

Association of transmembrane helices: what determines assembling of a dimer?

Roman G. Efremov^{a,*}, Yana A. Vereshaga^{a,b}, Pavel E. Volynsky^a, Dmitry E. Nolde^a & Alexander S. Arseniev^a

^a*M.M. Shemyakin and Yu.A. Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Ul. Miklukho-Maklaya, 16/10, V-437, 117997 GSP, Moscow, Russia;* ^b*Moscow Institute of Physics and Technology, State University, Institutskii per. 9, Dolgoprudny, 141700, Moscow Region, Russia*

Received 15 August 2005; accepted 17 January 2006
© Springer 2006

Key words: glycophorin A dimer, implicit membrane, mutagenesis *in silico*, transmembrane, electrostatic potential, two-stage model of membrane protein folding

Summary

Self-association of two hydrophobic α -helices is studied via unrestrained Monte Carlo (MC) simulations in a hydrophobic slab described by an effective potential. The system under study represents two transmembrane (TM) segments of human glycophorin A (GpA), which form homo-dimers in membranes. The influence of TM electrostatic potential, thickness and hydrophobicity degree of lipid bilayer is investigated. It is shown that the membrane environment stabilizes α -helical conformation of GpA monomers, induces their TM insertion and facilitates inter-helical contacts. Head-to-head orientation of the helices is promoted by the voltage difference across the membrane. Subsequent “fine-tuned” assembling of the dimer is mediated by van der Waals interactions. Only the models of dimer, calculated in a hydrophobic slab with applied voltage agree with experimental data, while simulations *in vacuo* or without TM voltage fail to give reasonable results. The moderate structural heterogeneity of GpA dimers (existence of several groups of states with close energies) is proposed to reflect their equilibrium dynamics in membrane-mimics. The calculations performed for GpA mutants G83A and G86L permit rationalization of mutagenesis data for them. The results of Monte Carlo simulations critically depend on the parameters of the membrane model: adequate description of helix association is achieved in the water–cyclohexane–water system with the membrane thickness 30–34 Å, while in membranes with different hydrophobicities and thickness unrealistic conformations of the dimer are found. The computational approach permits efficient prediction of TM helical oligomers based solely on the sequences of interacting peptides.

Abbreviations: GpA – hydrophobic segment 69–97 of human glycophorin A; GpA_{NMR} – NMR-derived model of the GpA dimer; GpA_{NMR-mem} – the lowest-energy states of GpA_{NMR} in implicit membrane; HP – helical “hairpin”; MC – Monte Carlo; MP – membrane protein; TM – transmembrane

Introduction

Hydrophobic and amphiphilic α -helices represent a dominant structural motif responsible for binding of membrane proteins (MPs) and membrane active peptides to membranes. Functional activity

*To whom correspondence should be addressed. Phone: +7-095-336-20-00; Fax: +7-095-336-20-00; E-mail: efremov@nmr.ru

of these molecules often requires helix association in lipid bilayers. Among them are integral and peripheral membrane receptors (e.g., proteins from the GPCR family, receptors of hormones, sensory systems), ion pumps and channels, receptor tyrosine kinases, defensins, and many others [1]. Also, helix–helix interactions are important for biological function of a wide variety of antimicrobial and channel forming peptides [2, 3]. That is why understanding of the factors that drive packing of α -helices in membranes has attracted considerable interest of researchers since a long time. In a case of globular proteins solution to the problem is facilitated by a large number of high-resolution protein structures available for the analysis. Based on this data, a number of important rules that govern helix–helix interactions have been delineated [4–6]. By contrast, for MPs the question still remains to be answered. What might be the role of heterogeneous lipid bilayer media in association of membrane-bound helices? The main impediment to the solution of the problem is a small number of atomic-resolution 3D structures of membrane α -helical complexes. This is because of limitations of modern experimental techniques and enormous difficulties related to isolation, purification, and handling of MPs in their “native-like” conformations [7]. Modern *in silico* methods represent a promising alternative, which can considerably extend and complement the possibilities of traditional biophysical approaches. Thus, interactions of individual α -helices with membranes and membrane-like environments have been extensively studied using both experimental and modeling techniques (reviewed in [8–10]). On the other hand, the principal obstacle in theoretical modeling of protein–protein interactions is a necessity to take into account media effects.

Homo-dimeric MPs containing a single TM helix represent attractive objects for the development of computational techniques to assess helix interactions in membranes. This is because of relative simplicity and stability of such systems. Several of them have been extensively studied so far. These include the glycoporphin A (GpA) [11, 12], the bacteriophage M13 major coat protein [13, 14], and the ErbB receptors [15, 16]. A number of motifs have been reported that promote packing of TM helices, including the GxxxG-like motifs [17], multiple Ser/Thr motifs (SxxxSSxxT, SxxSSxxT)

[18], and heptad repeat motif [19]. Detailed inspection of these motifs has shown that the helix packing may be determined either by one or several characteristics like complementarity of geometrical and/or hydrophobic properties [20], H-bonding, electrostatic interactions [17, 21]. At the same time, the question how “fine tuning” of TM helices proceeds in heterogeneous membrane environment is yet to be answered. This problem is only beginning to be explored systematically [22]. In this study we are investigating the association of two hydrophobic segments 69–97 of GpA. In membranes and detergent micelles GpA forms stable α -helical homo-dimers with parallel (head-to-head) orientation. Site-directed mutagenesis experiments [12] have revealed the following sequence pattern important for dimerization: LIX-xGVxxGVxxT, where x marks an arbitrary residue. Based on this data and the results of computer simulations, a spatial model of the dimer has been elaborated [23, 24]. It represents a right-handed helical supercoil with the angle $\sim 40^\circ$ between the monomers. Later, the model has been refined using several interhelical distance restraints obtained by NMR in detergent micelles [25] and lipid bilayers [26]. Even in these artificial environments the spatial structures of the dimer are somewhat different. Furthermore, it has been shown that association of TM helices may strongly depend on the membrane mimics used in the experimental setup [27, 28]. Real biological membranes are much more complex in terms of physico-chemical properties (heterogeneity of lipid composition, presence of TM voltage, etc.) and therefore may seriously affect helix interactions.

Nevertheless, the appearance of the 3D model of GpA dimer in membrane mimics (hereafter it is called “the native-like structure”) have stimulated the development of theoretical methods to model helix interactions in membrane domains of proteins. This gives a possibility of direct testing the simulation results against experimental data. In most of computational studies the simulation protocol resembles that proposed in [23, 24]. Thus, Ducarme et al. [29] considered GpA helix association in implicit membrane. The starting structures were the NMR-derived model of the dimer, and the α -helical monomers dissociated inside the membrane. It was found that the simulations with and without membrane may reach the “native-like” dimer structure, although the presence of the

membrane shortens this process and decreases the probability for unrealistic complexes. Association of GpA helices was also studied by Im et al. [30] with the help of implicit membrane. In that model electrostatic interactions were treated based on Generalized Born theory. MD simulations starting from non-interacting TM helices revealed two distinct families of dimers (right- and left-handed ones). Interestingly, the vast majority of the low-energy states corresponded to left-handed TM α -helical complexes. Similar approaches were later applied to elaborate 3D models of a number of α -helical dimers [31–33]. Although these studies provided considerable insight into the mechanisms of helix interactions in heterogeneous membrane-like medium, most of them possess a number of severe limitations. First, the simulations were carried out either *in vacuo*, or in a continuum dielectric with low permeability. Also, the α -helical monomers were always “rigid” and, hence, common occurrence of local distortions in TM helices, like kinks and bends, was not taken into account. Finally, it was proposed *a priori* that helices adopt a TM orientation. Therefore, the effects of membrane media on the secondary structure formation and/or stabilization, along with the events accompanying insertion of the peptides, cannot be assessed via such calculations. Naturally, this significantly reduces the number of degrees of freedom and facilitates prediction of the structure for helical oligomers. To have a predictive power, a method destined to studies of helical associates in lipid bilayer should not impose any restraints on the secondary structure and should not be based on *a priori* knowledge of the mode of membrane binding for the peptides. In this case it is believed to be useful for studies of new biologically relevant MPs. Alternative way to predict TM helix packing lies in employment of statistical analysis of α -helical oligomers in MPs with known spatial structure [34]. These techniques were shown to perform well for a number of helical pairs. At the same time, like for any statistically-based approach, predictive power of these algorithms critically depends on the training set of helical segments used to derive the potentials.

