Energetics of the Native and Non-native States of the Glycophorin Transmembrane Helix Dimer

Madhusoodanan Mottamal, Jinming Zhang, and Themis Lazaridis*
Department of Chemistry, City College of New York/CUNY, New York, New York

ABSTRACT **Using an implicit membrane model (IMM1), we examine whether the structure of the transmembrane domain of Glycophorin A (GpA) could be predicted based on energetic considerations alone. The energetics of native GpA shows that van der Waals interactions make the largest contribution to stability. Although specific electrostatic interactions are stabilizing, the overall electrostatic contribution is close to zero. The GXXXG motif contributes significantly to stability, but residues outside this motif contribute almost twice as much. To generate non-native states a global conformational search was done on two segments of GpA: an 18-residue peptide (GpA74-91) that is embedded in the membrane and a 29-residue peptide (GpA70–98) that has additional polar residues flanking the transmembrane region. Simulated annealing was done on a large number of conformations generated from parallel, antiparallel, left- and right-handed starting structures by rotating each helix at 20° intervals around its helical axis. Several crossing points along the helix dimer were considered. For 18-residue parallel topology, an ensemble of native-like structures was found at the lowest effective energy region; the effective energy is lowest for a right-handed structure with an RMSD of 1.0 Å from the solid-state NMR structure with correct orientation of the helices. For the 29-residue peptide, the effective energies of several left-handed structures were lower than that of the native, right-handed structure. This could be due to deficiencies in modeling the interactions between charged sidechains and/or omission of the sidechain entropy contribution to the free energy. For 18-residue antiparallel topology, both IMM1 and a Generalized Born model give effective energies that are lower than that of the native structure. In contrast, the Poisson-Boltzmann solvation model gives lower effective energy for the parallel topology, largely because the electrostatic solvation energy is more favorable for the parallel structure. IMM1 seems to underestimate the solvation free energy advantage when the CO and NH dipoles just outside the membrane are parallel. This highlights the importance of electrostatic interactions even when these are not obvious by looking at the structures. Proteins 2006;62:996–1009.**

Key words: membrane protein; glycophorin A; structure prediction; implicit solvation

INTRODUCTION

Integral membrane proteins are essential in mediating signal transduction and transport across cell membranes. This makes them crucial for our understanding of cellular processes as well as attractive targets for drug discovery. Because of technical problems related to purification and crystallization, determination of the structure of membrane proteins is very difficult; only about three dozen distinct folds have been solved to date at high resolution. Under these circumstances, computer modeling could play a major role in predicting the structure and understanding the function of these proteins.

In many integral membrane proteins, the transmembrane (TM) domain consists of one or more α-helices rich in hydrophobic residues. The function of many membrane proteins depends upon the formation of oligomers through interactions between their TM α-helices.[1,2] Assembly of some membrane receptor complexes, such as the T-cell receptor and Fc receptors, is driven by interactions between the TM domains of the constituent subunits.[3,4] Oligomerization is sequence specific, and is driven by van der Waals (packing) interactions, electrostatic interactions, and hydrogen bonding. However, the magnitude and relative importance of each of these forces is still unclear.

The simplest example of TM α-helix association is the dimeric TM domain of Glycophorin A (GpA), a small protein found in human erythrocytes. The structure of the transmembrane region of GpA in both DPC micelles[5] and in lipid bilayers[6] has been determined by NMR. The structure was consistent with earlier mutagenesis analysis.[7,8] The two helices form a symmetric right-handed homodimer stabilized by contact between two GXXXG sequence motifs. This motif occurs frequently, and is known to favor helix–helix interactions in membrane proteins[9–11] and also in soluble proteins.[12] Some other examples of homo-oligomerizing proteins are phospholam-

The Supplementary Material referred to in this article can be found online at http://www.interscience.wiley.com/jpages/0887-3585/suppmat/

Grant sponsor: the National Science Foundation; Grant number: MCB-0316667; Grant sponsor: the National Institute of Health; Grant number: 2S06GM008168.

*Correspondence to: Themis Lazaridis, Department of Chemistry, City College of New York/CUNY, 138th Street and Convent Avenue, New York, NY 10031. E-mail: tlazaridis@ccny.cuny.edu

Received 14 June 2005; Revised 8 September 2005; Accepted 5 October 2005

Published online 4 January 2006 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.20844

ban (pentamer), influenza virus A M2 (tetramer), and vpu from HIV-1 (pentamer).

Because of their symmetry and simplicity, homo-oligomeric transmembrane proteins offer an attractive system for computer modeling. One seminal study attempted structure prediction of the GpA TM by combination of simulated annealing and experimental mutagenesis data.[13] This study suggested a right-handed supercoil with extensive close packing along the dimer interface. A modified global search method later gave an improved prediction of the structure of GpA[14] and of the pentameric transmembrane domain of phospholamban.[15] Many studies followed, which used global conformational search using a vacuum energy function[16,17] or a van der Waals-only energy function.[18,19] However, in all studies some extraneous criterion had to be used to select the native state, such as evolutionary conservation,[16,20] experimental helix tilts,[17,21–24] boundary constraints,[25,26] or symmetry constraints.[19] Knowledge-based energy functions have also been developed.[27,28]

One recurring observation in the above studies is that vacuum energy alone cannot discriminate the correct conformation. This is not surprising; the membrane environment is expected to play a role in determining the native structure. In this work we approximate the effect of the membrane by using the implicit membrane model IMM1.[29] Several other models of this type have been developed recently.[30–35] We ask whether the correct, native structure of GpA can be predicted from first principles, based on free energy, and if so, what physical interactions are the determinants of the native structure. Thus, our goal is not only to predict the structure, but also to obtain physical insights in the process.

We first examine the native state and decompose the interhelical interaction energy into residue contributions. Then we generate a large number of non-native structures via simulated annealing. The first set of calculations is based on the18-residue segment, Thr74 through Ile91, which is most hydrophobic and should be well embedded in the membrane. To test the robustness of the results, we also use a 29-residue segment of GpA (Glu70 through Leu98), which includes some of the flanking polar and charged residues. The energies of all structures obtained were examined and compared to that of the native structure.

METHODS

All calculations were carried out using the program CHARMM.[36] The main effective energy function we use is IMM1,[29] which is an extension of EEF1[37,38] to membranes. Effective energy (W), or potential of mean force, is the sum of the intramolecular energy and the solvation free energy. In EEF1, the solvation free energy of a protein is assumed to be the sum of group contributions. The solvation free energy of each group is equal to a reference value determined from experimental solvation free energy data minus the amount of solvation lost due to exclusion of solvent by the surrounding protein atoms. In IMM1, the reference solvation free energies depend on the position of the atom with respect to the membrane, which lies parallel to the xy plane with its center at $z = 0$. The solvation parameters in the nonpolar core of the membrane were taken from experimental data for small model compounds in cyclohexane. The solvation parameters of all atoms are dependent on the vertical coordinate, z or $z' = |z|/(T/2)$, where T is the thickness of the nonpolar core of the membrane:

$$\Delta G_i(z') = f(z')\Delta G_i^{\text{water}} + (1 - f(z'))\Delta G_i^{\text{cyclohexane}} \quad (1)$$

The function f(z′) has the following form:

$$f(z') = \frac{z'^n}{1 + z'^n} \quad (2)$$

where n controls the steepness of the transition ($n = 10$). In addition, in IMM1, a modified dielectric screening function is used to account for the strengthening of electrostatic interactions in the membrane. This function has the following form:

$$\varepsilon = r^{(a+(1-a)\sqrt{f_i f_j})} \quad (3)$$

where f_i and f_j are obtained from Equation (2). The value 0.85 for "a" was found to give reasonable results for various systems studied earlier.[29] Depending upon the lipid, the thickness of a membrane can vary and it usually ranges between 20 and 30 Å. In this study, a thickness of 26 Å was chosen for the lipid membrane, as in previous work.[29]

Preliminary calculations on the 29-residue GpA peptide showed the Arg side chains sometimes interacting strongly with the backbone or with neighboring Arg and other polar residues. To mitigate this effect we used reduced partial charges on the Arg side chain atoms (Table SI in the Supplementary Material). The new partial charges also gave improved agreement with the Arg–Arg potentials of mean force calculated in explicit water.[39]

First, we analyzed the energetics of native GpA. We used two initial structures for the transmembrane domain: one obtained by solution NMR in detergent (pdb code: 1AFO),[5] and the other by solid-state NMR in lipid bilayers obtained from Prof. S.O. Smith.[6,40] The backbone RMSD of the 20 solution state NMR models from the solid-state structure ranges from 0.87 to 1.24 Å, with an average of 1.1 Å. For the solution NMR structure, we selected the first of 20 models. The backbone RMSD of this model from the others ranges from 0.48 to 1.20 Å, with an average of 0.74 Å. For both structures, residues 73 through 95 were included in the calculations. The initial structures were oriented with their principal axis perpendicular to the membrane surface, and then energy minimized for 800 steps by the "ABNR" method. Subsequently, 1-ns MD simulations starting from the minimized structures were performed with harmonic constraints on all the backbone atoms (force constant 1.0 kcal/mol/Å^2), to examine the effect of thermal motion while remaining close to the experimental structure.

