Supporting Information

Moon and Fleming 10.1073/pnas.1103979108

SI Methods

Analysis of Denaturant Titrations. Fig. S1*B* shows an example of a two-state folding/unfolding model (Eq. **1**) that relates to $\Delta G^{\circ}_{w,l}$ to the variation of observed fluorescence emission intensities (Y_{obs}) 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.

$$\begin{aligned} Y_{obs}([D]) = &((S_{unf}{}^{*}[D] + Y_{unf,w,l}) \\ &+ ((S_{fold}{}^{*}[D] \\ &+ Y_{fold,w,l})^{*}(\exp(-(\Delta G^{\circ}_{w,l} + m^{*}[D])/RT))))/(1 \\ &+ (\exp(-(\Delta G^{\circ}_{w,l} + m^{*}[D])/RT))), \end{aligned} \quad \textbf{[S1]}$$

where $Y_{unf,w,l}$and $Y_{fold,w,l}$ are fluorescence emission intensities in the absence of denaturant for the unfolded and folded conformations, respectively; S_{unf} and S_{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.

$$\begin{aligned} Y_{obs}([D]) = &(((S_{fold}{}^{*}[D] + Y_{fold,w,l}) \\ &+ ((\exp(-(DG^{\circ}_{1,w,l} + m_1{}^{*}[D])/RT))^{*}(S_{int}{}^{*}[D] + Y_{int,w,l})) \\ &+ ((\exp(-(\Delta G^{\circ}_{1,w,l} + m_1{}^{*}[D])/RT))^{*}(\exp(-(\Delta G^{\circ}_{2,w,l} \\ &+ m_2{}^{*}[D])/RT))^{*}(S_{unf}{}^{*}[D] + Y_{unf,w,l})))))/(1 \\ &+ (\exp(-(\Delta G^{\circ}_{1,w,l} + m_1{}^{*}[D])/RT)) \\ &+ ((\exp(-(\Delta G^{\circ}_{1,w,l} + m_1{}^{*}[D])/RT))^{*}(\exp(-(\Delta G^{\circ}_{2,w,l} \\ &+ m_2{}^{*}[D])/RT)))), \end{aligned} \quad \textbf{[S2]}$$

where $Y_{int,w,l}$ is the fluorescence emission intensity in the absence of denaturant for the intermediate conformation; S_{int} is the slope of the linear baseline in the intermediate region of the data; $\Delta G^{\circ}_{2,w,l}$ and $\Delta G^{\circ}_{2,w,l}$ are the free energies of the first and second structural transitions, respectively; and the m_1 and m_2-values describe how steeply $\Delta G^{\circ}_{1,w,l}$ and $\Delta G^{\circ}_{2,w,l}$, respectively, depend on [*D*].

An example of a three-state fit to a titration of the wild-type OmpLA is shown in Fig. S1*C*. 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 DTT reacts 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 501-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.

1. Latypov RF, Cheng H, Roder NA, Zhang J, Roder H (2006) Structural characterization of an equilibrium unfolding intermediate in Cytochrome c. *J Mol Biol* 357:1009–1025.
2. Huysmans GHM, Baldwin SA, Brockwell DJ, Radford SE (2010) The transition state for folding of an outer membrane protein. *Proc Natl Acad Sci USA* 107:4099–4104.
3. Dekker N, Tommassen J, Lustigi A, Rosenbusch JP, Verheij HM (1997) Dimerization regulates the enzymatic activity of *Escherichia coli* outer membrane phospholipase A. *J Biol Chem* 272:3179–3184.
4. Aarsman AJ, van den Bosch H (1979) A comparison of acyl-oxyester and acyl-thioester substrates for some lipolytic enzymes. *Biochim Biophys Acta* 572:519–530.
5. Stanley AM, Fleming KG (2007) The role of a hydrogen bonding network in the transmembrane β-barrel OMPLA. *J Mol Biol* 370:912–924.
6. Ubarretxena-Belandia I, Boots JWP, Verheij HM, Dekker N (1998) Role of the cofactor calcium in the activation of outer membrane phospholipase A. *Biochemistry* 37:16011–16018.
7. Lin MF, Williams C, Murray MV, Conn G, Ropp PA (2004) Ion chromatographic quantification of cyanate in urea solutions: Estimation of the efficiency of cyanate scavengers for use in recombinant protein manufacturing. *J Chromatogr B* 803:353–362.
8. Burgess NK, Dao TP, Stanley AM, Fleming KG (2008) β-Barrel proteins that reside in the *Escherichia coli* outer membrane in vivo demonstrate varied folding behavior in vitro. *J Biol Chem* 283:26748–26758.
9. Snider C, Jayasinghe S, Hristova K, White SH (2009) MPEx: A tool for exploring membrane proteins. *Protein Sci* 18:2624–2628.
10. Lomzie AL, Pogozheva ID, Lomzie MA, Mosberg HI (2006) Positioning of proteins in membranes: A computational approach. *Protein Sci* 15:1318–1333.
11. Gerstein M, Lynden-Bell RM (1993) What is the natural boundary of a protein in solution? *J Mol Biol* 230:641–650.

