Computational redesign of the lipid-facing surface of the outer membrane protein OmpA

James A. Stapleton^{a,b}, Timothy A. Whitehead^{a,c}, and Vikas Nanda^{b,1}

^aDepartment of Chemical Engineering and Materials Science, Michigan State University, East Lansing, MI 48824; ^bCenter for Advanced Biotechnology and Medicine and the Department of Biochemistry and Molecular Biology, Robert Wood Johnson Medical School, Rutgers University, Piscataway, NJ 08854; and ^cDepartment of Biosystems and Agricultural Engineering, Michigan State University, East Lansing, MI 48824

Edited by William F. DeGrado, School of Pharmacy, University of California, San Francisco, CA, and approved June 29, 2015 (received for review January 27, 2015)

Advances in computational design methods have made possible extensive engineering of soluble proteins, but designed β -barrel membrane proteins await improvements in our understanding of the sequence determinants of folding and stability. A subset of the amino acid residues of membrane proteins interact with the cell membrane, and the design rules that govern this lipid-facing surface are poorly understood. We applied a residue-level depth potential for β -barrel membrane proteins to the complete redesign of the lipid-facing surface of *Escherichia coli* OmpA. Initial designs failed to fold correctly, but reversion of a small number of mutations indicated by backcross experiments yielded designs with substitutions to up to 60% of the surface that did support folding and membrane insertion.

membrane proteins | protein design | OmpA | β -barrel | statistical potential

The β -barrel membrane proteins comprise one of the two structural classes of integral membrane proteins. They are found within the outer membranes of bacteria, mitochondria, and chloroplasts, where they perform a range of structural, transport, and catalytic functions (1). In addition to their biological interest they are increasingly relevant to biotechnology, serving as scaffolds for bacterial surface display (2, 3) and atomically precise pores for nanopore-based DNA sequencing. Although the suitability of natural β -barrel membrane proteins for biotechnology has been improved by protein engineering (3–10), the ability to design membrane proteins de novo would deliver tools customized to meet the demands of each application.

De novo design provides a stringent test of our understanding of the determinants of protein folding and stability. Protein design software [e.g., Rosetta (11, 12)] has made tremendous strides in addressing the design problem for small water-soluble proteins (13–15), and design of simplified model α -helical membrane proteins including single transmembrane helices and small bundles (16–20) has also been accomplished. In contrast, a designed β -barrel membrane protein has yet to be reported, perhaps as a consequence of the unique design challenges presented by the folding pathway and architecture of these proteins. Unlike the α -helical membrane proteins, nascent β -barrel membrane proteins must transit the periplasm to the outer membrane, where folding and membrane insertion are thought to occur in concert (21, 22). An extensive network of chaperones maintains the solubility of the unfolded barrel and guides membrane insertion. The C-terminal β -strand is known to interact with the BAM chaperone complex (23–25), which assists the folding of all β -barrel membrane proteins. However, despite recent progress (26–30), we do not fully understand how interactions between chaperones and transiting membrane proteins are directed by sequence-encoded information.

Further complicating design is the inside-out architecture of β -barrel membrane proteins. In place of a hydrophobic core is either a central water-filled pore or a solid core composed of polar side chains. The lipid bilayer becomes increasingly hydrophobic at greater depths within the membrane (31), and this

environmental anisotropy is reflected in the amino acid composition of the barrel surface. Aliphatic side chains are prevalent toward the center of the membrane, and aromatic side chains are common in the lipid head group regions, where they encircle the barrel in external- and periplasmic-side girdles (32).

Recently we developed $E_z\beta$, a membrane depth-dependent, residue-level potential calculated from an ensemble of experimentally determined outer membrane protein structures (33, 34). $E_z\beta$ can be used to estimate energetics of membrane insertion to predict transmembrane protein orientation within the bilayer, and to detect oligomerization sites on β -barrel surfaces (34). $E_z\beta$ and related statistical functions (35, 36) can recapitulate properties of natural outer membrane proteins (37, 38) and predict the effects of mutations on protein stability and oligomerization (39). Similar potentials have driven computational approaches that have fully redesigned α -helical membrane protein surfaces to convert membrane proteins into water-soluble ones (40–42).

Here, we considered whether the complete redesign of the lipid-facing surface of an outer membrane protein using a statistical potential such as $E_z\beta$ preserves its structure and function. This approach allowed us to investigate whether membrane insertion requires only a lipid-facing surface composed of depth-appropriate hydrophobic residues, or whether folding requires sequence-specific interstrand interactions, chaperone-recruiting sequences, evolutionarily optimized aromatic girdles, folding nucleation sites, or other design features lost during the population averaging inherent in parameter fitting of statistical potentials.

Significance

The ability to construct novel proteins from basic principles of molecular structure is the fundamental goal of protein design. This is particularly challenging in the case of the β -barrel outer membrane proteins, where our understanding of the rules governing structure and function lags behind that of other classes of proteins. Here, we took a step toward understanding β -barrel membrane protein architecture by focusing on the outward-facing amino acid positions that contact the cell membrane. We replaced the membrane-facing surface of OmpA with new surfaces designed to resemble natural β -barrel surfaces. We were able to design versions of OmpA with mutations at about two-thirds of all surface positions, indicating that β -barrel membrane protein surface design is achievable.

