

# UCLA

## UCLA Previously Published Works

### Title

YidC assists the stepwise and stochastic folding of membrane proteins

### Permalink

<https://escholarship.org/uc/item/57p936br>

### Journal

Nature Chemical Biology, 12(11)

### ISSN

1552-4450

### Authors

Serdiuk, Tetiana  
Balasubramaniam, Dhandayuthapani  
Sugihara, Junichi  
et al.

### Publication Date

2016-11-01

### DOI

10.1038/nchembio.2169

Peer reviewed

# YidC assists the stepwise and stochastic folding of membrane proteins

Tetiana Serdiuk<sup>1</sup>, Dhandayuthapani Balasubramaniam<sup>2</sup>, Junichi Sugihara<sup>2</sup>, Stefania A Mari<sup>1</sup>,  
H Ronald Kaback<sup>2-4\*</sup> & Daniel J Müller<sup>1\*</sup>

**How chaperones, insertases and translocases facilitate insertion and folding of complex cytoplasmic proteins into cellular membranes is not fully understood. Here we utilize single-molecule force spectroscopy to observe YidC, a transmembrane chaperone and insertase, sculpting the folding trajectory of the polytopic  $\alpha$ -helical membrane protein lactose permease (LacY). In the absence of YidC, unfolded LacY inserts individual structural segments into the membrane; however, misfolding dominates the process so that folding cannot be completed. YidC prevents LacY from misfolding by stabilizing the unfolded state from which LacY inserts structural segments stepwise into the membrane until folding is completed. During stepwise insertion, YidC and the membrane together stabilize the transient folds. Remarkably, the order of insertion of structural segments is stochastic, indicating that LacY can fold along variable pathways toward the native structure. Since YidC is essential in membrane protein biogenesis and LacY is a model for the major facilitator superfamily, our observations have general relevance.**

The biogenesis of membrane proteins in bacteria requires insertion into the cytoplasmic membrane by SecYEG and YidC, members of two highly evolutionarily conserved pathways<sup>1-3</sup>. The translocase SecYEG is critical for the insertion of proteins into the plasma membrane in bacteria and archaea, the endoplasmic reticulum in eukaryotes, and the chloroplast thylakoid membrane in plants<sup>4</sup>. The membrane protein insertase YidC can either couple with the SecYEG pathway to facilitate folding and assembly of membrane proteins inserted by the Sec translocase, or it can function as an independent pathway<sup>5</sup>. In the YidC pathway, insertase activity catalyzes insertion, folding and assembly of proteins into the bacterial cytoplasmic membrane<sup>2</sup>. Thus, YidC can function both as an insertase and as a chaperone. This YidC pathway and those of its family members, Oxa1 in mitochondria and Alb3 in chloroplasts, are highly conserved<sup>6,7</sup>. Chaperones promote folding of soluble proteins by suppressing misfolding and aggregation<sup>8</sup>, but it is not understood how chaperones and insertases assist insertion and folding of integral membrane proteins in general and cytoplasmic membrane proteins in particular. Furthermore, since basic elements of cytoplasmic membrane protein biogenesis are conserved in prokaryotes and eukaryotes, insight into how YidC facilitates insertion and folding of membrane proteins is broadly relevant.

Single-molecule force spectroscopy (SMFS) is the method of choice for the *in vitro* characterization of mechanical unfolding and refolding of single proteins<sup>9</sup>. In particular, atomic-force microscopy (AFM)-based SMFS can be applied to observe the unfolding and folding pathways of individual proteins in native or synthetic membranes at the level of secondary structural elements<sup>10-13</sup>. In the cellular context, unfolded proteins are exposed to various specific and nonspecific interactions with macromolecules. Among these, molecular chaperones are crucial for preventing unfolded polypeptides from misfolding and aggregating<sup>8</sup>. A few SMFS studies describe the effects of chaperones on the folding trajectories of soluble<sup>14-16</sup> and  $\beta$ -barrel outer membrane proteins<sup>17</sup>. However, it has not been possible so far to monitor how a chaperone and a translocon together guide the folding of cytoplasmic integral membrane proteins.

Here we have investigated the effect of YidC on the insertion and folding of single lactose permeases from *Escherichia coli* (LacY) under physiologically relevant conditions. LacY, a model for the major facilitator superfamily (MFS)<sup>18</sup>, has 12 mostly irregular transmembrane  $\alpha$ -helices organized into two pseudo symmetrical six-helix bundles connected by a relatively long cytoplasmic loop<sup>19,20</sup>. It has been shown that LacY requires YidC in order to fold correctly in the membrane after the SecYEG complex inserts LacY<sup>21,22</sup>. In our previous work we have applied SMFS to fully unfold native, functional LacY from phospholipid membranes<sup>23</sup>. Exposed to mechanical stress, LacY unfolds structural segments stepwise, giving rise to a highly reproducible unfolding pattern that is sensitive to the functional state<sup>23</sup> and to misfolding<sup>24</sup> of the transporter. Here we have used the SMFS-based assay to unfold a large part of LacY and then to follow the refolding of the unfolded substrate into the phospholipid membrane in the absence or presence of YidC. We found that LacY alone could insert structural segments into the membrane, but misfolding dominated this process and prevented the substrate from completing folding. YidC stabilized the unfolded LacY substrate, from which structural segments could insert one after the other until folding was completed. Surprisingly, the sequence in which this insertion and folding process occurred was stochastic.

## RESULTS

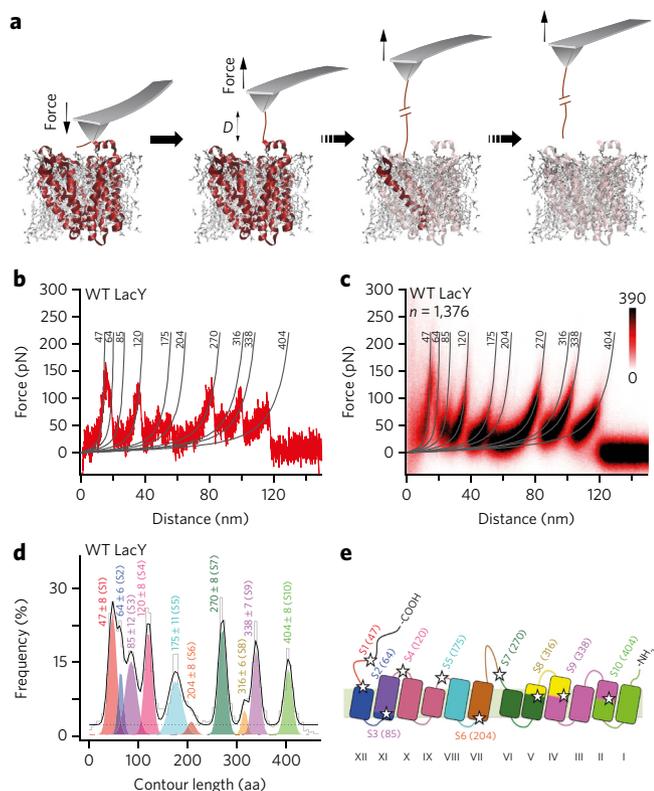
### LacY unfolds stepwise on exposure to mechanical stress

To set up the insertion and folding experiments, we reconstituted LacY into phospholipid membranes composed of 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoethanolamine (PE) and 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phospho-rac-(1-glycerol) (PG) at a ratio of 3:1 PE/PG<sup>25</sup> (Supplementary Results, Supplementary Fig. 1). At this phospholipid composition, which is close to the composition of the *E. coli* membrane (70–80% zwitterionic PE and 20–25% anionic PG plus cardiolipin), LacY is fully functional and adopts the native topology with the N and C termini on the cytoplasmic surface of the membrane<sup>19,26</sup>. For AFM imaging, the LacY proteoliposomes were adsorbed to mica in buffer solution, where they appear as

<sup>1</sup>Department of Biosystems Science and Engineering, Eidgenössische Technische Hochschule (ETH) Zurich, Basel, Switzerland. <sup>2</sup>Department of Physiology, University of California–Los Angeles, Los Angeles, California, USA. <sup>3</sup>Department of Microbiology, Immunology & Molecular Genetics, University of California–Los Angeles, Los Angeles, California, USA. <sup>4</sup>Molecular Biology Institute, University of California–Los Angeles, Los Angeles, California, USA.

