

E_z , a Depth-dependent Potential for Assessing the Energies of Insertion of Amino Acid Side-chains into Membranes: Derivation and Applications to Determining the Orientation of Transmembrane and Interfacial Helices

Alessandro Senes[†], Deborah C. Chadi[†], Peter B. Law
Robin F. S. Walters, Vikas Nanda and William F. DeGrado*

Department of Biochemistry and
Molecular Biophysics
School of Medicine
University of Pennsylvania
Philadelphia, PA 19104-6059
USA

We have developed an empirical residue-based potential (E_z potential) for protein insertion in lipid membranes. Propensities for occurrence as a function of depth in the bilayer were calculated for the individual amino acid types from their distribution in known structures of helical membrane proteins. The propensities were then fit to continuous curves and converted to a potential using a reverse-Boltzman relationship. The E_z potential demonstrated a good correlation with experimental data such as amino acid transfer free energy scales (water to membrane center and water to interface), and it incorporates transmembrane helices of varying composition in the membrane with trends similar to those obtained with translocon-mediated insertion experiments. The potential has a variety of applications in the analysis of natural membrane proteins as well as in the design of new ones. It can help in calculating the propensity of single helices to insert in the bilayer and estimate their tilt angle with respect to the bilayer normal. It can be utilized to discriminate amphiphilic helices that assume a parallel orientation at the membrane interface, such as those of membrane-active peptides. In membrane protein design applications, the potential allows an environment-dependent selection of amino acid identities.

© 2006 Published by Elsevier Ltd.

Keywords: membrane proteins; potential; hydrophobicity; membrane insertion; transmembrane helices

*Corresponding author

Introduction

The primary amino acid sequence of membrane proteins directs the proper positioning of the helices in the lipid bilayer following insertion *via* the translocon apparatus. Thus, considerable effort has been made to determine the thermodynamics of transfer of the amino acid side-chains from water to various regions of the membrane.¹ Schematically, the bilayer is generally considered to consist of distinct sectors representing the hydrophobic hydrocarbon core, a polar headgroup region, and the extra-membrane aqueous region. The free energies

of transfer of amino acid derivatives from water to the headgroup or hydrocarbon regions of a bilayer have been determined using a variety of experimental techniques.^{1–3} Also, amino acid sequence data for single-span and multi-span helical membrane proteins have been evaluated to determine the propensity of various amino acids to form helices at distinct sectors of the bilayer.^{4–7}

Although the multi-sector approach provides a good first approximation for the bilayer environment, it would be advantageous to determine the free energy profile for each residue type as a continuous function of its depth of insertion within a membrane. Here, we determine the frequency of occurrence of various amino acid types as a function of their position in a bilayer. Using the reverse-Boltzmann statistical approach, these propensities are converted to pseudo-energies, which are shown

[†] A.S. and D.C.C. contributed equally to this work.

E-mail address of the corresponding author:
wdegrado@mail.med.upenn.edu

to correlate well with experimental free energies of transfer of side-chains from water to distinct regions of the bilayer. The resulting method is used to predict the locations and orientations of helices in membrane proteins and has a variety of applications in the analysis of natural membrane proteins as well as in the design of novel structures. For example, there is keen interest in determining the extent to which polar side-chains, particularly Arg, can be accommodated stably within the bilayer, particularly as this relates to the mechanism of voltage-sensing in potassium channels,^{8–10} as well as the mechanism of membrane disruption by cationic membrane-permeabilizing and antimicrobial peptides.^{11–14} The energy function can help to define the orientation of a protein in the membrane, which cannot be inferred directly from a crystal structure because the positions of the phospholipids are generally not defined, although isolated lipids are visible in the electron density in some cases. Given the large amount of available sequence data, another important application is the estimation of the propensity for insertion of predicted transmembrane domains and their preferred angle with respect to the bilayer normal. Finally, a computationally efficient empirical energy function is useful in membrane protein design to guide the selection of specific side-chains at various positions of the peptide chain as a function of their environment.⁷

Results and Discussion

Depth-dependent propensity profiles for amino acid side-chains

To determine the propensities of the amino acid side-chains across the bilayer depth, we analyzed previously studied helical membrane proteins positioned in an implicit bilayer. Ideally, we would consider the cytoplasmic and extracellular ends of the helix differently, and would additionally discriminate the N-terminal ends from the C-terminal ends of the helices (because the residues have different rotamer distributions on either end of the helix).^{15,16} However, to maximize the signal-to-noise ratio, these features were not differentiated, and the distance of the residues C^B (C^α for Gly) from the bilayer center was measured.

Figure 1 illustrates the propensity for the side-chains to occupy consecutive 2 Å regions beginning at the center of the bilayer, and extending to 30 Å along a line that is normal to the bilayer plane. The z coordinate defines the distance from the center of the bilayer. We define the position-dependent propensity (P_z) for a given residue as the number of residues of this type in each 2 Å increment, divided by the mean value for all 2 Å increments for the entire 30 Å region considered. At a qualitative level, the resulting distributions are consistent with conventional wisdom concerning the partitioning of substances between apolar and

polar environments.¹ The hydrophobic residues (Figure 1(a)) have the highest propensity to occur near the center of the bilayer, and the polar residues (Figure 1(b)) show the highest propensities to occur on the exterior of the bilayer ($z > 20$ Å). Pro, which behaved similarly to the polar residues, has a hydrophobic side-chain but in a transmembrane helix it often causes one or more main chain hydrogen bonding groups to be unsatisfied, which are thermodynamically unfavorable in the non-polar core of the membranes.¹⁷ While it can be accommodated easily in a helix when embedded deep within a membrane,^{8,18} Pro is also the amino acid with the highest propensity to promote helical hairpin formation when present near the headgroup region.^{19,20}

Aromatic residues with a single polar group (Tyr and Trp, Figure 1(c)) have a high tendency to locate at the headgroup region, whereas, as expected, Phe behaves similarly to the other hydrophobic amino acids.^{21,22} Finally, relatively flat profiles are observed for the small residues Gly, Ser, Thr and Cys (Figure 1(d)), which are mildly polar but can be accommodated favorably also in apolar environments.²³

A continuous z -dependent potential can be derived from the amino acid propensities, which allows the calculation of pseudo-energies of insertion of a protein as a function of conformation, orientation, and depth in the membrane bilayer. The relationship of the amino acid propensities with energy can be expressed as:

$$P_z = P_{\text{aq}} e^{-\frac{\Delta E_z}{RT}} \quad (1)$$

where P_z is the propensity as a function of the depth coordinate, z ; P_{aq} is the propensity in pure water when z is extrapolated to infinity, and ΔE_z is the energy difference between the residue in water and at a given depth within the bilayer as defined by its z coordinate; R is the gas constant and T is the absolute temperature.

It is important to choose a functional form of ΔE_z to account for the differences in the free energy of solvation of the various residues at different portions of the bilayer. With the exception of Trp and Tyr, the P_z versus z distributions tend to be sigmoidal, with varying degrees of steepness. This behavior is captured well by the equation:

$$\Delta E_z = \frac{\Delta E_0}{1 + \left(\frac{z}{Z_{\text{mid}}}\right)^n} \quad (2)$$

where ΔE_0 is the pseudo-energy difference between center of the bilayer ($z=0$) and water (the standard state, $z=\infty$) for a residue; Z_{mid} is the z coordinate at which the energy is half-maximal; and n is a parameter that defines the steepness of the transition.

