Side-chain hydrophobicity scale derived from transmembrane protein folding into lipid bilayers

C. Preston Moon and Karen G. Fleming

Thomas C. Jenkins Department of Biophysics, Johns Hopkins University, 3400 North Charles Street, Baltimore, MD 21218

Edited* by Christopher Miller, Brandeis University, Waltham, MA, and approved April 29, 2011 (received for review March 11, 2011)

The transfer free energies of the twenty natural amino acid side chains from water to phospholipid bilayers make a major contribution to the assembly and function of membrane proteins. Measurements of those transfer free energies will facilitate the identification of membrane protein sequences and aid in the understanding of how proteins interact with membranes during key biological events. We report the first water-to-bilayer transfer free energy scale (i.e., a “hydrophobicity scale”) for the twenty natural amino acid side chains measured in the context of a native transmembrane protein and a phospholipid bilayer. Our measurements reveal parity for apolar side-chain contributions between soluble and membrane proteins and further demonstrate that an arginine side-chain placed near the middle of a lipid bilayer is accommodated with much less energetic cost than predicted by molecular dynamics simulations.

The transfer of amino acid side chains from water into phospholipid lipid bilayers is a fundamental energetic contribution to the total thermodynamic stability of transmembrane proteins (1, 2). A hydrophobicity scale that ranks the water-to-bilayer transfer energies of the twenty natural amino acid side chains could allow the identification of genes that code for membrane proteins (3–6) and aid the understanding of how proteins interact with membranes during key biological events, such as cell signaling, drug binding, and the gating of ion-channels. However, despite the abundance of available hydrophobicity scales, there have been no experimental measurements of water-to-bilayer transfer free energies of amino acid side chains that derive from the partitioning of a native transmembrane protein from water to the interior of a phospholipid bilayer. Existing estimations of water-to-bilayer partitioning of side chains were obtained from systems comprising small compounds to represent proteins and/or organic solvents to mimic bilayers. The only experimental hydrophobicity scale derived from proteins and bilayers (the “translocon scale”) described the translocon-to-bilayer transition (7, 8). Mediated membrane insertion by the translocon complex does not explain the stability of membrane proteins in lipid bilayers relative to a water-solvated state.

Here we achieve thermodynamic measurements of amino acid side-chain transfer into lipid bilayers in the context of a natively folded transmembrane protein spanning a phospholipid bilayer. Our measurements reveal the hydrophobic effect contributes equivalently to the stability of both membrane and soluble proteins (9). Our experiments further demonstrate that a membrane protein can accommodate an arginine side chain placed near the apolar middle of a lipid bilayer with much less cost in energy than has been previously predicted by molecular dynamics simulations (10–11). Moreover, we found the membrane partitioning of two arginine residues to be thermodynamically cooperative. Therefore, our measurements should inform the study of voltage-sensing ion channels, which concertedly move multiple arginines into the hydrophobic interior of membranes during their gating (12, 13).

Results and Discussion

We used the outer membrane phospholipase A (OmpLA) as a transmembrane scaffold on which to introduce amino acid side chains of our choice at various membrane depths. We selected OmpLA because: (a) spontaneously folds and inserts into lipid membranes from a solubilized unfolded state (14), (b) has a known three-dimensional structure (Fig. 1A) (15), and (c) has enzymatic activity that can be monitored to confirm native-like folding (16, 17). To measure side-chain hydrophobicity in a membrane, we selected the alanine at position 210 in OmpLA’s sequence as a host site for substitution to each of the other 19 amino acids as guests. Position 210 is a lipid-facing exterior residue whose α-carbon is located only 0.2 Å from the midplane between OmpLA’s membrane/water interfacial regions (Fig. 1A). Our membranes were large unilamellar vesicles (LUVs) of 1,2-dilauroyl-sn-glycero-3-phosphocholine (DLPC).

To determine the energetics of membrane partitioning of the guest amino acids, we adapted soluble protein chemical denaturation methods to measure the stability of folded proteins using tryptophan fluorescence spectroscopy (18, 19). This approach has only previously been attempted with two other transmembrane proteins folded into lipid vesicles (20–23) and it is only valid if folded and unfolded protein populations come to reversible equilibrium at all concentrations of denaturant (18). We verified reversible folding of OmpLA at pH 3.8 (Fig. S1E) and Fig. S1B and C). We do not yet understand why acidic pH promotes reversible folding of OmpLA or whether this is a common feature for other membrane proteins.

We confirmed the observed folded state of OmpLA was in a native-like transmembrane conformation by verifying that it had enzymatic activity, which requires an acyl-chain of a substrate in the bilayer to bind to a hydrophobic pocket on the transmembrane surface of OmpLA (16, 17). Further, we verified the folded state had its tryptophans embedded in the apolar lipid environment and was protected from protease digestion by being inserted across the membrane (Fig. S2). We determined the unfolded state of OmpLA was not embedded in the membrane and was instead solvated by water by verifying that it was enzymatically inactive, that its tryptophans were in a polar environment whether or not LUVs were present, and it was completely digestible by a protease (Fig. S2).

At equilibrium, OmpLA also adopted a thermodynamic intermediate state at a moderate range of denaturant concentrations, as observed from the fluorescence emission of its tryptophans (Fig. 1B–E and Fig. S1B and C). Therefore, we used a three-state linear-extrapolation model to extract a total standard-state

Author contributions: C.P.M. designed research; C.P.M. performed research; C.P.M. and K.G.F. analyzed data; and C.P.M. and K.G.F. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

See Commentary on page 10027.

