PROTEINS: Structure, Function, and Genetics 41:211–223 (2000)

# Dipole Lattice Membrane Model for Protein Calculations

**Alan Grossfield,[1] Jonathan Sachs,[2] and Thomas B. Woolf[1,2,3]***
[1]*Department of Biophysics and Biophysical Chemistry, Johns Hopkins Medical School, Baltimore, Maryland*
[2]*Department of Biomedical Engineering, Johns Hopkins Medical School, Baltimore, Maryland*
[3]*Department of Physiology, Johns Hopkins Medical School, Baltimore, Maryland*

***ABSTRACT*** **A dipole lattice model for lipid membranes and their interactions with peptides is presented. It uses the Langevin dipole method to calculate electrostatic interactions in the heterogeneous membrane environment. A series of test cases are presented, including spherical charges, dipoles, side chain analogs, and helical peptides. The model consistently produces qualitatively correct results. Proteins 2000;41:211–223.** © 2000 Wiley-Liss, Inc.

**Key words: Langevin dipoles; membranes; membrane proteins; electrostatics**

## INTRODUCTION

Membrane proteins are intimately involved in a wide variety of biological processes. Despite their importance, there is relatively little available structural information. Although solution nuclear magnetic resonance (NMR) and X-ray crystallography have yielded an enormous number of high-resolution structures of soluble proteins, these methods are more difficult to apply to membrane proteins. Moreover, membrane proteins are generally harder to overexpress and purify than soluble proteins. As a result, there are fewer than 20 high-resolution membrane protein structures in the protein databank.[1] The relative paucity of experimental data creates an environment in which computational methods can be very valuable.

Indeed, there have been several significant theoretical and computational contributions to the understanding of membrane protein structure. One of the earliest was the development of hydropathy analysis, which allowed the accurate prediction of the locations of membrane spanning helices.[2,3] More recently, several groups have computationally predicted the structure of the glycophorin A transmembrane domain dimer.[4–6] Other groups have used Poisson-Boltzmann or other electrostatic calculations to explore the origins of membrane protein stability.[7–12]

Most of these efforts used very simple membrane models, which complicates detailed analysis of their results. Specifically, they either ignore the membrane entirely,[6] or consider it to be a simple hydrophobic slab surrounded by water.[2,3,7–9] In a similar vein, one group recently presented a membrane model based on the "atomic solvation parameters",[13] simply adjusting the parameters to represent an isotropic hydrophobic environment.[14,15] Approximations of this nature assume that the membrane-water interface is vanishingly narrow and that any location in the membrane is adequately described as either bulk hydrocarbon or bulk water. However, these assumptions do not reflect the structure of real lipid membranes; the interfacial region contains a mixture of water, lipid headgroups, methylene groups, and even terminal methyls.[16] This "tumultuous region of chemical heterogeneity" has properties distinct from either the membrane interior or bulk water.[16] For example, small aromatic compounds, such as indole, bind preferentially in the membrane-water interface.[17] Since membrane proteins often have tryptophan residues in the interface, this may be relevant for membrane protein folding and stability.[17] Moreover, several biologically important peptides bind and fold in the interfacial region.[18–22]

Some calculations have attempted to capture the complexity of lipid membranes and their interactions with proteins at the atomic level. Several groups have performed all-atom molecular dynamics simulations of pure bilayers,[23–28] while others have focused on various membrane permeants,[29,30] peptides,[31–33] and proteins.[32–39] However, these are very expensive calculations, with the result that the time scales are typically on the order of nanoseconds, far too short to thoroughly explore conformational space. These problems exist in simulations of soluble proteins, but are greatly exacerbated by the slow relaxation times and long relevant length scales of the membrane environment.[40]

The present work describes a new method to calculate the interaction of molecules with lipid membranes. The membrane is modeled as a lattice of dipoles, using the Langevin dipoles method.[41,42] The dielectric heterogeneity of the membrane is represented by making the intrinsic moment of the dipoles a function of their position in the membrane. As a result, our method can capture the effects of a broad membrane-water interface while avoiding the computational costs of all-atom simulations. A variety of systems are treated, including spherical charges, dipole, side chain analogs, and helical peptides.

## METHODS

The key method used in this study is the representation of the membrane by a lattice of permanent, freely reorient-

Grant sponsor: American Heart Association; Grant sponsor: National Institutes of Health; Grant number: RR12600.

A. Grossfield's present address is Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine.

*Correspondence to: Thomas B. Woolf, Department of Biophysics and Biophysical Chemistry, Department of Physiology, Johns Hopkins Medical School, 725 N. Wolfe Street, Baltimore, MD 21205. E-mail: woolf@groucho.med.jhmi.edu

Received 28 January 2000; Accepted 22 May 2000

able dipoles. The electrostatic heterogeneity of the membrane is modeled by varying the magnitude of the permanent dipoles as a function of their position in the bilayer.

### Dipole Lattice Solvation Energies

The solvation free energy for a solute in a specific location in the membrane is estimated by calculating the response of the lattice dipoles to the solute partial charges. The solvation free energy is written as a sum of four terms: (1) electrostatic interaction between the solute's partial charges and the permanent dipoles of the lattice; (2) electrostatic interactions *between* the lattice dipoles, after polarization by the solute charges; (3) entropy lost by the lattice dipoles when oriented by the electric fields induced by the solute; (4) a cavitation penalty, calculated as the dipole-dipole interaction lost by inserting the solute.

The orientational polarization of the membrane dipoles is calculated using the Langevin response function[43,41,42]

$$\vec{\mu}^L = \mu_0 \hat{E}_0 \left( \coth y - \frac{1}{y} \right), \tag{1}$$

with

$$y = \frac{\mu_0 E_0}{k_B T}$$

In this notation, $E_0$ is the electric field, $\hat{E}_0$ is the unit vector parallel to $E_0$, $\mu_0$ is the dipole's permanent moment, and $\vec{\mu}^L$ is the thermally averaged moment. The Langevin function can be derived exactly from the partition function for a permanent dipole in an electric field at fixed temperature (see Appendix for derivation).

The electrostatic energy $U_{\text{elec}}$ for a solute is calculated in the following manner: (1) Place the solute in the lattice, removing any dipoles which overlap the solute. (2) Calculate the electric field at each dipole due to the solute's partial charges. (3) Calculate the dipoles induced by the electric field, using the intrinsic dipole moment $\mu_0$ and the Langevin response function (Eq. 1). (4) Calculate the electric fields due to the dipoles. (5) Calculate the total energy of the system, and return to step 3 if the change in the energy exceeds some tolerance. This procedure calculates the average electric field in a self-consistent manner. The local fluctuations of the dipoles—and the resulting fluctuations in the electric field—are not explicitly represented in this model. Instead, each dipole sees the average field due to its environment. Although explicit fluctuations of each dipole are replaced with its thermally averaged response, we can calculate the entropy lost due to this response as

$$\Delta S = k_B [\ln(\sinh y) - y \coth y - \ln y] \tag{2}$$

where $y$ is as defined in Eq. 1. Equation 2 can be derived exactly from the partition function for a permanent dipole in a constant field (see Appendix for derivation). The total entropy lost by the solvent is then calculated by summing the entropies lost by the individual dipoles. Although component entropies are not generally additive, this procedure is correct in the context of a Langevin dipole model (see Appendix).

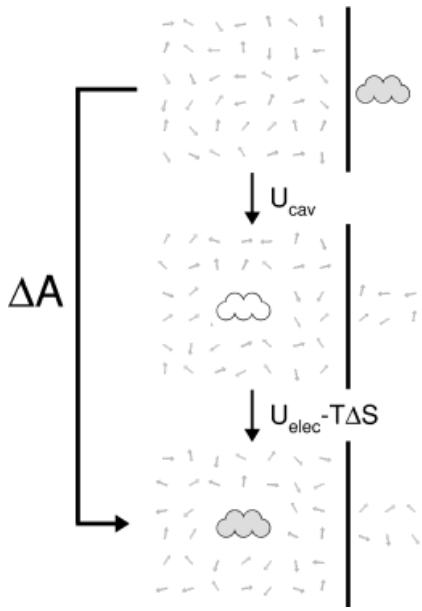


Fig. 1. A schematic for calculating solvation free energies with the lattice dipole membrane model. Interactions *between* dipoles that have been removed are not included, because these interactions are retained after removal. The vertical lines indicate infinite separation.

The final part of the calculation is the cavitation term. As described above, dipoles that overlap the solute are removed from the system before any energy calculations are performed. Because these are permanent dipoles, they interact favorably with the surrounding dipoles, even in the absence of solute partial charges to orient them. Accordingly, this interaction is estimated using the following expression

$$U_{\text{cav}} = k_{\text{cav}} \sum_i^{\text{Excl}} \sum_j^{\text{Incl}} \frac{\mu_{0_i}^2 \mu_{0_j}^2}{3 k_B T (4\pi\epsilon_0)^2 r_{ij}^6} \tag{3}$$

where $k_{\text{cav}}$ is an empirical constant, and the two summations are over the dipoles excluded from the calculation due to overlap and those remaining, respectively. The functional form in Eq. 3 is taken from the potential of mean force between two permanent dipoles at constant temperature, also known as the Keesom energy.[44] The constant is necessary because we are using a relatively weak term (the thermally averaged interaction of permanent dipoles) to account for behavior that is generally controlled by a significantly stronger term (dispersion forces between induced dipoles). We have done so in order to make maximal use of the quantities already present within our model (the permanent dipole moments) and to reduce the number of free parameters used to fit the model to experiment (see Parameterization section). Moreover, this formulation is easily generalizable to model heterogeneous environments.