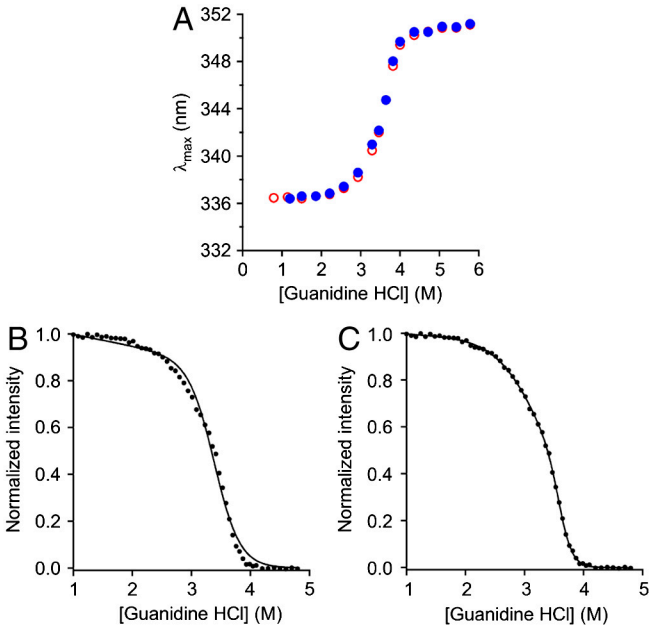

Fig. S1. OmpLA folds reversibly at pH 3.8 in a three-state structural transition across a titration of guanidine HCl. (*A*) Wavelength position of maximum fluorescence intensity is shown for samples of the wild-type OmpLA at pH 3.8 in different final concentrations of guanidine HCl. Samples were excited 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**).

