The Crystal Structure of the Periplasmic Domain of the *Escherichia coli* Membrane Protein Insertase YidC Contains a Substrate Binding Cleft

Stephanie Ravaud, Goran Stjepanovic, Klemens Wild and Irmgard Sinning

Access the most updated version of this article at doi: 10.1074/jbc.M710493200

Find articles, minireviews, Reflections and Classics on similar topics on the JBC Affinity Sites.

Alerts:

- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

Supplemental material:

http://www.jbc.org/content/suppl/2008/01/31/M710493200.DC1.html

This article cites 74 references, 30 of which can be accessed free at

http://www.jbc.org/content/283/14/9350.full.html#ref-list-1

The Crystal Structure of the Periplasmic Domain of the *Escherichia coli* Membrane Protein Insertase YidC Contains a Substrate Binding Cleft*[S]

Stephanie Ravaud, Goran Stjepanovic, Klemens Wild, and Irmgard Sinning[1]
From the Biochemie-Zentrum der Universität Heidelberg, Im Neuenheimer Feld 328, D-69120 Heidelberg, Germany

In bacteria the biogenesis of inner membrane proteins requires targeting and insertion factors such as the signal recognition particle and the Sec translocon. YidC is an essential membrane protein involved in the insertion of inner membrane proteins together with the Sec translocon, but also as a separate entity. YidC of *Escherichia coli* is a member of the conserved YidC (in bacteria)/Oxa1 (in mitochondria)/Alb3 (in chloroplasts) protein family and contains six transmembrane segments and a large periplasmic domain (P1). We determined the crystal structure of the periplasmic domain of YidC from *E. coli* (P1D) at 1.8 Å resolution. The structure of P1D shows the conserved β-supersandwich fold of carbohydrate-binding proteins and an α-helical linker region at the C terminus that packs against the β-supersandwich by a highly conserved interface. P1D exhibits an elongated cleft of similar architecture as found in the structural homologs. However, the electrostatic properties and molecular details of the cleft make it unlikely to interact with carbohydrate substrates. The cleft in P1D is occupied by a polyethylene glycol molecule suggesting an elongated peptide or acyl chain as a natural ligand. The region of P1D previously reported to interact with SecF maps to a surface area in the vicinity of the cleft. The conserved C-terminal region of the P1 domain was reported to be essential for the membrane insertase function of YidC. The analysis of this region suggests a role in membrane interaction and/or in the regulation of YidC interaction with binding partners.

Membrane proteins represent more than one-third of the gene-encoded proteome in most organisms and are essential for numerous fundamental biological processes (1). The insertion and assembly of membrane proteins is therefore of critical importance to all organisms. In *Escherichia coli*, the biogenesis of inner membrane proteins (IMPs)[2] is predominantly accomplished in a co-translational manner and involves three distinct steps (2): (i) membrane targeting, mediated by the signal recognition particle (SRP) and the SRP-receptor FtsY (3); (ii) insertion into the lipid bilayer by the Sec translocon, consisting of the protein-conducting channel SecYEG, the accessory complex SecDFYajC, and the ATPase SecA (4–6); and (iii) folding and final assembly into a lipid-embedded functional structure. Although the protein-conducting channel (SecYEG) is rather well characterized, the polytopic IMP YidC from *E. coli* was recently identified. It plays a central and versatile role during the integration, folding, and assembly of IMPs (7–12).

YidC associates with the Sec translocon (mainly with SecD-FYajC) and was suggested to operate downstream of the SecYEG channel, catalyzing the final anchoring of Sec-dependent substrates (*e.g.* FtsQ, Lep, MtlA, LacY) into the membrane and/or their folding into a physiological conformation (6, 13–15). However, YidC is present in excess over the Sec translocon (16) and has also been described as a Sec-independent membrane insertase (9). A number of IMPs utilize this alternative insertion pathway, including the small phage coat proteins M13 and Pf3 and the endogenous IMPs Foc and MscL (7, 17–19). YidC has also been shown to be involved in the targeting and translocation of lipoproteins (20). Therefore, YidC can be described as a chaperone or a channel (21, 22).

YidC is homologous to Alb3 and Oxa1, involved in the integration of proteins into the thylakoid membrane of chloroplasts and the inner membrane of mitochondria, respectively (11, 23). All YidC homologs contain five putative transmembrane (TM) segments that are thought to be essential for membrane protein insertion (21, 24–26). Unique regions at the N and C termini reflect specific requirements for the interaction with other targeting and insertion factors as well as for the topology of the protein substrates. Oxa1 contains a C-terminal domain for the interaction with the ribosome (27, 28). YidC from Gram-negative bacteria contains an additional N-terminal TM helix and a periplasmic domain (P1) between TM1 and TM2 (29). The P1 domain of *E. coli* YidC was shown to interact with SecF (30). Interestingly, deletion of TM1 and of 90% of the P1 domain does not abolish the insertase function (25). However, deletion of the C-terminal region of the P1 domain impairs cell viability and membrane insertion of a number of Sec-dependent and -independent substrates (30). The introduction of site-specific

The atomic coordinates and structure factors (code 3BS6) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Tables S1 and S2 and supplemental references.

[1] To whom correspondence should be addressed. Tel.: 49-6221-544780; Fax: 49-6221-544790; E-mail: irmi.sinning@bzh.uni-heidelberg.de.

[2] The abbreviations used are: IMP, inner membrane protein; TM, transmembrane; PEG, polyethylene glycol.

protease sites at the borders of the P1 domain results in a cold-sensitive YidC mutant with altered insertase activity (31). Taken together, the role of YidC in membrane protein insertion and, specifically, the role of the P1 domain is still not well understood. To date, *E. coli* YidC is the best characterized member of the Oxa1/Alb3/YidC family. Only a projection structure at low resolution is available from cryo-electron microscopy that suggests that the YidC monomer is a membrane pore that might be able to associate with itself (forming homodimers) or with the Sec translocon, depending on the protein substrate (32). The structural arrangement of YidC and specifically of the P1 domain is, however, not resolved.

