Callenberg et al., http://www.jgp.org/cgi/content/full/jgp.201110766/DC1

Supplemental text

Electrostatic energy

The electrostatic energy of the protein, G_{elec}, is highly dependent on the local dielectric environment, and there is a large energetic cost for moving charged and polar residues into the membrane. We determine the electrostatic potential, Φ, by solving the Poisson-Boltzmann equation:

$$-\nabla\cdot[\varepsilon(r)\nabla\phi(r)]+\kappa^2(r)\sinh[\phi(r)]=\frac{e}{k_BT}4\pi\rho(r), \quad \text{(S1)}$$

where $\phi = \Phi/k_BT$ is the reduced electrostatic potential, κ is the Debye-Huckel screening coefficient to account for ionic shielding, ε is the spatially dependent dielectric constant, e is the electron charge, and ρ is the charge density within the protein. The energy is then given by:

$$G_{elec}=\int \Phi\rho\, dxdydz. \quad \text{(S2)}$$

In aqueous solution, ε was set to the value of water for all points outside of the protein molecular surface, whereas ε is modified to take on values corresponding to the membrane for all spatial points between the upper surface, u^+, and lower surface, u^-, determined from solving Eq. 2 in the main text. Additionally, κ is set to zero for points between u^+ and u^-, indicating a lack of counterion penetration into the membrane. ΔG_{elec} is then given by the difference between the electrostatic energies calculated in bulk aqueous solution and in the presence of the membrane.

The headgroup regions were defined as all points 8 Å from the upper or lower membrane surfaces. For positions r within the headgroup regions, $\varepsilon(r)$ was set to a high dielectric of 80, whereas the ion accessibility $\kappa(r)$ was set to zero. A detailed description of the manipulations to the microenvironment of the protein in the presence of the membrane can be found in our original manuscript (Choe et al., 2008), and in our recent publication supporting our electrostatics software package APBSmem (Callenberg et al., 2010).

Because the points describing the membrane surface were on a different grid than those used for the APBS calculations, we used the cubic interpolation function in MATLAB to move between grids.

The linearized Poisson-Boltzmann equation was used for all calculations reported here; however, we performed test runs using the full nonlinear equation. The electrostatic component of the energy varied <0.2 kcal/mol for individual calculations, but when performing a search for optimal bending, the final electrostatic energy varied by up to 0.7 kcal/mol.

During the search, the electrostatics grid spacing was set to 0.77 Å for increased speed. Once the lowest energy membrane shape was identified, a final electrostatics calculation was performed using a grid spacing of 0.31 Å for greater accuracy. For searches that used the finer discretization from the start, final energy values differed by 0.3 kcal/mol or less.

Nonpolar energy

The nonpolar energy coming from the tendency of water to exclude molecules results in a large stabilizing force for proteins in the membrane. We model it by assuming that the energy difference of the protein in solution compared with the protein in the membrane, G_{np}, is proportional to the difference in the protein's solvent accessible surface area (SASA):

$$G_{np}=a\cdot(A_{mem}-A_{sol})+b, \quad \text{(S3)}$$

where A_{mem} is the protein SASA in the membrane, A_{sol} is the total protein SASA, a = 0.028 kcal/mol and b = −1.7 kcal/mol. These last two constants have been determined based on the partitioning of small nonpolar molecules between aqueous and organic phases (Sitkoff et al., 1994). SASA values are calculated with a modified Shrake-Rupley algorithm (Shrake and Rupley, 1973) using the solvent-accessible surface representation of the protein with a water probe radius of 1.4 Å. In the presence of the membrane, if the point on the surface of the protein lies between the upper, u^+, and lower, u^-, leaflets then it is considered occluded and does not contribute to A_{mem}. MATLAB's cubic interpolation function was used to navigate between the position of the point on the protein surface and the grid points describing the upper and lower membrane surfaces.

