

# Using the Wimley–White Hydrophobicity Scale as a Direct Quantitative Test of Force Fields: The MARTINI Coarse-Grained Model

Gurpreet Singh<sup>†</sup> and D. Peter Tieleman<sup>\*,†,‡</sup>

<sup>†</sup>Department of Biological Sciences, and <sup>‡</sup>Institute for Biocomplexity and Informatics, 2500 University Drive N.W., University of Calgary, Calgary, Alberta, Canada T2N 1N4

**ABSTRACT:** The partitioning of proteins and peptides at the membrane/water interface is a key step in many processes, including the action of antimicrobial peptides, cell-penetrating peptides, and toxins, as well as signaling. To develop a computational model that can be used to accurately represent such systems, the underlying model must be able to quantitatively represent the partitioning preferences of amino acids in the lipid membrane. The MARTINI model provides a consistent set of parameters for building coarse-grained models of systems involving lipids and proteins. Even though MARTINI is parametrized to reproduce the partitioning behavior of small molecules, its ability to reproduce partitioning preferences of amino acids at lipid/water interfaces has never been tested. In this study, we measured the partitioning free energies of side chains of amino acids using alchemical simulations and umbrella sampling. The pentapeptides of sequence Ac-WLXLL were simulated at the POPC/water and cyclohexane/water interfaces using MARTINI, and the computed free energies were compared with the Wimley–White hydrophobicity scale. The free energy values obtained using the free energy perturbation, thermodynamic integration, and umbrella sampling methods were compared to gain insight into the most efficient method and the degree of sampling required to obtain statistically accurate free energies for use with atomistic force fields in future work. With the standard MARTINI water model, the amino acids D, E, K, and R were found to be significantly too favorable in hydrophobic environments, whereas with the polarizable water model, the amino acids D, E, K, and R were found to give correct free energies of partitioning. The amino acids P and F showed significant deviations from the experimental values. This model system will be used in future improvements to the MARTINI model.

## 1. INTRODUCTION

Over the past few decades, computer simulations of biomolecules have become a useful tool in understanding the dynamics and mechanism of function with atomistic details.<sup>1</sup> New sampling methods coupled with increased computational resources have made it possible to simulate the folding of small peptides and proteins.<sup>2–4</sup> The underlying parameters or force fields used in atomistic representations of proteins have matured to the point that free energies calculated using computer simulation methods are in quantitative agreements with experiments.<sup>5</sup>

At the current state of simulation algorithms and computer hardware, the time scales of typical atomistic simulations are limited to (sub)microseconds, and the system size is limited to hundreds of thousands of atoms. Sampling all relevant degrees of freedom still remains a major challenge in biomolecular simulations, considering that many interesting biological processes occur on time scales beyond the reach of atomistic simulations.

To simulate processes such as vesicle fusion, pore formation in membranes, and the formation of protein complexes, several simplified models have been developed.<sup>6–8</sup> The use of these coarse models allows for simulations of larger systems, for longer time scales, while still providing significant structural detail. Marrink and co-workers developed one such coarse-grained model for simulations of lipids and surfactant, coined MARTINI.<sup>9,10</sup> In the MARTINI model, several atoms are grouped together in one bead that interacts with other beads through an effective potential. This model was successfully applied to study long-time-scale phenomena such as lipid monolayer collapse, domain formation in vesicles, and vesicle fusion.<sup>11–13</sup>

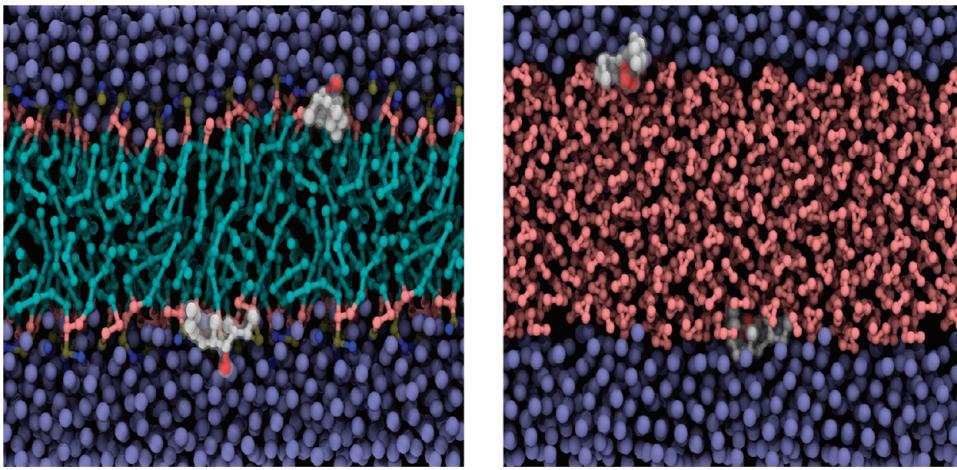
Consistent with the original design philosophy, the interaction parameters for the beads representing amino acids were chosen to reproduce the experimental partitioning free energies of amino acid analogues between water and cyclohexane.<sup>14</sup> The experimental data on partitioning free energies of amino acid analogues have also been used to parametrize recent versions of the GROMOS force field.<sup>15</sup> However, as free energy calculations are computationally intensive, small-molecule analogues rather than complete amino acids or peptides are generally used and compared with experimental data to obtain insight into the performance of various classical force fields.<sup>16–19</sup>

To test our ability to accurately model lipid–protein interactions, critical experimental data are essential, but such data are rare.<sup>20</sup> In addition, some accurate data are difficult to reproduce by simulations because of the time scales or system complexity that is required for an accurate comparison.<sup>21,22</sup> Wimley and White created a whole-residue hydrophobicity scale by measuring the partitioning free energies of model peptides of the sequence Ac-WLXLL, where X can be any of the 20 amino acids, at a phospholipid (1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine, POPC)/water interface and in a water/octanol mixture.<sup>23,24</sup> This scale is routinely used in hydropathy plots for the prediction of transmembrane regions in membrane proteins.<sup>25,26</sup> These peptides, because of their small size, can be simulated relatively easily and can be used to gain further insight into the thermodynamics of lipid–protein interactions. Some of these peptides have previously been studied using atomistic simulations at

Received: April 15, 2011

Published: June 06, 2011





**Figure 1.** Snapshots of simulations depicting WLELL peptides at the POPC/water interface (left) and WLFL peptides at the cyclohexane/water interface (right). Water beads are colored ice blue, cyclohexane is colored pink, peptide beads are colored white, and X3 beads are colored red.

cyclohexane/water, octanol/water, and 1,2-dioleoyl-*sn*-glycero-3-phosphocholine (DOPC)/water interfaces and were found to adsorb readily to the interface.<sup>27,28</sup>

The potential of mean force (PMF) profiles of amino acid analogues across the DOPC bilayer were studied by Macallum et al.<sup>29,30</sup> In that study, the computed bulk to interface transfer free energies were compared with free energies of partitioning of Wimley–White peptides at octanol/water and POPC/water interfaces, and moderate correlations (correlation coefficients of 0.84 and 0.61, respectively) were observed. These results were attributed to the fact that the Wimley–White scale represents the free energy of partitioning of side chains in the context of the pentapeptide rather than partitioning of small-molecule analogues.

