Contribution of Energy Values to the Analysis of Global Searching Molecular Dynamics Simulations of Transmembrane Helical Bundles

Jaume Torres,* John A. G. Briggs,† and Isaiah T. Arkin‡

*Cambridge Centre for Molecular Recognition, Department of Biochemistry, University of Cambridge, Cambridge CB2 1GA, United Kingdom; †Department of Biochemistry, University of Oxford, Oxford OX1 3QU, United Kingdom; and ‡The Alexander Silberman Institute of Life Sciences, Department of Biological Chemistry, The Hebrew University of Jerusalem, Givat-Ram, Jerusalem 91904, Israel

ABSTRACT Molecular interactions between transmembrane α-helices can be explored using global searching molecular dynamics simulations (GSMDS), a method that produces a group of probable low energy structures. We have shown previously that the correct model in various homooligomers is always located at the bottom of one of various possible energy basins. Unfortunately, the correct model is not necessarily the one with the lowest energy according to the computational protocol, which has resulted in overlooking of this parameter in favor of experimental data. In an attempt to use energetic considerations in the aforementioned analysis, we used global searching molecular dynamics simulations on three homooligomers of different sizes, the structures of which are known. As expected, our results show that even when the conformational space searched includes the correct structure, taking together simulations using both left and right handednesses, the correct model does not necessarily have the lowest energy. However, for the models derived from the simulation that uses the correct handedness, the lowest energy model is always at, or very close to, the correct orientation. We hypothesize that this should also be true when simulations are performed using homologous sequences, and consequently lowest energy models with the right handedness should produce a cluster around a certain orientation. In contrast, using the wrong handedness the lowest energy structures for each sequence should appear at many different orientations. The rationale behind this is that, although more than one energy basin may exist, basins that do not contain the correct model will shift or disappear because they will be destabilized by at least one conservative (i.e. silent) mutation, whereas the basin containing the correct model will remain. This not only allows one to point to the possible handedness of the bundle, but can be used to overcome ambiguities arising from the use of homologous sequences in the analysis of global searching molecular dynamics simulations. In addition, because clustering of lowest energy models arising from homologous sequences only happens when the estimation of the helix tilt is correct, it may provide a validation for the helix tilt estimate.

INTRODUCTION

Prediction of transmembrane domain structure is facilitated by its tendency to adopt (in most cases) an α-helical fold, limiting the number of possible conformations. Brunger and co-workers (Treutlein et al., 1992; Adams et al., 1995) have developed a procedure to explore transmembrane helix interactions based on global searching molecular dynamics simulations. In this method, multiple symmetric bundles of helices are constructed, each differing from the other by the rotation of the helices about their axes (Fig. 1). These are then used as starting positions for molecular dynamics and energy minimization protocols. The output structures from these simulations are compared and grouped into clusters that contain similar structures. An average of the structures forming a cluster represents a model with characteristic interhelical interactions and helix tilt.

The correct model is selected amongst the several different clusters, based on existing experimental data, either from mutagenesis (Lemmon et al., 1992a,b; Arkin et al., 1994) or orientational data from site specific infrared dichroism used as spatial restraints (Kukol et al., 1999; Torres et al., 2000). Alternatively, we have also used a purely computational approach where simulations are performed on close sequence variants that are likely to share the same structure (Briggs et al., 2001).

The value of energy as a discriminating tool is usually overlooked during the analysis due to the fact that the correct model is not necessarily the one with lowest energy according to the computational protocol. Recently however, in an exhaustive evaluation of the energy of energy-minimized helical bundles as a function of their helix tilt and rotational orientation, we obtained a picture of the energy landscape of α-helical bundles as a function of their interhelical interactions (Torres et al., 2001). The conclusion was that, although more than one energy basin was present, the model that was in agreement with experimental data, (i.e., the correct model), was always located at the bottom of one of these basins. In some cases, only two basins were found, one for each handedness, i.e., left ($\beta > 0$) or right ($\beta < 0$). In other cases, more basins were found, although the search for the correct model is limited to a slice of the energy surface due to the helix tilt constraint.

An example of this has been shown previously for phospholamban (Torres et al., 2000) and in the case of one particular variant of the M2 H^+ of Influenza A (Kukol et al.,

Submitted April 25, 2001, and accepted for publication February 12, 2002.

Address reprint requests to Isaiah T. Arkin, The Alexander Silberman Institute of Life Sciences, Department of Biological Chemistry, The Hebrew University of Jerusalem, Givat-Ram, Jerusalem 91904, Israel. Tel.: 972-0-2-658-4329; Fax: 972-0-2-658-4329; E-mail: arkin@cc.huji.ac.il.

© 2002 by the Biophysical Society

0006-3495/02/06/3063/09 $2.00

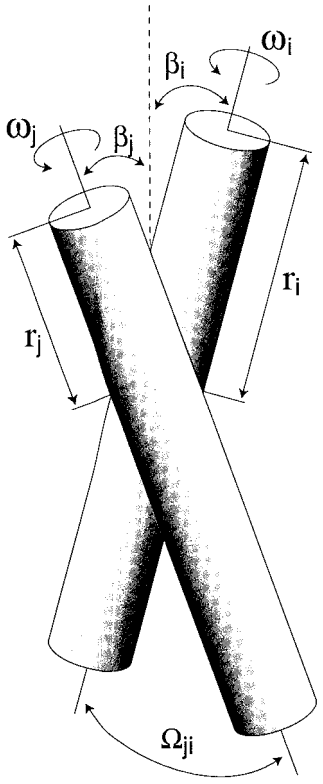

FIGURE 1 Schematic representation of the geometric parameters describing the conformation of an α-helical bundle. β, The helix tilt; r, the register; ω, the rotational angle; and Ω, the crossing angle.