Apart from difficulties of correct treatment of heterogeneous membrane environment, successful employment of these methods depends on availability of physically reliable and computationally efficient theoretical model describing influence of

TM electrostatic potentials ($\Delta\psi$) on protein's behavior in membrane. This is because real biological membranes do not possess a symmetry of their physico-chemical properties. In particular, most of lipid bilayer membranes reveal $\Delta\psi$ [35]. Occurrence of TM voltage is important for many processes in cell – e.g., for phenomena related to conservation and transformation of energy, for regulation of ion and molecular transport, for proteins' insertion into lipid bilayer, and so forth [35]. Appearance of $\Delta\psi$ in membranes of cells and organelles is caused by a nonuniform distribution of ions on different sides of lipid bilayer. The value of $\Delta\psi$ varies depending on the cell type, media conditions, presence of proteins and low-molecular weight compounds in membrane. In particular, in bilayer vesicles $\Delta\psi$ may reach $\sim 10^2$ mV – this corresponds to the strength of the electric field $\sim 10^6$ V/m. Obviously, in such cases TM voltage may play an important role in protein–membrane interactions and in assembling of MPs. It should be forthcoming that the effects of $\Delta\psi$ are especially meaningful for α -helical peptides which possess large dipole moments. For example, studies of a number of antibacterial peptides have revealed that TM voltage may change their mode of membrane binding – from peripheral to membrane-spanning [2, 3]. Changes of $\Delta\psi$ occurring, e.g., upon generation and transmission of nerve impulses, regulate functioning of voltage-dependent ion channels by switching them between open and closed states [36]. In addition, the voltage difference between the sides of outer membrane in *E. coli* acts as a trigger which initiates translocation of preproteins across bacterial membranes [37]. Molecular mechanisms of action of $\Delta\psi$ are still poorly understood. This is explained by serious difficulties in experimental studies of the influence of fast processes related to generation and action of weak potentials on structure and dynamics of large multi-component systems, like hydrated lipid–protein complexes. Additional insight into the problem may be gained via employment of molecular modeling techniques – computer simulations of such systems are rapidly developing during the last years [8, 10].

Several models taking into account $\Delta\psi$ have been reported in the literature. Thus, in a case of infinitive planar dielectric slab expression for $\Delta\psi$ can be obtained as a solution of the Poisson–Boltzmann equation with boundary conditions

[38, 39]. In this model the potential inside the hydrophobic layer linearly changes along the normal to the membrane (z -coordinate). Similar dependence of $\Delta\psi$ on z was used in studies of voltage-dependent interactions of a number of α -helical peptides with implicit membrane – the pore-lining segments $\alpha 5$ and $\alpha 7$ of endotoxin [40], antibacterial peptide cecropin P1 (CecP) [41], alamethicin [42], and others [8, 43]. In our previous works [44, 45] analogous model was applied to study the influence of TM voltage on spatial structure and mode of membrane binding of several peptides and proteins. For example, it was shown that at $\Delta\psi=0$ the monomer of GpA has two distinct energetically favorable states – TM α -helix and peripheral α -helical hairpin (HP), while TM voltage considerably promotes the first of them [44]. Simulations of the influence of $\Delta\psi$ on behavior of the *E. coli* prePhoE signal sequence – a 20-residue peptide extended with 7 amino acids of the mature protein – were described in [45]. It was demonstrated that $\Delta\psi$ stimulates insertion of the wild-type peptide by switching it between HP and TM states. The modeling results were found to be in excellent agreement with available experimental data. They helped to explain voltage-dependent interaction of the signal peptide with leader peptidase, as well as the role of helix-breaking residues (e.g., glycines) which often occur in the central part of signal sequences. The presence of such residues facilitates $\Delta\psi$ -induced translocation of the mature protein across the membrane.

Besides $\Delta\psi$, among the factors that may affect oligomerization of TM α -helices there are the thickness and the hydrophobicity degree of the membrane. It is known [46, 47] that the length of hydrophobic membrane-spanning regions in proteins fits well to the thickness of nonpolar layer of the membrane. This is the so-called effect of “hydrophobic match/mismatch”. Obviously, in a case of several interacting helices such phenomena may affect the geometry of the complex. Furthermore, it has been shown that the packing of helices may be sensitive to media effects [27, 28, 48]. Because the experimental structures of helical oligomers have been obtained not in real membranes, but in artificial membrane mimics (detergent micelles, lipid vesicles), this raises a question about the conformations of these systems under native conditions. It is therefore interesting to

explore different conformational possibilities for helix complexes: possibly, depending on the media properties, some of which may be realized in a cell. Finally, a number of recent studies show that the problem of conformational exchange in TM helical oligomers is quite serious [24, 26–28]. Therefore, the problem still exists and we are still far from being able to claim that we have reached a solution.

A conclusion can be made that delineation of the spatial structure of even the simplest membrane helical oligomers is not straightforward. Hence, development of new efficient algorithms to predict the way of helix association in lipid bilayers is believed to be very promising. The intention of the present study is to develop a computational approach to simulate association of α -helical fragments of proteins in a membrane-mimic environment. No restraints on the spatial structure of the monomers, their mutual disposition, and/or on their mode of membrane binding are employed. The approach is based on the Monte Carlo (MC) search of low-energy conformational states in implicit membrane [10]. Furthermore, the role of $\Delta\psi$, along with membrane thickness and hydrophobicity degree in association of TM helices of GpA is explored. The designed computational protocol permits rather more reliable description of the dimer’s structure, as compared to modeling *in vacuo* or in a symmetrical hydrophobic slab (at $\Delta\psi=0$). Principal objective of such calculations is to check, whether they are capable to predict the dimer structures in accord with the experiments. On the other hand, because the experimental data have been obtained not in real membranes, but in artificial membrane-mimics, it was interesting to explore other conformational possibilities for helix complexes – probably, some of them may be realized under native conditions.

Method of calculations

The system

The calculations were performed for one and/or two GpA segments with a sequence S⁶⁹EPEITLII-FGVMAGVIGTILLISYGIRR⁹⁷. All-atom starting structures of GpA were built in α -helical conformation – this is because previously we have

shown [44] that in implicit membrane the GpA monomer folds in TM α -helix from initial random coil conformation. To change during the simulation orientation of GpA with respect to the membrane, the fragment of 10 dummy residues was attached to its N-terminus. These “virtual” residues do not contribute to the energy of the system. First atom of the N-terminal dummy residue was always placed in the center of the hydrophobic layer with coordinates (0,0,0). In the case of two GpA monomers, 20 additional dummy residues were introduced between the monomers (Figure 1a). This allowed changing of orientation of the monomers with respect to each other and to the membrane. Inclusion of the “dummy” regions was dictated by the necessity of continuity of the protein backbone for simulations in dihedral angles space. The peptides termini and side chains were fully ionized. In simulations of the monomer its initial structure was arbitrary placed outside the hydrophobic layer of the membrane. As shown in Figure 1, simulations of two GpA segments were started from three independent states: both non-interacting α -helices 1 and 2 were arbitrary placed outside (start a) or inside (start b) the membrane; one of the helices was partly immersed in the hydrophobic slab, while the other stayed in polar phase (start c).

To compare energies of the states obtained via MC search with those of the NMR-derived model [25] of the GpA dimer (thereafter it is denoted as GpA_{NMR}), the last one was also simulated in the implicit membrane. Atomic coordinates of GpA_{NMR} (residues 69–97) were taken in the Brookhaven Protein Data Bank (PDB [49],

www.rcbs.org, entry 1afo). To perform MC-simulations for this model, dummy residues were introduced by a similar way as shown in Figure 1a. The starting structure was arbitrary placed in the aqueous phase (Figure 1d).

Simulation protocol

The peptides’ conformational space was explored via MC search in torsion angles space as described elsewhere [50, 51]. The membrane was represented by a “hydrophobic slab” approximated 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 [44]. All-atom potential energy function for the protein was taken in the form: $E_{\text{total}} = E_{\text{ECEPP}/2} + E_{\text{solv}}$. The term $E_{\text{ECEPP}/2}$ includes van der Waals, torsion, electrostatic, and hydrogen bonding contributions [52]. E_{solv} is the solvation energy:

$$E_{\text{solv.}} = \sum_{i=1}^N \Delta\sigma_i \text{ASA}_i \quad (1)$$

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 [50]. Interaction of the protein with both aqueous and membrane environments is given by Equation 2, where $\Delta\sigma_i$ depends on the z coordinate of atom i (axis z is normal to the membrane plane):

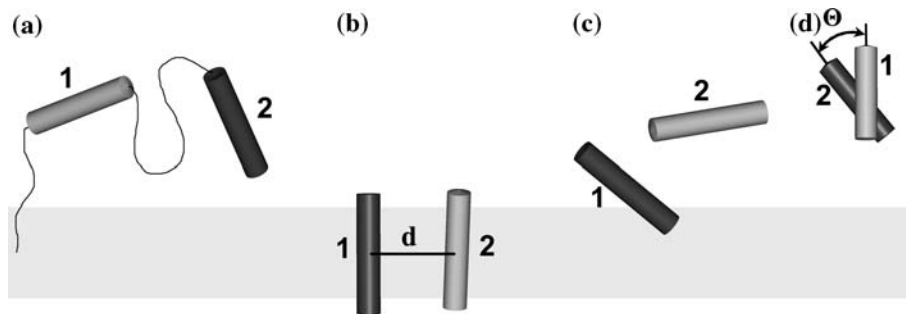


Figure 1. Starting configurations of two α -helices of GpA used in Monte Carlo simulations. (a–c) The arbitrarily chosen starts with non-interacting helices. (d) Initial disposition of the NMR-derived model of the GpA dimer. The helical monomers are displayed with cylinders marked “1” and “2”. The nonpolar layer of membrane is shown in gray. (a) Thin lines indicate regions with dummy residues. (b, d) Definition of the geometrical parameters for the α -helical dimer: d and Θ are respectively the distance and the angle between the helical axes.

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

where $\Delta\sigma_i^{\text{mem}}$ and $\Delta\sigma_i^{\text{wat}}$ 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, λ is a characteristic half-width of the water-membrane interface (in this study $\lambda=1.5$ Å). The simulations were carried out with $D = 24, 30, 34, 36$, and 40 Å.

