JCTC Journal of Chemical Theory and Computation

ARTICLE

pubs.acs.org/JCTC

Parameterization of PACE Force Field for Membrane Environment and Simulation of Helical Peptides and Helix−Helix Association

Cheuk-Kin Wan,[†] Wei Han,[†] and Yun-Dong Wu*[,†,‡,§]

[†]Department of Chemistry, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China

[‡]School of Chemical Biology and Biotechnology, Laboratory of Chemical Genomics, Peking University Shenzhen Graduate School, Shenzhen, 518055, China

[§]College of Chemistry, Peking University, Beijing, 100871, China

Supporting Information

ABSTRACT: The recently developed PACE force field was further parametrized so that it can be applied to the studies of membrane systems. Parameters for the interactions between united-atom protein particles and lipid hydrophobic tails were developed by reproducing the solvation free energies of small organic molecules in hexadecane. Interactions between protein particles and lipid heads were parametrized by fitting the potential of mean force of the corresponding all-atom simulation. The force field was applied to the study of five helical peptides in membrane environments. The calculated tilt angles of WALP and GWALP and their mutations are in good agreement with experimental data. The association of two glycophorin A (GpA) helices was simulated for 6 μs. Root-mean-square-deviation of the simulated dimer from the nuclear magnetic resonance structure was found to be 0.272 nm, better than all results obtained so far. These findings demonstrate the high accuracy and applicability of the PACE force field in studying membrane proteins.

INTRODUCTION

Membrane proteins are important for many cellular processes, including transport activities, signal transductions, and receptor functions.[1] Understanding structures, dynamics, and functions of membrane proteins is challenging.[2−4] For instance, although the first crystal structure of membrane proteins was obtained in 1985, only about 300 unique crystal structures have been determined in the 26 years since (see http://blanco.biomol.uci.edu/mpstruc/listAll/list for details).[5] Molecular dynamics (MD) simulation is a complementary tool to reveal the structural and dynamic details of proteins.[6−8] It can also help with the interpretation and evaluation of experimental findings.[9] All-atom force fields are commonly used but limited to relatively short simulation lengths. The folding and dynamics of membrane proteins are in the time scale of microseconds or even longer. Although continuous trajectories were achieved in 10 μs and 1 ms simulations by Schulten et al.[10] and Shaw et al.[11] for the folding of soluble proteins, they relied on the use of supercomputers and specially designed machines not available to most people.

Coarse-grained (CG) force fields have been actively developed to reduce computational cost for more efficient long time simulations. Klein et al. grouped three heavy atoms and their associated hydrogen atoms into one CG particle. This model was used to study the membrane insertion activity of antimicrobial polymers.[12,13] Sperotto et al. developed a similar CG model for the simulation of lipid-mediated protein−protein interactions.[14,15] Apart from coarse-graining three heavy atoms into one CG particle, a four-to-one mapping has also been used to simplify lipid molecules to investigate vesicle fission and fusion.[16,17] Voth et al. developed a solvent-free lipid bilayer model using a multiscale coarse-graining approach to enhance computational efficiency.[18]Another four-to-one CG model—the MARTINI force field—that has found extensive applications was developed by Marrink and co-workers.[19−21] It has been applied to biomolecular simulations, such as membrane fusion and the spontaneous gating of channel.[22,23] This model has been extended to include different types of particles. The first extended protein force field was built by Schulten et al. based on the first version of MARTINI lipid model to study lipoproteins.[24] Sansom et al. developed another protein model for membrane protein simulations.[25] Tarek et al. parametrized cyclic peptides to describe the self-assembly of $[Trp\text{-}Leu]_4$.[26]

The common feature of these CG models is the similar degree of coarse graining of proteins and solvent molecules. Although these CG approaches can speed up MD simulations by 2−3 orders of magnitude, they may not be accurate enough due to oversimplification and the loss of atomistic details. Details of solvent molecules are thought to be less significant, so CG or even implicit solvent models are generally employed in studying proteins. Atomistic details of protein molecules are more important in maintaining and predicting protein structures. Coarse graining of proteins is therefore less desirable. To solve this problem, a finer protein model could be incorporated into a CG solvent model to build a hybrid-resolution protein force field. Voth et al. built a mixed all-atom and CG model for gramicidin A ion channel simulation.[27] We proposed a hybrid-resolution force field (PACE) to couple a united-atom (UA) protein model with the MARTINI CG environment. A similar approach was recently proposed by Marrink et al. which was implemented differently.[28] In our previous work, we optimized the PACE force field and coupled it with the CG water model.[29−32] We showed that the

Received: June 21, 2011
Published: November 08, 2011

force field could reproduce the statistical backbone and side chain potentials of all amino acids accurately.[33] We demonstrated that not only could it maintain the stability of the native structures of proteins with medium size (50−150 amino acids), it could also fold several α-helical, β-sheet, and mixed helical/coil peptides from first principles.

Encouraged by these results, we continue to extend the PACE force field to cover the CG membrane environment. In this paper, we report our attempt to incorporate MARTINI's lipid model into our PACE force field.[20] The parametrizations involved include interactions between protein UA particles and lipid CG hydrophobic tail groups and protein UA particles and lipid CG head groups. To evaluate the quality of the modified PACE force field, we applied it to the simulations of WALP, a designed helical peptide that has been widely studied both experimentally and theoretically,[34−43] and four other helical peptides.[44,45] We pay particular attention to the tilt angle of these peptides in two membrane environments. We also studied the dimerization of glcyophorin helix A (GpA) in membrane, which is currently a topic of intense interest.[46−69] Our simulations reproduce related experimental observations quite well, which implies that the extended PACE force field may have potential applications in the study of membrane proteins with the atomistic details.

MODEL AND METHODS

PACE Protein Model. In the PACE force field, proteins are represented at a UA level and embedded in a CG environment.[31] Figure 1a shows the schematic representation of our model. The water and lipid models are adopted from the MARTINI model.[20] Around four heavy atoms are represented by a CG particle. The total energy of the MARTINI model is expressed in eqs 1−3:

$$E_{\text{CG-total}} = E_{\text{CG-bonded}} + E_{\text{CG-nonbonded}} \quad (1)$$

$$E_{\text{CG-bonded}} = E_{\text{CG-bond}} + E_{\text{CG-angle}} \quad (2)$$

$$E_{\text{CG-nonbonded}} = E_{\text{CG-vdW}} + E_{\text{CG-electrostatic}} \quad (3)$$

Bond lengths and angles are modeled by harmonic potentials $E_{\text{CG-bond}}$ and $E_{\text{CG-angle}}$, respectively. Lennard-Jones potential energy function (eq 4) is used to describe van der Waals interactions ($E_{\text{CG-vdW}}$) between CG particles, while ε_{ij} indicates the strength of the interaction, and δ_{ij} indicates the distance between two interacting groups with zero interaction energy.

$$E_{\text{CG-vdW}} = \sum_{i\neq j} 4\varepsilon_{ij}\left(\frac{\delta_{ij}^{12}}{r^{12}} - \frac{\delta_{ij}^{6}}{r^{6}}\right) \quad (4)$$

Electrostatic interactions ($E_{\text{CG-electrostatic}}$) between charged groups are modeled by Coulombic potential energy function (eq 5), and q_i and q_j are charges of the charged groups. Relative dielectric constant ε_r is set to be 15 for explicit screening.

$$E_{\text{CG-electrostatic}} = \sum_{i\neq j} \frac{q_i q_j}{4\pi\varepsilon_0\varepsilon_r r} \quad (5)$$

eqs 1−5 are used for CG−CG interactions. All parameters are adopted from the MARTINI model and can be found in ref 20.

The protein model of PACE is UA based in that hydrogen atoms are implicitly incorporated into the attached heavy atoms.

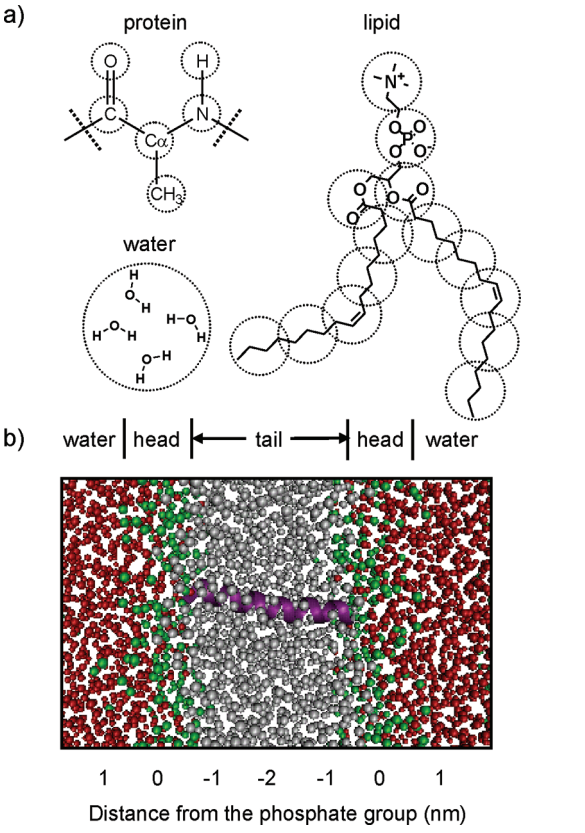

Figure 1. (a) Schematic representation of the PACE model. Protein particles and the environment (lipid bilayer and water) are represented by the UA and CG models respectively. (b) Snapshot of a peptide in the DOPC bilayer. The UA protein is shown in purple. The CG lipid tail, lipid head, and water particles are shown in gray, green, and red, respectively.

In order to have a good account of the hydrogen bonds, we represent explicitly the hydrogens that are attached to nitrogen for the backbone amide and side chains of Asn, Gln, Trp, and His. Detailed schematic representations of 20 amino acids can be found in ref 31. For UA−UA interactions, we used similar potential energy functions to describe bonded and nonbonded interactions. The total energy of our protein model is expressed in eqs 6 and 7:

$$E_{\text{UA-total}} = E_{\text{UA-bonded}} + E_{\text{UA-nonbonded}} + E_{\text{UA-HB}} \quad (6)$$

$$E_{\text{UA-bonded}} = E_{\text{UA-angle}} + E_{\text{UA-improper}} + E_{\text{UA-torsion}} + E_{\text{UA-14pair}} \quad (7)$$

All bond lengths are constrained by the LINCS algorithm.[70] Bond angles are constrained by a harmonic potential $E_{\text{UA-angle}}$ at their equilibrium. Planar geometries and chiral centers of molecules are maintained by $E_{\text{UA-improper}}$. $E_{\text{UA-torsion}}$ and $E_{\text{UA-14pair}}$ describe the potential energy of the dihedral angle of a rotatable bond. All optimized parameters can be found in ref 32.

None of the protein particles of PACE carry charges. A single potential energy function, eq 4 is used to describe nonbonded interactions ($E_{\text{UA-nonbonded}}$) between UA particles. The parameters for protein−water interactions were optimized by fitting the hydration free energies of 35 organic compounds in a previous study. The average deviation from experimental data

was 1.1 kJ/mol.[32] The interactions between protein particles were parametrized by reproducing the experimental density and self-solvation free energy of eight pure organic liquids. The average errors for density and self-solvation free energy were 3.2% and 0.7 kJ/mol, respectively. Side chain−side chain and side chain−backbone interaction potentials were parametrized to fit the potential of mean force of corresponding all-atom simulations using the OPLS-AA force field. The results and the optimized parameters can be found in ref 31.