1To whom correspondence should be addressed. E-mail: karen.fleming@jhu.edu.

This article contains supporting information online at www.pnas.orglookup/suppl/doi:10.1073/pnas.1103979108/DSupplemental.
are subtracted, the Table S1). Because the contributions of other residues of OmpLA
prise a unique whole-protein hydrophobicity scale (Fig. 2 and sequence variants (Fig. 1
shown in blue. Y are shown in orange. (A)
The correlation of the whole-protein scale and ASAs revealed
surface areas (ASAs) of those side chains and their values in our whole-protein scale (Fig. S3). Importantly, the ASAs were independent of the particular structural context of OmpLA because they were derived from exposure of the guest side chains in an extended model tripeptide. This independence supports the conclusion that our whole-protein scale is applicable to transmembrane proteins of diverse architecture. Indeed, a traditional hydropathy analysis (3) using our whole-protein scale with the Membrane Protein Explorer (MPEx) (5) application correctly predicted the location of all seven transmembrane α-helices of bovine rhodopsin (Fig. S4).

The correlation of the whole-protein scale and ASAs revealed that each Å² of surface area of a hydrophobic side chain removed from water and buried into a membrane added 23 cal mol⁻¹ to the stability of OmpLA. This value is a quantitative description of the hydrophobic effect and constitutes an atomic solvation parameter (ASP) for transferring nonpolar side chains from water into a lipid bilayer. Our ASP matches the ASP previously found for transferring nonpolar side chains from water into the nonpolar solvent octanol (24) as well as that for burying nonpolar side chains into the hydrophobic interior of a water-soluble protein (9). Therefore, the hydrophobic effect stabilizes both membrane and soluble proteins to an equivalent degree. We also used our ASP to calculate an insertion energy for the alanine side chain by multiplying it by alanine’s nonpolar ASA (69 Å²). By shifting the whole-protein scale by alanine’s insertion energy, we can express a scale for all twenty side chains (Table S1) that does not use any particular one of them as a reference. It should be noted that this side-chain scale does not include a contribution of the peptide bond.

Using our thermodynamically rigorous host-guest system, we were well positioned to investigate the energetic cost of water-to-membrane partitioning of arginine. This cost has been a matter of controversy because it concerns the ability of the arginine-rich S4 helix of KvAP and related ion channels to sense membrane depolarization and mediate channel gating (10, 11, 13, 25, 26). Earlier molecular dynamics simulations (10, 11) calculated the total cost of moving a single arginine to the middle of a membrane to be 14–17 kcal mol⁻¹. If the side chain snorkeled toward the water/bilayer interface, the calculated cost would still be
6–7 kcal mol$^{-1}$ (8). The energetics of water-to-membrane partitioning of arginine would also depend on possible rearrangement of other parts of the protein and/or a deformation of the membrane to reduce its hydrophobic thickness (10, 11, 27, 28). However, no such energetics have been measured with a whole protein until the present study. The previous translocon scale had a modest value for arginine (7), but it did not represent the water-to-bilayer transition (8). Our measurement for arginine here does represent a water-to-bilayer transition and it also is quite modest: 2.1 ± 0.1 kcal mol$^{-1}$ (Table S1). We speculate that at least part of the discrepancy between our experimental value and the previous theoretical values may be due to differences in the hydrophobic thicknesses of the lipid bilayers used in the different systems and may be resolved by further calculations that are matched to our experimental conditions.

To address the effects of bilayer depth on side-chain partitioning, we studied side chains at different positions on the OmpLA scaffold. In addition to position 210, we engineered both arginine and leucine at five additional locations in the OmpLA sequence (positions 120, 164, 212, 214, and 223) and compared the resulting variants to alanine variants at the same positions. Arginine partition energies were most unfavorable at the middle of the membrane (Fig. 3 and Table S2) and displayed a shape with a depth-dependence that recapitulated the trends observed in earlier molecular dynamics simulations (8, 10, 11, 27, 28) and in an earlier depth analysis of the translocon scale (6, 7). Leucine partition energies had an opposite response—they were more favorable closer to the middle of the membrane.

We also examined a double-arginine variant (A210R, G212R) to address the issue of tandem membrane insertion of multiple arginine side chains that could occur with KvAP channel gating. Fig. 4 and Table S2 show that the $\Delta \Delta G_{w;l}$ of the double-arginine variant is 1.6 kcal mol$^{-1}$ less than the sum of the $\Delta \Delta G_{w;l}$ for each single-arginine variant (A210R and G212R). Therefore, there is cooperativity between the two arginines that reduces their overall energetic cost to be in the membrane. The source of this cooperativity could be the ability of the two arginines to share access to a deformation or water penetration in the adjacent lipid bilayer. Therefore, our results suggest that the multiple arginines of the S4 helix could function together to suppress their energetic burden to pass across the apolar interior of the membrane during gating of the KvAP channel.

The power of our whole-protein hydrophobicity scale is that it unambiguously reflects the thermodynamics of water-to-bilayer partitioning of side chains in the context of a native transmembrane protein spanning a phospholipid bilayer. The earlier translocon scale (7) did not reflect a system verifiably at equilibrium, and its values are relevant to a different event: translocon-to-bilayer partitioning (8). The Wimley White water-to-octanol scale (24) did derive from equilibrium thermodynamic measurements, however, it represented the protein by short peptides and it approximated the lipid membrane environment with octanol. Our observations reveal that both of those previous scales would undervalue the energetics of most amino acids if they were used to represent water-to-bilayer side-chain partitioning (Fig. S5).