Author contributions: J.A.S., T.A.W., and V.N. designed research; J.A.S. performed research; J.A.S., T.A.W., and V.N. analyzed data; and J.A.S., T.A.W., and V.N. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: Source code for software, sequence and structures of models OR1-4 and variants are made available in the Github repository (https://github.com/jstapleton/2015_OmpA).

¹To whom correspondence should be addressed. Email: nanda@cabm.rutgers.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1501836112/-DCSupplemental.

Previous studies have explored the sensitivity of the β -barrel fold and its chaperone recognition mechanisms to mutations. The canonical eight-stranded β -barrel membrane protein OmpA tolerates a limited number of mutations to the lipid-facing surface, provided hydrophobicity is maintained (43, 44). More radically, the eight-stranded barrel OmpX has been duplicated to form a 16-stranded barrel capable of membrane insertion (45). However, the lipid-facing residues of transmembrane β -strands are conserved across homologous β -barrel membrane proteins beyond the extent expected from hydrophobicity alone (46, 47), implying a functional role that has yet to be elucidated.

To explore the sequence constraints on β -barrel membrane proteins, we extensively redesigned the lipid-facing surface of *E. coli* OmpA. We created a series of OmpA variants with entirely or partially redesigned lipid-facing surfaces and tested their ability to insert into the outer membrane of *E. coli*. Our results indicate that the surfaces of β -barrel membrane proteins are amenable to large-scale redesign, provided that energetically destabilizing substitutions are avoided.

Results

Redesign of the Lipid-Facing Surface. We began by constructing and testing OmpA mutants with redesigned surfaces. Forty-three sites comprising nearly the entire lipid-facing surface (Fig. 1) were selected for redesign on the basis of distance from the center of the membrane and degree of solvent exposure. Membrane depths of OmpA residues were calculated from an oriented structure optimized using the $E_z\beta$ potential as described previously (34). Lipid-facing residues obtained by this automated procedure were largely consistent with those specified by other studies of OmpA (48). Phe170, which is important for chaperone recognition (23), was left unchanged.

Early de novo transmembrane α -helical proteins presented simple, leucine-rich lipid-facing surfaces (17, 49). To test whether simple hydrophobicity is all that is required of the lipid-facing residues in OmpA, we mutated all 43 lipid-facing amino acids to leucine. We tested for correct protein folding by determining host cell susceptibility to the OmpA-dependent bacteriophage K3 (50), which uses the external surface-displayed loops of OmpA to recognize and invade *E. coli* cells. Spotting phage solution on an agar plate results in visible plaques on an otherwise confluent lawn, and the phage titer required for plaque formation provides a quantitative measure of susceptibility.

Fig. 1. $E_z\beta$ was used to generate a membrane-oriented structure of *E. coli* OmpA from PDB 1bxw (46). Forty-two lipid-facing amino acids that faced the protein exterior, had >20% solvent-accessible surface area, and lay within a 25-Å membrane-spanning region were selected for computational redesign (highlighted in orange). Wild-type identities were maintained for lipid-facing positions 168 and 170 given their role in targeting of OmpA to the outer membrane. Interstrand backbone hydrogen bonds are shown as dotted lines. Extracellular and periplasmic regions are omitted for clarity.

The OmpA deletion strain BL Δ ompA Δ ompF (51) is resistant to K3 infection unless OmpA is expressed from a plasmid and properly folded. K3 therefore provided a stringent, conservative assay for folding of OmpA redesigns, and should exclude variants that insert into the membrane yet fail to present external loops properly for phage recognition. *E. coli* transformed with a plasmid encoding the leucine-surfaced OmpA mutant expressed the protein (Fig. S1) yet was completely resistant to phage, indicating that the redesigned protein did not fold correctly in vivo.

We next redesigned the surface residues of four proteins (OmpA Redesigns 1–4, referred to as OR1–OR4) using a Monte Carlo algorithm that searched for optimal lipid-facing surfaces (software and model structures are available in the [Supporting Information](#)). The algorithm maximized a scoring function consisting of a weighted combination of an $E_z\beta$ energy term that rewarded depth-appropriate amino acid choices and a sequence complexity term that enforced a level of amino acid diversity similar to the wild-type surface. In the case of OR4, an additional term was added to the energy function to optimize the relative stability of the protein when centered within the membrane. The redesign process was fully automated: amino acids were selected without human intervention.

The ability of the four designed variants to insert into the outer membrane was tested in OmpA-knockout *E. coli*. In each case, we observed protein expression (Fig. S1) yet complete phage resistance in transformed *E. coli*, indicating that none of the redesigned proteins folded correctly in vivo. The designed proteins were observed in periplasmic and outer membrane fractions, suggesting that issues were primarily with folding rather than trafficking. Restoration of Tyr168, which along with Phe170 may be important for chaperone recognition (23), did not restore activity of any of the redesigns.

Computational analyses of structural models of the redesigned OmpA variants provided little insight into why correctly folded species were not observed. Both $E_z\beta$ and a second propensity-based β -barrel energy function (52) that additionally considers two-body propensities of hydrogen-bonding residue pairs on adjacent β -strands predicted all of the redesigned proteins to be more stable than wild-type OmpA (Table S1). Inspection of model structures with PyMOL and energy calculations using protCAD (53) and Rosetta (11, 12) did not reveal unresolvable clashes between any pair of redesigned sidechains or between redesigned and wild-type sidechains.

