LUND UNIVERSITY

Evaluating Force Fields for the Computational Prediction of Ionized Arginine and Lysine Side-Chains Partitioning into Lipid Bilayers and Octanol

Sun, Delin; Forsman, Jan; Woodward, Clifford E.

Published in:
Journal of Chemical Theory and Computation

DOI:
10.1021/ct501063a

2015

Link to publication

Citation for published version (APA):
Sun, D., Forsman, J., & Woodward, C. E. (2015). Evaluating Force Fields for the Computational Prediction of Ionized Arginine and Lysine Side-Chains Partitioning into Lipid Bilayers and Octanol. *Journal of Chemical Theory and Computation*, *11*(4), 1775-1791. https://doi.org/10.1021/ct501063a

Total number of authors:
3

General rights
Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.
• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: https://creativecommons.org/licenses/

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Evaluating Force Fields for the Computational Prediction of Ionized Arginine and Lysine Side-Chains Partitioning into Lipid Bilayers and Octanol

Delin Sun,[†] Jan Forsman,[‡] and Clifford E. Woodward*,[†]

[†]School of Physical, Environmental and Mathematical Sciences, University of New South Wales, Canberra ACT 2600, Australia

[‡]Theoretical Chemistry, Chemical Centre, Lund University, P.O. Box 124, S-221 00 Lund, Sweden

ABSTRACT

Abundant peptides and proteins containing arginine (Arg) and lysine (Lys) amino acids can apparently permeate cell membranes with ease. However, the mechanisms by which these peptides and proteins succeed in traversing the free energy barrier imposed by cell membranes remain largely unestablished. Precise thermodynamic studies (both theoretical and experimental) on the interactions of Arg and Lys residues with model lipid bilayers can provide valuable clues to the efficacy of these cationic peptides and proteins. We have carried out molecular dynamics simulations to calculate the interactions of ionized Arg and Lys side-chains with the zwitterionic 1, 2-dipalmitoyl-sn-glycero-3-phosphocholine (DPPC) lipid bilayer for ten widely-used lipid/protein force fields. These are: CHARMM36/CHARMM36; SLIPID/AMBER99SB-ILDN; OPLS-AA/OPLS-AA; Berger/OPLS-AA; Berger/GROMOS87; Berger/GROMOS53A6; GROMOS53A6/GROMOS53A6; non-polarizable MARTINI; polarizable MARTINI, and BMW MARTINI. We performed umbrella sampling simulations to obtain the potential of mean force for Arg and Lys side-chains partitioning from water to the bilayer interior. We found significant differences between the force fields, both for the interactions between side-chains and bilayer surface, as well as the free energy cost for placing the side-chain at the center of the bilayer. These simulation results were compared with the Wimley-White interfacial scale. We also calculated the free energy cost for transferring ionized Arg and Lys side-chains from water to both dry and wet octanol. Our simulations reveal rapid diffusion of water molecules into octanol whereby the equilibrium mole fraction of water in the wet octanol phase was ~25%. Surprisingly, our free energy calculations found that the high water content in wet octanol lowered the water-to-octanol partitioning free energies for cationic residues by only 0.6-0.7 kcal/mol.

1. INTRODUCTION

Understanding the structure-function relationship of cell membrane-associated proteins is critical to elucidating the mechanisms of a myriad of cellular processes. This is key to identifying the origins of many diseases and to the development of novel therapeutic agents. Thus far, clear physical models for the functioning of this important class of proteins remain in their infancy, despite a large amount of experimental and theoretical work in the area. This may reflect variability induced by evolutionary factors, though it is tempting to speculate that common motifs in these proteins indicate that they work through broadly similar or at least broadly classifiable mechanisms. Notwithstanding the reported attempts to establish such a set of common actions, the science remains controversial. A particular motif that has created a great deal of debate over the last decade features a large number of cationic arginine (Arg) and/or lysine (Lys) residues. Examples can be found in voltage-gated potassium ion channels as well as antimicrobial and cell-penetrating peptides. It has been suggested that the activity of these proteins and peptides is due to their enhanced ability to pass through cell membranes.[1-3] However, this is seriously challenged by the accepted view that cell membranes are essentially impermeable to charged molecules. Indeed, the mechanism via which proteins and peptides, rich in Arg and Lys residues, succeed in traversing the formidable free energy barrier of hydrophobic cell membranes remains largely unestablished. Precise thermodynamic studies (both theoretical and experimental) on the interactions of Arg and Lys residues with lipid bilayers can provide valuable clues to the efficacy of these membrane-associated cationic proteins.

There have been a number of experimental and computational studies aimed at developing "hydrophobicity scales" which rank the water-to-membrane partitioning free energies of the natural amino acids.[4-10] Wimley and White have determined two widely-used hydrophobicity scales.[9, 10] The first is the interfacial scale, obtained by measuring the partitioning of pentapeptides Ace-WLXLL (where X is the residue of interest) between water and the zwitterionic 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) lipid membrane. The second, so-called octanol scale was obtained

from the partitioning of the same pentapeptide species between water and octanol. The octanol scale is often used to estimate the fee energy cost for Arg- and Lys-rich antimicrobial peptides permeating through a lipid bilayer.[11] However, equivalence between bulk octanol and the interior of a lipid membrane is unlikely, especially given the high water content of the wet octanol phase and the inhomogeneous nature of lipid bilayer. In fact, the octanol scale predicts very small free energy costs for transferring ionized Arg (1.86 kcal/mol) and Lys (2.82 kcal/mol) from water to octanol[12] and it has been suggested that the octanol scale instead reflects the thermodynamics of amino acids at the bilayer/water interface instead of in the bilayer interior.[6, 13]Surprisingly, other biological hydrophobicity scales, proposed by Hessa *et al*.[4] and Moon and Fleming[7] also obtained very small free energy costs for inserting ionized Arg and Lys side-chains into the lipid bilayer interior: ~2.5 kcal/mol (from Hessa *et al*.) and ~3.7 kcal/mol (from Moon and Fleming) for the Arg side-chain and ~2.6 kcal/mol (Hessa *et al.*) and ~5.4 kcal/mol (Moon and Fleming) for the Lys side-chain. These seem consistent with the Wimley-White octanol scale for predicting the partitioning of ionized Arg and Lys into the interior of lipid bilayers. However, the apparently small free energy barriers that the experimentally determined hydrophobicity scales predict for bilayer insertion of ionized Arg and Lys side-chains are in stark contrast to the results reported by molecular dynamics (MD) simulations. MD simulations, using a variety of lipid and peptide models and force fields, have predicted that the water-to-bilayer partitioning free energy for Arg and Lys side-chains is in the range of 14-28 kcal/mol.[6, 14-16]

The physical origin of the discrepancy between experimental and computational measures for partitioning of these side-chains remains unresolved. One proposal is that the different molecular environments of the Arg and Lys residues in experiments and simulations are responsible. In this context, we note that molecular simulation studies by Johansson and Lindahl[17] found that the presence of trans-membrane proteins can help retain the water of hydration for Arg in membrane interiors. They estimated that the free energy cost for burying one Arg side-chain in the vicinity of

the SecY translocon membrane protein could be as low as ~3-5 kcal/mol. Gumbart *et al*.[18] designed a thermodynamic cycle and used a free energy perturbation method to obtain a value of only 6.5 kcal/mol for the free energy of Arg partitioning from the SecY translocon protein to the lipid bilayer center. They also observed that Arg in the bilayer center was hydrated. Gumbart and Roux[19] mimicked the experiments leading to the hydrophobicity scale of Moon and Fleming.[7] Here, a host-guest strategy was employed, based on an outer membrane phospholipase A (OmpLA) protein inserted into a thin 1, 2-dilauroyl-sn-glycero-3-phosphocholine (DLPC) lipid bilayer. The calculated free energy cost for burying one Arg at position 210 of the OmpLA protein in the DLPC bilayer center was only 2.7 kcal/mol. The Arg residue was found to "snorkel" towards the bilayer/water interface. A similar phenomenon was also reported in simulations by Fleming *et al*.[20] Thus, computer simulations can potentially provide an explanation for the apparent thermodynamic stability of cationic amino acids in lipid bilayers, provided the correct molecular environments are modelled. The lipid bilayer thickness and the availability of hydrating water molecules seem to be two critical factors for the lowering of the water-to-bilayer partitioning free energy. The low partitioning free energy for cationic side-chains predicted by the Wimley-White octanol scale, has been often attributed to the high water content in wet octanol.[13] However, to the best of our knowledge, a systematic simulation study on the water-to-octanol partitioning of ionized Arg and Lys amino acids has not been reported.

A major concern in any MD simulation study is the accuracy of the employed force field. The most widely used protein force fields include the all-atom OPLS-AA,[21] CHARMM,[22] and AMBER[23] force fields and the united-atom GROMOS[24] force field. All of these classical force fields are able to reproduce experimentally determined hydration free energies of amino acids with good accuracy.[25, 26] Among the most frequently used models for lipids are the all-atom CHARMM[27] and the united-atom Berger[28] models, which are able to reproduce some key structural properties of lipid bilayers. However, the validity of simply combining these protein and lipid force field models in order to

describe the interaction between peptides and lipid bilayers remains an open question. Such an approach has often been used in model studies of such systems. For example, MacCallum *et al*.[6] performed comprehensive simulations using a Berger/OPLS-AA combination (for lipid and peptide respectively) to derive the free energy profiles of 20 different amino acid side-chains permeating a zwitterionic 1, 2-dioleoyl-sn-glycero-3-phosphocholine (DOPC) lipid bilayer. From this study, a theoretical interfacial scale for residues was determined. While this scale can be compared with the Wimley-White interfacial scale, the caveats described earlier concerning such comparisons will apply. Indeed, for the cationic residues Arg and Lys, there is a large discrepancy between the two scales. The simulated interfacial scales for ionized Arg is -5.1 kcal/mol and -4.4 kcal/mol for Lys, while the Wimley-White values are 0.81 kcal/mol for Arg and 0.99 kcal/mol for Lys. While the different environments of the residues in the simulations and in the experiments may account for some (or even most) of the free energy disparity, it is not clear how sensitive the simulation results are to the force field used. In another simulation study using the all-atom CHARMM force field for both peptide and lipid, Li *et al*.[15] found that the interaction between ionized Arg and Lys side-chains with the zwitterionic 1, 2-dipalmitoyl-sn-glycero-3-phosphocholine (DPPC) lipid bilayer were almost negligible. Hence, it is important to establish the extent to which the theoretical description of peptide-lipid interactions is affected by the force field employed.

The aim of the work presented here is two-fold. Firstly, we use thermodynamic integration to calculate the water-to-octanol partitioning free energies for ionized Arg and Lys side-chains. In this case, we confined the simulations to the OPLS-AA all-atom force field, which was developed primarily for organic liquids. Secondly, we study the effect of lipid/protein force field models on the interactions of ionized Arg and Lys side-chains with the zwitterionic DPPC lipid bilayer. Here, we compared ten widely-used lipid/protein models, which are listed below and described in more detail later in this article. Three are all-atom models:

- CHARMM36/CHARMM36

- SLIPID/AMBER99SB-ILDN
- OPLS-AA/OPLS-AA.

Four are united-atom (or mixed united-atom/all-atom) models:

- Berger/OPLS-AA
- Berger/GROMOS87
- Berger/GROMOS53A6
- GROMOS53A6/GROMOS53A6.

Finally, we also investigated three coarse-grained models:

- non-polarizable MARTINI
- polarizable MARTINI
- big multipole water (BMW) MARTINI

With these force field combinations, we employ standard umbrella sampling simulations to compute the free energies for ionized Arg and Lys side-chains translocating through a DPPC lipid bilayer. This work will provide useful information for future molecular simulation studies of interactions between cationic peptides and lipid membranes. As well it will give some context to already published work, which have used these particular (or similar) force field combinations.

2. SIMULATION METHODS

2.1 Models. The lipid/protein models that we tested are listed in Table 1. The united-atom Berger lipid model[28] was initially developed to simulate the DPPC lipid bilayer. In this model, bonded interactions within the lipid were described with the GROMOS force field and Ryckaert-Bellemans dihedral potentials.[29] Lennard-Jones (LJ) parameters were taken from the OPLS-UA force field[30] and the geometric combination rules were used for mixed LJ interactions. Partial charges of the Berger

DPPC lipid head groups were obtained from *ab* initio quantum mechanical calculations by Chiu *et al*.[31] In published simulation studies of mixed membrane/protein systems, the Berger lipid model was mostly used in combination with the all-atom OPLS-AA or the united-atom GROMOS force field for the protein.[6, 32-34] In this study, we tested the combinations Berger/OPLS-AA, Berger/GROMOS87 and Berger/GROMOS53A6. The tested united-atom GROMOS53A6 DPPC lipid was developed by Kukol,[35] who used the GROMOS53A6 force field[24] to model bonded and non-bonded LJ interactions with partial charges also taken from Chiu *et al*.[31] We tested this lipid model in combination with the GROMOS53A6 protein model.[24] The CHARMM all-atom force field provides compatible parameters for membrane/protein simulations. In this work, we tested the CHARMM36 force field developed by MacKerell and co-workers.[27] The all-atom SLIPID model was developed by Lyubartsev and co-workers.[36] It has been suggested that this lipid model should be combined with the all-atom AMBER force field for proteins.[37] Here, we tested the SLIPID DPPC lipid model with the AMBER99SB-ILDN protein model.[38] We also tested an all-atom OPLS-AA DPPC model developed by Maciejewski *et al.*[39] in combination with the OPLS-AA protein model.[21] For these simulations, the method used to model Arg and Lys side-chains was similar to that used by MacCallum *et al.*:[6] the side-chains were truncated at the β-carbon; the α-carbon was replaced by a hydrogen atom and the partial charge for the β-carbon was adjusted to make the side-chains carry a net +1 charge. These modifications result in the propylguanidinium ion being the analogue for Arg and the butylammonium ion for Lys.