\*e-mail: rkaback@mednet.ucla.edu or daniel.mueller@bsse.ethz.ch





**Figure 1 | Unfolding steps and intermediates of native wild-type LacY.**

(a) Schematic of use of SMFS to mechanically unfold single LacY. After AFM imaging of the proteoliposome, the stylus of the cantilever is gently pressed onto LacY (red; PDB 1PV7) to nonspecifically attach the elongated C terminus. Then the cantilever is withdrawn to unfold LacY stepwise (horizontal arrows) until it has been completely unfolded and extracted from the phospholipid membrane<sup>23</sup>, while unfolding force and distance ( $D$ ) are recorded. (b) Force–distance curve (red) recorded upon mechanically unfolding a single wild-type (WT) LacY from the membrane. Fitting of each force peak with the WLC model<sup>27</sup> (gray lines) provides the contour length of the unfolded and stretched polypeptide (gray labels indicate contour length in amino acid residues). (c) Superimposition of 1,376 force–distance curves, each recorded during the unfolding of a single LacY. Color scale indicates density of curves in the superimposition in arbitrary units. The resulting fingerprint spectrum shows ten force peaks, each representing the unfolding of a stable structural segment of LacY. WLC curves and contour lengths are shown in gray as in (b). (d) Contour-length histogram of force peaks detected upon unfolding of LacY. Fitting the histogram with a Gaussian mixture model (black line) reveals the mean contour length of every force peak class (shown in amino acid residues (aa)  $\pm$  s.d.). The dashed line shows the uniform baseline noise. Each mean contour length of a force peak class determines the end of the previous structural segment and the beginning of the next structural segment. (e) The ten stable structural segments S1–S10, colored to match force peak classes in (d) and mapped to the secondary structure of LacY (PDB 1PV7). Each force peak class is marked by a star, and numbers indicate the mean contour lengths of force peak classes (in aa).

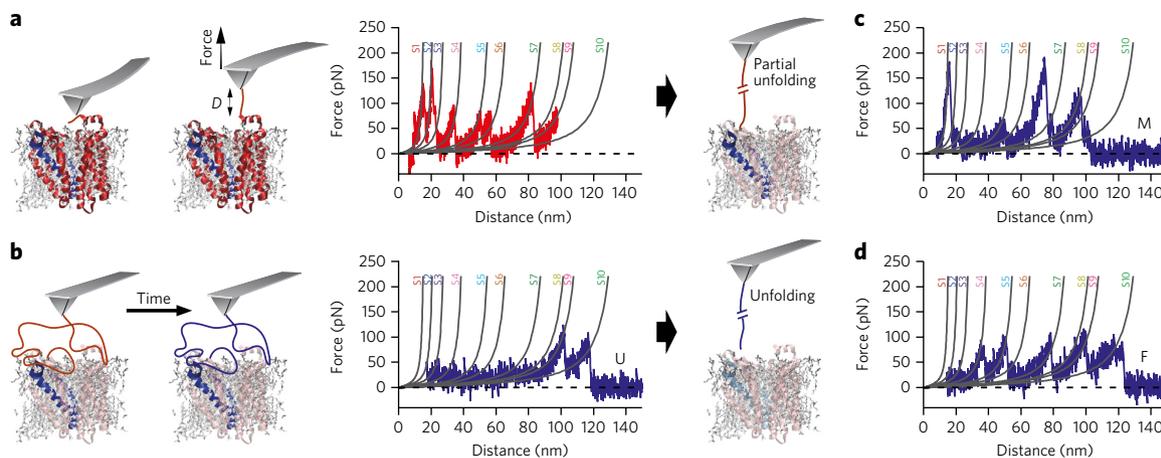
single-layered membranes (Supplementary Fig. 1b). Within the membranes, high-resolution AFM topographs showed evenly distributed LacY particles.

For SMFS, we proceeded as recently described for LacY<sup>23</sup>. Briefly, the AFM stylus was pushed gently (700 pN for 500 ms) onto LacY to nonspecifically attach individual LacY polypeptides to the stylus (Fig. 1a)<sup>27</sup>. To enhance the probability of attaching the C terminus, we extended the C terminus of LacY with a 36-residue unstructured

‘poly(Gly)’ polypeptide (GSM(G<sub>11</sub>)EAVEEAVEEA(G<sub>11</sub>)S) followed by an 8-residue His tag (His<sub>8</sub> tag; Supplementary Fig. 2a). The extension has little or no effect on LacY transport activity<sup>23</sup> and did not change the folding of LacY (Supplementary Fig. 2b,c). We then withdrew the stylus from the membrane while recording the deflection of the AFM cantilever as a force–distance curve (Fig. 1a,b). The force–distance curves typically exhibited a saw-tooth-like pattern with ten force peaks each indicating an unfolding step of the transporter. Repeating the experiment multiple times revealed a highly reproducible saw-tooth-like pattern (Fig. 1c). By fitting each of the ten force peaks with the worm-like-chain (WLC) model<sup>27</sup> we determined the contour lengths of the peptide segments unfolded in each step (Fig. 1b,c). Since we applied mechanical pulling force to the elongated C terminus of LacY, the C terminus was stretched until the first C-terminal bundle unfolded (Fig. 1a). This unfolding step elongated the polypeptide tethering the AFM stylus and the partially unfolded LacY molecule. Subsequent withdrawal of the stylus unfolded the transmembrane bundles stepwise in nine distinct structural segments formed by single or grouped  $\alpha$ -helices and their connecting polypeptide loops (Fig. 1d). Thus, the force–distance curves recorded a characteristic ‘fingerprint’ spectrum for native LacY, with single force peaks corresponding to unfolding of secondary structure segments (Fig. 1d,e). The summed length of the force–distance curves corresponds to the contour length of the fully unfolded and stretched LacY polypeptide<sup>23</sup>. It has been previously shown for LacY and for other membrane transporters and receptors that the force peaks in the native fingerprint spectrum are sensitive to the fold, functional state, ligand binding, lipid composition, mutations and supramolecular assembly of the membrane protein<sup>9,23,24,28–31</sup>. It is of particular interest for the present study that the characteristic fingerprint spectrum observed for native LacY is sensitive to misfolding of the transporter<sup>24</sup>. LacY inverts topology upon depletion of PE and becomes inactive<sup>32</sup>; when LacY inverts topology, the force peaks detected for the structural segments of native LacY change positions, and the shifted peaks can be used to identify the segments changing their fold<sup>24</sup>. Thus, the comparison of the native fingerprint spectrum of LacY with that recorded for single LacY molecules can identify misfolded transporters and structural segments.

### Partially unfolded LacY remains stably in the membrane

The highly reproducible fingerprint spectrum characterizing the unfolding steps of LacY (Fig. 1c) indicated that the unfolding intermediates of LacY remained stably embedded in the membrane. Similarly, stable unfolding intermediates formed by individual or grouped  $\alpha$ -helices and polypeptide loops have been observed for other transmembrane proteins, including members of the MFS family<sup>9,10,30,31,33</sup>. Molecular dynamics simulations also indicate that transmembrane  $\alpha$ -helical proteins exposed to mechanical force unfold stepwise and that the partially unfolded protein remains stably folded in the membrane<sup>34,35</sup>. To test more rigorously whether partially unfolded LacY behaves similarly, we conducted control experiments in which we unfolded LacY to different extents and left the partially unfolded protein in the membrane for 2 or 5 s (Supplementary Fig. 3). We then reapplied the pulling force to the partially unfolded  $\alpha$ -helical bundle, which exhibited the same unfolding pattern as observed for native LacY (Fig. 1c). Thus, partially unfolded LacY remained stable in the membrane and did not rearrange over a time course of seconds. The findings were consistent with *in vivo* experiments demonstrating that contiguous, nonoverlapping LacY fragments expressed together can self-assemble into functional LacY<sup>36,37</sup>. Having shown that the transmembrane  $\alpha$ -helical bundles of LacY unfold stepwise and leave stable unfolding intermediates in the membrane, we tested whether LacY pulled partially out of the membrane can reinsert and refold back into the membrane.



**Figure 2 | Refolding single lactose permeases.** (a) The initial unfolding step of the refolding experiment. After attaching the AFM stylus to the C terminus, we withdrew the cantilever (left schematic) to sequentially unfold LacY (red) while force and distance ( $D$ ) were recorded, leaving the last structural segment, S10 (blue), in the membrane (right schematic). The force–distance curve recording this partial unfolding process (represented as in **Fig. 1b**) shows force peaks matching the fingerprint spectrum of native LacY (**Fig. 1c**). (b) The refolding step of the assay, with a substrate that remained unfolded. After partial unfolding as in **a**, we brought the AFM stylus close to the membrane surface ( $d \sim 10\text{--}15\text{ nm}$ ) to relax the unfolded polypeptide (left schematic). After a refolding time of 2 s (blue line indicates polypeptide post-refolding time), we completely withdrew the stylus (right schematic) and measured the force–distance curve, allowing us to characterize refolding by comparison with the native fingerprint spectrum (**Fig. 1c** and **Supplementary Fig. 4**). The curve in **b** is typical of a polypeptide that remains unfolded (U). (c,d) Typical force–distance curves showing misfolding (M) or refolding (F) of some structural segments. Experiments were conducted in 50 mM potassium phosphate at pH 7.2 and 25 °C.

