Outer membrane protein folding and topology from a computational transfer free energy scale

Meishan Lin, Dennis Gessmann, Hammad Naveed, and Jie Liang*

University of Illinois at Chicago, Department of Bioengineering, Chicago, IL, 60607, USA

E-mail: jliang@uic.edu Phone: +1 (312)3551789

Accepted: J Am Chem Soc, 2016

http://dx.doi.org/10.1021/jacs.5b10307

Abstract

Knowledge of the transfer free energy of amino acids from aqueous solution to a lipid bilayer is essential for understanding membrane protein folding and for predicting membrane protein structure. Here we report a computational approach that can calculate the folding free energy of the transmembrane region of outer membrane β-barrel proteins (OMPs) by combining an empirical energy function with a reduced discrete state space model. We quantitatively analyzed the transfer free energies of 20 amino acid residues at the center of the lipid bilayer of OmpLA. Our results are in excellent agreement with the experimentally derived hydrophobicity scales. We further exhaustively calculated the transfer free energies of 20 amino acids at all positions in the TM region of OmpLA. We found that the asymmetry of the gram-negative bacterial outer membrane as well as the TM residues of an OMP determine its functional fold in vivo. Our results suggest that the folding process of an OMP is driven by the lipid-facing residues in its hydrophobic core, and its NC-IN topology is determined by the differential stabilities of OMPs in the asymmetrical outer membrane. The folding free

*To whom correspondence should be addressed

energy is further reduced by lipid A and assisted by general depth-dependent cooperativities that exist between polar and ionizable residues. Moreover, context-dependency of transfer free energies at specific positions in OmpLA predict regions important for protein function as well as structural anomalies. Our computational approach is fast, efficient and applicable to any OMP.

Introduction

Membrane proteins account for about 30% of all proteins in a typical genome[1] and serve a multitude of essential cellular functions such as immune response, metabolite transport and energy transduction.[2] Among the two major classes of membrane proteins, α-helical membrane proteins are predominantly located in the plasma membrane of eukaryotic cells, the inner membranes of eukaryotic organelles and prokaryotes. In contrast, β-barrel membrane proteins, or outer membrane proteins (OMPs), are found in the outer membranes of gram-negative bacteria, mitochondria, and chloroplasts. Both types of membrane proteins are involved in several life-threatening diseases either through altered function or dysfunction.[3] Understanding the principles governing membrane protein folding and stability, as well as identifying their functional form is therefore of fundamental importance.

Despite significant differences in their biogenesis,[4] the insertion of α-helical and β-strand TM segments is dictated by its partition free energy from the aqueous environment into the membrane.[5,6] A central question in membrane protein biology is therefore the assessment of the transfer free energy of amino acids from solution into a lipid bilayer. The transfer free energies of 20 amino acid residues, often called hydrophobicity scales, have been determined in several experimental systems[5–8] and have generate considerable insight. However, experimental measurement of transfer free energy is technically challenging, as identifying conditions of reversible folding can be time-consuming.[7,8] Conditions at which reversible folding were observed have been reported only for three OMPs.[9–13] It is therefore important to develop methods that allow fast generation of effective transfer free energies applicable to

membrane proteins in general.

Computational studies can complement experimental methods and expand our knowledge of the governing principles of membrane protein folding.[14,15] Knowledge-based hydrophobicity scales have been derived in several studies.[16–18] The $E_Z\alpha$ and $E_Z\beta$ empirical potentials can position and orient TM segments, discriminate side-chain decoys, and identify protein-lipid interfaces.[17,18] However, these scales do not consider physical interactions between residues either from neighboring helices/strands or from the same helix/strand. Such interactions are important for membrane protein insertion and folding.[19,20] Detailed molecular interactions in membrane proteins can be investigated using molecular dynamics (MD) simulations.[21–23] However, the choice of the reference unfolded state remains a challenging problem as reproducing transfer free energies of different amino acids require different reference states.[22] In this study, we describe a computational method that incorporates energies of depth-dependent membrane burial, intra- and inter-strand interactions, native as well as exhaustively enumerated non-native conformations with different strand registrations, which allows calculation of the transfer free energy of the TM section of OMPs. Inspired by the Moon-Fleming "whole-protein-scale",[12] we used OmpLA as a client OMP to derive the transfer free energy of 20 amino acid residues at the center of the bilayer. Our results are in excellent agreement with the experimentally derived scales. We have further expanded our studies to include residues at other depths in the bilayer to address key questions in OMP folding. Our findings suggest that the asymmetric nature of the gram-negative bacterial outer membrane as well as the specific amino acid composition in the TM region of an OMP are critical for adopting a functional form in the cell. Our results indicate that lipid-facing residues in the core region provide the driving force for OMP folding. Our method represents a new approach to evaluate the effects of single or multiple residue substitutions on the stability, structure and function of OMPs.

Results

Our goal was to develop a method for high-throughput computation of the transfer free energies of amino acid residues, accounting for key physical interactions in the TM region, positions of residues in the bilayer, and the type of OMP. Our approach is based on the TMSIP empirical energies of burial and inter-strand interactions developed from detailed combinatorial analysis of OMP structures,[24] with further improvement incorporating intra-strand side-chain interactions. The configuration space of OMPs is represented by a discrete state model,[25] in which each strand move up or down several positions from its native position. We enumerate all possible configurations in this reduced discrete state space for the TM β-strands of an OMP, with the probability of a specific configuration following the Boltzmann distribution. We calculated transfer free energies in the context of OMPs from the ensemble of enumerated conformations of the TM β-strands.

A general computational transfer free energy scale of OmpLA

We first examined if our computational method can reproduce the experimentally derived transfer free energy scales. Inspired by the work of Moon and Fleming,[12] we determined the transfer free energies of 20 amino acid residues by substituting the host residue A210 in OmpLA with the remaining 19 amino acid residues. Folding free energy of OmpLA with specific amino acid substitution was calculated and the difference between the wild type Ala and the substitution was assigned as the transfer free energy at the host residue A210. Importantly, key physical interactions in the TM region of OmpLA were taken into account in the calculations (Fig S1). While OmpLA has been reported to form a dimer for enzymatic activity,[26,27] the monomeric form of OmpLA in the outer membrane under normal condition[26] is used in deriving the transfer free energy scale. This computational transfer free energy scale, termed $\Delta\Delta G^{\mathrm{A210}}$, correlates well with the experimentally determined transfer free energy scales ($R^2 = 0.81$ for all 20 amino acids) with the Moon-Fleming scale[12] (Fig 1a).

Compared to Ala, Pro is more hydrophilic in our scale but more hydrophobic in the Moon-Fleming scale. Once Pro is excluded, the correlation between the two scales increases to 0.91. Furthermore, our scale also agrees well with Hessa *et al*'s biological scale ($R^2 = 0.86$ for all amino acids, $R^2 = 0.88$ after excluding Pro), and with the Wimley-White octanol scale ($R^2 = 0.67$ for all amino acids, $R^2 = 0.68$ after excluding Pro). This general agreement indicates that our reduced state model and the empirical energy function can capture key determinants of contribution of the amino acids to the stability of OMPs.

We then derived exhaustively transfer free energy scales using every lipid-facing TM residue as the host residue in OmpLA (52 hosts in addition to A210) (Fig S2 and Table S1). Overall, the transfer free energy values are very similar for residues at the same depth positions in the bilayer of the bacterial outer membrane, and a general transfer free energy scale emerges that is applicable to 41 out of the 53 TM residues. This general transfer free energy scale $\Delta\Delta G_{(i)}$ is then calculated by averaging the scales derived from host residues among the 41 residues that are at the same bilayer depth i (Fig 1b).

Strong context dependency of transfer free energies allows identification of residues important for enzymatic function of OmpLA

Host residues whose transfer free energy scales deviate from the general scale were detected by analyzing the correlation coefficient between the transfer free energy scales derived from host residues at the same depth positions (Table S2). If the average correlation coefficient increases significantly ($> 10\%$) upon excluding the scale of a host residue, this scale strongly deviates from the rest. Overall, we found that scales from 12 host residues, D36, N38, L40, Y42, Q94, H142, N156, P175, V235, L237, V241, and L265, strongly deviated from the remaining 41 host residues (Fig 2a-b). Six of them (D36, N38, L40, P175, L237, and L265) are located in the hydrophobic core region, while the other six are located in the headgroup region.

Except Q94, H142, and N156, all are located in a structurally deformed environment.

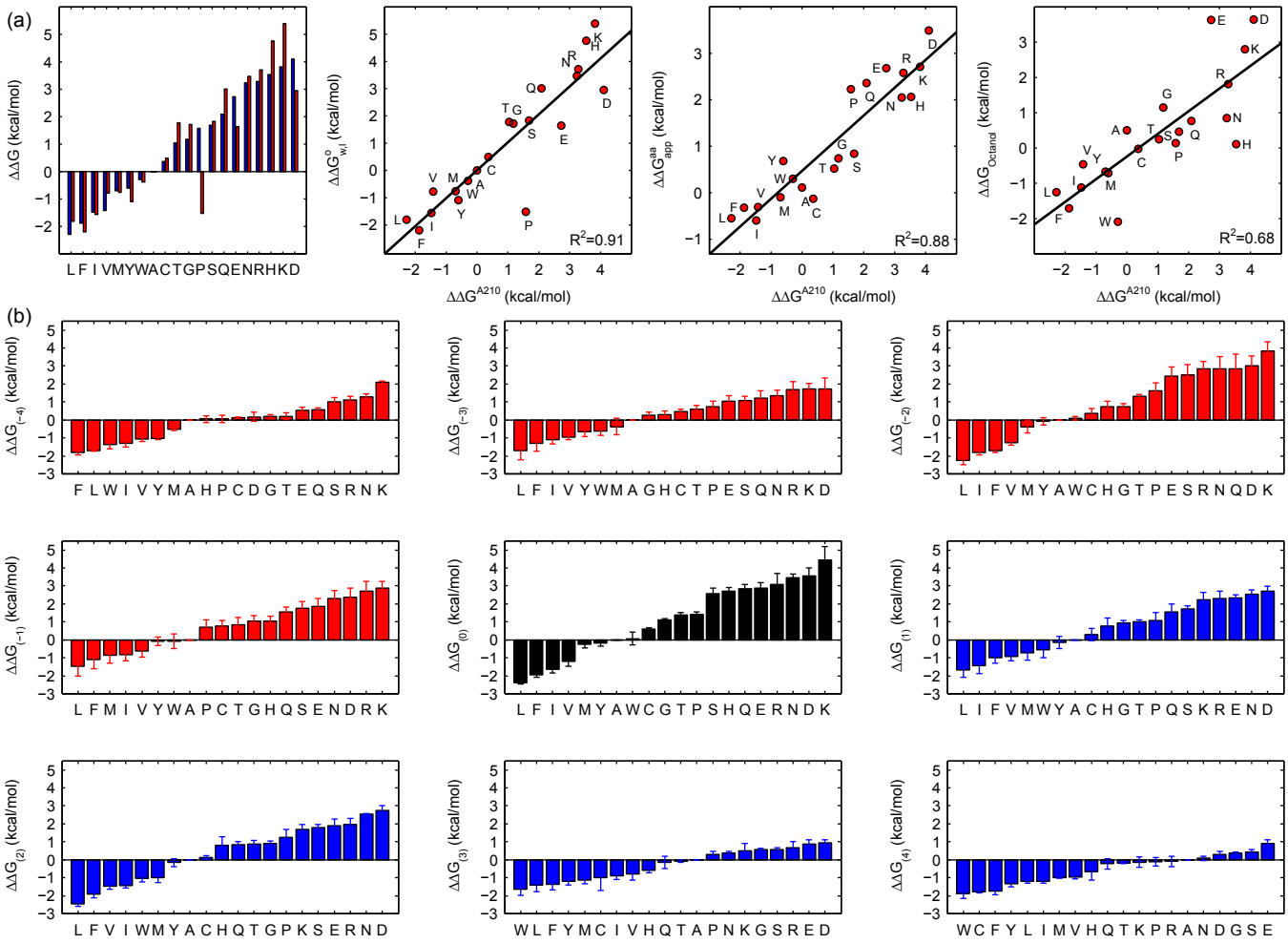


Fig 1. Transfer free energy $\Delta\Delta G^{\mathrm{A210}}$ and general position specific transfer free energy scales $\Delta\Delta G_{(i)}$ were derived using host residues in OmpLA. (a) $\Delta\Delta G^{\mathrm{A210}}$ (blue bar) calculated at host residue A210 of OmpLA was compared with the experimentally measured Moon-Fleming whole-protein scale[12] (red bar). $\Delta\Delta G^{\mathrm{A210}}$ correlates well with the whole protein scale $\Delta\Delta G^{\circ}_{w,l}$ [12] ($R^2 = 0.91$ excluding Pro), the biological scale $\Delta\Delta G^{aa}_{app}$ of Hessa *et al*[28] ($R^2 = 0.88$ excluding Pro), and the Wimley-White octanol scale $\Delta\Delta G_{\mathrm{Octanol}}$ [5] ($R^2 = 0.68$ excluding Pro). (b)$\Delta\Delta G_{(i)}$ are derived for all 9 positions in the TM region. For position i, $\Delta\Delta G$ derived from multiple host residues at this same depth-position are averaged to obtain $\Delta\Delta G_{(i)}$. Host residues with strong context dependency in their transfer free energy scales were excluded from calculation.

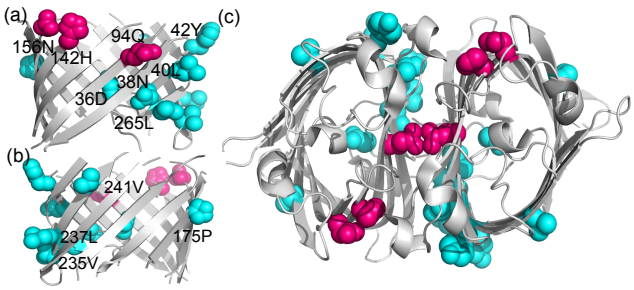


Fig 2. Strong context dependency of transfer free energies detects amino acid residues involved in OmpLA function. Side (a and b) view on and top (c) view on OmpLA dimer. Residues with strong context dependency are either located in a structurally deformed environment (cyan) or are functionally important residues (pink). These residues, except P175, V237, and V241, are located in the dimerization interface of OmpLA.

Residues Q94, H142, and N156 are functionally important residues. Active-site residues H142 and N156 are part of the catalytic triad of OmpLA.[29] Q94, V235, and L265 provide intermolecular interactions for OmpLA dimerization (Fig S4), which is required for enzymatic function.[26] Overall, most of these residues are found in the dimerization interface of OmpLA (Fig 2c). The significant deviation from the general transfer free energy scale implies that our computational hydrophobicity scale can be context dependent for certain host residues. We further suggest that such context dependency can be utilized to detect either possible structural anomaly or assign functional residues in OMPs of unknown tertiary structure.

The inner leaflet of the bacterial outer membrane imposes an energetic barrier to insertion of polar and ionizable amino acid residues

Experimental studies have shown that the energetic cost of transferring amino acids into a lipid bilayer is affected by the composition of the local membrane environment to which they are transferred as well as their depth in the bilayer.[6] As the gram-negative outer membrane is strongly heterogeneous in its composition, with the outer leaflet consisting solely of lipopolysaccharide (LPS) and the inner leaflet comprised of phospholipids,[30] an important question is how this asymmetric environment affects OMP folding.

We first examined the depth-dependent profiles of Leu and Arg and compared them to

the experiment values according to ref.[12] Here we assigned an index i to every host residue, with $i = 0$ corresponding to the center of the bilayer. Hosts located in the outer leaflet have $i > 0$, and hosts located in the inner leaflet have $i < 0$. Fig 3a demonstrates that the energetic cost to transfer Leu into any depth of the asymmetric bacterial OM is comparable to the symmetric DLPC bilayer used by Moon and Fleming.[12] In contrast, transfer of Arg into the outer leaflet (*i.e.*, the LPS containing leaflet) is much more favorable than insertion into the inner leaflet (*i.e.*, the phospholipid containing leaflet), as well as a DLPC leaflet. For example, $\Delta\Delta G$ of Arg at position $i = 2$ is 2.38 kcal/mol, which correlates well with the experimentally measured value of 2.35 kcal/mol. However, Arg at position $i = 3$ or $i = 4$ is much more favorable in an LPS than DLPC environment (0.19 kcal/mol in LPS vs 2.07 kcal/mol in DLPC at $i = 3$; 0.08 kcal/mol in LPS vs 0.61 kcal/mol in DLPC at $i = 4$). Our results therefore capture the effect of the asymmetric composition of the bacterial OM on OMP folding.

We then evaluated the effect of residue depth on the transfer free energy of all 20 amino acid residues. From the position specific general transfer free energy scale $\Delta\Delta G_{(i)}$, we can directly obtain the depth-dependent profiles of all 20 residues $\Delta\Delta G_{aa(i)}$ using an asymmetric Gaussian function (Fig 3b). As expected, aliphatic residues L, I, V, M, and aromatic residues F, Y, W exhibit favorable transfer free energies at all bilayer depths, whereas ionizable and polar residues R, K, H, D, E, N, Q, S, T, as well as G and P show unfavorable transfer free energies throughout the membrane. In general, hydrophobic residues are most favorable at the center of the bilayer, whereas hydrophilic residues are most unfavorable at the center of the membrane, consistent with previous experimental study.[6] However, we found that ionizable residues, R and K, and polar residues, N, Q, S and T, have higher free energy cost in the inner leaflet than the outer leaflet (Fig 3b), thus, exhibiting asymmetric depth-dependency. These findings suggest that the inner leaflet of the bacterial OM imposes an energetic barrier to the insertion of polar and ionizable residues, whereas the outer leaflet favors transfer of these residues.

