doi:10.1016/j.jmb.2012.02.043 J. Mol. Biol. (2012) 419, 89–101

Contents lists available at www.sciencedirect.com

Journal of Molecular Biology

journal homepage: http://ees.elsevier.com.jmb

Engineered Oligomerization State of OmpF Protein through Computational Design Decouples Oligomer Dissociation from Unfolding

Hammad Naveed [1], David Jimenez-Morales [1], Jun Tian [1], Volga Pasupuleti [1], Linda J. Kenney [2,3]* and Jie Liang [1]*

[1]*Department of Bioengineering, University of Illinois at Chicago, Chicago, IL 60607, USA*
[2]*Department of Microbiology and Immunology, University of Illinois at Chicago, Chicago, IL 60612, USA*
[3]*Mechanobiology Institute, National University of Singapore, Singapore 117411, Singapore*

Received 17 January 2012; received in revised form 24 February 2012; accepted 25 February 2012
Available online 3 March 2012

Edited by J. Bowie

Keywords:
porins;
β-barrel membrane proteins;
evolution;
membrane protein–protein interaction;
weakly stable regions

Biogenesis of β-barrel membrane proteins is a complex, multistep, and as yet incompletely characterized process. The bacterial porin family is perhaps the best-studied protein family among β-barrel membrane proteins that allows diffusion of small solutes across the bacterial outer membrane. In this study, we have identified residues that contribute significantly to the protein–protein interaction (PPI) interface between the chains of outer membrane protein F (OmpF), a trimeric porin, using an empirical energy function in conjunction with an evolutionary analysis. By replacing these residues through site-directed mutagenesis either with energetically favorable residues or substitutions that do not occur in natural bacterial outer membrane proteins, we succeeded in engineering OmpF mutants with dimeric and monomeric oligomerization states instead of a trimeric oligomerization state. Moreover, our results suggest that the oligomerization of OmpF proceeds through a series of interactions involving two distinct regions of the extensive PPI interface: two monomers interact to form a dimer through the PPI interface near G19. This dimer then interacts with another monomer through the PPI interface near G135 to form a trimer. We have found that perturbing the PPI interface near G19 results in the formation of the monomeric OmpF only. Thermal denaturation of the designed dimeric OmpF mutant suggests that oligomer dissociation can be separated from the process of protein unfolding. Furthermore, the conserved site near G57 and G59 is important for the PPI interface and might provide the essential scaffold for PPIs.

Introduction

Recent studies estimate that membrane proteins form 20–30% of all proteins in a genome,[1,2] yet they are sparsely represented in the protein structure data bank.[3] This is in part due to difficulties in the experimental determination of their three-dimensional structures.[4] Among the two classes of membrane proteins, β-barrel membrane proteins are found in the outer membrane of Gram-negative bacteria, mitochondria, and chloroplasts. β-Barrel

**Corresponding authors.* J. Liang is to be contacted at Department of Bioengineering, University of Illinois at Chicago, Chicago IL 60607, USA; L. J. Kenney, Department of Microbiology and Immunology, University of Illinois at Chicago, Chicago, IL 60612, USA.
E-mail addresses: kenneyl@uic.edu; jliang@uic.edu.

Abbreviations used: PPI, protein–protein interaction; OmpF, outer membrane protein F; TM, transmembrane; DPBS, Dulbecco's phosphate-buffered saline.

0022-2836/$ - see front matter © 2012 Published by Elsevier Ltd.

membrane proteins are known to participate in membrane anchoring, pore formation, enzyme activity, and bacterial virulence.[5–7] Extensive studies have been conducted to characterize their thermodynamic stability,[8,9] folding kinetics,[10–13] and biological functions.[14,15]

Newly synthesized β-barrel membrane proteins in Gram-negative bacteria must cross the cytoplasm, inner membrane, and periplasmic space to reach the outer membrane.[16] This process is facilitated by a machinery consisting of a number of biological complexes, including translocons (SecA),[17] chaperones (Skp, DegP, and SurA),[18,19] and transporters (BAM complex).[20] Folding and membrane insertion of β-barrel membrane proteins are thought to occur simultaneously, as insertion of individual β-strands is energetically unfavorable.[21,22] Moreover, in oligomeric β-barrels, quaternary structure formation may also be coupled with the process of folding and insertion.[23] Overall, the biogenesis of β-barrel membrane proteins is not yet well understood, as these coupling events are difficult to deconvolute.

The bacterial porin family is perhaps the best-studied protein family among β-barrel membrane proteins. Biologically, they allow diffusion of small solutes across the bacterial outer membrane. Porins can be divided into "general porins" and "specific porins." General porins, such as outer membrane protein F (OmpF), filter solutes based on their molecular weight. Specific porins, such as sucrose-specific porin (ScrY), have specific binding sites for certain solutes. Some porins form obligatory homotrimeric biological units, with significant in-plug domains in the interior of the barrel.[24]

OmpF is a bacterial porin from *Escherichia coli* whose native oligomerization state is thought to be trimeric.[24] However, dimeric structures have also been observed in both *in vitro* and *in vivo* experiments.[23,25–27] Similarly, despite the observed trimeric form of porin PhoE, a functional monomeric form of PhoE has been reported in *in vitro* and *in vivo* studies.[28,29] Although the protein–protein interaction (PPI) site for porins is known in many cases, the significance of preferring a particular oligomeric state over others is not clear.[30] Understanding the factors that determine the oligomerization state of these proteins will advance our understanding of their biogenesis and function.

The existence of an extensive regular hydrogen-bond network between transmembrane (TM) β-strands is thought to confer extreme stability on β-barrel membrane proteins.[31] However, recent studies showed the existence of weakly stable regions in the TM domains of these proteins.[9,32] These regions are found to be stabilized by four general mechanisms: (1) integration of small β-helices and β-strands, called in-plugs, inside the β-barrel;[33] (2) packing of non-barrel-embedded helices against TM β-strands;[34] (3) specific lipid binding, such as the binding of lipopolysaccharide molecules in FhuA, also stabilizes the protein;[35] and (4) multiple weakly stable regions on separate proteins may form interfaces that facilitate stabilization of PPIs.[33] These structurally stabilizing mechanisms often play important functional roles in β-barrel membrane proteins, including voltage sensing,[36] flux control of metabolites, and ion sensing.[33]

A complementary strategy to the study of stability constraints underlying β-barrel membrane proteins is characterizing their evolutionary patterns. The degree of sequence conservation often correlates with the importance of a particular position of residues in a protein. Slow protein evolution is a consequence of strong purifying selection, which varies among different proteins or even among different regions within the same protein, due to stability or functional constraints. The patterns of amino acid substitutions at different interfaces of the TM domains of β-barrel membrane proteins were recently characterized.[37] Compared with the barrel interior, twice as many amino acid substitutions were found to occur at the lipid-facing interface. However, the substitution pattern at the lipid interface is very narrow, suggesting that there are specific physical forces exerting a strong selection pressure on this interface for amino acids to maintain the same physical–chemical properties. An interesting possibility is whether the stability and oligomeric properties of the β-barrel membrane proteins can be altered by selecting different amino acid substitutions based on the estimated evolutionary patterns of substitutions.[37]

In this study, we explore strategies to reengineer the PPI interface in the TM domain of the β-barrel membrane protein OmpF. Several recent studies have focused on designing new PPI interfaces through computational protein reengineering and computational *de novo* interface design.[38–41] We have redesigned the PPI interface of OmpF by identifying weakly stable residues and by replacing them with energetically more favorable residues. Our goal was to obtain stable dimeric and monomeric forms of OmpF. A recent study showed that the elimination of weakly stable regions led to increased resistance of Tom40 protein (a mitochondrial β-barrel) to thermal and chemical denaturation.[9] In addition, by identifying a conserved region in the PPI interface of OmpF and by substituting the conserved residues following a substitution pattern different from that observed in β-barrel membrane proteins, we may obtain stable monomeric forms of β-barrel membrane proteins. Mutants engineered through site-directed mutagenesis following our design indeed showed stable dimeric and monomeric oligomerization states. We have explored further possible mechanisms of OmpF oligomerization through studies of the denaturation of a designed dimeric mutant. Our results suggest that oligomerization

occurs in a stepwise manner and that oligomer disassociation can be separated from protein unfolding.

