JCTC Journal of Chemical Theory and Computation

Helix Interactions in Membranes: Lessons from Unrestrained Monte Carlo Simulations

Yana A. Vereshaga, Pavel E. Volynsky, Dmitry E. Nolde, Alexander S. Arseniev, and Roman G. Efremov*

M.M. Shemyakin and Yu.A. Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Ul. Miklukho-Maklaya, 16/10, Moscow V-437, 117997 GSP, Russia

Received May 5, 2005

Abstract: We describe one of the first attempts at unrestrained modeling of self-association of α-helices in implicit heterogeneous membrane-mimic media. The computational approach is based on the Monte Carlo conformational search for peptides in dihedral angles space. The membrane is approximated by an effective potential. The method is tested in calculations of two hydrophobic segments of human glycophorin A (GpA), known to form membrane-spanning dimers in real lipid bilayers. Our main findings may be summarized as follows. Modeling in vacuo does not adequately describe the behavior of GpA helices, failing to reproduce experimental structural data. The membrane environment stabilizes α-helical conformation of GpA monomers, inducing their transmembrane insertion and facilitating interhelical contacts. The voltage difference across the membrane promotes "head-to-head" orientation of the helices. "Fine-tuning" of the monomers in a complex is shown to be regulated by van der Waals interactions. Detailed exploration of conformational space of the system starting from arbitrary locations of two noninteracting helices reveals only several groups of energetically favorable structures. All of them represent tightly packed transmembrane helical dimers. In overall, they agree reasonably well with mutagenesis data, some of them are close to NMR-derived structures. A possibility of left-handed dimers is discussed. We assume that the observed moderate structural heterogeneity (the existence of several groups of states with close energies) reflects a real equilibrium dynamics of the monomers—at least in membrane mimics used in experimental studies of GpA. The elaborated computational approach is universal and may be employed in studies of a wide class of membrane peptides and proteins.

1. Introduction

Membrane proteins (MP) constitute ∼30% of all proteins encoded by whole genomes.[1] Delineation of the structure−function relationships for MPs represents an intriguing challenge. Apart from fundamental importance, solving the problem would be invaluable in the optimization of these molecules' behavior for pharmaceutical applications. Many of MPs contain several transmembrane (TM) fragments or function as oligomers. It is well-established now that protein−protein interactions in lipid membranes are very important for a large number of crucial cell processes.[2,3] Membrane bound helix associates are the most appropriate objects to study such interactions owing to the relative simplicity of the systems and their stability. In addition, hydrophobic and amphiphilic α-helices represent a dominant structural motif responsible for the binding of MPs to membranes. Thus, TM helix interactions mediate the functional activity of a large number of integral and peripheral MPs: receptors, ion channels, and others.[4] Furthermore, it has recently been established that dimerization of TM α-helices may play a crucial role in the functioning of

* Corresponding author phone: (7-095) 336 20 00; e-mail: efremov@nmr.ru.

10.1021/ct0501250 CCC: $30.25
Published on Web 09/02/2005

receptor tyrosine kinases.[2] Interaction of individual α-helices with membranes and membrane mimics has been extensively studied using both experimental and modeling techniques (see refs 5−7 for reviews). Analysis of atomic resolution structures of MPs revealed a number of common principles of helix packing in membranes.[8,9] On the other hand, molecular mechanisms that drive such association are still poorly understood due to the difficulties of their experimental studies. Therefore, elaboration of independent techniques is especially well-timed. Molecular modeling represents a promising alternative, which considerably extends and complements traditional structural biology tools such as X-ray and NMR spectroscopy.

A suitable system for the development of computational techniques to assess helix interactions in membranes is the hydrophobic segment 69−97 (GpA) of human glycophorin A from erythrocyte membranes.[10] This relatively small system has been extensively studied in experiments. In membranes and detergent micelles GpA forms α-helical homodimers with parallel (head-to-head) orientation. Site-directed mutagenesis[11] has revealed that the LIxxGVxxG-VxxT sequence pattern (where x marks an arbitrary residue) is important for dimerization. A spatial model of the dimer has been elaborated on the basis of these data and the results of computer simulations.[12,13] This has been done for two helices in vacuo using global optimization technique combining molecular dynamics (MD) and simulated annealing. To prevent helix dissociation, a number of intermonomer distance restraints have been employed. The most plausible model of the dimer represents a right-handed helical supercoil. It has been selected in agreement with mutagenesis data and geometrical rules of helix packing in globular proteins. Later, the model has been refined using several interhelical distance restraints obtained by NMR in detergent micelles[14] and by solid-state NMR in lipid bilayers.[15] Despite the atomic resolution of the models, NMR experiments yield only five pairs of restraints on interhelical distances. Because of such a small number of restraints, the question of uniqueness of the model proposed for 3D structure of GpA still remains to be answered. Moreover, there is a wealth of NMR data[16,17] confirming that in many cases there is an equilibrium conformational exchange between different oligomeric states of helices in membrane-mimic environments. Such effects have been also recently reported based on MD simulations of hydrophobic helices in explicit hydrated octane slab.[18]

Noteworthy is that all the aforementioned experiments with GpA have been carried out in membrane mimics, like detergent micelles and lipid vesicles. Even in these artificial environments the spatial structures of the dimer are somewhat different. Furthermore, helix dimerization is mediated by the micellar composition.[19] Real biological membranes are much more complex in terms of physicochemical properties and therefore may strongly affect helix interactions. Nevertheless, the appearance of the 3D model of GpA dimer in membrane mimics (hereafter called "the nativelike structure", although the spatial structure under native conditions is yet to be solved) has stimulated the development of theoretical methods to study helix interactions in MPs. A number of approaches resembling that suggested in ref 12 have since been applied to predict 3D models of a number of α-helical dimers.[20−24] Most of them possess severe limitations. First, the simulations (except in refs 23 and 24) were carried out either in vacuo or in a continuum dielectric with low permeability. Additionally, the monomers were always "rigid", and hence the common occurrence of local distortions in TM helices, like kinks and bends, was not taken into account. Finally, it was a priori proposed that helices adopt a TM orientation. These studies show that the NMR structure of the GpA dimer can be predicted more or less correctly even without media effects—it is sufficient to introduce just a few restraints (helicity, TM orientation, parallel packing). This is because GpA is a well-studied system, and such restraints can be defined before the simulation. In this case the answer already lies (at least partially) in the input—one knows a priori what one wants to get. However, to proceed with the new oligomers, for which the structural experimental data are missing, the membrane effects on structure and/or insertion mechanisms should be taken into account. Thus, the modeling of even the simplest membrane helical oligomers is not straightforward, and the development of new efficient algorithms seems very promising. To have a predictive power, they should not impose any structural restraints and should not be based on a priori knowledge of the mode of membrane binding for the peptides. Previously we elaborated such a computational approach based on unrestrained Monte Carlo (MC) simulations in the presence of heterogeneous implicit membrane (reviewed in ref 7). Applied to a large number of α-helical peptides, this method was efficient in reproducing the main tendencies in peptide−membrane interactions in accord with experimental data.

Here we undertake one of the first attempts at unrestrained ab initio modeling of self-association of two GpA α-helices in implicit membrane. The objective was to check whether the calculations are capable of predicting the dimer structures close to those observed in experiments. In addition to this, because the experimental data on GpA have been obtained in artificial membrane mimics, not in real membranes, it was challenging to explore other conformational possibilities for helix complexes— possibly, depending on the media properties, some of them may be realized under native conditions.

2. Simulation Details

The System. The calculations were performed for two GpA segments with sequences $S^{(69)}$EPEITLIIFGVMAGVIGTIL-LISYGIRR$^{(97)}$. All-atom starting structures of GpA were built in α-helical conformation—previously we have shown[25] that in an implicit membrane the GpA monomer folds in the TM α-helix from the initial random coil. As shown in Figure 1A, 10 and 20 dummy residues were attached to the N-terminus of monomer-1 and introduced between the monomers, respectively. This was done to change the orientation of peptides with respect to each other and to the membrane during the simulation. These "virtual" residues were taken from the standard library of the FANTOM program.[26] They do not contribute to the energy of the system. First the atom of the N-terminal dummy residue was always positioned in the center of the hydrophobic layer with

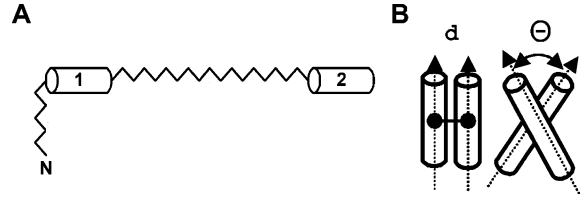

Figure 1. (A) Schematic drawing of the molecular system used in Monte Carlo simulations. Cylinders marked "1" and "2" represent α-helical monomers (segments 69−97) of GpA. Thin broken lines indicate dummy residues. Termini of the system are marked with symbols "N" and "C". (B) Definition of the geometrical parameters for two α-helices in a dimer: Θ and *d* are respectively the angle and the distance between the helical axes.