Backcross Hybrids Test Effects of the Folding Pathway and Aromatic Girdles.

We hypothesized that the failure of the redesigns to fold in vivo might be a result of perturbations to specific β -strand interactions important for folding. OmpA has been shown to fold less efficiently following circular permutation (54), suggesting that the design constraints differ among β -strands as a function of their order. In particular, β -strands 1 and 8 must come together to close the barrel, and β -strands 1 and 2 associate late in the folding pathway (21). We used site-directed mutagenesis to restore all permutations of β -strands 1, 2, and 8 to their wild-type sequences. We refer to these wild-type/redesign hybrids as backcrosses, where all positions on a given strand have either wild-type or redesigned sequences. For example, the H18 backcross hybrid would have wild-type sequences for strands 1 and 8 and redesigned sequences for strands 2–7. All backcrossing studies were performed on the OR4 design, chosen because its sequence was optimized both for energy of insertion and for specificity of position at the center of the lipid bilayer. Previous analyses of wild-type OMPs indicated that the $E_z\beta$ energy for various insertion depths had a funnel-like character with the minimum at the center of the bilayer (34). Therefore, we expected this would also be a favorable constraint on designed sequences. None of the seven backcrosses (H1, H2, H8, H12, H18, H28, and H128) restored bacteriophage susceptibility in an OmpA deletion

background (Table S2), suggesting that disrupted interactions between strands 1, 2, and 8 are not the primary cause of OR4 misfolding.

Another suspected issue was the perturbation of extracellular and periplasmic aromatic girdles in the all-leucine surface and the four computational OmpA redesigns. Aromatic residues are underrepresented in the redesigned proteins compared with the wild type because the sequence complexity constraint penalized designs with large numbers of these residues. To test this, we restored to OR4 all permutations of wild-type β -strands 2, 3, and 7 (H2, H3, H7, H23, H27, H37, and H237 in Table S2), the redesigned versions of which have fewer aromatic residues than the wild-type versions. None of the seven resulting OmpA backcrosses conferred phage susceptibility.

To more directly assess the contribution of aromatic girdles, we constructed OmpA variants in which only the girdles were exchanged (Table S3). Wild-type aromatic girdles (W7, W15, Y43, F51, Y55, W57, Y85, F123, Y129, Y141, W143, and Y168) were restored to all-Leu (L_{WT}girdles) and OR4 (OR4_{WT}girdles), but in neither case was phage susceptibility restored. Conversely, girdles in wild-type OmpA were replaced either with leucine (WT_{noGirdles}) or with residues from OR4 (WT_{OR4}girdles). WT_{OR4}girdles did not show phage susceptibility. Quite surprisingly, WT_{noGirdles} showed significant phage susceptibility (Fig. S2), suggesting it was able to insert into the outer membrane

despite having no aromatic residues. Together, the outcome of the strand 2/3/7 backcrosses and girdle swaps indicates that issues with aromatic girdles were not at the root of OR4 misfolding.

A Fractional-Factorial Set of Backcross Hybrids Reveals Destabilizing β -Strands. To more rigorously assess the contributions of each β -strand, but short of testing all possible backcrosses ($2^8 = 256$ possibilities; either wild-type or OR4 sequences for each of the eight strands), we constructed and tested a fractional-factorial set of 14 wild-type/redesign hybrids. Each hybrid was composed of a different set of four β -strands from OR4 and the remaining four from wild-type OmpA (named according to their wild-type β -strands; Table 1). Eight of the fourteen hybrid proteins rendered *E. coli* susceptible to phage (Fig. 2 and Fig. S2). Phage susceptibility measurements were in agreement with flow cytometry analysis using a fluorescently labeled antibody against a tag inserted in the fourth surface-displayed loop (Table 1, Fig. 2, and Fig. S2).

No single wild-type β -strand was necessary for insertion, and no mutant β -strand was prohibitive: each wild-type β -strand was present in at least one hybrid that failed to fold, and each redesigned β -strand was present in at least one hybrid that did fold (Table 1). ANOVA analysis of the fractional factorial hybrid dataset identified a statistically significant model (P value = 0.0014) that has factors that are the identities of β -strand 6 and β -strand 3 and two additional two-strand interactions. The power

