum was acquired in negative ion linear mode with an acceleration voltage of 20 kV and a delay time of 200 ns. Pure $^{13}C/^{15}N$-labeled PG-P was lyophilized and redissolved in NMR buffer either on its own or with unlabeled, peptidoglycan-free Pal.

Circular Dichroism. Circular dichroism measurements were performed using a Jasco spectropolarimeter, model J-720 and a water-jacketed quartz cell with a path length of 0.01 cm. The absorbance at 222 nm of 3 mg/mL of protein in NMR buffer was recorded as the temperature increased from 25 to 75 °C. Reversibility was confirmed by comparing full spectra between 200 and 250 nm taken at 25 °C before and after the melt.

NMR Spectroscopy. Spectra were collected at 298 K on a Bruker DRX-600 equipped with 3-axis gradient probes and recorded in States-TPPI mode. Pulsed-field gradients were used for coherence selection and solvent suppression. Data

were processed using nmrPipe (*16*) and analyzed with Sparky (T. D. Goddard and D. G. Kneller, UCSF). Backbone and side-chain assignments for Pal and bound PG-P were made using standard triple resonance experiments with a sample where both molecules were $^{13}C/^{15}N$-labeled (*17*). To assign free PG-P, a ^{15}N-labeled sample in H_2O was used to collect the two-dimensional (2D) ^{15}N-HSQC (*18*) and 2D HN−C′ plane of the three-dimensional (3D) HNCO (*19*). A $^{13}C/^{15}N$-labeled PG-P sample in D_2O was used to collect the 2D ^{13}C-HSQC and 2D TOCSY (*20*). Distance restraints in the Pal/PG-P complex were obtained from a 3D ^{15}N-edited NOESY (*21*) spectrum, 3D ^{13}C-edited NOESY (*22*) spectra in H_2O and D_2O, and a 2D NOESY (*23*) in D_2O. To assist in the assignment of intermolecular NOEs between ligand and protein, a 3D ^{13}C-edited NOESY was also acquired on a mixed sample ($^{13}C/^{15}N$-labeled PG-P with unlabeled Pal, 1:1). Mixing times for NOESY spectra were 120 ms and 150 ms. $^3J_{HNH}\alpha$ values were acquired from a 3D HNHA (*24*) experiment. Hydrogen−deuterium exchange data on the complex were obtained from a ^{15}N HSQC spectrum recorded after 17 min at 25 °C in D_2O. Residual dipolar coupling (RDC) data were collected for $N-H_N$ and $C'-N$ using the IPAP−HSQC (*25*) and quantitative HNCO (*26*) experiments on a sample with 18.7 mg/mL of Pf1 phage added (*27*). T_1 and T_2 measurements were made on the Pal/PG-P complex using the method of Farrow et al. (*28*).

Structure Calculations. Structures were calculated with CNS1.1 (*29*) using standard simulated annealing and torsion angle dynamics protocols. Prochiral groups were given floating stereospecific assignments until they could be unambiguously assigned from the structure. Initial NOE restraints were automatically assigned by NOEID (http://orban.umbi.umd.edu/NOEID) from a model based on the crystal structure of the truncated *E. coli* Pal protein (1oap). Distance restraints were based on peak intensities categorized as strong (1.8−3.0 Å), medium (1.8−4.0 Å), weak (2.7−5.3 Å), medium-weak (1.8−4.2 Å), medium-strong (1.8−3.5 Å), or very weak (3.0−6.0 Å). Dihedral restraints were obtained from TALOS (*30*) and $^3J_{HNH}\alpha$ values. H-bond restraints were 1.7−2.3 Å for r_{HN-O} and 2.5−3.1 Å for r_{N-O} and were assigned based on deuterium exchange data and calculated structures. Final force constants used were 1000 kcal mol^{-1} Å^2 for bond lengths, 500 kcal mol^{-1} rad^{-2} for angles and improper torsions, 40 kcal mol^{-1} Å^{-2} for experimental distance restraints, 200 kcal mol^{-1} rad^{-2} for dihedral restraints, and 4.0 kcal mol^{-1} Å^{-4} for the van der Waals repulsion term. A force constant of 0.3 kcal mol^{-1} Å^{-2} was used for the protein RDC restraints, while 0.5 kcal mol^{-1} Å^{-2} was used for the PG-P RDC restraints. The 20 best structures were chosen based on a low total energy, no NOE violations greater than 0.5 Å, no dihedral violations greater than 5°, and other diagnostic indicators of structure quality as summarized in Table 1. Structures were analyzed using PROCHECK (*31*) and displayed using PyMol (*32*).