### Outside the membrane LacY misfolds at high probability

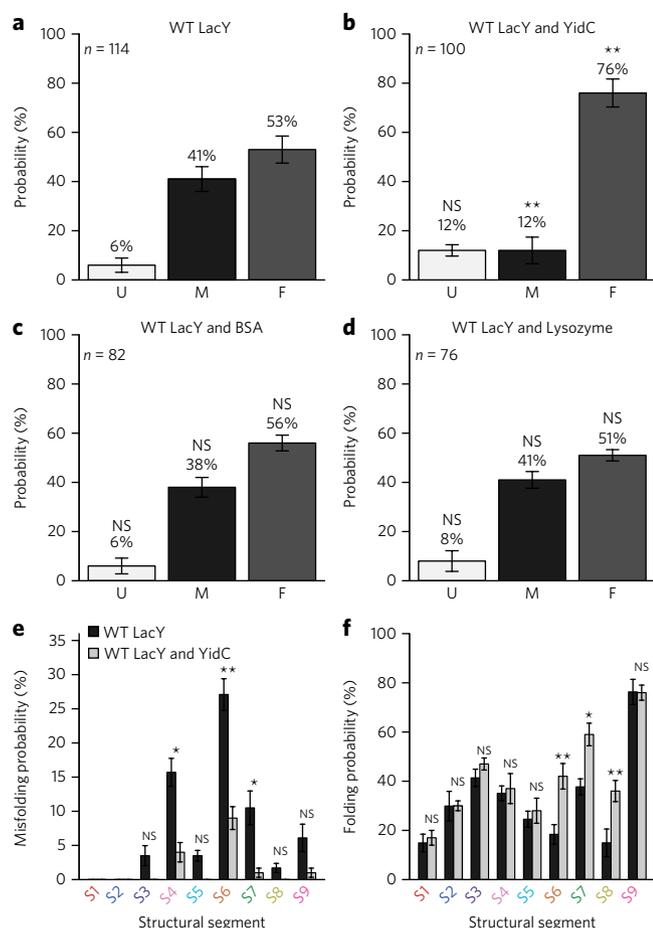
Insertion and folding of proteins into the cytoplasmic membrane starts from the N terminus<sup>1,3,38</sup>. Thus, in our *in vitro* refolding experiments we wanted to characterize the insertion and folding of LacY into the membrane proceeding from the N-terminal end. We used the SMFS technique described above to pick up the elongated C terminus, then unfolded and extracted from the membrane a large portion of LacY consisting of the C terminus, ten transmembrane  $\alpha$ -helices and the intervening loops, leaving the first two N-terminal transmembrane  $\alpha$ -helices in the membrane (**Fig. 2a**). We then lowered the AFM stylus into close proximity with the membrane ( $\sim 10\text{--}15\text{ nm}$ ), where the unfolded portion of LacY was allowed to insert and fold for 2 s. Finally, we probed the refolded polypeptide by pulling it from the membrane again (**Fig. 2b**). To interpret the force–distance curve from the latter pull, we used the native fingerprint spectrum of LacY as a template (**Fig. 1c**) to assess whether the polypeptide refolded native structural segments (**Supplementary Fig. 4**). Each measurement was classified into one of three categories: (i) if we detected no unfolding force peaks, we concluded that the LacY polypeptide did not adopt a stable fold and classified it as unfolded (**Fig. 2b**); (ii) if we observed one or more unfolding force peaks at contour lengths differing from the native fingerprint spectrum (**Fig. 1c**), we concluded that the polypeptide adopted a non-native fold and classified it as misfolded (**Fig. 2c**); (iii) if we recorded unfolding force peaks only at positions covered by the native fingerprint spectrum, the substrate was classified as refolding individual structural segments into the membrane (**Fig. 2d**). We then carried out 114 independent refolding experiments (**Fig. 3a** and **Supplementary Fig. 5**). After a refolding time of 2 s, 6% of the unfolded LacY substrates remained unfolded and 41% exhibited force peaks not corresponding to the native fingerprint spectrum of LacY, representing misfolding. 53% of the refolding experiments exhibited force peaks corresponding to the native fingerprint spectrum and were classified as having folded native structural segments. Close inspection of the force peaks in the single refolding experiments indicated that the folding of single native structural segments occurred as a stochastic process (**Supplementary Fig. 5**). However, in none of the experiments was a refolding spectrum with all the

force peaks of native LacY observed. Thus, although LacY refolded individual structural segments into the membrane, the polypeptide did not complete folding.

### YidC supports native folding intermediates

To test whether YidC affects insertion and folding of LacY, we purified and coreconstituted both proteins into PE/PG (3:1) membranes. Coreconstitution was verified by SDS–PAGE and AFM imaging (**Supplementary Fig. 1**) and by SMFS (**Supplementary Fig. 6**). Mechanical unfolding of single membrane proteins from the coreconstituted sample by SMFS yielded force–distance curves with two different patterns in equal frequency. One pattern was identical to that of native LacY, with ten force peaks (**Supplementary Fig. 6a**)<sup>23</sup>. The second pattern had only eight force peaks that extended toward much longer contour lengths than those observed for LacY alone (**Supplementary Fig. 6b**). The YidC polypeptide with a C-terminal His<sub>10</sub> tag is much longer than LacY (561 versus 461 residues), which suggests that the longer second pattern corresponds to the unfolding of YidC into a fully stretched conformation. Indeed, reconstitution of YidC alone produced force–distance curves of the second pattern (**Supplementary Fig. 7**). To determine how the coreconstituted LacY and YidC were distributed within the membrane, we conducted spatially resolved SMFS experiments (**Supplementary Fig. 8**). We found that LacY and YidC were in close proximity ( $<10\text{ nm}$ ). Such close proximity could result from a kinetically stable association of LacY and YidC; alternatively, an unstable association might have been facilitated by lateral diffusion, which is reduced but not abolished by adsorption to mica<sup>39</sup>. Thus, data from coreconstitution of LacY and YidC indicated that both proteins were present in the phospholipid membrane and in close proximity.

To investigate YidC-assisted insertion and folding, we repeated the LacY refolding experiments in the presence of YidC (**Supplementary Fig. 9**). Typically, we observed that LacY refolded a few but not all structural segments in 2 s time. Following the criteria described above, we classified each refolding experiment as unfolded, misfolded or refolded (**Fig. 3a,b**). The proportion of LacY refolding trajectories that showed folding of native structural elements increased from 53% to 76% in the presence of YidC;



**Figure 3 | YidC suppresses misfolding and assists folding of LacY.**

(a,b) Probability that the unfolded wild-type (WT) LacY polypeptide remains unfolded (U), misfolds (M) or refolds (F) in the absence (a) or presence (b) of YidC. Refolding experiments were conducted as described in **Figure 2** with a refolding time of 2 s. (c,d) Control experiments in which unfolded LacY was refolded in the presence of either 1  $\mu$ M BSA or 1  $\mu$ M lysozyme. (e,f) Probability that structural segments S1–S9 misfold (e) or fold (f) in the absence or presence of YidC. Refolding experiments analyzed in e,f are those from a,b. *n* is the number of single LacY refolding experiments analyzed. Error bars indicate the s.e.m. as estimated by resampling using 4,000 iterations. Significance of differences from LacY alone is indicated: NS,  $P \geq 0.01$ ; \* $P < 0.01$ ; \*\* $P < 0.001$ , as determined by two-tailed Z-test. Additional statistics obtained from  $\chi^2$  and two-tailed Fischer's exact tests are provided in **Supplementary Tables 1–4**. For single refolding experiments, see **Supplementary Figures 5 and 9**.

trajectories indicating the unfolded state increased slightly from 6% to 12%; and trajectories indicating the misfolded state decreased from 41% to 12%. Interestingly, the force required to unfold structural segments of native LacY ( $94.3 \pm 27.7$  pN; mean  $\pm$  s.d.,  $n = 614$ ) was similar to the force required to unfold segments that had been refolded in the absence ( $93.5 \pm 30.1$  pN;  $n = 408$ ) or presence ( $84.3 \pm 25.3$  pN;  $n = 493$ ) of YidC. Thus, structural segments of native and refolded LacY had similar mechanical stability.

