Mapping the Energy Surface of Transmembrane Helix-Helix Interactions

Jaume Torres, Andreas Kukol, and Isaiah T. Arkin
Department of Biochemistry, Cambridge Centre for Molecular Recognition, University of Cambridge, Cambridge CB2 1GA, United Kingdom

ABSTRACT Transmembrane helices are no longer believed to be just hydrophobic segments that exist solely to anchor proteins to a lipid bilayer, but rather they appear to have the capacity to specify function and structure. Specific interactions take place between hydrophobic segments within the lipid bilayer whereby subtle mutations that normally would be considered innocuous can result in dramatic structural differences. That such specificity takes place within the lipid bilayer implies that it may be possible to identify the most favorable interaction surface of transmembrane α-helices based on computational methods alone, as shown in this study. Herein, an attempt is made to map the energy surface of several transmembrane helix-helix interactions for several homo-oligomerizing proteins, where experimental data regarding their structure exist (glycophorin A, phospholamban, *Influenza* virus A M2, *Influenza* virus C CM2, and HIV *vpu*). It is shown that due to symmetry constraints in homo-oligomers the computational problem can be simplified. The results obtained are mostly consistent with known structural data and may additionally provide a view of possible alternate and intermediate configurations.

INTRODUCTION

Membrane proteins occupy a peculiar position in biological sciences. On the one hand it is widely recognized that they are by far the most biomedically important family of proteins, serving as the targets for the majority of pharmaceuticals, while on the other hand they have been resistive subjects to structural probing. Compounded by their genomic abundance, any computational tool that would provide structural insight into these proteins would obviously be useful.

Detailed interactions between transmembrane α-helices have been the subject of numerous studies (Lemmon and Engelman, 1994a,b). In many instances the specificity of the interaction is exquisite, such as in the case of the dimerizing glycophorin A (Lemmon et al., 1992, 1994) and pentamerizing phospholamban (Arkin et al., 1994). In both of these cases, subtle mutations (e.g., Ile to Leu) have resulted in abolition of oligomerization. Such specificity does lend hope to computational efforts aimed at determining the helix-helix interface based solely on energy calculations (i.e., molecular dynamics and energy minimization).

Molecular dynamics simulations of transmembrane α-helical bundles have been reported by many groups, and in general can be divided into two categories: 1) simulations in which a single starting position is subjected to a long molecular dynamics simulation usually in a fully hydrated lipid bilayer (Tieleman et al., 1999; Belohorcova et al., 1997; Sansom, 1998). Such simulations can be used to gain insight into multiple aspects of the protein under investigation including, among others, mechanism (Woolf, 1997) and stability (Woolf and Tychko, 1998; Forrest et al., 1999). 2) Multiple short in vacuo molecular dynamics simulations at different starting positions (Treutlein et al., 1992; Adams et al., 1995, 1996; Duneau et al., 1999), are used to identify possible structures for a particular helix bundle through comprehensive sampling of the interaction space.

In a bundle that contains n helices, $3n$ parameters are needed to describe the overall structure (see Fig. 1): 1) the tilt angle with respect to the bundle axis, β_i, related to the commonly used crossing angle Ω (Chothia et al., 1981); 2) the rotational angle about the helix director, ϕ_i, which defines which side of helix i is facing toward the bundle core; and 3) the helix register, r_i, which defines the relative vertical position of helix i. Thus, to search through the bundle configuration space efficiently, one would only have to vary the above three parameters for each helix (see Note 1 at end of text). Due to these three parameters, the CPU time will always increase exponentially in proportion to the oligomerization number. This increase is further compounded purely on the basis of the increased size of a larger complex.

Homo-oligomers, in contrast, offer an attractive system to investigate, since the number of degrees of freedom is reduced dramatically due to symmetry constraints. Because the register for all helices will be identical irrespective of the oligomerization number, only two parameters are needed to adequately describe the bundle configuration: a common helix tilt, $\beta = \beta_{i,j,\ldots,n}$ and a common rotational pitch angle, $\phi = \phi_{i,j,\ldots,n}$ about the helix axis. It is therefore possible to generate a two-dimensional surface depicting the variation of bundle energy as a function of β and ϕ.

The groups of Brünger and (Adams et al., 1995) and Genest (Sajot and Genest, 2000) have both undertaken an

Received for publication 28 January 2000 and in final form 3 May 2001.

Address reprint requests to (present address) Dr. Isaiah T. Arkin, Dept. of Biological Chemistry, Institute of Life Sciences, The Hebrew University, Givat-Ram, Jerusalem 91904, Israel. Tel.: 972-2-658-4329; Fax: 972-2-698-4329; E-mail: arkin@cc.huji.ac.il.

Andreas Kukol's present address is Dept. of Biological Sciences, University of Warwick, Coventry CV4 7AL, UK.

© 2001 by the Biophysical Society

0006-3495/01/11/2681/12 $2.00

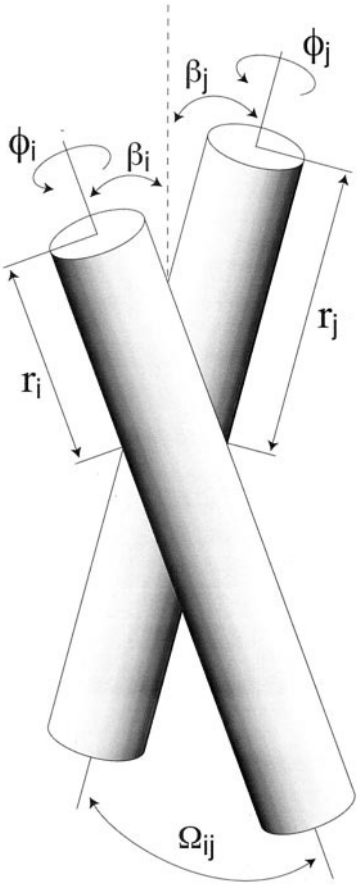

FIGURE 1 Schematic diagram of a helix bundle, depicting the parameters that define the bundle configuration. See text for details.

approach in which the sampling of the interaction scheme is focused on variation of the rotational pitch angle. Multiple starting positions at rotational increments of 10° (Adams et al., 1995, 1996) or 18° (Sajot and Genest, 2000) are subjected to a simulated annealing molecular dynamics protocol aimed at relaxing the structures. The endpoints of the simulations are then compared to find local energy minima to which multiple starting structures have converged. Variation of the tilt angle is achieved through starting positions of right- and left-handed crossing angles (corresponding to tilt angles of $\mp 25°$, respectively). This method, although instructive in terms of identifying possible candidate structures, is limited in its description of the overall energy surface. Furthermore, it relies upon the convergence of different starting structures to particular energy minima, which at times can be minimal without the aid of experimental data (Kukol et al., 1999).

Herein, a different approach is undertaken whereby the entire protein-protein interaction surface is mapped using several homo-oligomeric transmembrane α-helical bundles with available structural information regarding their native configuration. Instead of relying upon structure convergence, the energy of the bundle is calculated for every rotational pitch angle and tilt angle in increments of 1°. In this way a complete energy map of the protein-protein interactions is obtained.

One possible shortcoming of this approach, which is due to CPU time limitations, is the exclusion of a solvated lipid bilayer from the calculations. In this sense, although it is conceivable that contributions from protein-lipid interaction could modulate the energy surface in such a way that these calculations would result in a somewhat different representation, we believe that this is unlikely, and that the energetics of protein-protein interactions are the most critical driving force in the oligomerization process. That this is a reasonable assumption is justified based on the fact that in the majority of cases, the in vacuo calculations performed here, using a dielectric constant $\epsilon = 1$, do predict the presence of a large energy trough where the predicted structure should be located.

Finally, the energy surfaces calculated may provide insight into the stability of the structure, which would be proportional to the depth and volume of the energy basin. Furthermore, the shape of this energy trough may indicate possible folding pathways to the final structure. This point is further elaborated in the Discussion.