Equation (2) was inserted into equation (1) and non-linear least-squares analysis was used to obtain the values of P_{aq} , ΔE_0 , Z_{mid} , and n (Table 1), leading

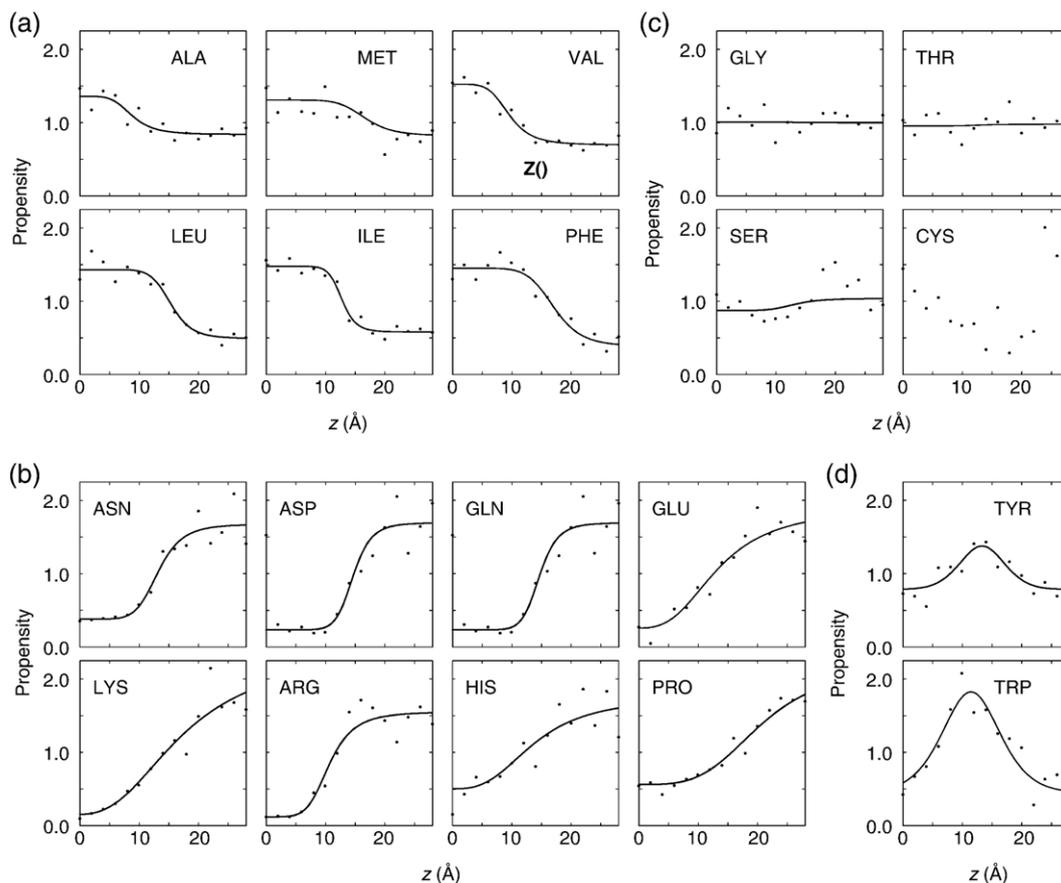


Figure 1. Propensity plots for the 20 amino acids as a function of depth in the bilayer. The z dimension represents the normal to the membrane plane, with the origin at the center of the bilayer. The data were obtained by counting each amino acid occurrence in bins of 2 Å. Panels are grouped by amino acid types: (a) hydrophobic residues; (b) very polar residues; (c) moderately polar residues; and (d) Tyr and Trp. (a) The hydrophobic group reaches a minimum at the center of the bilayer, while (b) the hydrophilic residues prefer to partition in water. Proline, which alters backbone hydrogen bonding in secondary structure, behaves similarly to the polar residues. (c) Gly, Ser and Thr show similar propensity in all environments, while Cys has a scattered distribution and was not fit (a flat zero energy profile was used). Groups (a), (b) and (c) were fit to a sigmoidal function. (d) The propensities of aromatic residue Trp and Tyr were fit to a Gaussian function that reaches the global minimum in the headgroup region.

to an excellent fit to the data (Figure 1). Together, these four parameters provide an accurate and readily interpretable depiction of the properties of the individual amino acids. For Trp and Tyr, whose propensities reach a maximum in the headgroup region and decrease in the membrane center (Figure 1), a properly scaled Gaussian distribution provided a good fit for ΔE_z :

$$\Delta E_z = \Delta E_{\min} e^{-\frac{(z-Z_{\min})^2}{2\sigma^2}} \quad (3)$$

where ΔE_{\min} is the energy minimum at coordinate Z_{\min} , and σ is the width of the transition.

Figure 2 illustrates the ΔE_z profiles for three typical amino acids in comparison with the membrane regions determined by the locations of specific lipid functional groups, as defined previously by diffraction studies.²⁴ The aromatic group of Trp has a strong preference for the carbonyl region of the bilayer; and similar behavior is observed for Tyr. The curve for Phe is typical for the hydrophobic group (Figure 2) with ΔE_z values that reach a limit as they

pass through the carbonyl region into the hydrocarbon core. The curves for highly polar residues, exemplified by Asp in Figure 2, show the opposite behavior.

ΔE_z profiles define barrier positions, energy gradients, and snorkeling potential

The parameters, ΔE_0 , Z_{mid} , and n provide a highly informative and intuitive picture of the behavior of the individual amino acids in a bilayer environment. For polar residues the middle point of the transition (Z_{mid}) gives an estimate of how deeply a group prefers to penetrate into a membrane (and *vice versa* for hydrophobic residues). The value of n describes the steepness of the transition from the water to the membrane center values; a group with a large n experiences a hard effective potential with a sudden transition, while a group with small n experiences a softer one that varies gradually over a larger range of z .

The large hydrophobic groups Leu, Phe, Ile, and Met all show highly favorable propensities in the

Table 1. Parameters of the pseudo-energy potential for the 20 amino acids

Functional form 1 (equation (2))			
Residue	ΔE_0	Z_{mid}	n
Ala	-0.29	10.22	4.67
Asp	1.19	14.25	8.98
Glu	1.30	14.66	4.16
Phe	-0.80	19.67	7.12
Gly	-0.01	13.86	6.00
His	0.75	12.26	2.77
Ile	-0.56	14.34	10.69
Lys	1.66	11.11	2.09
Leu	-0.64	17.34	8.61
Met	-0.28	18.04	7.13
Asn	0.89	12.78	6.28
Pro	0.83	18.09	3.53
Gln	1.21	10.46	2.59
Arg	1.55	9.34	4.68
Ser	0.10	13.86	6.00
Thr	0.01	13.86	6.00
Val	-0.47	11.35	4.97

Functional form 2 (equation (3))			
Residue	ΔE_{min}	Z_{min}	σ
Trp	-0.85	11.65	7.20
Tyr	-0.42	13.04	6.20

A sigmoidal fit was applied to the distribution of 17 amino acid types (equation (2)). For the aromatic residues Trp and Tyr, a Gaussian fit was applied (equation (3)). Cys had a scattered distribution profile and was fit to a constant zero energy curve.

range of $z=0$ to 13 \AA , which corresponds with the hydrocarbon core of the membrane. As z increases into the positions of the carbonyl groups of the fatty acid esters in the bilayer (Figure 2), ΔE_z becomes less favorable, reaching the limiting value at the transition through the headgroup region into water. The value of ΔE_z has already decayed to near zero close to the center of the region occupied by the choline headgroups. Interestingly, the propensities for the smaller hydrophobic residues Val and Ala tend to fall off at smaller values of z ($Z_{\text{mid}}=11.3 \text{ \AA}$ and 10.2 \AA , respectively), indicating that they are not as well accommodated in the headgroup region.

The differences in the values of ΔE_0 , Z_{mid} , and n for the strongly polar and charged residues is particularly interesting. By comparing Asp with Asn, and Glu with Gln, it becomes apparent that the carboxylate-containing Asp and Glu are excluded more effectively from the apolar region of the bilayer than their isosteric neutral counterparts ($Z_{\text{mid}}=14.7 \text{ \AA}$ for Glu *versus* 10.5 \AA for Gln; 14.3 \AA for Asp *versus* 12.8 \AA for Asn, Table 1).

In general, polar residues with long side-chains are able to snorkel;^{15,25} the ability to snorkel appears to translate to having softer potentials of mean force, as assessed by the value of n . Asp experiences a particularly hard potential with $n=6.3$, which decreases to 4.2 for Glu, which is one methylene group longer and thus it is more flexible. The same trend is observed for Asn *versus* Gln, whose values of n are 6.3 and 2.6, respectively. Indeed, the longest side-chain, Lys, has the softest potential of the amino acids ($n=2.0$).