All dihedral angles were sampled, except angles ω in “real” residues. The step of variation of each dihedral was chosen randomly on the range -180 to 180° . Nonbond interactions were truncated with a spherical cutoff of 30 Å. As discussed earlier [50], 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. Acceptance of MC-states was done according to the Metropolis criterion [53]. To cross the energy barriers between local minima, adaptive-temperature schedule protocol [54] was employed. Then several consecutive MC runs (5×10^3 steps each) with different seed numbers and sampled 5, 3, 2, and 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 inter-monomer distance restraints derived from the NMR-structure, was applied. The later MC-stages were performed without restraints. In total, $\sim 4 \times 10^4$ MC steps were done for GpA_{NMR}. To address the convergence problem in MC search, several independent simulations with different starting orientations of α -helices (Figure 1) were carried out. It is important to note, that the sets of low-energy states obtained in all the simulations are quite similar in total energy and its individual terms, as well as in structure and mode of membrane binding. This provides strong grounds to believe that the essential sampling of the GpA’s

conformational space was reached in each MC simulation.

The effect of TM potential ($\Delta\psi$) was taken into account by introducing into the potential energy function a special term ($E_{\Delta\psi}$) [39, 43]:

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

where q_i and z_i are partial charge and coordinate z of atom i , and F is Faraday’s constant. For $|z_i| > z_0$, $\psi(z) = \text{const}$. All calculations were carried out with $\Delta\psi = 300$ mV. Earlier we have shown [45] 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.

GpA mutants G83A (model GpA_{G83A}) and G86L (model GpA_{G86L}) were built by replacement of corresponding residues in the starting α -helical conformations of both monomers. This was done using the FANTOM program [54]. MC simulations of both mutant systems were performed with $D = 30$ Å starting from arbitrarily chosen orientations of helices with respect to one another and to the hydrophobic slab. Details of the simulations were similar to those used for the wild-type GpA (see above). The calculation and visualization of hydrophobic/hydrophilic properties of GpA and two its mutants were carried out using the molecular hydrophobicity potential (MHP) approach, as described elsewhere [55]. The MHP values were calculated on the solvent-accessible surfaces of α -helical segments Ser69-Arg98 in the resulting lowest-energy conformers of the models GpA_{NMR-mem.}, GpA_{G83A}, and GpA_{G86L}. GpA_{NMR-mem.} is the lowest-energy state found for the NMR-derived model of the dimer (GpA_{NMR}) in implicit membrane.

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 ASPs obtained on the basis of the gas-octanol free

energies of transfer [50]. In the latter case, this region was approximated with ASPs corresponding to a hypothetical solvent that is “more hydrophobic” than cyclohexane. (Hereafter this model is called “water–cyclohexane’–water”.) To do this, the gas–cyclohexane’ values of ASPs were calculated according to the formula: $\Delta\sigma_i^{\text{mem}'} = a\Delta\sigma_i^{\text{mem}} - (a-1)\Delta\sigma_i^{\text{wat}}$, where $\Delta\sigma_i^{\text{mem}}$ and $\Delta\sigma_i^{\text{wat}}$ are the original gas–cyclohexane and gas–water ASP values for the type i atom (see Equation 2), and $a = 1.4$. The modified membrane models were employed in MC simulations analogous to those described above. The simulations were performed with $D = 30 \text{ \AA}$ and $\Delta\psi = 300 \text{ mV}$.

Analysis of the results

Resulting states were analyzed using a set of auxiliary programs which were specially written for this. Only the low-energy states (with energy in the range $[E_{\text{min.}}, E_{\text{min.}} + \Delta E]$, where $E_{\text{min.}}$ is the minimal energy, $\Delta E = 15 \text{ kcal/mol}$) were used for further inquiry. Mutual disposition of the monomers with respect to each other and to the membrane was described in terms of the angle (Θ) and the distance (d) between the helix axes (Figure 1). If axial projection of a point of the closest approach between the axes falls outside the helical segments, the value d was defined as a distance between the axial projections of terminal C_α atoms of the monomers. The ASA values for residues and their secondary structure were

assessed using the DSSP program [56]. Dimerization interface was determined as follows. Residue i was considered to lie on the interface if the difference between its ASA values in dimer and in monomer exceeds 25 \AA^2 (10 \AA^2 for glycines). The same criteria were used to delineate intramonomer contacts between residues. Clustering of low-energy states into groups was done based on the parameters of helix packing (Θ , d) and the dimerization interface.

Results and discussion

Lessons from simulations in vacuo and in a symmetrical hydrophobic slab

Such simulations were carried out to assess the role of membrane effects in association of TM helices of GpA. Analysis of the resulting low-energy states shows that the neglect of the membrane-mimic media leads to inappropriate description of the system. Thus, initial entirely α -helical structure of the individual monomer is strongly distorted in the middle part due to energetically favorable van der Waals and electrostatic interactions between charged residues on the termini (Figure 2a). Hereinafter such conformers are called “hairpins” (HP). Although the presence of the second monomer *in vacuo* does stabilize to some extent the helices (Figure 2b), the packing interface and the overall structure of the dimer drastically differ from those observed in

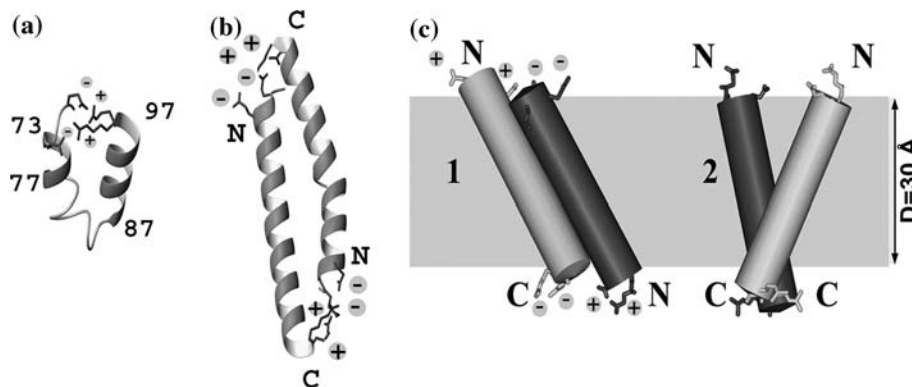


Figure 2. Low-energy states obtained via Monte Carlo simulations of one (a) and two (b) GpA α -helices *in vacuo*. In the later case initial configuration of the system corresponds to non-interacting helices. (c) The lowest-energy states obtained in a symmetrical implicit membrane ($\Delta\psi = 0 \text{ mV}$). Only two groups of states (1 and 2) are shown. The peptide's termini are marked with symbols “N” and “C”. Boundaries of α -helical segments in the hairpin-like structures (a) are indicated with residue numbers. Positively (R96, R97) and negatively (E70, E72) charged residues of GpA are shown with symbols “+” and “-”, respectively. Other details are the same as in the legend to Figure 1.

the experiments [12, 25] (data not shown). How does this picture change in a membrane-mimic environment? Analysis of the low-energy MC-states obtained in implicit membrane permits the following conclusions. α -Helical conformation of the monomers is well preserved. They insert in TM orientation and form stable dimers. The monomers interact either via their termini or via the residues which lie apart from the dimerization interface known from mutagenesis studies. In contrast to experimentally determined parallel (head-to-head) orientation, a large number of antiparallel ($\text{TM}_{\uparrow\downarrow}$) states are found, although some amount of parallel ($\text{TM}_{\uparrow\uparrow}$) orientations present as well (Figure 2c). One exception is provided by the group of $\text{TM}_{\uparrow\downarrow}$ states with $d \approx 9.5 \text{ \AA}$ and $|\Theta| \approx 160^\circ$ (Figure 2c, group 1) observed in a wide range of D (24–36 \AA). Dimerization interface in such structures is close to that in the NMR-derived models: LIXxGVxxGxxxT for monomer 1 and TxxxGxxVGxxI for monomer 2. So, although modeling in a hydrophobic slab performs much better than simulations *in vacuo*, it is difficult to extract the correct (“native-like”) dimer conformations from the ensemble of low-energy MC-states.

How to explain such a large population of misfolded states? We suppose that this is due to symmetrical nature of the membrane employed in simulations: there is no preferential direction for the dipole moments of α -helices, both types of the dimers are energetically favorable. In contrast, real biological membranes manifest TM potentials ($\Delta\psi$), present across lipid bilayers. The role of $\Delta\psi$ seems to be especially important for TM helices of GpA because they carry charges of opposite sign on their N- and C-termini and possess large dipole moments (~ 90 – 100 Debye). It can be assumed that the $\text{TM}_{\uparrow\downarrow}$ -states are dominant only in the absence of $\Delta\psi$. In addition, symmet-

rical nature of the membrane model makes possible occurrence of the HP-like states.