RESULTS

Ligand Purification and Identification. While making the backbone NMR assignments of Pal (*17*), signals due to five sequentially connected residues not belonging to the protein were found in the 3D HNCACB and CBCACONH experiments. A literature search identified this sequence as being part of UDP-*N*-acetylmuramyl-pentapeptide, a biosynthetic

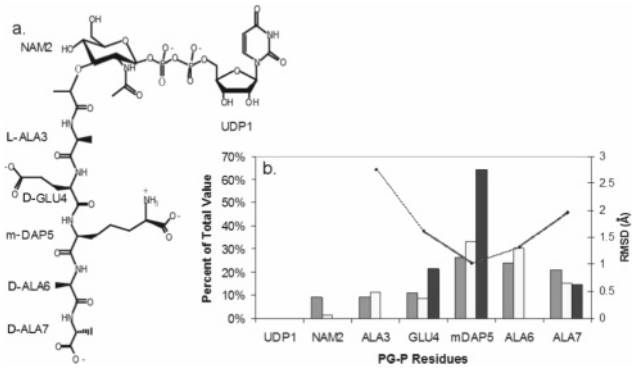

FIGURE 2: (a) Structure of PG-P. (b) Evidence supporting PG-P contact sites with Pal. Gray columns represent the ratio of ^{15}N-relaxation rates (R_2/R_1) in PG-P. White columns show the sum of the differences of N, H_N, C_α, H_α, C_β, H_β, and C′ chemical shifts between bound and unbound PG-P. To make them comparable to proton shifts, ^{15}N and ^{13}C shifts were adjusted by multiplying them by 0.154 and 0.276, respectively. Black columns represent the intermolecular NOE restraints per residue. For comparative purposes, values are shown as a percent of the total for each category (e.g., *m*-Dap has 18/28 = 64% of the total intermolecular restraints). The line plot shows the average backbone (C′, N, and C_α) rmsd values per residue.

peptidoglycan precursor (PG-P). The PG-P was purified from Pal for further characterization using a modified protocol that also enabled purification of peptidoglycan-free Pal without significant degradation of the protein. Previous methods to purify Pal from the mature PG layer required that the protein preparation be heated to 65 °C for 30 min then centrifuged (*33*). Here, Pal was heated at 47 °C while bound to an affinity column and washed with heated buffer. Pal melts reversibly at 47 °C, but about 15% of the protein does not refold correctly. Higher temperatures and longer exposure to heat noticeably decreased the amount of protein recovered from the column.

The assignments of UDP-NAM-L-Ala-γ-D-Glu-L-Lys-D-Ala-D-Ala from *Anabaena cylindrica*, a Gram-positive bacterium, have been reported previously (*34*), and these were used to identify the UDP-NAM part of the PG-P from the ^{13}C-HSQC spectrum. The pentapeptide portion of PG can differ between Gram-negative and Gram-positive bacteria in two ways. The second amino acid in the pentapeptide can be either γ-D-Glu or α-D-Glu. Also, the third amino acid is *m*-Dap in all Gram-negative bacteria and Gram-positive bacilli and L-Lys in most other Gram-positive bacteria (*35*). Since the backbone could be assigned sequentially by through bond connectivities to α-carbons, the second residue is unambiguously an α-D-Glu. Further, the molecular weight determined by mass spectrometry matches the calculated weight for PG-P with *m*-Dap instead of Lys (measured *m*/*z* 1200, calculated *m*/*z* 1202, for a ^{15}N-sample), consistent with expression in Gram-negative *E. coli*. NMR assignments of the third amino acid also indicated a Dap rather than Lys chemical shift pattern with the H_ϵ resonance at 4.23 ppm. On the basis of the combination of these data, the primary structure of the PG-P co-purified with Pal was identified as UDP-NAM-L-Ala-α-D-Glu-*m*-Dap-D-Ala-D-Ala (Figure 2a).

