Transmembrane Helix−Helix Interactions: Comparative Simulations of the Glycophorin A Dimer†

Jonathan M. Cuthbertson, Peter J. Bond, and Mark S. P. Sansom*

Department of Biochemistry, University of Oxford, South Parks Road, OX1 3QU, U.K.

Received June 1, 2006; Revised Manuscript Received October 2, 2006

ABSTRACT: The glycophorin helix dimer is a paradigm for the exploration of helix−helix interactions in integral membrane proteins. Two NMR structures of the dimer are known, one in a detergent micelle and one in a lipid bilayer. Multiple (4 × 50 ns) molecular dynamics simulations starting from each of the two NMR structures, with each structure in either a dodecyl phosphocholine (DPC) micelle or a dimyristoyl phosphatidylcholine (DMPC) bilayer, have been used to explore the conformational dynamics of the helix dimer. Analysis of the helix−helix interaction, mediated by the GxxxG sequence motif, suggests convergence of the simulations to a common model. This is closer to the NMR structure determined in a bilayer than to micelle structure. The stable dimer interface in the final simulation model is characterized by (i) Gly/Gly packing and (ii) Thr/Thr interhelix H-bonds. These results demonstrate the ability of extended molecular dynamics simulations in a lipid bilayer environment to refine membrane protein structures or models derived from experimental data obtained in protein/detergent micelles.

Membrane proteins play a central role in many aspects of cell biology (*1*), including transport, signalling, and energy transport. About 25% of all genes code for membrane proteins (*2*). Over the past few years, there has been considerable progress in membrane protein structural biology (*3*), which has resulted in high-resolution structures for ∼110 distinct membrane proteins (see http://blanco.biomol.uci.edu/Membrane_Proteins_xtal.html for a summary). The majority of membrane proteins are formed by bundles of transmembrane (TM)[1] α-helices. It is now well-established that folding of α-helical membrane proteins proceeds via two stages: in the first stage stable TM α-helices are inserted into a lipid bilayer; in the second stage, the TM helices self-assemble into a helix bundle. This model was initially derived from studies largely of in vitro refolding of membrane proteins (*4*, *5*). More recently, it has been shown that a similar model may operate for in vivo membrane protein folding (*6*), reflecting the mechanism of translocon-mediated membrane protein insertion (*7*−*10*).

Glycophorin A (GpA) is a relatively simple erythrocyte membrane protein composed of 131 amino acid residues found in three domains: a short hydrophilic cytosolic domain, a hydrophobic TM domain (of ∼25 residues), and a more substantial extracellular domain that bears extensive carbohydrate modification (*11*). The latter domain contains approximately 12 O-linked and a single N-linked glycan (*12*) and is a major carrier of sialic acid. GpA forms a symmetrical homodimer (*13*) and has been shown to dimerize both in detergent micelles (*14*−*17*) and in a membrane (*18*, *19*) as a consequence of specific interactions between the TM helices (*20*, *21*). The structure of the TM domain of the dimer has been determined by solution NMR in dodecyl phosphocholine (DPC) micelles (*15*) (pdb code 1AFO) and by solid-state NMR in dimyristoyl phosphatidylcholine (DMPC) and palmitoyl oleoyl phosphatidylcholine (POPC) bilayers (*22*). The helices pack together in a right-handed manner, with a crossing angle of ∼40°.

GpA is of some biomedical interest as the cellular receptor for the malarial parasite *Plasmodium falciparum*, via binding of parasite ligands to sialic acid on GpA (*23*−*25*). Interestingly, GpA content can be reduced or absent, due to genetic defects, without known clinical consequences (*26*), although absence of GpA in red cells does affect the structure and anion transport properties of Band 3 (*27*).

GpA is a well-studied model system for investigating helix−helix interactions within membranes. The dimer is stable in sodium dodecyl sulfate (SDS), and hence mutagenesis studies were performed using SDS polyacrylamide gel electrophoresis to assess dimer stability (*14*). More recent studies have used the TOXCAT assay, which employs an in vivo expression system to measure TM helix association in the *Escherichia coli* inner membrane (*18*, *19*, *28*). The sensitivity of GpA dimerization to mutation was found to be similar in membranes and in detergent micelles. Nonetheless, several significant differences were noted including mutations to polar residues, which are generally disruptive in SDS but exhibit sequence specificity in membranes (*19*). This suggests that environment may influence TM helix oligomerization of GpA.

Molecular dynamics (MD) simulations offer a method for studying the conformational dynamics of membrane proteins

† This work was supported by grants from the Wellcome Trust and the BBSRC (as part of the MPSI consortium). J.C. was supported by an MRC studentship.

* Corresponding author. E-mail: mark.sansom@bioch.ox.ac.uk; phone: 44-1865-275371; fax: 44-1865-275273.

[1] Abbreviations: DMPC, dimyristoyl phosphatidylcholine; DPC, dodecyl phosphocholine; GpA, glycophorin A; MAS, magic angle spinning; MD, molecular dynamics; MSF, mean square fluctuation; POPC, palmitoyl oleoyl phosphatidylcholine; RMSD, root mean square deviation; TM, transmembrane.

10.1021/bi0610911 CCC: $33.50
Published on Web 11/09/2006

Table 1: Summary of Simulations[a]

environment	GpA model	temp (K)	no. of detergents/ lipids	no. of waters	box size (nm^3)
DMPC bilayer	GpA-S	310	115	3678	$6.4 \times 5.2 \times 7.5$
	GpA-E	310	117	3617	$6.3 \times 5.3 \times 7.4$
DPC micelle	GpA-S	313	60	14,960	$8.9 \times 8.7 \times 6.5$
	GpA-E	313	60	15,000	$8.9 \times 8.7 \times 6.5$

[a] All simulations were of 50 ns duration.

Table 2: Structural Drift of the Two Models of GpA[a]

environment	GpA model	secondary structure Cα rmsd (Å)	seven-residue dimerization motif Cα rmsd (Å)
DMPC bilayer	GpA-S	1.3 ± 0.1	0.7 ± 0.1
	GpA-E	1.3 ± 0.1	1.1 ± 0.1
DPC micelle	GpA-S	0.9 ± 0.1	0.7 ± 0.1
	GpA-E	1.3 ± 0.1	1.1 ± 0.1

[a] Structural drift of the two models of GpA, from their starting structures for secondary structural elements and for the seven-residue dimerization motif, averaged over the last 10 ns of the simulations.

(*29*, *30*). For example, MD simulations have been used to compare the dynamics of membrane proteins in a detergent micelle vs a lipid bilayer (*31*, *32*). A number of simulation studies of glycophorin have been performed, focusing on, for example, its interactions with lipid bilayers (*33*), self-assembly of glycophorin/detergent micelles (*34*, *35*), and estimation of the energetics of helix dimerization (*36*). More recently, it has also been studied by coarse-grained MD simulations (*37*), focusing on mechanisms of helix-dimer self-assembly in a lipid bilayer environment.

Here we perform extended (50 ns) atomistic MD simulations of the GpA helix dimer. Simulations based on two starting structures (the NMR structures in a detergent micelle and in a lipid bilayer) are performed. For each starting structure, two simulations are performed, one in a detergent micelle and one in a lipid bilayer. The results indicate convergence of the simulations to a common structure, stabilized by Thr/Thr side chain H-bond interactions between the two monomers.

METHODS

Initial Structures. Simulations were initiated from two structures (see Table 1), namely, GpA-E, i.e., the solution NMR structure in DPC micelles (*15*) (pdb code 1AFO) and GpA-S, the solid-state NMR structure determined in DMPC bilayers (*22*) (coordinates supplied by S. Smith). To enable a fair comparison only the TM sections of GpA were used. Hence the GpA-E structure, which contains disordered residues 62−71 that are believed to lie outside the bilayer, was truncated to the same size as the GpA-S model (residues 73−95). Since all 20 structures of the GpA-E model are virtually identical in the TM region, model 4 was chosen as a representative model on the basis that it has the highest surface area buried between the two monomers, as calculated using NACCESS (http://wolf.bi.umist.ac.uk/naccess). The Cα rmsd between the two initial structures was 1.4 Å.

Micelle Simulations. Preformed protein/detergent micelle systems were generated using a previously devised protocol (*31*). For both GpA models, the preformed micelles were constructed using 60 DPC detergents on the basis that this is comparable to the aggregation number of 50−60 detergents for pure DPC micelles (*38*). The aromatic side chain belts, which for GpA consist of a phenyalanine (close to the N-terminus) and a tyrosine (close to the C-terminus) residue, were used as a guide to construct an expanded micelle-like torus around the helix bundle. This torus consisted of semicircular planes of DPC detergent molecules in the all-trans configuration (having undergone energy minimization and a brief MD simulation in vacuo) which radiated out like spokes from the protein surface (Figure 1B). The DPC molecules were randomly rotated about their tail axes and placed with the terminal methyl groups of their tails a minimum of 0.5 nm from the protein exterior and relative to each other so as to avoid steric clashes. Each detergent head group was placed approximately equidistant from its nearest neighbors.

Each system was solvated with simple point charge (SPC) waters (*39*) by superimposing a box of water followed by removal of any water molecules that were too close to any of the other system components. Equal numbers of sodium and chloride counter ions were added to the systems by random replacement of water molecules, to a final concentration of ∼0.1 M NaCl. The systems were then subjected to relatively short MD runs of 1 ns duration to allow relaxation of water molecules and DPC molecules to pack around the protein. During this equilibration period, the positions of all non-H GpA atoms were harmonically restrained with a force constant of 1000 kJ mol^{-1} nm^{-2}. Finally position restraints were released and 50 ns production runs were carried out for both systems.

Bilayer Simulations. Each GpA model was embedded in pre-equilibrated DMPC bilayers using the previously described protocol (*40*). Similar to the micelle systems the aromatic belts of GpA were used to guide insertion into the bilayer. The resulting protein bilayer systems (Figure 1C) contained 115 DMPC for GpA-S and 117 DMPC molecules for GpA-E. Following insertion, SPC water and ions equivalent to ∼0.1 M NaCl were added giving ca. 33 water molecules per lipid molecule and ensuring complete solvation of the bilayer. The systems were then energy minimized and underwent 1 ns of MD during which the protein coordinates were restrained in the same fashion as for the micelle simulations. Then fully unrestrained production runs of 50 ns were performed for both systems.