Control experiments conducted in the presence of lysozyme or BSA did not increase the probability that the unfolded LacY polypeptide would insert and fold, indicating that nonspecific interactions did not affect refolding outcomes (**Fig. 3c,d**). The observations suggest that YidC protected LacY from misfolding and supported the refolding of individual structural segments in the manner observed for native LacY. However, we also observed that the refolding LacY substrate did not produce all of the force peaks observed with native LacY.

### YidC protects $\alpha$ -helices V–VII and IX–X from misfolding

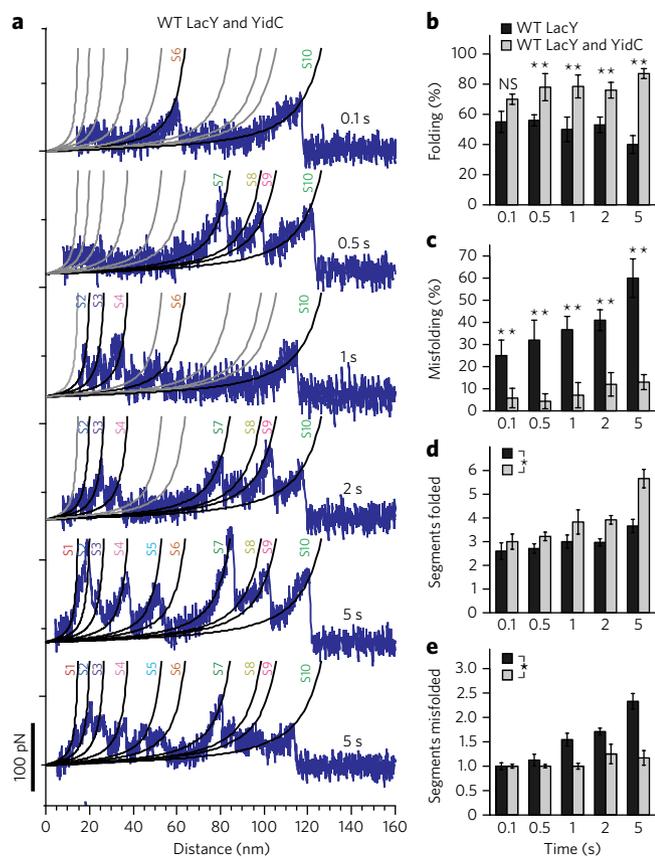
Which structural regions of LacY are prone to misfolding in the absence of YidC, and which are preferentially inserted and folded in the presence of YidC? To answer this query, we analyzed the LacY refolding experiments carried out in the absence of YidC to obtain the positions of every misfolded force peak. These positions were compared with those in the native fingerprint spectrum of LacY. To approximate the position of a misfolded structural segment, we assigned the force peak indicating misfolding to the closest force peak detected in the native fingerprint spectrum. In the absence of YidC, structural segments S4, S6 and S7 had the highest probability of misfolding (**Fig. 3e**). Segments S6 and S7 contain transmembrane  $\alpha$ -helices V, VI and VII and the long cytoplasmic loop<sup>19</sup>. Within these secondary structures, helix VII is weakly hydrophobic and is considered to be the 'weakest link' in LacY folding, stability and topogenesis<sup>32,40</sup>. Asp237 and Asp240 in helix VII form salt bridges with Lys358 (helix XI) and Lys319 (helix X), respectively<sup>41</sup>, and they become negatively charged as the salt bridge is broken when the partner helix is pulled from the membrane. Such mutual stabilization of secondary structures from S4, S6 and S7, which is crucial for LacY folding and function, may explain the high misfolding probability observed. The situation changed markedly in the presence of YidC, which suppressed misfolding of these structural segments and enhanced the probability of folding properly into the membrane (**Fig. 3e,f**).

In the absence of YidC, structural segment S6, which encompasses transmembrane  $\alpha$ -helix VII, had the highest probability of misfolding (27%) among all structural segments of LacY (**Fig. 3e**). As shown previously,  $\alpha$ -helix VII tends to misfold upon depletion of PE<sup>24</sup>. Misfolding of  $\alpha$ -helix VII may thus disturb the folding of  $\alpha$ -helix X, which interacts with  $\alpha$ -helix VII via the 240–319 salt bridge<sup>41</sup>. Interestingly, structural segment S4, which contains  $\alpha$ -helices IX and X, had the second-highest misfolding probability (16%). Furthermore, the refolding experiments in which misfolding of segment S4 was detected also showed that segment S6 was misfolded, whereas in only 6% of all cases segment S4 misfolded and segment S6 folded correctly. The data thus suggest that misfolding of the structural segment with  $\alpha$ -helices IX and X and misfolding of the segment with  $\alpha$ -helix VII are coupled. Possibly, misfolding of  $\alpha$ -helix VII destabilizes  $\alpha$ -helices IX and X because the stabilizing salt bridges and hydrogen bond network among Tyr236 and Asp240 in  $\alpha$ -helix VII, Arg302 in  $\alpha$ -helix IX, and Lys319, His322 and Glu325 in  $\alpha$ -helix X cannot be formed<sup>19</sup>. The chaperoning function of YidC thus appears to preferentially prevent misfolding of the structural segments encompassing  $\alpha$ -helices VII, IX and X.

### YidC supports insertion and folding steps to completion

In the absence or presence of YidC, unfolded LacY could insert and fold structural segments in the membrane. Our single-molecule refolding experiments indicated that in both cases, the individual structural segments inserted and folded in a stochastic process (**Supplementary Figs. 5 and 9**). This result surprised us, as we expected a priori that YidC would determine the sequence of insertion and folding.

Within the refolding time of 2 s, our refolding experiments yielded the correct folding of a few but not all of the structural segments detected in native LacY. Therefore, we investigated whether increasing the refolding time would increase the number of structural segments that refold properly. Unfolded LacY was allowed to refold for times ranging from 0.1 to 5 s in the absence or presence of YidC, and the folding intermediates were characterized (**Fig. 4** and **Supplementary Fig. 10**). In the absence of YidC, increasing the refolding time increased the number of force peaks and thus the number of refolding steps (**Supplementary Fig. 10**). However, with increasing refolding time the fraction of correctly folded curves decreased (**Fig. 4b**),

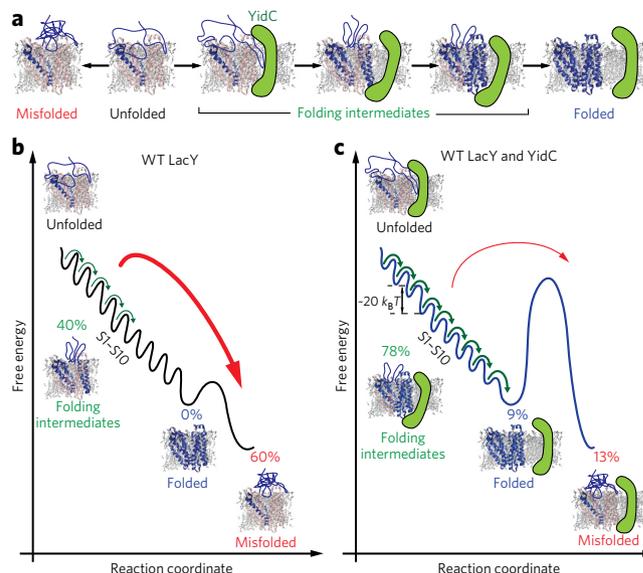


**Figure 4 | LacY shows variable folding intermediates toward full refolding.**

(a) Force–distance curves for unfolded wild-type (WT) LacY refolding in the presence of YidC, with refolding times varying from 0.1 to 5 s (experiments were conducted as described for Fig. 2). WLC curves with the structural segment indicated are taken from the native fingerprint spectrum of LacY (Fig. 1c). WLC curves matching the refolding force peaks in terms of mean  $\pm$  s.d. (Supplementary Fig. 4) are colored black; those not matching any peak are colored gray. (b,c) Probability that the unfolded LacY substrate correctly folded structural segments (b) or misfolded (c) after the indicated refolding times, in the absence (black; 204 experiments analyzed) or presence (gray; 206 experiments analyzed) of YidC. For single refolding experiments, see a and Supplementary Figure 10. (d,e) Number of folded segments (d) and misfolded segments (e) per substrate after the indicated refolding times, in the absence (black) or presence (gray) of YidC. Error bars indicate s.e.m. as estimated by resampling using 4,000 iterations. In b,c, significance of differences from LacY alone is indicated: NS,  $P \geq 0.01$ ; \* $P < 0.01$ ; \*\* $P < 0.001$ , as determined by two-tailed Z-test (Supplementary Tables 1 and 2). Over all refolding times, the number of folded (d) or misfolded segments (e) differed significantly in the presence versus the absence of YidC (\* $P < 0.01$  as determined by analysis of covariance (ANCOVA)).

while the probability of detecting misfolding events increased greatly (Fig. 4c). Consequently, at 5 s refolding time, misfolding of the LacY substrate dominated (~60%). No LacY remained unfolded, and the number of correctly folded structural segments remained constant for all refolding times (Fig. 4d), while with increasing refolding time the number of misfolded segments increased by ~130% to ~2.3 (Fig. 4e). Thus, with increasing refolding time, force peaks corresponding to misfolded segments did not transform into force peaks of the fingerprint spectrum of native LacY. Moreover, none of the refolding experiments conducted over extended times yielded a complete set of force peaks like that of native LacY.