Here we report the crystal structure of the P1 domain of YidC from *E. coli* at 1.8 Å resolution. The structure provides insights into the role of the P1 domain for YidC function.

EXPERIMENTAL PROCEDURES

Structure Determination and Refinement—The cloning, overexpression, purification, and crystallization of P1D (residues 56 –329 of *E. coli* YidC) and its selenomethionine (SeMet) derivative were performed as described (73). The structure of P1D was determined using single wavelength anomalous diffraction data collected on a SeMet crystal at the peak wavelength ($\lambda = 0.9790$ Å) and the PHENIX (Python-based Hierarchical Environment for Integrated Xtallography) program suite (33, 34). Details of data collection and processing are given elsewhere (73). The eight selenium sites (4 selenium/molecule; 2 molecules in the asymmetric unit) present in the asymmetric unit and the non-crystallographic symmetry relations were identified and the phases were calculated. Density modification, phase extension, and automatic model building were carried out using a remote data set collected at 1.8 Å resolution and previously scaled to the peak wavelength data using SCALEIT (35). The AUTOBUILD routine from PHENIX was able to build 473 residues (and the Cα traces of 16 additional amino acids) for the two molecules of the asymmetric unit. The model was completed using iterative cycles of model building in Coot (36), and refinement was done with REFMAC5 (35, 37). The two protein chains were refined independently. The model quality was analyzed with PROCHECK (38). The refinement statistics are summarized in supplemental Table 1.

Structure Analysis—The secondary structure was analyzed with DSSP (39); the multiple sequence alignments were prepared with ClustalW (40) with sequences retrieved from a PSI BLAST search and rendered with ESPript (41). Surface potentials were calculated with GRASP (42). Conservation scores were determined using the ConSurf server on the basis of 40 sequences of P1 domains from Gram-negative bacteria. Residues interacting with the polyethylene glycol (PEG) molecule are according to LIGPLOT analysis (43). The PEG/protein and protein/protein interactions within the crystal packing were analyzed with the PISA server (44). Figures were generated with PyMOL.

Characterization of P1D in Solution—The apparent molecular mass was analyzed by size exclusion chromatography using a Superdex 200 HR 10/30 column (GE Healthcare) in 10 mM Hepes, pH 7.5, 200 mM NaCl (buffer A) with ± 10 mM CaCl₂. The absolute molecular mass was determined by static light scattering (mini Dawn Tristar; Wyatt Technologies) and refractive index analysis (WGE Dr. Bures Δn1000, Dallgow, Germany). The oligomeric state of P1D in solution was determined by analytical ultracentrifugation experiments using a Beckman Optima XL-A ultracentrifuge equipped with absorbance optics and an An60 Ti rotor (Beckman Coulter, Fullerton, CA). Sedimentation velocity studies were carried out at 35,000 rpm at 20 °C in three different buffers (buffer A, buffer A + 1 mM CaCl₂, and buffer A + 50 mM CaCl₂).

RESULTS AND DISCUSSION

Overall Structure of P1D—*E. coli* YidC has a molecular mass of 61 kDa and consists of six transmembrane helices. The P1 domain comprises residues 24 –342 and is inserted in the periplasmic loop between TM1 and TM2 (29) (Fig. 1*A*). Primary structure analysis revealed that the N-terminal region of the P1 domain (residues 26 –55) contains a low complexity region predicted to be unstructured according to multiple sequence alignments and secondary structure predictions. The C-terminal region of the P1 domain could not be assigned unambiguously (see below). Therefore, two constructs of the P1 domain were expressed and purified, representing residues 56 –329 and 56 –342, respectively (73). Although the longer one suffered from aggregation problems, the shorter construct (P1D, residues 56 –329) yielded soluble protein and diffraction quality crystals (73). The crystal structure was determined at 1.8 Å resolution using single wavelength anomalous diffraction and refined using standard procedures. The refinement statistics are given in the online supplemental Table S1.

The P1D structure is well ordered (overall B-factor of 17.2 Å²) with the exception of residues 204 –216, which are part of a flexible loop. P1D has an overall globular shape with approximate dimensions of 60 × 50 × 40 Å (Fig. 1, *B* and *C*). The major part (residues 56 –310) consists of two twisted anti-parallel β-sheets, S1 and S2. S1 (11-stranded β-sheet, major) and S2 (8-stranded β-sheet, minor) pack against each other and form a β-supersandwich fold (Structural Classification of Proteins data base) (45). The curvature of the two β-sheets creates a concave surface on the S1 side and a convex surface on the S2 side. A number of irregular loops as well as a 3₁₀ helix (η1) between β5 and β6 and one α helix (α1) between β12 to β13 connect the β-strands. Sequence analysis using different bioinformatic tools did not allow prediction of this fold. The C terminus of the P1D structure (residues 311 –329) consists of the helices α2 and α3, which are packed against the S2 layer of the β-supersandwich by hydrophobic interactions (see below, Fig. 6*B*). The helix α3 is truncated and forms a 3₁₀ helix in P1D, but it is predicted to extend to residue 338 toward TM2 in full-length YidC (see below, Fig. 6*A*). Therefore, we refer to the C-terminal, helical part of P1D as the linker region. P1D is conserved in YidC of Gram-negative bacteria (Fig. 1*D*). More than 40 homologous sequences were identified by PSI-BLAST analysis, with sequence identities varying from 20 to 99%. Structure-based sequence alignments suggest that they all share the topology of the P1D structure with a protein core adopting a β-supersandwich fold and a helical linker region at the C terminus.

