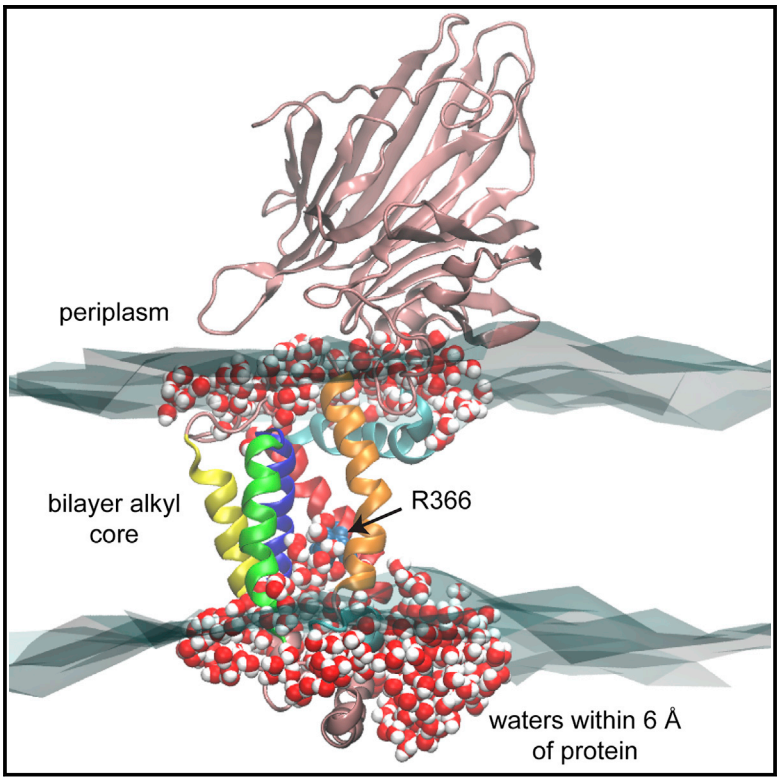
Article

Structure

YidC Insertase of *Escherichia coli*: Water Accessibility and Membrane Shaping

Graphical Abstract



Authors

Yuanyuan Chen, Sara Capponi, Lu Zhu, Patrick Gellenbeck, J. Alfredo Freites, Stephen H. White, Ross E. Dalbey

Correspondence

stephen.white@uci.edu (S.H.W.), dalbey@chemistry.ohio-state.edu (R.E.D.)

In Brief

Chen et al. have measured the *in vivo* water accessibility of the *E. coli* YidC insertase and carried out MD simulations of YidC in POPE:POPG bilayers. Water accessibilities are similar for the two approaches. The simulations reveal important structural features that stabilize YidC in membranes.

Highlights

- *In vivo* water accessibility of *E. coli* YidC was determined using NEM labeling
- Accessibility of YidC determined from all-atom MD simulations agrees with NEM results
- Compared with crystal structure, YidC structure in POPE:POPG is much more compact
- YidC thins the bilayer locally due to interface aromatic residues and salt bridges

Chen et al., 2017, Structure *25*, 1403–1414
September 5, 2017
http://dx.doi.org/10.1016/j.str.2017.07.008

CellPress

Structure
Article

YidC Insertase of *Escherichia coli*: Water Accessibility and Membrane Shaping

Yuanyuan Chen,[1,4] Sara Capponi,[2,4,5] Lu Zhu,[1,6] Patrick Gellenbeck,[1] J. Alfredo Freites,[3] Stephen H. White,[2,7,*] and Ross E. Dalbey[1,*]
[1]Department of Chemistry and Biochemistry, The Ohio State University, Columbus, OH 43210, USA
[2]Department of Physiology and Biophysics, University of California, Irvine, CA 92697, USA
[3]Department of Chemistry, University of California, Irvine, CA 92697, USA
[4]These authors contributed equally
[5]Present address: Cardiovascular Research Institute, Department of Pharmaceutical Chemistry, University of California, San Francisco, CA 94158, USA
[6]Present address: Department of Molecular Biology and Genetics and Weill Institute for Cell and Molecular Biology, Cornell University, Ithaca, NY 14853, USA
[7]Lead Contact
*Correspondence: stephen.white@uci.edu (S.H.W.), dalbey@chemistry.ohio-state.edu (R.E.D.)
http://dx.doi.org/10.1016/j.str.2017.07.008

SUMMARY

The YidC/Oxa1/Alb3 family of membrane proteins function to insert proteins into membranes in bacteria, mitochondria, and chloroplasts. Recent X-ray structures of YidC from *Bacillus halodurans* and *Escherichia coli* revealed a hydrophilic groove that is accessible from the lipid bilayer and the cytoplasm. Here, we explore the water accessibility within the conserved core region of the *E. coli* YidC using *in vivo* cysteine alkylation scanning and molecular dynamics (MD) simulations of YidC in POPE/POPG membranes. As expected from the structure, YidC possesses an aqueous membrane cavity localized to the membrane inner leaflet. Both the scanning data and the MD simulations show that the lipid-exposed transmembrane helices 3, 4, and 5 are short, leading to membrane thinning around YidC. Close examination of the MD data reveals previously unrecognized structural features that are likely important for protein stability and function.

INTRODUCTION

In bacteria, the YidC insertase functions in membrane protein insertion and assembly (Dalbey et al., 2014). YidC can promote insertion of membrane proteins independently (Chen et al., 2002; Samuelson et al., 2000, 2001; Serek et al., 2004; van Bloois et al., 2004; van der Laan et al., 2004; Yi et al., 2004) or cooperatively with the SecYEG complex (Celebi et al., 2006, 2007; Yi et al., 2003; Zhu et al., 2012). The single-span pf3 coat protein carrying negative charges in the N-terminal tail is the prototypical substrate for Sec-independent insertion (Serek et al., 2004). Structures of YidC were recently determined by X-ray crystallography at high-resolution for *Bacillus halodurans* (Kumazaki et al., 2014a) and *Escherichia coli* (Kumazaki et al., 2014b). The crystallographic structure of a YidC-like protein (DUF106) from *Methanocaldococcus jannaschii* demonstrated that YidC-like proteins are present in all branches of life (Borowska et al., 2015).

These YidC structures along with extensive *in vivo* water accessibility studies of the YidC homolog SpoIIIJ in *Bacillus subtilis* (Shimokawa-Chiba et al., 2015) showed that the conserved, membrane-embedded region of YidC contains a U-shaped hydrophilic cavity dominated by a strictly conserved arginine. The cavity is exposed both to the lipid bilayer and cytoplasm, but is inaccessible from the extracytoplasmic side of the membrane. Although this conserved arginine is essential for both bacterial growth and membrane protein insertion in *B. subtilis*, it is not essential in *E. coli* (Chen et al., 2014). Cys crosslinking studies show that the substrate-binding site of the *E. coli* YidC includes transmembrane helix 3 (TM3) (Klenner and Kuhn, 2012; Klenner et al., 2008; Yu et al., 2008), as well as TM1, TM4, and TM5 (Klenner and Kuhn, 2012). Major contacts between these TM segments of YidC and the hydrophobic segment of a single-span membrane protein substrate occur in the respective YidC/substrate hydrophobic region(s) spanning the membrane (Klenner and Kuhn, 2012). Cryoelectron microscopy studies of ribosome-YidC-nascent chain complexes (Kedrov et al., 2016; Wickles et al., 2014) show model substrates bound parallel to TM3 and TM5 of the hydrophilic cavity, consistent with the crosslinking studies.

Here we examine the *in vivo* water accessibility of *E. coli* YidC using Cys-scanning alkylation (Guan and Kaback, 2007; Kimura-Someya et al., 2000; Tamura et al., 2001). In parallel, we carried out a comprehensive 185 ns molecular dynamics (MD) simulation of *E. coli* YidC embedded in a palmitoyloleoylphosphatidylethanolamine:palmitoyloleoylphosphatidylglycerol (POPE:POPG) (75:25) lipid bilayer with the goal of determining water accessibilities for comparison with the *in vivo* measurements. The *in vivo* and simulation results both confirm that the *E. coli* YidC possesses a solvent-exposed region that is localized within the inner leaflet of the membrane. Our data reveal significant thinning of the bilayer caused by the short lengths of TM3, TM4, and TM5. The hydrophilic cavity of YidC and the membrane thinning around it likely reduce the energy barrier of TM helix insertion by decreasing the length of the hydrophobic core of the membrane. The protein appears to be stabilized and tightly linked to

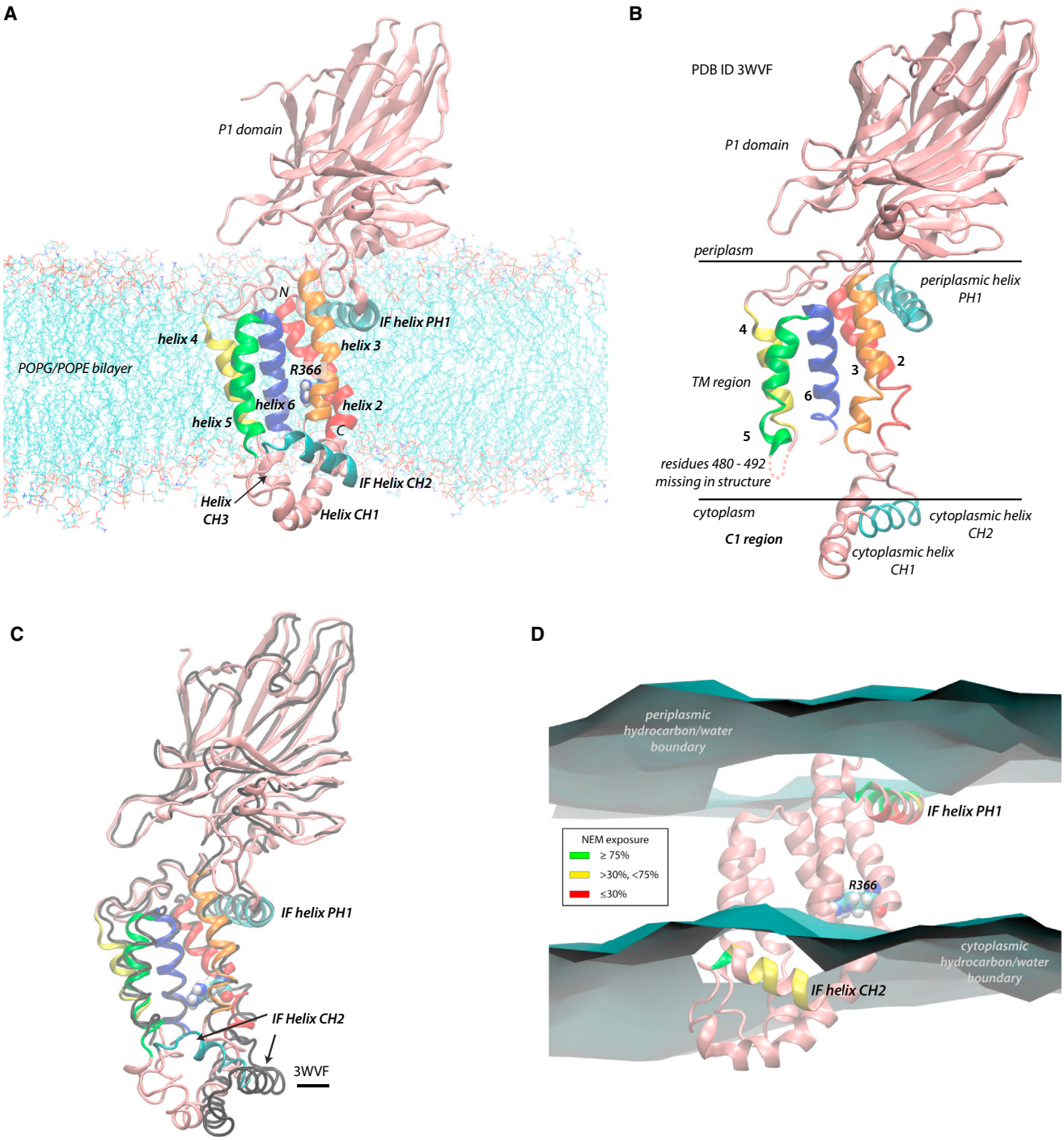


Figure 1. The Transmembrane Region of the YidC Insertase Becomes More Compact and Ordered in POPE/POPG Bilayers

E. coli YidC has six TM helices, but TM1 near the N terminus, which is not required for normal functioning, was disordered and not observed in the structure. Consequently, the observable TM helices in the structure are numbered 2 through 6. The colors of the helices are preserved for all figures in the paper (TM2, red; TM3, orange; TM4, yellow; TM5, green; TM6, blue).

(A) A snapshot of the equilibrated structure of *E. coli* YidC in a POPE:POPG (75:25) bilayer. The front half of the bilayer facing the viewer has been removed to show more clearly the arrangement of the YidC helices.

(B) The crystal structure of *E. coli* YidC determined by Kumazaki et al. (2014b); structural regions designated by Kumazaki et al. are shown. Because of crystal disorder, residues 480–492 in the TM4–TM5 loop were missing in the structure (dashed pink line). For the MD simulations, the missing loop residues were modeled into the protein using Phyre2 (Kelley et al., 2015).

(legend continued on next page)

the lipid bilayer by interfacial aromatic rings, a cytoplasmic salt bridge cluster, and a periplasmic interfacial helix enriched in aromatic residues. A cluster of aromatic residues above R366 may provide a polarizable hydrophobic surface for guiding nascent chains across and into the lipid bilayer.