1999). When the helix tilt was restrained to the experimentally determined value by Fourier transform infrared, site-specific infrared dichroism, the correct model was also the lowest energy model when the simulation was performed using the correct handedness (known a priori). This is despite that the lowest energy model with the opposite, wrong handedness, could have had even lower energy and could have been equally compatible with experimental data.

Herein, we have attempted to discriminate between the energy basin that contains the correct model and those that do not by using homologous sequences. We have shown previously (Briggs et al., 2001) that silent amino-acid substitutions, present in close homologues or unveiled by mutagenesis studies, do not affect the stability of the native structure during the simulation, although they destabilize some of the non-native structures also found. Global searches carried out on enough variants produce a set of similar structures that appear in all of the searches, forming a "complete set." We argued that the structure represented by this "complete set" should be the native one, because it had not been destabilized by any of the conservative/silent mutations.

Similarly, we expect that a particular basin that does not contain the correct model, regardless of how low its energy is, will shift or disappear in at least one of the simulations (of the different homologues), whereas the basin containing the correct model will remain in all simulations.

When helix tilt and handedness are correct, therefore, upon use of different homologous sequences, the lowest energy structures should cluster around a particular orientation. Conversely, lowest energy structures should spread over many different orientations if either of these two parameters is incorrect. We show this to be true for the glycophorin A homodimer for which a structure has been reported (MacKenzie et al., 1997).

The use of the model's energy values using close homologues is an essential tool to resolve ambiguities that arise in the new approach reported recently (Briggs et al., 2001) that also uses close variants. For example, if the number of close homologues available is not sufficient, more than one model may form a "complete set." Alternatively, if the correct model is not found amongst the candidate structures and the number of sequences available is too small, a wrong model may still be found to persist in all of the simulations tested. Finally, when mutagenesis data are used, results are not always clear cut, because function is not always the parameter under monitoring, thereby creating ambiguity when evaluating these results. This can have the effect of precluding the formation of any "complete set," even when the models arising from global searching molecular dynamics contain the correct structure. In all of these cases, it should be possible to select the correct model by looking at the energy of each model and determining where clustering of lowest energy models occurs around certain orientations.

We provide examples of, and overcome, this ambiguity with a homotetramer and a homopentamer. In one case, we use the α-helical (Duff et al., 1992; Kukol et al., 1999) transmembrane domain of M2, a protein from the Influenza A virus that forms homotetrameric H^+-selective ion channels. For M2 global searching molecular dynamics simulations (GSMDS), we have used in the simulations close homologues that are likely to have the same native structure. However, we have deliberately included some variants that are known to confer resistance to the antiviral drug amantadine (Hay et al., 1985). Since it has been shown that the transmembrane part of the protein is necessary and sufficient for the functional and structural properties of the protein (Duff and Ashley, 1992; Duff et al., 1994), it is possible that the variants of M2 from amantadine-resistant viruses possess structural properties that are somewhat different from the amantadine-sensitive variants. This is incompatible with the main premise in the analysis of GSMDS by using close variants in which all sequences should share the same structure.

The other system used, the homopentameric phospholamban (Plb), a 52-residue protein located in the cardiac sarcoplasmic reticulum, contains a hydrophobic domain (residues 26–52) that forms a left-handed α-helical bundle (Arkin et al., 1994; for recent reviews see Arkin et al., 1997; Simmerman and Jones, 1998). No evolutionary data is available

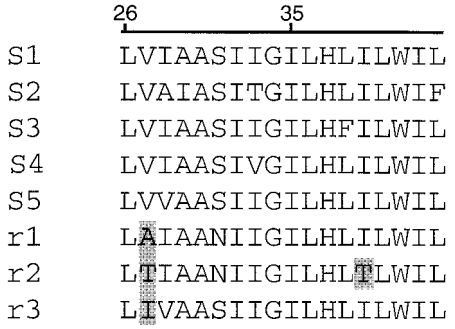

FIGURE 2 Sequences of the transmembrane segment of M2 (TM-M2) used in the simulations. The corresponding accession numbers are: s1, p36348; s2, m63533; s3, p05778; s4, q9w986; s5, p21430; r1, p06821; r2, u08863; and r3, p03492. The variants s1 to s5 are amantadine sensitive mutants, whereas r1, r2, and r3 are amantadine resistant mutants (with mutations at residues 27 and 39, see gray rectangles) (Bleasby et al., 1994).

for Plb, although a wealth of information exists concerning the residues that are important in maintaining the pentameric structure, mainly based on mutagenesis data (Fujii et al., 1989; Arkin et al., 1994; Simmerman et al., 1996; Kimura et al., 1997). Ambiguity here is introduced by the results of mutagenesis studies, which did not define conservative mutations as those preserving function but those that did not prevent pentamerization. Thus, some of the conservative changes reported may had affected function.

The correct orientations and handedness for M2 and Plb are known. For M2, the rotational orientation and the helix tilt have been experimentally determined by two independent methods: solid-state nuclear magnetic resonance (NMR) (Kovacs and Cross, 1997) and site-directed infrared dichroism (Kukol et al., 1999). For Plb, two models were capable of "accommodating" the mutagenesis data (Arkin et al., 1994; Simmerman et al., 1996), although the leucine zipper model (Simmerman et al., 1996) has been confirmed by labeling experiments performed in sodium dodecyl sulphate (SDS) (Karim et al., 1998) and a site-directed infrared dichroism study (Torres et al., 2000). Herein, new ways in which examination of energy values can contribute to the analysis of results of GSMDS are presented, using the aforementioned examples.

