Introducing an Implicit Membrane in Generalized Born/Solvent Accessibility Continuum Solvent Models

Velin Z. Spassov,* Lisa Yan, and Sándor Szalma

Accelrys Inc., 9685 Scranton Road,, San Diego, California 92121

Received: March 11, 2002; In Final Form: May 21, 2002

A new empirical approach to model the solvent effects in protein−membrane complexes is proposed. The generalized Born (GB) approximation is extended by including an implicit membrane (IM) in the calculation of the electrostatic contribution to the solvation free energy (GB/IM model). In addition, nonpolar solvation energy terms are calculated on the basis of the solvent-accessible surface approximation including the effect of membrane (SA/IM model). The generalized Born−solvent-accessible surface area (GBSA) model with implicit membrane (GBSA/IM) is implemented in the CHARMm package and is applicable for energy calculations and molecular dynamics simulations. The potential of the new method for studying large molecular systems is demonstrated with the example of two transmembrane proteins, bacteriorhodopsin and rhodopsin. The results show a clear directional asymmetry of the solvation energy during translocation of the proteins through the membrane, which is suggested to be an alternative explanation of the known "positive-inside" rule. The method is also tested in nanosecond molecular dynamics (MD) simulations of the influenza virus HA2 (1:20) fusion peptide. Interestingly, when starting from two different initial positions of peptide, during the 3 ns runs, the helical fragment consistently adopts a tilted (∼20−25°) orientation with respect to the membrane in very close agreement with the known electron paramagnetic resonance (EPR) data. We also found an excellent agreement between the p*K*s of the N-terminal amino group computed for the final MD structures and the known NMR titration data.

1. Introduction

In molecular modeling, the treatment of the interactions of macromolecules with biomembranes is a difficult problem. Despite the remarkable progress in simulation techniques employing explicit lipid and solvent molecules, their use in many cases is impractical because of the high computational cost and size limits. The recent explosion in structural genomics and bioinformatics increases the demand for simpler methods. The use of continuum electrostatic methods based on the solution of the Poisson−Boltzmann (PB) equation has shown that macroscopic models in many cases provide a reasonable approximation of the polar contribution to the solvent effects.[1] Several PB solvers[2−6,51−53] include options where the membrane is represented as an infinite dielectric slab. However, the PB solvers are still not fast and convenient enough for intensive simulations. A decade ago a generalized Born (GB) approximation was introduced by Still et al.[7] as a practical method of calculating the electrostatic contributions to the atomic solvation effects in molecular models using implicit solvent. Since then, several modifications of the GB method have been applied successfully in molecular mechanics and molecular dynamics simulations,[8−11] quantum-mechanical solvation models,[12] p*K* calculations,[13] ligand docking,[14] etc. (see ref 15 for a recent review). The generalized Born−surface area (GBSA) approximation[7] is based on the combination of the GB method with a nonpolar term linearly dependent on the solvent-accessible surface area. Although the GBSA method is a simplification of the real physical picture, it allows for the two major but opposite contributions to the solvation forces: the general hydrophobic effect that acts in a direction to reduce the exposure of all solute atoms and the electrostatic interactions between solute and solvent that attract the polar and charged groups to the water environment. However, in contrast to PB solvers, none of the known implementations of the GB method has an option to allow for the effect of the interactions between solute and membrane.

Here we present the formalism of an extension of the GBSA model.[7] It allows the membrane, represented implicitly (IM) as an infinite planar dielectric slab, to be included in GBSA calculations of solvation energy. in the frame of a new, generalized Born−surface area−implicit membrane (GBSA/IM) approach. For this purpose we modified the methods to calculate both polar and nonpolar contribution to solvation energy. First we extended the GB approximation[7] by introducing implicit membrane in the calculations of electrostatic contribution to solvation energy, what we will call the GB/IM method. We included membrane also in the calculations of solvent-accessible surface used in the approximation of the nonpolar term (SA/IM method). The calculations of solvation energy are based on the combined use of GB/IM and SA/IM methods, and we will call it the GBSA/IM approach. The most significant modification concerns the expression for the effective atomic Born radii, where we propose that the contribution of the membrane (or any other dielectric body of ideal geometry) should be approximated by a simple empirical function. The new method is implemented in the c27b4 version of the CHARMm program package[16,17] by modifying the modules Genborn[9] and Aspenr,[18] and we found it convenient to use for energy minimization and molecular dynamics simulations of peptides or proteins interacting with membranes. The parametrization of the GB/IM method and tests of its accuracy based on comparison to the results of

* To whom correspondence should be addressed.

10.1021/jp020674r CCC: $22.00 © 2002 American Chemical Society
Published on Web 07/31/2002

finite-difference solutions to the Poisson−Boltzmann equation (FDPB) are described under Methods. Using the example of two transmembrane proteins, bacteriorhodopsin and rhodopsin, we show the potential of the GBSA/IM method to study the dependence of solvation energy on the position and orientation of large macromolecules with respect to the membrane. In addition to the overall better computational efficiency with respect to the grid-based PB solvers, another advantage of the generalized Born approximation is that the speed of calculations depends only on the number of atoms, not on the linear dimensions of the molecular system. In the above case it allowed us to calculate the polarization energy at any distance of the protein from the center of the membrane without costly extensions of the grid size. This advantage will be even more important if one models proteins in an extended conformation or long nucleic acid chains.

Finally, we demonstrate the applicability of the GBSA/IM method to molecular dynamics simulations. We describe the results of several 3 ns simulations of a model peptide representing the N-terminal 20-amino acid fragment of influenza hemagglutinin HA2 fusion peptide[19] starting from different initial positions of an ideal α-helical structural model. We show the effect of membrane on the evolution of peptide conformation and orientation. The results of the MD simulations are compared to the experimental NMR and electron paramagnetic resonance (EPR) data[20] and theoretical FDPB calculations[21] of the same peptide.

2. Theory

Continuum Solvent Model with Implicit Membrane. In common with most continuum solvent models, the protein molecule in our consideration is treated in full atomic detail where each atom i is assigned a partial charge q_i and the molecular interior is assumed to have a uniform dielectric constant ϵ_m. The solvent is present as a uniform polar medium with a dielectric constant ϵ_w. The nonpolar region of the membrane is regarded as a planar dielectric slab with thickness L, and the dielectric constant inside the slab is assumed to have the same value ϵ_m as inside the protein. It will be seen below that the latter approximation was necessary to include the effect of membrane in the generalized Born method. In the case of a protein−membrane complex, the low dielectric constant (∼2) inside the hydrophobic part of lipid bilayer is consistent with the suggestion for a similar value inside the protein (∼2), if most electrostatic interactions, but not the induced polarization, are treated explicitly.[55] Note that although the two-dielectric approximation is a standard for continuum electrostatic models, the PB solvers, in principle, allow the assignment of more than two values of dielectric constants to the different parts of the system. An interesting implementation is in the recent version of DelPhi,[53] where several geometrical objects can be treated as bodies having different dielectric properties.

The total solvation energy is expressed as a sum of the polar and the nonpolar interaction terms:

$$\Delta G_{slv} = \Delta G_{pol} + \Delta G_{np} \quad (1)$$

For the nonpolar term, including the energy of cavity formation and van der Waals interactions between solute and solvent, we use the same approximation as in the GBSA method:[7]

$$\Delta G_{np} = \sigma SA \quad (2)$$

where SA is the total solvent accessible surface of the solute, calculated according to the Lee and Richards definition.[22] In a more extensive model, eq 2 can be extended by including individual atomic parameters mimicking the specific interactions between protein atoms and water molecules from first hydration shell as well as by including the effect of curvature[55] on the surface tension. If both energy terms in eq 1 are referenced to an infinite uniform environment with the properties of the molecular interior, ΔG_{slv} can be added to the potential of mean force in a standard program for atomistic calculations, such as CHARMm, AMBER etc. When the electrostatic contribution to the energy of solvation is calculated in the framework of the generalized Born approximation, a combined use of eqs 1 and 2 corresponds to the original GBSA model.[7] In this study we present an extension of the GBSA approach by including the effect of implicit membrane, IM, in the calculations of both ΔG_{pol} (GB/IM) and ΔG_{np} (SA/IM) terms. We will call this the GBSA/IM approach.

GB/IM: Generalized Born Model with Implicit Membrane. The electrostatic contribution to the energy of solvation is calculated in the framework of the generalized Born approximation. In the initial formulation of the GB method, Still et al.[7] proposed to calculate the electrostatic contribution to the free energy of solvation, ΔG_{pol}, when the molecular environment is changed from vacuum to a solvent with dielectric constant ϵ_{slv}. The vacuum environment can be easily generalized to any reference medium,[13,23] for example, a medium with the dielectric constant of the molecular interior, ϵ_m. Then the free energy of solvation of a constellation of N point charges, q_i, embedded in one or more dielectric bodies of arbitrary shape and surrounded by a solvent with a dielectric constant ϵ_{slv} is given by

$$\Delta G_{pol} = -166\left(\frac{1}{\epsilon_m} - \frac{1}{\epsilon_{slv}}\right)\sum_i\sum_j \frac{q_i q_j}{\sqrt{r_{ij}^2 + \alpha_i\alpha_j \exp(-r_{ij}^2/4\alpha_i\alpha_j)}} \quad (3)$$

In eq 3 the effect of molecular size and shape is absorbed in the set of effective Born radii, α_i. By definition, the α_i radii are related to the polarization energy $G_{pol,i}$ of a positive unit charge in the same environment but in the absence of the rest of the charge constellation:

$$\alpha_i = -\left(\frac{1}{\epsilon_m} - \frac{1}{\epsilon_{slv}}\right)\frac{166}{G_{pol,i}} \quad (4)$$

The method of calculating the effective Born radii is crucial for the computational efficiency of GB implementations. In the past few years a number of fast analytical approximations have been proposed.[8,23,24] Most of the practical methods for α_i calculations are based on the Coulomb field approximation,[15,23] where the dielectric displacement created by a charge q_i is replaced by a Coulomb field. It leads to a convenient expression[15,23,25] for $G_{pol,i}$:

$$G_{pol,i} = \left(\frac{1}{\epsilon_m} - \frac{1}{\epsilon_{slv}}\right)\left[-\frac{166}{R_i} + \frac{166}{4\pi}\int_{\substack{in,\\ r>R_i}} \frac{1}{r^4}\,dV\right] \quad (5)$$

where the integration is taken over all intramolecular space but not inside the van der Waals sphere of atom i with radius R_i.