Transmembrane electrostatic potential facilitates correct assembling of α -helices

The MC-search from three independent starting positions of GpA helices in relation to one another and to the membrane ($D = 30 \text{ \AA}$) revealed that the low-energy states are characterized as follows. Like in a symmetric ($\Delta\psi = 0 \text{ mV}$) membrane, the monomers retain well their initial α -helical conformation (Figure 3). Each α -helix adopts a TM-orientation, in which almost all hydrophobic amino acid residues are immersed into the non-polar layer of the membrane, whereas charged residues on the termini are exposed to water. Moreover, the monomers constitute tightly packed complexes inside the membrane. At the same time, a number of significant differences in the system behavior can be observed comparing with the case when $\Delta\psi = 0$. For instance, the dipole moments of α -helices are oriented in the direction of the TM electric field. As a result, dimers with parallel mutual disposition of the peptides are formed ($\text{TM}_{\uparrow\uparrow}$, Figure 3). We shall notice that when $\Delta\psi = 0$, the population of such states is significantly lower than that of $\text{TM}_{\uparrow\downarrow}$ ones. Moreover, the applied potential affects the packing of TM α -helices. Thus, at $\Delta\psi = 0$ a high level of conformational heterogeneity could be observed in the position of helices with respect to one another. On the contrary, analysis of the results at $\Delta\psi \neq 0$ (Table 1) shows that they agree much better with the NMR data [25]. For instance, head-to-head packing of monomers typical of the GpA in membrane-mimic media is well reproduced in the calculations. Besides that, less degree of conformational heterogeneity of the resulting states leads to a restricted number of

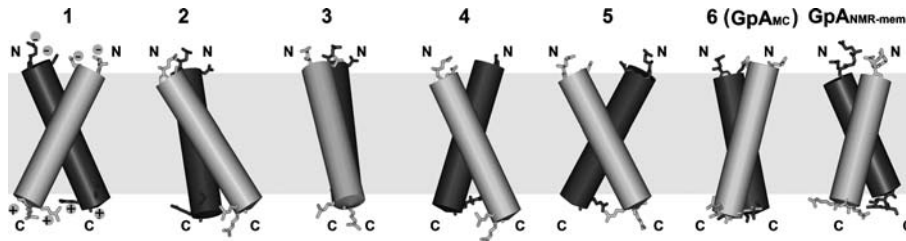
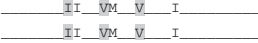
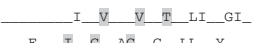

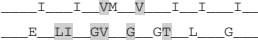

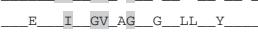
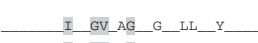
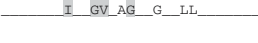
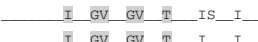
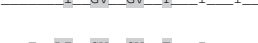
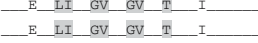

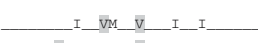
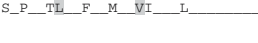
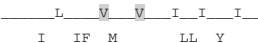
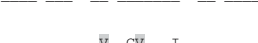
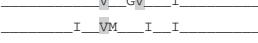

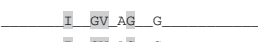
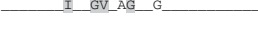
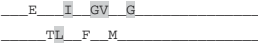

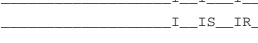

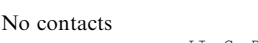
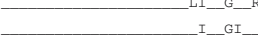
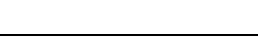


Figure 3. Low-energy states found via MC simulations of two GpA helices in implicit membrane with applied TM potential. (1–6) Representative conformers from the groups 1–6. Numbering of the groups is that as in Table 1. $\text{GpA}_{\text{NMR-mem}}$ is the lowest-energy state of the NMR-derived model of GpA. The membrane thickness is 30 \AA . For other details see the legend to Figure 2.

Table 1. Helix packing and energetic parameters for the groups of low-energy states of the GpA dimer found via Monte Carlo simulations in implicit membrane and in the modified membrane-mimic slab model (see “Methods”) with applied transmembrane voltage.

No group ^a	d^b	Θ^b	Dimerization interface ^c	E_{total}^d	$E_{\text{elect.}}$	E_{vdW}	E_{solv}
Water–cyclohexane–water membrane							
1	10.1 (0.1)	−50.1 (0.5)	SEPEITLIIIFGVMAVGIGTILLISYGIRR 	−629.8 (0.5)	18.2 (0.2)	−370.4 (0.3)	−191.8 (0.4)
2	9.2 (0.1)	22.1 (1.1)	 	−632.4 (1.7)	18.9 (0.2)	−382.3 (1.5)	−185.9 (0.6)
3	9.0 (0.1)	5.9 (0.2)	 	−626.9 (2.7)	18.8 (0.3)	−379.3 (1.9)	−188.0 (1.3)
4	6.9 (0.1)	42.6 (0.8)	 	−629.3 (1.6)	21.6 (0.4)	−392.3 (1.4)	−174.8 (0.7)
5	5.5 (0.2)	60.2 (2.0)	 	−626.5 (2.9)	19.3 (1.3)	−378.6 (3.0)	−190.2 (1.2)
6 (GpA _{MC}) ^e	8.5 (0.0)	−27.8 (0.3)	 	−622.7 (0.4)	17.2 (0.2)	−376.5 (0.2)	−186.0 (0.3)
GpA _{NMR-mem.} ^f	7.1 (0.1)	−39.8 (0.3)	 	−622.3 (1.0)	20.6 (0.2)	−382.5 (3.7)	−187.0 (2.8)
Water–octanol–water membrane ^g							
1'	10.2 (0.1)	8.2 (0.9)	 	−587.3 (3.3)	16.4 (0.4)	−375.4 (2.3)	−150.8 (1.0)
2'	Hairpin–Hairpin		 	−581.2 (0.6)	21.6 (0.1)	−377.2 (0.6)	−154.8 (0.8)
3'	9.1 (0.1)	−53.2 (0.3)	 	−581.7 (0.9)	17.3 (0.1)	−370.1 (0.5)	−153.9 (0.4)
Water–cyclohexane–water membrane							
1'	6.5 (0.3)	65.2 (4.2)	 	−619.3 (2.5)	18. (1.2)	−367.1 (4.0)	−214.6 (3.5)
2'	Helix–Hairpin or Hairpin–Hairpin		 	−616.8 (1.8)	21.7 (0.9)	−363.2 (2.3)	−224.6 (1.6)
3'	8.9 (0.8)	−43.8 (25.4)	 	−622.9 (3.9)	18.2 (0.5)	−356.7 (5.2)	−224.2 (4.3)
4'	31.1 (4.9)	−7.4 (22.3)	No contacts	−617.8 (1.3)	16.6 (1.4)	−344.8 (2.5)	−228.2 (2.7)
5'	Hairpin–Hairpin		No contacts	−615.5 (0.4)	18.9 (0.6)	−344.6 (1.2)	−234.5 (0.7)
6'	Hairpin–Hairpin		 	−617.2 (1.6)	20.4 (0.5)	−354.1 (0.9)	−229.5 (1.3)

^aThe groups' numbers are the same as in Figure 3.

^b d and Θ are the distance and the angle between helix axes, respectively.

^cResidues 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 [12].

^d E_{total} , $E_{\text{elect.}}$, E_{vdW} , $E_{\text{solv.}}$ are the total, electrostatic, van der Waals, and solvation energies, respectively (in kcal/mol).

^eGpA_{MC} is the calculated model most close to the NMR-structure of the dimer.

^fGpA_{NMR-mem.} is the lowest-energy state found for the NMR-derived model of the dimer [25] in implicit membrane.

^gResults of MC simulations with the modified membrane-mimic slab model (see “Methods”).

solutions. To define how adequate the dimeric MC-structures are, it is necessary to conduct a detailed analysis of the mutual positions of the helices. Furthermore, it is important to understand what interactions are the most significant for dimerization in the membrane.

Inspection of the total energies and their components for the entire complexes, individual monomers and residues, along with analysis of the dimerization interfaces, permits the following conclusions. In the resulting states the inter-monomer interface comprises 6–10 residues and the contact area is $370 \pm 50 \text{ \AA}^2$. All these conformers have low values of E_{vdW} and $E_{\text{solv.}}$ terms (Table 1) along with low energies ($-60 \pm 15 \text{ kcal/mol}$) of intermolecular interactions (the sum of E_{vdW} and $E_{\text{elec.}}$). Furthermore, these states reveal quite a similar interface (Table 1, groups 1–6), and therefore the helix interactions are specific. As the oligomeric states of MPs are usually functionally active [57], the packing of α -helices in the complexes has to be quite specific.

In the result, a nonredundant set of ~ 800 structures was delineated. The conformers demonstrate somewhat diverse types of helix packing (Table 1). Accordingly, the composition of residues forming 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 (Table 1) and therefore only a limited number of possible packing geometries should be used for future analysis. No outliers of the clusters were found. 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 by Lemmon et al. [12] (Figure 4). Discrepancies are mainly observed for residues near the C-terminus: G86, I88, L90, I91. Possible reasons for this are discussed later. We should note that such an overall comparison may be done only in a qualitative manner because of two principal reasons. First, the mutagenesis results are somewhat ambiguous. Thus, as indicated by Lemmon