RESULTS

The Structure of YidC in a POPE:POPG Bilayer Becomes More Compact Compared with the Crystallographic Structure

Figure 1A shows a snapshot of fully equilibrated YidC in a 75:25 POPE:POPG lipid bilayer. Interestingly, the key R366 side chain is located very near the center of mass of the bundle of five helices (TM2–TM6). Compared with the X-ray structure (Figure 1B), the helix bundle has become more compact. The superposition of the crystal structure with the MD structure (Figure 1C) shows that the main differences are a modest reorientation of helix CH1 and a very large reorientation of the cytoplasmic helix CH2 (re-labeled here as IF helix CH2) toward the membrane interface, a likely consequence of the interaction of the C1 region (Figure 1B) with phosphatidylethanolamine (PE) and phosphatidylglycerol (PG) head groups. Residues 480–492 (not seen in the crystal structure) have formed a short third cytoplasmic helix CH3. Together, CH1, CH2, and CH3 form what we refer to as the cytoplasmic cap. Importantly, IF helix PH1 is buried in the bilayer hydrocarbon core just beneath the lipid head groups, while IF helix CH2 is located at the interface outside of the hydrophobic core. Significantly, YidC-bilayer interactions have caused the bilayer to thin in the neighborhood of the protein as though riveting the two monolayers together.

Although the crystal structure was determined by the lipidic cubic phase method to produce type I (lamellar) crystals, the only lipid present was monoölein. We suggest that the absence of PE and PG head groups in the crystallization is responsible for the differences between the crystallographic and simulation structures.

Water Accessibility of YidC as Probed by Site-Directed Alkylation

The structures of YidC from *B. halodurans* and *E. coli* suggested that YidC does not contain a membrane-spanning channel, but rather a U-shaped hydrophilic groove closed on the periplasmic surface but exposed to the cytoplasm and the bilayer (Kumazaki et al., 2014a, 2014b). To examine the water exposure of the groove within TM2–TM6 of the *E. coli* YidC in intact bacteria, we used the Cys alkylation approach of Shimokawa-Chiba et al. (2015), described in detail in STAR Methods. In brief, single Cys residues were introduced one at a time within the TM segments or in the flanking polar regions in a Cys-less YidC, and their accessibility to membrane-permeable N-ethylmaleimide (NEM) determined. Maleimides react with thiolate anions, which require water exposure. They will not react with Cys residues in the low-dielectric (i.e., lipid exposed) membrane environment, because the thiol group is protonated and unreactive. Consequently, NEM reacts with Cys residues within the membrane region only if it has water access. Thus, no change in the position of YidC on a gel when NEM is added prior to the shifting reagent methoxypolyethylene glycol maleimide (Mal-PEG) (molecular weight ~5,000 Da), which labels unreacted Cys residues, indicating a solvent-exposed Cys residue (see YidC E24C, Figure 2A, right panel). In contrast, YidC with a TM Cys residue exposed to lipid (see YidC A11C, Figure 2A, middle panel) shows a shift with Mal-PEG even when NEM is added. As a control, we determined that treatment of YidC C24 with Mal-PEG results in a gel shift when NEM is not added beforehand. In addition, we verified that a Cys-less YidC does not show a shift with Mal-PEG (Figure 2A, left panel).

To confirm that the YidC Cys mutants were functional, we assayed the mutants for activity using a YidC complementation assay involving the YidC depletion strain JS7131 (Jiang et al., 2003). Figure 2B shows that YidC Cys-less, YidC 24C, and YidC 11C all fully complement the YidC depletion strain at all the tested dilutions. All the Cys mutants we studied were assayed by this complementation method; except for a few mutants (T362C, M380C, Y412C, N418C, and G512C) all were fully active (see Table S1). The experimental data for 103 Cys mutants, covering the five conserved TM helices, are presented in summary form in Figure 2C; complete datasets are shown in Figures S1 and S2. Water exposures were determined from the optical densities of the Mal-PEG-shifted and the non-shifted YidC bands, as described in STAR Methods. All experiments were carried out twice.

The mean NEM accessibilities for the Cys mutants are summarized in Figures 3A and 3B. The light-blue bars indicate the NEM accessibility of each mutant. The colored horizontal bars show the locations of the TM helices color coded as in Figure 1. Broadly, non-helix residues are highly exposed, whereas the exposure of helical residues varies from very high to very low. There are only a few residues that have less than 5% exposure (e.g., P431, L436, T503, and L515). Some exposure is expected for all residues, because the *in vivo* experiments probe YidC in a working state that undoubtedly involves significant functional and thermal fluctuations. Cognizant of this fact, we divided exposures into three groups: 30% or less, greater than 30% but less than 75%, and 75% or greater. These results are similar to those obtained by Shimokawa-Chiba et al. (2015) in their NEM accessibility study of the SpoIIIJ YidC homolog from *B. subtilis*.

(C) The crystal structure in (B), shown in this panel in black, has been overlaid on the simulation structure shown in (A) to reveal that the most significant differences between the two structures occur in the C1 region (B). We attribute the compaction of the C1 region to interactions between POPE and POPG that are missing in the crystals, which contain only glycerol monoölein.

(D) Summary of NEM exposure of residues in helices PH1 and CH2. We arbitrarily define inaccessibility as 30% or less modified by NEM, partially accessible by 30%–75% labeling, and fully accessible by greater than 75% labeling. Notice that the amphipathic character of PH1 buried in the membrane interface is clearly revealed by the accessibilities (unexposed on the surface facing the bilayer hydrocarbon core; highly exposed on the surface facing the periplasm). The labeled residues on the CH2 helix reveal only partial exposure for all labeled residues. High exposures would be expected based on the crystal structure alone, because CH2 is not able to engage a membrane interface.

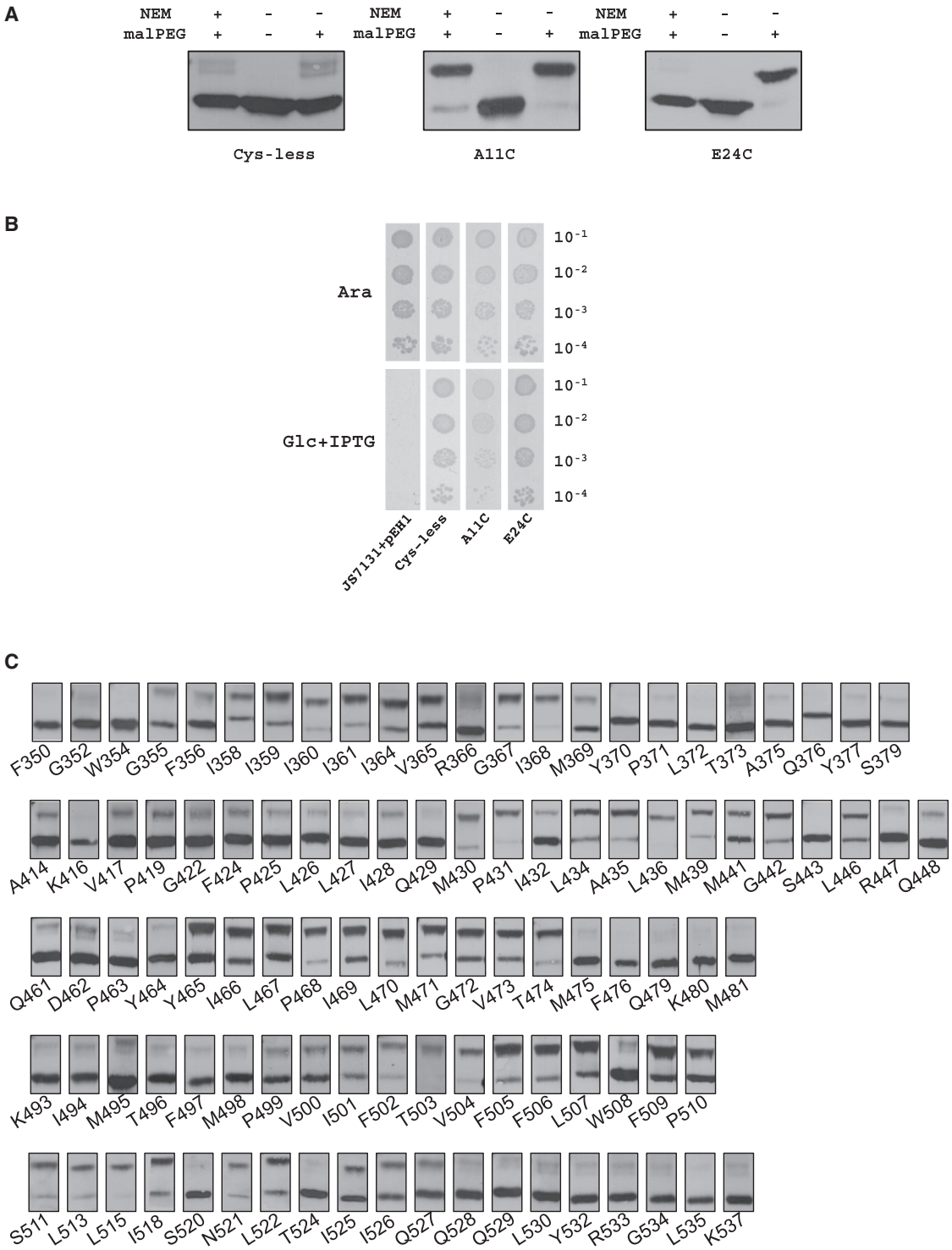


Figure 2. Cys-Based Alkylation Method to Map the Lipid- and Solvent-Exposed Membrane Regions of YidC

(A) The gel-shift assay employed to examine whether a Cys residue is in a solvent- or lipid-exposed environment. The solvent-exposed YidC 24C and lipid-exposed 11C, as well as Cys-less control, were analyzed in this example. BL21 cells expressing the indicated YidC mutants were treated with or without NEM, washed, and analyzed as described in the STAR Methods. Where indicated, the samples were treated with Mal-PEG, which modifies unreacted Cys residues, and then analyzed by SDS-PAGE and western blot. Solvent exposures were quantitated using optical scans of the blots (STAR Methods).

(legend continued on next page)

Water Accessibility of YidC as Probed by Molecular Dynamics

Water accessibilities of the amino acid residues probed using NEM exposures were also determined by counting the average number of waters within 6 Å of each residue during the last 34 ns of the 185-ns equilibrium simulation. Recognizing that large residues inherently have greater exposure than small ones, we normalized the data according to the standard-state accessible surface areas A^0 of fully exposed residues of soluble proteins determined by Rose et al. (1985). A^0 varies between 88 $Å^2$ for Gly and 266 $Å^2$ for Trp. The normalization allowed accessibility to be expressed as a fraction ranging from 0.0 to 0.85. The fractional accessibilities are shown in Figures 3A and 3B as vertical red bars.

The agreement between the simulation exposures and NEM exposures is quite good, which generally validates the MD simulations. Residues that have low or no exposure according to the simulation correspond to residues with less than 30% NEM exposure. But there are a few significant discrepancies (vertical arrows in A and B of Figure 3), specifically L427, I432, L434, and V500. Interestingly, all of these residues are on either helix 3 (L427, I432, L434) or helix 5 (V500) that form the entrance to the hydrophilic cavity from the bilayer interior (Figure 3C).

Residues with Low Water Accessibility Form a Narrow Band in the Center of YidC

Figure 4A shows a snapshot of the waters within 6 Å of the protein near the midpoint of the equilibrated simulation. Broadly, waters can penetrate to the level of R366, having a funnel-shaped distribution, the narrow end of which is located at R366, and are highly dynamic (Movie S1). The broad end of the funnel is located roughly at the cytoplasmic lipid head group/hydrocarbon core boundary defined by the glycerol C2 carbon surface. There is little water penetration into the protein at the periplasmic surface.

In Figure 4B and Movie S2, we show the results of the NEM water accessibility of the amino acids in TM segments combined with a snapshot of the residues having waters within 6 Å, as determined in the simulation. The residues are shown in wire-frame surface representations and color coded according to NEM accessibility. The residues with low accessibility (30% or less) occupy roughly the central one-third of the helical bundle.

YidC Thins the Bilayer Uniformly but Asymmetrically in Its Vicinity

Asymmetric thinning of the POPE:POPG bilayer in the vicinity of YidC is strikingly apparent in Figure 1A. The periplasmic bilayer surface is relatively flat compared with the indented cytoplasmic surface. The general shape of the bilayer within 20 Å of the protein's center of mass at R366 is presented more clearly in Figure 5A. A periplasmic view of the protein and surrounding lipids is presented in Figure 5B. The carbonyl groups of the phospholipids are represented as colored spheres: lipids within 20 Å are colored blue, within 40 Å purple, and greater than 40 Å pink. Phosphates on the opposite (cytoplasmic) monolayer are represented by the same colors, but faded.

We examined the transbilayer distributions of the lipid carbonyl groups along the bilayer normal relative to the protein center of mass, within each of the three zones illustrated in Figures 5B and S6. The carbonyl groups were chosen because they are excellent markers of the head group/hydrocarbon boundary of the bilayer (Wiener and White, 1992). The statistical distributions of the carbonyls are shown in Figure 5C. The mean hydrocarbon thickness in the 20-Å shell is about 28 Å, whereas in the outermost shell it increases to about 33 Å (Table S3). A striking and significant feature of carbonyl distributions within the 20-Å shell is their asymmetries (arrows), which mirror the flatness of the bilayer in the vicinity of YidC. In addition, the periplasmic peaks are all coincident, whereas the cytoplasmic peaks move progressively outward, which also mirrors the flatness of the bilayer in YidC's vicinity. Within the 20-Å zone, the periplasmic carbonyl peak is located 16 Å above the protein center of mass, whereas the cytoplasmic peak is located 12 Å below the center of mass, consistent with the dimpling apparent in Figure 1A. These data suggest that YidC and the bilayer in its immediate vicinity are locked together structurally, causing the periplasmic and cytoplasmic surfaces to be remarkably flat.