Here we use eq 5 as a basis for extending the GB approximation including the interactions of solute with a membrane, represented by an infinite dielectric slab. Although the analysis given below is oriented toward the inclusion of a membrane in the molecular system, the formalism can be easily generalized for any low-dielectric body with ideal geometry, for example, oil droplets, spherical or cylindrical micelles, etc.

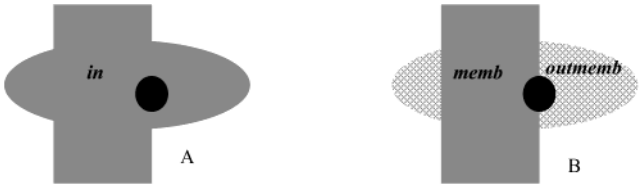

Figure 1. Integration domains according to eq 5 (A) and eq 6 (B). The black circles represent the van der Waals sphere of the atom *i*.

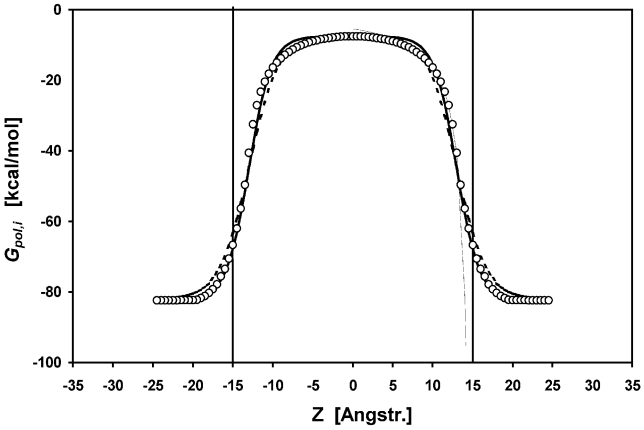

Figure 2. Polarization energy of a spherical ion of radius $R_i = 2$ Å and one unit electric charge computed as function of the distance from the center of an infinite planar dielectric slab with 30 Å thickness and $\epsilon_m = 1$. (○) FDPB values; lines represent GBIM energy at (—) $\gamma = 0.72$ Å^{-1} (best fit) and (---) $\gamma = 0.55$ Å^{-1}. The thin line represents the analytical solution for a point charge near a planar dielectric boundary between two infinite media.

If the membrane is regarded as a medium having the same dielectric constant as the solute, the integral in eq 5 can be split into two parts:

$$G_{\mathrm{pol},i} = -\left(\frac{1}{\epsilon_m} - \frac{1}{\epsilon_{slv}}\right)\left[\frac{166}{R_i} - \frac{166}{4\pi}\int_{\substack{\mathrm{memb},\\ r>R_i}} \frac{1}{r^4}\,\mathrm{d}V - \frac{166}{4\pi}\int_{\substack{\mathrm{outmemb},\\ r>R_i}} \frac{1}{r^4}\,\mathrm{d}V\right] \quad (6)$$

The first integral is taken over the space inside the membrane, while the second is over the rest of the solute interior, which is outside the membrane (see Figure 1).

Note that, according to the Coulomb approximation, the sum of the first two terms inside the brackets of eq 5 is nothing more than the polarization energy, $G^S_{\mathrm{pol},i}$, of a single ion in the presence only of the membrane. For an ideal membrane, represented by an infinite planar dielectric slab with thickness L, then, because of the symmetry, $G^S_{\mathrm{pol},i}$ depends only on one single geometrical variable—the distance Z from the membrane midplane:

$$G_{\mathrm{pol},i} = G^S{}_{\mathrm{pol}}(Z_i, R_i, L) + \left(\frac{1}{\epsilon_m} - \frac{1}{\epsilon_{slv}}\right)\frac{166}{4\pi}\int_{\substack{\mathrm{outmemb},\\ r>R_i}} \frac{1}{r^4}\,\mathrm{d}V \quad (7)$$

The value of G^s_{pol} can be estimated (tabulated) by solving the Poisson equation. Figure 2 shows the value of G^s_{pol} as a function of Z for a spherical ion with 2.0 Å radius in the presence of a membrane of 30 Å thickness calculated by the MEAD[2,3] program Solvate using the finite-difference Poisson−Boltzmann method at $\epsilon_m = 1$. We found (see below, Figure 2) that for slabs that are not very thin, i.e., $L \geq 20$ Å, the FDPB values of G^s_{pol} can be quite well approximated by a simple analytical function Γ:

$$G^S{}_{\mathrm{pol}}(Z_i, R_i, L) = \left(\frac{1}{\epsilon_m} - \frac{1}{\epsilon_{slv}}\right)\Gamma(Z_i, R_i, L)$$

where Γ depends only on one single adjustable parameter γ (in reciprocal angstroms):

$$\Gamma(Z_i, R_i, L) = g_i^{slv} + \frac{g^{cntr} - g_i^{slv}}{1 + \exp[\gamma(Z_i + R_i - L/2)]} \quad (8)$$

The term g^{slv} is the limiting value of Γ at infinite separation of atom i from the membrane and is therefore proportional to the Born energy of a single ion surrounded only by the solvent:

$$g_i^{slv} = \frac{G^S{}_{\mathrm{pol}}(Z_i \to \infty)}{\left(\frac{1}{\epsilon_m} - \frac{1}{\epsilon_{slv}}\right)} \cong \frac{166}{R_i}$$

while the $g^{cntr}(L)$ term is proportional to the solvation energy of a unit positive charge positioned exactly in the membrane midplane. At $\epsilon_m \ll \epsilon_{slv}$, $g^{cntr}(L)$ depends only on the membrane thickness. Following Parsegian,[26] we use the approximation

$$g^{cntr} = -\frac{332 \ln 2}{L}$$

Because the maximal slope of Γ function (eq 8) is proportional to the value of γ, a change of this parameter can also be used to mimic the sharpness of dielectric boundary. Figure 2 shows that if an ion is buried by more than one atomic radius inside the membrane, its polarization energy does not differ significantly from the analytical solutions[15] for a point charge at the same distance from a planar boundary between two infinite dielectric media. The hydrophobic part of lipid membranes, ~30 Å, is thick enough that a buried charge will feel the effect of only one boundary at a time. The analysis up until this point is valid for any practical realization of the GB method based on eq 6. In principle, the suggested Γ function can be used to represent implicitly the membrane in any GB technique, irrespective of whether it is based on pairwise interaction schemes or numerical volume integration.

Recently, Dominy and Brooks[9] proposed a linear version of the asymptotic pairwise GB approach of Qui at al.[8] The method of Dominy and Brooks, implemented in CHARMm module Genborn, was subjected to very intensive parametrization[9,27] to make it consistent with different CHARMM force fields. The authors used a linear fit procedure over large databases of peptides, proteins, and nucleic acids. The results of the tests show quite good agreement between the calculated GB solvation energy terms and the corresponding FDPB energies. It is demonstrated that the GB method makes the inclusion of the macroscopic polarization model into MD simulations feasible,[9] even in studies of protein folding thermodynamics.[10] According to the linear version[9] of the pairwise approach, eq 6 is approximated by

$$\frac{G_{\mathrm{pol},i}}{\frac{1}{\epsilon_m} - \frac{1}{\epsilon_{slv}}} = -\frac{166}{\lambda R_i} + P_1\frac{166}{R_i^2} + \sum_j^{\mathrm{bond}} P_2\frac{V_j}{r_{ij}^4} + \sum_j^{\mathrm{angle}} P_3\frac{V_j}{r_{ij}^4} + \sum_j^{\mathrm{nonbond}} P_4\frac{V_j}{r_{ij}^4}C(r_{ij}, R_i, R_j, P_5) \quad (9)$$

where V_j are atomic volumes, r_{ij} are interatomic distances, λ and P_1-P_5 are empirical parameters, and C is an empirical function.

By analogy with eq 7, we modified eq 9 to allow for the effect of the membrane on atomic polarization energies:

$$\frac{G_{\mathrm{pol},i}}{\frac{1}{\epsilon_m}-\frac{1}{\epsilon_{slv}}} = \Gamma(Z_i, R_i, L) + \sum_j^{\mathrm{bond}} P_2 \frac{V_j}{r_{ij}^4} + \sum_j^{\mathrm{angle}} P_3 \frac{V_j}{r_{ij}^4} + \sum_j^{\mathrm{nonbond}} P_4 \frac{V_j}{r_{ij}^4} C(r_{ij}, R_i, R_j, P_5) \quad (10)$$

Note that the sums in eq 10 are taken only over the atoms outside the membrane slab, while the contributions of the internal atoms are absorbed in the empirical function Γ, here calculated by eq 8. Note also, to be consistent with the Dominy and Brooks approximation, the g_i^{slv} term does not correspond exactly to the Born formula but is given by

$$g_i^{\mathrm{slv}} = -\frac{166}{\lambda R_i} + P_1 \frac{166}{R_i^2} \quad (11)$$

We used eqs 8 and 10 to modify the Genborn[9] subroutine of CHARMm. The derivatives of the function Γ were expressed analytically and distributed in the Genborn code as well. The preliminary tests performed on several transmembrane peptides and proteins showed quite a good agreement (less than 4−5% difference) between the GB and FDPB polarization energies. However, the CHARMm test for continuity of the first derivatives failed for certain atoms situated too close to the membrane boundaries. The reason for such a discontinuity is that atomic contributions appear or disappear in a stepwise manner in the sums in eq 10 when the atomic centers cross the boundary. Therefore, we made an additional modification of eq 10 to avoid the discontinuity of the analytical gradients and, in this way, to make it appropriate for use in molecular dynamics or minimization protocols:

$$\frac{G_{\mathrm{pol},i}}{\frac{1}{\epsilon_m}-\frac{1}{\epsilon_{slv}}} = \Gamma(Z_i, R_i, L) + \sum_j^{\mathrm{bond}} P_2 \frac{V_j(Z_j)}{r_{ij}^4} + \sum_j^{\mathrm{angle}} P_3 \frac{V_j(Z_j)}{r_{ij}^4} + \sum_j^{\mathrm{nonbond}} P_4 \frac{V_j(Z_j)}{r_{ij}^4} C(r_{ij}, R_i, R_j, P_5) \quad (12)$$

In contrast to eq 10, in eq 12 if an atom is partially buried in the membrane, its contribution is taken into account but with a volume corresponding only to the outside part of the atomic sphere and expressed as a simple continuous analytical function $V_j(Z_j)$ of the distance from membrane midplane. In the final GBIM version of the CHARMm module Genborn, eq 10 was replaced with eq 12 and the new analytical derivatives of the Born radii were distributed in the code as well. After incorporation of eq 12, the CHARMm tests show a practically zero difference between the analytical and numerical gradients, obtained for large sets of polar atoms from the bacteriorhodopsin structure.