$E_{\text{UA-HB}}$ is used to control the strength of backbone−backbone H-bond interactions and is given by eq 8. Parameters are optimized by reproducing the experimental α-helical and β-sheet content of AK17 ($(AAKAA)_3GY$) and GB1m2 (GEWTYNPATGKFTVTE) peptides.[31] Note that we use a slightly different potential to describe side chain H-bonds, as the H-bond potentials for backbones can only handle hydrogen donors containing only one hydrogen atom. These parameters are obtained by fitting the all-atom potential of mean forces (PMFs). The schematic representation of the different H-bond types is shown in Figure S1, Supporting Information.

$$E_{\text{UA-HB}} = \sum_{|i-j|>2}\left[4\varepsilon_{\text{attr}}\left(\frac{\delta^{12}_{O_i-NH_j}}{r^{12}_{O_i-NH_j}} - \frac{\delta^{6}_{O_i-NH_j}}{r^{6}_{O_i-NH_j}}\right) + 4\varepsilon_{\text{rep}}\frac{\delta^{12}_{O_i-C_{\alpha j}}}{r^{12}_{O_i-C_{\alpha j}}} + 4\varepsilon_{\text{rep}}\frac{\delta^{12}_{O_i-C_j-1}}{r^{12}_{O_i-C_j-1}} + 4\varepsilon_{\text{rep}}\frac{\delta^{12}_{C_i-NH_j}}{r^{12}_{C_j-NH_j}}\right] \quad (8)$$

To incorporate the UA protein model into the CG lipid model, parametrizations are carried out for nonbonded interactions between protein−lipid tail and protein−lipid head groups. eq 4 is used to describe UA−CG interactions. Therefore, ε_{ij} and δ_{ij} are carefully optimized in this work.

As most of the computation time is spent on simulations of the solvent and the lipid bilayer, simplification of the solvent and lipid model can reduce the number of degrees of freedom and greatly enhance the computation efficiency. Note that the effective times of the MARTINI model are normally interpreted as four times the simulation times.[21] Our protein model is UA based, and the effective times can only be interpreted after further studies. In this work, all reported simulation lengths are actual simulation times.

Simulation Parameters. All of the simulations were performed with the GROMACS software package, version 3.3.1.[71] The van der Waals interactions were shifted to zero between distances of 0.9 and 1.2 nm. The electrostatic forces were shifted from 0.0 to 1.2 nm with a dielectric constant of 15 for explicit screening. The neighbor list was updated every 10 steps with a searching cutoff distance of 1.4 nm. The temperature of each group (protein, lipid and water) was kept constant using the Berendsen thermostat[72] with a time constant of 0.1 ps. The pressure on the z-axis and that on the xy plane were each coupled using the Berendsen barostat[72] with a time constant of 0.5 ps and a compressibility of 4.5×10^{-5} bar^{-1}. In all simulations, the normal of the membrane was aligned along the z-axis. All of the systems were kept at 1 atm. As explicit hydrogen atoms are represented by dummy atoms, a larger time step of 6 fs can be used, which has been shown to satisfy energy conservation during simulations.[73] In the preparation stage, 5000 steps of steep descent optimization were performed, followed by a 1 ns pre-equilibrium simulation at 300 K and 1 atm. The generated coordinates and velocities were used for simulation runs.

Solvation Free Energy Calculations. The solvation free energy is the change in free energy of a molecule from the ideal gas state to a state where the molecule is immersed in the solvent. It is calculated by introducing a coupling parameter (λ) with interaction potentials between the solute and the solvent using the free energy perturbation method. In the simulations, as λ gradually changed from zero to unity, the solute−solvent interactions were gradually turned off. The solvation free energy was then calculated as follows:[74]

$$\Delta G = G_1 - G_0 = \int_{\lambda=0}^{\lambda=1} d\lambda \left\langle \frac{\partial U(\lambda)}{\partial \lambda} \right\rangle_{\lambda} \quad (9)$$

The solute molecule was put into a dodecahedron box containing 400 solvent molecules. For each solvation free energy calculation, 24 intermediate stages were applied (λ = 0.0, 0.06, 0.15, 0.25, 0.35, 0.4, 0.45, 0.5, 0.53, 0.56, 0.59, 0.62, 0.65, 0.68, 0.71, 0.73, 0.75, 0.78, 0.81, 0.84, 0.88, 0.92, 0.96, 1.0). A 6 ns simulation was performed for each intermediate stage giving a total of 144 ns of simulations. The temperature was kept at 300 K. The reproducibility of the calculations was tested. The average deviation of the solvation free energy calculations in several runs is about 1 kJ/mol.

PMF Calculations. For each PMF calculation, a small molecule was put in the center of the dioleoylphosphatidylcholine (DOPC) bilayer containing 70 lipid molecules and 1100 CG water molecules. PMFs were calculated using the free energy perturbation method. The position of the center of mass of a molecule was fixed by a potential $k/2(x - x_0)^2$. The x_0 can be perturbed by introducing a λ. As λ varied from zero to unity, the solute was moved from the center of the lipid bilayer to the aqueous phase in a straight line along the bilayer normal. The free energy in the aqueous phase (λ = 1) was chosen as the reference state. The free energy difference between any intermediate point (λ_i) and the starting point ($\lambda = 0$) can be estimated by accumulating $dU/d\lambda$ until $\lambda = \lambda_i$. In each simulation, the free energies of 100 intermediate states were chosen for PMF plots. Each simulation was carried out for 400 ns. For each PMF calculation, 10 perturbation simulations were performed to generate an average PMF curve, leading to a total of 4 μs of simulations. All simulations were kept at 300 K throughout the entire duration.

Transmembrane Peptide Simulations. Transmembrane peptides were inserted into the pre-equilibrated lipid bilayer, dilauroylphosphatidylcholine (DLPC), dipalmitoylphosphatidylcholine (DPPC), or DOPC, of 128 lipids with 1500 CG water molecules. The starting structures of the peptides were fully α-helical and in orientation of about 0° tilt relative to the bilayer normal. To have the peptides vertically inserted into the lipid bilayer for MD simulations, different setup methods have been proposed.[75−77] In this work, the simulation system was first subjected to 5000 steps of steep descent optimization. If there were crashes between peptide and lipid molecules, then the corresponding lipid molecules were removed from the system. Interestingly, crashes rarely occurred. The whole system was then pre-equilibrated for 1 ns with the peptide molecules constrained. The generated coordinates are used for simulation runs with the peptides fully relaxed at 323 K and 1 atm. To validate this simulation procedure, we compared the results with the self-assembling bilayer. Lipid molecules were randomly distributed in

Table 1. Experimental and Calculated Solvation Free Energies of UA Small Molecules in CG Hexadecane and CG Water

	solvation free energy in hexadecane (kJ/mol)			solvation free energy in water (kJ/mol)[b]		transfer free energy from hexadecane to water (kJ/mol)	
compound	CG	exptl[a]	experimental log P (hexadecane/gas)	CG	exptl	CG	exptl
ethane	−3.0	−2.8	0.5	7.3	7.4	10.3	10.2
propane	−6.2	−6.0	1.1	8.2	8.3	14.4	14.3
butane	−8.7	−9.3	1.6	9.2	9.0	17.9	18.3
methanol	−4.8	−5.6	1.0	−20.1	−20.2	−15.4	−14.6
ethanol	−8.1	−8.6	1.5	−20.5	−21.0	−12.4	−12.4
1-propanol	−10.6	−12.1	2.1	−20.1	−20.4	−9.5	−8.3
2-propanol	−10.4	−10.5	1.8	−22.0	−19.9	−11.6	−9.5
ethylamine	−9.3	−9.7	1.7	−19.4	−18.8	−10.1	−9.2
acetone	−10.3	−9.7	1.7	−14.1	−16.1	−3.8	−6.4
butanone	−12.3	−13.2	2.3	−15.2	−15.2	−2.9	−2.1
acetamide	−14.1	−14.0	2.4	−39.7	−40.6	−25.6	−26.6
benzene	−14.5	−16.0	2.8	0.0	−3.6	14.5	12.4
toluene	−21.8	−19.1	3.3	−3.7	−3.7	18.1	15.4
naphthalene	−26.7	−30.7	5.3	−7.2	−10.0	19.5	20.7
methylindole	−38.2	–	–	−31.0	−24.6	7.2	–
p-cresol	−28.7	−24.8	4.3	−29.0	−25.6	−0.3	−0.8
acetic acid	−10.1	−10.1	1.8	−27.4	−28.0	−27.4	−28.0
dimethyl sulfide	−13.2	−12.9	2.2	−8.5	−6.4	4.7	6.5
average error	1.0			1.1		1.1	

[a] The solvation free energy in hexadecane is calculated from experimental partition coefficients between gas phase and hexadecane (log P).[78,79] [b] Ref 32.

the system box. The lipid bilayer then self-assembled around the peptides. The tilt angles of WALP23 in DOPC were calculated using both methods. We found that there was no significant difference in the tilt angle between the two methods (data not shown). This showed that small peptides could easily be fitted into the CG lipid model without the need for complicated setup procedures. Note that this probably applies to small simple peptides only. For large proteins, this simple scheme may not work.

RESULTS AND DISCUSSIONS

Parameterization of Protein−Lipid Tail Interactions. Each lipid molecule comprises two major parts: the nonpolar tail group and the polar headgroup. The nonpolar tail part of the lipid molecule is the core of the lipid bilayer. A large portion of transmembrane peptides is exposed to the nonpolar environment. This could be considered as solvating in the nonpolar solvent. One of the more promising parametrization strategies is to reproduce the experimental thermodynamic properties of organic molecules.[19−21,32] To parametrize protein UA particles and CG lipid tail particles, we fitted the solvation free energies of UA small organic molecules in CG hexadecane, which is a better model than cyclohexane, for the membrane environment. The experimental partition coefficients (log P) of small organic molecules between the gas phase and hexadecane are available.[78,79] The solvation free energy can be calculated from the partition coefficient by:

$$\Delta G_{sol} = -2.303\,RT \log P \qquad (10)$$

where R is the universal gas constant and T = 300 K. Seventeen small molecules from eight classes of organic compounds (alkane, alcohol, ketone, amide, amine, aromatics, carboxylic acid, and sulfide) covering different types of amino acid side chains were used to optimize the parameters. Ionizable amino acids are normally charged in an aqueous environment. But they may prefer to exist in the neutral form to avoid a large desolvation penalty when they are in a nonpolar environment. In the PACE force field, side chains of Asp, Glu, Lys, and Arg are assumed to be neutral inside the membrane but charged in the aqueous medium.[31] Therefore, carboxylic acid was used for the calculation of the solvation free energies in hexadecane of Asp and Glu, while amine was used for Lys.

