Parameters of Monovalent Ions in the AMBER-99 Forcefield: Assessment of Inaccuracies and Proposed Improvements

Alan A. Chen[†] and Rohit V. Pappu*[,†,‡]

Molecular Biophysics Program & Center for Computational Biology, Washington University School of Medicine, St. Louis, Missouri 63110, and Department of Biomedical Engineering & Center for Computational Biology, Washington University in St. Louis, One Brookings Drive, Campus Box 1097, St. Louis, Missouri 63130

The monovalent ion parameters used by the AMBER-99 forcefield are shown to exhibit physically inaccurate behavior in molecular dynamics simulations of strong 1:1 electrolytes. These errors arise from an ad hoc adaptation of Åqvist's cation parameters. The result is the rapid formation of large, unphysical clusters at concentrations that are well below solubility limits. The observed unphysical behavior poses a serious challenge for simulating ions around highly charged polymers such as nucleic acids. In this communication, we explain the source of this unphysical behavior. To facilitate the continued use of the popular AMBER parameters, we prescribe a simple fix whereby Åqvist's cations and anions are used in conjunction with the AMBER forcefield for nucleic acids. A preliminary test of this strategy suggests that the proposed fix is reasonable and is likely to be generalizable for simulating diffuse and specific ion binding to nucleic acids.

Introduction

Biological reactions, particularly those involving nucleic acids, are known to depend on both concentration and type of salt.[1] Atomistic simulations play an important role in the refinement/development of theories for explaining salt-dependent phenomena.[2−6] However, accurate modeling of ions in aqueous electrolytes remains a challenge in the design of reliable molecular mechanics forcefields. Close attention must be paid to the accuracy of both the ion parameters and the empirical rules for combining these parameters with those for nucleic acids.

The purpose of this communication is to notify the community of a subtle yet significant inaccuracy in the cation parameters that are part of the AMBER-99 forcefield,[7] which is commonly used in simulations of nucleic acids.[6] The AMBER-99 cation parameters are designated as being adapted from Åqvist. This adaptation merits close scrutiny. Åqvist[8] used the geometric mixing rule to generate parameters for both σ (hard sphere diameter) and ϵ (well-depth) to model Lennard-Jones (L-J) interactions between all ion pairs. Conversely, the AMBER forcefield uses Lorentz−Berthelot mixing rules, which preserve geometric mixing for generating ϵ values but switches to arithmetic mixing for generating σ values. The result is the rapid formation of large, unphysical clusters in strong 1:1 electrolytes at concentrations that are well below solubility limits. This can be seen in Figure 1, which shows a snapshot from a simulation of CsCl (400 mm) in TIP3P water using default AMBER-99
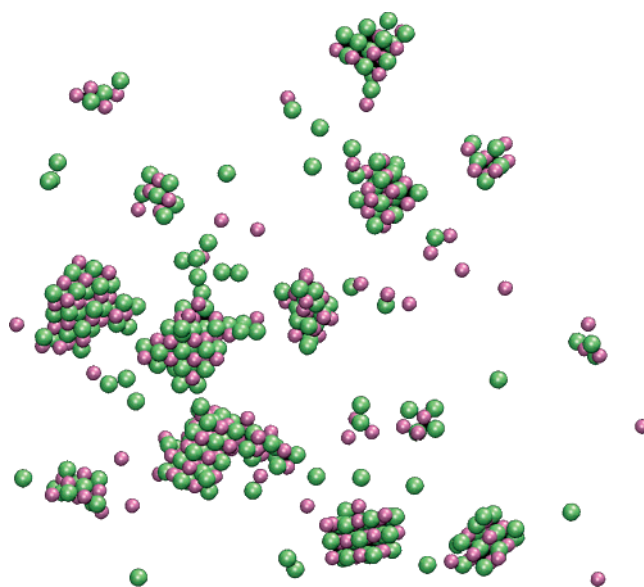

Figure 1. Snapshot from a simulation of 400 mm CsCl solvated in TIP3P water (not shown) using the default AMBER-99 ion parameters under ambient conditions. These large clusters form within a few nanoseconds of simulation time.

parameters. Here, we demonstrate that the conversion between mixing rules is the underlying cause for the observed unphysical behavior.

Recently, we showed that simulations based on Åqvist's original parameters for monovalent ions (as implemented in the OPLS forcefield)[9] reproduce with quantitative accuracy the ion pairing and clustering propensities in strong 1:1 electrolytes.[10] Accuracy was judged based on comparison to estimates of ion

* To whom correspondence should be addressed. E-mail: pappu@wustl.edu.
† Molecular Biophysics Program & Center for Computational Biology.
‡ Department of Biomedical Engineering & Center for Computational Biology.