In summary, the total solvation free energy is written

$$\Delta A = U_{\text{elec}} - T\Delta S + U_{\text{cav}} \quad (4)$$

Figure 1 is a graphic representation of the solvation free energy calculation. The solute is represented as a rigid body in a vacuum, which is inserted into the membrane lattice. The dipoles that overlap it are displaced into the vacuum, *still in their lattice configuration.* As a result, interactions between these dipoles do not change and can be neglected. The energy change upon their removal is proportional to the average interaction between them and the dipoles remaining in the system (see Fig. 1).

## Implementation Details

The membrane lattice model was implemented as part of the Molecular Modeling Toolkit, MMTK, using a combination of the Python and C languages.[45] All models were constructed using bond, angle, and dihedral parameters from the AMBER94 forcefield, and AMBER94 partial charges were used in all electrostatic calculations.[46] The electrostatics calculations were typically performed using an $8 \times 8 \times 16$ cubic lattice (except as noted), with 4 Å spacing, yielding a periodic system with a unit cell $32 \times 32$ Å in the membrane plane and 64 Å along the membrane normal. We used a 12 Å cutoff for dipole-dipole and charge-dipole interactions. For the purposes of identifying dipoles overlapping the solute, all dipoles were assigned a radius of 1 Å, while atomic radii were taken from the AMBER94 parameters.[46] The electric fields and induced dipoles were iterated until the energy change was less than 0.1%.

Calculations in a dipole lattice are vulnerable to lattice artifacts; to compensate, one generally performs multiple calculations, offsetting the lattice origin (or the solute) by a fraction of the lattice interval.[10] These artifacts are exacerbated in the present calculations by the heterogeneity of the lattices in the $z$ dimension. Previous calculations considered an isotropic environment, where any position dependence in the calculation was a lattice artifact.[42,10] By contrast, there is real positional dependence in the present model; correctly assessing the dependence of the solvation free energy on the solute's location in the membrane is the primary goal of our calculations. Accordingly, we took care to properly average away lattice artifacts. Each free energy calculation was repeated multiple times; typically, the solute was offset by 0 Å, 1.5 Å, and 3 Å in each dimension (27 calculations in all). The intrinsic dipole moments on the lattice were offset with the solute, such that the position of the solute along the membrane normal was preserved. This procedure is physically equivalent to translating the lattice in the opposite direction. Since the various offsets represent physically equivalent states, the resulting quantities are arithmetically averaged.

The current choice of offsets is a compromise between computational speed and precision; there is some oscillation in the calculated free energy curves due to lattice artifacts. These oscillations are worst when the solute is located in a polar environment and when the solute is locally very polar. However, these oscillations do not prevent qualitative analysis of the curves, and only moderately diminish their precision. Moreover, if greater precision is desired, one can simply calculate more offsets in the region of interest.

## Parameterization

As currently defined, the model has three free parameters, once the lattice spacing and type are defined: $\mu_{\text{hc}}$ (the intrinsic dipole moment for dipoles in the hydrophobic core of a membrane), $\mu_{\text{wat}}$ (the intrinsic dipole moment for dipoles in the water region), and $k_{\text{cav}}$, the multiplier for the cavitation term. These parameters were chosen to reproduce transfer free energies for side chain analogs from bulk hydrocarbons to water.[47] It must be recalled that $\mu_{\text{hc}}$ and $\mu_{\text{wat}}$ are not molecular parameters, and as such, there is no direct connection between the dipole moment of a single water molecule and $\mu_{\text{wat}}$. Rather, $\mu_{\text{wat}}$ (when combined with a specific choice of lattice spacing) reflects an average dipole density, which allows the model to reproduce experimental results.

The parameterization was performed as follows: First, side chain analogs were constructed for 16 of the 20 amino acids. Proline was excluded because of its different backbone, glycine and alanine were excluded because of a lack of available parameters to describe their respective analogs, and arginine was excluded because parameters for its neutral form were unavailable. The remaining charged residues (glutamic acid, aspartic acid, and lysine) were represented in their neutral forms. The additional free energy to transform the analog from its uncharged to its charged form was estimated from the $pK_a$.[48] No correction was applied for histidine.

Each side chain analog structure was subjected to up to 1,000 steps of conjugate gradient minimization, to ensure that it was relaxed. We then generated a library of structures from a molecular dynamics trajectory. First, the velocities of each molecule were initialized and scaled to a temperature of 50 K. The temperature was then raised to 300 K over the course of 5 ps of dynamics, using a 1 fs timestep. The structures were re-minimized, using 1,000 steps of steepest descent minimization, and reheated to 300 K. Finally, velocities were reassigned at 300 K, and 400 ps of dynamics were run. Structures were saved every 20 ps. All calculations were performed using the AMBER94 forcefield, as implemented in MMTK.[46,45]

Next, we calculated the solvation energy for each structure in our library in a series of homogeneous lattices, where the intrinsic dipole moment of each dipole was varied from 0.1 D to 1.0 D, in steps of 0.01 D, using an $8 \times 8 \times 8$ lattice with 4 Å spacing and a 12 Å cutoff. We calculated the transfer energies between all pairs of lattices, with the restraint that $\mu_{\text{hc}} < \mu_{\text{wat}}$. For each combination, we chose $k_{\text{cav}}$ to minimize the average error between our calculated transfer energies and the experimental values.[47] We found the best overall choice of parameters was $\mu_{\text{hc}} = 0.1D$, $\mu_{\text{wat}} = 0.87D$, $k_{\text{cav}} = 4.19 \cdot 10^9$. These values gave an average error of 1.18 kcal/mol per residue, with a standard deviation of 1.67 kcal/mol; for comparison, the absolute values of the transfer energies (including charge neutralization, where appropriate) ranged from 6.6

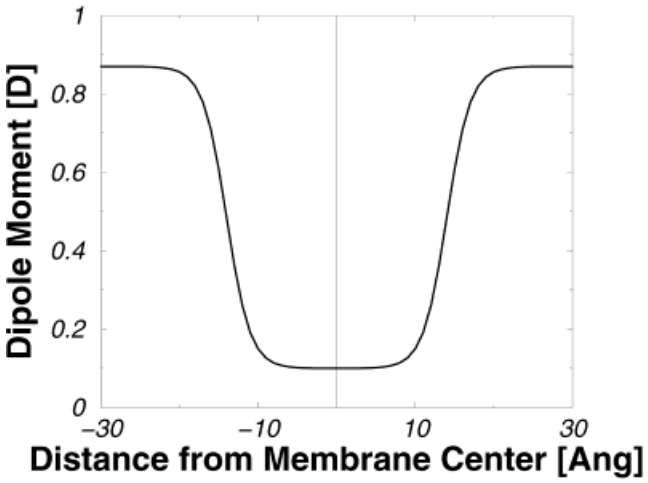


Fig. 2. Intrinsic dipole moment of a membrane lattice, as a function of position in the membrane.

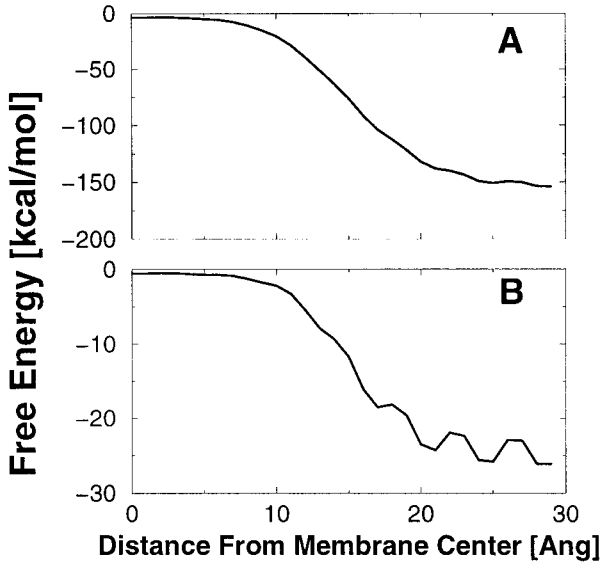


Fig. 3. Free energy profile for (**A**) a spherical charge and (**B**) a physical dipole. The sphere has charge + 1e and radius 2 Å. The dipole is constructed from two spheres, with charge ± 1e, separation 1 Å, and radius 2 Å, yielding a dipole moment of 1 D.

kcal/mol (acetamide, the analog for aspargine) to 0.14 kcal/mol (cresol, the analog for tyrosine), with an average of 3.14 kcal/mol. The largest error was 2.3 kcal/mol, for 4-methyl imadizole, the analog for histidine. The average standard error per residue was 0.1 kcal/mol, indicating that 20 structures per residue were sufficient to allow statistical convergence.

### Membrane Model

The dielectric heterogeneity of the membrane environment is modeled by varying the intrinsic moment of the dipoles as a function of their position in the lattice, as shown in Figure 2. Taking the membrane normal to be the $z$ axis, and the origin to be the center of the membrane, we used the following function

$$\mu_{0,i} = \frac{\mu_{\mathrm{wat}} - \mu_{\mathrm{hc}}}{2}\left[\tanh\left(\frac{|z_i| - w_{\mathrm{mem}}}{w_{\mathrm{int}}}\right) + 1\right] + \mu_{\mathrm{hc}} \qquad (5)$$

where $z_i$ is the position of dipole $i$, $w_{\mathrm{mem}} = 14$ Å is half the membrane width, and $w_{\mathrm{int}} = 3$ Å is the width of the interface. As shown in Figure 2, the polarity changes smoothly from low values ($\mu_{\mathrm{hc}}$) in the membrane interior to higher values in the water ($\mu_{\mathrm{wat}}$). The primary advantage of this functional form is the ready control of membrane and interface width. With these parameters, the spatial dependence for the dipole moment roughly parallels the experimentally derived water density in real membranes.[16]

## RESULTS

The purpose of this study is to demonstrate that a dipole lattice membrane model can be used to describe the energetics of small molecules and peptides interacting with lipid bilayers, without incurring the computational cost of all-atom molecular dynamics simulations. The long-term goal of this work is to generate a model that could be used to represent the membrane in predictions of membrane protein structure. Accordingly, we will present a series of test cases, to demonstrate the accuracy of the model in a variety of circumstances.