Search strategies

Searches that are initiated from a flat membrane must overcome a large nonpolar energy barrier to expose a central charged residue to aqueous solution. The red curve in Fig. S1 illustrates the difficulty that the original search algorithm has in crossing this barrier to solvate a central lysine residue on a hydrophobic leucine-alanine TM segment. At iteration 0, the membrane is flat and the insertion energy for the segment is −15 kcal/mol, whereas we know that the true minimum energy is below −40 kcal/mol (see blue and green curves). Even after ∼3,000 iterations, the initial strategy fails to promote sufficient bending of the membrane to expose the lysine, resulting in unfavorable electrostatics. Our next approach was to ignore the nonpolar energy term in the cost function for the first 50 iterations, thereby removing the nonpolar barrier and minimizing the electrostatic energy component. Within the first 50 iterations, the search identifies large bent configurations that drop the total energy down below −20 kcal/mol by exposing the lysine (green curve). By 1,000 iterations, the search has dropped below −40 kcal/mol and shows little improvement over the next 2,000 iterations. The large spikes in the energy every 250 iterations are caused by the search algorithm attempting to cover the charged residue. Lastly, we started the search from two initial guesses that exposed the buried charged residue. First, the lower leaflet remained flat and the shape of the upper leaflet was set to a pure sinusoidal curve with a period of 2π and an amplitude and phase that placed the charged atoms in the polar headgroup region or solvent. This was then repeated for the bottom leaflet, and the shape that produced the lowest energy was used to initiate the Powell's-based search. The initial guess method significantly outperforms the other strategies in both its speed and ability to identify global minima as indicated by the blue curve in Fig. S1. For all of the results presented in this manuscript, the initial guess strategy was used and produced the most energetically favorable solutions; however, in many cases we ran all search strategies as a precaution.

Comparison to existing residue insertion scales

The construction of a hydrophobicity scale allows us to compare our method to a range of previous computational and experimental work. Our expectation was that optimizing the membrane configuration would lead to lower insertion energies compared with our previous calculations, which posited simple contact curves (Choe et al., 2008). This is important because experiments reveal relatively low insertion energy values compared with computational scales. As in previous studies (Hessa et al., 2005; Choe

et al., 2008), we matched each helix containing a unique central amino acid with an identical helix whose central residue had been replaced with leucine. We determined the apparent free energy of insertion for each amino acid by finding the difference in the two peptides' insertion energies and subtracting the insertion energy of an individual leucine, −3.6 kcal/mol, which was calculated as described in our previous work (Choe et al., 2008).

As shown in Fig. S2, our revised scale is consistent with the overall trend of experiments that finds hydrophobic residues most stable in the membrane and polar and charged residues less stable. The magnitude and spread of the energy values remain the same between our current results and our original work (Choe et al., 2008), but importantly, our new algorithm outperforms the original model for all amino acids, predicting charged residue insertion energies to be 1–2 kcal/mol lower. Although the rank ordering of our scale is similar to a recent experiment on peptide insertion in the inner mitochondrial membrane (Botelho et al., 2011), our predicted values are still more in line with computational results (MacCallum et al., 2007) than the low values reported by the translocon studies (Hessa et al., 2005, 2007). Visualization of the system geometry revealed that the membrane significantly bends around central charged residues but remains flat when polar and hydrophobic amino acids are inserted. As we previously reported, charged lysine and polar asparagine have comparable insertion energies, although the cost of inserting asparagine is primarily electrostatic because the membrane does not bend, whereas the cost of inserting lysine is largely nonpolar because the membrane bends to expose large regions of the TM segment.

We also tested the effect that the material properties of the membrane had on the scale. In separate calculations, we reduced the compression modulus (K_a), bending modulus (K_c), and stretch (α) parameters to 1/2 of their normal value and, in each case, the insertion energy scale was only mildly impacted. The charged residue values were most affected, becoming 0.5 kcal/mol easier to insert (data not shown).