Analysis of Crystal Packing—YidC has been reported to exist as a monomer or dimer *in vivo*; however, its oligomeric state is
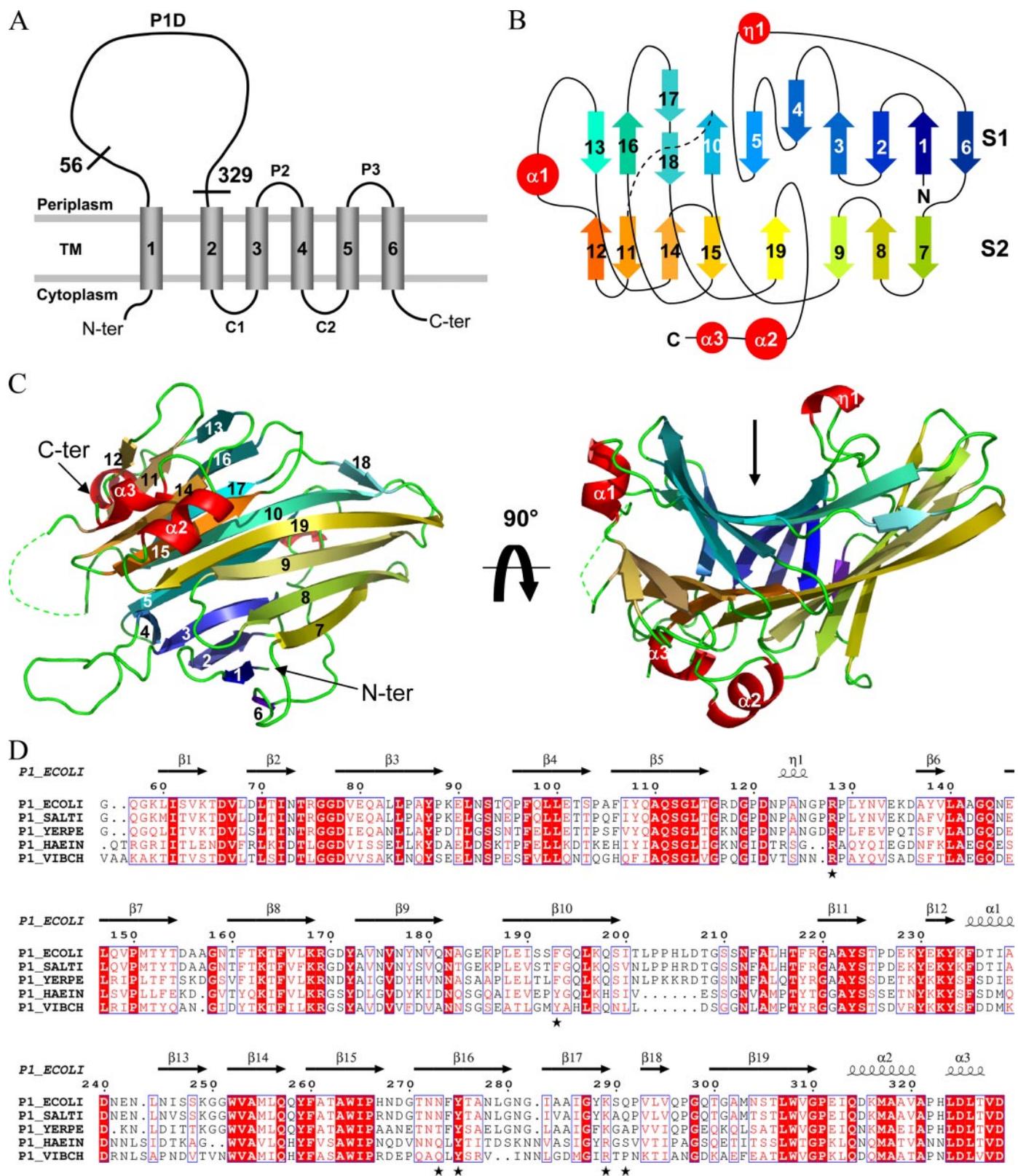
FIGURE 1. **Structure of the periplasmic domain (P1D) of YidC.** *A*, schematic of full-length YidC according to Sääf *et al.* (29). The domain boundaries of P1D (*residues 56 –329*) are indicated. *B*, topology diagram of P1D. β-strands are represented as *arrows*, helices as *red circles*, and loops as *lines*. The S1 and S2 layers of the β-supersandwich motif are colored in *blue/green* and *yellow/green* shading, respectively. The β10–β11 loop is depicted as *dotted lines*. *C*, ribbon representation of P1D. The colors are the same as in *B*. The view on the right is rotated by 90° along a horizontal axis with respect to the view on the left. The position of the P1D cleft on the concave side (S1) of the structure is indicated by an *arrow*. N and C termini are labeled. The missing part of the flexible β10–β11 loop is indicated as *dotted lines*. *D*, sequence alignment of selected periplasmic domains of YidC from Gram-negative bacteria (*E. coli*, ECOLI; *Salmonella typhimurium*, SALTI; *Yersinia pestis*, YERPE; *Haemophilus influenzae*, HAEIN; and *Vibrio cholerae*, VIBCH). Secondary structure elements in P1D of *E. coli* are indicated above. P1D residues interacting with the PEG molecule are marked by *stars*. Numbering is according to the sequence of *E. coli* YidC.
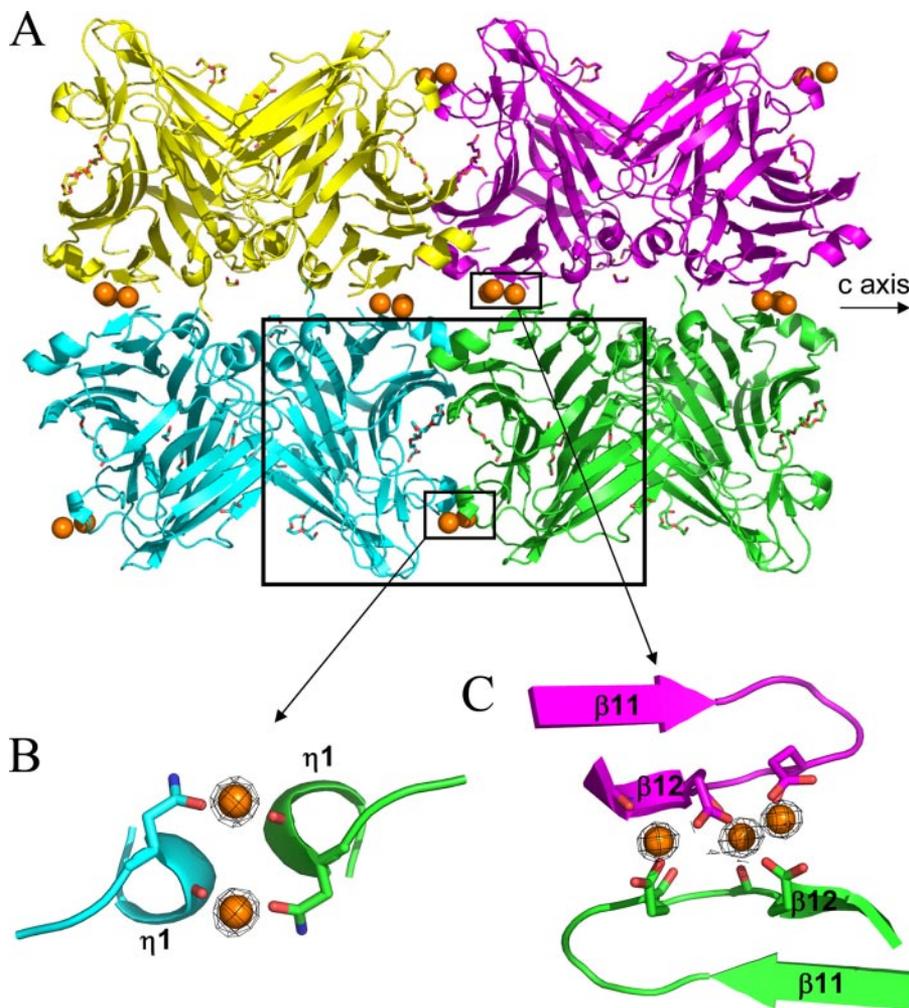
FIGURE 2. **Crystal packing of P1D.** *A*, ribbon representation of part of the crystal-packing contacts. Each *color* represents one asymmetric unit. Calcium ions are displayed as *orange balls* and the PEG molecules as *sticks*. The crystallographic *c* axis is indicated by an *arrow*. The energetically favorable dimer interface is highlighted by a *box* at the bottom, containing two monomers (*blue* and *green*) from different asymmetric units. *B* and *C*, detailed views on two different calcium binding sites. Final $2 F_o - F_c$ map contoured at 3σ shows the electron densities for the calcium ions.