MATERIALS AND METHODS

Global search molecular dynamics

Glycophorin

The protocols for the simulations and the homologues used for glycophorin homodimers were as described previously (Briggs et al., 2001).

M2

Simulations with M2 were performed using the transmembrane sequence from residues 26 to 43. M2 variants (Fig. 2) with at least 60% identity were obtained searching the OWL data base (Bleasby et al., 1994). These were either variants from amantadine sensitive (s1–s5) or amantadine resistant (r1–r3) viral strains. The identity between the sequences from the amantadine resistant strains was at least 72%.

Plb

For Plb, conservative mutations in the wild-type transmembrane sequence (residues 35–52) FCLILICLLLICIIVMLL (see text) were used from either Engelman and co-workers (Arkin et al., 1994) or Jones and co-workers (Simmerman et al., 1996).

GSMDS protocol

All calculations were performed using parallel crystallography and NMR system, the parallel-processing version of the *Crystallography and NMR System* (CNS Version 0.3) (Brunger et al., 1998), with the optimized potential for liquid simulations parameter set and united atom topology (Jorgensen and Tirado-Rives, 1988), explicitly describing only polar and aromatic hydrogens. A global search was carried out in vacuo as described elsewhere, using CHI (CNS Helical Interactions), assuming a symmetrical interaction between the helices in the homooligomer (Adams et al., 1995). Briefly, trials were carried out starting from either left or right crossing angles (i.e., $\pm 25°$, respectively). For each of these cases, the helices were rotated a total of 350° about their helical axes in 10° increments (Adams et al., 1995), so that all possible interhelical interactions were explored. Four trials were carried out from each starting configuration using different initial random velocities, making a total of $36 \times 2 \times 4 = 288$ trials, each producing a final structure. Clusters of output structures were identified, defined as those that contain a minimum number of structures (typically 10). Any structure belonging to a particular cluster was typically within 1.0 Å α-carbon root mean square deviation (RMSD) from any other structure within that cluster. Some clusters therefore overlap, and output structures may be members of more than one cluster. The structures belonging to each cluster were averaged and subjected to a further simulated annealing protocol. This final structure was taken as the representative of the cluster.

Finally, the side chain rotamer angles are free to rotate during the course of the molecular dynamics run as well as the energy minimization protocol. The initial values are those found most prevalent in the rotamer library (Ponder and Richards, 1987).

Analysis of the simulations

The results from the global searching molecular dynamics simulations were represented graphically by plotting each cluster representative as a function of two parameters, the helix tilt β and the rotational orientation ω, as described previously (Briggs et al., 2001) and shown schematically in Fig. 1. The tilt angle of the model, β, was taken as the average of the angles between each helix axis and the bundle axis, which is coincident with the normal to the bilayer and was calculated by CHI (Adams et al., 1995, 1996). The helical axis is a vector with starting and end points above and below a defined residue, where the points correspond to the geometric mean of the coordinates of the five α carbons N-terminal and the five α carbons C-terminal to the defined residue.

The rotational orientation ω, is defined relative to an arbitrarily chosen, specified residue. The angle ω is defined by the angle between a vector perpendicular to the helix axis, oriented towards the middle of the peptidic C═O bond of the residue, and a plane that contains both the helical axis and the normal to the bilayer (Fig. 1). This angle is 0° when the residue is located in the direction of the tilt (Arkin et al., 1997). The structures identified were plotted against the ω angle of residue 83 for glycophorin A, 35 for M2, and 42 for Plb.

Precise comparisons between similar clusters obtained from different variants were made by calculating the RMSD between their α-carbon

backbones. In the simulations, the handedness of the bundle is indicated by the helix tilt sign, positive or negative, which corresponds to left and right handed bundles, respectively. Note that in a helix dimer, the sum of both tilt angles is equal to the commonly used helix crossing angle, Ω. Such a relationship does not necessarily exist for higher order oligomers.

M2 simulations in which the helix tilt was restrained were performed as described previously (Torres et al., 2000). Specifically, the angle between the vectors connecting every C_α of residue i and C_α of residue $i + 7$ and the z axis was restrained to 31.5°, which is the experimental value obtained for the helix tilt using site-directed dichroism (Kukol et al., 1999).

Energy calculation

The energies calculated correspond to the total energy of the system, including both bonded (e.g. bond, angle, dihedral, improper) and non-bonded (i.e. Van der Waals, electrostatic) terms (Brünger et al., 1998). No account is made for contributions from the lipids to the energetics of the system (Torres et al., 2001).

RESULTS AND DISCUSSION

Glycophorin homodimer

The results of the simulations for various homologous sequences of GpA are shown in Fig. 3. The lowest energy structures for each of the sequences at left handed configurations (Fig. 3 *a*) appear at various orientations (Fig. 3 *b*). In contrast, the right handed structures (Fig. 3 *c*) show a clustering of the lowest energy structures around $\omega \sim -80°$ (Fig. 3 *d*), although the lowest energy structures from pig and gibbon are 25° and 50° apart from this orientation. These results are consistent with a previous approach in which a structure at $\omega = -80°$ and $\beta = -10°$ (right handed) was identified at the bottom of an energy basin (Torres et al., 2001). This model is also indistinguishable to that obtained computationally using the effect of conservative substitutions (Briggs et al., 2001) and experimentally by solution NMR in detergent micelles (MacKenzie et al., 1997).

