

A Hydrophilic Microenvironment in the Substrate-Translocating Groove of the YidC Membrane Insertase is Essential for Enzyme Function

Yuanyuan Chen¹, Marcos Sotomayor¹, Sara Capponi^{2,3}, Balasubramani Hariharan¹, Indra D. Sahu^{4,7}, Maximilian Haase⁵, Gary A. Lorigan⁴, Andreas Kuhn⁵, Stephen H. White⁶, and Ross E. Dalbey^{1*}

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

²Department of Industrial and Applied Genomics, IBM AI and Cognitive Software Organization, IBM Almaden Research Center, San Jose, CA

³NSF Center for Cellular Construction, University of California in San Francisco, San Francisco, CA, USA

⁴Department of Chemistry and Biochemistry, Miami University, Oxford, Ohio 45056

⁵Institute of Microbiology and Molecular Biology, University of Hohenheim, Stuttgart, Germany, 70599

⁶Department of Physiology and Biophysics, University of California, Irvine, CA 92697, USA

⁷ Natural Science Division, Campbellsville University, Campbellsville, KY 42718

*correspondence: dalbey@chemistry.ohio-state.edu (Ross E. Dalbey)

Running Title: Role of the Polar Groove of YidC in Membrane Insertion

Keywords: YidC; *S. mutans*; molecular dynamics; membrane biogenesis; YidC; membrane transport

ABSTRACT

The YidC family of proteins are membrane insertases that catalyze the translocation of the periplasmic domain of membrane proteins via a hydrophilic groove located within the inner leaflet of the membrane. All homologs have a strictly conserved, positively charged residue in the center of this groove. In *B. subtilis*, the positively charged residue has been proposed to be essential for interacting with negatively charged residues of the substrate, supporting a hypothesis that YidC catalyzes insertion via an early-step electrostatic attraction mechanism. Here we provide data suggesting that the positively charged residue is important not for its charge, but for increasing the hydrophilicity of the groove. We found that the positively charged residue is dispensable for *E. coli* YidC function when an adjacent residue at position 517 was hydrophilic or aromatic, but was essential when the adjacent residue was apolar. Additionally, solvent accessibility studies support the idea that the conserved positively charged residue functions to keep the top and middle of the groove sufficiently hydrated. Moreover, we demonstrate that both the *E. coli* and *S. mutans* YidC homologs are functional when the strictly conserved arginine is replaced with a negatively charged residue, provided proper stabilization from neighboring residues. These combined results show that the positively charged residue functions to maintain a hydrophilic microenvironment in the groove necessary for the insertase activity, rather than to form electrostatic interactions with the substrates.

Introduction

The YidC/Oxa1/Alb3 proteins are found in bacteria, mitochondria, and chloroplast where they play a pivotal role in membrane protein biogenesis (1,2). More recently, homologs have also been found in the eukaryotic ER membrane (3-5). This includes the Get1, EMC3 and TMCO1 proteins that function within large membrane complexes that play a key role in ER membrane protein biogenesis (5). In bacteria, YidC can function on its own (6-11) or in concert with the Sec system to insert proteins into the cytoplasmic membrane (12-14). YidC helps facilitate the removal of transmembrane proteins from the Sec channel (15) and assists in the formation of α -helical bundles of membrane proteins (16,17).

In 2014, the structures of YidC from *Bacillus halodurans* (18) and *Escherichia coli* (19) were solved at high resolution using X-ray crystallography. These structures showed that the YidC protein possesses a five transmembrane segment core domain (TM2-TM6 in the case of the *E. coli* YidC) with an unusual hydrophilic cavity within the inner leaflet of the membrane that is accessible from the cytoplasm and lipid bilayer, but not from the periplasm. In addition, YidC possesses a conserved coiled-coil region in the cytoplasm comprised of the two α -helices CH1 and CH2 that most likely is involved in substrate binding of YidC substrates such as Pf3 coat (20), as well as binding of SRP and the SRP receptor FtsY that targets ribosome nascent chains to YidC (21). Interestingly, CH1 forms a continuous helix with the first TM region of the core domain (TM2 of *E. coli* YidC), which is kinked by two conserved prolines, one in TM1 of the core (TM2 of *E. coli* YidC) and one in front of CH1. These prolines along with the disordered loop (Pro-Leu-Gly-Gly-Cys-Phe-Pro in *E. coli* YidC) connecting CH2 to second TM core segment (TM3 of *E. coli* YidC) likely function to move the substrate from the coiled-coil domain into the hydrophilic groove.

The *E. coli* YidC residues that have been determined as contacting the transmembrane (TM) segment of substrates during insertion were found to cluster on the same face of TM3 and TM5 that forms a 'greasy slide' (1,22) while the hydrophilic region of the substrate was incorporated transiently into the hydrophilic groove prior to translocation across the membrane (23). The cavity contains a strictly conserved positively charged residue that is essential in *Bacillus subtilis* (19). This residue has been proposed to participate in an electrostatic step to translocate the negatively charged N-terminal tail region of the MifM protein across the membrane (19). However, the conserved positive charge is not essential for the insertase activity of the *E. coli* YidC or an *Arabidopsis thaliana* Alb3 derivative (24).

A molecular dynamics simulation of *E. coli* YidC embedded in a POPE:POPG phospholipid bilayer showed the protein to be more compact than the crystal structure and reveals significant thinning in the vicinity of the protein (25). The thinning of the membrane bilayer and the presence of the hydrophilic groove reduces the energy cost for translocation, which is a new paradigm in biology (26). The strictly conserved arginine 366 in the hydrophilic groove of the *E. coli* YidC is in proximity to tyrosine 516 and 517 at the top of the groove (25) (Fig. 1), and is likely hydrogen bonded via water molecules. These tyrosine residues are part of an aromatic cluster at the interface between the aqueous environment in the groove and the embedded hydrophobic portion of YidC in the outer leaflet of the membrane.

To gain further insights into the role of R366 in the YidC insertase activity, we investigated the function of this conserved positively charged residue in the hydrophilic groove of the *E. coli* YidC and *Streptococcus mutans* YidC2. We found that the positive charge (R366) is important for the *E. coli* YidC when a nearby tyrosine at position 517 is changed to an amino acid with an apolar side chain. However, it is not essential when the tyrosine is changed to a hydrophilic or aromatic amino acid. Similar results were also found with *Streptococcus mutans* YidC2. Solvent accessibility assays show that the arginine is required to keep the upper part of the groove solvent accessible when Y517 is replaced with an apolar residue. Strikingly, both *E. coli* YidC and *S. mutans* YidC2 can function in membrane protein insertion when the arginine is replaced with a negatively charged glutamic acid residue, provided that a suppressor mutation is also present, which most likely stabilizes the groove region. Taken together, these results support the hypothesis that the positively charged residue is employed to keep the groove hydrophilic and water exposed rather than to attract the negative charge region of the substrate to be translocated.

Results

The conserved positively charged residue R366 of E. coli YidC becomes functionally indispensable by mutation of tyrosine 517

Previous structural (18) and MD simulation studies (25) revealed that the aromatic residues at 516 and 517 in the *E. coli* YidC are located above arginine 366 at the top of the hydrophilic groove (Fig. 1B). Our first question concerned the necessity of these two aromatic residues. Y516, Y517, or both Y516 and Y517 were mutated to alanines. The activities of these mutants were determined utilizing the YidC depletion strain JS7131, with the endogenous *yidC* knocked out and a new copy of *yidC* gene under the control of the *araBAD* promoter introduced at the lambda attachment site (6). The JS7131 cells harboring wild-type (WT) YidC or the YidC mutants

were analyzed by serial dilution on a LB glucose or arabinose plate at 37°C. We found that the Ala substitution of Y517 (Fig. 2A, Fig. S1), Y516 (Fig. S1), or Y516/Y517 (Fig. S1) complemented the YidC depletion strain. However, when R366 was mutated to Ala, then Y517 was important for activity but not Y516. Ala substitutions of both R366 and Y517 did not complement the YidC depletion strain (Fig. 2A; Fig. S1), while the single Ala substitutions of R366 (Fig. 2A) and double R366/ Y516 did complement (Fig. S1). Next, we examined directly the activity of the mutants by testing the membrane insertion of two YidC-dependent proteins. First, we analyzed the model protein Procoat-Lep (Fig. 1C). Procoat-Lep corresponds to procoat with its C-terminal cytoplasmic region extended by 101 amino acids of the periplasmic domain of leader peptidase (Lep) (9). If procoat-Lep (PC) inserts across the membrane, it is cleaved by signal peptidase 1 and converted to coat-Lep (C). The R366A and Y517A YidC single mutants were fully active in inserting procoat-Lep while the double R366A/Y517A mutant was inhibited leaving PC partially uncleaved (Fig. 2B, -Ara condition). Similar results were seen with the Pf3-23Lep substrate, which has the Lep residues 23-323 fused after the transmembrane segment of Pf3 coat and an arginine introduced after the TM of Pf3-23Lep to prevent translocation of the Lep C-terminal domain (24) (Fig. 1C). When Pf3-Lep (P) inserts across the membrane, proteinase K (P.K.) digests the N-tail of Pf3-Lep and converts it to a smaller fragment (F). Figure 2B shows that Pf3-23Lep was inserted efficiently with either the R366A or Y517A single mutant, while insertion was blocked with the R366A/Y517A double mutant (Fig. 2B, - Ara condition). It is worth noting that unlike Pf3-23Lep, which showed almost complete loss of insertion with the double mutant, procoat-Lep was only partially inhibited. This has also been observed previously (24), possibly due to the different insertion mechanisms employed by the two substrates.

For all the mutants described here, the protein was stably expressed as determined by Western blotting (Fig. S5). The requirement of the positive charge was confirmed by substituting R366 with neutral (Ala, Cys, Asn), positively charged (Lys) or negatively charged (Asp) residues in the presence of Y517A mutation. As shown in Figure S1, only the lysyl mutant, R366K/Y517A, is active as it complemented the YidC depletion strain. The combined results show that we were able to make the arginine 366 in *E. coli* YidC essential by mutating the nearby tyrosine 517 residue.

Functional importance of the E. coli YidC conserved positively charged residue depends on the identity of residue 517

To address what characteristics of residue 517 determines whether R366 is required for YidC function, we substituted tyrosine 517 with an apolar, aromatic or polar residue and tested whether

YidC is functional when the positive charge (R366) is removed. Complementation studies revealed that the R366 neutral mutant (R366A or R366N) did not complement the YidC depletion strain when combined with an apolar substitution for Y517 (Y517G, Y517C, Y517A Y517V or Y517I) while very poor complementation was observed for the Y517T and Y517M mutants (Fig. 3A and 3C). In contrast, when the residue at position 517 had a hydrophilic (517N, 517S) or aromatic side chain (517W and 517F), we observed full complementation under YidC depletion conditions (Fig. 3A and 3C). For the 366 residue no difference was observed if there was a slightly hydrophobic alanine or a hydrophilic asparagine residue (data not shown). As a second test of YidC activity, we performed membrane insertion assays using PC-Lep and Pf3-23Lep (Fig. S2) and determined the amount of inserted protein by quantitation (Fig. 3D). Membrane insertion of PC-Lep and Pf3-23Lep was inhibited when the 517 tyrosine was changed to an apolar, non-aromatic amino acid and arginine 366 was changed to either an Ala or Asn residue corroborating the complementation results (Fig. S2).

The essential arginine R73 can be made dispensable by mutation of W234 in S. mutans YidC2

Previously, we found that the strictly conserved arginine R73 in *S. mutans* 247YidC2 (equivalent to R366 in *E. coli*) is essential as the protein was inactive when the R73 residue in the hydrophilic groove was substituted with a neutral or negatively charged residue (24). YidC2₂₄₇ consists of residues 1–247 of the *E. coli* YidC fused to residues 25–310 of *S. mutans* YidC2 containing the conserved arginine at position 73. Since the hydrophobicity at residue 517 can determine whether the R366 is essential or not in the *E. coli* YidC, we went on to test whether mutation of the W234 residue (corresponding to the Y517 residue in *E. coli*) in the *S. mutans* YidC2 could affect the requirement of R73 (corresponding to R366 in *E. coli*). Remarkably, when the W234 residue at the top of the groove of 247YidC2 was substituted with amino acids with hydrophilic side chains (Asn or Ser), YidC2 could complement the YidC depletion strain even without the positively charged residue in the groove (R73Q), while it did not complement with an apolar amino acid at 234 (Val or Ile) (Fig. 3B). As expected, the membrane insertion activities of the *S. mutans* R73Q/W234V and R73Q/W234I YidC2₂₄₇ mutants were inhibited for the PC-Lep substrate, and was completely abolished for the Pf3-Lep substrate (Fig. S2 and Fig. 3D).

Taken together, these results reveal a pattern that applies to both *E. coli* YidC and *S. mutans* YidC2: the conserved positively charged residue is required for activity when there is an apolar residue at position 517 in the *E. coli* YidC or at position 234 in the *S. mutans* YidC2, but not when there is a hydrophilic residue at these sites.

The conserved positively charged residue R366 keeps the E. coli YidC hydrophilic cavity hydrated when Y517 is substituted with an apolar residue

The results reported above show a clear correlation between the requirement of the conserved positive charge and the hydrophobicity near the top of the hydrophilic groove in both *E. coli* YidC and *S. mutans* YidC2. We hypothesize that the positively charged residue is required to maintain a hydrophilic environment in the groove.