exclusion chromatography and analytical ultracentrifugation experiments (data not shown). The interfaces between molecules A and B in the crystal bury around 1000 Å^2, which is in the lower range of values described for oligomeric proteins. However, the analysis of the structural and chemical properties and of the probable dissociation pattern of the different assemblies revealed that one of them is energetically favorable. In this case, the concave sides of two A and B molecules are facing each other, forming a large central cavity (Fig. 2*A*). The biological relevance of such an interaction needs to be investigated.

P1D Shares the Fold of Carbohydrate-binding Proteins—Analysis of the P1D structure using the Dali server (47) identified a number of proteins from the galactose mutarotase-like family as structural relatives (supplemental Table S2). Comparison of the three-dimensional structures and topologies shows that P1D is similar to the lectin-like domains found in calnexin, an endoplasmic reticulum chaperone involved in quality control of protein folding (48); in neurexin, a putative cell recognition molecule (49); and also in the carbohydrate recognition domain of p58/ERGIC-53, an animal lectin involved in glycoprotein export from the endoplasmic reticulum (50). Number and length of the β-strands and/or of the connectivities differ between P1D and these structures as reflected in the high root mean square deviation of $\sim$3.5–4 Å, and the sequence identities are very low (<12%). However, the main characteristics of the fold are maintained: a pair of antiparallel β-sheets curved to form a concave and convex side (Fig. 3). The β-sandwich motif was reported to have two putative functions. First, in maltose or chitobiose phosphorylases, rhamnogalacturonase, glucoamylase, and galactosidase (supplemental Table S2), it participates in intramolecular interactions and was related to transglycosylation, thermostability, or fold correction (51–54). Second, it is involved in sugar recognition (lectin-like domains) and plays a role in catalysis (galactose mutarotase, aldose epimerase, OpgG). Carbohydrate recognition occurs on the concave side of the domain, in a large open cleft in the lectin family, whereas in enzymes like galactose mutarotase or OpgG the cleft is narrower and more shielded by loops. The carbohydrate binding residues are not strictly conserved among the different families but typically involve a highly negatively charged surface (Fig. 3*C*). Acidic residues are involved in hydrogen bonding interac-

still a matter of debate (8, 23, 32, 46). P1D is monomeric in solution according to size exclusion chromatography, static light scattering, and sedimentation velocity experiments (data not shown). However, two molecules are present in the asymmetric unit of the crystal (molecules A and B), related by a 2-fold non-crystallographic symmetry axis. They have very similar structures with a root mean square deviation of 0.5 Å and interact by the following regions: $\alpha 2$, $\beta 19$, the loop between $\beta 6$ and $\beta 7$, and the loop between $\beta 16$ and $\beta 17$ (Fig. 2*A*). P1D interactions between adjacent asymmetric units involve different regions and are favored by calcium ions (Fig. 2). Two calcium ions are bound between the $\eta 1$ helix of chain A and the $\eta 1$ helix of chain B of an adjacent asymmetric unit (Fig. 2*B*). The three other calcium ions were observed at the interfaces between the $\beta 11$–$\beta 12$ loops of two molecules, A and B, from two adjacent asymmetric units (Fig. 2*C*). Because crystals were only obtained (or stable) in the presence of calcium ions, they are likely to promote crystallization by tightening the interaction between two P1D molecules. However, calcium ions are unable to promote dimerization in solution as tested by size

FIGURE 3. **Structural comparison of P1D (*middle*) with galactose mutarotase (*left*) and the carbohydrate recognition domain of the animal lectin p58/ERGIC-53 (*right*).** *A* and *B*, the structures (*1*, Protein Data Bank entry 1L7K; *2*, 3BS6; and *3*, 1GV9) have been superimposed and are shown side by side in a ribbon representation with β-strands in *yellow* and helices in *red*. Glucose and PEG molecules are bound in the cleft of the galactose mutarotase and P1D, respectively. They are shown in *stick representation* and colored as elements. In *A* the same orientation is used as in Fig. 1C, *right panel*. *B* is rotated by 90° along a horizontal axis. *C*, same view as in *B*. The molecular surface is colored *blue* and *red* according to positive and negative electrostatic potential, respectively.

tions with the substrate whereas aromatic residues participate in stacking interactions with the carbohydrate ring.