In this study, we computed the partitioning free energies of side chains of all 20 amino acids as part of the WLXLL peptide at the cyclohexane/water and POPC/water interfaces using the MARTINI model. The latter set of values can be directly compared with the experimental results of the Wimley–White scale. The main motivations of this study were to investigate the feasibility of a direct calculation of the Wimley–White scale for use in the development and testing of force fields, to test and possibly improve the MARTINI model's ability to describe lipid–protein interactions, and to establish a well-defined experimental test system and simulation protocol that will be useful in the future development of the MARTINI model.

## 2. METHODS

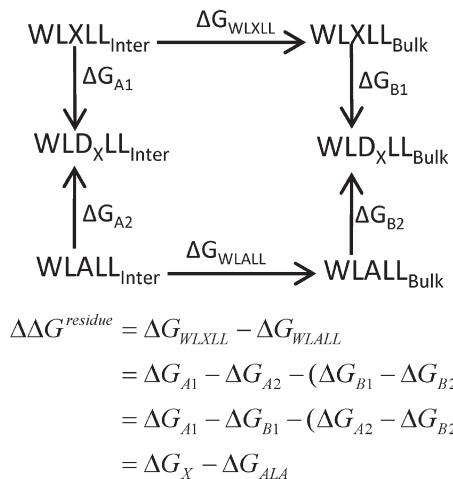
**2.1. Coarse-Grained Simulations.** All simulations were performed with the GROMACS software (version 4.0 and above).<sup>31</sup> In all simulations, a leapfrog integrator was used. All systems were simulated at a temperature of 300 K and a pressure of 1 bar using periodic boundary conditions. Peptides with the sequence WLXLL, where X can be any of the 20 amino acids, were simulated at the cyclohexane/water interface, the POPC/water interface, and in bulk water, using MARTINI force field v2.1<sup>14</sup> for protein parameters and v2.0<sup>10</sup> for lipid parameters. In the MARTINI model, four heavy atoms are generally represented by a single interaction site. Polar (P), nonpolar (N), apolar (C), and charged (Q) are the four main types of interaction sites. These primary interaction types are further divided into subtypes based on either hydrogen-bonding capabilities or degree of

polarity. To model the N-terminal cap of the Wimley–White peptides, the default N-terminal bead was replaced with the P<sub>5</sub> bead.

Amino acids were represented in their default protonations state for pH 8; that is, the D and E side chains were negatively charged, the K and R side chains were positively charged, and the rest were neutral. The C terminus was modeled with a negative charge to mimic the deprotonated state at pH 8, except for X = K or R, which were represented by a neutral C-terminus bead (P<sub>5</sub>) to mimic pH 2. Additional simulations for the system with X = A were carried out with a neutral C terminus.

For the systems containing either the cyclohexane/water or POPC/water interface, the initial configuration was prepared by inserting two peptides in the water phase and deleting the overlapping water molecules. In the case of the system containing the POPC/water interface, the initial adsorption of the peptides at opposite leaflets of the bilayer was facilitated by subjecting the peptides to a constant acceleration of 0.05 nm ps<sup>-2</sup> in opposite directions, for a duration of 30 ns. The system was further simulated without any acceleration for 500 ns, during which time the temperature was maintained by coupling the system to a Nosé–Hoover<sup>32,33</sup> thermostat with a coupling constant of 1.5 ps and the pressure was maintained by coupling the system semi-isotropically to a Parrinello–Rahman barostat<sup>34</sup> with a coupling constant of 2.5 ps. The systems containing cyclohexane/water were coupled to a Berendsen pressure bath,<sup>35</sup> in only the Z dimension (perpendicular to the cyclohexane/water interface), with a coupling constant of 1 ps and were simulated for an additional 2  $\mu$ s. Snapshots of the cyclohexane/water and POPC/water systems with peptides are shown in Figure 1 and were prepared using VMD software.<sup>36</sup>

**2.1.1. Free Energy Calculations.** The free energies of partitioning of the side-chain residues can be calculated using the thermodynamic cycle shown in Figure 2. Simulations in which a chemical species is transformed into another through an unphysical pathway are referred to as alchemical simulations.<sup>37</sup> Alternatively, a system can be transformed from one state into another by performing simple displacements. Thus, the free energies of partitioning can be obtained by either following the vertical lines using alchemical simulations or by following the horizontal lines and computing the potential of mean force as a function of distance from the interface, for example, by using



**Figure 2.** Thermodynamic cycle for the calculation of free energies of partitioning of amino acids,  $\Delta\Delta G^{\text{residue}}$ , at interfaces.  $\text{WLXLL}_{\text{Inter}}$  and  $\text{WLXLL}_{\text{Bulk}}$  represent the peptide at the interface and in the bulk, respectively.  $\text{WLD}_X\text{LL}$  represents a peptide in which the side-chain atoms of residue X have been converted to dummy atoms that have no Lennard-Jones or Columbic interactions and the backbone bead is converted to  $P_5$ .

umbrella sampling (US).<sup>38</sup> Free energies from alchemical simulations can be computed using the thermodynamic integration (TI) method, in which the derivative of the Hamiltonian is computed along a reaction coordinate describing the alchemical transformation, which is then integrated numerically to obtain the free energy difference,<sup>39</sup> or using the free energy perturbation (FEP) method, in which the free energy difference is estimated from the exponential average of the difference in energy between the states.<sup>40</sup> To calculate the free energy estimates from the data obtained using either the FEP or US approach, methods such as the weighted histogram analysis method (WHAM) or the multiple Bennett acceptance ratio (MBAR) can be employed.<sup>41,42</sup>

The partitioning free energies for amino acids were calculated at the cyclohexane/water interface and POPC/water interface using the thermodynamic cycle shown in Figure 2. At the cyclohexane/water interface, free energies were computed using both alchemical simulations and umbrella sampling. For alchemical simulations, the free energies were computed using both FEP and TI. The MBAR method implemented in the pyMBAR<sup>41</sup> program was used to estimate free energies and uncertainties from FEP data, and for TI, the uncertainties were calculated using the method described by Hess, which involves fitting an analytic function to a standard error estimate of a measured observable ( $\partial H / \partial \lambda$ , in this case) as a function of block size.<sup>43</sup>