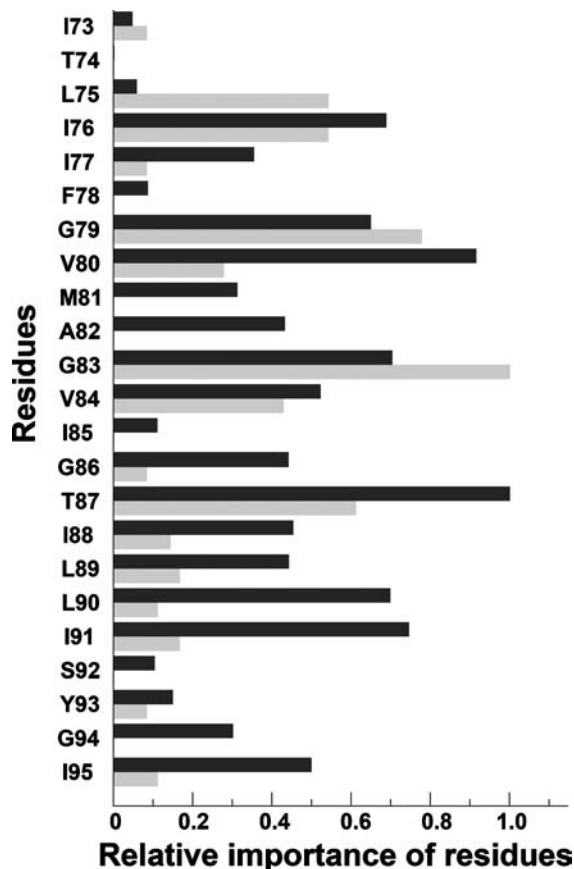


Figure 4. Involvement of GpA residues in helix–helix contacts. The residues’ names and numbers are indicated along the y-axis. The value on the x-axis corresponds to the relative importance of residues for dimerization. Black bars: mutagenesis data, taken from [12]; Gray bars: Frequency of occurrence 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). Both distributions are normalized on the range [0, 1].

et al. [12], 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 [12]) was built using an arbitrary chosen scale. Furthermore, similar histogram for the calculated MC-states represents just a summation over the ensemble of low-energy structures. Our data may be insufficient to get the real distribution of states, although inspection of the convergence problem shows that the essential sampling of the GpA’s conformational space was reached in several independent MC simulations. Nevertheless, reasonably good agreement between the experimental and computational results demonstrates eligibility of

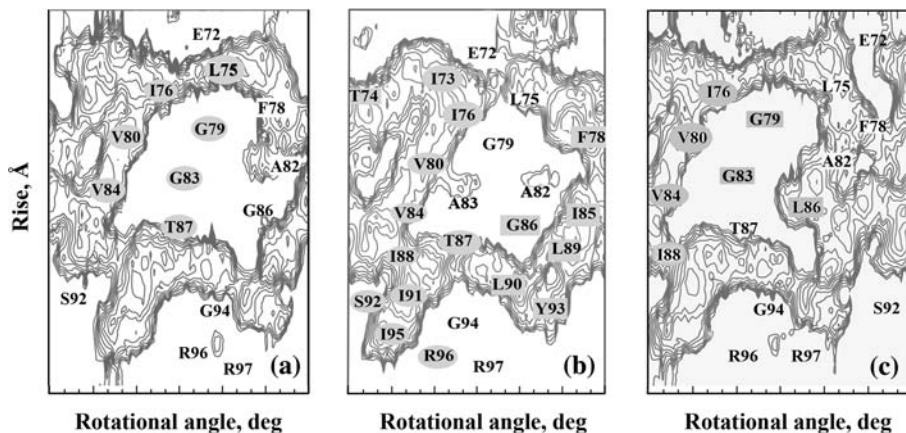


Figure 5. Hydrophobicity and packing of the NMR-derived and calculated models of mutant GpA dimer. Hydrophobic properties of α -helices. 2D isopotential map of the MHP on the peptide surface calculated as described in [55]. The value on the x -axis is the rotation angle about the helix axis; the parameter on y -axis is the distance along the helix axis. MHP is given in octanol–water logP units. Only the hydrophobic areas with MHP > 0.05 are shown. Contour intervals are 0.015. The positions of residues are indicated by letters and numbers. (a) GpA_{NMR}: Gray ovals show residues on the helix–helix interface; (b) GpA_{G83A}: Gray ovals and rectangles indicate interfacial residues for monomers 1 and 2 in non-symmetrical right-handed dimer of the mutant, respectively (see text); (c) GpA_{G86L}: Gray ovals and rectangles show interfacial residues in symmetrical right- and left-handed dimers, respectively (note that I76, V80 are present in both structures) (see text).

the theoretical approach. It is principally important that, despite certain conformational heterogeneity, in all the calculated structures the helices contact each other with the same side, and it is exactly the side revealed in experiments.

As seen in Figure 3, the geometry of membrane insertion for resulting dimers and the degree of exposure of their termini to water depend on the details of inter-helical interactions. A pictorial example is provided by the two different conformers from groups 1 and 3 (Table 1). For the former one the dimerization interface contains six residues in the middle part of helices, and the monomers are considerably tilted ($\Theta \approx -50^\circ$). For the latter conformer the interface is extended over the entire sequences, thus making helices almost parallel to each other ($\Theta \approx 6^\circ$). Interestingly, in both cases the complexes are well-adapted to the thickness of the hydrophobic layer, although by a different way. (The term “adaptation” means that hydrophobic and hydrophilic parts of the peptides are exposed respectively to nonpolar and polar layers of the membrane.) In the former case the symmetry axis of the dimer is perpendicular to the membrane plane, while in states of the group–3 the dimer fits to the hydrophobic medium by tilting this axis up to $75\text{--}80^\circ$ with respect to the membrane plane. In both structures favorable inter-helical contacts are well-preserved. As seen in Table 1, occurrence of glycines on the interface between monomers (e.g.,

in models GpA_{NMR} and GpA_{MC}) increases their packing density (small values of d).

Most often, in the considered structures α -helices pack in a symmetrical (or nearly symmetrical) manner with respect to one another, i.e. the same amino acid residues of each monomer participate in dimerization (Table 1). Because the α -helices have identical sequences, this result seems quite probable, although asymmetrical homodimers have also been predicted for some other TM peptides [58]. We should note that the NMR-model has been built as a symmetrical dimer as well. For the resulting ensemble of low-energy states the largest population is observed when $d \sim 9 \text{ \AA}$. In addition, there exist conformations with an even tighter packing of monomers – for example, in groups 4 and 5. This result agrees with the data on statistical analysis of high resolution spatial structures of α -helical complexes in MPs [5, 59]: in the majority of them the helices are located at a distance $d \sim 9.6 \text{ \AA}$, while conformations with $d \sim 7.0 \text{ \AA}$ can rarely be observed. Among the states with $\Theta < 0$ is one (highly populated) group of the dimers (Table 1) which have the packing parameters close to the model GpA_{NMR}. The best correspondence is observed for the conformers of 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 in zwitterionic micelles.

NMR-derived model of the dimer is stable in implicit membrane

Besides the just considered models, it is also interesting to assess the energetic characteristics of the GpA_{NMR} model in the implicit membrane ($\text{GpA}_{\text{NMR-mem.}}$). Firstly, this is indispensable for evaluation of the computational efficiency of the approach: whether is it possible to use the energy-related criteria for selection of correct solutions and how this should be done? Secondly, comparison of the results obtained starting from the “correctly-packed” dimer structure with those found from arbitrary starts is important to understand the quality of sampling of the GpA ’s conformational space – is it sufficient enough to provide unambiguous solutions? To address these issues, we performed MC simulations with the NMR-derived starting structure (GpA_{NMR} model). Analysis of the resulting low-energy states revealed that the complex adopts a TM orientation with the termini accessible to water phase (Figure 3, Table 1). The structure and the mutual location of the monomers are well retained – for residues 71–95 the backbone RMSDs ($\text{GpA}_{\text{NMR-mem.}}$) with the starting NMR-model do not exceed 1.2 Å. We should note that no restraints were used on the last stages of the simulation. The model $\text{GpA}_{\text{NMR-mem.}}$ 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 3) and dimerization interface (Table 1). One exception is provided by residues L75 – in the model GpA_{MC} their side chains do not form such tight inter-monomer contact, as in the experimental structure. Despite a relatively high backbone RMSD value (2.5 Å on the region 71–95) and somewhat over-estimated distances for a number of helix contacts (data not shown), in overall, the model GpA_{MC} represents a rather good approximation for the structure observed by NMR in micelles. Here we should mention that, unlike the GpA_{NMR} model, the helices were not restrained to their ideal conformations.

Behavior of the GpA_{NMR} model was also studied via MD simulations in all-atom lipid bilayers [60] and SDS micelles [61]. It was concluded that the dimer retains well its overall structure, although some conformational

rearrangements were observed as well. In particular, during the early phase of micelle formation around the GpA dimer, Braun et al. [61] detected significant bending of one of α -helices. Also, the tilt angle between helices changed from -37° in the starting NMR structure to -44° over the course of the simulation. Such values were comparable to those seen in relatively short (1.5 ns) MD simulations of GpA carried out in a variety of lipid bilayers [60]. Despite such structural adjustments, the residues G79 and G83 were always found on the dimerization interface and represented a sort of knuckle in the centre of the dimer, which allows the rotation around them. In our GpA_{MC} model the angle between monomers is -28° and even in this case the dimerization interface is almost the same as in the GpA_{NMR} model (Table 1). Therefore, the results of unrestrained MC search agree reasonably well with those obtained using independent computational approaches.

Is the alternative packing possible for the dimer?

Analysis of the resulting low-energy MC-states shows that among them there are conformers with $\Theta > 0$ (e.g., groups 2 and 5 in Table 1). They correspond to left-handed helical dimers. Indeed, packing parameters of helices of the group-3 are frequently observed in left-handed TM helical pairs [5], while the densely packed complexes of the group-5 seem to be quite similar to the left-handed dimers proposed earlier based on mutagenesis data [12]. 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. In our opinion, the presence of such type of structures raises an issue about the possibility of dimeric models, which are alternative to those obtained in the result of NMR-data interpretation [25]. It should be noted that the model GpA_{NMR} was built based on mutagenesis studies, on NMR-data confirming α -helical conformation of the monomers, and on a set of 10 inter-helical NOE-contacts observed by NMR. Whether is it possible a dimeric structure with $\Theta > 0$, which also satisfies the restraints imposed by mutagenesis and NMR results?