Two exceptions are aspartic acid and glutamic acid, whose energetics are underrepresented in our whole-protein scale because the pH of our experiments was 3.8, which is close to the pK$a$ values for model compounds of Asp and Glu side chains (29). It is reasonable to expect that a significant population of the guest Asp and Glu side chains in our system were protonated and thereby more easily partitioned into the membrane than they would have at normal physiological pH. Therefore, our results for Asp and Glu may help explain why certain membrane active peptides, such as bacterial toxins (30), can partition across membranes only at low pH. This phenomenon of acidic membrane partitioning should be considered in the study of how protein–lipid interactions affect the many types of tumors that feature acidic extracellular pH (31).

Our measurements here provide water-to-bilayer transfer free energies of amino acid side chains determined from the spontaneous insertion of a whole transmembrane protein into membranes. Therefore, our whole-protein free energy measurements are highly applicable to models of spontaneous changes in protein–lipid interactions, such as models of channel gating. To completely describe such events, however, more than just equilibrium energies should be known. Channel gating is a dynamic and kinetic event, so the activation energy barriers to amino acid side-chain insertion into membranes should be determined. Our system of using a scaffold transmembrane protein for guest amino acids should also allow us to measure those kinetic barriers.
Methods

Protein Folding denaturant Titrations. The wild-type OmpLA and its sequence variants were engineered, expressed, and purified as previously described (17). Prior to each experiment, we made fresh unfolded protein stocks dissolved in 8 M guanidine HCl (UltraPure powder, Invitrogen). The background buffer for all fluorescence experiments was 100 mM citrate (Sigma) and 2 mM EDTA (Sigma), pH 3.8. For the variant A210C, we included 1 mM dithiothreitol (DTT, from Sigma) and prepared LUVs as previously described (14), but with 21 passes through two stacked polycarbonate filters having a 0.1 μm pore size.

We set up protein folding and unfolding reactions in three steps. The first step was a dilution of the unfolded protein stocks to a final guanidine concentration of 2.5 M and a final protein concentration of 6.0 μM with 1.4 M 3-N,N-Dimethylmyristyl-ammonio [propanesulfonate (SB3-14, from Sigma)]. The SB3-14 prevented protein aggregation before the LUVs were added. The second step was a threefold dilution of the protein/detergent mixture into LUVs of DLP: to attain a 2,000:1 lipid/protein ratio and to take the SB3-14 below its critical micelle concentration. For titrations for free energy measurements (Fig. 1), the second step kept the protein unfolded at 5.0 M guanidine HCl. For reversible folding verification (Fig. S1A), we prepared two reactions at the second step: one allowing the protein to fold at 2.5 M guanidine HCl and the other keeping the protein unfolded at 5.0 M guanidine HCl. We incubated all samples from the second step at 37 °C with gentle rotation for 5 h. The third step was a fivefold dilution into buffer/guanidine HCl mixtures to attain different final concentrations of guanidine HCl. The final protein concentration after the third step was 400 nM and the final lipid concentration was 2.0 μM. We incubated the titration samples from the third step at 37 °C with gentle rotation for 40–50 h before fluorescence measurements. For free energy measurements, we carried out three titrations of the wild-type OmpLA and two titrations for each of the sequence variants. Each titration was an independent experiment with its own protein, lipid, denaturant, and buffer stocks.

Tryptophan Fluorescence Emission Spectroscopy. Tryptophan fluorescence emission was monitored on an ISS PC1 photon counting spectrofluorometer. The excitation wavelength was 295 nm and the path length was 1 cm. We used an excitation polarizer at 90° with 2.4 mm slits and an emission polarizer at 0° with 2.0 mm slits (32, 33). For samples used for free energy measurements, we collected 350 readings of emission intensity at 330 nm with 0.3 s of signal averaging for each reading. For samples used to show full tryptophan spectral properties (Figs. S1A and S28), we averaged four emission scans from 280 to 400 nm at 1 nm resolution. We found the wavelength position of maximum intensity (λmax) by fitting spectra to a log-normal function (32, 33).

Free Energy Measurements. We globally fit all sets of titration data for all proteins to a three-state linear-extrapolation model (34) using Igor Pro v6.12 (WaveMetrics) to find a common measure of the two m-values in the model (see SI Text). To find total free energies of unfolding for the proteins, we summed the two ΔGm cell values for each set of titration data included in the fit. Our measured free energy of unfolding for the wild-type OmpLA (32.4 ± 0.1 kcal mol−1) is about two- to threefold higher than previously reported values for OmpA (35) and Pgp (23), which are smaller proteins. Thus, the unfolded states of OmpLA and OmpA may be favored considering the different sizes of the proteins and also the variation in the lipid systems used to make the measurements.

Membrane Depth Measurements. We determined where the membrane midplane should intersect OmpLA by using its atomic coordinates in PDB ID code 1do5 (Fig 1A). First, we defined a boundary plane for each side of OmpLA’s hydrophobic surface (extracellular and intracellular) by averaging the set of planes that were formed by bisecting the transmembrane helix of each protein. We incubated the titration samples from the third step at 37 °C with gentle rotation for 40–50 h before fluorescence measurements. For free energy measurements, we carried out three titrations of the wild-type OmpLA and two titrations for each of the sequence variants. Each titration was an independent experiment with its own protein, lipid, denaturant, and buffer stocks.