For the conformational search, two segments of GpA containing 18 residues (T_{74}LIIPGV MAGVIGTILLI$_{91}$) and 29 residues (E_{70}PEITLIIPGVMAGVIGTILLISYGIRRL$_{98}$)

TABLE I. Total Energies and Interhelical Interactions of the Two NMR GpA Structures

	Solution NMR structure		Solid-state NMR structure	
	Total effective energy	Interhelical interaction effective energy	Total effective energy	Interhelical interaction effective energy
After 800 steps ABNR minimization	−1108.1 (−247.0, −748.4, −170.2)	−23.6 (−36.2, 1.7, 10.8)	−1106.5 (−254.0, −737.4, −169.1)	−29.3 (−42.9, 1.5, 12.1)
Average over the last 0.9 ns	−760.8 (−220.3, −729.1, −172.8)	−21.8 (−34.4, 1.7, 10.9)	−764.9 (−224.0, −729.2, −171.4)	−24.9 (−39.5, 2.2, 12.4)

All energies are in kcal/mol. The values in parentheses are van der Waals, electrostatic, and solvation energies, in this order.

were considered. Ideal α-helices containing the above sequences were generated and placed parallel to each other at a helical axis distance of 9 Å. After orienting the principal axis of the dimer along the z axis, 324 structures were generated by rotating each helix by 20° around its helical axis, defined by the centers of mass of the five Cα atoms from the N- and C-termini. All structures were energy minimized for 300 ABNR steps and then subjected to 50-ps simulated annealing followed by 30-ps Nose-Hoover dynamics at 300 K. For annealing, the system was first heated to 500 K and then brought to room temperature (300 K), cooling by 50 K every 4 ps with a time step of 0.002 ps. Distance constraints of 3.25 Å (with a maximum/minimum deviation of ±0.25 Å) were applied between O_i and N_{i+4} atoms to maintain the α-helical conformation. The force constants for minimum and maximum deviations were 1 and 7 kcal/mol/Å^2, respectively. Stronger force constants were also tried, with similar results. Similarly, to prevent the helices from separating, a distance constraint was applied on the centers of mass of the Cα atoms of five consecutive amino acids from each monomer at the same helix region. This constraint fixes the crossing point of the helices. Average energies at room temperature were obtained from the last 20 ps of the 30-ps Nose-Hoover dynamics, starting from the final simulated annealing structures at 300 K with constraints. The effective energies of these final structures (50 ps annealing + 30-ps Nose-Hoover dynamics + 300 steps of ABNR minimization) or the average effective energy over the last 20 ps of the simulation with and without the constraint energies were used to evaluate conformations. We also calculated the backbone RMSD of all the structures from the solid-state NMR structure. Thus, unless specified all the RMSDs reported in this article are from the solid-state NMR structure.

For the 18mer we considered four crossing points at residues 76–80, 79–83, 82–86, and 85–89 from each chain. An additional three crossing points (73–77, 88–93, and 91–95) were considered for the 29mer. A distance constraint of 5.7 Å between the helical axes was employed in all simulations. A force constant of 5 kcal/mol/Å^2 allowed the distance to span about 2 Å and accommodate bigger side chains at the interface. A force constant of 1 kcal/mol/Å^2 gave a very similar overall energy distribution. A search with an interhelical distance constraint of 7 Å at the experimental crossing point produced no structure with RMSD smaller than 2.2 Å and gave lowest energy structures that tended to be more parallel ($-10 < \Omega < 10$). Thus, when the interhelical distance is too large, the search cannot generate native-like structures.

Additional simulated annealing runs started from right and left-handed (RH and LH) conformations of GpA generated from an initial supercoiled structure. From the coordinates of a standard α-helix, the initial Cα, N, and C coordinates of right-and left-handed supercoils were generated from their respective Cα, N, and C atoms using Crick's formula.[41] The radius of the supercoil was set to 5.2 Å and the repeat distance to 186 Å. From the initial Cα, N, and C coordinates of the supercoils, right- and left-handed supercoils were obtained by building the internal coordinates using CHARMM. These structures were then rotated around an axis defined by the centers of mass of the Cα of the first and last seven amino acids at an interval of 20°. For the RH and LH conformations only the experimental crossing point was considered. The simulated annealing protocol was as described above.

Although the GpA helix dimer is known to be parallel, we also considered the 18-residue peptide in antiparallel conformation. We considered only two crossing points: residues 6–10 from helix 1 with residues 9–13 from helix 2; and residues 3–7 from helix 1with residues 12–16 from helix 2. The simulation protocol was the same as above. In addition to IMM1, we performed MD simulations with the Poisson-Boltzmann (PB) model and the Generalized Born model with a simple smoothing function (GBSW).[34,35] For PB, we used a membrane thickness of 26 Å and a dielectric constant of 1 in the protein interior and in the membrane. For these simulations, we did not use any cutoff for the nonbonded interactions and the systems studied were two low energy antiparallel, five low energy parallel and the solid-state NMR structures.

The CPU time for the global search on the 18mer with IMM1 (a total of 155 ns) was 265 h on a Xeon 3.06 GHz processor. The same for the 29mer was 620 h (181 ns). IMM1 (9 Å cutoff) is four times slower than vacuum, 12 times faster than GBSW without cutoff, and 3200 times faster than PB without cutoff.

RESULTS

Energetics of the Native Structure

The minimized and average total and interhelical interaction energies of the solid-state and solution NMR GpA structures are presented in Table I. The solid-state NMR structure has a stronger interhelical interaction probably due to two reasons. First, the OH group of Thr87 makes an intrahelical hydrogen bond with the backbone in the

TABLE II. Residue–Residue Interactions in the Energy-Minimized Solution NMR GpA Structure

	A75LEU	A76ILE	A79GLY	A80VAL	A82ALA	A83GLY	A84VAL	A87THR	A88ILE
B75LEU		−1.8 (−2.3,0.2,0.3)							
B76ILE	−1.8 (−2.3,0.1,0.3)	−1.4 (−2.6,0.4,0.8)	−1.1 (−0.9,−0.4,0.2)						
B79GLY		−0.9 (−0.8,−0.3,0.2)		−0.9 (−1.5,0.0,0.6)					
B80VAL			−0.7 (−1.4,0.0,0.6)	−0.5 (−1.3,0.4,0.4)	−0.7 (−0.9,0.0,0.2)	−1.6 (−1.4,−0.6,0.5)			
B82ALA				−0.6 (−0.8,0.0,0.2)					
B83GLY				−1.5 (−1.4,−0.6,0.4)			−1.0 (−1.5,0.0,0.4)		
B84VAL						−1.0 (−1.5,0.0,0.4)		−1.5 (−2.0,−0.2,0.7)	
B87THR							−1.5 (−2.0,−0.2,0.7)	−0.6 (−1.0,−0.2,0.6)	−0.6 (−0.8,−0.1,0.2)
B88ILE								−0.6 (−0.7,−0.1,0.2)	

All energies are in kcal/mol. The values in parentheses are van der Waals, electrostatic, and solvation energies, in this order. Only the residue–residue interactions with absolute values greater than 0.5 kcal/mol are shown.
The small differences observed in some symmetric positions are due to a slight asymmetry in the minimized structure.

solution NMR structure but an interhelical hydrogen bond in the solid-state NMR structure (during the MD simulation, Thr87 rapidly reverts to an intrahelical hydrogen bond). Second, the solid-state NMR structure has a smaller crossing angle. The largest overall contribution to the interhelical energy is the van der Waals interaction. The desolvation energy contribution is positive, reflecting the loss of protein–lipid interactions upon dimer formation. The electrostatic contribution is very small.