Recently we have demonstrated [62] that such a hypothetical model can be built based on the structural scaffold provided by the left-handed

MC-conformers. This model agrees reasonably well with the NMR distance restraints. Interestingly, residues I76, G79, V80, and G83, whose importance for helix association was proved by the mutagenesis results [12], form a helix–helix interface in the model with $\Theta > 0^\circ$. Besides them, the residues G86, L89, L90 are also involved in helix association. According to mutagenesis data, replacement of each of the last ones significantly alters dimerization of GpA, although in the model GpA_{NMR} they lie apart from the interface. It is important that the hydrophobic organization of right- and left-handed complexes is quite similar and does not contradict to the known principles of packing of TM helices in proteins: α -helices of GpA interact via residues which form hydrophilic stretches on their surfaces [62]. In this case there is a strong complementarity between polarity regions of the monomers, whereas the most hydrophobic surfaces are exposed to nonpolar lipid environment (Figure 5a).

The feasibility of left-handed dimers has been discussed previously in the mutagenesis work of Lemmon et al. [12]. Also, this question was considered in the recent paper by Im et al. [30]. Indeed, MD simulations of GpA in implicit membrane starting from non-interacting TM helices revealed two distinct families of conformers: a right-handed ($\Theta \sim -50^\circ$), and a left-handed ($\Theta \sim 40^\circ$) dimers. Interestingly, the left-handed configuration showed 94% occupancy, whereas the right-handed one only occurred 4% of the MD-time. Moreover, an unexpected result of the work [30] was that the left-handed dimers demonstrated tighter packing as compared with the right-handed models.

The factors important for helix association

Two-stage model of folding

How do the helices self-associate in membrane adopting the “native-like” conformation? (The dimer models which agree with mutagenesis and NMR data in membrane-mimic media are referred to as the “native-like” ones, although the spatial structure of the GpA dimer under native conditions is yet to be solved.) The results of unrestrained MC search are consistent with the famous two-stage model of folding of TM domains in proteins [63]. In the beginning of simulations significant gain in energy is caused by the

independent insertion of each monomer into the hydrophobic layer of the membrane. This corresponds to the first stage, or insertion. The media effects play crucial role in this process. During the later MC-steps the helices self-associate with subsequent “fine tuning” of the dimer. This represents the second stage, or dimerization. Such events are mainly driven by appearance of favorable van der Waals inter-monomer contacts. Applied TM voltage favors to a parallel (“head-to-head”) disposition of α -helices. Such a configuration of the system represents a suitable starting point for obtaining the most adequate (“native-like”) final solutions via MC search – tightly packed parallel TM α -helical dimers. This is because the number of possible variants of monomers’ association is seriously limited as compared, for example, with two conformationally-labile peptides *in vacuo*. 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. At the second stage the main energy gain is achieved upon the establishment of favorable contacts between the monomers. On the other hand, 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. An indirect support of this is provided by destabilization of the GpA_{NMR} model in MC simulations *in vacuo* (data not shown).

Diverse lipid composition of biomembranes affects their hydrophobic properties and geometrical characteristics. This, in turn, may influence protein–membrane and protein–protein interactions. In the next sections the role of the thickness and the hydrophobicity degree of the implicit membrane is considered in conformity with the helix association.

The role of membrane thickness

How does the thickness (D) of the hydrophobic layer of a membrane affect the inter-monomer interactions? Analysis of the low-energy states obtained at different values of D in the range 24–40 Å shows that their structures depend on the membrane thickness. Thus, increasing of D leads to decreasing of the tilt (Θ) between the monomers. Particularly, in “thin” ($D = 24$ Å) and in “thick” ($D = 40$ Å) membranes the most populated

are the states with $60^\circ < |\Theta| < 100^\circ$ and $|\Theta| < 20^\circ$, respectively. In the intermediate case the angle Θ varies from $(-50^\circ, 60^\circ)$ at $D=30$ Å up to $(-20^\circ, 20^\circ)$ at $D=36$ Å (data not shown). It should be noted that the model GpA_{NMR} is found among the low-energy states only at $D \leq 34$ Å. Future thickening of the membrane is not energetically favorable for this structure because polar termini of the monomers become to be exposed to the hydrophobic environment.

Detailed inspection of the MC-states obtained from arbitrary starts (including GpA_{NMR} and a number of MC-structures obtained at $D=30$ Å) permits a conclusion that the “hydrophobic match/mismatch” effects are important for each of non-bonded or weakly-bonded monomers, i.e. before their tight association into a complex. As soon as the most energetically favorable mutual orientation of α -helices is achieved, the spatial structure of the dimer is defined. Later, only reorientation of the entire complex with respect to the membrane is possible without noticeable changes of its conformation. This serves as an additional confirmation of the two-stage hypothesis mentioned above. To summarize, we can conclude that the thickness of the hydrophobic layer of the membrane strongly affects dimerization. The “native-like” models were obtained at the values of D in the range 30–34 Å, which correspond well to the length of the hydrophobic part of each helix of GpA.

Influence 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 the models of “less hydrophobic” and “more hydrophobic” membrane (see Methods). These computational experiments are believed to be very useful to estimate the reasonable weighting factor for the solvation term included into the potential energy function. Also, they may give ideas how the helix association proceeds in membranes of different composition.

Water–octanol–water membrane. The low-energy states obtained in the MC-search represent TM α -helical dimers, the majority of which (group 1' in Table 1) have the parameters $d \approx 10$ Å, $\Theta < 10^\circ$,

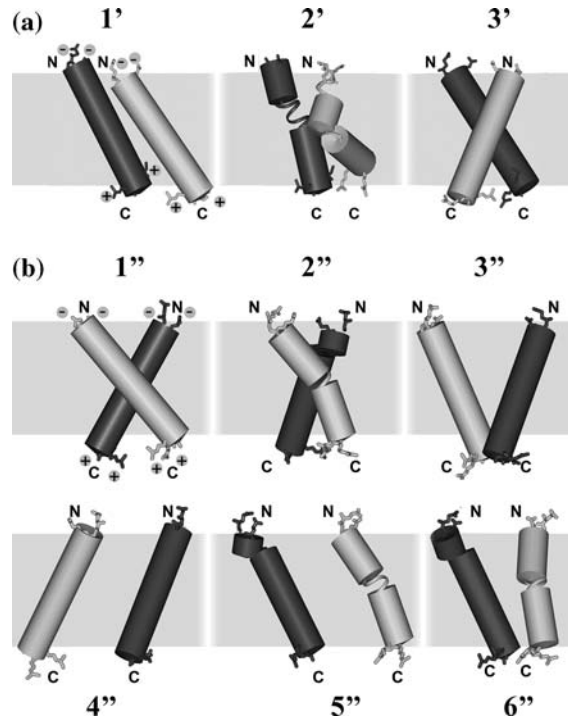


Figure 6. Low-energy states found via MC simulations of two GpA helices in implicit membranes with different degree of hydrophobicity. The results obtained in the “less hydrophobic” (water–octanol–water) (a) and in the “more hydrophobic” (water–cyclohexane–water, see text) (b) membranes. Numbering of the groups is that as in Table 1. Other details are as in the legend to Figure 1.

and a less numerous groups of states (e.g., 3') have the parameters $d \approx 9$ Å, $\Theta \approx -50^\circ$ (Figure 6a). A number of HP-like structures with the helices broken on residues GxxxG were observed as well (group 2'). Interestingly, in the “less hydrophobic” membrane a number of monomers settle themselves in such a way that their charged amino acid residues (E70, for instance) are buried in the nonpolar environment. This was not observed in calculations with the original membrane model. Analysis of the low-energy states shows that α -helices interact mainly by means of bulky hydrophobic side chains, and the dimerization interface differs greatly from the experimentally defined (Table 1). Therefore, the “more polar” environment promotes contacts between the most hydrophobic residues. Moreover, the states close in their structure to those observed in the water–cyclohexane–water membrane (e.g., GpA_{MC}) were not found.

Water–cyclohexane–water membrane. In this hypothetical model the central part of the nonpolar slab is “more hydrophobic” than cyclohexane (in terms of the free energy of transfer from water). One might expect that in such a medium TM helices may tend to form tightly bound complexes via contacts between the hydrophilic patterns G79/G83/T87 or G79/G83/G86 on their surfaces, thus avoiding exposure of these motifs to apolar environment. However, MC simulations did not support this assumption. In particular, about 150 of the ~ 1000 low-energy states are the left-handed dimers (Table 1, group-1'). Their conformations are close to the models of the group-5 obtained in the non-modified membrane. The packing parameters and helix–helix interfaces for them are shown in Table 1. Interestingly, the region of intermonomer contacts contains three glycines, which are known to be important for dimerization. At the same time, the vast majority of the other energetically favorable states reveal large degree of conformational heterogeneity. Therefore, their clustering is seriously impeded. Roughly, three modes of monomers' interactions may be delineated. Mode-1: Regardless of the secondary structure (all-helical and HP-like), GpA monomers do not form dimers (Figure 6b, Table 1, groups 4' and 5'). This is the most populated group of states (~ 600 conformers). Mode-2: “V-shaped” (Figure 6, groups 3' and 6') conformers revealing weak intermolecular contacts only near their termini (~ 150 states). Like in the former case, the monomers may have α -helical or distorted helical conformation (Table 1, groups 3' and 6'). Mode-3: Densely packed dimers (Figure 6b, group 2') with destabilized α -helical structure (Table 1, group 2'). It is worth noting that most of the conformers from the groups 2'–6' have no specificity in helix–helix contacts.