METHODS

Computational methods

The energy surface mapping was undertaken with a modified version of the CHI (CNS Helix Interaction) software suite (Adams et al., 1995). All calculations were performed with the parallel processing version of the Crystallography and NMR System (CNS Version 0.3) (Brunger et al., 1998) modified by Dr. Greg McMullan to run on a Hitachi SR2201, 256-node parallel computer. The OPLS parameter set with a united atom topology was used, explicitly representing all polar hydrogen and aromatic side chain atoms (Jorgensen & Tirado-Rives, 1988). All calculations were carried out in vacuo with the initial coordinates of a canonical α-helix (3.6 residues per turn). The dielectric constant was set to 1.0 to mimic the effect of a the low dielectric environment of the lipid bilayer, as previously used in similar simulations (Adams et al., 1995). It is noted that setting the dielectric constant to 2.0, on a small subset of the simulations, produced indistinguishable results. All calculations used a nonbonded cutoff of 13 Å, and a switching function was applied to van der Waals interactions between 10 Å and 12 Å.

Symmetric, canonical helical bundles were constructed (see sequences below) by replicating a helix and rotating it by $\phi = 360°/n$ about its helix axis, whereby n represents the size of the oligomers. The initial distances of the N—H· · ·O ═ C hydrogen bond was set to 2.1 Å. The helices were then radially translated from the initial position (which is consistent with the bundle axis) by a distance of 10 Å at the direction of the rotational angle that was used to rotate the helix, ϕ (the new x and y coordinates of each atom would therefore be changed by 10 cos ϕ Å and 10 sin ϕ Å, respectively).

This initial position was then used to generate multiple starting positions by changing the helix tilt angle from $-45°$ to 45°, whereby positive and negative values indicate a left- and right-handed helical bundle, respectively. Furthermore, each of these different starting positions was used as an initial point for further variation through rotation about the helix director, from 0° to 359°. This resulted in a total number of structures analyzed of $\beta \times \phi = 91 \times 360 = 32{,}760$.

TABLE 1 Sequences used in mapping the energy surface of transmembrane helix-helix interactions

Protein	Oligomeric Size	Sequence
Human glycophorin A	2	CH_3-CO-TLIIFGVMAGVIGTILLI-NH-CH_3
Influenza virus A M2	4	CH_3-CO-LVVAASIIGILHLILWIL-NH-CH_3
Influenza virus C CM2	4	CH_3-CO-YMLTLASLGLGIITMLYLLV-NH-CH_3
HIV *vpu*	5	CH_3-CO-IAIVALVVAIIIAIVVWSIVII-NH-CH_3
Human phospholamban	5	CH_3-CO-FCLILICLLLICIIVML-NH-CH_3

Each of these structures was then energy-minimized using the Powell energy minimization as implemented in CNS (Brunger et al., 1998) with the following protocol. Initially, 350 steps of minimization were undertaken with electrostatic interactions turned off and the REPEL function turned on to rapidly remove any steric clashes. Subsequently, 500 steps of standard minimization were undertaken, at the end of which the energy of the system was evaluated and recorded. Finally, a three-dimensional plot was obtained listing the energy at each of the 32,760 different structures as a function of the helix tilt angle β and rotational angle ϕ. The results were smoothed, averaging the energy at every point $E(\beta_i, \phi_j)$ according to the following equation:

$$E(\beta_i, \phi_j) = \frac{1}{(1 + 2m)^2} \sum_{k=i-m}^{i+m} \sum_{l=j-m}^{j+m} E(\beta_k, \phi_l)$$

whereby m, the smoothing factor, is equal to 1 unless otherwise indicated. The CPU time involved in each of the calculations was roughly 11,520 h per processor in the case of a tetramer when the program was run using 64 processors at a time.

Protein sequences

Several helical bundles were simulated in which experimental data pertaining to the structure was available. These proteins covered a range of different oligomeric sizes from a dimer to a pentamer. The sequences used in the simulations are given in Table 1. In each instance, to mimic a peptide bond, the amino terminus was acetylated and the carboxyl terminus was methylaminated.

Surface area calculation and helix property vectors

Surface areas were calculated using the program DSSP (Kabsch and Sander, 1983). The interaction surface area was calculated by subtracting the accessible surface area of the oligomer from that obtained by multiplying the accessible surface area of the protamer times the oligomerization number.

The hydrophobic vectors were calculated for each helix according to the geometrical average of the vector of every amino acid according to the helical periodicity using the GES scale (Engelman et al., 1986). The surface area was calculated in the same way using the values for the surface area of every amino acid using a probe with a diameter of 1.4 Å as described in Stevens and Arkin (1999).

RESULTS AND DISCUSSION

General

The purpose underlining this study was to attempt to map the energy surface of transmembrane helix-helix interactions. Such interaction between helices in the lipid bilayer are the cornerstone of the two-stage model for membrane protein folding and oligomerization proposed by Popot and Engelman (Popot et al., 1987; Popot & Engelman, 1990). To assess the relevance of the results obtained, the calculations were performed using helical bundles with known structures or where sufficient structural information is present. These include human glycophorin A (MacKenzie et al., 1997), the only homo-oligomerizing helical bundle structure that has been solved so far; *Influenza* virus A M2 H^+ channel (Kukol et al., 1999); HIV *vpu* (Kukol & Arkin, 1999), human phospholamban (Torres et al., 2000); and *Influenza* C CM2 (Kukol & Arkin, 2000), all of which have been analyzed using spatial restraints derived from site-directed dichroism (Arkin et al., 1997), a technique that provides information enabling one to determine the interacting surfaces of the helices. Also, in the case of phospholamban, exhaustive mutagenesis data exist (Arkin et al., 1994).

Human glycophorin A

Glycophorin was the first membrane protein to be sequenced in which a hydrophobic stretch of amino acids was identified. Glycophorin was also one of the first clear instances of specific oligomerization of proteins that was driven by the transmembrane domain. In fact, a chimera formed by a glycophorin transmembrane domain fused to a heterologous water-soluble protein was found to dimerize as well (Lemmon et al., 1992). The fact that dimerization was not abolished in SDS gels facilitated mutagenesis studies in which the exquisite sensitivity of particular residues within the transmembrane domain toward substitution was identified. These residues were then shown to be the first identifiable dimerization motif within a lipid bilayer (Lemmon et al., 1994). A global-search molecular dynamics study coupled with the information from the mutagenesis results was able to produce a model for the transmembrane domain (Treutlein et al., 1992; Adams et al., 1996). This model was later shown to be remarkably similar to the structure solved by NMR spectroscopy in dodecylphosphocholine detergent micelles (MacKenzie et al., 1997). The structure of the glycophorin transmembrane domain dimer is a right-handed coiled-coil in which the helices are in contact with one another via the residues identified by the earlier mutagenesis study.

The calculated surface energy of glycophorin as a function of the helix tilt β and the rotational pitch angle ϕ is shown in Figure 2. The graph depicts a large minimum centred at a tilt angle of $-23°$ and a rotational angle of 260°. Remarkably, these values are virtually identical to those obtained experimentally by solution NMR (MacKenzie et al., 1997). Note that in the case of a dimer the interhelix crossing angle, Ω, can be derived directly from the tilt angle of the helices ($\Omega = 2\beta$). The energy basin covers more than half of the surface area, and the difference between the absolute minimum (-72 kcal/mol) and maximum (-19 kcal/mol) is only 53 kcal/mol. Although the energy difference is relatively small (to that obtained for oligomers of larger order, see below), the shape of the energy trough is very smooth, pointing to a possible explanation why glycophorin is an exceptionally stable transmembrane helix dimer.

Clues to the driving force behind the large interaction energy basin of GPA may come from analyzing the helix amphipathicity and the helix lopsidedness (the preferential distribution of amino acid of similar sizes on one side of the helix). The bottom two panels in Fig. 2 depict: 1) the interaction helical wheel diagram of GPA at 3.9 amino acids per turn corresponding to a right-handed coiled-coil (MacKenzie et al., 1997) color-coded by hydrophobicity, and 2) a canonical helical wheel representing the amphipathicity and lopsidedness vectors. The average hydrophobicity of GPA, -2.4 $\Delta G_{\text{Water}\Rightarrow\text{Oil}}$ kcal/mol per residue (according to the GES scale (Engelman et al., 1986)), is nearly identical to the average observed in a large database of putative transmembrane α-helices (Arkin and Brunger, 1998). As shown previously (MacKenzie and Engelman, 1998) what is more significant is the lopsidedness. Both the amphipathicity and lopsidedness vectors are parallel and are pointing to opposite the protein-protein contact region. This side of helix is thus more polar and less bulky, due to several Gly residues, which experimental evidence has shown to be essential for dimerization (Lemmon et al., 1992). As shown below, it is the magnitude of the lopsidedness vector of GPA that distinguishes it from the other sequences analyzed.