Table 1. Composition and properties of hybrid OmpA constructs

Name	Strand composition 12345678*	Phage score [†]	Signal in FC [‡]	Ez β score (34)	TMsip score (52)	$N_{\text{mutations}}$	$N_{\text{aromatics}}$
WT	WWWWWWWW	8	yes	−51.2	2.9	0	13
OR4hybrid [§]	mmmmmmmm	0	no	−52.8	3.6	33 [¶]	6
H1346	WmWWmWmm	7	yes	−53.5	3.2	14	11
H1256	WWmmWWmm	8	yes	−48.9	2.9	15	9
H2345	mWWWWmmm	8	yes	−51.9	2.9	16	10
H3567	mmWmWWWm	7	yes	−52.8	3.0	17	12
H2368	mWWmmWmW	7	inconclusive	−52.6	3.2	17	10
H1678	WmmmmWWW	7	yes	−53.8	3.3	17	9
H1237	WWWmmmWm	0	inconclusive	−55.5	3.2	17	13
H2578	mWmmWmWW	0	no	−52.2	2.9	19	8
H3478	mmWWmmWW	0	no	−56.8	3.4	18	10
H1248	WWmWmmmmW	0	no	−52.8	3.2	16	7
H2467	mWmWmWWm	0	no	−58.6	3.2	16	10
H1457	WmmWWmWm	0	no	−53.1	2.9	16	9
H4568	mmmWWWmW	0	no	−50.2	3.0	16	6
H1358	WmWmWmmW	0	no	−47.2	2.8	17	9
H16	WmmmmWmm	5	yes	−52.5	3.3	23	8
H36	mmWmmWmm	7	yes	−53.9	3.3	24	10
H136	WmWmmWmm	7	yes	−52.2	3.2	19	11
H146	WmmWmWmm	7	yes	−53.8	3.2	18	8
H346	mmWWmWmm	7	inconclusive	−55.2	3.3	19	10
OR4cons	mmmmmmmm	8	yes	−53.0	3.0	25	8
OR4cons_G41A	mmmmmmmm	8	—	−53.8	3.2	26	8
OR4cons_G99I	mmmmmmmm	7	—	−54.4	3.2	26	8
OR4cons_P121F	mmmmmmmm	8	—	−56.6	3.0	26	9
OR4cons_G125L	mmmmmmmm	8	—	−55.0	3.1	26	8
OR4cons_Y129H	mmmmmmmm	7	—	−51.5	3.2	26	7
OR4cons_Y129L	mmmmmmmm	8	—	−53.4	3.2	26	7

*The origin of β -strands 1–8 in a hybrid, where W indicates a wild-type β -strand and m indicates a mutant β -strand.

[†]A solution containing the OmpA-dependent bacteriophage K3 was serially diluted in tenfold steps. Phage susceptibility was defined as the maximum number of dilution steps at which phage plaques were observed after spotting the dilutions on a bacterial lawn. Scores represent the highest dilution at which plaque formation was observed out of four replicates. A score of 0 indicates no susceptibility at any titer. See Fig. S2 for photographs of phage assay plates.

[‡]See Fig. S3 for flow cytometry fluorescence histograms.

[§]The version of OR4 used as a parent in the creation of hybrids; some mutations at the ends of strands were reverted to facilitate hybrid assembly (Fig. S4).

[¶]This number is lower than the 43 redesign sites because the redesign algorithm independently selected wild-type residues at a few OR4 surface positions, and because some OR4 mutations were omitted from the hybrids to facilitate construct assembly (Fig. S4).

Fig. 2. OmpA membrane insertion was determined with two experimental methods. Data for wild-type OmpA, H2578, H1678, and OR4cons are shown. Data for all variants are provided as [Supporting Information](#). (A) Phage susceptibility was tested by spotting successive tenfold dilutions of phage solution on a confluent plate of *E. coli*. Bacteriophage K3 infection depends upon the presence of properly folded OmpA. (B) Flow cytometry of *E. coli* cells induced with IPTG and incubated with FITC-labeled (indicated by a star) anti-HA antibodies. A fluorescent population indicates the presence of OmpA loops on the exterior of the cell.

of the factorial experiment was inadequate to uniquely resolve these two-strand interactions, but the only interactions among the set of aliased possibilities that involved the two individually significant β -strands (β -strands 6 and 3) linked those strands to either β -strand 1 or β -strand 4. By focusing on these β -strands identified by ANOVA we were able to identify insertion-competent hybrid barrels containing as few as two wild-type β -strands. H16, H36, H136, H146, and H346 all insert into the membrane, as verified by phage infectivity and flow cytometry (Table 1 and Fig. S3). H3 and H6 were not phage susceptible, and introducing designed strands 3 or 4+6 in a wild-type background (H1245678, H123578) maintained phage susceptibility (Table S3), consistent with the observation that no one strand determined the ability of OR4 to fold.

Reversion of Potentially Destabilizing Residues Restores Folding. A key finding in the fractional factorial set was the high significance of β -strand 6 to design success, suggesting mutations at this β -strand in OR4 might be significantly destabilizing. At position 121, which Pautsch and Schulz (46) did not include in the lipid-facing surface but which was included in our automated selection process, proline is changed to phenylalanine. Proline residues often play important structural roles in proteins, and are frequently left untouched in protein redesigns. At position 125, glycine is mutated to leucine. Like proline, glycine often serves a structural role and is preserved in redesigns. At position 129, tyrosine is mutated to histidine. There are no histidines in the wild-type OmpA surface, although they do occur in other β -barrel membrane protein surfaces.

We therefore constructed a conservative version of OR4 in which we restored the wild-type residues at the two designed histidine positions (H129Y on β -strand 6 and H143W on β -strand 7), three of the four mutated glycine positions (G41 on β -strand 2, G99 on β -strand 5, and G125 on β -strand 6), F121 to proline, and Q35 to leucine (restoring hydrophobicity at this position). The β -strand 8 was taken from the all-leucine design, preserving Y168. Folding of this conservative redesign (OR4cons) was comparable to the wild type in our phage assay (Table 1). This design features 25 mutations to the lipid-facing surface (Fig. S4). Site-directed mutagenesis to individually reintroduce some of the potentially destabilizing mutations did not abolish phage sensitivity (OR4cons point mutants: G41A, G99I, P121F, G125L, Y129H, and Y129L; Table 1).