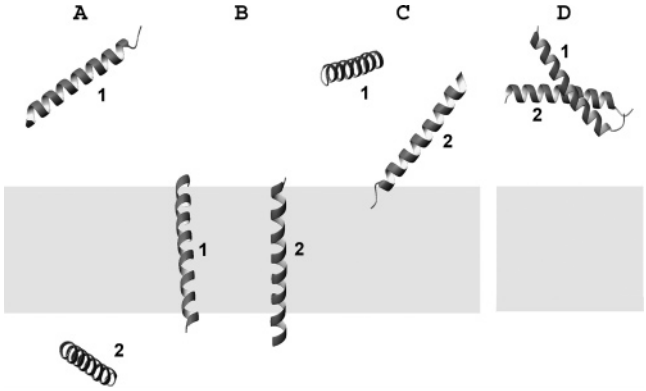

Figure 2. (A−C) The arbitrary chosen starting configurations of two noninteracting α-helices of GpA used in Monte Carlo simulations. (D) The starting structure of the NMR-derived model of the GpA dimer. The peptides are displayed with ribbons. The nonpolar layer of membrane is shown in gray. The monomers are marked "1" and "2".

coordinates (0,0,0). The inclusion of the "dummy" regions was dictated by the necessity of continuity of the protein backbone for simulations in dihedral angles space. Charges were assigned to the peptide atoms at pH 7. Simulations were started from three independent positions of noninteracting α-helices (Figure 2). Atomic coordinates of the NMR-derived model[14] of the dimer (GpA_{NMR}, residues 69−97) were taken from the Brookhaven Protein Data Bank,[27] entry 1afo. MC search for GpA_{NMR} in an implicit membrane was performed with dummy residues introduced as shown in Figure 1A. The starting structure was arbitrarily placed in an aqueous phase (Figure 2D). The lowest-energy state found for GpA_{NMR} in a membrane is referenced as $GpA_{NMR-mem}$.

Simulation Protocol. The peptides' conformational space was explored via MC search in torsion angles space as described elsewhere.[28] The membrane was represented by a "hydrophobic slab" described by an effective solvation potential. This was done using atomic solvation parameters (ASP) for gas−cyclohexane and gas−water transfer, which mimic the hydrophobic membrane core, lipid−water interface, and aqueous solution.[25] All-atom potential energy function for the protein was taken in the following form: $E_{total} = E_{ECEPP/2} + E_{solv} + E_{\Delta\psi}$. The term $E_{ECEPP/2}$ includes van der Waals, torsion, electrostatic, and hydrogen bonding contributions.[29] E_{solv} is the solvation energy

$$E_{solv} = \sum_{i=1}^{N} \Delta\sigma_i \, ASA_i \tag{1}$$

where ASA_i and $\Delta\sigma_i$ are accessible surface area (ASA) and ASP of atom i, respectively, and N is the number of atoms. The values of ASPs were taken from ref 28. Interaction of the protein with both aqueous and membrane environments is given by eq 2, where $\Delta\sigma_i$ depends on the z coordinate of atom i (the axis Z is normal to the membrane plane)

$$\Delta\sigma_i(z) = \begin{cases} \Delta\sigma_i^{mem} - 0.5\cdot(\Delta\sigma_i^{mem} - \Delta\sigma_i^{wat})\cdot e^{(|z|-z_0)/\lambda} & \text{if } |z| < z_0 \\ \Delta\sigma_i^{wat} + 0.5\cdot(\Delta\sigma_i^{mem} - \Delta\sigma_i^{wat})\cdot e^{-(|z|-z_0)/\lambda} & \text{if } |z| \geq z_0 \end{cases} \tag{2}$$

where $\Delta\sigma_i^{mem}$ and $\Delta\sigma_i^{wat}$ are ASP values for the type i atom in aqueous (wat) or nonpolar (mem) environments, respectively; z_0 is a half-width of the membrane (i.e., the hydrophobic layer is restricted by the planes given by the equation $|z| = z_0$); $D = 2z_0$ is the membrane thickness (30 Å); and λ is a characteristic half-width of the water−membrane interface (in this study $\lambda = 1.5$ Å).

Nonbond interactions were truncated with a spherical cutoff of 30 Å. As discussed earlier,[28] electrostatic interactions were treated with distance-dependent dielectric permeability $\epsilon = 4 \times r$. Prior to MC simulations the structures were subjected to 80−150 cycles of conjugate gradients minimization. Then several consecutive MC runs (5×10^3 steps each) with different seed numbers and sampled 5, 3, 2, 1 randomly chosen torsion angles were performed without restraints. At each MC step the structures were minimized via 50−150 conjugate gradients iterations. In each run the initial conformation was the lowest-energy one found in previous runs. In sum, $\sim 2.5 \times 10^4$ MC steps were performed for all systems in one complete MC simulation. To preserve the structure of the GpA_{NMR} model, in the beginning of the MC search (first 3.3×10^3 MC steps), a set of intermonomer distance restraints derived from the NMR structure was applied (Table 1 in the Supporting Information). The later MC stages were performed without restraints. In total, $\sim 4 \times 10^4$ MC steps were done for GpA_{NMR}. One of the resulting low-energy conformers (model GpA_{MC}) was subjected to constant-temperature MC simulation without minimization. This was done using 5×10^5 MC steps at $T = 300$ K.

The effect of TM potential ($\Delta\psi = 300$ mV) was taken into account using a special energy term ($E_{\Delta\psi}$)[7,30,31]

$$E_{\Delta\psi} = (F\Delta\psi/D)\sum_{i=1}^{N} q_i z_i \tag{3}$$

where q_i and z_i are partial charge and coordinate z of atom i, and F is Faraday's constant. For $|z_i| > z_0$, $\psi(z) = \text{const}$. Earlier we have shown[31] that modeling with $\Delta\psi = 100$ mV reproduces the behavior of a signal peptide adequately well as compared to experimental observations made under $\Delta\psi = 30$ mV. So, we assume that the magnitude of $\Delta\psi =$

300 mV used in simulations (later also denoted as "$\Delta\psi \neq 0$") roughly corresponds to $\Delta\psi \sim 100$ mV in real membranes.

The models of both "less hydrophobic" and "more hydrophobic" membranes were constructed as follows. In the former case the membrane's nonpolar part ($|z| \leq 15$ Å) was described by the ASP values obtained on the basis of the gas-octanol free energies of transfer.[28] In the latter case, this region was approximated with ASPs corresponding to a hypothetical solvent that is "more hydrophobic" than cyclohexane. To do this, the gas-cyclohexane values of ASPs were calculated according to the following formula: $\Delta\sigma^{mem'} = a \times \Delta\sigma^{mem}(a - 1) \times \Delta\sigma^{wat}$, where the coefficient a was taken to be equal to 1.2. The modified membrane models were employed in MC simulations analogous to those described above. Other details will be given elsewhere (Efremov et al., manuscript in preparation).

A Hypothetical Left-Handed Model of the GpA Dimer. Such the model (hereafter indicated as GpA_L) was built as follows. The low-energy MC state with parameters ($d \approx 5.7$ Å; $\Theta \approx 58°$) was used as a structural template. d and Θ are the distance and the angle between helical axes of the monomers (Figure 1B). Two ideal α-helices of GpA were fitted into this model over C_α atoms using the least-squares criterion, thus preserving the helix packing inherent in the MC model. The resulting structure was subjected to energy minimization in vacuo using 3×10^2 steepest descent iterations followed by 2×10^4 conjugate gradients steps. Five pairs of NMR distance restraints[14] were employed in the minimization protocol. The calculation and visualization of hydrophobic/hydrophilic properties of GpA were carried out using the molecular hydrophobicity potential (MHP) approach, as described elsewhere.[32] The MHP values were calculated on the solvent-accessible surfaces of α-helical segments Glu72-Ile95 in the models $GpA_{NMR-mem}$, GpA_{MC}, and GpA_L.

Analysis of the Results. Resulting states were analyzed using a set of auxiliary programs specially written for this. Only the low-energy states (in the range $[E_{min}, E_{min}+\Delta E]$, where E_{min} is the minimal energy, $\Delta E = 15$ kcal/mol) were considered. Mutual disposition of the monomers was described in terms of the values d, Θ, and the dimerization interface. The helix axes were calculated using the least-squares fit to coordinates of backbone atoms. Accessible surface areas (ASA) of residues and their secondary structure were assessed using the DSSP program.[33] Residue i was considered to lie on the dimerization interface if the difference between its ASA values in dimer and in monomer exceeds 25 Å^2 (10 Å^2 for glycines). Clustering of low-energy states was done based on the parameters of helix packing (Θ, d) and the composition of dimerization interface.