For a more detailed attribution of association energy, the interhelical interactions were decomposed into contributions between each pair of interhelical residues. A similar decomposition has been performed previously but without the desolvation contribution.[42] The matrices of interaction pairs are shown in Tables II and III for the solution and solid-state NMR structures, respectively. In most residue pairs, van der Waals interactions make the largest contribution and electrostatic interactions the smallest. The desolvation energies are always positive, except in one case where the pair is just outside the membrane. There are some differences in these two matrices. (1) the largest distance in sequence for two interhelical residues interacting strongly is three residues in the solution NMR structure and four residues in the solid-state NMR structure. (2) In the solution NMR structure, the first residue with strong interaction with a residue on the other helix is Leu75 and the last one Ile88, but in the solid-state NMR structure, they are Ile73 and Ile91, respectively. That is because the solid-state NMR structure has a longer interface between two helices and a smaller crossing angle than the solution NMR structure. (3) The solid-state NMR structure has a positive interaction between two Thr87 residues due to an unfavorable dipole–dipole interaction between the OH groups.

The interhelical residue–residue interaction energies shown in Tables II and III account for about 85.0% of the total interhelical interaction energy. Interhelical interactions from the GXXXG motif are ~−6.5 kcal/mol in both minimized structures, which corresponds to about 30% of the total interhelical interaction. The van der Waals, electrostatic, and solvation contributions are about −8.2, −1.0, and +2.7 kcal/mol, respectively. Thus, there are some important interactions outside this motif, such as Ile76: Leu75 and Val84: Gly83. The 6 Cα-H···O hydrogen bonds in GpA contribute ~−6.0 kcal/mol to stability, largely due to favorable electrostatic interactions.[43] However, the overall electrostatic interaction in GpA is close to zero, which means that there must exist repulsive interactions that cancel out the hydrogen bond contribution. In both NMR structures, the electrostatic interactions between the interhelical residues Gly79 and Gly79 and Gly83 and Gly83 are repulsive (~1.5 kcal/mol). In the solid-state structure, the electrostatic interaction between the interhelical residues Thr87 & Thr87 is also positive (~2.2 kcal/mol) mainly due to the interactions between the hydroxyl group atoms. Other weak repulsive electrostatic interactions are between Val80 and Val80 and Ile76 and Ile76. The results in Tables II and III are consistent with recent mutational studies that show that interactions outside the GXXXG motif are important for helix association.[44,45] For example, a mutant that lacks the GXXXG motif can still dimerize, while a mutant containing the GXXXG motif failed to dimerize.[46]

Conformational Search on the 18-Residue Segment of GpA

We first simulated residues Thr74 through Ile91, which should be well within the membrane boundaries. This segment was originally selected for structure prediction by simulated annealing in vacuum.[13] Figure 1(A) shows the minimized energy of all GpA conformations against their RMSD from the native structure (a 3D representation of the effective energy versus the crossing angle and RMSDs from the NMR structure is given in Fig. S1 in the Supplementary Material). The lowest effective energy conformations are very close to the native structure. The

TABLE III. Residue–Residue Interactions in the Energy-Minimized Solid-State NMR GpA Structure

	A73ILE	A75LEU	A76ILE	A79GLY	A80VAL	A82ALA	A83GLY	A84VAL	A87THR	A88ILE	A91ILE
B73ILE	−1.2 (−1.1,0.0,−0.1)	−0.6 (−0.8,0.1,0.1)									
B75LEU	−0.6 (−0.8,0.0,0.1)		−1.9 (−2.4,0.1,0.4)								
B76ILE		−1.9 (−2.4,0.1,0.4)	−0.9 (−2.0,0.4,0.7)	−1.2 (−1.1,−0.4,0.3)							
B79GLY			−1.1 (−1.0,−0.4,0.3)		−0.8 (−1.4,0.0,0.6)						
B80VAL				−0.8 (−1.4,0.0,0.6)		−0.7 (−0.9,0.0,0.2)	−1.6 (−1.5,−0.6,0.5)				
B82ALA					−0.7 (−0.9,0.0,0.2)						
B83GLY					−1.6 (−1.5,−0.6,0.5)			−1.0 (−1.5,0.1,0.4)	−0.6 (−0.3,−0.4,0.1)		
B84VAL							−1.0 (−1.5,0.1,0.5)		−2.5 (−2.0,−1.3,0.8)		
B87THR							−0.5 (−0.3,−0.3,0.1)	−1.8 (−2.0,−0.6,0.8)	0.9 (−1.1,1.3,0.7)	−0.5 (−0.9,0.1,0.3)	
B88ILE									−0.6 (−1.1,0.1,0.4)		
B91ILE											−1.4 (−1.9,0.0,0.4)

All energies are in kcal/mol. The values in parentheses are van der Waals, electrostatic, and solvation energies, in this order. Only the residue–residue interactions with absolute values greater than 0.5 kcal/mol are shown.
The small differences observed in some symmetric positions are due to a slight asymmetry in the minimized structure.

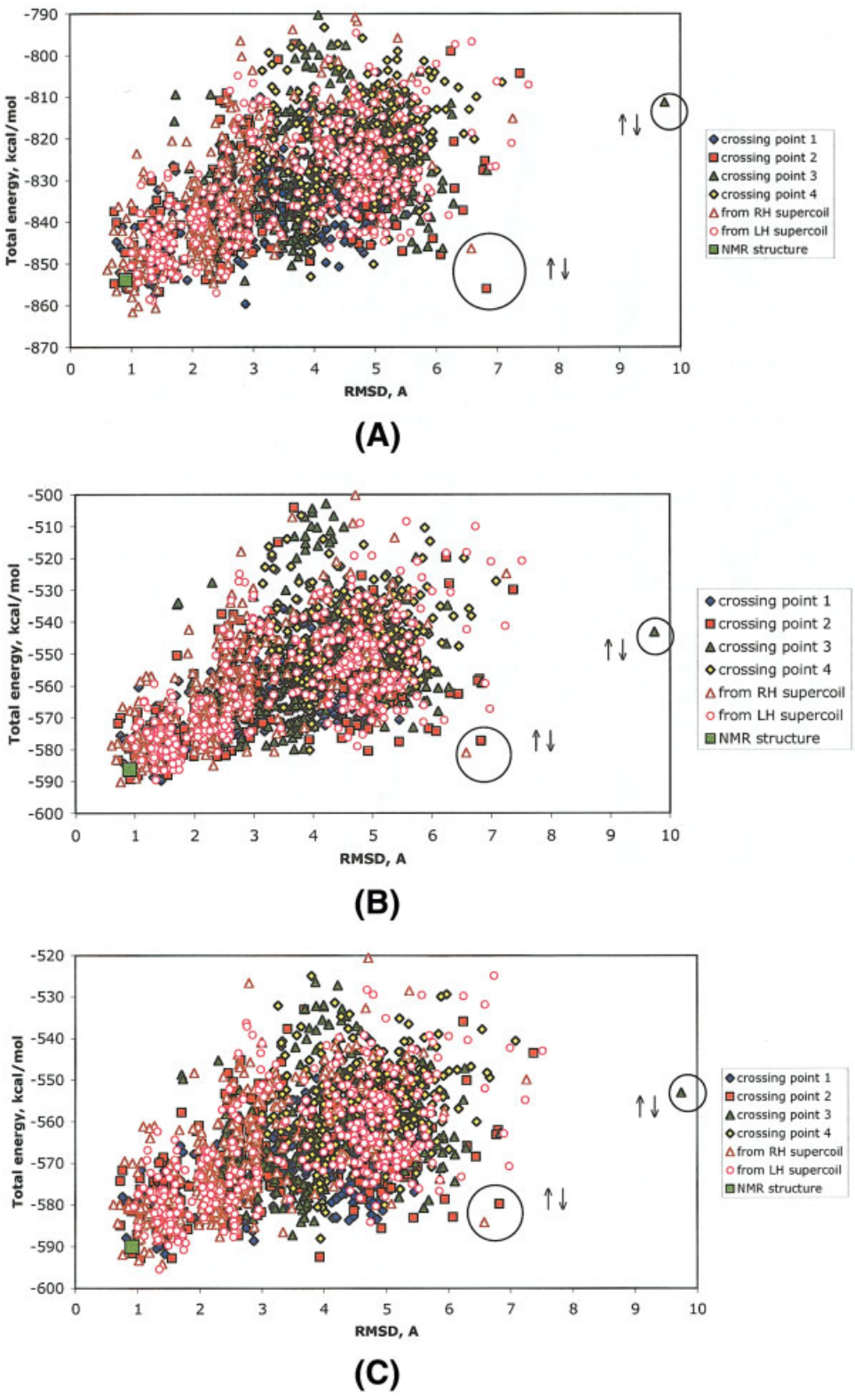

Fig. 1. IMM1 energy of all conformations of 18-residue GpA versus their RMSD from the NMR structure. Encircled points marked by ↑ ↓ are the only three antiparallel structures. (**A**) Minimized energy, (**B**) average effective energy with constraint energy, and (**C**) average effective energy without constraint energy. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