SA/IM Calculations of Nonpolar Contribution to Solvation Energy. On the basis of simple geometrical considerations, we modified the subroutine Anarea[28] in the CHARMm module Aspenr[18] to include the effect of the membrane on the solvent-accessible surface of the solute. In the present study we used Aspenr for calculations of ΔG_{np} according to eq 2 where the nonpolar term is related only to the total solvent-accessible area. However, the extended Aspenr module, without any limitations, can also be used for energy calculations based on individual atomic solvation parameters.[8,18,46] One of the problems in the implementation of the GBSA/IM approach was the choice of numerical value of the σ parameter. The MD simulations of a globular protein[10] show that the molecular conformation is not too sensitive to the surface-dependent term, which, in practice, can be omitted. However, when the protein interacts with the membrane, the nonpolar interactions must be present in the model, because they form the force attracting the molecule to the membrane interior. In all calculations presented below we used $\sigma = 28$ cal/mol·Å^2, taken from a similar solvation model based on FDPB/SA.[21] This value is in the range of the experimental hydrocarbon−water transfer energies.[29]

3. Methods

Structural Models. The protein model of bacteriorhodopsin was built on the basis of the 1.55 Å X-ray structure,[30] and the rhodopsin model was based on the 2.8 Å structure[31] published recently. The coordinates of the heavy atoms that are missing in the X-ray crystallographic files were taken from ref 32 for bacteriorhodopsin and constructed for rhodopsin with the InsightII loop-building facility. The coordinates of all polar hydrogen atoms were generated by the CHARMm HBUILD routine and finally the structures were relaxed by ABNR minimization and the standard polar hydrogen param19 parameter set.[33]

The calculations for the amino terminus of the influenza virus fusion peptide are based on the sequence GLFGAIAGFIENG-WEGMIDG-amide. This represents the sequence of the first 20 amino acids of the N-terminal fragment of the HA2 influenza virus fusion peptide and is the same as the sequence of the synthetic peptide used in the EPR and NMR experiments of Zhou et al.,[20] as well as in the model study of Bechor and Ben-Tal.[21] For the purpose of energy calculations and MD simulations, we wrote several CHARMm scripts to build an initial model of the peptide in an α-helical backbone conformation and to optimize the interactions between side chains. A procedure of full search and CHARMm ABNR minimization in the conformational space of all possible Glu11, Glu15, and Asp19 rotamers was used to find a conformer corresponding to a minimum electrostatic repulsion between negatively charged groups. The structure of this conformer was additionally optimized by searching for an optimal orientation of selected neighboring nonpolar side chains. The final conformation of the peptide was suggested as a basic model for the solvation energy calculations as well as a starting conformation in MD simulations.

Solvation Energy Calculations. To calculate the GBSA/IM solvation energy terms in the presence of membrane we used our extended version of the CHARMm routines Genborn and Aspen. The param19 parameter set[33] and the corresponding protein GB parameters[9] were used in all energy calculations and MD simulations. The thickness of the membrane was set to 30 Å, a value corresponding to the thickness of the hydrophobic core[29] of lipid bilayers. For the GBIM empirical parameter in eq 8, the value of $\gamma = 0.55$ Å^{-1} was used. $\sigma = 28$ cal/mol·Å^2 was taken for the surface tension parameter.[21,34] A probe radius of 1.4 Å was used in SA calculations according to eq 2. The value of the dielectric constant for the solute and

membrane interior was set to $\epsilon_m = 2$ and to $\epsilon_{slv} = 80$ for the solvent. We used the MEAD program suite[2,3] for all energy and pK calculations based on the finite-difference Poisson−Boltzmann method. Because the ionic concentration is not taken into account, the FDPB results actually correspond to a solution of the Poisson equation. The MEAD package was found to be appropriate for the present study not only because it allows for inclusion of the membrane in the system but also because it uses the method of images for a physically correct estimation of potential on the outer boundaries of the finite-difference lattice. The computations of all electrostatic terms necessary to calculate the pH-dependent protonation of the influenza virus fusion peptide were carried out with the MEAD program[3] Multiflex. The protonation of the ionizable groups was calculated by use of the full Boltzmann statistics of the states of protonation.[35] For all structures that were studied, we did not find any anomalies in the ionization curves of the individual sites. The results presented below are, therefore, limited to the values of p$K_{1/2}$ that were obtained. The latter is defined as the pH at which the site is half-protonated. We used the standard values of 4.0 and 4.4 for model-compound pKs of Asp and Glu, respectively, and 8.0 for the N-terminal amino group of Gly.[36] The parameters of the FDPB lattices were set according to the size of the solute. After focusing, the minimal spacing between grid points varies between 0.1 Å for single ion and 0.5 Å for protein solvation energy calculations. A standard 1.4 Å radius of the spherical "water" probe was used to calculate the boundary between solute and solvent, in this case defined not by the center of the water probe, as it is used in eq 2 for SA definition, but by the solvent-accessible molecular surface.

Molecular Dynamics Simulations. To examine the suitability of GBSA/IM for molecular dynamics calculations, we studied the trajectories produced by several standard 3 ns simulations at room temperature of the HA2 fusion peptide in the presence of a 30 Å thick membrane. The starting structures were relaxed by ABNR energy minimization before the dynamics runs. The simulations were carried out according to a standard protocol including heating, 40 ps equilibration, and a 3 ns production run. The solvation energy term, according to eq 1, was added to the potential of mean force. The Coulombic term for electrostatic interactions as well as the GB polarization energy term were calculated with param19 partial charges and $\epsilon_m = 2$ for the molecular dielectric constant. $\sigma = 28$ cal/mol·Å^2 was used as above for the parameter of nonpolar energy in eq 2. A time step of 1 fs was used in all simulations, as was the SHAKE algorithm[37] to avoid the vibrations in bond distances between hydrogen and heavy atoms. Because of the relatively small size of the peptide, the noncovalent interactions were calculated without the use of any cutoff distances. In all simulations the peptide was assumed to be in a standard state of ionization characterized by a protonated N-terminus and deprotonated Glu and Asp residues. We would like to note that, in the real interactions of peptides with membranes, even at normal pH, conformational changes can be accompanied by changes of ionization state and vice versa. Modeling such a process of conjugated chemical and structural changes is not easy and it is beyond the methods of standard molecular dynamics. However, it will be shown below that the "all-groups-ionized" approximation used at normal pH is consistent with the results for the protonation of acidic and basic groups obtained for the final MD conformations.

Parametrization of GBIM Model. Initially, $\gamma = 0.72$ Å^{-1} was obtained by searching for minimum rms deviation between GBIM and FDPB polarization energies (Figure 2) for a single 2 Å ion in the presence of a 30 Å dielectric slab with $\epsilon_m = 1$. A slightly different $\gamma = 0.92$ Å^{-1} was found to be optimal for an ion of smaller, 1 Å, radius. This indicates that instead of a single-value γ, a set of parameters dependent on the atomic type might be more appropriate in order to minimize the differences between FDPB and GB results. However, we found in all test calculations of the molecular polarization energies that a single γ parameter seems to be enough to keep the GBIM error within the same few percent margin, as in the tests of the linear pairwise approximation[9] for proteins in water. Therefore, at this stage of the development of the method, we used the simpler but computationally more efficient variant where only one single-value γ parameter is added to the standard set of GB parameters.

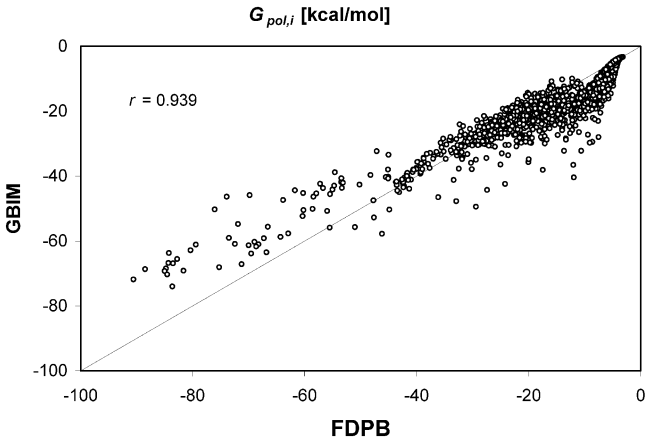

Figure 3. Comparison of FDPB and GBIM atomic polarization energies $G_{pol,i}$ in the bacteriorhodopsin structure computed in the presence of a 30 Å membrane.

To find an optimal value of γ reproducing the effective GB radii of the atoms of a macromolecular object, the individual $G_{pol,i}$ were estimated by calculating the FDPB atomic polarization energies for the bacteriorhodopsin X-ray structure.[30] Bacteriorhodopsin was chosen because it is a typical transmembrane protein having a high-resolution structure with a statistically relevant number of membrane-embedded and solvent-exposed atoms. The molecule was centered in a 30 Å low-dielectric slab and oriented such that the principal axis of inertia was perpendicular to the membrane planes. The membrane thickness and the molecular dielectric constant $\epsilon_m = 2$ were chosen to be close to the properties of the central hydrophobic region of biological membranes.[29] A unit positive charge was placed on each of the 2087 atoms consecutively, while the rest of the molecule was kept neutral. The FDPB $G_{pol,i}$ were calculated with the MEAD program suite.[2,3] The results were obtained by focusing on the charge center with a final grid size of 0.25 Å. The best fit between GBIM and FDPB atomic energies was found at $\gamma = 0.55$ Å^{-1} with an rms deviation of 4.9 kcal/mol. Note that the estimated rms deviation in atomic polarization energies of bacteriorhodopsin atoms in the presence of membrane is slightly better than the reported average 11.3 kcal/mol rmsd for proteins in water.[9] The comparison between FDPB and GBIM $G_{pol,i}$ is shown in Figure 3. The calculated correlation coefficient ($r = 0.94$) is similar to the value estimated for globular proteins[8] ($r = 0.92$). Thus the $G_{pol,i}$ results suggest that the inclusion of the membrane does not decrease the accuracy of calculation of the effective Born radii in comparison to the standard GB method. Almost all significant differences between FDPB and GBIM values are between $G_{pol,i}$ of hydrogen atoms, because of the complex effect of the small atomic radius on the accuracy of both FDPB and GB calculations.