The initial configurations of production runs were used as the starting configurations for each intermediate state, and the peptides were accelerated toward the interface for 200 ps. Nineteen equally spaced intermediate states were defined using the coupling parameter  $\lambda$ , which linearly switches off Lennard-Jones (LJ) and Columb interactions of the perturbed side chain as  $\lambda$  goes from 0 to 1. To avoid singularities, soft-core interactions were used for LJ interactions.<sup>44</sup> All simulations included 10 ns of equilibration, followed by production runs of 0.5  $\mu$ s for each  $\lambda$  value. For TI, the numerical integration of  $\partial H / \partial \lambda$  was carried out using Simpson's rule. The potential energy differences between all intermediate states are required in MBAR and were obtained

by recalculating the potential energies from the trajectories using the -rerun option of the GROMACS mdrun program. For peptides containing charged residues, free energy calculations were also performed using the polarizable MARTINI water model, and the charges were decoupled separately.<sup>45</sup>

The PMF profiles for pulling the entire peptide from an interface to the bulk were computed using the distance between the center of mass (COM) of cyclohexane and the COM of the peptide as the reaction coordinate. Thirty-one equally spaced windows (0.1 nm) were used. The COM of the peptide was held at its position by applying a harmonic potential with a force constant of 1000 kJ nm<sup>-2</sup>. After initial equilibration, data were collected for 250 ns for each window, and the entire simulation was repeated twice. The data were analyzed using the weighed histogram method (WHAM) as implemented in GROMACS.<sup>46</sup> The free energy of adsorption at the interface from the bulk was calculated as  $\Delta G_X = -RT \ln[\int_{z_f}^{z_s} e^{-\Delta G(z)/RT} dz]$ , where  $z_s = 1.5$  nm and  $z_f = 4.0$  nm represent the coordinates of the cyclohexane/water interface and bulk water, respectively. The integrations were carried out with numerical integration methods using Simpson's rule. The averages and standard errors were computed using the two data sets and are reported.

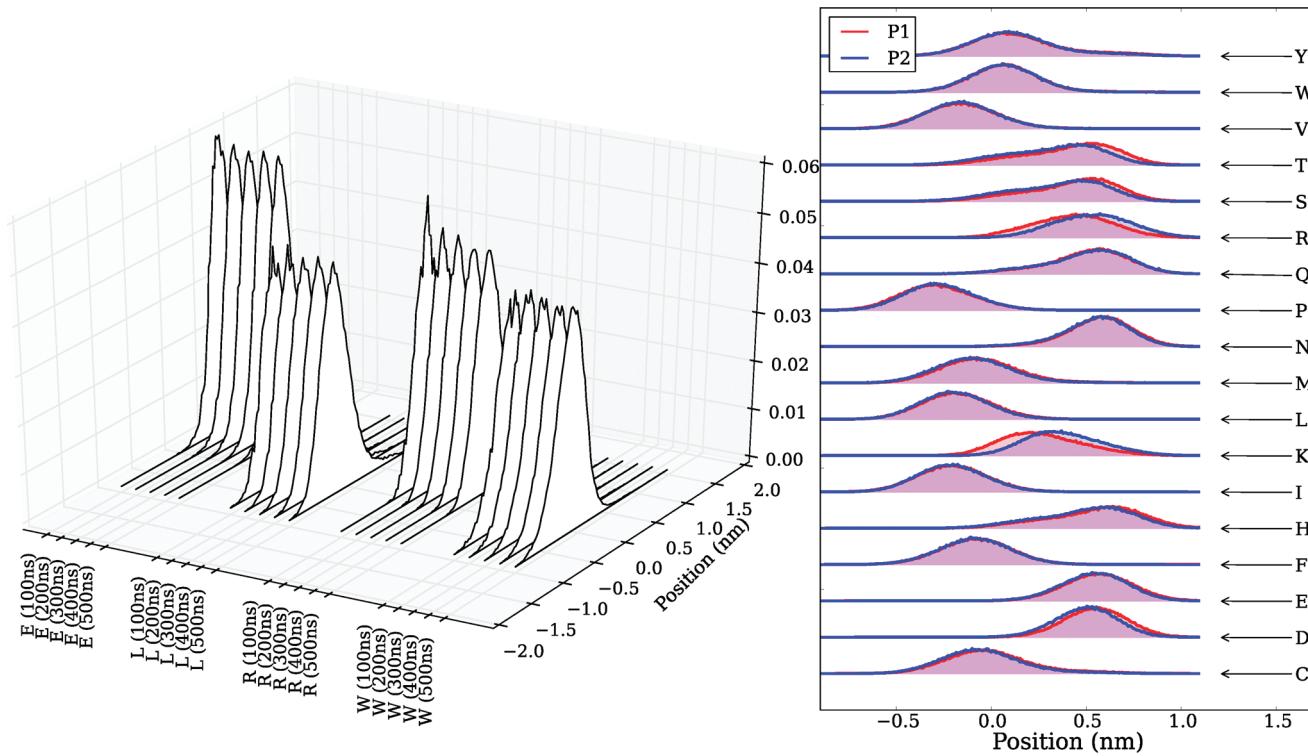
All peptides were modeled without any dihedral potential on consecutive backbone beads. However, to determine the effects of different pseudodihedral parameters defining the backbone geometry, two peptides (with X = I or E) were chosen, and the free energy of partitioning at the cyclohexane/water interface was computed with different backbone parameters representing helix, extended, turn, and bend structures and with a pseudodihedral potential derived from an atomistic simulation (only WLELL). In all cases, the backbone was represented by a  $P_5$  bead.

**2.2. All-Atom Simulations.** All atomistic simulations were performed using the OPLS force field and SPCE water model.<sup>47,48</sup> The N terminus of the peptide was capped with an acetyl group, and the C terminus was deprotonated. To obtain the backbone dihedral potential for the coarse-grained WLELL peptide, 40 atomistic simulations of 40-ns duration were carried out. The probability densities of pseudodihedrals between four consecutive C- $\alpha$  carbons were measured, and the dihedral potential was derived by the Boltzmann inversion of the average probability density.

### 3. RESULTS AND DISCUSSION

Experimental studies of these peptides in POPC/water have shown that the peptides partition at the interface, without deep penetration into the hydrocarbon core.<sup>24</sup> This behavior was attributed to the lack of secondary structure in the peptides, resulting in a high energy cost of partitioning of the non-hydrogen-bonded peptide backbone into the hydrocarbon core. In agreement with the experimental observations, in our simulations at both interfaces, the peptides were adsorbed at the interface, with the charged C terminus oriented toward water.