Simulation Methods. All simulations were performed using the GROMACS 3.1.4 simulation package (*41*) (http://www.gromacs.org). The bilayer simulations used a force field based on GROMOS87 (*42*). Micelle simulations used an extended united atom version of the GROMOS96 force field (*43*). Prior to MD simulations, each system was energy minimized (using <1000 steps of the steepest descents method) to relax any steric conflicts generated during the setup. All MD simulations were performed using constant temperature, pressure, and number of particles (NPT) and periodic boundary conditions. Electrostatics were calculated using particle mesh Ewald (*44*, *45*), with a 1 nm cutoff for the real space calculation. A cutoff of 1 nm was also used for van der Waals interactions. The DPC micelle simulations were run at a temperature of 313 K to match the temperature used to solve the structure of GpA-E in the same environment. For the DMPC bilayer simulations, a temperature of 310 K was used to keep DMPC above the gel−liquid

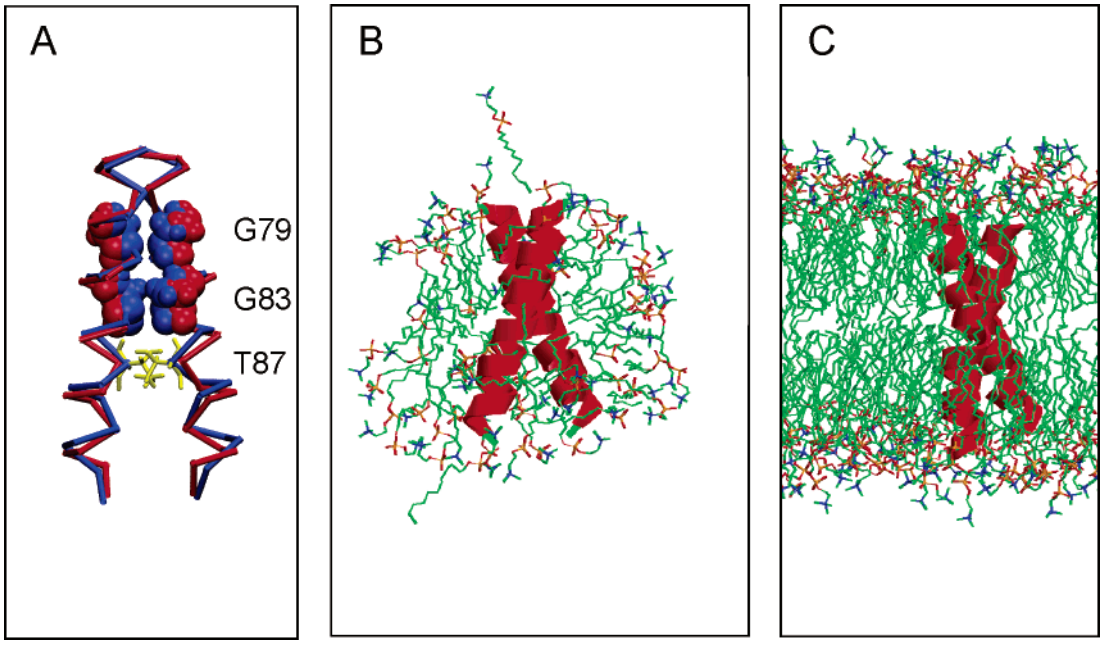

FIGURE 1: (A) NMR structures of GpA. The protein chains are drawn in backbone format (GpA-S in blue and GpA-E in red). The glycine residues (79 and 83) at the dimer interface are shown in spacefill mode. In the GpA-S structure, glycine residues are in van der Waals contact, which also enables interhelical H-bonding by the side chains of Thr87 (yellow bonds). (B) Snapshot from the GpA-S micelle simulation. The protein (red) is surrounded by 60 DPC detergents arranged in a toroidal fashion around the central GpA molecule. The detergents are shown in bonds format. Waters are omitted for clarity. (C) Snapshot from the GpA-S bilayer simulation. The protein (red) molecule is inserted in a bilayer of 115 DMPC detergents. The lipids are shown in bonds format. Again, waters are omitted for clarity.

crystalline phase transition. The temperatures of the protein, the DMPC or DPC, and solvent were each coupled separately, using the Nosé-Hoover thermostat (*46*, *47*) with a coupling constant of $\tau_T = 0.5$ ps. The pressure was coupled at 1 bar using the Parinello-Rahman barostat (*48*, *49*) with coupling constant $\tau_P = 5$ ps. For DMPC and DPC simulations, the compressibility was set to 4.5×10^{-5} bar^{-1} in all box dimensions. The LINCS algorithm (*50*) was used throughout to restrain bond lengths. The time step for integration was 2 fs, and coordinates and velocities were saved every 5 ps. Simulations were performed on dual Pentium III Linux workstations or on a 68-node Linux cluster composed of dual Xeon 4 processors, or on a 32 node cluster, containing 64 750 Mhz Pentium III processors.

RESULTS

Conformational Drift and Flexibility. Four simulations were performed (Table 1) to explore the influence of starting model (GpA-S and GpA-E) and of environment (detergent micelle vs lipid bilayer) on the conformational dynamics of the GpA TM helix dimer. This amounts to a total of 0.2 μs of simulation time, for systems of ~17 000 (bilayers) and ~47 000 (micelles) atoms. The two starting structures are broadly similar, with the TM helices packed together in a right-handed (RH) fashion. However, they differ in the details of their contact residues. Thus, an early model for the GpA TM helix dimer based on solid-state magic angle spinning (MAS) NMR measurements in bilayers suggested that the methyl groups of the sidechains of Val80 and Val84 packed against Gly79 and Gly83, respectively, on the opposing helix (*21*). The complete solution NMR structure of the dimeric TM domain in DPC micelles (GpA-E) (*15*) is consistent with the earlier MAS NMR data (*21*) and exhibits Val−Gly contacts but does not include direct Gly−Gly contacts. A more recent solid-state NMR-derived model of GpA in membrane bilayers (GpA-S) (*22*) shows direct Gly79−Gly79 and Gly83−Gly83 contacts and also additional interhelical interactions between Ile76 and Gly79, and between Val80 and Gly83. These additional interactions restrict the rotational orientation and crossing angle of the helices. Thus, the major difference between the structure determined in membranes (GpA-S) and that determined in detergent micelles (GpA-E) is a rotation of the interacting helical faces by ~25°, which places the glycine residues in direct van der Waals contact in the GpA-S structure allowing interhelical hydrogen bonding (H-bonding) involving Thr87 across the dimer interface (*51*). The solid-state NMR derived membrane structure is consistent with the interhelical NOE constraints from solution NMR measurements, but the detergent structure is not consistent with the high-resolution constraints established by MAS NMR measurements (*22*). According to Smith et al., the crossing angle for the membrane structure is slightly smaller (~35°) than for that in DPC detergent (~40°). They suggest a smaller crossing angle in the bilayer structure allows a more extensive packing interface.

The protocols for both micelle and bilayer simulations had already been developed and tested using other membrane proteins, including OmpA (*31*) (a β-barrel protein) and GlpF (*32*) (an α-helical bundle). For these more complex proteins, extended simulations revealed only subtle changes in dynamics between the two environments, with little change in structure, other than in the extramembraneous loops. However, given the differences between the two NMR structures, one might expect a greater degree of conformational change in the GpA dimer in response to a change in environment.

The progress of the four simulations may be assessed visually (Figure 2). In all four simulations, the GpA helix dimer remains intact throughout the simulation, and little

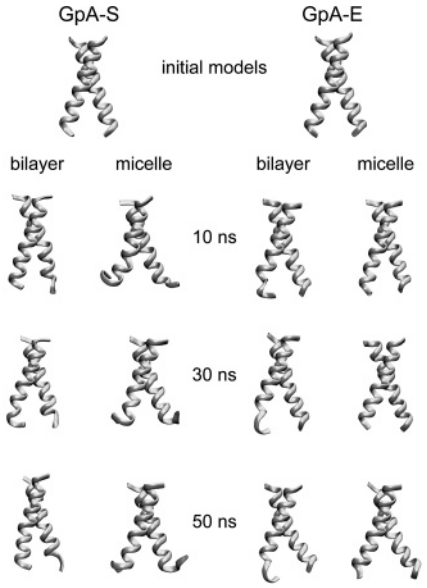

FIGURE 2: Snapshots of the helix dimer (gray ribbons) taken from the simulations of GpA-S and GpA-E in bilayer and micelle environments at the start of each simulation, and at 10, 30, and 50 ns.

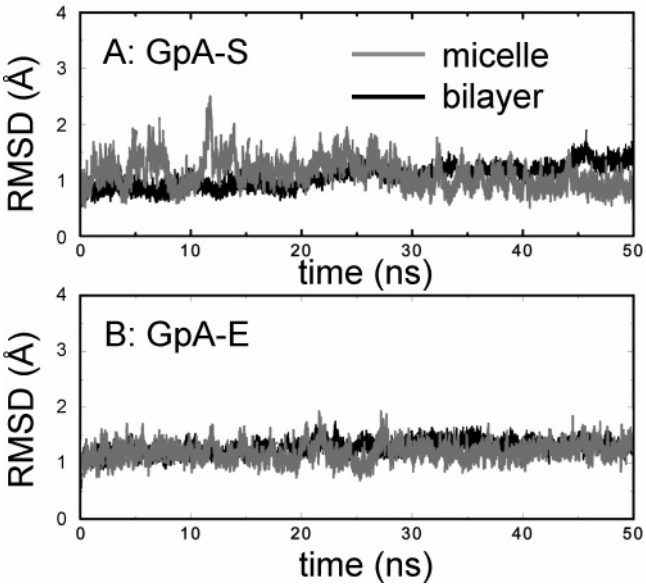

FIGURE 3: Cα rmsd values of the secondary structure elements (i.e., the TM helices) as a functions of time for (A) the GpA-S dimer and (B) the GpA-E model, in a DMPC bilayer (black line) and a DPC micelle (gray line).

helix distortion occurs, other than occasional helix unwinding at the termini which is to be expected. However, it is evident that there is some difference in behavior between the bilayer and the micelle environments, with a greater degree of variation in the packing of the helices in the micelle simulations. In particular, there appears to be a degree of drift to larger helix crossing angles in the two micelle simulations.