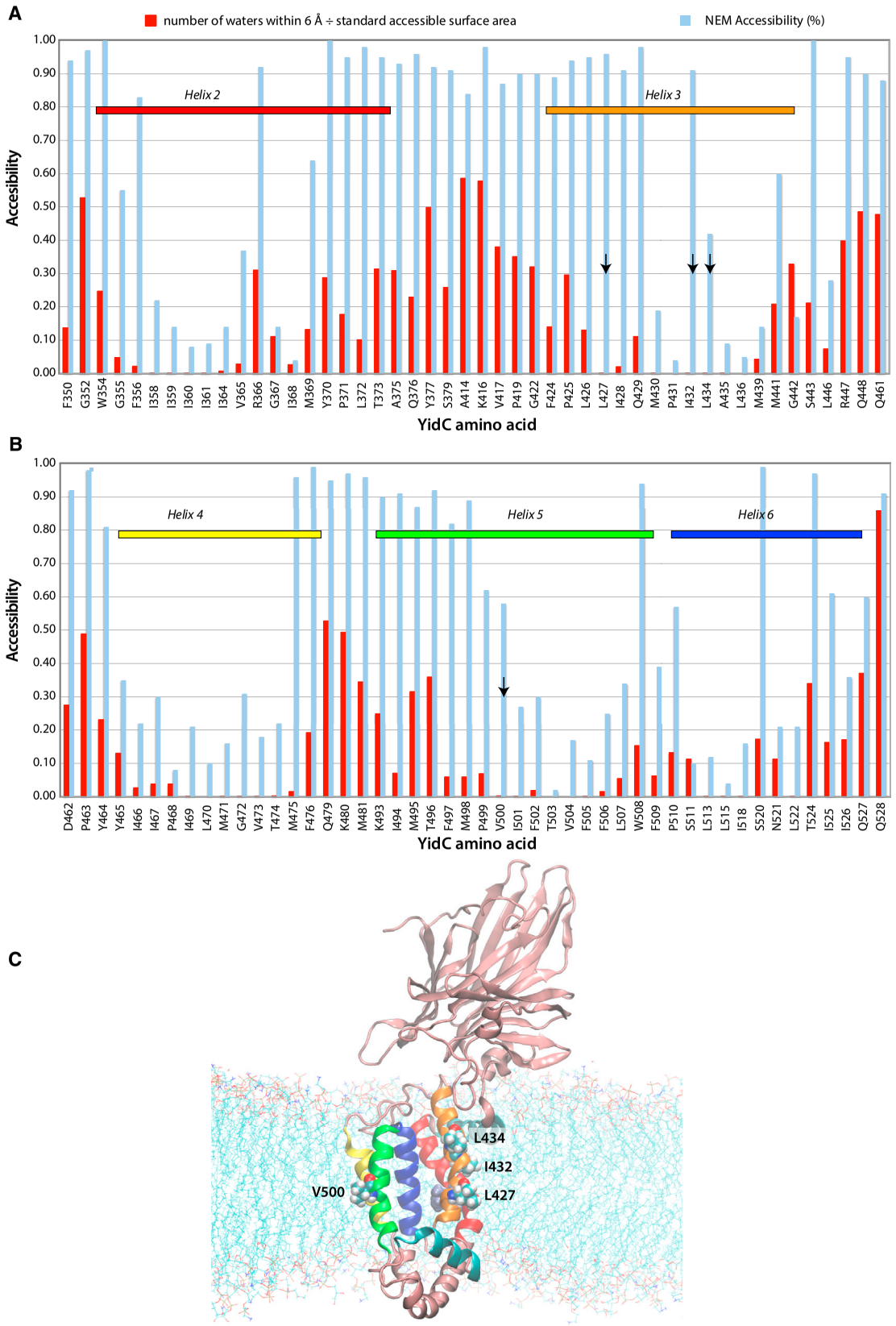
Rings of Aromatic Residues and a Salt Bridge Cluster Stabilize YidC in the Membrane

Aromatic residues, particularly Tyr and Trp, are found at the membrane interface of virtually all membrane proteins (Killian and von Heijne, 2000; Ulmschneider et al., 2005; Wallin et al., 1997), as first reported for the photosynthetic reaction center (Schiffer et al., 1992). The preference of Trp (and probably Tyr) for interfaces is most likely due to its flat rigid shape that limits access to the hydrocarbon core and its electronic structure and associated quadripolar moment, which stabilize the protein in the bilayer interface by complex electrostatic interactions (Yau et al., 1998). It is thus not surprising that YidC is highly enriched in aromatics at the membrane interfaces (Figure 6; Movie S3). There are 14 aromatics (including Phe) at the periplasmic surface. Of these, six located on IF helix PH1 (Figures 6B; Movie S3) likely provide strong stabilization. Two lysine residues (K342 and K345) that snorkel into the interface to interact with lipid phosphates may provide additional stability (Movie S3).

Fewer aromatics are found at the cytoplasmic interface, where only seven are found distributed at the cytoplasmic termini of the TM helices except for TM6 (Figure 6C). These seven aromatics and the cluster of salt bridges (the cytoplasmic cap) presumably provide additional stability to YidC (Figures 6A and 6E; Movie S3). The acidic and basic residues involved in salt bridging are heavily concentrated on IF helix CH2 (Figure 6D). As shown in Figure 6E and Movie S4, inter-residue salt bridges are

(B) Complementation assay for monitoring the activity of the YidC mutants. JS7131 cells harboring the indicated pEH1YidC Cys mutant were grown under YidC expression condition (top panel) and YidC depletion (bottom panel). For controls, JS7131 cells with vector encoding the Cys-less YidC (C423S) mutant or with the empty vector (pEH1) were analyzed.

(C) Summary of the accessibility of Cys residues introduced into helices TM2 through TM6 to map the solvent- and lipid-exposed residues following the procedures summarized in (A). In all cases, the samples were treated with NEM followed by treatment with Mal-PEG, as discussed for (A). The entire dataset is shown in Figures S1 and S2.



(legend on next page)

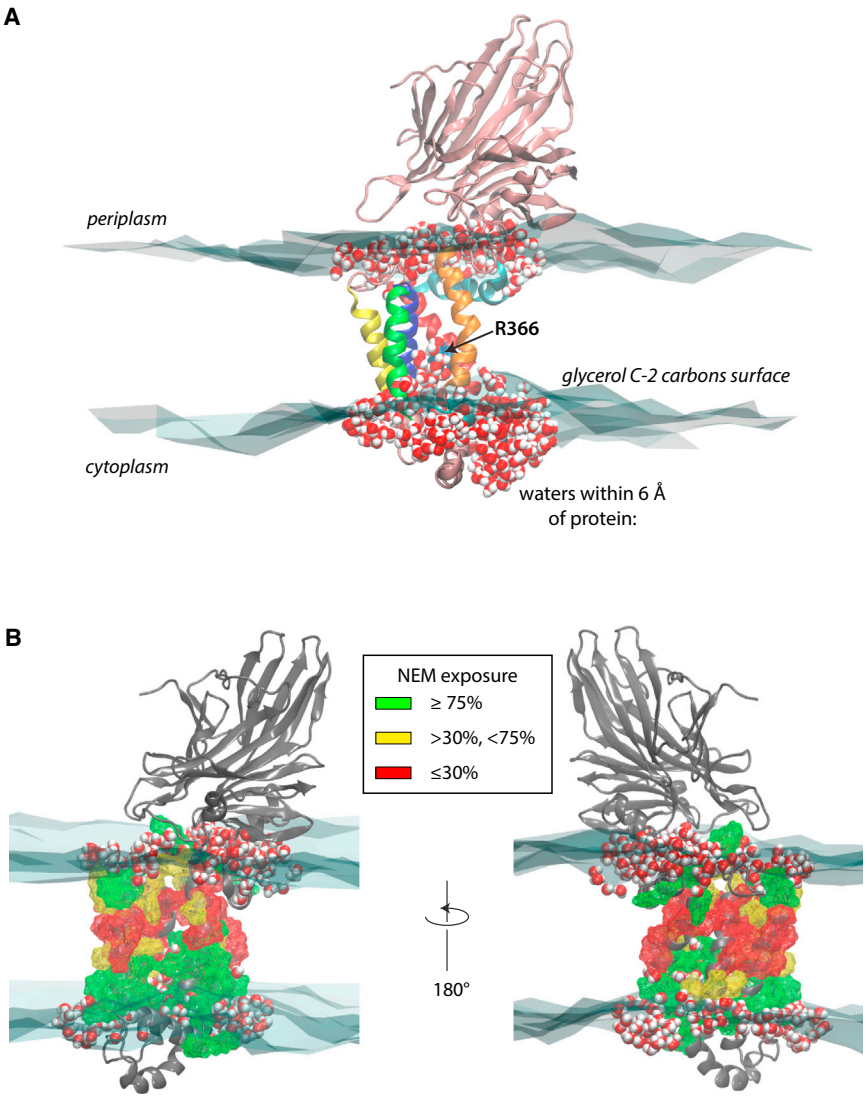


Figure 4. Water Accessibilities in the Context of the In-Bilayer YidC Structure Reveal a Narrow Band of Inaccessibility Near the Bilayer Midplane

(A) A snapshot from the MD simulation showing the waters within 6 Å of the protein. The head group/hydrocarbon boundary defined by the glycerol C2 carbons (Wiener and White, 1992) is shown as a transparent blue surface constructed using Delaunay triangulation. Waters can penetrate roughly to the middle of the protein due to hydration of the strictly conserved Arg366 residue.

(B) NEM water accessibility of TM-segment amino acids represented by wire-frame surfaces colored according to NEM exposure.

concentrated in the interfacial extension of TM2 (helix CH1, pink) and IF helix CH2 (cyan). Some of these also form salt bridges to the phospholipid head groups. Interestingly, there are four basic residues (K416, K382, R384, and K401) that are not hydrogen bonded to acidic residues or phosphates. The critical R366 sits above these residues toward the periplasm. We speculate that these residues might serve as guideposts for movement of the negatively charged N-terminal amino acids of substrates (Serek et al., 2004) into the YidC/bilayer interface for passage across the membrane.

A Cluster of Aromatic Residues Shields the Critical Arginine and May Provide Additional Stability

The aromatic cluster shown in Figure 6A stood out glaringly in our initial examination of membrane-embedded YidC. The cluster in other YidC species, however, is not as dramatic, based upon the sequence alignment of Kumazaki et al. (2014b). Such clusters are known to stabilize soluble proteins and likely play a similar role in stabilizing *E. coli* YidC (Burley and Petsko, 1985; Kannan and Vishveshwara, 1999). The sequence alignment of Kumazaki et al. (2014b) reveals three positions that are invariably occupied by aromatics across species: Y437, Y516, and Y517. The aromatics at positions 516 and 517 are interesting, because they sit directly above R366 toward the periplasmic surface (Figures 7A and 7B), acting as a periplasmic lid over R366. The Tyr residues interact with R366 via H bonds mediated by water molecules. It is notable that aromatics can interact favorably with positive charges through cation-π interactions (Dougherty, 1996). R366 and its tyrosine lid do not appear to be directly exposed to the TM segments of nascent chains, but rather sit well back from the critical TM3 and TM5 helices, which form the so-called greasy slide (Dalbey and Kuhn, 2014) (Figures 7C–7E) that has been established as providing the principal interaction zones of incoming nascent chains (Klenner and Kuhn, 2012).

The high sequence conservation of R366, Y516, and Y517 led us to examine the effects of mutating these residues.

Figure 3. The Water Accessibilities of YidC Residues Determined by the NEM Assays Generally Agree with Water Accessibilities Determined from the MD Simulations of YidC in POPE/POPG Bilayers

Accessibilities determined by the NEM assays are shown as light-blue bars; accessibilities determined from the MD simulation are shown as red bars. The vertical arrows indicate four residues for which significant discrepancies exist between the NEM and MD data. We arbitrarily define inaccessibility as 30% or less modified by NEM, partially accessible by 30%–75% labeling, and fully accessible by greater than 75% labeling. Vertical arrows indicate residues with high NEM exposure but low MD-determined accessibility.

(A) Accessibility results for residues 350–461.

(B) Accessibility results for residues 462–528.

(C) Structure of YidC in the lipid bilayer showing the four residues (vertical arrows) in (A and B) for which significant discrepancies exist between the NEM and MD results. L427, I432, L434, and V500 have significant exposure (greater than 30%) according to NEM labeling but virtually no exposure according to the MD results.