Influenza virus A M2 and *Influenza* virus C CM2

Influenza virus A M2 and *Influenza* virus C CM2 are members of a new family of small hydrophobic viral membrane proteins, which are characterized by a single transmembrane domain sufficient for homo-oligomerization and ion channel activity (Carrasco, 1995). *Influenza* virus A M2 H^+ has been characterized extensively and its function in the virion is twofold: 1) it enables acidification of the virion upon acidification of the endosome, thereby releasing the RNA from the viral matrix proteins (Bui et al., 1996); and 2), it ensures that the acidification machinery along the exocytic pathway does not result in a pH lower than that required for hemagglutinin, the major *Influenza* spike glycoprotein, to undergo an irreversible conformational change (Lamb and Pinto, 1997). Structurally, *Influenza* virus A M2 is known to be a homotetramer in which disulfide bonds stabilize the interaction formed by the transmembrane domains (Holsinger & Lamb, 1991; Lamb et al., 1985; Sugrue & Hay, 1991). This protein is a target for the anti-influenza drug amantadine, which blocks the ion-channel activity of M2 (Belshe et al., 1989).

Much less is known about *Influenza* virus C CM2, but it is a tetramer (Hongo et al., 1994), and is assumed to be an ion channel. This assumption is based on the fact that, although M2 does not have a clear homologous sequence in *Influenza* virus C, it is of similar organization, which leads one to consider it as an ortholog of M2.

The energy surface diagrams of transmembrane helix-helix interactions for *Influenza* virus A M2 and *Influenza* virus C CM2 (depicted in Figs. 3 and 4, respectively) are dissimilar from that obtained for glycophorin A in both the energy range and other features. The difference between minimum and maximum energy, for example, is dramatically bigger in both the tetrameric structures (-32 to -295 kcal/mol) than that found in the dimeric glycophorin A (-19 to -72 kcal/mol). This results in part from the increased interaction surface found in a tetrameric structure relative to that found in a dimeric structure. For example, the interaction surface area of the glycophorin A dimer is 1008 Å^2 as opposed to 3249 Å^2 for the M2 tetramer.

The surface features found in the tetrameric structures are also distinct from those found in glycophorin. Whereas in glycophorin there was one dominating trough, whose energy minimum coincided with the actual structure, both M2 and CM2 exhibit multiple energy minima.

For example, as seen in Fig. 3, the energy surface diagram obtained for the *Influenza* virus A M2 transmembrane domain contains two large minima. The first is located at $\phi = 290°$, $\beta = 35°$, and corresponds precisely to the structure obtained from spatial restraints (Kukol et al., 1999). This minimum is at the bottom of a large and long basin, which may indicate that different tilt angles are possible. This particular energy landscape might indicate that the protein, while retaining the same rotational angle, might allow some variability in the tilt of the helices, perhaps providing a gating mechanism. The second minimum is located at the left-hand bottom of the plot and contains two regions separated by a high-energy region. The significance of this right-handed structure is not clear.

The interaction energy landscape obtained for *Influenza* virus C CM2 depicted in Fig. 4 is more complex. Multiple minima exist, the largest of which is centered at $\phi = 90°$, $\beta = 41°$. This minimum is in a large basin that contains multiple low-energy regions. The structure we have obtained by spatial restraints (Kukol and Arkin, 2000) is located in one of the smaller energy minima, at $\phi = 31°$, $\beta = 15°$.

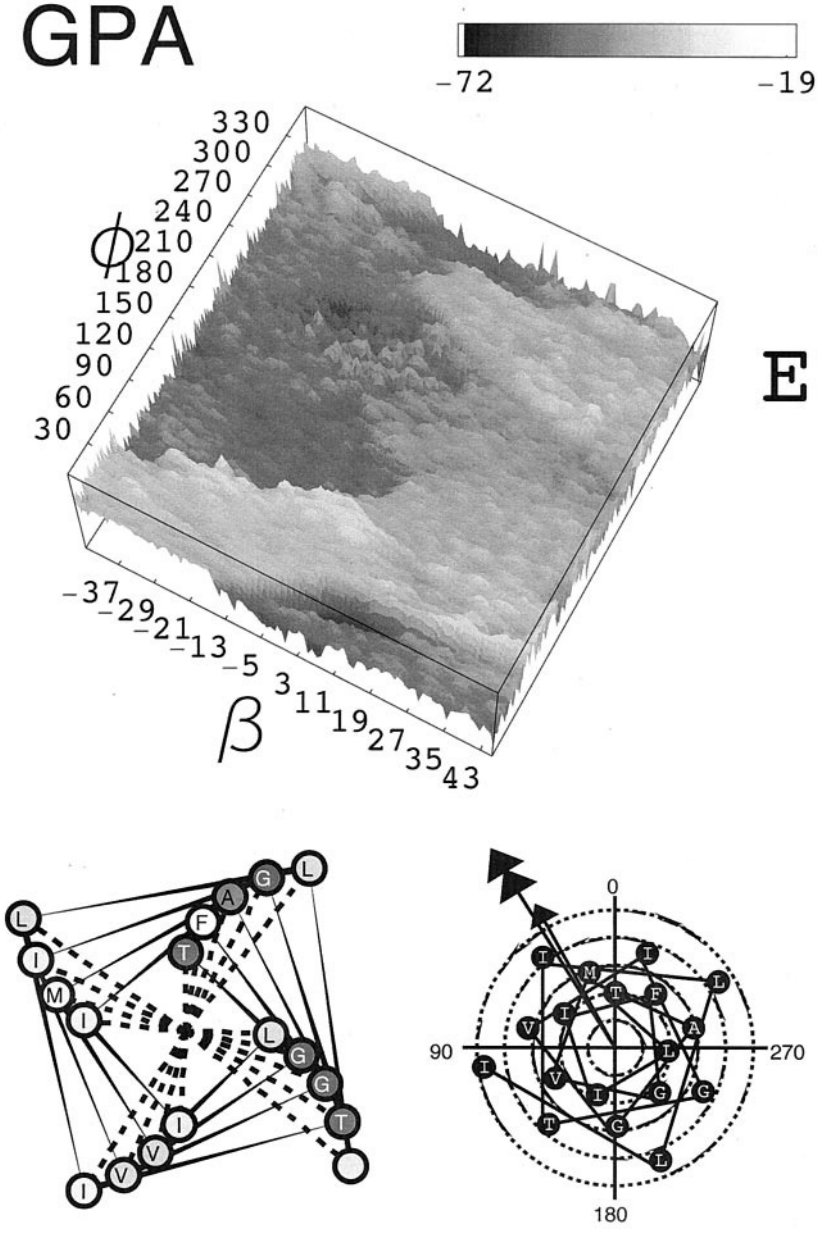

FIGURE 2 *Top two panels:* Energy, E, surface diagram of transmembrane helix-helix interactions in the dimeric human glycophorin A transmembrane domain as a function of the helix tilt, β, and the rotational pitch angle ϕ. The color index corresponds to energy in units of kcal/mol. *Middle panel:* 3.9 amino acids per turn (right-handed coiled-coil) interaction helical wheel diagram corresponding to sequence of glycophorin A simulated, indicating the rotational pitch angle ϕ. The color coding is a function of the residue hydrophobicity according to the GES scale (Engelman et al., 1986), whereby white is the most hydrophobic amino acid (Phe at -3.7 $\Delta G_{\text{Water} \Rightarrow \text{Oil}}$ kcal/mol) and black is the most hydrophilic (Tyr at 0.7 $\Delta G_{\text{Water} \Rightarrow \text{Oil}}$ kcal/mol). *Bottom panel:* Canonical helical wheel (3.6 amino acids per turn) depicting the hydrophobic (cal/mol $\Delta G_{\text{Oil} \Rightarrow \text{Water}}$) and surface area (double arrow, Å^2) helix vectors (see Material and Methods). The dotted line represents an increment of 50 cal/mol $\Delta G_{\text{Oil} \Rightarrow \text{Water}}$ and 20 Å^2.