The overall extent of conformational drift with respect to time can be assessed via measurement of the root-mean-square deviation (rmsd) of the Cα atoms from the initial structure for each simulation (Figure 3). This comparison reveals a relatively low final Cα rmsd ($<\sim$1.5 Å) for each of the simulations, confirming the overall stability of the TM

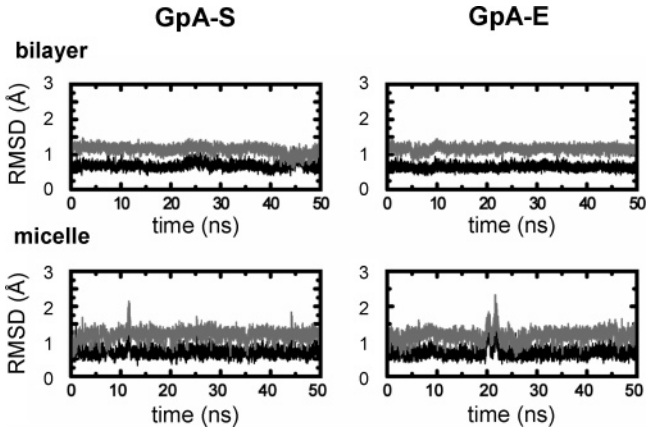

FIGURE 4: Conformational drift of the seven-residue dimerization motif for the GpA-S (left) and GpA-E (right) simulations. Each graph shows the Cα rmsd of this motif for a given simulation relative to both the GpA-S initial structure (black line) and the GpA-E initial structure (gray line).

dimer structures. However, there are subtle differences in behavior for the different environments. For both of the initial structures, there are significant fluctuations in the rmsd for the micelle simulation, suggesting that the dimers may sample different conformational states. In contrast, in the lipid bilayer environment following an initial equilibration period the Cα rmsd does not exhibit any substantial fluctuations, reflecting a lesser degree of conformational flexibility of the GpA dimer in this environment.

One may compare the relative flexibility of the protein in each simulation by calculating the mean square fluctuation (MSF) of the Cα atoms as a function of time window (see refs *52* and *53*) for a detailed discussion of this method). The results of such analysis (data not shown; see Supporting Information) indicate that for all four simulations the *relative* values of the MSFs remain more or less constant over the analysis time windows. On average the ratio of Cα MSFs for micelle vs bilayer environment is 1.9:1. Thus, the GpA dimer is $\sim$2$\times$ more flexible in a detergent micelle than in a phospholipid bilayer. This is consistent with previous results for OmpA and GlpF, although the difference in flexibility between the micelle and the bilayer was only $\sim$1.5$\times$ for the two more complex proteins.

Convergence of Simulations. Mutagenesis and other studies indicate that the GpA TM helix contains a seven-residue motif that is essential for dimerization ($L^{75}I^{76}XXG^{79}V^{80}$-$XXG^{83}V^{84}XXT^{87}$) (*5*, *54*). The conformational drift of this dimerization motif can be estimated in terms of the rmsd for the Cα atoms of the seven residues. The resultant rmsd values for the dimerization motif are relatively low ($\sim$1 Å or less), indicative of the stability of this region. In addition to estimating the conformational drift from the initial structure of each simulation, it is also informative to calculate the Cα rmsd of one simulation (e.g., GpA-S in the micelle simulation) relative to the starting structure of the other (i.e., GpA-E). Comparing the RMSDs of the seven-residue motif from the two starting structures of GpA, it is found that for all four simulations, the drift from the initial GpA-E structure is greater than from the initial GpA-S structure (Figure 4). Thus, by this metric, GpA-S is a more representative structure for the dimer not only in a DMPC bilayer environment but also in a DPC micelle environment. Even the simulation of GpA-E in DPC micelles drifts further from the initial GpA-E

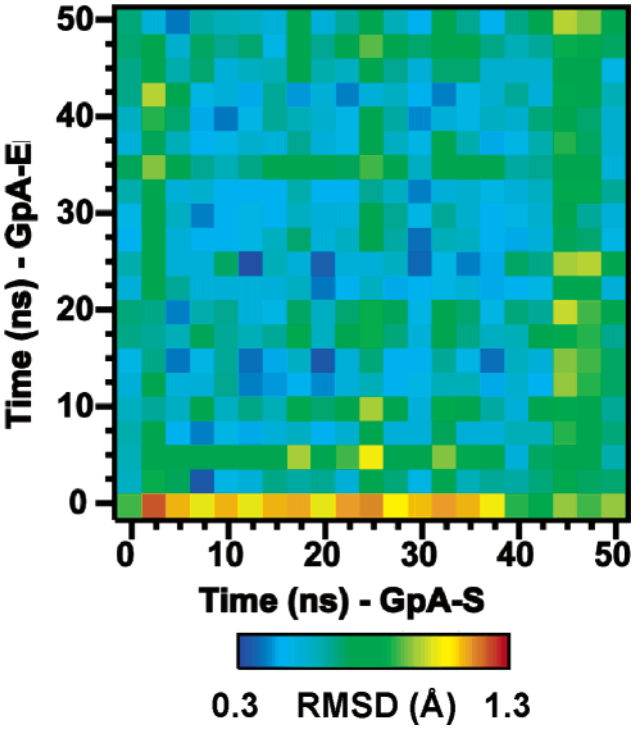

FIGURE 5: Matrix showing the Cα rmsd (for the seven essential residues) for each pair of structures in the GpA-S and GpA-E bilayer simulations.

structure than from the initial GpA-S structure. Interestingly, in the GpA-E simulations, the seven-residue motif very quickly (in <20 ps) drifts further from the initial GpA-E than from the initial GpA-S structure, suggesting that the GpA-S structure may more accurately model the most stable contacts at the TM helix interface region.

On the basis of this, we may ask if the pairs of simulations in the same environment converge to a common structure. This may be achieved by evaluating a matrix containing the seven-residue Cα RMSDs of all pairs of structures from the GpA-S bilayer and GpA-E bilayer simulations (Figure 5). For the first ~0.2 ns of the GpA-E trajectory, the rmsd from structures throughout the GpA-S trajectory is ~1 Å for the seven Cα atoms. Subsequently, the GpA-E simulation structures are between ~0.3 and ~0.8 Å from all GpA-S simulation structures. Thus, while in the first ~0.2 ns the GpA-S simulation actually drifts further away from the initial GpA-E structure, in the GpA-E simulation the seven-residue dimerization motif actually converges toward the GpA-S simulation.

Helix Crossing Angles. The two initial models of GpA differ in the crossing angle (as defined in Figure 6) between the two helices. Both dimer structures contain right-handed, parallel helices, but according to Smith et al. (*22*) the GpA-S dimer has a crossing angle of ~35° compared to an angle of ~40° for the GpA-E structure in DPC micelles.

Crossing angles were calculated for both initial models, with the start and ends of each helix defined using DSSP (*55*). The GpA-E model (PDB id 1AFO) consists of an ensemble of 20 structures, with a mean crossing angle of 42° ± 4°. Model 4 of the ensemble was used in this study, which has a crossing angle of 42°. The corresponding angle for the initial GpA-S model calculated using our method was 43°, suggesting that the starting structures actually have almost identical crossing angles.

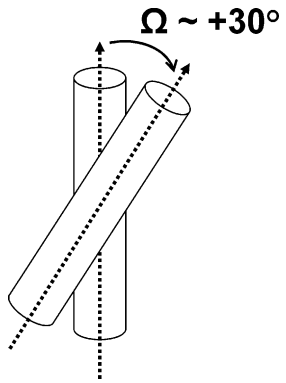

FIGURE 6: Helix−helix crossing angle. Two parallel helices pack in a right-handed manner (as shown in the figure) if the interhelical angle Ω is between 0° and +90°.

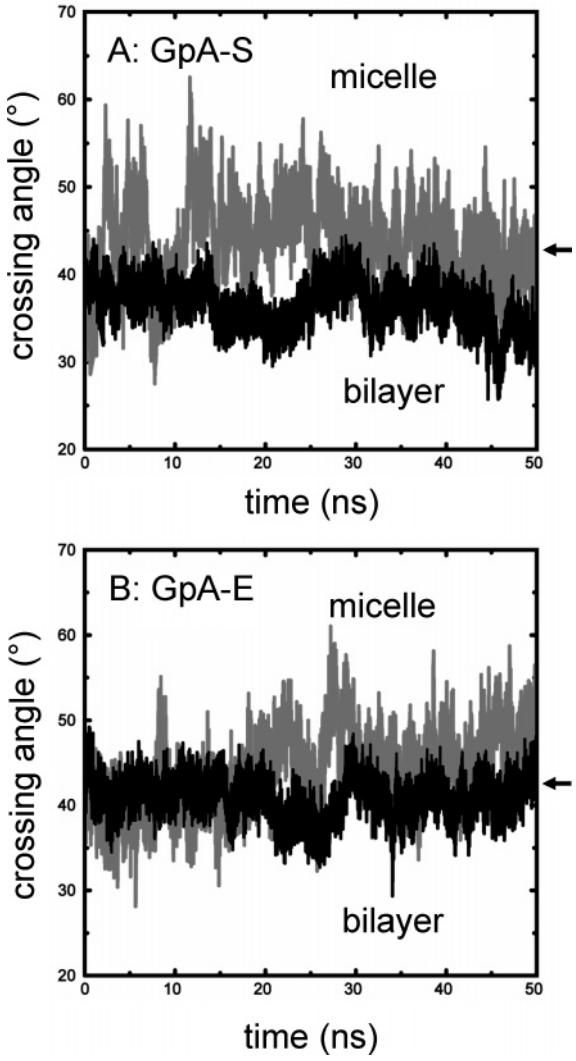

FIGURE 7: (A) Helix−helix crossing angle as a function of time for GpA-S in a DPC micelle (gray line) and a DMPC bilayer (black line). (B) Helix−helix crossing angle as a function of time for GpA-E in a DPC micelle (gray line) and a DMPC bilayer (black line). In each graph, the crossing angle of the starting structure is shown by a small horizontal arrow.

For the GpA-S simulations, larger crossing angles are sampled in the DPC micelle environment than in DMPC bilayers (Figure 7A). The fluctuations in helix crossing angle are also greater in the micelle environment. Thus, for GpA-S the average crossing angle (and associated standard deviation)

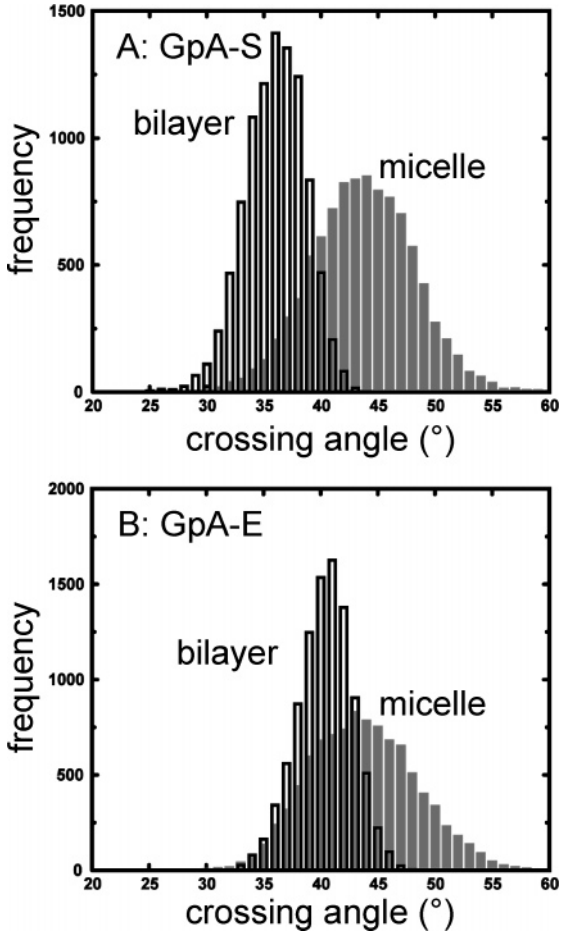

FIGURE 8: (A) Distribution of crossing angles for GpA-S in a micelle (gray bars) and bilayer (black bars) environment. (B) Distribution of crossing angles for GpA-Eng in a micelle (gray bars) and bilayer (black bars) environment.

in a bilayer (Figure 8A) is 37° ± 3°, compared with 44° ± 5° in a micelle.