This changed when we refolded the LacY substrate in the presence of YidC (Fig. 4a,b). Again, the number of force peaks detected



**Figure 5 | Schematic model of the folding free-energy landscape of LacY sculpted by YidC.**

(a) LacY partially unfolded and extracted from the membrane misfolds with high probability. YidC prevents LacY from misfolding and enables the unfolded substrate to adopt the native folded state via numerous folding intermediates. (b) In the absence of YidC, the unfolded LacY substrate inserts single structural segments into the membrane. After 5 s none of the LacY substrates remain unfolded, and because the structural segments have a higher probability of misfolding than folding correctly, the substrate predominantly misfolds (~60%). (c) YidC stabilizes the unfolded LacY substrate and reduces misfolding from ~60% to ~13%. From this stabilized unfolded state, ~87% of the LacY substrates insert and fold structural segments into the membrane (~78% yielding folding intermediates and ~9% completing folding). The unfolded LacY substrate inserts and folds one structural segment after another until all structural segments S1–S10 fold, thereby reaching the natively folded state. However, the sequence in which the individual segments insert appears to be stochastic. After 5 s, ~9% of all substrates have completed folding into the native LacY structure. Probabilities shown are from refolding experiments conducted with 5 s refolding time (Fig. 4). As previously revealed by dynamic SMFS<sup>23</sup>, the average free energy stabilizing one folded structural segment of LacY is  $-20 k_B T$ , which sums to  $-200 k_B T$  for the entire natively folded LacY.

for refolded LacY increased with the refolding time. Again, the force peaks indicating the folding steps occurred stochastically for all refolding times (Fig. 4a and Supplementary Fig. 10b), indicating that LacY inserted and folded structural segments in variable sequences. However, in the presence of YidC the force peaks of the refolded structural segments were at the same positions as in native LacY. The probability of detecting LacY with correctly folded segments increased with refolding time from ~70% to ~87%, while the probability of misfolding increased from ~6% to ~13% (Fig. 4b,c). No unfolding was detected at 5 s, and with increasing refolding times the number of insertion and folding steps increased until only force peaks typical of native LacY were observed (Fig. 4d). Thus, refolding of structural segments dominated over misfolding. Finally, ~9% of these refolding experiments showed all the force peaks observed in the fingerprint spectrum of native LacY. In summary, the refolding experiments recorded in the presence of YidC showed that the stepwise insertion of structural segments proceeded until LacY fully refolded into the native structure.

## DISCUSSION

We developed an SMFS-based assay to study how YidC assists the insertion and folding of LacY into a phospholipid membrane under

physiologically relevant conditions. The assay, which allows characterization of insertion and folding of individual structural segments within single membrane proteins, contrasts with biophysical experiments carried out in bulk that refold polypeptides from a chaotrope-denatured state into detergent micelles or phospholipid bilayers<sup>42,43</sup>. Importantly, the method can be readily applied to study the folding of a wide range of membrane proteins.

With our SMFS-based assay, we observed that partially unfolded LacY can self-insert and fold structural segments into a phospholipid membrane (Fig. 5). The process is prone to misfolding, and with increasing refolding time misfolding dominated over the many folding steps required to complete the operation. Consequently, we did not observe the complete refolding of LacY in the absence of YidC. Given that the number of LacY substrates showing correctly folded structural segments decreased with folding time while misfolding increased, it is unlikely that misfolded LacY substrates can transform into natively folded proteins. YidC protected LacY from misfolding by stabilizing the unfolded substrate and promoting stepwise insertion and folding of  $\alpha$ -helix-containing segments into the membrane. These insertion and folding steps proceeded until the fingerprint spectrum of native LacY was obtained. Surprisingly, the folding steps assisted by YidC did not occur in a specific order. The results highlight that YidC acts as a chaperone, stabilizing unfolded LacY by suppressing misfolding, as well as an insertase that facilitates stepwise insertion and folding of structural segments into the membrane until LacY completes folding.

The phospholipid bilayer is equivalent to a free-energy sink that determines how unfolded LacY substrates prevented from misfolding by YidC insert and fold (Fig. 5). Once a structural segment is inserted into the lipid bilayer, it is thermodynamically stable and remains folded (Supplementary Fig. 3). Such long-term stability has been demonstrated for fragments of LacY and of other membrane proteins<sup>10,11,17,36,37,44</sup>. Our refolding experiments further show that after the first structural segment has been inserted, the remaining stretches of the polypeptide are maintained in an unfolded state by YidC until the entire LacY substrate completes folding. Via this unique coordination of insertion and folding steps, YidC and the membrane prevent misfolded species and promote folding (Fig. 5). Energetically, perhaps YidC protects structural regions of LacY (for example,  $\alpha$ -helices VII, IX and X) by increasing the free-energy barrier to misfolding and lowering the free-energy barrier to folding. The increase in correct refolding with time is a direct reflection of altered free-energy barriers.

It is not precisely clear how YidC prevents LacY from misfolding and supports correct folding. The X-ray structure of YidC suggests that folding of a polypeptide is facilitated by transient interactions with a hydrophilic groove in YidC<sup>22,45</sup>. This groove, which is opened toward the lipid bilayer and the cytoplasmic side<sup>45,46</sup>, contains several conserved hydrophilic residues and the highly conserved Arg366, which generates a positively charged surface<sup>47</sup>. We observed that YidC was particularly important for the correct folding of structural segment S6, containing  $\alpha$ -helix VII. This result suggests that interaction of the positively charged hydrophilic groove of YidC with the negatively charged, weakly hydrophobic  $\alpha$ -helix VII reduced its misfolding propensity considerably. Subsequently, the properly inserted and folded  $\alpha$ -helix VII could establish the salt bridges and hydrogen network needed to stabilize structural segment S4, encompassing  $\alpha$ -helices IX and X. Notably, a weakly hydrophobic  $\alpha$ -helix VII appears to be common among sugar permeases in the MFS<sup>32,40</sup>. This common property suggests that the YidC-assisted insertion and folding of  $\alpha$ -helix VII is of general importance.

Our finding that the stepwise insertion and folding appears to be a stochastic process surprised us, as current models expect translocases like SecYEG to act processively. A sequential insertion model seems logical in spite of the structural details known from translocons<sup>38,48</sup>. This model is currently challenged by models

assuming the stochastic insertion of structural segments, which after insertion anneal to the functional structure of the membrane protein<sup>38,49,50</sup>. Our experiments, in which we observed the insertase YidC to assist the stochastic insertion of structural segments until finally the polypeptide assembled the native fold of LacY, support the latter model.

Insertion of cytoplasmic membrane proteins *in vivo* involves SecYEG and/or YidC<sup>2</sup>. Interestingly, recent structures of YidC reveal a structural arrangement<sup>45</sup> lacking a polypeptide-conducting channel like that seen in SecYEG<sup>3</sup>. Presumably, this structural feature allows YidC to guide the insertion of individual structural segments such as  $\alpha$ -helices and loops via various coexisting folding pathways toward the native LacY structure. Since the YidC insertase is an essential component of the biogenesis pathway for many cytoplasmic membrane proteins, our mechanistic findings likely have general relevance. In this regard, our experiments may be seen as a milestone toward studying the mechanisms by which YidC, SecYEG and the SecYEG–YidC complex insert and fold membrane proteins.

Received 16 April 2016; accepted 14 June 2016;  
published online 5 September 2016

## METHODS

Methods and any associated references are available in the [online version of the paper](#).