To test this hypothesis, we examined the solvent accessibility of nine previously determined groove residues under WT and mutated conditions in the intact bacterial cells at 30°C using a well-developed cysteine alkylation assay (6). Using this technique, each of the groove residues were mutated to single cysteine residues. If the cysteine is in a water-exposed environment, it will be reactive toward N-ethyl maleimide (NEM) because the cysteine thiol is deprotonated. If the cysteine is in a lipid-exposed environment, then the cysteine will not react with NEM because the thiol is protonated. To determine whether the YidC cysteine is modified by NEM, we treated the samples in a later step with Mal-PEG, which leads to a 5 kD shift in the molecular weight of YidC, only if it had not reacted with NEM (lipid exposed). Therefore, a change in the position of YidC on the gel when NEM is added prior to the addition of Mal-PEG indicates exposure to lipid. If there is no shift with Mal-PEG when NEM is added, the Cys is solvent exposed. The WT condition starts with the Cys-less YidC (C423S) (Fig. 4A) and refers to the circumstance where Y517 and R366 are left un-mutated except when their own solvent accessibility is studied (top panels). Under the mutated condition, Y517 is exchanged to an apolar isoleucine residue and R366 is substituted with a neutral asparagine amino acid simultaneously.

Figure 4A shows that the solvent exposure of Y517C is decreased considerably when the R366 is changed to a neutral asparagine. Similarly, R366C is much less water accessible when Y517 is mutated to apolar isoleucine residue. Likewise, the solvent accessibility is decreased for 432C when Y517 and R366 are substituted with an Ile and Asn, respectively. A less dramatic effect on water exposure is observed further down in the groove. With the Y517I and R366N mutations introduced, the solvent exposures of S520C and I428C residues are decreased, but to a lesser extent, compared to the cysteine mutants in the upper region of the groove. A negligible effect was observed when the water accessibility was examined toward the cytoplasmic side of the groove at positions M475, Y370, M498, and T524 when the Y517I and R366N mutations were made. A summary of the results is shown in Fig. 4B, with the top panels representing the WT and the bottom panels the mutated construct. Cyan indicates residues that are completely water exposed, yellow moderate accessibility, and red residues that are not solvent accessible. The top

panel shows that the groove is mostly aqueous under WT condition, while a cluster of red and yellow residues can be observed at the dome of the groove under mutated condition as revealed by the bottom panel, implying that the top part of the groove is not hydrated in these cases.

To confirm these results, we also performed EPR power saturation studies to examine solvent accessibility of a spin label incorporated into the hydrophilic groove. His-tagged YidC mutants listed in Fig. 4C were purified, spin-labeled with (1-Oxyl-2,2,5,5-tetramethyl- Δ 3-pyrroline-3-methyl) Methanethiosulfonate MTSL, reconstituted into DOPC liposomes, and subjected to EPR power saturation experiments at 22°C as described in “Materials and Methods”. The saturation curves for each mutant in the presence of either nitrogen, oxygen or NiEDDA are shown in Fig. 4C, from which the saturation parameters $P_{1/2}$ were obtained and the immersion depth parameters Φ were calculated using equation 1 in “Materials and Methods”.

As a control, a spin label was introduced at position 366 and determined to be solvent accessible as indicated by the highly negative Φ value (-0.8). When Y517 was substituted with a hydrophilic residue (Y517N), position 366 remained accessible (Φ = -0.8). However, the spin label at 366 was no longer water accessible (Φ = 0.9) when the apolar residue isoleucine was introduced at position 517 (Fig. 4C).

A spin label was also attached to the top of the hydrophilic groove (Y517C) and studied with different mutations at R366. Figure 4C shows that in all cases studied (Y517C, R366N/Y517C, R366I/Y517C), position 517 was accessible to solvent. However, a closer investigation of the Φ values suggested that when the positive charge was removed (R366N/Y517C or R366I/Y517C), 517 became much less solvent accessible (-0.2) as compared to when it was present (-0.9), indicating that the positive charge at 366 contributes to keeping the groove hydrated.

Additionally, we carried out a series of comprehensive 200-ns long molecular dynamics (MD) simulations of WT *E. coli* YidC and the Y517I/R366N mutant at ~27°C (300 K) embedded in a palmitoyl-oleoyl-phosphatidylethanolamine:palmitoyl-oleoyl-phosphatidylglycerol (POPE:POPG)(75:25) lipid bilayer. To evaluate the level of hydration at various locations within the groove, the number of water molecules within 6 Å of C α atoms of various groove residues during the last 40 ns was counted (25) and normalized to the standard-state accessible surface areas A^0 of the fully exposed residues of soluble proteins as previously reported (25). Fig. S3 shows that the water accessibility decreases surrounding residues 366, 373, 520, 521 and 527 in the middle and bottom part of the groove for the construct Y517I/R366N compared to the wild-type YidC.

The combined results support our hypothesis that the positive charge at site 366 in *E. coli* is required to maintain a hydrophilic microenvironment in the groove when Y517 is substituted with an apolar residue. We note that 517C appears to be more solvent accessible in EPR studies compared to in alkylation studies, possibly due to the size or polarity of the MTSL spin label.

The YidC hydrophilic groove can function with a negatively charged residue

If the arginine in the groove serves to make the groove sufficiently hydrated, then a negative charge should also function in this capacity. However previous studies showed that the negative charge inactivated YidC (22). To determine if it is possible to obtain a functional *E. coli* YidC with a negative charged residue replacing the arginine at site 366, we searched for suppressor mutations of the inactive YidCR366E mutant. Intriguingly, we did obtain one suppressor mutation, R366E/F433S, after randomly mutagenizing YidC in the XL1 red mutator strain. This secondary mutation allowed the R366E YidC to complement the YidC depletion strain under glucose condition (Fig. 5A). The mutated serine side chain (F433S) is located above site 366 in the hydrophilic groove and may interact with the tyrosine 517 residue (see Fig. 6A). As expected, the R366E/F433S YidC mutant promotes membrane insertion of PC-Lep and of Pf3-23Lep (Fig. 5B and C; Fig. S4). Further, in a reconstituted *in vitro* system, we examined the binding of R366E and R366E/F433S YidC to a fluorescently labeled Pf3 coat protein substrate (Fig. 5D). To fluorescently label YidC, a cysteine was introduced in the cytoplasmic loop C1 at position 405 in the cysteine-less wild-type, R366E and R366E/F433S YidC proteins. The mutants were purified and Atto 647N was attached to residue 405C. The fluorescently labeled YidC proteins were then reconstituted with POPC to generate proteoliposomes. After the purified Pf3 coat protein was labeled at residue 48 with the Atto520 dye in the cytoplasmic tail region, the Pf3 coat protein substrate was added to the proteoliposomes containing the YidC proteins in increasing amounts (see Materials and Methods). The binding efficiency of Pf3 coat protein to the YidC mutants was determined by Förster resonance energy transfer (FRET) with a K_d of 241 ± 75 nM (standard deviation of 10 independent measurements) for the wild-type and $349 \text{ nM} \pm 26 \text{ nM}$ for R366E/F433S (experiments performed at room temperature). No K_d could be determined for the R366E mutant. Clearly, the R366E mutant was defective in the binding of the Pf3 coat protein, whereas wild-type YidC and the double mutant R366E/F433S were capable of efficient substrate binding.

We were able to identify a suppressor mutation (W234M) that rescued the activity of the *S. mutans* YidC2 (R73D) which corresponds to R366 in *E. coli* (22) (Fig. 5A). This *S. mutans* double mutant R73D/W234M had a mutation of the W234 residue corresponding to the *E. coli* Y517 YidC

residue. Interestingly, when we made the corresponding Ser mutant in the *S. mutans* YidC2 R73D background, we found that this R73D/F151S YidC2 mutant (the *S. mutans* F151 residue aligns with F433 in *E. coli* YidC) was also fully functional at room temperature but inactive at 37°C (Fig. 5A). Moreover, a mutant with a substitution of the nearby W234 residue, namely R73D/W234Y, was also found to be temperature-sensitive with little growth at 37°C but functional at room temperature. The temperature-sensitive phenotype of these mutants suggests that the negatively charged residue in the groove destabilizes the protein at higher temperature. Figure 5B and 5C summarizes the membrane insertion results of PC-Lep and Pf3-23Lep (also see Fig. S4) that is generally consistent with the complementation results. To our surprise, the insertion of PC-Lep with *S. mutans* R73D/W234M was inhibited, while Pf3-23Lep was almost completely inserted with the same YidC2 mutant. Possibly this is due to the different insertion mechanisms used by the two substrates.

To explore the molecular mechanism that maintains the functionality of R366E/F433S YidC mutant we performed a 100 ns simulation at 27 °C (Fig. 6A). Remarkably, the F433S suppressor mutation hydrogen bonds to the tyrosine 517 most of the time in the MD simulations (Fig.S6), which may stabilize the groove region. The electrostatic map computed from the simulation trajectory (Fig. 6B) predicts that the interior of the groove of the R366E/F433S mutant is indeed negatively charged (red color), while for the WT YidC, the groove is slightly positively charged (blue color).

We conclude from these results that YidC is functional with a negatively charged residue in the groove when there is a suppressor mutation, indicating that the electrostatic attraction between the arginine in the groove and the negatively charged residues in the substrate region is not required for YidC-mediated insertion.

Discussion

The hydrophilic groove of all YidC family members possess a strictly conserved positively charged residue that has been proposed to interact electrostatically with the hydrophilic peptide chain of the substrate (19). Evidence for this electrostatic attraction model comes from the finding that both the positively charged residue in the Gram-positive *B. subtilis* SpoIIJ and the acidic residues in the N-tail of the substrate MifM are required for insertion (19). In addition, a positively charged residue in the groove is required for the insertase activity of the Gram-positive *S. mutans* YidC2 (24). However, the positive charge of this residue is not essential for the *E. coli* YidC or *A. thaliana* chloroplast Alb3 (24). This suggest that an electrostatic attraction mechanism due to the

conserved arginine residue is not a general process required for substrate insertion into the groove prior to translocation.

Here, we show that the chemical nature of residue 517 (normally a tyrosine residue) of *E. coli* YidC determines whether the strictly conserved arginine is essential or not (Fig. 2). If an apolar residue is present at position 517, the arginine at 366 is essential, but when an aromatic or polar residue is present at 517, the arginine is no longer required (Fig. 3C). Similar results were observed for the Gram-positive *S. mutans* YidC2. By substituting W234 (analogous to *E. coli* Y517) with a more hydrophilic (serine or asparagine) residue, we were able to make the originally essential R73 dispensable (Fig. 3). Note that this *S. mutans* construct has amino acids 1–247 of the *E. coli* YidC fused to amino acids 25–310 of *S. mutans* YidC2. Fusing of *E. coli* YidC region possessing the non-conserved TM1 and large periplasmic region is necessary for correct membrane insertion and efficient activity of the *S. mutans* YidC2 in *E. coli* (24).

A closer look at the position of Y517 in the structure of *E. coli* YidC shows that it is located at the dome of the hydrophilic groove and is part of an aromatic cluster at the interface between the aqueous environment in the groove and the hydrophobic outer leaflet embedded portion of YidC. Interestingly, the water accessibility in the top one third of the groove is affected when Y517 is substituted with an apolar residue, and is significantly decreased when the positive charge at 366 in the middle of the groove is simultaneously removed (Fig. 4A-C). This was determined by examining the water accessibility in the groove by reactivity of introduced cysteine residues to NEM, and by EPR power saturation studies with a cysteine-incorporated spin label in the groove and monitoring whether the residue is lipid exposed or solvent exposed using paramagnetic O₂ and NiEDDA quenchers, respectively. Finally, a decrease in the exposure of water surrounding the groove residues S520, N521, T373 and Q527 was corroborated by examining the number of waters in a fully equilibrated YidC in a lipid bilayer by MD simulations. We cannot rule out that the changes in solvation are caused by small conformational changes in the groove due to the interruption of the aromatic cluster. Alternatively, the water molecules may just fail to reach the top of the groove because of the additional hydrophobicity introduced with the hydrophobic sidechains at 517.

Taken together, our results suggest that the primary purpose of positively charged residue in the groove is to maintain the hydrophilic environment of the groove (Fig. 4), which is crucial for YidC function possibly by accommodating for the polar region of substrates during the process of insertion. This is in line with recent crosslinking results showing that the hydrophilic N-tail of the single-spanning Pf3 coat protein gets transiently incorporated into the hydrophilic groove and the

C-region of N-tail interacts with the groove residues 517, 520, and 524 (23). The importance of hydrophilic microenvironment in the groove was also reported by Shimokawa-Chiba et al. (27) by substituting six polar residues in SpoIIJ with Ala or Leu, both resulting in a defect in insertase activity.

Strikingly, YidC is functional when the strictly conserved arginine is replaced with a negatively charged residue (Fig. 5) with proper stabilization. We have found two suppressor mutations (F433S in *E. coli* YidC and W234M in *S. mutans* YidC2) that allow YidC to be functional with a negatively charged residue in the groove. In addition, *S. mutans* YidC2 variants with an aspartic acid substituting for R73 in combination with W234Y or F151S are fully active at room temperature, although impaired at 37°C. Interestingly, in all cases, the suppressor mutations are part of the aromatic cluster, which we believe stabilizes the negatively charged residue in the YidC groove through indirect interactions. In the case of the R366E/F433S YidC mutant, MD simulations show that F433S forms a hydrogen bond to tyrosine 517 (Fig. 6A; Fig. S6). The combined results show negatively charged residues can function in the groove, suggesting that the electrostatic attraction is not generally required for YidC mediated insertion.

