*

30, 2007, — — 12, 2007 , BC , — 19, 2007, 10.1074/ . 708936200

1

From the Department of Molecular Biology and Biochemistry, Simon Fraser University, Burnaby, British Columbia V5A 1S6, Canada

The essential bacterial membrane protein YidC facilitates insertion and assembly of proteins destined for integration into the inner membrane. It has homologues in both mitochondria and chloroplasts. Here we report the crystal structure of the *Escherichia coli* YidC major periplasmic domain ($\mathrm{YidC_{EC}P1}$) at 2.5 Å resolution. This domain is present in YidC from Gram-negative bacteria and is more than half the size of the full-length protein. The structure reveals that $\mathrm{YidC_{EC}P1}$ is made up of a large twisted β-sandwich protein fold with a C-terminal α-helix that packs against one face of the β-sandwich. Our structure and sequence analysis reveals that the C-terminal α-helix and the β-sheet that it lays against are the most conserved regions of the domain. The region corresponding to the C-terminal α-helix was previously shown to be important for the protein insertase function of YidC and is conserved in other YidC-like proteins. The structure reveals that a region of YidC that was previously shown to be involved in binding to SecF maps to one edge of the β-sandwich. Electrostatic analysis of the molecular surface for this region of YidC reveals a predominantly charged surface and suggests that the SecF-YidC interaction may be electrostatic in nature. Interestingly, $\mathrm{YidC_{EC}P1}$ has significant structural similarity to galactose mutarotase from *Lactoca*[illegible]

---

*

tase activity, amino acid sequence alignments reveal that the C-terminal ∼

yethylene glycol 3350, and 20% glycerol. Crystals were incubated in cryo-solution for ~5 min before being flash-cooled in liquid nitrogen.

Data Collection—Diffraction data were collected on selenomethionine-incorporated crystals at beamline 8.2.2 of the Advance Light Source, Lawrence Berkeley Laboratory, University of California at Berkeley using a Quantum 315 ADSC area detector. The crystal-to-detector distance was 320 mm. Data were collected with 1° oscillations, and each image was exposed for 3 s. The diffraction data were processed with the program HKL2000 (24). See Table 1 for data collection statistics.

Structure Determination and Refinement—The YidC$_{EC}$P1 structure was solved by single wavelength anomalous dispersion using a data set collected at the peak wavelength (0.9794 Å), the program SHELX (25) within ccp4i (26), and Autosol within PHENIX version 1.3 (27). SHELXC found eight of the possible ten selenium sites. The program Autobuild within PHENIX version 1.3 (27) automatically constructed ~90% of the polypeptide chain and performed density modification. The rest of the model was built using the program Coot (28). The structure was refined using the program Refmac5 (29) and the program CNS (30). The final models were obtained by restrained refinement in Refmac5 with Translation Liberation Screw Rotation (TLS) restraints obtained from the TLS motion determination server (31). The data collection, phasing, and refinement statistics are summarized in Table 1.

Structural Analysis—Secondary structural analysis was performed with the programs DSSP (32), HERA (33), and Promotif (34). The programs SUPERIMPOSE (35) and SUPERPOSE (36) were used to overlap coordinates for structural comparison. The program CONTACT within the program suite CCP4 (26) was used to measure the hydrogen bonding and van der Waals contacts. The program CASTp (37) was used to analyze the molecular surface and search for potential substrate binding sites. The program SURFACE RACER 1.2 (38) was used to measure the solvent-accessible surface of the protein and individual atoms within the protein. A probe radius of 1.4 Å was used in the calculations. The Protein-Protein Interaction Server (39, 40) was used to analyze the interactions between the molecules in the asymmetric unit. The stereochemistry of the structure was analyzed with the program PROCHECK (41). The DALI server was used to find proteins with similar protein folds (42).

Figure Preparation—Figures were prepared using PyMOL (43). The alignment figure was prepared using the programs ClustalW (44) and ESPript (45).

RESULTS

Structure of YidC$_{EC}$P1—We have produced a C-terminal His$_6$-tagged soluble construct of the major periplasmic domain of *E. coli* YidC (YidC$_{EC}$P1) that spans residues Asp^{26}-Leu^{340} (Fig. 1*A*). High resolution size-exclusion chromatography analysis and multiangle light scattering analysis of the YidC$_{EC}$P1 reveal that the protein is very soluble in the absence of detergents and is monomeric in nature (data not shown). YidC$_{EC}$P1 was crystallized, and the structure was solved by single wavelength anomalous diffraction and refined to 2.5 Å resolution. There are two molecules in the asymmetric unit, and the refined structure includes residues 57–340. In addition, there is electron density for 3 residues at the C terminus that corresponds to the affinity tag used to purify the protein. There is no visible electron density observed for a presumably mobile loop that spans residues 207–216. Additionally, no electron density is observed for the N-terminal residues 26–56. To facilitate crystallization and improve the diffraction quality of the crystals the following mutations were introduced into YidC$_{EC}$P1: E228A, K229A, E231A, K232A, and K234A. These lysine and glutamate residues were targeted for mutation to alanine in an attempt to reduce the degree of conformational entropy associated with longer side chains that may impede crystallization (46, 47). The mutant YidC$_{EC}$P1 procould accommodate interactions with the acyl chains of the membrane lipids but could possibly interact with the lipid head groups. This is consistent with the solubility of YidC$_{EC}$P1 in the absence of detergents.