We first optimized parameters for aliphatic carbons, as these are the common part of all amino acid side chains. Parameters for $-CH_3$ were obtained from ethane, and $-CH_2$ could then be parametrized by propane and butane. Benzene was used to parametrize aromatic carbon. Then, other functional groups were parametrized using the corresponding small organic molecules. For simplicity, we adopted the Lorentz−Berthelot combination rule ($\delta_{ij} = (\delta_{ii} + \delta_{jj})/2$) to derive δ_{ij}. In hexadecane, ε_{ij} was optimized to reproduce experimental solvation free energy. The results and parameters are shown in Tables 1 and 2, respectively. Due to a lack of the experimental data, the S of Cys and Met side chains share the same parameters, so do the N of Lys and Arg side chains. Backbone N, the N of Asn, Gln, Trp, and His side chains share the same parameters. The absolute average error for the solvation free energies of UA small molecules in CG hexadecane is about 1.0 kJ/mol. The transfer free energies of these UA small molecules from CG hexadecane to CG water were obtained by calculating the difference between solvation free energy in water and in hexadecane:

$$\Delta G_{hex \rightarrow water} = \Delta G_{water} - \Delta G_{hex} \qquad (11)$$

Table 2. Small Organic Molecules Used in the Calculation of Solvation Free Energy in Hexadecane for Protein−Lipid Tail Parameterization and the Resulting Optimized Parameters

		parameter	
protein UA group	compound used for parametrization	ε_{ij} (kJ/mol)	δ_{ij} (nm)
$-\underline{CH_3}$	ethane	1.7	0.43
$-\underline{CH_2}$	propane butane	1.3	0.43
$-\underline{CH}$	2-propanol	1.0	0.43
aromatic C[a]	benzene toluene naphthalene	0.9	0.4225
$-N$[b]	ethylamine	2.2	0.40
$O{=}C-\underline{NH_2}$[c]	acetamide	2.1	0.40
$-\underline{C}{=}O$[d]	acetone	1.0	0.415
$-C{=}\underline{O}$[d]	butanone	2.0	0.375
$-\underline{C}OO^-$[e]	acetic acid	1.0	0.415
$-C\underline{OO}^-$[e]		1.7	0.375
$-\underline{OH}$[f]	methanol ethanol 1-propanol	2.0	0.38
$-\underline{OH}$[g]	*p*-cresol	2.0	0.38
$-S-$[h]	dimethyl sulfide	3.0	0.405

[a] Aromatic C for Phe, Tyr, Trp and His side chains. [b] −N for Lys and Arg side chains. [c] −N for backbone amide, Trp, and His side chains. [d] −C═O for backbone amide, Asn, and Gln side chains. [e] $-COO^-$ for Asp and Glu side chains. [f] −OH for Ser and Thr side chains. [g] −OH for Tyr side chain. [h] −S− for Cys and Met side chains.

The absolute average error is about 1.1 kJ/mol. This shows that the PACE force field can reproduce the free energy of solvation and the free energy of partitioning between the oil and the aqueous phase well.

Parameterization of Protein−Lipid Head Interactions. It has been found that some of the amino acid side chains have favorable interactions with lipid head groups which influence the orientation of transmembrane peptides.[80,81] Therefore, it is important to carefully parametrize the interactions between protein and lipid head groups. One of the successful parametrization strategies is to reproduce the partitioning free energies of amino acid side chains in a lipid bilayer using all-atom simulations. We have adopted this strategy for our parameter development.[21] PMFs of amino acid side chain analogues in the DOPC bilayer were calculated and fitted with all-atom results.[82,83] Figure 2 shows the PMFs of 18 amino acid side chain analogues (except glycine and proline) calculated using the OPLS-AA force field[82] and the PACE force field. Table 3 gives the optimized parameters. Note that DOPC contains unsaturated CG carbon particles which are absent in CG hexadecane molecules. We tentatively assume that parameters for the interactions between UA particles and CG saturated carbon particles could be transferred to the CG unsaturated carbon particles in DOPC. The PMF results show that this assumption is valid in this case. Modification may be needed in the future.

Figure 1b shows the graphical representation of the DOPC system for PMF calculation. The far right of the PMF curves in Figure 2 indicates that the side chains stay in water and have very little or even no interactions with the lipid bilayer. This corresponds to the solvation free energy of side chains in aqueous phase and is taken as the reference point. The far left of the curves indicates that the side chains stay at the center of the lipid bilayer. As the hydrophobic thickness of the DOPC bilayer is large (2.96 nm determined experimentally),[84] the side chains can be considered to be solvated in the nonpolar environment. As the last point of the PMF curve is taken as the reference point, the first point of the PMF curve could be regarded as the transfer free energy of the side chain from water to the hydrophobic environment. The accuracy of the calculated transfer free energies indicates the transferability of our parameters. The results show that the transferability of our model is reasonably good and can be transferred to different types of lipid bilayer. The middle region of the PMF curves shows the free energy of interactions between the side chain and lipid head groups relative to the reference state. To obtain correct descriptions of protein−lipid head interactions, the free energy profiles of this region must be accurately fitted.

We used the similar strategy as the parametrization of protein−lipid tail interactions. Combination rule was used to obtain δ_{ij} and ε_{ij} was optimized by fitting the all-atom PMFs. Parameters for $-CH_3$ were obtained from PMF of Ala. Then, PMFs of Val and Ile were used for the parametrization of $-CH_2$. Parameters of −CH and aromatic C could be optimized in a similar way using Leu and Phe, respectively. Other functional groups were parametrized by fitting PMFs of all amino acid side chain analogues. Lipid head composes of three CG particles: glycerol, phosphate, and choline. As they are at different positions of the lipid bilayer, we found that change of the parameters would have a different effect to the PMF curves. Phosphate is at the center of the PMF curve, the PMF at ∼0 nm is the most sensitive to the change of the parameters. Glycerol is nearer to the lipid tail than phosphate. It mainly affects the PMF at about −0.5 nm. Choline is closest to water, and the effect of the parameters is mainly reflected at ∼0.5 nm of the PMF. Interactions between UA particles and these three CG particles were parametrized by carefully fitting the whole PMF curves.

Alkane compounds have a high transfer free energy from hexadecane to water (∼10−18 kJ/mol in Table 1). It is expected that hydrophobic residues, Ala, Val, Ile, and Leu prefer to stay in the lipid core. Our results are in very good agreement with the all-atom results. When the size of the hydrophobic side chains increases, the interaction with nonpolar lipid tails increases, and the free energy difference becomes larger. The free energy difference of hydrophobic residues between water and the center of the bilayer is about 8.4 to 22.1 kJ/mol. CG results for Ala, Val, and Leu are in excellent agreement with the all-atom results. The deviation is about 2 kJ/mol for Ile. Although both the side chains of Leu and Ile consist of four carbon atoms with the same interaction parameters with lipid tails, the relative free energy at the center of bilayer for Ile is 7 kJ/mol more favorable than that for Leu. This is because the branched Leu side chain packs into the lipid bilayer less efficiently than the linear Ile side chain. This effect can be observed in our model where every heavy atom of amino acids is represented explicitly. This phenomenon is absent when these side chains are represented by the same CG particles, indicating the need for finer protein models to describe the interactions between proteins and lipid heads. The free energy barrier at the headgroup region in our model is present in these cases. The deviation from the all-atom results is about 1 kJ/mol for Val, Ile, and Leu. The barrier is underestimated by about

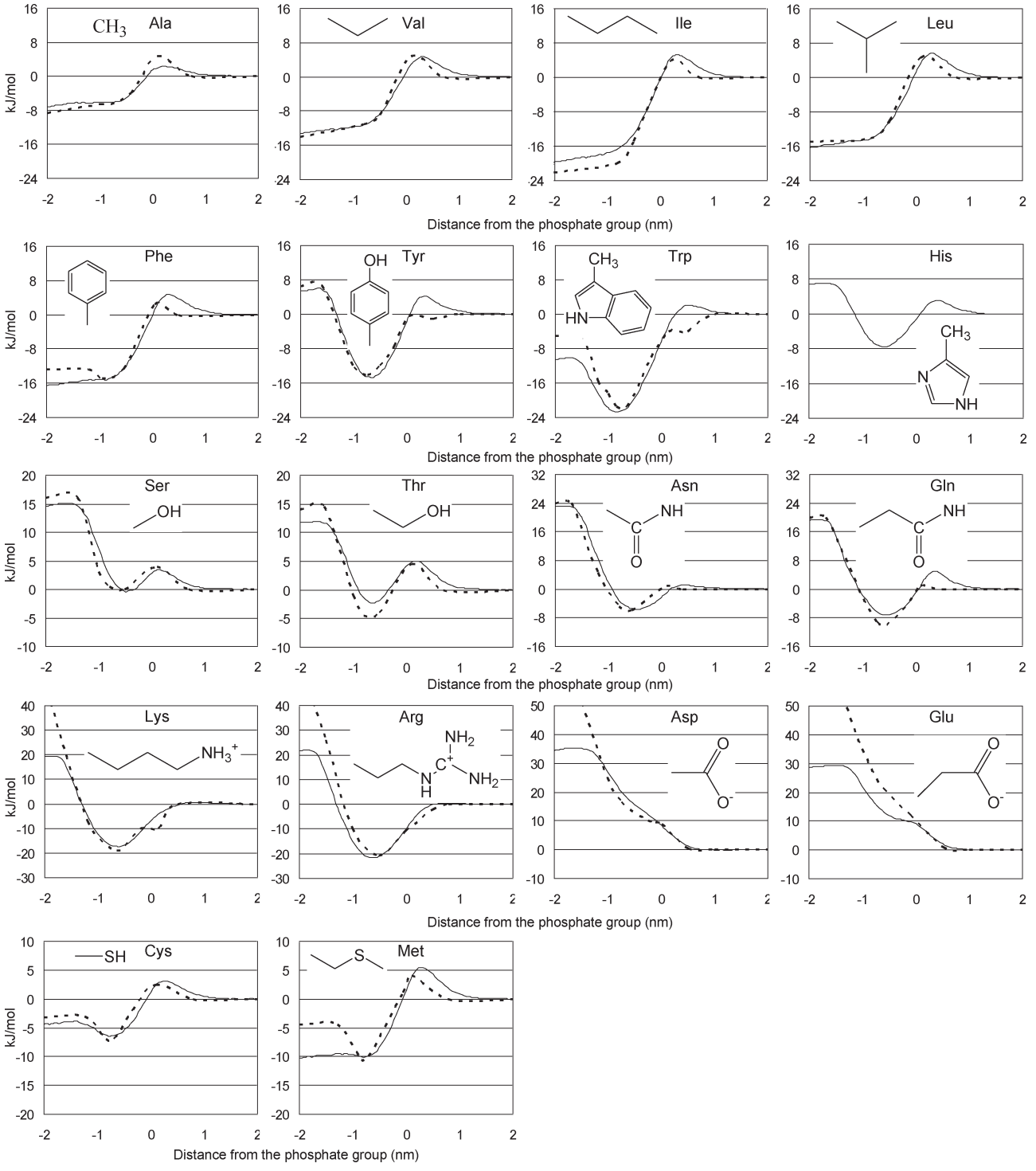

Figure 2. PMFs of 18 amino acid side chain analogues. CG results and all-atom results[82] are shown in solid lines and dotted lines, respectively.