We conclude from our biochemical, biophysical, and simulation studies that the hydrophilic characteristics of the membrane cavity surrounding the 366 residue are crucial to the protein translocation mechanism. This allows the hydrophilic region of the substrate to move into the membrane at least half way within a hydrophilic water containing cavity (25). In addition, another important feature based on previous work is that YidC thins the bilayer in its vicinity (25). Therefore, this would reduce the energy cost of translocation since the distance by which a hydrophilic region would need to cross is reduced. Membrane thinning has been shown to operate in other transport systems such as BamA (28,29), TatA (30) and the Hrd1 complex (for review see (26)). We hypothesize that both features, membrane thinning and the presence of the hydrophilic membrane cavity, are mechanistically coupled in YidC family members to reduce the membrane barrier to promote translocation.

Experimental Procedures

Materials, strains and plasmids

Lysozyme, EDTA, L-(+)-Arabinose, D-(+)-Glucose, and Methoxypolyethylene glycol maleimide (malPEG) were purchased from Sigma-Aldrich. PEG6000 was purchased from Hampton Research. Proteinase K was from Qiagen. Isopropyl 1-thio- β -D-galactopyranoside was from Research Products International Corp. PMSF was purchased from United States Biochemical

(Affymetrix). Phosphate Buffered Saline, pH 7.2 (PBS) and N-Ethylmaleimide (NEM) were purchased from Thermo Scientific. Tran^[35S]-label, a mixture of 85% [^{35S}]-methionine and 15% [^{35S}]-cysteine, 1000 Ci/mmol, was from PerkinElmer Life Sciences. Antiserum to leader peptidase (anti-Lep) was from our own laboratory collection. Anti-6X His tag antibody (HRP) was purchased from Abcam.

JS7131, the YidC depletion strain (from our collection), has its endogenous *yidC* inactivated and has a new copy of the *yidC* gene inserted at the lambda attachment site, under the control of the *araBAD* promoter (6). The low copy number plasmid pACYC184, which encodes chloramphenicol resistance, was used to express mutants of the *E. coli* YidC and the *S. mutans* 247YidC2 (31,32). Expression of these pACYC184-encoded proteins is under control of the native *yidC* promoter in order to ensure close to chromosomal levels of protein expression. The 247YidC2 protein is comprised of residues 1-247 of the *E. coli* YidC fused to residues 25-310 of *S. mutans* YidC2 (31). The high copy number plasmid pMS119 (ampicillin resistant) was used to express Pf3-23Lep (33) or Procoat-Lep [PC-Lep, (34)] under the control of the *lac* promoter. BL21(DE3) strain and pEH1 plasmid were used to express YidC single mutants in the Cys-alkylation assay. Mutator strain XL1-red (#200129) was purchased from Agilent for the randomized mutagenesis studies. All of the mutants of YidC homologs were constructed by site directed mutagenesis using PCR. All of the mutations were verified by DNA sequencing.

Bacterial growth and pulse labeling

E. coli JS7131 was transformed with pMS119 encoding either Pf3-23Lep or PC-Lep and then further transformed with pACYC184 encoding YidC or 247YidC2. The cells were grown in LB media in the presence of ampicillin (final concentration, 100 µg/mL), chloramphenicol (final concentration, 25 µg/mL) and 0.2% arabinose. Depletion of the chromosomally-encoded YidC was achieved by growth in the presence of 0.2% glucose for 4 h at 37°C. The cells were subsequently washed twice with and transferred into M9 medium + 19 amino acids minus Met. After growing at 37°C for another 30 min, expression of YidC-dependent membrane proteins was induced with 1 mM isopropyl-1-thio-D-galactopyranoside (IPTG) for 5 min, followed by pulse-labeling with [^{35S}]-methionine (70 µCi/mL) for 1 min.

Signal peptide processing and protease mapping

To analyze membrane insertion of PC-Lep, a signal peptide cleavage assay was used. After [^{35S}]-labeling at 37°C, an equal volume of 20% TCA was added to precipitate total proteins. The

precipitate was washed with ice-cold acetone and resuspended in Tris-SDS (pH 8.0) buffer. The samples were then subjected to immunoprecipitation with antiserum against Lep to examine membrane insertion by signal peptide processing of PC-Lep to C-Lep.

Protease mapping was employed to assay membrane insertion of Pf3-23Lep. Following radiolabeling, cells were collected by centrifugation at 4°C and resuspended in 33 mM Tris-HCl (pH 8.0) with 40% sucrose. To prepare spheroplasts, lysozyme (5 µg/mL) and 1 mM EDTA (pH 8.0) were added and the sample was kept on ice for 30 min. Where indicated, proteinase K was added (0.5 mg/mL final concentration) and incubated on ice for 1 h, followed by the addition of PMSF to quench the proteinase K reaction. Immunoprecipitation of Pf3-23Lep was performed using Lep antiserum and the samples were analyzed by SDS-PAGE and phosphorimaging.

The translocation efficiencies were determined by quantitation of the appropriate bands using Image J, developed at NIH. The percent of translocation across the membrane was determined as previously described (24). The Pf3-23Lep study was done in triplicates while the PC-Lep were done in duplicates.

Complementation and Western blot assays

E. coli JS7131 cells bearing the respective pACYC184 plasmids were grown overnight at 37°C in LB medium with 0.2% arabinose and 25 µg/mL chloramphenicol. In the following morning the cells were washed once and back-diluted 1:100 into LB medium lacking arabinose. The cells were grown for 3 h at 37°C and subsequently serially diluted (1:10, 1:100, 1:1,000 and 1:10,000). Aliquots of each dilution (4 µL) were spotted on LB plates containing chloramphenicol and 0.2% arabinose or 0.2% glucose, respectively. The plates were incubated overnight at 37°C.

The same overnight culture of JS7131, bearing pACYC184 encoding YidC or 247YidC2, was washed twice with and resuspended in LB containing glucose and chloramphenicol. After 4h of growing at 37°C, 600 µL of cells were precipitated with same volume of 20% TCA on ice for 1 h, washed with ice-cold acetone and solubilized in Tris-SDS (pH 8.0). The samples were analyzed by Western blot with polyclonal antiserum against C-terminal peptides of *E. coli* YidC or *S. mutans* YidC2. YidC antibody was from our lab collection, while the antibody against *S. mutans* YidC2 was a gift from Jeanine Brady (University of Florida).

Cys-alkylation assay

Cys-alkylation assays were performed as previously described (25) and were done in duplicates. An overnight culture of BL21(DE3) bearing pEH1-YidC single-Cys mutant was back-diluted 1:100 into fresh LB medium containing kanamycin (50 µg/ml) and grown to mid-log phase ($OD_{600} \approx 0.55$). 1 mM IPTG was added to induce YidC mutants for 30 min. The cells were subsequently washed twice with PBS and the cell density was adjusted to $OD_{600} \approx 2.0$. Three aliquots (0.2 mL each) were made and the first aliquot was treated with NEM (0.5 mM final concentration) while the second and third were left untreated. These aliquots were then incubated at 30°C for 30 min and subsequently washed twice with PBS buffer (with 5mM DTT) and three times with PBS buffer (without DTT). To remove YidC aggregates, cells were sonicated 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 TCA for 1h on ice. The pellets were washed with ice-cold acetone and solubilized with 100 µl Tris-SDS-Urea buffer (15 mM Tris-HCl, 6 M Urea, 2% SDS, pH 7.5). The first and third aliquots were treated with Mal-PEG (M.W.~5000 Da, 5 mM final concentration) for 40 min at 37°C, while aliquot 2 was mock-treated with 5 mM PEG6000. The protein samples were then analyzed by Western blot using anti-6X His tag antibody (HRP).

Randomized mutagenesis

A pACYC184 plasmid encoding *E. coli* YidC R366E mutant or *S. mutans* YidC2 R73D mutant was transformed into XL1-red and grown on LB plate with chloramphenicol overnight at 37°C. The colonies were pooled and plasmids were isolated. The mixture of plasmids was then transformed into JS7131 and screened for active mutants on LB plates containing chloramphenicol and 0.2% glucose. The plates were incubated at 37°C for two days. Plasmid DNA was extracted from picked colonies and analyzed by sequencing.

Overexpression and Purification of YidC

The overexpression and purification of pEH1 expressing YidC was carried out in *E. coli* BL21 cells. Individual colonies were picked and placed in 5 mL cultures for overnight growth in 50 µg/ml Kanamycin antibiotic. Overnight culture was used to inoculate 1L LB culture and was grown at 37°C until it reached $OD_{600} \approx 0.6$, at which point protein expression was induced using 1 mM IPTG. Cells were harvested 3 h post induction by centrifugation (at 3200 g for 20 min, 4°C) and stored at -80°C.

Cell pellets were thawed on ice and re-suspended in PBS buffer pH 7.2 containing Lysozyme (1 mg/ml). The cells were sonicated on ice for several cycles at an output of 65% power to break open the membranes. The samples were centrifuged at 40,000 g for 50 min at 4°C to remove unlysed cells and inclusion bodies. To isolate the membrane pellets, the supernatant was then subjected to ultracentrifugation at 160,000 g for 50 min at 4°C. Membrane pellets were then solubilized in PBS containing 1% DDM (Anatrace) overnight by stirring with magnetic beads. The sample was again centrifuged for 25 min to remove all non-solubilized components and the supernatant was incubated with the Co2+-NTA matrix (Qiagen) for 3 h at 4°C. After washing with low imidazole buffer [PBS (pH 7.2), 20 mM imidazole, 10% glycerol and 0.2% (w/v) DDM], 3 mg of the spin label MTSL was added to the YidC sample on the column and shaken for 24 h at 4°C. YidC was eluted from the column with high imidazole buffer [PBS (pH 7.2), 400 mM imidazole containing 0.2% (w/v) DDM] and collected in 1 mL fractions. The samples were analyzed by 15% SDS-PAGE and the fractions containing pure YidC were pooled and dialyzed with buffer [PBS pH 7.2, 0.02% (w/v) DDM] overnight to remove glycerol, imidazole and any free spin label. The YidC protein was concentrated by centrifugation using a 50 kDa Amicon spin concentrator.

Reconstitution of YidC into Proteoliposomes

1,2-Dioleoyl-sn-glycero-3-phosphocholine (DOPC) was purchased from Avanti Polar Lipids (Alabaster, AL) and lipid vesicles prepared. The dry lipid film was resuspended in HEPES buffer (pH 8.0) containing 50 mM KCl, 0.02% (w/v) DDM and incubated at 37°C for 1 h. The sample was vortexed at a low speed to form multilamellar vesicles. Unilamellar vesicles were generated by the extrusion technique (Mini-Extruder, Avanti Polar Lipids Inc). Specifically, 1 mL of the lipid suspension was extruded 7-11 times through a membrane with a pore size of 0.4 µm until a semi-clear solution is achieved. For preparing proteoliposomes, a concentrated spin labeled YidC sample (at a final concentration of 20 µM) was added to the lipid sample and extruded as described above. The final protein:lipid molar ratio was set to 1:400. SM2-Biobeads were prepared according to the manufacturers protocol and added to the proteoliposomes to remove excess DDM as described in Kusters et al. (35) The removal of DDM in the lipid sample with the Biobeads aids in further reconstitution and is achieved by slow rotary shaking of the sample overnight at 4°C. The Biobeads are removed by centrifugation at 10,000 rpm for 5 min. The proteoliposomes are concentrated by centrifugation at 200,000 g to form a pellet and resuspended in 15 µl of HEPES buffer for EPR measurements.

Simulated Systems and Simulation Parameters

A system with an equilibrated (~185 ns), crystallography-based model of the WT YidC protein embedded in a lipid bilayer (75% POPE and 25% POPG) (25) was used as a starting conformation for the WT simulation. The same system was used along with the psfgen VMD (36) plugin to create five mutated versions of the equilibrated YidC and to neutralize the corresponding systems (Table 1). MD simulations using periodic boundary conditions were performed using NAMD 2.12 (37), the CHARMM36 force field for proteins and lipids with the CMAP correction and the TIP3P model for water (38-40). A cutoff of 12 Å with a force-based switching function starting at 10 Å was used for van der Waals interactions. A 2 fs integration time step was used together with SHAKE. The Particle Mesh Ewald method was used to compute long-range electrostatic forces every other time step without cutoff and with a grid point density of $>1 \text{ Å}^{-3}$. Langevin dynamics was utilized to enforce constant temperature $T = 300 \text{ K}$ with a damping coefficient of 0.1 ps^{-1} . Constant pressure simulations (NpT) at 1 atm were conducted using the hybrid Nosé-Hoover Langevin piston method with a 200 fs decay period and a 100 fs damping time constant. Each system was energy-minimized and equilibrated in the NpT ensemble (Table 1). Harmonic constraints ($k_s = 1 \text{ kcal mol}^{-1} \text{ Å}^{-2}$) to backbone atoms were applied when indicated. Coordinates of all atoms in the system were saved every 2 ps. Computation of the electrostatic potential was carried out using the VMD PMEpot plugin with an Ewald factor of 0.25 Å^{-1} using conformations saved every 80 ps for simulations in which YidC's backbone was constrained. Molecular images were created with the molecular graphics program VMD (36).

CW Power Saturation Experiments

Power saturation experiments were performed on a Bruker EMX X-band CW-EPR spectrometer consisting of an ER 041XG microwave bridge coupled with an ER 4123D CW-Resonator (Bruker BioSpin)(41). Samples were loaded into gas permeable TPX capillary tubes with a total volume of 4-6 μL at a spin label concentration of 40–80 μM . EPR data collection was carried out using a modulation amplitude of 1 G and a varying microwave power of 0.4–100 mW. The scan range of all spectra was 90 G, and the final spectra were obtained by signal averaging 10 scans.