At the outset of this study we hypothesized that recapitulating the depth-dependent surface hydrophobicity of OmpA using rationally or computationally selected sequences would be sufficient to ensure proper trafficking to and folding in the outer membrane. Although localization of nearly all designs to the outer membrane was observed, none were susceptible to bacteriophage infection, indicating that correct folding resulting in nativelike presentation of extracellular loops was not attained. Reversion of a small subset of rationally selected amino acids in OR4 to wild-type identities did recover wild-type phage susceptibility. Our results support using depth-dependent surface hydrophobicity as a design parameter in engineering outer membrane proteins, and highlight the importance of identifying critical residues that perturb structure.

Discussion

In water-soluble globular proteins, large numbers of surface residues can be mutated without disrupting protein stability as long as surface hydrophilicity is maintained (55), and hydrophobic α -helical membrane protein surfaces have been shown to be similarly tolerant. It had been suggested that the same was true for β -barrel membrane protein surfaces (44), a prediction that makes intuitive sense based on the favorable energetic contributions made by lipid-facing hydrophobic residues (56), and known evolutionarily allowed mutation patterns (57).

The results of our complete redesign of the OmpA surface establish that de novo design of β -barrel membrane protein surfaces is possible provided that highly destabilizing residue choices are avoided. Computational redesign of surface-exposed amino acids based solely on favoring depth-appropriate amino acids did not perturb trafficking of OmpA variants to the outer bacterial membrane. In this aspect, the designs were quite tolerant, and even an all-leucine surface OmpA variant was shown to partition into the outer membrane. However, to restore correct folding, it was necessary to revert specific positions back to their wild-type identities. The backcrossing approach evaluating combinations of designed and wild-type β -barrel strands proved invaluable in helping guide our chemical intuition on which positions should be reverted. Finally, in the 43 lipid facing positions of the OR4cons design, 32 were selected computationally, resulting in 25 mutations, and the remaining 9 were manually fixed as wild type. Progress toward fully automated selection of amino acid substitutions awaits a better structural and thermodynamic

Fig. 3. Structural role of proline 121. P121 is unable to participate in a hydrogen bond with the backbone carbonyl of N145, ending the beta-strand hydrogen bond network between strands 6 and 7. Mutation to any other amino acid could propagate the strand further, altering the conformation locally. Other than for P121, only backbone atoms are depicted for clarity.

understanding of how specific wild-type reversions from OR4 to OR4cons restored native structure.

In hindsight, we can rationalize why changing certain positions may have affected folding. For example, P121 is the first amino acid not part of the beta-sheet hydrogen bonding network. This may be due to its backbone imine, which cannot participate in hydrogen bonding, but is positioned to interact with the N145 carbonyl (Fig. 3). The proline acts as a structural “punctuation mark,” ending the interactions between strands 6 and 7. Mutation to another amino acid would presumably allow the hydrogen bond network between strands 6 and 7 to extend farther, altering the conformation of extracellular loops that contribute to bacteriophage binding. Likewise, at other positions in OR4cons, histidine and tyrosine can serve structural roles in nucleating or stabilizing beta-turns (58). Differences in positional and turn-type preferences of these two amino acids could promote misfolding of interstrand connections. Alternatively, rotameric conformations that position the polar moieties of these side chains into the amphiphilic lipid headgroup region may have been unavailable, resulting in unfavorable interactions with the lipid bilayer. Although no one amino acid was critical for stability and folding, collectively these reversions were able to promote a successful design with 25 surface mutations that inserts into the *E. coli* outer membrane.

Although the ubiquity of aromatic girdles in β -barrel membrane proteins suggests a benefit to fitness (by increasing the thermodynamic stability of the protein or contributing to cell viability in a functional role), our results fail to demonstrate that benefit in our OmpA system. In addition, we found no evidence that interactions between β -strands involved in barrel closure (β -strands 1, 8, and 2; ref. 21) were important.

Redesigned OmpA surfaces present an opportunity to explore the stability contributions of a variety of surface features. For example, tyrosine has been thought to be a particularly suitable amino acid for β -barrel aromatic girdles. Its amphiphilic character is ideal for the interfacial environment, it is statistically overrepresented and highly conserved in aromatic girdles, and it has been experimentally observed to make a larger contribution to β -barrel stability than would be predicted from hydrophobicity scales (59). In natural β -barrels, tyrosine is predominantly found at the C-terminal ends of β -strands (60); at this end the available χ_1 dihedral angles better position the sidechain to “snorkel” into the lipid headgroup region. However, although all of the aromatic residues in OR4 happen to be at the N-terminal ends of their respective β -strands, they seem to be as effective as the wild-type C-terminal aromatic residues in stabilizing hybrid β -barrels.

The extensive redesign of the lipid-facing surface presented here is only the first step toward a fully de novo β -barrel membrane protein design. We have demonstrated that despite their high degree of conservation, the majority of the amino acids on

the lipid-facing surface of OmpA can be substituted without disrupting chaperone binding, folding pathways, or interstrand interactions. Additional work will be required to clarify the roles of structural keystone positions such as proline 121 and to determine the design rules governing the residues that comprise the interior of the barrel.