The helices of *Influenza* virus A M2 and *Influenza* virus C CM2 differ in their amphipathicity and lopsidedness. M2 is markedly amphipathic, mostly due to the presence of the pore-lining His residue. Hence, it is not surprising that the amphipathicity vector is pointing to the opposite face of the bundle core. Whether the role of the His residue extends beyond channel gating (Wang et al., 1995) (e.g., a tetramerizing driving force) is difficult to state at present. The lopsidedness of M2 is much smaller than that found for GPA, indicating that the helix is more like a uniform cylinder.

CM2, however, is more similar to GPA in that both vectors are roughly parallel and are substantial. However, they do not point opposite to the suggested protein-protein interaction surface, indicating that other driving forces, which are not encompassed in these graphical vector repre-

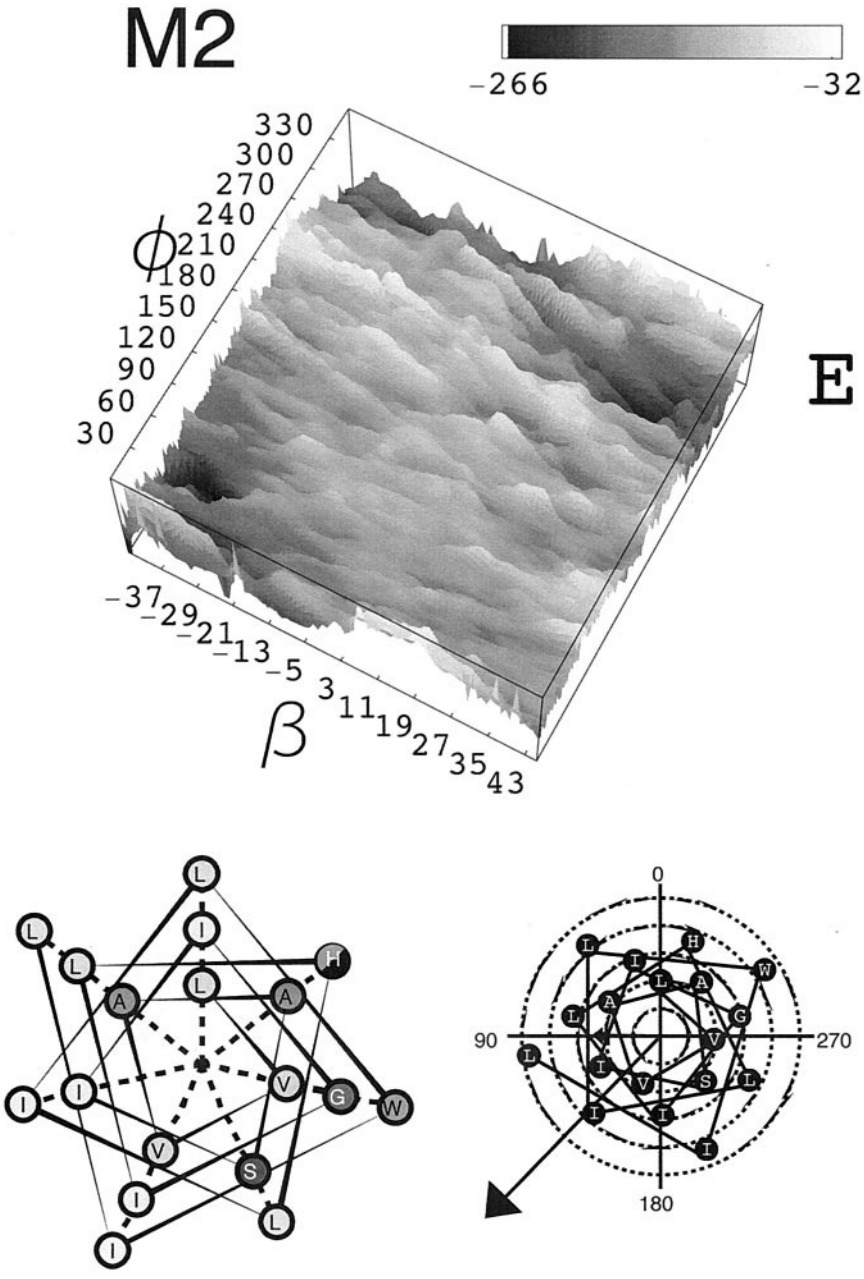

FIGURE 3 Same as Fig. 2 except for the tetrameric *Influenza* virus A M2 transmembrane domain. The interaction helical wheel is depicted with a helical periodicity of 3.5 amino acids per turn corresponding to a left-handed coiled-coil. The amino acid His is depicted in half circle gray and black, indicating its marked hydrophilicity (3 $\Delta G_{\text{Water}\Rightarrow\text{Oil}}$ kcal/mol).

sentations (e.g., maximization of van der Waals interactions (Stevens and Arkin, 1999)), are driving the interaction.

Human phospholamban and HIV *vpu*

Human phospholamban has been the subject of much research, which has focused on its function as a regulator of the cardiac sarcoplasmic reticulum Ca^{2+} ATPase (Arkin et al., 1997). Structurally, phospholamban forms pentamers that persist in SDS gels, a feature that provided a useful oligomerization assay. Saturation mutagenesis studies (Arkin et al., 1994) have pointed to the important residues in pentamerization, and have led to a model for the pentameric transmembrane α-helical bundle (Adams et al., 1995). An alternative model has been suggested (Simmerman et al., 1996), and recently we were able to show using spatial restraints (Torres et al., 2000) that the latter model was correct, proposing a structure for the complex. Functionally,