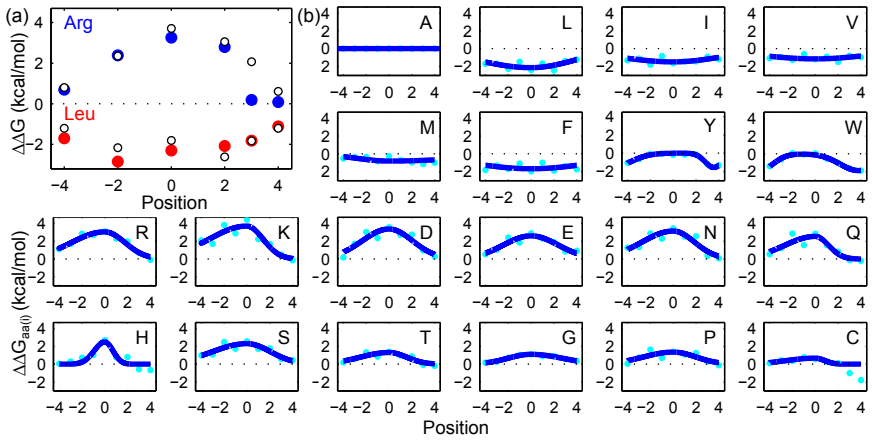


Fig 3. Depth-dependent profiles of amino acid transfer free energies exhibit asymmetric distribution in the bacterial OM. (a) $\Delta\Delta G$s of Leu (red closed circle) calculated using six host residues (L120, L164, A210, G212, Y214, and A223) of OmpLA are similar to measured values (open circle).[12] However, $\Delta\Delta G$s of Arg (blue closed circle) are lower than measured values at outer leaflet ($i > 0$) of the bacterial OM, suggesting LPS favors insertion of Arg. (b) $\Delta\Delta G_{aa(i)}$ is plotted against the position index i for each individual residue type (cyan dots). Ionizable residues Arg and Lys, polar residues Gln, Asn, Ser, and Thr, aromatic residues Tyr and Trp have lower $\Delta\Delta G_{aa(i)}$ in the outer leaflet ($i > 0$) compared to that in the inner leaflet ($i < 0$). Blue lines are fitted single or double asymmetric Gaussian curves. i is set to 0 at the bilayer center. It increases towards outer leaflet headgroup, and decreases towards the inner leaflet headgroup. Details of $\Delta\Delta G_{aa(i)}$ can be found in Table S3.

Presence of Lipid A lowers the folding free energy of OMPs

The asymmetry in the depth-dependent profiles of ionizable and polar amino acid residues demonstrated a clear correlation between the transfer free energies of these residues and the heterogeneity of the bacterial OM. To assess the effect of this phenomenon on the folding of bacterial OMPs, we built a model of a symmetric OM by implicitly substituting lipid A, the membrane part of LPS, in the outer leaflet with the phospholipids of the inner leaflet (Fig 4). Symmetric membranes are biologically relevant, as bacterial OMPs are capable of folding into the symmetric lipid bilayer of the mitochondrial OM.[31] We approximated the folding free energy of OMPs with the summation of the transfer free energy of lipid-facing TM residues of the OMPs and evaluated the effect of the asymmetric composition of the OM on the stability of OMPs. For the asymmetric bacterial OM, we used the depth-dependent profiles $\Delta\Delta G_{aa(i)}$ to calculate the total transfer free energy ($\Delta\Delta G_{\mathrm{asym}}$) of a given client OMP. For the symmetric OM model, we first modified the asymmetric depth-dependent profiles to a symmetric profile $\Delta\Delta G^{*}_{aa(i)}$, in which the values in the outer leaflet ($i > 0$) were replaced by those in the inner leaflet ($i < 0$), and the respective total transfer free energy of a client OMP ($\Delta\Delta G_{\mathrm{sym}}$) were then calculated.

With the exception of OmpA, all OMPs are more stable in the asymmetric membrane by an $\mid \Delta\Delta G \mid > 1$kcal/mol (Fig 4). The same results are reached when folding free energies of the TM regions instead of transfer free energies are used. To rule out the possibility that our observations arise from the asymmetry of the empirical potential function, itself derived from OMP structures in asymmetric outer membranes, we evaluated protein stabilities using the only asymmetric term in our energy function, namely, the single body burial term.[24] We found that only 15 out of the 24 OMPs are more stable in the asymmetric membrane with an $\mid \Delta\Delta G \mid > 1$kcal/mol (Fig 4). Therefore, we conclude that the asymmetry in the potential function is insufficient to determine the asymmetric stabilities of OMPs. Rather, physical interactions of TM residues in the ensemble of native and non-native configurations collectively give rise to the asymmetric thermodynamic stabilities. We further conclude that

the presence of LPS in the outer leaflet not only serves the Gram-negative cell as a defense mechanism against solubilization by hydrophobic molecules, but further plays an important role in OMP biogenesis as it contributes to their thermodynamic stability. This conclusion is consistent with a number of experimental findings. For example, PhoE was found to refold more efficiently in the phospholipid/LPS bilayer than in phospholipid/phospholipid bilayer,[32] and LPS accelerates OmpA folding and insertion into lipid vesicles.[33]

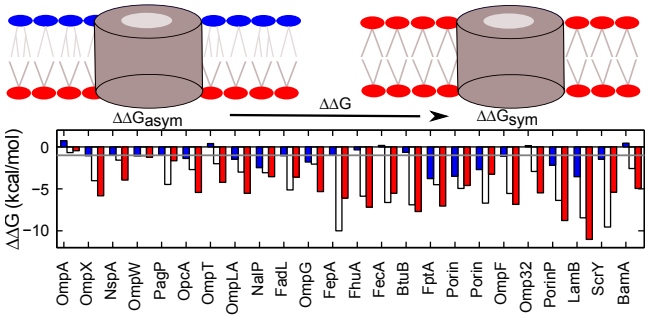


Fig 4. Lipid A contributes to the thermodynamic stability of bacterial OMPs. The stability differences $\Delta\Delta G$s of 24 bacterial OMPs in the asymmetric bacterial outer membrane (OM) and a symmetric model OM were determined. With $\mid \Delta\Delta G \mid >$ 1kcal/mol as the cutoff (grey line), all except OmpA have higher stability in the asymmetric membrane than in the symmetric membrane using either the total transfer free energy of all lipid-facing TM residues (white bar), or the folding free energy of the TM region (red bar). Only 15 OMPs show higher stability in the asymmetric OM if the total single burial energy of all lipid-facing TM residues is used (blue bar).

Differential stabilities of TM regions in determining native topology of OMPs

The observation that LPS contributes to the thermodynamic stability of an OMP leads us to an interesting question about the native topology of OMPs. All bacterial OMPs adopt a membrane topology where the N- and C- terminus are located in the periplasmic space, termed here as the "NC-IN topology". In contrast, the orientation of mitochondrial OMPs in the symmetric mitochondrial OM is highly debated and likely of dual-topology.[34] We examined this issue by first approximating the folding free energy in the TM region following an additive model based on the computed transfer free energies of lipid-facing residues. In

this additive model, the stability of an OMP is calculated as the summation of the transfer free energy of lipid-facing residues in the native configuration. Fig 5 shows that out of 24 representative bacterial OMPs examined, 17 of them exhibit higher stability in the native NC-IN topology.

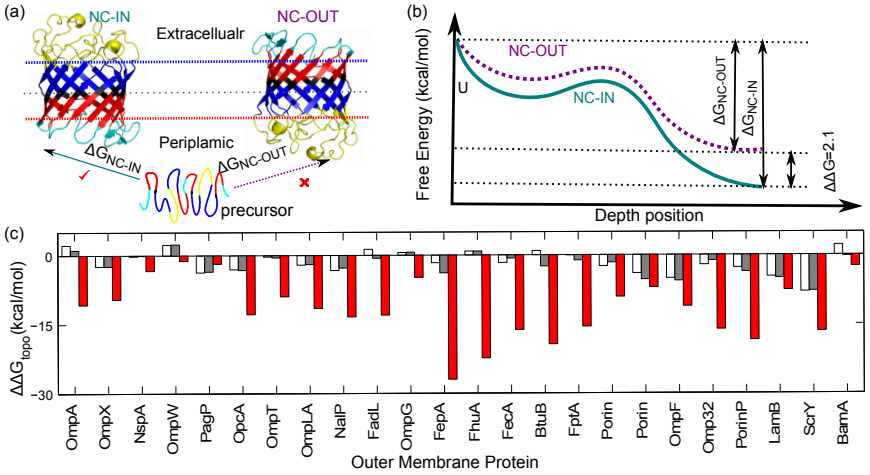


Fig 5. Native bacterial OMP topology is driven by the asymmetric nature of the OM as well as the lipid-facing residues in the head-group region. (a) Precursor sequence of an unfolded OMP inserts into the bacterial OM to either the native NC-IN topology or the non-native NC-OUT topology. OmpLA is used here as an example at different discrete depth position. (b) Difference of folding free energy of OmpLA in two topologies is calculated as the difference of total transfer free energy of all lipid-facing TM residues. (c) Stability differences of OMPs adopting native NC-IN topology and non-native NC-OUT topology are calculated. 17 OMPs have higher stability with NC-IN topology using the total transfer free energy of all lipid-facing TM residues (white bar) under the additive model. All 24 OMPs have higher stability in native NC-IN topology (red bar) using the non-additive model, which directly calculates the folding free energy of the TM region. 20 OMPs show higher stability in native topology if the total single burial energy of all lipid-facing residues is used (grey bar).

As this additive model failed to account for the native topology of 7 bacterial OMPs, we next adopt an alternative approach and directly estimate the thermodynamic stability of OMPs by computing the folding free energy of the TM region. In this case, non-additive effects that depend on local inter- and intra-strand interactions in native and non-native configurations are taken into account. Our results determine that all 24 bacterial OMPs in our analysis have lower folding free energy, thus higher thermodynamic stability, with NC-IN topology than that of NC-OUT topology (Fig 5c). Our findings suggest that the native

NC-IN topology of bacterial OMPs is driven by the interplay of membrane asymmetry and interactions between amino acids in the TM region. Nevertheless, a few OMPs show very small differences in folding free energy between the two topologies. This suggests that in addition to the stability of the TM-region, there are other factors influencing OMP topology. For example, OmpW has long extended β-strands in the extracellular domain to form a hydrophobic channel.[35] PagP has a periplasmic α-helix that out-clamps the barrel.[25,36] These special structural features not incorporated in our model may also contribute to the native topologies of OMPs in the membrane.

Cooperativity among ionizable and polar residues

It is well documented that charged residues in OMPs are preferred in the extracellular side.[24,30] This is different from helical membrane proteins, where charged residues are preferred in the cytoplasmic side.[37] Non-additive effects may play an important role in transferring these amino acid residues across or into the lipid bilayer.[38,39] For example, the cooperative nature of inserting multiple Arg residues has long been of interests in the studies of membrane protein biophysics.[12,40,41] Cooperativities between aromatic residues have also been recently examined.[39] However, the mechanism of inserting or translocating multiple ionizable or polar residues into or across highly hydrophobic membrane bilayer is not clear, and the role of cooperativities between ionizable and polar residues beyond Arg is not known in general. To investigate non-additive effects between residues on insertion into the TM region, we derived the cooperativity between any two ionizable or polar residues at different position of the bilayer. The sum of the energy costs of the single variants at a host pair is compared to the energy cost of the double variant, and the difference quantifies the cooperativity. Basic residues Arg, Lys, and His, acidic residues Asp and Glu along with two highly polar residues Asn and Gln were included in this analysis. We chose two adjacent lipid-facing residues in strands β4, β8, and β9 as the host residue pairs, and derived the cooperativity at the respective positions (Fig 6a). These strands are chosen as their lipid

facing residues do not show strong context dependency. We further averaged cooperativity values of ionizable or polar residue at host pairs of the same position from all three strands.

We found cooperativity between two ionizable or highly polar residues exist at most of the positions (Fig 6b). The strongest cooperativity was observed in pairs located in the inner leaflet core region (position (0, −2)), followed by pairs in the inner leaflet headgroup region (position (−2, −4)). Overall, the cooperativity decreases as the host pair moves away from the core region. Lys-Lys has the highest cooperativity ($\Delta\Delta\Delta G$ = 4.20 kcal/mol) at position (0, −2), while cooperativity in the outer leaflet headgroup region (positions 4, 2) is negligible (Fig 6b). Our results suggest that positive and depth-dependent cooperativity is a general property for ionizable or highly polar residue pairs. Located in the TM or the extracellular loop region, these residues encounter an energetic barrier during insertion into or translocate the hydrophobic lipid bilayer. Due to this cooperativity, clustering of these residues may reduce the energetic cost during their insertion, providing an additional driving force for the folding and stability of OMPs. Coopeartivity may play a similar role in the translocation of Arg-rich cell penetrating peptides.

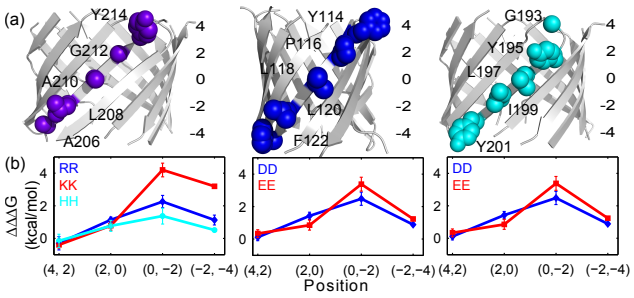


Fig 6. Positive cooperativity is general, asymmetric, and depth-dependent for ionizable or polar residue pairs'. (a) Four host pairs are selected from β9, β4, and β8 to derive the cooperativity. Here $(i+2, i)$ are the locations indices of the host pairs. *i.e.*, Y214-G212 in strand β9 has location index (4, 2). (b) Positive cooperativity was observed for most residue pairs (Table S5). The average cooperativity at a specific position is plotted with its standard error. The strongest cooperativity is observed in the inner leaflet core region at (0, −2).

Spontaneous insertion of OMPs is driven by the lipid-facing residues in the hydrophobic core

It is well known that bacterial OMPs can fold spontaneously without an external energy source to their native state.[42] However, it is unclear what occurs thermodynamically during insertion as the inner leaflet imposes an energetic barrier to insertion of ionizable and polar residues. To address this question, we used a simplified folding model with 9 discrete steps (Fig 7a), based on the concerted folding mechanism proposed by Kleinschmidt *et al.*[43] We calculated the total transfer free energy of lipid-facing residues inserted into the bilayer in each step. While pore-facing residues and their interactions also contribute to the overall protein stabilities, our analysis shows that lipid-facing residues contribute approximately 5 times more that the pore-facing residues in OmpLA. Therefore, we focus on lipid-facing residues that are directly involved in insertion and folding in the membrane environment. To identify the origin of the favorable free energy, we decomposed the overall free energy into contributions from the hydrophobic core residues and from the headgroup residues (Fig 7b).

Indeed, we found that the folding free energy of all investigated OMPs during the insertion of the TM region is overall favorable with $\Delta\Delta G < 0\,(\text{kcal/mol})$ (red line in Fig 7b), in agreement with many studies indicating spontaneous OMP insertion and folding. If the head-group region residues face an energetic barrier during OMP insertion, the hydrophobic core residues compensate their costs by reducing the overall folding free energy, thus enabling spontaneous OMP insertion (Fig 7b). Therefore, we conclude that lipid-facing residues in the hydrophobic core region of OMPs provide the main driving forces for OMP folding into the lipid membrane. Our results suggest that assembly machinery proteins such as BAM complex,[44] or periplasmic chaperons such as skp[45] may accelerate the folding process but are unlikely to affect the overall mechanism of the process. These results are in agreement with the experimental folding study on OmpA by Kleinschmidt *et al.*[43]

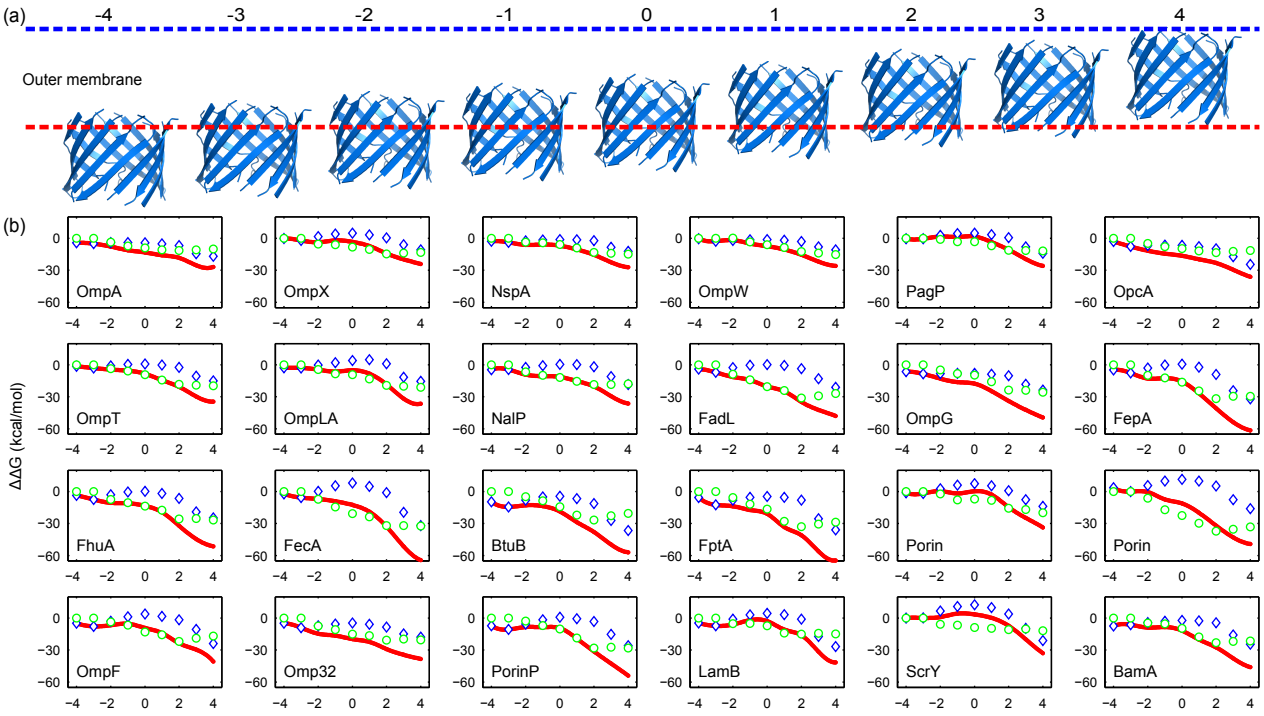


Fig 7. Lipid-facing residues in the hydrophobic core region drive spontaneous OMP insertion. (a) OMP insertion is described by 9 sequential discrete steps. The depth position (−4 to 4) up to where the TM region inserts into is regarded as the reaction coordinate. (b) Folding free energies of bacterial OMPs are approximated as the total transfer free energy of lipid-facing residues inserted into the bilayer (red line). The energetic barrier caused by the lipid-facing residues in the headgroup region (blue diamond) is compensated by the favorable folding free energy of the lipid-facing residues in the hydrophobic core (green circles).