## References

- Driessen, A.J. & Nouwen, N. Protein translocation across the bacterial cytoplasmic membrane. *Annu. Rev. Biochem.* **77**, 643–667 (2008).
- Dalbey, R.E., Wang, P. & Kuhn, A. Assembly of bacterial inner membrane proteins. *Annu. Rev. Biochem.* **80**, 161–187 (2011).
- Luirink, J., Yu, Z., Wagner, S. & de Gier, J.W. Biogenesis of inner membrane proteins in *Escherichia coli*. *Biochim. Biophys. Acta* **1817**, 965–976 (2012).
- Stephenson, K. Sec-dependent protein translocation across biological membranes: evolutionary conservation of an essential protein transport pathway (review). *Mol. Membr. Biol.* **22**, 17–28 (2005).
- van Bloois, E. *et al.* The Sec-independent function of *Escherichia coli* YidC is evolutionary-conserved and essential. *J. Biol. Chem.* **280**, 12996–13003 (2005).
- Wang, P. & Dalbey, R.E. Inserting membrane proteins: the YidC/Oxa1/Alb3 machinery in bacteria, mitochondria, and chloroplasts. *Biochim. Biophys. Acta* **1808**, 866–875 (2011).
- Saller, M.J., Wu, Z.C., de Keyser, J. & Driessen, A.J. The YidC/Oxa1/Alb3 protein family: common principles and distinct features. *Biol. Chem.* **393**, 1279–1290 (2012).
- Hartl, F.U., Bracher, A. & Hayer-Hartl, M. Molecular chaperones in protein folding and proteostasis. *Nature* **475**, 324–332 (2011).
- Engel, A. & Gaub, H.E. Structure and mechanics of membrane proteins. *Annu. Rev. Biochem.* **77**, 127–148 (2008).
- Kedrov, A., Janovjak, H., Ziegler, C., Kuhlbrandt, W. & Müller, D.J. Observing folding pathways and kinetics of a single sodium-proton antiporter from *Escherichia coli*. *J. Mol. Biol.* **355**, 2–8 (2006).
- Kessler, M., Gottschalk, K.E., Janovjak, H., Müller, D.J. & Gaub, H.E. Bacteriorhodopsin folds into the membrane against an external force. *J. Mol. Biol.* **357**, 644–654 (2006).
- Damaghi, M., Köster, S., Bippes, C.A., Yildiz, O. & Müller, D.J. One  $\beta$  hairpin follows the other: exploring refolding pathways and kinetics of the transmembrane  $\beta$ -barrel protein OmpG. *Angew. Chem. Int. Edn Engl.* **50**, 7422–7424 (2011).
- Bosshart, P.D. *et al.* The transmembrane protein KpOmpA anchoring the outer membrane of *Klebsiella pneumoniae* unfolds and refolds in response to tensile load. *Structure* **20**, 121–127 (2012).
- Bechtluft, P. *et al.* Direct observation of chaperone-induced changes in a protein folding pathway. *Science* **318**, 1458–1461 (2007).
- Kaiser, C.M. *et al.* Tracking UNC-45 chaperone-myosin interaction with a titin mechanical reporter. *Biophys. J.* **102**, 2212–2219 (2012).
- Nunes, J.M., Mayer-Hartl, M., Hartl, F.U. & Müller, D.J. Action of the Hsp70 chaperone system observed with single proteins. *Nat. Commun.* **6**, 6307 (2015).
- Thoma, J., Burmann, B.M., Hiller, S. & Müller, D.J. Impact of holdase chaperones Skp and SurA on the folding of  $\beta$ -barrel outer-membrane proteins. *Nat. Struct. Mol. Biol.* **22**, 795–802 (2015).
- Marger, M.D. & Saier, M.H. Jr. A major superfamily of transmembrane facilitators that catalyse uniport, symport and antiport. *Trends Biochem. Sci.* **18**, 13–20 (1993).

19. Abramson, J. *et al.* Structure and mechanism of the lactose permease of *Escherichia coli*. *Science* **301**, 610–615 (2003).
20. Guan, L., Mirza, O., Verner, G., Iwata, S. & Kaback, H.R. Structural determination of wild-type lactose permease. *Proc. Natl. Acad. Sci. USA* **104**, 15294–15298 (2007).
21. Nagamori, S., Smirnova, I.N. & Kaback, H.R. Role of YidC in folding of polytopic membrane proteins. *J. Cell Biol.* **165**, 53–62 (2004).
22. Zhu, L., Kaback, H.R. & Dalbey, R.E. YidC protein, a molecular chaperone for LacY protein folding via the SecYEG protein machinery. *J. Biol. Chem.* **288**, 28180–28194 (2013).
23. Serdiuk, T. *et al.* Substrate-induced changes in the structural properties of LacY. *Proc. Natl. Acad. Sci. USA* **111**, E1571–E1580 (2014).
24. Serdiuk, T., Sugihara, J., Mari, S.A., Kaback, H.R. & Müller, D.J. Observing a lipid-dependent alteration in single lactose permeases. *Structure* **23**, 754–761 (2015).
25. Viitanen, P., Newman, M.J., Foster, D.L., Wilson, T.H. & Kaback, H.R. Purification, reconstitution, and characterization of the lac permease of *Escherichia coli*. *Methods Enzymol.* **125**, 429–452 (1986).
26. Herzlinger, D., Viitanen, P., Carrasco, N. & Kaback, H.R. Monoclonal antibodies against the lac carrier protein from *Escherichia coli*. 2. Binding studies with membrane vesicles and proteoliposomes reconstituted with purified lac carrier protein. *Biochemistry* **23**, 3688–3693 (1984).
27. Müller, D.J. & Engel, A. Atomic force microscopy and spectroscopy of native membrane proteins. *Nat. Protoc.* **2**, 2191–2197 (2007).
28. Kedrov, A., Krieg, M., Ziegler, C., Kuhlbrandt, W. & Müller, D.J. Locating ligand binding and activation of a single antiporter. *EMBO Rep.* **6**, 668–674 (2005).
29. Sapra, K.T., Besir, H., Oesterhelt, D. & Müller, D.J. Characterizing molecular interactions in different bacteriorhodopsin assemblies by single-molecule force spectroscopy. *J. Mol. Biol.* **355**, 640–650 (2006).
30. Ge, L., Perez, C., Waclawska, I., Ziegler, C. & Müller, D.J. Locating an extracellular K<sup>+</sup>-dependent interaction site that modulates betaine-binding of the Na<sup>+</sup>-coupled betaine symporter BetP. *Proc. Natl. Acad. Sci. USA* **108**, E890–E898 (2011).
31. Bippes, C.A. *et al.* Peptide transporter DtpA has two alternate conformations, one of which is promoted by inhibitor binding. *Proc. Natl. Acad. Sci. USA* **110**, E3978–E3986 (2013).
32. Dowhan, W. & Bogdanov, M. Lipid-dependent membrane protein topogenesis. *Annu. Rev. Biochem.* **78**, 515–540 (2009).
33. Oesterhelt, F. *et al.* Unfolding pathways of individual bacteriorhodopsins. *Science* **288**, 143–146 (2000).
34. Fanelli, F. & Seeber, M. Structural insights into retinitis pigmentosa from unfolding simulations of rhodopsin mutants. *FASEB J.* **24**, 3196–3209 (2010).
35. Kappel, C. & Grubmüller, H. Velocity-dependent mechanical unfolding of bacteriorhodopsin is governed by a dynamic interaction network. *Biophys. J.* **100**, 1109–1119 (2011).
36. Bibi, E. & Kaback, H.R. In vivo expression of the lacY gene in two segments leads to functional lac permease. *Proc. Natl. Acad. Sci. USA* **87**, 4325–4329 (1990).
37. Zen, K.H., McKenna, E., Bibi, E., Hardy, D. & Kaback, H.R. Expression of lactose permease in contiguous fragments as a probe for membrane-spanning domains. *Biochemistry* **33**, 8198–8206 (1994).
38. Cymer, F., von Heijne, G. & White, S.H. Mechanisms of integral membrane protein insertion and folding. *J. Mol. Biol.* **427**, 999–1022 (2015).
39. Müller, D.J. *et al.* Observing membrane protein diffusion at subnanometer resolution. *J. Mol. Biol.* **327**, 925–930 (2003).
40. Bogdanov, M., Xie, J., Heacock, P. & Dowhan, W. To flip or not to flip: lipid-protein charge interactions are a determinant of final membrane protein topology. *J. Cell Biol.* **182**, 925–935 (2008).
41. Guan, L. & Kaback, H.R. Properties of a LacY efflux mutant. *Biochemistry* **48**, 9250–9255 (2009).
42. Bowie, J.U. Solving the membrane protein folding problem. *Nature* **438**, 581–589 (2005).
43. Harris, N.J. & Booth, P.J. Folding and stability of membrane transport proteins in vitro. *Biochim. Biophys. Acta* **1818**, 1055–1066 (2012).
44. Engelman, D.M. *et al.* Membrane protein folding: beyond the two stage model. *FEBS Lett.* **555**, 122–125 (2003).
45. Kumazaki, K. *et al.* Structural basis of Sec-independent membrane protein insertion by YidC. *Nature* **509**, 516–520 (2014).
46. Shimokawa-Chiba, N. *et al.* Hydrophilic microenvironment required for the channel-independent insertase function of YidC protein. *Proc. Natl. Acad. Sci. USA* **112**, 5063–5068 (2015).
47. Kumazaki, K. *et al.* Crystal structure of *Escherichia coli* YidC, a membrane protein chaperone and insertase. *Sci. Rep.* **4**, 7299 (2014).
48. Rapoport, T.A., Goder, V., Heinrich, S.U. & Matlack, K.E. Membrane-protein integration and the role of the translocation channel. *Trends Cell Biol.* **14**, 568–575 (2004).
49. Van Lehn, R.C., Zhang, B. & Miller, T.F., III. Regulation of multispanning membrane protein topology via post-translational annealing. *eLife* **4**, e08697 (2015).
50. White, S.H. The messy process of guiding proteins into membranes. *eLife* **4**, e12100 (2015).