Because we used two peptides (one at each interface), we looked at the convergence of the simulations by comparing the time evolution of the average probability density of the center of mass (COM) of side chains at the interface. Figure 3 shows the time evolution of side chain X for a few selected peptides. The converged distributions were obtained within 300 ns of simulation, and further sampling did not change the shape of the distributions. Also, similar distributions were obtained from the simulations of peptides at the cyclohexane/water interface, with



**Figure 3.** (Left) Probability density of the side-chain COM as a function of the position at the interface and the simulation length at the POPC/water interface for a few selected peptides. (Right) Distributions of the side-chain COM of amino acids at the cyclohexane/water interface for simulations with (P2) and without (P1) an angle potential. For comparison of the distributions of different amino acids, the *x* axis was adjusted such that the intersection of the densities of cyclohexane and water was always at 0. In the case of the POPC/water interface, the center between  $Q_0$  and COM of the first two  $C_1$  positions was set as 0.

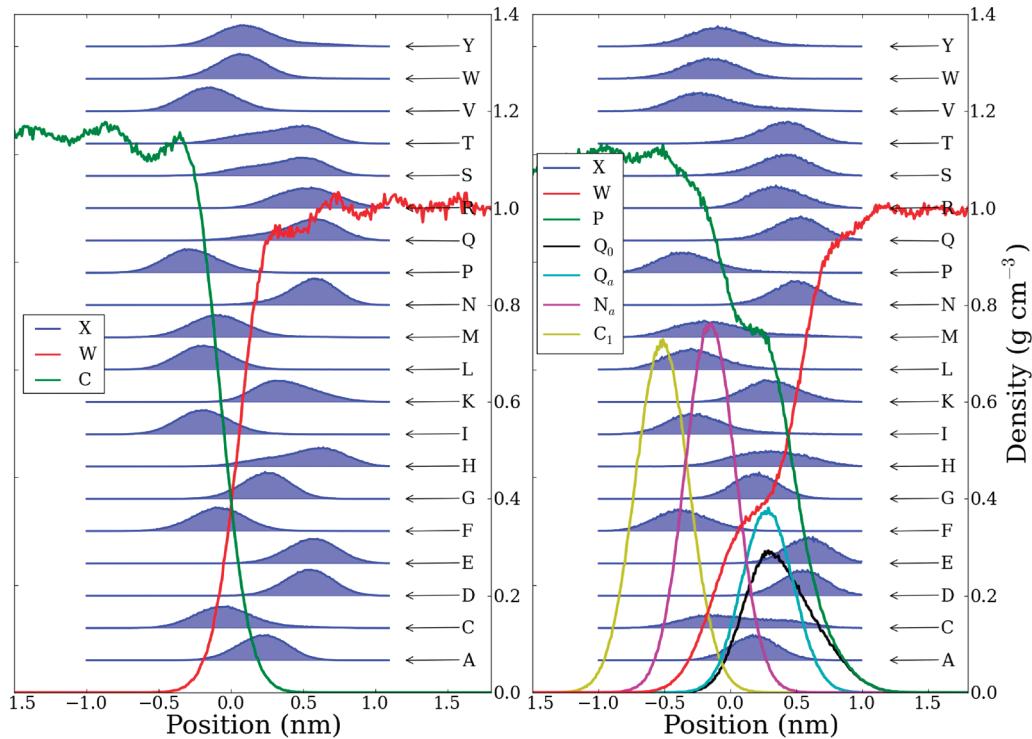
or without angle potential parameters. Figure 4 shows the distributions of side chain X for all peptides. At the cyclohexane/water interface, the density of water drops from its bulk value to almost 0 within  $\sim 1$  nm. The change in density of water is more gradual near the membrane interface, with the decrease in bulk water density marked by an increase in density of the choline group ( $Q_0$ ), followed by that of the phosphate group ( $Q_a$ ), both of which interact most favorably with the water bead ( $P_4$ ). The region of the glycerol ester group ( $N_a$  beads) is chemically most diverse, as all of the beads have a presence in that region. The decrease in density of water from its bulk value to almost 0 occurs in  $\sim 1.7$  nm near POPC. At the POPC/water interface, the density distributions of D and E coincide with those of beads  $Q_0$  and  $Q_a$ , whereas the F, I, L, and P density distributions coincide with those of the  $C_1$  and  $N_a$  beads. Even though the POPC/water interface is far more complex and chemically diverse than the cyclohexane/water interface, certain similarities between the side-chain distributions at the two interfaces can be seen. All of the distributions have a single maximum, and the position of this maximum correlates with the hydrophobicity of the amino acids, with polar side chains preferring water and apolar side chains oriented away from bulk water. The similarity in side-chain distributions at the cyclohexane/water and POPC/water interfaces indicates that the partial density of water at the interface has the strongest effect on the side-chain positions.

The side-chain distributions of the rest of the amino acids in the pentapeptide (W1, L2, L4, and L5) for a few selected peptides are shown in Figure 5. The type of amino acid at X does not significantly affect the positions of the rest of the amino

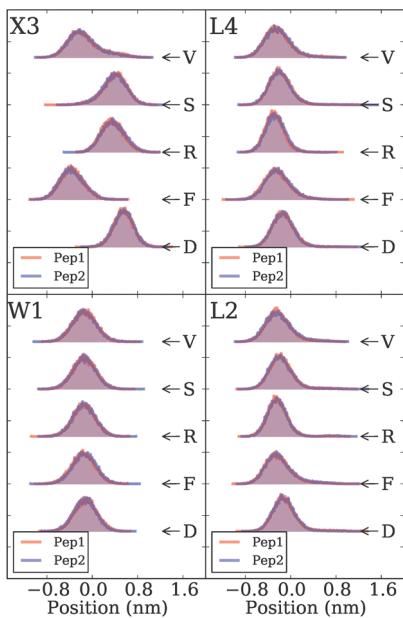
acids, indicating that the peptides have enough conformational flexibility to allow almost independent orientation of the individual X amino acid. The distributions obtained from the two peptides in the system are very similar, indicating that the amount of sampling was adequate.