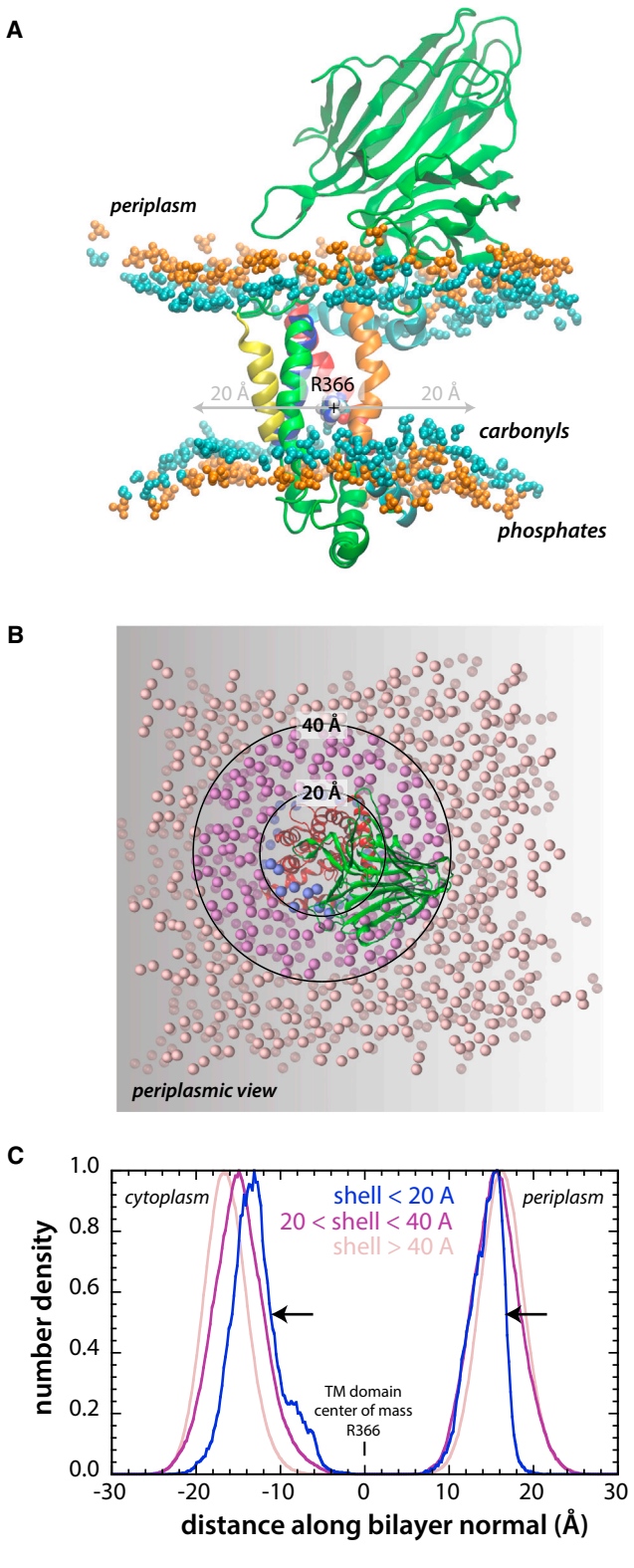


Figure 5. YidC Thins and Flattens the Membrane in Its Vicinity

(A) A snapshot of *E. coli* embedded in the POPE:POPG membrane. The lipid carboxyl (cyan) and phosphate (orange) groups are drawn as van der Waals spheres. Except for the interfacial and TM helices colored as in Figure 1, the protein is colored green. The image shows that the membrane is thinnest and quite flat within a radius of 20 Å of R366 due to interactions described in Figure 6. R366 is located at the center of mass of the TM helices and is about 11 Å above the cytoplasmic lipid carbonyl plane.

(B) View of the membrane-embedded protein from the periplasm along an axis normal to the membrane plane. The spheres represent the lipid carbonyl groups located within 20 Å (blue), between 20 and 40 Å (purple), and greater than 40 Å (pink) from R366 as projected onto the bilayer plane. The faded spheres are on the cytoplasmic surface.

(C) Distributions of lipid carbonyl positions projected onto the bilayer normal as a function of distance from the protein center of mass at R366. Within the 20-Å shell, the bilayer thickness is about 28 Å, whereas in the outer shell the average thickness is about 33 Å; the change is asymmetric with respect to R366. Of the approximate 5-Å decrease in thickness, about 3 Å can be attributed to cytoplasmic half and 2 Å to the periplasmic half. The distributions were determined over the last 80 ns of the simulation (Figures S5A and S5D; Table S3).

Surprisingly, Cys substitutions for each of these residues separately complemented the YidC depletion strain (Table S1), as did Ala and Ser substitutions (Figure S4), which is at odds with the finding that mutation of Y517, but not Y516, to alanine inactivated YidC (Wickles et al., 2014). Because waters are shared between R366, Y516, and Y517 (Figures 4A, 7A, and 7B), we determined the NEM accessibilities of the Cys-substituted residues in addition to MD-determined water accessibilities (Table S2). The data show that R366 is 90% exposed, whereas Y516 and Y517 are exposed only 8% and 34%, respectively. The accessibilities determined from the MD simulation are consistent with the NEM-determined accessibilities.

DISCUSSION

We have used site-directed cysteine alkylation (Figure 2) and MD simulations (Figures 3 and 4) to examine the water exposure of the hydrophilic groove within the membrane-embedded TM sequences of *E. coli* YidC (Kumazaki et al., 2014a, 2014b). The agreement between the MD data and the alkylation data is striking (Figures 3A and 3B). The only serious disagreement is the water accessibility of residues L427, I432, and L434 on TM3, and V500 on TM5 (Figure 3C). These are quite accessible according to the alkylation data, but inaccessible according to the MD data. We suggest that this reflects in part dynamic motions of TM3 and TM5 that occur *in vivo* in the functioning protein that are not possible in the simulation. A conformational change in the TM2, TM3, and PH1 region was observed in a cryo-electron microscopy study of a Foc substrate complexed with a YidC-nanodisc-ribosome complex (Kedrov et al., 2016). The fully inserted TM segment of Foc was found to be located at the interface of the TM3–TM5 greasy slide region of YidC.

We examined the water accessibility of the protein in the context of the residues on TM3 and TM5 that contact incoming nascent chains (Figure 7F). If we think of the nascent chain-contacting residues on TM5 (F502, F505, P468, and W508) and TM3 (F424, L427, and M430) as marking the insertion pathway, Figure 7F suggests a handoff of the nascent chain from TM3 to TM5, because substrate-interaction sites TM3 are near the cytoplasmic membrane surface, while interaction sites on TM5 are near the periplasmic surface. The aromatic cluster (Figure 6A) shared between TM3 and TM5 might provide an electrically polarizable hydrophobic pathway, allowing it to interact favorably with both non-polar and polar amino acids on the nascent chain.

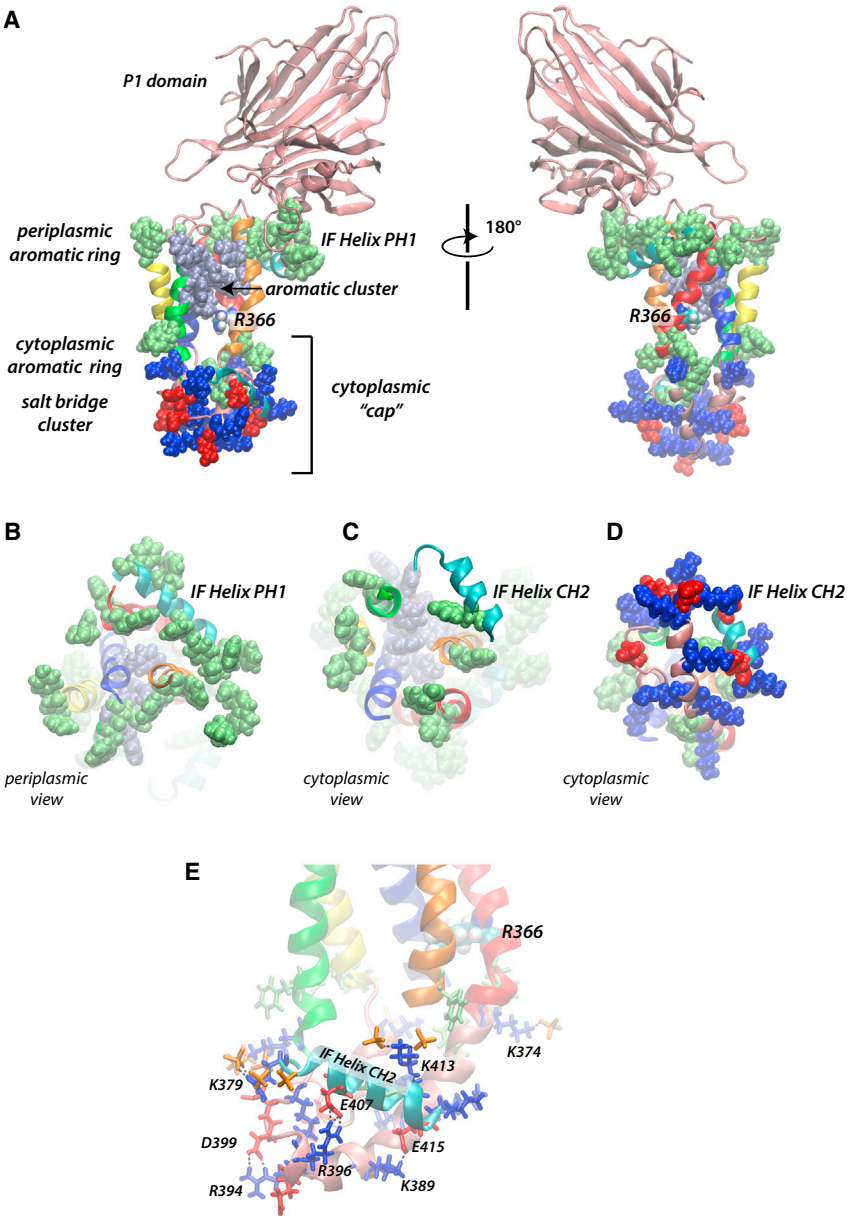


Figure 6. Several Significant Structural Features of YidC Stabilize It in the POPE:POPG Bilayer

(A) Rings of aromatic residues on both the periplasmic and cytoplasmic surfaces stabilize the protein in the bilayer interfaces. Beneath the cytoplasmic aromatic ring is a salt bridge cluster that, together with the aromatic ring, forms a tightly packed cytoplasmic cap. A striking feature is the nine-residue aromatic cluster (F433, Y437, F502, F505, F506, W508, F509, Y516, and Y517) located near the periplasmic surface. Notice that R366 is situated toward the cytoplasm directly beneath the aromatic cluster.

(B) The periplasmic aromatic ring, represented by van der Waals spheres. The amphipathic aromatic-rich interfacial helix PH1 on the periplasmic surface should provide strong anchoring in the periplasmic membrane surface (Schibli et al., 2002; Yau et al., 1998).

(C) The cytoplasmic aromatic ring is composed of seven aromatic residues.

(D) Cytoplasmic view of the protein that includes the salt bridge cluster sitting above the cytoplasmic aromatic ring. Basic residues are colored blue and acidic residues red.

(E) The cytoplasmic cap viewed parallel to the membrane, including H bonds (black dashes). Lipid phosphates (orange), basic residues (blue), and acidic residues (red) are represented in stick format. The salt bridges are shown more clearly in Movie S4. The significant salt bridge pairs are R394–D399, K389–E415, and R396–E407. The residues that H bond to lipid phosphates include K374, R403, K413, and K493.

Compared with the YidC X-ray structure (Figure 1B), the equilibrium structure of YidC in a POPE:POPG (75:25) lipid bilayer in excess water is more compact, involving principally reorientation of the cytoplasmic interfacial helix CH2 (Figure 1C). The missing residues in the crystal structure (480–492), which were modeled into the structure prior to the simulation, form a short helix (CH3) that is part of the cytoplasmic cap and salt bridge cluster (Figure 6A). Significantly, helix PH1 is buried in the bilayer beneath the lipid head groups, whereas helix CH2, the interfacial extension of TM2 (helix CH1), and helix CH3, are located outside the bilayer hydrocarbon core. These structural rearrangements apparently cause the uniform but asymmetric thinning of the bilayer of about 5 Å that produces the dimpling effect apparent in Figure 1A. The 28-Å thickness of the bilayer surrounding YidC is quite uniform within a 20-Å radius (Figures 5A and 5C). Wickles et al. (2014) reported membrane thinning of up to 8 Å around YidC in an MD simulation, but their simulation was of a clever YidC computational model rather than the native protein. Kedrov et al. (2016) noted a strong evolutionary coupling between PH1 and helix 2, suggesting a tight structural connection. This suggestion is supported by our study in that the geometry of PH1 and helix 2 is essentially identical in the crystal structure and the MD structure (Figure 1C).

We suggest that our MD structure is an accurate representation of the *in vivo* structure, because the packing and orientation of the TM helices and IF helix PH1 are essentially the same as in the crystal structure (Figure 1C). The reorientation of IF helix CH2 in the simulation makes perfect sense in the context of aromatic/membrane interface interactions and lipid/protein salt bridges that are expected to occur at the POPE/POPG interfaces. These interfaces are not available in the crystal, because the bilayer in the crystal lacks phosphate head groups.