Notably, P1D exhibits a large cleft at the corresponding position (~30 × 10 Å, with a depth of 3–7 Å; Figs. 1C and 3). The shape and dimensions of the P1D cleft are similar to those observed in the lectin-like domain. However, the negative surface potential is not maintained in P1D (Fig. 3C) and the residues involved in carbohydrate binding are not conserved. The analysis of the P1D cleft shows a high conservation within the P1 domains of YidC (Fig. 4A) with a major contribution of hydrophobic residues (Fig. 3C). Lys-289 introduces a positive charge in the middle of the cleft. From the molecular details of the P1D cleft, it is unlikely to accommodate polysaccharides.

Although P1D shares the protein fold and the overall position and architecture of the binding pocket with its structural homologs, it seems designed for different ligands. The dimensions and electrostatic properties of the cleft make it well suited to accommodate an extended and predominantly unpolar molecule.

The P1D Cleft Contains a Ligand—The structure of P1D contains an elongated electron density in the binding cleft (Fig. 4). It was interpreted as an ordered polyethylene glycol molecule (PEG 400) that arises from the crystallization buffer. The PEG/protein interface corresponds to a buried surface area of 807 Å^2 and is formed by 22 residues mainly from the S1 layer. The interactions are predominantly hydrophobic (including Phe-
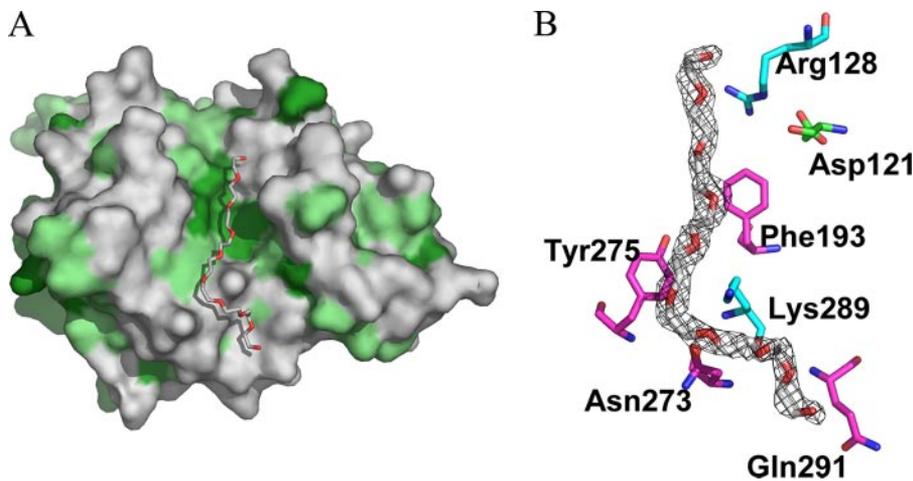
FIGURE 4. **The binding cleft of P1D.** *A*, the degree of sequence conservation within the P1 domain in Gram-negative bacteria is mapped onto the surface of the P1D structure. *Dark green* and *light green* indicate residues that are highly or partially conserved, respectively. *B*, the binding cleft of P1D contains polyethylene glycol. The PEG400 molecule is shown in *stick* representation. Hydrophobic residues and H bond donors are colored in *magenta* and *cyan*, respectively. The final $2F_o - F_c$ electron density contoured at 1σ is shown for the PEG molecule.
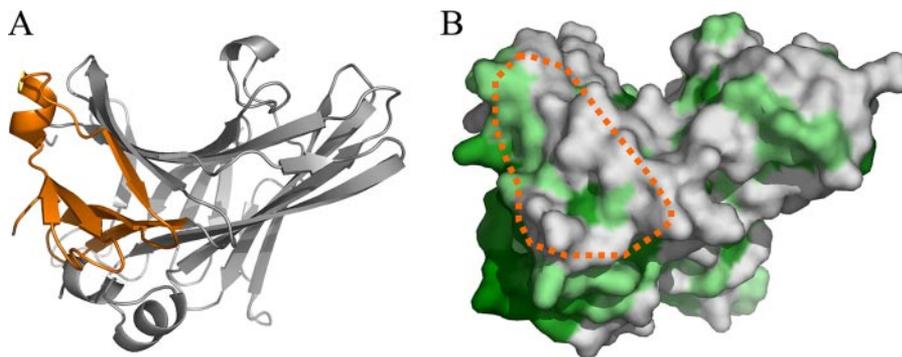
FIGURE 5. **Proposed interaction site between P1D and SecF.** *A*, the interaction region according to Xie *et al.* (30) is colored in *orange* while the rest of the P1D structure is colored in *gray*. *B*, same view as *A*. The degree of sequence conservation is mapped onto the surface of the P1D structure as for Fig. 4*A*. The proposed interaction site is contoured.

chain from an interacting protein or from an unfolded polypeptide and therefore act as chaperone. All the Sec-independent YidC substrates identified to date contain short periplasmic tails that from their primary structure could interact with the P1D cleft (19, 64, 65). Other potential ligands include Sec-dependent YidC substrates, *e.g.* lipoproteins (20) or periplasmic proteins such as periplasmic folding factors (66). Because deletion of the major part of P1D corresponding to the β-supersandwich domain does not impair the insertase function of YidC (30), it is tempting to assign an independent function to the β-supersandwich region present only in Gram-negative bacteria. However, only a small number of YidC substrates are known and have been tested for P1D requirement.