In the case of the GpA-E simulation, again the trajectory of the crossing angle as a function of time (Figure 7B) and the crossing angle distribution (Figure 8B) reveal that the helices sample a significantly larger range of crossing angles in the micelle environment compared to the bilayer. Thus, in DPC the mean crossing angles 44° ± 5°, i.e., the same as for the GpA-S micelle simulation. However, in the bilayer environment the mean crossing angle of 41° ± 2° for GpA-E is a little higher than the corresponding average of 37° ± 3° for GpA-S.

To analyze the statistical difference between simulations, the trajectories over the 25−50 ns period were split into blocks of 5 ns windows, allowing the calculation of the mean and standard errors over the fully equilibrated latter halves of each simulation. This confirmed that for both models of GpA, the micelle environment resulted in greater crossing angles than the bilayer, with means and associated standard errors of 43° ± 2° (micelle) and 36° ± 2° (bilayer) for GpA-S, and similarly, 46° ± 2° (micelle) and 41° ± 1° (bilayer) for GpA-E.

Interhelix Distances—Defining the Interface. The nature of the difference between the two initial models can be seen if one examines the inter-residue distances across the helix−helix interface (Table 3). It can be seen that at the contacts of Gly79 with Gly79, and of Gly83 with Gly83 the helices approach more closely in the GpA-S model than in the GpA-E model, whereas the contacts made by the Val80 and Val84 sidechains are closer for GpA-E than for GpA-S.

Table 3: Interhelical Distances from NMR[a]

helix A	helix B	GpA-S distance (Å)	GpA-E distance (Å)
Gly79 (Cα)	Gly79 (C═O)	4.1	5.2
Ile76 (C═O)	Gly79 (Cα)	4.8	5.3
Gly83 (Cα)	Gly83 (C═O)	4.3	5.2
Gly83 (Cα)	Val80 (C═O)	4.2	4.5
Gly79 (C═O)	Val80 (CH_3)	4.0	2.8
Gly83 (C═O)	Val84 (CH_3)	4.0	3.4

[a] Interhelical distances from solid-state NMR measurements (GpA-S) of the GpA TM dimer and from the model of the dimer based on solution NMR NOE data (GpA-E).

The corresponding set of average interhelix distances are summarized for all four simulations in Table 4. The difference from the initial structure distances (Table 3) is quite striking. For each pair of atoms, the mean interhelical distances are essentially the same across all four simulations, regardless of starting model or environment. This adds weight to the earlier structural drift analysis that revealed stable contacts for the seven essential residues involved at the dimer interface.

That the mean interhelical distances for these key residues are very similar in both DMPC bilayer and DPC micelle simulations suggests that the residues at the dimer interface favor a specific set of contacts independent of (membrane-like) environment. On the basis of these average values, all distances, apart from the Gly83 Cα to Val80 C═O distance (which is closer to the GpA-E distance), are closer to the values quoted for the GpA-S model. This would explain why when comparing the Cα RMSDs of the seven essential residues across all simulations, the dimer always drifts further from the GpA-E initial structure than from the GpA-S initial structure.

To strengthen this conclusion, we examined individual distributions of the interhelix distances (Figure 9). Comparing both within a particular environment, and between environments, for all but the Ile76 C═O and Gly79 Cα interaction (in the bilayer environment), both initial structures yielded very similar distributions, again suggesting a convergence of the simulations.

For simulations based on both initial structures, and in both DPC micelles and DMPC bilayers, the modes of five out of the six interhelix distances measured are closer to the GpA-S initial distance than to the GpA-E initial distance. The exception is the Gly83 Cα and Val80 C═O distance where both models have very similar starting distances. For both simulations in a bilayer, the mode for this distance is 4.5 Å, i.e., the same as the value seen in the GpA-E starting structure. For the micelle simulations the modal distances are slightly closer to the 4.2 Å seen in the initial GpA-S model.

In summary, the simulations suggest, both on the basis of these six atom-pair distances and of the Cα rmsd of the seven essential residues, that across all simulations (both micelle and bilayer) the GpA-S model is a better representative of the range of structures sampled than the GpA-E structure.

Energetics and Hydration. It is informative to compare the potential energies of interaction between the helices in

Table 4: Mean Interhelical Distances[a]

helix A	helix B	bilayer: GpA-S (Å)	bilayer: GpA-E (Å)	micelle: GpA-S (Å)	micelle: GpA-E (Å)
Gly79 (Cα)	Gly79 (C═O)	4.1 ± 0.3	4.0 ± 0.3	4.0 ± 0.4	4.0 ± 0.4
Ile76 (C═O)	Gly79 (Cα)	4.9 ± 0.5	4.5 ± 0.3	4.7 ± 0.6	4.7 ± 0.6
Gly83 (Cα)	Gly83 (C═O)	4.5 ± 0.4	4.3 ± 0.4	4.1 ± 0.3	4.1 ± 0.4
Gly83 (Cα)	Val80 (C═O)	4.6 ± 0.3	4.7 ± 0.3	4.4 ± 0.4	4.5 ± 0.4
Gly79 (C═O)	Val80 (CH_3)[b]	4.1 ± 0.3	4.1 ± 0.3	4.2 ± 0.4	4.2 ± 0.3
Gly83 (C═O)	Val84 (CH_3)[b]	4.6 ± 0.4	4.5 ± 0.4	4.5 ± 0.5	4.5 ± 0.5

[a] Mean interhelical distances for GpA-Smith (red) and GpA-Eng (black parentheses) in bilayer and micelle environments. [b] The Val side chain CH_3 used to calculate the distance was always the one that was closest to the Gly carbonyl group.

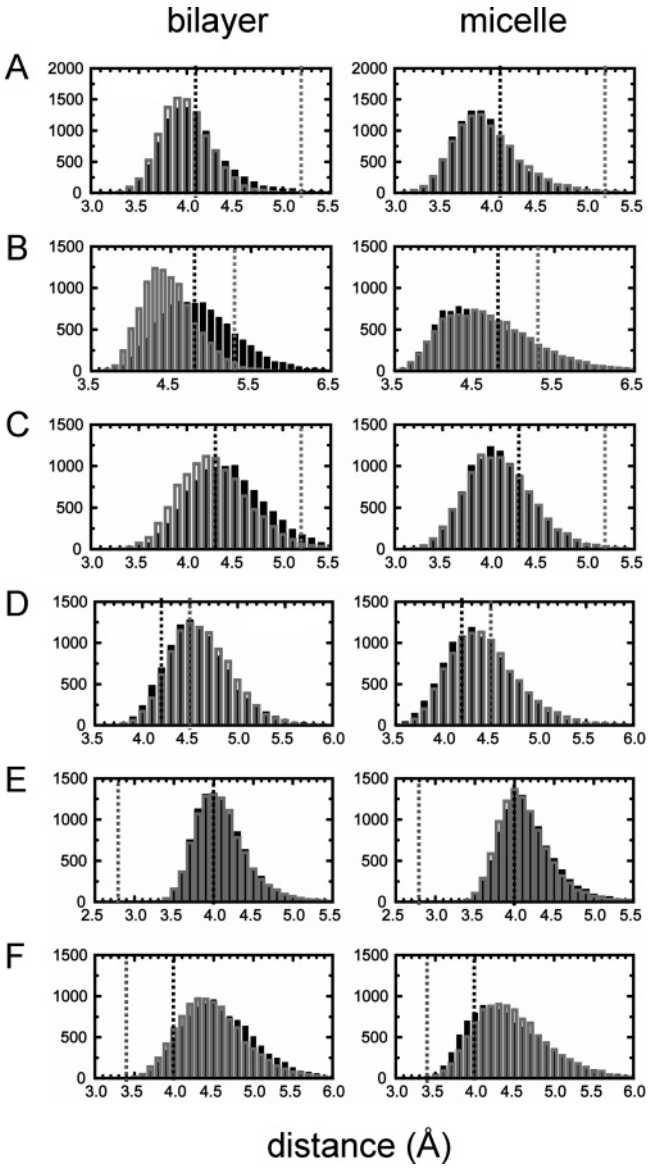

FIGURE 9: Distributions of interhelical distances in bilayer and micelle environments between (A) Gly79 Cα and Gly79 C═O; (B) Ile76 C═O and Gly79 Cα; (C) Gly83 Cα and Gly83 C═O; (D) Gly83 Cα and Val80 C═O; (E) Gly79 C═O and Val80 CH_3; and (F) Gly83 C═O and Val84 CH_3. For E and F the valine CH_3 group used is always the one that yields the smaller distance. The distributions for GpA-S simulations are in black and for GpA-E are in gray. The interhelical distances for the two initial structures are shown with vertical dashed lines (black and gray respectively).

the different models and simulations (Table 5). The most striking difference is between the simulations in a bilayer and in a micelle. In the micelle, the van der Waals interactions are noticeably weaker than in the bilayer. This mirrors the great flexibility and more pronounced fluctuations in crossing angle for the micelle simulations noted above. When in the same environment, the total interaction energies (both van der Waals and electrostatic) between helices are very similar for GpA-S and GpA-E. This is consistent with convergence of each model to a common structure. For both GpA-S and GpA-E, the electrostatic interaction energies between helices are approximately constant between the bilayer vs micelle environments. The main contribution to this unchanging interaction is H-bonding of the Thr87 residues on opposing helices to one another.

The van der Waals interaction energies are more favorable for the bilayer than micelle simulations, by ~60 kJ mol^{-1}. Part of this is due to changes in interaction around the nonpolar residues of the dimerization motif. However, a significant source of the overall change in van der Waals interaction between pairs of helices in different environments is from residues outside of the dimerization motif. This is consistent with the larger crossing angle observed in the micelle relative to the bilayer simulations, which corresponds to reduced packing between residues furthest away from the dimerization motif.