**3.1. Free Energies.** The partitioning free energies of side chains at the cyclohexane/water interface obtained using the FEP, TI, and US methods are reported in Table 1. Although similar free energy estimations were obtained for all three methods, the estimated errors were smallest for the free energy estimates from FEP using the MBAR method. As both the TI and US methods involve numerical integration of the primary results ( $\partial H / \partial \lambda$  and an average force as a function of distance, respectively), the resulting values have comparatively larger uncertainties. However, PMF profiles along physically relevant reaction coordinates can provide additional insight; for example, the PMF profiles at the cyclohexane/water interface of one representative amino acid from each group are shown in Figure 6. All of these profiles have a single minimum at the interface, indicating that the adsorption of the peptide from bulk water is barrierless, and the free energy increases rapidly while moving away from the interface in either direction. Very similar PMF profiles for Q, E, and A also indicate that the standard MARTINI model does not sufficiently distinguish between water and hydrophobic environments for charged particles.

The free energies of partitioning of side-chain residues at the POPC/water interface,  $\Delta\Delta G^{\text{residue}}$ , as computed using the FEP method with the thermodynamic cycle in Figure 2, are shown in Figure 7. Based on experimental free energies of partitioning, the amino acids can be roughly divided into five groups, as shown in



**Figure 4.** Distributions of the side-chain COM of amino acids at the cyclohexane/water (left) and POPC/water (right) interfaces. The side-chain distributions (legend label X) are filled. The partial densities of water, cyclohexane, and POPC are labeled as W, C, and P respectively. In the case of the POPC/water interface, the density distributions of  $Q_0$ ,  $Q_a$ ,  $N_a$ , and  $C_1$  beads are also shown. For the  $N_a$  and  $C_1$  distributions, the COMs of the two  $N_a$  beads and the first two  $C_1$  beads, respectively, were used. For comparison of the distributions of different amino acids, the x axis was adjusted such that the intersection of the densities of cyclohexane and water was always at 0. In the case of the POPC/water interface, the center between  $Q_0$  and the COM of the first two  $C_1$  positions was set as 0. The average densities of water, cyclohexane, and POPC as functions of the Z dimension are also plotted.



**Figure 5.** Distributions of the side-chain COM of amino acids at the POPC/water interface. Pep1 and Pep2 are the two peptides. The distributions obtained from the two peptides (one at each interface) are shown for W1 (lower left panel), L2 (lower right panel), X3 (top left panel), and L4 (top right panel) side chains.

Figure 7. Group I contains the amino acids that are most hydrophobic on the Wimley–White scale, in particular those

with aromatic side chains; group II consists of amino acids with aliphatic side chains such as L and I and side chains containing sulfur (C and M); group III consists of amino acids whose partitioning free energies are very similar to that of alanine (A, S, T, V, H, and G); group IV includes side chains containing amide groups and proline (N, Q, and P); and group V consists of the charged amino acids (D, E, K, and R). Considering the coarse-grained description of the system, the experimental and computed values are in good agreement for groups II and III. F, P, and the amino acids in group V have the largest deviations from the experimental values.

The free energies obtained at the POPC/water and cyclohexane/water interfaces are compared in Figure 8. The side chains of amino acids P and L are the most hydrophobic at the cyclohexane/water interface, whereas W is the most hydrophobic at the POPC/water interface. Also, the cyclohexane-to-water transfer free energies are generally more positive than the transfer free energies for POPC/water, except for W, Y, and H. In membrane proteins, the W and Y residues are known to be preferentially located at the regions corresponding to the membrane interfaces. This preference is generally attributed to several factors, including their flat rigid shape, cation–π interactions, and hydrogen bonding.<sup>49–51</sup>

**3.1.1. Group V.** For side chains containing a net charge, Figure 4 shows that at the cyclohexane/water interface, the peptides orient themselves to position these charged side chains toward the bulk water, thus correctly preferring the water phase over cyclohexane. However, the calculated free energies of partitioning are close to 0, as the free energies of annihilation of charged

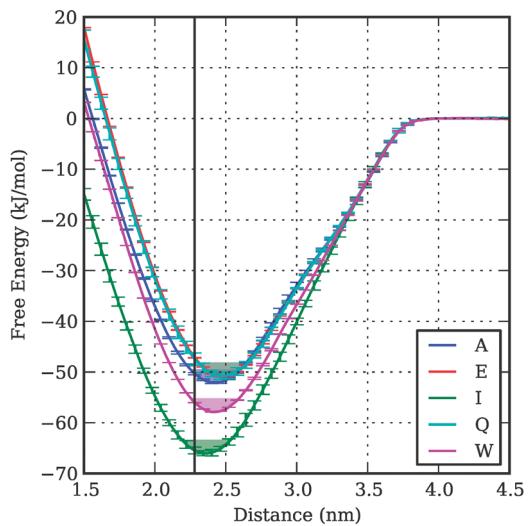
**Table 1.** Values of  $\Delta\Delta G$  (kJ mol<sup>-1</sup>) at the Cyclohexane/Water Interface Calculated Using the FEP, TI, and US Methods

AA	$\Delta\Delta G_{\text{FEP}}^{\text{res}}$	$\Delta\Delta G_{\text{TI}}^{\text{res}}$	$\Delta\Delta G_{\text{US}}^{\text{res}}$
C	6.09 ± 0.03	6.1 ± 0.1	6.2 ± 0.1
D	0.68 ± 0.04	0.7 ± 0.1	-0.6 ± 0.1
E	0.34 ± 0.04	0.3 ± 0.1	-1.4 ± 0.2
F	10.45 ± 0.05	10.5 ± 0.1	10.4 ± 0.2
G	-1.011 ± 0.001	-1.02 ± 0.02	-1.5 ± 0.2
H	-2.20 ± 0.05	-2.2 ± 0.2	-1.9 ± 0.2
I	14.24 ± 0.03	14.2 ± 0.1	13.8 ± 0.3
K	1.76 ± 0.05	1.8 ± 0.1	2.6 ± 0.2
L	14.68 ± 0.03	14.7 ± 0.1	14.5 ± 0.2
M	7.83 ± 0.04	7.8 ± 0.1	7.9 ± 0.1
N	-1.38 ± 0.04	-1.4 ± 0.1	-1.2 ± 0.4
P	20.24 ± 0.04	20.3 ± 0.1	20.0 ± 0.1
Q	-1.37 ± 0.04	-1.4 ± 0.1	-1.3 ± 0.3
R	-1.12 ± 0.05	-1.2 ± 0.1	0.4 ± 0.5
S	-0.27 ± 0.03	-0.3 ± 0.1	-0.4 ± 0.1
T	-0.25 ± 0.04	-0.2 ± 0.1	-0.1 ± 0.3
V	11.52 ± 0.03	11.5 ± 0.1	11.1 ± 0.3
W	5.43 ± 0.06	5.4 ± 0.2	5.6 ± 0.1
Y	2.15 ± 0.05	2.2 ± 0.1	2.0 ± 0.3