Interaction of P1D with SecF— The association of *E. coli* YidC with the Sec translocon seems unique within the YidC/Oxa1/Alb3 family, and the P1 domain might play a role in this interaction (21). The P1 domain was shown to interact with SecF based on copurification experiments with the SecDFyajC complex (67). Residues 215–265 of the P1 domain were identified as the region of interaction by analysis of deletion mutants (30). However, the

193, Tyr-275, Asn-273, and Gln-291) except for hydrogen bonds formed between the PEG molecule and Arg-128 and Lys-289, and two hydrogen bonds mediated by water molecules (Fig. 4*B*). The center of the cleft and especially the two aromatic residues interacting with the PEG molecule (Phe-193 and Tyr-275) are highly conserved in the P1 domains of YidC. Positively charged residues (Lys, Arg, or Gln) are mainly found at position 289 but are replaced in some bacteria by Ile, Val, or Thr, further enhancing the hydrophobic character of the cleft. The high degree of conservation suggests that all P1 domains bind similar ligands (Figs. 1*D* and 4*A*).

PEG molecules are known to occupy clefts that naturally bind elongated polymers such as either polypeptides in chaperones (55, 56), in peptide deformylase (57, 58), and in the neuronal calcium sensor (59) or long acyl chains in enzymes (59–61) and in a periplasmic lipoprotein localization factor (62). The interaction network described for these examples is similar to the one observed in the P1D structure. The PEG molecule described here could therefore mimic a natural P1D ligand. P1D may bind acyl chains of peptidoglycans or lipopolysaccharides as described for the periplasmic folding factor Skp (63). Alternatively, P1D could accommodate an elongated peptide

P1D structure clearly indicates that these deletions severely affect or even destroy the β-supersandwich fold. The proposed interaction site localizes to an exposed edge of the β-supersandwich fold and includes the end of the flexible loop $\beta10$– $\beta11$; helix $\alpha1$, which is involved in a crystallographic contact; and the β-strands 11–15 (Fig. 5*A*). This region is not well conserved in YidC of Gram-negative bacteria (Fig. 5*B*) with the exception of β-strands $\beta14$ and $\beta15$, which are involved in hydrophobic interactions with the C-terminal linker region. Although the proposed interaction site localizes in the vicinity of the cleft, it does not overlap with it. Therefore, P1D might be able to interact with different binding partners independently but also simultaneously. In an effort to analyze the SecF/P1D interaction in more detail, we used one-dimensional NMR titrations and isothermal titration calorimetry experiments. However, a specific interaction of the periplasmic domain of SecF (residues 48–140) with P1D was not observed (data not shown). If this interaction occurs *in vivo* it might have a low binding affinity or it might be transient as also previously suggested (30).

Membrane Interaction of the P1 Domain—The region covering residues 323–346 of *E. coli* YidC was shown to be essential

FIGURE 6. **Analysis of the C-terminal linker region.** *A*, sequence alignment and secondary structure prediction for the YidC/Oxa1/Alb3 family (*E. coli* YidC for Gram-negative bacteria, *Bacillus subtilis* SpoIIIJ for Gram-positive bacteria, *Saccharomyces cerevisiae* Oxa1 for mitochondria, and *Arabidopsis thaliana* Alb3 for chloroplast). A multiple alignment was originally generated with 160 YidC sequences. Secondary structure elements in P1D structure (*YIDC_ECOLI*) are indicated at the *top*. The *cylinders* represent predicted α helices (PredictProtein server) (68). Potential starting points of the transmembrane segment are indicated by *arrows*. They were predicted using four algorithms (*1*, HMMTOP; *2*, TopPred/TMHMM1; *3*, PHD). *B*, interaction of the helices α2 and α3 with the β-supersandwich region. The hydrophobic residues in the interface are colored in *magenta*. *C*, analysis of the C-terminal part of the P1 domain by the Amphipaseek server (69). The *upper panel* shows the prediction score, and the *lower panel* represents this region as a helical wheel. *D*, models for the orientation of P1D toward the membrane. The position of the cleft is indicated by an *arrow*.

for cell growth and membrane insertion of both Sec-dependent and Sec-independent substrates (25, 30). It locates at the C terminus of the periplasmic domain and is conserved within the YidC/Oxa1/Alb3 family (Figs. 1*D* and 6*A*). Residues 323–329 form the α3 helix in the P1D structure and together with the helix α2 pack against the β-supersandwich via a hydrophobic and highly conserved interface (Fig. 6*B*). The region including

residues 330–346 is not unambiguously defined because it is either connecting P1D to TM2 or it is already part of TM2, as suggested by the standard programs (Fig. 6*A*). We therefore analyzed this sequence in more detail by different prediction tools (68, 69) (Fig. 6, *A* and *C*). The analysis shows that α3 is likely to extend to residue 338 with a hydrophobic C terminus favorable for membrane interaction. The adjacent sequence

(residues 340–355) is predicted as an amphipathic helix (Fig. 6*C*) that could attach or even insert into the membrane (70, 71). This region is conserved in YidC homologs that do not have a P1D, suggesting an important and conserved role of this region for the function of YidC (see above). In addition, we tested whether P1D contributes to the membrane interaction. Flotation assays indicated that P1D alone is interacting with *E. coli* membrane lipids (data not shown) (experimental conditions as in Parlitz *et al.*; Ref. 72). For the proposed functions of P1D it would be interesting to know how P1D, and especially the binding cleft, orient with respect to the TM part of YidC. Based on the analyses reported here, we derived two models (Fig. 6*D*). In Model 1, the cleft of P1D is accessible from the membrane pore and therefore could participate in folding of YidC substrates (see above). In Model 2, the cleft is oriented more away from the membrane and could serve as an interaction site for periplasmic molecules (peptidoglycans or lipopolysaccharides) or proteins (see above). More experiments are needed to test these models.

CONCLUSIONS

The structure of P1D consists of a β-supersandwich fold similar to carbohydrate binding proteins and an α-helical linker region. P1D contains an elongated binding cleft occupied by a PEG molecule that could mimic a natural ligand. The P1D cleft could be utilized by substrate proteins according to a membrane chaperone function of YidC, which would require a location of the cleft in close proximity of the membrane pore (Fig. 6*D*, *Model 1*). The cleft could also bind periplasmic molecules or proteins pointing to an additional function (Fig. 6*D*, *Model 2*). The C-terminal part of the linker region is important for the YidC insertase function and is conserved in YidC homologs in chloroplasts and mitochondria, suggesting a functional role conserved within the YidC/Oxa1/Alb3 family. In *E. coli* YidC, the C-terminal region of the P1 domain might be involved in orienting P1D with respect to the membrane, probably through an amphipathic region adjacent to TM2. The structure presented here provides the basis for a detailed analysis of the periplasmic domain and its importance for the multiple functions of YidC proteins.