### Spherical Charges and Physical Dipoles

Perhaps the simplest test of an electrostatic model is its treatment of a spherical charge. Figure 3A shows the free energy for a spherical charge as a function of its position in the membrane. Not surprisingly, the membrane lattice model shows that it is enormously unfavorable (≈ 154 kcal/mol) to embed a bare charge in a membrane. The largest contribution to this free energy change is electrostatic; however, there is a significant entropy loss (≈ 40 kcal/mol) upon transferring the charge into water. Physically, this finding indicates that the dipoles surrounding the charge are strongly oriented by its electric field. As an aside, it is interesting to note that continuum electrostatics also predicts a significant entropic penalty for solvating a charge in water, once the temperature dependence of the dielectric constant is taken into account.[44]

Figure 3B shows the analogous curve for a 1 Debye dipole, composed of two spheres of the same charge magnitude and radius as in Figure 3A. Comparing the two curves, it is obvious that forming a dipole greatly reduces the penalty to move charges into the membrane. These results, while unsurprising, indicate at the simplest level that the model is performing as expected.

It is also important to note that, although both curves have noticeable oscillations due to lattice artifacts in the water region (≈ ± 1.5 kcal/mol), they are quite small relative to the total free energy change. Given their small size and large partial charges, these two systems should be more vulnerable to lattice effects than most biomolecules, indicating that this protocol can be used to calculate free energy profiles without significant artifacts.

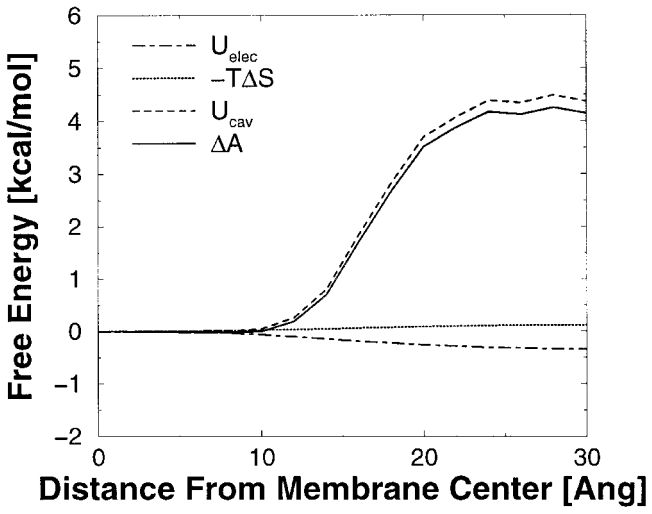


Fig. 4. Free energy profile for butane, the analog for a isoleucine side chain. Each point on the curves is a Boltzmann average of free energies calculated for five different butane configurations. The free energy is lowest in the membrane interior, primarily because $U_{cav}$ opposes transfer into water.

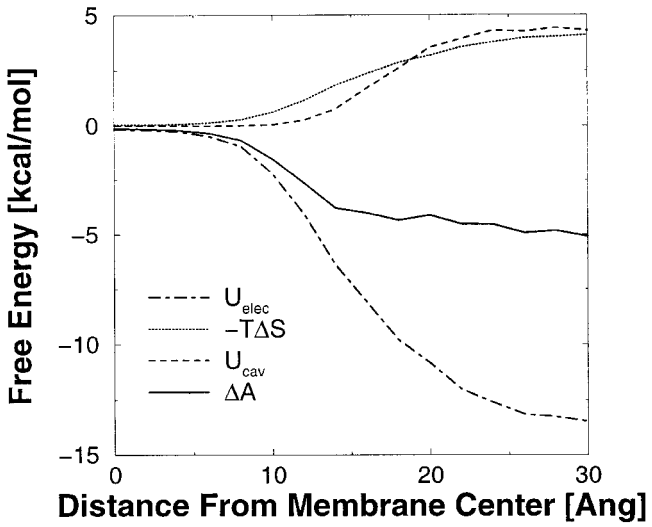


Fig. 5. Free energy profile for 4-methylimadizole, the analog for a histidine side chain. Even though this curve was calculated using the neutral form of the molecule, the free energy minimum is clearly in the water, driven primarily by $U_{elec}$.

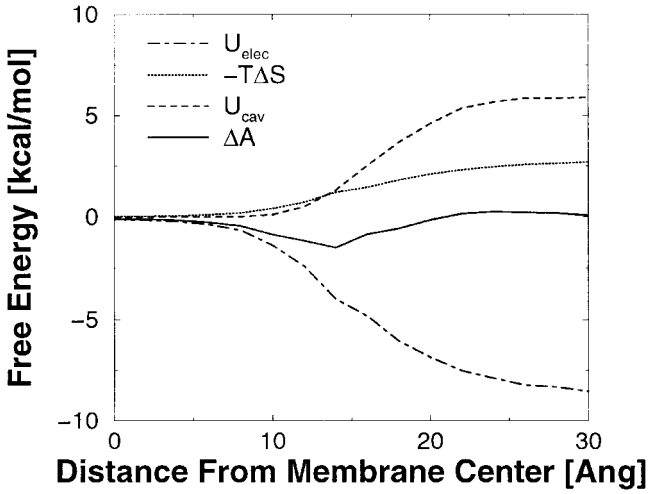


Fig. 6. Free energy profile for 3-methylindole, the analog for a tryptophan side chain. The free energy minimum is in the membrane interface (≈ 15 Å from the membrane center), in agreement with experiment [17]. All free energy components make significant contributions.

### Side Chain Analogs

The next step in validating the model is to calculate free energy curves for various side chain analogs. Since the model was parameterized to reproduce bulk hydrocarbon to bulk water transfer energies for these analogs, it should produce reasonable free energy differences for transfer from the membrane interior to water. However, it is still instructive to examine some of these curves, to understand which free energy components are important in different circumstances.

The free energy curves for each side chain analog were calculated as follows: First, five structures were generated for each analog, using the procedure described in the Parameterization section. Next, a free energy curve was calculated for each structure, using the general method described in the Implementation Details section. A $6 \times 6 \times 16$ lattice was used for these calculations, with offsets of 0.0 Å, 0.75 Å, 1.5 Å, 2.25 Å, 3.0 Å, and 3.75 Å in the $z$ direction, and 0 Å, 1.5 Å, and 3.0 Å in the $x$ and $y$ directions (54 total calculations). The five free energy values were then Boltzmann averaged at each location. The free energy change between membrane interior and water and the overall shape of the free energy curves do not depend significantly on the lattice size.

Figures 4, 5, and 6 show the free energy curves for the isoleucine, histidine, and tryptophan side chain analogs, and the contributions of the different terms. Some basic properties are immediately obvious. For example, $U_{\text{elec}}$ always favors the more polar environments. This finding is as expected, since induced polarization will by definition result in favorable interaction, and there are larger dipoles in the more polar environments. Conversely, the entropic term favors the less polar environment; a larger intrinsic dipole feels a larger force in a given field, and as a result is more strongly oriented and loses more entropy. The entropic free energy change is always smaller in magnitude than $U_{\text{elec}}$. Finally, the cavity term also favors the nonpolar environments, because it is proportional to the square of the magnitude of the dipoles displaced by the solute and is, thus, larger in a polar environment.

These three analogs were chosen to illustrate the relative importance of the free energy terms for different solutes. For example, the leucine analog (Fig. 4) has very small partial charges in the AMBER potential; as a result, its $U_{\text{elec}}$ and entropy losses are small, regardless of the

environment. However, it is a somewhat bulky molecule, so there is a significant cavity term, which drives it from the water into the membrane interior. The histidine analog (Fig. 5), however, has a large electrostatic component ($U_{\mathrm{elec}}^{\mathrm{water}} - U_{\mathrm{elec}}^{\mathrm{hc}} \approx -13.3$ kcal/mol), which drives it into the water region.

Perhaps the most interesting example is 3-methylindole, the tryptophan side chain (Fig. 6). The free energy minimum for this analog is about 14 Å from the membrane center, in the membrane-water interface, as a result of compensation between $U_{\mathrm{elec}}$, $U_{\mathrm{cav}}$, and the entropy change. The Boltzmann averaged location for 3-methylindole is 13.4 Å from the membrane center; this result agrees with recent NMR measurements showing that it partitions into the head group region of lipid bilayers.[17] Experimentally, 3-methylindole favors binding to POPC vesicles over bulk water by about 8.6 kcal/mol, while in the present calculation the free energy difference is only 1.4 kcal/mol.[49] Clearly, the present model significantly underestimates the favorability of interfacial binding, but it must be recalled that no effort was made to parameterize this model to reproduce interfacial properties. Moreover, the present model contains no representation of the lipid headgroup dipoles, which would almost certainly further stabilize interfacial binding. Finally, it is difficult to imagine methods that do not represent the finite width of the interfacial region producing this result.