Context dependence of hydrophobicity scales

A biological hydrophobicity scale was reconstructed using a TM segment with different flanking residues (yellow bars in Fig. S2). We replaced the N- and C-terminal flanking glycine residues, similar to the H-segment studied by Hessa et al. (2005), with a charged sequence, NNKK ... KKNN, typical of amino acid sequences at the membrane–water interface. As shown in the bar chart, most of the Helix 2 insertion energies are within 1 kcal/mol of the original values (dark green); however, glutamate, aspartate, lysine, and arginine all destabilize the new segment 2–7 kcal/mol more than they destabilize the original segment.

This destabilization is predominantly caused by a large increase in the solvent-accessible surface area. In the new reference peptide, in which the central amino acid is a leucine, the membrane remains flat and the lysine residues are buried in the headgroup region. This is an energetically favorable configuration because it reduces the SASA of the TM segment while allowing the charged lysines to interact favorably with polar lipid headgroups. When the central amino acid is replaced by a charged residue, the membrane bends to expose the NNKK sequence to water, resulting in a much greater nonpolar energy penalty than incurred for the original GGGG sequence, due to the larger size of lysines and asparagines.

Comparison with Generalized Born

We wanted to compare our continuum method for computing the energetics of membrane proteins to other continuum methods that use Generalized Born methods along with modifications to model the membrane. We chose the CHARMM GBSW module, which is an extremely popular method in the field (Im et al., 2003a,b). Insertion energies for TM helices harboring varying numbers of arginine residues were calculated in CHARMM c32b2 using the molecular surface representation, a bilayer thickness of 42 Å, a membrane dielectric of 2, a solvent dielectric of 80, and an effective surface tension coefficient of 0.005. We used the same set of parameter values to carry out the comparison calculations using our method. The peptide's atomic radii were taken from the radius_gbsw set, which was calibrated for GBSW calculations (Nina et al., 1997; Chen et al., 2006). A membrane-switching length of 2.5 Å was used. The default values were used for all other GBSW parameters. Briefly, we performed an electrostatic point calculation in the presence of the membrane, and then we performed the same point calculation in solution using an external dielectric value of 80 everywhere outside the protein. The electrostatic component of the insertion energy was recorded as the difference between these two energies.

SUPPLEMENTAL REFERENCES

Botelho, S.C., M. Osterberg, A.S. Reichert, K. Yamano, P. Björkholm, T. Endo, G. von Heijne, and H. Kim. 2011. TIM23-mediated insertion of transmembrane α-helices into the mitochondrial inner membrane. *EMBO J.* 30:1003–1011. http://dx.doi.org/10.1038/emboj.2011.29

Callenberg, K.M., O.P. Choudhary, G.L. de Forest, D.W. Gohara, N.A. Baker, and M. Grabe. 2010. APBSmem: a graphical interface for electrostatic calculations at the membrane. *PLoS ONE.* 5:e12722. http://dx.doi.org/10.1371/journal.pone.0012722

Chen, J., W. Im, and C.L. Brooks III. 2006. Balancing solvation and intramolecular interactions: toward a consistent generalized Born force field. *J. Am. Chem. Soc.* 128:3728–3736. http://dx.doi.org/10.1021/ja057216r

Choe, S., K.A. Hecht, and M. Grabe. 2008. A continuum method for determining membrane protein insertion energies and the problem of charged residues. *J. Gen. Physiol.* 131:563–573. http://dx.doi.org/10.1085/jgp.200809959

Hessa, T., H. Kim, K. Bihlmaier, C. Lundin, J. Boekel, H. Andersson, I. Nilsson, S.H. White, and G. von Heijne. 2005. Recognition of transmembrane helices by the endoplasmic reticulum translocon. *Nature.* 433:377–381. http://dx.doi.org/10.1038/nature03216

Hessa, T., N.M. Meindl-Beinker, A. Bernsel, H. Kim, Y. Sato, M. Lerch-Bader, I. Nilsson, S.H. White, and G. von Heijne. 2007. Molecular code for transmembrane-helix recognition by the Sec61 translocon. *Nature.* 450:1026–1030. http://dx.doi.org/10.1038/nature06387

