PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

Structure of YraM, a protein essential for growth of *Haemophilus influenzae*

J. Vijayalakshmi,[1] Brian J. Akerley,[2] and Mark A. Saper[1,3]*

[1] Biophysics, University of Michigan, Ann Arbor, Michigan 48109-1055

[2] Department of Molecular Genetics and Microbiology, University of Massachusetts Medical School, Worcester, Massachusetts 01655

[3] Department of Biological Chemistry, University of Michigan, Ann Arbor, Michigan 48109-1055

ABSTRACT

Nontypeable* Haemophilus influenzae *is an obligate human parasite that often causes middle ear infections in children and exacerbates chronic obstructive pulmonary disorder, the fourth leading cause of death in the United States. There are no effective vaccines available for this strain. The lipoprotein YraM (gene HI1655) was identified as essential for the growth and viability of* H. influenzae *but its function is unknown. Sequence comparisons showed that YraM is a fusion of two protein modules. We grew crystals of the carboxyl-terminal module of YraM comprising residues 257–573 (YraM-C), phased the diffraction data by the multiwavelength anomalous diffraction technique, and refined the model to a crystallographic R-factor of 0.16 (R_{free} = 0.19) with data to 1.35 Å resolution. The two-domain structure of YraM-C adopts a fold similar to that observed for the open, unliganded forms of several periplasmic binding proteins (PBPs) involved in bacterial active transport. Sequence alignments of YraM homologues from other Gram-negative species showed that the most conserved residues of YraM-C cluster between the two domains in the location where other PBPs bind their cognate ligand. Modeling of YraM-C into a closed conformation similar to the leucine-bound form of the Leu/Ile/Val-binding protein (LIVBP) shows a putative binding pocket larger than the leucine-binding site in LIVBP. The pocket has both polar and nonpolar surfaces, with the latter located in the same area where a leucine side chain binds to LIVBP. We discuss possible biological functions of YraM considering its predicted location in the outer membrane, a novel place for such a binding protein.

Proteins 2008; 73:204–217.

***Key words:* X-ray crystallography; ABC transporter; lipoprotein; LppC; HI1655; chronic obstructive pulmonary disorder; otitis media; tetracopeptide repeat.**

INTRODUCTION

Chronic obstructive pulmonary disorder (COPD) is one of the major diseases that afflict smokers and the fourth leading cause of death in the United States and Europe.[1] Infection with a nontypeable strain of *Haemophilus influenzae* (NTHi) often exacerbates COPD. Furthermore, the NTHi strain is responsible for ~30% of acute otitis media cases (middle ear infections), which afflict greater than 75% of children at some time during childhood. NTHi are Gram-negative bacteria with an outer membrane containing short lipopolysaccharides, but lacking the extracellular capsule found in the more virulent type b strain of *H. influenzae*. As obligate parasites of humans, NTHi bacteria commonly inhabit the respiratory tracts of healthy individuals and have no identified growth environment outside of the host. Because strains of *H. influenzae* have relatively small genomes (about one-third the size of the *E. coli* K12 genome), and lack genes essential for heme biosynthesis, they must transport heme and NAD from the host.

Increasing resistance to antimicrobials as well as the lack of an NTHi vaccine are stimulating the effort to find new drug targets or antigens suitable for a potential vaccine. Preferably, target antigens should be surface-exposed and functionally essential to minimize antigenic drift. After screening the genome of the nonencapsulated *H. influenzae* Rd strain by *in vitro* transposon mutagenesis, Akerley *et al.* identified 478 open reading frames, representing 38% of the genome, that were critical for growth or viability on rich medium.[2] Of these genes, 259 were annotated as encoding proteins of unknown function of which 159 were also found in other bacterial species. Such genes may encode proteins important for basic cellular metabolism.

One of the essential genes is *yraM* (HI1655) that encodes a 575-residue secreted lipoprotein of unknown function. The

Grant sponsor: NIH NIAID; Grant number: 1RO1-AI49437; Grant sponsor: American Heart Association; Grant number: 023508N; Grant sponsor: Philip Morris USA Inc., Philip Morris International.
J. Vijayalakshmi's current address is College of Pharmacy, University of Michigan, Ann Arbor, MI 48109-1065.
*Correspondence to: Mark A. Saper, Biophysics, University of Michigan, 930 N University Ave, Ann Arbor, MI 48109-1055. E-mail: saper@umich.edu.
Received 18 June 2007; Revised 18 January 2008; Accepted 8 February 2008
Published online 15 April 2008 in Wiley InterScience (www.interscience.wiley.com).
DOI: 10.1002/prot.22033

PFAM[3] 22.0 database entry PF04348 aligns YraM homologues from 116 distinct species of proteobacteria. The homologues are 20–62% identical to full-length YraM at the protein level. Following a hydrophobic signal sequence is the characteristic motif of -LAGCS- found in bacterial lipoproteins.[4] Lipoproteins of this type are translocated into the periplasm, modified with diacylglycerol on the conserved cysteine thiol (Cys26 in YraM), and proteolyzed by signal peptidase II just before the modified cysteine. The new N-terminus is then acylated to tether the protein to the membrane. In YraM, the serine following Cys26 directs the protein for translocation to the outer membrane.[5] Membrane fractionation and LC/MS analysis of the *E. coli* proteome showed that YraM is expressed and localizes to the outer membrane.[6] Although we detect *H. influenzae* YraM in membrane fractions (unpublished data), precisely which membrane has yet to be confirmed.

The growth dependency of *H. influenzae* Rd on *yraM* was verified by controlled expression of the gene under the control of the xylose promoter.[4] Substitution of Gly for Asp522, an invariant residue in all YraM homologues, renders the resulting *H. influenzae* mutant strain temperature sensitive. This temperature sensitive mutation is recessive, indicating that the substitution conditionally inactivates YraM function.[4] Experiments described for *Pseudomonas aeruginosa* and *E. coli* mutants showed that the *yraM* orthologs were not essential in these species, suggesting that bacteria with larger genomes might encode other proteins with complementary function.[7,8]

Although an effective vaccine exists for the encapsulated type b serotype of the bacterium, there is none for the nontypeable strains. The outer membrane location and essential function of YraM make it a potential vaccine antigen for *H. influenzae*. In fact, the antigenicity of a putative YraM homologue from *Mannheimia haemolytica* A1 (GS60, 33% identical to *H. influenzae* YraM) was correlated with the protective efficacy of a commercial bovine vaccine against this bacterium.[9] Additionally, the YraM homologue in *Edwardsiella tarda* is antigenic in fish[10] and LppC (67% identical to NTHi YraM) from *Haemophilus somnus* was reactive to bovine serum from infected animals.[11]

We report here that YraM is comprised of two distinct structural modules. With the goal of understanding the function of *H. influenzae* YraM, we describe the crystal structure to 1.35 Å resolution of the carboxyl-terminal portion of YraM (residues 257–573), and show it to be very similar to the family of bacterial periplasmic binding proteins (PBPs).

METHODS

Expression and purification of residues 257–573 of YraM

Restriction enzymes and polymerases for PCR were purchased from Invitrogen, Promega, and New England BioLabs and used according to the manufacturer's directions. Oligonucleotides were synthesized by Sigma Genosys, and plasmids and PCR fragments were purified with kits from Qiagen. Inserts cloned into plasmids were sequenced at the University of Michigan DNA sequencing facility in both directions to ensure error-free amplification.

Results of PSI-Blast sequence comparisons (see Results) suggested that the fragment corresponding to the PBP-like fold was between residues 257 and 575 of *H. influenzae* YraM. The 951 base pair gene fragment corresponding to Ser257-Val573 was amplified by PCR from genomic DNA isolated from the nonencapsulated *H. influenzae* Rd strain derivative of ATCC #9008. This gene sequence was almost identical to that present in a clinical isolate of a true NTHi strain, 86-029NP. Sequences of the primer oligonucleotides were 5′-CATGCCATGGTGTCACAAATTGGTTTACTCTTG-3′ (forward) and 5′-GCCGACGTCGACAACTGGTACAATTGCACCATC-3′ (reverse), which added NcoI and SalI sites to the 5′ and 3′ ends, respectively. The amplified fragment was cleaved with NcoI and SalI and ligated with T4 DNA ligase into the pETBlue^{TM}-2 vector (Novagen/EMD) that had been opened with NcoI and XhoI. Ligation products were transformed into *E. coli* NovaBlue competent cells. The resulting plasmid encoded the sequence Met-Val-YraM(Ser257-Val573)-Val-Asp-6(His), referred within as YraM-C.

To express soluble, folded protein, the plasmid was transformed into *E. coli* Origami^{TM}(DE3)-pLacI competent cells (Novagen), a strain that contains mutations to enhance disulfide bond formation in the cytoplasm (YraM-C has one disulfide). One liter of LB medium supplemented with ampicillin (100 μg/mL) was inoculated with 10 mL of an overnight culture, and incubated at 37°C with shaking until OD_{600} = 0.7. IPTG was added to a final concentration of 1 m*M* and growth was continued overnight at 25°C. Cells were harvested by centrifugation at 3000*g* for 10 min. To make protein with selenomethionine (SeMet) in place of methionine, YraM-C was expressed as above except that before induction, when the OD_{600} reached 0.6, cells were centrifuged, washed, and resuspended in M9 minimal media[12] supplemented with amino acids (except Met) and 80 μg/mL SeMet.