3. Results

The Choice of the Studied Systems and the Simulation Protocols. To understand to what extent the membrane environment determines the spatial structure of monomers and drives their association MC simulations of two α-helices of GpA were performed in vacuo and in a hydrophobic slab with and without TM voltage. (The term "vacuum" means that the solvation energy was omitted.) Two types of starting configurations were used in each medium, these being the randomly oriented helices and the GpA_{NMR} model (see Methods). Furthemore, in the former case several independent simulations with different random starts were carried out with a view of checking, whether the unrestrained MC search may be used to predict the states that agree with the experimental observations. In the latter case the following questions were addressed: "Whether the experimental structure is stable in membrane mimic?"; "Whether the yielded low-energy states compete in energy with the "nativelike" structures?"; and "How do they bind to the membrane (geometry of insertion)?" Comparison of the results obtained in both cases is required to assess the quality of sampling of the GpA's conformational space: is it sufficient enough to provide solutions that agree with the experimental data. It is important that the low-energy states obtained in simulations of the same type (see below) revealed similar energies, structures, and modes of membrane binding. This gives strong grounds to believe that the essential sampling of the conformational space was reached in MC simulations.

Simulations in Vacuo: Wrong Way to the "Nativelike" Structure of the Dimer. The low-energy states obtained from random starts represent either "hairpin (HP)-helix" structures or antiparallel (head-to-tail) dimers ($TM_{\uparrow\downarrow}$) (not shown). They demonstrate large conformational heterogeneity—HP-helices may be destabilized on different residues, and spatial disposition of the monomers varies in a wide range. Furthermore, interfaces between the monomers drastically differ from those in the GpA_{NMR} structure. A similar situation was seen with the simulations starting from the GpA_{NMR} model. In this case the initial "realistic" structure significantly loses in energy to the conformers found from random starts. In vacuo the "nativelike" model is unstable—the monomers rapidly change their mutual orientation, and the structure converges into the ones obtained from random starts (not shown). Therefore, unrestrained modeling in vacuo cannot be used for adequate simulations of TM helix interactions. Below we present the results obtained in a symmetrical membrane. (The term "symmetrical" means that both sides of the membrane are equivalent in the mathematical sense.)

Two Helices in a Symmetrical Hydrophobic Slab: Predominance of Misfolded Dimers. As seen in Figure 3A, the presence of a hydrophobic layer stabilizes the α-helical structure of both monomers and forces them to adopt TM orientations. Also, in their energetically favorable states the helices pack together and form antiparallel ($|\Theta| > 90°$, $TM_{\uparrow\downarrow}$) or parallel ($|\Theta| < 90°$, $TM_{\uparrow\uparrow}$) dimers. The population of the former ones is considerably higher. Helix packing parameters for some of the low-energy states are shown in Table 1. It can be seen that the states $TM_{\uparrow\downarrow}$ and $TM_{\uparrow\uparrow}$ have low energies of interaction with the environment (E_{solv}): the hydrophobic residues in the central part of helices are exposed to apolar membrane core, while the hydrophilic ones (on the GpA termini) are accessible to water. Also, the overall stability of such dimers is determined by the saturation of their H-bonding potencies (maximal number of residues in α-helix) and by favorable van der Waals interhelical contacts.

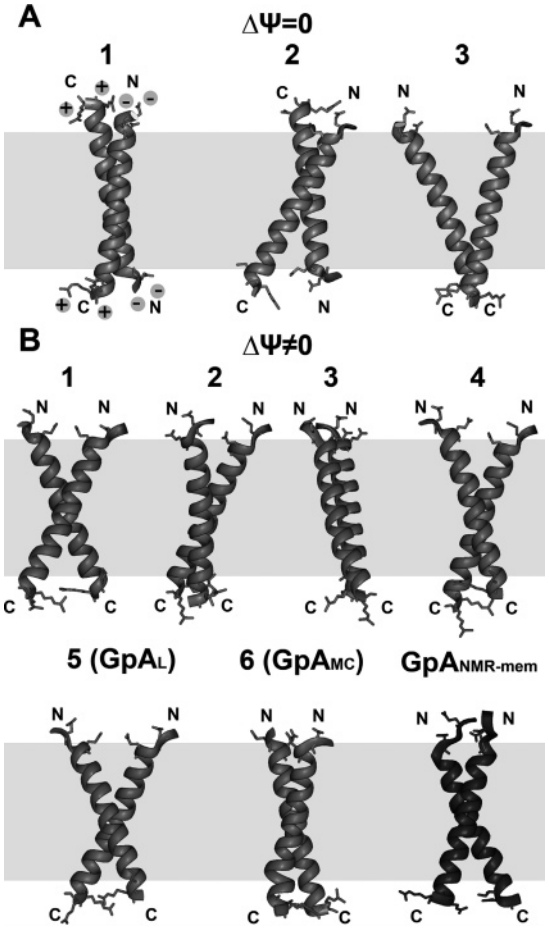

Figure 3. Monte Carlo simulations of two α-helices of GpA in implicit membrane without (A) and with (B) applied TM potential. Groups of the low-energy states are indicated with numbers (as in Table 1). $GpA_{NMR-mem}$ is the lowest-energy state of the NMR-derived model of GpA. The peptide's termini are marked with symbols "N" and "C". Side chains of positively (Arg96, Arg97) and negatively (Glu70, Glu72) charged residues are shown in stick presentation. In the structure A-1 they are also marked with symbols "+" and "−", respectively. The peptides are displayed with ribbons. The nonpolar layer of membrane is shown in gray.

It is remarkable that in all TM dimers occurrence of the motif GxxxG on the helix−helix interface correlates with low values of the E_{vdW} energy term. $TM_{⇅}$ structures reveal smaller values of the electrostatic term (E_{elect}) due to interactions between opposite charges on the monomers' termini. In the calculations starting from the GpA_{NMR} model, the low-energy states ($GpA_{NMR-mem}$) represent TM α-helical dimers, and the monomers almost do not change their packing as compared to the initial structure. Tight packing of parallel helices in $GpA_{NMR-mem}$ (low E_{vdW}) compensates for the unfavorable electrostatic contacts between their termini.

Electrostatic Potential on the Membrane: The Way To Correct Folding of the Dimer? *Unrestrained Simulations from Arbitrary Starts.* The low-energy states obtained from random starts in a membrane with applied TM voltage are characterized as follows. Like in a symmetric ($\Delta\psi = 0$ mV) membrane, the monomers retain well their initial α-helical conformation (Figure 3B). Both α-helices adopt a

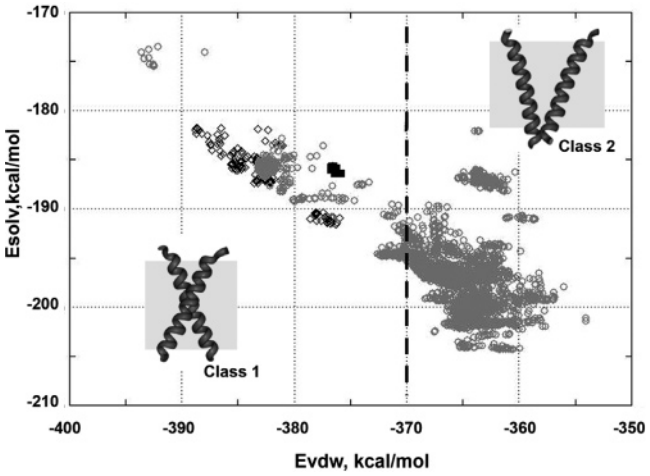

Figure 4. Two classes of the low-energy states found via Monte Carlo simulations in implicit membrane with applied TM potential. Distribution of the dimers over their solvation (E_{solv}) and van der Waals (E_{vdW}) energy terms. Class 1: tightly packed dimers with symmetrical interface; class 2: dimers with helices interacting by the terminal parts. Typical examples of structures from both classes are shown with ribbons. Dashed vertical line indicates an approximate boundary between the two classes. The symbols "◇*" and "■" mark respectively $GpA_{NMR-mem}$ and GpA_{MC} states. For other details see the legend to Figure 3.