2.5 kJ/mol for Ala. Note that in the case of Ala, a single UA particle of CH_3 instead of CH_4 is used for parametrization. This may account for the small deviation of the PMF curves of Ala.

For aromatic residues, Tyr, Trp, and His have strong interactions with lipid head groups, as they are able to form H-bonds with lipid head groups. The minima for Tyr and Trp are well reproduced. The relative free energies at the center of the bilayer are reasonably reproduced for Tyr. For Phe and Trp, our model shows 4−5 kJ/mol more stabilization than the all-atom model. All-atom simulations for the solvation free energy calculations of 15 neutral amino acid side chains using the AMBER(ff94), CHARMM22, and OPLS-AA force fields show an average error of 4.6−6.8 kJ/mol.[85] All-atom calculations of transfer free energies from hexadecane to water for all amino acid side chains using the OPLS-AA force field have an average error of about 4 kJ/mol.[86] Considering the errors associated with the solvation parameters of all-atom force fields, no further optimization was done to minimize the transfer free energy difference between our

Table 3. Optimized Parameters for Protein−Lipid Head Interactions

	CG lipid headgroup					
	glycerol (N_a)		phosphate (Q_a)		choline (Q_0)	
protein UA group	ε_{ij} (kJ/mol)	δ_{ij} (nm)	ε_{ij} (kJ/mol)	δ_{ij} (nm)	ε_{ij} (kJ/mol)	δ_{ij} (nm)
$-\underline{CH_3}$	1	0.43	0.4	0.43	0.4	0.43
$-\underline{CH_2}$	1	0.43	0.4	0.43	0.4	0.43
$-\underline{CH}$	0.4	0.43	0.4	0.43	0.4	0.43
aromatic C[a]	0.7	0.4225	0.7	0.4225	0.7	0.4225
$-N$[b]	10	0.4	17	0.4	15[k]	0.4
$O{=}C-\underline{NH_2}$[c]	2.1	0.4	2.1	0.4	2.1	0.4
$O{=}C-\underline{NH_2}$[c]	5	0.375	5	0.375	0	0.375
$-\underline{C}{=}O$[c]	1	0.415	0.4	0.415	0.4	0.415
$-C{=}\underline{O}$[c]	2	0.375	20[k]	0.375	2	0.375
$-\underline{NH}$[d]	2.1	0.4	2.1	0.4	2.1	0.4
$-\underline{NH}$[d]	6	0.375	6	0.375	0	0.375
$-C{=}\underline{N}$[e]	2.1	0.4	2.1	0.4	2.1	0.4
$-N$[f]	6	0.4	6	0.4	15[k]	0.4
$-C^{+}$[f]	1	0.415	6	0.415	15[k]	0.4
$-\underline{COO^{-}}$[g]	1	0.415	0.4	0.415	0.4	0.415
$-\underline{COO^{-}}$[g]	1.7	0.375	20[k]	0.375	7	0.375
$-\underline{OH}$[h]	6	0.38	2	0.38	2	0.38
$-\underline{OH}$[i]	8	0.38	2	0.38	2	0.38
$-S-$[j]	4	0.405	3	0.405	3	0.405

[a] Aromatic C for Phe, Tyr, Trp and His side chains. [b] −N for Lys and Arg side chains. [c] −C═O and O═C−NH_2 for backbone amide, Asn, and Gln side chains. [d] −NH for Trp and His side chains. [e] −C═**N** for His side chain. [f] −N and $-C^{+}$ for Arg side chain. [g] $-COO^{-}$ for Asp and Glu side chains. [h] −OH for Ser and Thr side chain. [i] −OH for Tyr side chain. [j] −S− for Cys and Met side chains. [k] The following potential is used: $E_{\text{nonbonded}} = \Sigma_{i\neq j}(4\varepsilon_{ij}\delta_{ij}{}^{12})/r^{12}$.

model and the all-atom model. Our model shows free energy barriers of about 2−4.5 kJ/mol but only about 2.5 kJ/mol (Phe) or even no barrier (Tyr and Trp) in the all-atom results. The existence of the barrier may be due to the lipid model, as it is also found in the MARTINI model. There is room for improving our force field. As there is no all-atom result for His, a comparison cannot be made.

For polar residues, Ser, Thr, Asn, and Gln each contain a hydroxyl group or amide group that is able to form H-bonds with lipid head groups. Our model can reproduce the minima of these residues at the lipid head region accurate to ±3 kJ/mol. The heights of the barrier for Ser and Thr are in very good agreement with the all-atom results. Asn shows no barrier in both our model and the all-atom model. But there is a barrier for Gln in our calculation that is absent from the all-atom simulation.

For all-atom PMFs, the relative free energies of ionizable residues at the center of the bilayer is very large (>50 kJ/mol). This is because charged molecules have large desolvation penalties when they are moved from the aqueous medium to the nonpolar environment. Our results show smaller relative free energies at the center of the bilayer as these residues are considered to be neutral in the membrane, but they were assumed to be charged in the all-atom calculations. Tieleman et al. calculated pK_a values of ionizable groups in DOPC using all-atom model. Asp and Glu were found to be neutral in the lipid core, whereas Lys and Arg prefer charged state.[82] It indicates that our assumption may be valid for Asp and Glu only. Introduction of charged Lys and Arg will be considered in the future. The ionizable residues are more likely to enter the lipid core in our model than in the all-atom models. The middle region of the PMF curves, which was our main focus and indicates the interactions between protein particles and the lipid head, was fitted to the all-atom results. For Lys and Arg, the free energy minima indicating strong polar−polar interactions with lipid head groups are well reproduced. In the case of negatively charged side chains, the PMF of Asp was fitted to the all-atom result, and the optimized parameters were transferred to Glu. Both residues show unfavorable interactions with the lipid bilayer. However our model is somewhat less unfavorable for Glu in the middle region than the all-atom calculation. This may indicate that Asp and Glu should not share the same parameters and that further improvement should be made in the future.

Both the PMFs of Cys and Met match the all-atom results except that Met shows more favorable interactions with the lipid tail in our model than in the all-atom model. Our model shows that Met is 10 kJ/mol more stable in the lipid tail, whereas it is only 4.4 kJ/mol more stable in the lipid tail in the all-atom calculation. The experimental transfer free energy of Met from cyclohexane to water is 9.8 kJ/mol.[87] These results show that our model is comparable to the all-atom simulations in describing the partitioning of amino acid side chains in the lipid bilayer.

Modification of Backbone Hydrogen-Bonding Potential in Membrane Environment. Before we present the detailed results obtained with the optimized parameters for lipid−protein interactions, the hydrogen-bonding (HB) potential of PACE needs to be further investigated. The original HB parameters of PACE were obtained by reproducing experimental α-helical and β-sheet contents of model peptides in aqueous simulations.[31] As the simulations were carried out with the MARTINI water model that bears no dipole moment, the screening effect of water was

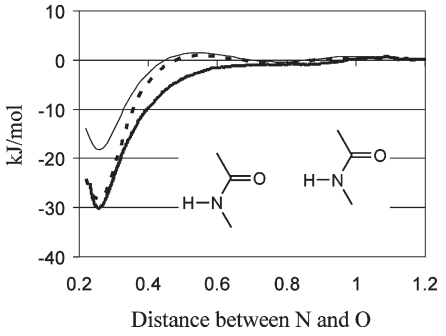

Figure 3. PMFs of two amides in pure octane. CG calculations using original and modified HB potentials and all-atom calculation are shown in thin, dotted, and thick lines, respectively.

implicitly taken into the original HB parameters. Hence, the electrostatic part of HB may be underestimated in membrane environment. We tried to estimate the underestimation of the electrostatic part by calculating the dimerization free energy of *N*-methylacetamide in octane with both the original HB parameters of PACE and OPLS-AA.[88] As shown in Figure 3, the HB parameters of PACE lead to the dimerization energy of 18.2 kJ/mol in octane, whereas OPLS-AA gives dimerization energy of 29.5 kJ/mol. To account for the difference between the PACE and OPLS-AA results, we generated a modified PACE in which ε_{attr} (eq 8) was increased so that the dimerization energy of *N*-methylacetamide in octane was reproduced.

To examine the effect of the modification, we performed the simulations of WALP-based peptides and glycophorin A helix dimer in membrane with both the original and modified PACEs. The comparison reveals that all the WALP-based peptides have stable helices and similar tilt angles for both HB potentials of PACE (Table S1, Supporting Information). For the glycophorin A helix, the helical structure in the region of GVIG becomes unstable in the simulation with the original PACE, owning to the helix-breaking propensity of the two glycine in the middle of helix. The helical structure in the same region is however stable in the simulation with the modified PACE. The comparison suggests that the original PACE can be used to simulate the membrane helical peptides without glycine in the middle of helix. For the peptide like the glycophorin A helix, the modified PACE is needed. In the following sections, only the results in the simulations with the modified PACE will be discussed in detail.

It should be noted that the major limitation of our way of modifying the HB parameters is that all HB will be strengthened, although the dielectric environment across membrane is quite different[89] and thus HBs located at different places of membrane should have distinct strength. Moreover, when HBs dynamically change their positions during simulations, their strength should also change accordingly. Therefore, a more desirable model should be able to adjust HB strength on the fly, which is under our investigation. Alternatively, a recent polarizable CG water model in MARITNI may also be a good choice.[90]

Tilting of Helical Peptides. Different transmembrane peptides show significant variations in the tilt angle in lipid bilayers. For example, the influenza A M2 channel was reported to have a tilt angle of 38°, whereas the channel-lining M2 segments from the d-subunit of the nicotinic acetylcholine receptor shows a tilt angle of 14°.[91,92] Artificial transmembrane peptides were designed to systematically study the factors affecting the orientation of transmembrane peptides. It is generally accepted that hydrophobic mismatch affects the orientations and therefore functions of membrane proteins. When the hydrophobic length of the peptide is larger than the hydrophobic width of the lipid (positive mismatch), the peptide tilts to allow it to have better interactions with the lipids. When the hydrophobic length of the peptide is smaller than the hydrophobic width of the lipid (negative mismatch), the tilting of the peptide is smaller.[93]

Table 4. Amino Acid Sequences of the Peptides Used in the Literature and in Simulation

peptide	sequence
WALP23	acetyl-GWW$(LA)_8$LWWA-amide
WALP19	acetyl-GWW$(LA)_6$LWWA-ethanolamine
WALP19-P10	acetyl-GWW$(LA)_3$P$(AL)_3$WWA-ethanolamine
GWALP23	acetyl-GGALW$(LA)_6$LWLAGA-amide
KWALP23	acetyl-GKALW$(LA)_6$LWLAKA-amide

In this work, we calculated the tilt angles of WALP and GWALP peptides and their mutants. Table 4 shows the amino acid sequences of the peptides that were studied. WALP peptides were chosen because they have been widely studied in experiments and computer simulations.[34−43] It is interesting to compare our results with theirs. We found that the dynamic model should be used instead of the static model to determine the tilt angle. We also studied GWALP and KWALP peptides as they have recently been examined using the dynamic model.[44] Moreover, we report the first simulation result of WALP19-P10 which shows a different tilting preference from WALP19.[45]

The peptides were vertically inserted into the pre-equilibrated lipid bilayers with initial fully α-helical structures. Each simulation was performed for 300 ns, and the last 250 ns were used for analysis. The average tilt angles of the simulated peptides and the calculated hydrophobic widths of lipids and lengths of peptides are shown in Table 5. The tilt angle of a peptide was calculated as the angle between the helical axis of the hydrophobic segment of the peptide and the normal of the lipid bilayer. As the bilayer normal was aligned along the *z*-axis, the *z*-axis was taken so there was no need to calculate the bilayer normal. The hydrophobic width of the lipid bilayer is defined as the average distance between the first hydrophobic beads in the two leaflets. The peptide hydrophobic length is the distance of all the hydrophobic residues between the anchoring groups. For ideal helical peptides, the hydrophobic length for each residue is 0.15 nm. The deviations of the calculated hydrophobic width of lipids and length of peptides from experimental data were smaller than 0.05 nm. The calculated tilt angles of WALP19 and WALP23 were in the range of 7.5° to 17.5° which shows the hydrophobic mismatch effect. The tilting of WALP19 was determined using the static model in experiment. We found a larger tilting than experimental results. This suggests that the dynamic model should be used. GWALP and KWALP also show a hydrophobic mismatch effect, and the calculated tilt angles are in good agreement with experimental results which used the dynamic model. Interestingly, WALP19-P10 did not show a hydrophobic mismatch effect and has the same tilting in both DOPC and DLPC. Our simulation supports the experimental finding that the Pro10 affects the tilting.