Wild type YraM-C and YraM-C(SeMet) were then purified by the same protocol. The cell pellet was resuspended in 50 mL of buffer A (50 m*M* sodium phosphate pH 8.0, 300 m*M* NaCl, 1 m*M* imidazole) containing an EDTA-free protease inhibitor cocktail tablet (Roche). Cells were sonicated and centrifuged at 4°C for 1 hr at 30,000*g*. The supernatant was added to Ni^{2+}-NTA agarose (Qiagen) and incubated with gentle rocking for 1 hr at 4°C. The lysate/Ni^{2+}-NTA mixture was loaded into a column and washed with buffer A, with increasing concentrations of imidazole up to 20 m*M*. Protein was eluted with buffer containing 300 m*M* imidazole and analyzed by SDS-PAGE. The purified protein was dialyzed

overnight at 4°C against a large volume of buffer (20 m*M* Tris pH 8.0, 1 m*M* EDTA, and 0.1 m*M* benzamidine) containing 0.1% β-mercaptoethanol (BME). The dialyzed protein was further purified on a MonoQ-10/100 column (Amersham) and eluted with a 0–1.0*M* NaCl gradient in the above buffer containing 2 m*M* dithiothreitol (DTT). Protein eluted between 60 and 100 mM NaCl. The protein was concentrated to 10 mg/mL in 20 mL Apollo™ (10-kDa cutoff) concentrators (Orbital Biosciences) and stored at −20°C.

Crystallization and data collection

Native YraM-C crystals grew as thin, needle clusters at 22°C in a drop containing 2 μL of 10 mg/mL protein solution and precipitant comprised of 2 μL 3.4*M* Na formate and 0.4 μL 0.1*M* DTT. The drop was equilibrated by vapor diffusion against 1 mL of precipitant containing 0.05% BME. Crystals grew to a maximum size of 0.8 × 0.1 × 0.1 mm^3 and were in the hexagonal space group $P6_5$, with cell dimensions $a = b = 117.2$ Å, $c = 115.1$ Å and two molecules per asymmetric unit. For cryoprotection, 20 μL of a precipitant containing 3.4*M* Na formate, 45 m*M* NaCl, and 0.05% BME with 50% saturated Li formate solution was added to the crystals immediately before freezing in liquid nitrogen.

The YraM-C(SeMet) protein crystallized at 22°C in a hanging drop containing 2 μL of protein and 2 μL of precipitant containing unbuffered 1.0*M* lithium sulfate, 2% (w/v) polyethylene glycol 8000, and 0.05% BME. The thin plate-like crystals reached a maximum size of 0.8 × 0.3 × 0.04 mm^3 and belonged to the monoclinic space group *C*2, with cell dimensions $a = 109.9$ Å, $b = 50.65$ Å, $c = 63.66$ Å, and $\beta = 106.8°$ and one molecule per asymmetric unit. For data collection, crystals were transferred to precipitant containing 25% glycerol and frozen in liquid nitrogen.

Diffraction intensities were measured for both crystal forms at DND-CAT (beamline 5-ID), Advanced Photon Source, Argonne, IL. Friedel pairs at three wavelengths were measured from a single SeMet crystal by the inverse beam method. Reflection intensities were integrated and scaled by d*TREK[13] (Rigaku/MSC). Intensity statistics for native and SeMet crystals are shown in Table I

Structure solution and refinement

Data from the monoclinic SeMet data were phased to 1.8 Å by the multiwavelength anomalous diffraction method as implemented in SOLVE[16] (Table I). The RESOLVE program[17] improved electron density maps by maximum-likelihood density modification, and located several α-helices of the structure. The remainder of the structure was modeled with O.[18] The SeMet-containing YraM-C model was refined by slow-cool simulated annealing (torsion dynamics, starting temperature = 3000 K) in CNS,[19] and finally with Cartesian dynamics (1500 K), with a bulk solvent correction applied. Because of its higher resolution, we chose to refine the model against the peak wavelength anomalous data diffraction data (0.97843 Å, $f' = -8.13$ and $f'' = 5.05$). After each round of refinement, σ_A-weighted $2mF_O$-DF_C and mF_O-DF_C maps were calculated with model phases for interactive model building. Upon reaching convergence at $R_{work} = 0.207$ ($R_{free} = 0.227$) with CNS, the model was refined with restrained anisotropic thermal displacement parameters (ADPs) to a final R_{work} of 0.16 ($R_{free} = 0.19$) for all data to 1.35 Å using the programs REFMAC[20] and Coot.[21] Weights were adjusted to obtain a root mean square deviation from ideality of 0.016 Å for bond lengths. All non-glycine residues were found in the allowed regions of a Ramachandran plot. Difference peaks were assigned to water molecules if the density in $2mF_O$-DF_C maps was above 1 σ over the mean, and formed at least one hydrogen bond to another atom. The final structure includes residues 257–574, 333 water molecules, one sulfate anion, and one unknown molecule modeled as BME (details in Table I). Eight of the residues had poor electron density; atoms from most of these were set to zero occupancy (see Results). Alternate conformations for 13 side chains were also observed. The mean anisotropy of the ADPs for all protein atoms was 0.67.[15] The model coordinates and structure factors are available from the Protein Data Bank[22] as entry 3CKM.

Diffraction data from the hexagonal crystal form of native YraM-C was phased by molecular replacement with CNS[19] and the SeMet structure as the search model, and partially refined to 2.7 Å (results not shown). Refinements were discontinued at this point since the model showed the same open conformation as observed in the high-resolution monoclinic structure.

RESULTS

Sequence analyses of YraM

YraM, encoded by the *Haemophilus influenzae* Rd HI1655 gene, is a 575-residue secreted polypeptide that, after processing, is a 60.6 kDa lipoprotein acylated at the N-terminal Cys26. PSORTb predicts that the protein is not retained on the inner membrane but is tethered to the outer membrane because the residue following Cys26 is Ser and not Asp.[23] PSI-Blast[24] searches of the non-redundant protein sequence library with the YraM sequence (Genbank accession code AAC23299) showed homology to two different families of molecules, suggesting that YraM may contain two distinct protein modules. A search with the amino-terminal region, comprising residues 33–256 (YraM-N), detected YraM orthologs but also several bacterial proteins of known function at a lower significance level ($E > 10^{-4}$): the amino-terminal portion of HemY, a membrane associated protein involved in protoheme biosynthesis;[25] the amino-termi-

Table I
Crystallographic data for YraM-C.

Data Collection

	YraM-C SeMet			YraM-C
Crystal	(beamline 5ID, DND-CAT, APS)			(DND, APS)
Unit cell (Å)	$a = 109.9$, $b = 50.65$, $c = 63.66$, $\beta = 106.8^\circ$			$a = b = 117.2$, $c = 115.08$
Space Group	$C2$			$P6_5$
Wavelength (Å)	0.97857	0.97843 (peak)	0.96321	1.0
Resolution[a] (Å)	47.1–1.4 (1.45–1.4)	47.1–1.35 (1.4–1.35)	47.1–1.45 (1.5–1.45)	46.4–2.7 (2.8–2.7)
Unique reflections	128,689[b] (12,927)	143,557[b] (14,350)	116,038[b] (11,558)	23,645 (2,254)
Observed reflections	405,018 (40,277)	451,051 (44,554)	366,438 (36,051)	76,978 (7,208)
Completeness[a] (%)	99.4 (99.4)	99.4 (99.3)	99.6 (99.7)	95.6 (91.6)
R_{merge}(%)[a,c]	4.5 (42.0)	4.4 (42.8)	4.0 (36.9)	4.8 (37.9)
$I/<\sigma_I>$[a]	12.4 (2.7)	12.2 (2.8)	13.4 (3.1)	16.7 (3.1)

Multiwavelength anomalous diffraction phasing with SOLVE[d]

Resolution (Å)	Number of Se sites	Correlation of local rms density in native Fourier	Mean figure of merit[e]	Overall *Z*-score
1.8	9	0.39	0.57	81.9

REFMAC refinement statistics of monoclinic SeMet crystal

Wavelength (Å)	0.9784
Resolution (Å)	38.95–1.35
R_{work}[f] (num. of reflections)	0.155 (66,014)
R_{free}[g] (num. of reflections)	0.190 (7,201)
Number of atoms[h]	
All	2879 (115)
Protein	2458 (92)
Water	356 (23)
Heterogen atoms	9
Se atoms	9
Number of amino acids	318
With alternate conformations	13
With poor electron density[i]	
Root mean square deviation from ideality	
Bonds (Å)	0.016
Angles (°)	1.6
Mean B_{iso}-factor (Å^{-2})[j]	
All	21.8
Protein	19.6
Main-chain	18.2
Side chains	21
Waters and heterogens	31.5
Mean atom B-factor anisotropy[k]	
Protein	0.67
Water	0.71