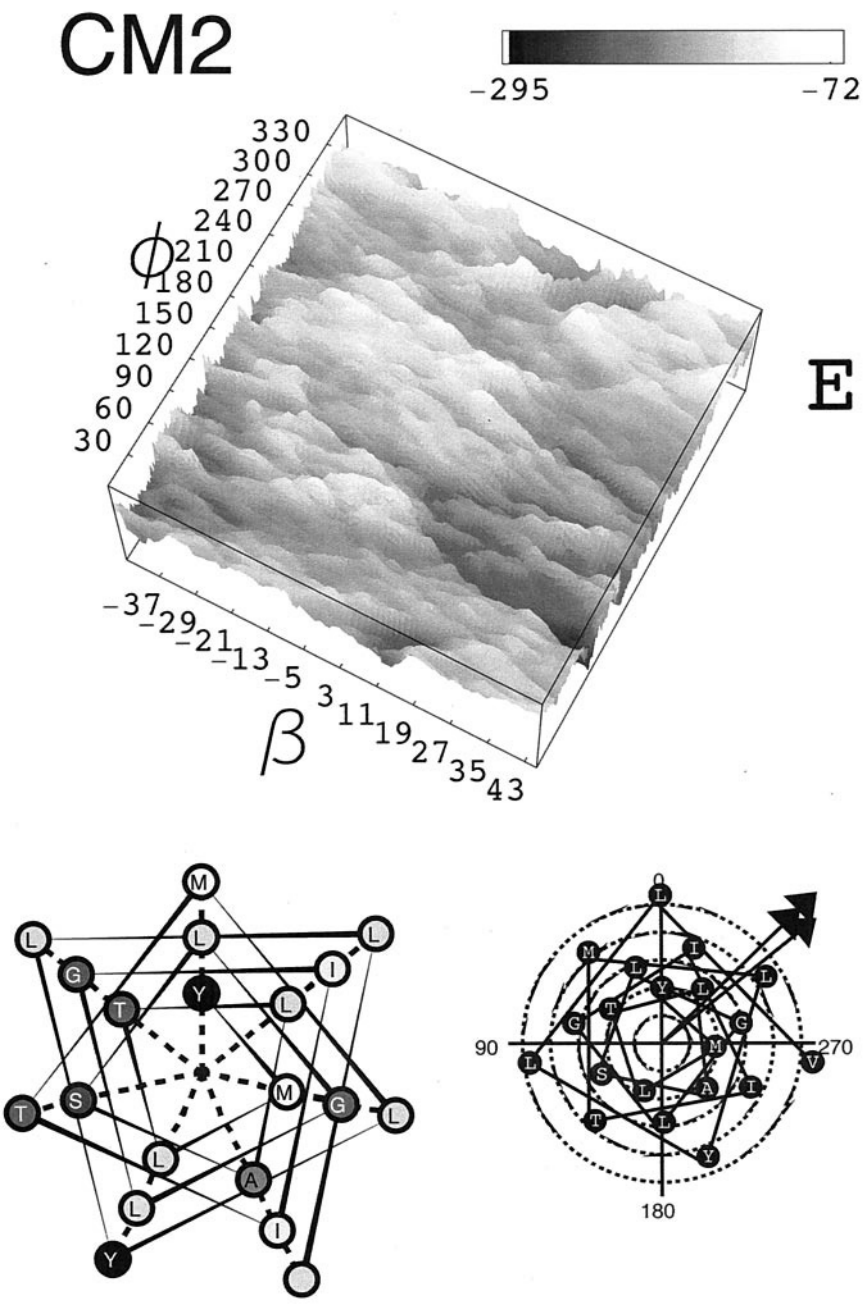

FIGURE 4 Same as Fig. 2 except for the tetrameric *Influenza* virus C CM2 transmembrane domain. The interaction helical wheel is depicted with a helical periodicity of 3.5 amino acids per turn corresponding to a left-handed coiled-coil.

some experiments suggest that phospholamban exists in a pentamer-monomer equilibrium in which the monomer is the inhibitory species, and the pentamer is stabilized by phosphorylation (the off-switch of the protein) (Karim et al., 1998; Li et al., 1998). We note, however, that other experiments do not support these hypotheses (Toyofuku et al., 1994; Kimura et al., 1998; Chu et al., 1997, 1998).

The energy surface diagram obtained for the helix-helix interactions for phospholamban is depicted in Fig. 5. As in *Influenza* virus C CM2 (Fig. 4) multiple minima exist, situated at $\phi = 300°$, $\beta = -35°$; $\phi = 140°$, $\beta = 43°$; and $\phi = 320°$, $\beta = 28°$. All of the above tilt angles are significantly higher than that obtained from spatial restraints (Torres et al., 2000), $\beta = \pm(11° \pm 7°)$ (see Note 2). In this region of the graph ($-18° \leq \beta \leq 18°$, taking into account the broadest error ranges) two prominent minima are observed: (A) $\phi = 255°$, $\beta = 8°$, and (B) $\phi = 220°$, $\beta = 12°$. These structures, A and B, correspond to the two models based on mutagenesis studies (A, Adams et al., 1995 and B, Simmerman et al., 1996). Interestingly, these two structures

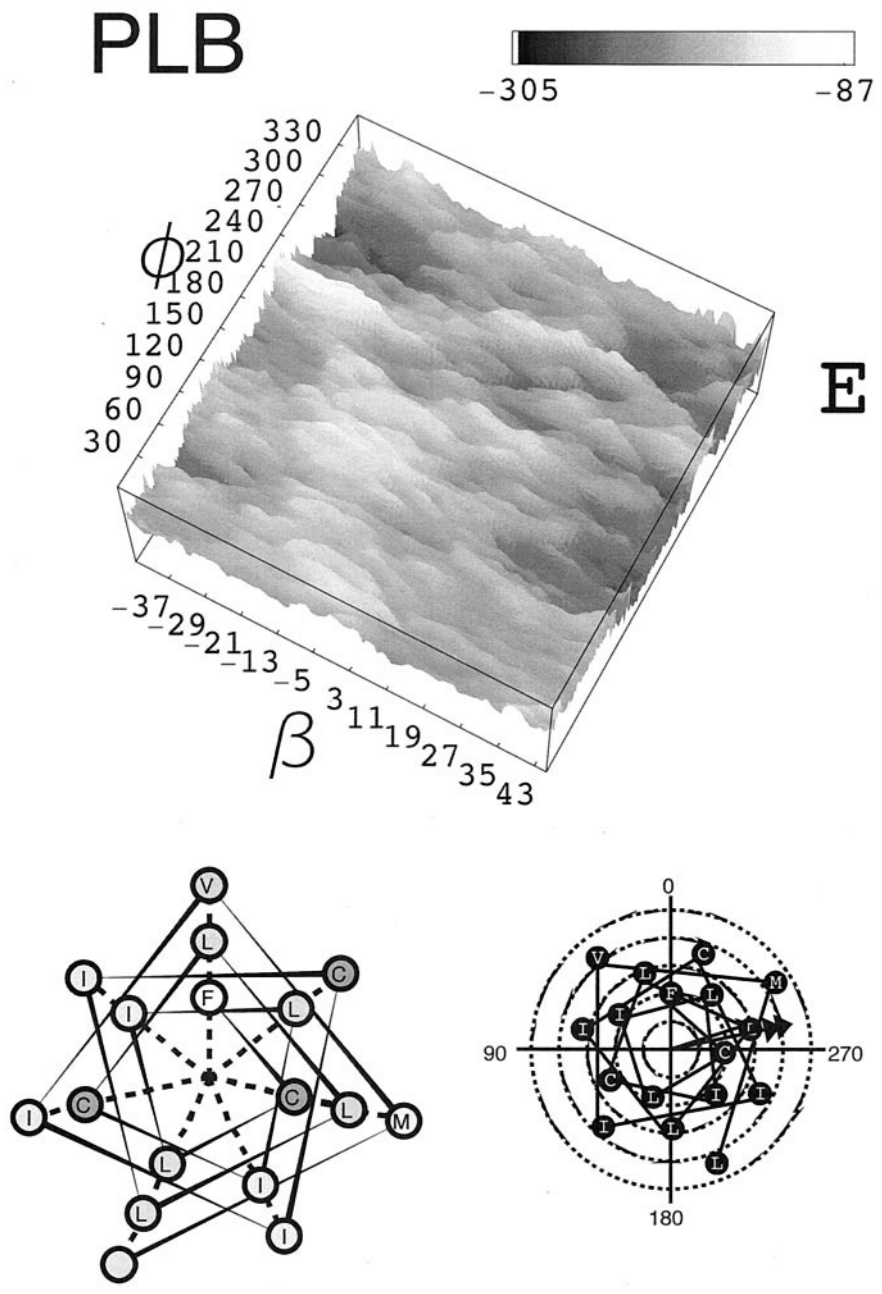

FIGURE 5 Same as Fig. 2 except for the pentameric human phospholamban transmembrane domain. The interaction helical wheel is depicted with a helical periodicity of 3.5 amino acids per turn corresponding to a left-handed coiled-coil.

are connected via a low-energy path that implies a reduction of the helix tilt. Whether interconversion indeed exists in the case of phospholamban is not known, but as phospholamban is presumed to exist in a monomer-pentamer equilibrium, one can speculate that one of these models may be less stable, and represent an intermediate in an equilibrium as follows:

$$A \rightleftharpoons B \rightleftharpoons 5\mathrm{M}$$

whereby M represents the monomer. Further experiments are undoubtedly needed to confirm this hypothesis.