One of the most extensively studied artificial transmembrane peptides is WALP.[34−43] WALP is a poly-(Leu-Ala) peptide with two Trp residues at both ends. It has been used to develop and validate CG force fields.[21,101] The exact tilting of helical peptides

Table 5. Average Tilt Angles of WALP, GWALP Peptides and Their Mutants and Calculated Hydrophobic Widths of Lipids and Lengths of Peptides[a]

	tilt angle (°)				calculated lipid hydrophobic width[g] (nm)		calculated peptide hydrophobic length[h] (nm)	
	DOPC		DLPC					
peptide	sim.	exptl.	sim.	exptl.	DOPC	DLPC	DOPC	DLPC
WALP23	7.5 (5.0)	11[c]	17.5 (7.6)	15[d]	3.01 (0.05)	1.93 (0.04)	2.50 (0.03)	2.49 (0.03)
WALP19	8.0 (5.7)	4.0[e]	13.5 (7.2)	4.0[e]	2.99 (0.05)	1.93 (0.04)	1.94 (0.02)	1.94 (0.02)
WALP19-P10[b]	15.0 (8.0)	11.6[e]	15.5 (8.8)	11.9[e]	2.99 (0.05)	1.92 (0.05)	1.97 (0.04)	1.97 (0.04)
tilt difference	7.0	7.6	2.0	7.9				
GWALP23	6.5 (5.3)	6.0[f]	15.5 (7.5)	18.6[f]	2.99 (0.05)	1.93 (0.04)	1.94 (0.02)	1.94 (0.02)
KWALP23	7.5 (5.7)	7.3[f]	18.0 (8.0)	18.0[f]	3.00 (0.05)	1.93 (0.04)	1.94 (0.02)	1.94 (0.02)
tilt difference	1.0	1.3	2.5	−0.6	exptl value: 2.96[i]	exptl value: 1.95[j]		

[a] All the experimental tilt angles were determined using the dynamic model except for WALP19 and WALP19-P10. Standard deviations are shown in parentheses. [b] Average tilt is calculated for the C-terminal segment of WALP19-P10. [c] Ref 43. [d] Ref 98. [e] Ref 45. [f] Ref 44 [g] The lipid hydrophobic width is the average distance between the first hydrophobic beads in the two leaflets. [h] The peptide hydrophobic length is the distance between the anchoring groups for all the hydrophobic residues. For ideal helical peptides, the hydrophobic length for each residue is 0.15 nm. [i] Ref 84. [j] Ref 113.

is still debated, as experimental and simulation results differ greatly.[94] WALP23 was found to have a tilt angle of about 12° in dimyristoylphosphatidylcholine (DMPC) by ATR-FTIP spectroscopy.[95] Recent studies by fluorescence spectroscopy found large tile angles for WALP23 in DOPC (23.6°).[96] Koeppe et al. developed another method that uses solid-state ^{2}H NMR based on geometric analysis of labeled alanines (GALA) to study the orientation of a transmembrane peptide. This technique allows a higher resolution of tilt angles (<1°), but small tilt angles were determined for WALP23 in DOPC (4.8°) and DLPC (8.1°).[38,80]

The small angles found by the GALA method have been explained by computer simulation studies.[9,97] In the conventional GALA method, the tilt angle of a peptide is calculated by fitting two parameters, τ (tilt angle) and ρ (rotation angle) to the experimentally measured quadrupolar splitting of labeled alanines. Fluctuation of the peptide leads to an averaging effect so the tilt angle would be underestimated. To take the fluctuation of the peptide into account, the global order parameter S could be used as a fitting parameter instead of taken as a constant. With this parameter, the effects of additional internal vibrations, rotations, and wobbling of the peptide are included (model 3 in ref 98). Tilt angles of WALP23 in DMPC and DLPC have been found to be 7° and 15°, respectively.[98] This dynamic model has been used in recent studies to determine the tilt angles of the GWALP peptide and its mutants.[44,81] Strandberg et al. proposed to explicitly consider the fluctuations of τ and ρ by introducing the Gaussian distribution to them. A total of four parameters including two additional parameters, σ_τ and σ_ρ, were used to fit the experimental data (model 6 in ref 98). Larger tilt angles were obtained for WALP23 in DMPC (14°) and DLPC (29°).[98]

All-atom MD simulations show larger tilt angles compared with the results from solid-state ^{2}H NMR. In a study using the CHARMM22 force field with the GBSW implicit membrane, tilt angles of 32.7° and 15.5° were found for WALP23 and WALP19, respectively, in a 0.23 nm thick membrane hydrophobic core that corresponds to DMPC.[41] A similar result was obtained when the ffgmx force field was used with the explicit membrane (33.5° for WALP23 in DMPC).[9] Slightly smaller tilt angles were determined for WALP23 (28.1°) and WALP19 (12.1°) in DMPC with the CHARMM22 force field and the explicit membrane.[99,100] A tilt angle of 14° was found for WALP23 in DPPC using the MARTINI CG force field and Bond and Sansom's protein model. But these models showed larger tilting for shorter peptides (22° for WALP19).[101] Another study using the MARTINI protein and lipid model showed a tilt angle of 11.4° in DOPC and 23.7° in DLPC for WALP23.[102]

In our work, WALP23 shows moderate tilting in DOPC (7.5°) and DLPC (17.5°). The results are in good agreement with that determined by the GALA method using the dynamic model. This supports the argument that the dynamic motion of the peptide should be considered. But the values are smaller than those obtained using the Gaussian distribution. The question of whether the Gaussian distribution is suitable for describing the motion of peptides has been raised.[102] The fluctuations of tilt and rotation angles are assumed to follow a Gaussian distribution. But both all-atom simulations and long CG simulations showed non-Gaussian distributions. In particular, the rotation angle was found to follow a distribution that is far from the Gaussian one.[81,102,97] Our results do not show a strict Gaussian distribution for the tilt angle either (Figure 4). When a uniform distribution was applied instead of a Gaussian distribution, a larger tilt angle was found for WALP23 in DMPC (21° rather than the 14° in ref 98).[103] The most suitable functional form to describe the motion of peptides has not yet been determined. Our results suggest that introducing S as a free parameter is enough to account for the motion of peptides.

WALP19 in DOPC and DLPC was found to have tilting angles of 8.0° and 13.5°, respectively (Table 5). The GALA method without considering the dynamic motion of peptides showed smaller tilt angles in DOPC and DLPC (4°). WALP19 and WALP23 in DOPC show small tilt angles as expected under negative mismatch. WALP19 in DLPC is under the hydrophobic match condition. As results determined by the dynamic model are unavailable, we suspect that WALP19 may have a tilt angle of about 13.5° in DLPC. Similar tilt angles were also obtained for peptides under the hydrophobic match condition by all-atom (KALP23 in DMPC) and CG (WALP23 in DPPC) simulations.[93,102] The hydrophobic mismatch effect is mainly due to favorable helix-lipid interactions.[100] The anchoring groups of WALP at the terminals, Trp, were mainly situated at the lipid head region. This is because Trp side chains interact favorably

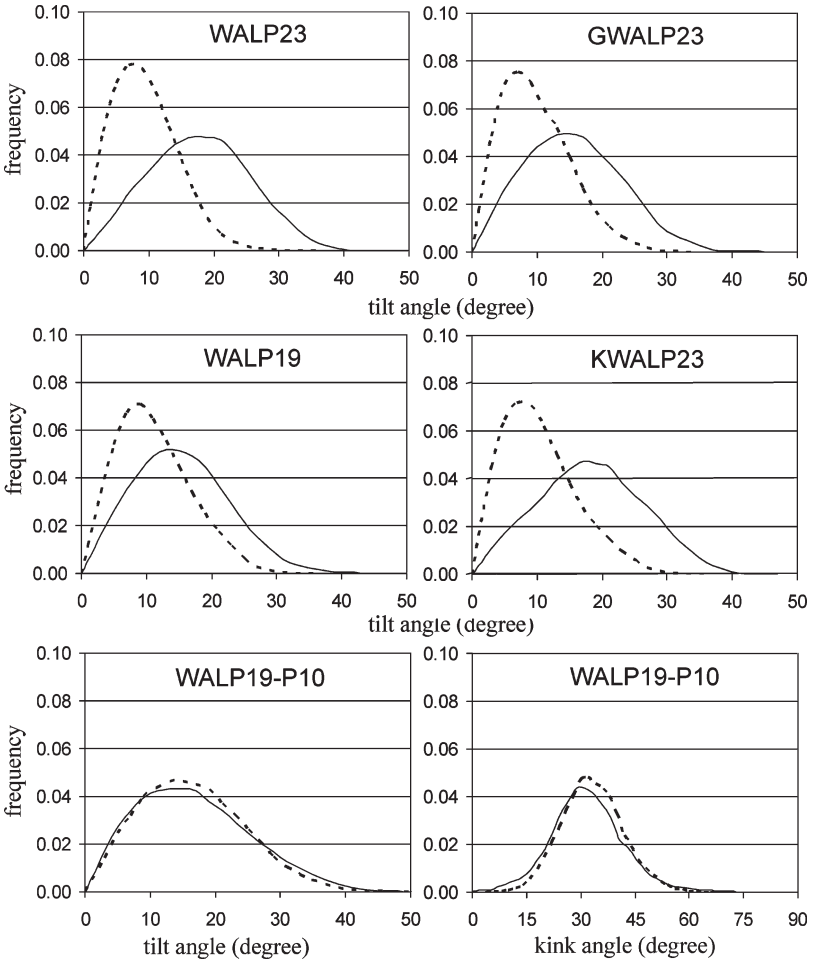

Figure 4. Distribution of tilt and kink angles of WALP and GWALP peptides and their mutants in the DOPC (dotted line) and DLPC (solid line) bilayers.