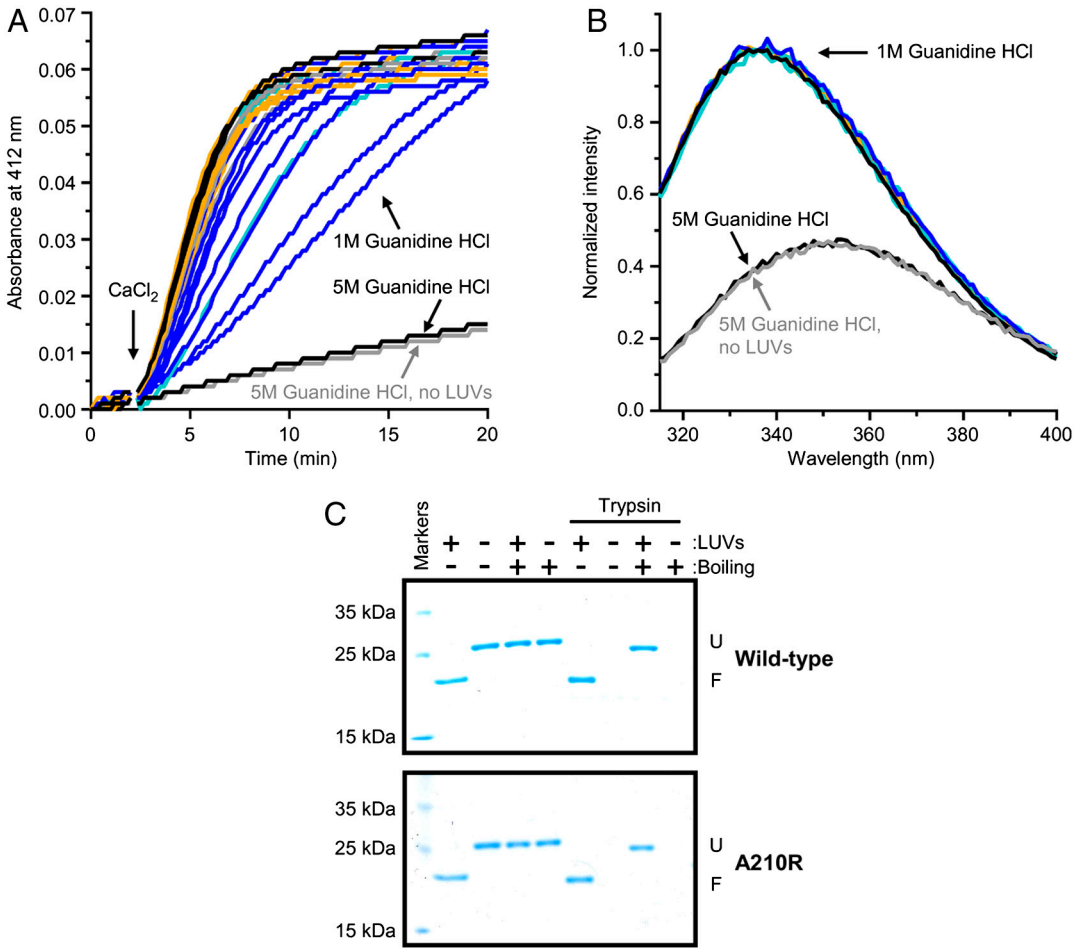

Fig. S2. OmpLA and its sequence variants fold and function in lipid bilayers. (*A*) Enzymatic activity is shown by a change in absorbance at 412 nm over time for samples of every OmpLA variant in this study (except for A210C). The traces for the folded variants are colored with the same scheme as in 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 1M 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.

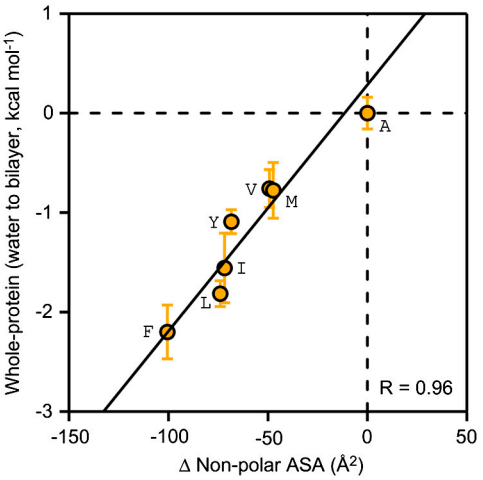

Fig. S3. Membrane partition energies of hydrophobic residues strongly correlate with the amount of nonpolar surface area buried in the membrane. The difference in nonpolar accessible surface area (ASA) between alanine and the hydrophobic residues F, L, I, Y, V, and M is plotted on the horizontal axis. 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 kcal mol^{-1} $Å^{-2}$ and intercepting the vertical axis at 0.164 kcal mol^{-1}.

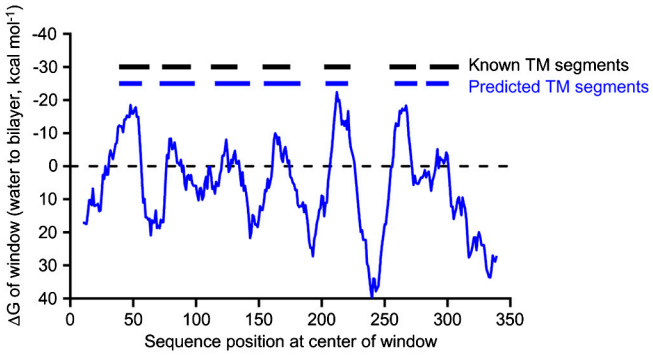

Fig. S4. Hydropathy analysis using the whole-protein scale successfully predicts the transmembrane segments of bovine rhodopsin. Hydropathy plot for bovine rhodopsin (3cap.pdb) was prepared using MPEx (9) with the ΔCONH and window values set to their defaults of 0 kcal mol^{-1} and 19, respectively. All histidine, aspartic acid, and glutamic acid residues were considered to be protonated. Hydropathy values are for transfer from water into bilayer.