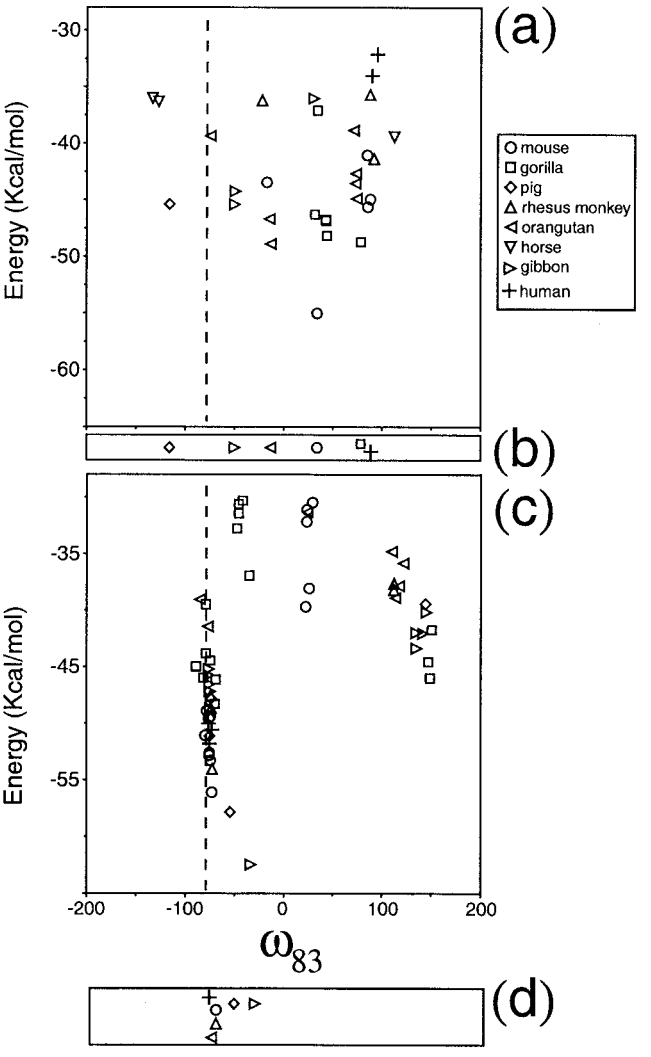

FIGURE 3 Results obtained for the homologues of GpA. The energies of the models have been plotted as a function of the orientation, represented by ω_{83}. (*a*) Left handed structures. (*b*) Orientation of the lowest energy structures for each of the sequences represented in *a*. (*c* and *d*) Same as *a* and *b*, but for right-handed structures. The vertical broken line indicates the position at which ω_{83} should be located according to solution NMR (MacKenzie et al., 1997). Representations such as those in *b* and *d* are only intended to compare the orientation (*x* axis) of the lowest energy models for each sequence. The slight shift of the symbols along the *y* axis, when occurs, is only for the sake of clarity to avoid overlapping.

M2

No restraints

Fig. 4 depicts the result of the simulations using both left- and right-handed configurations of TM-M2, without restraining the helix tilt. Fig. 4 *a* shows that no structure could be found with a helix tilt β of 31.5° or $\omega_{35} = \sim -100°$ compatible with that obtained experimentally (see vertical broken line (Kukol et al., 1999; Kovacs and Cross, 1997). In fact, in all simulations the helix tilt of all the structures identified was smaller than 20°. Interestingly however, a left-handed bundle (positive helix tilt, above the horizontal broken line) at $\omega_{35} = 25°$ (see gray squares) is present in all the simulations carried out with the amantadine sensitive variants (s1–s5) and also with two of the amantadine resistant variants, r1 and r3. However, even relaxing the clustering parameters so that clusters should contain a minimum of six structures instead of 10, r2 failed to produce a model at $\omega_{35} = \sim 25°$ (not shown). This is a clear example of how ambiguity can easily arise when using the approach reported previously (Briggs et al., 2001). In this case, if r2 had not been used, the model at $\omega_{35} \sim 25°$ might have been taken as correct. Alternatively, one may have thought that r2 possesses a structure that is somewhat different and should not have been included in the simulation in the first place, which again would justify the model at $\omega_{35} \sim 25°$ as correct for the other sequences. Thus, due to a limited number of sequences or to uncertainty regarding the structural identity of the variants used, a wrong model is obtained. We show