Results

We have used two computational approaches to identify residues that contribute significantly to the PPI interface of OmpF. The first approach identified weakly stable residues in the PPI interface. We replaced these weakly stable residues with energetically more favorable residues. The second approach identified a conserved patch in the PPI, and we replaced the key residues with substitutions not found in natural β-barrel membrane proteins. We first describe the results of both computational methods. We then experimentally assess the contributions of the residues identified by these computational methods to the PPI interface of OmpF.

Detecting weakly stable regions by empirical potential function

OmpF exists predominantly as a trimer, although stable dimeric forms have been observed *in vitro* and *in vivo*.[23,25–27] After summing up the energetic contribution of individual residues in the TM domain according to the empirical potential function TmSIP, we found that the eight strands with the highest overall energies are strands 1–6 and 15–16. These are predicted as weakly stable strands. The observed oligomeric interface of OmpF in the crystal structure consists of strands 1–5 and 16. The weakly stable strands detected by the empirical potential function largely coincide with the oligomeric interface of OmpF.

We have examined the energetic contributions of individual amino acids facing the lipids in the high-energy strands of the TM domain (Fig. 1a). The stability of the residues is based on their depth in the bilayer and side-chain orientation. We find that R100 is the most unstable residue in the PPI interface, along with G19, G135, and N141 (Fig. 1a). R100 interacts with the loop between strands 3 and 4 from another chain; this loop between strands 3 and 4 is thought to be the major contributor to interchain interactions,[33] while G19 is located in strand 1, and G135 and N141 are located in strand 6 (Fig. 1b). Based on these calculations, we designed six mutants (G19W, R100L, G135W, N141W, R100L/G135W, and R100L/G19W) to verify the

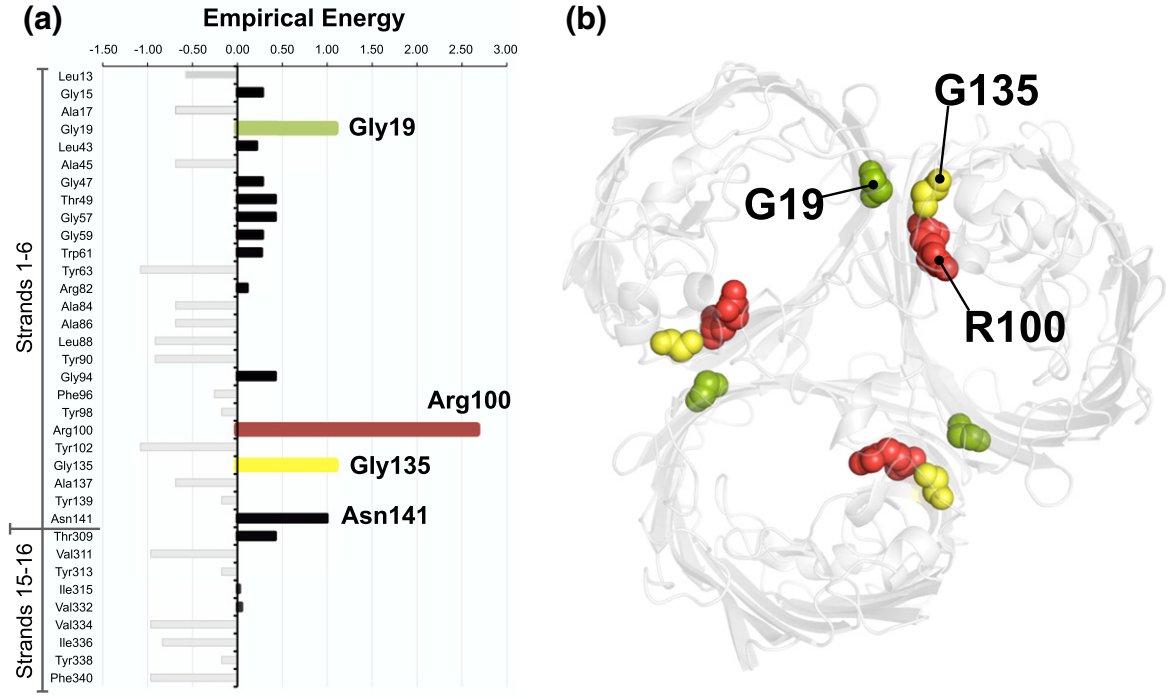

Fig. 1. The energetic contributions of individual residues in the TM domain of OmpF. OmpF exists in a trimeric state, with strands 1–5 and 16 forming the interaction surface. Among the 16 strands, strands 1–6 and 15–16 are found to have the highest overall energy values based on the empirical potential function TmSIP. These strands largely coincide with the PPI interface. (a) The empirical energy profile of all residues facing the lipids in strands 1–6 and 15–16. (b) Residues R100, G19, G135, and N141 are unstable compared to the rest of the residues. R100 interacts with the loop between strands 3 and 4 from another chain, which is thought to be a major contributor to interchain interactions.[33] G19, G135, and N141 (data not shown) are located in strands 1, 6, and 6, respectively.

contribution of these high-energy residues to the PPI interface. These residues were replaced by the most stable residue in the specific region of the protein, as predicted by the empirical potential function TmSIP.[42]

Evolutionarily conserved residues in the PPI interface of OmpF

The overall picture of the residue conservation of OmpF is in agreement with a previous study.[37] The out-facing residues located at the protein–protein interface are found to be better conserved than those facing the lipids. Among them, a group of well-conserved residues is found to form a structurally contiguous surface patch (Fig. 2a), including the C-terminal end of the second strand (Gly47 and Thr49) and the N-terminal end of the third strand (Leu55-Gly59) (Fig. 2b and c).

To verify the roles of this well-conserved surface patch in maintaining OmpF stability and promoting OmpF oligomerization, we designed three mutants: G57A/G59A, G57S/G59S, and G57I/G59L. These substitutions were chosen based on an analysis of the evolutionary pattern of amino acid substitutions at the TM segments of β-barrel membrane proteins.[37] According to the previous study, glycine is well conserved in the out-facing lipid interface. The few substitutions detected for glycine were replacements with serine, threonine, and alanine. Long aliphatic and hydrophobic residues (L, V, I, and A) were found to readily substitute among themselves in this interface, but not with glycine.