## WALP

The next model system we considered is the WALP-peptide (WALP). The WALPs are short peptides ($\approx$ 16–25 residues) containing alternating alanine and leucine residues, with two tryptophans at each terminus. They have been used experimentally to investigate hydrophobic matching in transmembrane helices and have been simulated recently using molecular dynamics.[50–53] The current calculations use WALP-19 (19 residues total length) to demonstrate that our model can correctly recognize a stable transmembrane helix.

WALP-19 was constructed in an ideal $\alpha$-helical conformation, with all side chains in their extended states. The N-terminus was blocked with a formyl group, and the C-terminus was blocked with ethanolamine, consistent with the experimental work and simulations.[52,53] The structure was then minimized, to allow the side chains to find more favorable orientations. For clarity, the "position" of the WALP is defined to be the location of its center of mass relative to the center of the membrane, and its tilt angle is defined relative to the membrane normal. For example, a position of $z = 0$ and a tilt of 0° would be a standard transmembrane configuration.

The free energy was calculated with WALP-19 in a variety of positions and orientations. The first calculation simply dragged WALP-19 through the bilayer, moving the center of mass from −22 Å to 22 Å, at tilts of 0°, 10°, 20°, and 90°. The transmembrane configurations (0°, 10°, and 20°, $0 < z < 5$) had the lowest free energies, as expected (data not shown). By contrast, the 90° configuration had the least favorable free energies over the same range of positions, roughly 30 kcal/mol higher than the transmembrane configurations (data not shown). This is reasonable, because burying the entire helix in the membrane core should be very unfavorable.

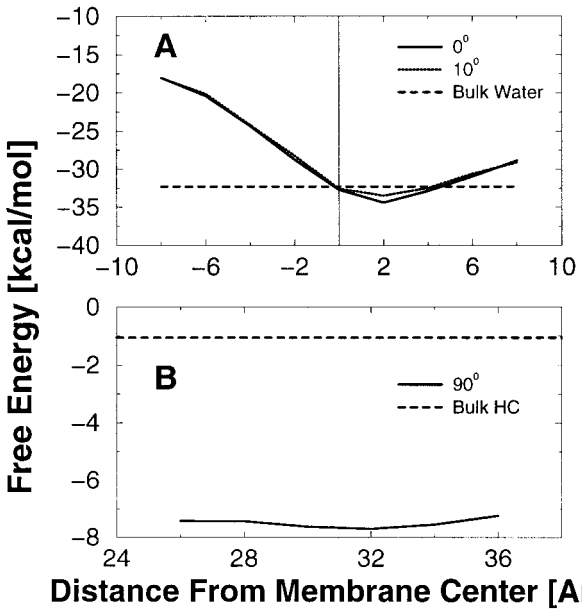


Fig. 7. Free energy profile for WALP in different configurations. **A:** The free energy for two transmembrane orientations, tilted 0° and 10° with respect to the membrane normal and centered near the center of the membrane. Each point is a Boltzmann average of free energies calculated for rotations about the helical axis, at 30° intervals. For comparison purposes, the free energy for the same helix in bulk water is also shown. **B:** The free energy for a peripheral interfacial location, with the helix tilted 90° with respect to the membrane normal. The free energy for the helix in bulk hydrocarbon is shown for comparison. The transmembrane orientations have the lowest overall free energy, when located 2 Å above the membrane center.

However, this simple calculation neglected variations of free energy with rotations about the helical axis, which can significantly effect the free energy when the helical tilt is nonzero. Accordingly, we recalculated the free energies for a subset of these configurations, considering all rotations about the helical axis in 30° increments. The free energies for the different rotations were then Boltzmann averaged, at each $z$ position. Similar calculations were performed for pure hydrocarbon and water lattices, for comparison purposes. The results are reasonable (see Fig. 7): WALP-19 is most stable in a transmembrane orientation. The next most favorable state is in pure water, approximately 2.1 kcal/mol less stable. The interfacial and pure hydrocarbon states are far less favorable. These free energies correspond to a situation where roughly 97% of the WALP is found in the transmembrane state. In actuality, the model most likely underestimates the preference for the transmembrane states, because, as described in the Side Chain Analogs section, it underestimates the favorability of placing the terminal tryptophans in the membrane interface. Still, the model correctly identifies WALP as a transmembrane helix.

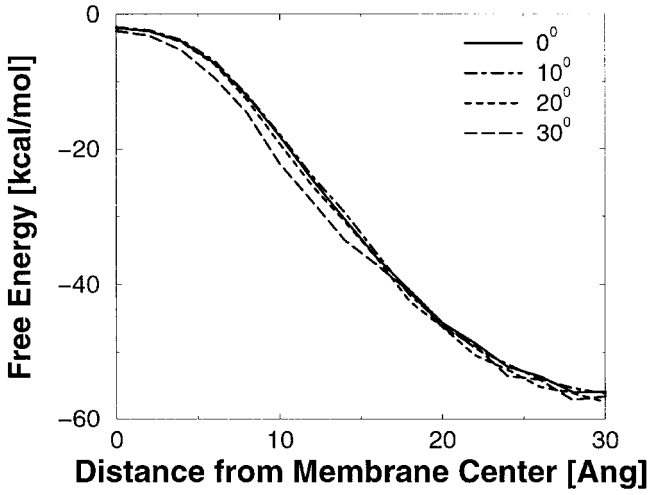


Fig. 8. Free energy profile for the neutral form of alamethicin. Each data point is a Boltzmann average of free energies calculated for rotations about the long axis. Curves are shown for alamethicin tilted by 0°, 10°, 20°, and 30° with respect to the plane of the membrane.

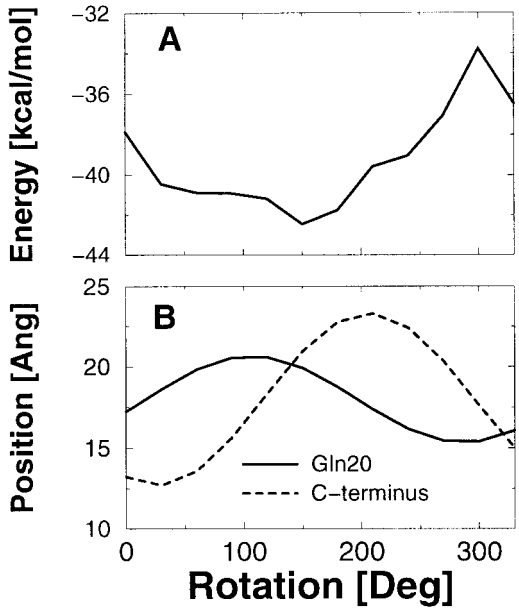


Fig. 9. Free energy for the neutral form of alamethicin in the membrane-water interface, 18 Å from the membrane center. **A:** The free energy as the molecule is rotated about its long axis. **B:** The center of mass positions for the side chain of Gln19 and the C-terminus. The free energy is lowest when both Gln19 and the C-terminus are oriented away from the membrane, even though this conformation partially buries Glu18.

### Alamethicin

The final test case is the channel-forming peptide alamethicin. Alamethicin is a 20 residue peptide, rich in α-amino butyric acid, which at low concentrations and transmembrane voltages binds to membrane interfaces in a mostly helical conformation.[18] The sequence is BPB-ABAQBVBGLBPVBBEQP, where B is used to represent α-amino butyric acid. The N-terminus is acetylated, while the final residue is an α-amino alcohol. There has been a great deal of experimental work on this peptide, including determination of its structure in nonaqueous environment[54] and measurements of its channel forming and conductance properties.[18] The structure and dynamics of alamethicin in methanol has been investigated using NMR and molecular dynamics by the Dempsey group.[55,56] More recently, Tieleman et al. have published several articles that report molecular dynamics simulations of this peptide in a variety of environments.[32,36,57]

For these calculations, we used the coordinates from a crystal structure solved in a methanol-acetonitrile solution.[54] There are three alamethicin molecules in the unit cell; therefore, we (arbitrarily) used the first one. It is possible that this structure, solved in a homogeneous environment, does not reflect the true structure of alamethicin when bound to a membrane. However, molecular dynamics simulations of alamethicin in a variety of environments and experimental NMR work in methanol have implied that this structure is probably not far from the membrane-bound structure.[32,36,55–57] There is a single ionizable residue, Glu18; we performed calculations using both the charged and uncharged forms.

First, we calculated the free energy for the neutral form of alamethicin, as a function of distance from the center of the membrane. Since alamethicin is expected to bind interfacially, tilt angles are defined relative to the plane of the membrane. Alamethicin was oriented parallel to the plane of the membrane, with tilts of 0°, 10°, 20°, and 30°. For each tilt and distance from the membrane center, we performed solvation calculations for all rotations about the long axis of the molecule, in 30° increments. As shown in Figure 8, the free energy curves drop smoothly as alamethicin moves from the membrane interior to the water. Although these calculations correctly indicate that completely burying the molecule in the membrane interior is unfavorable, they do not show a free energy minimum in the interface, as one would expect for interfacial binding. This is most likely due to the absence of favorable interactions between lipid head groups and the peptide in our model.

However, we can examine interfacial conformations to see which are preferred, and why. Figures 9 and 10 focus on the free energy for alamethicin when oriented parallel to the membrane, 18 Å from the center of the bilayer. In each figure, part A shows the free energy as a function of rotation about the molecule's long axis, while part B shows the location, relative to the membrane center, of specific parts of the molecule. Figure 9 shows the results of calculations performed using the neutral form of alamethicin, while Figure 10 shows analogous results obtained using the charged form. It should be noted that the two free energy curves cannot be directly compared, because they use different reference states. Rather, the changes in free energy upon rotation should be compared.

Perhaps unsurprisingly, the two free energy curves are quite different: the location of the free energy minimum is shifted by roughly 100°, and the range of free energy values is radically different (8.7 kcal/mol vs. 54 kcal/mol).