TM-orientation, and their hydrophobic parts strongly interact with the medium imitating lipid bilayer—the corresponding impact to the solvation energy of the dimer (E_{solv}) being ~ -35 kcal/mol for the hydrophobic membrane core and ~ -15 kcal/mol for the interfacial region. The peptides' termini reveal highly favorable contacts with water, E_{solv} being ~ -135 kcal/mol. Moreover, in all low-energy states the monomers constitute tightly packed TM complexes, i.e., the formation of protein−protein contacts in a nonpolar environment is an energetically more beneficial process than giving to each of the peptides all the accessible area for contacts with the medium. At the same time, a number of significant differences were observed comparing with the case $\Delta\psi = 0$. For instance, the dipole moments of the α-helices are oriented in the direction of the TM electric field. As a result, $TM_{⇈}$-dimers typical of GpA in membrane are well reproduced in the calculations (Figure 3B). Besides, the presence of $\Delta\psi$ affects the geometry of helix packing: unlike the case $\Delta\psi = 0$, the resulting states agree better with the GpA_{NMR} model (Table 1). Additionally, a lesser degree of conformational heterogeneity of the MC states results in a restricted number of solutions. To define how adequate the dimeric MC structures are, it is necessary to inspect the mutual positions of the helices and to pinpoint the interactions driving the dimerization in the membrane.

With this aim in view, the total energy and its components for both the entire complexes and individual monomers and residues were considered, and the dimerization interfaces were delineated. The results reveal two highly populated classes of conformers (Figure 4). In the first of them the area of intermonomer contacts is located in the central part of the dimer, while in the second class the subunits interact mostly by their termini. States from both classes have

Table 1. Parameters of the Low-Energy States of the GpA Dimer Found via Monte Carlo Simulations in Implicit Membrane with and without Applied Transmembrane Voltage

Group[1]	d [2]	Θ [2]	dimerization interface [3]	E_{total} [4]	$E_{elect.}$	E_{vdW}	E_{solv}
Membrane with $\Delta\psi = 0$							
1 ($TM_{\uparrow\downarrow}$)	9.2 (0.4)	165.3 (2.7)	SEPEITLIIFGVMAGVIGTILLISYGIRR ___E__LI__GV__G__T__LI___I__ R__GY_IL__T___G__VG__I___E___	-584.5 (2.4)	10.3 (1.3)	-373.1 (8.4)	-173.0 (5.4)
2 ($TM_{\uparrow\downarrow}$)	9.8 (0.1)	142.2 (0.4)	_____I__V__G__T__LI__GI_R RR___S___I__I___M___I__T_____	-585.6 (3.6)	13.4 (0.5)	-362.5 (0.6)	-181.9 (0.6)
3 ($TM_{\uparrow\uparrow}$)	8.9 (0.2)	-33.8 (1.3)	_____V___I__I___I__ _____I__IS_____	-577.5 (1.3)	14.1 (0.2)	-357.4 (1.5)	-183.5 (0.5)
Membrane with $\Delta\psi \neq 0$ Class-1 structures							
1 ($TM_{\uparrow\uparrow}$)	10.1 (0.1)	-50.1 (0.5)	SEPEITLIIFGVMAGVIGTILLISYGIRR _____II__VM__V___I_____ _____II__VM__V___I_____	-629.8 (0.5)	18.2 (0.2)	-370.4 (0.3)	-191.8 (0.4)
2 ($TM_{\uparrow\uparrow}$)	9.2 (0.1)	22.1 (1.1)	____I___I__V__GV__T___I__G__ _____I__V__GV__T__LI_____	-632.4 (1.7)	18.9 (0.2)	-382.3 (1.5)	-185.9 (0.6)
3 ($TM_{\uparrow\uparrow}$)	9.0 (0.1)	5.9 (0.2)	____I___I__VM__V___I__I___I__ ___E__LI__GV__G__GT__L___G___	-626.9 (2.7)	18.8 (0.3)	-379.3 (1.9)	-188.0 (1.3)
4 ($TM_{\uparrow\uparrow}$)	6.9 (0.1)	42.6 (0.8)	_____LI__GV_AG__G__LL__Y__R_ ___E___I__GV_AG__G__LL__Y____	-629.3 (1.6)	21.6 (0.4)	-392.3 (1.4)	-174.8 (0.7)
5 ($TM_{\uparrow\uparrow}$)	5.5 (0.2)	60.2 (2.0)	_____I__GV_AG__G__LL__Y____ _____I__GV_AG__G__LL_____	-626.5 (2.9)	19.3 (1.3)	-378.6 (3.0)	-190.2 (1.2)
6 (GPA_{MC}) [5] ($TM_{\uparrow\uparrow}$)	9.1 (0.4)	-25.9 (1.4)	___E__LI__GV__GV__T___I_____ ___E__LI__GV__GV__T___I_____	-628.4 (4.4)	17.4 (0.4)	-372.7 (2.5)	-192.0 (3.7)
$GpA_{NMR\text{-}mem}$ 6 ($TM_{\uparrow\uparrow}$)	7.1 (0.1)	-39.8 (0.3)	___E__LI__GV__GV__T___I_____ ___E__LI__GV__GV__T___I_____	-622.3 (1.0)	20.6 (0.2)	-382.5 (3.7)	-187.0 (2.8)
Class-2 structures [7]							
a ($TM_{\uparrow\uparrow}$)	7.1 (0.9)	-39.9 (4.2)	_____I__GI__ _____T__LI__GI__	-623.1 (2.1)	17.0 (1.5)	-359.7 (2.4)	-196.5 (3.5)
b ($TM_{\uparrow\uparrow}$)	9.5 (0.8)	47.0 (5.9)	__PE__L_____ _____L_____	-624.1 (1.7)	18.1 (0.4)	-361.1 (2.6)	-195.5 (1.3)
c ($TM_{\uparrow\uparrow}$)	10.9 (0.5)	47.8 (4.2)	_____LI__G__G_____ __P___L__F_____	-622.5 (1.4)	19.0 (0.6)	-359.4 (3.9)	-196.5 (2.2)

[1] The group numbers are the same as in Figure 3. The symbols $TM_{\uparrow\downarrow}$ and $TM_{\uparrow\uparrow}$ indicate antiparallel and parallel TM dimers, respectively. [2] d and Θ are respectively the mean distance (in Å) and the mean angle (in deg) between helix axes for a given group of states. Standard deviations are given in brackets. [3] Residues found on the helix−helix interface are indicated with their one-letter code. Gray hatching shows, whether these residues also contribute to dimerization according to mutagenesis data.[11] Sequences of the monomers 1 and 2 are shown either in antiparallel or in parallel orientations—for $TM_{\uparrow\downarrow}$ and $TM_{\uparrow\uparrow}$ states, respectively. [4] E_{total}, E_{elect}, E_{vdW}, and E_{solv} are the mean values of total, electrostatic, van der Waals, and solvation energies, respectively (in kcal/mol). Standard deviations are given in brackets. [5] GpA_{MC} is the calculated model most close to the NMR structure of the dimer.[14] [6] $GpA_{NMR-mem}$ is the lowest-energy state found starting from the NMR-derived model (GpA_{NMR}) of the dimer[14] in implicit membrane. The values Θ and d in $GpA_{NMR-mem}$ and in GpA_{NMR} are very similar. [7] These class-2 states are not shown in Figure 3B.

comparable values of total energy, but the terms E_{solv} and E_{vdW} differ. Thus, the states of class-1 are characterized by a tighter packing of helices (Figure 4). As a result, E_{vdW} for them is ~10 kcal/mol lower than in class-2. The states of class-2 have a smaller area of intermonomer contacts (370 ± 50 Å^2 and 250 ± 70 Å^2 in classes 1 and 2, respectively). Because of this the subunits interact with the membrane to a greater extent, thus resulting in lower values of E_{solv}.

As argued below (see Discussion), the class-1 of states was selected for further inquiry. In total, about 800 such structures were delineated. They have diverse types of helix

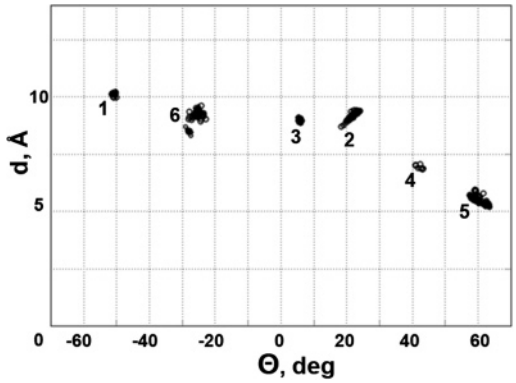

Figure 5. Distribution of the low-energy states of class-1 found via Monte Carlo simulations in implicit membrane with applied TM potential over their helix packing parameters *d* and Θ (distance and angle between helix axes, respectively). Numbering of groups is that as in Table 1.