With these observed patterns of substitutions, we hypothesized that substitution of the glycines located in the well-conserved surface patch (i.e., G57 and G59) with leucine, valine, or isoleucine would affect the stability of the OmpF trimer. In

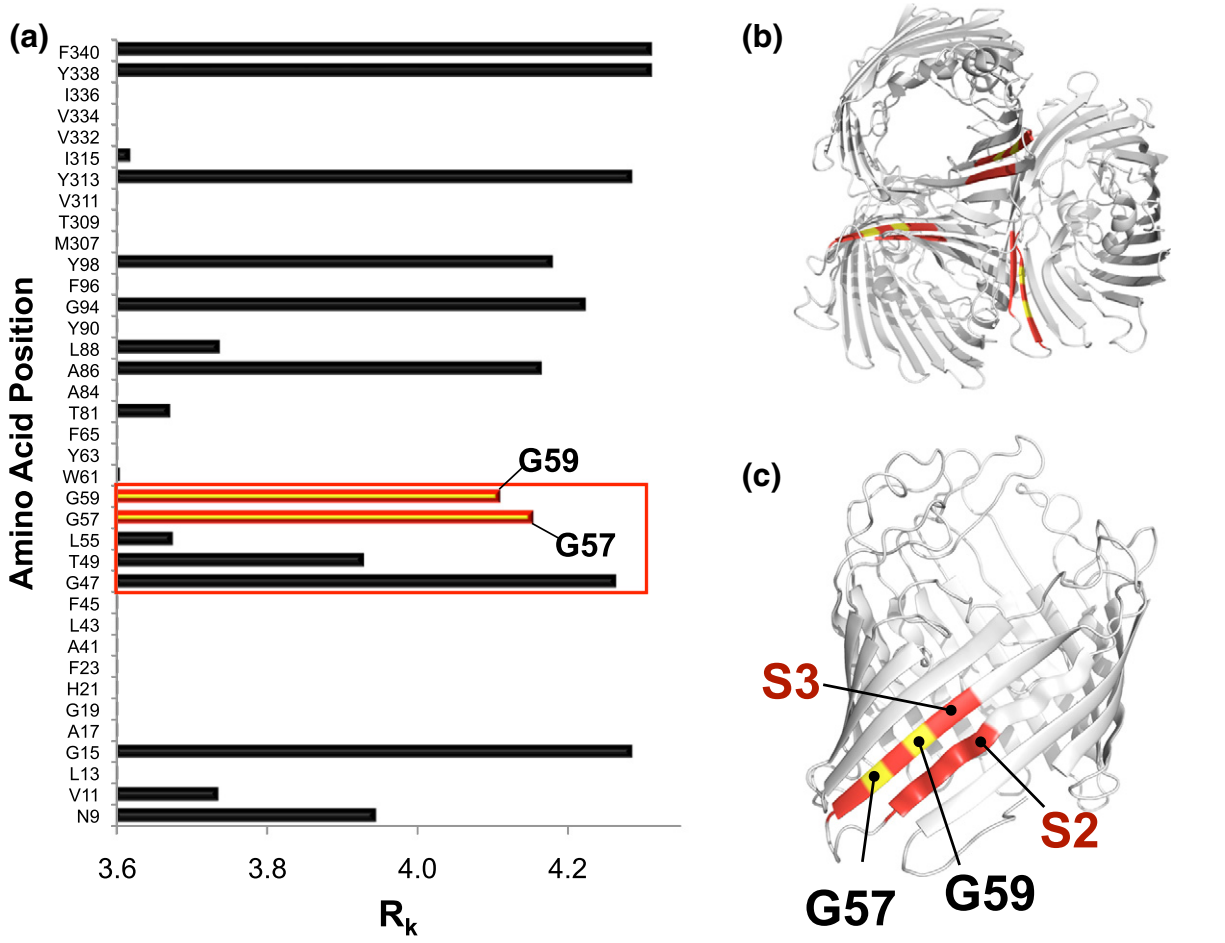

Fig. 2. Evolutionarily conserved positions of amino acids are found at the protein–protein interface of OmpF. (a) Conservation entropy (in bits) of TM residues. For illustration, we show $R_k = \log_2 20 - H_k$, where well-conserved columns show the longest bars. The red box highlights a well-conserved structurally contiguous surface patch at the oligomerization interface (strands S2 and S3). (b and c) Structural locations of strands S2 and S3 in the trimer and monomer, respectively.

contrast, substitutions of G57 and G59 with serine or alanine would be expected not to have a significant impact on the stability of the trimer of OmpF.

Secondary and tertiary structures of wild-type and mutant OmpF proteins

In order to determine to what extent the predicted residues contribute to the overall oligomerization state, we first assessed the folding of wild-type and mutant OmpF porins. Wild-type and mutant OmpF proteins were expressed in *E. coli* and purified from inclusion bodies under denaturing conditions. Proteins were refolded by dilution of the denaturant into the refolding buffer, as reported by Visudtiphole *et al.*[25] SDS-PAGE and Coomassie blue staining indicated that all of the mutants and the wild-type protein were successfully expressed and purified. CD spectra analyses confirmed that the mutants had the same typical β-barrel spectra as wild-type OmpF, with a peak at around 217 nm (Fig. 3a; Figs. S4–S8). In addition, the protein samples were free of higher-order aggregates, as indicated by the CD spectra, approaching ellipticity values close to 0 at wavelengths of >250 nm. The spectral characteristics of the wild-type and mutant proteins are summarized in Table S2.

The tertiary structure of the wild-type and mutant OmpF proteins was analyzed via tryptophan fluorescence spectroscopy. Wild type, G57A/G59A mutant, G57S/G59S mutant, and G57I/G59L mutant have two tryptophan residues at positions 61 and 214. The emission spectra of the wild type and of mutants G57A/G59A, G57S/G59S, and G57I/G59L were similar (Fig. 3b), with intensity maxima at approximately 318 nm and with an unchanged width of emission spectra. The mutants R100L/G19W and R100L/G135W have an extra tryptophan at position 19 or 135, respectively, in addition to positions 61 and 214. The intensity maxima for these mutant proteins is around 327 nm, with a similar width of emission as the wild-type protein. The

