Local Bilayer Hydrophobicity Modulates Membrane Protein Stability

Dagan C. Marx and Karen G. Fleming*

Thomas C. Jenkins Department of Biophysics, Johns Hopkins University, Baltimore, MD.

Keywords: membrane protein folding, transfer free energies, nonpolar solvation parameter, translocon, protein stability.

ABSTRACT: Through the insertion of nonpolar side chains into the bilayer, the hydrophobic effect has long been accepted as a driving force for membrane protein folding. However, how the changing chemical composition of the bilayer affects the magnitude side chain transfer free energies (ΔG_{sc}°) has historically not been well understood. A particularly challenging region for experimental interrogation is the bilayer interfacial region that is characterized by a steep polarity gradient. In this study we have determined the ΔG_{sc}° for nonpolar side chains as a function of bilayer position using a combination of experiment and simulation. We discovered an empirical correlation between the surface area of nonpolar side chain, the transfer free energies, and the local water concentration in the membrane that allows for ΔG_{sc}° to be accurately estimated at any location in the bilayer. Using these water-to-bilayer ΔG_{sc}° values, we calculated the interface-to-bilayer transfer free energy ($\Delta G_{i,b}^{\circ}$). We find that the $\Delta G_{i,b}^{\circ}$ are similar to the "biological", translocon-based transfer free energies, indicating that the translocon energetically mimics the bilayer interface. Together these findings can be applied to increase the accuracy of computational workflows used to identify and design membrane proteins, as well as bring greater insight into our understanding of how disease-causing mutations affect membrane protein folding and function.

Introduction

The cellular functions of a protein are regulated by the intrinsic thermodynamic stability of the three-dimensional structure adopted by its polypeptide chain.[1–4] The stabilities of many water-soluble proteins have been experimentally measured *in vitro*, resulting in an understanding of the forces that regulate structure and function of this class of proteins.[5] For membrane proteins, which make up approximately one third of the eukaryotic proteome,[6] clarity on the balance of forces that determine their thermodynamic stabilities is much less mature.[7] This lag in knowledge derives from the fact that membrane proteins are aggregation prone in aqueous solutions and must embed themselves into a hydrophobic environment such as detergent micelles or phospholipid bilayers to fold and function.[8–10] The aggregation propensity renders traditional experimental approaches for measuring protein stabilities untenable for most membrane proteins.[7]

The stabilities of membrane proteins are directly linked to the energetics of burying amino acid side chains into the phospholipid bilayer, also referred to as the side chain transfer free energy change (ΔG_{sc}°).[7,10–12] Both the magnitude and sign of the water-to-bilayer ΔG_{sc}° for a side chain in a transmembrane domain (TMD) depend on the exact chemical compositions of the two end-point solvents (water and bilayer).[13–22] There is the additional complexity in that a bilayer is not a simple end-point solvent. Biological membranes contain two chemically distinct regions: the largely desolvated hydrocarbon core and the chemically-heterogeneous interface.[23,24] Along the z-coordinate of the bilayer normal, each interface is an approximately 15 Å wide region separating bulk water from the dehydrated center, and the transition between these two end-points is characterized by a continuously changing water concentration.[15,24] This water gradient should result in a position-dependence on the free energies of transfer for side chains along the interface. Arguably, these membrane interfaces represent half of the bilayer volume, and a full understanding of membrane protein folding forces depends on the energetics imposed by this steeply changing polarity of this region.

The hydrophobic effect has empirically been related to the (water) solvent accessible surface area (ASA) of the side chain through an energy termed the nonpolar solvation parameter, σ_{NP}.[16,18,19,25–28] This relationship quantifies the energy (cal mol^{-1}) gained per Å^2 of nonpolar surface area removed from water and buried in a nonpolar solvent. Across a wide range of model-protein systems and membrane mimics, the σ_{NP} value shows a remarkably precise value, ranging from -23 to -25 cal mol^{-1} Å^{-2}.[16,18,19] In the vast majority of studies, this singular σ_{NP} has traditionally been the basis for estimating side chain-bilayer partitioning as well as for identifying polypeptide sequence stretches that are likely to adopt transmembrane location. An exception to this is a single water-to-bilayer-interface study that found a σ_{NP} value for that reaction equal to -12 cal mol^{-1} Å^{-2}.

One challenge with all of these studies is a precise mapping of the z-position in the membrane to which each σ_{NP} corresponds.[17] Moreover, how to relate partitioning energies to a bilayer location is arguably impossible to know for those studies employing organic solvents as the bilayer mimic. Even with these drawbacks, usage of the nonpolar solvation parameter and accompanying transfer free energy changes for

nonpolar side chains have proven to be very powerful for identification of transmembrane segments from sequence.

Nevertheless, these energy values are still a somewhat crude tool for estimating the stabilities of these polypeptide regions at different locations in the chemically-heterogeneous bilayer. Here we take advantage of the naturally-occurring water gradient in the bilayer to ask how this changing water concentration affects the energetics of nonpolar side chains and accordingly, the nonpolar solvation parameter. We accomplish these measurements using a natively-folded protein that possesses a very favorable folding free energy and that is firmly anchored in its position within the experimental bilayer.[29] Using this setup, we have resolved the bilayer *z*-position dependence of σ_{NP} using a combination of protein folding titrations, thermodynamic cycles and molecular dynamics simulations.

We discovered a linear correlation between σ_{NP} and the local water concentration across the bilayer normal. This relationship essentially quantifies the hydrophobic effect across the bilayer interface and enables prediction of the functional form for a depth-dependent solvation parameter, $\sigma_{NP}(z)$, and more generally, functions describing side-chain transfer free energies, $\Delta G^{\circ}_{sc}(z)$, across the bilayer. These energy values should improve the accuracy of membrane protein structure prediction and design algorithms. Additionally, we use our findings to better understand the thermodynamic basis for how the Sec translocon aids α-helical membrane protein folding *in vivo*.

Materials and Methods

OmpLA Variant Expression, Purification, Activity Assay and Folding Titrations. The OmpLA variants were engineered, expressed, and purified from inclusion bodies as previously described.[15,16] Each variant was verified to have enzymatic activity while folded into DLPC large unilamellar vesicles (LUV) using a previously developed colorimetric assay.[15,16] Folding titrations were also set up and performed as previously described in detail.[15,16,18,30,31] Briefly, unfolded OmpLA variants added to DLPC LUVs (diameter = 100nm) in 5 M guanidine hydrochloride (GdnHCl). After overnight incubation at 37 ºC, samples were diluted to varying GdnHCl concentrations ranging from 1-5 M at 0.08 M increments (51 total samples). Following at least a 40-hour incubation in a rotating incubator at 37 ºC, intrinsic tryptophan fluorescence was measured at 330 nm on a PC1 Fluorometer with cross-polarization (ISS, Champaign, IL, USA).

Fitting OmpLA variant $\Delta G^{\circ}_{w,l}$ and σ_{NP}. The resulting titration data were fit to a three-state linear extrapolation model to extract the stability ($\Delta G^{\circ}_{w,l}$) of each variant using previously determined *m*-values using the Igor Pro 8 software package (Wavemetrics, Portland, OR, USA).[15,16] Host (alanine, Ala) variant $\Delta G^{\circ}_{w,l}$ were previously determined and used to calculate $\Delta\Delta G^{\circ}_{w,l}$ for each nonpolar guest variant.[15] The σ_{NP} for each site on OmpLA was determined by weighted linear regression using Igor Pro, with weights corresponding to the error (standard deviations) in $\Delta\Delta G^{\circ}_{w,l}$:

$$\Delta\Delta G^{\circ}_{w,l} = \sigma_{NP} * \Delta ASA + b \quad (1).$$

where

$$\Delta ASA = ASA_{Guest} - ASA_{Ala} \quad (2).$$

and σ_{NP} and *b* are fitted parameters. The surface area for each nonpolar side chain was previously determined in the context of a Gly-X-Gly peptide.[16]

Molecular Dynamics Simulations of OmpLA variants. All-atom molecular dynamics simulations of each OmpLA variant using the CHARMM36 force field were performed to determine the position of each guest side chain in the experimental DLPC bilayer. The monomeric crystal structure of OmpLA (1qd5) was used as the starting structure, and the remaining components of the system and scripts were created using CHARMM-GUI.[32–34] A system was built for each OmpLA variant embedded in a DLPC bilayer and 0.2M NaCl at 37ºC, resulting in 33 individual simulations. To approximate the experimental pH, some ionizable residues were protonated as described previously.[15] Systems were initially equilibrated using the protocol provided by CHARMM-GUI, followed by an additional 50ns of equilibration to allow for the system to fully relax (Figure S1). All simulations used NAMD and were run on the Maryland Advanced Research Computing Center super computer.[35]

VMD plug-ins and homemade scripts were used to analyze all trajectories.[36] To determine the position of the Cα for each variant relative to the phosphate plane, we first calculated the average *z*-position for the phosphate atoms in each leaflet of the bilayer. We then calculated the distance of the Cα for the variant side chain to the closest phosphate plane. These calculations were performed for each frame in the last 100ns of the trajectory, and the distances from the phosphate plane were binned into 0.1Å bins to create the histograms in Figure S2. Histograms were fit to Gaussian distributions using Igor Pro software to determine average positions and standard deviations found in Table S1.