Therefore, the contribution of the solvation effects was overestimated in such a hypothetical apolar membrane. This is pictorially illustrated by the large number of non-interacting monomers where inter-molecular contacts are less energetically favorable than solvation of individual peptides with the hydrophobic medium. It is remarkable, that no significant changes of the complexes occur in both membranes when starting MC search from tightly packed structures (GpA_{NMR-mem.} model and MC dimers of groups 1–6, Table 1). Energy of such states is comparable with those obtained from random

starts. Once again, this agrees well with the two-stage model of MPs' folding

To summarize, *ab initio* predictions of the dimer's structure and its mode of membrane binding are reliable only in a case, when an adequate membrane model is employed. Thus, alterations of the original (water–cyclohexane–water) membrane model [44] do not provide correct description of TM helix interactions. This clearly demonstrates the importance of the balance between $E_{\text{solv.}}$ and other energy terms: underestimation as well as overestimation of the media effects may considerably affect the resulting low-energy states.

Mutagenesis in silico: effects of point mutations on dimerization of GpA

Whether the proposed computational approach is capable to reproduce the point-mutagenesis data on the role of particular residues in dimerization of GpA? To address this issue, we performed MC simulations of two mutant homo-dimers composed respectively of α -helices with G83A and G86L amino acid replacements. The former mutant is known to disrupt dimerization of GpA [12], while the later one possesses the wild-type level of activity. Analysis of the low-energy states obtained in the water–cyclohexane–water system shows that in both cases the most energetically favorable conformers represent TM α -helical dimers. At the same time, the packing of monomers is quite sensitive to the point amino acid replacements. Thus, the structures obtained for the mutant G83A may be clustered into two groups. The first one is more populated and corresponds to a right-handed helical complex with the parameters $d \approx 10$ Å, $\Theta = -30$ to -5° and asymmetrical interface:

monomer-1: I T V V TI IS IR
monomer-2: S T F IG LL Y

Here the gray hatching shows, whether a given residue contributes to dimerization according to mutagenesis data. The conformers from the second group are left-handed dimers with the parameters $d \approx 8$ Å, $\Theta \approx 40^\circ$. They also reveal asymmetrical interface:

monomer-1: TI V AV T LI G R
monomer-2: SE LI FG A G LL Y R

It is therefore seen that the dimeric structures from the both groups are drastically different from the wild-type NMR-derived models and MC-states (Table 1). Furthermore, in contrast to simulations with the wild-type peptides, the helix interfaces found in the GpA_{G83A} models do not correlate with mutagenesis data on dimerization. Most probably, the loss of the wild-type packing is explained by distortion of the “hydrophilic pattern” G79, G83, T87 on the helix surface important for association. Thus, the replacement G83A leads to creation of a hydrophobic zone in the center of such pattern (Figure 5b). As a consequence, dispositions of the helix–helix interfaces are significantly different as compared with the wild-type dimer (Figure 5a). An important question arises: why do the simulations still predict dimers for the GpA_{G83A} model, while mutagenesis results do not? We assume that one possible reason for this may be related to limitations of our modeling approach, where helix association is more favorable than existence of TM helices in the monomeric state. But even in this case analysis of MC results permits a conclusion that dimerization is not as strong as for the wild-type: the mutant helices pack via different sides, and there is a large structural difference between the families of the found conformers. Therefore, based on such results it is possible to realize that the mutation G83A induces destabilization of the dimer. This agrees well with the experimental observations.

Behavior of the mutant GpA_{G83A} was also studied via MD simulations in all-atom hydrated SDS micelles [61]. In such environment the mutant was much less stable as compared with the wild type GpA dimer. For instance, the protein backbone root-mean-square deviation continued to increase past 3.5 Å after 2.5 ns of MD, indicating non-equilibrium character of the system. Also, the helices rotated with respect to one another. These structural rearrangements were accompanied by serious distortions of the micellar system. So, both modeling techniques (MC and MD) lead to the same general conclusion – the replacement G83 → A significantly decreases ability of TM segments to associate in membrane.

Drastically different picture is observed in simulations of the G86L mutant – resulting low-energy states reveal packing parameters resembling those found for the wild-type GpA. Thus, several groups of states with $\Theta \approx -50, -30, 20, 45,$

and 65° were delineated. Moreover, the conformers from two of these groups were very close to the structures from the groups 1 and 5 observed for the wild-type peptides (Table 1). The former ones have packing parameters $d \approx 10$ Å, $\Theta \approx -50^\circ$ and dimerization interface I⁷⁶IXXVMXXV⁸⁴ (it does not contain mutated residue). The conformers from another group have parameters $d \approx 6.5$ Å, $\Theta \approx 65^\circ$ and the interface I⁷⁶XXGVXXGXXL⁸⁶. (Here the replaced residue L86 falls in the area of inter-helical contacts.) Interestingly, such left-handed dimers may be formed despite the fact that the mutation G86L distorts the “hydrophilic pattern” created by residues G79, G83, and G86 on the helix surface. In this case occurrence of the motif G⁷⁹XXXG⁸³ is sufficient enough to allow helix packing, although it is possible that the presence of G86 may give an additional stabilization to the dimer. We should note that the role of G86 in dimerization is still not well understood – for example, the mutations G86 → T and G86 → Y disrupt the dimer [12] but the NMR-derived model of GpA [25] fails to explain such effects. Therefore, the MC-results obtained for the mutant G86L are well consisted with the data for the wild-type GpA. This also agrees with the conclusions made in the mutagenesis work by Lemmon et al. [12].

To summarize, testing of the developed computational approach on two GpA mutants clearly shows that it provides an adequate description of the role of point mutations (at least for those considered in this study) in association of TM helices. It therefore can be further applied to assess the effects of replacements of particular amino acids in GpA and in other membrane bound helical oligomers.

Conclusions

This work describes modeling of two α -helices of GpA *in vacuo*, in implicit membrane without and with applied TM voltage ($\Delta\psi$). The role of different factors promoting formation of homodimer in heterogeneous membrane environment is considered. Among them are: medium effects, $\Delta\psi$, thickness and hydrophobicity degree of membrane, choice of the starting configuration of the system. It was shown that the membrane-mimic medium considerably stabilizes α -helical

conformation of the GpA monomers, facilitates their insertion into membrane, and mediates helix interactions. The voltage difference between the membrane sides plays an important role in the “native-like” association of TM α -helices providing their “head-to-head” orientation. In addition, at the initial “rough tuning” stage of dimerization the “hydrophobic match” effects and hydrophobic properties of membrane seem to be very important. Subsequent “fine-tuning” of the complex is mediated by van der Waals interactions. The calculations performed for GpA mutants G83A and G86L permit rationalization of mutagenesis data. Thus, the former mutant is unable to form dimers observed in experiments, while the later one reveals dimerization interface close to that in the wild-type GpA. In overall, the results of calculations agree reasonably well with the mutagenesis data. Also, several groups of the found low-energy dimeric structures are close to those obtained by NMR. Finally, MC simulations in membranes of different thickness and hydrophobicity degree demonstrate that the most adequate behavior of GpA is achieved at $D=30\text{--}34\text{ \AA}$, and in the medium defined as water-cyclohexane-water three-layer system.

The proposed computational approach possesses a number of shortcomings and moot points. As shown in the present study, the choice of the “native-like” structure of the dimer is not straightforward if there is no *a priori* knowledge about exact helix-helix interface. Although most of the predicted conformations of the GpA dimer have quite similar overall patterns of inter-monomer contacts, their detailed structures may differ. This depends on a number of factors, like membrane thickness and hydrophobicity degree, presence of TM voltage, and so forth. On the other hand, in many cases, the objective is not elaboration of a very precise spatial structure of a protein oligomer. Instead, the problem often lies in establishment of a general mode of membrane insertion for a complex, along with delineation of some crucial residues on the dimerization interface. Such a rough model may be future employed to rationalize experimental observations and to design new experiments. Furthermore, 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 an

equilibrium dynamic behavior of the monomers in membrane-mimic environments used in the experimental studies. We also propose that some fraction of the left-handed α -helical dimers may present as well. As discussed above, such a conformational heterogeneity may be caused by the fact that the exact packing of helices is quite sensitive to media effects and the geometry of membrane binding. Partially, such a hypothesis is corroborated by somewhat vague results of mutagenesis studies (see discussion in [12]), as well as by NMR [27, 48] and MD [30, 58, 60] data that demonstrate the importance of media effects for stability of helical oligomers and provide examples of their multi-state equilibrium in lipid bilayers and membrane mimics. In real biological membranes the situation may be more complex – due to nonhomogeneous content of lipid bilayers, their domain structure, variations of physico-chemical characteristics, presence of small molecules (e.g., cholesterol), etc. As a result, a predominant conformation(s) 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 TM voltage, an optimal membrane thickness and hydrophobicity degree, etc., only a limited number of spatial structures are energetically favorable. Some of them are very close to those observed by NMR in detergent micelles. We assume that the others may also be realized in real lipid bilayers and in artificial membrane-mimic environments with different properties. In our opinion, future MD simulations starting from these representative conformations in explicit bilayers and micelles will provide an additional insight into the equilibrium behavior of helical oligomers. This work is currently in progress in our group. Therefore, the proposed computational approach represents an indispensable step towards the development of efficient theoretical method to study helix interactions in membranes.