In addition to the lipid/protein force field, the water model we employ is also expected to have an effect on the calculated thermodynamics. The choice of water models are listed in Table 1 and are in accord with the original lipid model development papers.[27, 28, 35, 36, 39] For united-atom lipid bilayer systems (Berger and GROMOS53A6), the simple point charge (SPC) water model[40] was used; for the all-atom lipid bilayer systems (CHARMM36, SLIPID, OPLS-AA), the TIPS3P[41] or TIP3P[42]

water model was used. For our studies on water-to-octanol partitioning, we used the OPLS-AA force field developed by Caleman *et al*[43] to model octanol. In this case, the TIP3P[42] water model was used.

MARTINI is a coarse-grained force field developed by Marrink and co-workers for large-scale lipid bilayer-protein simulations.[44-46] Of particular note is the MARTINI protein model,[44, 46] which was parameterized to reproduce the atomistic simulations of MacCallum *et al.* for free energy profiles of amino acid side-chains permeating through a zwitterionic DOPC lipid bilayer.[6] The simulations of MacCallum *et al.* predicted strong interactions of Arg (-5.1 kcal/mol) and Lys (-4.4 kcal/mol) side-chains with the DOPC lipid bilayer/water interface.[6] In the standard (non-polarizable) MARTINI model, a cluster of four waters is treated as one large neutral bead. The Arg and Lys side-chains were represented as two beads. The standard MARTINI model predicts relatively weak interactions between Arg and Lys side-chains with the DOPC bilayer surface.[44, 46] This weak interaction was attributed to the relatively large dielectric constant (ε=15) used in the standard MARTINI mode1.[46] The development of the polarizable water model (polarizable MARTINI) by Yesylevskyy *et al.*[47] and the big multipole water model (BMW MARTINI) by Wu *et al.*[48] ameliorates this effect, by giving electrostatic interactions to the water molecules and a reduced dielectric constant. It should be noted, however, that the Wimley-White interfacial scale predicts quite weak interactions between ionized Arg and Lys with zwitterionic lipid bilayers.[10] Li *et al.*'s all-atom simulations using the CHARMM force field also showed that interactions of ionized Arg and Lys with a zwitterionic bilayer are weak.[15]

Table 1. Lipid, Protein and Water Models Used in the Lipid Bilayer/Amino Acid Side-Chain Systems.

Lipid models	Protein models	Water models
Berger	OPLS-AA	SPC
Berger	GROMOS87	SPC
Berger	GROMOS53A6	SPC

GROMOS53A6	GROMOS53A6	SPC
CHARMM36	CHARMM36	TIPS3P
SLIPID	AMBER99SB-ILDN	TIP3P
OPLS-AA	OPLS-AA	TIP3P
Non-pol MARTINI[a]	Non-pol MARTINI	Non-pol MARTINI
Pol[b] MARTINI	Pol MARTINI	Pol MARTINI
BMW[c] MARTINI	BMW MARTINI	BMW MARTINI

[a] Non-polarizable MARTINI (version 2.2); [b] Polarizable; [c] big multipole water

2.2 Simulation Systems.

2.2.1 Arg and Lys residues with bilayers. Our lipid bilayer simulation systems using the united-atom lipid models (Berger and GROMOS53A6) contained 64 DPPC lipids and 3800 water molecules. Simulation systems using the all-atom lipid models (CHARMM, SLIPID and OPLS-AA) contained 64 DPPC lipids and 3200 water molecules. The coarse-grained simulations used 128 DPPC lipids and 3000 (coarse-grained) waters. The aim of these simulations was to calculate the water-to-bilayer partitioning free energy of the cationic side-chain (Arg or Lys). In all cases, a chloride ion was added to neutralize the system, but there was no additional salt. There is continuing debate over the accuracy of force fields with respect to interactions between electrolytes and the lipid bilayer.[49] In our comparison of the different force fields, our focus will be on the differences that follow from the interactions between amino acid side-chains and the lipid bilayer within the appropriate water model. Including additional salt into our simulations would only further complicate our comparisons. The size of the simulated systems was held constant within the class of force field. A smaller lipid area was used for the atomistic models, and the coarse-grained models used bilayers twice as large. Hu *et al.* [50] explored the effect of lipid bilayer area on the calculated free energy cost for Arg side-chain permeating though a 1,2-dimyristoyl-sn-glycero-3-phosphocholine (DMPC) lipid bilayer. They found that using a larger membrane area (288 lipids)

reduced the free energy cost for Arg permeating through the bilayer by as much as 6.5 kcal/mol.[50] This was shown to be largely due to electrostatic contributions from the neutralizing counterion, reflecting an implicit ionic strength effect. Our consistent use of a bilayer patch of 64 DPPC lipids (for atomistic models) should allow us to investigate the influence of force field parameters on the calculated water-to-bilayer partitioning free energies. Furthermore, we are also able to compare our simulation results with the published studies of similar systems, which generally used 64 [5, 6]or 48 lipids.[15, 51] On the other hand, comparisons between the coarse-grained and atomistic model results may be subject to the size-dependencies discussed above. In all cases, the systems were firstly equilibrated for 100 ns at the temperature of 323 K, before free energy evaluations were carried out using umbrella sampling.

2.2.2 Water to Octanol Transfer of Arg and Lys. Water molecules in wet octanol are believed to play an important role in lowering the water-to-octanol partitioning free energies for charged residues.[13] In order to calculate partitioning free energies for Arg and Lys side-chains from water into wet octanol, the saturated water content in the octanol has to be determined. To this end, we modeled the octanol water interface, wherein a dry octanol (bulk) phase was initially placed adjacent to a bulk water phase. This biphasic simulation used a periodic system consisting of a "slab" of 512 octanol molecules, sandwiched between 4500 water molecules. The system was equilibrated for 200 ns at a temperature of 323 K and pressure of 1 atm. During the equilibration, we found rapid diffusion of water molecules into the octanol, which eventually reached an equilibrium value. Thus, the saturated water content in octanol could be determined. The water-to-(wet)octanol partitioning free energies for the Arg and Lys side-chains were calculated using a thermodynamic integration method using two isolated bulk phase simulations. In one simulation, the box contained a single cationic residue (Arg or Lys) and 2500 TIP3P water molecules, while the other simulation box consisted of a single amino acid side-chain, 95 TIP3P water molecules and 300 octanol molecules. This number of water molecules in the wet octanol was determined from the 200 ns equilibration

simulations of water-octanol-water system. To quantify the role of water molecules in wet octanol, we also constructed a cubic simulation box consisting of one side-chain and 300 octanol molecules (dry octanol). These systems were generally equilibrated for 5 ns at a temperature of 323 K and pressure of 1 atm before the solvation free energies of Arg and Lys side-chains in water and wet/dry octanol were calculated via thermodynamic integration.

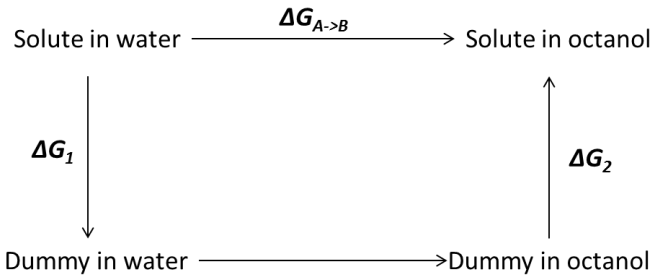
2.3 Free Energy Calculations. Steered MD[52] and umbrella sampling[53] simulations were carried out to obtain the water-to-bilayer potentials of mean force (PMF) for ionized Arg and Lys side-chains. In steered MD simulations, an Arg or Lys side-chain together with one counterion was inserted into the water phase and a harmonic potential with a force constant of 1000 kJ/mol/nm^2 was applied between the center of mass of the side-chain and the center of mass of the lipid bilayer. The side-chain was then pulled along the z-axis (perpendicular to the bilayer/water interface) at a rate of 0.01 nm/ps to generate a number of configurations. The generated configurations were selected every 0.1 nm along the z-axis and used as starting points for subsequent umbrella sampling simulations. In this way, 31 umbrella sampling windows were used to construct the free energy profiles for Arg and Lys side-chains partitioning into lipid bilayer. In each window, we carried out either 50 ns (for all-atom and united-atom simulations) or 100 ns (for coarse-grained simulations). During umbrella sampling simulations, a biased harmonic potential with a force constant of 3000 kJ/mol/nm^2 was used to confine the side-chain within the sampling window. The 0.1 nm spacing, combined with the force constant of the biasing potential gave sufficient overlap between adjacent windows along the reaction coordinate (Figure S1). The free energy profiles and error estimates were determined using the weighted histogram analysis method (WHAM).[54] Generally, the first 20 ns within each window was discarded to allow for appropriate equilibration at the new solute position, and the free energy profile was determined from the remaining 30 ns. For all the atomistic force fields, the potentials of mean force, when calculated over 10 ns windows within this final 30 ns period, showed no

systematic drift (Figure S2), thus indicating they had come to equilibrium. The same is true for the coarse-grained force fields, except in this case the sampling windows were twice as long (Figure S2).

The water-to-octanol partitioning free energies for ionized Arg and Lys side-chains were determined using a thermodynamic integration approach. In this approach, the interaction between the inserted residue and the surrounding fluid is treated as a function of a coupling parameter λ. The free energy difference between states A and state B, characterized by λ_A and λ_B respectively, is given by

$$\Delta G_{A\to B} = \int_{\lambda A}^{\lambda B} \langle {}^{\partial H(\lambda)}/_{\partial \lambda} \rangle_{\lambda'} \, d\lambda' \qquad (1)$$

In our calculations, not the absolute solvation free energies (vacuum as a reference state) of Arg and Lys side-chains in water and octanol but only the water-to-octanol partitioning free energies were determined. The thermodynamic cycle used in our simulations is illustrated in Scheme 1.



Scheme 1. The thermodynamic cycle used in the water-to-octanol partitioning free energy calculations.

The free energy difference, calculated in this way, for water-to-octanol partitioning of ionized Arg and Lys side-chains does not include the contribution due to the electrostatic potential of the water/octanol solvent interface. It should be noted that in actual partitioning experiments this potential will play no role, as electro-neutral combinations of ions will be transferred across the interface. Furthermore, in the presence of a finite concentration of ionic species, the interfacial

potential is properly determined by the excess chemical potentials of the ionic species and has a value that maintains electro-neutrality in the bulk regions of both solvents (via a so-called Donnan equilibrium); it is not the bare solvent interfacial potential. Nevertheless, in the results below we will include a "correction term" due to the bare solvent potential to obtain the transfer free energies for single ions at infinite dilution the so-called real transfer free energies. A similar quantity was defined and used in work by Deng and Roux.[55] The bare solvent interfacial potential was determined using an equation reported by Sachs *et al*.[56] and is obtained as

$$\emptyset(z) = -\frac{1}{\epsilon_0}\sum_i q_i \left[\int_0^z (z-u)\rho_i(u)\,du - \frac{z}{L}\int_0^L (L-u)\rho_i(u)du\right] \quad (2)$$

where Φ is the potential, z is the direction perpendicular to the water/octanol interface, ε_0 is the permittivity of free space, q_i is the (partial) charge of atom i, ρ_i is the density of atoms of type i, and L is the simulation box size in the z direction. This equation enforces periodic boundary conditions such that Φ (0) = Φ (L) (by addition of a linear term), which is appropriate for our periodic octanol slab simulations. This expression effectively averages the potential change across the two water/octanol interfaces and thus leads to essentially equivalent results to that obtained by symmetrizing the charge density. The water-to-octanol partitioning free energy for ionized Arg and Lys is obtained as,

$$\Delta G_{A\to B} = \Delta G_1 + \Delta G_2 + \Delta G_{corr} \quad (3)$$

where the free energy components ΔG_1 and ΔG_2 are defined by the thermodynamic cycle described in Scheme 1, and ΔG_{corr} denotes the component due to the interfacial potential. The simulations were carried out at a constant isotropic pressure of 1 atm and a temperature of 323 K. Integration of equation of motion was performed using the stochastic dynamics. A series of 51 windows with the λ equally spaced between 0 and 1 were constructed for both water and wet/dry octanol phases. Each window was simulated for 800 ps and the first 100 ps simulation was discarded for equilibration. A

soft core potential with α=1, σ=0.3 nm and λ-power of 1 was used in order to avoid singularities. The two-step procedure for decoupling LJ and electrostatic interactions independently was employed.

2.4 Other Simulation Parameters. All simulations were performed using the GROMACS 4.5.5 package.[57] The Berger, GROMOS53A6, SLIPID, OPLS-AA and MARTINI lipid models were all implemented in the GROMACS package in accordance with the lipid model development papers.[28, 35, 36, 39, 45-48] The CHARMM36 lipid models were originally developed within the CHARMM package.[27] Using the switching function as in the CHARMM package to treat the LJ interactions in the DPPC lipid bilayer, resulted in an area per lipid (APL) of 0.629 nm^2 at temperature of 323 K,[27] in reasonable agreement with experiment. However, using the native switching functions in the NAMD and GROMACS packages to treat LJ interactions gave a smaller APL (0.591 nm^2 in the NAMD package[27] and 0.586 nm^2 in the GROMACS package[58]). Since the APL will likely have an impact on the calculated free energies for side chain permeation through the lipid bilayer, we used another cut-off strategy to treat the non-bonded interactions in this work. Here, we used cut-offs of 0.9 nm for the LJ interactions and 0.9 nm for electrostatic interactions, which were evaluated using the particle-mesh Ewald (PME) method.[59] As well, a 0.9 nm neighbour list was updated every 5 time steps. This treatment of non-bonded interactions was able to accurately predict the APL of the DPPC bilayer when using the CHARMM36 force field in GROMACS package. It should be noted that reducing the cut-off distance to treat LJ interactions, when implementing the CHARMM force field, may cause artefacts in systems containing proteins (or residues). Piana *et al* studied the effect of cut-off distance (using CHARMM force field) on the folding thermodynamics and structural properties of small proteins.[60] They found that a cut-off distance of 0.9 nm has little effect on the calculated thermodynamic and structural properties of proteins.[60] Hence, for our simulations of amino acid side-chain interacting with lipid bilayer, a cut-off distance of 0.9 nm is expected to have little effect as well.