Discussion

Our results demonstrated that the computational model presented here is capable of capturing key factors determining the stability of the TM region of bacterial OMPs and therefore enabled us to reproduce experimentally measured transfer free energy scales for studying thermodynamic properties of bacterial OMPs. While several hydrophobicity scales specific for OMPs have been obtained by converting observed abundance of a TM residue at a particular depth position into empirical free energy,[17,24,30] these empirical scales agree poorly with the whole protein Moon-Fleming 20 amino acid transfer free energy scale (Fig S4a), as the protein-specific details of strong interactions between residues on neighboring strands are not considered in these scales. In contrast, our computational scale has excellent agreement with the Moon-Fleming scale. This is due to the fact that in addition to the single-body burial energy term,[24] we also consider intra- and inter-strand interactions. Furthermore, we examine all energies of configurations of different strand registrations through enumeration, with Boltzmann contributions from both native and non-native configurations treated alike in calculating ensemble properties. Detailed configuration enumeration has been successfully applied to predict and engineer oligomerization states of OMPs,[25,46] to predict protein-protein interactions interface in the TM region,[25] and to predict three-dimensional structures of the TM region of OMPs.[47] By computationally substituting specific residues in the TM region of an OMP, we can determine desired hydrophobicity scales at any position in the lipid bilayer. We used this approach to determine systematically the transfer free energies of 20 amino acids at all possible TM lipid-facing positions of the OMP OmpLA. We noted both the Moon-Fleming scale as well as our computational scale derived in the context of an OMP correlate reasonably well with the biological scale derived in the context of a TM helix ($R^2 = 0.73$ and $R^2 = 0.86$ after excluding Pro, respectively). This suggests that important aspects of the overall thermodynamic properties of residue side-chains are preserved in both types of membrane proteins.

Our results show that there exists a general depth-dependent transfer free energy scale

that is applicable to most lipid-facing positions of OmpLA. We also found that deviations from the general transfer free energy values correlated with lipid-facing amino acid residues important for enzymatic function of OmpLA (*i.e.*, dimerization as well as formation of the active site), or for structural anomalies. Deformed structures are known to be relevant for the function of OMPs. For example, BamA, a conserved OMP responsible for the assembly of OMPs in bacterial outer membrane, induces the thinning of the bilayer around the short β-strands 1 and 16, thus creating a local membrane defect. This defect is hypothesized to facilitate OMP folding.[44] Our computational method can therefore be used to predict the locations of residues that are important for OMP function.

Success in detecting such a residue requires examination of the profile of the 19 transfer free energy values at the position of this residue. A naive approach such as examining the only depth-dependent single-body burial energy term E_b in the potential function does not work, as this term only depends on the depth position and does not contribute to the observed context dependency. While evaluating the full empirical energy of the native structure incorporating additional terms can successfully identify unstable strands implicated in protein-protein interactions,[25] it is inadequate as this approach detects only 2 (D36 and N38 on the weakly stable β1 strand) of the 12 residues (see Fig S5). Neither D36 in the lumen of the barrel nor N38 before the beginning of strand β1 participates in forming β-strands, and are not at lipid-facing positions as would be expected from a canonical model of a TM strand (Fig S1). While our single-body burial energy term E_b is sufficient to detect the structurally deformed local environments of these two residues, overall it is the collective effects of the burial, the intra-strand and inter-strand interaction terms, the ensemble of $7^N = 7^{12}$ native and non-native configurations involving the host strand, as well as the effects of mutating the host position to all 19 other amino acid residues that distinguish these deformed or functional residues. Furthermore, our method can be applied to OMPs with no known structures to identify residues in deformed environment and residues involved in function, as the calculation of transfer free energies requires only knowledge of the TM

sequences.

We also demonstrated that the depth-dependent profiles of amino acids can be used to determine the correct topology of OMPs in the cell outer membrane, and further to evaluate the folding free energy contributed by the TM segments during the insertion of OMPs. While asymmetric distributions of residues in different regions of the membrane bilayer are apparent from statistical analysis,[24] profiles of empirical energy values of Arg and Leu at different depth positions using either potential converted from their asymmetric distributions or the full energy values evaluated on the native structure[25] were unable to accurately reproduce the experimentally measured energy costs of transferring these residues to different positions of the bilayer (Fig S4b). In contrast, profiles of computed transfer free energies of Arg and Leu have excellent agreement with those obtained from experimental measurements. In addition, our results show that the energy cost of ionizable and polar residues in the inner leaflet is higher than that in the outer leaflet, suggesting that the inner leaflet creates an energy barrier for OMP folding. Further analysis showed that lipid-facing TM residues located in the headgroup region need to overcome an energy barrier to translocate through the hydrophobic core region during the insertion process. Experimental folding studies have indeed demonstrated that folding OMPs in the native lipids of *E. Coli* is very inefficient.[44] In vivo, OMPs folding and insertion is accomplished with the existence of BAM complex. BamA may reduce the energy barrier created by the outer membrane through destabilizing the local membrane.[48]

Our analysis also suggests that Lipid A of the LPS in the outer leaflet of bacterial outer membrane contributes to the thermodynamic stability of OMPs, as the folding free energy of the TM region of OMPs is lower in the asymmetric membrane formed by phospholipids inner leaflet and LPS outer leaflet than the folding free energy in the symmetric membrane whose leaflets are both formed by phospholipids. Although the difference in thickness of inner and outer leaflet is not incorporated in this study, LPS outer leaflet with thinner hydrophobic core[49] may further increase the stability of the TM region. In addition, the O-antigen

polysaccharides chain of LPS provides a densely confined space in the extracellular domain of the cell membrane, further increasing the stability of the OMPs.[50] The stabilization effect may be further enhanced by the favorable interaction between the LPS and the extracellular loops.[49,51] Therefore, asymmetric bacterial outer membrane plays an important role in the OMP folding through stabilizing both TM and non-TM region of OMPs.

Approximation of the stability of OMPs in the membrane bilayer using the total transfer free energy of lipid-facing residues was sufficient to capture the preference of OMPs in the asymmetric OM. While asymmetric Gaussian functions were used to represent the depth-dependent transfer free energy profiles of lipid-facing residues, alternative parametric models based on 2nd- and 3rd-degree polynomials also showed higher OMP stabilities in the asymmetric outer membrane (Fig S6). The importance of lipid-facing residues in maintaining the stability of OMPs is also reflected in their substitution patterns. A detailed evolutionary analysis of the substitution rates showed that lipid-facing residues have a conserved pattern of allowed and forbidden substitutions across different OMP families.[52] While pore-facing residues are also under strong selection pressure as they perform important biological functions such as channel conductance, substrate binding, and substrate transport, the pattern of conservation is more specific to individual protein families whose members are of similar functions.

As the asymmetric bacterial outer membrane provides a stabilizing environment to bacterial OMPs, one important question is why such asymmetry is not observed in eukaryote outer membrane. This can be explained by the difference in the biogenesis of bacterial OMPs and eukaryote OMPs. In bacteria, OMP is translocated to the periplasmic space after its synthesis in the cytoplasm.[4] Energy sink in the bacterial outer membrane is necessary to sort the OMPs precursors to the outer membrane.[13] This indicates that OMP folding in bacteria is controlled by the physical nature of the OM. This is critical, as a porin would destroy the proton motive force if spontaneously fold into the inner membrane (IM), resulting in ions and small molecules diffusing through its nonspecific channel, which will lead to cell death.

In mitochondria, on the other hand, OMP folding is already tightly controlled throughout it biogenesis, as folding is directly coupled to OMP synthesis in the cytosol and translocation across the OM by the TOM-SAM supercomplex.[53] This well-controlled process of folding would ensure that OM is the only environment a mitochondria OMP would encounter. Unlike bacterial OMP, energy sink may not be necessary for sorting eukaryotic OMP to the mitochondria OM.

In our study, cooperativity is broadly observed in pairs of ionizable and polar residues when exposed to the lipids. The depth-dependent cooperativity is stronger when the residue pairs are closer to the hydrocarbon core of the bilayer. Studies of the non-additive effect of the translocation of Arg-Arg pair through molecular dynamic simulations by MacCallum *et al*[40] suggest that water-filled defect appears when an isolated Arginine residue is inserted into the center of the hydrocarbon core, which drastically reduces the energetic cost of inserting the second Arg residue.[21,40] Similar observations on water-filled defect were found for TM helix and strand.[22,54,55] Using a continuum mechanical model of membrane, a study on the insertion of a single TM helix suggests that the cooperativity between Arg residues arise from the fact that once membrane bends to accommodate the first charged residue, no further bending is needed for the second charged residue.[56] However, MD simulation studies showed that Arg residue is special, as it is the only ionizable residue that can maintain water-filled defects when placed in the center of the bilayer, whereas such defects dissipate when other ionizable residues are placed close to the bilayer center.[21] It is unclear whether non-additivity of other ionizable residues beyond Arg in the TM segment of β-barrel membrane protein observed in this study is governed by the same mechanism as found in ref.[40] Our study suggests that cooperativity between ionizable residues may partly arise from favorable neighboring interactions in an overall depth-dependent membrane environment. It is possible that such favorable interactions are related mechanistically to water-filled defects and membrane deformation, and contribute to the cooperativities of ionizable residue pairs in β-barrel outer membrane proteins.

We also studied the OMP membrane-insertion and folding as a discrete process based on the reported synchronized translocation of β-hairpins and the concerted folding model of OmpA proposed in reference.[43] We calculated the energetic cost of inserting the OMPs at each step of the insertion process. This model is a highly simplified model and important details such as how membrane would reorganize during each insertion step are not accounted for. Nevertheless, our results enabled a number of useful findings. We found that folding of the TM-region of an OMP into the bacterial outer membrane is energetically favorable, and there is no significantly energy barrier observed during the insertion. This agrees well with the fact that OMPs can fold spontaneously both in vitro and in vivo without any external energy source. The ability of spontaneous folding is also preserved in the OMPs forming oligomers in the outer membrane. In fact, similar folding efficiency of the trimeric OmpF and the engineered monomeric OmpF was observed in in vitro study,[46] which indicates that folding an individual subunit and oligomerization are two independent processes. All these observations support that OMPs behave as autonomous folding domains, in agreement with the viewpoint of Popot and Engelman.[57] Interestingly, spontaneous insertion into bacterial outer membrane is also predicted for mitochondrial outer membrane protein VDAC (Fig S7). This is consistent with experimental *in vivo* refolding study of VDAC into bacterial outer membrane[58] and *in vitro* refolding into PC bilayer.[59]

To summarize, we first validated our method by computationally reproducing the measured transfer free energy scale of Ala210 (Fig 1a) and the depth-dependent profiles of Arg and Leu (Fig 3a).[12] We then applied our method and predicted a new set of transfer free energy scales for all other 52 lipid-facing residues in OmpLA as hosts, which currently have no experimental measurements (Fig 1b). These scales lead to the depth-dependent profiles of the 20 amino acids (Fig 3b), which complement the two known profiles of Arg and Leu.[12] Our method can also predict TM residues either in a structural deformed environment or are related to functions of OmpLA (Fig 2). As only TM sequences are required, our method can be applied to other OMPs without requiring knowledge of their structures.

Another prediction is the general cooperativity between all charged TM residue pairs and their depth-dependency, complementing current knowledge of cooperativity between Arg pairs.[12] Our method also predicts that the asymmetric nature of the bacterial OM stabilizes OMPs (Fig 4), which can be experimentally tested once techniques for generating asymmetric lipid bilayer become more practical. Our method also explains several known facts, including the basis of the correct NC-IN topology of OMPs, as well as the mechanism of the spontaneous folding of OMPs in the bacterial OM. In addition, our method can be used to study the behavior of eukaryotic outer membrane proteins, such as the spontaneous insertion of mitochondrial protein VDAC in bacterial OM.[58,59]

In conclusion, this work introduced a new computational approach to derive transfer free energy scale in the context of an OMP and to evaluate the thermodynamic stability of OMPs, which can be used to reveal important biological insight. Future directions include deriving transfer free energy scales for other OMPs and relate to their properties. As OMPs are found in gram-negative bacteria, mitochondria, and chloroplasts, our method can aid in understanding the general biophysical principles of structure, stability, and function of this important class of membrane proteins. Because of its computational efficiency, our approach can also be useful in *de novo* design of outer membrane proteins as nanodevices for biotechnological applications such as DNA and RNA sequencing[60,61] and single molecule sensing.[62–65] Our current method is designed for β-barrel membrane proteins, as the empirical potential function was developed specifically for outer membrane proteins in their native membrane environments. As the lipid bilayer is not modeled explicitly, the effectiveness of evaluating the thermodynamic properties of OMPs in non-natural membranes with lipid compositions different from that of their native host membranes is unknown. It is possible to generalize our approach to α-helical membrane proteins,[14,15,66,67] which would require the development of an appropriate discrete state model and an enumeration method or an effective sampling method, with empirical potential function specifically constructed for the state model upon removal of confounding effects.[66–68]

Methods

The state space Ω of the native and non-native configurations of the TM region of an OMP is defined following reference,[25] where each strand has 16 residues and can slide up to 3 residues away from its center position. Each TM strand residue interacts with neighboring strand residues through strong H-bond, non H-bond, or weak H-bond following previous studies.[24,69,70] In addition, neighboring TM residues on the same strand facing the same side of the OMPs (lipid-facing or pore-facing) interacts. For a specific configuration $\mathbf{d}$, an empirical strand energy function incorporating single residue burial energy E_{B}, strong H-bond interaction energy E_{SH}, non H-bond interaction energy E_{NH}, weak H-bond interaction energy E_{WH}, and intra-strand interaction energy E_{Intra} with proper weights is used to calculate the energy of the TM region $E(\mathbf{d})$. The folding free energy G_{lipid} of the TM segment of an OMP was then calculated from the partition function Z_{lipid}, which is the summation of the the Boltzmann factors $e^{-\frac{E(\mathbf{d})}{k_B T}}$ of residues in the TM region over all discrete configurations in the state space $G_{\mathrm{lipid}} = -k_B T \ln Z_{\mathrm{lipid}} = -k_B T \ln \sum_{\mathbf{d}\in\Omega} e^{-\frac{E(\mathbf{d})}{k_B T}}$. Given a lipid-facing host residue of OmpLA, the differences in the TM folding free energy of the Ala substitution compared to the other 19 amino acid substitutions were used to construct the transfer free energy scale of that host residue. Following an additive model, the total transfer free energy of lipid-facing residues was used to approximate the folding free energy and the thermodynamic stability of an OMP. The calculated stability of an OMP in the native asymmetric membrane is then compared to the calculated stability in a symmetric membrane. The total transfer free energy was also used to test whether it is sufficient to determine the native NC-IN topology of OMPs. Furthermore, membrane protein insertion was analyzed using the total transfer free energy of lipid-facing residues upon insertion. More details of our method can be found in *SI Method and Material*.

Acknowledgement

This work was supported by NIH grant GM079804, NSF grants DMS-0800257 and DBI-1062328, and Chicago Community Trust grant Catalyst-2013-9.

Supporting Information Available

This material is available free of charge via the Internet at `http://pubs.acs.org/`.

References

(1) Wallin, E.; von Heijne, G. *Protein Sci.* **1998**, *7*, 1029–1038.

(2) Fairman, J.; Noinaj, N.; Buchanan, S. *Curr Opin Struct Biol* **2011**, *21*, 523–531.

(3) Sanders, C.; Nagy, J. *Curr Opin Struct Biol* **2000**, *10*, 438–442.

(4) Hagan, C.; Silhavy, T.; Kahne, D. *Annu Rev Biochem* **2011**, *80*, 189–210.

(5) Wimley, W. C.; Creamer, T. P.; White, S. H. *Biochemistry* **1996**, *35*, 5109–5124.

(6) Hessa, T.; Meindl-Beinker, N. M.; Bernsel, A.; Kim, H.; Sato, Y.; Lerch-Bader, M.; Nilsson, I.; White, S. H.; von Heijne, G. *Nature* **2007**, *450*, 1026–1030.

(7) Moon, C.; Kwon, S.; Fleming, K. *J Mol Biol* **2011**, *413*, 484–494.

(8) Otzen, D.; Andersen, K. *Arch Biochem Biophys* **2013**, *531*, 34–43.

(9) Hong, H.; Tamm, L. *Proc Natl Acad Sci U S A* **2004**, *101*, 4065–4070.

(10) Hong, H.; Szabo, G.; Tamm, L. *Nat Chem Biol* **2006**, *2*, 627–635.

(11) Huysmans, G.; Baldwin, S.; Brockwell, D.; Radford, S. *Proc Natl Acad Sci U S A* **2010**, *107*, 4099–4104.

(12) Moon, C.; Fleming, K. *Proc Natl Acad Sci U S A* **2011**, *108*, 10174–10177.

(13) Moon, C.; Zaccai, N.; Fleming, P.; Gessmann, D.; Fleming, K. *Proc Natl Acad Sci U S A* **2013**, *110*, 4285–4290.

(14) Liang, J. *Curr Opin Chem Biol* **2002**, *6*, 878–884.

(15) Liang, J.; Naveed, H.; Jimenez-Morales, D.; Adamian, L.; Lin, M. *Biochim Biophys Acta* **2012**, *1818*, 927–941.

(16) Koehler, J.; Woetzel, N.; Staritzbichler, R.; Sanders, C.; Meiler, J. *Proteins* **2009**, *76*, 13–29.

(17) Hsieh, D.; Davis, A.; Nanda, V. *Protein Sci.* **2012**, *21*, 50–62.

(18) Schramm, C. A.; Hannigan, B. T.; Donald, J. E.; Keasar, C.; Saven, J. G.; Degrado, W. F.; Samish, I. *Structure* **2012**, *20*, 924–935.

(19) Moore, D.; Berger, B.; DeGrado, W. *Structure* **2008**, *16*, 991–1001.

(20) Hedin, L. E.; Ojemalm, K.; Bernsel, A.; Hennerdal, A.; Illergard, K.; Enquist, K.; Kauko, A.; Cristobal, S.; von Heijne, G.; Lerch-Bader, M.; Nilsson, I.; Elofsson, A. *J. Mol. Biol.* **2010**, *396*, 221–229.

(21) MacCallum, J.; Bennett, W.; Tieleman, D. *J Gen Physiol* **2007**, *129*, 371–377.