Solution Structure of Pal. Pal forms a monomeric α/β sandwich with the secondary elements arranged in the order $\alpha-\beta-\alpha-\beta-\alpha-\beta-\alpha-\beta$ (Figure 1b,c). The first three strands, ordered β1, β4, and β2, are antiparallel with β3 lying parallel to β2. Helices α1 and α4 are each seven-residues

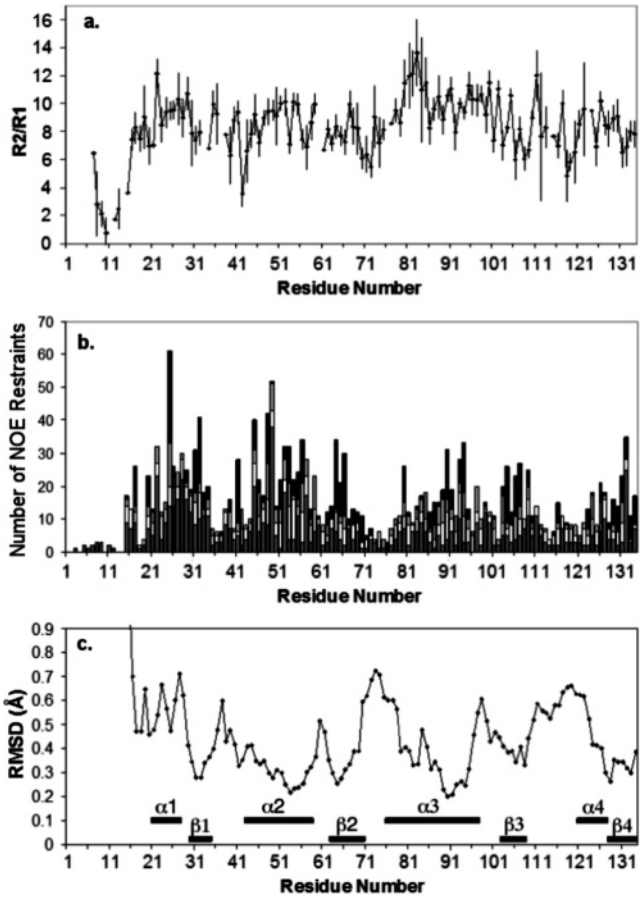

FIGURE 3: (a) Ratio of ^{15}N-relaxation rates (R_2/R_1) for backbone amides in Pal. Error bars represent ± 1 SD. (b) Number of NOE restraints per residue for Pal. Restraint types are shown from the bottom of each column as follows: intraresidue (gray), sequential (white), medium range (light gray), and long range (black). (c) Average rmsd per residue for the backbone N, C_α, and C′ atoms. The secondary structure is shown at the bottom of the plot.

long and located on the same side of the β-sheet at opposite ends of β1. The longer α2 and α3 helices angle across the β-sheet, α2 from the C-terminus of β1 to the N-terminus of β2 and α3 from the C-terminus of β3 to the N-terminus of β2. The first 15 residues are flexible and unstructured as evidenced by the low ^{15}N R_2 values and a lack of NOE restraints (Figure 3a,b). The β1−α2, β2−α3, and β3−α4 loops appear to be slightly more mobile than the regions of secondary structure based on decreased ^{15}N R_2 values in these regions (Figure 3a). Consistent with these data is the observation that some residues in these loops do not give strong NOE signals. Residues in the loop regions had fewer NOE restraints due to their surface exposure or difficulty in making assignments due to overlapping chemical shifts (Figure 3b). Because of this, loop residues, particularly in β2−α3 and β3−α4, have a higher rmsd than residues in α-helices or β-strands (Figure 3c). The two smaller helices (α1 and α4) have higher rmsd values than the two larger helices (α2 and α3) due to fewer NOE restraints and their surface exposure. In addition, α1 is in close proximity to the base of a large flexible tail that is probably contributing to its higher rmsd values. On average, 17.3 NOE restraints/residue were used for residues 16−134 (Figure 3b). The structure statistics are summarized in Table 1.