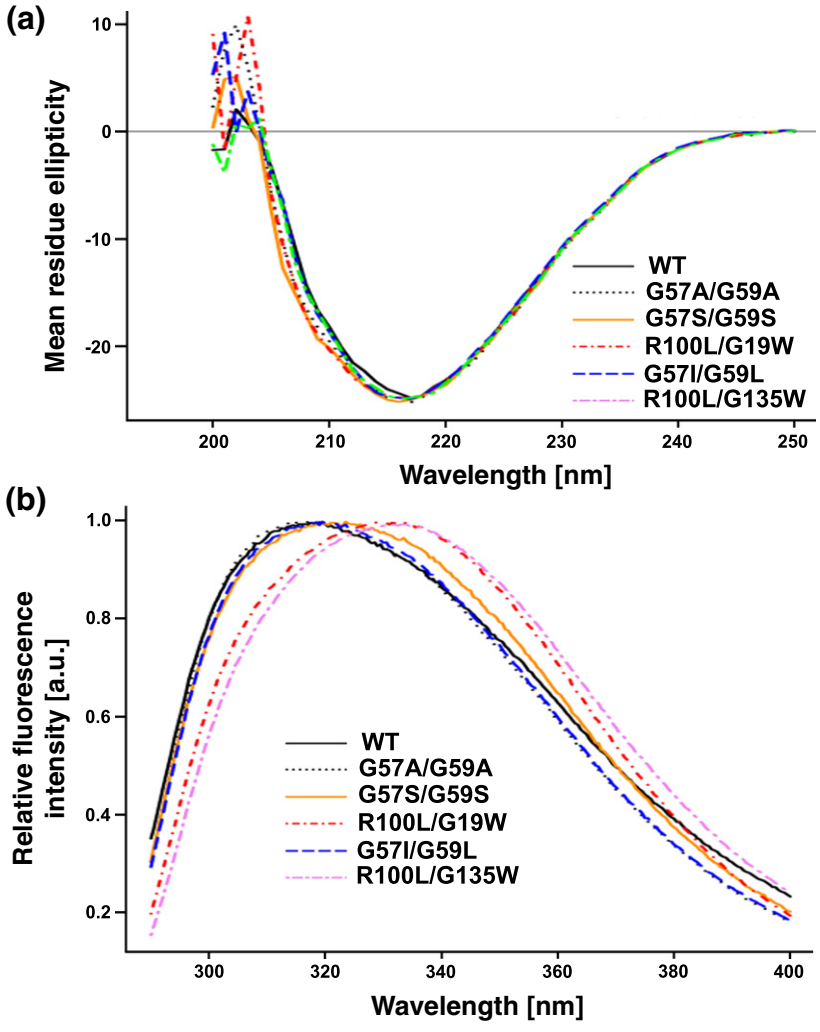

Fig. 3. UV-CD spectra and tryptophan fluorescence emission spectra of wild-type and mutant OmpF. (a) Comparison of UV-CD spectra between wild-type and mutant OmpF proteins. Measurements were recorded for a protein concentration of 0.2 mg/ml. An average of three scans at 50 nm/min was acquired at a bandwidth of 0.2 nm and a response time of 1 s using three independent protein preparations. The final CD spectrum was then corrected for background by subtracting the spectrum of protein-free samples recorded under the same conditions. Noisy data below 200 nm have been removed. (b) Comparison of the tryptophan fluorescence emission spectra of wild-type and mutant OmpF proteins. Measurements were recorded using samples containing 0.2 mg/ml protein held in a 1-mm path length cuvette, with an excitation wavelength of 290 nm. All mutant and wild-type OmpF proteins exhibit similar characteristics in UV-CD and tryptophan fluorescence emission spectra, given that mutants R100L/G19W and R100L/G135W have one more tryptophan compared to the wild-type protein. Thus, the secondary structure formation and the environment of the tryptophan residues in wild-type and mutant OmpF proteins are similar.

change in intensity maxima may be attributed to the extra tryptophan in the R100L/G19W and R100L/G135W mutants. Overall, these observations can be interpreted as an indication of an unaltered environment of the tryptophan residues compared to the wild-type. Any observed differences in thermal stability and oligomeric state are therefore unlikely to be due to altered protein structures.

Engineered dimeric and monomeric oligomerization states

To compare the oligomerization states of the wild-type and OmpF mutants, we analyzed these proteins by SDS-PAGE and Coomassie blue staining (Fig. 4). Mutants R100L/G19W (Fig. 4a, lane 6) and G57I/G59L (Fig. 4a, lane 7), which are designed to alter the PPI interface, indeed were found to exist in a folded monomeric form instead of the trimeric form, in addition to the unfolded form. Mutant R100L/G135W (Fig. 4a, lane 5) exists in folded dimeric and monomeric forms, as well as in an unfolded state. The wild type (Fig. 4a, lane 2), as well as mutants G57A/G59A and G57S/G59S (Fig. 4a, lanes 3 and 4), was found to be present in the folded trimeric state and in an unfolded state. This was also in agreement with our experimental design, in which the two substitutions (G57A/G59A and G57S/G59S) follow the substitution pattern observed in β-barrel membrane proteins and were not expected to have an altered oligomerization state.

Trypsin resistance has been used in previous studies as a useful indicator of protein folding.[23,44] We observed that the folded oligomeric and monomeric bands (Fig. 4a) were resistant to trypsin digestion, while the unfolded band disappeared after trypsin treatment (Fig. 4b). This indicates that the folded oligomeric and monomeric species are compact and well folded. Moreover, when the samples were heated to 95 °C, both folded oligomeric and monomeric bands disappeared, and the density of the unfolded band increased (Fig. S1). Furthermore, mass spectrometric analysis also revealed that only OmpF protein was present in these samples (data not shown). Overall, our results from SDS-PAGE, trypsin digestion, thermal experiments, and mass spectrometric analysis confirmed that the oligomerization state of mutants R100L/G19W and G57I/G59L was changed from a trimeric form to a monomeric form, and that R100L/G135W was altered from a trimeric form to a dimeric form.

Residues G19, G57, G59, R100, and G135 contribute significantly to the stability of the oligomerization state

We further assessed the stability of the oligomeric state of wild-type and mutant OmpF proteins. In thermal denaturation experiments, wild-type OmpF trimers were found to dissociate into monomers at 73 ± 1 °C. Mutants G19W, R100L, G135W, N141W, R100L/G135W, and R100L/G19W (designed by improving the stability of the weakly stable regions, which would lead to more stable monomers) showed a significant decrease at the temperature at which the trimers dissociated into monomers, except N141W (Fig. 5 and Table 1). For example, the

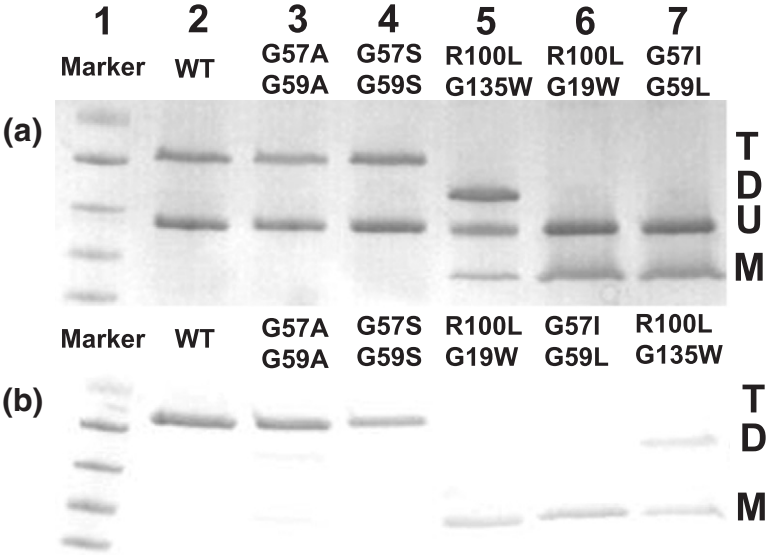

Fig. 4. Oligomerization state of wild-type and mutant OmpF. OmpF proteins were expressed, purified, and refolded from inclusion bodies as described in Methods. Folding reactions were quenched by adding 5× SDS gel-loading buffer[43] to a final dilution of 1× SDS gel-loading buffer. Fifty microliters of sample was loaded onto a 4–20% acrylamide continuous gradient precast gel to resolve the folded and unfolded populations.[12] (a) The wild type and the mutants (G57A/G59A and G57S/G59S) were only present in either the folded trimeric state (T) or the unfolded state (U). The mutants R100L/G19W and G57I/G59L have folded monomeric (M) and unfolded species. The mutant R100L/G135W exists in folded dimeric (D), monomeric, and unfolded states. (b) After overnight refolding, subsequent degradation of the unfolded protein was induced by the addition of trypsin (trypsin/protein, 1:100 wt/wt).[44] The final protein was purified and concentrated using centrifugal filters. The folded oligomeric and monomeric bands were resistant to trypsin digestion, while the unfolded band disappeared after treatment with trypsin.

dissociation temperature for mutant G19W decreased by 26 °C and that for mutant R100L/G135W decreased by 29 °C. These results indicate that residues G19, R100, and G135 contribute significantly to the stability of the oligomeric interface. Moreover, mutants R100L/G135W and R100L/G19W were found to be predominantly dimeric and monomeric at room temperature, respectively. Residue R100 interacts with loop 2 from another subunit.[33] In a previous study by Phale *et al.*, which demonstrated the stability conferred by these interactions to the PPI interface, a number of single and double substitutions have been performed by replacing R100.[33] However, none of these single and double mutants produced a change in the oligomerization state. Our energy function not only identifies weakly stable residues but also suggests stable substitutions based on the depth of the weakly stable residue in the bilayer and the orientation of its side chain. As a result of these stable substitutions, we are able to observe stable monomeric and dimeric oligomerization states for OmpF mutant proteins. As expected, the double mutants G57A/G59A and G57S/G59S designed through evolutionary analysis did not significantly decrease the temperature of trimer disassociation. In