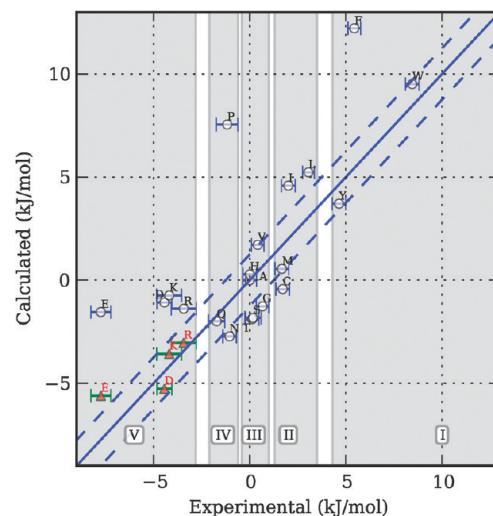
side chains in the bulk and at the interface are very similar, suggesting that the charged side chains are effectively in bulk water. At the POPC/water interface, side chains D and E are close to positively charged choline beads, whereas K and R have maxima close to the negatively charged phosphate bead of POPC. Even in this case, the calculated free energy values are close to 0. The water molecule in the standard MARTINI model is represented by a Lennard-Jones bead, and the interaction level of charged beads with the water bead is the same as for other polar beads, resulting in underestimation of the hydration free energies of charged side chains. This problem was recently fixed by introducing point charges into the water model to make the water molecules polarizable and by modifying the interactions of Q-type particles.<sup>45</sup> Using this polarizable water model in the free energy calculations of charged side chains resulted in significant improvements (see Figure 7) at the POPC/water interface. In Figure 9, the side-chain COM distributions obtained using the polarizable water model are compared with those obtained from standard MARTINI. The side-chain COM distributions with respect to the interface in the two cases are similar.

In comparison to POPC, cyclohexane is far more apolar than the lipid head groups. Therefore, it can be argued that the free energy of partitioning of charged residues should be more negative; alternatively, the cyclohexane/water interface is narrower than the POPC/water interface, and reorientation of peptides at cyclohexane/water interface could place the charged side chains effectively in bulk water, resulting in low partitioning free energies for the charged residues as compared to those at the POPC/water interface. Because of the unavailability of experimental data for these peptides at the cyclohexane/water interface, it is not possible to validate the computed values.

**3.1.2. F and P.** In the MARTINI force field, the aromatic side chains are modeled using ring particle beads that have lower masses than other beads, with ring–ring interactions that are scaled ( $\sigma = 0.43$  and  $\epsilon$  is scaled to 75% of its original value) compared to interactions with the rest of the beads. From Figure 4, it can be seen that the side chain of amino acid F predominantly interacts with the tail beads (bead type C<sub>1</sub>) of POPC molecules. The computed free energy of partitioning of F

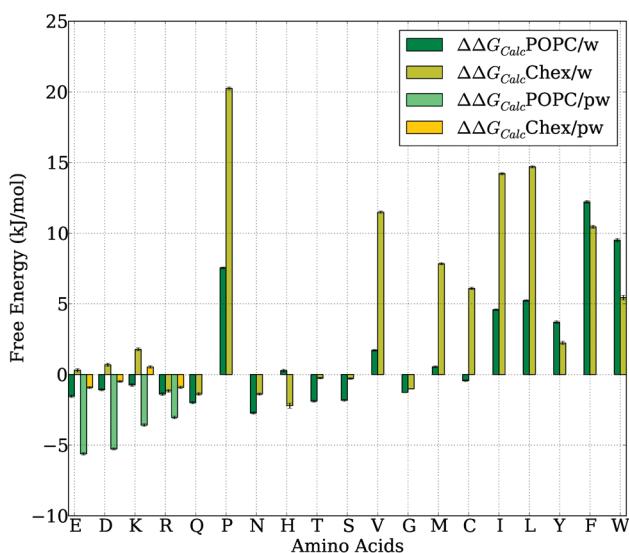


**Figure 6.** PMF profiles for peptides with X = A, E, I, Q, or W at the cyclohexane/water interface as a function of the distance between the COMs of cyclohexane and the peptide. The vertical black line represents the interface, that is, where the densities of water and cyclohexane are equal. The minimum for each profile is filled up to 1RT. The error bars represent the variance multiplied by 10 so that they can be distinguished from the lines.

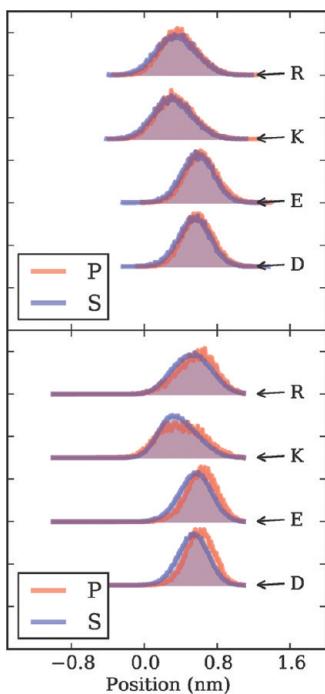


**Figure 7.** Comparison of experimental (horizontal axis) and calculated (vertical axis) partitioning free energies at the POPC/water interface as obtained using the standard MARTINI FF (circles). Residues are indicated by a corresponding single-letter amino acid code. The partitioning free energies calculated using the polarizable water model for side chains E, D, K, and R are represented by triangles. The solid blue line indicates perfect agreement with the experimental data, and the dashed lines indicate  $\pm 0.5RT$ .

indicates that the C<sub>1</sub>–SC<sub>4</sub> interactions are strongly attractive, resulting in large positive values as compared to the experimental data. Proline is modeled using C<sub>2</sub> and N<sub>a</sub> beads, where the N<sub>a</sub> bead represents the backbone. This mapping leads to a higher free energy of partitioning. Using beads that have more attractive interactions with water molecules could improve the free energy of partitioning compared to experimental values; however, because of the lack of experimental solvation data for the proline



**Figure 8.** Calculated side-chain partitioning free energies at the POPC/water and cyclohexane/water interfaces as obtained using standard MARTINI are labeled as  $\Delta\Delta G_{\text{calc}}^{\text{POPC}/\text{w}}$  and  $\Delta\Delta G_{\text{calc}}^{\text{Chex}/\text{w}}$ , respectively. Free energies calculated using MARTINI with the polarizable water model (for side chains D, E, R and K) are labeled as  $\Delta\Delta G_{\text{calc}}^{\text{POPC}/\text{pw}}$  and  $\Delta\Delta G_{\text{calc}}^{\text{Chex}/\text{pw}}$ , respectively.