As a further test of our MD structure, we carried out additional NEM accessibility studies of interface-buried IF helix PH1 and non-buried IF helix CH2 (Figure S3). The results are

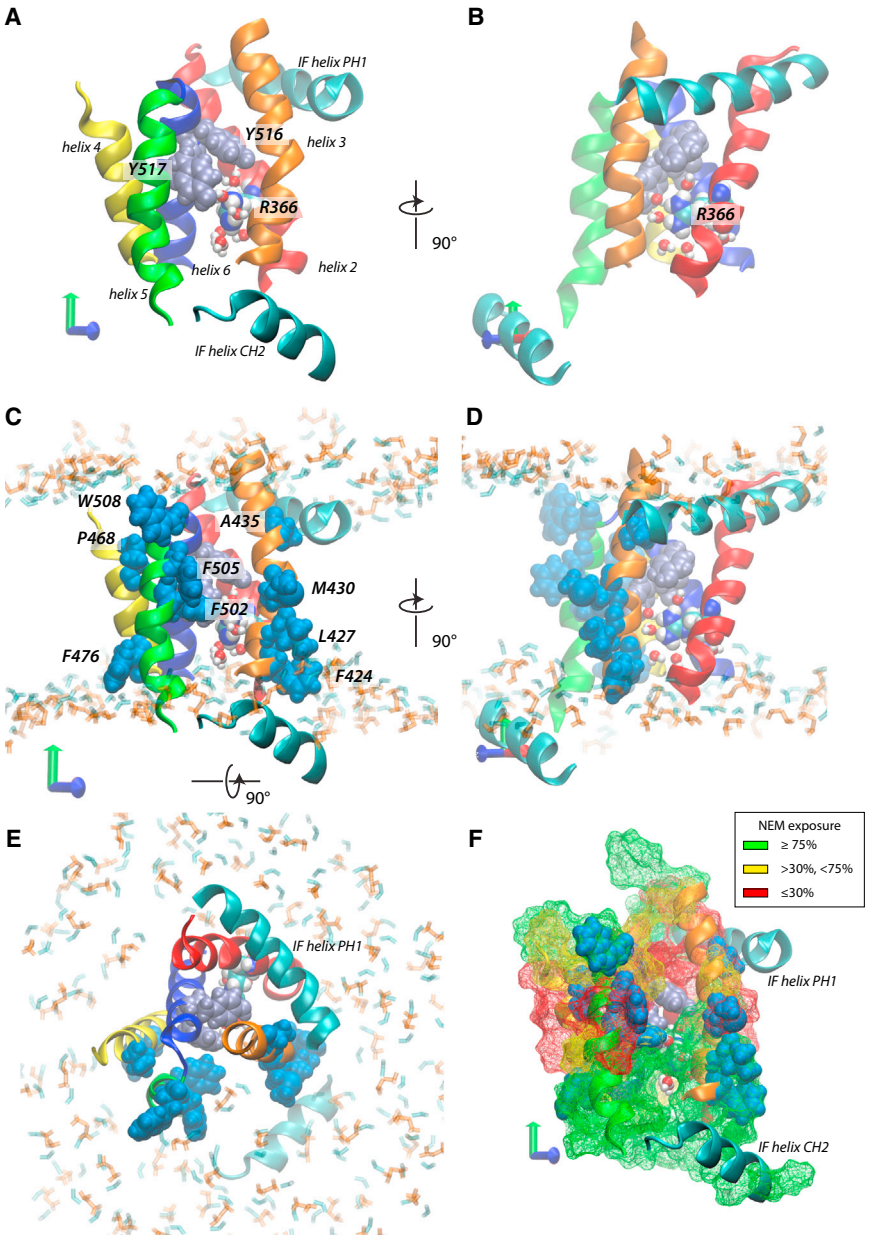


Figure 7. Two Highly Conserved Tyr Residues, Y516 and Y517, form a Hydrophobic Lid Facing toward the Periplasm above the Critical Arginine Residue (R366)

(A) The tyrosines are hydrogen bonded to R366 via intervening waters. We suggest that these tyrosines play two important roles: they stabilize the position of R366 and they provide a hydrophobic transition to the hydrophobic core of the protein, which in the case of YidC is comprised of the aromatic cluster (Figure 6A).

(B) The stabilizing tyrosine-arginine hydrogen bonding causes R366 on TM2 to be set well back into the protein toward TM5 and TM6.

(C and D) Residues, represented by blue van der Waals spheres, that have been shown to interact strongly with the single-span capsid protein pf3 during YidC-mediated insertion (Klenner and Kuhn, 2012) have been added to the structures of (A and B). Transparent stick models of the carbonyl (cyan) and phosphate (orange) groups define the bilayer surfaces. The structure is consistent with the idea that TM3 and TM5 act as a "greasy slide" for nascent single-span TM helices (Dalbey and Kuhn, 2014).

(E) This periplasmic view shows that all of the strongly acting residues are located on the membrane-facing surfaces of TM3 and TM5, well away from R366.

(F) The water accessibility (Figure 4) is shown superimposed on the structure of (C). Waters apparently can penetrate from the cytoplasmic side into the protein only up to F505 and M430.

summarized in Figure 1D. As expected for an amphipathic helix at a membrane interface, the residues of PH1 buried in the bilayer hydrocarbon have little water exposure (red highlighting), whereas residues facing away from the bilayer hydrocarbon have high exposure (green highlighting). Helix CH2, on the other hand, is not buried in the membrane, but rather is located at the interface where one expects significant water exposure, which we observed. Interestingly, however, the exposure is rated as intermediate (yellow highlighting). This likely reflects the complex environment of CH2 (Figure 6E).

The C1 region of YidC in the crystal structures had rather high crystallographic B factors (Kumazaki et al., 2014b). However, the increased B factor may actually arise in part from looser packing in the crystal due to the absence of the phospholipid bilayer (compare B with A, Figure 1). Unlike a simulation of the *B. halodurans* YidC in a lipid bilayer (Kumazaki et al., 2014a), the C1 region of our *E. coli* YidC was comparatively stable with root-mean-square fluctuations (RMSFs) of less than 3 Å (Figure S5D). This might be because the *B. halodurans* YidC simulation was carried out in a pure POPE lipid bilayer. Nevertheless, the C1 region in our simulation did have higher RMSFs compared with the membrane-buried TM helices. Significantly, helix PH1 had an exceptionally low RMSF, consistent with its apparent tight coupling to TM helix 2.

The compact and stable shape of YidC within the bilayer originates in part from impressive rings of aromatic residues on its cytoplasmic and periplasmic ends located at the hydrocarbon-head group boundaries of the bilayer (Figure 6). Beneath the cytoplasmic aromatic ring, a salt bridge cluster stabilizes the C1–C2 loop through internal salt bridges and salt bridges to the bilayer phosphate groups (Figure 6E). The NEM water accessibility data are generally consistent with the structure determined in the bilayer environment (Figures 1D, 4B, and 7F).

Our study clarifies the nature of the U-shaped hydrophilic cavity that penetrates roughly halfway into the membrane from the cytoplasmic surface (Kumazaki et al., 2014a,

2014b; Shimokawa-Chiba et al., 2015). The MD simulation shows that waters penetrate to the level of R366 and are highly dynamic (Movie S1), with little stable water penetration into the protein at the periplasmic surface in the simulation (Figure 4A). This dynamic water penetration allows solvent access observed in both the NEM and MD measurements (Figures 3 and 4B).

The MD simulation reveals novel features associated with the burial of R366 in the membrane environment. This strictly conserved arginine has an aromatic lid on its periplasmic side formed by tyrosines at positions 516 and 517 (Figures 7A and 7B). The Tyr residues interact with R366 via H bonds mediated by water molecules. Additional stabilization likely arises from cation-π interactions (Dougherty, 1996). We think it is significant that R366 and its tyrosine lid do not appear to be directly exposed to the TM segments of nascent chains, but rather sit well back from the critical TM3 and TM5 helices (Figures 7C–7E) that have been established as providing the principal interaction zones of incoming nascent chains (Klenner and Kuhn, 2012). It has been proposed that the hydrophilic region of a YidC substrate is accommodated in the hydrophilic cavity prior to its translocation, meaning that it could approach R366 (Kumazaki et al., 2014b). An alternative hypothesis is that R366 and its tyrosine lid provide a positively charged membrane indentation extending to the center of YidC, which shapes the local electric field within the protein. Such shaping of the local electric field might preclude the necessity for a direct approach by the nascent chain. Given the intriguing juxtapositions of R366, Y516, and Y517, and their very high sequence conservation, it is mysterious that each can be mutated to Cys or Ala without affecting *E. coli* viability. We suppose the high conservation results from evolutionary stresses that are not apparent under laboratory conditions.

The water accessibility of the protein in the context of the residues on TM3 and TM5 that contact incoming nascent chains (Figure 7) suggests an insertion pathway, in which there is a handoff of nascent chains from TM3 to TM5 via the aromatic cluster situated between TM3 and TM5. These aromatic residues may provide an electrically polarizable hydrophobic pathway that allows it to interact favorably with both non-polar and polar amino acids on the nascent chain on the periplasmic side of the membrane.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACTS FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Bacterial Growth, Cys Alkylation Assay, and Western Blots
  - Complementation Assay
  - Quantification of NEM Modification of YidC Single Cysteine Mutants
  - Molecular Dynamics Simulations

SUPPLEMENTAL INFORMATION

Supplemental Information includes six figures, four tables, four movies, and one PDB model and can be found with this article online at http://dx.doi.org/10.1016/j.str.2017.07.008.

AUTHOR CONTRIBUTIONS

R.E.D. and S.H.W. designed the experiments. S.C. carried out the MD simulations under the supervision of S.H.W. Y.C., L. Z., and P.G. performed the cysteine alkylation studies under the supervision of R.E.D. S.C., J.A.F., and S.H.W. analyzed the MD data. S.H.W. and R.E.D. wrote the paper, which was reviewed and approved by all authors. Both S.H.W. and R.E.D. are designated as corresponding authors because of the distinctly different methodologies used in the two labs.

ACKNOWLEDGMENTS

This work was supported by National Science Foundation grant MCB-1052033 (to R.E.D.) and NIH grants GM74637 and GM86685 (to S.H.W.). The simulations were performed on the HPC cluster at the University of California, Irvine. We thank Joseph Farran for excellent technical support.

Received: February 23, 2017
Revised: May 31, 2017
Accepted: July 21, 2017
Published: August 24, 2017

REFERENCES

Borowska, M.T., Dominik, P.K., Anghel, S.A., Kossiakoff, A.A., and Keenan, R.J. (2015). A YidC-like protein in the archaeal plasma membrane. Structure *23*, 1715–1724.

Burley, S.K., and Petsko, G.A. (1985). Aromatic-aromatic interaction: a mechanism of protein structure stabilization. Science *229*, 23–28.

Celebi, N., Yi, L., Facey, S.J., Kuhn, A., and Dalbey, R.E. (2006). Membrane biogenesis of subunit II of cytochrome *bo* oxidase: contrasting requirements for insertion of N-terminal and C-terminal domains. J. Mol. Biol. *357*, 1428–1436.

Celebi, N., Dalbey, R.E., and Yuan, J. (2007). Mechanism and hydrophobic forces driving membrane protein insertion of subunit II of cytochrome bo_3 oxidase. J. Mol. Biol. *375*, 1282–1292.

Chen, M., Samuelson, J.C., Jiang, F., Muller, M., Kuhn, A., and Dalbey, R.E. (2002). Direct interaction of YidC with the sec-independent Pf3 coat protein during its membrane protein insertion. J. Biol. Chem. *277*, 7670–7675.

Chen, Y., Soman, R., Shanmugam, K., Kuhn, A., and Dalbey, R.E. (2014). The role of the strictly conserved positively charged residue differs among the Gram-positive, Gram-negative, and chloroplast YidC homologs. J. Biol. Chem. *289*, 35656–35667.

Dalbey, R.E., and Kuhn, A. (2014). How YidC inserts and folds proteins across a membrane. Nat. Struct. Mol. Biol. *21*, 435–436.

Dalbey, R.E., Kuhn, A., Zhu, L., and Kiefer, D. (2014). The membrane insertase YidC. Biochim. Biophys. Acta *1843*, 1489–1496.

Darden, T., York, D., and Pedersen, L. (1993). Particle mesh Ewald: an $N\cdot\log(N)$ method for Ewald sums in large systems. J. Chem. Phys. *98*, 10089–10092.

Dougherty, D.A. (1996). Cation-pi interactions in chemistry and biology: a new view of benzene, Phe, Tyr, and Trp. Science *271*, 163–168.

Essmann, U., Perera, L., Berkowitz, M.L., Darden, T., Lee, H., and Pedersen, L.G. (1995). A smooth particle mesh Ewald method. J. Chem. Phys. *103*, 8577–8593.

Feller, S.E., Zhang, Y., Pastor, R.W., and Brooks, B.R. (1995). Constant pressure molecular dynamics simulation: the Langevin piston method. J. Chem. Phys. *103*, 4613–4621.

Frishman, D., and Argos, P. (1995). Knowledge-based protein secondary structure assignment. Proteins *23*, 566–579.

Grubmüller, H., Heller, H., Windemuth, A., and Schulten, K. (1991). Generalized Verlet algorithm for efficient molecular dynamics simulations with long-range interactions. Mol. Simul. *6*, 121–142.

Guan, L., and Kaback, H.R. (2007). Site-directed alkylation of cysteine to test solvent accessibility of membrane proteins. Nat. Protoc. *2*, 2012–2017.

Humphrey, W., Dalke, W., and Schulten, K. (1996). VMD: visual molecular dynamics. J. Mol. Graph. *14*, 33–38.

Jiang, F., Chen, M., Yi, L., de Gier, J.-W., Kuhn, A., and Dalbey, R.E. (2003). Defining the regions of *Escherichia coli* YidC that contribute to activity. J. Biol. Chem. *278*, 48965–48972.

Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., and Klein, M.L. (1983). Comparison of simple potential functions for simulating liquid water. J. Chem. Phys. *79*, 926–935.

Kalé, L., Skeel, R., Bhandarkar, M., Brunner, R., Gursoy, A., Krawetz, N., Phillips, J., Shinozaki, A., Varadarajan, K., and Schulten, K. (1999). NAMD2: greater scalability for parallel molecular dynamics. J. Comput. Phys. *151*, 283–312.

Kannan, N., and Vishveshwara, S. (1999). Identification of side-chain clusters in protein structures by a graph spectral method. J. Mol. Biol. *292*, 441–464.

Kedrov, A., Wickles, S., Crevenna, A.H., van der Sluis, E.O., Buschauer, R., Berninghausen, O., Lamb, D.C., and Beckmann, R. (2016). Structural dynamics of the YidC:ribosome complex during membrane protein biogenesis. Cell Reports *17*, 2943–2954.

Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J.E. (2015). The Phyre2 web portal for protein modeling, prediction, and analysis. Nat. Protoc. *10*, 845–858.

Killian, J.A., and von Heijne, G. (2000). How proteins adapt to a membrane-water interface. Trends Biochem.Sci. *25*, 429–434.

Kimura-Someya, T., Iwaki, S., Konishi, S., Tamura, N., Kubo, Y., and Yamaguchi, A. (2000). Cysteine-scanning mutagenesis around transmembrane segments 1 and 11 and their flanking loop regions of Tn10-encoded metal-tetracycline/H^+ antiporter. J. Biol. Chem. *275*, 18692–18697.

Klauda, J.B., Venable, R.M., Freites, J.A., O'Connor, J.W., Tobias, D.J., Mondragon-Ramirez, C., Vorobyov, I., MacKerell, A.D., Jr., and Pastor, R.W. (2010). Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. J. Phys. Chem. B *114*, 7830–7843.

Klenner, C., Yuan, J., Dalbey, R.E., and Kuhn, A. (2008). The Pf3 coat protein contacts TM1 and TM3 of YidC during membrane biogenesis. FEBS Lett. *582*, 3967–3972.