(22) Gumbart, J.; Roux, B. *Biophys. J.* **2012**, *102*, 795–801.

(23) Ulmschneider, M.; Ulmschneider, J.; Schiller, N.; Wallace, B.; von Heijne, G.; White, S. *Nat Commun* **2014**, *5*, 4863.

(24) Jackups, R., Jr.; Liang, J. *J Mol Biol* **2005**, *354*, 979–993.

(25) Naveed, H.; Jackups, R., Jr.; Liang, J. *Proc Natl Acad Sci U S A* **2009**, *106*, 12735–12740.

(26) Dekker, N.; Tommassen, J.; Lustig, A.; Rosenbusch, J.; Verheij, H. *J Biol Chem* **1997**, *272*, 3179–3184.

(27) Snijder, H.; Ubarretxena-Belandia, I.; Blaauw, M.; Kalk, K.; Verheij, H.; Egmond, M.; Dekker, N.; Dijkstra, B. *Nature* **1999**, *401*, 717–721.

(28) Hessa, T.; Kim, H.; Bihlmaier, K.; Lundin, C.; Boekel, J.; Andersson, H.; Nilsson, I.; White, S. H.; von Heijne, G. *Nature* **2005**, *433*, 377–381.

(29) Kingma, R.; Fragiathaki, M.; Snijder, H.; Dijkstra, B.; Verheij, H.; Dekker, N.; Egmond, M. *Biochemistry* **2000**, *39*, 10017–10022.

(30) Slusky, J.; Dunbrack, R., Jr. *Bioinformatics* **2013**, *29*, 2122–2128.

(31) Walther, D.; Papic, D.; Bos, M.; Tommassen, J.; Rapaport, D. *Proc Natl Acad Sci U S A* **2009**, *106*, 2531–2536.

(32) Hagge, S.; de Cock, H.; Gutsmann, T.; Beckers, F.; Seydel, U.; Wiese, A. *J Biol Chem* **2002**, *277*, 34247–34253.

(33) Bulieris, P.; Behrens, S.; Holst, O.; Kleinschmidt, J. *J Biol Chem* **2003**, *278*, 9092–9099.

(34) Ujwal, R.; Cascio, D.; Chaptal, V.; Ping, P.; Abramson, J. *Channels (Austin)* **2009**, *3*, 167–170.

(35) Hong, H.; Patel, D.; Tamm, L.; van den Berg, B. *J Biol Chem* **2006**, *281*, 7568–7577.

(36) Ahn, V.; Lo, E.; Engel, C.; Chen, L.; Hwang, P.; Kay, L.; Bishop, R.; Prive, G. *EMBO J* **2004**, *23*, 2931–2941.

(37) von Heijne, G. *J Mol Biol* **1992**, *225*, 487–494.

(38) Hristova, K.; Wimley, W. *J Membr Biol* **2011**, *239*, 49–56.

(39) Hong, H.; Park, S.; Jimenez, R.; Rinehart, D.; Tamm, L. *J Am Chem Soc* **2007**, *129*, 8320–8327.

(40) MacCallum, J.; Bennett, W.; Tieleman, D. *Biophys J* **2011**, *101*, 110–117.

(41) Marks, J.; Placone, J.; Hristova, K.; Wimley, W. *J Am Chem Soc* **2011**, *133*, 8995–9004.

(42) Burgess, N.; Dao, T.; Stanley, A.; Fleming, K. *J Biol Chem* **2008**, *283*, 26748–26758.

(43) Kleinschmidt, J.; den Blaauwen, T.; Driessen, A.; Tamm, L. *Biochemistry* **1999**, *38*, 5006–5016.

(44) Gessmann, D.; Chung, Y.; Danoff, E.; Plummer, A.; Sandlin, C.; Zaccai, N.; Fleming, K. *Proc Natl Acad Sci U S A* **2014**, *111*, 5878–5883.

(45) Denoncin, K.; Schwalm, J.; Vertommen, D.; Silhavy, T.; Collet, J. *Proteomics* **2012**, *12*, 1391–1401.

(46) Naveed, H.; Jimenez-Morales, D.; Tian, J.; Pasupuleti, V.; Kenney, L. J.; Liang, J. *J. Mol. Biol.* **2012**, *419*, 89–101.

(47) Naveed, H.; Xu, Y.; Jackups, R.; Liang, J. *J. Am. Chem. Soc.* **2012**, *134*, 1775–1781.

(48) Noinaj, N.; Kuszak, A.; Gumbart, J.; Lukacik, P.; Chang, H.; Easley, N.; Lithgow, T.; Buchanan, S. *Nature* **2013**, *501*, 385–390.

(49) Wu, E.; Fleming, P.; Yeom, M.; Widmalm, G.; Klauda, J.; Fleming, K.; Im, W. *Biophys J* **2014**, *106*, 2493–2502.

(50) Zhou, H.; Dill, K. *Biochemistry* **2001**, *40*, 11289–11293.

(51) Edrington, T.; Kintz, E.; Goldberg, J.; Tamm, L. *J Biol Chem* **2011**, *286*, 39211–39223.

(52) Jimenez-Morales, D.; Liang, J. *PLoS One* **2011**, *6*, e26400.

(53) Qiu, J.; Wenz, L.; Zerbes, R.; Oeljeklaus, S.; Bohnert, M.; Stroud, D.; Wirth, C.; Ellenrieder, L.; Thornton, N.; Kutik, S.; Wiese, S.; Schulze-Specking, A.; Zufall, N.; Chacinska, A.; Guiard, B.; Hunte, C.; Warscheid, B.; van der Laan, M.; Pfanner, N.; Wiedemann, N.; Becker, T. *Cell* **2013**, *154*, 596–608.

(54) Dorairaj, S.; Allen, T. W. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 4943–4948.

(55) Fleming, P. J.; Freites, J. A.; Moon, C. P.; Tobias, D. J.; Fleming, K. G. *Biochim. Biophys. Acta* **2012**, *1818*, 126–134.

(56) Callenberg, K.; Latorraca, N.; Grabe, M. *J Gen Physiol* **2012**, *140*, 55–68.

(57) Popot, J.; Engelman, D. *Biochemistry* **1990**, *29*, 4031–4037.

(58) Walther, D.; Bos, M.; Rapaport, D.; Tommassen, J. *Mol Biol Evol* **2010**, *27*, 887–895.

(59) Shanmugavadivu, B.; Apell, H.; Meins, T.; Zeth, K.; Kleinschmidt, J. *J Mol Biol* **2007**, *368*, 66–78.

(60) Branton, D.; Deamer, D.; Marziali, A.; Bayley, H.; Benner, S.; Butler, T.; Di Ventra, M.; Garaj, S.; Hibbs, A.; Huang, X.; Jovanovich, S.; Krstic, P.; Lindsay, S.; Ling, X.; Mastrangelo, C.; Meller, A.; Oliver, J.; Pershin, Y.; Ramsey, J.; Riehn, R.; Soni, G.; Tabard-Cossa, V.; Wanunu, M.; Wiggin, M.; Schloss, J. *Nat Biotechnol* **2008**, *26*, 1146–1153.

(61) Ayub, M.; Hardwick, S.; Luisi, B.; Bayley, H. *Nano Lett* **2013**, *13*, 6144–6150.

(62) Gu, L.; Braha, O.; Conlan, S.; Cheley, S.; Bayley, H. *Nature* **1999**, *398*, 686–690.

(63) Braha, O.; Gu, L.; Zhou, L.; Lu, X.; Cheley, S.; Bayley, H. *Nat Biotechnol* **2000**, *18*, 1005–1007.

(64) Rotem, D.; Jayasinghe, L.; Salichou, M.; Bayley, H. *J Am Chem Soc* **2012**, *134*, 2781–2787.

(65) Fahie, M.; Chen, M. *J Phys Chem B* **2015**, *119*, 10198–10206.

(66) Adamian, L.; Jackups, R., Jr.; Binkowski, T.; Liang, J. *J Mol Biol* **2003**, *327*, 251–272.

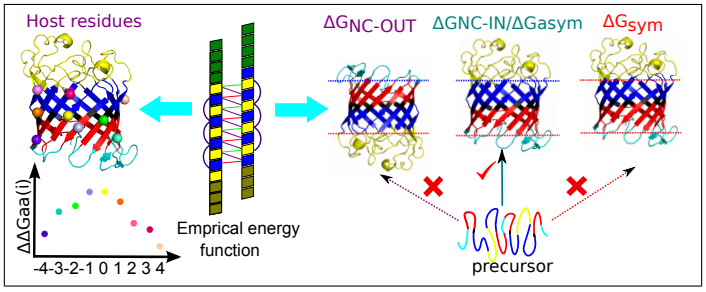
(67) Adamian, L.; Nanda, V.; DeGrado, W.; Liang, J. *Proteins* **2005**, *59*, 496–509.

(68) Jackups, R., Jr.; Liang, J. *IEEE/ACM Trans Comput Biol Bioinform* **2010**, *7*, 524–536.

(69) Wouters, M. A.; Curmi, P. M. *Proteins* **1995**, *22*, 119–131.

(70) Ho, B.; Curmi, P. *J Mol Biol* **2002**, *317*, 291–308.

Graphical TOC Entry



Outer membrane protein folding and topology from a computational transfer free energy scale

Meishan Lin, Dennis Gessmann, Hammad Naveed, and Jie Liang*

University of Illinois at Chicago, Department of Bioengineering, Chicago, IL, 60607, USA

E-mail: jliang@uic.edu

Phone: +1 (312)3551789

Supporting information

Reduced discrete state model of conformational space

The reduce discrete state space Ω for the transmembrane (TM) region of an outer membrane β-barrel protein (OMP) is defined following reference.[1] Briefly, each TM strand has a length $L+1=16$. Each residue on a TM strand interacts with neighboring strands through strong H-bond, non H-bond, or weak H-bond.[2–4] One strand i can slide $d_i \in \{-l, \cdots, 0, \cdots, l\}$ residues away from its center position. A specific conformation of the TM region of an OMP with N TM strands is defined by a vector $\mathbf{d}$:

$$\mathbf{d} = (d_1, \cdots, d_N),$$

in which d_i denotes the offset position of strand i. The full configuration space Ω is:

$$\Omega = \{\mathbf{d} | \mathbf{d} = (d_1, \cdots, d_N) \in \mathbb{Z}^N\}.$$

*To whom correspondence should be addressed

Because each offset position d_i can have a value from $-l$ to l, the size of the configuration space Ω is $(2l+1)^N$. We used $l = 3$ in this study.

TM strands, side chain orientations, and center positions

Sequences of TM strands are extracted from PDB structures of OMPs. The side-chain orientations of TM residues are determined by visualization using pymol.[5] The inter-strand interaction type between two residues on neighboring strands, *i.e.*, strong H-bond, non H-bond and weak H-bond, is determined using the DSSP software.[6] The midplanes of the lipid bilayers of OMPs are taken from the OPM server.[7] The center residue of a TM strand is defined as the residue with the shortest distance to the mid-plane and is assigned the position index $i = 0$. Position indice for other residues in the TM strands are assigned according to their relative positions to the center residue.

Energy calculation of an OMP in a specific configuration

The empirical potential function for calculating the energy of a TM strand is based on our previous empirical MSIP function,[4] which was derived from structures of OMPs using the combinatorial permutation model as the reference state to resolve the coupling effects in short TM strands. For a specific configuration d, the empirical energy function consists of four terms. First, each residue contributes to the single-residue burial energy $E_{\mathrm{B}}(\mathbf{d})$ based on the region in which the residue is located (*i.e.*, hydrophobic core, headgroup, and cap region) and its side-chain orientation (see reference[4] for numerical details). Detailed analysis on the distribution of amino acid residues in different regions of membrane bilayers was described in our previous study.[4] Second, each residue in the TM region interacts with residues in the neighboring strands through strong H-bond interaction (E_{SH}) and non H-bond interaction (E_{NH}) (see reference[4] for numerical details). The summation of E_{SH} and E_{NH} is the inter-strand interaction energy $E_{\mathrm{Inter}}(\mathbf{d})$. Third, each residue in the TM region of a strand interacts with two other residues on the neighboring strand through weak H-

bond ($E_{\mathrm{WH}}(\mathbf{d})$). Fourth, each residue interacts with the nearest residues on the same TM strand, which are of the same side-chain orientation. This term is the intra-strand interaction energy $E_{\mathrm{Intra}}(\mathbf{d})$ (see reference[8] for numerical details) . Following Wouters and Curmi,[2] the strand-strand interaction network is that of the ideal anti-parallel β-sheet (Fig S1). Residues that do not form β-pairing interactions in the native structures usually have less favorable interactions in the canonical model and contribute little in the Boltzmann factor.

The strand energy $E(i;\,\mathbf{d})$ for a TM strand i in configuration $\mathbf{d}$ is then calculated as:

$$
\begin{aligned}
E(i;\,\mathbf{d}) &= w_{\mathrm{B}} \cdot E_{\mathrm{B}}(i;\,\mathbf{d}) + w_{\mathrm{Intra}} \cdot E_{\mathrm{Intra}}(i;\,\mathbf{d}) + w_{\mathrm{Inter}} \cdot E_{\mathrm{Inter}}(i;\,\mathbf{d}) + w_{\mathrm{WH}} \cdot E_{\mathrm{WH}}(i;\,\mathbf{d}) \\
&= w_{\mathrm{B}} \cdot \sum_{k_i} E_{\mathrm{B}}(k_i;\,d_i) + w_{\mathrm{Intra}} \cdot \sum_{k_i} E_{\mathrm{Intra}}(k_i;\,d_i) \\
&\quad + w_{\mathrm{Inter}} \cdot [\sum_{k_i} \sum_{k_{i-1}} (E_{\mathrm{Inter}}(k_i,\,k_{i-1};\,d_i,\,d_{i-1})) + \sum_{k_i} \sum_{k_{i+1}} (E_{\mathrm{Inter}}(k_i,\,k_{i+1};\,d_i,\,d_{i+1}))] \\
&\quad + w_{\mathrm{WH}} \cdot [\sum_{k_i} \sum_{k_{i-1}} (E_{\mathrm{WH}}(k_i,\,k_{i-1};\,d_i,\,d_{i-1}) + \sum_{k_i} \sum_{k_{i+1}} (E_{\mathrm{WH}}(k_i,\,k_{i+1};\,d_i,\,d_{i+1})], \quad (1)
\end{aligned}
$$

where k_i is the position of a residue in strand i, w_{B}, w_{Inter}, w_{WH}, and w_{Intra} are weights for the energy terms. We assigned weight values of 1.0, 0.6, 0.4, and 0.8 for w_{B}, w_{Inter}, w_{WH}, and w_{Intra}, respectively. Our results suggest that position-dependent residue burial term plays the most important role. As long as the rank order of $w_{\mathrm{B}} \geq w_{\mathrm{Intra}}, w_{\mathrm{Inter}} > w_{\mathrm{WH}}$ is maintained, our model is not sensitive to specific values of the weights of these terms. Specifically, 38 sets of weight parameters were tested and the transfer free energy scales were derived at residue A210 using each parameter set. The correlation coefficient (R^2) of the derived scales with the Moon-Fleming scale is 0.81, with a small standard deviation of ± 0.04, indicating our model is robust with respect to the choice of weights. Previous studies have shown that the strength of a weak H-bond is slightly stronger than half of that of a conventional H-bond,[9–13] consistent with an assignment of $w_{\mathrm{Inter}} > w_{\mathrm{WH}}$.

The energy of the TM region of an OMP in a specific configuration $\mathbf{d}$ is then calculated

as the summation of strand energies of all TM strands:

$$E(\mathbf{d}) = \sum_i E(i; \mathbf{d}). \tag{2}$$

Calculating the folding free energy of the TM region of an OMP

Considering each configuration $\mathbf{d}$ as a microstate, the partition function Z_{lipid} of the TM region of an OMP in the lipid bilayer is then calculated by summing the Boltzmann factor of TM strands over all microstates:

$$Z_{\text{lipid}} = \sum_{\mathbf{d}\in\Omega} e^{-\frac{E(\mathbf{d})}{k_B T}}. \tag{3}$$

The free energy of the protein in the lipid bilayer is then computed as:

$$G_{\text{lipid}} = -k_B T \ln Z_{\text{lipid}}. \tag{4}$$

The folding free energy of the TM region of a protein is therefore:

$$\Delta G = G_{\text{water}} - G_{\text{lipid}}, \tag{5}$$

where G_{water} is assign a constant C. That is, the free energy of the ensemble of conformations of an OMP in the solution phase is not hugely affected by a point mutation. This assumption is reasonable, as OMPs in solution are unlikely to be in an extended linear form. For example, a significant amount of residual structures were reported for denatured OmpX by NMR and MD studies, even under the strong condition of 8 M urea.[14,15]

Calculating transfer free energy scale of a host residue in OmpLA

Following the work of reference,[16] we take differences in the ΔG of a mutant at the host residue of OmpLA and the ΔG of Ala at the same host residue as the transfer free energy

from water to bilayer $\Delta\Delta G$ of the mutated amino acid.

$$\Delta\Delta G^{\text{Host}}_{\text{mutant}} = \Delta G^{\text{Host}}(\text{Ala}) - \Delta G^{\text{Host}}(\text{mutant}). \tag{6}$$

We further simplify Equation 6 using Equation 5:

$$\begin{aligned}
\Delta\Delta G^{\text{Host}}_{\text{mutant}} &= \Delta G^{\text{Host}}(\text{Ala}) - \Delta G^{\text{Host}}(\text{mutant}) \\
&= [G^{\text{Host}}_{\text{water,Ala}} - G^{\text{Host}}_{\text{lipid,Ala}}] - [G^{\text{Host}}_{\text{water,mutant}} - G^{\text{Host}}_{\text{lipid,mutant}}] \\
&= G_{\text{lipid,mutant}} - G_{\text{lipid,Ala}},
\end{aligned} \tag{7}$$

assuming $G^{\text{Host}}_{\text{water,Ala}} = G^{\text{Host}}_{\text{water,mutant}} = C$. That is, the free energy of the ensemble of conformations of an OMP in the solution phase is not hugely affected by a point mutation. This assumption is reasonable as OMPs in solution are unlikely to be in an extended linear form. For example, a significant amount of residual structures were reported for denatured OmpX by NMR and MD studies, even under the strong condition of 8 M urea.[14,15]

Detecting transfer free energy scales with strong context dependency

For n_i number of host residues at the depth position i whose transfer free energy scales have been calculated, we further calculate the pairwise correlation coefficients between scales obtained for each of the the $\binom{n_i}{2}$ pair of residues, and then take the average. We then exclude each of the k-th residue in turn and recalculate the average correlation among the $\binom{n_i-1}{2}$ scales. If this new averaged correlation coefficient increases by $\geq 10\%$, we do not include this k-th residue in the calculation of the final averaged general scale. Therefore, both the average scale and the identification of deviating residues do not depend on the sequence in which residues were chosen for calculation.