Addendum—Another structure of the E.coli YidC periplasmic domain has been recently published (74).

REFERENCES

1. Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L. (2001) *J. Mol. Biol.* **305,** 567–580
2. Luirink, J., von Heijne, G., Houben, E., and de Gier, J. W. (2005) *Annu. Rev. Microbiol.* **59,** 329–355
3. Wild, K., Halic, M., Sinning, I., and Beckmann, R. (2004) *Nat. Struct. Mol. Biol.* **11,** 1049–1053
4. Veenendaal, A. K., van der Does, C., and Driessen, A. J. (2004) *Biochim. Biophys. Acta* **1694,** 81–95
5. Dalbey, R. E., and Chen, M. (2004) *Biochim. Biophys. Acta* **1694,** 37–53
6. Van den Berg, B., Clemons, W. M., Jr., Collinson, I., Modis, Y., Hartmann, E., Harrison, S. C., and Rapoport, T. A. (2004) *Nature* **427,** 36–44
7. Samuelson, J. C., Chen, M., Jiang, F., Moller, I., Wiedmann, M., Kuhn, A., Phillips, G. J., and Dalbey, R. E. (2000) *Nature* **406,** 637–641
8. Chen, M., Xie, K., Jiang, F., Yi, L., and Dalbey, R. E. (2002) *Biol. Chem.* **383,** 1565–1572
9. Serek, J., Bauer-Manz, G., Struhalla, G., van den Berg, L., Kiefer, D., Dalbey, R., and Kuhn, A. (2004) *EMBO J.* **23,** 294–301
10. van der Laan, M., Nouwen, N. P., and Driessen, A. J. (2005) *Curr. Opin. Microbiol.* **8,** 182–187
11. Yi, L., and Dalbey, R. E. (2005) *Mol. Membr. Biol.* **22,** 101–111
12. Kiefer, D., and Kuhn, A. (2007) *Int. Rev. Cytol.* **259,** 113–138
13. Urbanus, M. L., Scotti, P. A., Froderberg, L., Sääf, A., de Gier, J. W., Brunner, J., Samuelson, J. C., Dalbey, R. E., Oudega, B., and Luirink, J. (2001) *EMBO Rep.* **2,** 524–529
14. Beck, K., Eisner, G., Trescher, D., Dalbey, R. E., Brunner, J., and Muller, M. (2001) *EMBO Rep.* **2,** 709–714
15. van Bloois, E., Haan, G. J., de Gier, J. W., Oudega, B., and Luirink, J. (2006) *J. Biol. Chem.* **281,** 10002–10009
16. Urbanus, M. L., Froderberg, L., Drew, D., Bjork, P., de Gier, J. W., Brunner, J., Oudega, B., and Luirink, J. (2002) *J. Biol. Chem.* **277,** 12718–12723
17. van der Laan, M., Bechtluft, P., Kol, S., Nouwen, N., and Driessen, A. J. (2004) *J. Cell Biol.* **165,** 213–222
18. Chen, M., Samuelson, J. C., Jiang, F., Muller, M., Kuhn, A., and Dalbey, R. E. (2002) *J. Biol. Chem.* **277,** 7670–7675
19. Facey, S. J., Neugebauer, S. A., Krauss, S., and Kuhn, A. (2007) *J. Mol. Biol.* **365,** 995–1004
20. Froderberg, L., Houben, E. N., Baars, L., Luirink, J., and de Gier, J. W. (2004) *J. Biol. Chem.* **279,** 31026–31032
21. Kuhn, A., Stuart, R., Henry, R., and Dalbey, R. E. (2003) *Trends Cell Biol.* **13,** 510–516
22. Dalbey, R. E., and Kuhn, A. (2004) *J. Cell Biol.* **166,** 769–774
23. Luirink, J., Samuelsson, T., and de Gier, J. W. (2001) *FEBS Lett.* **501,** 1–5
24. Jiang, F., Yi, L., Moore, M., Chen, M., Rohl, T., Van Wijk, K. J., De Gier, J. W., Henry, R., and Dalbey, R. E. (2002) *J. Biol. Chem.* **277,** 19281–19288
25. Jiang, F., Chen, M., Yi, L., de Gier, J. W., Kuhn, A., and Dalbey, R. E. (2003) *J. Biol. Chem.* **278,** 48965–48972
26. Yuan, J., Phillips, G. J., and Dalbey, R. E. (2007) *J. Bacteriol.* **189,** 8961–8972
27. Jia, L., Dienhart, M., Schramp, M., McCauley, M., Hell, K., and Stuart, R. A. (2003) *EMBO J.* **22,** 6438–6447
28. Preuss, M., Ott, M., Funes, S., Luirink, J., and Herrmann, J. M. (2005) *J. Biol. Chem.* **280,** 13004–13011
29. Sääf, A., Monne, M., de Gier, J. W., and von Heijne, G. (1998) *J. Biol. Chem.* **273,** 30415–30418
30. Xie, K., Kiefer, D., Nagler, G., Dalbey, R. E., and Kuhn, A. (2006) *Biochemistry* **45,** 13401–13408
31. Chen, M., Xie, K., Nouwen, N., Driessen, A. J., and Dalbey, R. E. (2003) *J. Biol. Chem.* **278,** 23295–23300
32. Lotz, M., Haase, W., Kuhlbrandt, W., and Collinson, I. (2008) *J. Mol. Biol.* **375,** 901–907
33. Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K., and Terwilliger, T. C. (2002) *Acta Cryst. Sect. D Biol. Crystallogr.* **58,** 1948–1954
34. Adams, P. D., Gopal, K., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Pai, R. K., Read, R. J., Romo, T. D.,

Sacchettini, J. C., Sauter, N. K., Storoni, L. C., and Terwilliger, T. C. (2004) *J. Synchrotron. Radiat.* **11,** 53–55