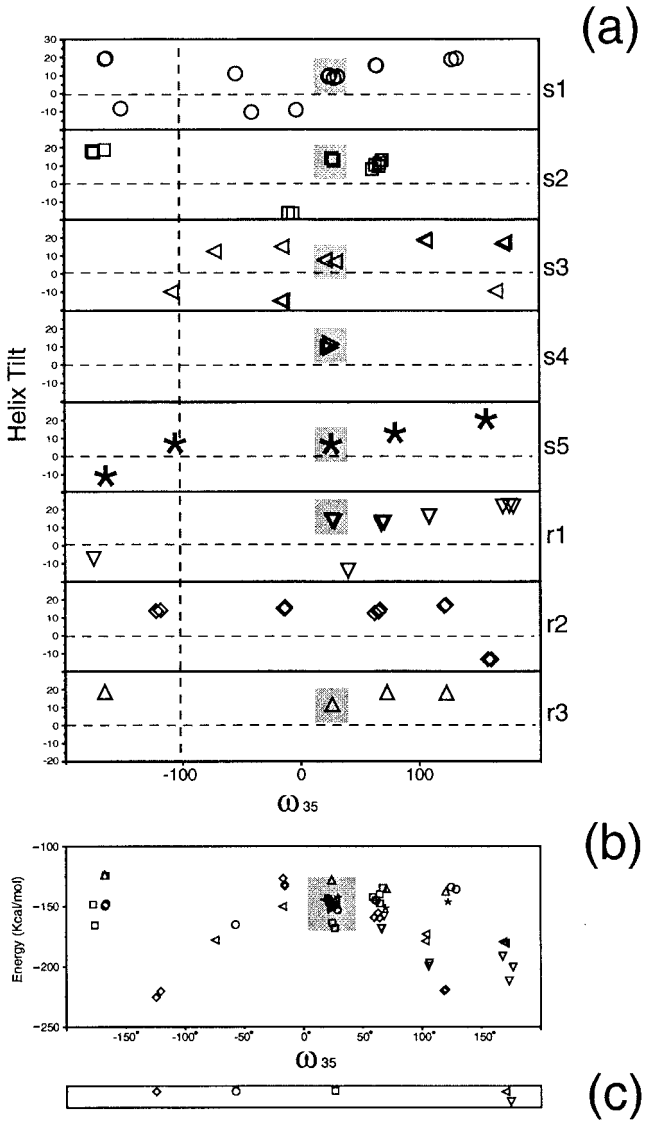

FIGURE 4 Results obtained for the homologues of M2 (Fig. 2) without restraining the helix tilt. (*a*) Plot corresponding to the models obtained from the MD simulations as a function of the helix tilt and ω_{35}. The horizontal broken line in each sequence separates left-handed bundles (above line) from right-handed ones (below line). The vertical broken line indicates the position at which ω_{35} should be located according to previous reports (Kukol et al., 1999; Kovacs and Cross, 1997). (*b*) Plot of the energy as a function of ω_{35} for the left-handed structures. (*c*) Lowest energy clusters for each sequence.

here that this can be overcome by simply examining the relative energy of the models generated.

Indeed, examination of the relative energies of the left-handed structures for each of the simulations (Fig. 4 *b*) shows that the structures at $\omega_{35} = 25°$ (see gray square) are not the lowest energy structures within a particular sequence for any of the simulations. On the contrary, it has almost the highest energy in each of the simulations, suggesting that this model is wrong. In addition, as no other model forms a "complete set," we can infer that the correct structure is not present amongst the low energy clusters generated by the global searching molecular dynamics protocol.

When the helix tilt is incorrect, it is not possible to generate a "complete set" where the components are all lowest energy structures for a certain handedness. Therefore, the lack of results described above may indirectly help to identify when a helix tilt is incorrect and can be used as a feedback on the validity of experimental measurements.

Use of homologues restraining the helix tilt

Fig. 5 shows the results of the global search molecular dynamics protocol for different M2 variants when the helix was restrained to 31.5°, the experimentally determined value (Kukol et al., 1999), as described in Materials and Methods. We note that this restraint is not completely strict, and the actual helix tilt at the end of the simulation can vary up to $\pm 5°$ from the value used in the restraint. Comparing the results of all of these global searches reveals that at two positions, $\beta \sim 31°$, $\omega_{35} \sim -100°$, and $\beta \sim 31°$, $\omega_{35} \sim 100°$ (both left handed) clusters are found for all the homologues that are amantadine sensitive (first five variants). No structure within either of the "complete sets," at $\omega_{35} \sim -100°$ and $\omega_{35} \sim 100°$, differs from any other in the set by more than 0.8 Å Cα RMSD. Therefore, these two orientations are in principle possible according to this computational approach. One of these models, however (see gray rectangles), is equivalent to the model previously determined using site-directed infrared dichroism (Kukol et al., 1999) and solid-state NMR (Kovacs and Cross, 1997). We would expect that if sufficient variants could be simulated, the incorrect structure would not persist, as it would be destabilized by at least one conservative mutation (Briggs et al., 2001).

Examination of the dependency of energy on the orientation in left-handed models (Fig. 5, lower panel) shows that the model that is consistent with the experimental results (see dotted line) is also the one with lowest energy in each of the simulations. As this structure was not identified for any of the amantadine resistant strains, i.e., r1, r2, and r3, the clustering parameters were then relaxed by imposing the condition that the clusters should contain at least six structures instead of 10. These results are shown in Fig. 6, where it is shown that the structure in agreement with the data is represented also in two of the three amantadine resistant variants, r1 and r3. As before, for these two variants, these structures are also the lowest energy structures (see Fig. 6, lower panel). Sequence r2 however, which contains the amantadine insensitivity mutations at residues 27 and 39, did not converge to this orientation, although a cluster was found close to that model (at $\sim -125°$, see arrow) that is also the lowest energy cluster. In fact, the α-carbon RMSD between this cluster and any other model belonging to the "complete set" at $\omega_{35} \sim -100°$ is less than 1.4 Å, which