It is particularly interesting to compare the curves for the acidic residues Asp and Glu to the strongly basic residues Arg and Lys. While the basic amino acids have slightly more unfavorable values of ΔE_0 for transfer from water to the center of the bilayer (Table 1), they nevertheless have smaller values of Z_{mid} ; which are 9.3 \AA and 11.1 \AA for Arg and Lys *versus* 14.3 \AA and 14.7 \AA for Asp and Glu, respectively. Thus, while it is energetically more difficult to transfer the Arg and Lys side-chains fully to the center of the bilayer (as can be inferred from the values of ΔE_0 , Table 1), they can nevertheless penetrate approximately 4–5 \AA more deeply than the acidic side-chains without incurring too large a hydration penalty. Also, the values of n for these residues indicates that Lys, in particular, experiences a surprisingly soft potential as it penetrates into the bilayer. Presumably, these differences between the acidic and basic side-chains reflect two features: the first is that Lys and Arg are longer and more flexible than Asp and Glu, providing opportunities for snorkeling towards the surface. Additionally, these residues can interact favorably with the negatively charged phosphate group of phospholipids. The phosphate group is located more deeply than the corresponding positively charged groups (e.g. the choline moiety of phosphatidylcholine lipids) even in zwitterionic phospholipids.

Analysis of the profiles for Trp and Tyr shows the transfer of these residues to the headgroup/interfacial region of the bilayer to be highly favorable ($Z_{\text{mid}}=11.6 \text{ \AA}$ and 13.0 \AA) in agreement with the results of previous studies.^{21,22,26,27} As expected

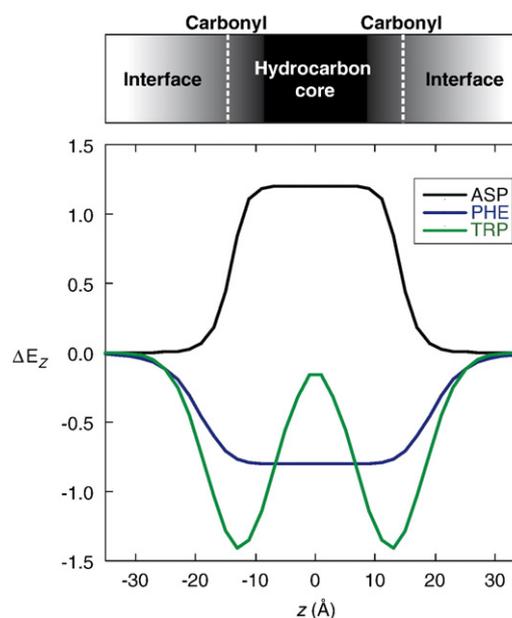


Figure 2. Representative energy curves for a polar (Asp, black), a hydrophobic (Phe, blue) and an interfacial aromatic residues (Trp, green). The hydrocarbon core (black), headgroup interfacial region (gray) and water (white) are highlighted in the top panel. The standard state was set to bulk water ($z=\infty$).

from its chemical structure, the transfer of Trp is more favorable and it penetrates, on average, slightly more deeply into the bilayer.

Correlations with experimental transfer free energies

The values of ΔE_0 obtained from equations (2) and (3), is the pseudo free energy of transfer from water to the center of the bilayer. As expected, the values of ΔE_0 correlate well with various hydrophobicity scales (Figure 3), including that of Eisenberg (slope=1.19, intercept=-0.37 kcal/mol, $R=0.94$)²⁸ and White (slope=0.87, intercept=-0.13 kcal/mol, $R=0.78$)²⁹ better than these two hydrophobicity scales agree with one another (slope=0.60, intercept=0.14 kcal/mol, $R=0.69$). The plots of ΔE_0 versus the experimental scales have slopes near unity, indicating that the pseudo-chemical approximation used in this work (with an effective temperature of 25 °C) provides reasonable scaling of pseudo-energies. The correlation with the scale used by Eisenberg is particularly good, and only Arg falls off the line significantly. This residue has the largest counting error near the center of the bilayer; if it is removed from the plot, the correlation coefficient of Figure 3(a) becomes 0.95.

An advantage of a continuous potential is that it makes it possible to estimate the transfer energies to any depth within the bilayer. For example, it is possible to compute the free energy of transfer from water to the headgroup or "interface" region of the bilayer, allowing a comparison with another experimental energy scale obtained by White and co-workers, for transfer to the membrane interface.³⁰ They measured the transfer free energies for a series of peptides, which were known to localize in 1-palmitoyl-2-oleoyl-phosphatidylcholine (POPC) bilayers near the headgroup region of the bilayer. We calculated the correlation between their interface scale and ΔE_z evaluated at various values of z . A plot of the correlation coefficients for the regression analyses versus the value of z at which ΔE_z was

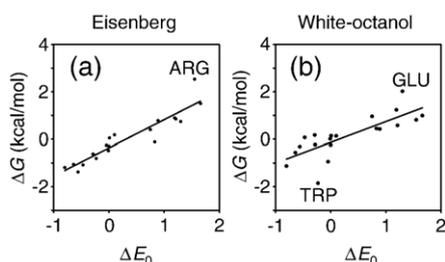


Figure 3. Comparison of the E_z water to hydrocarbon transfer energies (ΔE_0) with experimental transfer free energies for the 20 amino acid types. (a) Correlation with the Eisenberg scale: slope=1.19, intercept=-0.37, $R=0.94$. The major outlier, Arg, is marked on the graph. (b) Correlation with the White-Octanol scale: slope=0.87, intercept=-0.13, $R=0.78$. The major outliers, Trp and Glu, are marked.

computed shows a maximum near $z=12$ Å (Figure 4(a)). Figure 4(b) illustrates the excellent correlation between the ΔE_{12} (the pseudo-energy of transfer of a residue from bulk water to $z=12$ Å) and White interface scale (slope=2.54, intercept=0.26 kcal/mol, $R=0.91$). This is consistent with the conclusion reached by White and co-workers, that the model peptides, upon which the scale was devised, are bound to the headgroup region of the bilayer.

Correlation with translocon-mediated helix insertion

Overall, there is reasonable agreement between the ΔE_z scale and a "biologically" defined scale obtained by Von Heijne and co-workers (slope=0.6; intercept=-0.38 kcal/mol; $R=0.88$; data not shown). The scale was obtained by measuring the efficiency of insertion of model sequences into the endoplasmic reticulum membrane via the Sec61 translocon.³ A hydrophobic segment was introduced into the leader peptidase so that it could either span the membrane or be translocated to the ER lumen. One guest amino acid was varied in the middle of the segment, and the relative amount of glycosylation that occurred at specific sites was taken as a topological indicator of the fraction inserted in a transmembrane orientation. Apparent equilibrium constants and free energy of insertions were thus derived. While biological insertion is a complex and active process operated by a cellular machinery, the good agreement of the data with hydrophobicity-based scales suggests that interaction with the lipids is probably involved in the process, and that the efficiency of insertion is related to the thermodynamic stability in the membrane.³ There is good consistency with the observed rank, including the behavior of Pro, which is a non-polar residue that tends to avoid the center of the bilayer unless required for function.^{17,18,31} One difference is in the value of Phe, which is the most hydrophobic of the non-polar residues in the ΔE_z scale, but third most apolar after Leu and Ile in the biological scale.

The same authors performed a series of experiments to investigate the effect of positioning of the guest amino acids, using constructs in which a pair of residues of the same type were moved symmetrically from the center toward the ends of the segment.³ We made a direct comparison to their data by calculating equivalent pseudo free energies of insertion with the E_z potential (ΔG_{Ez}). The free energies were calculated as the difference between two ensembles: a transmembrane state that includes all the conformations in which the helix spans the bilayer; and one non-transmembrane state, composed of conformations in which the helix is completely exposed to water or associated with the membrane interface. We chose a thickness of the water phase of 50 Å and a temperature of 298 K to simulate a helical segment constrained to be in the vicinity of the membrane (these choices affect the absolute value but do not change the shape of the

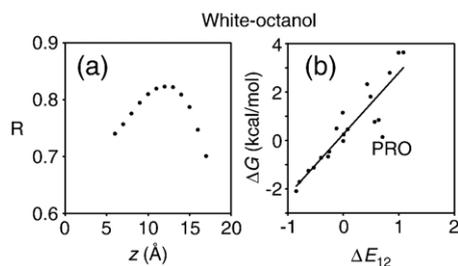


Figure 4. Comparison of the E_z water to headgroup region transfer energies (ΔE_{12}) with experimental transfer free energies of the White interface scale. (a) Correlation coefficients (R values) of the two transfer energies as a function of chosen depth for the E_z potential data. The correlation reaches a maximum at $z=12$ Å. (b) XY plot of the transfer energies for the ΔE_{12} (E_z transfer energy from bulk water to 12 Å from the membrane center) and experimental energies for the 20 residue types. Slope = 2.54; intercept = 0.26; $R = 0.91$.