TABLE IV. Energetics of the Relaxed NMR and 5 Lowest Effective Energy Parallel Structures of 18-Residue GpA with IMM1

Structure	Total energy	van der Waals	Electrostatic	Covalent	Solvation	RMSD
Minimized energies						
P1	−861.6	−190.4	−582.4	43.4	−132.3	1.02
P2	−860.6	−191.2	−584.1	46.6	−131.9	1.39
P3	−860.2	−194.3	−578.5	45.0	−132.4	1.11
P4[a]	−859.7	−180.3	−581.6	45.2	−142.9	2.87
P5	−858.8	−185.3	−581.0	44.6	−137.1	1.35
relaxed NMR	−854.0[b]	−188.2	−581.8	45.6	−129.6	0.92
Average energies including constraint energy						
P6	−590.0	−161.0	−550.9	253.2	−133.2	0.75
P7[a]	−589.7	−166.8	−552.4	262.3	−134.9	1.44
P5	−589.5	−160.5	−552.6	257.2	−139.5	1.35
P8	−589.5	−160.6	−554.1	255.0	−131.9	0.92
P1	−588.9	−168.1	−549.2	262.2	−138.2	1.02
relaxed NMR	−586.3[b]	−160.1	−552.4	259.8	−137.3	0.92
Average energies excluding constraint energy						
P5	−595.4	−160.5	−552.6	257.2	−139.5	1.35
P2	−594.2	−160.1	−554.6	261.7	−141.1	1.40
P1	−593.4	−168.1	−549.2	262.2	−138.2	1.02
P9	−592.8	−160.0	−552.2	258.7	−139.4	1.55
P10	−592.8	−164.0	−553.6	259.1	−134.2	1.05
relaxed NMR	−590.0[b]	−160.1	−552.4	259.8	−137.3	0.92

All energies are in kcal/mol and the RMSDs are in Å.
[a]Structures that do not have native orientation of the helices.
[b]Average over four runs with different initial random number.

effective energy of the NMR structure subjected to the same simulated annealing protocol is a few kcal higher than the lowest energy structure. The lowest energy conformations come from parallel, left-handed or right-handed starting structures. Out of 1944 structures in Figure 1, three structures marked with ↑ ↓ were found to adopt somewhat antiparallel structures. The effective energies of these three structures are 7 to 50 kcal/mol higher than the lowest observed effective energies. In Figure 1(A) one structure whose energy is only 2 kcal/mol higher than the lowest energy structure has an RMSD of 2.9 Å. The crossing angle of this structure is ~10° smaller than the native-like structure and the helices are slightly rotated with respect to the NMR structure. The average effective energy of this structure (discussed below) is significantly higher than the lowest average effective energy.

Average energy is a better indicator of stability because it includes the effect of thermal motion (energy-minimized conformations may have low vibrational entropy). The average effective energies of all simulated annealing structures are shown in Figure 1. The results in Figure 1(B) include the constraint energy and those in Figure 1(C) do not. In both cases the lowest energy structures are very close to the solid-state NMR structure. When average energies are considered, the difference between the lowest energy structure and the NMR structure is much smaller than that between the energy minimized structures. Without constraint energy one structure with relatively large RMSD (~4 Å) is 3 kcal/mol higher than the lowest average effective energy structure. Although it adopts a right-handed conformation, the crossing angle of this structure is ~20° smaller and the orientation of the helices deviates from the NMR structure. The minimized energy of this structure is ~12 kcal/mol higher than the lowest minimized effective energy.

Table IV shows the energetics of the five lowest energy structures from each of the Figures 1(A), 1(B), and 1(C) and the NMR structure. For the NMR structure, simulations were done with four different sets of initial velocities. For these four structures the minimized effective energies range from −849.5 to −856.2, with an average of −854.0. The lowest W structure (P1) has an effective energy 7.6 kcal/mol lower than the relaxed NMR structure. This energy difference is comparable to the variation in energy among the four NMR structure runs. The lowest effective energy structures are very similar to the native structure, but P4 and P7 exhibit a slight rotation of the helices (Fig. 2). In all structures Thr87 makes an intrahelical hydrogen bond with Gly80, except for P7 where it makes an interhelical Thr87–Thr87 hydrogen bond. This hydrogen bond is facilitated by the rotation of the helices.

The variation in energy among the low energy structures obtained by the calculations is similar to the variation observed among models of the NMR structure. Energy minimization of the 20 solution NMR models for GpA gave energies that differ within a range of ~8 kcal/mol. When our native-like structures were superimposed they were found to lie perfectly on each other and the only main difference seen among them is a slight variation in the position of the sidechains. Thus, all low energy structures except P4 and P7 can be considered as part of the native ensemble.

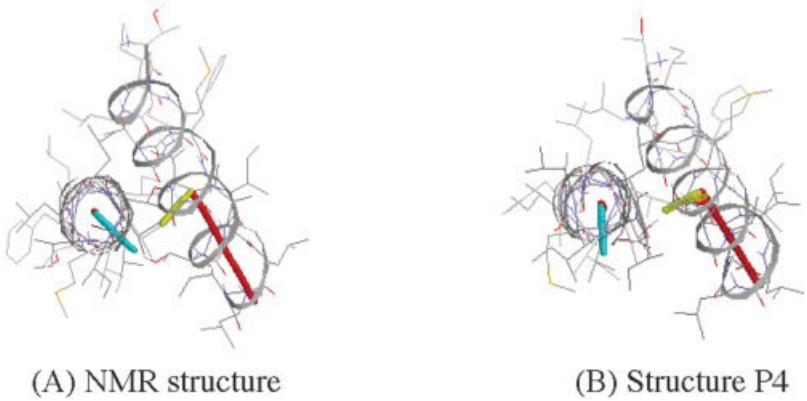

Fig. 2. Orientation of the helices in the NMR and P4 structures. Vectors shown in red, yellow, and cyan are the helix axis, a vector that is perpendicular to the helix axis and passes through G83.Cα of helix 1 and a vector that is perpendicular to the helix axis and passes through V84.Cα of helix 2, respectively. (P7 is very similar to P4.)

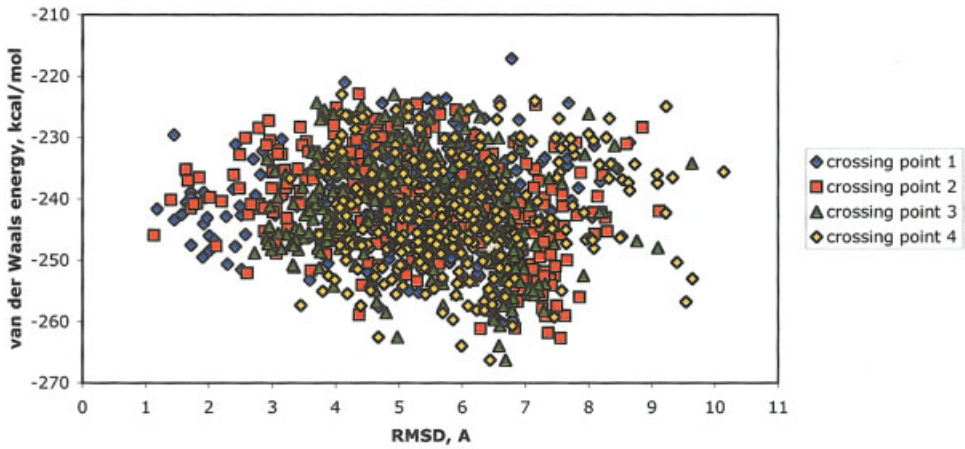

Fig. 3. The van der Waals energy versus RMSD from the NMR structure for the minimized structures obtained via conformational search with an energy function that excludes the electrostatic and solvation energies.