Letters

J. Phys. Chem. B, Vol. 111, No. 41, 2007 **11885**

TABLE 1: Comparison of Equilibrium Pairing Constants[a]

salt	AMBER-99	OPLS	conductometric measurement
NaCl	0.96 ± 0.14	0.76 ± 0.05	0.82
KCl	13 ± 1.9	0.57 ± 0.04	0.53
CsCl	19 ± 2.9	0.49 ± 0.07	0.62

[a] The pairing equilibrium constant is defined as $[CA]/[C^+][A^-]$ where the quantities in square brackets refer to molal concentrations of pairs (CA), free cations (C), and free anions (A), respectively. Pairing constants from molecular dynamics simulations are compared to conductometric measurements of Fuoss.[11]

TABLE 2: Comparison of Cation−Anion Contact Distances Extracted from RDFs[a]

salt	molten liquid (Å)	crystal lattice (Å)	preferred contact distances obtained from analysis of simulated RDFs	
			AMBER-99 (Å)	OPLS (Å)
NaCl	2.72−2.78	2.76−2.83	**2.74**	2.74
KCl	3.06	3.14−3.19	**3.00**	3.10
CsCl	3.38−3.40	3.48−3.50	**3.25**	3.48

[a] The locations of the first peak of the cation−anion RDFs extracted from simulation are compared with the crystallographic measurements of Shannon[17] and neutron diffraction measurements of molten ionic liquids compiled by Neilson and Adya.[18] The distances extracted from RDFs based on the AMBER-99 parameters are shown in boldface to emphasize the fact that they tend to be smaller than or equivalent to values tabulated for molten liquids, which we view as a lower limit that should not be crossed.

pairing obtained from conductometric measurements.[11] Ion pairing is a balance between favorable unshielded coulomb interactions and unfavorable partial dehydration of ions. For monovalent ions, the former is dictated by the ionic radius, and the latter is determined by the free energy of hydration. Favorable comparison of ion pairing and clustering obtained using Åqvist's parameters suggests that the original parameters correctly capture the tug of war between the two relevant processes. Our goal is to understand the consequences of nonidealities observed in strong 1:1 electrolytes on counterion-dependent conformational and binding equilibria of nucleic acids. This motivated us to assess the accuracy and reliability of ion parameters in the AMBER-99 forcefield, which is the most commonly used forcefield for simulating nucleic acids.[6,12−16]

Results and Discussion

In Table 1, we compare equilibrium constants for ion pairing obtained from simulations of alkali chlorides with the AMBER ion parameters to those from previous work.[10,11] The degree of ion pairing observed using AMBER deviates significantly from previous observations. Similar large deviations were quantified for the formation of triplets and higher-order clusters (data not shown). The magnitude of the discrepancy increases with increase in cation size. Additionally, the overall trend in ion pairing is also incorrect with AMBER. Experiments and our previous results show that the order of preference for ion pairing in alkali chlorides is NaCl > KCl > CsCl and this trend is inverted in simulations based on the AMBER ion parameters.

We also analyzed differences in cation−anion contact distances obtained using Åqvist/OPLS versus AMBER. These data are summarized in Table 2 where we quantify the location of the first peak of cation−anion radial distribution functions (RDFs). Experimental data for cation−anion RDFs are unavailable because alkali chlorides are weak scatterers (compared to water). However, X-ray and neutron scattering data are available for crystal lattices[17] and molten ionic liquids.[18] These data are

useful because they place bounds on what constitutes a reasonable interionic distance in aqueous solution.

Preferred interionic contact distances in solution under ambient conditions should be larger than those observed in molten ionic liquids. Table 2 shows that the preferred interionic distances of aqueous alkali chlorides obtained using Åqvist/OPLS parameters are comparable to those measured in crystals and larger than the corresponding values measured in molten ionic liquids. However, the preferred interionic distances computed using the default AMBER parameters are smaller in KCl and CsCl when compared to those measured in corresponding molten ionic liquids. The results summarized in Tables 1 and 2 indicate that the adaptation of Åqvist's parameters in AMBER is erroneous. We now discuss the source of these errors and propose a solution that should allow the continued use of AMBER-99 nucleic acid parameters.

The procedure for adaptation of Åqvist's parameters is posted on the AMBER website.[19] The Åqvist paradigm is constrained by the experimentally measured free energies of hydration and the location of the first peak in cation−water RDFs. To match Åqvist's numbers, the cation σ values in AMBER were parametrized to reproduce the Lennard-Jones (L-J) pair interaction between the cation and the TIP3P water oxygen atom under Lorentz−Berthelot mixing rules, thereby preserving the σ_{ij} value for one specific interaction. However, the AMBER adaptation does not preserve the σ_{ij} cross types for cation−cation and cation−anion pairs. We show the L-J parameters used for cations and anions in both forcefields in Table 3. The resulting σ_{ij} values for all cross types and the differences between the AMBER and OPLS values are shown in Table 4. Not surprisingly, the σ_{ij} values for cation−TIP3P water oxygen pairs are similar to those in Åqvist/OPLS. Conversely, the σ_{ij} values for all other cross types vary to different degrees with the largest, intermediate, and smallest deviations corresponding to Cs^+, K^+, and Na^+ ions, respectively. The data shown in Table 4 provide a clear explanation for the discrepancy between simulations using the two parameter sets. To complete this discussion, we note that the AMBER forcefield uses parameters developed by Smith and Dang[20] for Cl^- ions, whereas the OPLS forcefield uses parameters developed by Chandresekhar et al.[21] The former is based on the Lorentz−Berthelot mixing rules, and the latter uses geometric mixing rules.