Calculating depth-dependent transfer free energy profiles of amino acid residues

All depth-dependent transfer free energy profiles $\{\Delta\Delta G_{aa(i)}\}$ of amino acid residues, except those for Tyr and Trp, are approximated by an asymmetric stepwise Gaussian function:

$$\Delta\Delta G_{aa(i)} = \begin{cases} a_0 e^{(-i^2/2a_1^2)}, & \text{if } i \geq 0 \\ a_0 e^{(-i^2/2a_2^2)}, & \text{if } i < 0 \end{cases} \tag{8}$$

where $i \in \{-p, \cdots, 0, \cdots, p\}$ denotes the position index of the TM residue, with $i = 0$ corresponding to the center of the bilayer. We used $p = 4$ in this study. For Trp and Tyr, a double Gaussian function is used:

$$\Delta\Delta G_{aa(i)} = a_0 e^{(-(i-a_2)^2/2a_1^2)} + a_0 e^{(-(i+a_4)^2/2a_3^2)}. \tag{9}$$

The fitted $\Delta\Delta G_{aa(i)}$ values and the values of a_0, a_1, a_2, a_3, a_4 can be found in Table S3.

Comparing thermodynamic stability of OMPs in the asymmetric outer membrane and in the hypothetical symmetric outer membrane

The thermodynamic stability of the TM region of an OMP is approximated by the total transfer free energy of the lipid-facing TM residues. Denote the amino acid residue in the j-th position on the i-th strand as $A_{i,j}$, the stability of an OMP in the asymmetric outer membrane is calculated as:

$$\Delta\Delta G_{\text{asym}} = \sum_{i=1}^{N} \sum_{j=-p}^{p} \Delta\Delta G_{A_{i,j}(j)}, \tag{10}$$

where N is the number of strands in the OMP, and p is defined as above.

The stability of an OMP in the hypothetical symmetric outer membrane is calculated as:

$$\Delta\Delta G_{\text{sym}} = \sum_{i=1}^{N} \sum_{j=-p}^{p} \Delta\Delta G^{*}_{A_{i,\,j}(j)}, \tag{11}$$

with

$$\Delta\Delta G^{*}_{A_{i,\,j}(j)} = \begin{cases} \Delta\Delta G_{A_{i,\,j}(j)}, & \text{if } j \leq 0 \\ \Delta\Delta G_{A_{i,\,j}(-j)}. & \text{if } j > 0 \end{cases} \tag{12}$$

By subtracting Equation 10 from Equation 11, we obtain the difference in thermodynamic stability between the TM region of an OMP in the native asymmetric outer membrane and in the hypothetical symmetric outer membrane.

Comparing thermodynamic stabilities of OMPs in NC-IN and NC-OUT topologies

Additive model. When an additive model is used, we calculate the thermodynamic stability of the TM region as the total transfer free energy of lipid-facing TM residues of a given OMP. We denote $A_{i,\,j}$ as the amino acid in the j-th position of the i-th strand. The thermodynamic stability of an OMP in the NC-IN topology is calculated as:

$$\Delta\Delta G_{\text{NC-IN}} = \sum_{i=1}^{N} \sum_{j=-p}^{p} \Delta\Delta G_{A_{i,\,j}(j)} \cdot \delta_{ij}, \tag{13}$$

with

$$\delta_{ij} = \begin{cases} 1, & \text{if } A_{i,\,j} \text{ is lipid facing,} \\ 0, & \text{if } A_{i,\,j} \text{ is pore facing,} \end{cases} \tag{14}$$

Similarity, the thermodynamic stability of an OMP in NC-OUT topology is calculated as:

$$\Delta\Delta G_{\text{NC-OUT}} = \sum_{i=1}^{N} \sum_{j=-p}^{p} \Delta\Delta G_{A_{i,\,j}(-j)} \cdot \delta_{ij}, \tag{15}$$

with

$$\delta_{ij} = \begin{cases} 1, & \text{if } A_{i,j} \text{ is lipid facing,} \\ 0. & \text{if } A_{i,j} \text{ is pore facing.} \end{cases} \tag{16}$$

By subtracting Equation 13 from Equation 15, we obtain the difference in stability when adopting the two different topologies:

$$\begin{aligned} \Delta\Delta G_{\text{topo}} &= \Delta\Delta G_{\text{NC-OUT}} - \Delta\Delta G_{\text{NC-IN}} \\ &= \begin{cases} \geq 0, & \text{OMPs in NC-IN topology,} \\ < 0, & \text{OMPs in NC-OUT topology.} \end{cases} \end{aligned} \tag{17}$$

The predicted topology is chosen to be the one with lower $\Delta\Delta G$.

Nonadditive model. Alternatively, we directly calculate the folding free energies of the TM region of an OMP using Equation 5 by directly summing up the Boltzman factors of residues in the TM regions over all enumerated miscrostates. The difference in thermodynamic stability $\Delta\Delta G_{\text{topo}}$ between the two topologies is then:

$$\begin{aligned} \Delta\Delta G_{\text{topo}} &= \Delta G_{\text{NC-OUT}} - \Delta G_{\text{NC-IN}} \\ &= \begin{cases} \geq 0, & \text{OMPs in NC-IN topology,} \\ < 0, & \text{OMPs in NC-OUT topology.} \end{cases} \end{aligned} \tag{18}$$

Calculating cooperativity of residue pairs

The cooperativity of two residues at the positions (WT, WT) of a pair of host positions in OmpLA is calculated as:[16]

$$\Delta\Delta\Delta G^{\text{Host Pair}}_{\text{mutant}} = \Delta\Delta G^{\text{Host Pair}}_{(\text{WT, mutant})} + \Delta\Delta G^{\text{Host Pair}}_{(\text{mutant, WT})} - \Delta\Delta G^{\text{Host Pair}}_{(\text{mutant, mutant})}. \quad (19)$$

References

(1) Naveed, H.; Jackups, R., Jr.; Liang, J. *Proc Natl Acad Sci U S A* **2009**, *106*, 12735–12740.

(2) Wouters, M. A.; Curmi, P. M. *Proteins* **1995**, *22*, 119–131.

(3) Ho, B.; Curmi, P. *J Mol Biol* **2002**, *317*, 291–308.

(4) Jackups, R., Jr.; Liang, J. *J Mol Biol* **2005**, *354*, 979–993.

(5) Schrödinger, LLC,

(6) Kabsch, W.; Sander, C. *Biopolymers* **1983**, *22*, 2577–2637.

(7) Lomize, M. A.; Lomize, A. L.; Pogozheva, I. D.; Mosberg, H. I. *Bioinformatics* **2006**, *22*, 623–625.

(8) Miyazawa, S.; Jernigan, R. *J Mol Biol* **1996**, *256*, 623–644.

(9) Vargas, R.; Garza, J.; Dixon, D. A.; Hay, B. P. *Journal of the American Chemical Society* **2000**, *122*.

(10) Scheiner, S.; Kar, T.; Gu, Y. *J. Biol. Chem.* **2001**, *276*, 9832–9837.

(11) Scheiner, S. *J Phys Chem B* **2005**, *109*, 16132–16141.

(12) Scheiner, S. *J Phys Chem B* **2006**, *110*, 18670–18679.

(13) Bowie, J. U. *Curr. Opin. Struct. Biol.* **2011**, *21*, 42–49.

(14) Tafer, H.; Hiller, S.; Hilty, C.; Fernandez, C.; Wuthrich, K. *Biochemistry* **2004**, *43*, 860–869.

(15) Krautler, V.; Hiller, S.; Hunenberger, P. *Eur Biophys J* **2010**, *39*, 1421–1432.

(16) Moon, C.; Fleming, K. *Proc Natl Acad Sci U S A* **2011**, *108*, 10174–10177.

(17) Kingma, R.; Fragiathaki, M.; Snijder, H.; Dijkstra, B.; Verheij, H.; Dekker, N.; Egmond, M. *Biochemistry* **2000**, *39*, 10017–10022.

(18) Li, X.; Hu, C.; Liang, J. *Proteins* **2003**, *53*, 792–805.

(19) Hsieh, D.; Davis, A.; Nanda, V. *Protein Sci.* **2012**, *21*, 50–62.

(20) Slusky, J.; Dunbrack, R., Jr. *Bioinformatics* **2013**, *29*, 2122–2128.

(21) Hessa, T.; Kim, H.; Bihlmaier, K.; Lundin, C.; Boekel, J.; Andersson, H.; Nilsson, I.; White, S. H.; von Heijne, G. *Nature* **2005**, *433*, 377–381.

(22) Wimley, W. C.; Creamer, T. P.; White, S. H. *Biochemistry* **1996**, *35*, 5109–5124.

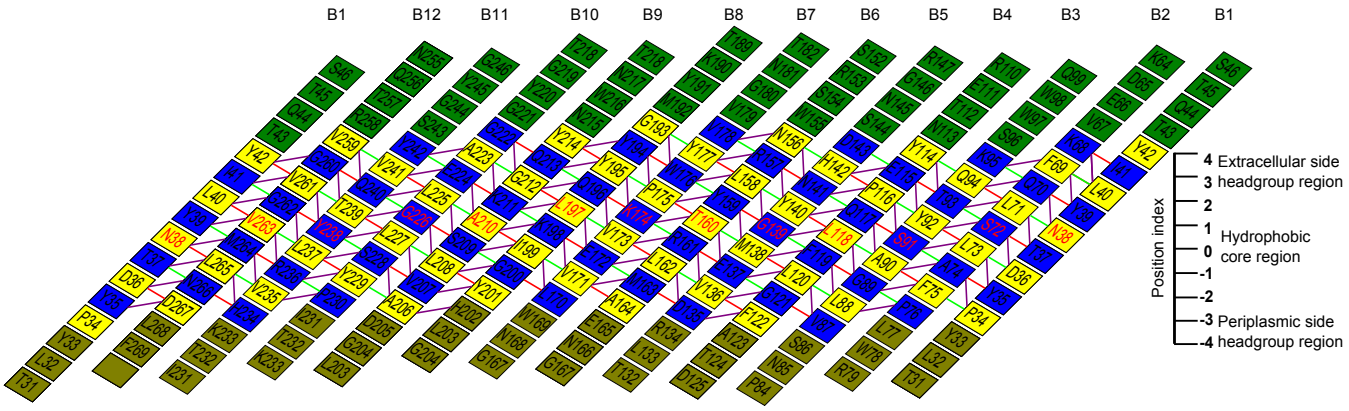


Fig. S1. The discrete model of the transmembrane region of outer membrane phospholipase A (OmpLA). In this model, a β-strand has 16 residues: 4 residues in the extracellular cap region (green), 9 residues in the TM region (pore-facing residues in blue, lipid-facing residues in yellow), and 3 residues in the periplasmic cap region (olive). Position index i of TM residues as integers are listed on the vertical line. Residues at position $i = 0$ are highlighted in red. Important inter-strand interactions between two neighboring TM residues are drawn in red (strong H-bond), green (non H-bond), and brown (weak H-bond) lines.

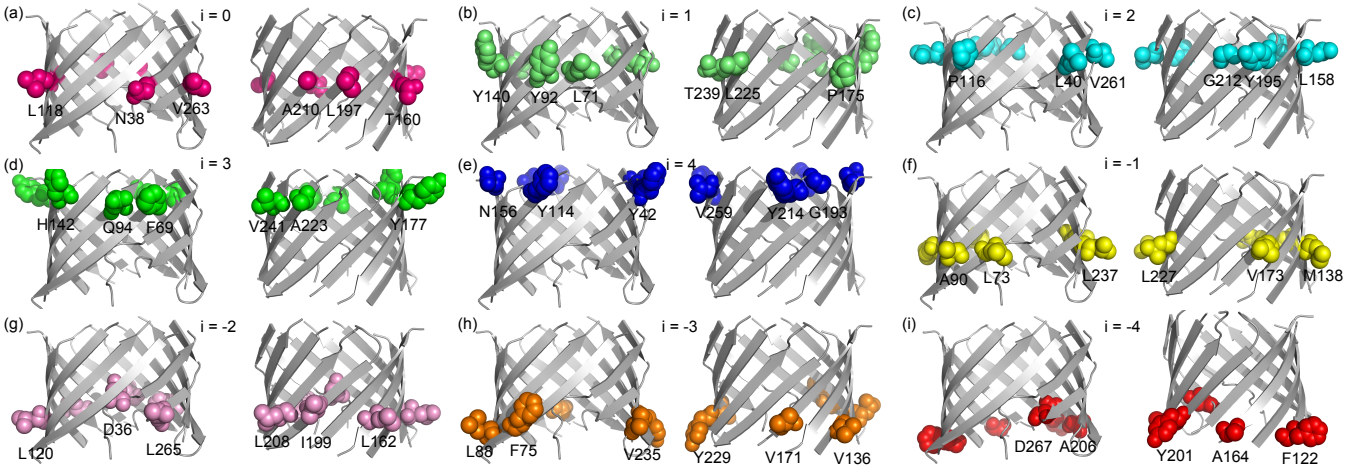


Fig. S2. 53 lipid-facing TM residues are used in deriving transfer free energy scales in OmpLA. For depth-position i, lipid-facing residues with the same position index i are used as host residues. Side-chains of the 53 host residues are shown in spheres.

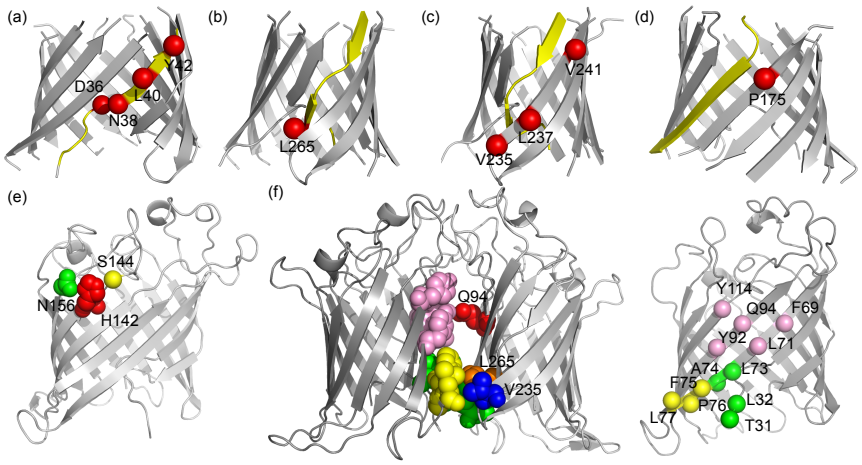


Fig. S3. Host residues demonstrating strong context dependency in transfer free energy scales are located in structurally deformed environment (a-d), active-site residues (e), or forming inter-molecular interactions for dimerization (f). (a) D36, N38, L40, and Y42 are located in β1, which contains a TM coil (residue 35-38). (b) L265 is located in β12, which also contains a TM coil (residue 260-263) (c) V235, L237, and V241 are next to a deformed strand β12 as their neighboring strand. (d) P175 in β7 is neighboring the deformed portion of strand β8. (e) H142 (red) and N156 (green) are part of the catalytic triad (H142-S144-N156) of OmpLA for its enzymatic function.[17] (f) Q94 (red), V235 (blue), and L265 (orange) provide inter-molecular interactions for the OmpLA dimerization: Q94 makes contacts with F69, L71, Y92, and Y114 (pink); V235 with F75, P76, and L77 (yellow); L265 with T31, L32, L73, A74, F75, P76, and L77 (yellow and green). Inter-molecule contacts are identified using a simplicial edge criteria[18]

(a)

	$\Delta\Delta G^{o}_{w,l}$	$\Delta\Delta G^{aa}_{app}$	$\Delta\Delta G_{Octanol}$
$Ez\beta$[a]	NA[c] / NA[d] / 0.65[e]	NA / NA / 0.84	NA / NA / 0.55
Slusky-Dunbrack[b]	NA / 0.31 / 0.30	NA / 0.44 / 0.43	NA / 0.52 / 0.51
E_b	0.47 / 0.64 / 0.64	0.45 / 0.78 / 0.76	0.50 / 0.68 / 0.68
$\Delta\Delta G^{A210}$	0.81 / 0.81 / 0.81	0.86 / 0.88 / 0.89	0.67 / 0.67 / 0.66

a Data of Cys, Met, and Thr not reported
b Data of Cys not reported
c R^2 is calculated for all 20 residues
d R^2 is calculated with Cys excluded
e R^2 is calculated with Cys, Met, and Thr excluded

(b)

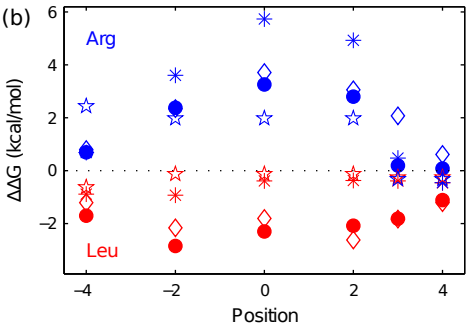


Fig. S4. (a) Three scales based on probabilities of finding residues in the TM strands (values see Table S6), $Ez\beta$,[19] Slusky-Dunbrack scale,[20] and E_b,[4] were compared with three experimental hydrophobicity scales. (b) Computed transfer free energies (closed circle) of Arg and Leu agree well with the experimentally measured values (diamond).[16] However, profiles of empirical energy values either from statistical distribution (empty star)[4] or from the earlier empirical energy function (asterisk)[1] were unable to reproduce the experimental values (diamond).