Materials and Methods

Design Details. Forty-three lipid-facing amino acids identified from the OmpA crystal structure (46) were selected for redesign. Redesigned surfaces were generated by a Metropolis Monte Carlo search that optimized the surface with respect to a scoring function. For OR1–OR3, the score was a weighted sum of the independent $E_z\beta$ energies of each amino acid (34) and the sequence complexity, defined as the multinomial coefficient (61):

$$SC = \frac{(\sum_{i=1}^{20} n_i)!}{\prod_{i=1}^{20} n_i!},$$

where n_i is the number of instances of the i th amino acid. The final design function for OR1–3 was: $E_z\beta + 5.06 \times 10^{-70} \times (3.23 \times 10^{35} - SC)^2$. Details on the approach for choosing the weights of each component of the scoring function are described in more detail in the supplementary materials (Fig. S5). The three top-scoring surfaces from ~50 runs were selected.

For OR4, the score was a weighted sum of the $E_z\beta$ energy, the sequence complexity, and the probability of the barrel being centered in the membrane. This last term was calculated as a Boltzmann probability from the $E_z\beta$ energies of the barrel when centered in the membrane or translated up or down by up to 15 Å:

$$P = \frac{\exp(-\beta E_{z=0})}{\sum_z \exp(-\beta E_z)}.$$

The final design function for OR4 was: $E_z\beta + 5.06 \times 10^{-70} \times (3.23 \times 10^{35} - SC)^2 - 10 \log P$. The best-scoring surface from approximately twenty runs was chosen for OR4.

The soluble C-terminal domain of OmpA was not included in the constructs. The first 171 amino acids of the mature OmpA, which encode the membrane-spanning β -barrel, were preceded by an N-terminal 21 amino acid periplasmic targeting-sequence fusion. A FIASH tag (CCPGCC) (62) and a hexa-histidine tag were fused to the C terminus. A hemagglutinin tag (YPYDVPDYA) was inserted in the fourth external loop of the wild-type/mutant hybrids and the second-generation designs for use in flow cytometry experiments. Details of the fraction factorial design of experiments are presented in the [Supporting Information](#) and Fig. S6.

Phage Susceptibility Assay. Phage K3 (strain number NCCB 3448) was obtained from the Centraalbureau voor Schimmelcultures, Fungal Biodiversity Centre, Institute of the Royal Netherlands Academy of Arts and Sciences, Utrecht, the Netherlands.

Phage susceptibility assays were performed following a published protocol (54). Here, 2-mL aliquots of liquid top agar (0.7% agar in LB medium) were held at 50 °C. IPTG was added to a final concentration of 0.03 mM. Cells were added and the mixtures were immediately spread onto LB-agar plates containing 100 μ g/mL ampicillin or carbenicillin. After the top agar had solidified, 2- μ L droplets of 10 $\times$ serial dilutions of phage K3 in LB were applied to the plates. The plates were then incubated at 37 °C for 16 h.

Flow Cytometry. Cultures grown in LB media at 37 °C were induced with 0.1 mM IPTG at an OD between 0.2 and 0.4 and grown for an additional hour at 37 °C. Cells were harvested by centrifugation and incubated at 4 °C for 45 min in PBS with FITC-conjugated anti-HA antibodies (catalog number H7411; Sigma-Aldrich). Labeled cells were pelleted by centrifugation and the buffer was removed. Cells were resuspended in PBS and subjected to flow cytometry on an Accuri C6 benchtop flow cytometer. Events corresponding to single cells were gated by forward and side scatter.

ACKNOWLEDGMENTS. The authors thank Daniel Hsieh and Alexander Davis for helpful discussions, Lili Mao and Masayori Inouye for providing the *E. coli* BL Δ ompA Δ ompF strain, and Hammad Naveed and Jie Liang for providing their relative melting temperature code. This research was supported by the National Institutes of Health under Ruth L. Kirschstein National Research Service Award F32-GM-099291 from the National Institute of General Medical Sciences (to J.A.S.), and by R01-GM-089949.