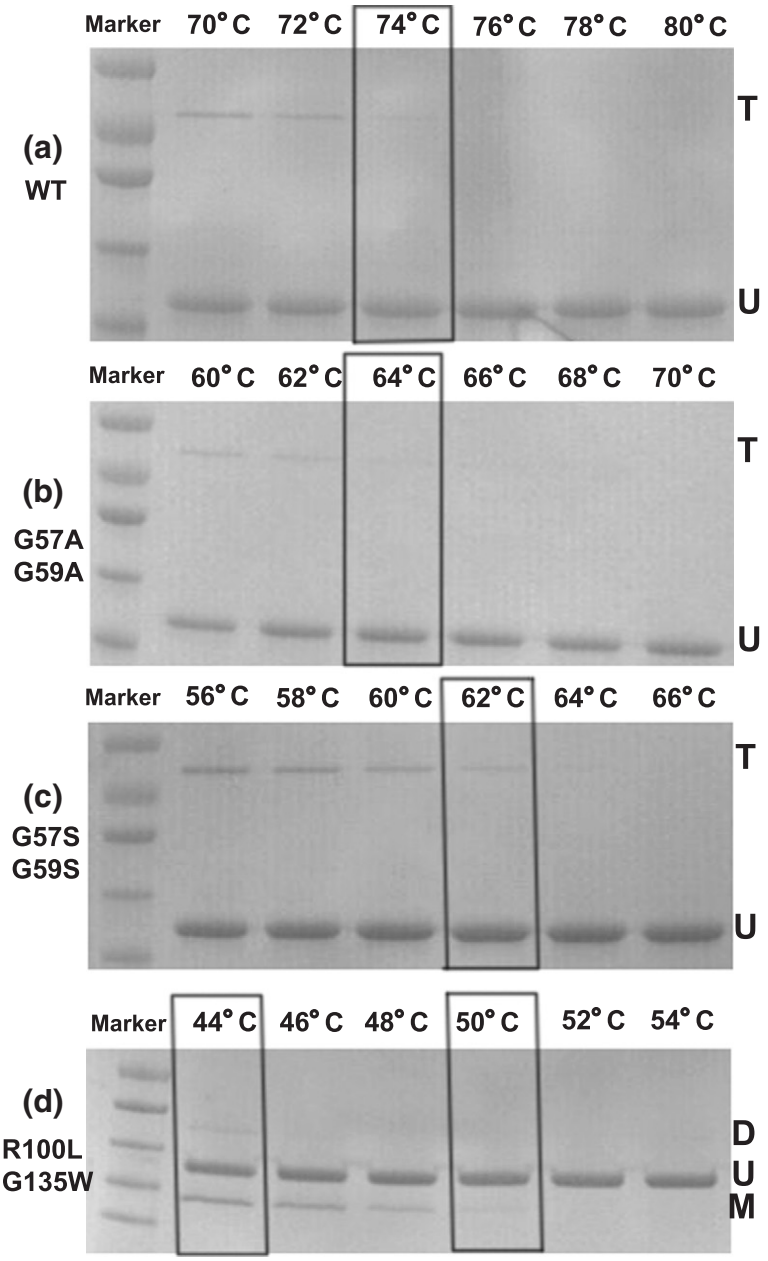

Fig. 5. Dissociation temperature during SDS-PAGE separation for (a) wild type, (b) mutant G57A/G59A, (c) mutant G57S/G57S, and (d) mutant R100L/G135W. Wild-type and mutant OmpF proteins were expressed, purified, and refolded from inclusion bodies as described in Methods. Folding reactions were quenched by adding 5× SDS gel-loading buffer[43] to a final dilution of 1×SDS gel-loading buffer. Fifty microliters of sample was loaded onto a 4–20% acrylamide continuous gradient precast gel to resolve the folded and unfolded populations.[12] Wild-type and mutant proteins were subjected to temperature increases of 2 °C from 20 to 90 °C. Gel lanes at the transition temperature are shown in boxes. The dissociation of dimeric species takes place first, followed by the unfolding of the monomeric species. Trimeric (T), dimeric (D), monomeric (M), and unfolded (U) species are labeled.

Table 1. The temperature at which oligomers (trimers/dimers) dissociate into monomers, as well as the oligomeric state of the wild type and mutants as designed

Mutant	Dissociation temperature (°C)	ΔT (°C)	Oligomerization state
Wild type	73 ± 1	NA	Trimer
G19W	47 ± 2	−26	Trimer
R100L	52 ± 2	−21	Trimer
G135W	52 ± 2	−21	Trimer
N141W	67 ± 2	−6	Trimer
R100L/G135W	44 ± 1	−29	Dimer
R100L/G19W	NA	NA	Monomer
G57A/G59A	63 ± 1	−10	Trimer
G57S/G59S	61 ± 1	−12	Trimer
G57I/G59L	NA	NA	Monomer

Wild-type and mutant proteins were subjected to temperature increases of 2 °C from 20 to 90 °C. The dissociation temperature was measured through SDS-PAGE separation. ΔT is the difference in the dissociation temperatures of the respective mutant OmpF and the wild type: $\Delta T = T_{\text{mutant}} - T_{\text{wild type}}$.

contrast, the G57I/G59L mutant, which was designed with an expected altered oligomerization state, existed primarily in monomeric form at room temperature.

Oligomer dissociation precedes protein unfolding

The mutant R100L/G135W predominantly forms dimeric species at room temperature (Fig. 4a, lane 5). Thermal denaturation of the dimeric mutant R100L/G135W indicates that although this mutant can exist in monomeric form, it is significantly stabilized by oligomerization (Fig. 5d). When the temperature was raised to 44 °C, the dimeric form of this mutant dissociated into a folded monomeric form, which was then unfolded at 50 °C. This behavior suggests that the dissociation of oligomeric species occurred prior to protein unfolding. Furthermore, we have confirmed by temperature-titrated CD spectroscopy that unfolding of the R100L/G135W mutant does not occur around the dimer dissociation temperature (44 °C) (Fig. S2).

We also examined the long-term stability of the dimeric mutant R100L/G135W, as a previous study had reported a transient dimeric form of OmpF.[25] In contrast to the previous study, the dimeric oligomerization state of mutant R100L/G135W was stable and did not convert into a trimeric or monomeric state even after 30 days (Fig. 6). Moreover, the detection of a folded monomer in the thermal denaturation experiments (Fig. 5d), even after the disappearance of the dimeric band, indicated that dimer dissociation and unfolding are not coupled events.

This decoupling of oligomer dissociation from protein unfolding is only seen in the mutant R100L/G135W and not in the wild type or other mutants. This is possibly because the monomeric form of mutant R100L/G135W is sufficiently more stable than those of the wild type and other mutant proteins. It is possible that the wild-type OmpF protein also undergoes trimer dissociation before protein unfolding but our experiments are not sensitive enough to capture the monomeric form that may exist for a short interval of time, unlike the relatively long-lived state of the R100L/G135W mutant protein.

Discussion

Folding efficiency is independent of oligomerization

Previous studies on β-barrel membrane protein folding suggest that OmpF has a lower yield of folded protein than OmpA due to the extra step of oligomerization.[23] Our results showed that after the process of oligomerization had been eliminated by converting OmpF into a monomeric form through amino acid substitutions, the yield of folded protein did not show a significant difference: it changed from 47 ± 6% (12 samples) to 41 ± 4% (6 samples). This suggests that the oligomerization of OmpF does not reduce the folding efficiency of the protein.