Klenner, C., and Kuhn, A. (2012). Dynamic disulfide scanning of the membrane-inserting Pf3 coat protein reveals multiple YidC substrate contacts. J. Biol. Chem. *287*, 3769–3776.

Kumazaki, K., Chiba, S., Takemoto, M., Furukawa, A., Nishiyama, K.-I., Sugano, Y., Mori, T., Dohmae, N., Hirata, K., Nakada-Nakura, Y., et al. (2014a). Structural basis of Sec-independent membrane protein insertion by YidC. Nature *509*, 516–520.

Kumazaki, K., Kishimoto, T., Furukawa, A., Mori, H., Tanaka, Y., Dohmae, N., Ishitani, R., Tsukazaki, T., and Nureki, O. (2014b). Crystal structure of *Escherichia coli* YidC, a membrane protein chaperone and insertase. Sci. Rep. *4*, 7299.

MacKerell, A.D., Jr., Bashford, D., Bellott, M., Dunbrack, R.L., Jr., Evanseck, J.D., Field, M.J., Fischer, S., Gao, J., Guo, H., Ha, S., et al. (1998). All-atom empirical potential for molecular modeling and dynamics studies of proteins. J. Phys. Chem. B *102*, 3586–3616.

MacKerell, A.D., Jr., Feig, M., and Brooks, C.L., II (2004). Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing conformational distributions in molecular dynamics simulations. J. Comput. Chem. *25*, 1400–1415.

Martyna, G.J., Tobias, D.J., and Klein, M.L. (1994). Constant-pressure molecular-dynamics algorithms. J. Chem. Phys. *101*, 4177–4189.

Phillips, J.C., Braun, B., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D., Kalé, L., and Schulten, K. (2005). Scalable molecular dynamics with NAMD. J. Comput. Chem. *26*, 1781–1802.

Rose, G.D., Geselowitz, A.R., Lesser, G.J., Lee, R.H., and Zehfus, M.H. (1985). Hydrophobicity of amino acid residues in globular proteins. Science *229*, 834–838.

Ryckaert, J.-P., Ciccotti, G., and Berendsen, H.J.C. (1977). Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes. J. Comput. Phys. *23*, 327–341.

Samuelson, J.C., Chen, M., Jiang, F., Möller, I., Wiedmann, M., Kuhn, A., Phillips, G.J., and Dalbey, R.E. (2000). YidC mediates membrane protein insertion in bacteria. Nature *406*, 637–641.

Samuelson, J.C., Jiang, F., Yi, L., Chen, M., de Gier, J.-W., Kuhn, A., and Dalbey, R.E. (2001). Function of YidC for the insertion of M13 procoat protein in *Escherichia coli*. J. Biol. Chem. *276*, 34847–34852.

Schibli, D.J., Epand, R.F., Vogel, H.J., and Epand, R.M. (2002). Tryptophan-rich antimicrobial peptides: comparative properties and membrane interactions. Biochem. Cell Biol. *80*, 667–677.

Schiffer, M., Chang, C.H., and Stevens, F.J. (1992). The functions of tryptophan residues in membrane proteins. Protein Eng. *5*, 213–214.

Serek, J., Bauer-Manz, G., Struhalla, G., van den Berg, L., Kiefer, D., Dalbey, R., and Kuhn, A. (2004). *Escherichia coli* YidC is a membrane insertase for Sec-independent proteins. EMBO J. *23*, 294–301.

Shimokawa-Chiba, N., Kumazaki, K., Tsukazaki, T., Nureki, O., Ito, K., and Chiba, S. (2015). Hydrophilic microenvironment required for the channel-independent insertase function of YidC protein. Proc. Natl. Acad. Sci. USA *112*, 5063–5068.

Tamura, N., Konishi, S., Iwaki, S., Kimura-Someya, T., Nada, S., and Yamaguchi, A. (2001). Complete cysteine-scanning mutagenesis and site-directed chemical modification of the Tn10-encoded metal-tetracycline/H^+ antiporter. J. Biol. Chem. *276*, 20330–20339.

Ulmschneider, M.B., Sansom, M.S.P., and Di Nola, A. (2005). Properties of integral membrane protein structures: derivation of an implicit membrane potential. Proteins *59*, 252–265.

van Bloois, E., Haan, G.J., de Gier, J.-W., Oudega, B., and Luirink, J. (2004). F_1F_0 ATP synthase subunit c is targeted by the SRP to YidC in the *E. coli* inner membrane. FEBS Lett. *576*, 97–100.

van der Laan, M., Bechtluft, P., Kol, S., Nouwen, N., and Driessen, A.J.M. (2004). F_1F_0 ATP synthase subunit c is a substrate of the novel YidC pathway for membrane protein biogenesis. J. Cell Biol. *165*, 213–222.

Wallin, E., Tsukihara, T., Yoshikawa, S., von Heijne, G., and Elofsson, A. (1997). Architecture of helix bundle membrane proteins: an analysis of cytochrome *c* oxidase from bovine mitochondria. Protein Sci. *6*, 808–815.

Wickles, S., Singharoy, A., Andreani, J., Seemayer, S., Bischoff, L., Berninghausen, O., Soeding, J., Schulten, K., van der Sluis, E.O., and Beckmann, R. (2014). A structural model of the active ribosome-bound membrane protein insertase YidC. Elife *3*, e03035.

Wiener, M.C., and White, S.H. (1992). Structure of a fluid dioleoylphosphatidylcholine bilayer determined by joint refinement of X-ray and neutron diffraction data. III. Complete structure. Biophys. J. *61*, 434–447.

Wu, E.L., Cheng, X., Jo, S., Rui, H., Song, K.C., Dávila-Contreras, E.M., Qi, Y., Lee, J., Monje-Galvan, V., Venable, R.M., et al. (2014). CHARMM-GUI *Membrane Builder*. Toward realistic biological membrane simulations. J. Comput. Chem. *35*, 1997–2004.

Yau, W.-M., Wimley, W.C., Gawrisch, K., and White, S.H. (1998). The preference of tryptophan for membrane interfaces. Biochemistry *37*, 14713–14718.

Yi, L., Jiang, F.L., Chen, M.Y., Cain, B., Bolhuis, A., and Dalbey, R.E. (2003). YidC is strictly required for membrane insertion of subunits a and c of the F_1F_0ATP synthase and SecE of the SecYEG translocase. Biochemistry *42*, 10537–10544.

Yi, L., Celebi, N., Chen, M., and Dalbey, R.E. (2004). Sec/SRP requirements and energetics of membrane insertion of subunits a, b, and c of the *Escherichia coli* F_1F_0 ATP synthase. J. Biol. Chem. *279*, 39260–39267.

Yu, Z., Koningstein, G., Pop, A., and Luirink, J. (2008). The conserved third transmembrane segment of YidC contacts nascent *Escherichia coli* inner membrane proteins. J. Biol. Chem. *283*, 34635–34642.

Zhu, L., Klenner, C., Kuhn, A., and Dalbey, R.E. (2012). Both YidC and SecYEG are required for translocation of the periplasmic loops 1 and 2 of the multispanning membrane proteinTatC. J. Mol. Biol. *424*, 354–367.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-6X His tag® antibody (HRP)	Abcam	Ab1187
Bacterial and Virus Strains		
E. coli: BL21(DE3) Competent Cells	Agilent Technologies	200131
E. coli: JS7131	Samuelson et al., 2000	Lab collection
Chemicals, Peptides, and Recombinant Proteins		
N-Ethylmaleimide	Thermo Scientific	23030
Methoxypolyethylene glycol maleimide	Sigma-Aldrich	63187
Polyethylene Glycol 6000	Hampton Research	HR2-513
DTT	Anatrace	D2000
IPTG	Anatrace	I1003
Kanamycin Sulfate	Teknova	K2150
Lysozyme from Chicken Egg White	Sigma-Aldrich	L6876
BupH^{TM} Phosphate Buffered Saline Packs	Thermo Scientific	28372
EDTA	OmniPur	4010
Sucrose Crystal	Mallinckrodt	8360
L-(+)-Arabinose	Sigma-Aldrich	A3256
D-(+)-Glucose	Sigma-Aldrich	G8270
Critical Commercial Assays		
QuickChange Site Directed Mutagenesis Kit	Agilent Technologies	200515
NucleoSpin® Plasmid Purification Kit	Clontech Laboratories	740499.250
HyGLO Chemiluminescent HRP Antibody Detection Reagent	Denville Scientific Inc.	E-2500
Deposited Data		
Crystal Structure of YidC from *E. coli*	Kumazaki et al., 2014b	PDB: 3WVF
PDB coordinates for 1 frame of the YidC simulation in a POPE/POPG lipid bilayer are available upon request to stephen.white@uci.edu		
Experimental Models: Organisms/Strains		
E. coli: BL21(DE3) Competent Cells	Agilent Technologies	200131
E. coli: JS7131	Samuelson et al., 2000	Lab collection
Oligonucleotides		
See Table S4 for a full list of oligonucleotides used.	This paper	N/A
Recombinant DNA		
pEH1-Cys-less YidC (C423S) with C-terminal His_6-tag	Jiang et al., 2003	Lab collection
Software and Algorithms		
ImageJ	NIH	https://imagej.nih.gov/ij/
NAMD 2.9 and VMD 1.9.3 software	Univ. of Illinois, Urbana- Champaign	http://www.ks.uiuc.edu/Research/namd/
Delaunay Triangulation scripts for VMD	Univ. of Illinois, Urbana- Champaign	http://www.ks.uiuc.edu/Research/vmd/script_library/scripts/delaunay_triangulate/
Other		
Nitrocellulose Membrane	Bio-Rad	162-0115
HyBlot CL® Autoradiography Film	Denville Scientific Inc.	E3012

CONTACTS FOR REAGENT AND RESOURCE SHARING

Regarding molecular dynamics simulations, contact Stephen H. White at stephen.white@uci.edu. Regarding NEM labeling experiments, contact Ross E. Dalbey at dalbey@chemistry.ohio-state.edu

EXPERIMENTAL MODEL AND SUBJECT DETAILS

BL21 (DE3) cells used for Cys alkylation assay were purchased from Agilent Technologies (200131). They were grown in LB media at 37°C with shaking.

JS7131 cells used in complementation assay were from our lab collection. They were grown in LB media supplemented with 0.2% arabinose at 37°C with shaking.

METHOD DETAILS

Bacterial Growth, Cys Alkylation Assay, and Western Blots

Buffer A consisted of 10mM potassium phosphate buffer (pH 6.5), 1mM EDTA, 20% (w/v) sucrose, and 1 mg/ml lysozyme. Buffer B consisted of 10mM potassium phosphate buffer (pH 6.5) and 1mM EDTA. PBS buffer consisted of 0.1M $NaPO_4$, 0.15M NaCl (pH 7.2).

Single Cys mutants of YidC were made using the pEH1-YidC (C423S, Cys-less) vector as a template. An overnight culture of BL21 (DE3) bearing the single Cys pEH1-YidC expression vector was back-diluted 1:100 into fresh LB media containing kanamycin (50 μg/ml final concentration). Cells were grown to mid-log phase (OD600 ≈ 0.55) and IPTG was added to 1mM final concentration to induce expression of the YidC protein for 30min. The cells were subsequently washed twice with PBS buffer and placed on ice.

For the Cys alkylation assay, the cell density was adjusted to OD600 ≈ 2.0 and split into 3 aliquots (0.2ml for each tube). The first aliquot was treated with NEM (0.5 mM final concentration) while the second and third aliquots were left untreated. These aliquots were then held at 30°C for 30min and subsequently washed twice with PBS buffer (with 5mM DTT) and three times with PBS buffer (without DTT). To remove YidC aggregates, cells were harvested, resuspended in 20 μl buffer A and incubated for 30 min on ice. Buffer B (380 μl) was then added, followed by sonication with Sonic Dismembrator Model 500 (Fisher Scientific) (microtip, 40% duty, eight cycles) while cooling. Intact cells and aggregates were removed by centrifugation at 16,000 g for 30 min at 4°C. Supernatants were collected and precipitated with 10% (final concentration) ice-cold trichloroacetic acid for 1h on ice. The pellets were washed with ice-cold acetone and solubilized with 100μl Tris-SDS-Urea buffer [15 mM Tris–HCl, 6M Urea, 2% SDS, pH 7.5]. The first and third aliquots was treated with Mal-PEG (M.W.∼5000 Da, 5mM final concentration) for 40 min at 37 °C, while aliquot 2 was mock-treated with 5mM PEG6000. The protein samples were then analyzed by 10% SDS-PAGE followed by Western blotting. The YidC proteins with a C-terminal His-tag were detected using anti-6His antibody (HRP conjugated, from Abcam) and a Chemiluminescent detection kit (Pierce).

Complementation Assay

The YidC depletion strain, JS7131 bearing pEH1-YidC encoding the single cysteine mutants were cultured in LB media containing 0.2% arabinose and 50μg/ml Kanamycin at 37°C. Overnight cultures were washed twice with fresh LB media and back diluted 1:100 into LB media without arabinose. After 2h of growth, the cells were serially diluted (1:10, 1:100, 1:1000 and 1:10,000 fold) in LB medium. Aliquots of the diluted cells were then spotted on LB plates containing 50μg/ml Kanamycin, and 0.2% arabinose or 0.2% glucose, respectively. Where indicated, 50μM IPTG was added and the plates were incubated overnight at 37 °C.

Quantification of NEM Modification of YidC Single Cysteine Mutants

The optical density of the Mal-PEG-shifted and the non-shifted YidC bands were quantified using ImageJ software developed at the National Institutes of Health (https://imagej.nih.gov/ij/) using modification efficiency E = (Intensity of non-shifted YidC protein)/[(Intensity of non-shifted YidC protein) + (Intensity of shifted YidC protein)].