CW-EPR power saturation curves were obtained for all the spin labeled YidC Cys mutants under three conditions: (1) equilibrated with nitrogen as a control; (2) equilibrated with a lipid-soluble paramagnetic reagent: air (20% oxygen); and (3) equilibrated with nitrogen in the presence of a water-soluble paramagnetic reagent nickel(II) ethylenediaminediacetate (NiEDDA) chelate (2 mM), as described (41). The samples were purged with gas for at least 60 min at a rate of 10

mL/min before performing each EPR measurement. High purity nitrogen and house supply compressed air lines were used. The resonator remained connected to the gas line during all measurements, and the sample temperature was held at 295 K. The peak-to-peak amplitude (A) of the first derivative $m_I = 0$ resonance line was measured and plotted against the square root of the incident microwave power. The data points were then fit using a Matlab software script using equation 1:

Eq 1:
$$A = I \sqrt{P} \left[1 + (2^{1/\varepsilon} - 1) P / P_{1/2} \right]^{-\varepsilon}$$

where I is a scaling factor, $P_{1/2}$ is the power where the first derivative amplitude is reduced to half of its unsaturated value, and ε is a measure of the homogeneity of saturation of the resonance line. In the above equation, A , I , ε , and P are adjustable parameters and yield a characteristic $P_{1/2}$ value. The corresponding Φ depth parameters were calculated using the following equation:

Eq. 2:
$$\Phi = \ln \left(\frac{\Delta P_{1/2}(\text{O}_2)}{\Delta P_{1/2}(\text{NiEDDA})} \right)$$

where $\Delta P_{1/2}(\text{O}_2)$ is the difference in the $P_{1/2}$ values for air and nitrogen-exposed samples, and $\Delta P_{1/2}(\text{Ni-EDDA})$ is the difference in the $P_{1/2}$ values for Ni-EDDA and nitrogen-exposed samples (41).

Substrate binding to YidC by a Förster resonance energy transfer (FRET) assay

A cysteine was first incorporated at position 405 of a cysteine-less YidC at the cytoplasmic face of the membrane into the wild-type and the mutants 366E and 366E/433S. The 405C residues of the YidC mutants were used to attach a fluorescent label with the maleimide dye Atto520 (donor). Residue 48 of the substrate Pf3 was labeled with Atto 647N (acceptor). For the labeling reaction, 1.3 fold excess of the dye was added to the protein and incubated for 1.5 h at room temperature. The free dye was separated from the protein by size exclusion chromatography with a Superdex

200 10/30 column and the YidC proteins were reconstituted in DOPC liposomes. The fluorescence measurements to examine binding were performed on a Fluorolog 2.3 (Jobin Yvon-Spex). The extinction wavelength was set at 500 nm to excite the donor dye (Atto 520) and the emission wavelength was scanned from 510 to 700 nm. The fluorescence signal of the acceptor was monitored at 669 nm. The Atto647N-Pf3 coat protein was added in steps, each after 1 min, to the proteoliposomes containing 300 nM Atto520-YidC. The concentrations in each of the steps were 60, 60, 150, 150, 150, 300, and 300 nM, reaching a final concentration of 1.17 μ M. The measurement started after an incubation period of 1 min after the addition of the Pf3 coat substrate. For our analysis of the binding of Pf3 coat to YidC we plot the $\Delta F_{\text{Acceptor}}$ signal:

$$\text{Eq. 3} \quad \Delta F_{\text{Acceptor}} = |F_{A_i} - F_{A_0}|$$

F_{A_0} is the initial fluorescence values at 669 nm of YidC without labeled Pf3 and F_{A_i} is the value of the fluorescence signal at the i-th titration step of labeled Pf3.

The measurements were background-corrected by a buffer and labeled Pf3 only spectrum. The values of the fluorescence signal were corrected by a titration factor of the added Pf3.

For the calculation of the K_D the binding curve was fitted against the following equation:

$$\text{Eq. 4} \quad \Delta F = \Delta F_{\text{max}} \frac{[Pf3]}{[Pf3] + K_D}$$

$$\text{Eq. 5} \quad K_D = \frac{\Delta F_{\text{max}} * [Pf3]}{\Delta F} - [Pf3]$$

where ΔF_{max} is the maximum fluorescence signal change reached at saturation.

Data availability

All data generated for this study are included within this manuscript.

Supporting Information

This article contains supporting information

Funding and additional information

This work was supported by grants from National Science Foundation Grant MCB-1052033 to RED and from the Deutsche Forschungsgemeinschaft Ku 749/6 to AK. Simulations were performed at the Ohio Supercomputer Center (PAS1037 to M.S.). Gary A. Lorigan would like to acknowledge NSF CHE-1807131 grant, NSF (MRI-1725502) grant, NIGMS/NIH Maximizing Investigator's Research Award (MIRA) R35 GM126935 award the Ohio Board of Regents, and Miami University. Gary A. Lorigan would also like to acknowledge support from the John W. Steube Professorship. Indra D. Sahu would also like to acknowledge support from the NSF MCB-2040917 award. Stephen White is pleased to acknowledge the support of NIH grants RO1 GM-74637 and R35 GM-139652.

References

1. Dalbey, R. E., and Kuhn, A. (2014) How YidC inserts and folds proteins across a membrane. *Nat Struct Mol Biol* **21**, 435-436. doi: 10.1038/nsmb.2823
2. Hennon, S. W., Soman, R., Zhu, L., and Dalbey, R. E. (2015) YidC/Alb3/Oxa1 Family of Insertases. *J Biol Chem* **290**, 14866-14874. doi: 10.1074/jbc.R115.638171
3. Chen, Y., and Dalbey, R. E. (2018) Oxa1 Superfamily: New Members Found in the ER. *Trends Biochem Sci* **43**, 151-153. DOI: [10.1016/j.tibs.2017.12.005](https://doi.org/10.1016/j.tibs.2017.12.005)
4. Dalbey, R. E., and Kuhn, A. (2015) Membrane Insertases Are Present in All Three Domains of Life. *Structure* **23**, 1559-1560. doi: 10.1016/j.str.2015.08.002
5. McDowell, M. A., Heimes, M., and Sinning, I. (2021) Structural and molecular mechanisms for membrane protein biogenesis by the Oxa1 superfamily. *Nat Struct Mol Biol* **28**, 234-239. doi: 10.1038/s41594-021-00567-9
6. Samuelson, J. C., Chen, M., Jiang, F., Moller, I., Wiedmann, M., Kuhn, A., Phillips, G. J., and Dalbey, R. E. (2000) YidC mediates membrane protein insertion in bacteria. *Nature* **406**, 637-641. doi: 10.1038/35020586
7. 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. doi: 10.1074/jbc.M110644200
8. Van Der Laan, M., Bechtluft, P., Kol, S., Nouwen, N., and Driessen, A. J. (2004) F1F0 ATP synthase subunit c is a substrate of the novel YidC pathway for membrane protein biogenesis. *The Journal of cell biology* **165**, 213-222. doi: 10.1083/jcb.200402100
9. Samuelson, J. C., Jiang, F., Yi, L., Chen, M., and de Gier, J. W. (2001) Function of YidC for the insertion of M13 procoat protein in E. coli: translocation of mutants that show differences in their membrane potential dependence and Sec requirement. *J. Biol. Chem.* **276**, 34847-34852. doi: 10.1074/jbc.M105793200
10. Serek, J., Bauer-Manz, G., Struhalla, G., van den Berg, L., and Kiefer, D. (2004) Escherichia coli YidC is a membrane insertase for Sec-independent proteins. *EMBO J.* **23**, 294-301. doi: 10.1038/sj.emboj.7600063
11. van Bloois, E., Jan Haan, G., de Gier, J. W., Oudega, B., and Luirink, J. (2004) F(1)F(0) ATP synthase subunit c is targeted by the SRP to YidC in the E. coli inner membrane. *FEBS Lett* **576**, 97-100. doi: 10.1016/j.febslet.2004.08.069

12. Yi, L., Jiang, F., Chen, M., Cain, B., Bolhuis, A., and Dalbey, R. E. (2003) YidC is strictly required for membrane insertion of subunits a and c of the F(1)F(0)ATP synthase and SecE of the SecYEG translocase. *Biochemistry* **42**, 10537-10544. doi: 10.1021/bi034309h
13. 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. DOI: [10.1016/j.jmb.2006.01.030](https://doi.org/10.1016/j.jmb.2006.01.030)
14. Zhu, L., Klenner, C., Kuhn, A., and Dalbey, R. Both YidC and SecYEG Are Required for Translocation of the Periplasmic Loops 1 and 2 of TatC, a 6-Membrane-Spanning Protein. *J Mol Biol* **424**, 354-367. doi: 10.1016/j.jmb.2012.09.026 Epub 2012 Oct 8
15. Sachelaru, I., Winter, L., Knyazev, D. G., Zimmermann, M., Vogt, A., Kuttner, R., Ollinger, N., Siligan, C., Pohl, P., and Koch, H. G. (2017) YidC and SecYEG form a heterotetrameric protein translocation channel. *Sci Rep* **7**, 101. doi: 10.1038/s41598-017-00109-8
16. Zhu, L., Kaback, H. R., and Dalbey, R. E. (2013) YidC protein, a molecular chaperone for LacY protein folding via the SecYEG protein machinery. *J Biol Chem* **288**, 28180-28194. doi: 10.1074/jbc.M113.491613.
17. Wagner, S., Pop, O., Haan, G. J., Baars, L., Koningstein, G., Klepsch, M. M., Genevaux, P., Luirink, J., and de Gier, J. W. (2008) Biogenesis of MalF and the MalFGK(2) maltose transport complex in Escherichia coli requires YidC. *J Biol Chem* **283**, 17881-17890. doi: 10.1074/jbc.M801481200
18. Kumazaki, K., Chiba, S., Takemoto, M., Furukawa, A., Nishiyama, K., Sugano, Y., Mori, T., Dohmae, N., Hirata, K., Nakada-Nakura, Y., Maturana, A. D., Tanaka, Y., Mori, H., Sugita, Y., Arisaka, F., Ito, K., Ishitani, R., Tsukazaki, T., and Nureki, O. (2014) Structural basis of Sec-independent membrane protein insertion by YidC. *Nature* **509**, 516-520. DOI: [10.1038/nature13167](https://doi.org/10.1038/nature13167)
19. Kumazaki, K., Kishimoto, T., Furukawa, A., Mori, H., Tanaka, Y., Dohmae, N., Ishitani, R., Tsukazaki, T., and Nureki, O. (2014) Crystal structure of Escherichia coli YidC, a membrane protein chaperone and insertase. *Scientific reports* **4**, 7299. DOI: [10.1038/srep07299](https://doi.org/10.1038/srep07299)
20. Laskowski, P. R., Pluhackova, K., Haase, M., Lang, B. M., Nagler, G., Kuhn, A., and Muller, D. J. (2021) Monitoring the binding and insertion of a single transmembrane protein by an insertase. *Nat Commun* **12**, 7082. DOI: [10.1038/s41467-021-27315-3](https://doi.org/10.1038/s41467-021-27315-3)
21. Petriman, N. A., Jauss, B., Hufnagel, A., Franz, L., Sachelaru, I., Drepper, F., Warscheid, B., and Koch, H. G. (2018) The interaction network of the YidC insertase with the SecYEG translocon, SRP and the SRP receptor FtsY. *Sci Rep* **8**, 578. doi: 10.1038/s41598-017-19019-w
22. Klenner, C., and Kuhn, A. (2012) Dynamic disulfide scanning of the membrane-inserting Pf3 coat protein reveals multiple YidC substrate contacts. *The Journal of biological chemistry* **287**, 3769-3776. DOI: [10.1074/jbc.M111.307223](https://doi.org/10.1074/jbc.M111.307223)

23. He, H., Kuhn, A., and Dalbey, R. E. (2020) Tracking the Stepwise Movement of a Membrane-inserting Protein In Vivo. *J Mol Biol* **432**, 484-496. DOI: [10.1016/j.jmb.2019.10.010](https://doi.org/10.1016/j.jmb.2019.10.010)
24. Chen, Y., Soman, R., Shanmugam, S. 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. DOI: [10.1074/jbc.M114.595082](https://doi.org/10.1074/jbc.M114.595082)
25. Chen, Y., Capponi, S., Zhu, L., Gellenbeck, P., Freitas, J. A., White, S. H., and Dalbey, R. E. (2017) YidC Insertase of Escherichia coli: Water Accessibility and Membrane Shaping. *Structure* **25**, 1403-1414 e1403. doi: 10.1016/j.str.2017.07.008
26. Wu, X., and Rapoport, T. A. (2021) Translocation of Proteins through a Distorted Lipid Bilayer. *Trends Cell Biol* **31**, 473-484. DOI: [10.1016/j.tcb.2021.01.002](https://doi.org/10.1016/j.tcb.2021.01.002)
27. 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 U S A* **112**, 5063-5068. DOI: [10.1073/pnas.1423817112](https://doi.org/10.1073/pnas.1423817112)
28. Noinaj, N., Kuszak, A. J., Gumbart, J. C., Lukacik, P., Chang, H., Easley, N. C., Lithgow, T., and Buchanan, S. K. (2013) Structural insight into the biogenesis of beta-barrel membrane proteins. *Nature* **501**, 385-390. doi: 10.1038/nature12521
29. Lundquist, K., Bakelar, J., Noinaj, N., and Gumbart, J. C. (2018) C-terminal kink formation is required for lateral gating in BamA. *Proc Natl Acad Sci U S A* **115**, E7942-E7949. DOI: [10.1073/pnas.1722530115](https://doi.org/10.1073/pnas.1722530115)
30. Rodriguez, F., Rouse, S. L., Tait, C. E., Harmer, J., De Riso, A., Timmel, C. R., Sansom, M. S., Berks, B. C., and Schnell, J. R. (2013) Structural model for the protein-translocating element of the twin-arginine transport system. *Proc Natl Acad Sci U S A* **110**, E1092-1101. DOI: [10.1073/pnas.1219486110](https://doi.org/10.1073/pnas.1219486110)
31. Dong, Y., Palmer, S. R., Hasona, A., Nagamori, S., Kaback, H. R., Dalbey, R. E., and Brady, L. J. (2008) Functional overlap but lack of complete cross-complementation of Streptococcus mutans and Escherichia coli YidC orthologs. *J Bacteriol* **190**, 2458-2469. • DOI: [10.1128/JB.01366-07](https://doi.org/10.1128/JB.01366-07)
32. Jiang, F., Yi, L., Moore, M., Chen, M., and Rohl, T. (2002) Chloroplast YidC homolog Albino3 can functionally complement the bacterial YidC depletion strain and promote membrane insertion of both bacterial and chloroplast thylakoid proteins. *J. Biol. Chem.* **277**, 19281-19288. doi: 10.1074/jbc.M110857200
33. Zhu, L., Wasey, A., White, S. H., and Dalbey, R. E. (2013) Charge composition features of model single-span membrane proteins that determine selection of YidC and SecYEG translocase pathways in Escherichia coli. *J Biol Chem* **288**, 7704-7716. doi: 10.1074/jbc.M112.429431