Protein-Protein Interactions Observed in the YidC$_{EC}$P1 Crystals—Superposition of the two molecules in the asymmetric unit shows that the only significant structural difference between molecule A and molecule B is a shift in the orientation of the C-terminal α-helix 3 (Fig. 2). The difference in the orientation of helix 3 is likely due to crystal-packing interactions. As mentioned earlier, to facilitate crystallization, five mutations were introduced into a region of YidC$_{EC}$P1 to replace a cluster of lysine and glutamate residues. The structure shows that these residues are located on or near β-strand 12. Molecule A makes residues(molecu04(orien-0s)-219(are)1-390(the)1-395(intr1-39s)]TJeol

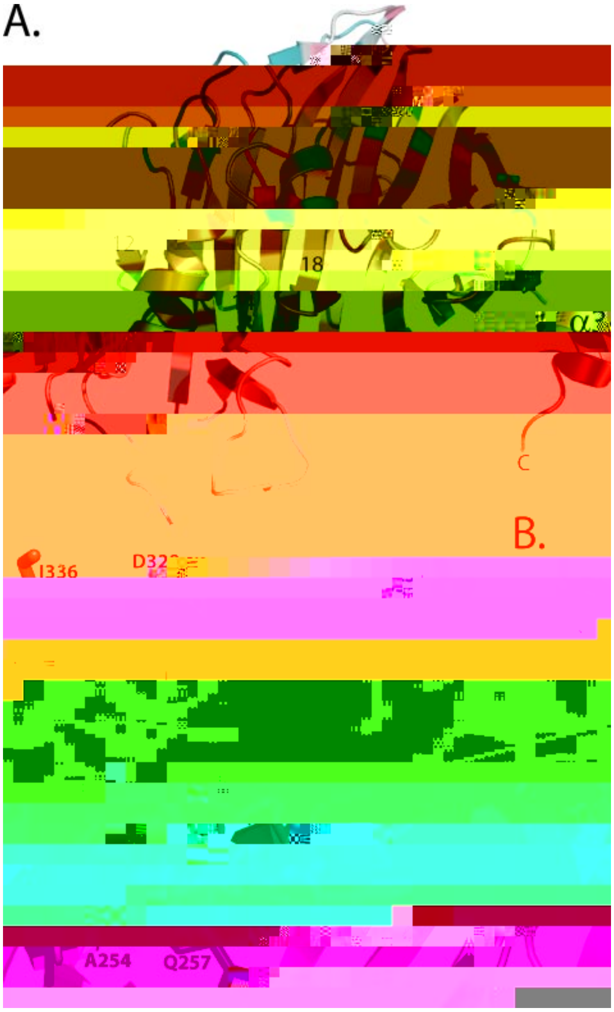

FIGURE 4. **Conserved amino acids within the large periplasmic domain of YidC in Gram-negative bacteria mapped onto the structure of YidC$_{EC}$P1.** *A*, [illegible] *purple*, [illegible] *blue*. [illegible] α-helix 3 [illegible] β-sheet 2, [illegible] α-helix 3 [illegible] *B*, [illegible] α-helix 3. [illegible] *sticks*, [illegible]

ecule A. This type of interaction is not observed in molecule B, giving a possible explanation for the differences seen in the orientation for the C-terminal α-helix 3.

Conserved Regions of YidC$_{EC}$P1—Amino acid sequence alignment of eight YidC variants from various Gram-negative bacterial species reveals a number of conserved residues located throughout Yid$_{EC}$P1 (Fig. 3). Most notably, the region at the extreme C terminus of the construct corresponding to α-helix 3 is well conserved. PFAM (48) analysis reveals that the residues 61–350 of YidC define a conserved domain PFAM-B_1222 that is remarkably consistent with the region of YidC$_{EC}$P1 observed in the electron density (residues 57–340). Sixty-one YidC variants were extracted from the domain PFAM-B_1222, aligned using ClustalW (44), and analyzed using the program CONSURF (49) that maps conserved residues onto a three-dimensional structure. As shown in Fig. 4, a significant number of conserved residues map to α-helix 3 and to β-strands 11, 12, 14, 15, 18 that cluster on the face of β-sheet 2 and pack against α-helix 3 (Fig. 4*A*). Closer inspection reveals that many of the