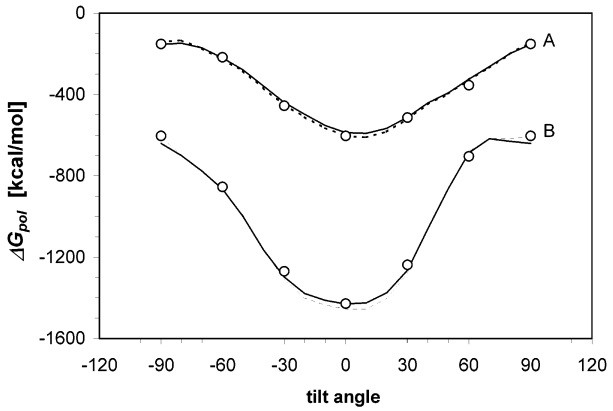

Figure 4. Molecular polarization energies ΔG_{pol} calculated at different tilt angles relative to the membrane normal for bacteriorhodopsin (A) and rhodopsin (B). (○) FDPB values; lines represent GBIM values at (—) $\gamma = 0.55$ and (---) $\gamma = 0.72$.

Finally, we studied the accuracy of the GBIM method in calculations of total molecular polarization energies. For that purpose the GBIM and FDPB ΔG_{pol} values were computed for different orientations of bacteriorhodopsin and rhodopsin in the membrane. These proteins were chosen as representatives of the wide class of seven-helical transmembrane proteins. To create the set of test structures, the molecules were initially oriented with the first axis of inertia collinear with the membrane normal and after that tilted to the membrane plane by stepwise rotation around the second axis of inertia. The tilt angle is defined as the angle between the membrane normal and the molecular principal axis of inertia. Such an approach makes it possible to test the method on a number of structures having quite different ratios of atoms buried inside the membrane to atoms exposed to water. In Figure 4 it can be seen that the molecular GBIM polarization energy terms are in a very close agreement with the FDPB results. The absolute errors for both bacteriorhodopsin and rhodopsin data sets are, on average, 2.5% for $\gamma = 0.55$ and 2.6% for $\gamma = 0.72$ Å^{-1}. It reflects our observation (data not shown) that the calculated molecular polarization energies are not too sensitive to the value of the parameter γ in the interval 0.5−0.8 Å^{-1}. Therefore, for the calculations in this study, the lower value of the parameter, $\gamma = 0.55$ Å^{-1}, was chosen because, while reproducing the FDPB results well, it models a less sharp membrane interface, i.e., closer to the real case.[38]

The results of all tests show that the inclusion of the membrane in GB calculations does not decrease but even slightly improves the accuracy of the calculations in comparison to the overall accuracy of the param19 GB parametrization.[9] It makes us believe that the proposed GBIM method is a reasonable approximation of the effect of the membrane on the solvation energies. One of the advantages of the GB approximation, besides that it makes the programs significantly faster than the grid-based Poisson−Boltzmann solvers, is that the computer time and memory limits are dependent only on the number of atoms, not on the geometrical dimensions of the system. In the GBIM approximation the computations are even faster than for the same molecule surrounded only by the solvent because the presence of the membrane reduces the number of elements in the sums in eq 12. In the limiting case of all atoms inside the membrane, eq 12 will include only the Γ term. In comparison, the inclusion of the membrane increases the computer time for the FDPB solvers.

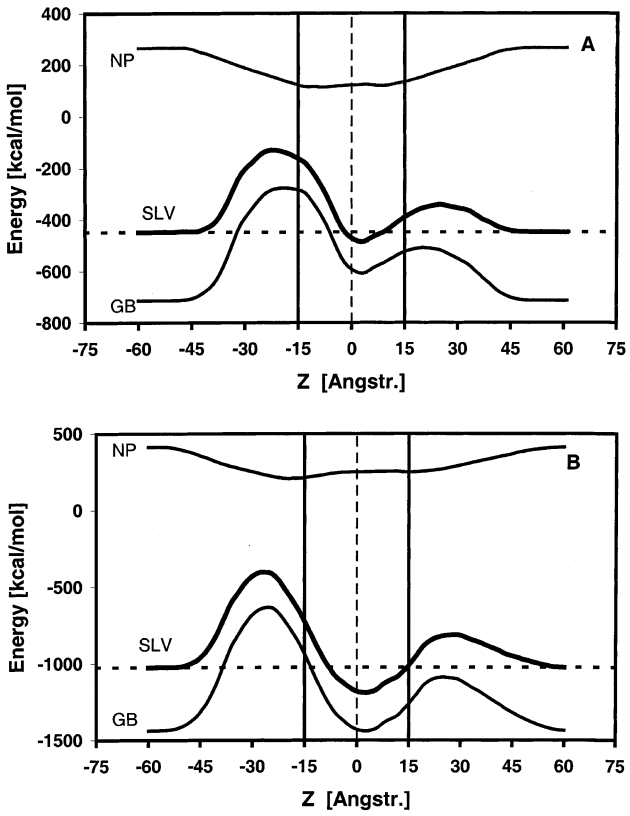

Figure 5. Solvation energy terms in eq 1, ΔG_{slv} (SLV), ΔG_{pol} (GB), and ΔG_{np} (NP), as a function of the distance Z of the protein center of mass from the center of the membrane, computed for the structural models of bacteriorhodopsin (A) and rhodopsin (B). Positive distances correspond to movement into the cytoplasmic space, as defined by the orientation of the proteins in cells. The vertical lines represent the membrane boundaries. The dotted line corresponds to ΔG_{slv} at infinite separation of the protein from membrane.

4. Results and Discussion

Bacteriorhodopsin and Rhodopsin. In one of the first tests of the GBIM method, we calculated the changes in the solvation energy obtained by moving the bacteriorhodopsin and rhodopsin structures across a 30 Å dielectric slab. For comparison, the same calculations with a conventional FDPB program at a moderate 0.5 Å grid size are only possible if a large box of about 300^3 grid points is used. In this test the protein orientation, defined by the principal axis of inertia, was kept strongly perpendicular to the membrane plane. The changes in solvation energy computed as a function of the distance of the protein center of mass from the center of the membrane are shown in Figure 5. The maximum lengths of bacteriorhodopsin and rhodopsin, along the direction of their principal axes, are ~60 and 70 Å, respectively.

A very interesting feature is seen in both the bacteriorhodopsin and rhodopsin curves. The position dependence of the solvation energy is strongly asymmetric across the membrane, showing a significantly higher barrier if the protein crosses the membrane into the extracellular rather than into the cytoplasmic space. The minimum of the solvation energy for bacteriorhodopsin corresponds to a position of the protein center shifted about 3 Å from the center of the membrane toward the cytoplasmic side. A similar, 2 Å, shift is seen in the rhodopsin curve. It is remarkable that a slightly larger, but in exactly the same direction, 5 Å displacement of the protein center from the middle of the lipid bilayer is estimated from the analysis of bacteriorhodopsin X-ray data by Luecke et al.[30] The asymmetry

in the ΔG_{slv} curve explains well the experimental observation[30] that the cytoplasmic portion of bacteriorhodopsin is less embedded in the hydrophobic bilayer than the extracellular portion. This indicates that the sinking of the cytoplasmic polar cap has a higher energy penalty than that of the extracellular cap. It is tempting to speculate about the possible biological reasons for such an asymmetry. Luecke et al. suggested that the observed shift of the bacteriorhodopsin center may be related to the fact that the large-scale conformational changes occur in the cytoplasmic portion of the protein. On the basis of the less steep slope of ΔG_{slv} in Figure 5, it can be suggested that not only the static-shifted position of the protein but also the less restricted mobility of the entire protein in the direction of the cytoplasm may be relevant to the structural reorganization in the cytoplasmic part during the photochemical cycle. A possible explanation of the source of the asymmetrical shape of ΔG_{pol} is that the solvation forces are used as a factor during evolution to design the structure in a way that prevents the protein from crossing the membrane in the extracellular direction and that prevents loss of protein mass from the cell. In structural terms, the asymmetry in ΔG_{pol} is a function of the imbalance in the spatial distribution of charged groups, i.e., the structures of both proteins have many more charged residues in the cytoplasmic region than in the extracellular region. This imbalance is caused almost entirely by the distribution of the positively charged Lys and Arg residues. It reflects the well-known "positive-inside" rule, which is valid for most of the transmembrane proteins.[39] Because of the considerable difference in ΔG_{pol} barriers, it could be suggested that the solvation forces must be an important, even major factor for the directional orientation of proteins in membranes, restricting more strongly the translocation of the polar cap that has more charged groups. If so, the question is still open as to why the charge imbalance is caused only by the imbalance of positively charged groups, i.e., why the rule is "positive inside" and not "charged inside". One can speculate that the cationic residues are chosen during evolution because their longer side chains provide more potential to interact with the solvent than do those of anionic residues. Such a solvation energy-based mechanism does not exclude the role of electrostatic interactions between cationic Arg and Lys residues and the anionic phospholipids in the directional orientation of membrane proteins[40] but can be complementary to it.

Figure 5 sheds some light on the general conditions for a seven-helical protein to be incorporated into the membrane $\Delta\Delta G_{np}$ between the configuration with minimum ΔG_{slv} and the configuration in which the protein is fully outside the membrane is -143.1 kcal/mol for bacteriorhodopsin and -162.4 kcal/mol for rhodopsin. The corresponding values for $\Delta\Delta G_{pol}$ are $+105.0$ and $+3.0$ kcal/mol, respectively.

This demonstrates that nonpolar interactions are the major force that attracts the protein into the membrane, overcoming the penalty of polar interactions around the minimum of solvation energy. However, the nonpolar interactions are not too sensitive to the exact position of the protein. As expected, the two flat minima of ΔG_{np} are at positions where the membrane crosses not the hydrophobic seven-helical domain but more developed polar caps. In contrast, ΔG_{pol} has a much sharper minimum, giving rise to the global minimum of the solvation energy. Thus, if one wanted to manipulate the affinity or the orientation of the protein in the membrane, the most promising targets for design would be the polar and ionizable groups, which are situated close to the protein surface.

A comprehensive search[21] of the solvation energy minimum in the space of rotational and translational degrees of freedom was used to locate more precisely the optimal position of the two proteins in the membrane. The energy was calculated for the grid points representing the distances between the protein center and membrane midplane, as well as the rotation angle around the first axis of inertia and the tilt angle with respect to the membrane normal. The final step sizes were 2° for the rotation and tilt angles and 0.5 Å for the translation. The energy minimum that was located for each protein corresponds to an almost vertical orientation with a tilt angle of 4° for bacteriorhodopsin and 8° for rhodopsin. The energy difference between the optimally tilted and vertical positions is negligible for bacteriorhodopsin ($\sim$1 kcal/mol), while the slightly tilted optimal orientation of rhodopsin is preferred by $\sim$10 kcal/mol.