Im, W., M. Feig, and C.L. Brooks III. 2003. An implicit membrane generalized born theory for the study of structure, stability, and interactions of membrane proteins. *Biophys. J.* 85:2900–2918. http://dx.doi.org/10.1016/S0006-3495(03)74712-2

Im, W., M.S. Lee, and C.L. Brooks III. 2003. Generalized born model with a simple smoothing function. *J. Comput. Chem.* 24:1691–1702. http://dx.doi.org/10.1002/jcc.10321

MacCallum, J.L., W.F. Bennett, and D.P. Tieleman. 2007. Partitioning of amino acid side chains into lipid bilayers: results from computer simulations and comparison to experiment. *J. Gen. Physiol.* 129:371–377. http://dx.doi.org/10.1085/jgp.200709745

Nina, M., D. Beglov, and B. Roux. 1997. Atomic radii for continuum electrostatics calculations based on molecular dynamics free energy simulations. *J. Phys. Chem. B.* 101:5239–5248. http://dx.doi.org/10.1021/jp970736r

Shrake, A., and J.A. Rupley. 1973. Environment and exposure to solvent of protein atoms. Lysozyme and insulin. *J. Mol. Biol.* 79:351–371. http://dx.doi.org/10.1016/0022-2836(73)90011-9

Sitkoff, D., K. Sharp, and B.H. Honig. 1994. Accurate calculation of hydration free energies using macroscopic solvent models. *J. Phys. Chem.* 98:1978–1988. http://dx.doi.org/10.1021/j100058a043

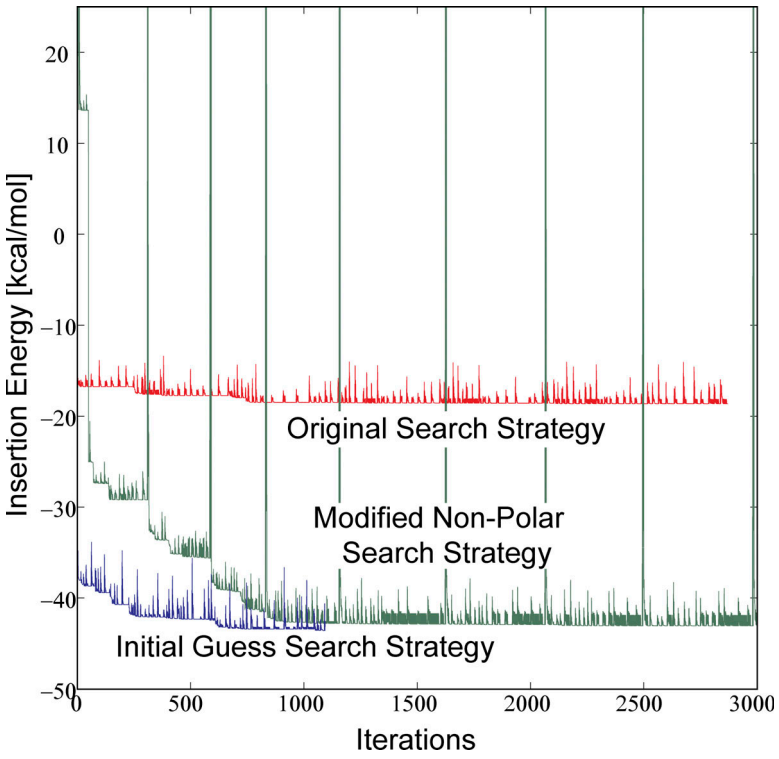

Figure S1. Three different search strategies. We attempted to minimize the total insertion energy of a hydrophobic helix with a central charged lysine. The original search strategy has the membrane start from a flat, unstressed state. The height value of the contact curve nodes shown in Fig. 1 C were then used as independent parameters in the Powell's search strategy. After nearly 3,000 iterations of the search strategy, the total energy has been reduced by only a few kcal/mol because of the inability to cross a high-energy barrier of exposing hydrophobic residues to solvent before uncovering the lysine (red curve). The modified nonpolar search strategy disregards the nonpolar energy for the first 50 iterations to overcome the barrier and expose the central charged residue to water after ∼1,000 iterations (green curve). The initial guess search strategy starts from a distorted contact curve that already exposes the buried charged residue to water (blue curve). This method quickly identifies a membrane configuration that is a few kcal/mol more stable than the nonpolar method.