11. Burgess NK, Dao TP, Stanley AM, Fleming KG (2008) Using tryptophan fluorescence to measure the stability of wide-type OmpLA and two titrations for each of the sequence variants. Each titration was an independent experiment with its own protein, lipid, denaturant, and buffer stocks.

Supporting Information

Moon and Fleming 10.1073/pnas.1103979108

SI Methods

Analysis of Denaturant Titrations. Fig. S1B shows an example of a two-state folding/unfolding model (Eq. 1) that relates to \( \Delta G^* \) to the concentrations of guanidine HCl \((D)\) fit to a titration of the wild-type outer membrane phospholipase A (OmpLA). This model does not describe our data well.

\[
Y_{\text{obs}}([D]) = \left( \left( S_{\text{fold}}^*[D] + Y_{\text{unfolded}}^* \right) + \left( S_{\text{fold}}^*[D] + Y_{\text{folded}}^* \right) \right) + \left( \exp\left( -\left( \Delta G_{w,j}^* + m^*[D]/RT \right) \right) \right) / (1 + \left( \exp\left( -\left( \Delta G_{w,j}^* + m^*[D]/RT \right) \right) \right)),
\]

where \( Y_{\text{unfolded}}^* \) and \( Y_{\text{folded}}^* \) are fluorescence emission intensities in the absence of denaturant for the unfolded and folded conformations, respectively; \( S_{\text{fold}} \) and \( S_{\text{fold}} \) are the slopes of linear baselines in the unfolded and folded regions of the data, respectively; the \( m^* \)-value is a constant that describes how steeply the protein’s free energy depends on \([D]\); \( R \) is the gas constant; and \( T \) is the temperature in Kelvin.

Instead, we used a three-state linear extrapolation model (Eq. 2) (1) to fit our titration data.

\[
Y_{\text{obs}}([D]) = \left( \left( S_{\text{fold}}^*[D] + Y_{\text{unfolded}}^* \right) + \left( \exp\left( -\left( \Delta G_{w,j}^* + m_1^*[D]/RT \right) \right) \right) \left( S_{\text{intermediate}}^*[D] + Y_{\text{intermediate}} \right) \right) + \left( \exp\left( -\left( \Delta G_{m,j}^* + m_2^*[D]/RT \right) \right) \right) \left( S_{\text{fold}}^*[D] + Y_{\text{folded}}^* \right) / (1 + \left( \exp\left( -\left( \Delta G_{m,j}^* + m_2^*[D]/RT \right) \right) \right) \left( S_{\text{intermediate}}^*[D] + Y_{\text{intermediate}} \right)),
\]

where \( Y_{\text{intermediate}} \) is the fluorescence emission intensity in the absence of denaturant for the intermediate conformation; \( S_{\text{fold}} \) is the slope of the linear baseline in the intermediate region of the data; \( \Delta G_{w,j}^* \) and \( \Delta G_{m,j}^* \) are the free energies of the first and second structural transitions, respectively; and the \( m_1 \) and \( m_2 \)-values describe how steeply \( \Delta G_{w,j}^* \) and \( \Delta G_{m,j}^* \), respectively, depend on \([D]\).

An example of a three-state fit to a titration of the wild-type OmpLA is shown in Fig. S1C. Because the intermediate baseline region is not well resolved in the data from some of the variants, the two \( m \)-values from individual fits of those datasets were poorly determined. Therefore, we made the assumption that our sequence substitutions would not appreciably alter the two \( m \)-values of OmpLA (2) and we globally fit all titration data from every sequence variant with Eq. 2 to find common measures of the two \( m \)-values shared by all variants. From the global fit, the \( m \)-value of the first transition was determined to be 2.03 kcal mol\(^{-1}\) M\(^{-1}\) and the \( m \)-value of the second transition was determined to be 7.18 kcal mol\(^{-1}\) M\(^{-1}\). All other parameters in Eq. 2 were determined locally for each dataset. We used Igor Pro v6.12 (www.wavemetrics.com) for all model fitting routines.

Enzymatic Activity Assay. We measured activity in a similar way to a previously reported method (3), except that we aimed to preserve the bilayer structure of our large unilamellar vesicles (LUVs) instead of working with mixed micelles. Being a serine hydrolase, OmpLA is not active at the acidic pH of our normal reversible folding experiments, so we performed all activity measurements at pH 8.0. At that pH, the behavior of the folded state of OmpLA in 1.0 M guanidine HCl is the same as it is at pH 3.8, as judged by tryptophan fluorescence emission and SDS-PAGE. The buffer for our samples at pH 8.0 was 100 mM glycine-glycine, 2 mM EDTA.

We started the activity assay with protein samples from the third step described above that were incubated for 24 h at 37°C in 1.0 M guanidine HCl. We diluted these samples with a buffer/guanidine mixture such that the final protein concentration during measurements would be 48 nM and the final guanidine concentration would remain at 1.0 M. We dried an aliquot of the substrate 2-hexadecanoyl-1-ethylphosphorylcholine (HEPC, from Cayman) briefly under nitrogen and then hydrated it to 25 mg mL\(^{-1}\) in buffer. We added enough hydrated HEPC to the protein samples such that the final HEPC concentration would be 2.5 μM, which is below its reported critical micelle concentration of 3.5–4.5 μM (4). The folded protein/HEPC mixtures were then incubated for at least 12 h at room temperature (22–24°C) to allow full incorporation of the substrate into the lipid bilayers. We also prepared blank samples: one containing protein unfolded in 5 M guanidine HCl in the presence of LUVs and the other containing protein unfolded in 5 M guanidine HCl with no LUVs. To begin activity measurements, we added the secondary substrate 5,5'-dithiobis(2-nitro-benzoic acid) (DTNB, from Cayman) such that it would be at a final concentration of 0.8 mM. For each measurement, we blanked a Beckman Coulter DU 730 spectrophotometer on the mixture following addition of DTNB. We then monitored absorbance at 412 nm over time. After 2 min of baseline collection, we added CaCl\(_2\) to reach a final calcium concentration of 20 mM. Calcium initiates OmpLA’s activity on HEPC because it mediates dimerization of the protein (5, 6). The hydrolysis of HEPC releases a product that then cleaves DTNB producing a yellow moiety. Because DT reactivity with the DTNB, we did not carry out activity measurements on the A210C variant.