34. Soman, R., Yuan, J., Kuhn, A., and Dalbey, R. E. (2014) Polarity and charge of the periplasmic loop determine the YidC and sec translocase requirement for the M13 procoat lep protein. *The Journal of biological chemistry* **289**, 1023-1032. DOI: [10.1074/jbc.M113.522250](https://doi.org/10.1074/jbc.M113.522250)
35. Kusters, I., van den Bogaart, G., de Wit, J., Krasnikov, V., Poolman, B., and Driessen, A. (2010) Purification and functional reconstitution of the bacterial protein translocation pore, the SecYEG complex. *Methods Mol Biol* **619**, 131-143. doi: 10.1007/978-1-60327-412-8_8
36. Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: visual molecular dynamics. *Journal of molecular graphics* **14**, 33-38, 27-38. doi: 10.1016/0263-7855(96)00018-5
37. Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kale, L., and Schulten, K. (2005) Scalable molecular dynamics with NAMD. *Journal of computational chemistry* **26**, 1781-1802. doi: 10.1002/jcc.20289
38. Buck, M., Bouguet-Bonnet, S., Pastor, R. W., and MacKerell, A. D., Jr. (2006) Importance of the CMAP correction to the CHARMM22 protein force field: dynamics of hen lysozyme. *Biophysical journal* **90**, L36-38. DOI: [10.1529/biophysj.105.078154](https://doi.org/10.1529/biophysj.105.078154)
39. 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. *The journal of physical chemistry. B* **114**, 7830-7843. DOI: [10.1021/jp101759q](https://doi.org/10.1021/jp101759q)
40. Huang, J., and MacKerell, A. D., Jr. (2013) CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. *Journal of computational chemistry* **34**, 2135-2145. DOI: [10.1002/jcc.23354](https://doi.org/10.1002/jcc.23354)
41. Coey, A. T., Sahu, I. D., Gunasekera, T. S., Troxel, K. R., Hawn, J. M., Swartz, M. S., Wickenheiser, M. R., Reid, R. J., Welch, R. C., Vanoye, C. G., Kang, C., Sanders, C. R., and Lorigan, G. A. (2011) Reconstitution of KCNE1 into lipid bilayers: comparing the structural, dynamic, and activity differences in micelle and vesicle environments. *Biochemistry* **50**, 10851-10859. DOI: [10.1021/bi2009294](https://doi.org/10.1021/bi2009294)
42. Moon, C. P., and Fleming, K. G. (2011) Side-chain hydrophobicity scale derived from transmembrane protein folding into lipid bilayers. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 10174-10177. DOI: [10.1073/pnas.1103979108](https://doi.org/10.1073/pnas.1103979108)

Figure Legends

Figure 1. Structure of the *E. coli* YidC and domain architecture of substrates. **A.** Structure of the *E. coli* YidC depicting its transmembrane and cytoplasmic helices. **B.** A closeup view of the hydrophilic groove (PDB 3WVF) (25). Residues studied in this paper are indicated, including the conserved positively charged R366, as well as the tyrosine and groove residues. Side chains are shown as stick for R366 and the aromatic residues mutated to study the requirement of the conserved arginine, while C α atoms are shown as spheres for groove residues studied in Cys-alkylation assay. **C.** YidC substrates Pf3-23Lep and PC-Lep used in this study (see Material and Methods for details) and their membrane topologies. The red arrow in PC-Lep depicts the cleavage site for signal peptidase.

Figure 2. The conserved R366 residue in *E. coli* YidC becomes essential in the presence of the Y517A mutation. **A.** Complementation assay to test the effect of mutating R366 and Y517 on *E. coli* YidC function. pACYC184 encoding R366A, Y517A and R366A/Y517A YidC mutants were constructed using site-directed mutagenesis and transformed into YidC depletion strain JS7131. After depletion of endogenous YidC for 3 h, serial dilutions of the cells were spotted and incubated under YidC expression conditions (left panels) and depletion conditions (right panels), respectively (37°C). Plasmid encoding WT YidC and empty plasmid were also expressed, as positive and negative controls. **B.** Same *E. coli* YidC mutants were tested for their ability to insert PC-Lep (left panel) and Pf3-23Lep (right panel). JS7131 harboring pACYC184 encoding WT YidC, empty plasmid, or YidC mutants were co-transformed with pMS119 encoding PC-Lep or Pf3-23Lep. Cells were grown under YidC expression conditions (ara+) or depletion conditions (ara-) for 4 h (37°C). Expression of substrates was induced by 1 mM IPTG for 5 min and labeled with [³⁵S]-methionine for 1 min. Signal peptide processing assay was used to assess the insertion of PC-Lep (left panel), and the insertion of Pf3-23Lep was examined by protease mapping (right panel), both as described in “Materials and Methods”. Red numbers indicate molecular weight marker values in kD.

Figure 3. The polarity at the dome of the YidC hydrophilic groove determines the necessity of the conserved positive charge. **A.** Complementation assay to examine the importance of R366 (by mutating R366 to Ala or Asn) for *E. coli* YidC when Y517 was substituted with Asn, Ser, Gly, Thr, Ala, Cys, Val, Ile, Met, Trp or Phe. YidC depletion strain JS7131 was transformed with pACYC184 encoding these double mutants and a spot test at 37°C was performed to test complementation, as described in Fig. 2A. Note the data for R366A/Y517A (see Fig. 2A) is

included for comparison. **B.** Complementation assay to examine the role of R73 (by substituting R73 with Q) of *S. mutans* YidC2 when W234 was mutated to Asn, Ser, Val or Ile. Note that in *S. mutans* YidC2, W234 aligns with Y517 in *E. coli* YidC, while R73 is equivalent to R366. **C.** Summary of complementation results for the *E. coli* YidC 517 mutants. The hydrophobicity panel shows the standard free energy in kcal/mol for each amino acid tested (42). The “+” indicates that the mutant complemented the YidC depletion strain, indicating the arginine is not essential. The “-” means it did not complement showing the arginine is required for function. **D.** *E. coli* YidC and *S. mutans* YidC2 mutants were tested for their ability to insert PC-Lep (blue bars) and Pf3-23Lep (orange bars). Plasmids pACYC184 encoding YidC or YidC2 mutants were co-transformed with pMS119 encoding PC-Lep or Pf3-23Lep into JS7131. After expression of YidC substrates and labeling, the membrane insertion of PC-Lep and Pf3-23Lep was tested as described in Fig. 2B. The results (one representative trial shown in Fig. S2) were quantified as previously described (24) and summarized in panel D.

Figure 4. The conserved positive charge R366 in *E. coli* keeps the top part of the groove hydrated when Y517 is substituted with an apolar residue. **A.** Cys-alkylation assay at 30°C to examine the solvent accessibility of 9 previously determined groove residues under WT conditions (lanes 1-3 in each panel) or mutated conditions (lanes 4-6). BL21 cells expressing the single Cys mutants were analyzed for modification by NEM using an indirect gel shift assay involving Mal-PEG, as described in “Materials and Methods”. Proteins were analyzed by SDS-PAGE and YidC was detected by Western blotting using anti-6X His tag antibody (HRP). WT condition refers to the condition when Y517 and R366 were maintained, except for the cases when 517C and 366C were tested. Mutated condition refers to the condition when Y517 was substituted with an apolar residue (Ile or Cys) and R366 was mutated to a noncharged residue (Asn or Cys). Red numbers indicate molecular weight marker values in kD. **B.** The solvent accessibility is summarized for the 9 groove residues studied in panel A as previously described (25). Cyan means completely accessible to water; yellow represents partially solvent accessible; red codes for complete inaccessible to water. Side and cytoplasmic views are shown. **C.** EPR power saturation assay at 22°C (295K) to determine the water accessibility of residues 366 and 517 under WT conditions (R366C and Y517C) or mutated conditions (Y517N/R366C, Y517I/R366C, Y517C/R366N and Y517C/R366I). The YidC mutants were overexpressed from BL21 cells, purified, spin-labeled with MTSL, reconstituted into DOPC liposome, and subjected to CW-EPR power saturation experiment as described in “Materials and Methods”. Water-soluble NiEDDA or insoluble air was applied as paramagnetic quencher, while diamagnetic nitrogen was

used as a control. $P_{1/2}$ values (the saturation parameter) were obtained from the saturation curves and Φ values (the immersion depth parameter) were calculated using equation 1 in “Materials and Methods”.

Figure 5. YidC can function with a negatively charged residue substituting for the conserved positive charge in the groove. **A.** Complementation assay to examine the activity of *E. coli* YidC and *S. mutans* YidC2 mutants with a negative charge in the groove. The suppressor mutants F433S for *E. coli* YidC and W234M for *S. mutans* YidC2 were identified using mutator strain XL1-Red as described in “Materials and Methods”. F151 in *S. mutans* YidC2 aligns with *E. coli* YidC F433. All spot tests were performed at 37°C unless indicated (RT). **B. C.** The insertase activity (results from one representative trial shown in Fig. S4) of *E. coli* YidC and *S. mutans* YidC2 negatively charged mutants were tested for insertion of PC-Lep (B. blue bar) and Pf3-23Lep (C. orange bar) and performed as described in Fig. 3D. JS7131 used in all insertion assays were grown at 37°C unless specifically indicated (RT). **D.** Pf3 coat protein binding to *E. coli* YidC assayed by Förster resonance energy transfer (FRET). The Pf3 coat protein was labeled with the acceptor dye Atto647N at position 48 in the C-terminal tail. After a cysteine was introduced at position 405, the YidC mutants were expressed, purified and labeled with the donor dye Atto520 at 405C in the cytoplasmic loop C1 and reconstituted in DOPC liposomes. The starting concentration of *E. coli* YidC was 0.3 μ M and the Pf3 coat was added in increasing amounts in 1 min steps. The fluorescence signal at the acceptor signal ($\Delta F_{\text{acceptor}}$) was plotted against the substrate concentration. YidC (black), YidC R366E (red) and the double mutant YidC R366E/F433S (blue) were analyzed.

Figure 6. The *E. coli* YidC R366E/F433S protein with a negative charge in its groove is functional. **A.** Snapshot of the equilibrated *E. coli* YidC R366E/F433S system simulated in a hydrated POPE:POPG (75:25) bilayer (phospholipids and water are not shown; simulation S3b at ~28°C in Table 1). An interaction between the side chain of F433S (suppressor mutation) and the tyrosine residue at site 517 is observed during the simulation trajectory (Fig. S6). This interaction might stabilize YidC’s groove. **B,C.** Two-dimensional contour plot of the averaged electrostatic potential computed during 100-ns long simulations of *E. coli* YidC at ~28°C with the protein backbone constrained (WT and R366E/F433S systems in **B** and **C**, respectively; simulations S1a and S3a in Table 1). The map corresponds to the electrostatic potential of a slice perpendicular to the membrane plane passing through the center of the protein near site 366.

Twenty contour lines are drawn over the range of voltages. All values above or below the scale limits (color scale) are shown at the same level. The electrostatic potential for the R366E/F433S system is negative near the 366 site (**C**).

Table 1. Summary of Simulations

Label	System	t_{sim} (ns)	Type	Start	Constraints	Size (#atoms)	Initial Size (nm ³)
S1a	WT	100	EQ ^a	–	backbone	197,865	11.3 × 12.2 × 14.1
S1b		103	NpT	S1a	–		11.8 × 12.3 × 13.3
S2a	R366E	100	EQ ^a	–	backbone	197,852	11.3 × 12.2 × 14.1
S2b		109	NpT	S2a	–		11.2 × 12.6 × 13.6
S2c ^b		100	EQ ^a	S2a	backbone		11.2 × 12.6 × 13.6
S2d		82	NpT	S2c	–		11.3 × 12.8 × 13.4
S3a	R366E/F433 S	100	EQ ^a	–	backbone	197,843	11.3 × 12.2 × 14.1
S3b		104	NpT	S3a	–		12.4 × 11.5 × 13.4
S4a	R366N/Y517 I	100	EQ ^a	–	backbone	197,851	11.3 × 12.2 × 14.1
S4b		104	NpT	S4a	–		11.0 × 13.2 × 13.3
Total		1,002					

^a EQ indicates simulations that consisted of 1,000 steps of minimization, and 100 ns of dynamics with the protein backbone constrained ($k = 1$ kcal/mol/Å²) in the *NpT* ensemble ($\gamma = 0.1$ ps⁻¹).