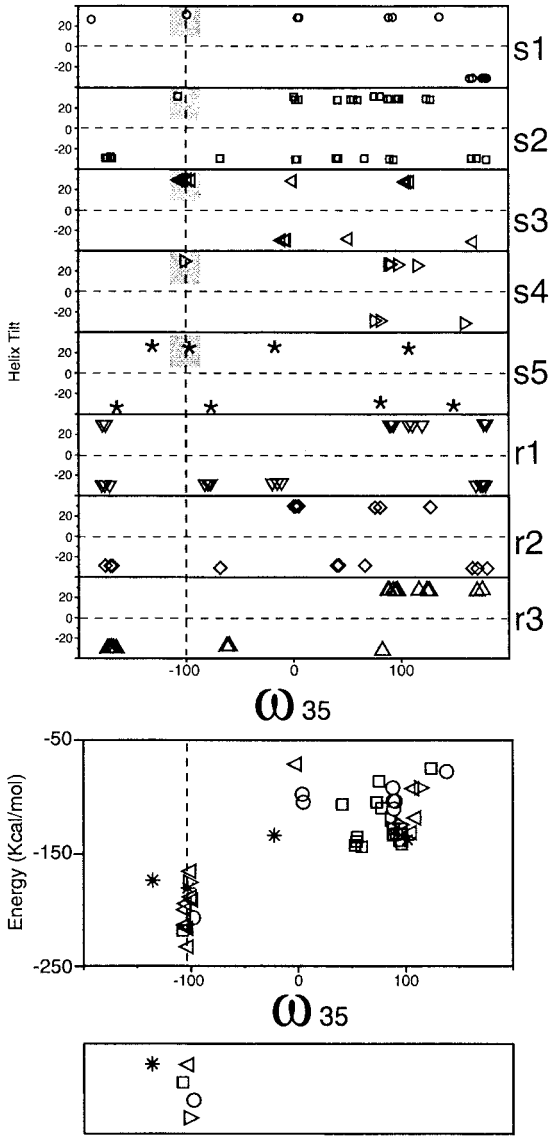

FIGURE 5 (*Top panel*) Results obtained with the sequences in Fig. 4 when the helix tilt was restrained to 31.5° in the left-handed configuration (RMSD, 1 Å; $n = 10$). The horizontal broken line in each case separates left-handed bundles (above line) from right-handed ones (below line). The vertical broken line indicates the position at which ω_{35} should be located according to previous reports (Kukol et al., 1999; Kovacs and Cross, 1997). (*Middle panel*) Rotational orientation (ω_{35}) for the lowest energy clusters in each simulation. (*Bottom panel*) Plot of the energy as a function of the orientation of ω_{35} for the sequences s1-s5, using only the left-handed structures.

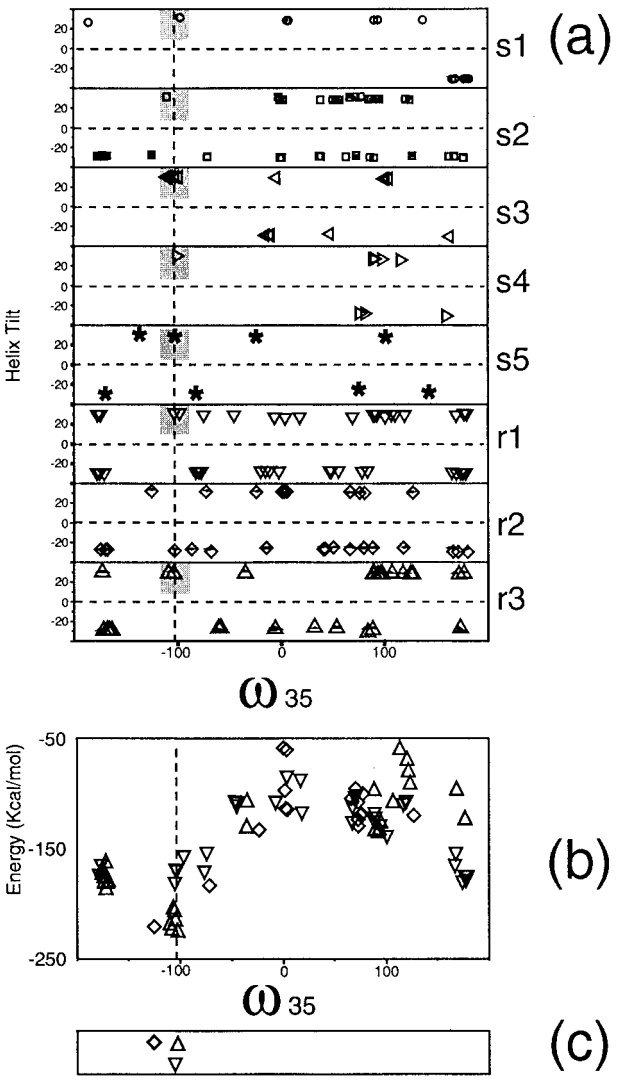

FIGURE 6 (*a*) Results obtained for the homologues of M2 (Fig. 2) restraining the helix tilt to 31.5° (RMSD, 1 Å; $n = 6$). The horizontal broken line in each case separates left-handed bundles (above line) from right handed ones (below line). The vertical broken line indicates the position at which ω_{35} should be located according to previous reports (Kukol et al., 1999; Kovacs and Cross, 1997). (*b*) Plot of the energy as a function of the orientation of ω_{35} for sequences r1, r2, and r3 for the left-handed structures. (*c*) Rotational orientation (ω_{35}) for the lowest energy clusters in each simulation.

shows that this structure is in fact very similar to those corresponding to the "complete set."

In this case, therefore, we have used energy considerations to determine that neither of the models obtained without helix tilt restraints were correct. Then, once the tilt was restrained to that determined experimentally, the use of energy values allowed us to discriminate between equally probable models.

We note that, apart from the two substitutions at residues 27 and 39, s2 is otherwise identical to r1. According to previous experimental data, residue 39 faces the lipids (Kukol et al., 1999; Kovacs and Cross, 1997). It is therefore likely to be the substitution at residue 27 that is changing the conformational space in such a way as to prevent the molecular dynamics simulations identifying the same structure for s2.