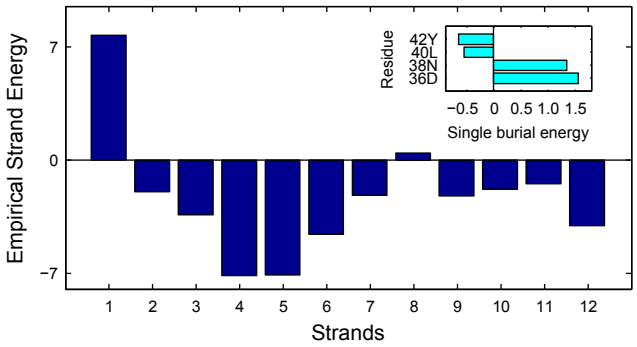


Fig. S5. Strand stability analysis of OmpLA detects unstable strand $\beta 1$ and its unstable residues. Strand energy of each TM strand of the native structure is calculated based on ref[1] . Strand $\beta 1$ is a relatively unstable TM strand, as it has significantly higher strand energy than other TM strands. Lipid-facing residues D36 and N38 are detected as the unstable residues of strand $\beta 1$ based on their high single burial energy (inset figure).

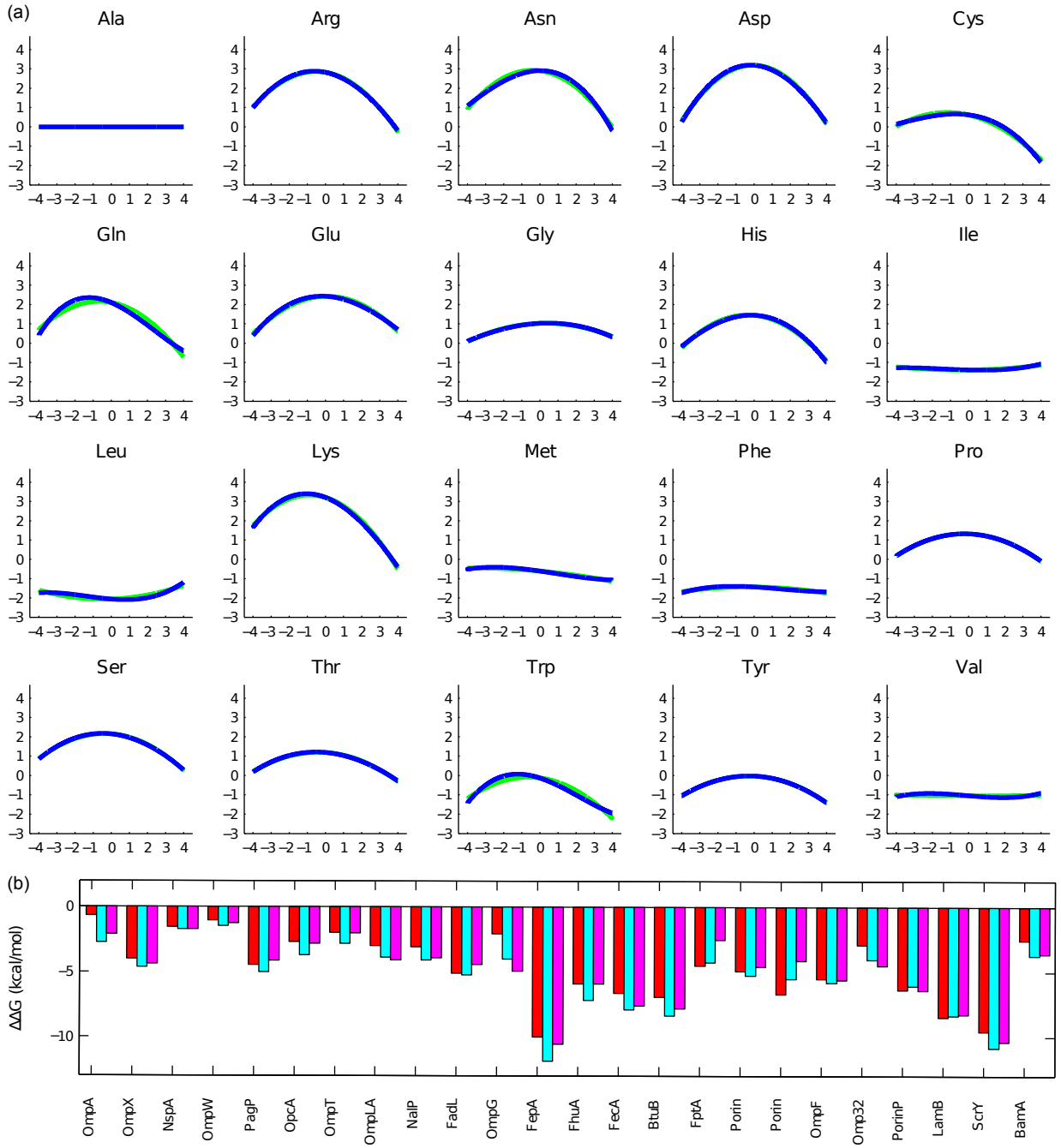


Fig S6. Approximation on thermodynamic stability is not sensitive to parameters of depth-dependent profiles. (a) Polynomial functions were used to derive depth-dependent transfer free energy profiles. Asymmetry in the profiles was observed in both second-degree (green) and third-degree (blue) polynomial function. (b) The stability differences of 24 OMPs in the asymmetric and a symmetric OM were determined using profiles derived from asymmetric Gaussian (red), second-degree polynomial (cyan), and third-degree polynomial (blue). All OMPs showed higher stability ($\Delta\Delta G < 0$) in the asymmetric OM.

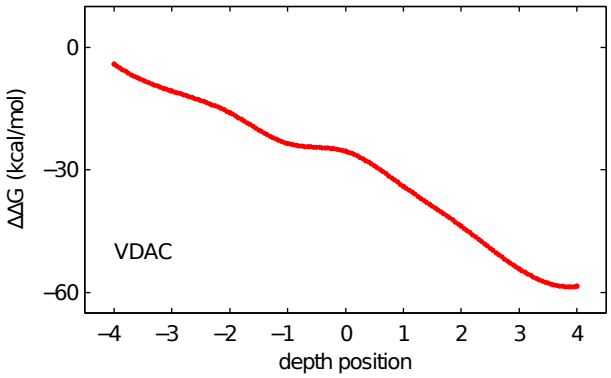


Fig. S7. Spontaneous insertion is predicted for mitochondria outer membrane protein VDAC. Folding energy of VDAC (PDB id: 3EMN) is approximated as the total transfer free energy of lipid-facing residues inserted into the bilayer. No energy barrier is observed during the folding process.

Table S1. Transfer free energy $\Delta\Delta G^{\text{host}}$ (kcal/mol) at different host residues of OmpLA and the general transfer free energy $\Delta\Delta G_{(i)}$ calculated from non-context dependent hosts at each depth-position i

A.A	$\Delta\Delta G(kcal/mol)$					
	$\Delta\Delta G^{\text{F122}}$	$\Delta\Delta G^{\text{A164}}$	$\Delta\Delta G^{\text{Y201}}$	$\Delta\Delta G^{\text{A206}}$	$\Delta\Delta G^{\text{D267}}$	$\Delta\Delta G_{(-4)}$
A	0	0	0	0	0	0±0
R	1.86	0.69	1.16	0.71	1.12	1.11±0.21
N	1.87	1.03	0.85	1.45	1.14	1.27±0.18
D	1.07	0.30	0.03	−0.4	−0.12	0.18±0.25
C	0.25	0.12	0.07	0.03	0.12	0.12±0.04
Q	0.82	0.57	0.30	0.73	0.36	0.56±0.10
E	1.14	0.14	0.24	0.71	0.42	0.53±0.18
G	0.62	0.12	0.25	−0.01	−0.03	0.19±0.12
H	0.20	−0.47	−0.32	0.62	0.22	0.05±0.20
I	−1.70	−1.14	−1.33	−0.65	−1.74	−1.31±0.20
L	−1.63	−1.70	−1.57	−1.78	−1.80	−1.70±0.04
K	2.13	1.89	2.08	2.18	2.24	2.10±0.06
M	−0.48	−0.38	−0.69	−0.48	−0.57	−0.52±0.05
F	−2.17	−2.01	−1.40	−1.63	−1.85	−1.81±0.14
P	−0.37	0.43	0.41	0.36	−0.49	0.07±0.20
S	1.75	0.90	0.76	0.24	1.36	1.00±0.26
T	0.29	0.28	0.23	−0.43	0.69	0.21±0.18
W	−1.86	−1.59	−1.53	−0.50	−1.37	−1.37±0.23
Y	−1.20	−1.10	−1.10	−1.00	−0.78	−1.04±0.07
V	−1.33	−0.84	−0.7	−1.15	−1.29	−1.06±0.12

	$\Delta\Delta G^{\text{V235}}$	$\Delta\Delta G^{\text{F75}}$	$\Delta\Delta G^{\text{L88}}$	$\Delta\Delta G^{\text{V136}}$	$\Delta\Delta G^{\text{V171}}$	$\Delta\Delta G^{\text{Y229}}$	$\Delta\Delta G_{(-3)}$
A	0	0	0	0	0	0	0±0

Continued on next page

Table S1 continued

A.A	$\Delta\Delta G\,(kcal/mol)$						
R	−0.66	0.67	1.4	3.27	1.41	1.70	1.69±0.43
N	0.34	0.33	1.06	2.27	1.59	1.49	1.35±0.32
D	1.10	0.21	0.72	3.17	3.17	1.31	1.72±0.62
C	0.28	0.15	0.22	0.81	0.83	0.26	0.45±0.15
Q	−0.13	0.79	0.41	0.58	2.22	2.15	1.23±0.39
E	−0.72	0.40	0.39	1.87	1.07	1.50	1.05±0.29
G	−0.32	0.26	−0.02	0.08	0.29	0.80	0.28±0.14
H	−0.90	0.20	0.79	−0.26	0.09	0.67	0.30±0.19
I	−0.31	−1.41	−1.15	−0.57	−0.65	−1.76	−1.11±0.23
L	−0.61	−2.16	−2.95	−1.06	−0.11	−2.25	−1.71±0.50
K	0.84	0.60	1.67	2.02	2.39	1.89	1.71±0.30
M	−0.39	−1.27	−1.58	0.64	0.61	−0.27	−0.37±0.46
F	−0.46	−1.00	−2.42	−0.77	−0.26	−2.11	−1.31±0.41
P	0.45	−0.02	0.94	0.33	1.76	0.75	0.75±0.30
S	−0.16	0.59	0.59	1.33	1.08	1.81	1.08±0.23
T	−0.05	0.15	0.86	0.57	0.32	1.17	0.61±0.18
W	−0.34	−0.23	−1.41	−0.59	−0.29	−0.63	−0.63±0.21
Y	−0.74	−0.48	−1.64	−0.16	−0.80	−0.17	−0.65±0.27
V	−0.45	−0.96	−0.92	−0.75	−0.70	−1.46	−0.96±0.13

	$\Delta\Delta G^{\mathrm{D36}}$	$\Delta\Delta G^{\mathrm{L265}}$	$\Delta\Delta G^{\mathrm{L120}}$	$\Delta\Delta G^{\mathrm{L162}}$	$\Delta\Delta G^{\mathrm{I199}}$	$\Delta\Delta G^{\mathrm{L208}}$	$\Delta\Delta G_{(-2)}$
A	0	0	0	0	0	0	0±0
R	−0.24	1.49	2.38	4.04	2.56	2.36	2.84±0.40
N	0.06	0.72	2.04	4.62	3.11	1.66	2.86±0.66
D	0.51	0.28	2.12	4.47	3.15	2.38	3.03±0.53

Continued on next page

Table S1 continued

A.A	$\Delta\Delta G\,(kcal/mol)$						
C	0.37	1.02	−0.34	0.82	0.60	0.41	0.37±0.25
Q	−0.37	1.30	1.87	4.83	3.40	1.34	2.86±0.79
E	−0.03	0.17	1.89	3.71	2.82	1.28	2.42±0.53
G	−0.96	0.42	0.33	1.11	0.92	0.58	0.74±0.17
H	−0.65	1.23	−0.05	1.27	0.47	1.22	0.73±0.32
I	0.28	−0.97	−2.02	−2.00	−1.73	−1.42	−1.79±0.14
L	−0.52	−1.36	−2.85	−1.89	−2.02	−2.27	−2.26±0.21
K	−0.09	1.83	4.36	4.82	3.67	2.46	3.83±0.51
M	−0.28	0.48	−1.29	0.06	0.24	−0.46	−0.36±0.34
F	−0.24	−1.53	−1.47	−1.96	−1.74	−1.59	−1.69±0.11
P	1.07	0.02	2.70	0.96	1.99	0.89	1.64±0.44
S	−0.49	1.43	2.06	4.01	2.70	1.20	2.49±0.59
T	−0.21	0.80	1.31	1.39	1.52	1.04	1.32±0.10
W	−0.51	0.14	0.25	0.16	0.19	−0.24	0.09±0.11
Y	−0.30	0.09	−0.41	−0.15	−0.25	0.49	−0.08±0.20
V	−0.29	−1.07	−1.43	−1.34	−1.40	−0.92	−1.27±0.12

	$\Delta\Delta G^{\mathrm{L237}}$	$\Delta\Delta G^{\mathrm{L73}}$	$\Delta\Delta G^{\mathrm{A90}}$	$\Delta\Delta G^{\mathrm{M138}}$	$\Delta\Delta G^{\mathrm{V173}}$	$\Delta\Delta G^{\mathrm{L227}}$	$\Delta\Delta G_{(-1)}$
A	0	0	0	0	0	0.00	0±0
R	1.98	2.28	3.43	1.76	4.47	1.53	2.69±0.55
N	2.34	1.96	3.56	1.06	3.14	1.71	2.29±0.46
D	1.82	1.86	3.51	1.53	3.72	1.19	2.36±0.52
C	1.25	0.27	1.31	0.81	1.52	0.01	0.78±0.29
Q	2.44	1.60	2.01	0.70	2.13	1.37	1.56±0.26
E	1.54	1.54	3.08	0.66	2.65	1.33	1.85±0.44

Continued on next page

Table S1 continued

A.A	$\Delta\Delta G\,(kcal/mol)$						
G	0.17	1.17	2.05	0.20	1.24	0.49	1.03±0.32
H	1.87	0.38	1.27	0.87	1.96	0.67	1.03±0.27
I	0.76	−1.60	−1.38	−0.08	−0.02	−1.06	−0.83±0.33
L	−0.10	−2.04	−2.20	−0.14	−0.35	−2.68	−1.48±0.52
K	2.42	2.78	2.98	2.05	4.15	2.46	2.88±0.35
M	1.22	−1.18	−1.85	−0.42	0.70	−1.46	−0.84±0.45
F	0.64	−1.63	−2.07	−0.02	0.22	−2.01	−1.10±0.50
P	0.73	−0.08	2.07	0.12	0.34	1.09	0.71±0.39
S	1.31	1.65	2.97	0.64	2.17	1.30	1.75±0.39
T	0.42	0.94	1.87	−0.35	1.33	0.50	0.86±0.38
W	1.87	−0.81	0.04	0.84	0.75	−1.16	−0.07±0.40
Y	0.50	−0.57	0.33	0.11	0.49	−0.71	−0.07±0.24
V	−0.09	−1.62	−1.22	0.39	−0.16	−0.43	−0.61±0.36

	$\Delta\Delta G^{\mathrm{N38}}$	$\Delta\Delta G^{\mathrm{L118}}$	$\Delta\Delta G^{\mathrm{T160}}$	$\Delta\Delta G^{\mathrm{L197}}$	$\Delta\Delta G^{\mathrm{A210}}$	$\Delta\Delta G^{\mathrm{V263}}$	$\Delta\Delta G_{(0)}$
A	0	0	0	0	0	0	0±0
R	1.04	3.29	5.08	1.76	3.26	2.08	3.09±0.58
N	0.87	3.96	3.54	3.74	3.10	2.88	3.44±0.20
D	1.14	4.30	3.78	1.83	4.04	3.89	3.57±0.44
C	0.85	0.57	0.47	0.64	0.38	0.91	0.59±0.09
Q	0.28	3.67	2.89	2.85	2.25	2.56	2.84±0.24
E	0.60	3.35	3.00	3.41	2.89	1.76	2.88±0.30
G	−0.06	1.05	0.88	1.20	1.15	1.32	1.12±0.07
H	0.83	2.82	3.04	2.00	3.23	2.43	2.70±0.22
I	0.50	−2.25	−1.30	−1.89	−1.53	−1.27	−1.65±0.19

Continued on next page

Table S1 continued

A.A	$\Delta\Delta G\,(kcal/mol)$						
L	0.15	−2.47	−2.46	−2.25	−2.30	−2.49	−2.39±0.05
K	0.16	4.60	6.88	4.57	3.97	2.15	4.43±0.76
M	0.18	−0.50	−0.18	0.48	−0.62	−0.45	−0.25±0.20
F	0.54	−1.79	−2.28	−2.24	−1.82	−1.65	−1.96±0.13
P	0.26	1.22	1.61	1.46	1.81	0.92	1.40±0.15
S	0.47	3.19	3.14	2.14	1.73	2.73	2.59±0.28
T	0.20	1.75	1.46	1.61	1.07	1.11	1.40±0.13
W	0.35	−0.69	1.37	0.22	−0.34	−0.23	0.07±0.36
Y	0.25	−0.09	−0.36	−0.41	−0.45	0.42	−0.18±0.16
V	0.52	−1.57	−0.11	−1.42	−1.39	−1.52	−1.20±0.27

	$\Delta\Delta G^{\mathrm{P175}}$	$\Delta\Delta G^{\mathrm{L71}}$	$\Delta\Delta G^{\mathrm{Y92}}$	$\Delta\Delta G^{\mathrm{Y140}}$	$\Delta\Delta G^{\mathrm{L225}}$	$\Delta\Delta G^{\mathrm{T239}}$	$\Delta\Delta G_{(1)}$
A	0	0	0	0	0	0	0±0
R	1.15	2.56	1.26	3.51	1.77	2.47	2.31±0.38
N	1.96	3.01	1.61	2.90	2.46	2.70	2.54±0.25
D	1.54	2.83	1.90	3.54	2.71	2.62	2.72±0.26
C	0.87	0.03	−0.61	1.27	−0.05	0.92	0.31±0.34
Q	0.98	1.53	1.19	2.06	0.23	2.84	1.57±0.44
E	0.72	2.80	2.30	2.57	1.84	2.10	2.32±0.17
G	0.78	0.78	1.14	1.24	1.05	0.43	0.93±0.15
H	1.82	0.37	0.11	1.88	−0.18	1.74	0.78±0.43
I	−0.09	−2.22	−0.93	−0.02	−2.39	−1.64	−1.44±0.44
L	−0.02	−2.84	−1.24	−0.36	−2.08	−1.77	−1.66±0.42
K	1.48	2.23	1.69	3.37	0.98	2.86	2.23±0.42
M	−0.56	−1.64	−1.00	0.68	−1.39	−0.25	−0.72±0.42