^b A potassium ion found inside YidC was exchanged with a bulk water molecule before starting simulation S2a.

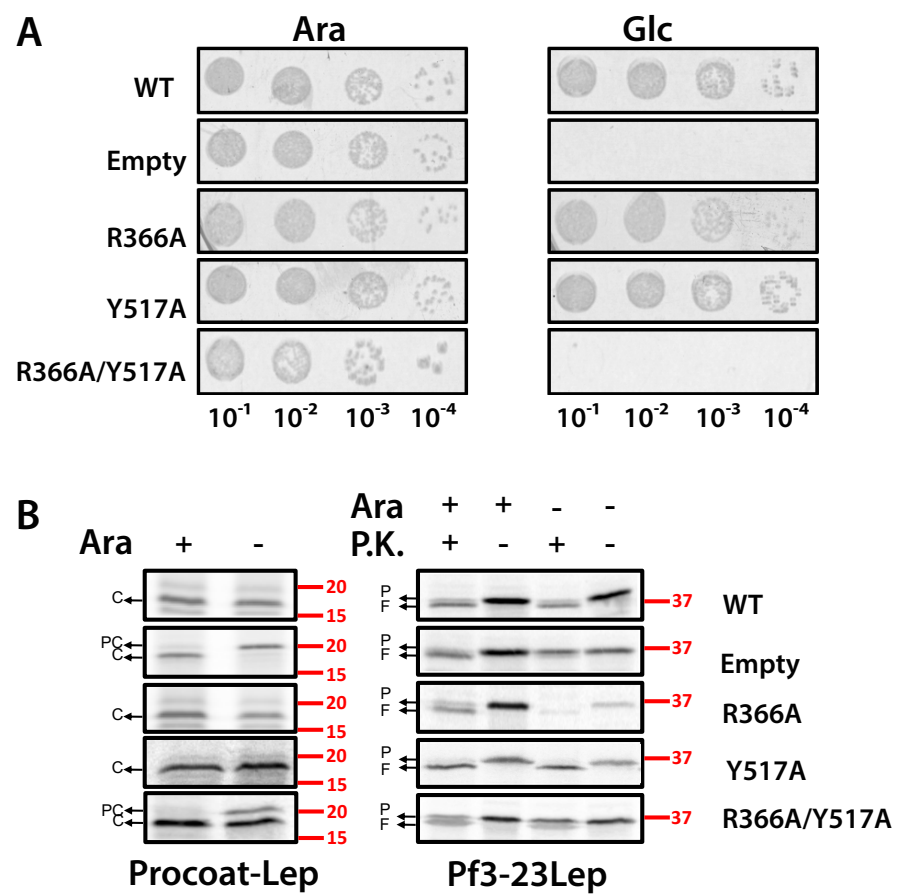


Figure 2

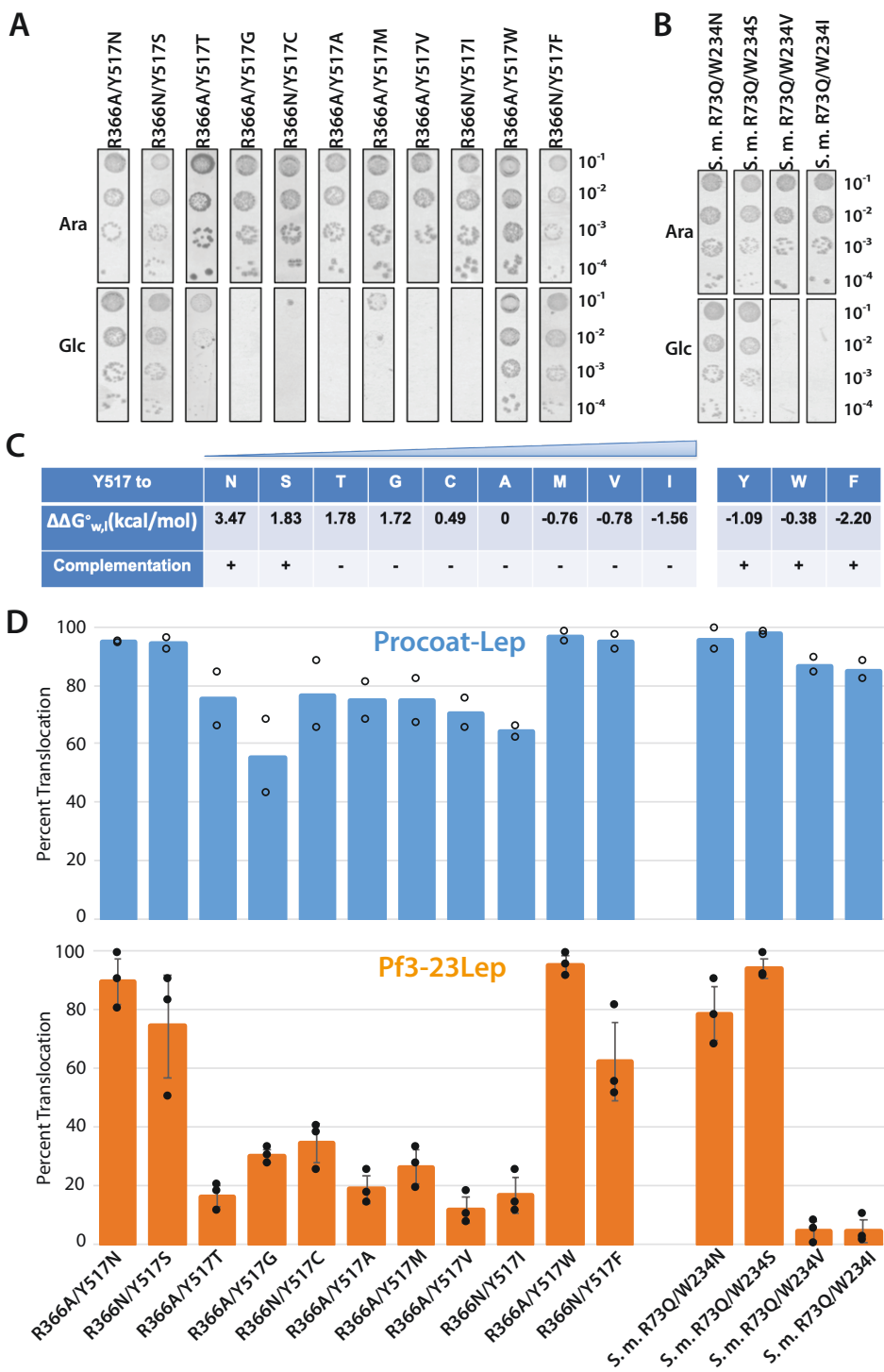


Figure 3

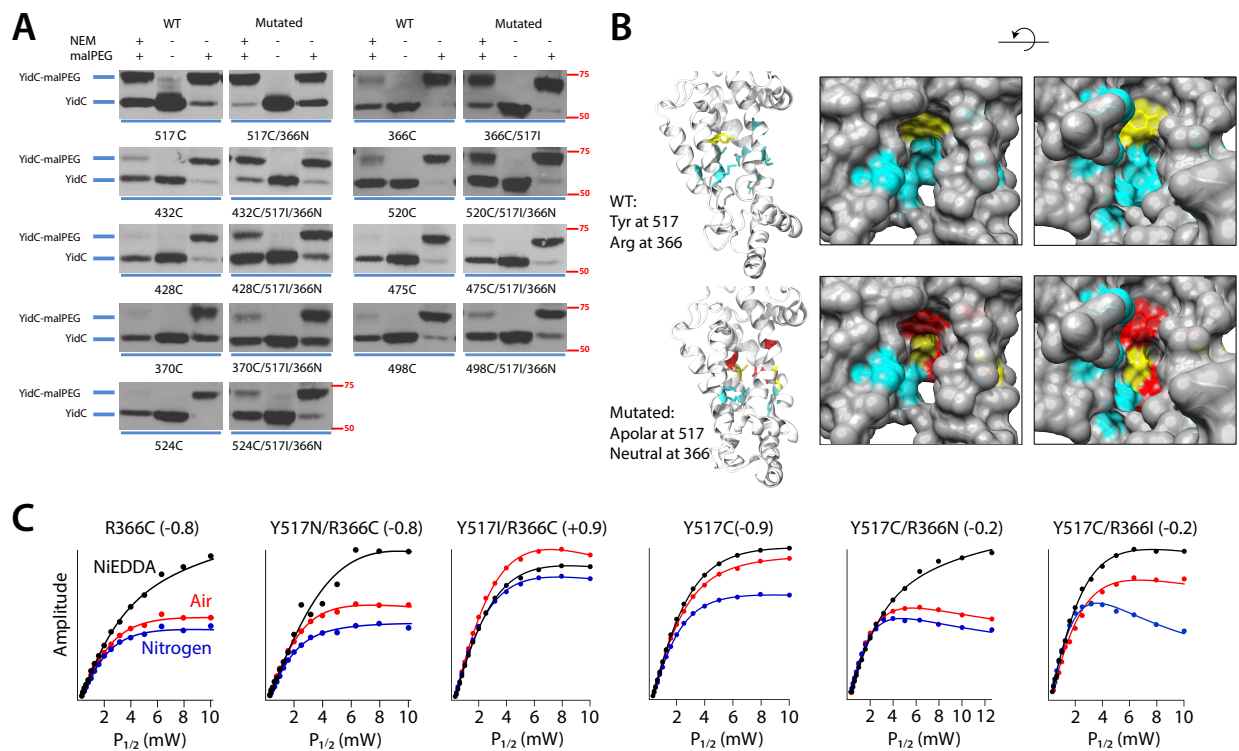


Figure 4

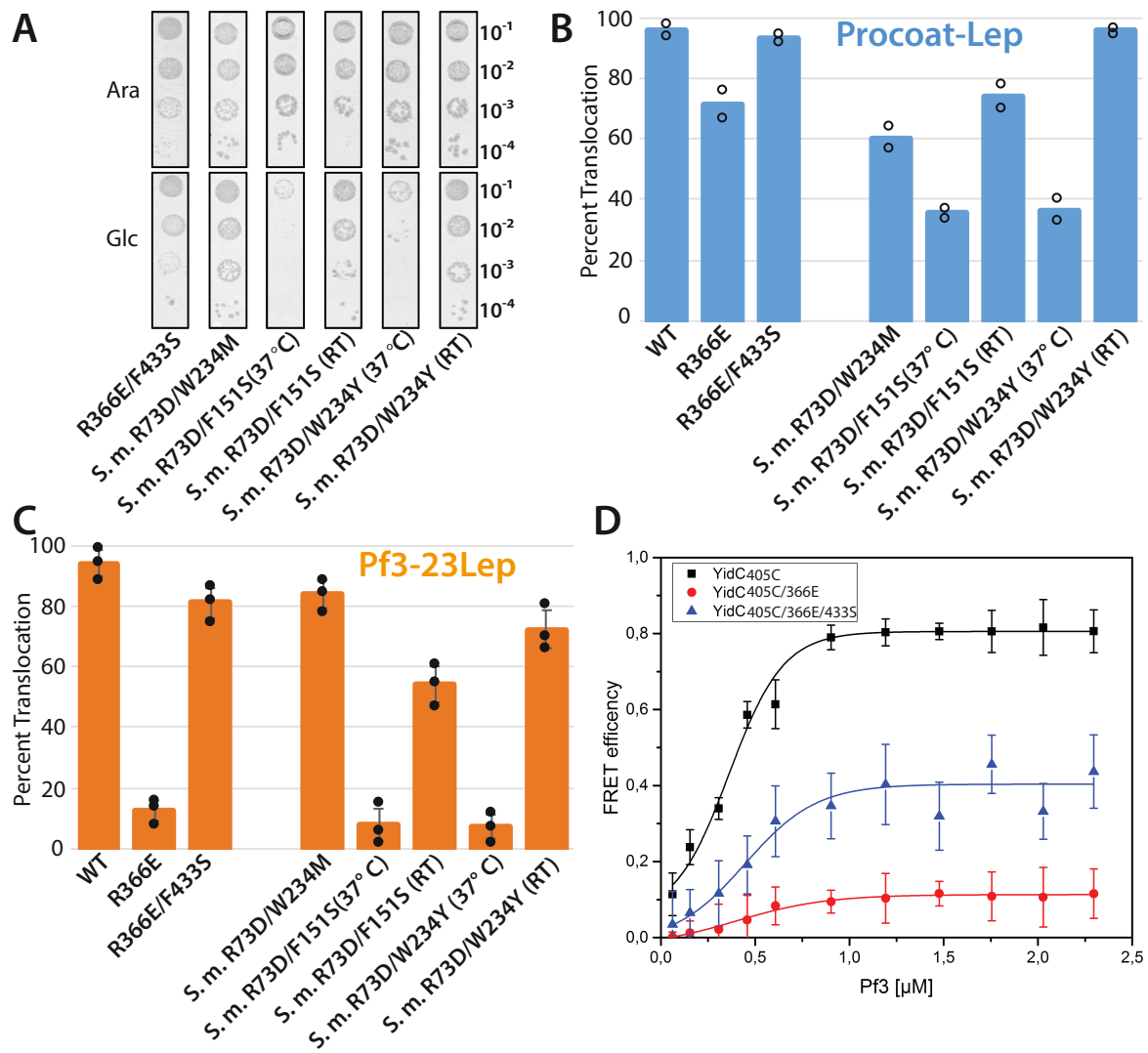


Figure 5

**The hydrophilic microenvironment in the substrate-translocating groove of the
YidC membrane insertase is essential for enzyme function**