free energy profiles). Our results are shown in Figure 5 with the same layout as Figure 4 of Hessa *et al.*³ While the energy ranges in the simulation tend to be reduced in magnitude and not as rich in fine features, the trends are in good agreement with most aspects of the experimental results, in particular for hydrophobic and aromatic residues (Figure 5(a)). We did not observe the same stabilization of the

transmembrane conformation when more Ala residues were substituted by Leu (Figure 5(b)), which seems counter intuitive considering that the ΔE_0 of Leu is twice as “hydrophobic” as that of Ala (Table 1). The explanation is that Leu has a transition that occurs quite steeply in the head group region ($Z_{\text{mid}}=17.3$ Å, Table 1) while the transition for Ala is closer to the membrane center ($Z_{\text{mid}}=10.2$ Å). Therefore, a Leu residue will provide similar energy to transmembrane and membrane interfacial conformations, while Ala stabilizes preferentially the transmembrane conformations.

Effects of amphipathicity analogous to those observed experimentally can be recognized in the ΔG_{E_z} profiles of strongly polar residues, such as Asn and Lys (Figure 5(d)), but not for the mildly polar Ser (Figure 5(d)), as a consequence of its rather constant propensity across all environments (Figure 1). Finally, further support for the applicability of the E_z potential is provided by the comparison of the two series involving Pro, either as a pair (Figure 5(e)) or a single residue (Figure 5(f)). The striking similarity between the experimental and computational profiles suggests that while the insertion efficiency *in vitro* is influenced, in part, by helix and turn formation propensity,^{19,20} which cannot be accounted for in a rigid helix simulation, their effects are still captured by the statistically derived propensities.

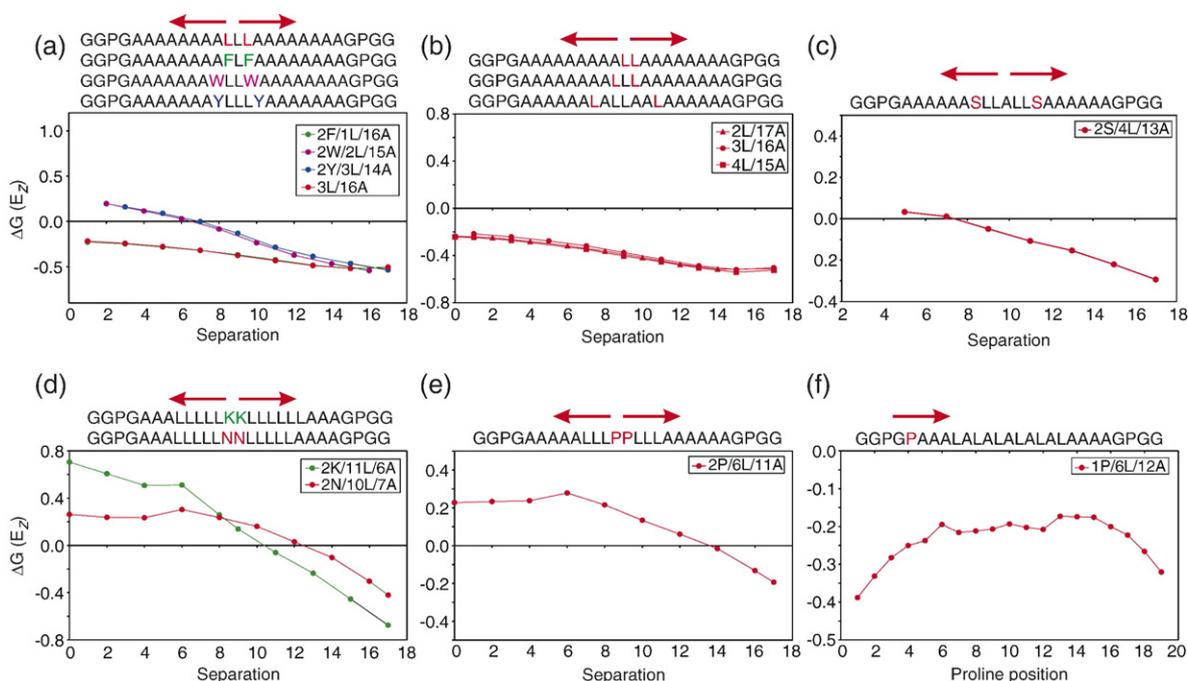


Figure 5. Free energy of insertion of helices in transmembrane form using the E_z potential. The experiment is aimed to computationally recapitulate the series in Figure 4 of Hessa *et al.*³ and the panels here are reproduced with the same layout. In the original experiments the sequences added to the C terminus of the membrane protein leader peptidase so that they can either be inserted as transmembrane helices or be translocated across the membrane. The fraction of transmembrane/translocated molecules was measured and converted to an apparent ΔG of biological insertion. The calculations were performed on a rigid single helix and the sequence is terminated by an Asn residue on both ends. The ΔG_{E_z} of insertion is calculated as the difference in ensemble free energy of two subsets: helices in transmembrane orientation (the termini are on opposite sides of the bilayer center and at $z \geq 13$ Å), and the set of helices that were in water-exposed and interfacial conformations.

Orientation of known structures of membrane proteins using E_z

To verify the extent to which the residue-based potential is able to predict transmembrane helix orientation, we have applied it to a set of known membrane protein structures. The proteins were rigidly rotated and translated to find their global minima positions. Since a definitive orientation of a protein in the membrane cannot be determined directly from its crystal structure, we compared the results to those obtained with the semi-automated method that we utilized to orient the structures used to generate the statistical potential (training set).³² This method is dependent on manual definition of the transmembrane segments. In Table 2 we show the average differences in depth (Δz) and angle ($\Delta\theta$) between the structures oriented with both methods. For the training set of structures, the application of the potential reproduced the original depth in the bilayer with an average difference from the original orientation $\Delta z = 0.2(\pm 2.0)$ Å (Figure 5(a)). The tilt with respect to the membrane plane was also in very good agreement, with an average difference $\Delta\theta = 8(\pm 6)^\circ$.

When the same test was applied to each individual α -helix, the agreement decreased, but was still good, with an average difference, $\Delta z = 1(\pm 6)$ Å, $\Delta\theta = 16(\pm 20)^\circ$. The major outliers included helices that are unlikely to be thermodynamically stable without the aid of the surrounding environment, such as the short helices in the potassium channel and aquaporins that do not span the transmembrane region completely, or helices that include several polar contacts near the center of large bundles.

Finally, we compared the orientation of a set of recently solved membrane protein structures, which were not included in the training set. The average difference in orientation between the two methods is similar to that observed with the training set (Table 2). Figure 6 shows the final orientation obtained with the E_z potential of four proteins in the virtual bilayer. Visually, the proteins appear properly oriented in the membrane. The polar side-chains (the same as Figure 1(a)) are explicitly shown in red. As expected, these side-chains tend to cluster above the dotted lines that mark the 15 Å distance from the membrane center. They also tend to snorkel toward the bulk aqueous phase. The side-chains of the

residues Trp and Tyr, which are enriched in the interfacial region, are shown in yellow in the Figure. It is interesting to notice how the amphiphilic helices of the Na/K channel (2AHY) lay parallel with the plane of the membrane and are located exactly at the edge of the hydrophobic region.