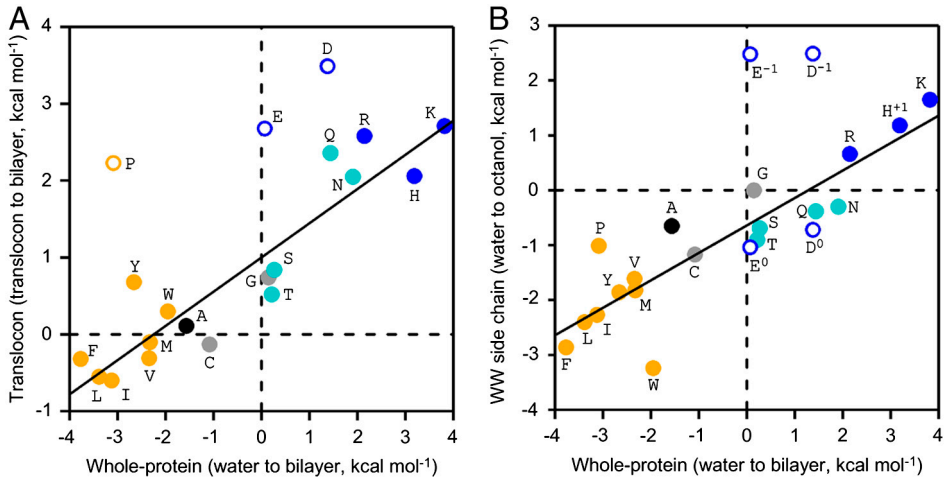

Fig. S5. Comparison of two existing hydrophobicity scales to our whole-protein hydrophobicity scale. The existing scales are plotted on the vertical axis. Our side chain transfer free energy scale from Table S1 (ΔG^{sc}_{wbi}) is plotted on the horizontal axis. Data points are colored by the same scheme 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 α-helical transmembrane segments. The slope of the line is 0.45 and the line intercepts the vertical axis at 1.00 kcal mol^{-1}, and $R = 0.92$. (*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 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 $R = 0.89$.

1 Hessa T et al. (2005) Recognition of transmembrane helices by the endoplasmic reticulum translocon. *Nature* 433:377–381.
2 Wimley WC, Creamer TP, White SH (1996) Solvation energies of amino acid side chains and backbone in a family of host-guest pentapeptides. *Biochemistry* 35:5109–5124.

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

Variant	First transition $\Delta G^{\circ}_{w,l}$ (kcal mol^{-1})*	Second transition $\Delta G^{\circ}_{w,l}$ (kcal mol^{-1})†	Total $\Delta G^{\circ}_{w,l}$ (kcal mol^{-1})	$\Delta\Delta G^{\circ}_{w,l}$ (kcal mol^{-1})$\Delta G^{\circ}_{w,l}$‡	ΔG^{sc}_{wbi} (kcal mol^{-1})§
WT (A)	6.49 ± 0.12	25.97 ± 0.04	32.45 ± 0.11		−1.57¶
A210C	6.31 ± 0.04	25.66 ± 0.13	31.97 ± 0.09	0.49 ± 0.15	−1.08
A210D	5.02 ± 0.12	24.49 ± 0.02	29.51 ± 0.14	2.95 ± 0.18	1.38
A210E	5.71 ± 0.02	25.11 ± 0.01	30.82 ± 0.03	1.64 ± 0.12	0.07
A210F	5.94 ± 0.35	28.71 ± 0.10	34.65 ± 0.24	−2.20 ± 0.27	−3.77
A210G	6.02 ± 0.16	24.72 ± 0.13	30.74 ± 0.03	1.72 ± 0.12	0.15
A210H	3.86 ± 0.28	23.83 ± 0.05	27.69 ± 0.24	4.76 ± 0.26	3.19
A210I	5.37 ± 0.35	28.64 ± 0.02	34.01 ± 0.33	−1.56 ± 0.35	−3.12
A210K	3.15 ± 0.50	23.92 ± 0.00	27.07 ± 0.51	5.39 ± 0.52	3.82
A210L	6.09 ± 0.09	28.18 ± 0.16	34.27 ± 0.07	−1.81 ± 0.13	−3.32
A210M	6.45 ± 0.01	26.76 ± 0.14	33.21 ± 0.15	−0.76 ± 0.19	−2.33
A210N	4.95 ± 0.19	24.03 ± 0.07	28.98 ± 0.26	3.47 ± 0.28	1.91
A210P	6.95 ± 0.01	27.02 ± 0.06	33.97 ± 0.05	−1.52 ± 0.12	−3.09
A210Q	5.08 ± 0.07	24.37 ± 0.06	29.45 ± 0.02	3.01 ± 0.11	1.44
A210R	4.53 ± 0.03	24.21 ± 0.09	28.74 ± 0.07	3.71 ± 0.13	2.14
A210S	6.10 ± 0.07	24.52 ± 0.13	30.62 ± 0.19	1.83 ± 0.22	0.26
A210T	6.18 ± 0.26	24.49 ± 0.06	30.67 ± 0.32	1.78 ± 0.34	0.21
A210V	5.61 ± 0.32	27.62 ± 0.06	33.23 ± 0.26	−0.78 ± 0.28	−2.34
A210W	5.49 ± 0.09	27.35 ± 0.09	32.84 ± 0.17	−0.38 ± 0.21	−1.95
A210Y	6.71 ± 0.10	26.83 ± 0.07	33.55 ± 0.03	−1.09 ± 0.12	−2.66

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 $\Delta\Delta G^{\circ}_{w,l}$ 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

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

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 G212A variant.

‖Change in stability with respect to the Y214A variant.