All lipid bilayer simulations were performed at the temperature of 323K and the molecular species were independently coupled to the Nosé-Hoover thermostat[61, 62] with a coupling time constant of 0.5 ps. The system's volume was allowed to fluctuate according to the semi-isotropic pressure coupling method, i.e., the lateral and perpendicular pressures (both 1atm) were independently coupled to the Parrinello-Rahman barostat[63] with a coupling time constant of 2 ps and compressibility of 4.5×10^{-5} bar^{-1}. In thermodynamic integration calculations, the system's pressure was controlled using an isotropic coupling method. The system's temperature was also kept constant at 323 K using the velocity-rescaling thermostat.[64] The PME method was used for electrostatic calculations. Periodic boundary conditions were employed in all our simulations. All bonds lengths in molecules apart from water were constrained using the LINCS algorithm.[65] Water molecule bonds were constrained using the SETTLE algorithm.[66] The simulation time step was 2 fs.

3. RESULTS

3.1 Lipid Bilayer Structure. The lipid bilayer fluidity and thickness can markedly influence the partitioning free energy profiles for the charged amino acid side-chains.[51] Hence, it is instructive to examine the simulated structural properties of the DPPC bilayer, as predicted by the various lipid force fields. We begin by considering the equilibrium values of the APL and density distribution profiles of the various lipid components. From the latter we can obtain the bilayer thickness. The accepted experimental value for the APL of the DPPC bilayer at 323 K is in the range of 0.63-0.64 nm^2.[67, 68] Figure 1 shows the simulated APL of the DPPC bilayer as a function of simulation time for the various force field models. It can be seen that all the lipid models we tested predict an APL in reasonable agreement with experiments. We note particularly the good result we obtain with the non-bonded cut-off strategy used for the CHARMM36 implementation in the GROMACS package. It is interesting that the fluctuations in APL are significantly larger in the atomistic (all-atom and united-atom) simulations, compared with the coarse-grained simulations. This is to be expected if we

interpret the time-step in the coarse-grained models as effectively larger than in atomistic simulations. In the original development papers, it was suggested the effective time scale in coarse-grained simulations to be 2-10 times larger than in atomistic simulations.[45, 46] The fluctuations in the atomistic models occur at high frequency, which when averaged over a larger effective time-step would lead to more damped profiles, as is observed in the coarse-grained results. We also note that the atomistic simulations utilize molecular models for water (SPC, TIP3P or TIPS3P) while their coarse-grained counterparts map several water molecules onto a much larger effective aqueous species. The smaller fluctuations of the coarse-grained bilayers imply a reduced planar compressibility. This is due to the relatively rigid effective lipid (resulting from the four atoms to one mapping of the MARTINI model). There is a general tendency for coarse-grained potentials to supress fluctuations. By their nature these potentials are essentially free energies, as they combine enthalpic and entropic contributions that usually compensate one another when fluctuations occur. For example, clusters of lipids may form at short wavelengths decreasing both energy and entropy for the molecules within the cluster, but also creating more free volume at length-scales larger than the cluster sizes. In addition, fluctuations in the lipid area will involve some penetration of water into the head-group region of the bilayer interface and this will be suppressed in coarse-grained models, where the free energy for insertion of the effective water molecules is expected to be higher, due to their larger radius and their inability to increase their hydrogen bonding in hydrophobic environments. Suppressed fluctuations will also manifest themselves in the inability of the coarse-grained models to form significant water defects in the presence of imbedded charged Arg and Lys side-chains (see below).

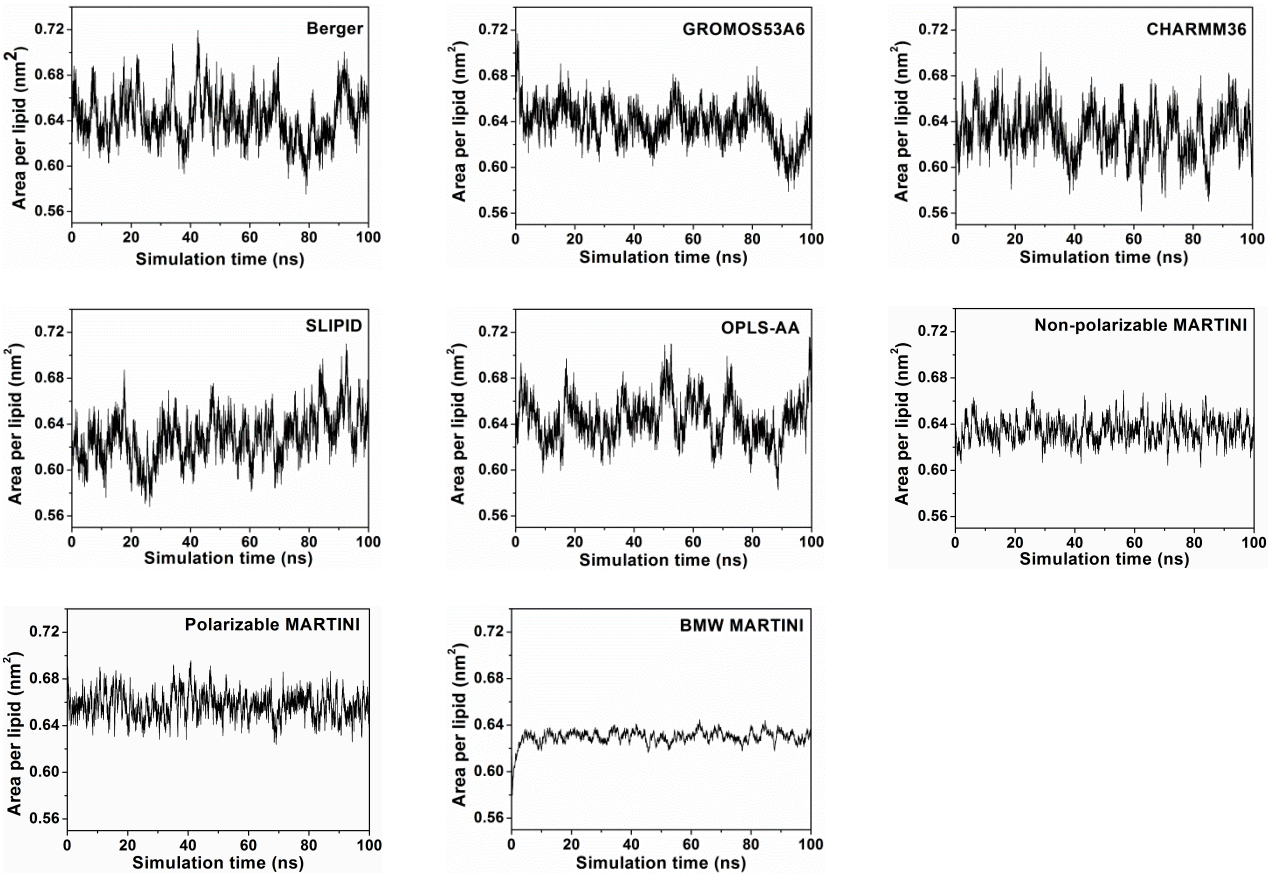


Figure 1. Time evolution of the area per lipid of DPPC lipid bilayer predicted by different lipid models.

Figure 2 shows the simulated density profiles for various sites of the lipid molecule: phosphate, glycerol, and the two terminal alkyl chain ends of the DPPC lipid bilayer. The lipid bilayer thickness was determined by doubling the distance from the lipid phosphate peak to the bilayer centre. This measure corresponds to the experimentally determined head-to-head distance (D_{HH}). The lipid density profiles of the atomistic models are qualitatively alike and their predicted bilayer thicknesses are also quite similar. The generally accepted (experimental) value of D_{HH} for the DPPC bilayer at 323K is in the range 3.78-3.86 nm.[69, 70] The all-atom CHARMM36 (D_{HH}=3.90 nm) and SLIPID (D_{HH}=3.74 nm) models predict average thicknesses in good agreement with experiments. On the other hand, the all-atom OPLS-AA (D_{HH}=3.58 nm), the united-atom Berger (D_{HH}=3.56 nm) and GROMOS53A6 (D_{HH}=3.62 nm) lipid models underestimate the experimental value somewhat. However, the width of the phosphate peak, see Figure 2, would suggest that these discrepancies are

not significant. The profiles predicted by the coarse-grained models are qualitatively different from the atomistic simulations, due mainly to the mapping protocol used to replace a many-bead atomistic model with a smaller number of molecular sites. MARTINI coarse-grained models also tend to overestimate the D_{HH} (4.0 nm for the non-polarized MARTINI, 4.1 nm for polarized MARTINI and 4.16 nm for BMW MARTINI), but given the approximate nature of the modelling, the agreement with experiment is reasonably good.

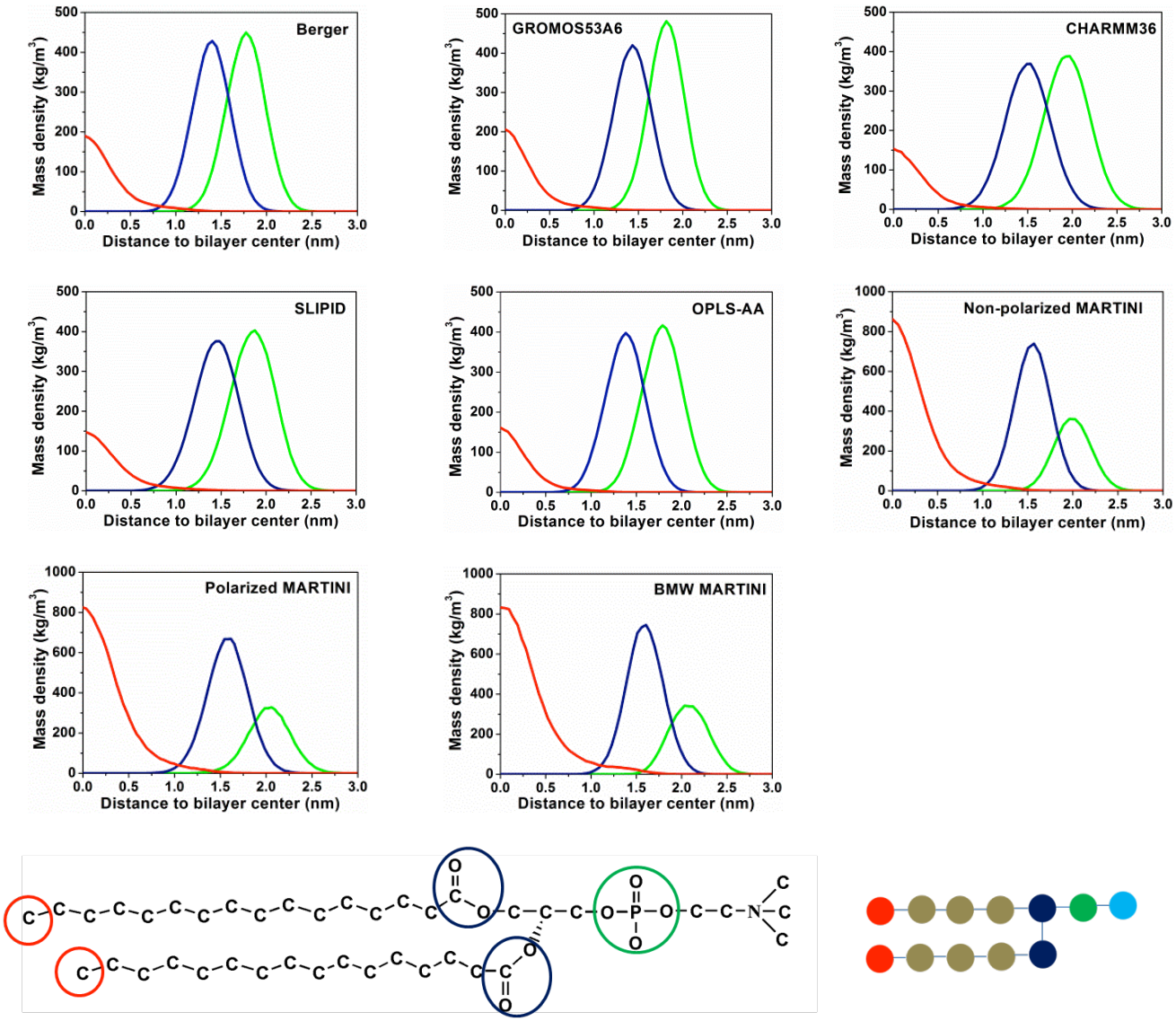


Figure 2. Simulated density distribution profiles for lipid phosphate groups (green), lipid glycerol groups (blue) and lipid terminal carbon atoms (red).