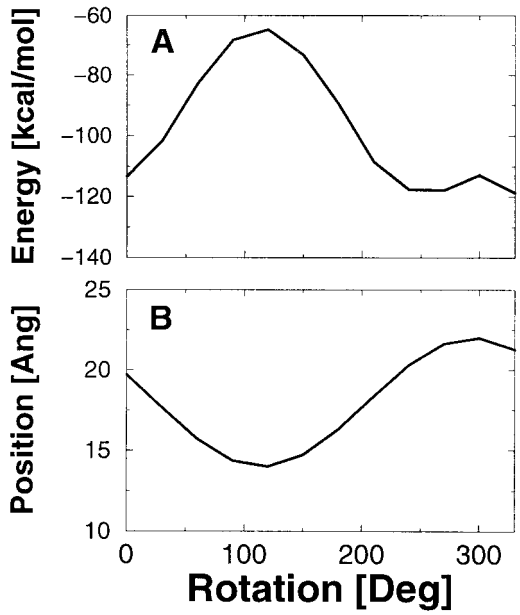


Fig. 10. Free energy for the charged form of alamethicin in the membrane-water interface, 18 Å from the membrane center. **A:** The free energy as the molecule is rotated about its long axis. **B:** The center of mass positions for the side chain of Glu18. The free energy is lowest when both Glu18 is oriented away from the membrane.

Part B of each figure explains the physical origin of these effects. In Figure 9, the center of mass positions of the Gln19 side chain and the C-terminus backbone are shown; the low free energy states occur when these groups are exposed to water. In this conformation, Glu18 is mostly buried in the membrane, but the solvation penalty is not particularly large because the neutral form of the side chain is used. By contrast, in Figure 10 the conformational free energy is dominated by the location of the charged Glu18 side chain. If the real system were forced to bury the glutamate, the molecule's conformation would change, and the glutamate would likely protonate, or bind a counterion or water molecule. We calculate unphysically large barriers to rotation ($\approx$ 54 kcal/mol), because these relaxation mechanisms are not currently represented in the model.

## DISCUSSION

### Comparison With Continuum Electrostatics Calculations

Significant effort has been invested in modeling the electrostatics of membrane-protein interactions. Most of the work has used continuum electrostatic theory.[7–9,11,58–60] The most common approach has been to solve the Poisson-Boltzmann equation numerically.[61,9] In these calculations, the membrane is represented as a low dielectric slab (typically $\epsilon = 2$ or 4) surrounded by high dielectric region ($\epsilon = 80$). In this sort of model, the membrane-water interface is infinitesimally narrow; every point in space is either hydrocarbon-like, water-like, or inside the molecule. This finding is quite different from the situation in real membranes, where the interfacial region accounts for up to half the width of the bilayer.[16,19] However, it is difficult to know how to generalize a continuum model to represent a smoothly varying environment. A dielectric constant is inherently a bulk property, so it is not physically meaningful to talk about local variations on the angstrom length scale. These limitations have not prevented a number of calculations from successfully exploring the behavior of transmembrane helices.[7,8] Clearly, "dielectric slab" models capture some membrane properties quite effectively. However, it is difficult to use them to explore the importance of the width and shape of the membrane-water interface. This is in contrast to dipole lattice models; since the intrinsic moment of a dipole is a local property, local variations of electrostatic properties can be explored in a physically consistent manner.

This finding does not mean that interfacial binding cannot be explored using dielectric models. On the contrary, two recent papers by Murray et al. have used a dielectric slab membrane model, combined with an explicit all-atom representation for the lipid head groups, to explore the behavior of charged peptides and myristolated proteins with bilayers.[59,60] However, these calculations explore peripheral membrane binding; it is not clear how they could be used to construct a full free energy profile, because there is no mechanism by which the lipid headgroups can reorient in response to a solute. These papers demonstrate the importance of the lipid dipoles (and charges) for peptide binding. They also highlight another historical strength of Poisson-Boltzmann approaches, namely their ability to capture the effects of ionic strength on electrostatic interactions. Current dipole lattice models do not represent salt effects, although work is ongoing in our lab to introduce them, as described below.

### Parameterization Rationale

A new solvent model almost always raises the issue of parameterization. In the present case, our choice of parameterization method reflects both the anticipated applications of the model and our assessment of what physical aspects are and are not captured by it. Because we intend to use the model to describe the behavior of membrane proteins, we parameterized the model to reproduce side chain analog properties. As described in the Parameterization section, the parameterization does not directly include any membrane specific information. Rather, it makes the assumption that the interior of the membrane is chemically similar to bulk hydrocarbon. This assumption is supported by molecular dynamics simulations, which use hydrocarbons to parameterize lipid chains,[62,63] and by Poisson-Boltzmann calculations, which typically use a dielectric of 2 or 4, to match that of most bulk hydrocarbons.[7,8]

Furthermore, we chose to parameterize $\mu_{hc}$ and $\mu_{wat}$ simultaneously to reproduce free energies of transfer between hydrocarbon and water, although other choices are certainly plausible. For example, $\mu_{hc}$ and $\mu_{wat}$ could be estimated separately, using vacuum to liquid transfer. However, this approach would certainly fail, if the present model were used unchanged. Specifically, $U_{cav}$ is not sophisticated enough to appropriately capture all nonelectrostatic aspects of solvation. Experimental solvation free

energies are in principle the sum of many terms, including solute-solvent dispersion forces, solvent structure, solute entropy, and long-range electrostatics. Assuming that the Langevin dipole methodology correctly captures the electrostatic component of solvation, this leaves $U_{cav}$ to capture the other components. This is completely analogous to the role of surface area terms in most Poisson-Boltzmann and Generalized Born calculations.[64–67]

Virtually all of the compounds considered favor transfer from vacuum to both water and hydrocarbon.[47] This is a phenomenon $U_{cav}$ is fundamentally incapable of capturing, since it is unfavorable by construction. It states that displacing dipoles is unfavorable, and that displacing strong dipoles is more unfavorable than displacing weak dipoles. Its main purpose in our model is to ensure that nonpolar compounds have lower free energies in less polar environments. No purely electrostatic solvation model can produce this behavior, because the solute partial charges will always interact favorably with the solvent dipoles (or dielectric), and will interact more favorably with stronger dipoles (a higher dielectric).

At this point, we could have added a second nonelectrostatic term, to represent attractive solute-solvent interactions. However, we chose instead to assume that these forces are relatively nonspecific, and do not change significantly with solvent composition. With this in mind, we parameterized against hydrocarbon to water transfer data, and used $U_{cav}$ to represent the difference in cavitation cost between different media.

Moreover, it would perhaps be desirable to use membrane and membrane binding experiments to parameterize the lattice membrane model. Unfortunately, there are several problems with using membrane binding data directly. First, there is the issue of transferability. Binding properties of small molecules and peptides can depend strongly on the salt concentrations and lipid compositions used.[60,68] This could greatly complicate parameterization, especially if data had to be combined from multiple sources. Second, a combination of thermodynamic and structural information would be needed. For example, the White group has performed binding measurements of a series of peptides (WL-X-LL) to POPC vesicles, where all 20 amino acids were substituted for X.[69] In principle, these binding data could be used to calibrate a membrane model. However, a pentapeptide has significant conformational flexibility, and the membrane environment is heterogeneous. As a result, calculations would have to be performed for a large number of solute conformations, in a variety of positions and orientations. In essence, the entire free energy curve would be needed for every set of parameters, because the experiments represent an ensemble average over all conformations. Finally, this kind of approach would couple the shape of the membrane with the other parameters, making it more difficult to explore the effects of varying membrane profile. By contrast, the present model is quite flexible. For example, one could construct a crude model for a micelle simply by making the intrinsic dipole moment a function of distance from the origin, instead of position along the membrane normal.

As was stated above, $U_{cav}$ serves much the same purpose in our calculations as the surface area terms do in Poisson-Boltzmann or Generalized Born calculations.[64,65,66,67] Indeed, two recent papers by Efremov et al. attempted to model the behavior of membrane proteins using nothing but surface area terms.[14,15] However, there are some significant drawbacks to such an approach. For example, surface area approaches alone do not correctly capture the electrostatics of solvation.[67] Moreover, surface area terms do not readily capture environmental heterogeneity. Specifically, the model used by Efremov et al. treated the membrane as a homogeneous hydrocarbon.[14,15]

Finally, our calculation of the entropy lost due to dipole orientation is quite different from any found in the literature. This sort of term is not relevant in the context of continuum electrostatics; it is only when dielectric response is discretized into a finite number of dipoles that one can reasonably speak of the solvent entropy. However, even in the context of dipole lattice models, this effect has not been widely discussed. A recent article by Florián and Warshel addressed this issue.[70] The authors used the Langevin response function to estimate the average angle between the external electric field and the permanent dipole moment, then related the entropy to the conformational volume available to the dipole while restrained at this angle. They found that this functional form overestimated the entropy loss at small fields, and applied an empirical correction. The functional form used in the present work produces low-field entropy losses larger than their modified entropy but lower than their uncorrected functional form. At large fields, it is significantly larger than either of their terms. Although it is not clear which approach will perform better in practice, the present method has the advantage that it is derived exactly from the physics of a dipole in a field (see Appendix).

### Planned Improvements

There are several areas where the model could be improved, to represent membrane-protein interactions in a more realistic way. The present article was intended to demonstrate that lattices of dipoles could be used to qualitatively describe the behavior of membrane proteins. As such, the simplest plausible form of the method was used.