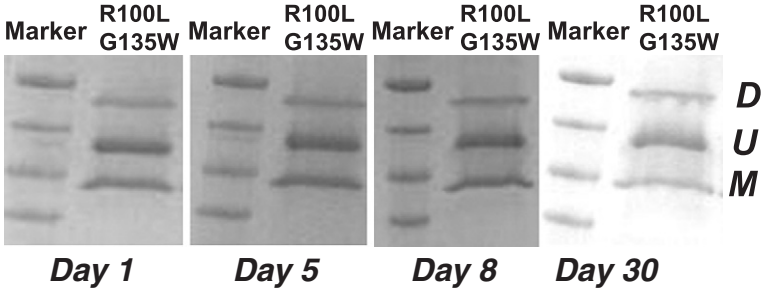

Fig. 6. The mutant R100L/G135W has a stable dimeric state. Mutant R100L/G135W protein was expressed, purified, and refolded from inclusion bodies as described in Methods. Folding reactions were quenched by adding 5× SDS gel-loading buffer[43] to a final dilution of 1× SDS gel-loading buffer. Fifty microliters of sample was loaded onto a 4–20% acrylamide continuous gradient precast gel to resolve the folded and unfolded populations.[12] The populations of folded monomeric (M) and dimeric (D) species, as well as unfolded (U) species, remain constant during the course of 30 days, as measured by SDS-PAGE separation.

The minor difference in folding efficiency might be attributed to the amino acid substitutions in the mutant proteins.

Insight into the β-barrel folding process

We have identified three residues (G19, R100, and G135) in the extensive PPI interface spanning six β-strands out of a total of 16 β-strands. Our results suggest that oligomerization occurs in a stepwise manner. Dimeric species are formed first along the interface near G19, followed by the formation of trimeric species through additional interactions at the interface near G135. Alteration of the interface near G19 (by replacing it with a Trp) inhibits this process. In this case, only the monomeric form of OmpF exists, irrespective of the nature of the interface near G135 (i.e., mutant R100L/G19W was monomeric). The region around G57 and G59 is relatively stable and therefore not likely to be the initiation point of the oligomerization process. However, this region might provide a scaffold for PPIs, and disrupting this scaffold results in monomeric protein (i.e., mutant G57I/G59L was monomeric).

Thermal denaturation of the dimeric mutant R100L/G135W revealed that the folded monomeric form of the mutant protein can be detected even after the disappearance of the dimeric form. This indicates that the dissociation of the dimer was not coupled with the unfolding of the protein. This phenomenon is consistent with recent unfolding studies on LamB, another β-barrel membrane protein, where the trimeric protein disassociated into folded monomeric species at low pH values, suggesting that trimer disassociation and unfolding of the protein may not occur simultaneously.[44]

Trimeric *versus* dimeric interface

The wild-type OmpF protein has two distinct high-energy regions. The first is composed of strand 1, and the second is composed of strands 5 and 6. To date, the oligomerization of bacterial porins has not been linked to any biological process, and the rationale for their preference for a particular oligomerization state (trimer for OmpF) is not clear. We hypothesize that bacterial porins such as OmpF prefer the trimeric oligomerization state to stabilize their two high-energy regions. This is supported by the observation that the R100L/G135W mutant—where the second region (strands 5 and 6) is fully stabilized by the substitutions—forms dimers. The mutant R100L/G135W has only one unstable region (region 1, strand 1) left and likely participates in only one PPI. This would give rise to a dimeric state. We expect the dimeric interface to be significantly different from the trimeric interface, possibly accompanied by a number of conformational changes. However, further studies are required to fully elucidate the dimeric interface of the R100L/G135W mutant.

Relevance of oligomerization to porins

The complex and multistep process of the biogenesis of porins makes it difficult to fully elucidate the mechanism of oligomerization.[30] Even though the PPI interfaces of these proteins are known, their importance to the function of porins is largely unexplored. In this study, we showed that three mutants (R100L/G19W, R100L/G135W, and G57I/G59L) can exist in nonpreferred oligomeric states (dimer and monomer). These constructs can be useful in studying the process and mechanism of oligomerization. For example, the effect of oligomerization on membrane insertion can be studied by reconstituting these mutant proteins in artificial lipid membranes. It will also be interesting to study the effect of oligomerization on transport from the cytoplasm to the membrane and to assess whether and how different oligomerization states affect the mechanisms of translocation, chaperone interaction, and transporter binding.

Application of biological pores in nanobiotechnology

A significant number of biological pores are β-barrel membrane proteins (e.g., porins and α-hemolysin).[45] With strong substrate specificity and a multitude of control points, biological nanopores are promising devices for reagentless DNA sequencing, bioalarm systems, monitoring of single-molecule chemical reactions, bio-inspired batteries, and nanotransistors.[45–48] However, these applications have only been explored under controlled laboratory conditions. The limitations of biological nanopores, such as lack of stability, nonspecific binding, and undesirable oligomeric states, hamper their applications in the uncontrolled environment of the real world, where extreme temperature and denaturing conditions are often encountered. By altering the oligomeric state of β-barrel membrane proteins, we have provided a useful computational addition to the nanobiotechnician's toolbox, with the promise of accelerating efforts for designing novel biological nanopores.

Conclusion

We have shown that computational prediction of weakly stable regions in β-barrel membrane proteins can be used to design β-barrel membrane proteins that exist in a desired oligomerization state. Moreover, evolutionary analysis of the PPI interface can be used to identify important residues that are

required for stable PPIs in the TM domain. Combining experimental and computational approaches to engineer β-barrel membrane proteins with different oligomerization states and structural properties can facilitate further studies on the structural biology of membrane proteins and the design of novel nanopores in nanobiotechnology.

Methods

Identification of weakly stable regions in OmpF

We assess the stability of individual TM β-strands by calculating their empirical energy values using the methodology described by Naveed *et al.*[32] Briefly, the energy of each residue in the predicted native conformation was calculated using an empirical potential function, TmSIP, which was derived from a combinatorial analysis of β-barrel membrane protein structures.[42] The energy for each residue consists of two terms. First, each residue is assigned an energy value of burying this residue type at a particular depth in the lipid bilayer and with the specific orientation of its side chain. There are two possible orientations, namely side chains facing the lipid environment or side chains facing inside the barrel. This is termed the "single-body term." In general, burying a hydrophobic residue facing the lipid is favorable and results in low-energy value, while burying a charged residue facing the lipids is unfavorable and results in high-energy value. Second, each residue interacts with two residues on separate neighboring strands through strong backbone H-bond interaction, side-chain interactions, and weak H-bond interactions, which collectively make up the two-body energy term. For example, Phe-Phe hydrophobic interaction (also known as π-stacking) is a favorable interaction, hence lowering the energy value of the TM β-barrel. The overall strand energy is the summation of both single-body and two-body energy terms over all residues in the strand. Strands with energy higher than the mean energy of all the strands are regarded as "weakly stable strands." In order to identify the most unstable residue in a strand, we calculated the empirical energy required to insert that residue at its proper depth in the lipid bilayer with the corresponding side-chain orientation, as seen in the crystal structure. The residues that require a relatively high free energy to insert into the lipid bilayer are termed "weakly stable residues."

Identification of evolutionarily conserved residues in the PPI interface of OmpF

Quantification of variability at each position of the TM domains of the OmpF protein was made by collecting homologous sequences from a nonredundant database. A multiple sequence alignment was generated, and the entropy of sequence variability was then calculated for every position.

Briefly, using the sequence of OmpF (Protein Data Bank ID: 2OMF) as query, we carried out a Blast search against the nonredundant National Center for Biotechnology Information protein database (default parameters, except for a word size of 2, for increased sensitivity).[49] We selected pairwise sequence alignments of >30% sequence identity that cover at least three-fourths of the OmpF sequence. There are 869 alignments that matched these criteria. A simple multiple sequence alignment was then built based on individual pairwise sequence alignment (i.e., by stacking homologous sequences based on their alignment to OmpF).

The information entropy H_k of sequence variations (in bits) was then calculated for every column k of the multiple sequence alignment using the formula: $H_k = -\sum_{i=1}^{20} f_i \log_2 f_i$. Here f_i is the frequency of amino acid i in the column k. Low entropy indicates that the specific column is very well conserved. The maximum entropy is $\log_2 20 = 4.32$ bits. For illustration, we show $R_k = \log_2 20 - H_k$, where well-conserved columns show the longest bars (Fig. 2).