3.2 Water/Octanol Structure. For the biphasic water/octanol system, we performed MD simulations by starting with a configuration consisting of a pure octanol phase sandwiched by two water layers, as is shown in Figure 3 A. Periodic boundary conditions apply in all directions. During equilibration, we found that water molecules spontaneously and rapidly diffused into the octanol phase. In the octanol phase, the water molecules were not uniformly distributed but instead formed pockets of several molecules, see Figure 3 B. Similar aqueous structures in octanol were also observed in the work of Wick and Chang,[71] in which a polarizable model was used for both water and octanol molecules. Figure 3 C plots the time evolution of the number of water molecules in the octanol phase over 200 ns of equilibration time. It indicates that the process of water partitioning into the octanol phase reaches equilibrium within 150 ns. The resulting saturated mole fraction of water in octanol was approximately 0.25 (no octanol was observed to move into the water layer). This value agrees with that obtained by Wick and Chang,[71] as well as with experimental data.[72] Hence, the all-atom OPLS-AA force field appears to provide a reliable set of potential parameters with which to simulate the water/octanol biphasic system. Following equilibration, we plotted the density distribution profiles for water molecules, the total octanol site density, as well as those of the octanol hydroxyl group (octanol head). These are plotted in Figure 3D. The density profiles indicate that the thickness of the water/octanol interface is 1.7 nm while the thickness of the bulk octanol phase is 7.9 nm. The electrostatic potential profile along the direction perpendicular to the water/octanol interface (z direction) is also plotted in Figure 3 D. The water/octanol interface electrostatic potential is calculated to be 0.5 V, which is close to the electrostatic potential at the water/air interface (0.54 V).[73]

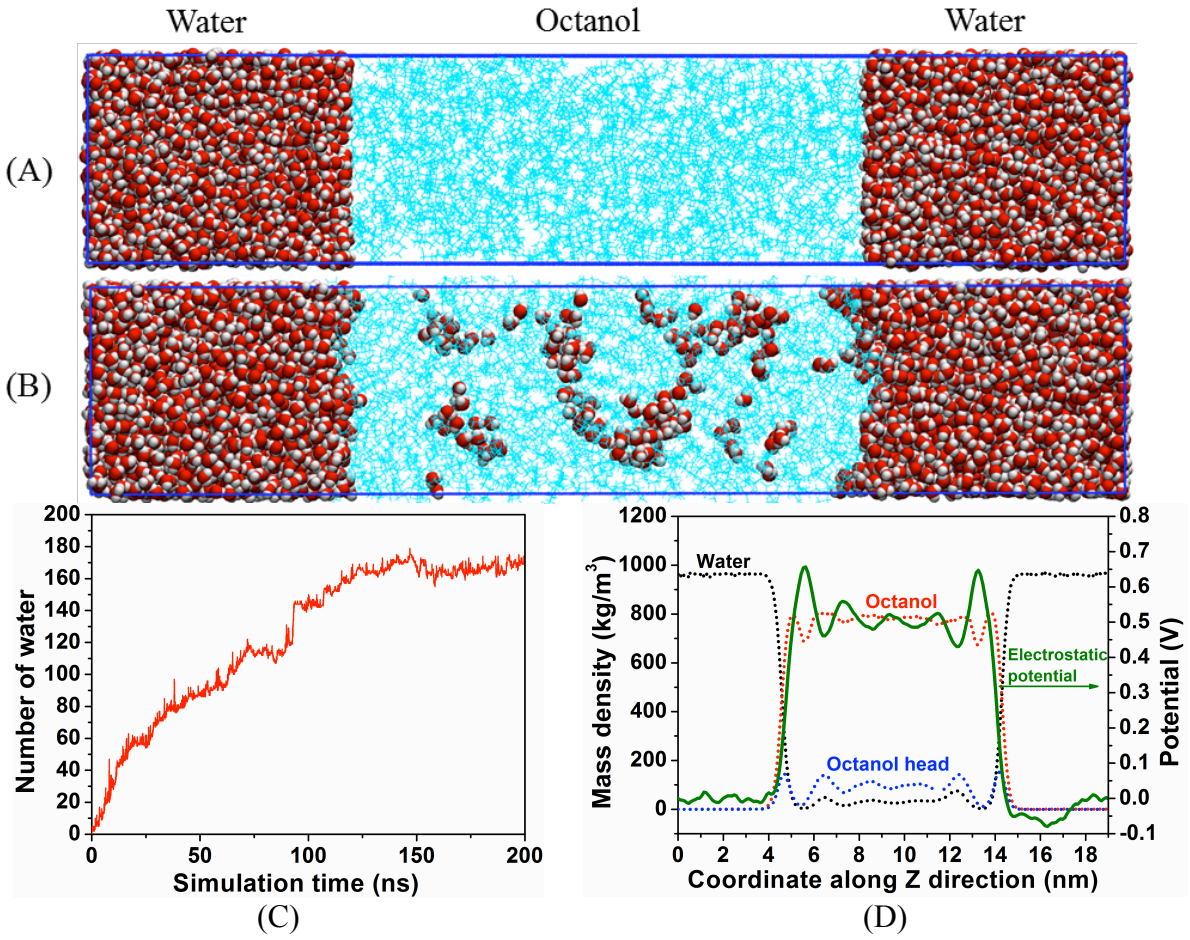


Figure 3. (A) The starting configuration and (B) the equilibrium configuration of the water/octanol biphasic system. (C) The simulated number of water molecules in octanol as a function of simulation time. (D)The density distribution profiles for water, octanol, octanol head and the electrostatic potential along the direction perpendicular to water/octanol interface (z direction).

3.3 Partitioning of Ionized Arg and Lys Side-Chains into Octanol. One would have expected that the water pockets formed in wet octanol have considerable effect on the water-to-octanol partitioning free energies for the ionized Arg and Lys side-chains. However, our free energy calculations found that the water in octanol makes only a small contribution to the cation solvation in wet octanol. Table 2 shows the calculated water-to-octanol (wet and dry) partitioning free energies. The free energies, ΔG_2 and ΔG_2^*, correspond to transferring the cations from vacuum into the dry and wet octanol

respectively. We see the presence of water in the wet octanol contributes only about a 1% decrease in the free energy of cation solvation in octanol. The specific contributions of the water pockets solvation in octanol for Arg and Lys are -0.63 kcal/mol and -0.69 kcal/mol respectively. As the net free energies of transfer are the result of subtracting two large numbers, the water makes a relatively larger contribution to the free energy of transfer. The net free energy costs for transferring Arg and Lys side-chains from water to wet octanol are 0.65±0.5 kcal/mol and 3.26±0.3 kcal/mol, respectively; whereas these values increase to 1.28 kcal/mol and 3.95 kcal/mol in the absence of water molecules in octanol.

Table 2 Calculated Solvation Free Energies and Water-to-Octanol Partitioning Free Energies for Arg and Lys Side-Chains. The Unit for the Free Energy is kcal/mol. Asterisk Represents the Wet Octanol Phase.

Amino acids	ΔG_1[a]	ΔG_{2*}[a]	ΔG_2[a]	ΔG_{corr}[b]	$\Delta G_{W\text{-}O*}$	$\Delta G_{W\text{-}O}$	WWOS[c]
Arg	49.46±0.4	-60.34±0.5	-59.71±0.4	11.53	0.65±0.5	1.28±0.4	1.81±0.13
Lys	58.65±0.3	-66.92±0.3	-66.23±0.2	11.53	3.26±0.3	3.95±0.3	2.80±0.11

[a] *The calculated solvation free energies were obtained without considering the correction terms due to the electrostatic potentials at the water/air and octanol/air interfaces.* [b] *The water/octanol electrostatic potential is 0.5 V in our simulations and this value corresponds to a free energy contribution of 11.53 kcal/mol due to the electrostatic potential.* [c] *The experimental values from Wimley-White octanol scales for the whole residues .*

To better understand why the presence of water in octanol has only a small effect on the solvation of the cationic Arg and Lys residues, we plotted the coordination numbers of oxygen atoms from the octanol or water molecules surrounding the central carbon atom of the guanidinium ion and the central nitrogen atom of the ammonium ion. Figure 4 indicates that in the wet octanol phase, Arg and Lys side-chains coordinate oxygen atoms from both water and octanol molecules whereas in the dry octanol phase, the side-chains only coordinate the octanol oxygen. However, the total numbers of oxygen atoms surrounding the side-chains are only slightly larger in the wet octanol phase. This

observation is consistent with our solvation free energy calculations. That is, water molecules largely displace the octanol molecules in the wet octanol phase. Interestingly, our results also reveal that the guanidinium ion mainly coordinates water molecules whereas the ammonium ion mainly coordinates the octanol molecule in the wet octanol (Figure 4).

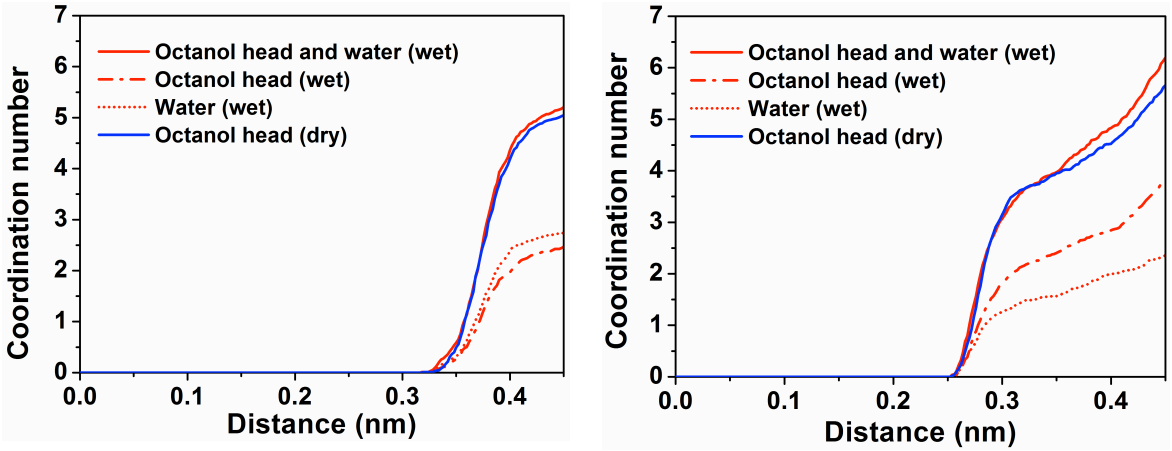


Figure 4. The distance-dependent coordination numbers of oxygen atoms on octanol heads or/and on water molecules surrounding the central carbon atom of guanidinium ion (A) and the central nitrogen atom of ammonium ion (B).

3.4 Partitioning of Ionized Arg and Lys Side-Chains into Lipid Bilayer. The simulated water-to-bilayer partitioning free energy profiles, using different force field combinations, are shown in Figure 5. From these profiles, we can extract the water-to-bilayer interface partitioning free energy, $\Delta G_{W\text{-}I}$ (free energy minimum) and the water-to-bilayer center partitioning free energy, $\Delta G_{W\text{-}C}$ (free energy maximum). The results for the different force field models are summarized in Table 2.

The free energy, $\Delta G_{W\text{-}I}$, purportedly measures a similar property to the Wimley-White interfacial scale. The Wimley-White value for ionized Arg is 0.81 kcal/mol and for ionized Lys it is 0.99 kcal/mol.[10] It should be noted, however, that these are so-called whole residue scales, Wimley and White also estimated a free energy contribution of 1.2 kcal/mol from the amino acid backbone.[10] As

a result, the side-chain contribution to the interfacial scale should be reduced by this quantity. However, as with the octanol scales discussed above, a direct comparison of these absolute quantities is probably not justifiable on physical grounds, given the different environments of the residues in the simulations and experiments. Furthermore, the Wimley-White interfacial scale was determined for the zwitterionic POPC lipid. However, it is probably reasonable to conclude that, according to the Wimley-White scale, the Arg side-chain appears to have a slightly greater affinity to a zwitterionic bilayer interface than Lys and that the association for both residues appears to be relatively weak with respect to k_BT.

From Table 2, we see that the Berger lipid model tends to predict larger $\Delta G_{W\text{-}I}$ and smaller $\Delta G_{W\text{-}C}$. With the Berger/OPLS-AA model, we obtained $\Delta G_{W\text{-}I}$=-7.1±0.2 kcal/mol, $\Delta G_{W\text{-}C}$=7.0±0.2 kcal/mol for the Arg side-chain and $\Delta G_{W\text{-}I}$=-4.7±0.2 kcal/mol, $\Delta G_{W\text{-}C}$=6.5±0.3 kcal/mol for the Lys side-chain. The association of these side-chains to the interface amount to several k_BT, which appears to contradict the Wimley-White interfacial scales. Though, the caveats which arise from considering the different molecular environments in the experiments compared to the simulations need to be considered.[10] MacCallum *et al.*[6] used the Berger/OPLS-AA combination to carry out a similar simulation study, but for the zwitterionic DOPC lipid bilayer instead. They report $\Delta G_{W\text{-}I}$=-5.1kcal/mol and $\Delta G_{W\text{-}C}$=13.9 kcal/mol for Arg side-chain and $\Delta G_{W\text{-}I}$=-4.4 kcal/mol, $\Delta G_{W\text{-}C}$=11.7 kcal/mol for Lys side-chain.[6] Thus, a very sizeable interfacial association is also predicted for the DOPC bilayer, with the Arg side-chain showing a ~$3k_BT$ stronger association free energy than Lys. However, there is some question as to whether the simulations of MacCallum *et al.*[6] have properly converged. Neale *et al.* performed extensive umbrella sampling simulations (205 ns per window) on essentially the same system as MacCallum *et al.*[6], also using the Berger/OPLS-AA model and they obtained $\Delta G_{W\text{-}I}$=-6.7 kcal/mol for Arg side-chain interacting with the DOPC bilayer.[74] While this value is close to our simulation result for Arg side-chain interacting with the DPPC bilayer, a more detailed review of these comparisons is warranted, and is presented in the Supporting Information.

The Berger/GROMOS87, Berger/GROMOS53A6 and GROMOS53A6/GROMOS53A6 models all predict strong attractions between the side chains and the zwitterionic DPPC lipid bilayer. They also give the somewhat suspicious result that Lys interacts more strongly than Arg with the DPPC bilayer. In particular, the Berger/GROMOS87 model predicts $\Delta G_{W\text{-}I}$=-9.8±0.3 kcal/mol for Arg and $\Delta G_{W\text{-}I}$=-10.8±0.2 kcal/mol for Lys. While the Berger/GROMOS53A6 model gives $\Delta G_{W\text{-}I}$=-5.2±0.2 kcal/mol for Arg and $\Delta G_{W\text{-}I}$=-12.9±0.3 kcal/mol for Lys. Strikingly, the Berger/GROMOS53A6 combination predicts a negative $\Delta G_{W\text{-}C}$ value for Lys. This surprising result appears to be due to the particularly strong interactions between Lys and the lipid head groups, which persist as the Lys is pulled into the membrane center. The GROMOS53A6/GROMOS53A6 model gives smaller association free energies for both cations, but Lys is still predicted to be more attractive to the bilayer than Arg by ~4 k_BT. These results disagree with those obtained from the Berger/OPLS-AA model and (as we shall see) the all-atom simulation models using the OPLS-AA/OPLS-AA, CHARMM36/ CHARMM36 and SLIPID/AMBER99SB-ILDN force fields. Furthermore, they are inconsistent with the Wimley-White interfacial hydrophobicity scales,[10] as well as other experimental studies which have compared the membrane affinity of Arg-rich and Lys-rich peptides.[75, 76] The general experimental evidence suggests that Arg has a stronger affinity to lipid membranes than Lys.[75, 76] This is often linked to the enhanced ability of Arg-rich peptides to penetrate cell membranes compared with Lys-rich peptides. We conjecture that the difference in affinity between Arg and Lys is presumably not large, as the cell penetrating ability of poly-Arg is only manifested in oligomers of length greater than 8-9 residues. This conjecture is also supported by the Wimley-White interfacial scale.