[a]Values for highest resolution shell are in parentheses.
[b]Scaling procedure treated the Bijvoet-related reflections independently.
[c]$R_{merge} = \Sigma|I-<I>|/\Sigma<I>$, where I is measured intensity of each reflection and $<I>$ is the intensity averaged from multiple observations of symmetry-related reflections.
[d]Statistics from SOLVE[14] for phasing of the MAD data set collected from the SeMet YraM-C crystal.
[e]Mean figure of merit after *RESOLVE* = 0.82.
[f]$R_{work} = \Sigma\ |Fo-Fc|/\Sigma|Fc|$, where Fo and Fc are observed and calculated structure factors respectively. Bijvoet pairs were not used in refinement by *REFMAC*.
[g]R_{free} is the *R*-factor calculated from 10% of the reflections not included in the refinement.
[h]Values in parentheses are the number of atoms added in alternate conformations. Water molecules were considered 'alternate conformations' when the distance to a neighboring water was less than 2.5 Å yet made good hydrogen bonds to the protein or other waters.
[i]Side- or main-chain atoms with electron density <1.0 sigma above the mean of the map.
[j]All atoms were refined with anisotropic displacement parameters (ADPs). These are the isotropic equivalent B-factors.
[k]Anisotropy is the square of the axial ratio of an ellipsoid representing the atom calculated from the ADPs.[15]

nal domains of the cellulose synthase operon C protein, an outer membrane protein;[26] a periplasmic protein annotated as a soluble lytic murein transglycosylase; and PilF involved in pilus formation.[27] All of these are periplasmic proteins and contain tetracopeptide-like repeats (TPRs). TPRs are 34-residue motifs, each forming two antiparallel helices, and often arranged sequentially to form a superhelix of helices.[28] Such domains mediate protein-protein interactions in vesicle transport, protein translocation, protein phosphorylation, and glycosyla-

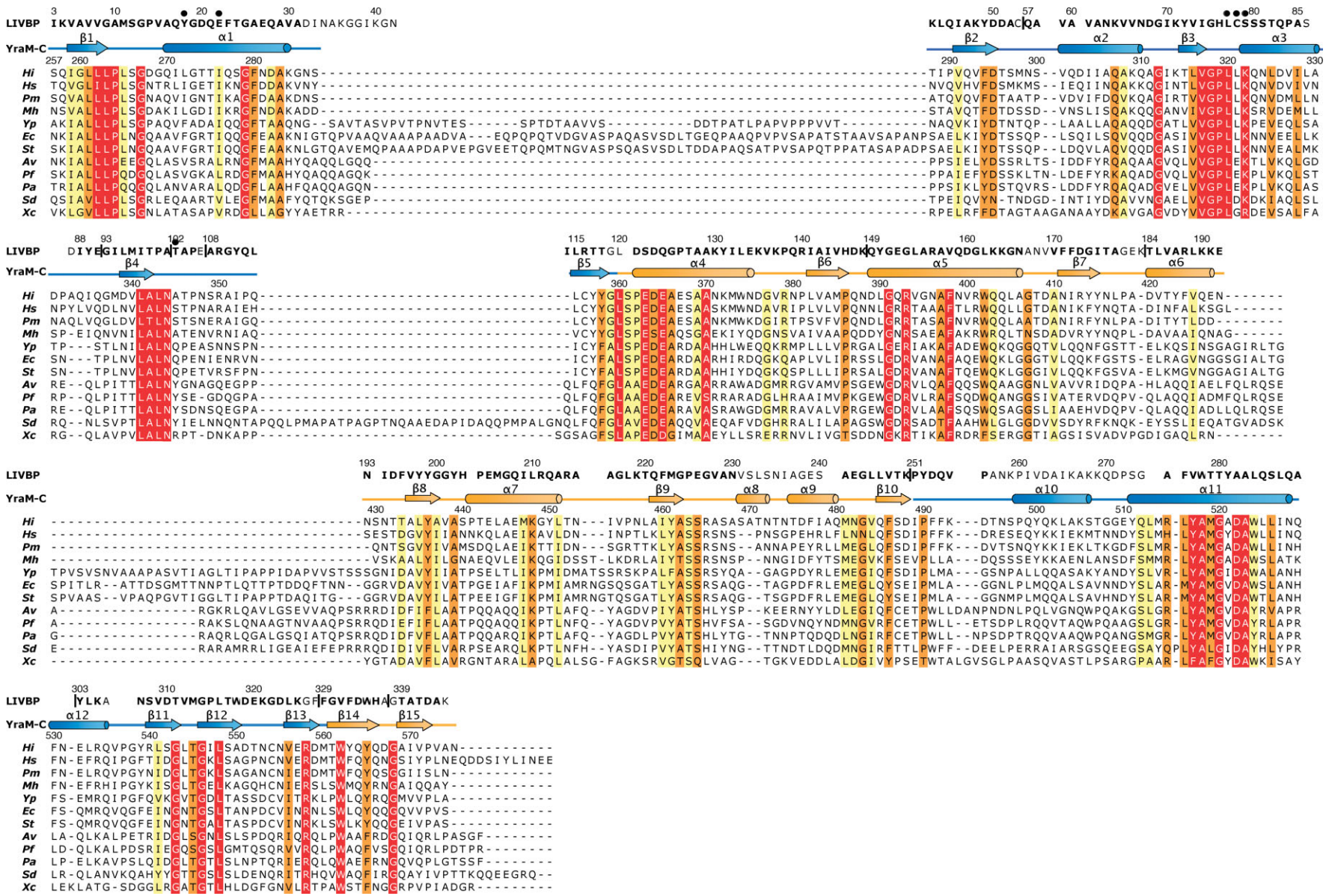

Figure 1

tion.[29] The presence of TPRs in YraM-N was also identified by the protein fold recognition server 3D-PSSM[30] that scored the TPR fold of the yeast vesicular transport protein sec17 as the best match with 80% confidence.

PSI-Blast searches with the carboxyl-terminal module YraM-C, comprised of residues 257–575 of YraM, showed low (10–16% identity) but very significant similarity to PBPs required for small molecule transport systems in bacteria. These were annotated as "putative branched chain amino acid binding proteins" and included the *E. coli* leucine/isoleucine/valine-binding protein (LIVBP, 11% identity) whose structures (with and without bound amino acid) have been solved.[31,32] The LIVBP fold was also predicted by the 3D-PSSM server. Figure 1 shows the alignment of YraM-C sequences from a diverse subset of YraM homologues. The percent identity between *H. influenzae* YraM-C and the other sequences in this alignment ranges from 23 to 62% (30 to 86% similar). A structure-based alignment with LIVBP is also shown.

Structure of YraM-C determined at 1.35 Å resolution

The native and selenium-containing proteins, with sequence Met-Val-YraM(257–573)-Val-Asp-His_6 (M_r = 35,800), were expressed and purified by metal affinity and ion exchange chromatography (see Materials and Methods). Multiwavelength diffraction data of the Se-containing protein were phased to 1.8 Å resolution by SOLVE and the complete model was refined to 1.35 Å to R_{work} = 0.16 and R_{free} = 0.19 (Table I) against diffraction data collected at the Se absorption peak wavelength (0.97843 Å). The final model contains YraM-C residues Ser257–Val573 and Val574, derived from the expression vector. Thirteen residues were modeled with alternate conformations. Because electron density was poor for residues 322, 346–348, 351, 429, 430, and 574, the residues are presumed to be disordered.

YraM-C is a putative periplasmic binding protein in an open configuration

As predicted, YraM-C(257-575) adopts a type 1 PBP-like fold[35] comprised of two distinct globular domains (N- and C-domains) connected by three interdomain crossovers. The N-domain includes residues 257–359 and 489–559, and the C-domain contains residues 360–488 and 560–573 [Figs. 1 and 2(A)]. The first portion of the N-domain forms an α/β structure with a five-stranded parallel β-sheet and three of the four flanking α-helices before crossing to the C-domain at residue 359 [Fig. 2(B)]. After the polypeptide returns to the N-domain at residue 490, it forms a solvent-exposed α-helix (498–506). This is followed by a long helix (511–535) containing a 120° kink at residue 528 that lies on the N-domain sheet to complete the α/β fold. To complete the domain, a three-stranded antiparallel β-sheet (536-559) joins the N-domain parallel sheet through two main-chain hydrogen bonds. A disulfide bond between residues Cys554 and Cys356 further anchors the three-stranded sheet to the parallel sheet.

The C-domain is comprised of a seven-stranded, mixed β-sheet flanked by helices [Fig. 2(C)]. The first segment (361–488) forms a five-stranded parallel β-sheet and all of the helices. After crossing to the N-domain to complete its fold, the polypeptide returns to the C-domain to form a final β-hairpin (β14-15) that is antiparallel to the parallel sheet.