We measured the interhelical distances for the six pairs of hydrogens that gave NOE signal in the solution NMR experiment.[5] They are V80(HG2)–G79(HN), L75(HD1)–I76(HB), V84(HG2)–T87(HG1), V84(HG2)–G83(HA), T87(HG2)–I88(HG1), and V84(HA)–T87(HG1). Because nonpolar hydrogen atoms are not treated explicitly in IMM1, we built hydrogens on all structures using the CHARMM22 all-atom force field. Among chemically equivalent hydrogens, the one that gave the shortest HH distance was selected. All HH pair distances were below 3 Å in our low energy structures except for V80(HG2)–G79(HN), L75(HD1)–I76(HB), which gave distances between 3 and 5 Å. The same observation was made in another simulation of GpA.[47] The HH distances for Leu75-HD1:Ile76-HB and Val80-HG2:Gly79:HN in the 20 solution NMR models are between 4 and 4.4 Å. In the solid-state NMR structure they both measure ~5 Å. In our lowest energy models (P1–10 and P11–P19 in supplementary Table SII), these distances are between 3 and 5 Å. Similarly, we measured the interhelical distances for the six pairs of carbon atoms that were restrained in the solid-state NMR structure of GpA.[6] They are G79(CH_2)–G79(CO), I76(CO)–G79(CH_2), G83(CH_2)–G83(CO), G83(CH_2)–V80(CO), G79(CO)–V80(CH_3), and G83(CO)–V84(CH_3). The six interhelical C—C distances in all the structures, except the four slightly rotated (P4, P7, P15, and P16) and one lowest energy van der Waals energy (P11) structure, are within the experimentally allowable range. In the slightly rotated structures and in P11, three or more interhelical C—C distances were >2 Å greater than the experimental values. Thus, our lowest energy structures are more similar to the solid-state NMR structure. They are also consistent with the solution NMR data, except for the two weak NOEs closer to the N terminus.

Previous studies have been able to discriminate the correct conformation of TM helix oligomers using only the van der Waals energy.[18,19] It is worth examining under what conditions this is possible. Among the structures in Figure 1, the one with the lowest van der Waals energy is 1.9 Å from the native structure (see Supplementary Mate-

rial, Fig. S2). Thus, for the 18-residue GpA the van der Waals energy alone can reasonably predict the native-like structure, if the full energy function is used for the global search. However, when the search was done without the electrostatic and solvation energy terms, the native structure could not be discriminated (Fig. 3). Some of the lowest energy structures have large RMSD, and in these structures the helices tend to wrap around each other. This comparison shows that the van der Waals term can discriminate the native structure only when certain constraints are placed on the native structure, as was done in previous studies.[18,19,26]

To examine the effect of the heterogeneous membrane environment in determining the native state, simulations were also done in pure implicit water and cyclohexane. We considered all four crossing points and performed annealing starting from parallel helices. The minimized effective energies are shown in Figure 4(A) and (B) for GpA74–91 in water and cyclohexane, respectively (average effective energy versus RMSD plots are similar). The native state of GpA is not the global effective energy minimum in pure water [Fig. 4(A)]. This agrees with a recent MD study showing that the GpA structure is not stable in explicit water.[47] In pure cyclohexane, the native structure is at a shallow global effective energy minimum, ~1 kcal/mol lower than the non-native structure [Fig. 4(B)]. The low-energy structures with large RMSD were found to adopt antiparallel or left-handed conformations. IMM1 seems to discriminate the native state better than pure cyclohexane (see Figs. 1 and 4). One reason for that is that the membrane provides a barrier that prevents the conformational change from parallel to antiparallel conformation (the energetic advantages of the antiparallel topology will be discussed below). Also, the membrane model destabilizes off-register structures due to burial of the termini in the hydrophobic core. In longer peptides it will be the burial of flanking polar residues that will prevent off-register structures.

Simulated annealing was also performed in vacuum. In this case, the N terminus was acetylated and the C terminus methyl-amidated to avoid the large Coulomb interactions of charged groups in vacuum. The lowest minimized energy structures are close in energy to the NMR structure but have large RMSD from the native structure [Fig. 4(C)]. These low-energy structures with large RMSD include some left-handed, some antiparallel, and some irregular structures. The studies above suggest that accounting for the heterogeneous membrane environment is useful in modeling transmembrane proteins.

Conformational Search on the 29-Residue Segment of GpA

The solution NMR experiments used 40-residue peptides (Val62–Lys101) and the solid-state NMR experiments used 29-residue peptides (70–98). To examine the robustness of the results and the influence of polar residues flanking the TM sequence, we also considered for annealing GpA70–98. The effective energies of all the minimized structures obtained at different crossing points are shown in Figure 5. Most of the low-energy conformations are obtained from the second and third crossing points and RMSDs ranging from 1.0 to 2.5 Å, which is very close to the native structure. However, the lowest energy structure obtained at this crossing point has a large RMSD from the NMR structure and is left-handed. In general, low-energy structures with large RMSDs tend to form either a left-handed helical pair or a conformation having lateral interaction with large side chains. Their effective energies are lower than the native-like conformation by ~7 kcal/mol. Average effective energies give similar results (see Supplementary Material, Fig. S3).

A detailed analysis of nine low energy conformations with large RMSD from the native structure showed that the interhelical residue–residue interactions between Y93–R96, Y93–R97, Y93–L89, Y93–Y93, L89–L89, and L89–L90 constitute about 20–45% of the total interhelical interaction energy. The strong interaction energies are mainly due to favorable van der Waals interaction between these residues. These interactions are zero in the native structure. Figure 6 shows the positions of Y93, R96, R97, L89, and L90 in the NMR and the lowest energy non-native (left-handed) structures.

One possibility for the lack of discrimination of native GpA in the 29-residue peptide is that the Arg–Arg interactions are still not correctly reproduced and are in reality more repulsive than the energy function predicts, or the favorable solvation of the Tyr side chain by the lipid headgroups may be underestimated. Another possibility is the omission of the side-chain entropy of the system. The side-chain entropy is expected to be higher for the native conformation. Thus, the *free energy* may indeed discriminate the native structure.

Conformational Search on Antiparallel 18-mers

In vivo, the parallel arrangement of the two helices is dictated by the topology of the entire protein. However, no such constraints exist in studies of peptides in vitro. Hence, we have also done a conformational search with antiparallel topology using the 18-residue segment. In contrast to experimental observations, the calculations find antiparallel conformations with lower energy than all parallel conformations. Table V gives the energetics of two lowest effective energy and two lowest average effective energy antiparallel topologies of GpA. The electrostatic energies of these four antiparallel structures are ~17–24 kcal/mol lower than that of the NMR structure. Except for AP2, the van der Waals energies of these structures are also lower than the vdW energy of the NMR but the solvation energy is higher for the antiparallel structures. A large contribution to the energetic advantage of the antiparallel structure comes from antiparallel CO–CO interactions (Fig. 7). In the antiparallel structure the total CO–CO dipole interaction energy contributes ~−14 kcal/mol, whereas in the NMR structure, the total CO–CO interaction energy is ~+2.3 kcal/mol.

To test a different effective energy function, we performed molecular dynamics simulations using a Generalized Born membrane model (GBSW)[34,35] starting from the