Estimation of the position of sequences of helical peptides in the membrane

We have further tested the E_z potential by calculating the energy landscape of previously studied membrane proteins and a series of peptides as a function of their depth and orientation (Figures 7 and 8). Hydrophobic matching is important for determining the behavior of insertion of a peptide, its orientation and depth in the membrane. This process is dependent on the peptide length, composition and presence of anchoring residues such as Lys and Trp.^{8,21,26} To illustrate the effect of variation of hydrophobic length with the E_z potential we have studied the behavior of a series of hydrophobic peptides formed by n Leu residues flanked by two Lys residue at both termini ($K_2L_nK_2$). Figure 7 shows the energy landscapes as a function of z depth and tilt angle from the membrane normal (θ) of transmembrane single helices in standard conformations. A third variable parameter, the rotation around the helical axis, cannot be shown explicitly in the graph, therefore it is represented by plotting the minimum energy around this rotation for any given z and θ . The minima of the $K_2L_nK_2$ for $n = 2-6$ lie at large z (aqueous environment) and are isotropic with respect to the tilt of the helix. The calculations predict that the helices start to have a tendency to partition to the interfacial zone of the membrane once they include eight to ten Leu residues, with minima around $z = 13$ Å and $\theta = 90^\circ$. They begin to be marginally stable in transmembrane orientation after $n = 12-14$. The perpendicular transmembrane conformation reaches maximum relative stability near $K_2L_{20}K_2$. This is in good agreement with experimental data obtained with poly-Leu peptides.³³ In agreement with experimental and theoretical predictions,³⁴ the helix acquires a tendency to tilt to accommodate more hydrophobic residues in the transmembrane region for peptides with larger number of Leu side-chains. The $n = 22, 24, 26$ and 28 peptides have minima with $\theta = 20^\circ, 30^\circ, 35^\circ$ and 44° , respectively. A saddle-point when the helices are surface-absorbed ($z = 13$ Å and $\theta = 90^\circ$) is in agreement with theoretical calculations on helix insertion.^{35,36}

The landscape of the glycoporphin A (GpA) monomer has features similar to those observed for the longer of the $K_2L_nK_2$ peptides, with a pronounced maximum near $\theta = 90^\circ$ and minimum at $\theta = 17^\circ$ in the vicinity of $z = 0$, which is in good agreement with the conformation inferred from the symmetry of the NMR dimeric structure.³⁷⁻³⁹ As for the poly-Leu peptides, a secondary shallow minimum is found for the interfacial parallel conformation.

Table 2. Average difference in orientation of known structures of transmembrane proteins with the E_z potential, compared to the semi-automated procedure described in Methods

Database	$\Delta z \pm S.D$ (Å)	$\Delta\theta \pm S.D$ (deg.)
Training set	0.2±2.0	8±6
Training set (individual helices)	1.1±6.2	16±20
Testing set	0.4±1.3	9±7

The testing set is formed by recent structures not included in the training set (PDB codes: 1xio, 2ahy, 2a65, 1xqf, and 1yew).

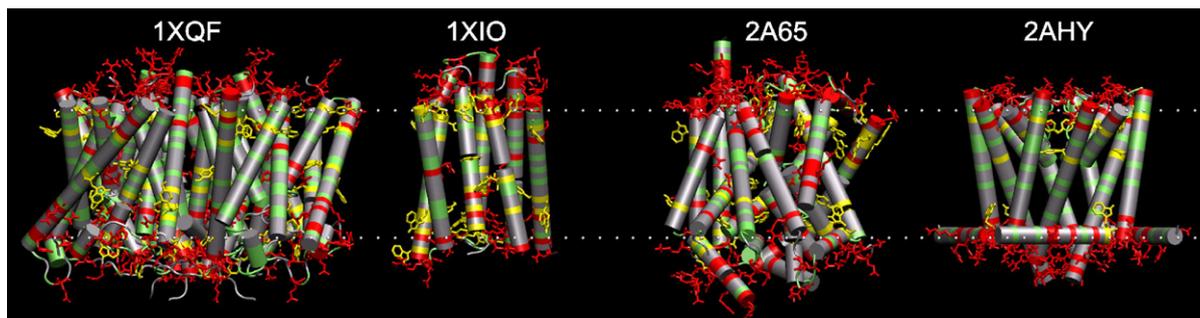


Figure 6. Orientation of transmembrane protein structures with the E_z potential. PDB codes: 1XQF, *E. coli* ammonia channel; 1XIO, *Anabaena* sensory rodhopsin; 2A65, *Aquifex aeolicus* leucine transporter; 2AHY, *Bacillus cereus* Na/K channel. The virtual membrane is represented as the dotted lines at $z = \pm 15$ Å. Hydrophobic (A, V, L, I, M, F) and moderately polar residues (G, S, T, C) are colored in gray and light green, respectively. For strongly polar residues and Pro (D, N, E, Q, H, K, R, P), in red, the side-chains are also shown to highlight snorkeling and the orientation of amphiphilic helices in 2AHY. The polar aromatic residues Trp and Tyr, which occur most frequently in the interfacial region, are represented in yellow.

We tested whether the ΔE_z potential could reliably determine the orientation of membrane-interactive peptides that can be surface-absorbed as well as vertically inserted (Figure 8). As representative examples, we considered the wasp venom peptide

mastoparan X,⁴⁰ the antimicrobial peptide pexiganan,¹⁴ a minimal synthetic antimicrobial peptide LKK (sequence (LKKLLKL)₂),¹² and a membrane-associated peptide from HIV gp41,⁴¹ which is involved in fusion of the virus envelope with the

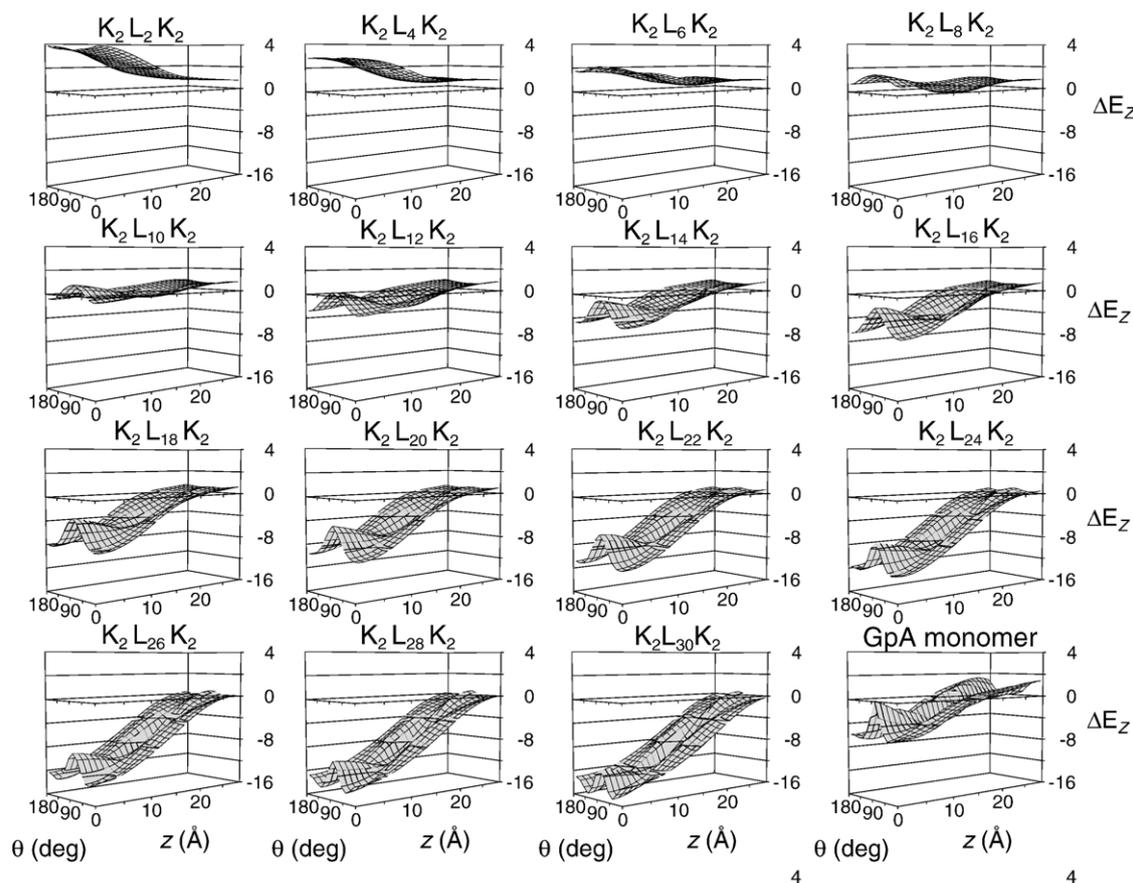


Figure 7. E_z energy landscapes for a series of poly-Leu helical peptide of increasing length, flanked by two Lys residues on each terminus ($K_2L_nK_2$). The energy of insertion is unfavorable (positive) for short helices and becomes more favorable with increasing hydrophobic length. At the membrane interface ($z = 13$ Å) the helices have local minima when parallel to the membrane ($\theta = 90^\circ$). Helices that are barely long enough to span the bilayer ($n = 14$ to 18) have energy minima for transmembrane conformations ($\theta = 0^\circ$ and 180°) that are similar to their interfacial conformation. Long helices ($n > 24$) tilt in the membrane ($\theta = 44^\circ$ for $n = 28$). The energy profile for a single helix of the well studied transmembrane segment of Glycophorin A (GpA) is shown as control. Its helix has a natural tilt near 20° .