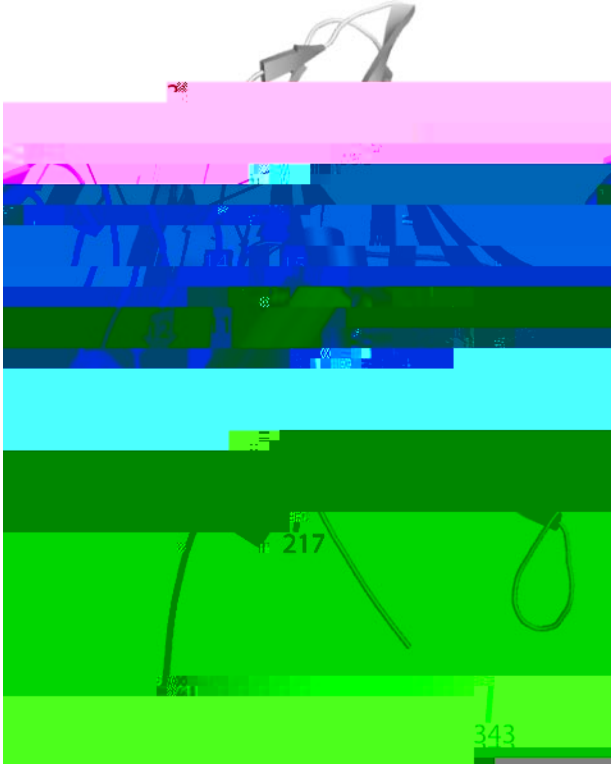

FIGURE 5. **The functional regions of YidC$_{EC}$P1.** [illegible] *blue region* (215 [illegible] 265) [illegible] SecF [illegible] SecDFYajC [illegible] (18). [illegible] *red region* (323 [illegible] 346) [illegible] (18, 21).

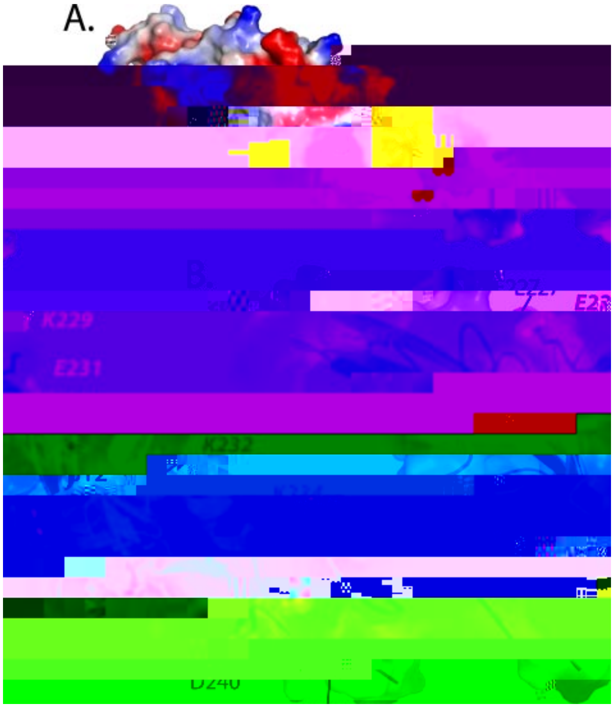

FIGURE 6. **Charged surface patches on the edge of the YidC$_{EC}$P1 β-sandwich correspond to a region that prior experiments suggest interacts with SecF (18, 21).** *A*, [illegible] YidC$_{EC}$P1. *B*, [illegible] *A* [illegible] YidC$_{EC}$P1 [illegible] β-[illegible] *sticks* [illegible] *red*. [illegible] β-[illegible] 12.

conserved residues are involved in interactions between α-helix 3 and β-sheet 2 (Fig. 4*B*), suggesting that the interaction may be biologically significant.

Mapping Functional Regions—Previous studies of YidC have mapped two functional regions to the YidC periplasmic domain. First, deletion analysis of YidC has revealed that residues 323–346 of the first periplasmic domain are essential for cell viability and insertase activity (18, 21). This region corresponds to the conserved α-helix 3 at the C terminus of $YidC_{EC}P1$ (Fig. 5

activity (18), correspond to α-helix 3. Remarkably, Jiang *et al.* (21) have performed alanine scanning mutagenesis experi-

. , . **45,** 1 01 1 0 3
1 . , ., , . ., , ., , ., , . .
(00) . , . , . .