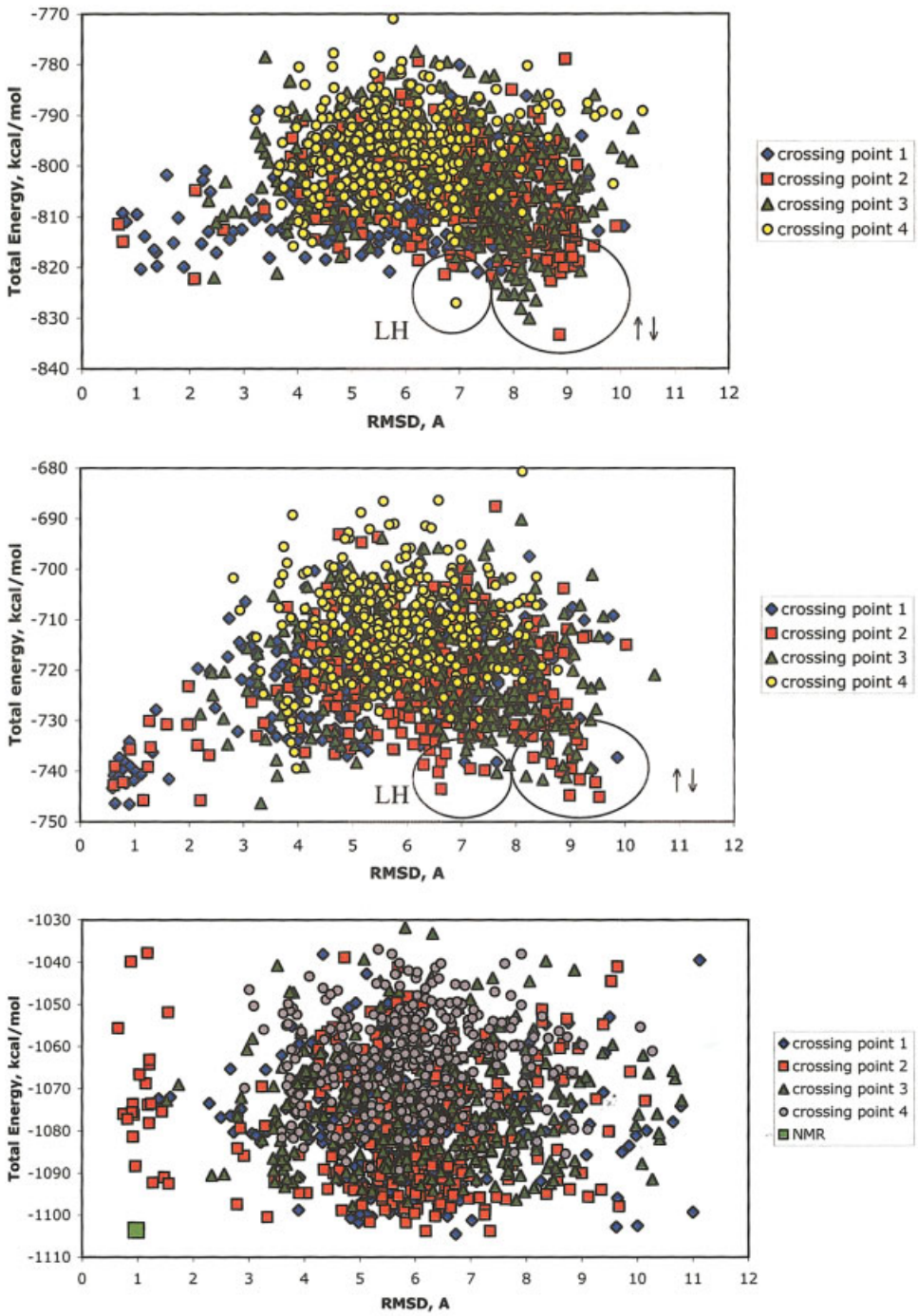

Fig. 4. Effective energy of all the conformations of 18-residue GpA in pure implicit solvents versus their RMSD from the NMR structure. Regions marked by LH and ↑ ↓ represent left-handed and antiparallel structures, respectively. (**A**) Water, (**B**) cyclohexane, and (**C**) vacuum. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

solid-state NMR, and several parallel and antiparallel low-energy structures. Table SIII in the Supplementary Material shows that GBSW also cannot discriminate the native-like from the antiparallel structures.

Finally, we performed MD simulations on the same structures using the Poisson-Boltzmann (PB) method. Because of the computational expense, only 2 ps of the dynamics was performed. Table VI gives the energy after PB molecular dynamics simulations followed by 300 steps of ABNR minimization of the four systems studied here. The lowest energy among the structures considered is observed for P5 followed by the solid-state NMR structure. Thus, the PB model correctly distinguishes the parallel structure of GpA as energetically superior. The deciding

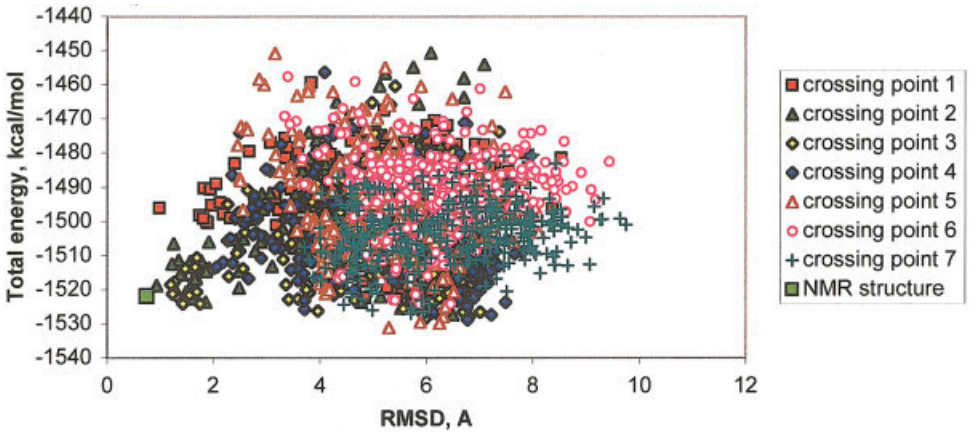

Fig. 5. IMM1 minimized energy of all conformations of 29-residue GpA versus their RMSD from the NMR structure. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

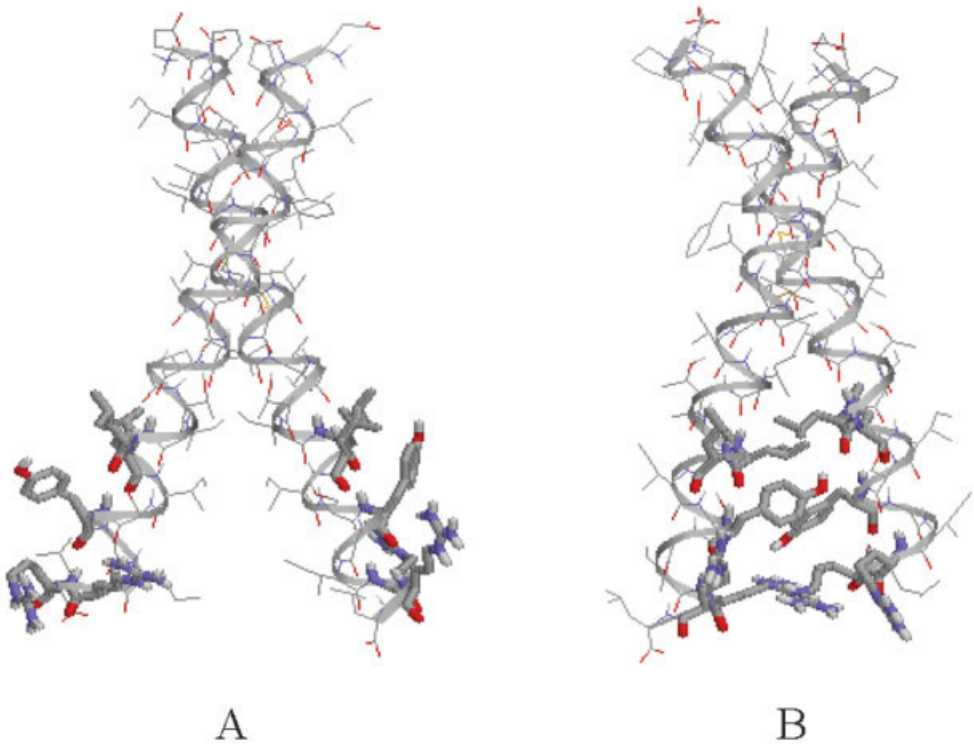

Fig. 6. Orientation of the amino acid residues L89, L90, Y93, R96, and R97 in 29-residue GpA. (**A**) NMR structure, (**B**) lowest effective energy left-handed parallel structure. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

factor seems to be the much more favorable solvation energy of the parallel structures.

To understand the origin of the more favorable solvation energy of the parallel structure we repeated single-point PB calculations for the solid-state NMR and the AP1 structures gradually turning off the charges starting from the N- and C-termini (Table SIV in the Supplementary Material). When the N- and C-termini were neutralized, the difference in total energy between the two structures was not diminished. The difference gradually diminishes when the charges on the first two or three residues from both sides are turned off. This shows that the residues, especially the backbone, outside the membrane are responsible for the differences seen between the energies of NMR and antiparallel structures.

DISCUSSION

GpA is the "hydrogen atom" of membrane protein structural biology. Yet, it still poses serious challenges to the theoretical biophysicist. A force field that neglects the heterogeneity of a lipid bilayer does not provide a realistic representation of transmembrane protein interactions. On

TABLE V. Energetics of the Lowest Effective Energy Antiparallel Structures of 18-Residue GpA with IMM1 Compared to That of the Best Parallel Structures

Structure	Total (vdW, elec, cov, solv)
Minimized energy	
P1	−861.6 (−190.3, −582.4, 43.4, −132.3)
AP1	−873.5 (−191.8, −604.0, 44.2, −126.9)
AP2	−869.7 (−190.1, −599.2, 42.9, −126.6)
Average energy	
P5	−595.4 (−160.5, −552.6, 257.2, −139.5)
AP3	−603.1 (−183.3, −571.0, 265.5, −126.2)
AP1	−602.7 (−164.6, −576.7, 263.9, −129.5)

All energies are in kcal/mol.