(22 kJ/mol) with lipid head groups. (Figure 2) Therefore, WALP23 has a larger tilt angle in DLPC than WALP19 in order to position the anchoring groups at the lipid head region.

The effect of hydrophobic mismatch is also observed for another family of peptide, GWALP. Unlike WALP peptides, GWALP peptides contain only one Trp residue at both terminals and positioned at the inner region. Tilt angles of GWALP23 and its mutant, KWALP23, in DOPC and DLPC were determined using a dynamic model similar to model 3 in ref 98. The tilt angle of GWALP23 in DLPC was analyzed by fitting ^{2}H quadrupolar splittings and ^{1}H−^{15}N dipolar coupling data. Tilt angles of 6.0° and 18.6° were found for GWALP23 in DOPC and DLPC, respectively.[44] Our results match the experimental findings very well. There is no significant difference in the tilt angle between KWALP23 with an Lys residue at the two ends and GWALP23.[44] Lys is also an anchoring group that prefers to position at the lipid interface and has been shown to affect the tilting of the transmembrane peptides.[80] In the case of KWALP23 containing both Lys and Trp residues, inner Trp seems to determine the tilting but outer Lys does not. Our results are in excellent agreement with the experimental data for KWALP23 and also support this finding from the experiment.

Apart from the studies of artificial transmembrane peptides with Leu and/or Ala in the hydrophobic core, mutation at the center of the peptides has been carried out to study the effect of a particular amino acid on the orientation of the peptide.[45,81] WALP19-P10 is a mutant of WALP19 with Leu10 replaced by Pro. Pro has no backbone amide hydrogen and is a helix-breaking residue. It was confirmed by circular dichroism spectroscopy that the helicity of WALP19-P10 is smaller than that of WALP19.[45] Our simulation result shows that the helix is broken at Pro10 and that two helical segments were observed instead of one. (Figure 5) A tilting angle of ∼12° was found for the C-terminal segment in both DOPC and DLPC by solid-state ^{2}H NMR spectroscopy.[45] A tilt angle of ∼15° was obtained in our calculations (Table 5). The tilting of WALP19-P10 in DOPC is apparently larger than that of WALP19. This indicates that the kink at Pro10 affects the tilt angle of the C-terminal segment. A kink angle of ∼19° was found experimentally for WALP19-P10.[45] Our simulation found average kink angles of 32° and 31° in DOPC and DLPC, respectively (Figure 4). Kink angles ranging from 9° to 41° were found for 48 Pro-containing helices in the crystal structures of soluble proteins.[104] Kink angles of 0°−70° were observed for 50 transmembrane helices containing proline in the crystal structures of membrane proteins.[105] Upon using the static model for the analysis of WALP19-P10 in the experiment, a kink angle of ∼19° was suggested to be the lower limit.[45] Our result suggests that the kink angle may be as large as ∼30° and that the kink also affects the tilting preference.

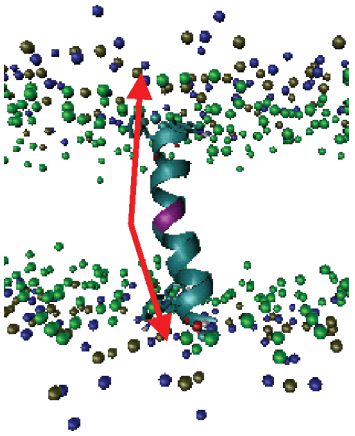

Figure 5. Snapshot of the WALP19-P10 peptide simulation in the DLPC bilayer. The residue Pro10 is shown in purple. The helical axes of N and C terminal segments are shown by the upper and lower arrows, respectively. The choline, phosphate, and glycerol particles of the DLPC molecules are shown in blue, brown, and green respectively.

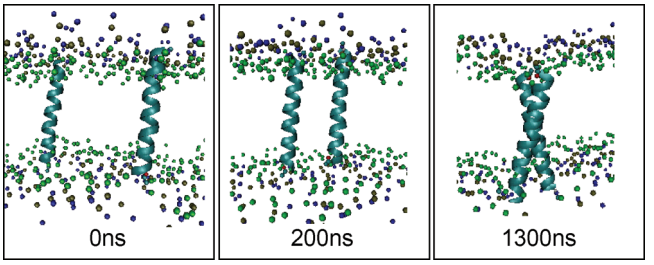

Figure 6. Snapshots of GpA helix associations in the DPPC bilayer at 0 ns, 200, and 1300 ns. The initial configuration of two GpA helices has a separation of ~3.5 nm. The choline, phosphate, and glycerol particles of the DPPC molecules are shown in blue, brown, and green, respectively.

Association of Glycophorin A Helix. Helix bundles are commonly found in membrane proteins. Understanding the folding and assembling of them enables us to predict their structures and to design new membrane proteins.[106,107] The two-stage folding model has been proposed for the assembling of transmembrane helices: Each transmembrane helix is inserted into the membrane independently followed by the association of helices.[108] In this work, we studied the association of the GpA helix. The GpA dimer contains a seven-residue motif ($L^{75}I^{76}xxG^{79}V^{80}xxG^{83}V^{84}xxT^{87}$) that has been found to be important in the packing and dimerization of GpA helices.[109] Previous CG simulations have attempted to study such processes.[67,68] Unfortunately, although a stable dimer was observed, significant deviations of the simulated structure from the experimental one were evident. It would be interesting to use the hybrid-resolution model to study such a system to validate our force field.

Two GpA helices (acetyl-EITLIIFGVMAGVIGTILLISYGIR-methylamide) were inserted into DPPC in a parallel fashion with a interhelix separation of ~3.5 nm. (Figure 6) After a 1 ns pre-equilibrium simulation with a position constraint on the peptides, the two helices were allowed to move freely in the system box for 2 μs. Three individual trajectories with different initial velocities were obtained, with a total of 6 μs of simulations. Interhelix separation, defined as the distance between the center of mass of the two GpA helices, as a function of time for three runs is shown in Figure S2, Supporting Information. The two helices met one another at ~200 ns, after which no disassociation was observed throughout the simulations. This indicates that the helix dimer is stable and is consistent with experiments and other CG simulations.[53,67,69,110,111] The ensemble after 300 ns of simulation was used for the analysis of the structural features of the GpA dimer in each run. Thermodynamic properties were not evaluated, as the trajectories had not fully converged.

Table 6. RMSD from NMR Structure in DPC Micelle and the Crossing Angle of GpA Dimer in Three Runs

	Cα RMSD (nm)[a]		crossing angle Ω (°)	
GpA simulation	all residues	seven-residue dimerization motif[b]	our work	expt
run 1	0.228 (0.051)	0.203 (0.045)	−37.0 (7.8)	−40[c], −35[d]
run 2	0.294 (0.032)	0.280 (0.029)	−43.5 (6.1)	
run 3	0.296 (0.092)	0.260 (0.052)	−32.5 (8.7)	
concatenated	0.272 (0.071)	0.249 (0.059)	−35.8 (9.8)	

[a] Average RMSDs were calculated relative to the DPC micelle NMR structure. [48] [b] Only seven residues in the dimerization motif (LIxxGVxx-GVxxT) were considered in the calculation. [c] Derived from solution NMR structure in DPC micelle. [48] [d] Derived from solid-state NMR structure in the DMPC bilayer.[52] Standard deviations are shown in parentheses.

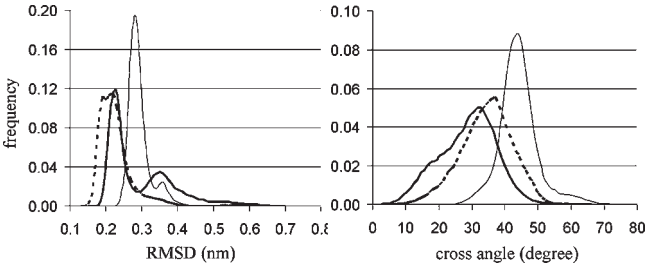

Figure 7. Distributions of RMSD relative to the NMR structure in DPC micelle and the crossing angle of GpA dimer in three runs. RMSD calculations were performed for all residues. Runs 1−3 are represented by dotted, thin, and thick lines, respectively.

Table 6 shows the Cα root-mean-square-deviation (RMSD) of the simulated dimers from the solution NMR structure in DPC micelle. The distributions of Cα RMSDs considering all residues for all three runs are shown in Figure 7. The most probable RMSDs are ~0.22 nm for two of the runs and ~0.28 nm for another run. On average, ~14% of all the simulated structures have RMSD ≤0.2 nm, and ~63% of structures have RMSD between 0.2 nm and 0.3 nm. When the seven-residue dimerization motif was considered only in the calculation, the Cα RMSDs were reduced by 0.014−0.036 nm, with the RMSD of the concatenated trajectory shrinking to just 0.249 nm. These values are smaller than that obtained using the MARTINI CG protein force field (0.36 nm).[47] The standard deviation for concatenated trajectory (0.071 nm for all residues and 0.059 nm for the seven-residue dimerization motif) was also lower than that obtained using the MARTINI CG protein force field (0.13 nm). Our model may give a better structural representation of proteins in membrane (Figure 8).

Another well-defined structural feature of the GpA dimer is the crossing angle between the two helices. The GpA dimer has

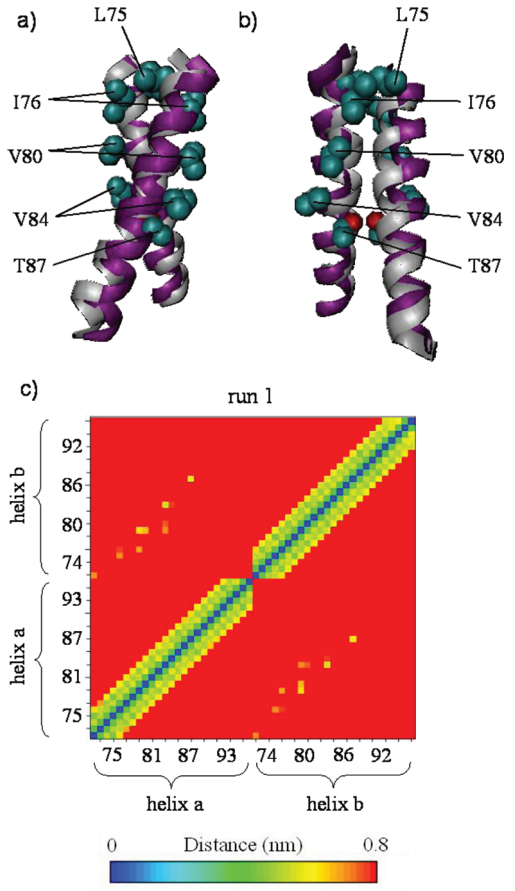

Figure 8. Representative structure of the simulated GpA dimer in purple in: (a) front and (b) side views. Side chains of key residues, L75, I76, V80, V84, and T87, are represented by green (hydrophobic particles) and red (hydrophilic particles) spheres. NMR structure in micelle is shown in gray for reference. (c) Contact map of the GpA dimer for the first run. The matrices correspond to the ensemble of the simulations with the cutoff distance of 0.8 nm. Seven common contacts were found in the three runs: L75−I76, I76−G79, G79−G79, G79−V80, G79−G83, G83−G83, and T87−T87.