**Figure 9.** Distributions of the side-chain COM of amino acids at the cyclohexane/water (bottom) and POPC/water (top) interfaces, as obtained using the polarizable water model (labeled as P and colored red) or standard MARTINI (labeled as S and colored blue).

residue, it is difficult to justify such mappings for proline. The side-chain mappings for these two residues will be addressed in a future version of MARTINI.

**3.2. Effects of Secondary Structure Parameters.** In the current MARTINI model, a protein can be modeled as having a helix, coil, extended, turn, or bend secondary structure. The

**Table 2.  $\Delta\Delta G^s$  for Various Secondary Structures**

secondary structure	$\Delta\Delta G^s$ ( $\text{kJ mol}^{-1}$ )	
	E	I
helix	0.9	18.7
extended	1.3	1.4
turn	1.1	9.4
bend	1.0	10.1
dih <sup>a</sup>	0.7	

<sup>a</sup> Dih represents simulations carried out using the dihedral potential extracted from atomistic simulations.

choice of secondary structure affects the type of backbone bead and associated bonded parameters. The current MARTINI version does not model changes in secondary structure. Even though the pentapeptides in this study do not have any preferred secondary structure, the conformations adopted by a peptide in the bulk water phase and at the interface could be significantly different. At the interface, the peptide is more likely to adopt conformations in which all of the hydrophobic side chains are oriented toward the cyclohexane and the charged side chains are oriented toward water, even when these conformations cause some strain in the backbone structure. As well, the addition of a dihedral potential will affect the conformational space explored by the peptide in the bulk solvent. These effects would be highly dependent on the peptide sequence, thus affecting partitioning free energies to different degrees. We investigated the effect of imposing a secondary structure on the free energy of partitioning by carrying out simulations with various secondary-structure parameters and computing the relative free energy of transfer of the peptide, with given secondary-structure constraints, to the bulk water as  $\Delta\Delta G^s = (\Delta G_{\text{inter}}^s - \Delta G_{\text{bulk}}^s) - (\Delta G_{\text{inter}}^c - \Delta G_{\text{bulk}}^c)$ , where  $\Delta G_{\text{inter}}^c$  and  $\Delta G_{\text{bulk}}^c$  are annihilation free energies of side-chain beads of E or I at the cyclohexane/water interface and in bulk water, respectively, for the peptide modeled as a coil and  $\Delta G_{\text{inter}}^s$  and  $\Delta G_{\text{bulk}}^s$  represent the corresponding free energies for the peptide represented with helix, turn, extended, or bend secondary structures. Furthermore, the dihedral potentials derived from bulk atomistic simulations of the same peptide were also employed. The resulting free energy values are reported in Table 2. The WLILL peptide favors the interface by  $18.7 \text{ kJ mol}^{-1}$  when modeled as a helix as compared to the coil structure, whereas for the WLELL peptide, the differences between various secondary structure representations are less than  $1.5 \text{ kJ mol}^{-1}$ . As all of the side chains in the WLILL peptide are hydrophobic, the helical conformations shield the polar backbone beads from cyclohexane beads, allowing the peptide to penetrate deeper into the cyclohexane region, resulting in better interactions between the side chains and the cyclohexane beads. On the other hand, the side-chain bead of E is always positioned toward the bulk water, regardless of the secondary-structure constraints.

Enforcing a particular secondary structure can have a significant impact on the partitioning free energy of a peptide at the interface, particularly for extreme cases such as enforcing an  $\alpha$  helix for a hydrophobic peptide. This finding highlights the importance of carefully considering the treatment of secondary structure when using MARTINI to model peptides or proteins and the need for further improvements in MARTINI such that secondary-structure transitions can be adequately incorporated.

## 4. CONCLUSIONS

We have calculated the partitioning free energy of amino acid side chains using the MARTINI model at cyclohexane/water and POPC/water interfaces for the Wimley–White hydrophobicity scale peptides WLXLL. The free energies obtained using TI, FEP, and umbrella sampling were found to be in good agreement with each other. Among the three methods employed in this study, the free energies obtained using the FEP method coupled with MBAR for free energy estimation had the smallest statistical uncertainties.

Comparison of the experimental free energies at the POPC/water interface with computed values revealed that the amino acids F and P and the charged amino acids have the largest deviation. Hydrophobic residues such as L, I, and V and those with a net charge on the side chain are biased toward the hydrophobic phase, whereas most of the polar amino acids have a slightly higher preference for bulk water. In the case of amino acids with a net charge, significant improvements in calculated free energies were obtained by using the polarizable MARTINI water model. P and F can be improved in future versions of MARTINI with different bead mappings. For peptides at the cyclohexane/water interface, the polarizable water model did not have a significant effect on the free energy of partitioning of charged residues. This could be due to limitations of the model, but no experimental data are available for this system.

The choice of secondary-structure constraints imposed on the peptides affects the partitioning free energies, highlighting the importance of improving the representation of secondary structure in MARTINI, because, for simulations of peptide adsorption at an interface, it is important that the coarse model of the peptide be able to model the secondary-structure changes that could occur during peptide adsorption.

Because the coarse-grained representation was used, simulations of  $0.2\text{--}0.5\ \mu\text{s}$  per  $\lambda$  point were feasible, resulting in good convergence. It can be estimated that, if all-atom models were to be employed, sampling of an order of magnitude higher would likely be required to obtain free energy values with reasonable uncertainty. Although this is currently a significant computational challenge, such simulations will be in easy reach in the near future and allow access to key thermodynamic data on lipid/peptide interactions to further improve simulations of biological membranes.

## AUTHOR INFORMATION

### Corresponding Author

\*E-mail: tieleman@ucalgary.ca.

## ACKNOWLEDGMENT

This work was supported by the Natural Sciences and Engineering Research Council (Canada). G.S. was supported by a postdoctoral fellowship from the Alberta Heritage Foundation for Medical Research (AHFMR). D.P.T. is an AHFMR Scientist. Calculations were performed in part on WestGrid/Compute Canada facilities.