The most obvious area for improvement is long-range electrostatics. The present version of the model uses a simple cutoff, which seems adequate for the present qualitative applications but is clearly not optimal for more quantitative calculations. For example, it is likely that such an approximation would adversely affect calculations of the free energy of an ion in a channel. There are several possible alternatives to the use of cutoffs. One possibility would be to implement Ewald summation.[71] This approach would have several advantages: it is a rigorously correct way to handle a truly periodic system, and the particle mesh Ewald method can be implemented very efficiently.[72] However, there is the possibility of artifacts because of the enforced periodicity.[73,74,75] Moreover, there are physically interesting systems, such as a bilayer with

an applied transmembrane potential, which are not fully periodic. Although Ewald summation can be applied to systems with only two dimensional periodicity, there is a significant performance penalty.[76] A second alternative would be to use a reaction field approach.[77,74,78] These methods avoid both cutoff and periodicity artifacts, and can also be implemented efficiently. However, it is not clear how to implement them correctly in the context of a dipole lattice model, especially one with heterogeneous dielectric properties. Finally, a spherical boundary condition method could be applied.[79] However, membrane systems do not possess spherical symmetry, requiring additional approximations.

Another limitation of the model is its treatment of charged chemical groups. In the present calculations, all amino acid residues and side chain analogs were assumed be in their neutral forms, except as specifically noted. This is clearly not physically correct for many realistic circumstances. However, this approximation is necessary for the success of the present model; inclusion of charged groups causes the model to drastically overestimate the favorability of solvation by water, compared with the membrane, as seen in Figure 10. Put another way, net charges have enormous solvation penalties when buried in the membrane (see Fig. 3). However, Figure 3 also shows that this cost is dramatically reduced by the formation of a dipole, roughly a factor of 6 in this specific calculation.

The primary lesson from this is that, if charged groups are to be handled correctly, some provision for neutralization must be made. There are several plausible ways to do this. One approach would be to manually scale down the partial charges used in the calculation. This is analogous to the approach used by Sham et al., where charge-charge interactions are screened by a large "dielectric constant".[80] Unfortunately, the appropriate value for this scaling factor is not immediately obvious, especially once the heterogeneity of the membrane environment is taken into account.

Alternatively, one could simply consider both charge states for each ionizable group. This would effectively require a $pK_a$ calculation for all ionizable groups *at all positions and orientations* in the membrane. While this, in principle, could be done, the computational expense would be very large when the molecule of interest has a significant number of titratable groups. Several approximate methods exist to deal with this combinatoric explosion,[80–85] but all of them are concerned for the most part with $pK_a$s of residues in soluble proteins and would not generalize readily to a membrane environment.

A third possibility would be to simply place a counterion near each ionizable group. At first glance, this approach has a certain appeal: the electrostatic solvation penalty would be greatly diminished by the formation of a dipole, in a manner consistent with the behavior of real proteins in salt solutions. However, appropriate placement of the counterion would be problematic; some form of sampling would be required, which would most likely have to be repeated for a variety of locations and orientations in the membrane.

The fourth, and perhaps most satisfying solution, would be to attempt to capture salt effects implicitly in the calculations. This is the approach taken in Poisson-Boltzmann calculations.[86,87] A similar approach should be possible within the dipole lattice framework.

The model's representation of the membrane itself could also be improved. The present model captures the changes in polarity of the membrane in a simple manner. A broad membrane-water interface is represented, but treated simply as a region with properties intermediate between the membrane core and water. Real membranes are more complex; zwitterionic lipid head groups have large dipole moments, which are not freely reorientable.[88] The resulting polarization of the head groups and interfacial water leads to very large electric fields in the interfacial regions, which can significantly effect biologically relevant properties such as ion permeation.[89] Moreover, it has been shown that peptide binding affinities are strongly dependent on bilayer lipid composition.[68]

Fortunately, there seem to be several ways to capture lipid head group effects in a dipole lattice membrane model. The simplest approach would be to increase intrinsic dipole moments in the interface. However, this approach would not capture critical interfacial properties: the orientation of the head group dipoles is largely determined by covalent interactions with the rest of the lipid, not by bilayer-solute interactions. On the other hand, there is little question that membrane permeants, especially ions, affect the local polarization profile of bilayers. Hence, the dipoles would have to be partially reorientable. This could be done by applying a harmonic restraint to them while performing the iterative Langevin dipole calculation. Alternatively, the restraint could be incorporated directly into the Langevin calculation, by adding a "restraining field" only to the interfacial dipoles, and using the summed fields for the purposes of iteration. The magnitude of the "restraining field" and the interfacial dipoles would be chosen to reproduce the potential profile for neat bilayers and the degree of polarization induced by permeants. A series of long molecular dynamics simulations are in progress in our group in an attempt to supply this information.[90,91]

## CONCLUSION

A dipole lattice membrane model has been developed. It produces qualitatively correct results when applied to a variety of molecules interacting with membranes, including simple charges and dipoles, side chain analogs, and helical peptides. The results indicate that this method can complement existing computational methods, such as molecular dynamics and continuum electrostatics. Further enhancements—such as representations of salt effects and inclusion of the lipid head group dipoles—are planned, to better capture the relevant physics of bilayers and their interactions with membrane proteins.

## ACKNOWLEDGMENTS

The authors wish to thank Dr. Horia Petrache, Dr. Daniel Zuckerman, Dr. Joanne Bright, and Dr. Rohit Pappu for careful reading of the manuscript and many helpful conversations. We further thank Dr. Jay Ponder for the use of computational resources.