To relate the Cα positions of each side chain to the local concentration of water ([water]) in the bilayer we derived an empirical relationship between [water] and *z*-position in the bilayer (Figure S3). These data were derived from simulations of a neat DLPC bilayer and extracted using the Density Profile Tool in VMD.[37] We found that the data were well described by a sigmoidal function:

$$[water](z_{Phos}) = 1.0 * 10^{-6} + \left\{\frac{31.563}{1+exp^{\frac{3.7015-z}{-1.8604}}}\right\} \quad (3).$$

where *z* is the position relative to the phosphate plane of the bilayer.

The function can alternatively be adapted to any reference plane in the membrane by refitting $[water](z)$ with an adjusted z-position axis. For example, using the lipid carbonyl plane as the reference point (z = 0), the function changes to:

$$[water](z_{Carb}) = 1.0 * 10^{-6} + \left\{\frac{31.563}{1+exp^{\frac{-0.2985-z}{-1.8604}}}\right\} \quad (4).$$

Fitting and Derivation of $\sigma_{NP}([water])$, $[water](z)$, $\sigma_{NP}(z)$, and $\Delta G^{\circ}_{sc}(z)$. The relationship between σ_{NP} and [water] was determined by weighted linear regression (errors derived from error of fitting σ_{NP} using Equation 1) using Igor Pro, resulting in the following relationship:

$$\sigma_{NP}([water]) = 0.87 * [water] - 25.6 \quad (5).$$

The $[water](z)$ (Equation 3) function allows for $\sigma_{NP}(z)$ to be derived:

$$\sigma_{NP}(z) = 0.87 * [water](z) - 25.6 \quad (6).$$

This function allows for $\sigma_{NP}(z)$ to be calculated at any z-position inside the phosphate plane of the bilayer. To derive $\Delta G_{sc}^{\circ}(z)$ in kcal mol^{-1} for any nonpolar side chain from $\sigma_{NP}(z)$ the following relationship is used:

$$\Delta G_{sc}^{o}(z) = \sigma_{NP}(z) * ASA/1000 \quad (7).$$

where ASA is the solvent accessible surface area for that side chain.

Simulations and Analysis of Neat POPC Bilayer. A neat POPC bilayer (75 lipids per leaflet) was created using CHARMM-GUI with 0.2M NaCl at 37 ºC. The CHARMM36m forcefield was used, and the simulation was equilibrated using CHARMM-GUI provided equilibration steps followed by 15 ns of additional equilibration time. The following 50 ns of simulation were used to measure the concentration of water molecules as a function of z-position in the bilayer and the z-position of the phosphate plane using the density profile plug in tool for VMD.[37] The relationship between z-position and [water] was fit to a sigmoid function as described above. The $\sigma_{NP}(z)$ for POPC was also determined as described above.

Results

Experimental determination of a z-dependent nonpolar solvation function. In this work we define a function, $\sigma_{NP}(z)$, that relates the value of the σ_{NP} to the z-position along the bilayer normal. This function is fundamentally derived from experimental protein folding titrations described below. Using a previously validated host-guest approach, we measured the thermodynamic stabilities of a series of host (Ala) and guest (nonpolar side chains isoleucine (Ile), methionine (Met), and valine (Val)) variants of the well-characterized membrane protein scaffold, E. *coli* outer membrane phospholipase A1 (OmpLA).[15,16,30,31] Host/guest sites were chosen at varying positions across the bilayer to allow side chains to experience a range of local chemical compositions created by the amphipathic phospholipids in the membrane. In total, five lipid-facing sites located on well-defined transmembrane β-strands were used in this study (residues 120, 164, 210, 212, and 214) and are shown in Figure 1A. Figure 1B shows the water concentration gradient across the experimental DLPC bilayer and highlights the 60-fold change in water concentration accessed by these sites.

The thermodynamic stabilities ($\Delta G_{w,l}^{\circ}$) of all variants were measured using intrinsic tryptophan fluorescence-based chemical denaturation titrations (Figure S4).[15,16,18,31] Figure S5 shows that each variant was enzymatically active, and Figure S6 shows that the folded and unfolded tryptophan fluorescence spectra for each variant overlay. Together these data indicate that single mutations do not affect the tertiary or quaternary structure of either the folded or unfolded states of OmpLA. We have previously shown that single site variants of OmpLA fold in a path-independent manner into large unilamellar DLPC vesicles, allowing for equilibrium thermodynamic parameters to be extracted by fitting the titration data to a three-state linear extrapolation model ($\Delta G_{w,l}^{\circ}$ values listed in Table S2).[15,16,30]

The thermodynamic contributions of the guest side chain to folding were isolated by taking the difference in the $\Delta G_{w,l}^{\circ}$ values of the host and guest variants at a given site on OmpLA as shown in the thermodynamic cycle in Figure S7. We interpret the resulting $\Delta\Delta G_{w,l}^{\circ}$ at each site to represent the water-to-bilayer transfer free energy change for each guest side chain relative to alanine at a particular site (Table S2). We find that our experimentally measured $\Delta\Delta G_{w,l}^{\circ}$ are remarkably similar to analogous, computationally estimated $\Delta\Delta G_{w,l}^{\circ}$ that were previously published by Liang and coworkers, though both Ile and Val deviated significantly at site 212 (Figure S8).[38]

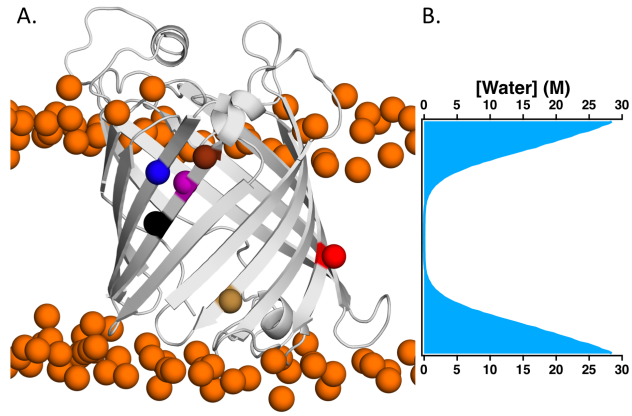


Figure 1. Host-sites on OmpLA exist in different water concentrations in the bilayer. (A) A snapshot of a molecular dynamics simulation of WT OmpLA in a DLPC bilayer is shown. Phosphate atoms of the DLPC bilayer are colored orange and six host sites on OmpLA used in this study are shown as colored spheres (black:210, yellow:164, red:120, purple:212, blue:223, brown:214). **(B)** The gradient of water concentrations inside the phosphate plane of the bilayer is plotted as a function of position in the bilayer. These values were obtained from previously published simulations of neat DLPC bilayers.[15] This water gradient is aligned with the structure of OmpLA at the two phosphate planes. The sites on OmpLA were chosen because they sample a wide range of z-positions positions and chemical compositions.