Influenza Hemagglutinin Fusion Peptide. The influenza virus hemagglutinin HA2 fusion peptide, which consists of the first 20−25 residues located at the N-terminus of the HA2 subunit, is a widely studied structure because of its key role in the process of membrane fusion between viral and host cell membranes.[41] Besides the biological importance, two further reasons motivated us to present the results of our study of the 20-residue N-terminal fragment of this peptide by the GBIM method. First, recent EPR and NMR experimental data[20,42] gives the conformation and orientation of the N-terminal fragment of the peptide with respect to the membrane, which can be used as a basis for verification of the theoretical approach. Second, the energetics of the HA2 N-terminal peptide in a low-dielectric layer have been studied in detail by Bechor and Ben-Tal[21] by the FDPB technique. This provides the opportunity to check the modeling potential of the GBSA/IM method by comparison to the alternative, grid-based FDPB technique.

The distribution of polar and hydrophobic groups along the sequence suggests an amphiphilic helical conformation of the first 20−25 N-terminal residues of the HA2 peptide.[43] Indeed, several spectroscopic studies[20,42,44,45] not only show a helical conformation of the N-terminal peptide but also suggest a tilted orientation of the helical axis when the peptide is bound to a membrane. Although a large number of synthetic analogues of the HA2 peptide have been studied, our analysis is limited to the 20 amino acid structure (for the sequence, see Methods) that was investigated recently in a combined EPR and NMR study by Zhou et al.[20] The measured immersion depths of a nitroxide EPR probe attached to several Cys mutants correspond to a helical conformation of the central 5−14-residue fragment and an almost pH-independent tilt angle of 25° (pH 7) and 28° (pH 5) (where the tilt leads to the N-terminal end of the peptide being buried deeper in the membrane). Bechor and Ben-Tal[21] investigated the possibility that the observed tilt angle was due to the effects of desolvation. The continuum solvation model used by Bechor and Ben-Tal is the same as ours, but the polarization term is computed by a modification of the FDPB program DelPhi.[47] The authors calculated the solvation energy terms by using a scheme of stepwise rotations and translations to find the optimal orientation of the peptide in the membrane. We adopted their scheme[21] for a preliminary search of the optimal orientation of the peptide, but instead of a FDPB program, the solvation energy was calculated by CHARMm with the new GBSA/IM method. Table 1 presents the GBSA/IM results for the position that corresponds to the minimum of the solvation energy obtained for the basic HA2 α-helical structure (see Methods), as well as the corresponding FDPB data.[21]

Similarly to the FDPB study,[21] the GBIM/SA minimum of solvation energy is found at a position and orientation in the membrane with most of the polar and charged groups exposed to water. However, the GBIM minimum corresponds to a

TABLE 1: Geometry and Energy Parameters Obtained for Helical Models of the N-Terminal HA2 Peptide in a Position Corresponding to the Minimum of Solvation Energy[a]

method and model	Z_{cntr} (Å)	tilt angle (deg)	$\Delta\Delta G_{pol}$	$\Delta\Delta G_{np}$	$\Delta\Delta G_{slv}$
FDPB/SA[b]	17	3	8.0	−19.0	−11.0
GBSA/IM	15	5	8.7	−35.5	−26.8

[a] Z_{cntr} is the distance of the center of the helix from the membrane midplane; the tilt angle is the angle between the helix axis and the membrane plane (a positive angle corresponds to the N-terminus being tilted down into the membrane). The energy terms (in kilocalories per mole) represent the differences between the values calculated in the presence of membrane and in water. [b] Reference 21.

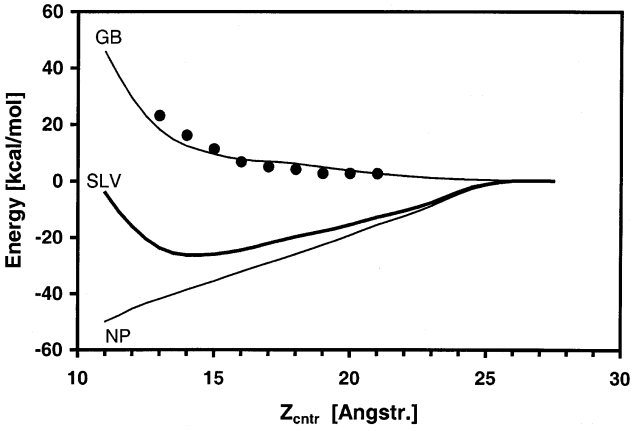

Figure 6. Solvation energy of binding, $\Delta\Delta G_{slv}(Z_{cntr}) = \Delta G_{slv}(Z_{cntr}) - \Delta G_{slv}(\infty)$, calculated as a function of the distance, Z_{cntr}, of the center of the helix from the center of the 30 Å membrane. The results are obtained for an orientation with the helical axis parallel to the membrane plane and the nonpolar region facing the membrane center. SLV, differences, in the total solvation energy, according to eq 1; GB, in generalized Born energies; NP, in the nonpolar contribution. (○) $\Delta\Delta G_{pol}$ energies computed by FDPB method.

position of the helix center situated exactly in the boundary plane, while according to the Bechor and Ben-Tal calculations,[21] the helix is shifted up 2 Å, exposing most of the molecule to the solvent. It is difficult to identify the main source of this difference—the computational method used, the differences in the side-chain conformations of the helical structures, or the different sets of atomic parameters. To investigate this we carried out a new set of FDPB calculations using the MEAD program suite[2] and param19 CHARMM parametrization of atomic radii and atomic partial charges instead of DelPhi and the PARSE charge set. The results of the MEAD calculations are compared to the corresponding GBIM calculations in Figure 6, where the solvation energy of binding to membrane is shown as a function of the distance of the helical center from the center of the membrane. As in Figure 4, it can be seen that the GBIM results are very close to the MEAD FDPB values. However, the polarization energy curves computed by both the GBIM and MEAD programs are almost two times less steep compared to those computed with DelPhi as shown in Figure 2 of Bechor and Ben-Tal's paper.[21] Thus, it can be suggested that the observed disagreement with the Ben-Tal calculations is not caused by the GBIM approximation but most probably reflects the differences in the parameter sets used or in the peptide conformation and orientation.

It is interesting to note that both methods predict almost the same tilt angle (Table 1), corresponding to a horizontal orientation of the α-helical construct. We repeated the calculations (data not shown) for a number of randomly selected rotamers of the peptide side chains, but in contrast to the experimental data, the optimal tilt angle remained close to 0°. Thus, the search procedure employing either computational technique fails to reproduce the experimentally observed tilted orientation of the peptide. A possible reason for this could be that the fixed geometry of the α-helical peptide makes a high polar penalty unavoidable if the charged N-terminal NH_3 group sinks into the low-dielectric slab. This motivated us to use the computational advantage of the GBIM approximation and to extend the study of the HA2 N-terminal peptide by using the method of molecular dynamics. We present below the results of five 3-ns standard molecular dynamics runs. As the starting structure for all simulations, the same α-helical model of the fusion peptide was used with different initial positions relative to the membrane. In the first run, which we will call the RO run, the helical structure was positioned at the "optimal" position in the membrane, with the tilt angle indicated in Table 1, the helical axis lying in the membrane plane and rotated by the same angle around the axis of the helix as obtained before for the energy minimum. In this orientation the polar face of the amphiphilic helical structure is entirely exposed to the solvent while most of the nonpolar groups are buried in the low-dielectric slab. The initial structure (0 ps) and the conformation at selected times along the MD trajectory are shown in Figure 7A.

Figure 8A shows the fluctuations of the root-mean-square deviation (RMSD) of the atomic coordinates obtained relative to the starting conformation during the 3 ns RO run. Figure 9 shows the time courses of both interaction and solvation electrostatic energy contributions to the potential of mean force. The results shown in both figures imply that near 1.5 ns of the simulation the molecule changes conformation. Indeed, the inspection of the peptide structures along the dynamics trajectory shows that during the first ∼1500 ps the structure fluctuates around an α-helical conformation of the entire 20 amino acid peptide. After 1500 ps the first turn of the helix, comprising 3−4 residues, unfolds and stretches, exposing more of the N-terminus to the solvent.

It is interesting that, at the same moment the peptide conformation changes, the orientation in the membrane changes too. This can be seen from the time dependence of the tilt angle in Figure 10 calculated for the residues 4−20 fragment. It is remarkable that in the new ensemble the tilt angle fluctuates around a value that is very close to the experimental tilt angle of 25° reported in the EPR studies.[20,42]

The time course of the individual energy terms shown in Figure 9 suggests that the change of the conformational ensemble is related mainly to some rearrangement of the network of electrostatic interactions reflected in the change of Coulombic energy of interatomic interactions. Figure 9 also demonstrates the high level of anticorrelation between the Coulomb and polarization energy electrostatic terms, making the total electrostatic contribution to the potential of mean force fluctuate around an almost constant value. Interestingly, the change of the tilt angle (Figure 10) is not accompanied by any visible change in the surface-dependent nonpolar term, which after the equilibration, on average, also stays constant.

To examine the extent to which the dynamics trajectory depends on the initial position of the peptide, we carried out three more 3 ns long simulations. In the first, RN, we put the helix in a horizontal position, similar to the RO run, but rotated the structure by 120° around the helical axis. The center of mass was shifted by 3 Å from the membrane plane in the direction of the solvent. In this position, obviously nonoptimal for both

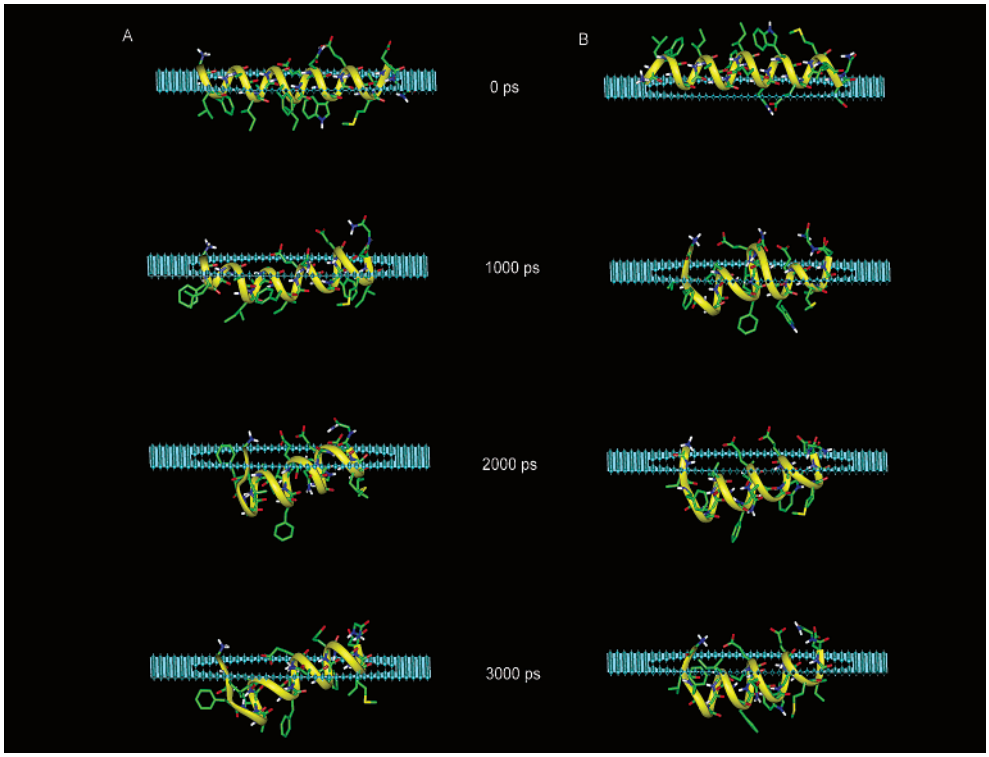

Figure 7. Changes in the conformation of the HA2 N-terminal peptide during a 3 ns molecular dynamics simulations. The snapshots are taken from the trajectories of (A) RO simulation and (B) RN simulation. The N-terminus of the peptide is at the left-hand end in these images. The blue plane represents the membrane surface, with the region above the plane being solvent.