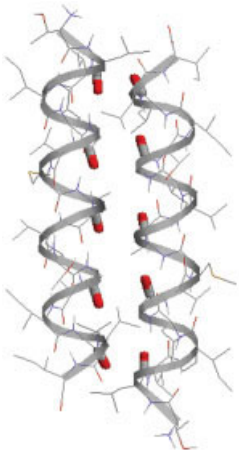

Fig. 7. Antiparallel CO–CO interactions in the lowest average effective energy antiparallel structure. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

TABLE VI. Energetics of the Solid-State NMR and Seven Decoys of GpA Using the PB Equation (PB Dynamics)

Structure	Total (vdW, elec, PBnp, PBelec, cov)	RMSD, Å
NMR	77.1 (−122.5, 376.6, 153.2, −542.5, 212.4)	1.56
P1	83.4 (−121.2, 343.9, 153.5, −503.9, 211.2)	1.47
P9	84.4 (−121.5, 377.8, 153.4, −535.0, 209.9)	1.03
P5	75.1 (−126.9, 375.2, 153.8, −533.6, 206.6)	1.82
P11	113.2 (−137.2, 371.5, 151.7, −515.9, 243.3)	2.11
P12	90.3 (−123.4, 382.9, 155.9, −539.0, 214.0)	1.79
AP1	80.5 (−114.1, 78.8,[a] 155.0, −252.2, 213.1)	8.53
AP3	83.6 (−123.9, 160.3, 154.4, −322.0, 214.8)	7.95

All energies are in kcal/mol and the RMSDs are in Å. PBelec and PBnp are the Poisson—Boltzmann electrostatic and nonpolar solvation energy contributions, respectively.

[a]The electrostatic energy of AP1 is much lower than all parallel and AP3 structures because AP1 T74:Oγ-Hγ was found to make strong interhelical hydrogen bond with the C terminus OT1 and OT2.

the other hand, treating membranes explicitly is not practical for structure prediction. Here we tested the ability of a physical, implicit membrane energy function to discriminate the native conformation of the transmembrane domain of GpA. Most of the previous work on GpA was done on an 18-residue peptide (GpA74–91), which is almost entirely immersed in the membrane. Here we also considered a 29-residue peptide (GpA70–98) that has additional flanking polar residues. For a true ab initio search, we considered both parallel and antiparallel topologies and different crossing points along the helix. The peptides were constrained to be α-helices and to remain in contact. Beyond these constraints, the coverage of conformational space was extensive.

We started from an analysis of the energetics of native GpA, using the two available experimental structures. Energetic analysis of these two NMR structures showed stronger interaction energy for the solid state structure due to differences in T87 hydrogen bonding and the crossing angles. However, the total energies were similar. The largest contribution to the interhelical energy was the van der Waals interaction. A large contribution to the interhelical energy came from the GXXXG motif. However, two-thirds of the interaction energy came from outside the GXXXG motif. Thus, non-GXXXG residues are also important for the stabilization of the helix dimer, in agreement with recent mutational studies on GpA.[44,45]

We used the minimized and the average energies as criteria to discriminate the native structure. For parallel 18mer GpA in the membrane, irrespective of the initial conformation, native-like conformations had the lowest IMM1 effective energy. Most of these lowest energy structures had an orientation of the helices identical to that of the NMR structure and satisfied the experimental distance constraints, although in many models the two HH pairs closer to the N terminus gave a weak NOE. This was also observed in other GpA simulations.[47] However, some low-energy structures were slightly rotated with respect to the NMR structures. Occasionally, some structures 3 to 4 Å away from native gave energies very close to the native-like structures. However, these were isolated, whereas the native-like structures formed a cluster. Also, the native-like structures were consistently low in energy whether one looked at the minimized or dynamics average energies.

The folding of GpA is believed to be driven by packing interactions at the dimer interface and to be largely independent of electrostatic forces.[7,19] Indeed, the overall contribution of electrostatic interactions to the interhelical energy in the native state is very small. However, this small value contains larger compensating favorable and unfavorable interactions. The favorable interactions come from the Cα-H· · ·O hydrogen bonds and the unfavorable ones from symmetric L76–L76, G79–G79, V80–V80, and G83–G83 interactions. In addition, in other systems polar interactions are known to be the major driving force for TM helix association.[48,49] Therefore, neglect of electrostatic interactions cannot be a generally viable strategy. Even here in GpA, conformational search with an energy function that lacks the electrostatic and solvation energy terms could not discriminate the correct structure. Consistent with previous work[13,16,17] a vacuum force field cannot unambiguously discriminate the native state. The heterogeneity of the membrane improves the discrimination by

raising the energy of structures that bury polar residues in the membrane or expose hydrophobic residues to water.

In the case of 29-residue GpA, several left-handed structures had somewhat lower energies than the NMR and native-like simulated structures. This is most likely due to deficiencies in modeling of the interactions between charged side chains, especially Arginine. The partial charges were chosen to reproduce the potential of mean force for Arg side chains surrounded by water. The potential of mean force may well be different next to a membrane. Another possibility is improper modeling of the interactions between charged and aromatic side chains with the lipid headgroups (in IMM1 the lipid headgroup area is treated like water). Future developments of the energy function will address these issues. Alternatively, it may be due to the omission of entropic factors, such as side-chain entropy, which would favor the native state (see below).

Both IMM1 and GBSW failed to discriminate the native parallel structure from antiparallel GpA configurations. Only the PB model was found to correctly favor the parallel structure. The solvation energy of the parallel structure in PB model is much more favorable than that of the antiparallel structure. Presumably, IMM1 (and perhaps the GBSW model) underestimates the decrease in solvation energy when the CO and NH dipoles just outside the membrane are parallel. Hints of the importance of interfacial water in stabilizing the parallel structure were obtained in a recent MD study of GpA in an explicit lipid bilayer.[47]

One fundamental difficulty when using atomistic models for structure predictions is that they do not give a precise value for either averaged or minimized energies. Side-chain orientations, helix orientations, or the exact position of the protein with respect to the membrane can fluctuate to produce a range of energy values within the native ensemble that is similar to the difference in free energy between the native and non-native ensembles. A second limitation is that effective energy values do not include protein entropy contributions, such as vibrational or side-chain entropy. In principle, native states are ensembles of conformations, and it is the free energy of the ensemble that should be at the global energy minimum. However, calculating side-chain and vibrational entropies is not trivial. For exact side-chain entropy calculations one needs to consider all rotamer combinations, which is practically impossible. Smart ways of avoiding the combinatorial explosion, such as dead end elimination[50,51] could provide a viable solution. Alternatively, empirical formulas for computing the entropy from MD simulations[52] could also be useful. These ideas will be explored in future work.

ACKNOWLEDGMENTS

We thank Dr. S.O. Smith for the solid-state NMR structure of GpA and Dr. I. Ubarretxena-Belandia for numerous discussions and comments on the manuscript.