The nonpolar solvation energy is not constant along the bilayer interface. To determine the σ_{NP} for each site on OmpLA we used the $\Delta\Delta G_{w,l}^{\circ}$ values for each nonpolar side chain. This dataset included previously published $\Delta\Delta G_{w,l}^{\circ}$ for leucine and phenylalanine.[15,16] Figure 2 shows nonpolar $\Delta\Delta G_{w,l}^{\circ}$ values plotted as a function of the nonpolar surface area of each side chain relative to alanine (ΔASA) for each site. The σ_{NP} at each site on OmpLA is the slope of the linear fit of the correlation between nonpolar $\Delta\Delta G_{w,l}^{\circ}$ and ΔASA (Equations 1 and 2, fit parameters in Table S3). The σ_{NP} values vary nearly two-fold from site to site with a range that spans from -27.3 to -16.1 cal mol^{-1} Å^{-2}.

This experimental setup requires a reference host side chain whose energetic contributions must be removed to obtain a reference-free scale. As alanine is the amino acid side chain in the host protein, we obtain a reference-free scale by removing the intrinsic contributions of Ala to $\Delta\Delta G_{w,l}^{\circ}$ at each site using the site-specific σ_{NP} values. The reference-free ΔG_{sc}° for alanine at each host site was determined by multiplying the corresponding σ_{NP} by the ASA of the alanine side chain.[16] The reference-free ΔG_{sc}° for the other nonpolar side chains were calculated by adding the ΔG_{sc}° for alanine to the $\Delta\Delta G_{w,l}^{\circ}$ at each site. The ΔG_{sc}° for each nonpolar side chain, including alanine, are listed in Table S4.

The nonpolar solvation parameter is linearly correlated with the amount of water in the bilayer. We next addressed

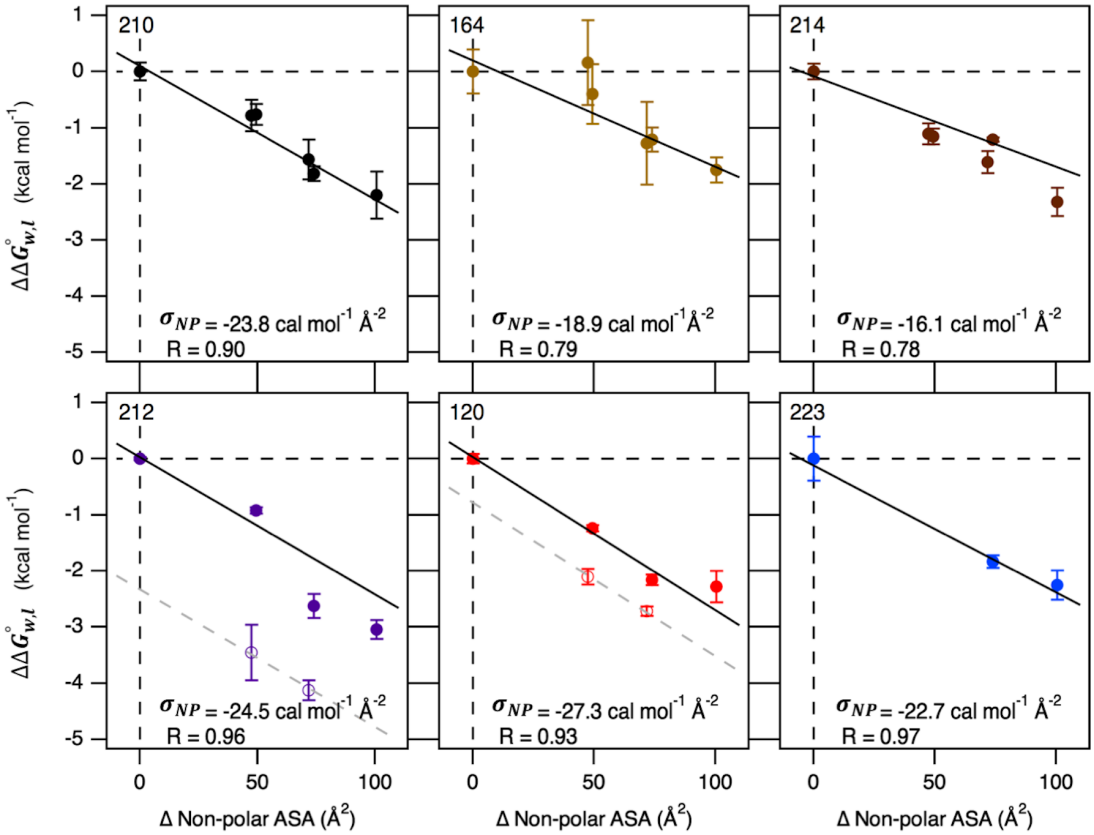


Figure 2. Nonpolar solvation parameter is not constant throughout the bilayer. The nonpolar solvation parameter, σ_{NP}, was determined for each site on OmpLA by calculating the slope of the linear relationship between $\Delta\Delta G^{\circ}_{w,l}$ for each nonpolar side chain at that site with the change in buried surface area compared to alanine in a Gly-X-Gly peptide (Equation 1).[16] Linear fits were weighted by the error in $\Delta\Delta G^{\circ}_{w,l}$ (standard deviations found in Table S2) for each site, with the slope of the line (σ_{NP}) and the Pearson correlation coefficient reported for each fit shown at the bottom of each panel. For sites 212 and 120, beta-branched side chains isoleucine and valine were omitted from σ_{NP} determination as they had anomalously greater $\Delta\Delta G^{\circ}_{w,l}$ compared to the other nonpolar side chains. The σ_{NP} determined from non-beta-branched side chains (solid line in 212 and 120 panels) fits the beta-branched side chains (dotted line), indicating that the increase in energy is derived from favorable local interactions that are restricted to beta-branched residues. By taking the difference in the y-intercept for the two lines, we estimate the energy gained due to local interactions for beta-branched residues at these sites to be 2.38 kcal mol^{-1} for site 212 and 0.80 kcal mol^{-1} for site 120. The σ_{NP} for site 223 was determined only using previously collected data.[16]

the question of how the local chemical composition of the bilayer modulates the σ_{NP} by deriving the relationship between the σ_{NP} and z-position along the bilayer normal. This results in the function $\sigma_{NP}(z)$. To accomplish this, we calculated the z-position of the side chain Cα of each OmpLA variant from all-atom molecular dynamics simulations conducted in the experimental lipid bilayer. Previous work showed that introduction of single side chains has no discernable effect on the z-position or tilt of OmpLA in the bilayer.[29] Figure S1 shows that each simulation was equilibrated after 50ns, and analyses were performed on the final 100 ns of each run. Instead of using the bilayer center as the reference point, ($z = 0$), we calculated the position of each Cα relative to the average position of the phosphate plane of the bilayer.

We chose the phosphate plane as our reference point for z because it can be more easily applied to bilayers with different lipid compositions and acyl chain lengths. The carbonyl plane is an alternative reference coordinate employed in the literature and could similarly be used. Figure S2 shows that variant Cα positions are generally well described by Gaussian distributions and that our sites capture the majority of the region between the bilayer center and phosphate plane. Table S1 lists the mean z-position and standard deviations for each variant Cα, which were averaged to calculate the z-position for the σ_{NP} at each site. We found that the bilayer z-position dependence of water concentration in our experimental DLPC bilayer was described by a sigmoidal function (Equation 3, Figure S3).[15] Using this correlation we calculates the corresponding local water concentration for each σ_{NP} (Table S5).

Figure 3A shows that the σ_{NP} is linearly correlated with the water concentration in the bilayer (Equation 5, $R^2 = 0.84$). The ~5,000-fold change in water concentration illustrates the large range of solvent conditions accessed in our experiments and necessitates presentation using a logarithmic abscissa. This novel functional relationship we derived from our measurements using a native protein and a phospholipid bilayer confirms our hypothesis that the water in a bilayer is a reporter on the hydrophobicity of the membrane and that the local hydrophobicity modulates membrane protein stabilities through the σ_{NP} directly.