A possible explanation for the difference in behavior between the micelle and the bilayer simulations was thought to reside in the penetration of water into the otherwise hydrophobic interior, which might be anticipated to perturb helix/helix interactions. We therefore analyzed the numbers of H-bonds between water molecules and each helix of the dimer in all four simulations (Figure 10). For a given environment (bilayer vs micelle), similar patterns of water penetration were seen for GpA-S and GpA-E, with H-bonds limited to the residues at either end of the α-helices. Perhaps surprisingly, a similar (neglible) extent of penetration of water around the core of the helix dimers was seen for the micelle and the bilayer environments. Hence, access of water to the dimerization motif is the same in both environments, consistent with similar $Thr^{87}-Thr^{87}$ electrostatic interaction energies in the simulations. Instead, the main difference between the two environments appears to be outside of the hydrophobic interior, with a greater number of H-bonds with water at the ends of helices for the micelle in comparison with the more constraining bilayer environment.

Nature of the Converged Model Structure. To characterize the stable GpA TM helix dimer, it is useful to measure distances of closest approach between the two helices. The mean minimum interhelical distance between the backbone atoms during the simulations is ~3.4 Å for both models in both environments. The smallest distance and therefore the closest approach between the backbone atoms observed during the simulations is ~3.0 Å and occurs between the two Gly79 residues, and between Val80 and Gly79. This is

Table 5: Energetics of Interactions between Helices[a]

interaction groups	Coulombic and Lennard-Jones Energies (kJ mol^{-1})			
	bilayer: GpA-S	bilayer: GpA-E	micelle: GpA-S	micelle: GpA-E
GpA1 & GpA2	-24 ± 7	-21 ± 7	-22 ± 7	-23 ± 7
	-214 ± 19	-212 ± 13	-151 ± 15	-156 ± 17
motif*(GpA1) & motif*(GpA2)	2 ± 1	2 ± 1	2 ± 1	2 ± 1
	-104 ± 9	-116 ± 7	-89 ± 9	-91 ± 9
motif*(GpA1) & T^{87}(GpA2)	-3 ± 1	-3 ± 2	-1 ± 2	-1 ± 2
	-9 ± 2	-9 ± 3	-7 ± 2	-7 ± 2
motif*(GpA2) & T^{87}(GpA1)	0 ± 3	0 ± 1	-1 ± 2	-2 ± 2
	-9 ± 3	-8 ± 2	-6 ± 2	-6 ± 2
T^{87}(GpA1) & T^{87}(GpA2)	-23 ± 6	-21 ± 7	-23 ± 8	-23 ± 6
	-1 ± 5	-2 ± 5	0 ± 5	-2 ± 4

[a] The potential energies of interaction presented are the mean short-range Coulombic and Lennard-Jones interaction energies between opposing GpA helices, and various subsets of residues, for GpA-Smith and GpA-Eng in bilayer and micelle environments, averaged over the final 10 ns of each simulation. *Critical nonpolar residues of the dimerization motif = L^{75}, I^{76}, G^{79}, V^{80}, G^{83}, and V^{84}.

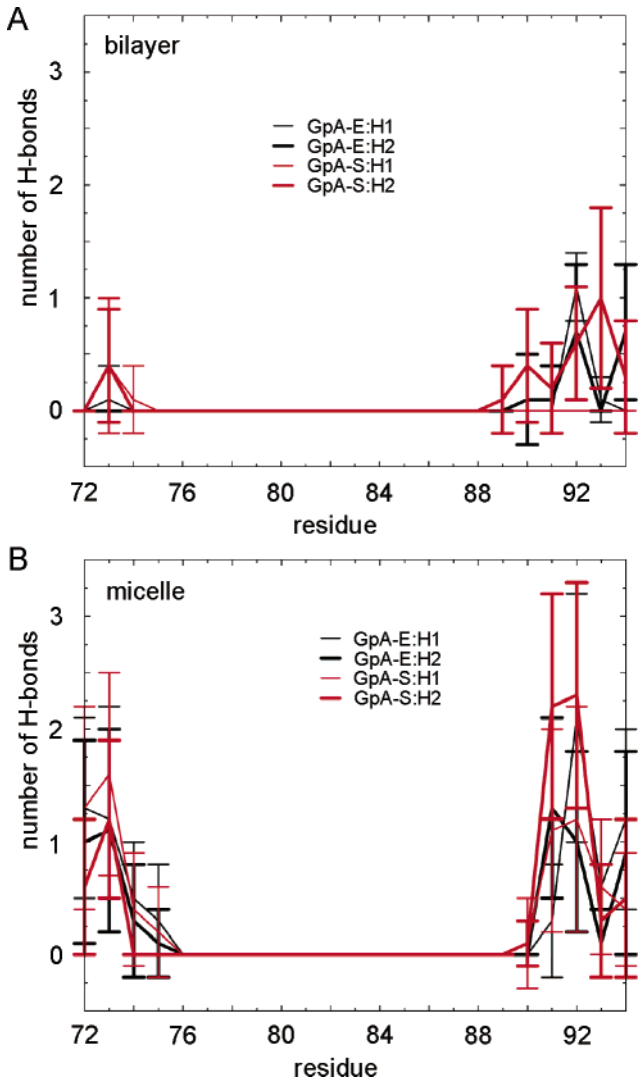

FIGURE 10: Mean number of hydrogen bonds between water molecules and helix 1 (H1; thin lines) or 2 (H2; thick lines) of GpA-E (black) and GpA-S (red), over the final 10 ns of (A) bilayer simulations and (B) micelle simulations.

in good agreement with MD simulations of GpA-E in SDS micelles, which show a closest approach of 3.5 Å at Gly79 and Val80 (*34*). There are also close contacts between the Gly83 residues. These observations are in agreement with the experimental observation (*14*) that Gly79 plays a critical role in GpA dimerization and confirms the view of Smith and colleagues (*22*) that the Gly79−Gly79 and Gly83−Gly83 residue pairs are in close van der Waals contact.

Such close contacts between glycine residues at the dimer interface are believed to allow Thr87 to hydrogen bond across the dimer interface (*51*). Indeed, the closest side chain contacts (~1.5 Å) are between the hydroxyl groups of the two Thr87 residues (Figure 1). Glycine−glycine contacts also support the hypothesis that a small side chain is necessary for close approach of the interacting helices facilitating stabilizing van der Waals interactions and possibly allowing Cα-mediated hydrogen bonds between Gly79, Val80, and Gly83 of the two helices to stabilize the dimer (*56*).

The TM domain of GpA contains mainly hydrophobic amino acids, and Thr87 is the only polar residue in the seven residue dimerization motif with the potential to stabilize the dimer through H-bonding interactions. Four of the NOE restraints used to model GpA-E involved this threonine. Interestingly, the refined structure does not show Thr87 forming an interhelical hydrogen bond. Instead, the β-hydroxyl group hydrogen bonds back to the carbonyl oxygen of Gly83 on the same helix (*15*). Indeed, threonines are among the most common polar residues found in TM domains precisely because they readily hydrogen-bond back to backbone carbonyls on the same helix (*57*).

The close Gly79−Gly79 and Gly83−Gly83 packing seen in the simulations presented here, and the presence of two threonines directly opposite each other in the parallel dimer, could allow Thr87 to form *inter*-helical hydrogen bonds. Mutagenesis studies certainly indicate that replacement of Thr87 with a valine results in partial disruption of the helix dimer (*14*, *58*). The T87V mutation replaces the Thr87 β-hydroxyl with a methyl group of roughly the same molecular volume, suggesting that hydrogen-bonding contributes to dimer stability.

According to Smith and colleagues, in both NMR models of GpA the closest interhelix side chain-backbone distance involving Thr87 is between the γ-methyl group of Thr87 and the amide nitrogen of Ile88. NMR measurements of this distance demonstrated close packing (4.0 ± 0.2 Å) across the dimer interface (*51*). These sites are ~5.0 Å apart in the GpA-E starting structure and ~4.0 Å apart in GpA-S (Figure 1A,B) but range from 3.5 to 4.0 Å apart in molecular dynamics simulations performed by Smith and co-workers (*22*). Smith and colleagues argue that this difference between the two structures is critical because the interhelical space is too large in the detergent structure (GpA-E) to allow direct