Several YraM-C homologues contain insertions on one or more of the exposed loops at the distal ends of the molecule [Fig. 1, arrows in Fig. 2(A)]. Two of the inserts are in the N-domain, between residues 286 and 288 (α1-β2 loop) and between 352 and 355 (β4-β5 loop). The former insert in *Yersinia*, *E. coli*, and *Salmonella* is 40–70 residues long and rich in Ala, Pro, Val, and Gln (see Fig. 1). There is another insert of 20–30 residues with similar amino acid composition in the C-domain between residues 427 and 431 (α6-β8 loop).

The folds of the YraM-C domains are most similar to the domains of periplasmic LIVBP, the homologous leucine-specific binding protein (LBP),[36] the cytoplasmic amide-binding domain of the AmiC repressor, and two eukaryotic binding protein-like domains: hormone binding domain of the atrial natriuretic peptide receptor and glutamate receptor (Table II).

Differences between YraM-C and LIVBP are primarily found in two loops located on the N-domain distal from

Figure 1

Sequence alignment of presumed YraM-C orthologs and structural alignment to LIVBP. Secondary structure representations for YraM are in blue for N-domain and gold for C-domain. Residues are colored based on degree of conservation calculated by SCORECONS.[14] Those in the top 10 percentile are highlighted in red, the next 10 percentile in orange, and the next 10 percentile in yellow. Numbering is for the H. influenzae *sequence. Also shown is a structure-based alignment of the LIVBP sequence with YraM-C with the structurally equivalent residues in* ***bold****. Vertical lines between residues in the LIVBP sequence represent residue insertions that were omitted for clarity. Black circles are LIVBP residues in the ligand-binding region of the N-domain (see text). The sequence and structure-based alignments were calculated with ClustalX[33] and LSQMAN,[34] respectively. Species, protein name, and Genbank accession numbers for sequences in the alignment: Hi,* Haemophilus influenzae *Rd KW20, HI1655 (16273542); Hs,* Haemophilus somnus, *LppC (4096758), Pm,* Pasteurella multocida, *LppC (15602511); Mh,* Mannheimia haemolytica, *GS60 antigen (62798901); Yp,* Yersinia pestis, *YP03548 (16123692); EcK12,* Escherichia coli *K12, YraM (7466039); St,* Salmonella typhimurium *LT2, YraM (16766562); Av,* Azotobacter vinelandii, *LppC (67086486); Pf,* Pseudomonas fluorescens, *LppC (77384908); Pa,* Pseudomonas aeruginosa *(9950656); Sd,* Saccharophagus degradans *2-40 (90022787); Xc,* Xanthomonas campestris, *XCC0711 (21230186).*

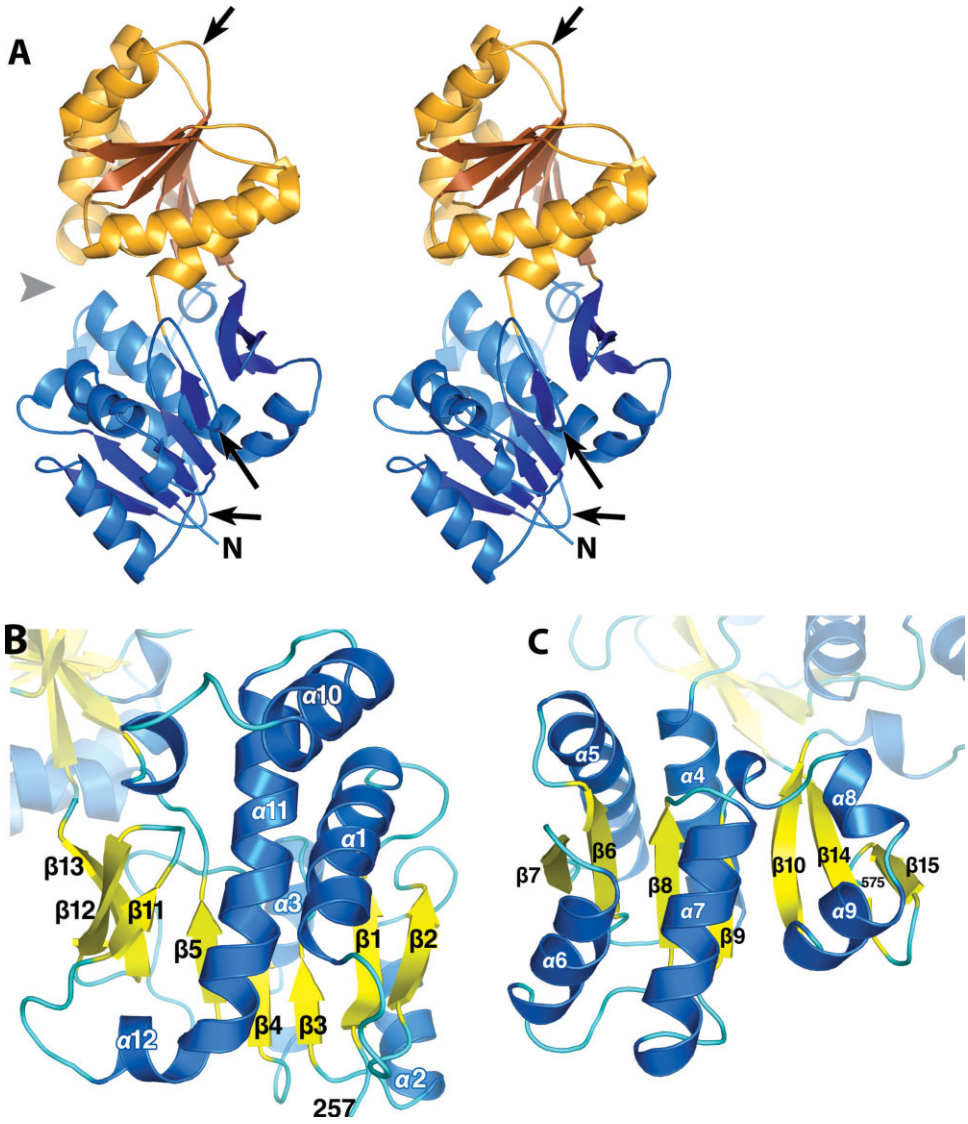

Figure 2

*Cartoon representation of YraM-C (residues 257–573). (**A**) Stereo view showing the two domains, N-domain in blue and C-domain in gold. Black arrows indicate the loops where insertions are found in the YraM homologues of several other species (see Fig. 1). The gray arrow is the view of Fig. 4, into the cleft between the domains. (**B**) and (**C**) N-domain and C-domain secondary structures, respectively, are labeled. This and subsequent figures were rendered with MacPyMOL (version 0.99, Delano Scientific).*

the interdomain cleft. The YraM-C α1 helix is shorter by two turns compared with LIVBP, resulting in a shorter loop (residues 286–289) to β2. This same loop in several YraM homologous contains the long insert discussed above. Helices α11 and α12 are shorter by two turns causing the loop between them to differ by up to 9 Å from LIVBP.

On the side of the domain facing the cleft the position of the β2-α2 loop (residues 296–300) also differs between the two proteins. This α2 helix is one turn shorter in YraM-C and lacks the disulfide bond present in LIVBP and LBP. Between the two domains of YraM-C is the α10 helix that is more solvent exposed and makes fewer interactions with the N-domain when compared with the corresponding region in LIVBP. Also, in the interdomain region (residues 551–558), LIVBP has a β-bulge (324–327) compared to the regular β-structure in YraM-C (556–558). This causes the direction of the polypeptides to diverge as they enter the C-domain, possibly causing the different orientation of the C-domain with respect to the N-domain in the two proteins (see below).

In YraM-C, the two domains are oriented to form a wide, open cleft that is similar to the unliganded, open

Table II
Superpositions of YraM-C domains with structural homologs

	PDB ID	YraM-C C-domain			YraM-C N-domain		
		Num equivalent	Number identical residues	r.m.s.d.[a]	Num equivalent	Number identical residues	r.m.s.d.
Leucine-binding protein (LBP), unliganded[36]	1USG	125	18	1.62	112	10	1.77
Leu/Ile/Val-binding protein (LIVBP), unliganded[31]	1Z15	122	17	1.67	115	9	1.77
Amide repressor AmiC, liganded[37]	1PEA	122	19	1.92	99	10	1.94
Ribose binding protein, unliganded[38]	1BA2	88	12	1.88	103	6	1.99
Atrial natriuretic peptide receptor, unliganded[39]	1DP4	113	20	1.74	94	12	1.68
Metabotropic glutamate receptor, unlliganded[40]	1EWT	116	13	1.93	92	9	1.92

Superpositions calculated with LSQMAN[34].
[a]r.m.s.d., root mean square of distances in Å between equivalent Cα atoms.

conformations observed in several other PBPs [Fig. 3(A)]. Such proteins are often dynamic and have flexible hinge regions connecting the two domains. Sometimes different degrees of hinge bending are observed in different crystal forms of the same PBP. Unfortunately, the hexagonal crystal form of YraM-C has an identical conformation to the monoclinic form discussed here (results not shown). The YraM-C domain arrangement is most similar to the "super-open" unliganded conformations of LIVBP[31] (Fig. 3(A), PDB 1Z15). When the N-domains of LIVBP and YraM-C are superimposed, the root mean square difference between equivalent residues is about 1.7 Å (Table II). In this superposition, the C-domain of YraM-C is rotated about 15–20° with respect to the C-domain of LIVBP along an axis perpendicular to the strands of the β-sheet (away from cleft, i.e., more open). The YraM-C C-domain is also translated closer to the N-domain (as much as 5 Å) in the direction perpendicular to the C-domain β-sheet [Fig. 3(A)].