Continued on next page

Table S1 continued

A.A	$\Delta\Delta G\,(kcal/mol)$						
F	0.07	−1.39	−1.57	0.18	−1.38	−0.78	−0.99±0.32
P	−0.14	2.35	0.50	0.30	2.02	0.24	1.08±0.46
S	0.57	2.01	1.45	2.15	1.40	1.66	1.73±0.15
T	−0.06	0.91	1.08	1.35	0.62	1.05	1.00±0.12
W	0.54	−0.93	−1.60	0.70	−1.17	0.27	−0.55±0.44
Y	0.15	−0.27	−1.30	0.48	−0.22	0.57	−0.15±0.34
V	−0.07	−1.70	−0.79	−0.16	−1.05	−0.83	−0.91±0.25

	$\Delta\Delta G^{\mathrm{L40}}$	$\Delta\Delta G^{\mathrm{P116}}$	$\Delta\Delta G^{\mathrm{L158}}$	$\Delta\Delta G^{\mathrm{Y195}}$	$\Delta\Delta G^{\mathrm{G212}}$	$\Delta\Delta G^{\mathrm{V261}}$	$\Delta\Delta G_{(2)}$
A	0	0	0	0	0	0	0±0
R	−0.19	2.65	1.70	1.54	2.80	1.20	1.98±0.32
N	−0.06	2.51	2.50	2.49	2.64	2.60	2.55±0.03
D	0.39	3.16	2.06	2.09	3.17	3.22	2.74±0.27
C	0.48	0.17	0.25	−0.06	0.42	−0.12	0.13±0.10
Q	−0.77	0.73	0.52	1.25	1.20	0.43	0.83±0.17
E	−0.82	2.55	1.81	2.37	2.23	0.55	1.90±0.36
G	−0.44	0.97	0.64	1.26	1.15	0.61	0.93±0.13
H	0.18	0.97	0.63	2.60	−0.02	−0.16	0.80±0.49
I	−0.04	−1.13	−1.86	−1.31	−1.24	−1.66	−1.44±0.14
L	0.03	−2.59	−2.13	−2.68	−2.08	−2.74	−2.44±0.14
K	0.72	1.06	1.36	2.65	1.88	1.57	1.70±0.27
M	−0.91	−1.54	−0.51	−0.34	−0.78	−1.73	−0.98±0.28
F	0.70	−2.47	−1.63	−1.24	−2.18	−1.93	−1.89±0.21
P	1.20	−0.34	0.98	2.27	1.51	1.85	1.25±0.45
S	−0.74	1.80	1.43	1.83	2.36	1.51	1.79±0.16

Continued on next page

Table S1 continued

A.A	$\Delta\Delta G\,(kcal/mol)$						
T	−0.61	1.07	0.72	1.23	1.27	0.12	0.88±0.21
W	0.18	−1.41	−1.12	−0.23	−1.08	−1.25	−1.02±0.21
Y	0.27	0.41	0.00	0.22	−0.65	−0.76	−0.16±0.23
V	0.10	−1.90	−1.53	−1.67	−1.22	−1.08	−1.48±0.15

	$\Delta\Delta G^{\mathrm{Q94}}$	$\Delta\Delta G^{\mathrm{H142}}$	$\Delta\Delta G^{\mathrm{V241}}$	$\Delta\Delta G^{\mathrm{F69}}$	$\Delta\Delta G^{\mathrm{Y177}}$	$\Delta\Delta G^{\mathrm{A223}}$	$\Delta\Delta G_{(3)}$
A	0	0	0	0	0	0	0±0
R	0.71	−0.80	−0.16	0.55	1.29	0.19	0.68±0.32
N	−0.23	−0.32	0.92	0.14	0.46	0.50	0.37±0.11
D	−0.29	−0.46	0.26	0.70	0.92	1.26	0.96±0.16
C	−0.90	−0.80	0.21	−1.62	0.40	−1.80	−1.01±0.71
Q	−0.11	−0.53	1.15	−0.16	−0.70	0.45	−0.14±0.33
E	0.85	−0.52	0.26	0.86	0.44	1.29	0.86±0.25
G	0.02	−0.39	−0.05	0.63	0.67	0.44	0.58±0.07
H	0.02	−1.74	−0.54	−0.61	−0.30	−0.81	−0.57±0.15
I	−0.05	0.18	−0.39	−1.26	−0.66	−0.78	−0.90±0.18
L	−0.38	−0.44	−0.45	−1.71	−0.65	−1.82	−1.39±0.37
K	0.68	−0.14	0.11	0.28	1.28	−0.09	0.49±0.41
M	−0.26	−0.12	0.43	−1.39	−0.73	−1.28	−1.13±0.20
F	−0.34	−0.44	−0.71	−1.61	−0.81	−1.71	−1.38±0.28
P	−0.12	0.02	−0.36	0.29	−0.01	0.59	0.29±0.17
S	−0.15	−0.18	0.36	0.61	0.38	0.75	0.58±0.11
T	0.04	−0.33	−0.10	−0.09	−0.18	0.05	−0.07±0.07
W	0.07	0.09	−0.35	−1.27	−1.34	−2.33	−1.65±0.34
Y	0.15	−0.53	−0.65	−1.30	−0.84	−1.49	−1.21±0.19

Continued on next page

Table S1 continued

A.A	$\Delta\Delta G\,(kcal/mol)$						
V	0.08	0.27	−0.42	−1.29	−0.91	−0.19	−0.80±0.32

	$\Delta\Delta G^{\mathrm{Y42}}$	$\Delta\Delta G^{\mathrm{N156}}$	$\Delta\Delta G^{\mathrm{Y114}}$	$\Delta\Delta G^{\mathrm{G193}}$	$\Delta\Delta G^{\mathrm{Y214}}$	$\Delta\Delta G^{\mathrm{V259}}$	$\Delta\Delta G_{(4)}$
A	0	0	0	0	0	0	0±0
R	−0.13	1.16	0.53	−0.88	0.08	−0.08	−0.09±0.29
N	−0.41	−1.05	−0.13	−0.06	0.17	0.38	0.09±0.12
D	−0.35	0.28	0.62	−0.21	0.37	0.40	0.30±0.18
C	−0.84	−2.22	−1.82	−1.91	−1.63	−1.75	−1.78±0.06
Q	−0.27	−0.62	−0.48	−0.73	−0.24	0.59	−0.22±0.29
E	−0.18	−0.07	1.33	0.57	0.60	1.18	0.92±0.20
G	−0.02	0.48	0.42	0.52	0.17	0.39	0.38±0.07
H	0.27	0.41	−0.79	−1.63	0.61	−0.84	−0.66±0.47
I	−0.08	−1.89	−1.13	−1.46	−1.11	−1.06	−1.19±0.09
L	−0.21	−1.62	−1.04	−1.24	−1.11	−1.45	−1.21±0.09
K	0.17	−0.19	−0.53	−0.65	0.64	0.01	−0.13±0.29
M	−0.18	−0.31	−0.96	−1.07	−0.99	−0.97	−1.00±0.03
F	−0.04	−1.94	−2.03	−2.07	−1.32	−1.57	−1.75±0.18
P	0.20	1.37	−0.75	0.35	−0.14	0.10	−0.11±0.24
S	−0.25	−0.07	0.55	0.13	0.43	0.69	0.45±0.12
T	0.00	−0.99	−0.24	−0.23	−0.11	−0.10	−0.17±0.04
W	0.16	−1.97	−1.88	−2.53	−1.27	−1.86	−1.89±0.25
Y	0.40	−2.11	−1.47	−1.57	−0.78	−1.50	−1.33±0.18
V	0.54	−1.12	−0.97	−1.24	−0.82	−0.75	−0.95±0.11

Table S2. Correlation coefficient (R^2) of transfer free energy scales from different host residues ($\Delta\Delta G^{\text{host}}$ in kcal/mol) of OmpLA at the same depth-position i

position $i = 4$

	$\Delta\Delta G^{F122}$	$\Delta\Delta G^{A164}$	$\Delta\Delta G^{Y201}$	$\Delta\Delta G^{A206}$	$\Delta\Delta G^{D267}$
$\Delta\Delta G^{F122}$	1.00				
$\Delta\Delta G^{A164}$	0.88	1.00			
$\Delta\Delta G^{Y201}$	0.87	0.95	1.00		
$\Delta\Delta G^{A206}$	0.71	0.76	0.76	1.00	
$\Delta\Delta G^{D267}$	0.90	0.87	0.89	0.76	1.00

position $i = 3$

	$\Delta\Delta G^{F75}$	$\Delta\Delta G^{L88}$	$\Delta\Delta G^{V136}$	$\Delta\Delta G^{V171}$	$\Delta\Delta G^{Y229}$	$\Delta\Delta G^{V235}$
$\Delta\Delta G^{F75}$	1.00					
$\Delta\Delta G^{F75}$	1.00					
$\Delta\Delta G^{V136}$	0.43	0.53	1.00			
$\Delta\Delta G^{V171}$	0.39	0.48	0.65	1.00		
$\Delta\Delta G^{Y229}$	0.85	0.78	0.59	0.57	1.00	
$\Delta\Delta G^{V235}$	0.11	0.23	0.23	0.51	0.14	1.00

position $i = 2$

	$\Delta\Delta G^{D36}$	$\Delta\Delta G^{L120}$	$\Delta\Delta G^{L162}$	$\Delta\Delta G^{I199}$	$\Delta\Delta G^{L208}$	$\Delta\Delta G^{L265}$
$\Delta\Delta G^{D36}$	1.00					
$\Delta\Delta G^{L120}$	0.08	1.00				
$\Delta\Delta G^{L162}$	0.01	0.79	1.00			
$\Delta\Delta G^{I199}$	0.04	0.89	0.95	1.00		
$\Delta\Delta G^{L208}$	0.03	0.83	0.84	0.86	1.00	
$\Delta\Delta G^{L265}$	0.01	0.55	0.64	0.64	0.71	1.00

Continued on next page

Table S2 continued

position $i = 1$						
	$\Delta\Delta G^{L73}$	$\Delta\Delta G^{A90}$	$\Delta\Delta G^{M138}$	$\Delta\Delta G^{V173}$	$\Delta\Delta G^{L227}$	$\Delta\Delta G^{L237}$
$\Delta\Delta G^{L73}$	1.00					
$\Delta\Delta G^{A90}$	0.90	1.00				
$\Delta\Delta G^{M138}$	0.53	0.48	1.00			
$\Delta\Delta G^{V173}$	0.82	0.71	0.75	1.00		
$\Delta\Delta G^{L227}$	0.87	0.87	0.47	0.65	1.00	
$\Delta\Delta G^{L237}$	0.44	0.37	0.58	0.63	0.38	1.00

position $i = 0$						
	$\Delta\Delta G^{N38}$	$\Delta\Delta G^{L118}$	$\Delta\Delta G^{T160}$	$\Delta\Delta G^{L197}$	$\Delta\Delta G^{A210}$	$\Delta\Delta G^{V263}$
$\Delta\Delta G^{N38}$	1.00					
$\Delta\Delta G^{L118}$	0.14	1.00				
$\Delta\Delta G^{T160}$	0.13	0.86	1.00			
$\Delta\Delta G^{L197}$	0.04	0.89	0.82	1.00		
$\Delta\Delta G^{A210}$	0.18	0.95	0.88	0.84	1.00	
$\Delta\Delta G^{V263}$	0.19	0.92	0.72	0.76	0.88	1.00

position $i = -1$						
	$\Delta\Delta G^{L71}$	$\Delta\Delta G^{Y92}$	$\Delta\Delta G^{Y140}$	$\Delta\Delta G^{P175}$	$\Delta\Delta G^{L225}$	$\Delta\Delta G^{T239}$
$\Delta\Delta G^{L71}$	1.00					
$\Delta\Delta G^{Y92}$	0.81	1.00				
$\Delta\Delta G^{Y140}$	0.70	0.67	1.00			
$\Delta\Delta G^{P175}$	0.42	0.35	0.66	1.00		
$\Delta\Delta G^{L225}$	0.94	0.76	0.59	0.34	1.00	
$\Delta\Delta G^{T239}$	0.79	0.65	0.86	0.64	0.64	1.00

Continued on next page

Table S2 continued

position $i = -2$

	$\Delta\Delta G^{L40}$	$\Delta\Delta G^{P116}$	$\Delta\Delta G^{L158}$	$\Delta\Delta G^{Y195}$	$\Delta\Delta G^{G212}$	$\Delta\Delta G^{V261}$
$\Delta\Delta G^{L40}$	1.00					
$\Delta\Delta G^{P116}$	0.05	1.00				
$\Delta\Delta G^{L158}$	0.02	0.89	1.00			
$\Delta\Delta G^{Y195}$	0.00	0.72	0.86	1.00		
$\Delta\Delta G^{G212}$	0.03	0.87	0.92	0.75	1.00	
$\Delta\Delta G^{V261}$	0.01	0.74	0.85	0.74	0.88	1.00

position $i = -3$

	$\Delta\Delta G^{F69}$	$\Delta\Delta G^{Q94}$	$\Delta\Delta G^{H142}$	$\Delta\Delta G^{Y177}$	$\Delta\Delta G^{A223}$	$\Delta\Delta G^{V241}$
$\Delta\Delta G^{F69}$	1.00					
$\Delta\Delta G^{Q94}$	0.26	1.00				
$\Delta\Delta G^{H142}$	0.00	0.00	1.00			
$\Delta\Delta G^{Y177}$	0.57	0.10	0.05	1.00		
$\Delta\Delta G^{A223}$	0.83	0.15	0.00	0.39	1.00	
$\Delta\Delta G^{V241}$	0.19	0.01	0.00	0.11	0.25	1.00

position $i = -4$

	$\Delta\Delta G^{Y42}$	$\Delta\Delta G^{Y114}$	$\Delta\Delta G^{N156}$	$\Delta\Delta G^{G193}$	$\Delta\Delta G^{Y214}$	$\Delta\Delta G^{V259}$
$\Delta\Delta G^{Y42}$	1.00					
$\Delta\Delta G^{Y114}$	0.03	1.00				
$\Delta\Delta G^{N156}$	0.01	0.49	1.00			
$\Delta\Delta G^{G193}$	0.01	0.72	0.46	1.00		
$\Delta\Delta G^{Y214}$	0.01	0.68	0.57	0.54	1.00	
$\Delta\Delta G^{V259}$	0.02	0.82	0.48	0.82	0.71	1.00

Table S3. Values of depth-dependent profile $\Delta\Delta G_{aa(i)}$ (kcal/mol) and fitting parameters

$\Delta\Delta G_{aa(i)}$	A	R	N	D	C	Q	E	G	H	I	L	K	M	F	P	S	T	W	Y	V
-4	0.00	1.20	1.08	0.78	0.19	0.83	0.66	0.15	0.00	−1.08	−1.49	1.77	−0.31	−1.27	0.41	0.96	0.36	−1.41	−1.09	−0.85
-3	0.00	1.80	1.72	1.48	0.33	1.36	1.20	0.36	0.01	−1.24	−1.75	2.44	−0.47	−1.44	0.69	1.41	0.64	−0.53	−0.55	−0.97
-2	0.00	2.42	2.41	2.34	0.49	1.92	1.85	0.68	0.24	−1.37	−1.95	3.08	−0.63	−1.56	1.01	1.86	0.95	−0.10	−0.19	−1.06
-1	0.00	2.88	2.94	3.08	0.63	2.37	2.39	0.98	1.39	−1.45	−2.09	3.54	−0.76	−1.65	1.26	2.19	1.22	−0.03	−0.05	−1.12
0	0.00	3.06	3.15	3.38	0.68	2.54	2.61	1.11	2.50	−1.48	−2.13	3.71	−0.81	−1.68	1.36	2.31	1.32	−0.13	−0.01	−1.15
1	0.00	2.61	2.73	3.01	0.33	1.76	2.39	1.04	0.93	−1.44	−2.06	2.89	−0.80	−1.65	1.20	2.06	1.07	−0.43	−0.01	−1.12
2	0.00	1.62	1.77	2.12	0.04	0.58	1.83	0.85	0.05	−1.33	−1.86	1.37	−0.77	−1.57	0.81	1.46	0.57	−1.03	−0.19	−1.05
3	0.00	0.73	0.86	1.19	0.00	0.09	1.17	0.61	0.00	−1.17	−1.57	0.40	−0.73	−1.44	0.43	0.82	0.20	−1.67	−1.20	−0.95
4	0.00	0.24	0.32	0.53	0.00	0.01	0.63	0.38	0.00	−0.97	−1.23	0.07	−0.67	−1.28	0.17	0.36	0.05	−1.86	−1.32	−0.82

Para	A	R	N	D	C	Q	E	G	H	I	L	K	M	F	P	S	T	W	Y	V
a_0		3.06	3.15	3.38	0.68	2.55	2.61	1.11	2.50	−1.48	−2.13	3.71	−0.81	−1.68	1.36	2.31	1.32	−1.88	−1.57	−1.15
a_1		1.77	1.87	2.07	0.82	1.17	2.37	2.73	0.71	4.35	3.83	1.42	6.63	5.55	1.97	2.08	1.54	1.19	1.67	4.90
a_2		2.91	2.73	2.34	2.50	2.67	2.41	2.01	0.92	5.00	4.74	3.29	2.87	5.41	2.57	3.02	2.49	4.90	5.43	5.17
a_3																		1.625	0.754	
a_4																		0.758	3.558	

Table S4. Differences in the folding free energy of the TM regions of OMPs in the NC-IN and the NC-OUT topology calculated using the additive model by summing up the transfer free energy of all lipid facing TM residues ($\Delta\Delta G_{topo}^{Additive}$ in kcal/mol) and by using the partition function as in Eq 5 ($\Delta\Delta G_{topo}^{Ensemble}$ in kal/mol). The folding free energy ($\Delta G_{NC-IN}^{Ensemble}$ in kcal/mol) calculated using Eq. 5 is also listed.