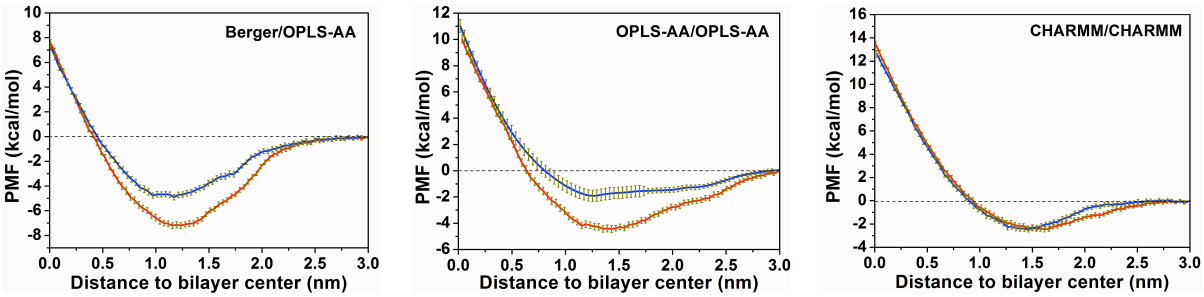
Using the all-atom CHARMM36/CHARMM36 model, we obtained essentially identical free energy profiles for Arg and Lys, with $\Delta G_{W\text{-}I}$=-2.4±0.2 kcal/mol, $\Delta G_{W\text{-}C}$=13.1±0.3 kcal/mol for Arg and $\Delta G_{W\text{-}I}$=-2.3±0.2 kcal/mol, $\Delta G_{W\text{-}C}$=12.5±0.3 kcal/mol for Lys. The all-atom SLIPID/AMBER99SB-ILDN model, gave $\Delta G_{W\text{-}I}$=-2.2±0.2 kcal/mol, $\Delta G_{W\text{-}C}$=14.0±0.2 kcal/mol for

Arg and $\Delta G_{W\text{-}I}$= -1.4±0.2 kcal/mol, $\Delta G_{W\text{-}C}$=12.1±0.2 kcal/mol for Lys. Thus, these all-atom force field combinations are in reasonable agreement with each other. Furthermore, the small $\Delta G_{W\text{-}I}$ value differences between Arg and Lys are also consistent with the Wimely-White interfacial scales.[10] Of note, our simulated $\Delta G_{W\text{-}C}$ values for both Arg and Lys side-chains are much smaller than those predicted by Li *et al*,[15] who also used the all-atom CHARMM force field for their free energy calculations. They reported $\Delta G_{W\text{-}C}$=21.6±0.1 kcal/mol for Arg side-chain and $\Delta G_{W\text{-}C}$=22.5±0.2 kcal/mol for a Lys side-chain permeating through the DPPC lipid bilayer. One primary difference between our simulations and those of Li *et al*. is the side-chain model. We used the longer and more hydrophobic propylguanidinium and butylammonium ions, whereas Li *et al.* used the shorter methylguanidinium and methylammonium ions.[15] The nontrivial effect of the side-chain length on the calculated free energies can be inferred from the simulation result of Li *et al*. They found that the calculated apparent free energy would be lowered by ~4 kcal/mol (compared with methylguanidinium and methylammonium ions) if propylguanidinium and butylammonium side-chains were attached to the poly-Leu transmembrane helix instead.[15] Furthermore, the movement of one Leu residue from the lipid bilayer to water will increases the free energy by ~3.4 kcal/mol as the propylguanidinium or butylammonium side chains are moved into the bilayer interior. This suggests that our simulation results and those of Li *et al*. and are somewhat in agreement.

The all-atom OPLS-AA/OPLS-AA combination estimated $\Delta G_{W\text{-}I}$=-4.4±0.3 kcal/mol, $\Delta G_{W\text{-}C}$=10.2 ±0.3 kcal/mol for the Arg side-chain and $\Delta G_{W\text{-}I}$=-1.6±0.4 kcal/mol, $\Delta G_{W\text{-}C}$=10.4±0.5 kcal/mol for the Lys side-chain. Compared with the Berger/OPLS-AA combination, the OPLS-AA/OPLS-AA model predicts a relatively weaker interaction between charged amino acids with the DPPC bilayer, which is more consistent with the Wimley-White interfacial scales.[10] However, the OPLS-AA/OPLS-AA model still predicts stronger Arg-lipid interactions than the other all-atom models (see above).

The three coarse-grained MARTINI force fields produced qualitatively different free energy profiles as well. Comparing the values of $\Delta G_{W\text{-}I}$, we see that both the non-polarizable and the

polarizable MARTINI force fields predict stronger interactions for the Lys side-chain with the bilayer than the Arg side chain. On the other hand, the BMW MARTINI force field predicts stronger interactions for Arg with the bilayer, Arg ($\Delta G_{W\text{-}I}$ =-4.8±0.3 kcal/mol) and Lys ($\Delta G_{W\text{-}I}$ =-3.6±0.3 kcal/mol). Indeed, these values are comparable with those calculated using the Berger/OPLS-AA combination. Of note is the fact that the non-polarizable MARTINI force field gives free energy profiles which flatten as the charged residues approach the bilayer center, whereas the BMW MARTINI and polarizable MARTINI force fields give a monotonically increasing free energy in this region. This is linked to the fact that latter models also predict a water-containing bilayer defect upon penetration by the cation, somewhat in agreement with the atomistic models. The non-polarizable MARTINI model shows no defect and this leads to the cation encountering an environment similar to a bulk hydrocarbon in the interior region of the bilayer. Hence, the free energy profile flattens. Compared with the atomistic models, the BMW MARTINI model predicts a very large membrane permeation free energy barrier for Arg ($\Delta G_{W\text{-}C}$=23.2±0.4 kcal/mol) and for Lys ($\Delta G_{W\text{-}C}$=23.4±0.5 kcal/mol). The other MARTINI versions give a $\Delta G_{W\text{-}C}$ in better agreement with the all-atom simulations. In general, the polarizable MARTINI force field seems to best reproduce the features of the free energy profiles as predicted by the all-atom models, CHARMM36/CHARMM36 and Slipid/AMBER99SB-ILDN. So if one assumes that the MARTINI models are constructed with the aspiration of mimicking all-atom results, it appears that the polarizable version is best suited to model the system under study.



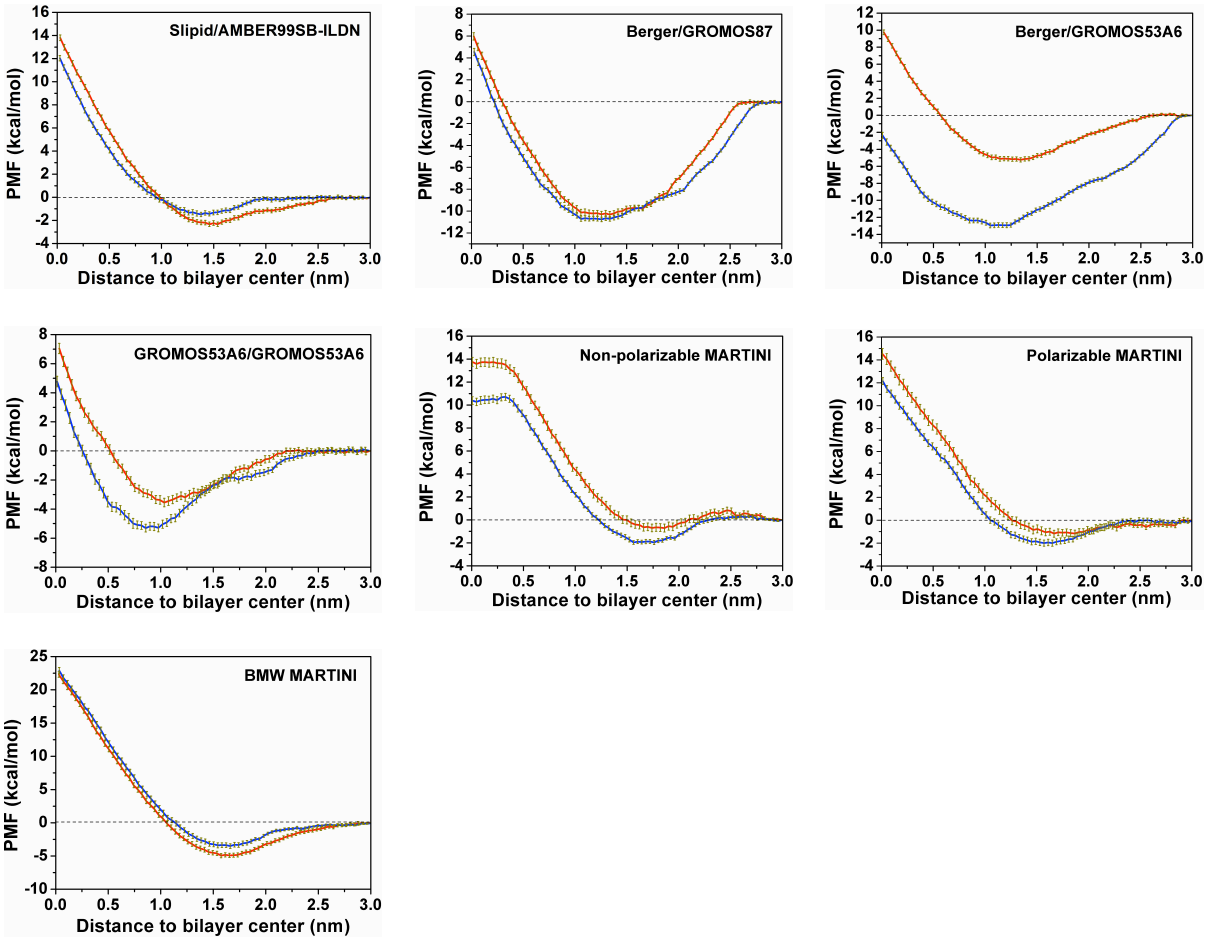


Figure 5. Simulated water-to-bilayer partitioning free energy profiles for Arg (orange) and Lys (cyan) side-chains. The free energy profiles were constructed using the last 30 ns simulation data in each umbrella sampling window. All free energy profiles are set to zero in the bulk water phase. Error bars in the figures represent the standard deviation of the mean.

The distribution profiles of the bilayer constituents (Figure 2) and the water-to-bilayer free energy profiles (Figure 5) suggest that, at the distance corresponding to the free energy minima, Arg and Lys cations are near the phosphate and glycerol regions of the lipid. Wu *et al.* [77] have used the all-atom CHARMM force field to compare the different interaction modes of Arg and Lys amino acids (with zwitterionic backbones) with the DPPC lipid bilayer. They found that Arg and Lys interact similarly with the lipid phosphate groups. However, Arg was found to interact much more strongly with the

lipid glycerol groups than Lys. This difference between Arg and Lys may possibly explain why Arg-rich peptides can induce negative Gaussian curvature (and membrane pores) in lipid membranes whereas Lys-rich peptides cannot.[78] We analyzed the interaction energies of Arg and Lys side-chains with the phosphate and glycerol groups for different force fields. The results are summarized in Table 3. These were obtained by constraining the side-chains at the position of the minimum of the free energy profile (where $\Delta G_{W\text{-}I}$ is calculated) for the corresponding force field model.

With the CHARMM36/CHARMM36 combination, we obtained results that are consistent with the simulations of Wu *et al.*[77] The calculated average interaction energies with phosphate are -30.3±3.8 kcal/mol and -14.8±2.5 kcal/mol for Arg and Lys respectively, whereas the interaction energy with the glycerol groups are -28.4±1.7 kcal/mol and -7.7±0.7 kcal/mol for Arg and Lys, respectively. The all-atom OPLS-AA/OPLS-AA, SLIPID/AMBER99SB-ILDN and the Berger/OPLS-AA models predict stronger interactions of Arg and Lys with the lipid phosphate groups than the lipid glycerol groups, whereas the united-atom models (Berger/GROMOS87, Berger/GROMOS53A6, GROMOS53A6/GROMOS53A6) predict stronger interactions of Arg and Lys with the lipid glycerol groups than the lipid phosphate groups (Berger/GROMOS53A6 model predicts stronger interactions of Arg with phosphate groups). Hence, our simulations suggest that the very strong interactions between Arg and Lys with the zwitterionic lipid bilayer, as predicted by the Berger/GROMOS87, Berger/GROMOS53A6 models are likely to be caused by the large electrostatic interactions between the lipid glycerol groups and the amino acids.

Table 2. Summarized Free Energy Calculation Results Predicted with Different Lipid and Protein Models.

Models	Free energies (kcal/mol)				Free energy minimum positions (nm)	
	$\Delta G_{W\text{-}I}$ (Arg)	$\Delta G_{W\text{-}C}$ (Arg)	$\Delta G_{W\text{-}I}$ (Lys)	$\Delta G_{W\text{-}C}$ (Lys)	Arg	Lys

Berger/OPLS-AA	-7.1±0.2	7.0±0.2	-4.7±0.2	6.5±0.3	1.2	1.2
Berger/GROMOS87	-9.8±0.3	5.7±0.3	-10.8±0.2	3.8±0.3	1.3	1.3
Berger/GROMOS53A6	-5.2±0.2	9.1±0.3	-12.9±0.3	-3.0±0.4	1.3	1.1
GROMOS53A6/GROMOS53A6	-3.3±0.3	6.9±0.3	-5.5±0.3	3.9±0.3	1.0	0.9
CHARMM36/CHARMM36	-2.4±0.2	13.1±0.3	-2.3±0.2	12.5±0.3	1.6	1.5
OPLS-AA/OPLS-AA	-4.4±0.3	10.2±0.3	-1.6±0.4	10.4±0.5	1.4	1.3
SLIPID/AMBER99SB-ILDN	-2.2±0.2	14.0±0.2	-1.4±0.2	12.1±0.2	1.5	1.4
Non-polarizable MARTINI	-0.7±0.3	13.5±0.4	-1.8±0.2	10.6±0.3	1.8	1.6
Polarizable MARTINI	-1.4±0.3	14.3±0.5	-2.1±0.3	12.3±0.3	1.7	1.6
BMW MARTINI	-4.8±0.3	23.2±0.4	-3.6±0.3	23.4±0.5	1.7	1.7

Table 3. Summarized Calculated Interaction Energies for Arg and Lys with Lipid Phosphate Groups (P) and Lipid Glycerol Groups (G). The Results Were Averaged over the Last 30 ns Umbrella Sampling Simulations.