The solution NMR structure of the 134-residue periplasmic domain of *H. influenzae* Pal is similar to the crystal structure of a truncated 108-residue homologue from *E. coli* reported in the PDB (1oap). These two structures have a backbone rmsd of 1.2 Å and 68% sequence identity over residues 29−134 (residues 48−152 in *E. coli*). The main difference is that a larger segment of the N-terminal region after the natural signal sequence cleavage site is deleted in the *E. coli* structure, so the α1 helix is not present. This does not affect the overall fold, however. The solution structure also has similarities to the crystal structure of the OmpA-like domain of RmpM from *N. meningitidis* (1r1m, (*15*)) with an rmsd of 2.3 Å over residues 30−134 (residues 76−203 in *N. meningitidis*) and a sequence identity of 30%. The structure of RmpM also lacks the N-terminal α1 helix.

Table 1: Statistics for the Ensemble of 20 Structures

	Pal	PG-P
(A) Experimental Restraints		
total NOE	2056	67
intraresidue	793	25
sequential ($\|i-j\|=1$)	489	10
medium-range ($\|i-j\|\leq 4$)	318	4
long-range ($\|i-j\|>4$)	493	0
intermolecular		28
hydrogen bond	160	8
total dihedral angle restraints	200	5
φ	101	3
ψ	99	2
total RDCs	212	12
(B) Structure Statistics		
average rmsd (Å) to mean[a]	residues 16−134	residues 4−7
all non-hydrogen atoms	0.99 ± 0.05	1.92 ± 0.15
backbone atoms (N, Cα, C′)	0.47 ± 0.03	1.60 ± 0.10
Ramachandran distribution (residues 16−134)[b]		
most favored (%)	82.7 ± 2.7	
additionally allowed (%)	17.5 ± 2.9	
generously allowed (%)	0.5 ± 0.6	
overall *G*-factor[b,c]	0.1 ± 0.2	
number of bad contacts for residues 16−134[b]	6.9 ± 1.7	

[a] Statistics quoted are for the mean ± 1 SD. [b] Obtained using PROCHECK. [c] Defined as the log-odds score based on each residue's $\varphi-\psi$, $\chi1-\chi2$, and $\chi1$ values. Low (negative) values suggest unusual geometry.

Interaction between Pal and PG-P. The binding interface between Pal and PG-P was initially characterized by comparing the chemical shifts of free and bound PG-P. No shift changes were seen for the signals of UDP and muramic acid, while the resonances of the pentapeptide exhibited increasingly large differences centered on the *m*-Dap residue (Figure 2b). ^{15}N R_2 values were higher for *m*-Dap5, D-Ala6, and D-Ala7 than for other residues in the PG-P, suggesting that these were the most ordered residues in bound PG-P. The combination of these results indicated that residues 4−7 of the pentapeptide make the most contact with Pal. Likewise, Pal residues in the β1−α2, β2−α3, and β3−α4 loops experienced chemical shift and peak intensity changes when PG-P was removed (data not shown), indicating they were involved in the interaction with PG-P.

The detailed structure of the binding interface was determined by identification of 28 intermolecular NOEs between Pal and PG-P. The majority of these NOE contacts were from *m*-Dap5 to Pal, consistent with the chemical shift perturbation and relaxation data (Figure 2b). The *m*-Dap5 residue has extensive NOEs to Y78, F36, and L82 (Figure 4a). Additionally, D-Glu4 has NOEs to Y78 and D-Ala7 contacts F36. There were a total of 18 NOE restraints from *m*-Dap5 to Pal, 6 from D-Glu4, and 4 from D-Ala7. Upon