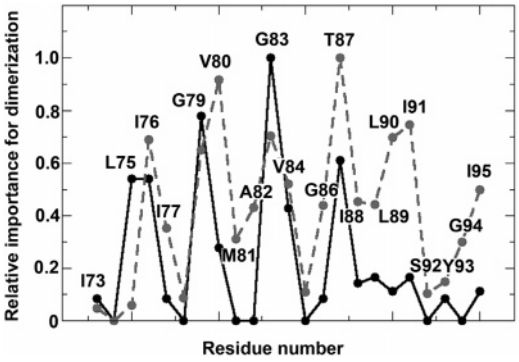

Figure 6. Relative importance of GpA residues in dimerization. Bold line: mutagenesis data, taken from ref 11. Dashed line: frequency of occurrence of residues on the helix−helix interface: results of summary analysis of the groups of low-energy states obtained via Monte Carlo simulations (listed in Table 1). The distributions are normalized on the range [0, 1].

packing (Table 1). Accordingly, the composition of the dimerization interface may vary as well. Nevertheless, in all these structures, most of the residues involved in helix−helix contacts were also found on the interface in experimental studies. The groups of low-energy states represent well-defined and compact clusters which do not overlap (Figure 5), and therefore only a limited number of possible packing geometries should be used for future analysis. No outliers of the clusters were found. As shown in Table 1, a complete conformity with the known motif of dimerization was not observed for the theoretically predicted states, although in some of them six of seven residues were determined correctly. Indeed, summary analysis of dimerization interfaces in all groups listed in Table 1 yields a distribution of residues most frequently involved in helix−helix contacts quite similar to that found via mutagenesis in ref 11 (Figure 6). Putative reasons for the discrepancies observed near the C-terminus are discussed later (see Discussion).

In the considered structures α-helices pack in a symmetrical (or nearly symmetrical) manner with respect to one another (Table 1). Because the α-helices have identical sequences, this result seems quite plausible, although asymmetrical homodimers have also been predicted for some other TM peptides (e.g., ref 18). We should note that the NMR model has been built as a symmetrical dimer as well. Interestingly, the formation of complexes with either negative (groups 1 and 6, Table 1) or positive (groups 2−5) values of the angle Θ is possible, corresponding to right- and left-handed double super-helix, respectively. For both types of structures the largest population of states is observed when $d \sim 9$ Å. This agrees with the data on statistical analysis of high-resolution structures of α-helical complexes in MPs:[8,34] in the majority of the helical pairs $d \sim 9.6$ Å, while those with $d \sim 7.0$ Å are rare. Among the states with $\Theta < 0$ is one (highly populated) group of dimers with packing parameters close to the model GpA_{NMR} (Table 1). The best correspondence is observed for the group-6 (GpA_{MC}). These are the states for which the dimerization interface is entirely symmetrical and agrees well with that found by NMR. To inspect the stability of the resulting dimeric states, one of the found low-energy conformers (GpA_{MC}) was subjected to constant-temperature MC simulation without minimization. It was shown that the GpA_{MC} model was stable during 5×10^5 MC steps: resulting all-atom RMSDs were within 1 Å from the initial structure.

Simulations with the GpA_{NMR} Start. Resulting low-energy states represent TM complexes with the termini accessible to water (Figure 3B, Table 1). The initial structure is well retained—for residues 71−95 the backbone RMSDs with the starting NMR model do not exceed 1.2 Å. The model $GpA_{NMR-mem}$ has the total energy close to that in other groups of states (Table 1). Comparison with one of such structures (GpA_{MC}) reveals that both dimers have similar geometry of membrane binding (Figure 3B) and dimerization interface (Table 1). One exception is provided by residues Leu75—in the model GpA_{MC} their side chains do not form such tight intermonomer contact, as in the experimental structure. Despite a relatively high backbone RMSD as compared to the GpA_{MC} structure (2.5 Å on the region 71−95) and somewhat overestimated distances for a number of helix contacts on the N-terminus (Table 2 in the Supporting Information), in overall, the model GpA_{MC} represents a rather good approximation for the structure observed by NMR in micelles. We should mention that, unlike the GpA_{NMR} model, the helices in the GpA_{MC} states were not restrained to their ideal conformations.

The Role of Hydrophobicity Degree of the Membrane. To estimate the sensitivity of the modeling results to changes of the parameters of the lipid bilayer, calculations analogous to those described above were conducted using "less hydrophobic" and "more hydrophobic" membranes. In the former case the membrane's nonpolar part was approximated with octanol, while in the latter study a hypothetical solvent that is "more hydrophobic" than cyclohexane (in terms of free energy of transfer from water) was employed (see "Methods"). It was shown that the application of such modified membranes leads to unrealistic conformations of the dimer: improper helix packing, helix destabilization, the appearance of nondimeric forms, etc. (data not shown,

manuscript in preparation). This clearly demonstrates the importance of the balance between E_{solv} and other energy terms: underestimation as well as overestimation of the media effects may considerably affect the resulting low-energy states.

Left-Handed α-Helical Dimer. As seen in Table 1, there are several groups of dimers with $\Theta > 0$. Two of them are especially interesting, namely groups 2 and 5. This is because the packing parameters in group-2 are frequently observed in left-handed TM helical pairs,[34] while the densely packed complexes of group-5 bear a resemblance to the left-handed dimers proposed earlier based on mutagenesis data.[11] Monomers in this group have favorable van der Waals contacts and demonstrate symmetrical tight packing of α-helices—mainly due to interactions in their middle parts, via the motif GVxAGxxG. Whether is it possible a dimeric structure with $\Theta > 0$, which also satisfies the constraints, imposed by mutagenesis and NMR results? To answer this question, a hypothetical structure was built based on one of MC conformers from the group-5 (see "Methods"). Detailed analysis of helix packing in structures GpA_L, $GpA_{NMR-mem}$, and GpA_{NMR} reveals that the model GpA_L satisfies reasonably well the two independent sets of NMR distance restraints.[14,15] In the former case, average violations of the restraints are 1.10, 0.86, and 0.74 ± 0.07 Å for GpA_L, $GpA_{NMR-mem}$, and an ensemble of 19 experimental models (GpA_{NMR}). In the latter case, such violations are 0.84, 0.18, and 0.47 ± 0.17 Å, respectively (Table 3 in the Supporting Information). So, the model GpA_L fits better to the solid-state NMR data. Interestingly, MC simulations of the model GpA_{NMR} in an implicit membrane lead to a very good agreement with the restraints measured in lipid vesicles as compared to those obtained in micelles: during MC simulations the "micellar" model converges to the structure observed in lipid vesicles. Therefore, overall, the hypothetical model GpA_L does not contradict NMR results. Analysis of helix−helix contacts in GpA_L demonstrates that the dimerization interface includes residues GXXXG (Table 1). Interestingly, residues Leu75, Ile76, Gly79, Val80, Gly83, and Thr87, the importance of which for helix association was proved by the mutagenesis results, form a helix−helix interface in the model with $\Theta > 0°$. Apart from those, the residues Gly86, Leu89, and Ile91 are also involved in helix association. According to mutagenesis data, replacement of each of the last ones seriously affects dimerization, even though they lie apart from the interface in the model GpA_{NMR}.

Complementarity of Hydrophobic/Hydrophilic Surfaces. Two-dimensional hydrophobicity maps for α-helices in models GpA_{NMR}, GpA_L, and GpA_{MC} are shown in Figure 7. Contour isolines display on their surfaces hydrophobic regions with high positive values of MHP, and the gray-hatched areas indicate dimerization interfaces. It can be seen that the characteristic feature of each α-helix is its strong hydrophobicity and presence of a relatively more polar "Λ-shaped" pattern. Its left and right "arms" are formed by residues Ser92, Thr87, Gly83, Gly79 and Gly94, Gly86, Gly79, respectively. In the right-handed complexes (GpA_{NMR}, GpA_{MC}) the dimerization interface fits to the first polar region (Figure 7A,B), while in the left-handed model (GpA_L)—to the second one (Figure 7C). As seen in Figure 7A,B, in the structure GpA_{MC} the interfacial contact area agrees well with that observed by NMR, although in the theoretical model the contact pattern has a somewhat smaller tilt with respect to the helix axis (and, therefore, smaller value of the angle Θ). In addition, in all the models the interface is composed of parallel hydrophilic and hydrophobic stretches. In the right- and left-handed dimers the last one is formed by residues Ile76, Val80, Val84, Ile91 and Val80, Val84, Leu89, Leu90, respectively. Analysis of the dimerization interface in both right- (GpA_{MC}) and left- (GpA_L) handed MC models shows that the interdigitation of the side chains ("grooves-into-ridges") observed in the NMR structures is well reproduced (Figure 7, bottom). Moreover, in the calculated structures χ_1 rotameric states of residues correspond well to those in the NMR models (data not shown).

4. Discussion

Simulations in Vacuo or in a Hydrophobic Slab? Unrestrained modeling in vacuo clearly demonstrates that media effects are very important for proper description of folding and assembling of TM α-helices. Thus, the simulations from random starts were unable to reproduce the "nativelike" parallel α-helical dimers observed in experiments. In addition, the well-packed GpA_{NMR} structure was unstable in vacuo. The last result points to the fact that the failure of MC search from arbitrarily chosen positions of helices is not related to the problem of insufficient sampling of the conformational space for the system. Instead, the reason is that the absence of heterogeneous polar/apolar membrane environment leads to overestimated electrostatic interactions between charged termini of the peptides. In addition, aliphatic side chains of a large number of residues tend to form hydrophobic contacts with each other but not to be exposed on the surface.