Models	Interaction energies (kcal/mol)			
	E_{Arg-P}	E_{Arg-G}	E_{Lys-P}	E_{Lys-G}
Berger/OPLS-AA	-31.5±6.0	-25.8±2.9	-28.5±6.2	-14.9±1.6
Berger/GROMOS87	-26.8±2.7	-30.2±1.4	-2.6±1.0	-53.2±4.5
Berger/GROMOS53A6	-38.9±3.7	-12.0±2.8	-30.1±5.1	-34.4±4.9
GROMOS53A6/GROMOS53A6	-4.6±0.7	-27.8±2.4	-2.3±0.2	-33.8±5.1
CHARMM36/CHARMM36	-30.3±3.8	-14.8±2.5	-28.4±1.7	-7.7±0.7
OPLS-AA/OPLS-AA	-57.7±6.5	-8.0±3.0	-34.6±7.6	-23.6±8.2
SLIPID/AMBER99SB-ILDN	-31.7±3.8	-12.7±1.1	-23.8±2.8	-13.1±1.9

3.5 Membrane Defect Analysis. Spontaneous defects in cell membranes can form via thermally driven undulations or lipid flip-flop, though their role in peptide bilayer interactions has not been extensively explored. Vamparys *et al.* [79] used simulations to establish that amphiphilic lipid packing sensor (ALPS) motifs can detect lipid defects. Proteins with ALPS motifs are able to preferentially adsorb to positively curved membranes, which are rich in lipid packing defects. Bennett *et al.* [80] also

used simulations to show that lipid flip-flop can induce the formation of a relatively stable water-filled membrane pore in thin lipid bilayers. Simulations by Johansson and Lindahl[17] found that Arg amino acid, when inserted in a membrane, could induce a water-filled defect which helps Arg position itself in the vicinity of a trans-membrane protein. Simulations by Sun *et al.* [34] found that membrane-bound Arg-rich peptide can significantly stabilize the water-filled membrane defect to potentially allow cooperative diffusion through membrane pores.

Similar to other reported work, [6, 15, 16, 51] we found that when Arg or Lys was pulled through the DPPC bilayer, a water-filled defect usually formed spontaneously. However, the size of the defect is force field dependent. We compared the membrane defects formed for the different force fields when the side-chain was constrained at the center of the bilayer. This was essentially the terminal point of the umbrella sampling simulations used to generate the free energy profiles shown in Figure 5. Snapshots of typical defects predicted in our simulations are shown in Figure 6. It is clear that qualitatively different kinds of membrane defect arise from different models. The Berger/OPLS-AA, Berger/GROMOS87, Berger/GROMOS53A6 and GROMOS53A6/GROMOS53A6 united-atom models predict relatively small membrane defects, containing few water molecules. On the other hand, the all-atom CHARMM36/CHARMM36, SLIPID/AMBER99SN-ILDN and OPLS-AA/OPLS-AA models all predict larger water-filled membrane defects. To quantify the sizes of these membrane defects, we calculated the time-averaged coordination numbers of water oxygen surrounding the central carbon atom on the guanidinium ion and the central nitrogen atom on ammonium ion within 0.5 nm. The results are listed in Table 4. The coordination number results are consistent with the configurations in Figure 6, which show that the all-atom models predict larger membrane defects than the united-atom models. It is tempting to attribute the difference between the united-atom and all-atom models to the cation-lipid interactions. As pointed out above, the united atom models give stronger interactions between the cations and the glycerol regions of the bilayer. This would allow the cations to imbed themselves further into the membrane, with a smaller cost in

free energy, and presumably a smaller water defect. However, another difference between the all-atom and the united-atom simulations is the associated water model. SPC water is used in the united-atom simulations whereas a TIP3P or TIPS3P model is used in the all-atom simulations. To test this conjecture, we repeated the simulations for the Berger/OPLS-AA model, but using instead the TIP3P rather than SPC water model. We subsequently found that a larger membrane defect was formed when the TIP3P model was used. The calculated coordination number increases from 2.3 to 6.5 for Arg and from 2.9 to 4.3 for Lys. This is likely due to the slightly larger dipole moment and dielectric constant of the TIP3P water model (μ=2.35 D and ε=82),[42] compared with the SPC water model (μ=2.27 D and ε=65).[40] That is the water-water cohesive energy within the higher dipole model is greater. Furthermore, a larger dipole means that the ability to form stronger hydrogen bonds in the more hydrophobic environment of the membrane defect is also increased.

The water model has a role to play in the formation of membrane defects predicted by the coarse-grained MARTINI force fields as well. For the non-polarizable MARTINI model, no membrane defect forms when the Arg or Lys side-chain is inserted into the bilayer center. This is the reason why the calculated free energy profiles become relatively flat when the cation approaches the bilayer center. That is, the cation experiences an essentially bulk-like environment in the central region of the bilayer. The polarizable MARTINI and the BMW MARTINI improve the situation somewhat. In these cases, we see a membrane defect forms when the side-chains are in the membrane center. However, the size of the membrane defect is still quite small compared with the all-atom simulation results. In the MARTINI models, the effective water species is mapped onto four actual water molecules making it more difficult to accommodate these larger species within a membrane defect, compared with the atomistic water models. Water-water interactions are also suppressed in the non-polarizable model, especially the enhancement of dipole-dipole interactions in the more hydrophobic environment of membrane defects. In addition, one would expect that coarse-grained lipids would be less flexible, decreasing the compressibility of the coarse-grained bilayer and making defects

more difficult to form, especially as they require greater curvature in the membrane. The situation becomes somewhat improved when the water model is allowed to undergo dipole fluctuations, as in the BMW and polarizable models.

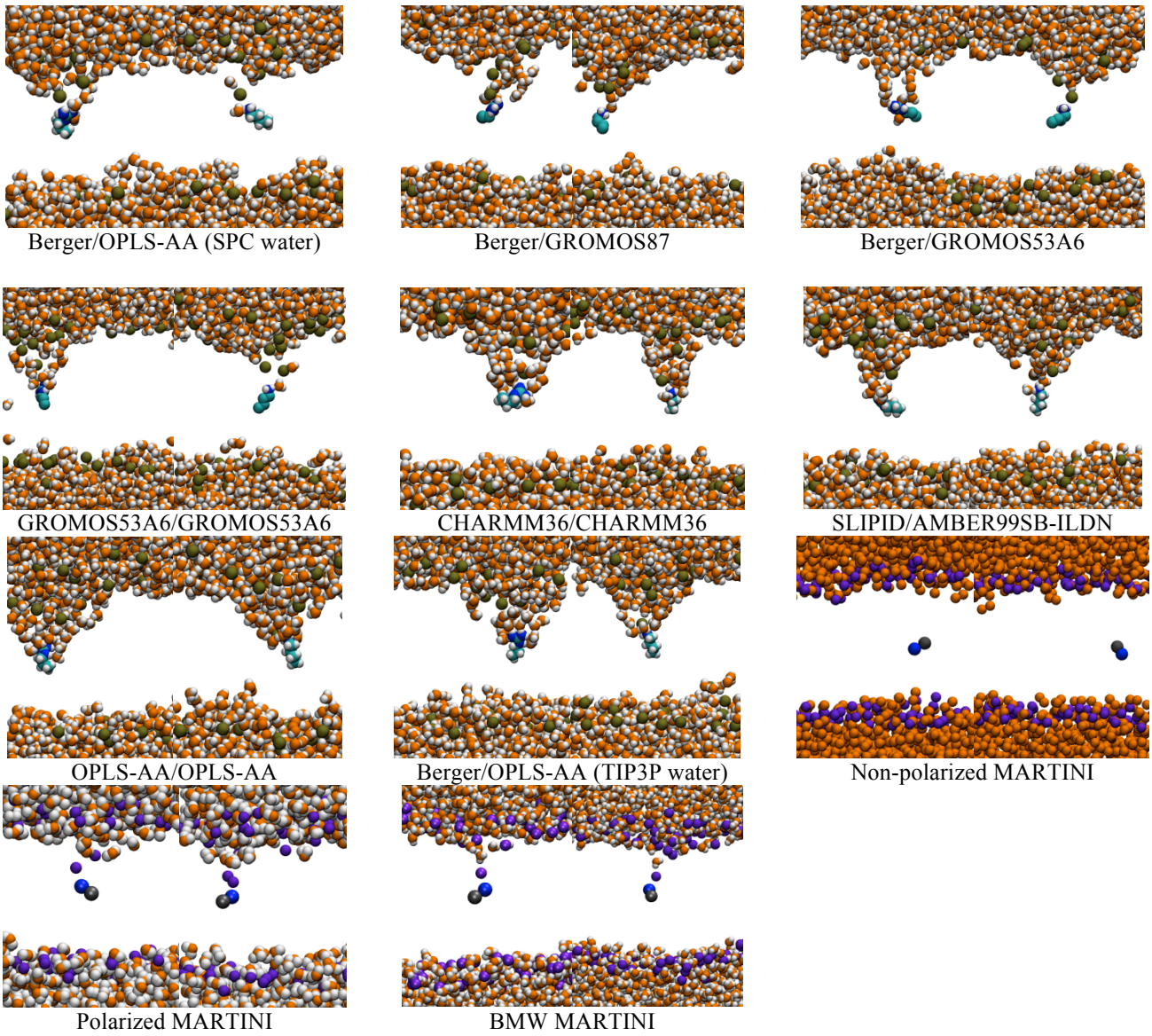


Figure 6. Water-filled membrane defects formed by constraining Arg (left) and Lys (right) side-chains in the lipid bilayer center. Only the side-chains, water molecules and phosphorus atoms (atomistic models) or phosphate groups (coarse-grained models) are shown for clarity purpose.

Table 4. Averaged Coordination Numbers of Water Oxygen Surrounding the Central Carbon Atom on Guanidinium Ion and the Central Nitrogen Atom on Ammonium Ion. The Solvation Shell Thickness Used to Calculate the Coordination Number Was Chosen to be 0.5 nm for both Arg and Lys Side-Chains.

Models	Coordination numbers	
	Arg	Lys
Berger/OPLS-AA (SPC water)	2.3	2.9
Berger/OPLS-AA (TIP3P water)	6.5	4.3
Berger/GROMOS87	3.7	2.0
Berger/GROMOS53A6	4.1	0.4
GROMOS53A6/GROMOS53A6	3.1	0.1
CHARMM36/CHARMM36	4.4	5.0
SLIPID/AMBER99SB-ILDN	4.5	5.4
OPLS-AA/OPLS-AA	5.8	6.2

4. DISCUSSION AND CONCLUSION

The Wimely-White octanol scale[9] is often used to estimate the ability of amino acid side-chains to be inserted into lipid bilayers. For example, Almeida and Pokorny have used this scale to determine the membrane-penetrating activity of antimicrobial peptides to be either "graded" or "all-or-none".[11] Using simulations, we are able to compare the free energies of transferring cationic amino acid side chains (Arg and Lys) from water to either octanol or the DPPC lipid bilayer. This allows us to directly evaluate the use of an octanol scale as a measure of the ability of a cationic species to penetrate lipid membranes, without the confounding factors that beset experiments.

In this work, we have used a thermodynamic integration method to calculate the free energies for transferring cationic Arg and Lys side-chains from water to either wet or dry octanol. It has been suggested that clusters of water molecules in wet octanol are responsible for the favorable solvation of cationic species, such as Arg and Lys.[13] Surprisingly, a high water content (0.25 mole fraction) in wet octanol lowered the solvation free energy free energies by only 0.63 kcal/mol and 0.69 kcal/mol for Arg and Lys, which accounted for about 1% of the solvation free energy (from vacuum). Although water molecules in wet octanol can dramatically alter the solvent microstructures[81, 82], the formation of water pockets in wet octanol does not significantly increase the number of polar groups (octanol head and water) coordinated to the side-chains. A similar phenomenon has also been found for other small polar and non-polar molecules.[82] Our simulations have also found that the free energies for transferring Arg and Lys side-chains from water to dry octanol are also very low.

The low water-to-octanol partitioning free energies for ionized Arg and Lys is a clear indication that water-to-octanol partitioning is different from water-to-lipid bilayer partitioning. Reported simulation results, using different lipid and peptide models and force fields, have predicted that the water-to-lipid bilayer partitioning free energies for Arg and Lys side-chains are of the order of 14-28 kcal/mol.[6, 14-16] On the other hand, a major issue that ultimately plagues simulation studies is the accuracy of the force field employed. In order to ascertain the kind of uncertainty to expect, we have carried out simulations on ten widely-used lipid/peptide force field combinations by plotting the water-to-bilayer potential of mean force profiles for the cationic Arg and Lys side-chains. Our key findings are summarized as follows.