right-handed packing with a negative crossing angle. The experimental crossing angles were −40° and −35°, as determined by solution NMR for DPC micelle and solid-state NMR for DMPC bilayer, respectively.[48,52] Different media may perturb the structure of the GpA dimer leading to slightly different values obtained in the micelle and lipid bilayer. The average crossing angles for the three runs were −37.0°, −43.5°, and −32.5° (Table 6). They are all consistent with the experimental findings. The result (−35.8°) of the concatenated trajectories is in excellent agreement with the experimental value determined for lipid bilayer (−35°). This is supported by all-atom simulations which find that the crossing angle in the lipid bilayer is smaller than that in micelle by 3°−7°.[60] The distributions of the crossing angle are shown in Figure 7. The distributions are larger than those obtained in all-atom simulations. The standard deviation of the crossing angle is 9.8° for the concatenated result and 2°−5° for all-atom results using the GROMOS force field.[60] This may be due to the short simulation length used by the all-atom force field (50 ns) or the less rugged free energy landscape of the CG force field. Large distributions of the crossing angle were also found using the MARTINI protein force field. The average crossing angle determined by previous CG simulation was −20° to −25°, smaller than the experimental values by 10°−15°.[67,69] Moreover, a positive crossing angle not observed in NMR structures was sampled. This indicates the limitation of the current CG protein force field and the higher accuracy and applicability of the PACE force field in studying helix assembly.

To further evaluate the structural features of the simulated GpA dimer, the contact maps for three runs were calculated and are shown in Figure 8c and Figure S3, Supporting Information. Seven common contacts were found in the three runs: L75−I76, I76−G79, G79−G79, G79−V80, G79−G83, G83−G83, and T87−T87. Residues involved in these contacts are in fact all the key residues of the $L^{75}I^{76}xxG^{79}V^{80}xxG^{83}V^{84}xxT^{87}$ motif. The contact analysis showed that binding of helices is mainly due to interactions among these residues. In a statistical study of the most frequently occurring motif that mediates helix−helix interactions, 13 606 transmembrane domains were analyzed.[112] GxxxG was found to be the most significant motif. Among seven common contacts found in our simulations, five of them involve Gly. This is because Gly has no side chain. This less bulky residue allows closer packing of helices. Even though Val84 was not one of the common contacts in the three runs, it is nevertheless one of the contacts in the runs 1 and 2. Previous simulations only showed three contacts (L75−I76, G83−G83, and T87−T87) using the same cutoff distance.[67] It may imply a less closely packed dimer and thus a larger RMSD value.

Interestingly, our calculations and other all-atom simulations also found T87−T87 contacts.[60,68] This type of contact is due to the interhelix polar interactions between the hydroxyl groups of Thr87. In the concatenated result, ∼50% of the time frames showed a distance between the hydroxyl groups of two Thr87 of less than 0.4 nm. Thr87 was in g− conformation in the initial simulated structure of GpA helices. In order to have side chain interaction between two Thr87, both Thr87 should rotate to adopt g+ conformation to bring their two hydroxyl groups closer. Our simulation results were consistent with the all-atom calculations which find that both Thr87 preferred g+ rather than g− conformation. As Thr was optimized to prefer g− conformation in our force field, a preference for g+ rotamer in this case is not an artifact of our model.[31] This result suggests that interhelix H-bonds between Thr may play a role in stabilizing the GpA dimer.

CONCLUSION

The PACE force field has been extended to include lipids. The interactions between protein particles and lipid tails were parametrized by reproducing experimentally the free energy of solvation in hexadecane using small organic molecules. The average absolute error is about 1.0 kJ/mol. The transfer free energies from hexadecane to water were also calculated with an error of 1.1 kJ/mol. The interactions between protein particles and lipid heads were parametrized by fitting the corresponding PMFs obtained using all-atom simulations. Tilt angles of WALP and GWALP peptides and their mutants were calculated in different lipid bilayers. The results are in good agreement with the experimental data without having to make assumptions about the motions of the peptides. Association of glycophorin A helices was performed for 6 μs. The simulated dimer in this work closely resembles the experimental structure. This is the first report showing the RMSD of the assembled GpA dimer lower than 0.3 nm compared to NMR structure. These results showcase the

high accuracy and efficacy of the extended PACE force field in studying membrane proteins.

■ ASSOCIATED CONTENT

Supporting Information. Figure S1shows schematic representation of hydrogen bonding potentials; Figure S2 shows interhelix separation of GpA helix dimer; Figure S3 shows contact map of the GpA dimer for the second and third runs; Table S1outlines calculated tilt angles of WALP, GWALP peptides, and their mutants with original and modified backbone hydrogen-bonding potentials. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
*E-mail: chydwu@ust.hk.

■ ACKNOWLEDGMENT

We are grateful to the Research Grants Council of Hong Kong (663509) and Peking University for financial support of this research.

■ REFERENCES

(1) Sachs, J. N.; Engelman, D. M. *Annu. Rev. Biochem.* **2006**, *75*, 707.
(2) Werten, P. J. L.; Rémigy, H.-W.; de Groot, B. L.; Fotiadis, D.; Philippsen, A.; Stahlberg, H.; Grubmüller, H.; Engel, A. *FEBS Lett.* **2002**, *529*, 65.
(3) von Heijne, G. *Nat. Rev. Mol. Cell Biol.* **2006**, *7*, 909.
(4) Phillips, R.; Ursell, T.; Wiggins, P.; Sens, P. *Nature* **2009**, *459*, 379.
(5) Bowie, J. U. *Nature* **2005**, *438*, 581.
(6) Roux, B.; Schulten, K. *Structure* **2004**, *12*, 1343.
(7) Gumbart, J.; Wang, Y.; Aksimentiev, A.; Tajkhorshid, E.; Schulten, K. *Curr. Opin. Struct. Biol.* **2005**, *15*, 423.
(8) Lindahl, E.; Sansom, M. S. P. *Curr. Opin. Struct. Biol.* **2008**, *18*, 425.
(9) Ozdirekcan, S.; Etchebest, C.; Killian, J. A.; Fuchs, P. F. J. *J. Am. Chem. Soc.* **2007**, *129*, 15174.
(10) Freddolino, P. L.; Liu, F.; Gruebele, M.; Schulten, K. *Biophys. J.* **2008**, *94*, L75.
(11) Shaw, D. E.; Maragakis, P.; Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Eastwood, M. P.; Bank, J. A.; Jumper, J. M.; Salmon, J. K.; Shan, Y.; Wriggers, W. *Science* **2010**, *330*, 341.
(12) Shelley, J. C.; Shelley, M. Y.; Reeder, R. C.; Bandyopadhyay, S.; Klein, M. L. *J. Phys. Chem. B* **2001**, *105*, 4464.
(13) Lopez, C. F.; Nielsen, S. O.; Srinivas, G.; DeGrado, W. F.; Klein, M. L. *J. Chem. Theory Comput.* **2006**, *2*, 649.
(14) Venturoli, M.; Smit, B.; Sperotto, M. M. *Biophys. J.* **2005**, *88*, 1778.
(15) de Meyer, F. J-M.; Venturoli, M.; Smit, B. *Biophys. J.* **2008**, *95*, 1851.
(16) Smeijers, A. F.; Pieterse, K.; Markvoort, A. J.; Hilbers, P. A. J. *J. Phys. Chem. B* **2006**, *110*, 13614.
(17) Markvoort, A. J.; Smeijers, A. F.; Pieterse, K.; van Santen, R. A.; Hilbers, P. A. J. *J. Phys. Chem. B* **2007**, *111*, 5719.
(18) Izvekov, S.; Voth, G. A. *J. Phys. Chem. B* **2009**, *113*, 4443.
(19) Marrink, S. J.; de Vries, A. H.; Mark, A. E. *J. Phys. Chem. B* **2004**, *108*, 750.
(20) Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H. *J. Phys. Chem. B* **2007**, *111*, 7812.
(21) Monticelli, L.; Kandasamy, S. K.; Periole, X.; Larson, R. G.; Tieleman, D. P.; Marrink, S. J. *J. Chem. Theory Comput.* **2008**, *4*, 819.
(22) Louhivuori, M.; Risselada, H. J.; van der Giessen, E.; Siewert J. Marrink, S. J. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 19856.
(23) Smirnova, Y. G.; Marrink, S. J.; Lipowsky, R.; Knech, V. *J. Am. Chem. Soc.* **2010**, *132*, 6710.
(24) Shih, A. Y.; Arkhipov, A.; Freddolino, P. L.; Schulten, K. *J. Phys. Chem. B* **2006**, *110*, 3674.
(25) Bond, P. J.; Sansom, M. S. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 2631.
(26) Khalfa, A.; Treptow, W.; Maigret, B.; Tarek, M. *Chem. Phys.* **2009**, *358*, 161.
(27) Shi, Q.; Izvekov, S.; Voth, G. A. *J. Phys. Chem. B* **2006**, *110*, 15045.
(28) Rzepiela, A. J.; Louhivuori, M.; Peter, C.; Marrink, S. J. *Phys. Chem. Chem. Phys.* **2011**, *13*, 10437.
(29) Han, W.; Wu, Y.-D. *J. Chem. Theory Comput.* **2007**, *3*, 2146.
(30) Han, W.; Wan, C.-K.; Wu, Y.-D. *J. Chem. Theory Comput.* **2010**, *6*, 3390.
(31) Han, W.; Wan, C.-K.; Jiang, F.; Wu, Y.-D. *J. Chem. Theory Comput.* **2010**, *6*, 3373.
(32) Han, W.; Wan, C.-K.; Wu, Y.-D. *J. Chem. Theory Comput.* **2008**, *4*, 1891.
(33) Jiang, F.; Han, W.; Wu, Y.-D. *J. Phys. Chem. B* **2010**, *114*, 5840.
(34) van der Wel, P. C. A.; de Planque, M. R. R.; Greathouse, D. V.; Koeppe, R. E.; Killian, J. A. *Biophys. J.* **1998**, *74* (2), A304.
(35) de Planque, M. R. R.; Bonev, B. B.; Demmers, J. A. A.; Greathouse, D. V.; Koeppe, R. E. II; Separovic, F.; Watts, A.; Killian, J. A. *Biochemistry* **2003**, *42*, 5341.
(36) Kol, M. A.; van Laak, A. N. C.; Rijkers, D. T. S.; Killian, J. A.; de Kroon, A. I. P. M.; de Kruijff, B. *Biochemistry* **2003**, *42*, 231.
(37) Weiss, T. M.; van der Wel, P. C. A.; Killian, J. A.; Koeppe, R. E. II; Huang, H. W. *Biophys. J.* **2003**, *84*, 379.
(38) Strandberg, E.; Ozdirekcan, S.; Rijkers, D. T. S.; van der Wel, P. C. A.; Koeppe, R. E. II; Liskamp, R. M. J.; Killian, J. A. *Biophys. J.* **2004**, *86*, 3709.
(39) Siegel, D. P.; Cherezov, V.; Greathouse, D. V.; Koeppe, R. E. II; Killian, J. A.; Caffrey, M. *Biophys. J.* **2006**, *90*, 200.
(40) Sparr, E.; Ash, W. L.; Nazarov, P. V.; Rijkers, D. T. S.; Hemminga, M. A.; Tieleman, D. P.; Killian, J. A. *J. Biol. Chem.* **2005**, *280*, 39324.
(41) Im, W.; Brooks, C. L., III *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6771.
(42) Holt, A.; de Almeida, R. F. M.; Nyholm, T. K. M.; Loura, L. M. S.; Daily, A. E.; Staffhorst, R. W. H. M.; Rijkers, D. T. S.; Koeppe, R. E. II; Prieto, M.; Killian, J. A. *Biochemistry* **2008**, *47*, 2638.
(43) Esteban-Martin, S.; Gimenez, D.; Fuertes, G.; Salgado, J. *Biochemistry* **2009**, *48*, 11441.
(44) Vostrikov, V. V.; Daily, A. E.; Greathouse, D. V.; Koeppe, R. E., II *J. Biol. Chem.* **2010**, *285*, 31723.
(45) Thomas, R.; Vostrikov, V. V.; Greathouse, D. V.; Koeppe, R. E., II *Biochemistry* **2009**, *48*, 11883.
(46) Treutlein, H. R.; Lemmon, M. A.; Engelman, D. M.; Brunger, A. T. *Biochemistry* **1992**, *31*, 12726.
(47) Langosch, D.; Brosig, B.; Kolmar, H.; Fritz, H. J. *J. Mol. Biol.* **1996**, *263*, 525.
(48) MacKenzie, K. R.; Prestegard, J. H.; Engelman, D. M. *Science* **1997**, *276*, 131.
(49) Brosig, B.; Langosch, D. *Protein Sci.* **1998**, *7*, 1052.
(50) Fisher, L. E.; Engelman, D. M.; Sturgis, J. N. *J. Mol. Biol.* **1999**, *293*, 639.
(51) Popot, J. L.; Engelman, D. M. *Annu. Rev. Biochem.* **2000**, *69*, 881.
(52) Smith, S. O.; Song, D.; Shekar, S.; Groesbeek, M.; Ziliox, M.; Aimoto, S. *Biochemistry* **2001**, *40*, 6553.
(53) Fisher, L. E.; Engelman, D. M.; Sturgis, J. N. *Biophys. J.* **2003**, *85*, 3097.
(54) Petrache, H. I.; Grossfield, A.; MacKenzie, K. R.; Engelman, D. M.; Woolf, T. B. *J. Mol. Biol.* **2000**, *302*, 727.
(55) Im, W.; Feig, M.; Brooks, C. L. *Biophys. J.* **2003**, *85*, 2900.