Cloning and strains

OmpF inclusion bodies were expressed from *E. coli* BL21(DE3) cells using a pET28A vector expression system (Novagen). The OmpF signal sequence (residues 1–22) was replaced by a single methionine residue using PCR. The OmpF cDNA was amplified by the OmpF forward and reverse primers and ligated into a BamHI/XhoI-digested pET28A vector. PrimerX† was used to design the primers for site-directed mutagenesis. Amplification of the product was performed using DNA polymerase (Pfu-Turbo Hotstart; Stratagene) in *E. coli* DH5 cells before sequencing and transformation into *E. coli* BL21 cells for protein expression. The primers are listed in Table S1.

Protein expression and purification

Transformed cells were grown at 37 °C in LB medium (LB Broth; Fisher Scientific) containing 50 μg/ml kanamycin. At an OD_{600} of 0.6, the culture was induced with 0.5 mM isopropyl β-D-thiogalactoside and grown for 3 h. The cells were then harvested by centrifugation at 3000*g* for 5 min. The cell pellet was resuspended in 1× Dulbecco's phosphate-buffered saline (DPBS). The solution was sonicated briefly to shear DNA, and the viscosity of the solution was reduced on ice. After sonication, the inclusion bodies were pelleted and washed two times with DPBS 1× and resuspended in 1% (vol/vol) Triton X-100, 20 mM Tris/HCl (pH 8.0), 0.1 mM ethylenediaminetetraacetic acid, and 1 mM dithiothreitol.[25] The suspension was then incubated at 37 °C for 1 h and washed with 1× DPBS.

Protein refolding and purification

The inclusion bodies were solubilized in denaturing buffer constituting of 50 mM HCl (pH 8.0) and 8 M urea at 55 °C for 30 min. Refolding was performed by a 20× dilution with thorough mixing of the pooled sample, whose concentration had been adjusted to 2 mg/ml. The refolding buffer contained 0.5% (wt/vol) *n*-dodecyl-β-D-glucopyranoside and 0.2% (wt/vol) *n*-dodecyl-β-D-maltoside in 50 mM Tris–HCl, 1 mM dithiothreitol, and 0.1 mM

† www.bioinformatics.org/primerx

ethylenediaminetetraacetic acid (pH 8).[25] The mixture was incubated overnight at 37 °C using an Isotemp Hybridization Incubator (Fisher Scientific). After overnight refolding, subsequent degradation of the unfolded protein was induced by the addition of trypsin (trypsin/protein, 1:100 wt/wt).[44] The final protein was purified and concentrated using Amicon Ultra-0.5-ml Centrifugal Filters 30K (Millipore). All refolded samples used in the analysis were incubated in 0.5% (wt/vol) *n*-octyl-oligooxyethylene and 10 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid buffer (pH 7.4).

Sds-page

Folding reactions of OmpF were quenched by adding 5× SDS gel-loading buffer[43] to a final dilution of 1× SDS gel-loading buffer. Fifty microliters of sample was loaded onto a 4–20% acrylamide continuous gradient precast gel (Bio-Rad) to resolve the folded and unfolded populations.[12]

Circular dichroism

Circular dichroism (UV-CD) spectroscopy measurements were recorded on a Jasco J-810 spectrometer (Jasco, Easton, MD) using cuvettes with 1-mm path lengths for a protein concentration of 0.2 mg/ml. The samples contained proteins before trypsin treatment, as the yields of all the proteins were similar (see Discussion). Every sample was scanned in the wavelength range 200–250 nm. An average of three scans at 50 nm/min was acquired with a bandwidth of 0.2 nm and a response time of 1 s using three independent protein preparations. The final CD spectrum was then corrected for background by subtracting the spectrum of protein-free samples recorded under the same conditions.

Tryptophan fluorescence measurements

Wild-type OmpF and all mutants, except for R100L/G19W and R100L/G135W, have two tryptophan residues, both located in the TM domain. Trp fluorescence measurements were recorded from 300 to 400 nm on a Fluoromax-3 spectrofluorimeter (Jobin-Yvon, Inc., Edison, NJ) using samples containing 0.2 mg/ml protein held in a 1-mm path length cuvette, with an excitation wavelength of 290 nm.[23]

Acknowledgements

The authors thank Dr. Timothy Keiderling for access to CD and fluorescence equipment and for generous advice, and Dennis Gessmann for helpful discussions. This work was supported by National Institutes of Health grants GM-079804 and GM-058746 (to L.J.K.) and National Science Foundation grants DMS-0800257 and DBI-1062328. H.N. was supported by the Fulbright Fellowship and the Higher Education Commission of Pakistan. D.J.-M. is very thankful for the support of Becas Talentia Excellence Grant (Andalusian Ministry of Innovation, Science, and Enterprise, Junta de Andalucia, Spain).

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2012.02.043

References

1. Wallin, E. & von Heijne, G. (1998). Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein Sci.* **7**, 1029–1038.
2. Arkin, I. & Brunger, A. (1998). Statistical analysis of predicted transmembrane alpha-helices. *Biochim. Biophys. Acta*, **1429**, 113–128.
3. Berman, H. M., Westbrook, J. D., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H. *et al.* (2000). The Protein Data Bank. *Nucleic Acids Res.* **28**, 235–242.
4. Baker, M. (2010). Making membrane proteins for structures: a trillion tiny tweaks. *Nat. Methods*, **7**, 429–434.
5. Delcour, A. (2002). Structure and function of pore-forming beta-barrels from bacteria. *J. Mol. Microbiol. Biotechnol.* **4**, 1–10.
6. Bishop, R. (2008). Structural biology of membrane-intrinsic beta-barrel enzymes: sentinels of the bacterial outer membrane. *Biochim. Biophys. Acta*, **1778**, 1881–1896.
7. Song, L., Hobaugh, M. R., Shustak, C., Cheley, S., Bayley, H. & Gouaux, J. E. (1996). Structure of staphylococcal alpha-hemolysin, a heptameric transmembrane pore. *Science*, **274**, 1859–1866.
8. Hong, H., Joh, N., Bowie, J. & Tamm, L. (2009). Methods for measuring the thermodynamic stability of membrane proteins. *Methods Enzymol.* **455**, 213–236.
9. Gessmann, D., Mager, F., Naveed, H., Arnold, T., Weirich, S., Linke, D. *et al.* (2011). Improving the resistance of a eukaryotic β-barrel protein to thermal and chemical perturbations. *J. Mol. Biol.* **413**, 150–161.
10. Tamm, L., Arora, A. & Kleinschmidt, J. (2001). Structure and assembly of beta-barrel membrane proteins. *J. Biol. Chem.* **276**, 32399–32402.
11. Tamm, L., Hong, H. & Liang, B. (2004). Folding and assembly of beta-barrel membrane proteins. *Biochim. Biophys. Acta*, **1666**, 250–263.
12. Burgess, N., Dao, T., Stanley, A. & Fleming, K. (2008). Beta-barrel proteins that reside in the *Escherichia coli* outer membrane *in vivo* demonstrate varied folding behavior *in vitro*. *J. Biol. Chem.* **283**, 26748–26758.
13. Danoff, E. & Fleming, K. (2011). The soluble, periplasmic domain of OmpA folds as an independent unit and displays chaperone activity by reducing the self-association propensity of the unfolded OmpA transmembrane beta-barrel. *Biophys. Chem.* **159**, 194–204.