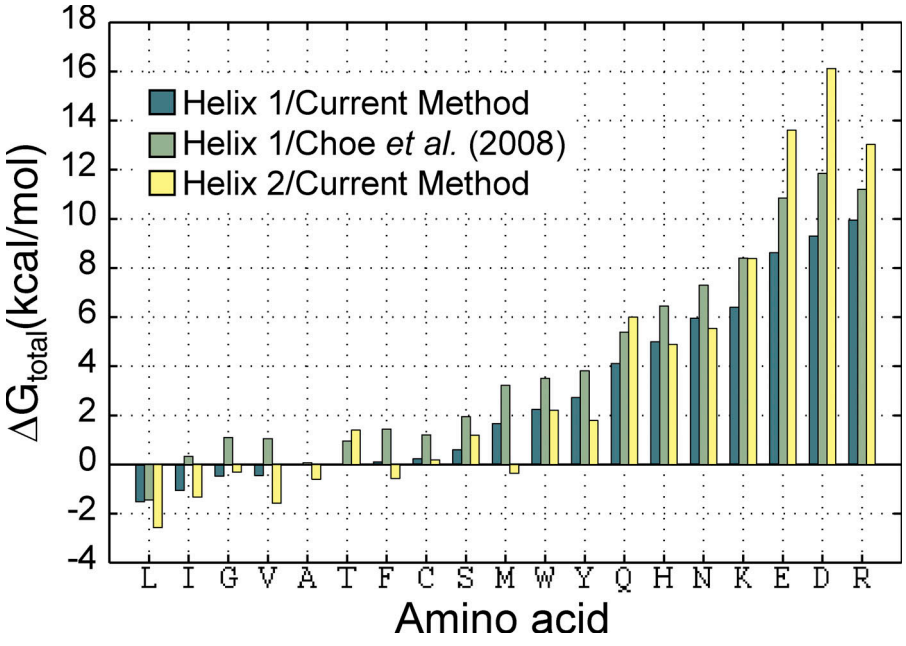

Figure S2. Biological hydrophobicity scale for inserting all-natural amino acids (except proline) in the center of a TM helix. Our search algorithm (dark green) identifies insertion energies that are lower than manual guesses from our previous work (light green) when using the same Helix 1 peptides (H-segments) flanked by four glycine residues (Choe et al., 2008). Charged amino acids (K, E, D, and R) are 2–7 kcal/mol more destabilizing to Helix 2 peptides (yellow bars), which are flanked by polar asparagines and charged lysines. All 3 scales were shifted by +2.04 kcal/mol to set the alanine insertion energy to zero for the current method.

Table S1

Electrostatics and system parameters for all calculations

Parameters	**Value**
Electrostatics grid deimensions	161 × 161 × 161 grid points
Coarse grid lengths	200 × 200 × 200 grid points
Medium grid lengths	100 × 100 × 100 grid points
Fine grid lengths	50 × 50 × 50 grid points
Counter-ions	0.1 M symmetric salt
Protein dielectric	2.0
Membrane dielectric	2.0
Headgroup dielectric	80.0
Solvent dielectric	80.0
Solution method	Linearized Poisson-Boltzmann equation
Solvent probe radius	1.4 Å
Surface sphere density	10.0 grid points/Å^2
Temperature	298.15 K
Membrane thickness	42.0 Å
Headgroup thickness	8.0 Å
Bending modulus (K_c)	2.85×10^{-10} N/Å
Compression modulus (K_a)	1.425×10^{-10} N/Å
Surface tension (α)	3×10^{-10} N/Å