It is also instructive to consider the σ_{NP} using bilayer z-coordinates. Figure 3B shows the bilayer z-position dependence of σ_{NP} in which the line of best fit in Figure 3A is transferred to units of distance from the phosphate plane of the

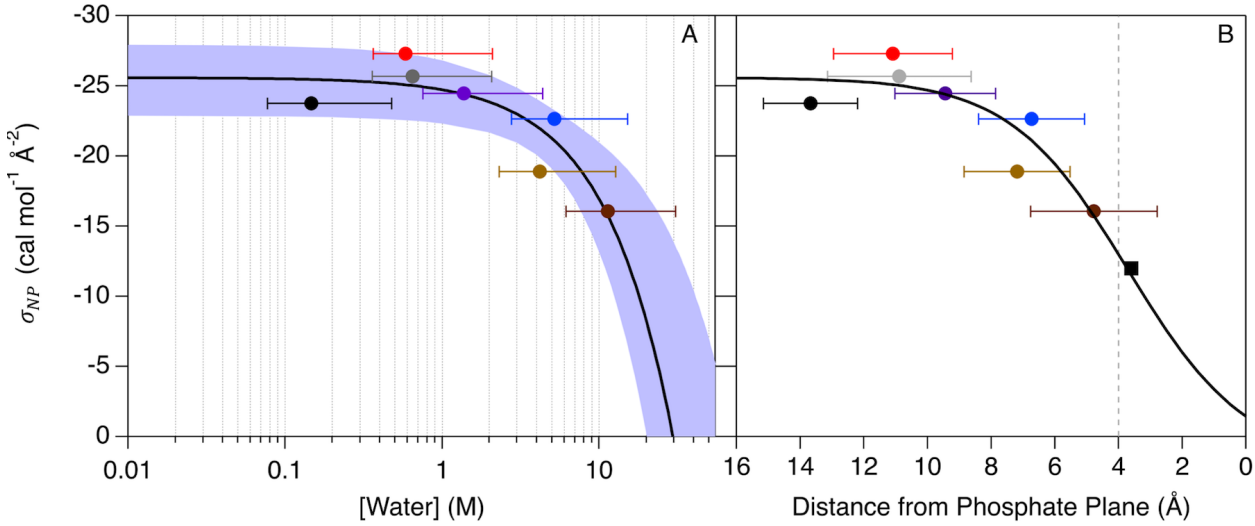


Figure 3. Nonpolar solvation parameter is linearly correlated with water concentration in the bilayer. The water-to-bilayer σ_{NP} calculated at each host site in OmpLA (colored as in Figure 1) and PagP (gray point) plotted as a function of water concentration (Table S4) (**A**) and distance from the phosphate plane (**B**).[18] The σ_{NP} are colored according to the scheme in Figure 1, with the σ_{NP} derived from PagP position 111 colored gray. The position of σ_{NP} for a given site is the average of the positions of each side chain at that site and the error bars reflect the standard deviation (Table S1, NP row). σ_{NP} and water concentration are linearly correlated with the equation shown in the bottom right hand corner of Panel **A** (black line: $R^2 = 0.84$; 95% confidence interval shaded in light blue). Using the derived relationship between bilayer position and water concentration shown in Figure S3, the direct relationship between σ_{NP} and bilayer position can be determined (Panel **B**). We were also able to assign a previously measured "water-to-interface" solvation parameter (-12 cal mol^{-1} Å^{-2}) to an exact position in the bilayer (3.75 Å from the phosphate plane) using this function (shown as a black square in Panel **B**) and find that it is reporting on the position of the lipid carbonyl plane in the bilayer (dashed vertical gray line).[17] The position-dependent function describing σ_{NP} allows for the ΔG°_{sc} of any nonpolar side chain to be determined for any position of the bilayer inside the phosphate plane (shown in Figure 4).

membrane. Using the resulting function, $\sigma_{NP}(z)$ (Equation 6), the σ_{NP} can easily and accurately be determined at any position in the bilayer. $\sigma_{NP}(z)$ can also be applied to any membrane system with a known [*water*] gradient along the bilayer normal. Figure S9 shows the [water](z) and $\sigma_{NP}(z)$ functions for a neat POPC bilayer, which is almost indistinguishable from DLPC indicating $\sigma_{NP}(z)$.

The $\sigma_{NP}(z)$ function plateaus to a value of -25.6 cal mol^{-1} Å^2 in the anhydrous core of the bilayer, which agrees with the σ_{NP} determined for this region of the bilayer in previous studies.[13,16,18,19] We find that the magnitude of the $\sigma_{NP}(z)$ decreases in the membrane interface to a predicted σ_{NP} equal to -1.4 cal mol^{-1} Å^{-2} at the phosphate plane, which is much less favorable than previous estimates of σ_{NP} for the interface (-12 cal mol^{-1} Å^{-2}).[17] Using $\sigma_{NP}(z)$, we calculate that this previously measured σ_{NP} for the water-to-interface transition corresponds to a z-position approximately 3.75 Å inside the phosphate plane. This *z*-position corresponds to the approximate position of the lipid carbonyl plane, which is a commonly used reference as it is approximately the boundary between hydrocarbon and interfacial regions of the bilayer.

Bilayer *z*-position dependence of reference-free ΔG°_{sc}. The large change in the σ_{NP} across the bilayer interface significantly impacts the contribution of nonpolar side chains to the overall stability of membrane proteins in a position-dependent manner. This can be visualized by using the $\sigma_{NP}(z)$ to calculate a reference-free, nonpolar ΔG°_{sc} as a function of *z*-position in the bilayer (Equation 7). Figure 4 shows the $\Delta G^{\circ}_{sc}(z)$ for each nonpolar side chain (black lines) and the ΔG°_{sc} for each site on OmpLA determined above (points). Side chains located in the bilayer interface can be several kcal mol^{-1} less favorable than side chains in the dehydrated core of the bilayer. Additionally, our findings indicate that interfacial side chains separated by a single turn of an alpha helix can have dramatically different contributions to the overall stability of a membrane protein.

The functional free energy dependence for the alanine side chain, $\Delta G^{\circ}_{Ala}(z)$, is particularly useful because the experimental setup necessitates that all side chain free energy perturbations be measured relative to Ala. Upon subtraction of the reference (Ala) energy, the *z*-position-dependence for any side chain can be place on a reference-free scale. Accordingly, the data herein allows conversion of previously measured $\Delta\Delta G^{\circ}_{w,l}$ values for arginine, tryptophan, and tyrosine to reference-free $\Delta G^{\circ}_{Ala}(z)$ scales.[15,16] Following the protocol above, a combination of new (Arg) and published (Tyr and Trp) all-atom molecular dynamics simulations were first used to determine the *z*-position of the Cα atom for each variant.[15] This ΔG°_{Ala} value at a particular *z*-position was calculated using the alanine functional free energy dependence ($\Delta G^{\circ}_{Ala}(z)$), and this value was added to the corresponding experimental value for the guest side chain.

Figures S9, S10 and Table S6 shows these adjustments applied to the arginine, tryptophan, and tyrosine side chains.[15,16] As expected, these data show that there is a larger energetic penalty for the placement of the arginine Cα within the bilayer ~13 Å from the phosphate plane (+2.5 kcal mol^{-1}). This observation is in contrast to a small, favorable energetic contribution when the side chain is located in the bilayer interfacial region (-0.5 kcal mol^{-1}). The $\Delta G^{\circ}_{Arg}(z)$ is remarkably similar to the reported translocon-mediated profile from Arg insertion in a α-helical TMD.[14] This similarity most like is derived from the fact that Arg can snorkel to the interface at any position in a TMD, which would change its thermodynamic "end state"