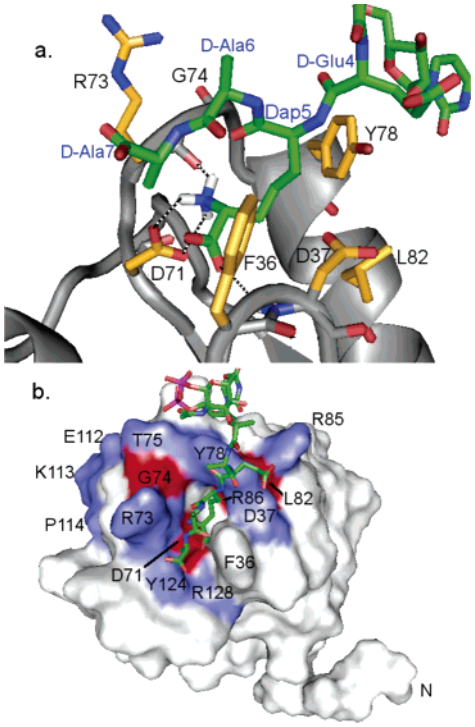

FIGURE 4: (a) Detailed view of the Pal/PG-P binding interface. Carbon atoms of Pal residues with NOEs or H-bonds to PG-P are shown as orange sticks with black labels. Carbon atoms of PG-P residues are in green with blue labels. The backbone of Pal is shown as a gray ribbon. H-bonds from *m*-Dap5 to Pal are represented by dashed lines. Oxygen atoms are red, nitrogen atoms are blue, and hydrogen atoms on the ϵ-ammonium group of *m*-Dap5 are white. (b) Surface representation of Pal with bound PG-P (green), demonstrating the conservation of residues around the binding site. The view is directly into the binding pocket (from the left side of Figure 1b). Residues colored in red are invariant in both Pal proteins and *E. coli* proteins containing the PG-binding motif. Residues in blue are conserved only in Pal proteins. The PG-P is oriented with the UDP/NAM groups at the top and D-Ala7 at the bottom of the figure.

application of the intermolecular NOE restraints in structure calculations, the side-chain oxygen atoms of *m*-Dap5 were consistently within H-bonding range of the main-chain amide group of the conserved D37 residue. The amide proton of D37 is in slow-exchange in D_2O, and as there are no other suitable acceptors within H-bonding distance in the structure, a hydrogen bond restraint was included between the D37 amide and the side-chain carboxyl group of *m*-Dap5 in the final stages of refinement. Similarly, the ϵ-ammonium group of *m*-Dap5 was consistently found to be within H-bonding distance of both the strictly conserved D71 side-chain carboxyl group and the backbone oxygen of conserved R73 (Figure 4a). This was supported by NOEs from the assigned *m*-Dap5 ϵ-ammonium group to the backbone amide protons of D71 and R73. Therefore, H-bond restraints were also included for these contacts in the later stages of structure refinement. It is possible that other H-bonds are formed, perhaps involving the side chain of R73. However, the NOE data were not sufficient to assign any further H-bond interactions between Pal and PG-P unambiguously.

The overall structure of the Pal/PG-P complex indicates that the PG-P binding site is located at a surface cavity in the Pal structure formed by the loop between β1 and α2, the loop between β2 and α3, and α4 (Figure 4b). The *m*-Dap5 residue of PG-P occupies this pocket in the complex and forms stabilizing interactions through H-bonding to its side chain ϵ-ammonium and ϵ-carboxylate moieties as well as through hydrophobic contacts to its side-chain methylene groups (Figure 4a). Residues D-Glu4 and D-Ala7 from PG-P also contribute to the binding interface with Pal through mostly hydrophobic interactions, but the majority of contacts come from *m*-Dap5.

DISCUSSION

The Pal family of proteins is widespread in Gram-negative bacteria with sequence homologues in more than 100 organisms. This sequence family includes other proteins such as the outer membrane protein, OmpA, and the inner membrane flagellar motor protein, MotB. Both contain two domains, one of which is a periplasmic PG-binding domain. Multiple alignment analysis indicates that six residues, F34, G68, D71, G74, L82, and R86, are invariant in the wider family of Pal-related proteins (Figure 1a). Three of these residues, D71, G74, and L82, are at or near the PG-P binding site (Figure 1b). The carboxylate group of D71 is involved in an H-bond interaction with the *m*-Dap5 side chain ammonium group, while L82 is one of several residues that have stabilizing hydrophobic contacts with PG-P (Figure 4a). As G74 does not interact directly with PG-P, its conservation may be due to steric reasons whereby substitution with any other amino acid would perhaps decrease the size of the binding pocket adversely. Residue R86 is buried in the hydrophobic core, extending from the middle of the α3 helix toward the PG-P binding surface, while neighboring F34 and G68 are also buried and likely to be important for maintaining the orientation of R86 (Figure 1b). Although the guanidinium group of R86 is close to the *m*-Dap5 carboxylate, it is not within H-bonding distance in any of the 20 lowest energy structures. It is noteworthy that the crystal structure of *E. coli* Pal reported in the PDB (1oap), which was determined in the absence of PG, has both the main-chain amide of D37 (numbered D55 in 1oap) and the side-chain guanidinium of R86 (numbered R104 in 1oap) H-bonded to a sulfate ion. The location of this sulfate ion is very similar to the position of the *m*-Dap5 ϵ-carboxylate group in the Pal/PG-P complex, which also H-bonds to the D37 backbone amide. In addition to these invariant amino acids, position 78 is always found to be either tyrosine or a hydrophobic residue.