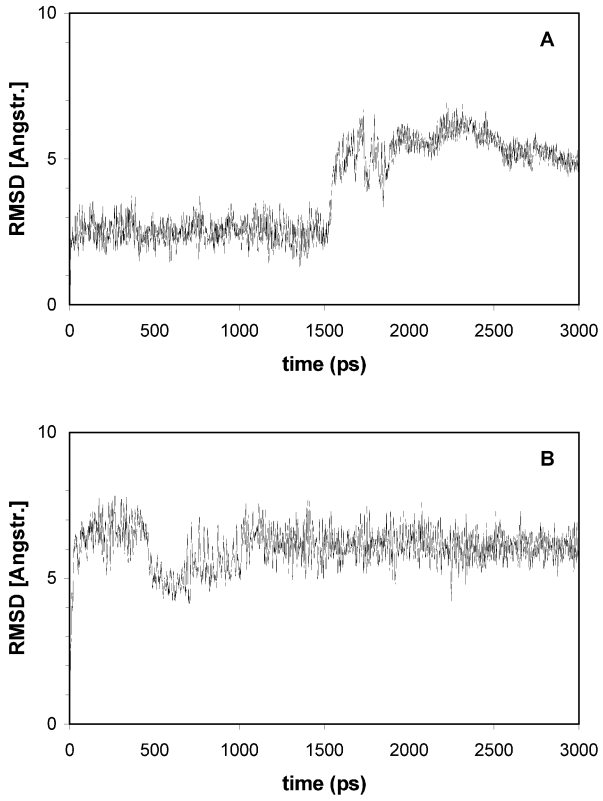

Figure 8. Time dependence of the root-mean-square deviation of atomic coordinates relative to the coordinates of starting structure: (A) data corresponding to the RO simulation; (B) RN simulation.

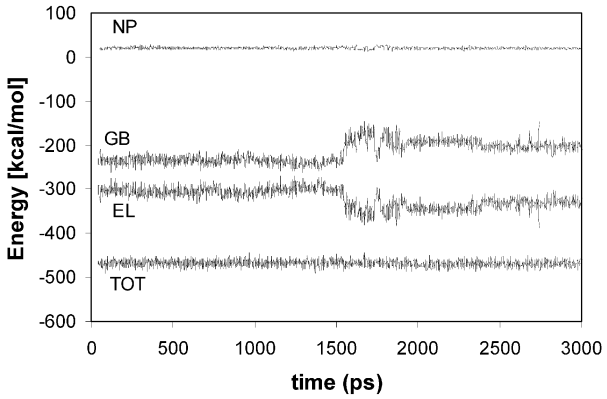

Figure 9. Time dependence of electrostatic and nonpolar contributions to the potential of mean force during the RO simulation: EL, interatomic Coulombic electrostatic term; GB, polarization energy term; TOT, total electrostatic contribution calculated as the sum of EL and GB; NP, nonpolar term.

polar and nonpolar interactions (Figure 7B, 0 ps), the side chains of all nonpolar groups are exposed to the solvent, while the charged and polar residues are either immersed into the membrane interior (Asn12, Glu15, Asp19) or situated close to the boundary (N-terminal NH_3 group, Glu11). In contrast to the RO run, during the first 1 ns, the evolution of RMSD fluctuations suggest a transition of both orientation and conformational state (Figure 8B). However, after ~1−1.2 ns the RN trajectory becomes even more stable than that of the RO run. Inspection of the structures shows that the peptide rotates along the long axis of inertia and very quickly, during the first 10 ps of equilibration, extracts the buried charged groups from the low-dielectric environment and exposes them to the water solvent. In Figure 11A the changes of the distances between the Glu11, Glu15, and Asp19 charged groups and the membrane plane are shown.

During this reorientation process the helical structure partially unfolds. At the end of the 1 ns period, however, most of the peptide (residues 4−20) folds back to a helical conformation that does not change until the end of the run. It is notable that,

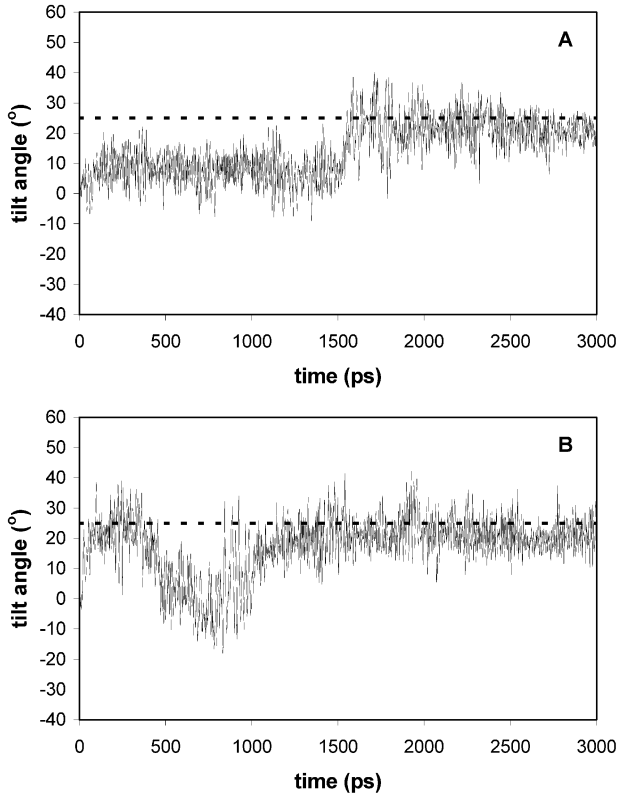

Figure 10. Time dependence of the tilt angle of the helical fragment of residues 4−20 of the HA2 fusion peptide calculated relative to the membrane plane. The dotted line corresponds to the experimental value[14] of the tilt angle at pH 7. (A) RO simulation; (B) RN simulation.

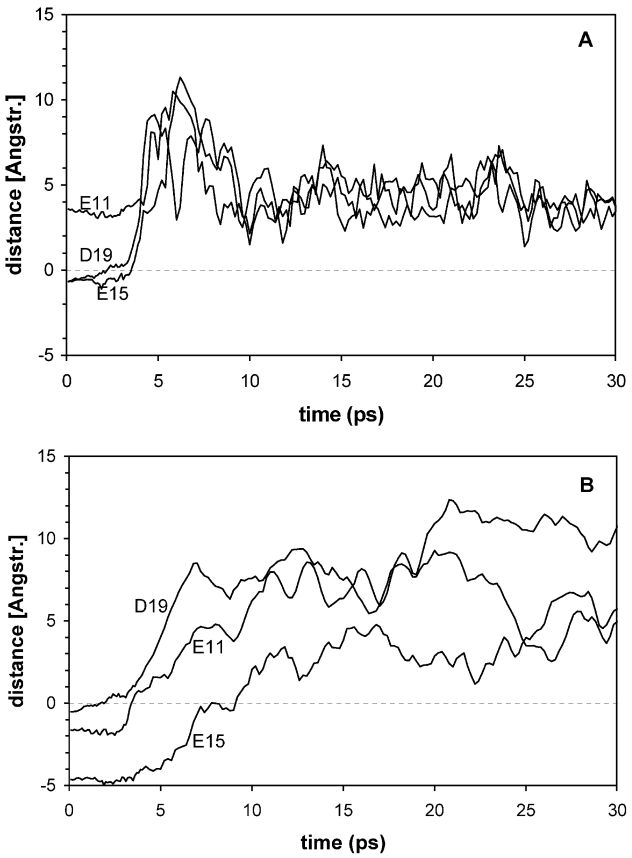

Figure 11. Time dependence of the distance between the centers of charged groups and the membrane plane: (A) RN simulation; (B) RN2 simulation.

during the stationary 1−3 ns period, the tilt angle of the 4−20 helix of the peptide fluctuates around the same value as in the RO run, which is very close to the experimental value.

In conclusion, both the RO and RN molecular dynamics simulations reproduce almost perfectly the tilt angle derived from the EPR experiments of Zhou et al.[20] and Mackosko et al.[42] Similarly to the EPR data[20] for this amphiphilic peptide, in the final RO and RN structures, the hydrophobic residues Phe3, Ile6, Phe9, Ile10, Trp14, and Ile18 are buried in the low-dielectric slab. The carboxyl groups of Glu11, Glu15, and Asp19 are mostly exposed to the solvent, which suggests that the solvation forces applied to these negatively charged groups are the major factor for the tilted orientation of the peptide. However, during both MD runs a short N-terminal fragment of three residues, Gly1−Phe3, adopts a nonhelical conformation that allows the charged amino terminus to stay in the water environment even when the nonpolar part of the helix is inserted into the membrane. This would be impossible if the peptide was in a strongly helical conformation. The nonhelical conformation of Gly1−Phe3 is not inconsistent with the experimental results, because in fact the EPR data[20,42] are obtained only for the residues of the Ala5−Trp14 fragment.

In the starting orientation for the next simulation, RN2, the helical peptide was also situated in a horizontal position but with the polar face inserted 4 Å deeper into the membrane than in the RN run. In another simulation, RV, the helical peptide was initially inserted in a vertical orientation into the dielectric slab with the C-terminal group close to the center of the membrane. The common feature between the results of these and the previous simulations is that all buried polar and charged residues become exposed to the solvent relatively quickly (∼10 ps), as can be seen in Figure 11B for the RN2 run. However, in both the RV and RN2 simulations, during the process of extracting the polar groups from the membrane interior, the helix unfolds from both ends and, at the end of the 3 ns trajectories, the peptide adopts an almost entirely nonhelical conformation. The final conformations for the RN2 and RV runs are shown in Figure 12.