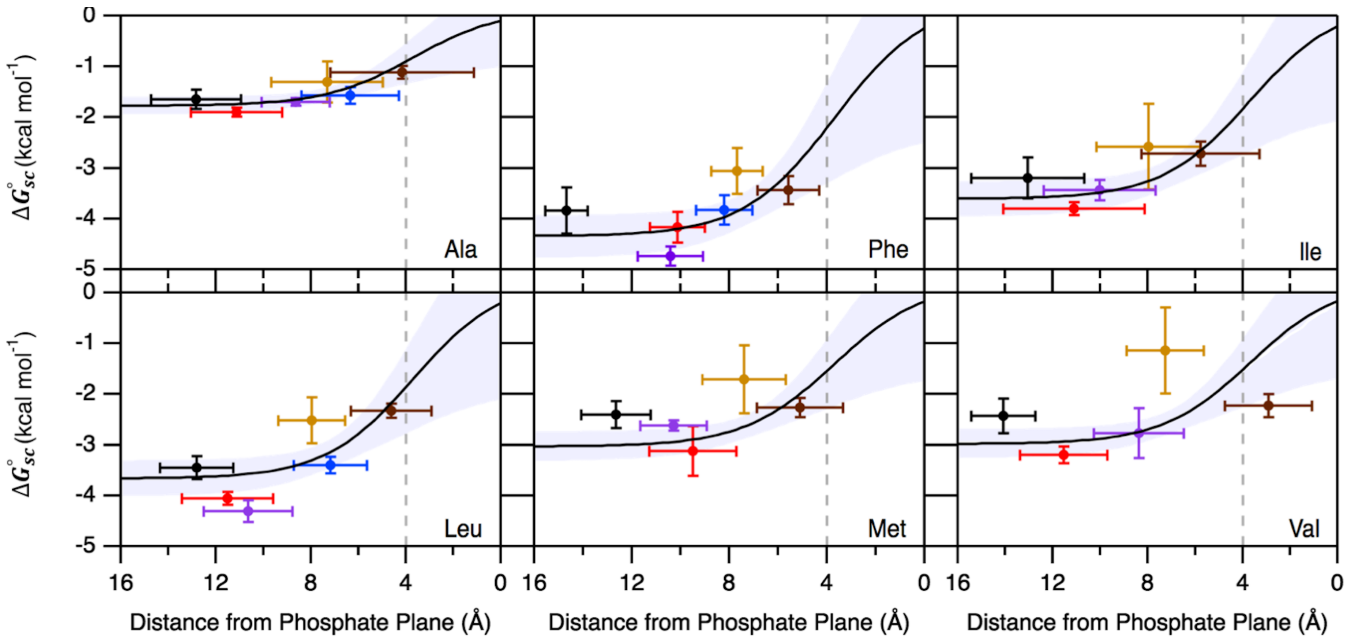


Figure 4. Bilayer position dependence of nonpolar ΔG°_{sc}. Reference-free side chain transfer free energies are plotted as a function of the average distance from the bilayer phosphate plane, as determined from molecular dynamics trajectories. ΔG°_{sc} for each host site in OmpLA are colored as in Figure 1. Error bars for both ΔG°_{sc} and bilayer position represent the standard deviations. The solid black line in each panel represents the simulated $\Delta G^{\circ}_{sc}(z)$ profile for each nonpolar side chain. This function is derived from $\sigma_{NP}(z)$ multiplied by the nonpolar surface area of each side chain (Equation 7). Vertical dotted lines represent the position of the lipid carbonyl groups.

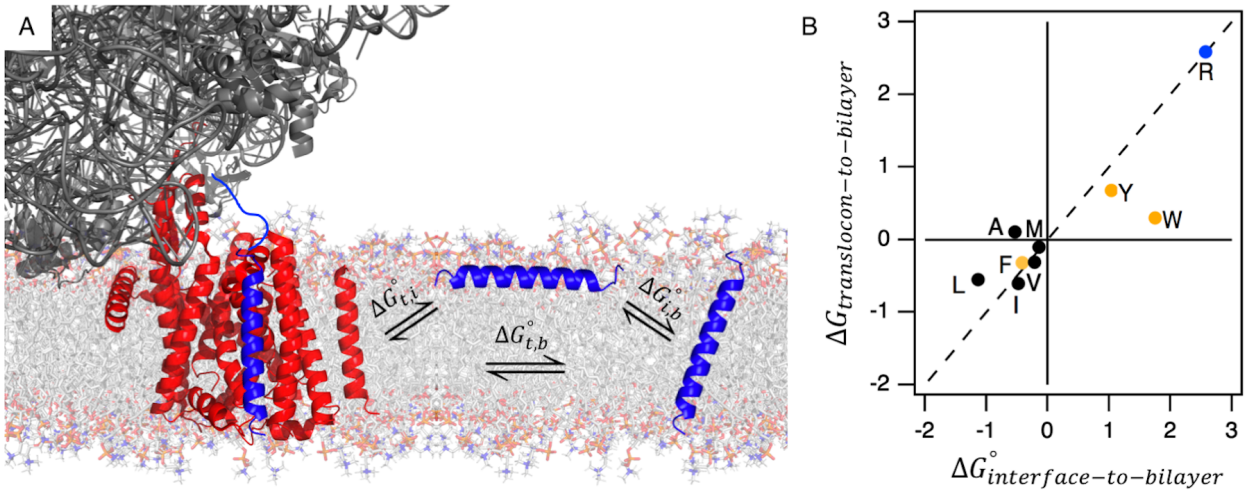


Figure 5. The translocon energetically mimics the bilayer interface. (A) A cartoon schematic of the co-translational insertion of a helix (blue) via the translocon (ribosome colored dark gray, translocon colored red; PDB code: 5GAE) is shown on the left. The helix can either partition to an interfacial conformation, which is energetically described by $\Delta G^{\circ}_{t,i}$, or to a transmembrane conformation, described by $\Delta G^{\circ}_{t,b}$ (t = translocon, i = interfacial, b = transmembrane). The interfacial-to-transmembrane conformation, which can occur in the absence of the translocon, is energetically described by $\Delta G^{\circ}_{i,b}$. **(B)** The values for $\Delta G^{\circ}_{t,b}$ and $\Delta G^{\circ}_{i,b}$ are plotted for nine side chains, with the dotted line representing energetic equivalence between the two equilibria (black = nonpolar, gold = aromatic, blue = ionizable).[41] For all side chains, except Trp, $\Delta G^{\circ}_{t,b} \approx \Delta G^{\circ}_{i,b}$, indicating that the translocon energetically mimics the interface ($\Delta G^{\circ}_{t,i} \approx 0$). The deviation for Trp is hypothesized to be due to the energetic preference of Trp to exist in the bilayer interface [15,42].

to be the interface in all cases. The energetic trends for the aromatic side chains $\Delta G^{\circ}_{Trp}(z)$ and $\Delta G^{\circ}_{Tyr}(z)$ reveal an opposite but also expected behavior: both side chains show an energetic preference for the bilayer interface as compared to the nonpolar central region of the bilayer.

The translocon energetically mimics the bilayer interface to facilitate TMD insertion into the membrane. With the reference-free, water-to-bilayer $\Delta G^{\circ}_{sc}(z)$ for nine amino acids determined as a function of z-position in the bilayer, we next asked whether our findings could be used to better

understand membrane protein folding by the Sec translocon machinery. The translocon acts as a protein channel in the bilayer that allows α-helical TMDs to partition into the bilayer while protecting hydrophilic loops and turns from exposure to the hydrophobic membrane.[10,39] All-atom molecular dynamics simulations have identified that the chemical properties of water in the pore of the translocon mimic the properties of water in the interfacial region of the membrane.[40]

Thus, a hypothesis for translocon function is that it mimics the thermodynamics of an interface-to-bilayer transition, not a water-to-bilayer transition.[10] Figure 5A shows the relevant free energy reactions that can be written for these processes. The water-to-bilayer-center nonpolar ΔG_{sc}° values reported here are much greater in magnitude than the $\Delta G_{app\ (t,b)}$ for the translocon-to-bilayer transition.[14,41] Because we have measured the water-to-bilayer ΔG_{sc}° across the bilayer normal, we can calculate the interface-to-center transition by taking the difference between ΔG_{sc}° at the center of the bilayer (A210) and the ΔG_{sc}° at the most interfacial site (Y214) (Figure 5A). Figure 5B shows that our interface-to-bilayer $\Delta\Delta G_{sc}^{\circ}$, or $\Delta G_{i,b}^{\circ}$, agree remarkably well with the $\Delta G_{app\ (t,b)}$.[41] Additionally, the σ_{NP} for the interface-to-center transition ($\sigma_{NP,int-cen} = \sigma_{NP,214} - \sigma_{NP,210}$) is equivalent to the σ_{NP} calculated for translocon-mediated folding (7.7 and 6-10 cal mol^{-1} Å^2, respectively).[26] The remarkable energetic equivalence between the interface-to-center and translocon-mediated transfer free energies provides thermodynamic backing to the hypothesis that the translocon aids TMD folding by mimicking properties of the interface.

Discussion

The thermodynamic stability of a membrane protein regulates its structure, conformational landscape, and ability to perform its biological function.[7,11] A key determinant of this stability is the ΔG_{sc}° of each side chain from water into the membrane. Historically, experiments measuring side chain transfer from water to organic solvents such as cyclohexane or octanol were used to approximate ΔG_{sc}°.[19,22] While these experiments accurately describe the water-to-bilayer-center transition, they failed to capture thermodynamic properties resulting from the dramatic chemical heterogeneity of the interfacial region of the bilayer.