Protein	PDB ID	$\Delta\Delta G_{topo}^{Additive}$	$\Delta\Delta G_{topo}^{Ensemble}$	$\Delta G_{NC-IN}^{Ensemble}$
OmpA	1BXW	−2.20	10.76	−85.80
OmpX	1QJ8	2.38	9.57	−80.80
NspA	1P4T	0.20	3.33	−85.03
OmpW	2F1T	−2.34	1.21	−79.58
PagP	1THQ	3.71	1.81	−78.90
OpcA	1K24	3.03	12.80	−101.65
OmpT	1I78	0.32	8.95	−97.57
OmpLA	1QD6	2.09	11.51	−115.17
NalP	1UYN	3.27	13.37	−119.08
FadL	1T16	−1.30	12.98	−140.39
OmpG	2F1C	−0.59	4.81	−144.21
FepA	1FEP	1.65	26.98	−201.88
FhuA	2FCP	−0.85	22.34	−205.09
FecA	1KMO	1.65	16.22	−206.17
BtuB	1NQE	−0.92	19.35	−189.49
FptA	1XKW	0.08	15.54	−203.87
Porin	1PRN	2.40	9.06	−138.77
Porin	2POR	3.95	7.03	−150.68
OmpF	2OMF	5.07	11.13	−141.27
Omp32	1E54	2.11	16.13	−146.27
PorinP	2O4V	2.73	18.44	−153.98
LamB	2MPR	4.66	7.60	−156.50
ScrY	1A0S	7.95	16.48	−157.27
BamA	4K3B	−2.17	8.32	−153.36

There is correlation between the folding free energy and the number of strands in the OMPs. However, OMPs of the same family can have different folding energy. Among the five OMPs belonging to the TonB-dependent transporter family, BtuB (PDB id: 1NQE) is less stable than the other four proteins (PDB id: 1FEP, 1KMO, 1NQE, and 1XKW). The folding free energy of BtuB is 7.2% lower than the average folding the other four members.

Table S5. Cooperativity of ionizable or polar residue pairs $\Delta\Delta\Delta G_{\mathrm{XX}}^{\mathrm{Hostpair}}$ (kcal/mol) of host pairs at different positions in OmpLA

XX	Host pair position $(4, 2)$				Host pair position $(2, 0)$			
	Y114P116	G193Y195	G212Y214	Average	P116L118	Y195L197	A210G212	Average
DD	0.44	−0.29	0.22	0.12	1.70	1.60	0.98	1.43
EE	0.82	0.15	0.05	0.34	1.13	1.21	0.25	0.86
RR	0.20	−1.11	0.02	−0.30	1.11	1.43	0.87	1.14
KK	−0.13	−0.81	−0.24	−0.39	0.29	1.25	0.81	0.78
HH	0.51	−0.90	−0.07	−0.15	1.18	1.03	0.11	0.77
NN	0.58	−0.32	0.14	0.13	1.73	1.49	0.32	1.18
QQ	0.06	−1.10	0.00	−0.35	0.65	1.37	0.10	0.71

XX	Host pair position $(0, -2)$				Host pair position $(-2, -4)$			
	L118L120	L197I199	L208A210	Average	L120F122	I199Y201	A206L208	Average
DD	2.37	1.84	3.22	2.48	0.84	1.78	0.04	0.89
EE	2.93	2.97	4.23	3.38	1.42	1.33	0.98	1.24
RR	2.61	1.51	2.66	2.26	1.36	1.50	0.53	1.13
KK	4.99	4.05	3.55	4.20	3.24	3.61	2.76	3.20
HH	0.78	1.00	2.34	1.37	0.51	0.34	0.68	0.51
NN	3.11	4.16	3.05	3.44	1.74	2.32	1.27	1.78
QQ	3.11	3.45	3.34	3.30	0.62	1.29	1.39	1.10

Table S6. Values of experimental and knowledge-based amino acid hydrophobicity scales in kcal/mol.

	experimental				knowledge-based	
A.A	$\Delta\Delta G^{o}_{w,l}$	$\Delta\Delta G^{aa}_{app}$	$\Delta\Delta G_{octanol}$	$E_z\beta$	Slusky-Dunbrack	E_b
A	0	0.11	0.50	−0.8	0.13	−0.69
R	3.71	2.58	1.81	1.4	−1.06	2.66
N	3.47	2.05	0.85	0.7	−0.69	2.30
D	2.95	3.49	3.64	1.3	−0.63	2.66
C	0.49	−0.13	−0.02	NA	NA	4.61
Q	3.01	2.36	0.77	0.7	−0.69	2.21
E	1.64	2.68	3.63	1.1	−2.48	4.61
G	1.72	0.74	1.15	0	0.04	0.27
H	4.76	2.06	0.11	1.2	−0.25	1.90
I	−1.56	−0.60	−1.12	−1.0	0.10	−0.83
L	−1.81	−0.55	−1.25	−2.0	0.35	−0.91
K	5.39	2.71	2.80	1.3	−0.98	3.51
M	−0.76	−0.10	−0.67	NA	0	−0.46
F	−2.20	−0.32	−1.71	−1.9	0.21	−0.25
P	−1.52	2.23	0.14	0.8	0.31	0.05
S	1.83	0.84	0.46	0.9	−0.11	1.31
T	1.78	0.52	0.25	NA	−0.30	0.40
W	−0.38	0.30	−2.09	−0.4	−0.44	0.26
Y	−1.09	0.68	−0.71	−0.0	−0.21	0.17
V	−0.78	−0.31	−0.46	−1.5	0.05	−0.96

$\Delta\Delta G^{o}_{w,l}$: Moon and Fleming's whole-protein scale[16]
$\Delta\Delta G^{aa}_{app}$: Hessa *et al*'s biological scale [21]
$\Delta\Delta G_{octanol}$: Wimley and White's octanol scale [22]
$E_z\beta$: Hsieh *et al*'s $E_z\beta$ scale[19]
Slusky-Dunbrack: Slusky and Dunbrack's scale[20]
E_b: Jackups and Liang's scale[4]

Table S7. Empirical strand interaction energy term values in kcal/mol.

E_{SH}	A	R	N	D	C	Q	E	G	H	I	L	K	M	F	P	S	T	W	Y	V
A	−0.17	−0.09	0.06	−0.08	−0.00	0.22	−0.09	0.01	0.18	0.02	0.15	0.12	0.56	0.17	0.36	−0.03	0.08	0.38	−0.07	−0.19
R	−0.22	0.07	0.01	−0.00	−0.16	0.04	0.21	−0.02	0.31	0.64	0.10	−0.23	0.75	−0.82	0.07	−0.19	−0.48	−0.01	−0.12	
N	−0.42	−0.54	−0.00	0.33	0.33	0.02	0.21	0.39	−0.06	−0.25	0.18	0.56	−0.47	−0.05	0.05	0.10	0.80	0.22		
D	1.78	−0.00	0.67	0.06	−0.11	−0.41	1.78	0.18	0.66	−0.45	1.78	−0.47	−0.08	0.03	−0.26	0.24	0.14			
C	−0.00	−0.00	−0.00	−0.00	−0.00	−0.00	−0.00	−0.00	−0.00	−0.00	−0.00	−0.00	−0.00	−0.00	−0.00	−0.00				
Q	−0.03	0.19	−0.04	0.20	−0.38	−0.12	−0.19	−0.14	−0.25	−0.00	0.04	−0.06	−0.61	0.15	0.35					
E	−0.29	0.15	1.78	1.78	−0.09	0.17	−0.08	0.56	−0.00	−0.18	−0.27	−0.39	0.09	0.22						
G	0.10	−0.13	0.21	0.06	0.09	0.34	−0.33	−0.60	0.07	0.15	0.15	−0.23	−0.12							
H	1.78	1.78	−0.01	−0.74	−0.41	−0.08	−0.95	−0.26	0.59	−0.65	1.78	−0.08								
I	0.13	0.06	−0.39	−0.43	0.02	1.78	−0.17	−0.15	0.15	−0.27	0.10									
L	0.08	0.67	−0.03	−0.07	0.21	0.03	−0.01	−0.26	−0.21	−0.02										
K	−0.01	−0.41	0.12	1.78	−0.29	−0.26	0.22	0.95	0.45											
M	1.78	0.14	1.78	−0.37	−0.10	0.57	0.23	−0.13												
F	0.09	0.11	0.22	−0.01	−0.15	0.10	−0.17													
P	−0.00	−0.06	−0.06	1.78	−0.31	1.78														
S	−0.06	0.03	0.73	0.30	0.34															
T	−0.08	0.42	0.12	0.15																
W	−0.11	0.14	−0.24																	
Y	0.67	−0.11																		
V	0.26																			

E_{NH}	A	R	N	D	C	Q	E	G	H	I	L	K	M	F	P	S	T	W	Y	V
A	−0.19	0.39	0.34	0.59	−0.00	0.15	−0.04	0.05	0.62	−0.11	−0.14	0.47	−0.29	−0.09	−0.08	0.04	0.23	0.60	0.23	−0.15
R	−0.13	−0.32	−0.10	−0.00	−0.24	−0.32	0.34	1.78	0.22	0.87	−0.06	1.78	−0.22	0.13	0.35	−0.09	0.46	−0.36	0.07	
N	−0.29	−0.15	−0.00	0.09	0.10	−0.16	−0.25	0.62	0.49	−0.08	0.15	1.78	1.78	−0.27	0.01	0.45	0.66	0.69		
D	−0.26	−0.00	−0.02	0.43	0.07	−0.11	1.78	0.34	0.01	0.15	0.39	−0.54	−0.28	−0.00	−0.33	−0.09	0.26			

Continued on next page

Table S7 continued

C	−0.00	−0.00	−0.00	−0.00	−0.00	−0.00	−0.00	−0.00	−0.00	−0.00	−0.00	−0.00	−0.00	−0.00	−0.00	−0.00
Q	0.47	0.59	−0.25	−0.06	1.78	0.01	−0.44	0.03	0.73	1.78	0.03	−0.18	−0.21	−0.03	1.78	
E	0.01	−0.12	−0.47	−0.06	0.23	−0.29	0.17	−0.56	1.78	0.66	0.09	−0.03	0.12	0.38		
G	0.16	0.28	−0.29	−0.17	0.73	−0.22	0.25	−0.11	−0.05	−0.01	0.12	0.33	−0.26			
H	1.78	0.19	0.13	0.08	1.78	−0.24	1.78	−0.40	−0.54	0.36	−0.17	0.08				
I	−0.01	−0.05	−0.06	−0.28	0.26	0.54	0.69	−0.10	0.15	0.08	−0.11					
L	−0.18	−0.09	0.11	0.18	−0.33	0.15	0.20	0.43	0.14	0.08						
K	−0.22	−0.02	0.32	0.18	−0.00	−0.21	−0.18	−0.14	0.31							
M	−0.26	−0.22	−0.30	0.11	0.06	0.32	0.27	0.53								
F	−0.02	0.31	−0.05	0.03	−0.18	−0.15	−0.05									
P	1.78	0.02	−0.10	0.26	0.19	−0.31										
S	−0.23	0.11	−0.08	0.02	0.47											
T	−0.24	0.31	0.15	0.08												
W	−0.08	−0.55	0.41													
Y	−0.09	−0.02														
V	−0.07															

E_{WH}	A	R	N	D	C	Q	E	G	H	I	L	K	M	F	P	S	T	W	Y	V
A	0.09	−0.24	−0.18	−0.00	−0.00	−0.06	−0.04	−0.09	−0.22	0.09	0.18	−0.24	0.25	−0.01	0.10	−0.01	0.21	0.12	0.01	−0.02
R	1.02	0.47	1.02	−0.00	0.22	0.66	0.27	0.29	−0.10	−0.30	0.67	−0.00	−0.35	0.12	0.47	−0.02	−0.18	0.03	−0.33	
N	0.25	0.92	−0.00	0.20	0.60	0.31	−0.03	−0.47	−0.05	0.35	−0.02	−0.27	−0.44	0.75	0.20	−0.30	−0.19	−0.24		
D	1.02	−0.00	0.02	0.87	0.31	1.02	−0.34	0.07	−0.07	0.15	−0.28	−0.84	0.75	−0.25	−0.38	−0.03	−0.41			
C	−0.00	−0.00	−0.00	−0.00	−0.00	−0.00	−0.00	−0.00	−0.00	−0.00	−0.00	−0.00	−0.00	−0.00	−0.00	−0.00				
Q	0.07	0.34	0.40	−0.32	−0.14	−0.15	0.56	0.28	−0.18	−0.35	0.13	0.25	−0.29	−0.23	−0.12					
E	0.06	−0.03	0.14	−0.27	−0.32	1.01	−0.40	−0.29	−0.06	1.02	0.12	0.26	−0.09	−0.13						
G	0.04	0.10	−0.11	−0.16	0.18	0.16	−0.05	−0.54	0.12	0.10	−0.05	0.04	−0.16							
H	1.02	−0.35	0.20	0.05	0.08	−0.11	1.02	−0.33	−0.24	−0.16	0.42	0.05								
I	1.01	0.53	−0.24	−0.06	0.49	1.02	−0.35	−0.09	0.11	0.23	0.19									

Continued on next page

Table S7 continued

L	0.12	−0.21	0.45	0.12	0.41	−0.23	−0.11	0.16	0.25	0.34
K	1.02	0.53	−0.26	1.02	0.18	0.25	−0.11	−0.16	−0.11	
M	−0.34	0.11	−0.29	−0.33	−0.03	0.13	−0.34	−0.01		
F	0.24	0.19	−0.16	−0.10	0.35	0.26	0.95			
P	1.02	0.53	1.02	1.02	−0.12	0.80				
S	0.67	0.31	−0.29	−0.26	−0.22					
T	0.08	−0.18	−0.30	−0.14						
W	0.31	0.43	0.06							
Y	0.12	0.18								
V	0.82									

E_{Intra}	A	R	N	D	C	Q	E	G	H	I	L	K	M	F	P	S	T	W	Y	V
A	−1.61	−1.08	−1.09	−1.01	−2.11	−1.12	−0.89	−1.37	−1.43	−2.71	−2.91	−0.78	−2.33	−2.85	−1.20	−1.19	−1.37	−2.26	−1.99	−2.39
R	−0.92	−0.97	−1.36	−1.52	−1.07	−1.34	−1.02	−1.28	−2.15	−2.38	−0.35	−1.85	−2.36	−1.01	−0.96	−1.12	−2.02	−1.87	−1.82	
N	−0.99	−0.99	−1.53	−1.01	−0.89	−1.03	−1.23	−1.92	−2.21	−0.72	−1.75	−2.22	−0.91	−0.94	−1.11	−1.82	−1.63	−1.67		
D	−0.72	−1.43	−0.86	−0.60	−0.94	−1.37	−1.88	−2.01	−0.99	−1.52	−2.06	−0.79	−0.96	−1.07	−1.68	−1.63	−1.47			
C	−3.22	−1.69	−1.34	−1.87	−2.13	−3.25	−3.45	−1.15	−2.95	−3.43	−1.82	−1.69	−1.84	−2.93	−2.46	−2.94				
Q	−0.91	−0.84	−0.98	−1.17	−2.17	−2.39	−0.76	−1.95	−2.43	−1.02	−0.88	−1.12	−1.84	−1.76	−1.82					
E	−0.54	−0.72	−1.27	−1.94	−2.12	−1.07	−1.71	−2.11	−0.75	−0.88	−1.03	−1.77	−1.65	−1.58						
G	−1.33	−1.27	−2.24	−2.46	−0.68	−2.01	−2.44	−1.11	−1.08	−1.23	−2.02	−1.78	−2.00							
H	−1.80	−2.45	−2.69	−0.80	−2.36	−2.82	−1.33	−1.25	−1.43	−2.36	−2.08	−2.12								
I	−3.87	−4.17	−1.78	−3.56	−4.05	−2.23	−2.08	−2.38	−3.42	−3.11	−3.58									
L	−4.36	−1.99	−3.79	−4.31	−2.49	−2.32	−2.57	−3.63	−3.36	−3.83										
K	−0.07	−1.47	−1.99	−0.57	−0.62	−0.78	−1.59	−1.54	−1.47											
M	−3.23	−3.88	−2.04	−1.79	−2.08	−3.28	−2.91	−3.15												
F	−4.30	−2.52	−2.38	−2.53	−3.65	−3.35	−3.72													
P	−1.04	−0.93	−1.12	−2.21	−1.89	−1.96														
S	−0.99	−1.16	−1.77	−1.65	−1.80															

Continued on next page

Table S7 continued

T	−1.25	−1.91	−1.78	−2.05
W	−2.99	−2.76	−3.07	
Y	−2.47	−2.73		
V	−3.27			

Table S8. Single burial energy term values in kcal/mol.

AA	E-CAP	E-HG-IN	E-HG-OUT	C-IN	C-OUT	P-HG-IN	P-HG-OUT	P-CAP
A	0.20	0.19	0.28	0.01	−0.41	0.19	0.39	0.03
R	0.12	−0.24	0.83	−0.30	1.57	−0.15	0.06	−0.23
N	−0.23	−0.14	0.59	−0.09	1.36	−0.01	0.66	−0.24
D	−0.34	−0.15	1.90	0.25	1.57	0.02	0.23	−0.30
C	0.83	0.83	0.83	0.83	0.83	0.83	0.83	−0.88
Q	0.25	−0.10	0.21	−0.21	1.31	−0.47	−0.11	−0.10
E	−0.16	−0.26	0.83	−0.32	0.83	−0.26	0.41	−0.08
G	−0.06	0.20	0.25	−0.33	0.16	−0.06	0.64	0.02
H	−0.09	−0.18	−0.46	0.20	1.12	0.64	−0.54	0.06
I	0.13	0.50	−0.32	0.60	−0.49	0.09	0.01	0.25
L	0.08	0.46	−0.34	0.54	−0.54	0.25	0.12	0.40
K	0.03	−0.51	1.21	−0.03	2.08	0.22	−0.19	−0.26
M	0.15	0.49	0.35	0.02	−0.27	−0.29	−0.05	0.07
F	−0.17	0.98	−0.57	0.54	−0.15	0.73	−0.36	0.26
P	−0.57	0.15	0.77	1.43	0.03	0.83	0.20	−0.10
S	0.13	−0.38	0.57	−0.22	0.77	−0.19	0.43	−0.16
T	−0.01	−0.35	0.25	−0.06	0.24	−0.27	0.19	0.06
W	0.20	0.60	−0.67	0.40	0.15	0.46	−0.72	0.25
Y	0.82	0.33	−0.54	0.06	−0.10	0.16	−0.64	0.30
V	0.28	0.53	−0.32	0.64	−0.57	0.03	0.02	0.35