On the contrary, MC search from random starts with noninteracting α-helices reveals that monomers of GpA retain α-helical conformation in the hydrophobic slab, adapting TM orientation and tightly packing together to form stable parallel and antiparallel dimers. In a symmetrical slab only a small fraction of the low-energy states resembles the GpA_{NMR} structure. Modeling in such a membrane does not possess high predictive power because it results in a large number of possible solutions, these being TM complexes with close energies but drastically different packing. Moreover, most of them are misleading, their "head-to-tail" topology not resembling that found in experiments. It is important that, independently of the starting configuration (random starts or GpA_{NMR} model), resulting low-energy states have close values of the total energy. This demonstrates high efficiency of the MC search and makes us confident that the developed protocol provides an essential sampling of states. Similar findings have also been reported in our previous studies of other peptides and proteins (reviewed in ref 7).

As for the surprisingly high population of misfolded ($TM_{\uparrow\downarrow}$) states in a symmetrical membrane, we suppose that it is due to the equivalence of the two membrane sides: since there is no preferential direction for the dipole moments of α-helices, both types of the dimers are energetically favor-

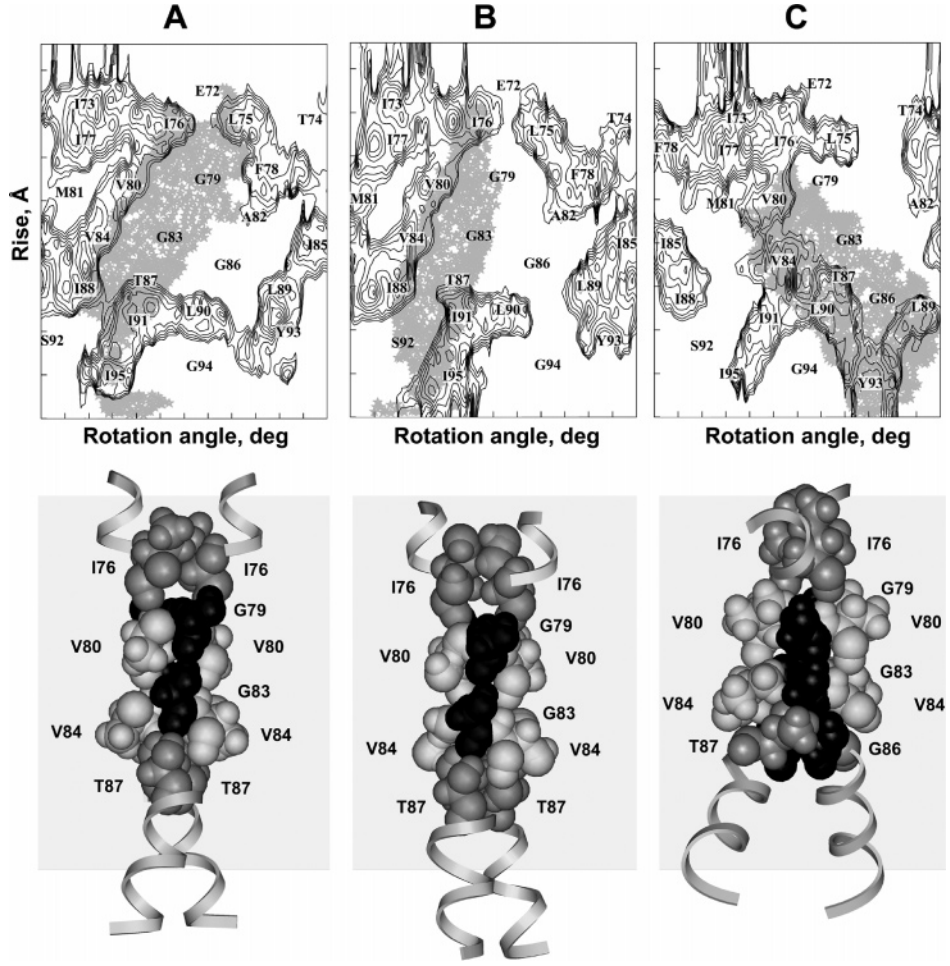

Figure 7. Hydrophobicity and packing of the NMR-derived and calculated models of GpA dimer: GpA_{NMR} (A), GpA_{MC} (B), and GpA_{L} (C). (Top) Hydrophobic properties of α-helices. 2D isopotential map of the molecular hydrophobicity potential (MHP) on the peptide surface calculated as described in ref 31. The value on the *X* axis is the rotation angle about the helix axis; the parameter on *Y*-axis is the distance along the helix axis. MHP is given in octanol−water logP units. Only the hydrophobic areas with MHP > 0.09 are shown. Contour intervals are 0.015. The positions of residues are indicated by letters and numbers. Gray-hatching marks helix−helix interface in the dimer. (Bottom) Molecular models of the dimers. The peptides are shown with ribbons. Residues on the dimerization interface are shown in ball presentation. Glycines are colored in black, valines—in light gray, Ile76 and Thr87—in medium gray. The nonpolar layer of membrane is shown in light gray.

able. In contrast, real biological membranes manifest TM potentials ($\Delta\psi$) present across lipid bilayers. They feed and control many biological processes, ranging from energy conversion and channel gating to protein insertion and translocation. The role of $\Delta\psi$ seems to be especially important for TM helices of GpA because they carry charges of opposite sign on their termini and possess large dipole moments (~90−100 D). It seems reasonable that the presence of $\Delta\psi$ destabilizes the $TM_{\uparrow\downarrow}$-states and sufficiently favors the $TM_{\uparrow\uparrow}$-ones, where the dipole moments of helices are oriented along the electric field. Therefore, to provide a more realistic treatment of helix association, the influence of $\Delta\psi$ should be taken into account.

Correct Assembling of Helices Requires a Hydrophobic Slab and a TM Voltage. In some aspects the results of simulations with applied TM voltage are similar to those in symmetrical membrane: the energetically favorable states represent membrane-spanning and densely packed α-helical dimers. At the same time, occurrence of $\Delta\psi$ provides a far more correct description of helix behavior. In particular, it is important for the "nativelike" association of TM α-helices maintaining their parallel orientation. As mentioned above, the found low-energy states are divided into two large classes (1 and 2, or "X"- and "V"-shaped conformers, respectively). The class-1 of states was selected for further inquiry. Such a choice is determined by the following notions. Primarily, conformers from the class-1 have significantly lower energy of intermolecular interactions as compared to the class-2: -60 ± 15 and -30 ± 5 kcal/mol, respectively. In the former case the contact area comprises 6−10 residues, and in the second one it is much less extended—only 3−6 residues. Second, the class-1 states have quite a similar interface (Table 1, groups 1−6), whereas in the class-2 it varies greatly (see examples in Table 1). Therefore, helix interactions in the class-1 are much more specific. As the oligomeric states of MPs are usually functionally active,[2] the packing of α-helices has to be quite specific. Finally, the average total energies in the groups of class-2 states (Table 1) are somewhat higher than those in the class-1. Taking all the above said into account, we assume that the complexes from class-1 are more

likely to represent the "nativelike" dimeric structures in membrane. However, we should note that the "V"-shaped dimers of TM helices (resembling our states of class-2) have been proposed for some other proteins,[35] and, moreover, the switch between "X"- and "V"-shaped conformations has been suggested as functionally important.[36] But at this stage we have no ideas about the significance and the putative role of such states found for GpA. As discussed below, we cannot also exclude the possibility that the "V"-shaped conformers may be stabilized in particular membrane-mimic media, which are different from those employed so far in the experiments.

Dimerization interfaces in the MC states show reasonable correlation with the mutagenesis data (Figure 6). The discrepancies are mainly observed near the C-terminus—for residues Gly86, Ile88, Leu90, and Ile91. We should note that such an overall comparison may be done only in a qualitative manner for two principal reasons. First, the mutagenesis results are somewhat ambiguous. Thus, as indicated in ref 11, the effects of GpA's residues substitutions with polar residues are not well understood. In addition, the final histogram of the relative degree of disruption of the dimer (Figure 5 in ref 11) was built using an arbitrary chosen scale. Furthermore, similar histogram for the calculated MC states represents just a summation over the ensemble (about 800) of low-energy structures of class-1, whereas the structures of the class-2 were omitted. (However, as seen in Table 1, the dimerization interfaces in the class-2 structures also contain residues that are important for helix association according to mutagenesis data.) Our data may be insufficient to get the real distribution of states, although the essential sampling of the GpA's conformational space was realized (see above). Nevertheless, reasonably good agreement between the experimental and computational results demonstrates eligibility of the theoretical approach. It is principally important that, despite certain conformational heterogeneity, in all of the calculated structures the helices contact each other with the same side, and it is exactly the side revealed in experiments.