REFERENCES

1. Lemmon MA, Engelman DM. Specificity and promiscuity in membrane helix interactions. Q Rev Biophys 1994;27:157–218.
2. Arkin IT. Structural aspects of oligomerization taking place between the transmembrane alpha-helices of bitopic membrane proteins. Biochim Biophys Acta 2002;1565:347–363.
3. Manolios N, Bonifacino JS, Klausner RD. Transmembrane helical interactions and the assembly of the T-cell receptor complex. Science 1990;249:274–277.
4. Kurosaki T, Gander I, Ravetch JV. A subunit common to an IGG FC-receptor and the T-cell receptor mediates assembly through different interactions. Proc Natl Acad Sci USA 1991;88:3837–3841.
5. Mackenzie KR, Prestegard JH, Engelman DM. A transmembrane helix dimer: structure and implications. Science 1997;276:131–133.
6. Smith SO, Song D, Shekar S, Groesbeek M, Ziliox M, Aimoto S. Structure of the transmembrane dimer interface of Glycophorin A in membrane bilayers. Biochemistry 2001;40:6553–6558.
7. Lemmon MA, Flanagan JM, Treutlein HR, Zhang J, Engelman DM. Sequence specificity in the dimerization of transmembrane alpha-helices. Biochemistry 1992;31:12719–12725.
8. Mingarro I, Whitley P, Lemmon MA, VonHeijne G. Ala-insertion scanning mutagenesis of the glycophorin A transmembrane helix: A rapid way to map helix–helix interactions in integral membrane proteins. Protein Sci 1996;5:1339–1341.
9. Russ WP, Engelman DM. The GxxxG motif: a framework for transmembrane helix association. J Mol Biol 2000;296:911–919.
10. Brosig M, Langosch D. The dimerization motif of the glycophorin A transmembrane segment in membranes: importance of glycine residues. Protein Sci 1998;7:1052–1056.
11. Senes A, Gerstein M, Engelman DM. Statistical analysis of amino acid patterns in transmembrane helices: the GxxxG motif occurs frequently and in association with beta-branched residues at neighboring positions. J Mol Biol 2000;296:921–936.
12. Kleiger G, Grothe R, Mallick P, Eisenberg D. GXXXG and AXXXA: common alpha-helical interaction motifs in proteins, particularly in extremophiles. Biochemistry 2002;41:5990–5997.
13. Treutlein HR, Lemmon MA, Engelman DM, Brunger AT. The Glycophorin A transmembrane domain dimer: seqeunce-sepcific propensity for a right-handed supercoil of helices. Biochemistry 1992;32:12726–12733.
14. Adams PD, Engelman DM, Brunger AT. Improved prediction for the structure of the dimeric transmembrane domain of glycophorin A obtained through global searching. Proteins 1996;26:257–261.
15. Adams PD, Arkin IT, Engelman DM, Brunger AT. Computational searching and mutagenesis suggest a structure for the pentameric transmembrane domain of phospholamban. Nat Struct Biol 1995;2:154–162.
16. Briggs JAG, Torres J, Arkin IT. A new method to model membrane protein structure based on silent amino acid substitutions. Proteins 2001;44:370–375.
17. Torres J, Briggs JAG, Arkin IT. Contribution of energy values to the analysis of global searching molecular dynamics simulations of transmembrane helical bundles. Biophys J 2002;82:3063–3071.
18. Pappu RV, Marshall GR, Ponder JW. A potential smoothing algorithm accurately predicts transmembrane helix packing. Nat Struct Biol 1999;6:50–55.
19. Kim S, Chamberlain AK, Bowie JU. A simple method for modeling transmembrane helix oligomers. J Mol Biol 2003;329:831–840.
20. Kochva U, Leonov H, Arkin IT. Modeling the structure of the respiratory syncytial virus small hydrophobic protein by silent-mutation analysis of global searching molecular dynamics. Protein Sci 2003;12:2668–2674.
21. Kukol A, Adams PD, Rice LM, Brunger AT, Arkin IT. Experimentally based orientational refinement of membrane protein models: a structure for the Influenza A M2 H+ channel. J Mol Biol 1999;286:951–962.
22. Kukol A, Arkin IT. vpu transmembrane peptide structure obtained by site-specific Fourier transform infrared dichroism and global molecular dynamics searching. Biophys J 1999;77:1549–1601.
23. Kukol A, Arkin IT. Structure of the influenza C virus CM2 protein transmembrane domain obtained by site-specific infrared dichro-

ism and global molecular dynamics searching. J Biol Chem 2000;275:4225–4229.
24. Torres J, Kukol A, Arkin IT. Use of a single glycine residue to determine the tilt and orientation of a transmembrane helix. A new structural label for infrared spectroscopy. Biophys J 2000;79: 3139–3143.
25. Kokubo H, Okamoto Y. Self-assembly of transmembrane helices of bacteriorhodopsin by replica-exchange Monte Carlo simulation. Chem Phys Lett 2004;392:168–175.
26. Kokubo H, Okamoto Y. Prediction of transmembrane helix configurations by replica-exchange simulations. Chem Phys Lett 2004;383: 397–402.
27. Fleishman SJ, Ben-Tal N. A novel scoring function for predicting the conformations of tightly packed pairs of transmembrane alpha helices. J Mol Biol 2002;321:363–378.
28. Dobbs H, Orlandini E, Bonaccini R, Seno F. Optimal potentials for predicting inter-helical packing in transmembrane proteins. Proteins 2002;49:342–349.
29. Lazaridis T. Effective energy function for proteins in lipid membranes. Proteins 2003;52:176–192.
30. Ducarme P, Thomas A, Brasseur R. The optimisation of the helix/helix interaction of a transmembrane dimer is improved by the IMPALA restraint field. BBA—Biomembranes 2000;1509:148–154.
31. Ducarme P, Rahman M, Brasseur R. IMPALA: a simple restraint field to simulate the biological membrane in molecular structure studies. Proteins 1998;30:357–371.
32. Spassov VZ, Yan L, Szalma S. Introducing an implicit membrane in generalized Born/solvent accessibility continuum solvent models. J Phys Chem B 2002;106:8726–8738.
33. Zheng WJ, Spassov VZ, Yan L, Flook PK, Szalma S. A hidden Markov model with molecular mechanics energy-scoring function for transmembrane helix prediction. Comput Biol Chem 2004;28: 265–274.
34. Im W, Feig M, Brooks III CL. An implicit membrane Generalized Born theory for the study of structure, stability, and interactions of membrane proteins. Biophys J 2003;85:2900–2918.
35. Im W, Lee MS, Brooks III CL. Generalized Born Model with a simple smoothing function. J Comput Chem 2003;24:1691–1702.
36. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. CHARMM: a program for macromolecular energy minimization and dynamics calculations. J Comput Chem 1983;4: 187–217.
37. Lazaridis T, Karplus M. "New view" of protein folding reconciled with the old through multiple unfolding simulations. Science 1997;278:1928–1931.
38. Lazaridis T, Karplus M. Effective energy function for proteins in solutions. Proteins 1999;35:133–152.
39. Masunov A, Lazaridis T. Potential of mean force between ionizable aminoacid sidechains in aqueous solution. J Am Chem Soc 2003;125:1722–1730.
40. Smith SO, Eilers M, Song D, Crocker E, Ying WW, Groesbeek M, Metz G, Ziliox M, Aimoto S. Implications of threonine hydrogen bonding in the glycophorin A transmembrane helix dimer. Biophys J 2002;82:2476–2486.
41. Crick FHC. Is alpha-keratin a coiled-coil. Nature 1952;170:882–883.
42. Petrache HI, Grossfield A, MacKenzie KR, Engelman DM, Woolf TB. Modulation of glycophorin A transmembrane helix interactions by lipid bilayers: molecular dynamics calculations. J Mol Biol 2000;302:727–746.
43. Mottamal M, Lazaridis T. The contribution of Cα-H.O hydrogen bonds to membrane protein stability depends on the position of the amide. Biochemistry 2005;44:1607–1613.
44. Schneider D, Engelman DM. Motifs of two small residues can assist but are not sufficient to mediate transmembrane helix interactions. J Mol Biol 2004;343:799–804.
45. Doura AK, Kobus FJ, Dubrovsky L, Hibbard E, Fleming KG. Sequence contest modulates the stability of a GXXXG-mediated transmembrane helix-helix dimer. J Mol Biol 2004;341:991–998.
46. Doura AK, Fleming KG. Complex interactions at the helix–helix interface stabilize the Glycophorin A transmembrane dimer. J Mol Biol 2004;343:1487–1497.
47. Nanda H, Sachs JN, Petrache HI, Woolf TB. Environmental effects on Glycophorin A folding and structure examined through molecular simulations. J Chem Theory Comput 2005;1:375–388.
48. Choma C, Gratkowski H, Lear JD, DeGrado WF. Asparagine-mediated self-association of a model transmembrane helix. Nat Struct Biol 2000;7:161–166.
49. Zhou FX, Merianos HJ, Brunger AT, Engelman DM. Polar residues drive association of polyleucine transmembrane helices. Proc Natl Acad Sci USA 2001;98:2250–2255.
50. Desmet J, De Maeyer M, Hazes B, Lasters I. The dead-end elimination theorem and its use in protein side-chain positioning. Nature 1992;356:539–542.
51. Lasters I, De Maeyer M, Desmet J. Enhanced dead-end elimination in the search for the global minimum energy conformation of a collection of protein side chains. Protein Eng 1995;8:815–822.
52. Schlitter J. Estimation of absolute and relative entropies of macromolecules using the covariance-matrix. Chem Phys Lett 1993;215:617–621.