The protein *vpu* is a homo-oligomeric membrane protein from HIV that has attracted attention due to its seemingly unrelated functions. One is the degradation of one of the HIV-1 co-receptor molecules, CD4 (Schubert et al., 1996; Schubert and Strebel, 1994), allowing the *env* glycoprotein to be transported to the cell surface. The other is related to virus particle release (Schubert et al., 1996). The molecular basis of these functions is unknown. It has also been shown to be an ion channel in various systems (Klimkait et al., 1990; Maldarelli et al., 1993; Schubert and Strebel, 1994; Strebel et al., 1989),

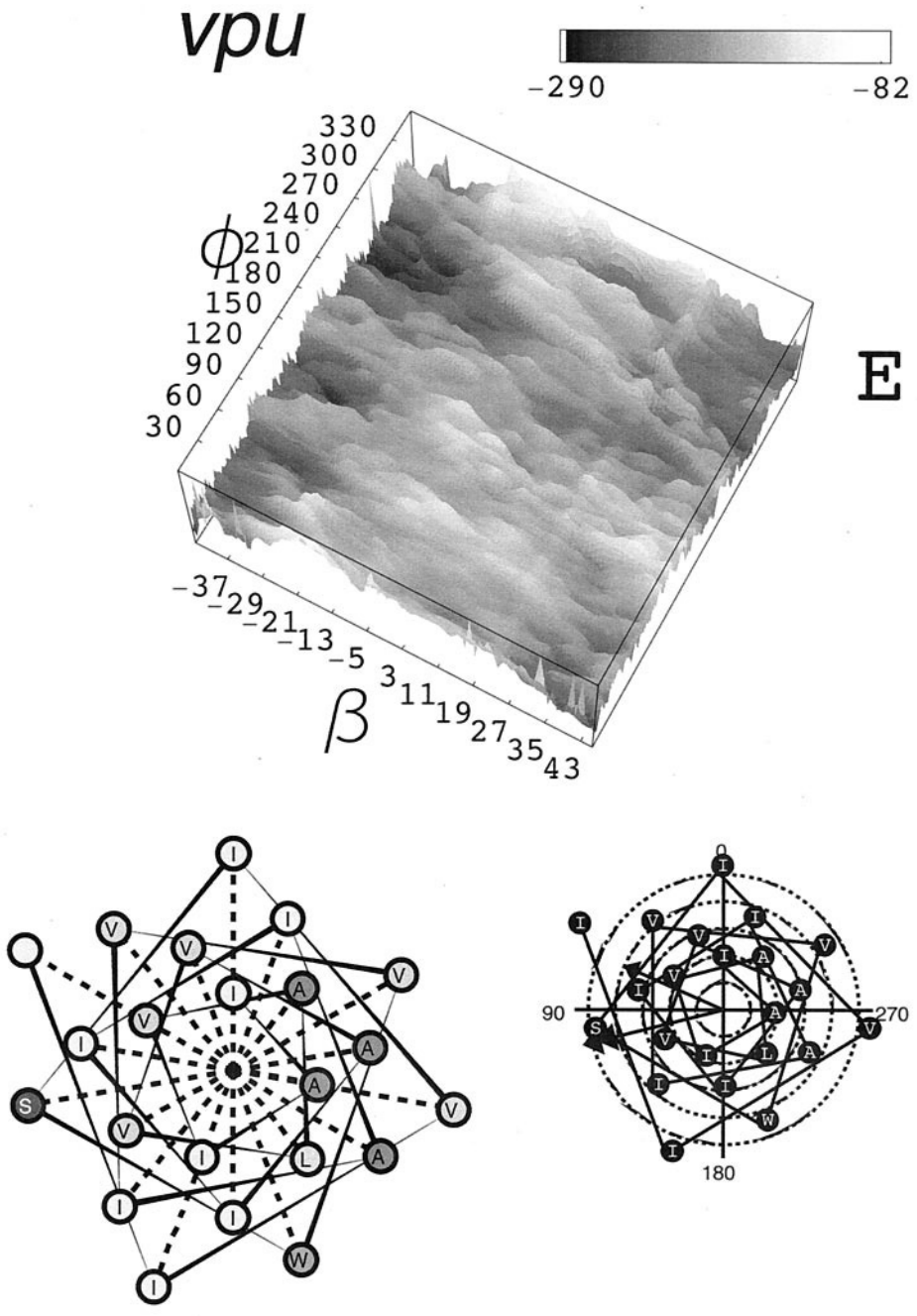

FIGURE 6 Same as Fig. 2 except for the pentameric HIV *vpu* transmembrane domain. The interaction helical wheel is depicted with a helical periodicity of 3.6 amino acids per turn corresponding to a canonical helix.

and the relation to the previous functions is still a matter of debate.

The oligomeric state of *vpu* is of at least four subunits as detected by gel electrophoresis (Maldarelli et al., 1993). Recently, using spatial restraints, we have studied *vpu* as tetramers, pentamers, and hexamers (Kukol and Arkin, 1999), and only a pentameric structure was in agreement with the orientational data.

The energy surface diagram for HIV *vpu* is depicted in Fig. 6, as well as the helical wheel diagram. In the case of *vpu* a canonical interaction helical pitch was used due to the very low tilt angle observed experimentally (Kukol and Arkin, 1999). Several large minima are present in the diagram, most of which correspond to large helix tilts ($\beta > 20°$). Because the experimentally derived tilt angle for *vpu* was relatively small, $\beta = \pm(6.5° \pm 1.7°)$ (Kukol and Arkin, 1999) the majority of these minima probably do not represent native stable structures. The structure obtained from spatial restraints, $\phi = 200°$ (see Note 3), $\beta = -5.5°$ does not reside in a well-defined deep energy minimum, but is

located in a shallow energy trough, which could be related to the ambiguity of reports describing the oligomeric structure of *vpu*.

Both HIV *vpu* and phospholamban are significantly more hydrophobic than the other sequences analyzed. Furthermore, the fact that the amphipathicity vectors are small indicates an even distribution of the hydrophobic residues. The lopsidedness vectors, while larger than those found in *Influenza* virus A M2, are smaller than those of *Influenza* virus C CM2 and much smaller than those of GPA. All this points to the fact that it is not surprising that it is difficult to correlate the orientation of these vectors to the protein-protein interaction surface.

Crossing angles in helix-helix interactions

In the case of HIV *vpu*, phospholamban, and *Influenza* virus C CM2 there are favored regions in which the helices are tilted much more than that measured experimentally. One reason for this finding might reside in the effect of helix topology in helix-helix interactions, as suggested by models that try to account for the statistical bias found in helix-helix crossing angles (Chothia et al., 1981; Walther et al., 1996). However, Bowie has recently shown that the statistical bias of helix-helix interaction can be accounted for simply based on geometrical grounds, removing the need for any particular model (Bowie, 1997a). In a more recent work, this author found that membrane helices are found in numerous crossing angles, the most prevalent being 20° (Bowie, 1997b). However, the conclusions for helices in water-soluble proteins might be different as, in most cases, membrane helices are juxtaposed to helices that are close in sequence space, while the analysis of soluble proteins does not take into account any helices in which the connecting loop is smaller than 20 amino acids (Chothia et al., 1981; Walther et al., 1996).

Validity of in vacuo interaction energy surface mapping

In any sort of atomistic simulation, certain assumptions must be made to undertake the calculations on a reasonable time scale. Whether the approximations are valid can be tested in light of the predictive power of the simulations in relation to a known system. The calculations undertaken in this study were exceptionally CPU-intensive, whereby each helix bundle took >11,000 h of a Hitachi SR2201 CPU. As such, certain assumptions were made, not the least of which is the simulation of the helical bundle in vacuo as opposed to a hydrated lipid bilayer.