Molecular Dynamics Simulations

We carried out all-atom molecular dynamics simulation of YidC embedded in a lipid bilayer using NAMD 2.9 software (Kalé et al., 1999; Phillips et al., 2005) with CHARMM 36 (Klauda et al., 2010) for the lipids and CHARMM 22 with the CMAP correction (MacKerell et al., 1998, 2004) for the protein. TIP3P model (Jorgensen et al., 1983) was used to describe water molecules. To start the simulation, we used the coordinates from the *E. coli* YidC crystal structure (Kumazaki et al., 2014b) (PDB code: 3WVF). Missing residues 480-492 between TM4 and TM5 were modeled using Phyre2 software (Kelley et al., 2015). We modeled all the His amino acids in the Nδ1 tautomeric state and maintained positive charge for Arg/Lys and negative charge for Asp/Glu. This complete YidC structure was then minimized in vacuum and used as a starting configuration to set up the atomistic simulation. Using VMD software (Humphrey et al., 1996), we aligned the principal axis of YidC with the x, y, and z space directions. We then embedded YidC in a bilayer formed by 500 lipid molecules (250 each leaflet) constructed with CHARMM-GUI software (Wu et al., 2014). The bilayer was formed with 75% of POPE and 25% of POPG lipid molecules, meaning 186 POPE and 64 POPG in each leaflet. We used these percentages of lipid species in order to mimic the real membrane environment in which YidC is found. We solvated the system with 89 waters per lipid and added 116 potassium ions in order to conserve electrical neutrality. Before starting the minimization and equilibration of the system,

we ensured that steric clashes were avoided. The simulation was carried out for ~185 ns, as described in detail in Supplemental Information.

We minimized the system energy using a conjugate gradient algorithm for 8000 steps heated it through a temperature ramp made of steps of 25 K starting from 25 K up to 300 K. In order to equilibrate the positions of the lipids, water molecules, ions, and the protein, we used harmonic restraints, which have been switched off in a series of consecutive runs of 1 ns length. Once the equilibration procedure was completed, we started the production run in the *NPT* (constant number of particle *N*, pressure *P*, and temperature *T*) ensemble. We used a time step of 1 fs and a reversible multiple time-step algorithm (Grubmüller et al., 1991) to integrate the equations of motion with time steps of 1 fs for the bonded forces, 2 fs for the short-range non-bonded forces, and 4 fs for the long-range electrostatic interactions. These last interactions were computed using the smooth particle mesh Ewald summation (Darden et al., 1993; Essmann et al., 1995). The short-range interactions were cut off at 12 Å by means of a switching function between 10 and 12 Å. The SHAKE algorithm (Ryckaert et al., 1977) was used to maintain the length of the bonds involving hydrogen atoms. The temperature was maintained at 300 K using a Langevin dynamics scheme, and a flexible cell in association with the Nosé-Hoover Langevin piston to keep the pressure constant at 1 atm (Feller et al., 1995; Martyna et al., 1994). The simulation was carried out for ~185 ns; trajectories were saved every 10 ps. We monitored the secondary structure by using STRIDE plugin of VMD (Frishman and Argos, 1995). We calculated the solvent exposure of the YidC TM domain by monitoring over the full length of the simulation the residence time of the water molecules within 6 Å of each residue. Then, we normalized each value using the standard accessible surface areas (Rose et al., 1985).

Structure, Volume *25*

Supplemental Information

YidC Insertase of *Escherichia coli*: Water Accessibility and Membrane Shaping

Yuanyuan Chen, Sara Capponi, Lu Zhu, Patrick Gellenbeck, J. Alfredo Freites, Stephen H. White, and Ross E. Dalbey

save date: 20 Jul 2017, 09:33

Supplementary Information

YidC Insertase of *Escherichia coli*: Water Accessibility and Membrane Shaping

Yuanyuan Chen[1†], Sara Capponi[2,3†], Lu Zhu[1,4], Patrick Gellenbeck[1], J. Alfredo Freites[5], Stephen H. White[2,6*], and Ross E. Dalbey[1*]

[1]Department of Chemistry and Biochemistry, The Ohio State University, Columbus, OH 43210

[2]Department of Physiology and Biophysics, University of California, Irvine, CA 92697

[3]Present address: Cardiovascular Research Institute, Department of Pharmaceutical Chemistry, University of California, San Francisco, CA 94158

[4]Present address: Department of Molecular Biology and Genetics and Weill Institute for Cell and Molecular Biology, Cornell University, Ithaca, NY 1485

[5]Department of Chemistry, University of California, Irvine, CA 92697

[†]These authors contributed equally. [6]Lead Contact.

*Correspondence:
Molecular dynamics simulations: stephen.white@uci.edu
NEM labeling experiments: dalbey@chemistry.ohio-state.edu

Running Title: YidC solvent accessibility and membrane shaping

Part 1. NEM Labeling Data

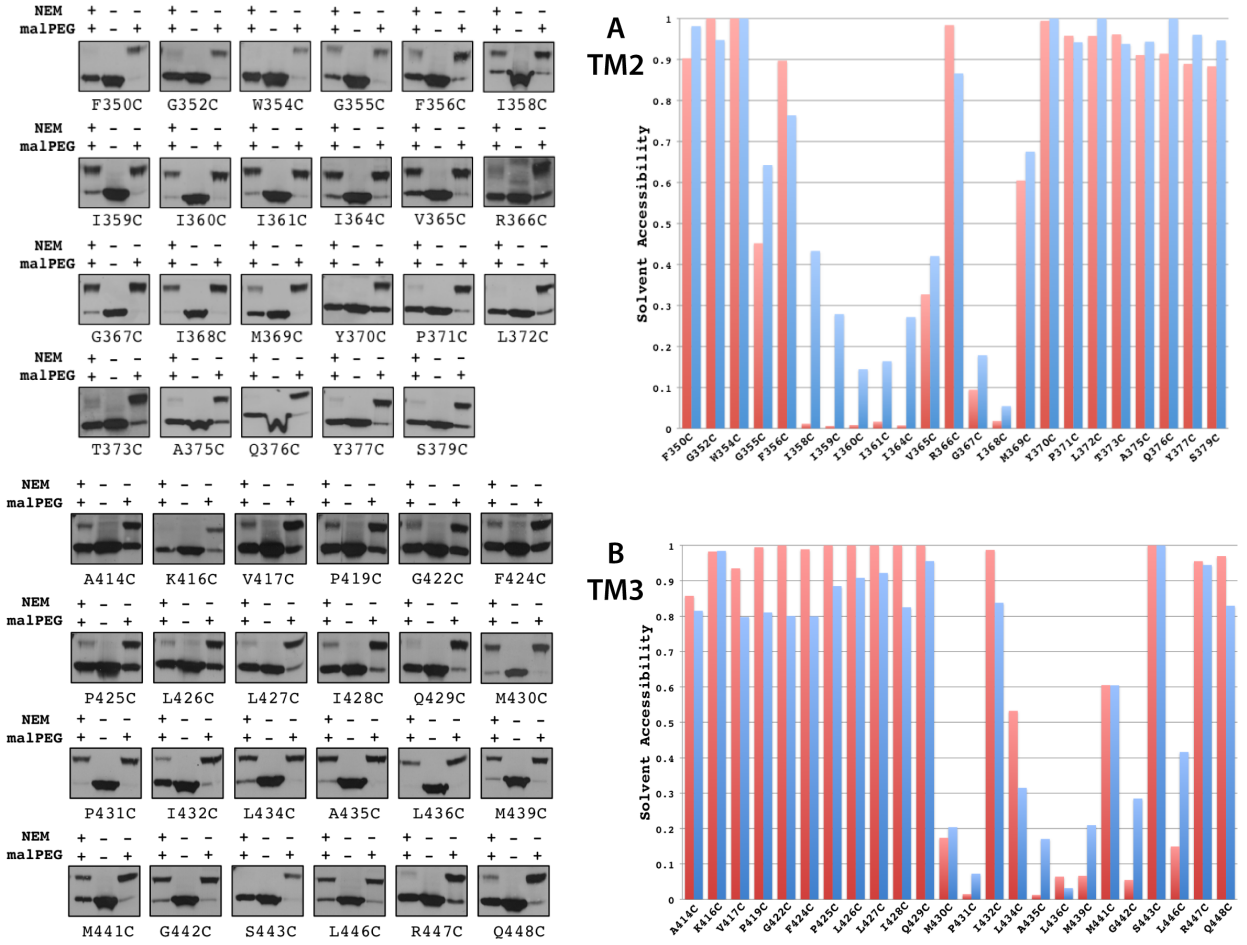


Figure S1 supports Figure2 and Figure 3A. Accessibility of Cys residues introduced into TM2 and TM3 to map the solvent and lipid exposed residues. NEM modification of Cys scanning mutants: (**A**) TM2 and (**B**) TM3 (left-hand panels). BL21 cells expressing the Cys mutants were analyzed for modification by NEM using the indirect gel shift assay involving Mal-PEG, as described in Figure 2A. The proteins were analyzed by SDS-PAGE and YidC was detected by Western blotting using the 6x His antibody. For this particular data set, blue bars indicate the percent modifications of Cys residues in TM2 and TM3 by NEM. The red bars summarize data from a second data set. The data were quantified as described in "Experimental Procedures". The average values of the two data sets are shown in Figure 3A.

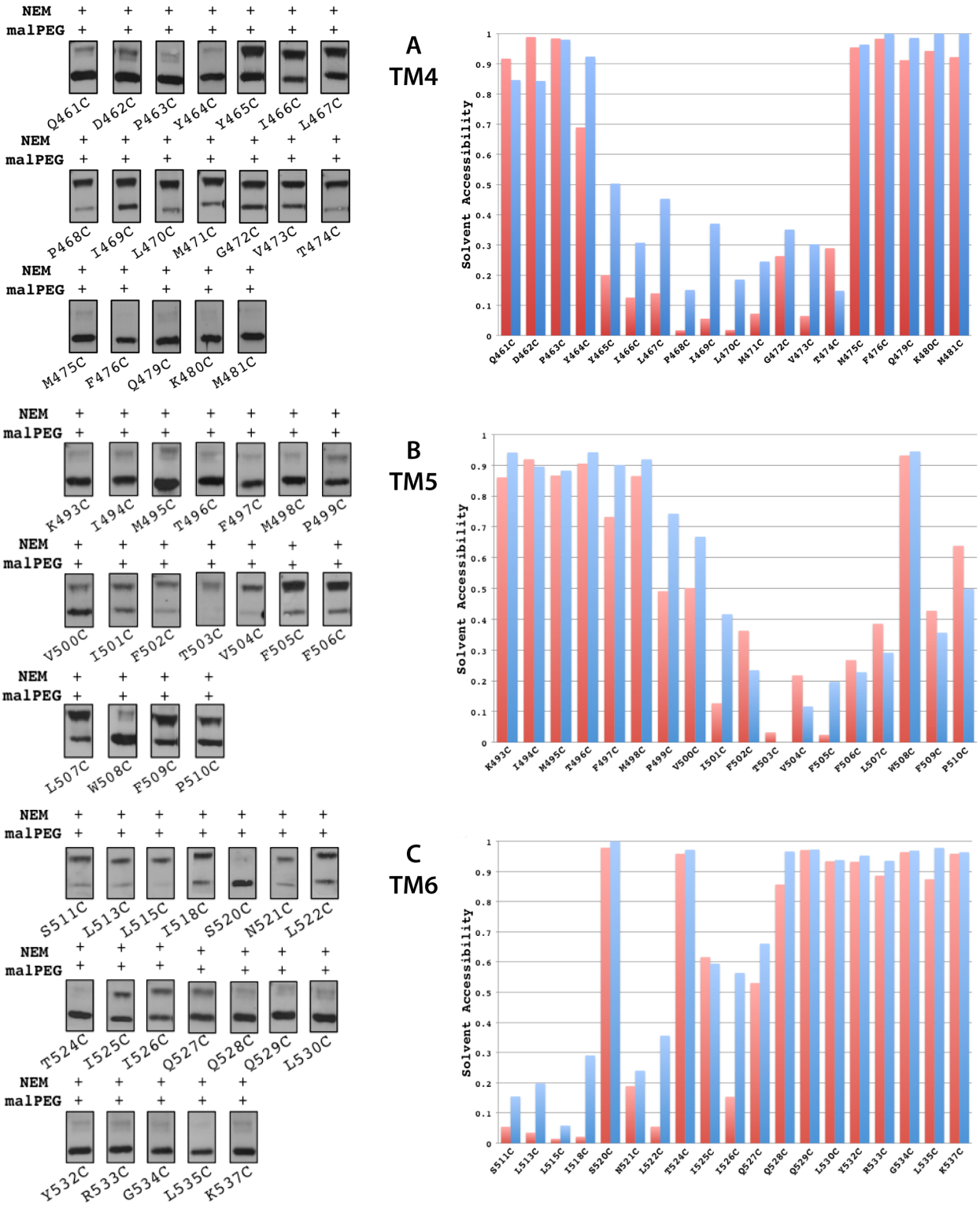


Figure S2 supports Figure2 and Figures 3A,B. Accessibility of Cys residues introduced into TM4 (**A**), TM5 (**B**), and TM6 (**C**) to map the solvent and lipid exposed residues. NEM modifications of Cys scanning mutants are shown in the left-hand panels. BL21 cells expressing the Cys mutants were analyzed for modification by NEM using the indirect gel shift

assay involving Mal-PEG, as described in Figure 2A. The proteins were analyzed by SDS-PAGE and YidC was detected by Western blotting using the 6x His antibody. For this particular data set, blue bars indicate the percent modifications of Cys residues in TM4, TM5, and TM6 by NEM. The red bars summarize data from a second data set. The data were quantified as described in “Experimental Procedures”. The average values of the two data sets are shown in Figure 3B.