Differences in the interdomain region may account for the difference in YraM-C's domain orientation when compared with LIVBP. The amino end of the α11 helix (512–529) in the YraM-C N-domain is longer by 1.5 helical turns compared to LIVBP. As a result, the α10-α11 loop is closer to the β9-α8 loop of the C-domain [Fig. 2(B)]. A consequence of the close approach of these loops is the formation of a pronounced pocket (not shown) in the interdomain hinge region near the C-domain. The pocket is formed from residues from both domains and includes four aromatics: Phe492 and Tyr511 from the N-domain and Phe486, Tyr563, and Ile570 from the C-domain. Ser467 and Asp488 are at the base of this pocket. In this pocket was difference electron density with peak height >10 standard deviations above the mean. We modeled the unknown ligand as BME because it is the only reagent in the crystallization mix close in size and shape to the difference density. We oriented the thiol group of the BME toward the bottom of the pocket to form hydrogen bonds of 2.7 Å to the Asp488 carboxylate and Ser567 OH. Similar short SH-OH hydrogen bonds have been observed previously.[41] Because only two of the residues (Phe486 and Asp488) in this pocket are conserved in the sequence alignment (see Fig. 1), it is unknown whether this pocket has functional significance.

Highly conserved residues are located in cleft between domains

To find regions of YraM-C that are potentially important for biological function, we colored the solvent-contact surface in Figure 4 with the most conserved residues of Figure 1. Most of the highly conserved and exposed residues are located in the wide cleft between the two domains. The cleft is ~20 Å long (vertical direction in Fig. 4) and 18 Å wide in the dimension between the two domains. Most notable is a small pocket at the base of the cleft surrounded by residues (clockwise) Pro362, Glu365, Arg558, Leu360, Arg515, and Ser464, with

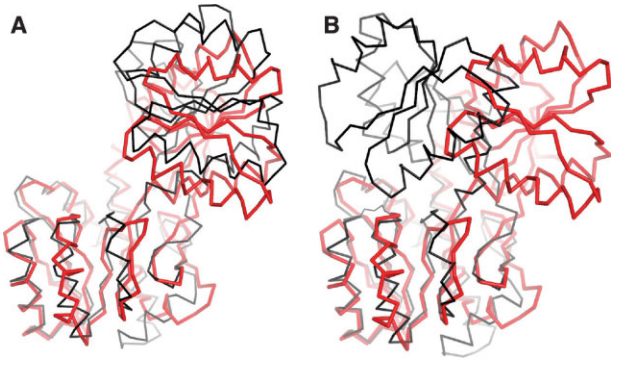

Figure 3
*Comparison of the relative domain orientation of YraM-C and periplasmic binding proteins. Red is the YraM-C Cα backbone. Black backbones are (**A**) LIVBP in unliganded, super-open conformation (PDB ID: 1Z15); and (**B**) LBP in closed, leucine-bound conformation (1USK).*

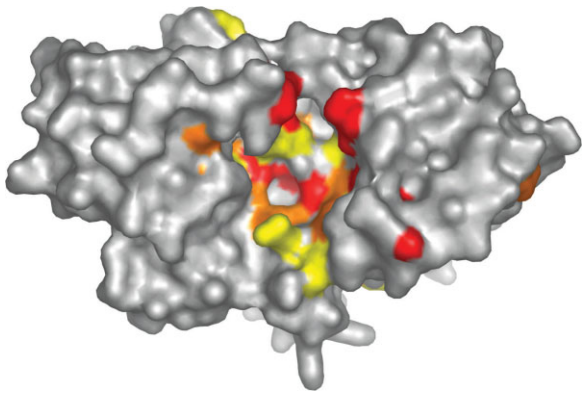

Figure 4

Conserved residues cluster in the putative ligand-binding cleft of YraM-C. In gray is the solvent contact surface of YraM-C. Surfaces of conserved side chains are colored as in Fig. 1. N-domain is on the right and the C-domain is on the left. See the gray arrow in Fig. 2(A) for the view orientation.

Asp488, Ile489, and Met 560 at the bottom [Fig. 5(A)]. Ser463, Tyr436, and Val438 form a small side crevice off the central pocket. The bottom of the cleft is mostly polar (Glu356, Arg558, Ser464, and Arg515) while the N-domain side [arrow in Fig. 5(A,C)] of the cleft is predominantly nonpolar (discussed later). The entire cleft is filled with arrays of water molecules. Near the hinge region is a bound sulfate hydrogen-bonded to Ser469 and Gln512 (not shown).

Model of YraM-C in a "closed" conformation

Crystal structures of many PBP-like proteins show a "closed" conformation of the cleft with a bound ligand between the two domains, effectively sequestering it from bulk solvent.[31] However, the degree of domain closure observed in ligand-bound crystal structures varies, and may be related to the size of the bound ligand. In the closed, leucine-bound forms of LBP (PDB 1USK) or LIVBP (PDB 1Z16), the cavity between the two domains is relatively small, and completely buried. In contrast, a subclass of PBPs typified by the ligand-bound structures of vitamin B-12 binding protein and the iron siderophore binding protein FhuD, have a wide open cleft that differs only slightly from the respective unliganded form crystal structures.[42,43] These proteins are characterized by a long α-helix between the two domains that may rigidify the structure.

YraM-C may also undergo a conformational change upon binding a putative ligand. To assess which YraM-C residues would interact with a small molecule ligand, we modeled a "closed" form of YraM-C by separately superposing the N- and C-domains of YraM-C on the corresponding domains of the leucine-bound form of LBP (PDB id 1USK, Fig. 6).[36] We also adjusted the main chain torsion angles in the hinge regions, residues 358–361 and 555–559. The closed YraM-C model has a distinct pocket or cavity in a region corresponding to the location of the bound leucine in LBP (magenta sphere in Fig. 6). The cavity is about 16 Å long in the direction perpendicular to the hinge, 8-Å deep, and 8-Å wide. We estimate that it is approximately three times larger than the leucine-binding site in the LBP closed structure. Residues lining the cavity from the N-domain are predominantly hydrophobic, and include Leu516, Leu272, Gln323, Leu320, Ile 276, and Met 519. However, four residues exposed to the cavity are polar and highly conserved: Asn344, Glu365, Ser464, and Arg515 (see Fig. 6).

Conserved residues Lys322 and Asn344 of the N-domain, and Arg393 and Glu363 from the C-domain, are located on the outer lips of the cleft, Comparison of the open structure and closed model suggests that these residues could form interdomain hydrogen bonds upon cleft closure. These interactions may serve to sequester the contents of the binding cavity from bulk solvent.

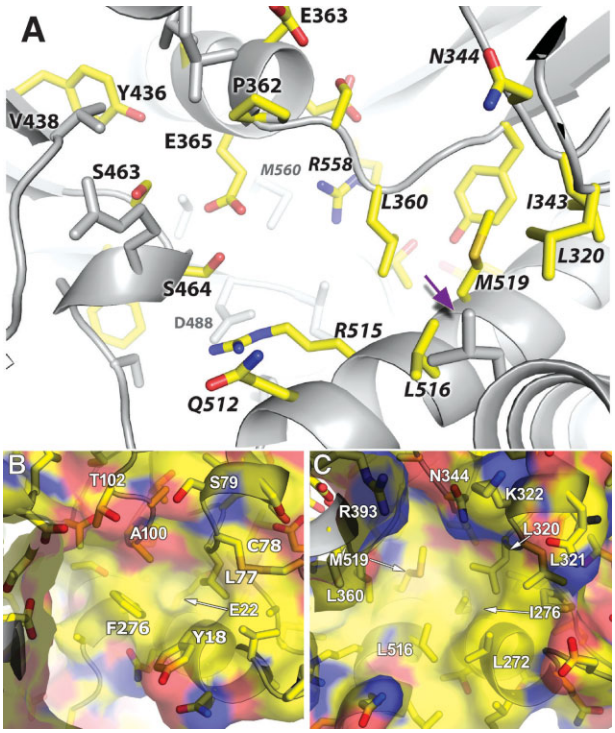

Figure 5

*Putative ligand-binding cleft of YraM-C. (**A**) At the base of the cleft is a pocket surrounded by residues L360, R515, S464, D365, and R558. Yellow-colored side chains are the most conserved residues. (**B**) and (**C**) show the structurally equivalent nonpolar surfaces (yellow) on the N-domains of LIVBP and YraM-C, respectively, viewed approximately in the direction of the violet arrow shown in (A). As in LIVBP, this surface in YraM-C may serve as an initial interaction site for a proposed amphipathic ligand. In LIVBP (B), the leucine side chain is positioned between A100, F276, Y18, and L77 and its amino and carboxylate groups form hydrogen bonds to T102 and S79.*

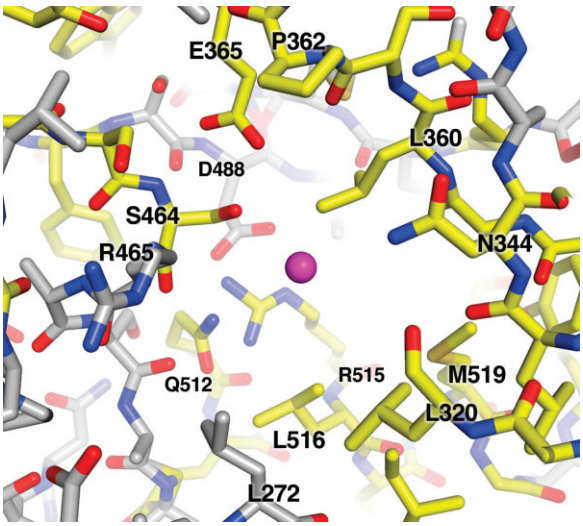

Figure 6

The YraM-C putative ligand binding site in a model of the "closed" form based on the closed, liganded structure of LBP. The conserved residues are colored in yellow and the N-terminal domain is on the right. The maroon sphere represents the corresponding location of bound leucine in the LBP structure.