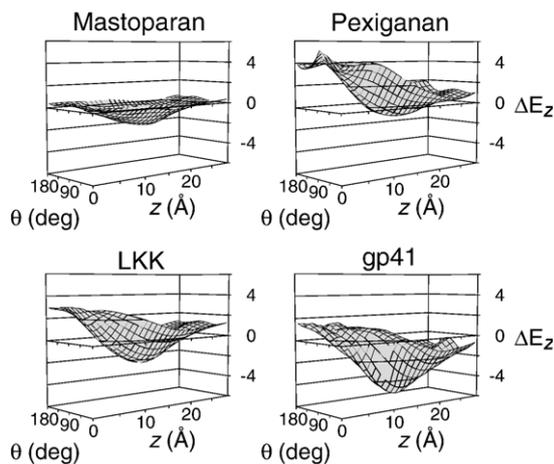


Figure 8. Energy landscapes for a series of interfacial amphiphilic peptides. Energy minima are expected and observed at the membrane interface (around $z = 13$ Å and $\theta = 90^\circ$).

cellular membrane. mastoparan X and pexiganan had a relatively flat minimum predictive of a surface-absorbed state. The difference between the surface-absorbed state and the inserted state was relatively small, which is consistent with data that indicate that peptides of this class adopt a surface-absorbed state at low peptide/lipid ratios, but that they can associate to form transmembrane species as the concentration of the peptide in the bilayer is increased.^{14,40,42–44} The minimal peptide (LKKLL-KL)₂ is too short to span the bilayer, and hence has a strong preference for the surface-absorbed state. As shown in Figure 8, only a single minimum was always observed in all cases, the most favorable orientation being interfacial, near 15 Å from the center of the bilayer and nearly parallel with the membrane plane. A shallower landscape is observed for mastoparan X, which has been observed in both transmembrane and parallel orientations in lipid bilayers. In excellent agreement with experiment,⁴¹ gp41 is predicted to adopt a surface-absorbed state, with the hydrophobic face defined by the Leu and Trp side-chain oriented toward the membrane interface. These results are consistent with the experimental observation, and indicate that the E_z potential can be applied to sequences characterized by different degrees of hydrophobicity and amphiphilicity.

Comparison with another empirical membrane potential

A related potential was reported recently by Ulmschneider *et al.*³⁶ While methodologically distinct, the two potentials are in good agreement overall, and here we will highlight similarities and difference. One of the most notable differences in the derivation of the potentials is the fact that Ulmschneider *et al.* treated the cytoplasmic and extracellular orientations independently. This is advantageous to

model residue propensities that are not symmetrical across the two sides of the membrane. The majority of the distributions, however, are very symmetrical (the most notable exception being Arg and, to a lesser extent, the other polar residues and Tyr), and the distinction comes at the cost of a reduced count size for an already limited structure database.

Ulmschneider *et al.* used different equation forms for their potential, i.e. a single Gaussian fit for all hydrophobic amino acids, and a sum of two Gaussians for the polar residues. Generally one is centered to produce a maximum near $z = 0$, and the second modulates a minimum on the cytosolic side of the membrane. Our method used a sigmoidal form for polar and non-polar amino acids, with the curves continuing symmetrically with mirror shape at $z = 0$ (Figure 2). Compared to a Gaussian, this allows variable steepness of the transition while still maintaining a relatively flat profile in the hydrophobic center of the bilayer. For Trp and Tyr profiles, both methods use a Gaussian function centered in the headgroup region. However, Ulmschneider *et al.* differentiation between cytosolic and extracellular data allowed them to fit pronounced asymmetry in the case of Tyr.

The major difference between the two methods is His. In our work, this amino acid follows a sigmoidal distribution that favors partition into water, which is very similar to all other polar residues. In Ulmschneider's potential, His behaves like Tyr and Trp, with Gaussian curves that have minima in the interfacial region. This significant disagreement might depend on differences in the databases. In our case, the raw counts do not show a decrease in frequency of the residue into the aqueous region, which is apparent in the Ulmschneider *et al.* distribution, particularly on the cytosolic side, where there is an increased tendency of His to penetrate deeper into the membrane.

Effect of lipid-exposure on E_z atom-based potentials

Different residue types have distinct position-dependent propensities to occur on the surface versus the interiors of membrane proteins. Thus, it is important to determine how the degree of exposure affects the E_z potentials of the specific side-chains. To facilitate this analysis, we consider the E_z profiles for specific atom types rather than the full side-chains. By considering specific atom types, we can better account for snorkeling, and it allows examination of the effect of burial on an atom-by-atom basis. A set of 52 atom types were chosen to represent the backbone and side-chain atoms. In general, the propensity profiles have similar values of Z_{mid} for the buried versus the exposed residues, although the exposed atom groups generally show larger differences in propensities as a function of z (and hence a larger magnitude for ΔE_0).

Figure 9 compares plots for typical buried versus exposed atomic groups, including representative backbone, polar and apolar side-chains. Amide

groups in protein α -helices have very limited accessibility to a probe of this size. Thus, exposed amide nitrogen atoms have a very low propensity to occur near the center of the bilayer, while the distribution for buried amide oxygen atoms is nearly isotropic with respect to z . Analysis of the curve for the exposed amide nitrogen atoms provided a value of $\Delta E_0 = 1.6$ kcal/mol, in agreement with the conclusion reached by Wimley and White,¹ that transfer of an amide group from water to a membrane is highly unfavorable. Most of the E_z profiles for other polar and non-polar groups tend to show a strong dependence on their degree of burial (Figure 9).

One exception is the terminal guanidino atoms of Arg, which are very infrequent near the center of the bilayer in both the exposed and the buried positions. Because of the limited number of counts, we can place only a lower limit for ΔE_0 of approximately 1.5 kcal/mol for both buried and exposed atoms, indicating that this is among the most difficult groups to bring fully to the center of the bilayer. The steepness of the curves and the value of Z_{mid} were not significantly different from that observed for the C^β of Arg.

The effect of snorkeling can be seen by comparing the position of the center of mass of the apolar aromatic six-membered ring of Trp to the corresponding position of the more polar indole N of the same side-chain. On average, the indole NH occurs $2.5(\pm 0.5)$ Å further away from the center of the bilayer than the center of the C6 ring.

While the atom-based potential captures details that are averaged in the residue-based E_z profiles (rotameric conformations, solvent exposure), the two are very similar overall. The atom-based potential could be more advantageous for detailed design applications that involve side-chain mobility, while for "lower resolution" purposes, such as rigid body searches, the residue-based E_z would be more

readily applicable, and appropriate when sequence but not structural information is available.

Conclusions

The presented empirical residue-based potential captures the essential aspects of the thermodynamics of membrane insertion. With the steady improvement of size and average resolution of the membrane protein structural database, some of the present limitations (such as symmetric treatment of both sides of the membrane, and the lack of inclusion of chemical details such as rotameric states, tertiary interactions and protonation states) will be overcome, presumably allowing greater accuracy. However, in the present form, the pseudo-energies associated with the individual types are remarkably consistent with the current understanding of partitioning of helices in the hydrophobic and headgroup regions. The potential performs well in comparison to data obtained experimentally. The results of water-membrane^{28,29} and water-interface³⁰ transfer energy scales are recapitulated with excellent correlation by the potential. Translocon-mediated apparent free energies of insertion³ are also in good agreement with our predictions. The E_z potential can orient single helices of varying length and composition with rigid body searches at global minima depths and angles that are sensitive to hydrophobic matching for transmembrane segments. The potential can discriminate and orient amphiphilic interface associate peptides correctly. The orientation algorithm can automatically place structures of multi-span membrane protein at orientations with an accuracy comparable to that of procedures that require guided identification of the transmembrane segments. The potential provides chemical insights, and the derived method could prove very useful for a variety of applications in the analysis and design of membrane proteins.