In this study we have measured the $\Delta\Delta G_{w,l}^{\circ}$ for Ile, Met, and Val as a function of z-position in the bilayer. Using these $\Delta\Delta G_{w,l}^{\circ}$, along with previously measured $\Delta\Delta G_{w,l}^{\circ}$ for Leu and Phe we were able to determine a novel functional form for the σ_{NP} across the bilayer normal, $\sigma_{NP}(z)$.[15,16] We discovered a direct correlation between the σ_{NP} - and thus the ΔG_{sc}° for nonpolar side chains - with the local water concentration in the bilayer. The $\sigma_{NP}(z)$ is a convenient numerical expression for assessing the magnitude of the hydrophobic effect, as it relates nonpolar surface area to the energy gained from transferring the surface area from bulk water to a dehydrated environment.

Importantly, the dependence of nonpolar ΔG_{sc}° on local water concentration differs starkly from the water-dependence of backbone hydrogen bond (bbHB) formation in membrane proteins. Although it was long expected that bbHB energies would also be strongly influenced by the steep water gradient intrinsic to bilayer interface, we recently observed bbHB energies to be constant throughout this region.[42]

The experimental data here offer an opportunity to consider how the translocon insertion process relates to the underlying physical chemistry of the reaction. Our data provides evidence that translocon-mediated membrane protein folding energetically mimics the interface-to-bilayer transition for ΔG_{sc}°. We find that $\Delta G_{i,b}^{\circ}$ is approximately equal to $\Delta G_{app\ (t,b)}$ for eight of the nine side chains investigated, with Trp being the only outlier; we speculate that this one outlier is most likely due to differential aromatic interactions in the OmpLA and translocon experimental setups.[15,41] This finding provides thermodynamic support for a recent model for translocon function in which transmembrane α-helices sample both interfacial and translocon associated conformations before they insert into the membrane.[10] Our findings indicate that these two states would be essentially energetically equivalent, indicating the energy barrier between these two conformations may be nonexistent with switching between the two states determined by the kinetics for a given TMD sequence.

Our results should improve the speed and accuracy of membrane protein structure prediction and design algorithms.[43–45] As membrane proteins account for over half of all therapeutic drug targets,[46] having accurate force fields to describe membrane protein energetics is essential. This simple, linear function $\sigma_{NP}(z)$ could be adapted to implicit membrane models or could alternatively be applied explicitly in systems such that a local value of σ_{NP} to be calculated over water concentrations ranging from millimolar to 30 molar. Subsequently, better estimates of the energetic contribution of any lipid-facing, nonpolar side chain to the stability of membrane protein.

By relating the σ_{NP} to the chemical properties of the bilayer, the $\sigma_{NP}(z)$ can be customized to any bilayer lipid composition by calculating the water concentration gradient of that particular bilayer (illustrated in Figure S9 for a POPC bilayer). We find that the $\sigma_{NP}(z)$ for both POPC and DLPC bilayers are almost indistinguishable, indicating that lipid acyl tail length and saturation have little effect on the shape and magnitude of $\sigma_{NP}(z)$ across the bilayer interface. Further work is needed to assess the impact of headgroups and other macromolecules found in bilayers such as cholesterol on the shape of $\sigma_{NP}(z)$. In addition, one caveat of $\sigma_{NP}(z)$ is that it relies on additive forcefields, which may underestimate the local water concentration in the bilayer interface. As polarizable forcefields become more readily accessible the $\sigma_{NP}(z)$ could be refined to even more accurately quantify the energetics of nonpolar side chain burial in membranes.

The $\sigma_{NP}(z)$ and the nine $\Delta G_{sc}^{\circ}(z)$ functions derived here also could be applied to predicting the contribution of the binding energy of proteins to the bilayer interface ($\Delta G_{sc}^{\circ}(z)$ functions found in Table S7). By relating energy to *z*-position of side chains in the bilayer, binding energies can be tuned based on structural information derived from experiment or computation. The simplest application of the $\Delta G_{sc}^{\circ}(z)$ functions would apply to proteins that do not undergo large structural changes upon membrane interaction, as the binding energy would be dominated by lipid-side chain interactions. In cases of conformational changes, our values may find utility in providing baseline values for the partitioning of many membrane binding proteins (such as antimicrobial peptides) even when the global binding energy contains additional contributions from other sources, such as the energy of the coil-to-helix transition that occurs upon binding.[16] Given the

complexity of membrane protein folding, even for apparently simple peptides, the safest application of our values is in understanding the effect of a point mutation on bilayer-protein interactions.

Additionally, our findings delineate how the different transfer free energies reported here should be applied to various biological functions. The critical parameter for consideration is the endpoints for the reaction at hand. The $\Delta G_{app\ (t,b)}$ or $\Delta G^{\circ}_{i,b}$ can be used to approximate the contribution of a side chain to the folding of alpha-helical TMDs.[14,41] On the other hand, water-to-bilayer ΔG°_{sc} are more applicable to modeling both changes in stability of different conformers of a membrane protein, the transition between interfacial and transmembrane conformations of antimicrobial peptides, and potentially unfolding processes, such as extraction of transmembrane regions from the bilayer by AAA-ATPases.[47] Experimentally determining these thermodynamic parameters for membrane proteins is extremely challenging and accurate computational modeling will allow for greater understand of essential processes such as the gating of a channel or the cycling of a transporter.[48–50] At the very least, we anticipate that $\Delta G^{\circ}_{sc}(z)$ can be used to better understand the effects of mutations that may impair membrane protein folding such as those occurring in cystic fibrosis and Charcot-Marie-Tooth disease.[51]

Conclusion

In summary, we have determined the relationship between nonpolar side chain transfer free energies, solvent accessible surface area, and the local water concentration in the bilayer. Using this relationship, we calculate $\Delta G^{\circ}_{sc}(z)$ functions for nonpolar side chains and remove the alanine dependence of previously measured $\Delta\Delta G^{\circ}_{w,l}$ for Arg, Trp, and Tyr. Our work shows that the translocon energetically mimics the bilayer interface for side chain transfer into the membrane. Together, these findings increase our understanding of the driving forces of protein stability in membranes by essentially quantifying the hydrophobic effect along the bilayer normal and should increase the accuracy of computational workflows to identify, design, and understand the energy landscapes of membrane proteins.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website.

The Supporting Information contains representative chemical denaturation titrations, spectroscopic and enzymatic confirmation of OmpLA variant folding, thermodynamic cycle, molecular dynamics simulation analysis, reference-free transfer free energy profiles for Arg, Trp, and Tyr, stabilities of OmpLA variants, side chain transfer free energy values (PDF).

AUTHOR INFORMATION

Corresponding Author

*Karen G. Fleming: Karen.Fleming@jhu.edu

Author Contributions

D.C.M. and K.G.F. designed research, D.C.M. performed experiments and analyzed data, D.C.M. and K.G.F. wrote the paper. All authors have given approval to the final version of the manuscript.

Funding Sources

This work was funded by National Institutes of Health grants R01 GM079440 and T32 GM008403.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Sarah McDonald and Margo Goodall for assistance with cloning the variants used in this study, Mark Kreutzberger for growth and purification of some OmpLA variants used in this study, and Dr. Sarah McDonald and Dr. Patrick Fleming for experimental and computational guidance.

ABBREVIATIONS

OmpLA, E. *coli* outer membrane protein phosphatase A1; PagP, E. *coli* lipid A palymitoyltransferase PagP; DLPC, 1,2 didodecanoyl-sn-glycero-3-phosphocholine; LUV, large unilamellar vesicle; TMD, transmembrane domain; σ_{NP}, nonpolar solvation parameter; $\Delta G^{\circ}_{w,l}$, OmpLA variant free energy of folding; $\Delta\Delta G^{\circ}_{w,l}$, alanine-referenced, water-to-bilayer side chain transfer free energy; ΔG°_{sc}, reference-free, water-to-bilayer side chain transfer free energy; $\Delta G_{app\ (t,b)}$, translocon-to-bilayer side chain transfer free energy; $\Delta G^{\circ}_{i,b}$, interface-to-bilayer side chain transfer free energy; $\Delta G_{t,i}$, translocon-to-interface side chain transfer free energy.