Acknowledgements

This work was supported by the Programme RAS MCB, by the Russian Foundation for Basic Research (Grant 04-04-48875-a, 05-04-49346-a), by the Russian Federation Federal Agency for Science and Innovations (The State contract

02.467.11.3003 of 20.04.2005, grant SS-1522.2003.4). R.G.E. is grateful to the Science Support Foundation (Russia) for the grant awarded. Access to computational facilities of the Joint Supercomputer Center (Moscow) is gratefully acknowledged.

References

1. Ubarretxena-Belandia, I. and Engelman, D.M., *Curr. Opin. Struct. Biol.*, 11 (2001) 370.
2. Bechinger, B., *Mol. Membr. Biol.*, 17 (2000) 135.
3. Shai, Y., *Biopolymers*, 66 (2002) 236.
4. Walther, D., Eisenhaber, F. and Argos, P., *J. Mol. Biol.*, 255 (1996) 536.
5. Bowie, J.U., *J. Mol. Biol.*, 272 (1997) 780.
6. Zhang, C., Hou, J. and Kim, S.H., *Proc. Natl. Acad. Sci. U.S.A.*, 99 (2002) 3581.
7. Torres, J., Stevens, T.J. and Samson, M., *Trends Biochem. Sci.*, 28 (2003) 137.
8. Forrest, L.R. and Sansom, M.S., *Curr. Opin. Struct. Biol.*, 10 (2000) 174.
9. Liang, J., *Curr. Opin. Chem. Biol.*, 6 (2002) 878.
10. Efremov, R.G., Nolde, D.E., Konshina, A.G., Syrtcev, N.P. and Arseniev, A.S., *Curr. Med. Chem.*, 11 (2004) 2421.
11. Furthmayr, H. and Marchesi, V.T., *Biochemistry*, 15 (1976) 1137.
12. Lemmon, M.A., Flanagan, J.M., Treutlein, H.R., Zhang, J. and Engelman, D.M., *Biochemistry*, 31 (1992) 12719.
13. Deber, C.M., Khan, A.R., Li, Z., Joensson, C., Glibowoska, M. and Wang, J., *Proc. Natl. Acad. Sci. U.S.A.*, 90 (1993) 11648.
14. Melnyk, R.A., Partridge, A.W. and Deber, C.M., *J. Mol. Biol.*, 315 (2002) 63.
15. Schlessinger, J., *Cell*, 103 (2000) 211.
16. Wells, J.A., *Proc. Natl. Acad. Sci. U.S.A.*, 93 (1996) 1.
17. Senes, A., Engel, D.E. and DeGrado, W.F., *Curr. Opin. Struct. Biol.*, 14 (2004) 465.
18. Dawson, J.P., Weiner, J.S. and Engelman, D.M., *J. Mol. Biol.*, 316 (2002) 799.
19. Langodch, D. and Heringa, J., *Proteins*, 31 (1998) 150.
20. Eisenberg, D., Weiss, R.M. and Terwilleger, T.C., *Nature*, 299 (1982) 371.
21. Senes, A., Ubarretxena, I. and Engelman, D.M., *Proc. Natl. Acad. Sci. U.S.A.*, 98 (2001) 9056.
22. Melnyk, R.A., Kim, S., Curran, A.R., Engelman, D.M., Bowie, J.U. and Deber, C.M., *J. Biol. Chem.*, 279 (2004) 16591.
23. Treutlein, H.R., Lemmon, M.A., Engelman, D.M. and Brünger, A.T., *Biochemistry*, 31 (1992) 12726.
24. Adams, P.D., Engelman, D.M. and Brünger, A.T., *Proteins*, 26 (1996) 257.
25. MacKenzie, K.R., Prestegard, J.H. and Engelman, D.M., *Science*, 276 (1997) 131.
26. Smith, S.O., Song, D., Shekar, S., Groesbeek, M., Ziliox, M. and Aimoto, S., *Biochemistry*, 40 (2001) 6553.
27. Gratkowski, H., Dai, Q., Wand, A.J., DeGrado, W.F. and Lear, J.D., *Biophys. J.*, 83 (2002) 1613.
28. Fisher, L.E., Engelman, D.M. and Sturgis, J.N., *Biophys. J.*, 85 (2003) 3097.
29. Ducarme, P., Thomas, A. and Brasseur, R., *Biochim. Biophys. Acta*, 1509 (2000) 148.
30. Im, W., Feig, M. and Brooks III, C.L., *Biophys. J.*, 85 (2003) 2900.
31. Pappu, R.V., Marshall, G.R. and Ponder, J.W., *Nat. Struct. Biol.*, 6 (1999) 50.
32. Smith, S.O., Smith, C., Shekar, S., Peersen, O., Ziliox, M. and Aimoto, S., *Biochemistry*, 41 (2002) 9321.
33. Sanguk, K., Chamberlain, A.K. and Bowie, J.U., *J. Mol. Biol.*, 329 (2003) 831.
34. Fleishman, S.J., Schlessinger, J. and Ben-Tal, N., *Proc. Natl. Acad. Sci. U.S.A.*, 99 (2002) 15937.
35. Gennis, R.B. *Biomembranes. Molecular Structure, Function*. Springer-Verlag, New York, Berlin, Heidelberg, Tokyo, 1989.
36. Beckstein, O., Biggin, P.C., Bond, P., Bright, J.N., Domene, C., Grottesi, A., Holyoake, J. and Sansom, M.S., *FEBS Lett.*, 555 (2003) 85.
37. van Dalen, A., Killian, A. and de Kruijff, B., *J. Biol. Chem.*, 274 (1999) 19913.
38. Lauger, P., Lesslauer, W., Marti, E. and Richter, J., *Biochim. Biophys. Acta*, 135 (1967) 20.
39. Roux, B., *Biophys. J.*, 73 (1997) 2980.
40. Gazit, E., Bach, D., Kerr, I.D., Sansom, M.S., Chejanovsky, N. and Shai, Y., *Biochem. J.*, 304 (1994) 895.
41. Gazit, E., Miller, I.R., Biggin, P.C., Sansom, M.S. and Shai, Y., *J. Mol. Biol.*, 258 (1996) 860.
42. Biggin, P.C. and Sansom, M.S., *Biophys. Chem.*, 76 (1999) 161.
43. La Rocca, P., Biggin, P.C., Tieleman, D.P. and Sansom, M.S., *Biochim. Biophys. Acta*, 1462 (1999) 185.
44. Efremov, R.G., Volynsky, P.E., Nolde, D.E. and Arseniev, A.S., *Theor. Chem. Acc.*, 106 (2001) 48.
45. Efremov, R.G., Volynsky, P.E., Nolde, D.E., van Dalen, A., de Kruijff, B. and Arseniev, A.S., *FEBS Lett.*, 526 (2002) 97.
46. Fattal, D. and Ben-Shaul, A., *Biophys. J.*, 65 (1993) 1795.
47. Webb, R.J., East, J.M., Sharma, R.P. and Lee, A.G., *Biochemistry*, 37 (1998) 673.
48. Li, M., Reddy, L.G., Bennett, R., Silva, N.D., Jones, L.R. Jr. and Thomas, D.D., *Biophys. J.*, 76 (1999) 2587.
49. Berman, H.M., Bhat, T.N., Bourne, P.E., Feng, Z., Gilliland, G., Weissig, H. and Westbrook, J., *Nat. Struct. Biol.*, 7 (2000) 957.
50. Efremov, R.G., Nolde, D.E., Vergoten, G. and Arseniev, A.S., *Biophys. J.*, 76 (1999) 2448.
51. Efremov, R.G., Volynsky, P.E., Nolde, D.E., Dubovskii, P.V. and Arseniev, A.S., *Biophys. J.*, 83 (2002) 144.
52. Némethy, G., Pottle, M.S. and Scheraga, H.A., *J. Phys. Chem.*, 87 (1983) 1883.
53. Metropolis, N., Rosenbluth, A.W., Teller, A.H. and Teller, E., *J. Chem. Phys.*, 21 (1953) 1087.
54. von Freyberg, B. and Braun, W., *J. Comp. Chem.*, 12 (1991) 1065.
55. Efremov, R.G. and Vergoten, G., *J. Phys. Chem.*, 99 (1995) 10658.
56. Kabsch, W. and Sander, C., *Biopolymers*, 22 (1983) 2577.
57. Arkin, I.T., *Biochim. Biophys. Acta*, 1565 (2002) 347.
58. Stockner, T., Ash, W.L., MacCallum, J.L. and Tieleman, D.P., *Biophys. J.*, 87 (2004) 1650.

59. Eilers, M., Patel, A.B., Liu, W. and Smith, S.O., *Biophys. J.*, 82 (2002) 2720.
60. Petrache, H.I., Grossfield, A., MacKenzie, K.R., Engelman, D.M. and Woolf, T.B., *J. Mol. Biol.*, 302 (2000) 727.
61. Braun, R., Engelman, D.M. and Schulten, K., *Biophys J.*, 87 (2004) 754.
62. Vereshaga, Y.A., Volynsky, P.E., Nolde, D.E., Arseniev, A.S. and Efremov, R.G., *J. Chem. Theory Comput.*, 1 (2005) 1252.
63. Popot, J.L. and Engelman, D.M., *Biochemistry*, 29 (1990) 4031.