Methods

Database

The initial database of twenty four membrane proteins (PDB ID: 1c3w, 1e12, ehk, 1eul, 1fx8, 1h2s, 1iwg, 1j4n, 1k4c, 1kb9, 1kf6, 1kpl, 1kqf, 1l7v, 1l9h, 1m3x, 1m56, 1msl, 1nek, 1ocr, 1pv6, 1pw4, 1q16, 1qla)⁴⁵ was prepared with a procedure similar to that described by Rees and co-workers.⁴⁶ Proteins were aligned along the z -axis perpendicular to the membrane plane, by calculating the projection of each transmembrane helix on the x - y plane as described.⁴⁶ The sum of the magnitude projections was minimized as a function of the three rotational degrees of freedom about its center of mass. Proteins were then centered in the membrane along the z -axis by calculating the average of the z coordinates of all non-backbone carbon atoms. This center of mass was then set to zero by translating the protein along the z -axis.

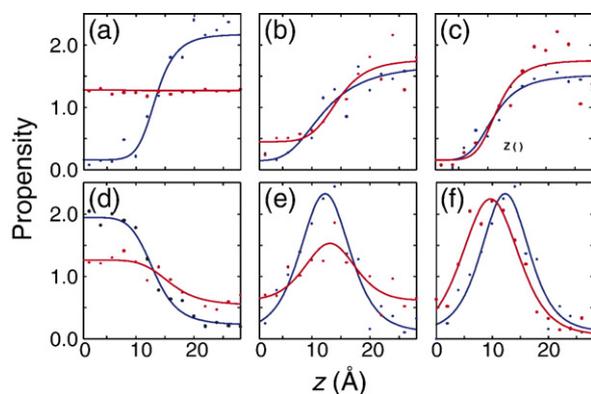


Figure 9. Propensity curves for specific atomic groups divided into solvent-accessible (red) and buried (black). (a) Backbone amide nitrogen. (b) Side-chain amide nitrogen. (c) Arg N^1/N^2 . (d) Phe ring carbon atoms. (e) Trp $N^{\epsilon 1}$. (f) Trp six-membered ring (blue) versus the indole $N^{\epsilon 1}$ (red).

Derivation of the potential

The derivation of the pseudo-potential was performed (as described in further details in the results section) by subdividing the aligned structural data into bins comprising 2 Å in the z dimension and counting the occurrence of each residue type (position taken at the C^β , or C^α for Gly). Due to the limited size of the database, we do not differentiate the cytoplasmic from the extra-cellular sides of the protein. Thus, z is always assumed to be positive, and the potential is symmetric across the two faces of the bilayer. The occurrence data was converted to a propensity using the equation:

$$P_{\text{res,bin}} = \frac{n_{\text{res,bin}}}{n_{\text{tot}} f_{\text{res}} f_{\text{bin}}} \quad (4)$$

Where $P_{\text{res,bin}}$ is the propensity for a given residue in a given bin, $n_{\text{res,bin}}$ is the bin's count for the residue, n_{tot} is the total count of all residue, f_{res} is the relative abundance of the residue, and f_{bin} is the total number of counts in a given bin (summed over all residue types) divided by $n_{\text{res,bin}}$. The pseudo-energy potential was obtained by fitting with non-linear least-squares analysis equations (2) and (3) ((2) and (4) for Trp and Tyr) with data defined between $z=0$ and 30 Å. The logarithmic form of the equation was used in fitting to avoid compression of the data as P_z approaches 0.

Calculation of state energies

The calculation of the E_z for a given structure was performed by summing the values of the individual residues (using the residue-based scale) for a given orientation. Mapping of the conformational energy landscapes was obtained by varying three parameters, the z coordinate and two orthogonal rotations on a grid of 9048 points in one translational (z translation from 0 Å to 28 Å with a 1 Å step) and two angular dimensions (0° to 180° tilting of the helix with respect to the membrane plane with a 15° step, and a 360° rotation around the helical axis with a 15° step). It was followed by a local Monte Carlo minimization around the grid search minimum.

Calculation of insertion pseudo free energies

The transmembrane insertion free energies (ΔG_{Ez}) were calculated using a grid of 9048 subdivided into three ensembles: the transmembrane segments that span the virtual bilayer (termini on opposite ends at $>13\text{Å}$ from the membrane center); embedded segments (both termini in the hydrocarbon region); and water exposed +interfacial and partially inserted, which comprised the remainder of the orientations.

The ensemble energy of a state was calculated from the difference of average state energy and entropy:

$$\Delta G_{Ez} = E_{Ez}^{\text{TM}} - E_{Ez}^{\text{WI}} + T(S_{Ez}^{\text{TM}} - S_{Ez}^{\text{WI}}) \quad (5)$$

$$E_{Ez}^{\text{state}} = \sum_i p_i E_i \quad (6)$$

$$S_{Ez}^{\text{state}} = \sum_i p_i \ln(p_i) \quad (7)$$

The probability p_i of each conformation i is:

$$p_i = \frac{e^{-\frac{E_i}{RT}} \int_{\theta_{\text{start}}}^{\theta_{\text{end}}} \sin\theta}{\sum_i e^{-\frac{E_i}{RT}} \int_{\theta_{\text{start}}}^{\theta_{\text{end}}} \sin\theta} \quad (8)$$

with θ the angle formed by the helical axis with the bilayer normal; the integral factor is a correction for the probability in the bin interval θ_{start} to θ_{end} , which is associated with the probability of each tilt state and follows a $\sin\theta$ law.

Web resources

The E_z potential parameters and utility programs are available on the internet[‡].

Acknowledgements

The authors acknowledge primary support from NIH (GM60610) and from the MRSEC program of NSF. The authors thank Doug Matcalf and Greg Caputo for helpful comments and discussion.

References

- White, S. H. & Wimley, W. C. (1999). Membrane protein folding and stability: physical principles. *Annu. Rev. Biophys. Biomol. Struct.* **28**, 319–365.
- Deber, C. M., Liu, L. P. & Wang, C. (1999). Perspective: peptides as mimics of transmembrane segments in proteins. *J. Pept. Res.* **54**, 200–205.
- Hessa, T., Kim, H., Bihlmaier, K., Lundin, C., Boekel, J., Andersson, H. *et al.* (2005). Recognition of transmembrane helices by the endoplasmic reticulum translocator. *Nature*, **433**, 377–381.
- Landolt-Marticorena, C., Williams, K. A., Deber, C. M. & Reithmeier, R. A. (1993). Non-random distribution of amino acids in the transmembrane segments of human type I single span membrane proteins. *J. Mol. Biol.* **229**, 602–608.
- Arkin, I. T. & Brunger, A. T. (1998). Statistical analysis of predicted transmembrane alpha-helices. *Biochim. Biophys. Acta*, **1429**, 113–128.
- Ulmschneider, M. B. & Sansom, M. S. (2001). Amino acid distributions in integral membrane protein structures. *Biochim. Biophys. Acta*, **1512**, 1–14.
- Yarov-Yarovoy, V., Schonbrun, J. & Baker, D. (2006). Multipass membrane protein structure prediction using Rosetta. *Proteins: Struct. Funct. Genet.* **62**, 1010–1025.
- Caputo, G. A. & London, E. (2003). Cumulative effects of amino acid substitutions and hydrophobic mismatch upon the transmembrane stability and conformation of hydrophobic alpha-helices. *Biochemistry*, **42**, 3275–3285.
- Hessa, T., White, S. H. & von Heijne, G. (2005). Membrane insertion of a potassium-channel voltage sensor. *Science*, **307**, 1427.