### Acknowledgments

We thank R.E. Dalbey (The Ohio State University) for providing plasmid pT7-7 encoding YidC with a His<sub>10</sub> tag at the C terminus; R. Newton, J. Thoma, S. Hiller and R.E. Dalbey for encouraging and constructive comments; and S. Weiser for assistance. This work was supported by the Eidgenössische Technische Hochschule Zürich (to D.J.M.), the Swiss National Science Foundation (grant 205320\_160199 to D.J.M.), the Swiss National Center of Competence in Research “NCCR Molecular Systems Engineering” (to D.J.M.), the European Union Marie Curie Actions program through the ACRITAS Initial Training Network (FP7-PEOPLE-2012-ITN, project 317348 to D.J.M.), U.S. National Institutes of Health grants DK51131, DK069463 and GM073210 (to H.R.K.), and U.S. National Science Foundation grant MCB-1129551 (to H.R.K.).

### Author contributions

T.S., D.J.M. and H.R.K. designed the experiments. J.S. and D.B. cloned, expressed, purified and reconstituted LacY and YidC. T.S. performed the SMFS experiments. S.A.M. recorded AFM images. All authors analyzed experimental data and wrote the paper.

### Competing financial interests

The authors declare no competing financial interests.

### Additional information

Any supplementary information, chemical compound information and source data are available in the [online version of the paper](http://www.nature.com/reprints/index.html). Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>. Correspondence and requests for materials should be addressed to H.R.K. or D.J.M.

## ONLINE METHODS

**LacY and YidC engineering, expression and purification.** *LacY*. A 36 amino acids (aa) long ‘poly(Gly)’ peptide followed by a His<sub>8</sub>-tag extension [GSM(G<sub>11</sub>)EAVEEAVEEA(Gly<sub>11</sub>)S(His<sub>8</sub>)] was added to the LacY C terminus using QuikChange® II PCR and plasmid pT7-5/LacY as template. Both poly(Gly) LacY and WT LacY were purified from *E. coli* XL1-Blue (Stratagene) transformed with pT7-5 plasmids harboring given mutant genes by using Co(II) affinity chromatography as prescribed<sup>31</sup>. LacY eluted from the Co(II)-Talon column was concentrated and washed with 50 mM sodium phosphate (NaP<sub>i</sub>), pH 7.5 and 0.01% (w/w) dodecyl-β-D-maltopyranoside (DDM, Maumee) on an Amicon Ultra-15 concentrator with a 30 kDa cut-off (Millipore).

*YidC*. Expression and purification of WT YidC with a His<sub>10</sub>-tag cloned in pT7-7 plasmid was performed similarly to LacY except that the expression strain used was *E. coli* BL21 (DE3) and the DDM concentration used was 0.03%. YidC eluted from the Co(II)-Talon column was concentrated and washed with 150 mM NaCl, 50 mM sodium phosphate (NaP<sub>i</sub>), pH 7.5 and 0.03% (w/w) DDM on an Amicon Ultra-15 concentrator with a 30 kDa cut-off. All protein preparations were at least 95% pure as judged by staining after sodium dodecyl sulfate polyacrylamide (SDS) gel electrophoresis (Supplementary Fig. 1).

The reconstitution of LacY and YidC into proteoliposomes containing 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoethanolamine (PE, Avanti Polar Lipids) and 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phospho-rac-(1-glycerol) (PG, Avanti Polar Lipids) (ratio 3:1) was accomplished by using the dilution method<sup>25</sup>. Briefly, LacY and YidC were mixed at equimolar (1:1) concentration and incubated for 30 min at room temperature before reconstitution. Protein concentrations of LacY, YidC (for the reconstitution of LacY or YidC) or the protein mixture of LacY and YidC (for the co-reconstitution of LacY and YidC) were dissolved in 0.01% DDM solution, then mixed with phospholipids dissolved in 1.2% octyl glucoside (OG, Maumee) to yield a lipid-to-protein ratio of 5 (w/w). The mixture was kept on ice for 20 min and then quickly diluted 50-fold in 50 mM NaP<sub>i</sub>, pH 7.5. Proteoliposomes were harvested by centrifugation for 1 h at 100,000g, suspended in the same buffer and flash frozen in liquid nitrogen.

**Single-molecule unfolding and refolding experiments.** Proteoliposomes containing WT LacY reconstituted in PE/PG (ratio 3:1), YidC reconstituted in PE/PG (3:1), or WT LacY co-reconstituted with YidC in PE/PG (3:1) were adsorbed to freshly cleaved mica in 50 mM potassium phosphate (KPI), pH 7.2 for 20 min. The samples were then gently rinsed with the same buffer to remove weakly attached proteoliposomes. For all AFM-based experiments we used the same AFM (Nanowizard Ultra, JPK Instruments AG) at room temperature and in buffer solution. Si<sub>3</sub>N<sub>4</sub> cantilevers (OMCL RC800PSA, Olympus) having a nominal spring constant of 0.05 N m<sup>-1</sup> were used for all measurements. Cantilevers were calibrated in SMFS buffer applying the equipartition theorem at the beginning and at the end of each experiment. Before SMFS, the reconstituted membrane proteins were imaged by AFM (Supplementary Fig. 1)<sup>27</sup>. These AFM images showed the proteins distributed evenly over the surface of the membrane. On average, the reconstitutions of LacY or/and YidC showed approximately the same low density of membrane proteins. To enhance the probability of attaching LacY molecules to the AFM stylus, we defined a grid (similar to that shown in Supplementary Fig. 8) of thousands of pixels distributed over the protein membrane. For each topographic pixel one or several force–distance curves were recorded. To record a force–distance curve, the AFM stylus was pushed to the membrane applying ~700 pN for 500 ms. To unfold LacY or YidC completely the cantilever was then withdrawn 0.7 μm s<sup>-1</sup> while recording a force–distance curve. The distance traveled by the AFM stylus was chosen to be sufficient to fully unfold and extract the polypeptide from the membrane. For refolding, the cantilever was retracted for 100 nm at 0.7 μm s<sup>-1</sup> to partially unfold LacY while recording a force–distance curve. The cantilever was then re-approached toward the membrane at 1 μm s<sup>-1</sup> and kept 10–15 nm above the membrane surface to relax the unfolded polypeptide chain. After controlled refolding times of 0.1–5 s, which allowed the unfolded polypeptide to reinsert into the membrane, the cantilever was retracted for 250 nm at 0.7 μm s<sup>-1</sup> to completely unfold the protein.

The probability of nonspecifically attaching the AFM stylus to the ‘poly(Gly)’ elongated C-terminal end of LacY was ~0.1% ( $n = 2,974$ ), which was ~10-fold higher than the probability of attaching the non-elongated C-terminal end<sup>23</sup>. This attachment is temporary and the polypeptide slipped from the AFM stylus after several seconds<sup>27,33</sup>. Thus, within a refolding time of 2 s the C-terminal end slipped from the AFM stylus in ~91% of all cases ( $n = 1,160$ ). Together the low attachment rate and the high slip off rate considerably lowered the throughput of the single-molecule refolding experiments. Therefore, in a normal experimental day we could record only from one to four successful refolding experiments. On average, recording more than 100 refolding experiments required more than 47 experimental days. For each experimental day we had to prepare a new LacY sample and AFM cantilever. Taken together several hundred fresh samples and AFM cantilevers were prepared to acquire the high number of force–distance curves and refolding experiments presented in the manuscript.