## REFERENCES


1. Berman H, Westbrook J, Feng Z, Gilliland G, Bhat T, Weissig H, Shindyalov I, Bourne P. The protein data bank. Nucleic Acids Res 2000;28:235–242.
2. Kyte J, Doolittle R. A simple method for displaying the hydropathic character of a protein. J Mol Biol 1982;157:105–132.
3. Engelman DM, Steitz TA, Goldman A, Identifying nonpolar transbilayer helices in amino acid sequences of membrane proteins. Annu Rev Biophys Biophys Chem 1986;15:321–353.
4. Treutlein HR, Lemmon MA, Engelman DM, Brünger AT. The glycophorin A transmembrane domain dimer: sequence-specific propensity for a right-handed supercoil of helices. Biochemistry 1992;31:12726–12732.
5. Lemmon MA, Treutlein HR, Adams PD, Brünger AT, Engelman DM. A dimerization motif for transmembrane α-helices. Nat Struct Biol 1994;1:157–163.
6. Pappu R, Marshall G, Ponder J. A potential smoothing algorithm accurately predicts transmembrane helix packing. Nat Struct Biol 1999;6:50–55.
7. Ben-Tal N, Ben-Shaul A, Nicholls A, Honig B. Free energy determinants of α-helix insertion into lipid bilayers. Biophys J 1996;70:1803–1812.
8. Ben-Tal N, Honig B. Helix-helix interactions in lipid bilayers. Biophys J 1996;71:3046–3050.
9. Forsten K, Kozack R, Lauffenburger D, Subramaniam S. Numerical solution of the nonlinear Poisson-Boltzmann equation for a membrane-electrolyte system. J Phys Chem 1994;98:5580–5586.
10. Åqvist J, Warshel A. Energetics of ion permeation through membrane channels. Biophys J 1989;56:171–182.
11. von Kitzing E, Soumpasis D. Electrostatics of a simple membrane model using Green's functions formalism. Biophys J 1996;71:795–810.
12. Woolf T, Grossfield A, Zuckerman D. Electrostatics of membrane systems—complex, heterogeneous environments. In: Pratt L, Hummer G, editors. Simulation and theory of electrostatic interactions in solution. no. 492 in AIP Conference Proceedings. American Institute of Physics; 1999. p 510–532.
13. Wesson L, Eisenberg D. Atomic solvation parameters applied to molecular dynamics of proteins in solution. Protein Sci 1992;1:227–235.
14. Efremov R, Nolde D, Vergoten G, Arseniev A. A solvent model for simulations of peptides in bilayers: I. Membrane-promoting α-helix formation. Biophys J 1999;76:2448–2459.
15. Efremov R, Nolde D, Vergoten G, Arseniev A. A solvent model for simulations of peptides in bilayers. II. Membrane-spanning α-helices. Biophys J 1999;76:2460–2471.
16. Wiener M, White S. Structure of a fluid dioloeoylphosphatidylcholine bilayer determined by joint refinement of x-ray and neutron diffraction data: III. Complete structure. Biophys J 1992;61:434–447.
17. Yau W-M, Wimley WC, Gawrisch K, White SH. The preference of tryptophan for membrane interfaces. Biochemistry 1998;37:14713–14718.
18. Sansom M. The biophysics of peptide models of ion channels. Prog Biophys Mol Biol 1991;55:139–236.
19. White S, Wimley W. Hydrophobic interactions of peptides with membrane interfaces. Biochim Biophys Acta 1998;1376:339–352.
20. Hristova K, Wimley W, Mishra V, Anantharamiah G, Segrest J, White S. An amphipathic α-helix at a membrane interface: A structural study using a novel x-ray diffraction method. J Mol Biol 1999;290:99–117.
21. Ladokhin A, White S. Folding of amphipathic alpha-helices on membranes: Energetics of helix formation by melittin. J Mol Biol 1999;285:1363–1369.
22. White S, Wimley W. Membrane protein folding and stability: Physical principles. Annu Rev Biophys Biomol Struct 1999;28:319–365.
23. Heller H, Schaefer M, Schulten K. Molecular dynamics simulation of a bilayer of 200 lipids in the gel and in the liquid-crystal phases. J Phys Chem 1993;97:8343–8360.
24. Egberts E, Marrink SJ, Berendsen HJC. Molecular dynamics simulation of a phospholipid membrane. Eur Biophys J 1994;22:423–436.
25. Pastor RW. Molecular dynamics and Monte Carlo simulations of lipid bilayers. Curr Opin Struct Biol 1994;4:486–492.
26. Chiu S-W, Clark M, Balaji V, Subramaniam S, Scott HL, Jakobsson E. Incorporation of surface tension into molecular dynamics simulation of an interface: a fluid phase lipid bilayer membrane. Biophys J 1995;69:1230–1245.
27. Feller SE, Zhang Y, Pastor RW. Computer simulation of liquid/liquid interfaces: II. Surface tension-area dependence of a bilayer and monolayer. J Chem Phys 1995;103:10267–10276.
28. Marrink S, Sok R, Berendsen H. Free volume properties of a simulated lipid membrane. J Chem Phys 1996;104:9090–9099.
29. Bassolino-Klimas D, Alper HE, Stouch TR. Solute diffusion in lipid bilayer membranes: an atomic level study by molecular dynamics simulation. Biochemistry 1993;32:12624–12637.
30. Marrink SJ, Berendsen HJC. Permeation process of small molecules across lipid membranes studied by molecular dynamics simulations. J Phys Chem 1996;100:16729–16738.
31. Damodaran KV, Merz KM Jr, Bager BP. Interaction of small peptides with lipid bilayers. Biophys J 1995;69:1299–1308.
32. Tieleman D, Berendsen H, Sansom M. Surface binding of alamethicin stabilizes its helical structure: molecular dynamics simulation. Biophys J 1999;76:3186–3191.
33. Forrest L, Kukol A, Arkin I, Tieleman D, Sansom M. Exploring models of the influenza a M2 channel—MD simulations in a phospholipid bilayer. Biophys J 2000;78:79–92.
34. Shrivastava I, Sansom M. Simulations of ion permeation through a K channel: molecular dynamics of KcsA in a phospholipid bilayer. Biophys J 2000;78:557–570.
35. Forrest L, Tieleman D, Sansom M. Defining the transmembrane helix of M2 protein from influenza A by molecular dynamics simulations in a lipid bilayer. Biophys J 1999;76:1886–1896.
36. Tieleman D, Sansom M, Berendsen H. An alamethicin channel in a lipid bilayer: Molecular dynamics simulations. Biophys J 1999;76:1757–1769.
37. Woolf TB, Roux B. Molecular dynamics simulation of the gramicidin channel in a phospholipid bilayer. Proc Natl Acad Sci USA 1994;91:11631–11635.
38. Woolf TB, Roux B. Structure, energetics and dynamics of lipid-protein interactions: a molecular dynamics study of the gramicidin a channel in a DMPC bilayer. Proteins 1996;24:92–114.
39. Woolf TB. Molecular dynamics of individual alpha-helices of bacteriorhodopsin in dimyristoyl phosphatidylcholine: I. Structure and dynamics. Biophys J 1997;73:2376–2392.
40. Feller S, Pastor R. Length scales of lipid dynamics and molecular dynamics. In: Altman R, Dunker A, Hunter L, Klein T, editors. Pacific symposium on biocomputing '97. World Scientific; 1996. p 142–150.
41. Warshel A, Levitt M. Folding and stability of helical proteins: carp myogen. J Mol Biol 1976;106:421–437.
42. Warshel A, Russell ST. Calculations of electrostatic interactions in biological systems and in solutions. Q Rev Biophys 1984;17:283–422.
43. McQuarrie DA. Statistical thermodynamics. Mill Valley, California: University Science Books; 1973.
44. Israelachvili J. Intermolecular and surface forces. 2nd ed. San Diego, California: Academic Press; 1991.
45. Hinsen K. The molecular modeling toolkit: a new approach to molecular simulations. J Comp Chem 2000;21:79–85.
46. Cornell WD, Cieplak P, Bayly C, Gould IR, Merz KM Jr, Ferguson D, Spellmeyer D, Fox T, Cauldwell J, Kollman P. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. J Am Chem Soc 1995;117:5179–5197.
47. Radzicka A, Wolfenden R. Comparing the polarities of the amino acids: side-chain distribution coefficients between the vapor phase, cyclohexane, 1-octanol, and neutral aqueous solution. Biochemistry 1988;27:1664–1670.
48. Wolfenden R, Andersson L, Cullis P, Southgate C. Affinities of amino acid side chains for solvent water. Biochemistry 1981;20:849–855.
49. Wimley WC, White SH. Membrane partitioning: distinguishing bilayer effects from the hydrophobic effect. Biochemistry 1993;32:6307–6312.
50. de Planque M, Greathouse D, Koeppe R II, Schäfer H, Marsh D, Killian J. Influence of lipid/peptide hydrophobic mismatch on the thickness of diacylphosphatidylcholine bilayers. a H-2 NMR and ESR study using designed transmembrane alpha-helical peptides and gramicidin A. Biochemistry 1998;37:9333–9345.
51. Killian J, de Planque M, van der Wel P, Salemink I, de Kruijff B, Greathouse D, Koeppe R II. Modulation of membrane structure

and function by hydrophobic mismatch between proteins and lipids. Pure Appl Chem 1998;70:75–82.
52. Killian J, Salemink I, de Planque M, Lindblom G, Koeppe R II, Greathouse D. Induction of nonbilayer structures in diacylphosphatidylcholine model membranes by transmembrane alpha-helical peptides: Importance of hydrophobic mismatch and proposed role of tryptophans. Biochemistry 1996;35:1037–1045.
53. Petrache H, Killian J, Koeppe R, Woolf T. Interactions of WALP peptides with lipid bilayers: Molecular dynamics simulations. Biophys J 2000;78:324a.
54. Fox R, Richards F. A voltage-gated ion channel model inferred from the crystal structure of alamethicin at 1.5 angstrom resolution. Nature 1982;300:325–330.
55. Gibbs N, Sessions RB, Williams PB, Dempsey CE. Helix bending in alamethicin: Molecular dynamics simulations and amide hydrogen exchange in methanol. Biophys J 1997;72:2490–2495.
56. Sessions RB, Gibbs N, Dempsey CE. Hydrogen bonding in helical polypeptides from molecular dynamics simulations and amide hydrogen exchange analysis: alamethicin and melittin in methanol. Biophys J 1998;74:138–152.
57. Tieleman D, Sansom M, Berendsen H. Alamethicin helices in a bilayer and in solution: molecular dynamics simulations. Biophys J 1999;76:40–49.
58. Honig B, Hubbell W. Stability of "salt bridges" in membrane proteins. Proc Natl Acad Sci USA 1984;81:5412–5416.
59. Murray D, Ben-Tal N, Honig B, McLaughlin S. Electrostatic interaction of myristolated proteins with membranes: simple physics, complicated biology. Structure 1997;15:985–989.
60. Murray D, Arbuzova A, Hangyás-Mihályné G, Gambhir A, Ben-Tal N, Honig B, McLaughlin S. Electrostatic properties of membranes containing acidic lipids and adsorbed basic peptides: theory and experiment. Biophys J 1999;77:3176–3188.
61. Gilson MK, Honig B. Calculation of the total electrostatic energy of a macromolecular system. Proteins 1988;4:7–18.
62. Venable R, Zhang Y, Hardy B, Pastor R. Molecular dynamics simulations of a lipid bilayer and of hexadecane: an investigation of membrane fluidity. Science 1993;262:223–225.
63. Chiu S-W, Clark M, Jakobsson E, Subramaniam S, Scoot H. Optimization of hydrocarbon chain interaction parameters: Application to the simulation of fluid phase lipid bilayers. J Phys Chem B 1999;103:6323–6327.
64. Sitkoff D, Ben-Tal N, Honig B. Calculation of alkane to water solvation free energies using continuum solvent models. J Phys Chem 1996;100:2744–2752.
65. Still W, Tempczyk A, Hawley R, Hendrickson T. Semianalytical treatment of solvation for molecular mechanics and dynamics. J Am Chem Soc 1990;112:6127–6129.
66. Qiu D, Shenkin P, Hollinger FP, Still W. The GB/SA continuum model for solvation: a fast analytical method for the calculation of approximate born radii. J Phys Chem A 1997;101:3005–3014.
67. Edinger S, Cortis C, Shenkin P, Friesner R. Solvation free energies of peptides: comparison of approximate continuum solvation models with accurate solution of the Poisson-Boltzmann equation. J Phys Chem B 1997;101:1190–1197.
68. Voglino L, McIntosh T, Simon S. Modulation of the binding of signal peptides to lipid bilayers by dipoles near the hydrocarbon-water interface. Biochemistry 1998;37:12241–12252.
69. Wimley WC, White SH. Experimentally determined hydrophobicity scale for proteins at membrane interfaces. Nat Struct Biol 1996;3:842–848.
70. Florián J, Warshel A. Calculations of hydration entropies of hydrophobic, polar, and ionic solutes in the framework of the Langevin dipoles solvation model. J Phys Chem B 1999;103:10282–10288.
71. Allen M, Tildesley D. Computer simulations of liquids. New York, New York: Oxford University Press; 1987.
72. Sagui C, Darden T. Molecular dynamics simulations of biomolecules: Long-range electrostatic effects. Annu Rev Biophys Biomol Struct 1999;28:155–179.
73. Warshel A, Papazyan A. Electrostatic effects in macromolecules: fundamental concepts and practical modeling. Curr Opin Struct Biol 1998;8:211–217.
74. Hünenberger P, van Gunsteren W. Alternative schemes for the inclusion of a reaction-field correction into molecular dynamics simulations: influence on the simulated energetic, structural, and dielectric properties of liquid water. J Chem Phys 1998;108:6117–6134.
75. Hünenberger P, McCammon J. Effect of artificial periodicity in simulations of biomolecules under ewald boundary conditions: a continuum electrostatics study. Biophys Chem 1999;78:69–88.
76. Yeh I, Berkowitz M. Ewald summation for systems with slab geometry. J Chem Phys 1999;111:3155–3162.
77. Tironi I, Sperb R, Smith P, van Gunsteren W. A generalized reaction field method for molecular dynamics. J Chem Phys 1995;102:5451–5459.
78. Kong Y, Ponder J. Calculation of the reaction field due to off-center point multipoles. J Chem Phys 1997;107:481–492.
79. Sham YY, Warshel A. The surface constraint all atom model provides size independent results in calculations of hydration free energies. J Chem Phys 1998;109:7940–7944.
80. Sham YY, Chu ZT, Warshel A. Consistent calculations of pKa's of ionizable residues in proteins: semi-microscopic and microscopic approaches. J Phys Chem B 1997;101:4458–4472.
81. You T, Bashford D. Conformation and hydrogen ion titration of proteins: a continuum electrostatic model with conformational flexibility. Biophys J 1995;69:1721–1733.
82. Baptista A, Martel P, Petersen S. Simulation of conformational freedom as a function of pH: constant pH molecular dynamics using implicit titration. Proteins 1997;27:523–544.
83. Schaefer M, Sommer M, Karplus M. pH-dependence of protein stability: absolute electrostatic free energy differences between conformations. J Phys Chem B 1997;101:1663–1683.
84. Beroza P, Fredkin D, Okamura M, Feher G. Protonation of interacting residues in a protein by a Monte Carlo method: applications to lysozyme and the photosynthetic reaction center of Rhodobacter spheroides. Proc Natl Acad Sci USA 1991;88:5804–5808.
85. Beroza P, Fredkin D, Okamura M, Feher G. Electrostatic calculations of amino acid titration and electron transfer, $Q_A^- Q_B \rightarrow Q_A Q_B^-$, in the reaction center. Biophys J 1995;68:2233–2250.
86. Gilson M, Sharp K, Honig B. Calculating the electrostatic potential of molecules in solution. J Comp Chem 1988;9:327–335.
87. Antosiewicz J, McCammon J, Gilson M. Prediction of pH-dependent properties of proteins. J Mol Biol 1994;238:415–536.
88. Beitinger H, Vogel V, Möbius D, Rahmann H. Surface potentials and electric dipole moments of ganglioside and phospholipid monolayers: contribution of the polar headgroup at the water/lipid interface. Biochim Biophys Acta 1989;984:293–300.
89. Chiu S-W, Clark M, Balaji V, Subramaniam S, Jakobsson E. Incorporation of surface tension into molecular dynamics simulations of an interface: a fluid phase lipid bilayer membrane. Biophys J 1995;69:1230–1245.
90. Wymore T, Woolf T. Molecular details of POPC bilayers: molecular dynamics calculations. Biophys J 2000;78:324a.
91. Grossfield A, Woolf T. Interactions between indole and lipid bilayers studied by molecular dynamics and free energy calculations. Biophys J 2000;78:324a.