A number of surface residues near the PG-P binding site are conserved in Pal sequences (Figure 1a, bottom sequence) but not in the broader Pal-related family (Figure 1a, top five sequences). These include D37 in the β1−α2 loop; R73 and T75 in the β2−α3 loop; Y78 and R85 from α3; E112, K113, and P114 in the β3−α4 loop; Y124 from α4; and R128 from the α4−β4 loop (colored blue in Figure 4b). At the PG-P binding site, only F36 is not conserved, suggesting that this residue is not critical for PG binding despite the observation of NOE contacts to the PG-P. The extent of surface conservation around the PG-P site present only in Pal sequences suggests that this particular family of proteins may have a further binding function in addition to PG recognition. This is consistent with results suggesting that another member of the Pal/Tol operon, TolB, interacts with Pal in a region that overlaps with the PG recognition site (*4*). The precise function of this Pal−TolB complex is not yet known.

The NMR structure of the Pal/PG-P complex demonstrates that Pal can bind to un-cross-linked PG. The structure provides further clues on how binding to cross-linked PG might also take place. Experimentally, it has been shown that about 30% of the peptides in the cell wall are cross-linked in *E. coli* (*36*). The *m*-Dap5 group is sequestered deep in the binding pocket and therefore cannot form a cross-link to another peptide. However, an extra pentapeptide fragment could be attached, through its *m*-Dap ϵ-amino group, to the solvent-accessible carbonyl of D-Ala6 without significantly altering the binding interface with Pal. In this model, D-Ala7 in the bound pentapeptide would be replaced with the second, cross-linking *m*-Dap moiety. Since D-Ala7 does not make extensive contacts to Pal, such cross-linking may not affect binding unfavorably. Thus, the structure presented here suggests that Pal may also be able to accommodate binding to cross-linked PG.

In conclusion, the first structure of a complex between Pal (or one of its related domains) and a peptidoglycan precursor is described here. The structure indicates that the majority of contacts are between conserved surface residues in Pal and the peptide region of the PG. In particular, the $\beta1-\alpha2$ and $\beta2-\alpha3$ loop regions of Pal provide a suitable binding pocket for the *m*-Dap residue from the PG-P, which has both stabilizing H-bond and hydrophobic interactions to the protein. This result is contrary to earlier suggestions from molecular docking to RmpM that PG might bind to Pal-like domains through the NAM or NAG carbohydrate groups (*15*). Pal homologues are highly conserved in Gram-negative bacteria, and so, it is noteworthy that the key interaction with PG is with an *m*-Dap residue, a component of the cell wall present in all Gram-negative bacteria. In contrast, most Gram-positive bacteria have L-Lys at the corresponding position. Since the Pal PG-binding domain is widely utilized in Gram-negative bacteria, the structure of the complex presented here will also be useful in understanding the mechanism of cell wall recognition by other proteins with PG-binding domains such as OmpA and MotB.

ACKNOWLEDGMENT

We thank Fenhong Song for assistance with mass spectrometry measurements and Alexander Grishaev and Ad Bax for advice with structure calculation.