Similar nonpolar surface in YraM-C and LIVBP clefts

The proposed ligand-binding site of YraM-C is comprised of residues from both domains and has both hydrophobic and polar regions. Despite fold similarity, the residues in the YraM-C binding site are not identical with those in LBP and LIVBP. Structures and kinetic studies of ligand binding to the maltose-binding protein (MBP) and LIVBP in various conformations suggest that ligands bind in two steps: recognition of ligand by one domain, followed by domain closure.[31] In MBP and LIVBP, the respective ligand binds "preferentially to the domain whose site is more nonpolar in character than that of the other domain."[31] In LIVBP, this nonpolar region is found on the surface of the N-domain centered above Glu22 and surrounded by Leu77, Cys78, Ala100, and Phe276 [Fig. 5(B)]. At one edge of this region are two polar residues, Thr102 and Ser79. The amide and carboxylate of a bound leucine molecule form hydrogen bonds to these two residues while the side chain is nestled in the hydrophobic pocket.

Interestingly, the N-domain of YraM-C [Fig. 5(C)] has a very similar region comprised of residues that are structurally equivalent to those residues in LIVBP mentioned earlier. In YraM-C, the hydrophobic pocket is centered above residue Ile276, and surrounded by Leu320, Leu321, Leu272, Leu516, and Met519 [Fig. 5(C)], most of which are conserved in YraM-C homologues. The polar residues are Asn344 and Lys322 and are structurally equivalent to the polar residues described for LIVBP above. Black dots in Figure 1 highlight the location of these residues in both proteins. All of the residues in this region are also in the ligand-binding site of the modeled closed form of YraM-C. We anticipate that YraM-C may bind an amphipathic ligand, and interact with both the nonpolar surface of the N-domain as well as the polar side chains contributed by the C-domain.

Location of residues responsible for the temperature sensitive mutant

The temperature sensitive mutant Asp522→Gly allowed *H. influenzae* cells to grow in rich medium at 30°C but not at 37°C.[4] Asp522, identical in all YraM homologues, is located on the N-domain's long α11 helix where it makes three hydrogen bonds to a β-turn (544–545) between the first and second strands (β11 and β12) of a three-stranded antiparallel sheet (536–559; Fig. 7). This small sheet forms part of the hinge region between the domains. The hydrogen bonds from Asp522 orient the β-turn and allow it to pack against a predominantly nonpolar loop, residues 489–491. Loss of these hydrogen bonds in the Asp522→Gly mutant may increase the dynamics of the β-sheet and possibly interfere with hinge movement or ligand binding.

DISCUSSION

The structure of the carboxyl-terminal module of YraM is very similar to that of type 1 PBPs.[44,45] Like YraM, the unliganded forms of PBPs show a wide cleft between the domains and a relatively flexible hinge

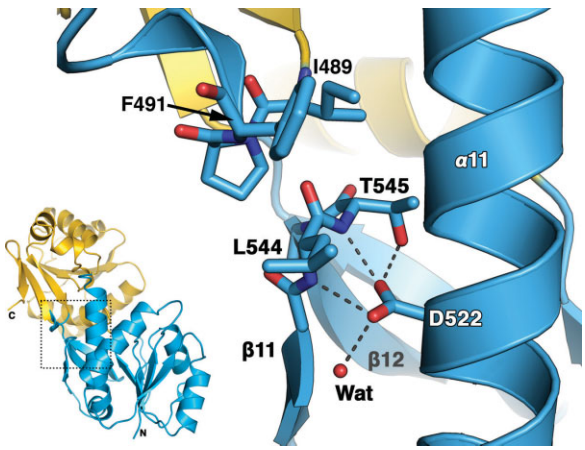

Figure 7

Location of Asp522 on α11 of the N-domain. Substitution of Asp522 with Gly produces a temperature-sensitive mutant that prevents H. influenzae *from growing at 37°C but grows normally at 30°C. Asp522 forms hydrogen bonds to the β-turn between strands β11 and β12 (also see Fig. 2(B)). Inset shows that this region is located between the two YraM-C domains.*

region. Upon binding their cognate ligands, most PBPs adopt a conformation in which the two domains close together to sequester the ligand molecule. For solute transport, the "closed" PBPs bind to a transmembrane permease, part of a ABC-type transporter. ATP hydrolysis by the associated cytoplasmic ATPase in this complex induces a conformational change that releases the ligand from the PBP for translocation into the cytoplasm.

PBPs and similarly folded proteins can have other functions besides transport. Some PBPs serve as initial chemoreceptors for bacterial chemotaxis.[46] Other PBP-like proteins bind small molecules that then interact with a bacterial transmembrane receptor to activate a cytoplasmic sensor kinase as part of two-component signaling systems.[47] Cytoplasmic proteins in bacteria and higher eukaryotes also have PBP folds, such as the ligand-binding regulatory domains of some transcription factors (e.g., *lac* repressor) and an inner membrane-associated glycosyltransferase.[48] In eukaryotes, the extracellular module of the transmembrane atrial natriuretic hormone receptor also has the PBP fold but forms dimers.[49] The hormone peptide does not bind in the receptor's cleft like other PBPs, but binds to a conserved surface between the two subunits of the dimer.

YraM-C is most similar to the leucine-specific- or Leu/Ile/Val-binding proteins in the super-open conformation observed for unliganded LIVBP.[31] Despite the similarity, YraM-C does not bind leucine (B. Tirupati, unpublished data). Although no bound ligand was observed in the YraM-C crystal structure, we noted that the majority of highly conserved residues in YraM orthologs are between the two domains in the same location where other PBPs bind their ligands.

PBPs often copurify with their respective ligand allowing identification in the crystal structure (e.g., Ref. 50). The most likely reason why a putative bound ligand was not observed in the YraM-C crystal structure is that either the appropriate ligand was not present in the *E. coli* cytoplasm or growth media, or the ligand dissociated during purification or crystallization. In support of the latter hypothesis, we noted that the packing of YraM-C in the monoclinc crystal form precludes the protein from assuming the closed form.

We have modeled a putative closed form of the protein, but the question remains if it is realistic. Normal modes derived from a molecular dynamics simulation of LIVBP showed the hinge-bending motion necessary for adopting the observed closed conformation.[31] To see if YraM-C might have similar dynamics necessary for closing the interdomain cleft, we analyzed the normal modes for unliganded LIVBP and YraM-C using the ElNémo server.[51] Both structures displayed the same interdomain hinge-bending motion necessary for domain closing (data not shown). From these observations we hypothesize that YraM-C is a solute binding protein and that ligand binding will stabilize a "closed" cleft conformation as observed for the LIVBP structure with bound amino acid [see Fig. 3(B)]. Ongoing experiments to identify this ligand will contribute to understanding the function of YraM.

Speculations on the function of YraM

Little has been published about the function of YraM or its homologues other than its identification as an essential protein in *H. influenzae*.[4] The location of *yraM* on the *H. influenzae* chromosome provides few clues to its function. *YraM* is the second in a cluster of five genes conserved in many Gram-negative genomes. The last gene in the cluster, HI1658 (*E. coli yraP*), is also an outer membrane lipoprotein and was shown to be essential for growth in *H. influenzae*.[2] In *E. coil* this protein is regulated by σ^E and was proposed to be involved in outer membrane integrity.[52]

The only gene in the cluster with a well-defined function is the cytoplasmic HI1657 (*E. coli yraO, diaA*). HI1657 was annotated as encoding a phosphoheptose isomerase similar to GmhA involved in lipopolysaccharide core biosynthesis. Recent studies of the orthologous *E. coli* DiaA protein show that the protein is important for initiation of chromosomal replication.[53] It forms a tight complex with DnaA, enabling the latter to bind and unwind *oriC*, the *E. coli* origin of replication. If this gene is coregulated with YraM, we might expect YraM to be involved in some process necessary for cell division.