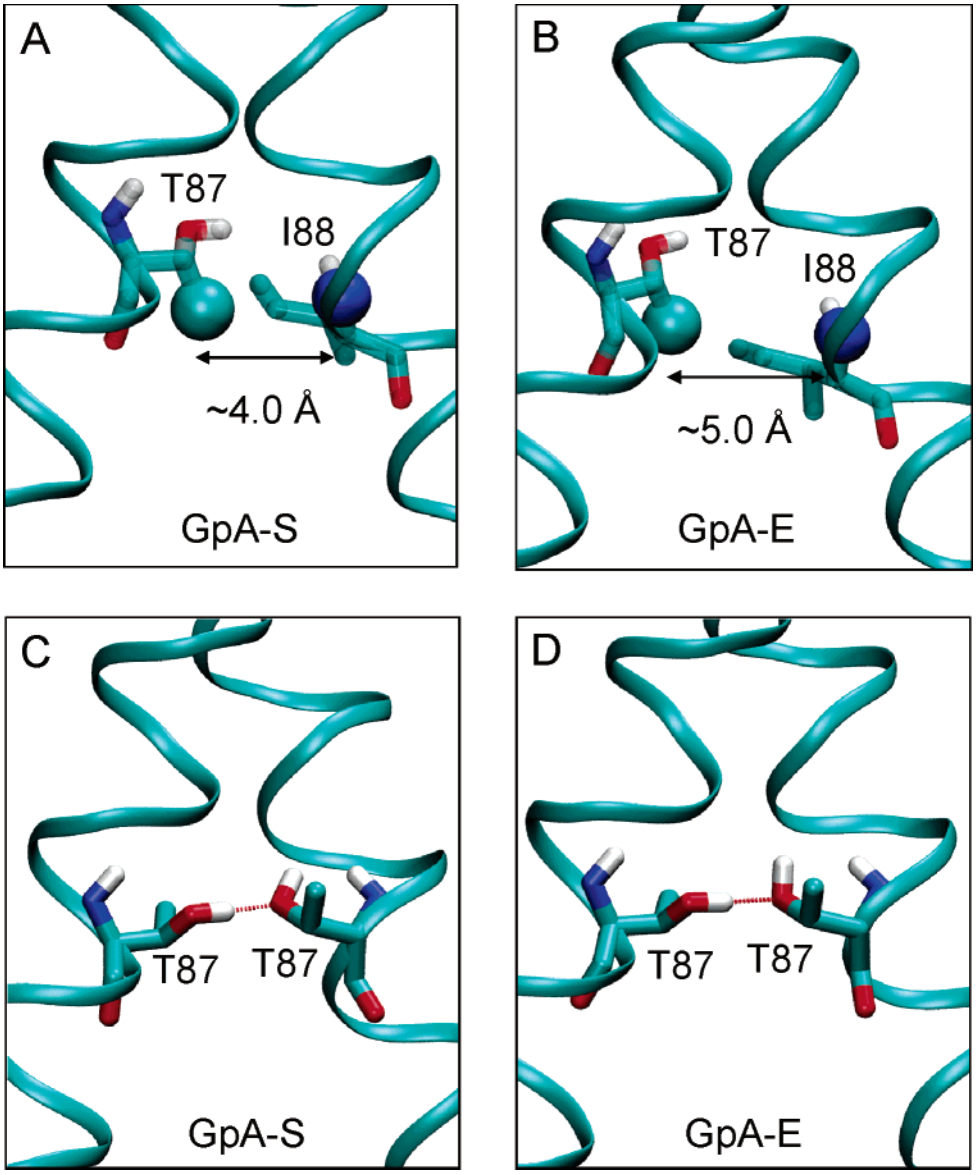

FIGURE 11: (A) Initial structure of GpA-S showing the γ-methyl group of Thr87 (cyan sphere) on helix A in close proximity (~4.0 Å) to the backbone amide nitrogen of Ile88 (blue sphere) on helix B. (B) Initial structure of GpA-E showing that the γ-methyl group of Thr87 on helix A and the backbone amide nitrogen of Ile88 on helix B are further apart (~5.0 Å). (C, D) Snapshots from the simulations: (C) the GpA-S bilayer simulation at 45 ps and (D) the GpA-E bilayer simulation at 230 ps. All Thr87 residues are in the $g+$ conformation. Both snapshots show the side chain of Thr87 (Helix A) forming an interhelical H-bond with that of Thr87 on helix B.

interhelical hydrogen bonding. The short interhelical distance seen by NMR in the bilayer environment places the beta-hydroxyl of Thr87 in GpA-S within H-bonding range of the backbone carbonyl of Val84 on the opposing helix (Figure 11A).

It is worth noting at this point that the model of GpA-S used in these simulations was developed using MD simulations and distance restraints involving Gly79 and Gly83 but not Thr87. Nonetheless, in the GpA-S starting structure the closest side chain-backbone distance involving Thr87 is actually 2.8 Å, between the threonine side chain hydroxyl and the backbone carbonyl carbon of Val84. For GpA-E, the shortest distance involving Thr87 is 3.0 Å between the threonine side chain hydroxyl and the α carbon of Val84. Consequently, in the GpA-E structure Thr87 forms an intrahelical hydrogen bond with the backbone carbonyl oxygen of Gly83 on the same helix (Figure 11B).

It is interesting that in both the GpA-Smith and GpA-Eng starting structures the $\chi 1$ torsion angle is $g-$, indicating that this conformation can accommodate both intrahelical (in the case of GpA-E) and interhelical H-bonding (in the case of GpA-S). At the start of the current simulations, with the threonine residues in a $g-$ conformation, the two threonine side chains do not form a side chain to side chain H-bond as the distance between the hydroxyl groups in both models is ca. 4.5 Å.

However, in all the simulations performed in this study the γ-methyl group of Thr-87 and the amide nitrogen of Ile88 do not remain tightly packed. This is because the $\chi 1$ torsion angle favored in the simulations is $g+$. As a consequence of this, when the Thr87 residues are in a $g+$ conformation the distance between the two Thr87 hydroxyl groups is reduced. During the simulations, the average distance between the center of mass of the two Thr87 hydroxyl groups

is ~2.8 Å bringing them within range to form a direct side chain to side chain interhelical H-bond. On average there is approximately one interhelical hydrogen bond (~0.97) per time frame in all simulations. The overwhelming majority of interhelical H-bonds formed in the simulations involve the threonine residues. Almost all interhelical H-bonds (99%) are between the two hydroxyl side chains of Thr87 ((Figure 11C,D). The remaining 1% of interhelical H-bonds formed involve one of the Thr87 hydroxyl groups as a hydrogen donor and either a Gly83 or Val84 carbonyl oxygen as the acceptor.

These results confirm the importance of the single polar residue present in the seven-residue dimerization motif and suggest that the two Thr87 residues favor forming interhelical hydrogen bonds with each other using their hydroxyl side chains rather than interhelical H-bonds with the protein backbone. By forming interhelical H-bonds they assist tight association and stability of the pair of helices at the dimer interface. The localization of a hydrogen bond between threonine residues located at the same position on both monomers in the dimer is significant since in vivo dimerization studies on the aspartate receptor also reveal that formation of symmetrical bonds enhances stability of the dimer (*59*).

It should be noted that the energetic stabilization from one such interhelical hydrogen bond is uncertain. Threonine can form an intrahelical H-bond in the unfolded (monomeric helical) state. In addition, studies on the introduction of polar side chains into model TM peptides found that residues containing two polar side chain atoms (such as asparagine) have a much greater tendency to drive TM helix association than residues containing only one polar side chain atom (serine or threonine) (*60*, *61*). Nonetheless, a recent study showed that the substitution of either of the two glycine residues in the $G^{79}XXXG^{83}$ motif with a serine could partly stabilize both the mutant homodimer and a heterodimer containing one wild type helix. This suggests that the disruptive packing effect of introducing a serine is partially negated due to serine's ability to form interhelical H-bonds (*62*). Recently, MD simulations were used to study the process of TM helix association of the model peptide MS1 which contains a single asparagine residue in the center of an otherwise hydrophobic helix (*63*). Only dimers with interhelical hydrogen bonds involving this polar residue formed stable structures.

While a single serine or threonine side chain may not be able to promote such helix association per se, it seems likely that the single threonine residue found in the seven-residue dimerization motif of GpA does have a complementary role in mediating tight association of the TM helices through formation of interhelical H-bonds. In addition to this, a portion of the interactive strength may originate from interhelix hydrogen bonds between Cα hydrogens and carbonyl oxygen atoms on the adjacent helix (*64*).

DISCUSSION

These studies indicate that extended atomistic MD simulations may be used to analyze and refine membrane protein models derived from, for example, NMR experiments. In particular, the convergence of the glycophorin dimer to a common structure, stabilized by interhelix H-bonds, suggests that simulations may be used to refine an initial model based on, for example, the structure of a simple membrane protein in a detergent micelle environment, to yield a model more representative of the structure of the protein in a lipid bilayer environment. This is of some general significance in the context of using computational methods to aid structure determination of (simple) membrane proteins. As has been discussed by a number of authors (e.g., refs *6*, *65*−*67*), a combination of indirect structural data and computational modeling may yield valuable insights into membrane protein structures. The current study suggests that extended MD simulations may be used to refine such models and to explore the conformational dynamics of a membrane protein structure in different environments (*31*, *32*, *37*). In combination with recent advances in more approximate MD simulations of membrane protein folding and self-assembly (*37*, *68*), the current results indicate that MD simulations are able to play a key role in prediction of membrane protein structure. More recently, MD simulations have been used to explore asparagine-mediated interhelix H-bonds between simple models of TM helices (*63*).

The GpA helix dimer has been the focus of a number of recent simulation studies. Thus, the free energy of interaction of the GpA TM helices in a membrane mimetic environment (a dodecane slab) has been estimated (*36*), and the GxxxG motif and Thr87 interactions shown to play a role in stabilizing the dimer. A number of atomistic simulations have focused on self-assembly of the glycophorin dimer/detergent micelle (*34*, *35*) and have indicated that the micellar environment helps to stabilize the structure of the helix dimer. More approximate simulation methods using either generalized Born (*69*), other implicit membrane models (*70*, *71*), knowledge-based potentials (*72*), or coarse-grained particle-based simulations (*37*) have all focused on the stability of the GpA helix dimer in a membrane environment. The key role of the membrane environment in the stability of the dimer has also been stressed in a recent modeling paper (*73*). All of these studies demonstrate the stability of the right-handed helix interaction for the GpA TM helices.

A key aspect of our refined structure of the GpA helix dimer is its stabilization by interhelix H-bonds. A number of studies have stressed the more general importance of such interactions in stabilizing transmembrane α-helix bundles (*74*). Thus, for example, aparagine-mediated interhelix H-bonding has been shown to promote self-association of model transmembrane helices based either on the GCN4 leucine zipper (*75*) or on a hydrophobic coiled-coil helix (*76*). These studies have been extended to show that a variety of other H-bonding sidechains can also support self-association of simple TM helices to varying degrees (*60*, *61*), although in these model systems Thr is at best a weak mediator of interhelix H-bonds (*77*).

The results of our simulations can be related to ongoing studies of the mechanisms of helix−helix assembly in membrane proteins, experimental techniques ranging from spectroscopy and mutagenesis to molecular biology and modeling (reviewed in, for example, refs *5*, *6*, *77*). Many of these studies are conducted on membrane proteins in a detergent micelle environment. It is likely that, at least in some cases, there may be subtle changes in helix packing upon transfer to a lipid bilayer environment. Although such refinement of initial micelle model structures can be achieved

by, for example, solid-state NMR studies (*78*−*80*) of proteins in a bilayer environment, we would suggest that molecular simulations may also play a key role in refining such models, in addition to having the potential of providing information on, for example, the nature of lipid−protein interactions (*81*, *82*).

It is important to consider the limitations of the simulations presented in this paper. One issue is that of incomplete sampling, even on timescales of 50 ns (*53*). To obtain complete sampling will probably need much longer (e.g., >1 μs) simulation times and multiple simulations. This can be achieved using, for example, coarse-grained simulations (*37*) or implicit bilayer methods (*69*) but with possible loss of accuracy in representation of the protein and/or its environment. Another possibility worthy of further exploration is the use of Monte Carlo simulations to explore helix dimers (*70*). However, the limitations of incomplete sampling may not be too severe, as indicated by the convergence of the GpA dimer simulations to a common structure shown in the current study.

A second limitation is that of the force-field. The GROMOS force-field treats aliphatic carbons via extended atoms, and thus there is no explicit treatment of possible Cα-H−O hydrogen-bonds (*64*). It might be possible to extend current simulations to use, for example, the OPLS or CHARMM force-fields to explore this further. Preliminary analysis (Samish, Cuthbertson, and Sansom, unpublished observations) of the existing simulations by "adding back" the Hα atoms and using a distance-based approach suggests that Cα-H−O H-bonds may be sustained during the simulations.