REFERENCES

(1) Bloom, J. D.; Labthavikul, S. T.; Otey, C. R.; Arnold, F. H. Protein Stability Promotes Evolvability. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103* (15), 5869–5874. https://doi.org/10.1073/pnas.0510098103.

(2) Osadchy, M.; Kolodny, R. Maps of Protein Structure Space Reveal a Fundamental Relationship between Protein Structure and Function. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108* (30), 12301–12306. https://doi.org/10.1073/pnas.1102727108.

(3) Reis, R.; Moraes, I. Structural Biology and Structure–Function Relationships of Membrane Proteins. *Biochemical Society Transactions*. Portland Press Ltd 2018, pp 47–61. https://doi.org/10.1042/BST20180269.

(4) Freire, E. The Thermodynamic Linkage between Protein Structure, Stability, and Function. *Methods in molecular biology (Clifton, N.J.)*. Methods Mol Biol 2001, pp 37–68. https://doi.org/10.1385/1-59259-193-0:037.

(5) Goldenzweig, A.; Fleishman, S. J. Principles of Protein Stability and Their Application in Computational Design. *Annual Review of Biochemistry*. Annual Reviews Inc. June 20, 2018, pp 105–129. https://doi.org/10.1146/annurev-biochem-062917-012102.

(6) Almén, M. S.; Nordström, K. J. V.; Fredriksson, R.; Schiöth, H. B. Mapping the Human Membrane Proteome: A Majority of the Human Membrane Proteins Can Be Classified According to Function and Evolutionary Origin. *BMC Biol.* **2009**, *7*, 50. https://doi.org/10.1186/1741-7007-7-50.

(7) Hong, H. Toward Understanding Driving Forces in Membrane Protein Folding. *Archives of Biochemistry and Biophysics*. Academic Press Inc. December 10, 2014, pp 297–313. https://doi.org/10.1016/j.abb.2014.07.031.

(8) Tan, A. E.; Burgess, N. K.; DeAndrad, D. S.; Marold, J. D.; Fleming, K. G. Self-Association of Unfolded Outer Membrane Proteins. *Macromol. Biosci.* **2010**, *10* (7), 763–767. https://doi.org/10.1002/mabi.200900479.

(9) Parton, D. L.; Klingelhoefer, J. W.; Sansom, M. S. P.

Aggregation of Model Membrane Proteins, Modulated by Hydrophobic Mismatch, Membrane Curvature, and Protein Class. *Biophys. J.* **2011**, *101* (3), 691–699. https://doi.org/10.1016/j.bpj.2011.06.048.

(10) Cymer, F.; Von Heijne, G.; White, S. H. Mechanisms of Integral Membrane Protein Insertion and Folding. *Journal of Molecular Biology*. Academic Press March 13, 2015, pp 999–1022. https://doi.org/10.1016/j.jmb.2014.09.014.

(11) White, S. H.; Wimley, W. C. Membrane Protein Folding and Stability: Physical Principles. *Ann. Rev. Biophys. Biomol. Str.* **1999**, *28*, 319–365. https://doi.org/10.1146/annurev.biophys.28.1.319.

(12) Popot, J. L.; Engelman, D. M. Membrane Protein Folding and Oligomerization: The Two-Stage Model. *Biochemistry* **1990**, *29* (17), 4031–4037. https://doi.org/10.1021/bi00469a001.

(13) Elazar, A.; Weinstein, J.; Biran, I.; Fridman, Y.; Bibi, E.; Fleishman, S. J. Mutational Scanning Reveals the Determinants of Protein Insertion and Association Energetics in the Plasma Membrane. *Elife* **2016**, *5*. https://doi.org/10.7554/eLife.12125.

(14) Hessa, T.; Meindl-Beinker, N. M.; Bernsel, A.; Kim, H.; Sato, Y.; Lerch-Bader, M.; Nilsson, I.; White, S. H.; Von Heijne, G. Molecular Code for Transmembrane-Helix Recognition by the Sec61 Translocon. *Nature* **2007**, *450* (7172), 1026–1030. https://doi.org/10.1038/nature06387.

(15) McDonald, S. K.; Fleming, K. G. Aromatic Side Chain Water-to-Lipid Transfer Free Energies Show a Depth Dependence across the Membrane Normal. *J. Am. Chem. Soc.* **2016**, *138* (25), 7946–7950. https://doi.org/10.1021/jacs.6b03460.

(16) Moon, C. P.; Fleming, K. G. Side-Chain Hydrophobicity Scale Derived from Transmembrane Protein Folding into Lipid Bilayers. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108* (25), 10174–10177. https://doi.org/10.1073/pnas.1103979108.

(17) Wimley, W. C.; White, S. H. Experimentally Determined Hydrophobicity Scale for Proteins at Membrane Interfaces. *Nat. Struct. Mol. Biol.* **1996**, *3* (10), 842–848.

(18) Marx, D. C.; Fleming, K. G. Influence of Protein Scaffold on Side-Chain Transfer Free Energies. *Biophys. J.* **2017**, *113* (3), 597–604. https://doi.org/10.1016/j.bpj.2017.06.032.

(19) Wimley, W. C.; Creamer, T. P.; White, S. H. Solvation Energies of Amino Acid Side Chains and Backbone in a Family of Host-Guest Pentapeptides. *Biochemistry* **1996**, *35* (16), 5109–5124. https://doi.org/10.1021/bi9600153.

(20) Ulmschneider, J. P.; Smith, J. C.; White, S. H.; Ulmschneider, M. B. The Importance of the Membrane Interface as the Reference State for Membrane Protein Stability. *Biochim. Biophys. Acta - Biomembr.* **2018**, *1860* (12), 2539–2548. https://doi.org/10.1016/j.bbamem.2018.09.012.

(21) Damodaran, S.; Song, K. B. The Role of Solvent Polarity in the Free Energy of Transfer of Amino Acid Side Chains from Water to Organic Solvents. *J. Biol. Chem.* **1986**, *261* (16), 7220–7222.

(22) Radzicka, A.; Wolfenden, R. Comparing the Polarities of the Amino Acids: Side-Chain Distribution Coefficients between the Vapor Phase, Cyclohexane, 1-Octanol, and Neutral Aqueous Solution. *Biochemistry* **1988**, *27* (5), 1664–1670. https://doi.org/10.1021/bi00405a042.

(23) Wiener, M. C.; White, S. H. Structure of a Fluid Dioleoylphosphatidylcholine Bilayer Determined by Joint Refinement of X-Ray and Neutron Diffraction Data. II. Distribution and Packing of Terminal Methyl Groups. *Biophys. J.* **1992**, *61* (2), 428–433. https://doi.org/10.1016/S0006-3495(92)81848-9.

(24) Wiener, M. C.; White, S. H. Structure of a Fluid Dioleoylphosphatidylcholine Bilayer Determined by Joint Refinement of X-Ray and Neutron Diffraction Data. III. Complete Structure. *Biophys. J.* **1992**, *61* (2), 434–447. https://doi.org/10.1016/S0006-3495(92)81849-0.

(25) Wesson, L.; Eisenberg, D. Atomic Solvation Parameters Applied to Molecular Dynamics of Proteins in Solution. *Protein Sci.* **1992**, *1* (2), 227–235. https://doi.org/10.1002/pro.5560010204.

(26) Öjemalm, K.; Higuchi, T.; Jiang, Y.; Langel, Ü.; Nilsson, I.; White, S. H.; Suga, H.; Von Heijne, G. Apolar Surface Area Determines the Efficiency of Translocon-Mediated Membrane-Protein Integration into the Endoplasmic Reticulum. *Proc Natl Acad Sci U S A* **2011**, *108* (31), E359-364. https://doi.org/10.1073/pnas.1100120108/-/DCSupplemental.

(27) Chothia, C. Hydrophobic Bonding and Accessible Surface Area in Proteins. *Nature* **1974**, *248* (5446), 338–339. https://doi.org/10.1038/248338a0.

(28) Reynolds, J. A.; Gilbert, D. B.; Tanford, C. Empirical Correlation Between Hydrophobic Free Energy and Aqueous Cavity Surface Area. *Proc. Natl. Acad. Sci.* **1974**, *71* (8), 2925–2927. https://doi.org/10.1073/pnas.71.8.2925.