We argue that in vacuo simulations are fully justifiable in this instance for the following reasons:

- Genest and co-workers (Duneau et al., 1999) have shown that proteins simulated in vacuo explored the same conformational space as proteins simulated in a hydrated lipid bilayer. Because, in the present study, due to the fine grid search, the conformational space search is limited, hence the probability of an accurate mapping is significantly high;
- The dielectric constant used in this study, $\epsilon = 1$, is a good approximation to that of of a membrane environment (Duneau et al., 1999);
- No use was made of a "bilayer potential" (Son & Sansom, 1999, 2000) because the simulations were only of the hydrophobic elements of the peptide and the bundle was symmetrical. Hence there is no need to ensure a matching between the hydrophobic segments of all of the helices in the bundle;
- Although lipid-protein interactions may represent an important factor in the stabilization of a transmembrane α-helical bundle, they are clearly not the driving force. This is exemplified by the fact that an otherwise conservative substitution (e.g., Ile to Leu) results in complete disruption of pentamerization of phospholamban, but only at a specific site (Arkin et al., 1994). In other words, substitutions of Ile to Leu or Leu to Ile (as an example) are readily accommodated in many sites in the proteins without any effect upon pentamerization; yet in the protein-protein interface they result in a complete disruption of the pentameric complex. This interaction scheme is best exemplified in the Popot and Engelman two-stage model for membrane protein folding an oligomerization (Popot et al., 1987; Popot & Engelman, 1990);
- Finally, the true test of any simulation procedure is its predictive powers. Earlier results from the Brünger group have shown that one can obtain an accurate structure for the glycophorin A dimerizing transmembrane α-helical bundle (Adams et al., 1996; MacKenzie et al., 1997). In the present study we identify prominent features in the interaction energy map that correspond to all of the structures and models of the protein investigated.

Based on the above arguments we conclude that in vacuo energy surface mapping is at present a useful methodology regardless of the omission of the lipid bilayer.

Our approach is distinct from that undertaken by Brünger and co-workers (Adams et al., 1995), whereby the configuration space is sparsely sampled: rotation intervals of 10° and two different crossing angles ($\beta = \pm 25°$) leading to a total of 72 sample points (as opposed to 32,760 points used in this study). Each of these starting positions is then subjected to a molecular dynamics, simulated annealing protocol, after which the convergence of different starting structures to a particular position is analyzed. In this approach one hopes to sufficiently sample the configuration space to ensure that through convergence one can detect the wild-type function. Although this may work for a variety of cases, it clearly has not worked for the M2 H^+ channel

(Kukol et al., 1999) and human phospholamban (Torres et al., 2000).

The approach taken herein is distinct in that it provides a complete view of the interaction energies for any ϕ and β pair. In doing so it is possible to simultaneously view the features of the interaction energy surface and to possibly gauge the stability of the complexes. Furthermore, one does not need to rely upon the convergence of the system to the wild-type structure rather than sample all possible configurations.

Our approach is based in part on the findings by Karplus and co-workers, which describe the advantages of running multiple short molecular dynamics trajectories over a single long run (Caves et al., 1998). Thus, one can view our work as a further extension of this principle, whereby we calculate the energy of nearly every possible conformation rather than use a single exhaustive molecular dynamics run. This is possible due to the limited configuration space of a transmembrane α-helical bundle, in comparison to a water-soluble protein of unknown topology (see Fig. 1).

CONCLUSION

In this study we have exhaustively searched the interaction space of transmembrane helix-helix interactions varying two parameters: the helix tilt (which is related to the helix crossing angle) and the rotational angle about the helix axis. As the calculations were undertaken in vacuo due to CPU time limitation, no contribution of lipid and/or solvent was taken into consideration.

We have presented calculations of energy surface diagrams for transmembrane helix-helix interactions that have proven to be useful even when a solvated lipid bilayer was excluded from the calculation (due to CPU time limitations). In most instances the experimentally determined structure corresponds to a recognizable local energy minimum on the surface, and in other cases it is the most prominent feature. Restriction to lower tilt angles, obtained from simple spectroscopic measurements (such as obtaining order parameters from FTIR), might enable better discrimination between true and false energy minima.

NOTES

1. Effects of helix curvature and bending are neglected here.

2. The helix tilt obtained from spatial restraints results in ambiguity of the sign of the angle because the absorption is proportional to $\cos^2(\beta)$.

3. The difference in the value of ϕ reported herein to that obtained previously (Kukol & Arkin, 1999), is a result of the very low tilt angle of *vpu*.

This work was supported by grants from the Wellcome Trust and the Biotechnology and Biological Sciences Research Council to I.T.A.

REFERENCES

Adams, P., I. Arkin, D. Engelman, and A. Brunger. 1995. Computational searching and mutagenesis suggest a structure for the pentameric transmembrane domain of phospholamban. *Nat. Struct. Biol.* 2:154–162.

Adams, P., D. Engelman, and A. Brunger. 1996. Improved prediction for the structure of the dimeric transmembrane domain of glycophorin A obtained through global searching. *Proteins.* 26:257–261.

Arkin, I., P. Adams, A. Brunger, S. Smith, and D. Engelman. 1997a. Structural perspectives of phospholamban, a helical transmembrane pentamer. *Annu. Rev. Biophys. Biomol. Struct.* 26:157–179.

Arkin, I., P. Adams, K. MacKenzie, M. Lemmon, A. Brunger, and D. Engelman. 1994. Structural organization of the pentameric transmembrane alpha-helices of phospholamban, a cardiac ion channel. *Embo J.* 13:4757–4764.

Arkin, I., and A. Brunger. 1998. Statistical analysis of predicted transmembrane α-helices. *Biochim. Biophys. Acta.* 1429:113–128.

Arkin, I., K. MacKenzie, and A. Brunger. 1997b. Site-directed dichroism as a method for obtaining rotational and orientational constraints for oriented polymers. *J. Am. Chem. Soc.* 119:8973–8980.

Belohorcova, K., J. Davis, T. Woolf, and B. Roux. 1997. Structure and dynamics of an amphiphilic peptide in a lipid bilayer: a molecular dynamics study. *Biophys. J.* 73:3039–3055.

Belshe, R., B. Burk, F. Newman, R. Cerruti, and I. Sim. 1989. Resistance of influenza A virus to amantadine and rimantadine: results of one decade of surveillance. *J. Infect. Dis.* 159:430–435.

Bowie, J. 1997a. Helix packing angle preferences. *Nat. Struct. Biol.* 4:915–917.

Bowie, J. 1997b. Helix packing in membrane proteins. *J. Mol. Biol.* 272:780–789.

Brunger, A., P. Adams, G. Clore, W. Gros, R. Grosse-Kunstleve, J. Jiang, J. Kuszewski, M. Nilges, N. Pannu, R. Read, L. Rice, T. Simonson, and G. Warren. 1998. Crystallography and the NMR system: a new software system for macromolecular structure determination. *Acta Crystallogr. D.* 54:905–921.

Bui, M., G. Whittaker, and A. Helenius. 1996. Effect of m1 protein and low pH on nuclear transport of influenza virus ribonucleoproteins. *J. Virol.* 70:8391–8401.

Carrasco, L. 1995. Modification of membrane permeability by animal viruses. *Adv. Vir. Res.* 45:61–112.

Caves, L., J. Evanseck, and M. Karplus. 1998. Locally accessible conformations of proteins: multiple molecular dynamics simulations of crambin. *Protein Sci.* 7:649–666.

Chothia, C., M. Levitt, and D. Richardson. 1981. Helix to helix packing in proteins. *J. Mol. Biol.* 145:215–250.

Chu, G., G. Dorn II, W. Luo, J. Hatter, V. Kadambi, R. Walsh, and E. Kranias. 1997. Monomeric phospholamban overexpression in transgenic mouse hearts. *Circ. Res.* 81:485–492.

Chu, G., L. Li, Y. Sato, J. Harrer, V. Kadambi, B. Hoit, D. Bers, and E. Kranias. 1998. Pentameric assembly of phospholamban facilitates inhibition of cardiac function in vivo. *J. Biol. Chem.* 273:33674–33680.

Duneau, J.-P., S. Crouzy, N. Garnier, Y. Chapron, and M. Genest. 1999. Molecular dynamics simulations of the ErbB-2 transmembrane domain within an explicit membrane environment: comparison with vacuum simulations. *Biophys. Chem.* 76:35–53.

Engelman, D., T. Steitz, and A. Goldman. 1986. Identifying nonpolar transbilayer helices in amino acid sequences of membrane proteins. *Annu. Rev. Biophys. Biophys. Chem.* 15:321–353.

Forrest, L., D. Tieleman, and M. Sansom. 1999. Defining the transmembrane helix of M2 protein from influenza A by molecular dynamics simulations in a lipid bilayer. *Biophys. J.* 76:1886–1896.

Holsinger, L., and R. Lamb. 1991. Influenza virus M2 integral membrane protein is a homotetramer stabilized by formation of disulfide bonds. *Virology.* 183:32–43.