The final conformation of a 3 ns simulation (RW) that was carried out without membrane, i.e., as if the peptide was immersed in water, is also shown in Figure 12. This simulation started from exactly the same conformation of an ideal helix. At the end of the trajectory, the peptide adopts a fold characterized by a short β-strand structure of Gly1−Ala5, interacting with a parallel helical fragment, Gly8−Ile18. Despite the small molecular mass of the peptide, in the final RW structure some properties, which are common for the native conformations of globular proteins, can be seen. The RW structure has compact packing; the nonpolar side chains of Leu2, Phe9, and Ile6 make a small nucleation and all polar and charged side chains are exposed to the solvent.

The comparison of the results of the RV, RN2, and RW runs with the RO and RN simulations suggests that a helical conformation of most of the residues of this amphiphilic peptide is stable only if the helix is situated close to the membrane interface as in the RO and RN runs. Moreover, at the end of the RO and RN runs the helical portion of the peptide adopts an oblique orientation very close to experiment.

However, forcing the polar groups deeper inside the low-dielectric environment results in an unfolding of the helical conformation, which has not returned by the end of the 3 ns simulations. If the membrane is removed from the system, as in the RW simulation, almost half of the helix unfolds, but an ∼10 amino acid fragment fluctuates around a helixlike conformation.

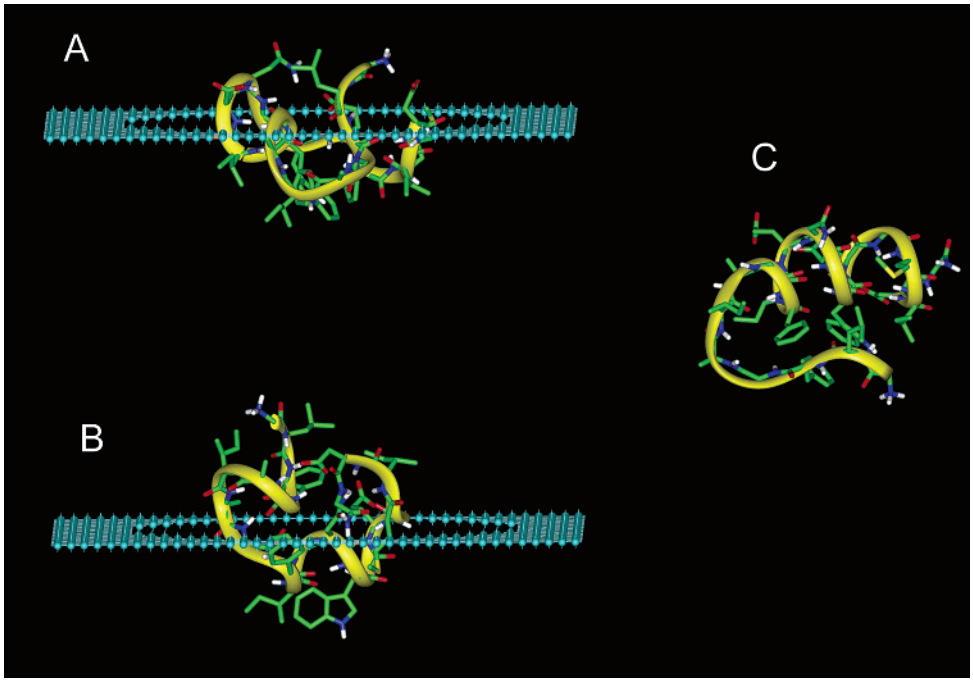

Figure 12. Final conformations of the N-terminal HA2 fusion peptide adopted after 3 ns molecular dynamics simulations: (A) RN2 run; (B) RV; (C) RW.

TABLE 2: p$K_{1/2}$ of the Ionizable Groups of the N-Terminal HA-2 Fusion Peptide Calculated for Structures Taken from the MD Trajectories

	p$K_{1/2}$			
structure	NH_2Gly1	Glu11	Glu15	Asp19
RO (0 ns)	7.00	6.44	5.99	4.34
RO (1 ns)	7.20	5.06	6.32	5.12
RO (3 ns)	8.67	3.51	6.62	4.38
RN (0 ns)	4.70	6.14	11.2	13.1
RN (1 ns)	8.94	−2.9	7.50	3.8
RN (3 ns)	8.56	2.81	4.70	4.73
RW (0 ns)	7.14	6.16	5.83	4.14
RW (3 ns)	7.49	5.60	5.73	3.95
experiment[a]	8.69			

[a] Reference 20.

In the study of Zhou et al.[20] the protonation of the α-amino group of Gly1 was measured as a function of pH. On the basis of the NMR data the authors reported a pK_a = 8.69 for the amino terminus, which is higher than the normal value[36] of ∼8.0. This observation motivated us to study theoretically the protonation of the ionizable groups of the peptide for several structures selected from the MD trajectories. The calculations were performed entirely in the framework of the macroscopic electrostatic model of pH-dependent electrostatic effects with the solute treated at full atomic detail in the presence of the membrane, as proposed by Bashford and Gerwert.[2] The details of the calculations are given under Methods. In Table 2 the calculated p$K_{1/2}$ values for the individual sites of protonation of the peptide, for several structures selected from the MD trajectories, are presented. All pK calculations were carried out after preliminary energy minimization of the MD structures.

In all 100% helical structures, such as RO (0 ns), RN (0 ns), and RW (0 ns), irrespective of the environment, the pK of the N-terminal amino group is calculated to be lower than 8.0, which is normally observed for the N-terminal Gly amino group.[36] The reason for this negative shift is complex, but besides the desolvation effect, a significant contribution comes from interactions with the peptide backbone because of the positive potential near the N-terminal ends of α-helical structures generated by peptide backbone dipoles.[48,49] A similar negative shift is found if the peptide is in a nonhelical conformation but entirely in a water environment, i.e., RW (3 ns).

In contrast to the previous results, the final MD structures of RO and RN show a positive shift, which is in almost perfect numerical agreement with the experimental pK value obtained for the Gly1 NH_2 group. The calculated p$K_{1/2}$ values of acidic and basic groups also suggest that the standard state of protonation, at normal pH, is reasonable for all final structures of the MD simulations. However, in some initial conformations, the calculated equilibrium is shifted to nonionized species because of the polar penalty the charged groups have to pay to enter the low-dielectric slab. For example, in the RN structure at 0 ps the p$K_{1/2}$ values of Glu15 and Asp19 are calculated to be far above 7, while the p$K_{1/2}$ of the Gly1 amino group is below 7, which means that in this orientation the "all-groups-charged" state will be populated to a very small extent. However, this standard state becomes dominant during the simulation as shown by the computed pKs for the RN 3 ns structure given in Table 2.

It was also shown above that, starting from different initial orientations during both the RO and RN runs, the peptide adopts a tilted orientation with respect to the membrane plane that is consistent with the EPR experiments. The final structures of both runs are characterized by a dynamically stable helical fold of the 4−20 fragment and differ only in the first 2−3 N-terminal residues, which adopt a nonhelical conformation. The very good agreement with the two independent sets of experimental data[20,42] makes us believe that the RO and RN simulations correctly predict the conformation of the first 20 residues of N-terminal HA2 influenza virus fusion domain. Note that the peptide adopts the tilted orientation only after minor, but important, changes of secondary structure of the first three residues that allow the terminal group to be shifted from the boundary and immersed in the water, keeping the helical shape for the rest of the peptide. We stress that this result was the product of the combined use of the MD technique and the new GBSA/IM method. The search procedure, based only on static

energy calculations, as used in our and Bechor and Ben-Tal[21] studies, failed to predict the correct tilt angle, mainly because it does not take into account the flexibility of the backbone. A more comprehensive search for the energy minimum in the space of both spatial and conformational degrees of freedom of even a short 20 amino acid peptide would face a serious combinatorial problem.

5. Conclusions

The initial goal of this work was to develop a fast and convenient computational method for modeling the interactions of protein molecules with lipid bilayers. For this purpose we extended the standard GBSA approach by implicitly including the membrane in the generalized Born approximation of polarization free energy as well as in the calculations of the surface-dependent term for nonpolar interactions. The empirical function that is used to modify the expression for the effective Born radii allows the gradients of solvation energy to be derived analytically, which makes the GBSA/IM method applicable to energy minimization and molecular dynamics simulations. We successfully implemented the GBSA/IM approach in CHARMm by modifying the modules Genborn and Aspenr. In all accuracy tests carried out on several protein and peptide examples (Figures 4 and 6) we found quite small (less than 5%) differences between the solvation energy values calculated by the GB/IM and FDPB methods. Thus, the accuracy of the GB/IM method is comparable to the reported accuracy of the original GB version of Dominy and Brooks[9] that we used as a basis for the implementation of the GBIM approximation. We also found that the presence of a membrane not only does not decrease the speed of calculations but actually makes them slightly faster than the standard GB implementations.

The presence of a membrane in solvent models makes the solvation forces affect not only the conformation but also the position and orientation of the molecule in the space.

We found that the GBSA/IM method offers a convenient way to study this dependence and avoids some of the principal limitations of the alternative grid-based solvers of the Poisson equation. Although PB models have been used successfully to study the effect of polar interactions on the association of peptides to membranes,[51,52,56] they are not effective for very large systems or MD simulations.

We tested the method for its applicability to solvation energy calculations of large molecular systems using two examples of transmembrane proteins, bacteriorhodopsin and rhodopsin. The results for these proteins show that the electrostatic solvation forces are the major factor responsible for the exact position of the protein, while the nonpolar interactions act in a much less specific way. This feature makes the polar and ionizable residues more promising targets for protein design than the nonpolar residues if one wants to modify the integration of the protein into membrane. An interesting result of the calculations is the clear asymmetry in polarization energy barriers when the proteins are moved in opposite directions across the membrane. On this basis we stress the importance of the solvation forces, which may be a complementary factor in the mechanism of the well-known "positive-inside" rule,[39] if not the major factor.