[‡] <http://degrado.med.upenn.edu/ez>

10. Long, S. B., Campbell, E. B. & Mackinnon, R. (2005). Voltage sensor of Kv1.2: structural basis of electromechanical coupling. *Science*, **309**, 903–908.
11. Huang, H. W., Chen, F. Y. & Lee, M. T. (2004). Molecular mechanism of Peptide-induced pores in membranes. *Phys. Rev. Letters*, **92**, 198304-1–198304-3.
12. DeGrado, W. F. (1988). Design of peptides and proteins. *Advan. Protein Chem.* **39**, 51–124.
13. Shai, Y. (2002). From innate immunity to de-novo designed antimicrobial peptides. *Curr. Pharm. Des.* **8**, 715–725.
14. Zasloff, M. (2002). Antimicrobial peptides of multicellular organisms. *Nature*, **415**, 389–395.
15. Chamberlain, A. K., Lee, Y., Kim, S. & Bowie, J. U. (2004). Snorkeling preferences foster an amino acid composition bias in transmembrane helices. *J. Mol. Biol.* **339**, 471–479.
16. Chamberlain, A. K. & Bowie, J. U. (2004). Analysis of side-chain rotamers in transmembrane proteins. *Bioophys. J.* **87**, 3460–3469.
17. Senes, A., Engel, D. E. & DeGrado, W. F. (2004). Folding of helical membrane proteins: the role of polar, GxxxG-like and proline motifs. *Curr. Opin. Struct. Biol.* **14**, 465–479.
18. Li, S. C., Goto, N. K., Williams, K. A. & Deber, C. M. (1996). Alpha-helical, but not beta-sheet, propensity of proline is determined by peptide environment. *Proc. Natl Acad. Sci. USA*, **93**, 6676–6681.
19. Nilsson, I., Saaf, A., Whitley, P., Gafvelin, G., Waller, C. & von Heijne, G. (1998). Proline-induced disruption of a transmembrane alpha-helix in its natural environment. *J. Mol. Biol.* **284**, 1165–1175.
20. Monne, M., Nilsson, I., Elofsson, A. & von Heijne, G. (1999). Turns in transmembrane helices: determination of the minimal length of a “helical hairpin” and derivation of a fine-grained turn propensity scale. *J. Mol. Biol.* **293**, 807–814.
21. Killian, J. A. & von Heijne, G. (2000). How proteins adapt to a membrane-water interface. *Trends Biochem. Sci.* **25**, 429–434.
22. Wallin, E., Tsukihara, T., Yoshikawa, S., von Heijne, G. & Elofsson, A. (1997). Architecture of helix bundle membrane proteins: an analysis of cytochrome c oxidase from bovine mitochondria. *Protein Sci.* **6**, 808–815.
23. Senes, A., Gerstein, M. & Engelman, D. M. (2000). Statistical analysis of amino acid patterns in transmembrane helices: the GxxxG motif occurs frequently and in association with beta-branched residues at neighboring positions. *J. Mol. Biol.* **296**, 921–936.
24. Wiener, M. C. & White, S. H. (1992). Structure of a fluid dioleoylphosphatidylcholine bilayer determined by joint refinement of x-ray and neutron diffraction data. III. Complete structure. *Biophys. J.* **61**, 437–447.
25. Segrest, J. P., De Loof, H., Dohlman, J. G., Brouillette, C. G. & Anantharamaiah, G. M. (1990). Amphipathic helix motif: classes and properties. *Proteins: Struct. Funct. Genet.* **8**, 103–117.
26. de Planque, M. R., Kruijtzter, J. A., Liskamp, R. M., Marsh, D., Greathouse, D. V., Koeppe, R. E., 2nd *et al.* (1999). Different membrane anchoring positions of tryptophan and lysine in synthetic transmembrane alpha-helical peptides. *J. Biol. Chem.* **274**, 20839–20846.
27. Seshadri, K., Garemyr, R., Wallin, E., von Heijne, G. & Elofsson, A. (1998). Architecture of beta-barrel membrane proteins: analysis of trimeric porins. *Protein Sci.* **7**, 2026–2032.
28. Eisenberg, D., Weiss, R. M. & Terwilliger, T. C. (1982). The helical hydrophobic moment: a measure of the amphiphilicity of a helix. *Nature*, **299**, 371–374.
29. Jayasinghe, S., Hristova, K. & White, S. H. (2001). Energetics, stability, and prediction of transmembrane helices. *J. Mol. Biol.* **312**, 927–934.
30. Wimley, W. C. & White, S. H. (1996). Experimentally determined hydrophobicity scale for proteins at membrane interfaces. *Nature Struct. Biol.* **3**, 842–848.
31. Yohannan, S., Yang, D., Faham, S., Boulting, G., Whitelegge, J. & Bowie, J. U. (2004). Proline substitutions are not easily accommodated in a membrane protein. *J. Mol. Biol.* **341**, 1–6.
32. Strop, P., Bass, R. & Rees, D. C. (2003). Prokaryotic mechanosensitive channels. *Advan. Protein Chem.* **63**, 177–209.
33. Ren, J., Lew, S., Wang, J. & London, E. (1999). Control of the transmembrane orientation and interhelical interactions within membranes by hydrophobic helix length. *Biochemistry*, **38**, 5905–5912.
34. Park, S. H. & Opella, S. J. (2005). Tilt angle of a transmembrane helix is determined by hydrophobic mismatch. *J. Mol. Biol.* **350**, 310–318.
35. Ben-Tal, N., Ben-Shaul, A., Nicholls, A. & Honig, B. (1996). Free-energy determinants of alpha-helix insertion into lipid bilayers. *Biophys. J.* **70**, 1803–1812.
36. Ulmschneider, M. B., Sansom, M. S. & Di Nola, A. (2005). Properties of integral membrane protein structures: derivation of an implicit membrane potential. *Proteins: Struct. Funct. Genet.* **59**, 252–265.
37. Smith, S. O., Jonas, R., Braiman, M. & Bormann, B. J. (1994). Structure and orientation of the transmembrane domain of glycophorin A in lipid bilayers. *Biochemistry*, **33**, 6334–6341.
38. Smith, S. O., Eilers, M., Song, D., Crocker, E., Ying, W., Groesbeek, M. *et al.* (2002). Implications of threonine hydrogen bonding in the glycophorin A transmembrane helix dimer. *Biophys. J.* **82**, 2476–2486.
39. MacKenzie, K. R., Prestegard, J. H. & Engelman, D. M. (1997). A transmembrane helix dimer: structure and implications. *Science*, **276**, 131–133.
40. Whiles, J. A., Brasseur, R., Glover, K. J., Melacini, G., Komives, E. A. & Vold, R. R. (2001). Orientation and effects of mastoparan X on phospholipid bicelles. *Biophys. J.* **80**, 280–293.
41. Chou, J. J., Kaufman, J. D., Stahl, S. J., Wingfield, P. T. & Bax, A. (2002). Micelle-induced curvature in a water-insoluble HIV-1 Env peptide revealed by NMR dipolar coupling measurement in stretched polyacrylamide gel. *J. Am. Chem. Soc.* **124**, 2450–2451.
42. Hori, Y., Demura, M., Iwadate, M., Ulrich, A. S., Niidome, T., Aoyagi, H. & Asakura, T. (2001). Interaction of mastoparan with membranes studied by ¹H-NMR spectroscopy in detergent micelles and by solid-state ²H-NMR and ¹⁵N-NMR spectroscopy in oriented lipid bilayers. *Eur. J. Biochem.* **268**, 302–309.
43. Munster, C., Spaar, A., Bechinger, B. & Salditt, T. (2002). Magainin 2 in phospholipid bilayers: peptide orientation and lipid chain ordering studied by X-ray diffraction. *Biochim. Biophys. Acta*, **1562**, 37–44.
44. Bechinger, B., Ruyschaert, J. M. & Goormaghtigh, E.

- (1999). Membrane helix orientation from linear dichroism of infrared attenuated total reflection spectra. *Biophys. J.* **76**, 552–563.
45. Adamian, L., Nanda, V., DeGrado, W. F. & Liang, J. (2005). Empirical lipid propensities of amino acid residues in multispans alpha helical membrane proteins. *Proteins Struct. Funct. Genet.* **59**, 496–509.
46. Spencer, R. H. & Rees, D. C. (2002). The alpha-helix and the organization and gating of channels. *Annu. Rev. Biophys. Biomol. Struct.* **31**, 207–233.

Edited by J. Bowie

(Received 16 June 2006; received in revised form 28 August 2006; accepted 3 September 2006)
Available online 12 September 2006