1. Wimley WC (2003) The versatile β -barrel membrane protein. *Curr Opin Struct Biol* 13(4):404–411.
2. Daugherty PS, Olsen MJ, Iverson BL, Georgiou G (1999) Development of an optimized expression system for the screening of antibody libraries displayed on the Escherichia coli surface. *Protein Eng* 12(7):613–621.
3. Chen M, Khalid S, Sansom MS, Bayley H (2008) Outer membrane protein G: Engineering a quiet pore for biosensing. *Proc Natl Acad Sci USA* 105(17):6272–6277.
4. Manrao EA, et al. (2012) Reading DNA at single-nucleotide resolution with a mutant MspA nanopore and phi29 DNA polymerase. *Nat Biotechnol* 30(4):349–353.
5. Mohammad MM, Howard KR, Movileanu L (2011) Redesign of a plugged beta-barrel membrane protein. *J Biol Chem* 286(10):8000–8013.
6. Mohammad MM, et al. (2012) Engineering a rigid protein tunnel for biomolecular detection. *J Am Chem Soc* 134(22):9521–9531.
7. Stoddart D, et al. (2010) Nucleobase recognition in ssDNA at the central constriction of the α -hemolysin pore. *Nano Lett* 10(9):3633–3637.
8. Butler TZ, Pavlenok M, Derrington IM, Niederweis M, Gundlach JH (2008) Single-molecule DNA detection with an engineered MspA protein nanopore. *Proc Natl Acad Sci USA* 105(52):20647–20652.
9. Cherf GM, et al. (2012) Automated forward and reverse ratcheting of DNA in a nanopore at 5-Å precision. *Nat Biotechnol* 30(4):344–348.
10. Clarke J, et al. (2009) Continuous base identification for single-molecule nanopore DNA sequencing. *Nat Nanotechnol* 4(4):265–270.
11. Leaver-Fay A, et al. (2011) ROSETTA3: An object-oriented software suite for the simulation and design of macromolecules. *Methods Enzymol* 487:545–574.
12. Richter F, Leaver-Fay A, Khare SD, Bjelic S, Baker D (2011) De novo enzyme design using Rosetta3. *PLoS One* 6(5):e19230.
13. Kuhlman B, et al. (2003) Design of a novel globular protein fold with atomic-level accuracy. *Science* 302(5649):1364–1368.
14. Dantas G, Kuhlman B, Callender D, Wong M, Baker D (2003) A large scale test of computational protein design: Folding and stability of nine completely redesigned globular proteins. *J Mol Biol* 332(2):449–460.
15. Koga N, et al. (2012) Principles for designing ideal protein structures. *Nature* 491(7423):222–227.
16. Whitley P, Nilsson I, von Heijne G (1994) De novo design of integral membrane proteins. *Nat Struct Biol* 1(12):858–862.
17. Choma C, Gratkowski H, Lear JD, DeGrado WF (2000) Asparagine-mediated self-association of a model transmembrane helix. *Nat Struct Biol* 7(2):161–166.
18. Cordova JM, Noack PL, Hilcove SA, Lear JD, Ghirlanda G (2007) Design of a functional membrane protein by engineering a heme-binding site in glycophorin A. *J Am Chem Soc* 129(3):512–518.
19. Korendovych IV, et al. (2010) De novo design and molecular assembly of a transmembrane diporphyrin-binding protein complex. *J Am Chem Soc* 132(44):15516–15518.
20. Joh NH, et al. (2014) De novo design of a transmembrane Zn²⁺-transporting four-helix bundle. *Science* 346(6216):1520–1524.
21. Kleinschmidt JH, den Blaauwen T, Driessen AJ, Tamm LK (1999) Outer membrane protein A of Escherichia coli inserts and folds into lipid bilayers by a concerted mechanism. *Biochemistry* 38(16):5006–5016.
22. Huysmans GH, Baldwin SA, Brockwell DJ, Radford SE (2010) The transition state for folding of an outer membrane protein. *Proc Natl Acad Sci USA* 107(9):4099–4104.
23. Struyvé M, Moons M, Tommassen J (1991) Carboxy-terminal phenylalanine is essential for the correct assembly of a bacterial outer membrane protein. *J Mol Biol* 218(1):141–148.
24. de Cock H, Struyvé M, Kleerebezem M, van der Krift T, Tommassen J (1997) Role of the carboxy-terminal phenylalanine in the biogenesis of outer membrane protein PhoE of Escherichia coli K-12. *J Mol Biol* 269(4):473–478.
25. Robert V, et al. (2006) Assembly factor Omp85 recognizes its outer membrane protein substrates by a species-specific C-terminal motif. *PLoS Biol* 4(11):e377.
26. Sklar JG, Wu T, Kahne D, Silhavy TJ (2007) Defining the roles of the periplasmic chaperones SurA, Skp, and DegP in Escherichia coli. *Genes Dev* 21(19):2473–2484.
27. Kim S, et al. (2007) Structure and function of an essential component of the outer membrane protein assembly machine. *Science* 317(5840):961–964.
28. Hagan CL, Kim S, Kahne D (2010) Reconstitution of outer membrane protein assembly from purified components. *Science* 328(5980):890–892.
29. Hagan CL, Silhavy TJ, Kahne D (2011) β -Barrel membrane protein assembly by the Bam complex. *Annu Rev Biochem* 80(1):189–210.
30. Heuck A, Schleiffer A, Clausen T (2011) Augmenting β -augmentation: Structural basis of how BamB binds BamA and may support folding of outer membrane proteins. *J Mol Biol* 406(5):659–666.
31. Wiener MC, White SH (1991) Fluid bilayer structure determination by the combined use of x-ray and neutron diffraction. II. “Composition-space” refinement method. *Biophys J* 59(1):174–185.
32. Schulz GE (2002) The structure of bacterial outer membrane proteins. *Biochim Biophys Acta* 1565(2):308–317.
33. Senes A, et al. (2007) E(z), a depth-dependent potential for assessing the energies of insertion of amino acid side-chains into membranes: Derivation and applications to determining the orientation of transmembrane and interfacial helices. *J Mol Biol* 366(2):436–448.