Protease Protection Assay. Because the protease experiments were analyzed by SDS-PAGE, we made changes to the normal folding protocol described above. Unfolded protein stocks were made in 10 M urea. Urea solutions were prepared from ultra pure grade powder (Amresco) and then preincubated with AG 50W-X8 resin (BioRad) for at least 1 h prior to addition of buffer ions. The buffer was 100 mM glycine-glycine (Sigma), 10 mM taurine (Sigma), and 2 mM EDTA, pH 8.0. Glycine-glycine and taurine were chosen for their ability to suppress formation of and/or scavenge cyanate ions in the urea (7). The folding reactions had two steps. The first step was a dilution of unfolded protein into LUVs to attain a urea concentration of 4.5 M, a protein concentration of 9.0 μM, and a lipid-protein ratio of 1,000:1. The second folding step was a further threefold dilution to yield a final urea concentration of 1.5 M and a final protein concentration of 3.0 μM, which is high enough to easily visualize as bands in gels. To achieve high folding efficiency and prevent aggregation at this high protein concentration (8), we used LUVs of 1,2-didecanoyl-sn-glycero-3-phosphocholine (DDPC). Following 24 h of incubation at 37°C with gentle rotation, two 100 μL aliquots of each folded protein sample were placed in fresh tubes at room temperature. We added 1.24 μL of trypsin (1 mg mL\(^{-1}\) in 2 mM CaCl\(_2\) and 1 mM HCl) into one of the two aliquots for each sample. The other aliquot was left undigested. All aliquots were then...
incubated for an additional 12 h at room temperature prior to SDS-PAGE. We quenched the samples by mixing them at a 4:1 ratio with 5× gel loading buffer (5) and then split each sample into two portions. For each sample, we immediately placed one portion on ice and then boiled the other portion for 7 min. Otherwise, the quenched samples were never frozen and were never warmed above 4°C until after electrophoresis. Gels were stained with GelCode Blue (Pierce) and were then imaged with an Epson 4490 flatbed scanner, using its transparency lamp.

Hydropathy Analysis. We used the Membrane Protein Explorer (MPEx) application (http://blanco.biomol.uci.edu/mpex) (9) to measure hydropathy values and to predict transmembrane segments of bovine rhodopsin (3cap.pdb). Hydropathy values were relative to the transfer of segments from water into lipid bilayers. The ΔCONH value was set to its default of 0 kcal mol⁻¹. The hydropathy window was 19 residues. All aspartic acid, glutamic acid, and histidine residues were considered to be protonated because our scale was measured at pH 3.8. The known transmembrane segments of rhodopsin were identified using the Orientations of Proteins in Membranes database (http://opm.phar.umich.edu/) (10).

Accessible Surface Area Calculations. Representations of Glycine-X-Glycine tripeptides (X = A, F, L, I, Y, V, and M) were constructed in PyMol (DeLano Scientific). The ASA for each resulting structure was calculated using the program calc-surface (11) using the default probe size of 1.4 Å.