Using the example of the influenza virus HA2 fusion peptide, we tested the GBSA/IM method for molecular dynamics simulations on a nanosecond scale. For this purpose we added the GBSA/IM solvation energy term to the standard CHARMM param19 force field. The fastest event found in the MD trajectories is the extraction of the charged groups from low-dielectric membrane interior during the first 10 ps by the solvation forces. In some cases this event dramatically changes both the peptide orientation and conformation. A stable trajectory is normally achieved after 1 ns of the run. The results also suggest that the helical conformation of this amphiphilic peptide is supported only when it lies close to the membrane interface. It is remarkable that, starting from a strictly horizontal position with respect to the membrane plane, the peptide adopts almost the same tilt angle between the helical axis and the membrane plane as is found in EPR experiments.[20,42] In a recent study Han et al.[50] reported the first experimentally determined atomic structures of HA2 fusion domain. The combined NMR and EPR results[50] were obtained on a host−guest peptide (P20H7) that has a different sequence from the HA2 peptide used in our calculations. The P20H7 has an additional fragment (H7) containing four positively charged Lys residues bound via a flexible linker to the C-terminus of the HA2 domain. Comparison of NMR[50] conformations with RO and RN as well as with the EPR[20] results show some interesting common features: in all structures the fragment of the first 10−11 residues is in a predominantly helical conformation and even shows a similar tilt angle (21° at pH 7.4, ref 50) with respect to the membrane. However, in contrast to the theoretical and experimental data of the HA2 peptide, the helical conformation in NMR structures[50] of P20H7 peptide is broken around residues 10−11 and the peptide is in a specific V-shape. The basic difference between the NMR structures[50] and RO, RN, and EPR[20,42] conformations is located at the polar C-terminal half of the HA2 peptide, where the three acidic residues of HA2 peptide are possibly interacting with the four Lys groups of the H7 fragment. Although the H7 part is most likely solubilized,[50] it is difficult to predict the extent of the influence of interaction of the four Lys residues with the three anionic groups of HA2 fragment on the conformation and protonation state of the peptide. A GBSA/IM study of the interaction of P20H7 molecule with membrane is in progress, which may provide more insight to the problem but is beyond the scope of this paper.

The suggestion that the final structures in the RO and RN runs correctly represent the conformation of the HA2 peptide and its orientation in the membrane was additionally confirmed by the calculated pH dependence of the protonation of the N-terminal amino group. It was found that the calculations were in excellent agreement with the NMR data[20] for the same peptide.

Though the continuum electrostatic models present a simplified picture of the complex interactions between biological molecules and lipid bilayers, the results of this study suggest that the GBSA/IM approximation will be a useful method for structural modeling of protein−membrane complexes.

Acknowledgment. We thank Professor Don Bashford for kindly providing us the MEAD program package and Dr. David Kitson for helpful discussions.

References and Notes

(1) Schaefer, M.; van Vlijmen, H.; Karplus, M. *Adv. Protein Chem.* **1998**, *51*, 1−57.

(2) Bashford, D.; Gerwert, K. *J. Mol. Biol.* **1992**, *224*, 473−86.

(3) Bashford, D. In *Scientific Computing in Object-Oriented Parallel Environment*; Ishikava, Y., Oldehoeft, R. R., Reynders, J. V. W., Tholburn, M., Eds.; Vol. 1343 of Lecture Notes in Computer Science, ISCOPE97; Springer: Berlin, 1997; pp 233−240.

(4) Karshikoff, A.; Spassov, V.; Cowan, S. W.; Ladenstein, R.; Schirmer, T. *J. Mol. Biol.* **1994**, *240*, 372−84.

(5) Im, W.; Beglov, D.; Roux, B. *Comput. Phys. Commun.* **1998**, *111*, 59−75.

(6) Rocchia, W.; Alexov, E.; Honig, B. *J. Phys. Chem. B* **2001**, *28*, 6507−6514.
(7) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson T. *J. Am. Chem. Soc.* **1990**, *112*, 6127−6129.
(8) Qiu, D.; Shenkin, P. S.; Hollinger, F. P.; Still, W. C. *J. Phys. Chem.* **1997**, *101*, 3005−3014.
(9) Dominy, B. N.; Brooks, C. L., III *J. Phys. Chem. B* **1999**, *103*, 3765−3773.
(10) Bursulaya, B. D.; Brooks, C. L., III *J. Phys. Chem. B* **2000**, *104*, 12378−12383.
(11) Schaefer, M.; Bartels, C.; Karplus, M. *Theor. Chem. Acc.* **1999**, *101*, 194−204.
(12) Giesen, D. J.; Storer, J. W.; Cramer, C. J.; Truhlar, D. G. *J. Am. Chem. Soc.* **1995**, *117*, 1057−1068.
(13) Onufriev, A.; Bashford, D.; Case, D. *J. Phys. Chem. B* **2000**, *104*, 3712−3720.
(14) Zou, X.; Sun, Y.; Kuntz, I. D. *J. Am. Chem. Soc*. **1999**, *121*, 8033−8043.
(15) Bashford, D.; Case, D. A. *Annu. Rev. Phys. Chem*. **2000**, *51*, 129−152.
(16) Brooks, B.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem*. **1983**, *4*, 187−217.
(17) CHARMm c27b4 is the commercial version of CHARMM c27b4 distributed by Accelrys Inc., 2001.
(18) Wesson, L.; Eisenberg, D. *Protein Sci.* **1992**, *1*, 227−235.
(19) Wiley, D. C.; Skehel, J. J. *Annu. Rev. Biochem.* **1987**, *56*, 365−394.
(20) Zhou, Z.; Macosko, J. C.; Hughes, D. W.; Sayer, B. G.; Hawes, J.; Epand, R. M. *Biophys. J.* **2000**, *78*, 2418−2425.
(21) Bechor, D.; Ben-Tal, N. *Biophys. J.* **2001**, *80*, 643−655.
(22) Lee, B.; Richards, F. M. *J. Mol. Biol.* **1971**, *55*, 379−400.
(23) Shaefer, M.; Karplus, M. *J. Phys. Chem.* **1996**, *100*, 1578−1599.
(24) Hawkins, G. D.; Cramer, C. J.; Truhlar, G. *J. Phys. Chem*. **1996**, *100*, 19824−19839.
(25) Scarsi, M.; Apostolakis, J.; Gaflisch, A. *J. Phys. Chem.* **1997**, *101*, 8098−8106.
(26) Parsegian, A. Energy of an Ion crossing a Low Dielectric Membrane: Solutions to Four Relevent Electrostatic Problems. *Nature* **1969**, *221*, 844−846.
(27) Dominy, B. N.; Brooks, C. L., III. *Parametrization of a Generalized Born Model for the MSI CHARMm (Momany and Rone) Force Field*, personal communication.
(28) Richmond, T. J. *J. Mol. Biol.* **1984**, *178*, 63−89.
(29) White, S. H.; Wimley, W. C. *Annu. Rev. Biophys. Biomol. Struct.* **1999**, *28*, 319−65.
(30) Luecke, H.; Schobert, B.; Richter, H. T.; Cartailler J. P.; Lanyi, J. K. *J. Mol. Biol.* **1999**, *291*, 899−911.
(31) Palczewski, K.; Kumasaka, T.; Hori, T.; Behnke, C. A.; Motoshima, H.; Fox, B. A.; Le Trong, I.; Teller, D. C.; Okada, T.; Stenkamp, R. E.; Yamamoto, M.; Miyano, M. *Science* **2000**, *289*, 739−745.
(32) Spassov, V. Z.; Luecke, H.; Gerwert, K.; Bashford, D. *J. Mol. Biol.* **2001**, *312*, 203−219.
(33) Neria, E.; Fischer, S.; Karplus, M. *J. Chem. Phys.* **1996**, *105*, 1902−1921.
(34) Sitkoff, D.; Ben-Tal, N.; Honig, B. *J. Phys. Chem.* **1996**, *100*, 2744−2752.
(35) Bashford, D.; Karplus, M. *Biochemistry* **1990**, *29*, 10219−10225.
(36) *CRC Handbook of Chemistry and Physics*; Weast, R. C., Ed.; CRC Press: Boca Raton, FL, 1986; D159−D160.
(37) Ryckaert, J. P.; Ciccotti, G.; Berenson, H. J. C. *J. Comput. Phys*. **1977**, *23*, 327.
(38) Wiener, M. C.; White, S. H. *Biophys. J.* **1992**, *61*, 437−447.
(39) von Heijne, G. *J. Mol. Biol.* **1992**, *225*, 487−94.
(40) van Klompenburg, W.; Nilsson, I.; von Heijne, G.; de Kruijff, B. *EMBO J.* **1997**, *16*, 4261−4266.
(41) Stegman, T.; Helenius, A. In *Viral Fusion Mechanism*; Benz, J., Ed.; CRC Press, Inc.: Boca Raton, FL, 1993; pp 89−111.
(42) Macosko, J. C.; Kim, C. H.; Shin, Y. K. *J. Mol. Biol.* **1997**, *267*, 1139−1148.
(43) Takahashi, S. *Biochemistry* **1990**, *29*, 6257−6264.
(44) Luneberg, J.; Martin, I.; Nussler, F.; Ruysschaert, J. M.; Herrmann, A. *J Biol. Chem*. **1995**, *270*, 27606−27614.
(45) Han, X.; Steinhauer, D. A.; Wharton, S. A.; Tamm, L. K. *Biochemistry* **1999**, *38*, 15052−15059.
(46) Eisenberg, D.; McLachlan, A. D. *Nature* **1986**, *319*, 199−203.
(47) Nicholls, A.; Honig, B. *J. Comput. Chem.* **1991**, *12*, 435−445.
(48) Hol, W. G. J.; van Duijnen, P. T.; Berendsen, H. J. C. *Nature* **1978**, *273*, 443−446.
(49) Spassov, V. Z.; Ladenstein, R.; Karshikoff, A. D. *Protein Sci.* **1997**, *6*, 1190−1196.
(50) Han, X.; Bushweller, J. H.; Cafiso D. S.; Tamm, L. K. *Nat. Struct. Biol.* **2001**, *8*, 715−720.
(51) Ben-Tal, N.; Honig, B.; Peitzsch, R. M.; Denisov, G.; McLaughlin, S. *Biophys. J.* **1996**, *71*, 561−575.
(52) Murray, D.; Honig, B. *Mol. Cell* **2002**, *9*, 145−154.
(53) Rocchia, W.; Sridaram, S.; Nicholls, A.; Alexov, E.; Chiabrera, A.; Honig, B. *J. Comput. Chem.* **2002**, *23*, 128−137.
(54) Warshel, A.; Papazyan, A. *Curr. Opin. Struct. Biol.* **1998**, *8*, 211−217.
(55) Sharp, K. A.; Nicholls, A.; Fine, R. F.; Honig, B. *Science* **1991**, *252*, 106−109.
(56) Bernèche, S.; Nina, M.; Roux, B. *Biophys. J.* **1998**, *75*, 1603−1618.