Left-Handed α-Helical Dimer—a Putative Alternative to the NMR-Derived Model? The feasibility of dimers with $\Theta > 0°$ has been discussed previously in the mutagenesis work of Lemmon et al.[11] In that study the left-handed models were omitted upon analysis of periodicity in the distribution of residues critical for dimerization. This has been done using Fourier transform processing of the peptide's sequence. On the other hand, such approaches simplify the picture to a huge extent, not taking into account the exact conformations of side chains. In our opinion, this criterion cannot serve as a basis for unambiguous decision about the impossibility of the left-handed α-helical complexes. Interestingly, two distinct families of dimers: right- ($\Theta \sim -50°$) and left- ($\Theta \sim 40°$) handed ones have recently been obtained in MD simulations of GpA in implicit membrane starting from noninteracting TM helices.[24] Although in that study the authors have not considered the membrane insertion of individual helices (like in this work), the results agree well with our findings. This proves that the question concerning the left-handed dimers remains to be answered. In our opinion, the presence of that type of structures raises issue of the possibility of dimeric models alternative to those obtained in the result of NMR data interpretation.[14]

Conformational Heterogeneity of Low-Energy States: A Shortcoming of the Method or an Indicator of Media-Dependent Equilibrium Distribution of Dimers? Several problems related to future applications of the developed computational approach should also be mentioned. The most severe of them is related to the unambiguity of MC solutions: the choice of the "nativelike" structure is not straightforward without a priori knowledge of the helix—helix interface. Although most of the predicted conformations of GpA dimer have quite similar overall patterns of intermonomer contacts, their detailed structures may differ. This depends on a number of factors, like the membrane thickness and hydrophobicity degree, presence of TM voltage, and so forth. For instance, changing ASP values which describe the hydrophobic slab (see above) may lead to incorrect structures of the dimer and even to nondimeric structures. In many cases the objective is not so much the elaboration of an extremely precise spatial structure of a protein oligomer as the establishment of a general mode of its insertion into membrane, along with the delineation of some crucial residues on the dimerization interface. Such a rough model may be employed in the future to rationalize experimental observations and to design new experiments. On the other hand, we assume that the aforementioned problem has a fundamental character. Namely, the moderate structural heterogeneity of predicted GpA dimers (existence of several groups of states with close energies) may reflect a real equilibrium dynamics in membrane-mimic media used in experimental studies. We also suppose that some fraction of left-handed α-helical dimers may present as well. As discussed above, such a conformational heterogeneity may be caused by the fact that the packing of helices is quite sensitive to media effects and the geometry of membrane binding. Such a hypothesis is partially corroborated by the vague results of mutagenesis studies (see discussion in ref 11) as well as by NMR[16,17] and MD[18,24] data that demonstrate media effects on the stability of helical oligomers and provide examples of their multistate equilibrium in membrane mimics. In real biological membranes the situation may be far more complex due to the heterogeneity of their physicochemical characteristics. As a result, one or more predominant conformations of the dimer may occur, although this assumption requires further investigation.

Despite this vagueness, in overall, the modeling results show that under certain conditions, like the presence of $\Delta\psi$, an optimal membrane thickness and hydrophobicity degree, only a limited number of structures are energetically favorable. Some of them are close to those observed by NMR. We assume that the others may also be realized in membranes and artificial membrane mimics with particular properties. In our opinion, future MD simulations in explicit bilayers and micelles starting from these representative conformations will help to gain an additional insight into the equilibrium behavior of helical oligomers. This work is currently in progress in our group. The proposed computational approach therefore represents an indispensable step

toward the development of efficient theoretical methods to study helix interactions in membranes.

Driving Forces for Helix Association. One of the criteria used to assess quality of the resulting models of the dimer is the distribution of hydrophobic/hydrophilic patterns on the surfaces of monomers. Thus, it is known that the helix−helix and helix−lipid interfaces in TM domains of proteins reveal complementarity of such properties.[37,38] To check, whether this is the case with the proposed models of GpA dimers, we applied the MHP approach, which has been previously employed to study a number of membrane-bound α-helices.[32] TM α-helices in MPs usually expose to lipids their nonpolar surfaces, and this is especially true for complexes with a small number of helices, e.g.—for dimers (e.g., ref 39). (In contrast, the stability of MPs with ≥ 4 TM helices is determined to a large extent by interhelical contacts.[40]) Therefore, judging by the MHP maps (Figure 7), one may propose that α-helices of GpA interact via the lengthy hydrophilic stretches on their surfaces. In this case there is a strong complementarity between polarity regions of the monomers, while the most hydrophobic surfaces are exposed to lipids. In principle, according to the proposed criteria, two types of helix packing are possible—via residues forming either a left or a right "arm" of the "Λ-shaped" pattern. Interestingly, both variants may be realized in real and in computational experiments—for both right- (GpA_{NMR}, GpA_{MC}) and left- (GpA_L) handed dimers.

To summarize, the comparison of geometrical and hydrophobic parameters of the predicted α-helical complexes allows the following conclusions to be drawn: (1) Despite minor differences in packing, the majority of structures with $\Theta < 0°$ reveal overall good agreement with NMR and mutagenesis data. (2) In models with the alternative fold ($\Theta > 0°$) the dimerization interface only partially corresponds to that in the GpA_{NMR} model, although these conformers satisfy reasonably well the NMR-derived geometrical restraints. (3) The hydrophobic organization of all types of complexes is quite similar and does not contradict the known principles of packing of TM helices. What drives helix association in the membrane? Analysis of various energy terms on different stages of MC search shows that the initial significant drop of the total energy is caused by the independent insertion of monomers into the membrane. As a result, total E_{solv} decreases by ∼70 kcal/mol. Then the energetically favorable intermonomer van der Waals contacts appear. In all complexes (including GpA_{NMR}) residues on the dimerization interface have low values of E_{vdW}. For example, distribution of E_{vdW} along the sequence in GpA_{MC} and GpA_{NMR} models is quite similar (not shown).

Therefore, the formation of TM dimers is mainly driven by two factors. At the first stage (insertion)—by interaction with membrane, while at the second one (helix association with subsequent "fine-tuning" of the complex)—by van der Waals contacts. Importance of the media effects lies in the fact that the hydrophobic peptides insert one by one into the membrane and adopt a TM orientation. Also, apolar medium considerably promotes their α-helical conformation. TM voltage favors a head-to-head disposition of α-helices, providing a suitable starting point for obtaining the most adequate ("nativelike") final solutions via MC search. This is because the number of possible variants of monomers' association is seriously limited as compared, for example, with two conformationally labile peptides in water or in a vacuum. Hence, at the initial stage the membrane plays a role of a peculiar matrix, which makes the system's components well-prepared for subsequent association. Then the main energy gain is achieved upon creation of favorable intermonomer contacts, although solvation effects still remain important because the optimal docking of helices occurs in cases when residues on the external surface of the dimer have energetically favorable interactions with the medium (see above). These observations agree well with the famous two-stage model of folding of TM domains in proteins.[41]

Several shortcomings are inherent in our method. They are mainly related to the limitations of the implicit membrane model. Thus, a number of important characteristics of biological membranes, like heterogeneity of dielectric properties, chemical composition, and microscopic nature of protein−lipid interactions (e.g., H-bonds), etc. are completely or partially omitted. Also, the influence of the protein on the lipid bilayer is neglected. Finally, the question about treatment of electrostatic interactions in membrane requires a special consideration (see ref 42 for recent review). Thus, during the last several years a number of implicit membrane models based on Generalized Born (GB) theory have appeared.[24,43] The models dealing with solving the Poisson−Boltzmann equation[44] and implementing the Gouy−Chapman term describing counterion-screened electrostatic interactions of a protein with anionic membranes[45] were reported as well. Earlier we also tried various schemes to treat electrostatics but did not get significant improvement of the results (in terms of their consistency with experimental data). Comparison of our results with those obtained with one of the recent GB models[24] shows that both models lead to quite similar conclusions (see above). Interestingly, Im et al.[24] ignored one Glu- on the N-terminus and two Arg+ on the C-terminus of GpA. They mentioned that this is related to the problem of "stabilization of the helical interface". In addition, in their study the start was chosen as two TM helices placed perpendicular to the membrane plane, while we used random starting configurations. One disadvantage of our simplified electrostatic screening model becomes apparent in studies of binding of positively charged proteins to membranes composed of anionic lipids. Thus, application of the Gouy−Chapman theory allowed better description of interactions of cardiotoxins from snake venom with negatively charged membranes,[45] although in zwitterionic bilayers and micelles (like here) the results are similar.