Fig. S1. OmpLA folds reversibly at pH 3.8 in a three-state structural transition across a titration of guanidine HCl. (A) Wavelength position of maximum fluorescence intensity is shown for samples of the wild-type OmpLA at pH 3.8 in different final concentrations of guanidine HCl. Samples were excited by light at 295 nm. Filled blue symbols represent a set of “folding” reactions where samples of protein initially unfolded in 5 M guanidine HCl with LUVs of 1,2-dilauroyl-sn-glycero-3-phosphocholine (DLPC) were diluted to lower final concentrations of guanidine HCl. Open red symbols represent a set of “unfolding” reactions where samples of protein initially folded in 1 M guanidine HCl with LUVs of DLPC were diluted into higher final concentrations of guanidine HCl. (B) Intrinsic fluorescence emission intensity at 330 nm for a similar titration as in Panel A, but with more data points. Solid line represents a two-state reversible equilibrium fit to the data (Eq. 1). (C) Same data points as in panel B, but the solid line represents a three-state reversible equilibrium fit to the data (Eq. 2).
Fig. S2. OmpLA and its sequence variants fold and function in lipid bilayers. (A) Enzymatic activity is shown by a change in absorbance at 412 nm over time for samples of every OmpLA variant in this study (except for A210C). The traces for the folded variants are colored with the same scheme as in Fig. 1. The variants were first folded into LUVs of DLPC with 1 M guanidine HCl at pH 8.0 and then mixed with the substrate HEPC and a secondary substrate DTNB. After 2 min of background data collection, calcium was added to initiate OmpLA’s cleavage of the HEPC substrate. The mixtures turned yellow upon the cleavage product reacting further with the secondary substrate DTNB. Also shown is a lack of activity for two samples of unfolded wild-type OmpLA in 5 M guanidine HCl, one with LUVs of DLPC (black trace) and the other without LUVs (gray trace). (B) Intrinsic tryptophan fluorescence spectra of a representative set of sequence variants (wild-type, A210L, A210M, A210Q, A210R, and A210S) folded into LUVs of DLPC with 1 M guanidine HCl at pH 3.8. The traces for these folded variants are colored with the same scheme as in Fig. 1. Also shown are fluorescence spectra for two samples of unfolded wild-type OmpLA in 5 M guanidine HCl, one with LUVs of DLPC (black trace) and the other without LUVs (gray trace). (C) Protection of folded OmpLA in liposomes from Trypsin digestion. Wild-type OmpLA (Upper) and the A210R variant (Lower) were initially folded in 1.5 M urea at pH 8.0 prior to the addition of Trypsin. Bands denoted “F” contain folded protein and bands denoted “U” contain unfolded protein. Trypsin digestion products appear 1–2 kDa smaller than undigested samples.
Fig. S3. Membrane partition energies of hydrophobic residues strongly correlate with the amount of nonpolar surface area buried in the membrane. The difference in nonpolar accessible surface area (ASA) between alanine and the hydrophobic residues F, L, I, Y, V, and M is plotted on the horizontal axis. The ASAs were calculated using a Gly-X-Gly peptide and a rolling probe with a radius of 1.4 Å. The partition energies from our whole-protein hydrophobicity scale for the same residues are plotted on the vertical axis. Error bars are standard errors of the mean. The solid line represents a linear fit to the data points, having a slope of $0.023 \text{ kcal mol}^{-1} \text{ Å}^{-2}$ and intercepting the vertical axis at $0.164 \text{ kcal mol}^{-1}$.

Fig. S4. Hydropathy analysis using the whole-protein scale successfully predicts the transmembrane segments of bovine rhodopsin. Hydropathy plot for bovine rhodopsin (3cap.pdb) was prepared using MPEx (9) with the $\Delta\text{CONH}$ and window values set to their defaults of $0 \text{ kcal mol}^{-1}$ and 19, respectively. All histidine, aspartic acid, and glutamic acid residues were considered to be protonated. Hydropathy values are for transfer from water into bilayer.
Fig. S5. Comparison of two existing hydrophobicity scales to our whole-protein hydrophobicity scale. The existing scales are plotted on the vertical axis. Our side chain transfer free energy scale from Table S1 ($\Delta G_{wbi}$) is plotted on the horizontal axis. Data points are colored by the same scheme as in Fig. 1. Solid lines are linear fits through the data points shown with filled symbols. Correlation coefficients for the fits are shown in the lower right of each panel. Data points shown with open symbols were left out of the linear fits. Data points for aspartic acid and glutamic acid were left out of all the linear fits because their protonation states are likely sensitive to the particular pH of our experiments (3.8) and that pH is not common to the other scales. (A) Translocon-to-bilayer transfer scale (1). The data point for helix-breaker proline was left out of the linear fit because the translocon scale used for the first structural transition was determined by a global fit to all sequence variants used in this study. (B) Wimley White side chain water-to-octanol transfer scale (2). Two different data points are shown for aspartic acid and for glutamic acid, one each for the deprotonated state from the WW scale determined at pH 9.0 and for the protonated state from the WW scale determined at pH 1.0. The slope of the line is 0.50, the line intercepts the vertical axis at 0.64 kcal mol$^{-1}$, and $R = 0.89$. (C) Wimley White side chain water-to-octanol transfer scale (3). Two different data points are shown for aspartic acid and for glutamic acid, one each for the deprotonated state from the WW scale determined at pH 9.0 and one each for the protonated state from the WW scale determined at pH 1.0. The slope of the line is 0.50, the line intercepts the vertical axis at 0.64 kcal mol$^{-1}$, and $R = 0.89$.


Table S1. Free energies of unfolding in water and lipids for OmpLA sequence variants at position 210 from the two equilibrium structural transitions during guanidine HCl titrations