Another possible limitation is the use of the partical mesh Ewald (PME) method to approximate long-range electrostatic interactions. This is generally assumed to be best practice for membrane simulations for membrane simulations (ref *83*, but see ref *84*). However, a number of studies have pointed out that this method is not without its limitations (*85*−*87*), and one may wish to examine the effects of cutoff (*88*) or reaction field (*89*) models, at least for the protein/detergent micelle simulations.

There are a number of future directions that could be pursued based on the current studies. From a methodological perspective, it would be attractive to combine coarse-grained (*37*) and atomistic (the current study) simulations of GpA dimers to develop a multi-scale (*90*, *91*) approach to membrane protein modeling. From a biological perspective, it would be of interest to extend these studies to other simple membrane proteins, such as helix-hairpin fragments from more complex membrane proteins (*92*). By combining methodological developments with a wider range of test systems, a simulation-based approach to membrane protein structure prediction may be achievable. This would provide a valuable addition to computational methods to enable structure prediction for those membrane proteins where, for example, X-ray studies are not possible (*93*).

ACKNOWLEDGMENT

Our thanks to our colleagues for their interest in the work, especially Dr. Declan Doyle.

SUPPORTING INFORMATION AVAILABLE

Figure of time-dependence of simulation B-factors of secondary and non-secondary structure elements, for GpA-S and GpA-E simulations in a DMPC bilayer, a DPC micelle, and a membrane-mimetic octane slab environments. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Nath, D. (2005) Membrane biology. *Nature 438*, 577.
2. Wallin, E., and von Heijne, G. (1998) Genome-wide analysis of integral membrane proteins from eubacterial, archean, and eukaryotic organisms. *Protein Sci. 7*, 1029−1038.
3. White, S. H. (2004) The progress of membrane protein structure determination. *Protein Sci. 13*, 1948−1949.
4. Popot, J. L., and Engelman, D. M. (1990) Membrane protein folding and oligomerization: the two-state model. *Biochem. 29*, 4031−4037.
5. Popot, J. L., and Engelman, D. M. (2000) Helical membrane protein folding, stability, and evolution. *Ann. Rev. Biochem. 69*, 881−922.
6. Bowie, J. U. (2005) Solving the membrane protein folding problem. *Nature. 438*, 581−589.
7. van den Berg, B., Clemons, W. M., Collinson, I., Modis, Y., Hartmann, E., Harrison, S. C., and Rapoport, T. A. (2003) X-ray structure of a protein-conducting channel. *Nature 427*, 36−44.
8. Clemons, W. M., Ménétret, J.-F., Akey, C. W., and Rapoport, T. A. (2004) Structural insight into the protein translocation channel. *Curr. Opin. Struct. Biol. 14*, 390−396.
9. Mitra, K., Schaffitzel, C., Shaikh, T., Tama, F., Jenni, S., Brooks, C. L., Ban, N., and Frank, J. (2005) Structure of the *E. coli* protein-conducting channel bound to a translating ribosome. *Nature 438*, 318−324.
10. Hessa, T., Kim, H., Bihlmaier, K., Lundin, C., Boekel, J., Andersson, H., Nilsson, I., White, S. H., and von Heijne, G. (2005) Recognition of transmembrane helices by the endoplasmic reticulum translocon. *Nature 433*, 377−381.
11. Tomita, M., Furthmayr, H., and Marchesi, V. T. (1978) Primary structure of human erythrocyte glycophorin A. Isolation and characterization of peptides and complete amino acid sequence. *Biochem. 17*, 4756−4770.
12. Zdebska, E., and Koscielak, J. (1999) A single-sample method for determination of carbohydrate and protein contents glycoprotein bands separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis. *Anal. Biochem. 275*, 171−179.
13. Furthmayr, H., and Marchesi, V. T. (1976) Subunit structure of human erythrocyte glycophorin A. *Biochem. 15*, 1137−1144.
14. Lemmon, M. A., Flanagan, J. M., Treutlein, H. R., Zhang, J., and Engelman, D. M. (1992) Sequence specificity in the dimerisation of transmembrane a helices. *Biochemistry 31*, 12719−12725.
15. MacKenzie, K. R., Prestegard, J. H., and Engelman, D. M. (1997) A transmembrane helix dimer: structure and implications. *Science 276*, 131−133.
16. Fisher, L. E., Engelman, D. M., and Sturgis, J. N. (1999) Detergents modulate dimerization, but not helicity, of the glycophorin A transmembrane domain. *J. Mol. Biol. 293*, 639−651.
17. Fisher, L. E., Engelman, D. M., and Sturgis, J. N. (2003) Effect of detergents on the association of the glycophorin a transmembrane helix. *Biophys. J. 85*, 3097−3105.
18. Langosch, D., Brosig, B., Kolmar, H., and Fritz, H. J. (1996) Dimerisation of the glycophorin A transmembrane segment in membranes probed with the ToxR transcription activator. *J. Mol. Biol. 263*, 525−530.
19. Russ, W. P., and Engelman, D. M. (1999) TOXCAT: a measure of transmembrane helix association in a biological membrane. *Proc. Natl. Acad. Sci. U.S.A. 96*, 863−868.
20. Challou, N., Goormaghtigh, E., Cabiaux, V., Conrath, K., and Ruysschaert, J. M. (1994) Sequence and structure of the membrane-associated peptide of glycophorin A. *Biochem. 33*, 6902−6910.
21. Smith, S. O., and Bormann, B. J. (1995) Determination of helix-helix interactions in membranes by rotational resonance NMR. *Proc. Natl. Acad. Sci. U.S.A. 92*, 488−491.
22. Smith, S. O., Song, D., Shekar, S., Groesbeek, M., Ziliox, M., and Aimoto, S. (2001) Structure of the transmembrane dimer interface of glycophorin A in membrane bilayers. *Biochem. 40*, 6553−6558.
23. Camus, D., and Hadley, T. J. (1985) A *Plasmodium falciparum* antigen that binds to host erythrocytes and merozoites. *Science 230*, 553−556.