Although the majority of YraM's conserved residues are found in YraM-C (data not shown), the structure of YraM-N (residues 26–256) may also give clues to YraM's function. We predicted that YraM-N contains helix-turn-helix motifs similar to the TPRs found in a large variety of other proteins. To the authors' knowledge, YraM is unique in having a structure in which this TPR fold is fused to the amino terminus of a PBP-like molecule. This novel fusion may be important for protein localization, or to enhance the interaction of YraM-C with other molecules. A similar idea was proposed for the periplasmic lytic transglycosylase Slt70.[28] Like YraM, SLT70 has an amino-terminal domain containing 11 TPR motifs, followed by a functional catalytic domain that is structurally similar to lysozyme. Considering SLT70's role in the hydrolysis of glycosidic bonds and turnover of peptidoglycan, researchers suggested that this amino-terminal helical domain may have affinity for peptidoglycan.[28] Although SLT70 is not a lipoprotein, there are similar murein hydrolases, such as MltA, MltC, and NlpD, which are outer membrane lipoproteins.

Because YraM orthologs are not found in Gram-positive organisms, we speculate that YraM may be involved in structures specific to Gram-negative bacteria, such as *meso*-diaminopimelate peptidocglycan or lipopolysaccharide. If we assume that full-length YraM is tethered on the inner leaflet of the outer membrane, it is likely within reach of the peptidoglycan estimated to be ~50 Å from

the membrane.[54] There is precedence for binding protein folds functioning in peptidoglycan biogenesis. For example, MppA[55] is a PBP that is essential for import of murein tripeptide and MurG is a cytoplasmic PBP-like glycosyltransferase required for the biosynthesis of peptidoglycan precursors.[48] Interestingly, a model of 1,6-anhydromuropeptide, which co-crystallized with SLT70,[28] will fit in the proposed binding cleft of YraM-C, but only in the open form (V. J., results not shown).

Although Gram-positive bacteria express PBPs as lipoproteins, only a few besides YraM have been described for Gram-negative bacteria. Interestingly, at least two of them may bind iron. HbpA is a heme-binding PBP in *H. influenzae*, but its role in heme utilization is not fully understood.[56,57] A putative iron-binding periplasmic lipoprotein SfbA (NlpA) from *Salmonella enteritidis* was induced in the presence of iron.[58] A recent dissertation, studying the outer membrane proteome of a clinically relevant *P. aeruginosa*, reports that YraM is expressed at low levels during normal growth in rich media but at high levels during iron starvation induced by addition of 1.1 m*M* 2,2-dipyridyl, an iron chelator, to the growth media.[59] This finding may be particularly relevant to *H. influenzae* because it must import iron and heme or porphyrin for aerobic growth. Interestingly, we note that HI1654 (*yraL* in *E. coli*), just upstream of *yraM,* encodes a putative tetrapyrrole methylase. Tetrapyrroles are precursors of porphyrin biosynthesis. Like YraM, HI1654 is essential in *H. influenzae,* but not in *E. coli* or *P. aeruginosa.*[2] Expression of the *P. aeruginosa* ortholog of HI1654 is upregulated in a *yraM* deletion mutant of *P. aeruginosa* (Wong and Akerley, unpublished results), suggesting that the function of these two genes are linked. These observations may imply a role for YraM in iron or porphyrin acquisition.

Is YraM part of a binding protein-mediated ligand transport system? In the ABC transporter system described earlier, binding proteins are in the periplasm (or anchored to the inner membrane) and physically interact with a membrane-bound permease in the cytoplasmic membrane. In contrast, YraM is anchored to the outer membrane. When MalE, a soluble periplasmic MBP, was modified to be an outer membrane lipoprotein, it failed to function in maltose uptake.[60] One explanation might be that the size of MalE and the close proximity of the peptidoglycan prevented its interaction with the appropriate membrane ABC permease. Results from electron microscopy[54] show the periplasm to be about 170 Å wide, but recent structures of the type III secretion system needle[61] and the exopolysaccharide translocation apparatus[62] suggest that the width may be much smaller (≈125–150 Å). Although the relative orientation of the two modules of YraM is not known, we estimate that the protein will be too short to interact with the membrane permease. To circumvent this problem, the YraM-C portion might be released by proteolysis, a mechanism occasionally observed for other lipoproteins. Of course, we cannot rule out that YraM may function in the transport of molecules across the *outer* membrane, perhaps by sequestering ligands in the periplasm. Addressing these possibilities represents a promising area for future research into the mechanism of YraM's function.

ACKNOWLEDGMENTS

We thank Dr. Sandy Wong (University of Massachusetts) for sharing unpublished data, Prof. Florante Quiocho for sharing the coordinates of LIVBP before release, and Drs. Linda Luck and Sherry Mowbray for discussions. We also thank the reviewers for their constructive comments. Portions of this work were performed at the DuPont-Northwestern-Dow Collaborative Access Team (DND-CAT) located at Sector 5 of the Advanced Photon Source (APS). DND-CAT is supported by E.I. DuPont de Nemours & Co., The Dow Chemical Company and the State of Illinois. Use of the APS was supported by the U. S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. DE-AC02-06CH11357. We appreciate the help of DND-CAT's expert staff.

REFERENCES

1. Celli BR, MacNee W. Standards for the diagnosis and treatment of patients with COPD: a summary of the ATS/ERS position paper. Eur Respir J 2004;23:932–946.
2. Akerley BJ, Rubin EJ, Novick VL, Amaya K, Judson N, Mekalanos JJ. A genome-scale analysis for identification of genes required for growth or survival of *Haemophilus influenzae*. Proc Natl Acad Sci USA 2002;99:966–971.
3. Bateman A, Coin L, Durbin R, Finn RD, Hollich V, Griffiths-Jones S, Khanna A, Marshall M, Moxon S, Sonnhammer EL, Studholme DJ, Yeats C, Eddy SR. The Pfam protein families database. Nucleic Acids Res 2004;32:D138–D141.
4. Wong SM, Akerley BJ. Inducible expression system and marker-linked mutagenesis approach for functional genomics of *Haemophilus influenzae*. Gene 2003;316:177–186.
5. Terada M, Kuroda T, Matsuyama SI, Tokuda H. Lipoprotein sorting signals evaluated as the LolA-dependent release of lipoproteins from the cytoplasmic membrane of *Escherichia coli*. J Biol Chem 2001; 276:47690–47694.
6. Lopez-Campistrous A, Semchuk P, Burke L, Palmer-Stone T, Brokx SJ, Broderick G, Bottorff D, Bolch S, Weiner JH, Ellison MJ. Localization, annotation, and comparison of the *Escherichia coli* K-12 proteome under two states of growth. Mol Cell Proteomics 2005;4:1205–1209.
7. Wong SM, Mekalanos JJ. Genetic footprinting with mariner-based transposition in *Pseudomonas aeruginosa*. Proc Natl Acad Sci USA 2000;97:10191–10196.
8. Moffat JC. Properties of conductance and inhibition of proton channels: M2 from influenza A virus and Fo from *Escherichia coli* ATP synthase. Master thesis. Brigham Young University; 2006.
9. Lo RY, Mellors A. The isolation of recombinant plasmids expressing secreted antigens of *Pasteurella haemolytica* A1 and the characterization of an immunogenic 60 kDa antigen. Vet Microbiol 1996;51:381–391.
10. Verjan N, Hirono I, Aoki T. Genetic loci of major antigenic protein genes of *Edwardsiella tarda*. Appl Environ Microbiol 2005;71:5654–5658.