(29) Fleming, P. J.; Freites, J. A.; Moon, C. P.; Tobias, D. J.; Fleming, K. G. Outer Membrane Phospholipase A in Phospholipid Bilayers: A Model System for Concerted Computational and Experimental Investigations of Amino Acid Side Chain Partitioning into Lipid Bilayers. *Biochim. Biophys. Acta - Biomembr.* **2012**, *1818* (2), 126–134. https://doi.org/10.1016/j.bbamem.2011.07.016.

(30) Moon, C. P.; Kwon, S.; Fleming, K. G. Overcoming Hysteresis to Attain Reversible Equilibrium Folding for Outer Membrane Phospholipase A in Phospholipid Bilayers. *J. Mol. Biol.* **2011**, *413* (2), 484–494. https://doi.org/10.1016/j.jmb.2011.08.041.

(31) Moon, C. P.; Fleming, K. G. *Using Tryptophan Fluorescence to Measure the Stability of Membrane Proteins Folded in Liposomes*, 1st ed.; Elsevier Inc., 2011; Vol. 492. https://doi.org/10.1016/B978-0-12-381268-1.00018-5.

(32) Jo, S.; Kim, T.; Iyer, V. G.; Im, W. CHARMM-GUI: A Web-Based Graphical User Interface for CHARMM. *J. Comput. Chem.* **2008**, *29* (11), 1859–1865. https://doi.org/10.1002/jcc.20945.

(33) Lee, J.; Cheng, X.; Swails, J. M.; Yeom, M. S.; Eastman, P. K.; Lemkul, J. A.; Wei, S.; Buckner, J.; Jeong, J. C.; Qi, Y.; Jo, S.; Pande, V. S.; Case, D. A.; Brooks, C. L.; MacKerell, A. D.; Klauda, J. B.; Im, W. CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *J. Chem. Theory Comput.* **2016**, *12* (1), 405–413. https://doi.org/10.1021/acs.jctc.5b00935.

(34) Wu, E. L.; Cheng, X.; Jo, S.; Rui, H.; Song, K. C.; Dávila-Contreras, E. M.; Qi, Y.; Lee, J.; Monje-Galvan, V.; Venable, R. M.; Klauda, J. B.; Im, W. CHARMM-GUI *Membrane Builder* toward Realistic Biological Membrane Simulations. *J. Comput. Chem.* **2014**, *35* (27), 1997–2004. https://doi.org/10.1002/jcc.23702.

(35) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, *26* (16), 1781–1802. https://doi.org/Doi 10.1002/Jcc.20289.

(36) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14* (1), 33-.

(37) Giorgino, T. Computing 1-D Atomic Densities in Macromolecular Simulations: The Density Profile Tool for VMD. *Comput. Phys. Commun.* **2014**, *185* (1), 317–322. https://doi.org/10.1016/j.cpc.2013.08.022.

(38) Lin, M.; Gessmann, D.; Naveed, H.; Liang, J. Outer Membrane Protein Folding and Topology from a Computational Transfer Free Energy Scale. *J. Am. Chem. Soc.* **2016**, *138* (8), 2592–2601. https://doi.org/10.1021/jacs.5b10307.

(39) Niesen, M. J. M.; Müller-Lucks, A.; Hedman, R.; von Heijne, G.; Miller, T. F. Forces on Nascent Polypeptides during Membrane Insertion and Translocation via the Sec Translocon. *Biophys. J.* **2018**, *115* (10), 1885–1894. https://doi.org/10.1016/j.bpj.2018.10.002.

(40) Capponi, S.; Heyden, M.; Bondar, A.-N.; Tobias, D. J.; White, S. H. Anomalous Behavior of Water inside the SecY Translocon. *Proc. Natl. Acad. Sci.* **2015**, 201424483. https://doi.org/10.1073/pnas.1424483112.

(41) Hessa, T.; Kim, H.; Bihlmaier, K.; Lundin, C.; Boekel, J.; Andersson, H.; Nilsson, I.; White, S. H.; von Heijne, G. Recognition of Transmembrane Helices by the Endoplasmic Reticulum Translocon. *Nature* **2005**, *433* (7024), 377–381. https://doi.org/10.1038/nature03216.

(42) Lessen, H. J.; Majumdar, A.; Fleming, K. G. Backbone Hydrogen Bond Energies in Membrane Proteins Are

Insensitive to Large Changes in Local Water Concentration. *J. Am. Chem. Soc.* **2020**, *142* (13), 6227–6235. https://doi.org/10.1021/jacs.0c00290.

(43) Alford, R. F.; Fleming, P. J.; Fleming, K. G.; Gray, J. J. Protein Structure Prediction and Design in a Biologically Realistic Implicit Membrane. *Biophys. J.* **2020**, *118* (8), 2042–2055. https://doi.org/10.1016/j.bpj.2020.03.006.

(44) Alford, R. F.; Koehler Leman, J.; Weitzner, B. D.; Duran, A. M.; Tilley, D. C.; Elazar, A.; Gray, J. J. An Integrated Framework Advancing Membrane Protein Modeling and Design. *PLoS Comput. Biol.* **2015**, *11* (9). https://doi.org/10.1371/journal.pcbi.1004398.

(45) Weinstein, J. Y.; Elazar, A.; Fleishman, S. J. A Lipophilicity-Based Energy Function for Membrane-Protein Modelling and Design. *PLoS Comput. Biol.* **2019**, *15* (8). https://doi.org/10.1371/journal.pcbi.1007318.

(46) Overington, J. P.; Al-Lazikani, B.; Hopkins, A. L. How Many Drug Targets Are There? *Nat. Rev. Drug Discov.* **2006**, *5* (12), 993–996. https://doi.org/10.1038/nrd2199.

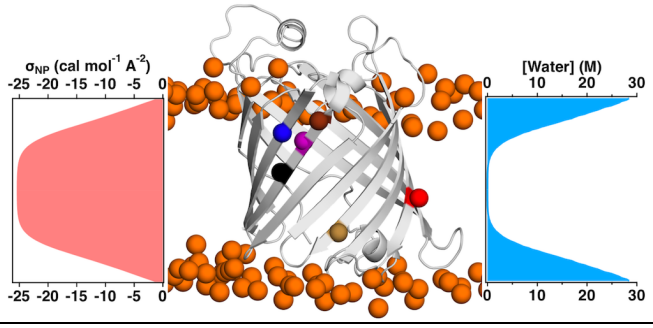
(47) Avci, D.; Lemberg, M. K. Clipping or Extracting: Two Ways to Membrane Protein Degradation. *Trends in Cell Biology*. Elsevier Ltd October 1, 2015, pp 611–622. https://doi.org/10.1016/j.tcb.2015.07.003.

(48) Sukharev, S. I.; Sigurdson, W. J.; Kung, C.; Sachs, F. Energetic and Spatial Parameters for Gating of the Bacterial Large Conductance Mechanosensitive Channel, MscL. *J. Gen. Physiol.* **1999**, *113* (4), 525–539. https://doi.org/10.1085/jgp.113.4.525.

(49) Haswell, E. S.; Phillips, R.; Rees, D. C. Mechanosensitive Channels: What Can They Do and How Do They Do It? *Structure*. NIH Public Access October 12, 2011, pp 1356–1369. https://doi.org/10.1016/j.str.2011.09.005.

(50) Hofmann, S.; Januliene, D.; Mehdipour, A. R.; Thomas, C.; Stefan, E.; Brüchert, S.; Kuhn, B. T.; Geertsma, E. R.; Hummer, G.; Tampé, R.; Moeller, A. Conformation Space of a Heterodimeric ABC Exporter under Turnover Conditions. *Nature* **2019**, *571* (7766), 580–583. https://doi.org/10.1038/s41586-019-1391-0.

(51) Marinko, J. T.; Huang, H.; Penn, W. D.; Capra, J. A.; Schlebach, J. P.; Sanders, C. R. Folding and Misfolding of Human Membrane Proteins in Health and Disease: From Single Molecules to Cellular Proteostasis. *Chemical Reviews*. American Chemical Society May 8, 2019, pp 5537–5606. https://doi.org/10.1021/acs.chemrev.8b00532.


σNP (cal mol-1 A-2)
-25 -20 -15 -10 -5 0
-25 -20 -15 -10 -5 0
[Water] (M)
0 10 20 30
0 10 20 30