REFERENCES

1. Park, J. T. (1996) The murein sacculus, in *Escherichia coli and Salmonella: Cellular and Molecular Biology*, ASM Press, Washington, DC, pp 48−57.
2. Guihard, G., Boulanger, P., Benedetti, H., Lloubes, R., Besnard, M., and Letellier, L. (1994) Colicin A and the Tol proteins involved in its translocation are preferentially located in the contact sites between the inner and outer membranes of *Escherichia coli* cells, *J. Biol. Chem. 269*, 5874−5880.
3. Walburger, A., Lazdunski, C., and Corda, Y. (2002) The Tol/Pal system function requires an interaction between the C-terminal domain of TolA and the N-terminal domain of TolB, *Mol. Microbiol. 44*, 695−708.
4. Ray, M. C., Germon, P., Vianney, A., Portalier, R., and Lazzaroni, J. C. (2000) Identification by genetic suppression of *Escherichia coli* TolB residues important for TolB−Pal interaction, *J. Bacteriol. 182*, 821−824.
5. Bernadac, A., Gavioli, M., Lazzaroni, J. C., Raina, S., and Lloubes, R. (1998) *Escherichia coli* tol-pal mutants form outer membrane vesicles, *J. Bacteriol. 180*, 4872−4878.
6. Cascales, E., Bernadac, A., Gavioli, M., Lazzaroni, J. C., and Lloubes, R. (2002) Pal lipoprotein of *Escherichia coli* plays a major role in outer membrane integrity, *J. Bacteriol. 184*, 754−759.
7. Llamas, M. A., Rodriguez-Herva, J. J., Hancock, R. E., Bitter, W., Tommassen, J., and Ramos, J. L. (2003) Role of *Pseudomonas putida* tol-oprL gene products in uptake of solutes through the cytoplasmic membrane, *J. Bacteriol. 185*, 4707−4716.
8. Vines, E. D., Marolda, C. L., Balachandran, A., and Valvano, M. A. (2005) Defective O-antigen polymerization in tolA and pal mutants of *Escherichia coli* in response to extracytoplasmic stress, *J. Bacteriol. 187*, 3359−3368.
9. Liang, M. D., Bagchi, A., Warren, H. S., Tehan, M. M., Trigilio, J. A., Beasley-Topliffe, L. K., Tesini, B. L., Lazzaroni, J. C., Fenton, M. J., and Hellman, J. (2005) Bacterial peptidoglycan-associated lipoprotein: a naturally occurring toll-like receptor 2 agonist that is shed into serum and has synergy with lipopolysaccharide, *J. Infect. Dis. 191*, 939−948.
10. Fortney, K. R., Young, R. S., Bauer, M. E., Katz, B. P., Hood, A. F., Munson, R. S., Jr., and Spinola, S. M. (2000) Expression of peptidoglycan-associated lipoprotein is required for virulence in the human model of *Haemophilus ducreyi* infection, *Infect. Immun. 68*, 6441−6448.
11. Murphy, T. F., Bartos, L. C., Rice, P. A., Nelson, M. B., Dudas, K. C., and Apicella, M. A. (1986) Identification of a 16,600-dalton outer membrane protein on nontypeable *Haemophilus influenzae* as a target for human serum bactericidal antibody, *J. Clin. Invest. 78*, 1020−1027.
12. Poolman, J. T., Bakaletz, L., Cripps, A., Denoel, P. A., Forsgren, A., Kyd, J., and Lobet, Y. (2000) Developing a nontypeable *Haemophilus influenzae* (NTHi) vaccine, *Vaccine 19* (Suppl. 1), S108−115.
13. Wu, T., Chen, J., Murphy, T. F., Green, B. A., and Gu, X. X. (2005) Investigation of nontypeable *Haemophilus influenzae* outer membrane protein P6 as a new carrier for lipooligosaccharide conjugate vaccines, *Vaccine 23*, 5177−5185.
14. Berenson, C. S., Murphy, T. F., Wrona, C. T., and Sethi, S. (2005) Outer membrane protein P6 of nontypeable *Haemophilus influenzae* is a potent and selective inducer of human macrophage proinflammatory cytokines, *Infect. Immun. 73*, 2728−2735.
15. Grizot, S. and Buchanan, S. K. (2004) Structure of the OmpA-like domain of RmpM from *Neisseria meningitidis*, *Mol. Microbiol. 51*, 1027−1037.
16. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes, *J. Biomol. NMR 6*, 277−293.
17. Parsons, L. M., and Orban, J. (2005) NMR assignment of the periplasmic domain of peptidoglycan-associated lipoprotein (Pal) from *Haemophilus influenzae*, *J. Biomol. NMR 32*, 93.
18. Mori, S., Abeygunawardana, C., Johnson, M. O., and van Zijl, P. C. (1995) Improved sensitivity of HSQC spectra of exchanging protons at short interscan delays using a new fast HSQC (FHSQC) detection scheme that avoids water saturation [published erratum appears in (1996) *J. Magn. Reson., Ser. B 110* (3), 321], *J. Magn. Reson., Ser. B 108*, 94−98.
19. Grzesiek, S., and Bax, A. (1992) Improved 3D triple-resonance NMR techniques applied to a 31 kDa protein, *J. Magn. Reson. 96*, 432−440.
20. Braunschweiler, L., and Ernst, R. R. (1983) Coherence transfer by isotropic mixing: application to proton correlation spectroscopy, *J. Magn. Reson. 53*, 521−528.
21. Fesik, S. W., and Zuiderweg, E. R. P. (1988) Heteronuclear three-dimensional NMR spectroscopy: a strategy for the simplification of homonuclear two-dimensional NMR spectra, *J. Magn. Reson. 78*, 588−593.
22. Ikura, M., Kay, L. E., Tschudin, R., and Bax, A. (1990) Three-dimensional NOESY−HMQC spectroscopy of a ^{13}C-labelled protein, *J. Magn. Reson. 86*, 204−209.
23. Kumar, A., Ernst, R. R., and Wuthrich, K. (1980) A two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete proton−proton cross-relaxation networks in biological macromolecules, *Biochem. Biophys. Res. Commun. 95*, 1−6.
24. Vuister, G. W., and Bax, A. (1993) Quantitative J correlation: a new approach for measuring homonuclear three-bond J(H^NH^α) coupling constants in ^{15}N-enriched proteins, *J. Am. Chem. Soc. 115*, 7772−7777.