11. Potter AA, Theisen M, Harland RJ, Rioux CR. Nucleic acid molecules encoding *Haemophilus somnus* proteins. U.S. Patent number 6,100,066; 2000.
12. Sambrook J, Fritsch EF, Maniatis T. Molecular cloning: a laboratory manual. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 1989.
13. Pflugrath JW. The finer things in X-ray diffraction data collection. Acta Crystallogr D Biol Crystallogr 1999;55:1718–1725.
14. Valdar WS. Scoring residue conservation. Proteins 2002;48:227–241.
15. Merritt EA. Expanding the model: anisotropic displacement parameters in protein structure refinement. Acta Crystallogr D Biol Crystallogr 1999;55:1109–1117.
16. Terwilliger TC, Berendzen J. Automated MAD and MIR structure solution. Acta Crystallogr D Biol Crystallogr 1999;55:849–861.
17. Terwilliger TC. Automated structure solution, density modification and model building. Acta Crystallogr D Biol Crystallogr 2002;58: 1937–1940.
18. Jones TA, Zou J, Cowan SW, Kjeldgaard M. Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr A Biol Crystallogr 1991; 47:110–119.
19. Brünger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Crystallogr D Biol Crystallogr 1998;54:905–921.
20. Murshudov GN, Vagin AA, Lebedev A, Wilson KS, Dodson EJ. Efficient anisotropic refinement of macromolecular structures using FFT. Acta Crystallogr D Biol Crystallogr 1999;55:247–255.
21. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 2004;60:2126–2132.
22. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. Nucleic Acids Res 2000;28:235–242.
23. Gardy JL, Laird MR, Chen F, Rey S, Walsh CJ, Ester M, Brinkman FS. PSORTb v. 2.0: expanded prediction of bacterial protein subcellular localization and insights gained from comparative proteome analysis. Bioinformatics 2005;21:617–623.
24. Altschul SF, Koonin EV. Iterated profile searches with PSI-BLAST–a tool for discovery in protein databases. Trends Biochem Sci 1998; 23:444–447.
25. Hansson M, Hederstedt L. *Bacillus subtilis* HemY is a peripheral membrane protein essential for protoheme IX synthesis which can oxidize coproporphyrinogen III and protoporphyrinogen IX. J Bacteriol 1994;176:5962–5970.
26. Zogaj X, Nimtz M, Rohde M, Bokranz W, Römling U. The multicellular morphotypes of *Salmonella typhimurium* and *Escherichia coli* produce cellulose as the second component of the extracellular matrix. Mol Microbiol 2001;39:1452–1463.
27. Kim K, Oh J, Han D, Kim EE, Lee B, Kim Y. Crystal structure of PilF: functional implication in the type 4 pilus biogenesis in *Pseudomonas aeruginosa*. Biochem Biophys Res Commun 2006;340:1028–1038.
28. van Asselt EJ, Thunnissen AM, Dijkstra BW. High resolution crystal structures of the *Escherichia coli* lytic transglycosylase Slt70 and its complex with a peptidoglycan fragment. J Mol Biol 1999;291:877–898.
29. D'Andrea LD, Regan L. TPR proteins: the versatile helix. Trends Biochem Sci 2003;28:655–662.
30. Kelley LA, MacCallum RM, Sternberg MJ. Enhanced genome annotation using structural profiles in the program 3D-PSSM. J Mol Biol 2000;299:499–520.
31. Trakhanov S, Vyas NK, Lüecke H, Kristensen DM, Ma J, Quiocho FA. Ligand-free and -bound structures of the binding protein (LivJ) of the *Escherichia coli* ABC leucine/isoleucine/valine transport system: trajectory and dynamics of the interdomain rotation and ligand specificity. Biochemistry 2005;44:6597–6608.
32. Sack JS, Saper MA, Quiocho FA. Periplasmic binding protein structure and function. Refined X-ray structures of the leucine/isoleucine/valine-binding protein and its complex with leucine. J Mol Biol 1989;206:171–191.
33. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 1997;25:4876–4882.
34. Kleywegt G, Jones TA. Detecting folding motifs and similarities in protein structures. Methods Enzymol 1997;277:525–545.
35. Fukami-Kobayashi K, Tateno Y, Nishikawa K. Domain dislocation: a change of core structure in periplasmic binding proteins in their evolutionary history. J Mol Biol 1999;286:279–290.
36. Magnusson U, Salopek-Sondi B, Luck LA, Mowbray SL. X-ray structures of the leucine-binding protein illustrate conformational changes and the basis of ligand specificity. J Biol Chem 2004;279:8747–8752.
37. Pearl L, O'Hara B, Drew R, Wilson S. Crystal structure of AmiC: the controller of transcription antitermination in the amidase operon of *Pseudomonas aeruginosa*. EMBO J 1994;13:5810–5817.
38. Bjorkman AJ, Mowbray SL. Multiple open forms of ribose-binding protein trace the path of its conformational change. J Mol Biol 1998;279:651–664.
39. van den Akker F, Zhang X, Miyagi M, Huo X, Misono KS, Yee VC. Structure of the dimerized hormone-binding domain of a guanylyl-cyclase-coupled receptor. Nature 2000;406:101–104.
40. Kunishima N, Shimada Y, Tsuji Y, Sato T, Yamamoto M, Kumasaka T, Nakanishi S, Jingami H, Morikawa K. Structural basis of glutamate recognition by a dimeric metabotropic glutamate receptor. Nature 2000;407:971–977.
41. Rajagopal S, Vishveshwara S. Short hydrogen bonds in proteins. FEBS J 2005;272:1819–1832.
42. Karpowich NK, Huang HH, Smith PC, Hunt JF. Crystal structures of the BtuF periplasmic-binding protein for vitamin B12 suggest a functionally important reduction in protein mobility upon ligand binding. J Biol Chem 2003;278:8429–8434.
43. Clarke TE, Rohrbach MR, Tari LW, Vogel HJ, Koster W. Ferric hydroxamate binding protein FhuD from *Escherichia coli*: mutants in conserved and non-conserved regions. Biometals 2002;15:121–131.
44. Locher KP. Structure and mechanism of ABC transporters. Curr Opin Struct Biol 2004;14:426–431.
45. Quiocho FA, Ledvina PS. Atomic structure and specificity of bacterial periplasmic receptors for active transport and chemotaxis: variation of common themes. Mol Microbiol 1996;20:17–25.
46. Baker MD, Wolanin PM, Stock JB. Signal transduction in bacterial chemotaxis. Bioessays 2006;28:9–22.
47. Neiditch MB, Federle MJ, Miller ST, Bassler BL, Hughson FM. Regulation of LuxPQ receptor activity by the quorum-sensing signal autoinducer-2. Mol Cell 2005;18:507–518.
48. Hu Y, Chen L, Ha S, Gross B, Falcone B, Walker D, Mokhtarzadeh M, Walker S. Crystal structure of the MurG:UDP-GlcNAc complex reveals common structural principles of a superfamily of glycosyltransferases. Proc Natl Acad Sci USA 2003;100:845–849.
49. Misono KS, Ogawa H, Qiu Y, Ogata CM. Structural studies of the natriuretic peptide receptor: a novel hormone-induced rotation mechanism for transmembrane signal transduction. Peptides 2005; 26:957–968.
50. Deka RK, Neil L, Hagman KE, Machius M, Tomchick DR, Brautigam CA, Norgard MV. Structural evidence that the 32-kilodalton lipoprotein (Tp32) of *Treponema pallidum* is an L-methionine-binding protein. J Biol Chem 2004;279:55644–55650.
51. Suhre K, Sanejouand YH. *ElNémo*: a normal mode web server for protein movement analysis and the generation of templates for molecular replacement. Nucleic Acids Res 2004;32:W610–W614.
52. Onufryk C, Crouch ML, Fang FC, Gross CA. Characterization of six lipoproteins in the σ^E regulon. J Bacteriol 2005;187:4552–4561.

53. Keyamura K, Fujikawa N, Ishida T, Ozaki S, Su'etsugu M, Fujimitsu K, Kagawa W, Yokoyama S, Kurumizaka H, Katayama T. The interaction of DiaA and DnaA regulates the replication cycle in *E. coli* by directly promoting ATP DnaA-specific initiation complexes. Genes Dev 2007;21:2083–2099.
54. Matias VR, Al-Amoudi A, Dubochet J, Beveridge TJ. Cryo-transmission electron microscopy of frozen-hydrated sections of *Escherichia coli* and *Pseudomonas aeruginosa*. J Bacteriol 2003;185:6112–6118.
55. Park JT, Raychaudhuri D, Li H, Normark S, Mengin-Lecreulx D. MppA, a periplasmic binding protein essential for import of the bacterial cell wall peptide L-alanyl-γ-D-glutamyl-*meso*-diaminopimelate. J Bacteriol 1998;180:1215–1223.
56. Hanson MS, Slaughter C, Hansen EJ. The *hbpA* gene of *Haemophilus influenzae* type b encodes a heme-binding lipoprotein conserved among heme-dependent *Haemophilus* species. Infect Immun 1992;60:2257–2266.
57. Morton DJ, Madore LL, Smith A, Vanwagoner TM, Seale TW, Whitby PW, Stull TL. The heme-binding lipoprotein (HbpA) of *Haemophilus influenzae*: role in heme utilization. FEMS Microbiol Lett 2005;253:193–199.
58. Pattery T, Hernalsteens JP, De Greve H. Identification and molecular characterization of a novel *Salmonella enteritidis* pathogenicity islet encoding an ABC transporter. Mol Microbiol 1999;33:791–805.
59. Saravanamuthu SS. Proteome analysis of differential adaptation of *Pseudomonas aeruginosa* morphotypes isolated from cystic fibrosis patients, to iron limiting conditions, oxidative stress and anaerobic conditions. Doctoral dissertation. Hannover, Germany: Hannover University; 2004.
60. Seydel A, Gounon P, Pugsley AP. Testing the '+2 rule' for lipoprotein sorting in the *Escherichia coli* cell envelope with a new genetic selection. Mol Microbiol 1999;34:810–821.
61. Marlovits TC, Kubori T, Sukhan A, Thomas DR, Galán JE, Unger VM. Structural insights into the assembly of the type III secretion needle complex. Science 2004;306:1040–1042.
62. Collins RF, Beis K, Dong C, Botting CH, McDonnell C, Ford RC, Clarke BR, Whitfield C, Naismith JH. The 3D structure of a periplasm-spanning platform required for assembly of group 1 capsular polysaccharides in *Escherichia coli*. Proc Natl Acad Sci USA 2007; 104:2390–2395.