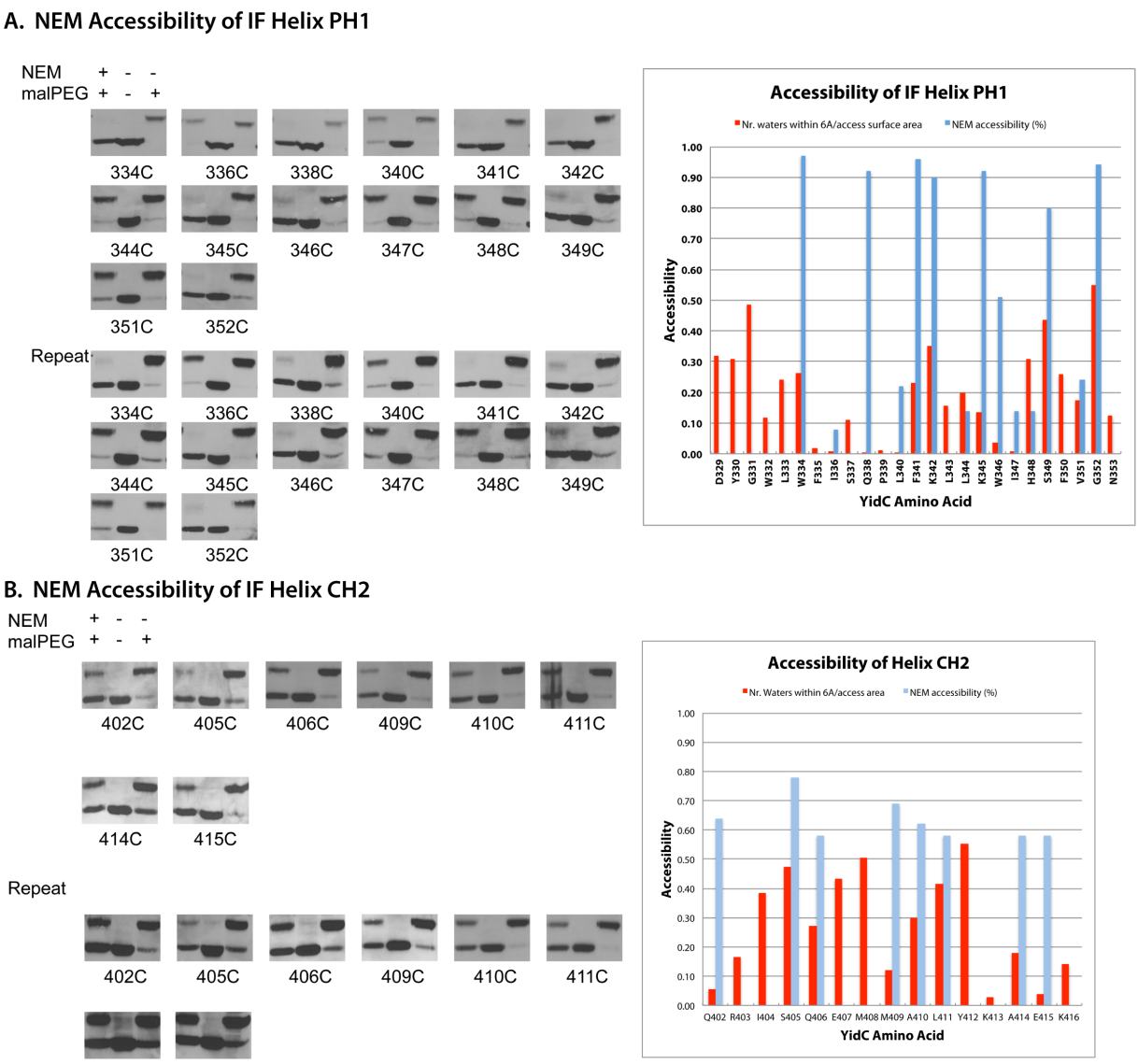


Figure S3 in support of Figure 1D. Accessibility of Cys residues introduced into IF Helix PH1 (**A**) and IF Helix CH2 (**B**). The same NEM labeling protocol described in Methods and Figure S1 and FigureS2 (above) was used. The right-hand panels summarize the exposure data from the NEM labeling (blue bars) and MD simulations (red bars). The values of the blue bars in panels A and B represent the means of the two NEM measurements.

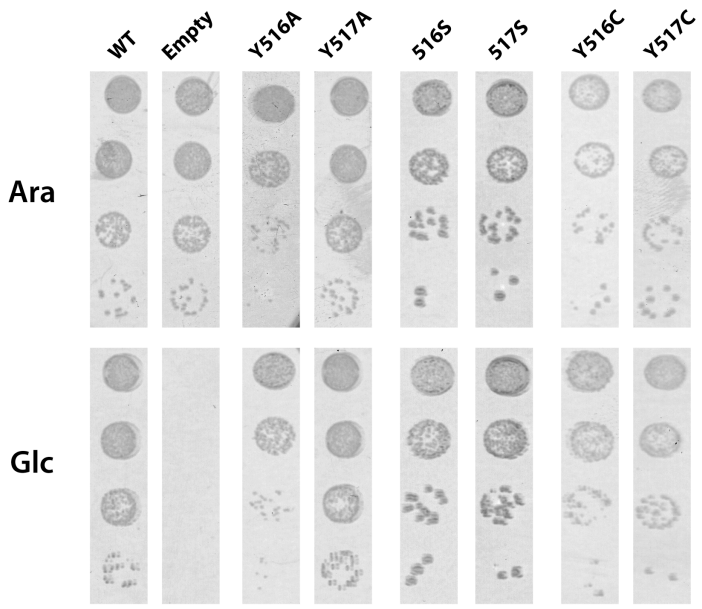


Figure S4 in support of Figure 7. Complementation assay for mutations of Y516 and Y517 to alanine, serine, and cysteine.

Part 2. Molecular Dynamics Data

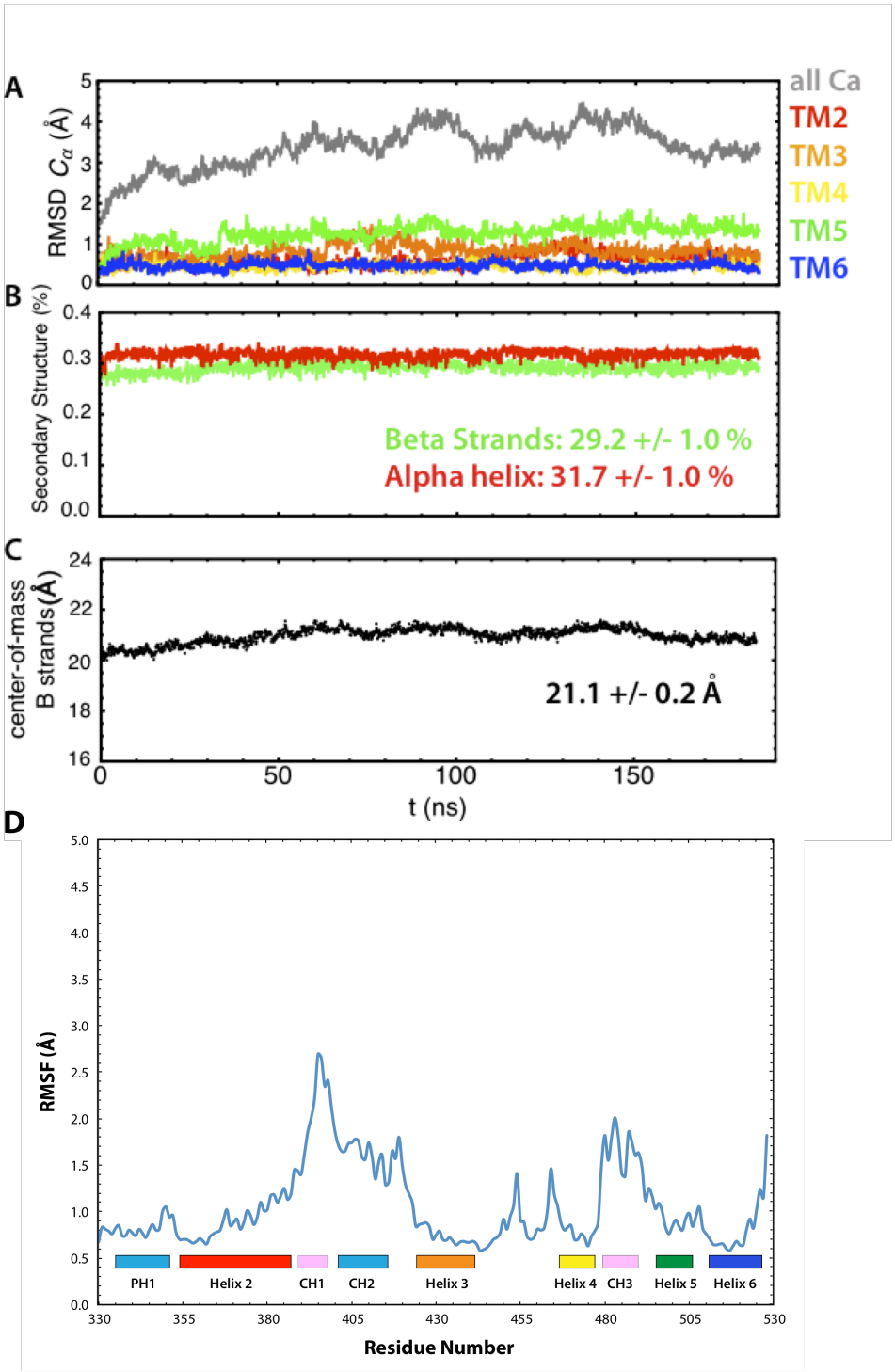


Figure S5 in support of Figure 1A. Stability of YidC structure over the course of the 185 ns simulation**.** The RMSD of the C_α atoms of the transmembrane helices remained smaller than 2

Å during the entire simulation (Figure S5A). This indicates that YidC transmembrane domain is rather rigid and stable. In order to check for potential drifting of the periplasmic domain, we monitored the center of mass of the β-strands of the P1 domain (Figure S5C), which we also found to be stable in time **A**. The time evolution of the root mean squared deviations (RMSD) of YidC TM helices. TM2 is colored in *red*, TM3 in *orange*, TM4 in *yellow,* TM5 in *green*, and TM6 in *blue*. **B**. Time evolution of the secondary structure of YidC. The protein was very stable structurally over the full course of the simulation. **C.** The stability of the P1 domain is indicated by the stable time evolution of the center of mass of the domain. **D**. Root mean-squared fluctuations (RMSF) of the YidC α-carbons over the full 185 ns course of the all-atom molecular dynamics simulation. The secondary-structure elements, represented as rectangles below the plot, are colored as in Figure 1A.

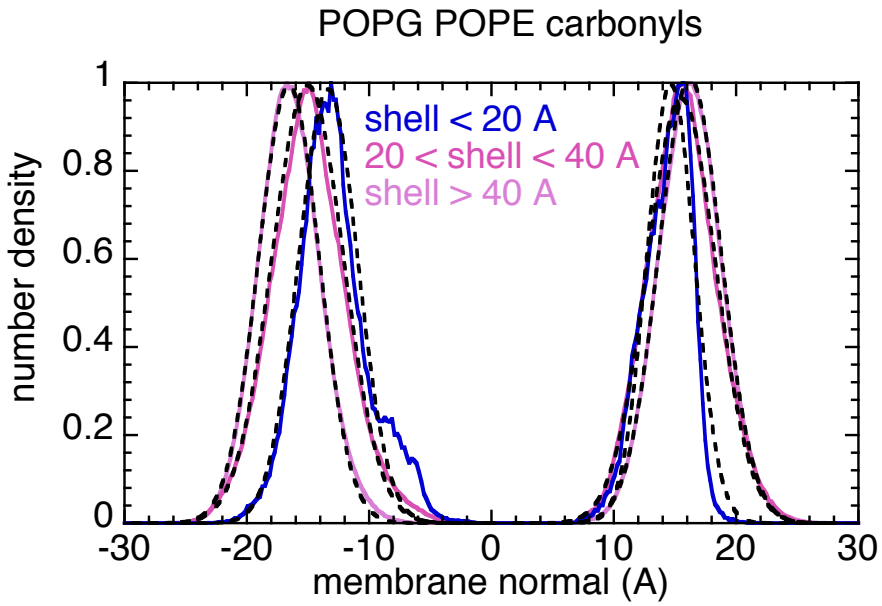


Figure S6 in support of Figure 5C. Time-averaged transbilayer distributions of the POPE and POPG carbonyl groups along the membrane normal. The distributions of the carbonyl groups were calculated for lipid molecules located within cylindrical shells centered on the YidC center of mass. Black dashed lines indicate gaussian fits to the distributions.

Supplementary Table S2

supporting Figure 4A and Figure 7

Accessibilities of E. coli YidC residues R366, Y516, & Y517

Residue	*Residue A^0 ($Å^2$)	†Average Nr. of waters within 6 Å	(Waters within 6 Å)÷A^0	NEM accessibility
R366	256.0	79.50	0.311	0.92
Y516	236.8	8.88	0.038	0.08
Y517	236.8	32.73	0.138	0.34

*A^0: Standard accessible surface area. †From molecular dynamics simulation.

Table S3
supporting Figure 5C

Membrane thicknesses calculated as the distance between the peaks of Gaussian functions fit to the distributions of the carbonyls groups belonging to the POPE and POPG lipid molecules (Figure S6). The difference in the membrane thickness calculated in the shell closest to YidC and faraway from YidC is 4.77 Å. However, relative to the YidC center of mass, this difference results from 3.2 Å on the cytoplasmic side of the membrane and of 1.57 Å on the periplasmic side.

shell	Membrane thickness (Å)
lipids < 20 Å	28.01 ± 0.25
20 < lipids < 40 Å	30.55 ± 0.31
lipids > 40 Å	32.78 ± 0.28