## REFERENCES

- (1) Scheraga, H. A.; Khalili, M.; Liwo, A. *Annu. Rev. Phys. Chem.* **2007**, *58*, 57–83.
- (2) Gnanakaran, S.; Nymeyer, H.; Portman, J.; Sanbonmatsu, K. Y.; Garcia, A. E. *Curr. Opin. Struct. Biol.* **2003**, *13*, 168–174.
- (3) Snow, C. D.; Nguyen, H.; Pande, V. S.; Gruebele, M. *Nature* **2002**, *420*, 102–106.
- (4) Simmerling, C.; Strockbine, B.; Roitberg, A. E. *J. Am. Chem. Soc.* **2002**, *124*, 11258–11259.
- (5) Fujitani, H.; Tanida, Y.; Ito, M.; Jayachandran, G.; Snow, C. D.; Shirts, M. R.; Sorin, E. J.; Pande, V. S. *J. Chem. Phys.* **2005**, *123*, 084108.
- (6) Smit, B.; Hilbers, P. A. J.; Esselink, K.; Rupert, L. A. M.; van Os, N. M.; Schlijper, A. G. *Nature* **1990**, *348*, 624–625.
- (7) Nielsen, S.; Lopez, C.; Srinivas, G.; Klein, M. *J. Phys.: Condens. Matter* **2004**, *6*, R481–R512.
- (8) *Coarse-Graining of Condensed Phase and Biomolecular Systems*, 1st ed.; Voth, G. A., Ed.; CRC Press: Boca Raton, FL, 2009.
- (9) Marrink, S. J.; de Vries, A. H.; Mark, A. E. *J. Phys. Chem. B* **2004**, *108*, 750–760.
- (10) Marrink, S. J.; Risselada, H. J.; Yefmov, S.; Tieleman, D. P.; de Vries, A. H. *J. Phys. Chem. B* **2007**, *111*, 7812–7824.
- (11) Baoukina, S.; Monticelli, L.; Risselada, H. J.; Marrink, S. J.; Tieleman, D. P. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 10803–10808.
- (12) Risselada, H. J.; Marrink, S. J. *Soft Matter* **2009**, *5*, 4531–4541.
- (13) Baoukina, S.; Tieleman, D. P. *Biophys. J.* **2010**, *99*, 2134–2142.
- (14) Monticelli, L.; Kandasamy, S. K.; Periole, X.; Larson, R. G.; Tieleman, D. P.; Marrink, S.-J. *J. Chem. Theory Comput.* **2008**, *4*, 819–834.
- (15) Oostenbrink, C.; Villa, A.; Mark, A. E.; van Gunsteren, W. F. *J. Comput. Chem.* **2004**, *25*, 1656–1676.
- (16) Shirts, M. R.; Pitera, J. W.; Swope, W. C.; Pande, V. S. *J. Chem. Phys.* **2003**, *119*, 5740–5761.
- (17) MacCallum, J. L.; Tieleman, D. P. *J. Comput. Chem.* **2003**, *24*, 1930–1935.
- (18) Shirts, M. R.; Pande, V. S. *J. Chem. Phys.* **2005**, *122*, 134508.
- (19) Hess, B.; van der Vegt, N. F. A. *J. Phys. Chem. B* **2006**, *110*, 17616–17626.
- (20) Tieleman, D. P.; MacCallum, J. L.; Ash, W. L.; Kandt, C.; Xu, Z.; Monticelli, L. *J. Phys.: Condens. Matter* **2006**, *18*, S1221.
- (21) Ozdirekcan, S.; Etchebest, C.; Killian, J. A.; Fuchs, P. F. *J. Am. Chem. Soc.* **2007**, *129*, 15174–15181.
- (22) Monticelli, L.; Tieleman, D. P.; Fuchs, P. F. *J. Biophys. J.* **2010**, *99*, 1455–1464.
- (23) Wimley, W. C.; Creamer, T. P.; White, S. H. *Biochemistry* **1996**, *35*, 5109–5124.
- (24) Wimley, W. C.; White, S. H. *Nat. Struct. Biol.* **1996**, *3*, 842–848.
- (25) White, S. H.; Wimley, W. C. *Annu. Rev. Biophys. Biomol. Struct.* **1999**, *28*, 319–365.
- (26) Snider, C.; Jayasinghe, S.; Hristova, K.; White, S. H. *Protein Sci.* **2009**, *18*, 2624–2628.
- (27) Aliste, M. P.; MacCallum, J. L.; Tieleman, D. P. *Biochemistry* **2003**, *42*, 8976–8987.
- (28) Aliste, M. P.; Tieleman, D. P. *BMC Biochem.* **2005**, *6*, 30.
- (29) MacCallum, J. L.; Bennett, W. F. D.; Tieleman, D. P. *J. Gen. Physiol.* **2007**, *129*, 371–377.
- (30) MacCallum, J. L.; Bennett, W. F. D.; Tieleman, D. P. *Biophys. J.* **2008**, *94*, 3393–3404.
- (31) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4*, 435–447.
- (32) Nose, S. *Mol. Phys.* **1984**, *52*, 255–268.
- (33) Hoover, W. *Phys. Rev. A* **1985**, *31*, 1695–1697.
- (34) Parrinello, M.; Rahman, A. *J. Appl. Phys.* **1981**, *52*, 7182–7190.
- (35) Berendsen, H. J. C.; Postma, J.; DiNola, A.; Haak, J. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
- (36) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graph.* **1996**, *14* (33–8), 27–8.
- (37) Shirts, M. R.; Mobley, D. L.; Chodera, J. D. *Annu. Rep. Comput. Chem.* **2007**, *3*, 41–59.
- (38) Torrie, G. M.; Valleau, J. P. *Chem. Phys. Lett.* **1974**, *28*, 578–581.
- (39) Straatsma, T. P.; McCammon, J. A. *J. Chem. Phys.* **1991**, *95*, 1175–1188.
- (40) Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22*, 1420–1426.

- (41) Shirts, M. R.; Chodera, J. D. *J. Chem. Phys.* **2008**, *129*, 124105.
- (42) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. *J. Comput. Chem.* **1992**, *13*, 1011–1021.
- (43) Hess, B. *J. Chem. Phys.* **2002**, *116*, 209–217.
- (44) Lindahl, E.; Hess, B.; van der Spoel, D. *J. Mol. Model.* **2001**, *7*, 306–317.
- (45) Yesylevskyy, S. O.; Schafer, L. V.; Sengupta, D.; Marrink, S. J. *PLoS Comput. Biol.* **2010**, *6*, e1000810.
- (46) Hub, J. S.; de Groot, B. L.; van der Spoel, D. *J. Chem. Theory Comput.* **2010**, *6*, 3713–3720.
- (47) Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657–1666.
- (48) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269–6271.
- (49) Yau, W. M.; Wimley, W. C.; Gawrisch, K.; White, S. H. *Biochemistry* **1998**, *37*, 14713–14718.
- (50) Killian, J. A.; von Heijne, G. *Trends Biochem. Sci.* **2000**, *25*, 429–434.
- (51) Sun, H.; Greathouse, D. V.; Andersen, O. S.; Koeppe, R. E. *J. Biol. Chem.* **2008**, *283*, 22233–22243.