34. Hsieh D, Davis A, Nanda V (2012) A knowledge-based potential highlights unique features of membrane α -helical and β -barrel protein insertion and folding. *Protein Sci* 21(1):50–62.
35. Jackups R, Jr, 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(4):979–993.
36. Slusky JS, Dunbrack RL, Jr (2013) Charge asymmetry in the proteins of the outer membrane. *Bioinformatics* 29(17):2122–2128.
37. Naveed H, Xu Y, Jackups R, Jr, Liang J (2012) Predicting three-dimensional structures of transmembrane domains of β -barrel membrane proteins. *J Am Chem Soc* 134(3):1775–1781.
38. Geula S, Naveed H, Liang J, Shoshan-Barmatz V (2012) Structure-based analysis of VDAC1 protein: Defining oligomer contact sites. *J Biol Chem* 287(3):2179–2190.
39. Naveed H, et al. (2012) Engineered oligomerization state of OmpF protein through computational design decouples oligomer dissociation from unfolding. *J Mol Biol* 419(1–2):89–101.
40. Zhao X, et al. (2014) Characterization of a computationally designed water-soluble human μ -opioid receptor variant using available structural information. *Anesthesiology* 121(4):866–875.
41. Perez-Aguilar JM, et al. (2013) A computationally designed water-soluble variant of a G-protein-coupled receptor: The human mu opioid receptor. *PLoS One* 8(6):e66009.
42. Slovic AM, Kono H, Lear JD, Saven JG, DeGrado WF (2004) Computational design of water-soluble analogues of the potassium channel KcsA. *Proc Natl Acad Sci USA* 101(7):1828–1833.
43. Klose M, Jähnig F, Hindennach I, Henning U (1989) Restoration of membrane incorporation of an Escherichia coli outer membrane protein (OmpA) defective in membrane insertion. *J Biol Chem* 264(36):21842–21847.
44. Koebnik R (1999) Membrane assembly of the Escherichia coli outer membrane protein OmpA: Exploring sequence constraints on transmembrane beta-strands. *J Mol Biol* 285(4):1801–1810.
45. Arnold T, Poynor M, Nussberger S, Lupas AN, Linke D (2007) Gene duplication of the eight-stranded beta-barrel OmpX produces a functional pore: A scenario for the evolution of transmembrane beta-barrels. *J Mol Biol* 366(4):1174–1184.
46. Pautsch A, Schulz GE (1998) Structure of the outer membrane protein A transmembrane domain. *Nat Struct Biol* 5(11):1013–1017.
47. Koebnik R, Locher KP, Van Gelder P (2000) Structure and function of bacterial outer membrane proteins: Barrels in a nutshell. *Mol Microbiol* 37(2):239–253.
48. Qu J, Behrens-Kneip S, Holst O, Kleinschmidt JH (2009) Binding regions of outer membrane protein A in complexes with the periplasmic chaperone Skp. A site-directed fluorescence study. *Biochemistry* 48(22):4926–4936.
49. Zhou FX, Cocco MJ, Russ WP, Bronger AT, Engelmann DM (2000) Interhelical hydrogen bonding drives strong interactions in membrane proteins. *Nat Struct Biol* 7(2):154–160.
50. Morona R, Krämer C, Henning U (1985) Bacteriophage receptor area of outer membrane protein OmpA of Escherichia coli K-12. *J Bacteriol* 164(2):539–543.
51. Vaiphei ST, Tang Y, Montelione GT, Inouye M (2011) The use of the condensed single protein production system for isotope-labeled outer membrane proteins, OmpA and OmpX in E. coli. *Mol Biotechnol* 47(3):205–210.
52. 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(31):12735–12740.
53. Summa CM (2002) *Computational Methods and their Applications for de novo Functional Protein Design and Membrane Protein Solubilization* (University of Pennsylvania, Philadelphia).
54. Koebnik R, Krämer L (1995) Membrane assembly of circularly permuted variants of the E. coli outer membrane protein OmpA. *J Mol Biol* 250(5):617–626.
55. Lawrence MS, Phillips KJ, Liu DR (2007) Supercharging proteins can impart unusual resilience. *J Am Chem Soc* 129(33):10110–10112.
56. Moon CP, Fleming KG (2011) Side-chain hydrophobicity scale derived from transmembrane protein folding into lipid bilayers. *Proc Natl Acad Sci USA* 108(25):10174–10177.
57. Jimenez-Morales D, Liang J (2011) Pattern of amino acid substitutions in transmembrane domains of β -barrel membrane proteins for detecting remote homologs in bacteria and mitochondria. *PLoS One* 6(11):e26400.
58. Guruprasad K, Rajkumar S (2000) Beta-and gamma-turns in proteins revisited: A new set of amino acid turn-type dependent positional preferences and potentials. *J Biosci* 25(2):143–156.
59. Hong H, Park S, Jiménez RH, Rinehart D, Tamm LK (2007) Role of aromatic side chains in the folding and thermodynamic stability of integral membrane proteins. *J Am Chem Soc* 129(26):8320–8327.
60. Chamberlain AK, Bowie JU (2004) Asymmetric amino acid compositions of transmembrane beta-strands. *Protein Sci* 13(8):2270–2274.
61. Slovic AM, Summa CM, Lear JD, DeGrado WF (2003) Computational design of a water soluble analog of phospholamban. *Prot Sci* 12(2):337–348.
62. Griffin BA, Adams SR, Tsien RY (1998) Specific covalent labeling of recombinant protein molecules inside live cells. *Science* 281(5374):269–272.
63. Quan J, Tian J (2009) Circular polymerase extension cloning of complex gene libraries and pathways. *PLoS One* 4(7):e6441.
64. Croarkin C, Tobias P, eds (2013) Summary tables of useful fractional factorial designs. *NIST/SEMATECH e-Handbook of Statistical Methods* (Available from www.itl.nist.gov/div898/handbook/).