<table>
<thead>
<tr>
<th>Variant</th>
<th>$\Delta G_{w,l}^{1}$ (kcal mol$^{-1}$)</th>
<th>$\Delta G_{w,l}^{2}$ (kcal mol$^{-1}$)</th>
<th>$\Delta G_{w,l}^{tot}$ (kcal mol$^{-1}$)</th>
<th>$\Delta G_{wbi}^{1}$ (kcal mol$^{-1}$)</th>
<th>$\Delta G_{wbi}^{2}$ (kcal mol$^{-1}$)</th>
<th>$\Delta G_{wbi}^{tot}$ (kcal mol$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT (A)</td>
<td>6.49 ± 0.12</td>
<td>25.97 ± 0.04</td>
<td>32.45 ± 0.11</td>
<td>-1.57†</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A210C</td>
<td>6.31 ± 0.04</td>
<td>25.66 ± 0.13</td>
<td>31.97 ± 0.09</td>
<td>0.49 ± 0.15</td>
<td>-1.08</td>
<td></td>
</tr>
<tr>
<td>A210D</td>
<td>5.02 ± 0.12</td>
<td>24.49 ± 0.02</td>
<td>29.51 ± 0.14</td>
<td>2.95 ± 0.18</td>
<td>1.38</td>
<td></td>
</tr>
<tr>
<td>A210E</td>
<td>5.71 ± 0.02</td>
<td>25.11 ± 0.01</td>
<td>30.82 ± 0.03</td>
<td>1.64 ± 0.12</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>A210F</td>
<td>5.94 ± 0.35</td>
<td>28.71 ± 0.10</td>
<td>34.65 ± 0.24</td>
<td>-2.20 ± 0.27</td>
<td>-3.77</td>
<td></td>
</tr>
<tr>
<td>A210G</td>
<td>6.02 ± 0.16</td>
<td>24.72 ± 0.13</td>
<td>30.74 ± 0.03</td>
<td>1.72 ± 0.12</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td>A210H</td>
<td>3.86 ± 0.28</td>
<td>23.83 ± 0.05</td>
<td>27.69 ± 0.24</td>
<td>4.76 ± 0.26</td>
<td>3.19</td>
<td></td>
</tr>
<tr>
<td>A210I</td>
<td>5.37 ± 0.35</td>
<td>28.64 ± 0.02</td>
<td>34.01 ± 0.33</td>
<td>-1.56 ± 0.35</td>
<td>-3.12</td>
<td></td>
</tr>
<tr>
<td>A210J</td>
<td>3.15 ± 0.50</td>
<td>23.92 ± 0.00</td>
<td>27.07 ± 0.51</td>
<td>5.39 ± 0.52</td>
<td>3.82</td>
<td></td>
</tr>
<tr>
<td>A210L</td>
<td>6.09 ± 0.09</td>
<td>28.18 ± 0.16</td>
<td>34.27 ± 0.07</td>
<td>-1.81 ± 0.13</td>
<td>-3.32</td>
<td></td>
</tr>
<tr>
<td>A210M</td>
<td>6.45 ± 0.01</td>
<td>26.76 ± 0.14</td>
<td>33.21 ± 0.15</td>
<td>-0.76 ± 0.19</td>
<td>-2.33</td>
<td></td>
</tr>
<tr>
<td>A210N</td>
<td>4.95 ± 0.19</td>
<td>24.03 ± 0.07</td>
<td>28.98 ± 0.26</td>
<td>3.47 ± 0.28</td>
<td>1.91</td>
<td></td>
</tr>
<tr>
<td>A210P</td>
<td>6.95 ± 0.01</td>
<td>27.02 ± 0.06</td>
<td>33.97 ± 0.05</td>
<td>-1.52 ± 0.12</td>
<td>-3.09</td>
<td></td>
</tr>
<tr>
<td>A210Q</td>
<td>5.08 ± 0.07</td>
<td>24.37 ± 0.06</td>
<td>29.45 ± 0.02</td>
<td>3.01 ± 0.11</td>
<td>1.44</td>
<td></td>
</tr>
<tr>
<td>A210R</td>
<td>4.53 ± 0.03</td>
<td>24.21 ± 0.09</td>
<td>28.74 ± 0.07</td>
<td>3.71 ± 0.13</td>
<td>2.14</td>
<td></td>
</tr>
<tr>
<td>A210S</td>
<td>6.10 ± 0.07</td>
<td>24.52 ± 0.13</td>
<td>30.62 ± 0.19</td>
<td>1.83 ± 0.22</td>
<td>0.26</td>
<td></td>
</tr>
<tr>
<td>A210T</td>
<td>6.18 ± 0.26</td>
<td>24.49 ± 0.06</td>
<td>30.67 ± 0.32</td>
<td>1.78 ± 0.34</td>
<td>0.21</td>
<td></td>
</tr>
<tr>
<td>A210V</td>
<td>5.61 ± 0.32</td>
<td>27.62 ± 0.06</td>
<td>33.23 ± 0.26</td>
<td>-0.78 ± 0.28</td>
<td>-2.34</td>
<td></td>
</tr>
<tr>
<td>A210W</td>
<td>5.49 ± 0.09</td>
<td>27.35 ± 0.09</td>
<td>32.84 ± 0.17</td>
<td>-0.38 ± 0.21</td>
<td>-1.95</td>
<td></td>
</tr>
<tr>
<td>A210Y</td>
<td>6.71 ± 0.10</td>
<td>26.83 ± 0.07</td>
<td>33.55 ± 0.03</td>
<td>-1.09 ± 0.12</td>
<td>-2.66</td>
<td></td>
</tr>
</tbody>
</table>

Standard errors of the mean are shown from independent titrations ($n = 2$).

*An equilibrium $m$-value of 2.03 kcal mol$^{-1}$ M$^{-1}$ for the first structural transition was determined by a global fit to all guanidine titrations for all sequence variants used in this study.

†An equilibrium $m$-value of 7.18 kcal mol$^{-1}$ M$^{-1}$ for the second structural transition was determined by a global fit to all guanidine titrations for all sequence variants used in this study.

‡Change in stability with respect to the wild type.

§Water-to-bilayer (wbi) transfer free energies for the amino acid side chains (sc) determined by subtracting the transfer free energy of alanine from the $\Delta G_{wbi}$ of the side chain variants.