As expected, the lowest energy right-handed models (Fig. 7 *a*) do not cluster around any particular orientation (Fig. 7 *b*).

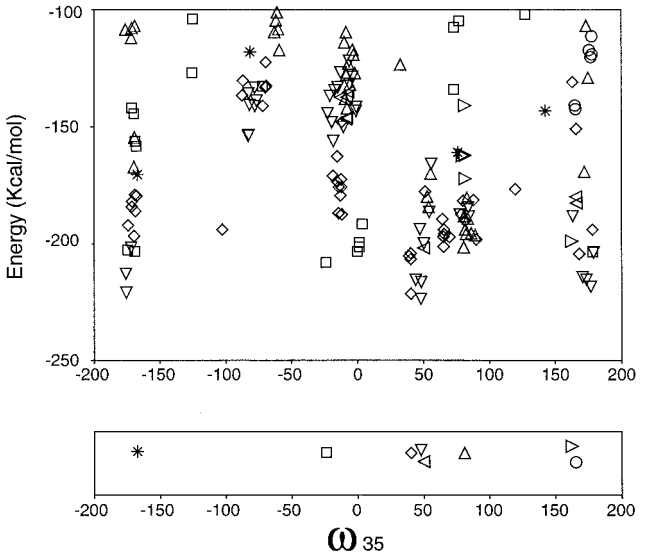

FIGURE 7 (*Top panel*) Results obtained with the sequences in Fig. 4 when the helix tilt was restrained to 31.5° in the right handed configuration (RMSD, 1 Å; $n = 10$). (*Bottom panel*) Rotational orientation (ω_{35}) for the lowest energy clusters in each simulation.

Phospholamban

For Plb, as no sequences are available from evolutionary data, variants resulting from mutagenesis data (Arkin et al., 1994; Simmerman et al., 1996) were used. Fig. 8 shows that when no restraints were used in the left-handed configuration no "complete set" was found. In contrast to the simulations on M2 in the absence of restraints, however, a number of structures are identified in the vicinity of the experimentally determined helix tilt (11°, Torres et al., 2000), so it is unlikely that the correct structure falls outside the conformational space searched.

Alternatively, some of the sequence variants used may contain some ambiguity regarding their ability to preserve the native structure, because the mutants were assayed for their ability to form pentamers and may not necessarily preserve function. Or in other words, they may pentamerize in a slightly different way than in the native sequence.

Clearly, not all mutations described in the literature can be conservative/silent. For example, the now accepted leucine zipper model (Simmerman et al., 1996), equivalent to $\omega_{35} \sim -55°$ in Fig. 8, displays C41 facing the lipidic environment, whereas C36 is located in the interhelical region and C46 is facing the lumen of the pore. The other alternative model, also in agreement with mutagenesis data (Arkin et al., 1994), at $\omega_{35} \sim -10°$, displays C41 and C46 in the interhelical contacts and C36 facing the pore. Only these two models, (see gray squares in Fig. 8) were present in all of the simulations but one. It is perhaps not surprising, therefore, that the results that involve C41V and C36F are similar. Whereas, the C41V simulation

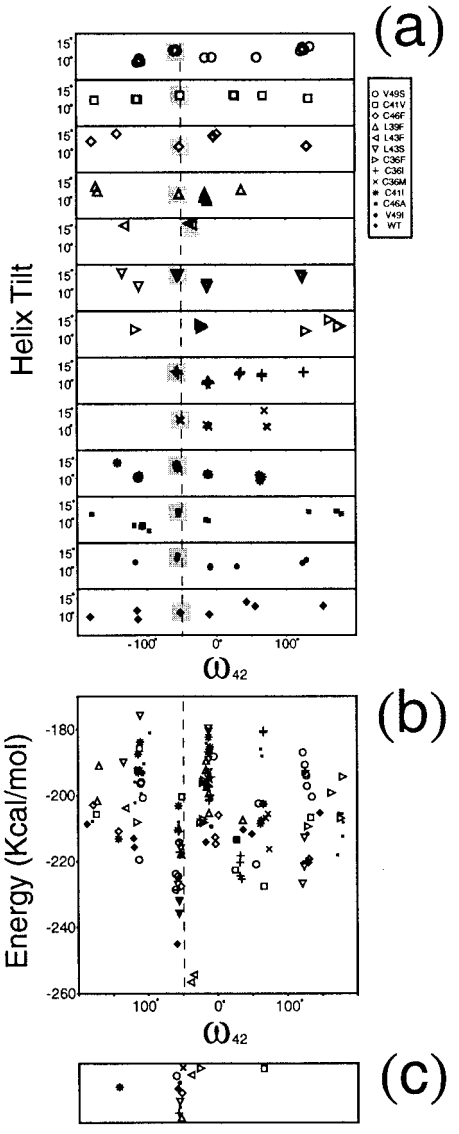

FIGURE 8 Results obtained without restraining the helix tilt for sequences of phospholamban that show pentamerization in SDS. Only the results assuming a left-handed bundle are shown. The symbols corresponding to each substitution are represented on the right hand side. (*a*) Plot corresponding to the low energy clusters obtained from the MD simulations as a function of the helix tilt and ω_{42}. Only the left-handed bundles are shown. (*b*) Plot of the energy as a function of ω_{42}. The vertical broken line indicates the position at which ω_{42} should be located according to previous reports (Torres et al., 2000). (*c*) Lowest energy clusters for each sequence.

does not display a model at $\omega \sim -10°$, C36F did not produce a model at $\omega \sim -55°$.