25. Ottiger, M., Delaglio, F., and Bax, A. (1998) Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra, *J. Magn. Reson. 131*, 373−378.
26. Jaroniec, C. P., Ulmer, T. S., and Bax, A. (2004) Quantitative J correlation methods for the accurate measurement of 13C′-13Calpha dipolar couplings in proteins, *J. Biomol. NMR 30*, 181−194.
27. Hansen, M. R., Mueller, L., and Pardi, A. (1998) Tunable alignment of macromolecules by filamentous phage yields dipolar coupling interactions, *Nat. Struct. Biol. 5*, 1065−1074.
28. Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by ^{15}N NMR relaxation, *Biochemistry 33*, 5984−6003.
29. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse, K. R., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 54*, 905−921.
30. Cornilescu, G., Delaglio, F., and Bax, A. (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology, *J. Biomol. NMR 13*, 289−302.
31. Laskowski, R. A., Rullmann, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR, *J. Biomol. NMR 8*, 477−486.
32. Delano, W. L. (2002) The PyMOL molecular graphics system, DeLano Scientific, San Carlos, CA.
33. Karalus, R. J., and Murphy, T. F. (1999) Purification and characterization of outer membrane protein P6, a vaccine antigen of non-typeable *Haemophilus influenzae*, *FEMS Immunol. Med. Microbiol. 26*, 159−166.
34. Kodani, S., Ishida, K., and Murakami, M. (1999) Occurrence and identification of UDP-*N*-acetylmuramyl-pentapeptide from the cyanobacterium *Anabaena cylindrica*, *FEMS Microbiol. Lett. 176*, 321−325.
35. Dziarski, R. (2004) Peptidoglycan recognition proteins (PGRPs), *Mol. Immunol. 40*, 877−886.
36. Caparros, M., Pisabarro, A. G., and de Pedro, M. A. (1992) Effect of D-amino acids on structure and synthesis of peptidoglycan in *Escherichia coli*, *J. Bacteriol. 174*, 5549−5559.
37. Thompson, J. D., Higgins, D. G., and 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.

BI052227I