(1) When the united-atom Berger lipid model is used in combination with the all-atom OPLS-AA or the united-atom GROMOS force field, interactions of Arg and Lys with the zwitterionic lipid bilayer surface ($\Delta G_{W\text{-}I}$) are predicted to be much stronger than the results predicted by all-atom models. The strong interactions are particularly striking if the Berger model is used in combination with the GROMOS87 (for both Arg and Lys) force field and with the GROMOS53A6 force field (for

Lys). The strong interactions between cationic residues and zwitterionic lipid bilayers are caused by the electrostatic interactions between the residues and the lipid glycerol groups. Numerous simulation studies of Arg- and Lys-rich peptides interacting with zwitterionic lipid bilayers have used the Berger/GROMOS87 combination.[32, 33, 83, 84] Taking into account the results of this work, it is possible that those studies may have overestimated peptide-bilayer interactions, but of course experiments must be the ultimate guide to their accuracy. Another concern regarding the Berger lipid model is whether current versions are suitable for modelling the range of lipid species to which they have been applied. Zhao *et al*.[85] developed the Berger palmitoyloleoyl-phosphatidylglycerol (POPG) lipid model and Mukhopadhyay *et al*.[86] developed the Berger palmitoyloleoylphosphatidylserine (POPS) lipid model, based on the Berger model for DPPC lipid. The calculated APL of POPG is 0.53 nm^2 in the absence of salt at 310 K.[85] The APL of POPS was calculated to be 0.55 nm^2 in the presence of salt at a temperature of 300 K.[86] These simulation results are much smaller than the experimental data for the APL of POPG (0.66 nm^2)[87] and POPS (0.63 nm^2)[88] bilayers. The smaller APL predicted by the Berger model is caused mainly by the significant number of so-called "intra-molecular hydrogen bonds". These hydrogen bonds are not observed in molecular simulations studies by Hénin et al.[89] who used the all-atom CHARMM force field. Clearly, significant intra-molecular hydrogen bonding, as predicted by the Berger anionic lipid models, would affect the interaction of Arg- and Lys-rich peptides with the lipid head groups.[90]

(2) The GROMOS87 and GROMOS53A6 force fields generally predict stronger interactions of Lys with the lipid bilayer surface than Arg. This result appears to be inconsistent with the Wimley-White interfacial scale for Arg and Lys, and is counter to the results of the all-atom CHARMM and AMBER combinations. Thus, it seems likely that the stronger interaction of Lys with the lipid bilayer is an artefact of the united-atom GROMOS protein models. Whether or not the newest version GROMOS54A8 force field[91] displays the same behaviour remains to be investigated.

(3) The all-atom CHARMM36/CHARMM36 and SLIPID/AMBER99SB-ILDN combinations produce consistent free energy profiles for Arg and Lys interacting with the DPPC bilayer. These force fields predict similar and weak interactions, $\Delta G_{W\text{-}I}$, between Arg and Lys side-chains and the DPPC lipid bilayer at the interface, which is in line with the Wimley-White interfacial scale.[10] The ~2 kcal/mol binding free energy difference between the simulation results and the Wimley-White interfacial scales could be due to the high concentration of electrolyte used in the experiments.[10] The major difference between these force fields is that the CHARMM combination predicts a stronger interaction between Arg and the lipid glycerol groups, compared to that of Lys. On the other hand, the SLIPID/AMBER99SB-ILDN combination predicts essentially identical interactions between the side-chains and the glycerol groups. The OPLS-AA/OPLS-AA force field predicts a much stronger interaction at the interface for Arg, compared with Lys, which is counter to the Wimley-White interfacial scales. This appears to be due to a much larger interaction between Arg and the lipid phosphate groups, compared with Lys. However, this force field has a much smaller interaction between Arg and the glycerol groups, compared with Lys. The good agreement with the qualitative aspects of the Wimley-White interfacial scales for Lys and Arg would suggest the CHARMM36/CHARMM36 and the SLIPID/AMBER99SB-ILDN models are the best to use in for molecular simulation work for this system at least.

(4) Not all of the coarse-grained MARTINI force fields accurately reproduced the qualitative features of the potentials of mean force predicted by the all-atom models CHARMM36/CHARMM36 and SLIPID-AMBER99SB-ILDN. The non-polarizable and polarizable MARTINI models both give weak interactions between the charged side-chains and the DPPC lipid bilayer, which are in agreement with all-atom simulations as well as the Wimley-White interfacial scales. However, they do predict that Lys interacts with the bilayer interface at least twice as strongly as Arg. In addition, the non-polarizable MARTINI model predicts that the free energy profiles flatten out when the charged residues approach the lipid bilayer center, in obvious contradiction to the all-

atom simulation results. On the other hand, the BMW MARTINI force field predicts stronger interactions for Arg with the bilayer interface. However, this model also predicts an extremely high free energy cost for placing Arg or Lys in the bilayer center. The polarizable MARTINI force field compares the most favourably with the all-atom models, though it would appear that future refinements of the MARTINI force field are necessary to better describe Arg and Lys interacting with a lipid bilayer.[92]

(5) Finally, we have examined the water-filled membrane defect that forms when Arg or Lys residues are positioned in the lipid bilayer center. The different force field classes give rise to different defect size, according to the following trend: all-atom models > united atom models> coarse-grained models. The replacement of the SPC water model by the TIP3P model in the Berger/OPLS-AA model caused the defect to grow significantly. More accurate quantum mechanics calculations are needed in order to clarify the differences among the atomistic models. There is also growing concern regarding the accuracy of classical atomistic force fields in predicting the free energies of solutes partitioning from water to a non-polar environment like the lipid bilayer interior due to the lack of explicit electronic polarization terms in the force fields.[93] For example, it has been reported that the CHARMM force field has serious problem in predicting the solvation free energy of the charged Arg side-chain in cyclohexane.[94] However, Vorobyov *et al*. found that the predicted free energies for transferring one Arg side-chain from water to the DPPC lipid bilayer center were almost identical for the classical non-polarizable CHARMM and the Drude polarizable CHARMM force fields.[94] This surprising result is due to the formation of the water-filled membrane defect as the charged amino acid side-chain translocates across the lipid bilayer, i.e. the charged amino acid is never really within the hydrophobic bilayer interior environment during the translocation process. Our simulations for water to octanol transfer show that the same is true in that case, i.e., the solvation environments of the side-chains in the water and octanol phases are very similar. Side-chain induced water-filled membrane defects are not observed in the non-polarizable MARTINI model, but small

defects are seen with the polarizable and BMW versions. A similar phenomenon was also found in Bennett and Tieleman's simulation work which used three versions of MARTINI force field to study the membrane defect formation induced by lipid flip-flop.[95] Further refinement of the coarse-grained water models are needed in order to accurately simulate the membrane defect, which is clearly a key element in peptide- membrane interactions and pore formation.

ASSOCIATED CONTENT

Supporting Information

Umbrella histogram overlaps and potential of mean force convergence analysis. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: c.woodward@adfa.edu.au

Notes

The authors declare no competing financial interests.

ACKNOWLEDGMENTS

An allocation time from the Lunarc Computing Center at Lund University are gratefully acknowledged. JF acknowledges financial support from the Swedish Research Council.

REFERENCES

(1) El-Andaloussi, S.; Holm, T.; Langel, U. Cell-penetrating peptides: Mechanisms and applications. *Curr. Pharm. Design* **2005,** *11*, 3597-3611.

(2) Jiang, Y. X.; Ruta, V.; Chen, J. Y.; Lee, A.; MacKinnon, R. The principle of gating charge movement in a voltage-dependent K+ channel. *Nature* **2003,** *423*, 42-48.

(3) Shai, Y. Mode of action of membrane active antimicrobial peptides. *Peptide Sci.* **2002,** *66*, 236-248.

(4) Hessa, T.; Kim, H.; Bihlmaier, K.; Lundin, C.; Boekel, J.; Andersson, H.; Nilsson, I.; White, S. H.; von Heijine, G. Recognition of transmembrane helices by the endoplasmic reticulum translocon. *Nature* **2005,** *433*, 377-381.

(5) Johansson, A. C. V.; Lindahl, E. Position-resolved free energy of solvation for amino acids in lipid membranes from molecular dynamics simulations. *Proteins: Struct. Funct. Bioinform.* **2008,** *70*, 1332-1344.

(6) MacCallum, J. L.; Bennett, W. F. D.; Tieleman, D. P. Distribution of amino acids in a lipid bilayer from computer simulations. *Biophys. J.* **2008,** *94*, 3393-3404.

(7) Moon, C. P.; Fleming, K. G. Side-chain hydrophobicity scale derived from transmembrane protein folding into lipid bilayers. *Proc. Natl. Acad. Sci. U. S. A.* **2011,** *108*, 10174-10177.

(8) Radzicka, A.; Wolfenden, R. Comparing the polarities of the amino acids: Side-chain distribution coefficients between the vapor phase, cyclohexane, 1-octanol, and neutral aqueous solution. *Biochemistry* **1988,** *27*, 1664-1670.

(9) Wimley, W. C.; Creamer, T. P.; White, S. H. Solvation energies of amino acid side chains and backbone in a family of host-guest pentapeptides. *Biochemistry* **1996,** *35*, 5109-5124.

(10) Wimley, W. C.; White, S. H. Experimentally determined hydrophobicity scale for proteins at membrane interfaces. *Nat.Struct. Biol.* **1996,** *3*, 842-848.

(11) Almeida, P. F.; Pokorny, A. Mechanisms of antimicrobial, cytolytic, and cell-penetrating peptides: From kinetics to thermodynamics. *Biochemistry* **2009,** *48*, 8083-8093.

(12) Wimley, W. C.; Gawrisch, K.; Creamer, T. P.; White, S. H. Direct measurement of salt-bridge solvation energies using a peptide model system: Implications for protein stability. *Proc. Natl. Acad. Sci. U. S. A.* **1996,** *93*, 2985-2990.

(13) Roux, B. Lonely arginine seeks friendly environment. *J. Gen. Physiol.* **2007,** *130*, 233-236.

(14) Dorairaj, S.; Allen, T. W. On the thermodynamic stability of a charged arginine side chain in a transmembrane helix. *Proc. Natl. Acad. Sci. U. S. A.* **2007,** *104*, 4943-4948.

(15) Li, L. B.; Vorobyov, I.; Allen, T. W. The different interactions of lysine and arginine side chains with lipid membranes. *J. Phys. Chem. B* **2013,** *117*, 11906-11920.

(16) Ou, S. C.; Lucas, T. R.; Zhong, Y.; Bauer, B. A.; Hu, Y.; Patel, S. Free energetics and the role of water in the permeation of methyl guanidinium across the bilayer-water interface: Insights from molecular dynamics simulations using charge equilibration potentials. *J. Phys. Chem. B* **2013,** *117*, 3578-3592.

(17) Johansson, A. C. V.; Lindahl, E. Protein contents in biological membranes can explain abnormal solvation of charged and polar residues. *Proc. Natl. Acad. Sci. U. S. A.* **2009,** *106*, 15684-15689.

(18) Gumbart, J.; Chipot, C.; Schulten, K. Free-energy cost for translocon-assisted insertion of membrane proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2011,** *108*, 3596-3601.

(19) Gumbart, J.; Roux, B. Determination of membrane-insertion free energies by molecular dynamics simulations. *Biophys. J.* **2012,** *102*, 795-801.

(20) Fleming, P. J.; Freites, J. A.; Moon, C. P.; Tobias, D. J.; Fleming, K. G. Outer membrane phospholipase A in phospholipid bilayers: A model system for concerted computational and experimental investigations of amino acid side chain partitioning into lipid bilayers. *Biochim. Biophys. Acta, Biomembr.* **2012,** *1818*, 126-134.

(21) Jorgensen, W. L.; Maxwell, D. S.; TiradoRives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* **1996,** *118*, 11225-11236.

(22) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **1998,** *102*, 3586-3616.

(23) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* **1995,** *117*, 5179-5197.

(24) Oostenbrink, C.; Villa, A.; Mark, A. E.; Van Gunsteren, W. F. A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. *J. Comput. Chem.* **2004,** *25*, 1656-1676.

(25) Shirts, M. R.; Pitera, J. W.; Swope, W. C.; Pande, V. S. Extremely precise free energy calculations of amino acid side chain analogs: Comparison of common molecular mechanics force fields for proteins. *J. Chem. Phys.* **2003,** *119*, 5740-5761.

(26) Villa, A.; Mark, A. E. Calculation of the free energy of solvation for neutral analogs of amino acid side chains. *J. Comput. Chem.* **2002,** *23*, 548-553.

(27) Klauda, J. B.; Venable, R. M.; Freites, J. A.; O'Connor, J. W.; Tobias, D. J.; Mondragon-Ramirez, C.; Vorobyov, I.; MacKerell, A. D.; Pastor, R. W. Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. *J. Phys. Chem. B* **2010,** *114*, 7830-7843.

(28) Berger, O.; Edholm, O.; Jähnig, F. Molecular dynamics simulations of a fluid bilayer of dipalmitoylphosphatidylcholine at full hydration, constant pressure, and constant temperature. *Biophys. J.* **1997,** *72*, 2002-2013.

(29) Ryckaert, J. P.; Bellemans, A. Molecular dynamics of liquid alkanes. *Faraday Discuss.* **1978,** *66*, 95-106.

(30) Jorgensen, W. L.; TiradoRives, J. The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.* **1988,** *110*, 1657-1666.

(31) Chiu, S. W.; Clark, M.; Balaji, V.; Subramaniam, S.; Scott, H. L.; Jakobsson, E. Incorporation of surface tension into molecular dynamics simulation of an interface: a fluid phase lipid bilayer membrane. *Biophys. J.* **1995,** *69*, 1230-1245.

(32) Dunkin, C. M.; Pokorny, A.; Almeida, P. F.; Lee, H. S. Molecular dynamics studies of transportan 10 (Tp10) interacting with a POPC lipid bilayer. *J. Phys. Chem. B* **2011,** *115*, 1188-1198.

(33) Herce, H. D.; García, A. E. Molecular dynamics simulations suggest a mechanism for translocation of the HIV-1 TAT peptide across lipid membranes. *Proc. Natl. Acad. Sci. U. S. A.* **2007,** *104*, 20805-20810.

(34) Sun, D. L.; Forsman, J.; Lund, M.; Woodward, C. E. Effect of arginine-rich cell penetrating peptides on membrane pore formation and life-times: A molecular simulation study. *Phys. Chem. Chem. Phys.* **2014,** *16*, 20785-20795.

(35) Kukol, A. Lipid models for united-atom molecular dynamics simulations of proteins. *J. Chem. Theory Comput.* **2009,** *5*, 615-626.

(36) Jämbeck, J. P. M.; Lyubartsev, A. P. Derivation and systematic validation of a refined all-atom force field for phosphatidylcholine lipids. *J. Phys. Chem. B* **2012,** *116*, 3164-3179.