(56) Kim, S.; Chamberlain, A. K.; Bowie, J. U. *J. Mol. Biol.* **2003**, *329*, 831.
(57) Braun, R.; Engelman, D. M.; Schulten, K. *Biophys. J.* **2004**, *87*, 754.
(58) Kokubo, H.; Okamoto, Y. *J. Chem. Phys.* **2004**, *120*, 10837.
(59) Hénin, J.; Pohorille, A.; Chipot, C. *J. Am. Chem. Soc.* **2005**, *127*, 8478.
(60) Cuthbertson, J. M.; Bond, P. J.; Sansom, M. S. P. *Biochemistry* **2006**, *45*, 14298.
(61) Efremov, R. G.; Vereshaga, Y. A.; Volynsky, P. E.; Nolde, D. E.; Arseniev, A. S. *J. Comput.-Aided Mol. Des.* **2006**, *20*, 27.
(62) Beevers, A. J.; Kukol, A. *J. Mol. Graphics Model* **2006**, *25*, 226.
(63) Bond, P. J.; Sansom, M. S. P. *J. Am. Chem. Soc.* **2006**, *128*, 2697.
(64) Elofsson, A.; Heijne, G. V. *Annu. Rev. Biochem.* **2007**, *76*, 125.
(65) Metcalf, D. G.; Law, P. B.; DeGrado, W. F. *Proteins: Struct., Funct., Bioinf.* **2007**, *67*, 375.
(66) Bu, L.; Im, W.; Brooks, C. L., III *Biophys. J.* **2007**, *92*, 854.
(67) Psachoulia, E.; Fowler, P. W.; Bond, P. J.; Sansom, M. S. P. *Biochemistry* **2008**, *47*, 10503.
(68) Psachoulia, E.; Marshall., D. P.; Sansom, M. S. P. *Acc. Chem. Res.* **2010**, *43*, 388.
(69) Sengupta, D.; Marrink, S. J. *Phys. Chem. Chem. Phys.* **2010**, *12*, 12987.
(70) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. *J. Comput. Chem.* **1997**, *18*, 1463.
(71) Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. *Comput. Phys. Commun.* **1995**, *91*, 43.
(72) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684.
(73) Feenstra, K. A.; Hess, B.; Berendsen, H. J. C. *J. Comput. Chem.* **1999**, *20*, 786.
(74) Mezei, M.; Beveridge, D. L. *Ann. N.Y. Acad. Sci.* **1986**, *482*, 1.
(75) Faraldo-Gomez, J. D.; Smith, G. R.; Sansom, M. S. P. *Eur. Biophys. J.* **2002**, *31*, 217.
(76) Kandt, C.; Ash, W. L.; Tieleman, D. P. *Methods* **2007**, *41*, 475.
(77) Wolf, M. G.; Hoefling, M.; Aponte-Santamaria, C.; Grubmuller, H.; Groengof, G. *J. Comput. Chem.* **2010**, *31*, 2169.
(78) Li, J.; Zhu, T.; Hawkins, G. D.; Winget, P.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G. *Theor. Chim. Acta* **1999**, *103*, 9.
(79) Abraham, M. H.; Whiting, G. S.; Fuchs, R.; Chambers, E. J. *J. Chem. Soc., Perkin Trans. 2* **1990**, 291.
(80) Ozdirekcan, S.; Rijkers, D. T. S.; Liskamp, R. M. J.; Killian, J. A. *Biochemistry* **2005**, *44*, 1004.
(81) Vostrikov, V. V.; Hall, B. A.; Greathouse, D. V.; Koeppe, R. E., II; Sansom, M. S. P. *J. Am. Chem. Soc.* **2010**, *132*, 5803.
(82) MacCallum, J. L.; Bennett, W. F. D.; Tieleman, D. P. *Biophys. J.* **2008**, *94*, 3393.
(83) MacCallum, J. L.; Bennett, W. F. D.; Tieleman, D. P. *J. Genet. Physiol.* **2007**, *129* (5), 371.
(84) Kucerka, N.; J. Nieh, P. M.-P.; Katsaras, J *Eur. Phys. J. E: Soft Matter Biol. Phys.* **2007**, *23*, 247.
(85) Shirts, M. R.; Pitera, J. W.; Swope, W. C.; Pande, V. S. *J. Chem. Phys.* **2003**, *119* (11), 5740.
(86) MacCallum, J. L.; Tieleman, D. P. *J. Comput. Chem.* **2003**, *24*, 1930.
(87) Radzicka, A.; Wolfenden, R. *Biochemistry* **1988**, *27*, 1664.
(88) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. *J. Phys. Chem. B* **2001**, *105*, 6474.
(89) Nymeyer, H.; Zhou, H.-X. *Biophys. J.* **2008**, *94*, 1185.
(90) Yesylevskyy, S. O.; Schafer, L. V.; Sengupta, D.; Marrink, S. J. *PLoS Comput. Biol.* **2010**, *6*, e1000810.
(91) Wang, J.; Kim, S.; Kovacs, F.; Cross, T. A. *Protein Sci.* **2001**, *10*, 2241.
(92) Inbaraj, J. J.; Laryukhin, M.; Lorigan, G. A. *J. Am. Chem. Soc.* **2007**, *129*, 7710.
(93) Kandasamy, S. K.; Larson, R. G. *Biophys. J.* **2006**, *90*, 2326.
(94) Andrea, H.; Killian, J. A. *Eur. Biophys. J.* **2010**, *39*, 609.
(95) de Planque, M. R. R.; Goormaghtigh, E.; Greathouse, D. V.; Koeppe, R. E.; Kruijtzer, J. A. W.; Liskamp, R. M. J.; de Kruijff, B.; Killian, J. A. *Biochemistry* **2001**, *40*, 5000.
(96) Holt, A.; Koehorst, R. B.; Rutters-Meijneke, T.; Gelb, M. H.; Rijkers, D. T.; Hemminga, M. A.; Killian, J. A. *Biophys. J.* **2009**, *97*, 2258.
(97) Esteban-Martin, S.; Salgado, J. *Biophys. J.* **2007**, *93*, 4278.
(98) Strandberg, E.; Esteban-Martin, S.; Salgado, J.; Ulrich, A. S. *Biophys. J.* **2009**, *96*, 3223.
(99) Im, W.; Lee, J.; Kim, T.; Rui, H. *J. Comput. Chem.* **2009**, *30*, 1662.
(100) Kim, T.; Im, W. *Biophys. J.* **2010**, *99*, 175.
(101) Bond, P. J.; Holyoake, J.; Ivetac, A.; Khalid, S.; Sansom, M. S. P. *J. Struct. Biol.* **2007**, *157*, 593.
(102) Monticelli, L.; Tieleman, D. P.; Fuchs, P. F. J. *Biophys. J.* **2010**, *99*, 1455.
(103) Holt, A.; Rougier, L.; Reat, V.; Jolibois, F.; Saurel, O.; Czaplicki, J.; Killian, J. A.; Milon, A. *Biophys. J.* **2010**, *98*, 1864.
(104) Sankararamakrishnan, R.; Vishveshwara, S. *Int. J. Pept. Protein Res.* **1992**, *39*, 356.
(105) Cordes, F. S.; Bright, J. N.; Sansom, M. S. P. *J. Mol. Biol.* **2002**, *323*, 951.
(106) Lear, J. D.; Stouffer, A. L.; Gratkowski, H.; Nanda, V.; DeGrado, W. F. *Biophys. J.* **2004**, *87*, 3421.
(107) Senes, A.; Engel, D. E.; DeGrado, W. F. *Curr. Opin. Struct. Biol.* **2004**, *14*, 465.
(108) Popot, J. L.; Engelman, D. M. *Biochemistry* **1990**, *29*, 4031.
(109) Russ, W. P.; Engelman, D. M. *J. Mol. Biol.* **2000**, *296*, 911.
(110) Fleming, K. G. *J. Mol. Biol.* **2002**, *323*, 563.
(111) Hong, H.; Blois, T. M.; Cao, Z.; Bowie, J. U. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 19802.
(112) Senes, A.; Gerstein, M.; Engelman, D. M. *J. Mol. Biol.* **2000**, *296*, 921.
(113) Lewis, B. A.; Engelman, D. M. *J. Mol. Biol.* **1983**, *166*, 211.