## APPENDIX: DERIVATIONS

In this section, we will derive the partition function for a permanent, freely reorientable dipole in a constant electric field. The partition function will then be used to derive the thermally averaged polarization of a dipole in a field, and the orientational entropy it loses upon polarization.

### Partition Function

We begin by writing the energy of a permanent dipole $\vec{\mu}$ in an electric field $\vec{E}$:

$$U = -\vec{\mu} \cdot \vec{E} \tag{A-1}$$

If we define $\mu_0 = |\vec{\mu}|$ and $E_0 = |\vec{E}|$, we can rewrite this as

$$U = -\mu_0 E_0 \cos\theta \tag{A-2}$$

where $\theta$ is the angle between $\vec{\mu}$ and $\vec{E}$. The partition function can then be written as the integral over all dipole conformations:

$$Q = \int e^{-U(\vec{E},\vec{\mu})/k_BT} d\vec{\mu}$$

$$= \int_0^{2\pi} \int_0^{\pi} e^{\mu_0 E_0 \cos\theta/k_BT} \sin\theta \, d\theta \, d\phi \quad \text{(A-3)}$$

Because the integrand is independent of the azimuthal angle $\phi$, integrating over $\phi$ gives $2\pi$. Changing variables to set $x = \cos\theta$, we get

$$Q = 2\pi \int_{-1}^{1} e^{\mu_0 E_0 x/k_BT} \, dx$$

$$= \frac{2\pi k_B T}{\mu_0 E_0} (e^{\mu_0 E_0/k_BT} - e^{-\mu_0 E_0/k_BT})$$

$$= \frac{4\pi k_B T}{\mu_0 E_0} \sinh\left(\frac{\mu_0 E_0}{k_B T}\right) \quad \text{(A-4)}$$

This expression for the partition function is exact for a permanent dipole with fixed location.[43]

**Average Energy**

The thermally averaged energy can be calculated from the canonical partition function using the standard thermodynamic relation[43]

$$\langle U \rangle = k_B T^2 \left(\frac{\partial \ln Q}{\partial T}\right) \quad \text{(A-5)}$$

Using Eqs. A-4 and A-5, and defining $y = \mu_0 E_0/k_B T$, we get

$$\langle U \rangle = -\mu_0 E_0 \left(\coth y - \frac{1}{y}\right)$$

$$= -\mu^L E_0 \quad \text{(A-6)}$$

which is the Langevin expression, used in Eq.[43,41,42]

**Entropy Loss due to Electric Field**

For a system with canonical (NVT) partition function $Q$, the entropy can be calculated as[43]

$$S = k_B \ln Q + k_B T \left(\frac{\partial \ln Q}{\partial T}\right) \quad \text{(A-7)}$$

Combining this with Equation A-4, and again defining $y = \mu_0 E_0/k_B T$, we can write

$$S = k_B[\ln(\sinh y) - y \coth y - \ln y + \ln 4\pi + 1] \quad \text{(A-8)}$$

Without loss of generality, we can define $S_0 = \ln 4\pi + 1$ as our standard state, such that

$$\lim_{\vec{E} \to 0} \Delta S(\vec{E}) = 0$$

$$lim_{\vec{E} \to 0} \Delta A(\vec{E}) = 0 \quad (9)$$

Physically, this choice of $S_0$ means we are using the dipole in the absence of an applied field as the reference state.

**Additivity of Dipole Entropies**

The partition function for a lattice of $N$ permanent dipoles can be written

$$Q_N = \int d\vec{\mu}_1 \ldots d\vec{\mu}_N \exp\left[\left(\sum_{i,j>i}^{N_{dip}} - U(\vec{\mu}_i, \vec{\mu}_j)\right)/k_BT\right] \quad \text{(A-10)}$$

where $d\vec{\mu}_i$ denotes integration over all orientations of dipole $i$, and $U(\vec{\mu}_i, \vec{\mu}_j)$ is the interaction energy between the dipole vectors $\vec{\mu}_i$ and $\vec{\mu}_j$. Furthermore, we can calculate the partition function for one of the dipoles in the lattice as

$$Q_{dip,i} = \int \rho_i(\vec{E}) d\vec{E} \int \exp(-U(\vec{E}, \vec{\mu}_i)/k_BT) d\vec{\mu}_i \quad \text{(A-11)}$$

where

$$\rho_i(\vec{E}_0) = \int d\vec{\mu}_1 \ldots d\vec{\mu}_N \left(\prod_j^{N_{dip}} \rho_j(\vec{\mu}_j)\right) \delta\left(\vec{E}_0 - \sum_j^{N_{dip}} \vec{E}_j\right) \quad \text{(A-12)}$$

In these equations, $\rho_i(\vec{E})$ is the probability distribution for the electric field at dipole $i$, $\rho(\vec{\mu}_j)$ is the probability distribution for orientations of dipole $j$, $\vec{E}(\vec{\mu}_j)$ is the electric field for $\vec{\mu}_j$, $U(\vec{E}, \vec{\mu}_i)$ is interaction energy between the dipole and electric field $\vec{E}$, and $\delta(x)$ is the Dirac delta function. The integration and product in Eq. A-12 exclude dipole $i$. Since $\rho(\vec{\mu}_j)$ depends on the orientations of all other dipoles in the lattice, this expression cannot be used to simplify $Q_N$.

However, in the present work, we calculate the electric field due to the average polarization of the environment and explicitly neglect fluctuations in the electric field. Effectively, we impose the condition

$$\rho_i(\vec{E}) = \delta(\vec{E} - \langle \vec{E}_i \rangle) \quad \text{(A-13)}$$

where $\langle \vec{E}_i \rangle$ is the average electric field at dipole $i$, calculated using the iterative approach described in the Dipole Lattice Solvation Energies section. Substituting Eq. A-13 into Eq. A-11 allows us to rewrite the partition function without explicit dependence on the state of the other dipoles. This, in turn, allows us to write

$$Q_N = \prod_i^{N_{dip}} Q_{dip,i} \quad \text{(A-14)}$$

If we substitute Eq. A-14 into Eq. A-7, we get

$$S = k \sum_i^{N_{dip}} \ln Q_{dip,i} + k_B T \sum_i^{N_{dip}} \frac{\partial}{\partial T} \ln Q_{dip,i} \quad \text{(A-15)}$$

$$= \sum_i^{N_{dip}} S_i \quad \text{(A-16)}$$

In words, it is correct to write the entropy for a lattice of dipoles as the sum of individual dipole entropies in the context of a Langevin dipole model.