24. Orlandi, P. A., Klotz, F. W., and Haynes, J. D. (1992) A malaria invasion receptor, the 175-kilodalton erythrocyte binding antigen of Plasmodium falciparum recognizes the terminal Neu5Ac(alpha 2-3)Gal- sequences of glycophorin A. *J. Cell. Biol. 116*, 901−909.
25. Mayor, A., Bir, N., Sawhney, R., Singh, S., Pattnaik, P., Singh, S. K., Sharma, A., and Chitnis, C. E. (2005) Receptor-binding residues lie in central regions of Duffy-binding-like domains involved in red cell invasion and cytoadherence by malaria parasites. *Blood. 105*, 2557−2563.
26. Cartron, J. P., and Rahuel, C. (1992) Human erythrocyte glycophorins: protein and gene structure analyses. *Transfus. Med. Rev. 6*, 63−92.
27. Bruce, L. J., Pan, R. J., Cope, D. L., Uchikawa, M., Gunn, R. B., Cherry, R. J., and Tanner, M. J. (2004) Altered structure and anion transport properties of band 3 (AE1, SLC4A1) in human red cells lacking glycophorin A. *J. Biol. Chem. 279*, 2414−2420.
28. Brosig, B., and Langosch, D. (1998) The dimerization motif of the glycophorin A transmembrane segment in membranes: importance of glycine residues. *Protein Sci. 7*, 1052−1056.
29. Ash, W. L., Zlomislic, M. R., Oloo, E. O., and Tieleman, D. P. (2004) Computer simulations of membrane proteins. *Biochim. Biophys. Acta. 1666*, 158−189.
30. Gumbart, J., Wang, Y., Aksimentiev, A., Tajkhorshid, E., and Schulten, K. (2005) Molecular dynamics simulations of proteins in lipid bilayers. *Curr. Opin. Struct. Biol. 15*, 423−431.
31. Bond, P. J., and Sansom, M. S. P. (2003) Membrane protein dynamics vs. environment: simulations of OmpA in a micelle and in a bilayer. *J. Mol. Biol. 329*, 1035−1053.
32. Patargias, G., Bond, P. J., Deol, S. D., and Sansom, M. S. P. (2005) Molecular dynamics simulations of GlpF in a micelle vs. in a bilayer: conformational dynamics of a membrane protein as a function of environment. *J. Phys. Chem. B. 109*, 575−582.
33. Petrache, H. I., Grossfield, A., MacKenzie, K. R., Engelman, D. M., and Woolf, T. B. (2000) Modulation of glycophorin A transmembrane helix interactions by lipid bilayers: Molecular dynamics calculations. *J. Mol. Biol. 302*, 727−746.
34. Braun, R., Engelman, D. M., and Schulten, K. (2004) Molecular dynamics simulations of micelle formation around dimeric Glycophorin A transmembrane helices. *Biophys. J. 87*, 754−763.
35. Bond, P. J., Cuthbertson, J. M., Deol, S. D., and Sansom, M. S. P. (2004) MD simulations of spontaneous membrane protein/detergent micelle formation. *J. Am. Chem. Soc. 126*, 15948−15949.
36. Hénin, J., Pohorille, A., and Chipot, C. (2005) Insights into the recognition and asscoiation of transmembrane α-helices. The free energy of α-helix dimerization in glycophorin A. *J. Am. Chem. Soc. 127*, 8478−8484.
37. Bond, P. J., and Sansom, M. S. P. (2006) Insertion and assembly of membrane proteins via simulation. *J. Amer. Chem. Soc. 128*, 2697−2704.
38. Le Maire, M., Champeil, P., and Moller, J. V. (2000) Interaction of membrane proteins and lipids with solubilizing detergents. *Biochim. Biophys. Acta. 1508*, 86−111.
39. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., and Hermans, J. (1981) *Intermolecular Forces*, Reidel, Dordrecht.
40. Faraldo-Gómez, J. D., Smith, G. R., and Sansom, M. S. P. (2002) Setup and optimisation of membrane protein simulations. *Eur. Biophys. J. 31*, 217−227.
41. Lindahl, E., Hess, B., and van der Spoel, D. (2001) GROMACS 3.0: a package for molecular simulation and trajectory analysis. *J. Molec. Model. 7*, 306−317.
42. van Gunsteren, W. F., and Berendsen, H. J. C. (1987) *Gromos-87 manual*, Biomos BV, Groningen.
43. van Gunsteren, W. F., Kruger, P., Billeter, S. R., Mark, A. E., Eising, A. A., Scott, W. R. P., Huneberger, P. H., and Tironi, I. G. (1996) *Biomolecular Simulation*: The GROMOS96 Manual and User Guide, Biomos & Hochschulverlag AG an der ETH Zurich, Groningen & Zurich.
44. Darden, T., York, D., and Pedersen, L. (1993) Particle mesh Ewald - an N.log(N) method for Ewald sums in large systems. *J. Chem. Phys. 98*, 10089−10092.
45. Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., and Pedersen, L. G. (1995) A smooth particle mesh Ewald method. *J. Chem. Phys. 103*, 8577−8593.
46. Nose, S. (1984) A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys. 52*, 255−268.
47. Hoover, W. G. (1985) Canonical dynamics: equilibrium phase-space distributions. *Phys. Rev. A31*, 1695-1697.
48. Parrinello, M., and Rahman, A. (1981) Polymorphic transitions in single-crystals - a new molecular-dynamics method. *J. Appl. Phys. 52*, 7182−7190.
49. Nose, S., and Klein, M. L. (1983) Constant pressure molecular-dynamics for molecular-systems. *Mol. Phys. 50*, 1055−1076.
50. Hess, B., Bekker, H., Berendsen, H. J. C., and Fraaije, J. G. E. M. (1997) LINCS: A linear constraint solver for molecular simulation. *J. Comp. Chem. 18*, 1463−1472.
51. Smith, S. O., Eilers, M., Song, D., Crocker, E., Ying, W., Groesbeek, M., Metz, G., Ziliox, M., and Aimoto, S. (2002) Implications of threonine hydrogen bonding in the glycophorin A transmembrane helix dimer. *Biophys. J. 82*, 2476−2486.
52. Faraldo-Gómez, J. D., Smith, G. R., and Sansom, M. S. P. (2003) Molecular dynamics simulations of the bacterial outer membrane protein FhuA: a comparative study of the ferrichrome-free and bound states. *Biophys. J. 85*, 1−15.
53. Faraldo-Gómez, J. D., Forrest, L. R., Baaden, M., Bond, P. J., Domene, C., Patargias, G., Cuthbertson, J., and Sansom, M. S. P. (2004) Conformational sampling and dynamics of membrane proteins from 10-nanosecond computer simulations. *Proteins: Struct. Func. Bioinf. 57*, 783−791.
54. Lemmon, M. A., Treutlein, H. R., Adams, P. D., Brunger, A. T., and Engelman, D. M. (1994) A dimerisation motif for transmembrane a helices. *Nature Struct. Biol. 1*, 157−163.
55. Kabsch, W., and Sander, C. (1983) Dictionary of protein secondary structure: pattern-recognition of hydrogen-bonded and geometrical features. *Biopolymers. 22*, 2577−2637.
56. Curran, A. R., and Engelman, D. (2003) Sequence motifs, polar interactions and conformational changes in membrane proteins. *Curr. Opin. Struct. Biol. 13*, 412−417.
57. Gray, T. M., and Matthews, B. M. (1984) Intrahelical hydrogen bonding of serine, threonine and cysteine residues within α-helices and its relevance to membrane-bound proteins. *J. Mol. Biol. 175*, 75−81.
58. Treutlein, H. R., Lemmon, M. A., Engelman, D. M., and Brunger, A. T. (1992) The glycophorin A transmembrane domain dimer: Sequence specific propensity for a right handed supercoil of helices. *Biochem. 31*, 12726−12733.
59. Sal-Man, N., Gerber, D., and Shai, Y. (2004) The composition rather than position of polar residues (QxxS) drives aspartate receptor transmembrane domain dimerization in vivo. *Biochem. 43*, 2309−2313.
60. Gratkowski, H., Lear, J. D., and DeGrado, W. F. (2001) Polar side chains drive the association of model transmembrane peptides. *Proc. Natl. Acad. Sci. U.S.A. 98*, 880−885.
61. Zhou, F. X., Merianos, H. J., Brunger, A. T., and Engelman, D. M. (2001) Polar residues drive association of polyleucine transmembrane helices. *Proc. Natl. Acad. Sci.* U.S.A. *98*, 2250−2255.
62. Schneider, D., and Engelman, D. M. (2004) Motifs of two small residues can assist but are not sufficient to mediate transmembrane helix interactions. *J. Mol. Biol. 343*, 799−804.
63. Stockner, T., Ash, W. L., MacCallum, J. L., and Tieleman, D. P. (2004) Direct simulation of transmembrane helix association: role of asparagines. *Biophys. J. 87*, 1650−1656.
64. Senes, A., Ubarretxena-Belandia, I., and Engelman, D. M. (2001) The Cα −H.O hydrogen bond: a determinant of stability and specificity in transmembrane helix interactions. *Proc. Natl. Acad. Sci. U.S.A. 98*, 9056−9061.
65. Arora, A., and Tamm, L. K. (2001) Biophysical approaches to membrane protein structure determination. *Curr. Opin. Struct. Biol. 11*, 540−547.
66. Torres, J., Stevens, T. J., and Samso, M. (2003) Membrane proteins: the 'Wild West' of structural biology *Trends Biochem. Sci. 28*, 137−144.
67. Kim, S., Chamberlain, A. K., and Bowie, J. U. (2003) A simple method for modeling transmembrane helix oligomers. *J. Mol. Biol. 329*, 831−840.
68. Im, W., and Brooks, C. L. (2005) Interfacial folding and membrane insertion of designed peptides studied by molecular dynamics simulations. *Proc. Nat. Acad. Sci. U.S.A. 102*, 6771−6776.
69. Im, W., Feig, M., and Brooks, C. L. (2003) An implicit membrane generalized Born theory for the study of structure, stability, and interactions of membrane proteins. *Biophys. J. 85*, 2900−2918.
70. Efremov, R. G., Vereshaga, Y. A., Volynsky, P. E., Nolde, D. E., and Arseniev, A. S. (2006) Association of transmembrane helices: what determines assembling of a dimer? *J. Comput.-Aided Mol. Des. 20*, 27−45.

71. Mottamal, M., Zhang, J., and Lazaridis, T. (2006) Energetics of native and non-native states of the glycophorin transmembrane helix dimer. *Proteins: Struct. Funct. Bioinf. 62*, 996−1009.
72. Chen, Z., and Xu, Y. (2006) Energetics and stability of transmembrane helix packing: a replica-exchange simulation with a knowledge-based membrane potential. *Proteins: Struct. Funct. Bioinf. 62*, 539−552.
73. Beevers, A. J., and Kukol, A. (2006) Systematic molecular dynamics searching in a lipid bilayer: application to the glycophorin A and oncogenic ErbB-2 transmembrane domains. *J. Mol. Graph. Model. 25*, in press.
74. Partridge, A. W., Melnyk, R. A., and Deber, C. M. (2002) Polar residues in membrane domains of proteins: Molecular basis for helix-helix association in a mutant CFTR transmembrane segment. *Biochem. 41*, 3647−3653.
75. Zhou, F. X., Cocco, M. J., Russ, W. P., Brunger, A. T., and Engelman, D. M. (2000) Interhelical hydrogen bonding drives strong interactions in membrane proteins. *Nat. Struct. Biol. 7*, 154−160.
76. Choma, C., Gratkowski, H., Lear, J. D., and DeGrado, W. F. (2000) Asparagine-mediated self-association of a model transmembrane helix. *Nat. Struct. Biol. 7*, 161−166.
77. Senes, A., Engel, D. E., and 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.
78. Marassi, F. M., Ma, C., Gratkowski, H., Straus, S. K., Strebel, K., Oblatt-Montal, M., Montal, M., and Opella, S. J. (1999) Correlation of the structural and functional domains in the membrane protein Vpu from HIV-1. *Proc. Nat. Acad. Sci. U.S.A. 96*, 14336−14341.
79. Opella, S. J., and Marassi, F. M. (2004) Structure determination of membrane proteins by NMR spectroscopy. *Chem. Rev. 104*, 3587−3606.
80. Tian, C. L., Tobler, K., Lamb, R. A., Pinto, L. H., and Cross, T. A. (2002) Expression and initial structural insights from solid-state NMR of the M2 proton channel from influenza A virus. *Biochem. 41*, 11294−11300.
81. Domene, C., Bond, P. J., Deol, S. S., and Sansom, M. S. P. (2003) Lipid-protein interactions and the membrane/water interfacial region. *J. Amer. Chem. Soc. 125*, 14966−14967.
82. Deol, S. S., Domene, C., Bond, P. J., and Sansom, M. S. P. (2006) Anionic phospholipids interactions with the potassium channel KcsA: simulation studies. *Biophys. J. 90*, 822−830.
83. Tobias, D. J. (2001) Electrostatics calculations: recent methodological advances and applications to membranes. *Curr. Opin. Struct. Biol. 11*, 253−261.
84. Bostick, D. L., and Berkowitz, M. L. (2003) The implementation of slab geometry for membrane-channel molecular dynamics simulations. *Biophys. J. 85*, 97−107.
85. Weber, W., Hunenberger, P. H., and McCammon, J. A. (2000) Molecular dynamics simulations of a polyalanine octapeptide under Ewald boundary conditions: Influence of artificial periodicity on peptide conformation. *J. Phys. Chem. B. 104*, 3668−3675.
86. Kastenholz, M. A., and Hünenberger, P. H. (2004) Influence of artificial periodicity and ionic strength in molecular dynamics simulations of charged biomolecules employing lattice-sum methods. *J. Phys. Chem. B. 108*, 774−788.
87. Hunenberger, P. H., and, J. A. M. (1999) Effect of artificial periodicity in simulations of biomolecules under Ewald boundary conditions: a continuum electrostatics study. *Biophys. Chem. 78*, 69−88.
88. Beck, D. A. C., Armen, R. S., and Daggett, V. (2005) Cutoff size need not strongly influence molecular dynamics results for solvated polypeptides. *Biochem. 44*, 609−616.
89. Anezo, C., de Vries, A. H., Holtje, H. D., Tieleman, D. P., and Marrink, S. J. (2003) Methodological issues in lipid bilayer simulations. *J. Phys. Chem. B. 107*, 9424−9433.
90. Lyubartsev, A. P. (2005) Multiscale modeling of lipids and lipid bilayers *Eur. Biophys. J. 35*, 53−61.
91. Chang, R., Ayton, G. S., and Voth, G. A. (2005) Multiscale coupling of mesoscopic- and atomistic-level lipid bilayer simulations. *J. Chem. Phys. 122*, 244716.
92. Choi, M. Y., Cardarelli, L., Therien, A. G., and Deber, C. M. (2004) Non-native interhelical hydrogen bonds in the cystic fibrosis transmembrane conductance regulator domain modulated by polar mutations. *Biochem. 43*, 8077−8083.
93. Fleishman, S. J., Unger, V. M., and Ben-Tal, N. (2006) Transmembrane protein structures without X-rays. *Trends Biochem. Sci. 31*, 106−113.

BI0610911