Yuanyuan Chen¹, Marcos Sotomayor¹, Sara Capponi^{2,3}, Balasubramani Hariharan¹, Indra D.
Sahu^{4,7} Maximilian Haase⁵, Gary A. Lorigan⁴, Andreas Kuhn⁵, Stephen H. White⁶, and Ross E.
Dalbey

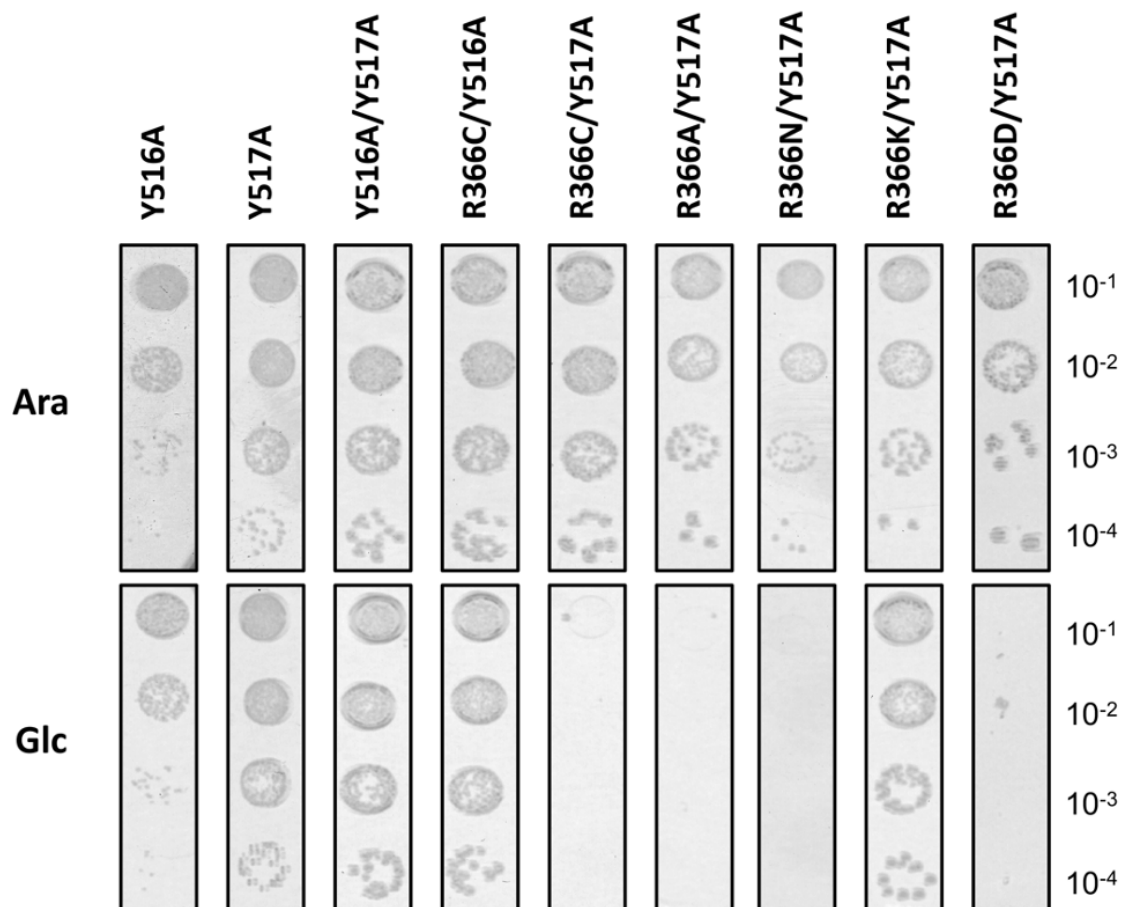


Figure S1. The *E. coli* YidC Y517A mutant requires a positive charge in the hydrophilic groove for function. Complementation assay was performed to examine the importance of the R366 residue for *E. coli* YidC function when Y517 was substituted with Ala. The R366 residue was mutated to Ala, Cys, Asn, Lys or Asp in pACYC184 in combination with the Y517A mutation, and transformed into JS7131. Y516 was also substituted with Ala, or both Y516 and Y517 were changed to Ala, Y516 was mutated to Ala in combination with R366 mutated to Cys and Y517 in combination with R366 Asn, Lys and Asp, respectively. The spot tests were performed as described in Fig. 2A. Note that the data for R366A/Y517A and Y517A (see Fig. 2A) is included here for comparison.

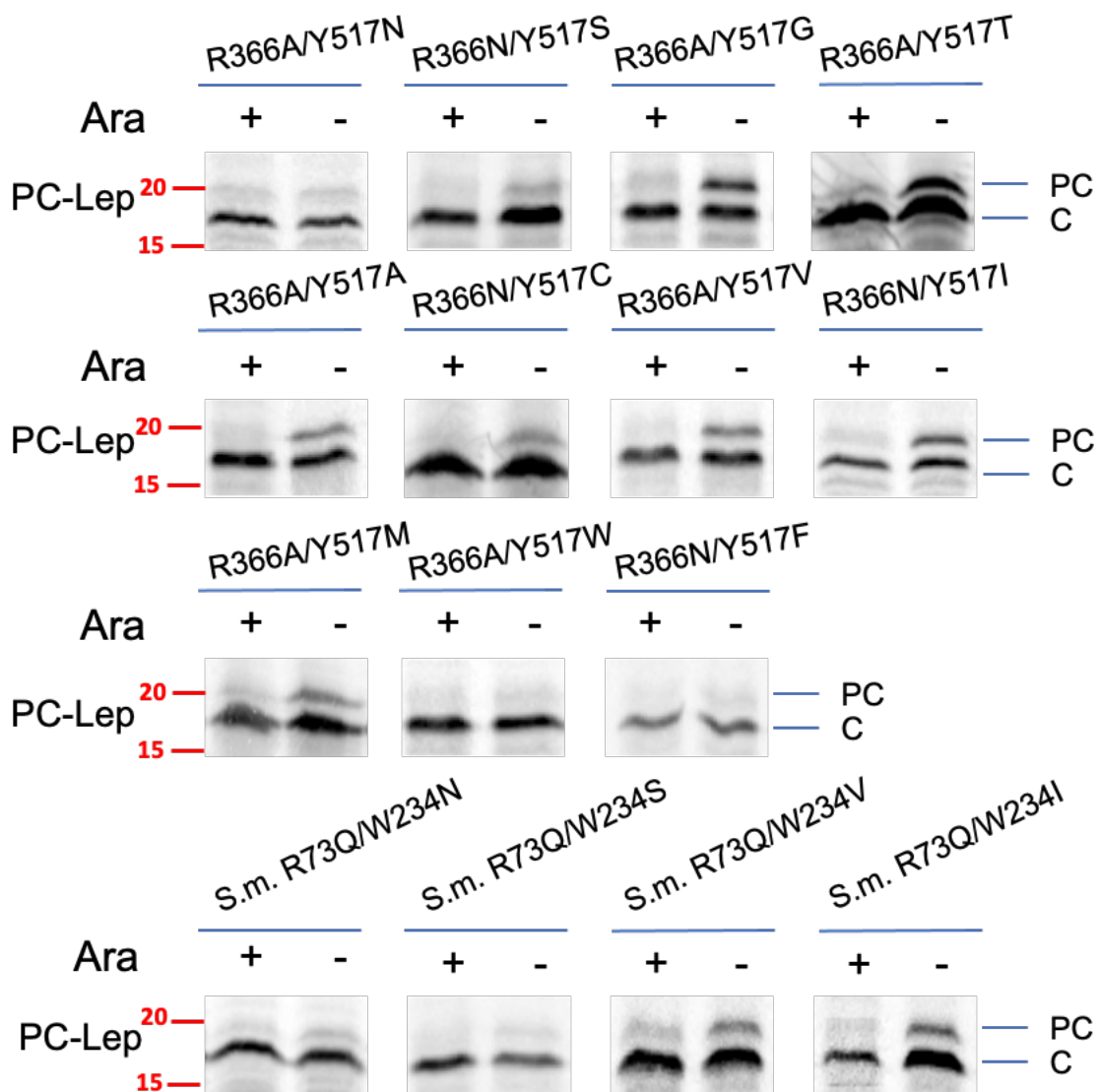


Figure S2A. Detailed results of insertion assay performed for Fig. 3D. A. Signal peptide processing assay was used to examine the insertion efficiency of PC-Lep as described in Fig. 2B. Red numbers indicate molecular weight marker values in kDa.

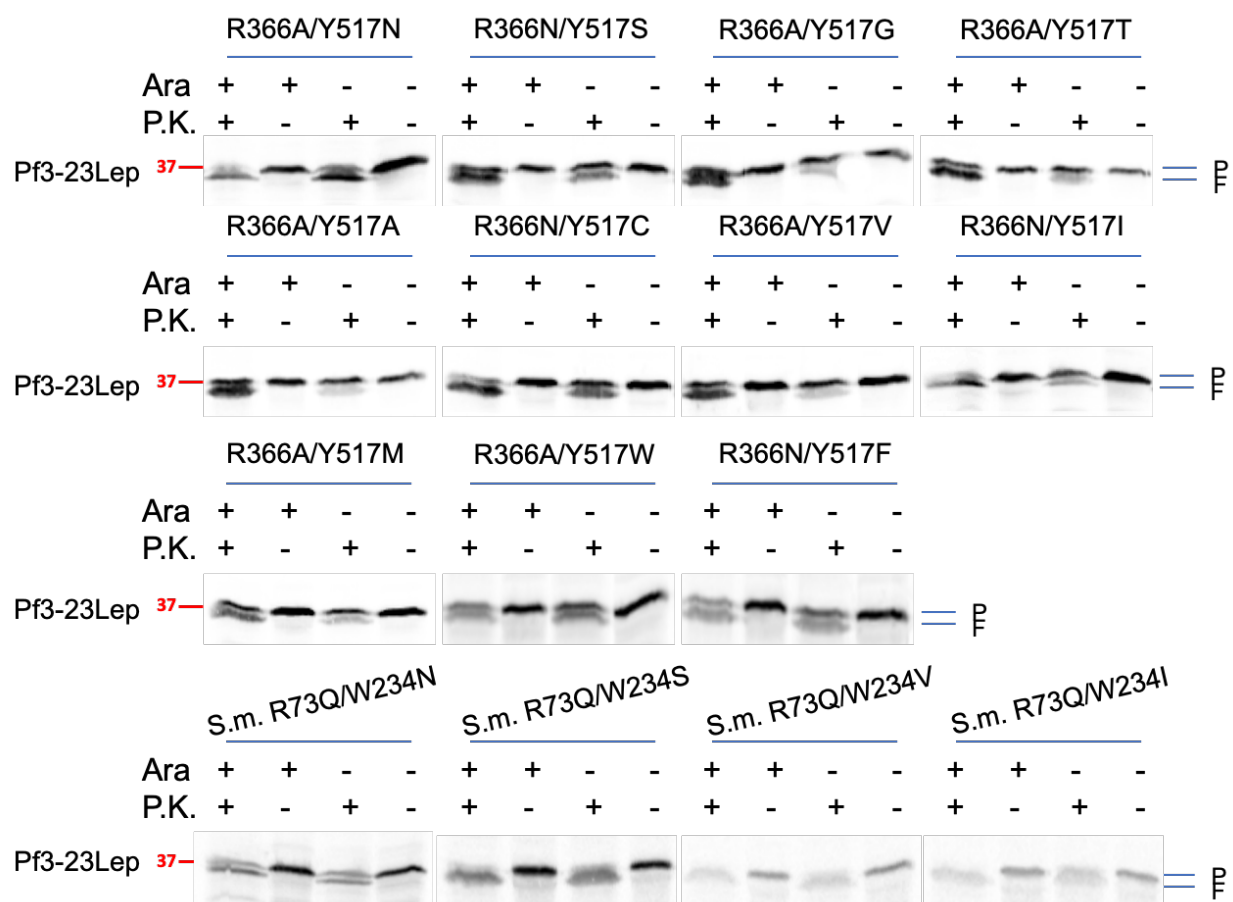


Figure S2B. Detailed results of insertion assay performed for Fig. 3D. B. The capability of *E. coli* YidC and *S. mutans* YidC2 mutants to insert Pf3-23Lep was tested by protease mapping assay as described in Fig. 2C. Red numbers indicate molecular weight marker values in kDa.