On the other hand, as for any theoretical model based on empirical parametrization (including ours, naturally), the only criterion of validity is its accord with experiments. Our membrane model was proved to be adequate for a large number of peptides and proteins with different fold (α-helical, *β*-structural) and mode of binding (TM and peripheral). These items were reviewed in ref 7. Being computationally efficient, the proposed technique permits exploration of a number of alternative scenarios that are too costly to be tested experimentally (e.g., comparative analysis of binding

for wild-type and mutant proteins). Finally, the method may provide good starting points for subsequent simulations in full-atom lipid bilayers and micelles. This computational approach is currently being tested on other TM α-helical complexes. We are also studying the sensitivity of the simulation results to the thickness and hydrophobicity degree of the membrane, as well as to some other factors (Efremov et al., manuscript in preparation). The approach will be refined in the future with the appearance of new experimental structural information concerning the nature of protein−protein interactions in membranes.

Abbreviations used: GpA, hydrophobic segment 69−97 of the human glycophorin A; MP, membrane protein; TM, transmembrane; MD, molecular dynamics; MC, Monte Carlo; NMR, nuclear magnetic resonance; DPC, dodecylphosphocholine; MHP, molecular hydrophobicity potential; RMSD, root-mean-square deviation; GpA_{NMR}, NMR-derived model of the GpA dimer; $GpA_{NMR-mem}$, the lowest-energy conformer found for GpA_{NMR} in membrane; GpA_{MC}, the calculated model most close to the NMR structure of the dimer; GpA_L, left-handed model of the GpA dimer.

Acknowledgment. This work was supported by the Program RAS MCB, by the Russian Foundation for Basic Research (Grant 04-04-48875-a), by the Russian Ministry of Education and Science (Lead 05, Living systems: Integrated project 3/001). R.G.E. is grateful to the Russian Science Support Foundation for the grant awarded. Access to computational facilities of the Joint Supercomputer Center (Moscow) is gratefully acknowledged.

Supporting Information Available: Distance restraints used in MC simulations of the GpA_{NMR} model in implicit membrane and interatomic distances in models calculated via MC simulations in implicit membrane and in models obtained by NMR spectroscopy (Tables 1−3). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Wallin, E.; von Heijne, G. *Protein Sci.* **1998**, *7*, 1029−1038.

(2) Arkin, I. T. *Biochim. Biophys. Acta* **2002**, *1565*, 347−363.

(3) Harder, T. *Philos. Trans. R. Soc. London B. Biol. Sci.* **2003**, *358*, 863−868.

(4) Ubarretxena-Belandia, I.; Engelman, D. M. *Curr. Opin. Struct. Biol.* **2001**, *11*, 370−376.

(5) Forrest, L. R.; Sansom, M. S. *Curr. Opin. Struct. Biol.* **2000**, *10*, 174−181.

(6) Liang, J. *Curr. Opin. Chem. Biol.* **2002**, *6*, 878−884.

(7) Efremov, R. G.; Nolde, D. E.; Konshina, A. G.; Syrtcev, N. P.; Arseniev, A. S. *Curr. Med. Chem.* **2004**, *11*, 2421−2442.

(8) Eilers, M.; Patel, A. B.; Liu, W.; Smith, S. O. *Biophys. J.* **2002**, *82*, 2720−2736.

(9) DeGrado, W. F.; Gratkowski, H.; Lear, J. D. *Protein Sci.* **2003**, *12*, 647−665.

(10) Furthmayr, H.; Marchesi, V. T. *Biochemistry* **1976**, *15*, 1137−1144.

(11) Lemmon, M. A.; Flangan, J. M.; Treutlein, H. R.; Zhang, J.; Engelman, D. M. *Biochemistry* **1992**, *31*, 12719−12725.

(12) Treutlein, H. R.; Lemmon, M. A.; Engelman, D. M.; Brünger, A. T. *Biochemistry* **1992**, *31*, 12726−12733.

(13) Adams, P. D.; Engelman, D. M.; Brünger, A. T. *Proteins: Struct., Funct., Genet.* **1996**, *26*, 257−261.

(14) MacKenzie, K. R.; Prestegard, J. H.; Engelman, D. M. *Science* **1997**, *276*, 131−133.

(15) Smith, S. O.; Song, D.; Shekar, S.; Groesbeek, M.; Ziliox, M.; Aimoto, S. *Biochemistry* **2001**, *40*, 6553−6558.

(16) Orekhov, V. Y.; Abdulaeva, G. V.; Musina, L. Y.; Arseniev, A. S. *Eur. J. Biochem.* **1992**, *210*, 223−229.

(17) Gratkowski, H.; Dai, Q.; Wand, A. J.; DeGrado, W. F.; Lear, J. D. *Biophys. J.* **2002**, *83*, 1613−1619.

(18) Stockner, T.; Ash, W. L.; MacCallum, J. L.; Tieleman, D. P. *Biophys. J.* **2004**, *87*, 1650−1656.

(19) Fisher, L. E.; Engelman, D. M.; Sturgis, J. N. *Biophys J.* **2003**, *85*, 3097−3105.

(20) Pappu, R. V.; Marshall, G. R.; Ponder, J. W. *Nat. Struct. Biol.* **1999**, *6*, 50−55.

(21) Smith, S. O.; Smith, C.; Shekar, S.; Peersen, O.; Ziliox, M.; Aimoto, S. *Biochemistry* **2002**, *41*, 9321−9332.

(22) Dobbs, H.; Orlandini, E.; Bonaccini, R.; Seno, F. *Proteins: Struct., Funct., Genet.* **2002**, *49*, 342−349.

(23) Ducarme, P.; Thomas, A.; Brasseur, R. *Biochim. Biophys. Acta* **2000**, *1509*, 148−154.

(24) Im, W.; Feig, M.; Brooks, C. L., III *Biophys. J.* **2003**, *85*, 2900−2918.

(25) Efremov, R. G.; Volynsky, P. E.; Nolde, D. E.; Arseniev, A. S. *Theor. Chem. Acc.* **2001**, *106*, 48−54.

(26) von Freyberg, B.; Braun, W. *J. Comput.Chem.* **1991**, *12*, 1065−1076.

(27) Berman, H. M.; Bhat, T. N.; Bourne, P. E.; Feng, Z.; Gilliland, G.; Weissig, H.; Westbrook, J. *Nat. Struct. Biol.* **2000**, *7*, 957−959.

(28) Efremov, R. G.; Nolde, D. E.; Vergoten, G.; Arseniev, A. S. *Biophys. J.* **1999**, *76*, 2448−2459.

(29) Némethy, G.; Pottle, M. S.; Scheraga H. A. *J. Phys. Chem.* **1983**, *87*, 1883−1887.

(30) Roux, B. *Biophys. J.* **1997**, *73*, 2980−2989.

(31) Efremov, R. G.; Volynsky, P. E.; Nolde, D. E.; van Dalen, A.; de Kruijff, B.; Arseniev, A. S. *FEBS Lett.* **2002**, *526*, 97−100.

(32) Efremov, R. G.; Vergoten, G. *J. Phys. Chem.* **1995**, *99*, 10658−10666.

(33) Kabsch, W.; Sander, C. *Biopolymers* **1983**, *22*, 2577−2637.

(34) Bowie, J. U. *Nat. Struct. Biol.* **1997**, *4*, 915−917.

(35) Kairys, V.; Gilson, M. K.; Luy, B. *Eur. J. Biochem.* **2004**, *271*, 2086−2092.

(36) Fleishman, S. J.; Schlessinger, J.; Ben-Tal, N. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 15937−15940.

(37) Rees, D. C.; DeAntonio, L.; Eisenberg, D. *Science* **1989**, *245*, 510−513.

(38) Efremov, R. G.; Vergoten, G. *J. Prot. Chem.* **1996**, *15*, 63−76.

(39) Popot, J. L.; Engelman, D. M. *Annu. Rev. Biochem.* **2000**, *69*, 881−922.

(40) Efremov, R. G.; Vergoten, G.; Arseniev, A. S. *Theor. Chem. Acc.* **1999**, *101*, 73−76.

(41) Popot, J. L.; Engelman, D. M. *Biochemistry* **1990**, *29*, 4031−4037.

(42) Tobias, D. J. *Curr. Opin. Struct. Biol.* **2001**, *11*, 253−261.

(43) Spassov, V. Z.; Yan, L.; Szalma, S. *J. Phys. Chem. B* **2002**, *106*, 8726−8738.

(44) Luo, R.; David, L.; Gilson, M. K. *J. Comput. Chem.* **2002**, *23*, 1244−1253.

(45) Lazaridis, T. *Proteins: Struct., Funct., Bioinf.* **2005**, *58*, 518−527.

CT0501250