35. Collaborative Computational Project, N. 4 (1994) *Acta Cryst. Sect. D Biol. Crystallogr.* **50,** 760–763

36. Emsley, P., and Cowtan, K. (2004) *Acta Cryst. Sect. D Biol. Crystallogr.* **60,** 2126–2132

37. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Cryst. Sect. D Biol. Crystallogr.* **53,** 240–255

38. Laskowski, R. A., Moss, D. S., and Thornton, J. M. (1993) *J. Mol. Biol.* **231,** 1049–1067

39. Kabsch, W., and Sander, C. (1983) *Biopolymers* **22,** 2577–2637

40. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res.* **22,** 4673–4680

41. Gouet, P., Robert, X., and Courcelle, E. (2003) *Nucleic Acids Res.* **31,** 3320–3323

42. Nicholls, A., Sharp, K. A., and Honig, B. (1991) *Proteins* **11,** 281–296

43. Wallace, A. C., Laskowski, R. A., and Thornton, J. M. (1995) *Protein Eng.* **8,** 127–134

44. Krissinel, E., and Henrick, K. (2007) *J. Mol. Biol.* **372,** 774–797

45. Lo Conte, L., Ailey, B., Hubbard, T. J., Brenner, S. E., Murzin, A. G., and Chothia, C. (2000) *Nucleic Acids Res.* **28,** 257–259

46. van der Laan, M., Houben, E. N., Nouwen, N., Luirink, J., and Driessen, A. J. (2001) *EMBO Rep.* **2,** 519–523

47. Holm, L., and Sander, C. (1996) *Methods Enzymol.* **266,** 653–662

48. Schrag, J. D., Bergeron, J. J., Li, Y., Borisova, S., Hahn, M., Thomas, D. Y., and Cygler, M. (2001) *Mol. Cell* **8,** 633–644

49. Rudenko, G., Nguyen, T., Chelliah, Y., Sudhof, T. C., and Deisenhofer, J. (1999) *Cell* **99,** 93–101

50. Velloso, L. M., Svensson, K., Schneider, G., Pettersson, R. F., and Lindqvist, Y. (2002) *J. Biol. Chem.* **277,** 15979–15984

51. Aleshin, A. E., Feng, P. H., Honzatko, R. B., and Reilly, P. J. (2003) *J. Mol. Biol.* **327,** 61–73

52. Hidaka, M., Honda, Y., Kitaoka, M., Nirasawa, S., Hayashi, K., Wakagi, T., Shoun, H., and Fushinobu, S. (2004) *Structure* **12,** 937–947

53. Imamura, H., Fushinobu, S., Yamamoto, M., Kumasaka, T., Jeon, B. S., Wakagi, T., and Matsuzawa, H. (2003) *J. Biol. Chem.* **278,** 19378–19386

54. Juers, D. H., Huber, R. E., and Matthews, B. W. (1999) *Protein Sci.* **8,** 122–136

55. Dollins, D. E., Immormino, R. M., and Gewirth, D. T. (2005) *J. Biol. Chem.* **280,** 30438–30447

56. Tanaka, S., Sawaya, M. R., Kerfeld, C. A., and Yeates, T. O. (2007) *Acta Cryst. Sect. D Biol. Crystallogr.* **63,** 1109–1112

57. Becker, A., Schlichting, I., Kabsch, W., Groche, D., Schultz, S., and Wagner, A. F. (1998) *Nat. Struct. Biol.* **5,** 1053–1058

58. Becker, A., Schlichting, I., Kabsch, W., Schultz, S., and Wagner, A. F. (1998) *J. Biol. Chem.* **273,** 11413–11416

59. Bourne, Y., Dannenberg, J., Pollmann, V., Marchot, P., and Pongs, O. (2001) *J. Biol. Chem.* **276,** 11949–11955

60. Qiu, W., Zhou, M., Mazumdar, M., Azzi, A., Ghanmi, D., Luu-The, V., Labrie, F., and Lin, S. X. (2007) *J. Biol. Chem.* **282,** 8368–8379

61. Oliveira, M. A., Guimaraes, B. G., Cussiol, J. R., Medrano, F. J., Gozzo, F. C., and Netto, L. E. (2006) *J. Mol. Biol.* **359,** 433–445

62. Takeda, K., Miyatake, H., Yokota, N., Matsuyama, S., Tokuda, H., and Miki, K. (2003) *EMBO J.* **22,** 3199–3209

63. Walton, T. A., and Sousa, M. C. (2004) *Mol Cell* **15,** 367–374

64. Chen, M., Xie, K., Yuan, J., Yi, L., Facey, S. J., Pradel, N., Wu, L. F., Kuhn, A., and Dalbey, R. E. (2005) *Biochemistry* **44,** 10741–10749

65. van Bloois, E., ten Hagen-Jongman, C. M., and Luirink, J. (2007) *Biochem. Biophys. Res. Commun.* **362,** 727–733

66. Mogensen, J. E., and Otzen, D. E. (2005) *Mol. Microbiol.* **57,** 326–346

67. Nouwen, N., and Driessen, A. J. (2002) *Mol. Microbiol.* **44,** 1397–1405

68. Rost, B., Yachdav, G., and Liu, J. (2004) *Nucleic Acids Res.* **32,** W321–W326

69. Combet, C., Blanchet, C., Geourjon, C., and Deleage, G. (2000) *Trends Biochem. Sci.* **25,** 147–150

70. Vogt, B., Ducarme, P., Schinzel, S., Brasseur, R., and Bechinger, B. (2000) *Biophys. J.* **79,** 2644–2656

71. Bechinger, B. (2001) *FEBS Lett.* **504,** 161–165

72. Parlitz, R., Eitan, A., Stjepanovic, G., Bahari, L., Bange, G., Bibi, E., and Sinning, I. (2007) *J. Biol. Chem.* **282,** 32176–32184

73. Ravaud, S., Wild, K., and Sinning, I. (2008) *Acta Cryst. F Struct. Biol. Crystalliz. Comm.* **64,** 144–148

74. Oliver, D. C., and Paetzel, M. (2008) *J. Biol. Chem.* **283,** 5208–5216