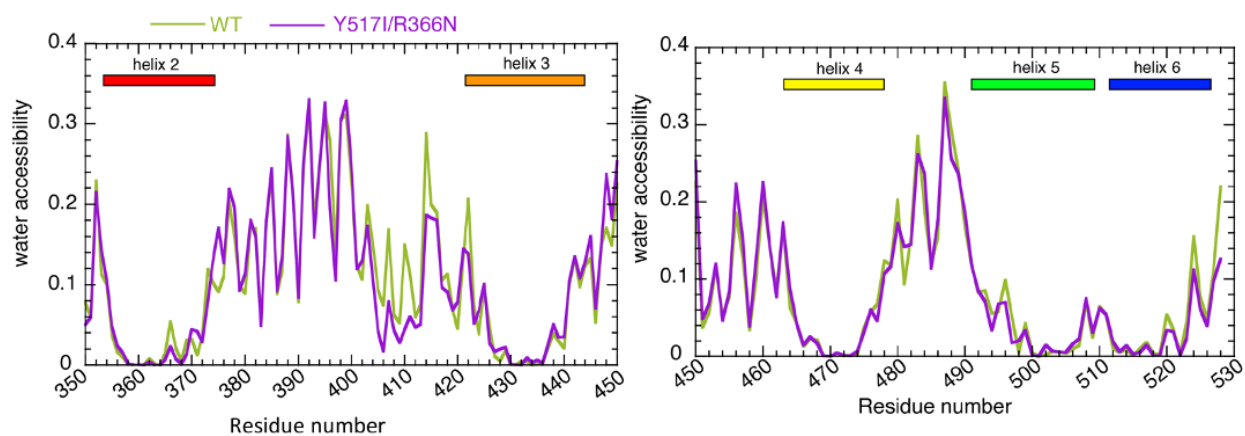


Figure S3. The water accessibilities of *E. coli* YidC wild-type and Y517I/R366N mutant determined from the MD simulations of YidC in POPE/POPG bilayers. A. Accessibility results from residues 350-450. B. Accessibility results for residues 450-528.

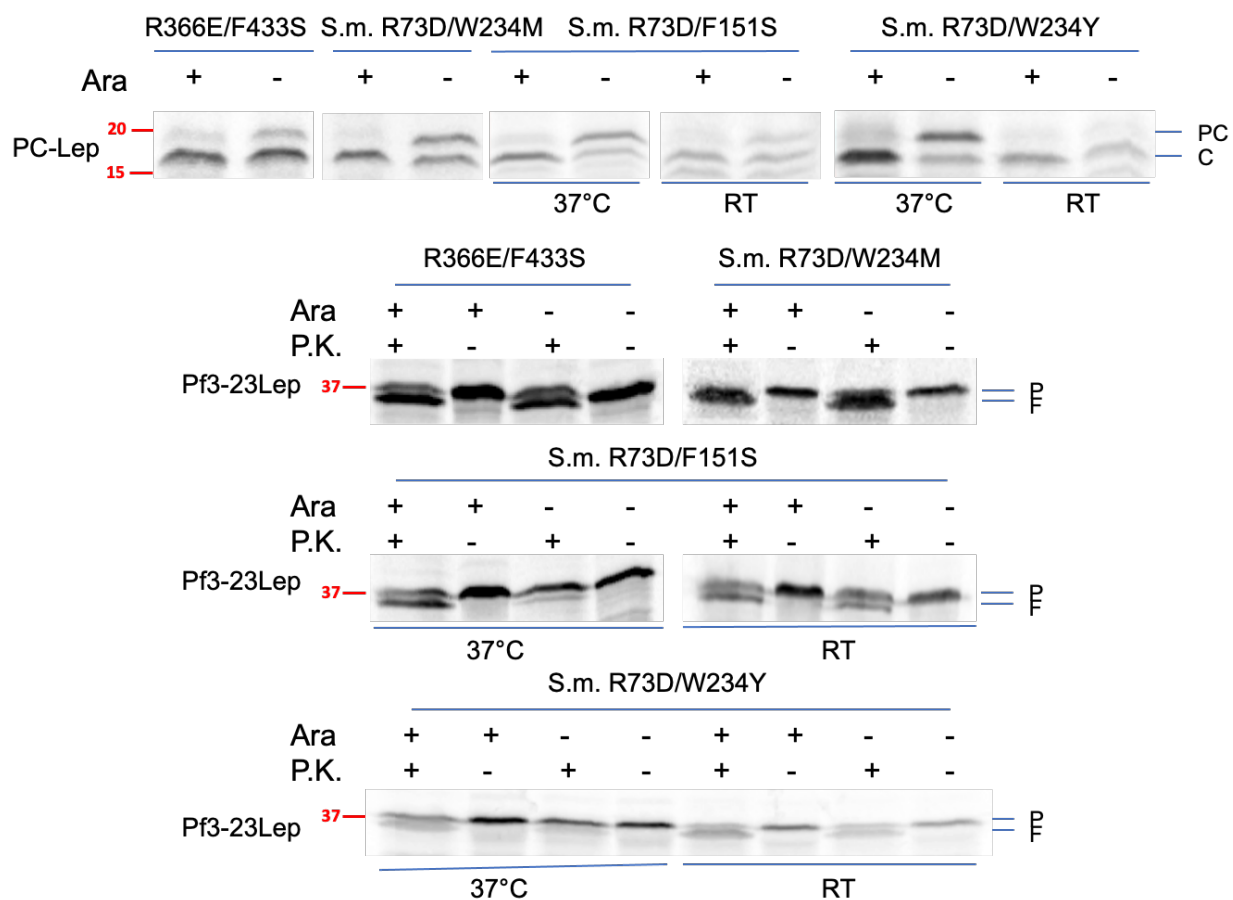


Figure S4. Detailed results of insertion assay with the negatively charged YidC mutants performed for Fig. 5B. Signal peptide processing (A) and protease mapping (B) assays were applied to examine the insertase activity of *E. coli* YidC and *S. mutans* YidC2 negative mutants for PC-Lep (A) and Pf3-23Lep (B), respectively. Red numbers indicate molecular weight marker values in kDa.

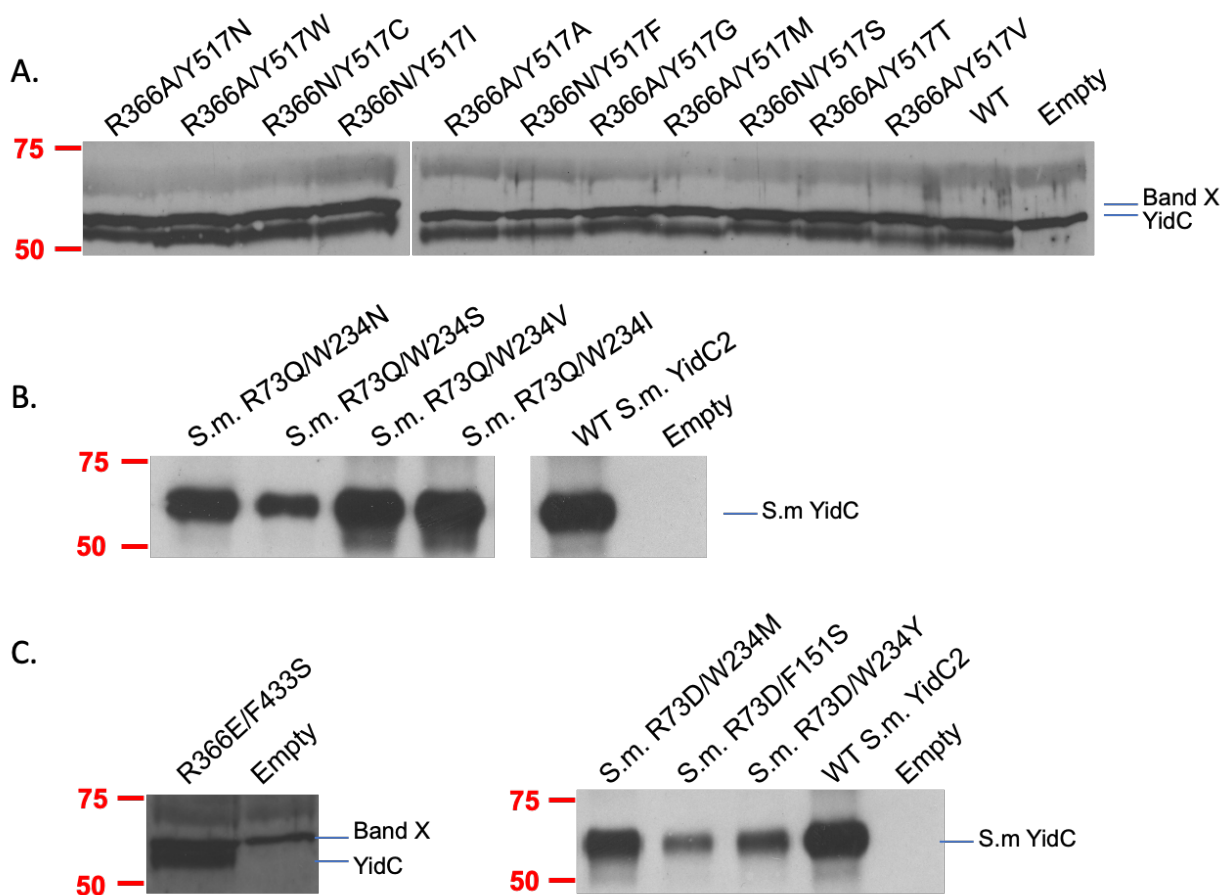


Figure S5. Western blots to check for expression levels of *E. coli* YidC and *S. mutans* YidC2 mutants. (A-B) Expression of *E. coli* YidC mutants (A) and *S. mutans* YidC2 mutants (B) for the studies in Fig. 3D. (C) Expression of the negatively charged *E. coli* YidC (left) and *S. mutans* YidC2 mutants (right) shown in Fig. S5B. In the panel for the *E. coli* YidC mutants, the top band is an unrelated protein recognized by our antiserum. The last two lanes in Fig. S5B are also used on the right-side of Fig. S5C as comparison controls. Red numbers indicate molecular weight marker values in kDa.

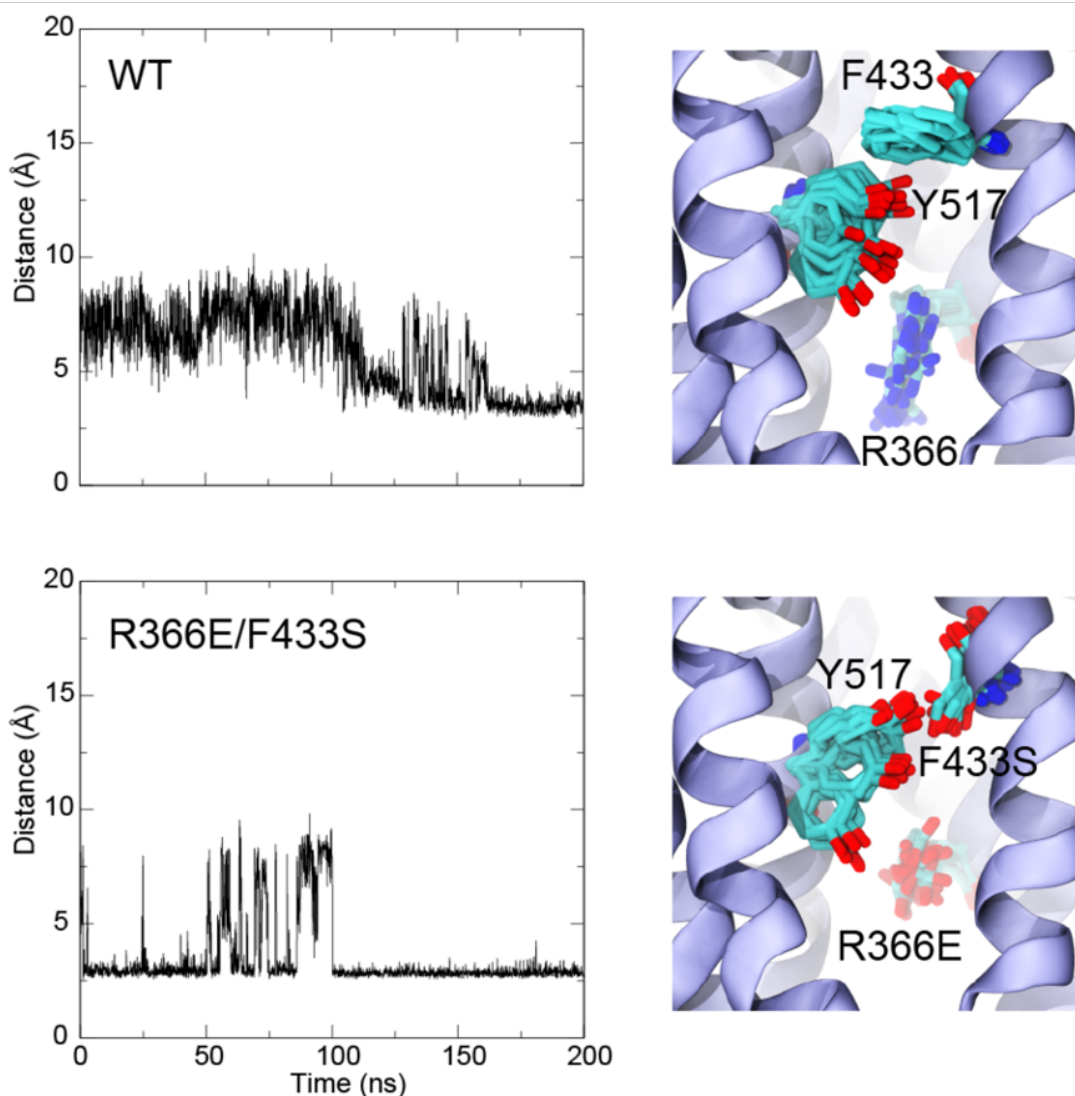


Figure S6. Dynamics of YidC residues controlling substrate translocation. Distance between atoms F433-C γ and Y517-O (top left) and atoms F433S-O γ and Y517-O (bottom left) versus time during simulations S1a-b and S3a-b, respectively. Right panels show detail of YidC's hydrophilic groove with key residues shown in stick representation for superposed conformations taken every 8 ns of simulation trajectory. Side chains of residues F433S and Y517.