We note that as with other systems simulated, although the ω value for a given residue may change somewhat, the structure can still be very similar. For example, the model for L43F at $\omega \sim -35°$ belongs in fact to the set at $\omega \sim -55°$ because is at less than 1 Å C_α RMSD from any structure within that set.

As in M2, however, the model with lowest energy for each of the simulations (see lower panel in Fig. 8 at $\sim -55°$) is also in agreement with previous experimental data (Torres et al., 2000). Therefore, our data suggest that C36F may preserve the pentameric properties of the protein but may not be functional due to slight changes in the pentameric structure.

CONCLUSION

In summary, the approach previously described (Briggs et al., 2001) is limited by the inability of the search protocol to include the correct structure among its output and by ambiguities present in the input sequences, either from evolutionary or mutagenesis data. We have shown that the incorporation of a simple restraint such as the helix tilt is sufficient to allow the native structure to be found. Once the helix tilt is correct, this in turn allows one to discriminate between the models according to their energy.

Examples of ambiguities have been provided and overcome by the use of energy values as an analysis tool. The main problems facing these methods are the imperfections in the simulation protocol, i.e., the force field does not reflect accurately the properties of the system and the simulations are performed in vacuo. We stress that it is because of the imperfections of the system referred to above, which may distort somewhat the energy landscape, that lowest energy models form a "loose" cluster, where orientations for some sequences can be as far as 50° from the main cluster.

It is likely that the contribution of lipid-protein interactions in the different models is similar, and therefore the correct model is also the one with lowest energy, even in the absence of lipids. This is important in three ways. It allows us to know when the correct model has not been found amongst the candidate structures, making necessary the incorporation of experimental restraints such as helix tilt. Also, it allows for the selection of the correct model when 1) more than one model form a "complete set," in a situation where the number of sequences is limited, or 2) when no "complete sets" are found due to ambiguous results from site-directed mutagenesis. Restraints such as the helix tilt can be relatively easy to obtain for homooligomers using techniques such as Fourier transfer infrared, and solid-state NMR or from low resolution structures from electron microscopy, although in the latter case, information on relative helix register, which can be obtained from electron paramagnetic resonance, is also necessary.

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

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

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

Bleasby, A. J., Y. Akrigg, and T. K. Attwood. 1994. OWL: a non-redundant composite protein sequence database. *Nucleic Acids Res.* 22:3574–3577.

Briggs, J. A. G., J. Torres, and I. T. Arkin. 2001. A new method to model membrane protein structure based on silent amino-acid substitutions. *Proteins Struct. Func. Genet.* 44:370–375.

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

Brünger, A. T., P. D. Adams, G. M. Clore, W. L. P. Gros, R. W. Grosse-Kunstleve, J. S. Jiang, J. Kuszewski, M. Nilges, N. S. Pannu, R. J. Read, L. M. Rice, T. Simonson, and G. Q. Warren. 1998. CNS Version 0.3. Yale University, New Haven, CT.

Duff, K. C., and R. H. Ashley. 1992. The transmembrane domain of influenza A M2 protein forms amantadine-sensitive proton channels in planar lipid bilayers. *Virology.* 190:485–489.

Duff, K. C., P. J. Gilchrist, A. M. Saxena, and J. P. Bradshaw. 1994. Neutron diffraction reveals the site of amantadine blockade in the influenza A M2 ion channel. *Virology.* 202:287–293.

Duff, K. C., S. M. Kelly, N. C. Price, and J. P. Bradshaw. 1992. The secondary structure of influenza A M2 transmembrane domain: a circular dichroism study. *FEBS Lett.* 311:256–258.

Fujii, J., K. Maruyama, M. Tada, and D. H. MacLennan. 1989. Expression and site-specific mutagenesis of phospholamban: studies of residues involved in phosphorylation and pentamer formation. *J. Biol. Chem.* 264:12950–12955.

Hay, A. J., A. J. Wolstenholme, J. J. Skehel, and M. H. Smith. 1985. The molecular basis of the specific anti-influenza action of amantadine. *EMBO J.* 4:3021–3024.

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.

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

Kimura, Y., K. Kurzydlowski, M. Tada, and D. H. MacLennan. 1997. Phospholamban inhibitory function is activated by depolymerization. *J. Biol. Chem.* 272:15061–15064.

Kovacs, F. A., and T. A. Cross. 1997. Transmembrane four-helix bundle of influenza A M2 protein channel: structural implications from helix tilt and orientation. *Biophys. J.* 73:2511–2517.

Kukol, A., P. D. Adams, L. M. Rice, A. T. Brunger, and I. T. 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.

Lemmon, M. A., J. M. Flanagan, J. F. Hunt, B. D. Adair, B. J. Bormann, C. E. Dempsey, and D. M. Engelman. 1992a. Glycophorin A dimerization is driven by specific interactions between transmembrane alpha-helices. *J. Biol. Chem.* 267:7683–7689.

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

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

Ponder, J. W., and F. M. Richards. 1987. Tertiary templates for proteins: use of packing criteria in the enumeration of allowed sequences for different structural classes. *J. Mol. Biol.* 193:775–791.

Simmerman, H. K. B., and L. R. Jones. 1998. Phospholamban: protein structure, mechanism of action, and role in cardiac function. *Physiol. Rev.* 78:921–947.

Simmerman, H. K. B., Y. M. Kobayashi, J. M. Autry, and L. R. 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.

Torres, J., P. D. Adams, and I. T. 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.

Torres, J., A. Kukol, and I. T. Arkin. 2001. Mapping the energy surface of transmembrane helix-helix interactions. *Biophys. J.* 81:2681–2692.

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