**SMFS data selection and analysis.** The data recorded in several hundred experiments (see details above) were pooled to obtain statistically sufficient amounts of data for each experimental condition. Every force–distance curve of the pooled data was corrected for deflection sensitivity and force offset by applying the standard AFM data processing software (JPK Instruments, software version 5.0.12). A fully unfolded and stretched LacY polypeptide (417 aa extended by a 36 aa long poly(Gly) tail and a His<sub>8</sub>-tag) has a contour length 461 aa and exhibits force peak patterns extending to a pulling distance of ~120 nm. To ensure that only force–distance curves recording the full unfolding of LacY from mechanically pulling the terminal end were analyzed, we selected force–distance curves showing force peak patterns extending to >110 nm.

A fully unfolded and stretched YidC polypeptide has a contour length 561 aa (551 aa extended by an His<sub>10</sub>-tag) and exhibits force peak patterns extending to ~150 nm. To ensure that only force–distance curves recording the full unfolding of YidC from mechanically pulling one terminal end were analyzed, we selected curves showing force peak patterns extending to >130 nm.

Every force peak of a force–distance curve corresponding to unfolding of LacY or YidC was fitted by the worm-like-chain (WLC) model using a persistence length of 0.4 nm and contour length of 0.36 nm per aa<sup>27,52</sup>. Each WLC gives the contour length (number of aa) of the polypeptide unfolded and stretched in an unfolding force peak. Histograms were generated showing how frequently force peaks were detected at given contour lengths. To calculate the frequency of the occurrence of force peaks at a given contour length the counts of force peaks were divided through the total number of force–distance curves analyzed. These histograms were fitted with the Gaussian mixture model<sup>53</sup>. Accordingly to this model, each  $i$ -th observed contour length  $l_i$  originates from one force peak class  $s = 1, \dots, M$  with probability  $\pi_s$ , or originates from background noise with probability  $\pi_0$ . We found 10 force peak classes for native LacY and eight force peak classes for YidC (Fig. 1 and Supplementary Fig. 7). The mean contour length  $\mu_s$  for a given force peak class  $s$  was described by a Gaussian distribution with variance  $\sigma_s^2$ . The probability density  $f$  of  $l_i$  can be presented as a mixture of Gaussians and background noise with weights  $\pi_s$  and  $\pi_0$  respectively:

$$f(l_i) = \sum_{s=1}^M \pi_s \phi(l_i, \mu_s, \sigma_s^2) + \pi_0 g(l_i) = \sum_{s=1}^M \pi_s \phi(l_i, \mu_s, \sigma_s^2) + \pi_0 g(l_i)$$

$\phi(l_i, \mu_s, \sigma_s^2)$  is the probability density of the Gaussian distribution and  $g(l_i)$  the background noise. The expectation optimization algorithm was applied to find parameters of the Gaussian mixture model ( $\pi, \mu, \sigma^2$ ). The most probable force peak class  $s_i$  was assigned to any given contour length  $l_i$  with the Bayes classifier by setting:

$$s_i = \text{argmax}_s (\pi_s \phi(l_i, \mu_s, \sigma_s^2) / \pi_0 g(l_i))$$

Every detected contour length was assigned to a force peak class, each of which was mapped to the secondary structure of LacY<sup>23</sup>. The length of the C-terminal poly(Gly) and His<sub>8</sub>-tag extension was taken into account for mapping the force peak classes. If the force peak class located the beginning or end of a stabilizing structural segment on the mica-facing side of the membrane or within the membrane, the thickness of the membrane was taken into account<sup>27</sup>.

**Control of YidC reconstitution.** To show the successful reconstitution of YidC in PE/PG membranes SMFS unfolding experiments were performed. Eight different membrane patches were characterized by SMFS. Recorded force–distance curves were superimposed revealing the unique force peak pattern of YidC (**Supplementary Fig. 7**). This force peak pattern served as the native fingerprint spectrum of YidC.

**Control of LacY and YidC co-reconstitution.** To show the successful co-reconstitution of LacY and YidC in PE/PG membranes SMFS unfolding experiments were performed. Twelve different membrane patches were imaged by AFM (**Supplementary Fig. 1**) and then characterized by SMFS. Recorded force–distance curves were classified in two classes and superimposed to show their common features. One class of force–distance curves revealed the native fingerprint spectrum recorded for WT LacY (**Fig. 1** and **Supplementary Fig. 6**) while the other class of force–distance curves showed the native fingerprint spectrum recorded for YidC (**Supplementary Figs. 6** and **7**). The ratio of force–distance curves representing the unfolding of LacY or YidC approached 1:1.

**Refolding experiments, data analysis and classification.** Unfolding and extraction of the last stable segment of LacY occurred at pulling distances >100 nm (force peak at 404 aa, **Fig. 1**); thus we stopped initial unfolding of LacY at distances of 100 nm and brought the unfolded polypeptide back into close proximity (~10 nm) to the membrane surface where the polypeptide was held in the relaxed state. After a defined relaxation time in which the unfolded protein was allowed to refold, the AFM stylus was fully withdrawn to unfold LacY completely. Force–distance curves recorded in these refolding experiments were analyzed as follows: The force–distance curve recording the initial partial unfolding of LacY was horizontally shifted to match the native fingerprint spectrum of WT LacY (**Fig. 1b,c**). If this initial unfolding force–distance curve did not match the native fingerprint spectrum the experiment was discarded. After this, the same horizontal shift that we applied to fit the initial force–distance curve to the native fingerprint spectrum was applied to the second unfolding force–distance curve recorded from the refolded LacY substrate. The force peaks in both unfolding force–distance curves were then fitted using the WLC model to determine their contour lengths and classified as having remained unfolded, having misfolded, or having inserted and folded structural segments. The classification criteria have been described in the Results and **Supplementary Figure 4**. Briefly, a force–distance curve was deemed to represent a LacY substrate remaining unfolded if no interactions were recovered after the refolding time passed. A force–distance curve was deemed to represent

a LacY substrate having misfolded if it detected at least one force peak at a position not matching the mean  $\pm$  s.d. of a force peak position detected in the native unfolding force peak pattern of WT LacY (**Fig. 1b,c**). A force–distance curve was deemed to represent a LacY having folded one or more structural segments if it detected at least one force peak at a position matching the mean  $\pm$  s.d. of a force peak position detected in the native unfolding force peak pattern of WT LacY (**Fig. 1**). Force–distance curves detecting both refolding and misfolding events were deemed to represent a misfolded LacY. To test whether nonspecific interactions play a role in determining the outcome of the refolding experiments, WT LacY was refolded as described above but in the presence of either 1  $\mu$ M BSA (Sigma-Aldrich) or 1  $\mu$ M lysozyme (Fluka Analytical). To examine the role of YidC in the outcome of the refolding experiments, WT LacY co-reconstituted with YidC was refolded as described above. All refolding experiments were analyzed following the same procedure and classified applying the same criteria described above.

**Statistical data analysis.** Approximately the same sample size of ~100 single refolding experiments was taken for each tested group (WT LacY, WT LacY & YidC, WT LacY & BSA, WT LacY & lysozyme). A sample size of ~100 provides a standard error of  $\leq 0.05$  (5%) for each of the three data classification categories, unfolded (U), misfolding (M), and folding (F) (**Figs. 3** and **4**). In order to test the significance of the difference between tested groups (WT LacY, WT LacY & YidC, WT LacY & BSA, WT LacY & lysozyme), two-tailed Z-tests, two-tailed Fisher's exact tests and Chi-square tests were performed for each refolding time and all categories (**Supplementary Tables 1–4**). Tested samples were independent and the data were categorical. For Fisher's exact tests and Chi-square tests  $2 \times 3$  contingency tables were analyzed (tested groups (WT LacY & YidC, WT LacY & BSA or WT LacY & lysozyme) vs control (WT LacY with possible outcomes U, M, and F)). There were no interactions between row and column classifications. The differences between categories/groups were considered not significant (NS) when  $P \geq 0.01$ , significant when  $*P < 0.01$  and highly significant when  $**P < 0.001$ . Error bars (s.e.m.) shown in **Figures 3** and **4** were estimated by resampling using 4,000 iterations. Statistical analysis was performed using R.

51. Smirnova, I. *et al.* Sugar binding induces an outward facing conformation of LacY. *Proc. Natl. Acad. Sci. USA* **104**, 16504–16509 (2007).
52. Bustamante, C., Marko, J.F., Siggia, E.D. & Smith, S. Entropic elasticity of lambda-phage DNA. *Science* **265**, 1599–1600 (1994).
53. Kawamura, S. *et al.* Kinetic, energetic, and mechanical differences between dark-state rhodopsin and opsin. *Structure* **21**, 426–437 (2013).