(37) Jämbeck, J. P. M.; Lyubartsev, A. P. An extension and further validation of an all-atomistic force field for biological membranes. *J. Chem. Theory Comput.* **2012,** *8*, 2938-2948.

(38) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins: Struct. Funct. Bioinform.* **2010,** *78*, 1950-1958.

(39) Maciejewski, A.; Pasenkiewicz-Gierula, M.; Cramariuc, O.; Vattulainen, I.; Rog, T. Refined OPLS all-atom force field for saturated phosphatidylcholine bilayers at full hydration. *J. Phys. Chem. B* **2014,** *118*, 4571-4581.

(40) Berendsen, H. J.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J., Interaction models for water in relation to protein hydration. In *Intermolecular Forces*, Pullman, B., Ed. 1981; pp 331-342.

(41) Neria, E.; Fischer, S.; Karplus, M. Simulation of activation free energies in molecular systems. *J. Chem. Phys.* **1996,** *105*, 1902-1921.

(42) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983,** *79*, 926-935.

(43) Caleman, C.; van Maaren, P. J.; Hong, M. Y.; Hub, J. S.; Costa, L. T.; van der Spoel, D. Force field benchmark of organic liquids: density, enthalpy of vaporization, heat capacities, surface tension, isothermal compressibility, volumetric expansion coefficient, and dielectric constant. *J. Chem. Theory Comput.* **2012,** *8*, 61-74.

(44) de Jong, D. H.; Singh, G.; Bennett, W. F. D.; Arnarez, C.; Wassenaar, T. A.; Schafer, L. V.; Periole, X.; Tieleman, D. P.; Marrink, S. J. Improved parameters for the Martini coarse-grained protein force field. *J. Chem. Theory Comput.* **2013,** *9*, 687-697.

(45) Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H. The MARTINI force field: Coarse grained model for biomolecular simulations. *J. Phys. Chem. B* **2007,** *111*, 7812-7824.

(46) Monticelli, L.; Kandasamy, S. K.; Periole, X.; Larson, R. G.; Tieleman, D. P.; Marrink, S. J. The MARTINI coarse-grained force field: Extension to proteins. *J. Chem. Theory Comput.* **2008,** *4*, 819-834.

(47) Yesylevskyy, S. O.; Schafer, L. V.; Sengupta, D.; Marrink, S. J. Polarizable water model for the coarse-grained MARTINI force field. *PLoS Comp. Biol.* **2010,** *6*, e1000810.

(48) Wu, Z.; Cui, Q.; Yethiraj, A. A new coarse-grained force field for membrane-peptide simulations. *J. Chem. Theory Comput.* **2011,** *7*, 3793-3802.

(49) Cordomi, A.; Edholm, O.; Perez, J. J. Effect of force field parameters on sodium and potassium ion binding to dipalmitoyl phosphatidylcholine bilayers. *J. Chem. Theory Comput.* **2009,** *5*, 2125-2134.

(50) Hu, Y.; Ou, S. C.; Patel, S. Free energetics of arginine permeation into model DMPC lipid bilayers: coupling of effective counterion concentration and lateral bilayer dimensions. *J. Phys. Chem. B* **2013,** *117*, 11641-11653.

(51) Li, L. B.; Vorobyov, I.; Allen, T. W. The role of membrane thickness in charged protein-lipid interactions. *Biochim. Biophys. Acta, Biomembr.* **2012,** *1818*, 135-145.

(52) Isralewitz, B.; Baudry, J.; Gullingsrud, J.; Kosztin, D.; Schulten, K. Steered molecular dynamics investigations of protein function. *J. Mol. Graph. Model.* **2001,** *19*, 13-25.

(53) Kastner, J.; Thiel, W. Bridging the gap between thermodynamic integration and umbrella sampling provides a novel analysis method: "umbrella integration". *J. Chem. Phys.* **2005,** *123*, 144104.

(54) Hub, J. S.; de Groot, B. L.; van der Spoel, D. g_wham-A free weighted histogram analysis implementation including robust error and autocorrelation estimates. *J. Chem. Theory Comput.* **2010,** *6*, 3713-3720.

(55) Deng, Y. Q.; Roux, B. Hydration of amino acid side chains: Nonpolar and electrostatic contributions calculated from staged molecular dynamics free energy simulations with explicit water molecules. *J. Phys. Chem. B* **2004,** *108*, 16567-16576.

(56) Sachs, J. N.; Crozier, P. S.; Woolf, T. B. Atomistic simulations of biologically realistic transmembrane potential gradients. *J. Chem. Phys.* **2004,** *121*, 10847-10851.

(57) Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; Hess, B.; Lindahl, E. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **2013,** *29*, 845-854.

(58) Piggot, T. J.; Piñeiro, Á.; Khalid, S. Molecular dynamics simulations of phosphatidylcholine membranes: A comparative force field study. *J. Chem. Theory Comput.* **2012,** *8*, 4593-4609.

(59) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A smooth particle mesh Ewald mthod. *J. Chem. Phys.* **1995,** *103*, 8577-8593.

(60) Piana, S.; Lindorff-Larsen, K.; Dirks, R. M.; Salmon, J. K.; Dror, R. O.; Shaw, D. E. Evaluating the effects of cutoffs and treatment of long-range electrostatics in protein folding simulations. *PLoS ONE* **2012,** *7*, e39918.

(61) Hoover, W. G. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A* **1985,** *31*, 1695-1697.

(62) Nosé, S. A. A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.* **1984,** *52*, 255-268.

(63) Parrinello, M.; Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J . Appl. Phys.* **1981,** *52*, 7182-7190.

(64) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007,** *126*, 014101.

(65) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **1997,** *18*, 1463-1472.

(66) Miyamoto, S.; Kollman, P. A. Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J. Comput. Chem.* **1992,** *13*, 952-962.

(67) Kučerka, N.; Nagle, J. F.; Sachs, J. N.; Feller, S. E.; Pancer, J.; Jackson, A.; Katsaras, J. Lipid bilayer structure determined by the simultaneous analysis of neutron and x-ray scattering data. *Biophys. J.* **2008,** *95*, 2356-2367.

(68) Nagle, J. F.; Tristram-Nagle, S. Structure of lipid bilayers. *Biochim. Biophys. Acta, Biomembr.* **2000,** *1469*, 159-195.

(69) Kučerka, N.; Neih, M. P.; Katsaras, J. Fluid phase lipid areas and bilayer thicknesses of commonly used phosphatidylcholines as a function of temperature. *Biochim. Biophys. Acta, Biomembr.* **2011,** *1808*, 2761-2771.

(70) Kučerka, N.; Tristram-Nagle, S.; Nagle, J. F. Closer look at structure of fully hydrated fluid phase DPPC bilayers. *Biophys. J.* **2006,** *90*, L83-L85.

(71) Wick, C. D.; Chang, T. M. Computational observation of pockets of enhanced water concentration at the 1-octanol/water interface. *J. Phys. Chem. B* **2014,** *118*, 7785-7791.

(72) Šegatin, N.; Klofutar, C. Thermodynamics of the solubility of water in 1-hexanol, 1-octanol, 1-decanol, and cyclohexanol. *Monatsh. Chem.* **2004,** *135*, 241-248.

(73) Lamoureux, G.; MacKerell, A. D.; Roux, B. A simple polarizable model of water based on classical Drude oscillators. *J. Chem. Phys.* **2003,** *119*, 5185-5197.

(74) Neale, C.; Bennett, W. F. D.; Tieleman, D. P.; Pomès, R. Statistical convergence of equilibrium properties in simulations of molecular solutes embedded in lipid bilayers. *J. Chem. Theory Comput.* **2011,** *7*, 4175-4188.

(75) Amand, H. L.; Boström, C. L.; Lincoln, P.; Nordén, B.; Esbjörner, E. K. Binding of cell-penetrating penetratin peptides to plasma membrane vesicles correlates directly with cellular uptake. *Biochim. Biophys. Acta, Biomembr.* **2011,** *1808*, 1860-1867.

(76) Su, Y. C.; Doherty, T.; Waring, A. J.; Puchala, P.; Hong, M. Roles of arginine and lysine residues in the translocation of a cell-penetrating peptide from (13)C, (31)P, and (19)F solid-state NMR. *Biochemistry* **2009,** *48*, 4587-4595.

(77) Wu, Z.; Cui, Q.; Yethiraj, A. Why do arginine and lysine organize lipids differently? Insights from coarse-grained and atomistic simulations. *J. Phys. Chem. B* **2013,** *117*, 12145-12156.

(78) Mishra, A.; Gordon, V. D.; Yang, L. H.; Coridan, R.; Wong, G. C. L. HIV TAT forms pores in membranes by inducing saddle-splay curvature: Potential role of bidentate hydrogen bonding. *Angew. Chem. Int. Ed.* **2008,** *47*, 2986-2989.

(79) Vamparys, L.; Vanni, S.; Gautier, R.; Drin, G.; Antonny, B.; Etchebest, C.; Fuchs, P. Molecular dynamics study of a protein motif that senses packing defects induced by membrane curvature. *Biophys. J.* **2013,** *104*, 662a.

(80) Bennett, W. F. D.; Sapay, N.; Tieleman, D. P. Atomistic simulations of pore formation and closure in lipid bilayers. *Biophys. J.* **2014,** *106*, 210-219.

(81) MacCallum, J. L.; Tieleman, D. P. Structures of neat and hydrated 1-octanol from computer simulations. *J. Am. Chem. Soc.* **2002,** *124*, 15085-15093.

(82) Chen, B.; Siepmann, J. I. Microscopic structure and solvation in dry and wet octanol. *J. Phys. Chem. B* **2006,** *110*, 3555-3563.

(83) Huang, K.; García, A. E. Free energy of translocating an arginine-rich cell-penetrating peptide across a lipid bilayer suggests pore formation. *Biophys. J.* **2013,** *104*, 412-420.

(84) Pourmousa, M.; Wong-ekkabut, J.; Patra, M.; Karttunen, M. Molecular dynamic studies of transportan interacting with a DPPC lipid bilayer. *J. Phys. Chem. B* **2013,** *117*, 230-241.

(85) Zhao, W.; Rog, T.; Gurtovenko, A. A.; Vattulainen, I.; Karttunen, M. Atomic-scale structure and electrostatics of anionic palmitoyloleoylphosphatidylglycerol lipid bilayers with Na+ counterions. *Biophys. J.* **2007,** *92*, 1114-1124.

(86) Mukhopadhyay, P.; Monticelli, L.; Tieleman, D. P. Molecular dynamics simulation of a palmitoyl-oleoyl phosphatidylserine bilayer with Na+ counterions and NaCl. *Biophys. J.* **2004,** *86*, 1601-1609.

(87) Kučerka, N.; Holland, B. W.; Gray, C. G.; Tomberli, B.; Katsaras, J. Scattering density profile model of POPG bilayers as determined by molecular dynamics simulations and small-angle neutron and X-ray scattering experiments. *J. Phys. Chem. B* **2012,** *116*, 232-239.

(88) Pan, J. J.; Cheng, X. L.; Monticelli, L.; Heberle, F. A.; Kučerka, N.; Tieleman, D. P.; Katsaras, J. The molecular structure of a phosphatidylserine bilayer determined by scattering and molecular dynamics simulations. *Soft Matter* **2014,** *10*, 3716-3725.

(89) Hénin, J.; Shinoda, W.; Klein, M. L. Models for phosphatidylglycerol lipids put to a structural test. *J. Phys. Chem. B* **2009,** *113*, 6958-6963.

(90) Vazdar, M.; Wernersson, E.; Khabiri, M.; Cwiklik, L.; Jurkiewicz, P.; Hof, M.; Mann, E.; Kolusheva, S.; Jelinek, R.; Jungwirth, P. Aggregation of oligoarginines at phospholipid membranes: Molecular dynamics simulations, time-dependent fluorescence shift, and biomimetic colorimetric assays. *J. Phys. Chem. B* **2013,** *117*, 11530-11540.

(91) Reif, M. M.; Hünenberger, P. H.; Oostenbrink, C. New interaction parameters for charged amino acid side chains in the GROMOS force field. *J. Chem. Theory Comput.* **2012,** *8*, 3705-3723.

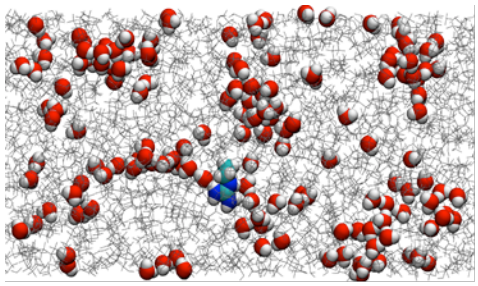
(92) Marrink, S. J.; Tieleman, D. P. Perspective on the Martini model. *Chem. Soc. Rev.* **2013,** *42*, 6801-6822.

(93) Leontyev, I.; Stuchebrukhov, A. Accounting for electronic polarization in non-polarizable force fields. *Phys. Chem. Chem. Phys.* **2011,** *13*, 2613-2626.

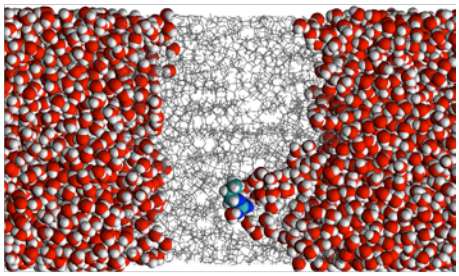
(94) Vorobyov, I.; Li, L. B.; Allen, T. W. Assessing atomistic and coarse-grained force fields for protein-lipid interactions: The formidable challenge of an ionizable side chain in a membrane. *J. Phys. Chem. B* **2008,** *112*, 9588-9602.

(95) Bennett, W. F. D.; Tieleman, D. P. Water defect and pore formation in atomistic and coarse-grained lipid membranes: Pushing the limits of coarse graining. *J. Chem. Theory Comput.* **2011,** *7*, 2981-2988.

Table of Contents Graphic



Arginine in octanol



Arginine in lipid bilayer