Hongo, S., K. Sugawara, H. Nishimura, Y. Muraki, F. Kitame, and K. Nakamura. 1994. Identification of a second protein encoded by influenza C virus RNA segment 6. *J. Gen. Virol.* 75:3503–3510.

Jorgensen, W., and J. Tirado-Rives. 1988. The OPLS potential function for proteins, energy minimization for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.* 110:1657–1666.

Kabsch, W., and C. Sander. 1983. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers.* 22:2577–2637.

Karim, C., J. Stamm, J. Karim, L. Jones, and D. Thomas. 1998. Cysteine reactivity and oligomeric structures of phospholamban and its mutants. *Biochemistry-US.* 37:12074–12081.

Kimura, Y., M. Asahi, K. Kurzydlowski, M. Tada, and D. MacLennan. 1998. Phospholamban domain Ib mutations influence functional interactions with the Ca^{2+}-ATPase isoform of cardiac sarcoplasmic reticulum. *J. Biol. Chem.* 273:14238–14241.

Klimkait, T., K. Strebel, M. Hoggan, M. Martin, and J. Orenstein. 1990. The human immunodeficiency virus type 1-specific protein vpu is required for efficient virus maturation and release. *J. Virol.* 64:621–629.

Kukol, A., P. Adams, L. Rice, A. Brunger, and I. Arkin. 1999. Experimentally based orientational refinement of membrane protein models: a structure for the influenza A M2 H^+ channel. *J. Mol. Biol.* 286:951–962.

Kukol, A., and I. Arkin. 1999. Structure of the HIV-1 Vpu transmembrane complex determined by site-specific FTIR dichroism and global molecular dynamics searching. *Biophys. J.* 77:1594–1601.

Kukol, A., and I. Arkin. 2000. Structure of the influenza C CM2 protein transmembrane domain obtained by site-specific infrared dichroism and global molecular dynamics searching. *J. Biol. Chem.* 78:55–69.

Lamb, R., and L. Pinto. 1997. Do Vpu and Vpr of human immunodeficiency virus type 1 and NB of influenza B virus have ion channel activities in the viral life cycles? *Virology.* 229:1–11.

Lamb, R., S. Zebedee, and C. Richardson. 1985. Influenza virus M2 protein is an integral membrane protein expressed on the infected-cell surface. *Cell.* 40:627–633.

Lemmon, M., and D. Engelman. 1994a. Specificity and promiscuity in membrane helix interactions. *FEBS Lett.* 346:17–20.

Lemmon, M., and D. Engelman. 1994b. Specificity and promiscuity in membrane helix interactions. *Q. Rev. Biophys.* 27:157–218.

Lemmon, M., J. Flanagan, H. Treutlein, J. Zhang, and D. Engelman. 1992. Sequence specificity in the dimerization of transmembrane alpha-helices. *Biochemistry.* 31:12719–12725.

Lemmon, M., H. Treutlein, P. Adams, A. Brunger, and D. Engelman. 1994. A dimerization motif for transmembrane alpha-helices. *Nat. Struct. Biol.* 1:157–163.

Li, M., R. Cornea, J. Autry, L. Jones, and D. Thomas. 1998. Phosphorylation-induced structural change in phospholamban and its mutants, detected by intrinsic fluorescence. *Biochemistry-US.* 37:7869–7877.

MacKenzie, K., and D. Engelman. 1998. Structure-based prediction of the stability of the transmembrane helix-helix interactions: the sequence dependence of glycophorin A dimerization. *Proc. Natl. Acad. Sci. USA.* 95:3583–3590.

MacKenzie, K., J. Prestegard, and D. Engelman. 1997. A transmembrane helix dimer: structure and implications. *Science.* 276:131–133.

Maldarelli, F., R. Willey, and K. Strebel. 1993. Human immunodeficiency virus type 1 Vpu protein is an oligomeric type I integral membrane protein. *J. Virol.* 67:5056–5061.

Popot, J., and D. Engelman. 1990. Membrane protein folding and oligomerization: the two-stage model. *Biochemistry.* 29:4031–4037.

Popot, J., S. Gerchman, and D. Engelman. 1987. Refolding of bacteriorhodopsin in lipid bilayers. A thermodynamically controlled two-stage process. *J. Mol. Biol.* 198:655–676.

Sajot, N., and M. Genest. 2000. Structure prediction of the dimeric neu/ErbB-2 transmembrane domain from multi-nanosecond molecular dynamics simulations. *Eur. Biophys. J.* 28:648–662.

Sansom, M. 1998. Ion channels: molecular modeling and simulation studies. *Methods Enzymol.* 293:647–693.

Schubert, U., S. Bour, A. Ferrer-Montiel, M. Montal, F. Maldarelli, and K. Strebel. 1996. The two biological activities of human immunodeficiency virus type 1 Vpu protein involve two separable structural domains. *J. Virol.* 70:809–819.

Schubert, U., and K. Strebel. 1994. Differential activities of the human immunodeficiency virus type 1-encoded Vpu protein are regulated by phosphorylation and occur in different cellular compartments. *J. Virol.* 68:2260–2271.

Simmerman, H., Y. Kobayashi, J. Autry, and L. Jones. 1996. A leucine zipper stabilizes the pentameric membrane domain of phospholamban and forms a coiled-coil pore structure. *J. Biol. Chem.* 271:5941–5946.

Son, H., and M. Sansom. 1999. Simulation of the packing of idealized transmembrane alpha-helix bundles. *Eur. Biophys. J.* 28:489–498.

Son, H., and M. Sansom. 2000. Simulation studies on bacteriorhodopsin alpha-helices. *Eur. Biophys. J.* 28:674–682.

Stevens, T., and I. Arkin. 1999. Are membrane proteins "inside-out" proteins? *Proteins.* 36:135–143.

Strebel, K., T. Klimkait, F. Maldarelli, and M. Martin. 1989. Molecular and biochemical analyses of human immunodeficiency virus type 1 vpu protein. *J. Virol.* 63:3784–3791.

Sugrue, R., and A. Hay. 1991. Structural characteristics of the M2 protein of influenza A viruses: evidence that it forms a tetrameric channel. *Virology.* 180:617–624.

Tieleman, D., H. Berendsen, and M. Sansom. 1999. Surface binding of alamethicin stabilizes its helical structure: molecular dynamics simulations. *Biophys. J.* 76:3186–3191.

Torres, J., P. Adams, and I. Arkin. 2000. Use of a new label, $^{13}C{=}{=}{}^{18}O$, in the determination of a structural model of phospholamban in a lipid bilayer. Spatial restraints resolve the ambiguity arising from interpretations of mutagenesis data. *J. Mol. Biol.* 300:677–685.

Toyofuku, T., K. Kurzydlowski, M. Tada, and D. MacLennan. 1994. Amino acids Glu2 to Ile18 in the cytoplasmic domain of phospholamban are essential for functional association with the Ca^{2+}-ATPase of sarcoplasmic reticulum. *J. Biol. Chem.* 269:3088–3094.

Treutlein, H., M. Lemmon, D. Engelman, and A. Brunger. 1992. The glycophorin A transmembrane domain dimer: sequence-specific propensity for a right-handed supercoil of helices. *Biochemistry.* 31: 12726–12732.

Walther, D., F. Eisenhaber, and P. Argos. 1996. Principles of helix-helix packing in proteins: the helical lattice superposition model. *J. Mol. Biol.* 255:536–553.

Wang, C., R. Lamb, and L. Pinto. 1995. Activation of the M2 ion channel of influenza virus: a role for the transmembrane domain histidine residue. *Biophys. J.* 69:1363–1371.

Woolf, T. 1997. Molecular dynamics of individual alpha-helices of bacteriorhodopsin in dimyristoyl phosphatidylcholine. I. Structure and dynamics. *Biophys. J.* 73:2376–2392.

Woolf, T., and M. Tychko. 1998. Simulations of fatty acid-binding proteins. II. Sites for discrimination of monounsaturated ligands. *Biophys. J.* 74:694–707.