¶Transfer free energy of alanine comes from its nonpolar ASA in a model tripeptide (69.1 Å$^2$) multiplied by the slope of the line in Fig. S3 (0.023 kcal mol$^{-1}$ Å$^2$).
Table S2. Free energies of unfolding in water and lipids for OmpLA sequence variants at different membrane depths and for a double arginine variant from the two equilibrium structural transitions during guanidine HCl titrations

<table>
<thead>
<tr>
<th>Position/Variant</th>
<th>First transition $\Delta G_{w,l}^\circ$ (kcal mol$^{-1}$)*</th>
<th>Second transition $\Delta G_{w,l}^\circ$ (kcal mol$^{-1}$)†</th>
<th>Total $\Delta G_{w,l}^\circ$ (kcal mol$^{-1}$)</th>
<th>$\Delta \Delta G_{w,l}^\circ$ (kcal mol$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>120</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L120A</td>
<td>5.20 ± 0.11</td>
<td>25.10 ± 0.08</td>
<td>30.03 ± 0.03</td>
<td></td>
</tr>
<tr>
<td>WT</td>
<td>6.49 ± 0.12</td>
<td>25.97 ± 0.04</td>
<td>32.45 ± 0.11</td>
<td>−2.16 ± 0.09†</td>
</tr>
<tr>
<td>L120R</td>
<td>3.34 ± 0.16</td>
<td>24.60 ± 0.05</td>
<td>27.95 ± 0.11</td>
<td>2.35 ± 0.12‡</td>
</tr>
<tr>
<td>164</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WT</td>
<td>6.49 ± 0.12</td>
<td>25.97 ± 0.04</td>
<td>32.45 ± 0.11</td>
<td>−2.16 ± 0.09†</td>
</tr>
<tr>
<td>A164L</td>
<td>6.84 ± 0.09</td>
<td>26.82 ± 0.10</td>
<td>33.66 ± 0.19</td>
<td>−1.21 ± 0.21†</td>
</tr>
<tr>
<td>A164R</td>
<td>6.00 ± 0.27</td>
<td>25.65 ± 0.02</td>
<td>31.66 ± 0.26</td>
<td>0.80 ± 0.27†</td>
</tr>
<tr>
<td>210</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WT</td>
<td>6.49 ± 0.12</td>
<td>25.97 ± 0.04</td>
<td>32.45 ± 0.11</td>
<td>−2.16 ± 0.09†</td>
</tr>
<tr>
<td>A210L</td>
<td>6.09 ± 0.09</td>
<td>28.18 ± 0.16</td>
<td>34.27 ± 0.07</td>
<td>−1.81 ± 0.10§</td>
</tr>
<tr>
<td>A210R</td>
<td>4.53 ± 0.03</td>
<td>24.21 ± 0.09</td>
<td>28.74 ± 0.07</td>
<td>3.71 ± 0.11§</td>
</tr>
<tr>
<td>212</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G212A</td>
<td>5.36 ± 0.34</td>
<td>27.70 ± 0.20</td>
<td>33.06 ± 0.14</td>
<td>−0.60 ± 0.16†</td>
</tr>
<tr>
<td>G212L</td>
<td>5.57 ± 0.18</td>
<td>30.11 ± 0.02</td>
<td>35.68 ± 0.16</td>
<td>−2.62 ± 0.21†</td>
</tr>
<tr>
<td>G212R</td>
<td>5.45 ± 0.01</td>
<td>24.55 ± 0.07</td>
<td>30.00 ± 0.06</td>
<td>3.06 ± 0.16†</td>
</tr>
<tr>
<td>214</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y214A</td>
<td>5.68 ± 0.00</td>
<td>24.40 ± 0.01</td>
<td>30.09 ± 0.01</td>
<td>2.37 ± 0.08§</td>
</tr>
<tr>
<td>Y214L</td>
<td>6.01 ± 0.05</td>
<td>25.29 ± 0.02</td>
<td>31.30 ± 0.03</td>
<td>−1.21 ± 0.04§</td>
</tr>
<tr>
<td>Y214R</td>
<td>5.39 ± 0.03</td>
<td>24.08 ± 0.01</td>
<td>29.47 ± 0.02</td>
<td>0.61 ± 0.02¶</td>
</tr>
<tr>
<td>223</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WT</td>
<td>6.49 ± 0.12</td>
<td>25.97 ± 0.04</td>
<td>32.45 ± 0.11</td>
<td>−2.16 ± 0.09†</td>
</tr>
<tr>
<td>A223L</td>
<td>6.33 ± 0.08</td>
<td>27.96 ± 0.01</td>
<td>34.29 ± 0.07</td>
<td>−1.83 ± 0.11†</td>
</tr>
<tr>
<td>A223R</td>
<td>5.50 ± 0.11</td>
<td>24.88 ± 0.11</td>
<td>30.38 ± 0.22</td>
<td>2.07 ± 0.23¶</td>
</tr>
<tr>
<td>Double ARG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G212A</td>
<td>5.36 ± 0.34</td>
<td>27.70 ± 0.20</td>
<td>33.06 ± 0.14</td>
<td>5.13 ± 0.14‖</td>
</tr>
<tr>
<td>A210R, 212R</td>
<td>4.47 ± 0.07</td>
<td>23.46 ± 0.06</td>
<td>27.93 ± 0.01</td>
<td></td>
</tr>
</tbody>
</table>

Standard errors of the mean are shown from independent titrations ($n=2$).

*An equilibrium $m$-value of 2.03 kcal mol$^{-1}$ M$^{-1}$ for the first structural transition was determined by a global fit to all guanidine titrations for all sequence variants used in this study.

†An equilibrium $m$-value of 7.18 kcal mol$^{-1}$ M$^{-1}$ for the second structural transition was determined by a global fit to all guanidine titrations for all sequence variants used in this study.

‡Change in stability with respect to the L120A variant.

§Change in stability with respect to the wild type.

¶Change in stability with respect to the Y214A variant.

‖Change in stability with respect to the G212A variant.

∥Change in stability with respect to the Y214A variant.