14. Wimley, W. (2003). The versatile beta-barrel membrane protein. *Curr. Opin. Struct. Biol.* **13**, 404–411.
15. Liang, J., Naveed, H., Jimenez-Morales, D., Adamian, L. & Lin, M. (2012). Computational studies of membrane proteins: models and predictions for biological understanding. *Biochim. Biophys. Acta*, **1818**, 927–941.
16. Walther, D., Rapaport, D. & Tommassen, J. (2009). Biogenesis of beta-barrel membrane proteins in bacteria and eukaryotes: evolutionary conservation and divergence. *Cell Mol. Life Sci.* **66**, 2789–2804.
17. Kusters, I. & Driessen, A. (2011). Seca, a remarkable nanomachine. *Cell Mol. Life Sci.* **68**, 2053–2066.
18. Behrens-Kneip, S. (2010). The role of sura factor in outer membrane protein transport and virulence. *Int. J. Med. Microbiol.* **300**, 421–428.
19. Sklar, J., Wu, T., Kahne, D. & Silhavy, T. (2007). Defining the roles of the periplasmic chaperones sura, skp, and degp in *Escherichia coli*. *Genes Dev.* **21**, 2473–2484.
20. Knowles, T., Scott-Tucker, A., Overduin, M. & Henderson, I. (2009). Membrane protein architects: the role of the BAM complex in outer membrane protein assembly. *Nat. Rev. Microbiol.* **7**, 206–214.
21. Tamm, L., Hong, H. & Liang, B. (2004). Folding and assembly of beta-barrel membrane proteins. *Biochim. Biophys. Acta*, **1666**, 250–263.
22. White, S. & Wimley, W. (1998). Hydrophobic interactions of peptides with membrane interfaces. *Biochim. Biophys. Acta*, **1376**, 339–352.
23. Surrey, T., Schmid, A. & Jahnig, F. (1996). Folding and membrane insertion of the trimeric beta-barrel protein OmpF. *Biochemistry*, **35**, 2283–2288.
24. Cowan, S. W., Schirmer, T., Rummel, G., Steiert, M., Ghosh, R., Pauptit, R. A. *et al.* (1992). Crystal structures explain functional properties of two *E. coli* porins. *Nature*, **358**, 727–733.
25. Visudtiphole, V., Thomas, M., Chalton, D. & Lakey, J. (2005). Refolding of *Escherichia coli* outer membrane protein F in detergent creates LPS-free trimers and asymmetric dimers. *Biochem. J.* **392**, 375–381.
26. Reid, J., Fung, H., Gehring, K., Klebba, P. & Nikaido, H. (1988). Targeting of porin to the outer membrane of *Escherichia coli*. Rate of trimer assembly and identification of a dimer intermediate. *J. Biol. Chem.* **263**, 7753–7759.
27. Watanabe, Y. (2002). Effect of various mild surfactants on the reassembly of an oligomeric integral membrane protein OmpF porin. *J. Protein Chem.* **21**, 169–175.
28. de Cock, H., Hendriks, R., de Vrije, T. & Tommassen, J. (1990). Assembly of an *in vitro* synthesized *Escherichia coli* outer membrane porin into its stable trimeric configuration. *J. Biol. Chem.* **265**, 4646–4651.
29. Van Gelder, P. & Tommassen, J. (1996). Demonstration of a folded monomeric form of porin PhoE of *Escherichia coli in vivo*. *J. Bacteriol.* **178**, 5320–5322.
30. Meng, G., Fronzes, R., Chandran, V., Remaut, H. & Waksman, G. (2009). Protein oligomerization in the bacterial outer membrane (review). *Mol. Membr. Biol.* **26**, 136–145.
31. Haltia, T. & Freire, E. (1995). Forces and factors that contribute to the structural stability of membrane proteins. *Biochim. Biophys. Acta*, **1228**, 1–27.
32. Naveed, H., Jackups, R., Jr & Liang, J. (2009). Predicting weakly stable regions, oligomerization state, and protein–protein interfaces in transmembrane domains of outer membrane proteins. *Proc. Natl Acad. Sci. USA*, **106**, 12735–12740.
33. Phale, P. S., Philippsen, A., Kiefhaber, T., Koebnik, R., Phale, V. P., Schirmer, T. & Rosenbusch, J. P. (1998). Stability of trimeric OmpF porin: the contributions of the latching loop L2. *Biochemistry*, **37**, 15663–15670.
34. Evanics, F., Hwang, P., Cheng, Y., Kay, L. & Prosser, R. (2006). Topology of an outer-membrane enzyme: measuring oxygen and water contacts in solution NMR studies of PagP. *J. Am. Chem. Soc.* **128**, 8256–8264.
35. Adamian, L., Naveed, H. & Liang, J. (2011). Lipid-binding surfaces of membrane proteins: evidence from evolutionary and structural analysis. *Biochim. Biophys. Acta*, **1808**, 1092–1102.
36. Van Gelder, P., Saint, N., Phale, P., Eppens, E. F., Prilipov, A., van Boxtel, R. *et al.* (1997). Voltage sensing in the PhoE and OmpF outer membrane porins of *Escherichia coli*: role of charged residues. *J. Mol. Biol.* **269**, 468–472.
37. Jimenez-Morales, D. & Liang, J. (2011). Pattern of amino acid substitutions in transmembrane domains of [beta]-barrel membrane proteins for detecting remote homologs in bacteria and mitochondria. *PLoS ONE*, **6**, e26400.
38. Kortemme, T. & Baker, D. (2004). Computational design of protein–protein interactions. *Curr. Opin. Chem. Biol.* **8**, 91–97.
39. Ben-Shimon, A. & Eisenstein, M. (2010). Computational mapping of anchoring spots on protein surfaces. *J. Mol. Biol.* **402**, 259–277.
40. Sharabi, O., Yanover, C., Dekel, A. & Shifman, J. (2011). Optimizing energy functions for protein–protein interface design. *J. Comput. Chem.* **32**, 23–32.
41. Fleishman, S. J., Whitehead, T. A., Strauch, E. M., Corn, J. E., Qin, S., Zhou, H. X. *et al.* (2011). Community-wide assessment of protein-interface modeling suggests improvements to design methodology. *J. Mol. Biol.* **414**, 289–302.
42. Jackups, R. & Liang, J. (2005). Interstrand pairing patterns in beta-barrel membrane proteins: the positive-outside rule, aromatic rescue, and strand registration prediction. *J. Mol. Biol.* **354**, 979–993.
43. Sambrook, J., Maniatis, T. & Fritsch, E. F. (1989). *Molecular Cloning: A Laboratory Manual*, 2nd edit. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY; B.25.
44. Baldwin, V., Bhatia, M. & Luckey, M. (2011). Folding studies of purified lamb protein, the maltoporin from the *Escherichia coli* outer membrane: trimer dissociation can be separated from unfolding. *Biochim. Biophys. Acta*, **1808**, 2206–2213.
45. Majd, S., Yusko, E. C., Billeh, Y. N., Macrae, M. X., Yang, J. & Mayer, M. (2010). Applications of biological pores in nanomedicine, sensing, and nanoelectronics. *Curr. Opin. Biotechnol.* **21**, 439–476.
46. Butler, T., Pavlenok, M., Derrington, I., Niederweis, M. & Gundlach, J. (2008). Single-molecule DNA detection with an engineered MspA protein nanopore. *Proc. Natl Acad. Sci. USA*, **105**, 20647–20652.
47. Adiga, S., Jin, C., Curtiss, L., Monteiro-Riviere, N. & Narayan, R. (2009). Nanoporous membranes for

medical and biological applications. *Wiley Interdiscip. Rev. Nanomed. Nanobiotechnol.* **1**, 568–581.
48. Banerjee, A., Mikhailova, E., Cheley, S., Gu, L. Q., Montoya, M., Nagaoka, Y. *et al.* (2010). Molecular bases of cyclodextrin adapter interactions with engineered protein nanopores. *Proc. Natl Acad. Sci. USA*, **107**, 8165–8170.
49. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410.