Prescription to Overcome Problems with the AMBER Ion Parameters. For simulations of nucleic acids in strong 1:1 electrolytes, we propose the use of Åqvist/OPLS parameters for ions (as shown in Table 4) in conjunction with AMBER-99 nucleic acid parameters. Although this represents a minimal modification, obvious questions will be raised regarding the use of ion parameters shown in Table 4 with AMBER-99 nucleic acid parameters. We present a preliminary test of our prescription to show that this is a reasonable approach.

Nucleic acids interact with monovalent ions primarily through the negatively charged polyphosphate backbone, resulting in phenomena such as diffuse and specific ion binding. Accordingly, we carried out simulations to test the cation−phosphate interactions obtained using the prescription outlined above. To this end, we carried out simulations of a single dimethylphosphate anion (DMP^-) in 1 m solutions of NaCl and KCl. The DMP^- anion is a small molecule mimic of the phosphate group in the phosphodiester linkage along nucleic acid backbones. The charges and atom types used to model DMP^-, shown in Figure 2, are the AMBER nucleotide backbone parameters.[22] We used the default AMBER-99 parameters for DMP^- interactions with anions, cations, and water molecules. However, for interactions

TABLE 3: Lennard-Jones Parameters Used for Cations and Anions[a]

atom (AMBER-99)	σ (Å)	ϵ (kJ/mol)	atom (OPLS)	σ (Å)	ϵ (kJ/mol)
Na$^+$	3.32840	1.15897×10^{-2}	Na$^+$	3.33045	1.15980×10^{-2}
K$^+$	4.72302	1.37235×10^{-3}	K$^+$	4.93463	1.37235×10^{-3}
Cs$^+$	6.04920	3.3723×10^{-4}	Cs$^+$	6.71600	3.38904×10^{-4}
Cl$^-$	4.40104	4.184×10^{-1}	Cl$^-$	4.41724	4.92833×10^{-1}
O (TIP3P)	3.15061	6.36396×10^{-1}	O (TIP3P)	3.15061	6.36396×10^{-1}
O2 (DMP$^-$)	2.95992	8.7864×10^{-1}	O2	2.96000	8.7864×10^{-1}
OS (DMP$^-$)	3.00001	7.11280×10^{-1}	OS	3.00000	7.11280×10^{-1}

[a] Åqvist's cation parameters as specified by the AMBER-99[7] and OPLS[9] forcefields. The parameters for the TIP3P water model are taken from Jorgensen et al.[31] Note that AMBER adopts the Cl$^-$ of Smith and Dang[20] while OPLS adopts the Cl$^-$ of Chandresekhar et al.[21] The last two rows show the AMBER-99 L-J parameters for the phosphate oxygen atoms of the DMP$^-$ anion.

TABLE 4: Derived Parameters for L-J σ Cross Types Using Default Mixing Rules[a]

interaction	AMBER-99 σ_{ij} (Å)	OPLS σ_{ij} (Å)	$\Delta\sigma_{ij}$ (Å) = $\lvert$OPLS − **AMBER-99**$\rvert$
Na$^+$−O (TIP3P)	3.23950	3.23928	0.00022
K$^+$−O (TIP3P)	3.93680	3.94287	0.00607
Cs$^+$−O (TIP3P)	4.59990	4.59990	0.00000
Cl$^-$−O (TIP3P)	3.77580	3.73054	0.04526
Na$^+$−Cl$^-$	3.86472	3.83554	0.02918
K$^+$−Cl$^-$	4.56203	4.66870	**0.10667**
Cs$^+$−Cl$^-$	5.24465	5.44660	**0.20195**
Na$^+$−Na$^+$	3.32840	3.33045	0.00205
K$^+$−K$^+$	4.72302	4.93463	**0.21161**
Cs$^+$−Cs$^+$	6.04920	6.71600	**0.66680**
Cl$^-$−Cl$^-$	4.40104	4.41724	0.01620
Na$^+$−O2 (DMP$^-$)	3.14416	3.13977	0.00439
K$^+$−O2 (DMP$^-$)	3.84147	3.82185	0.01962
Na$^+$−OS (DMP$^-$)	3.16421	3.16091	0.00330
K$^+$−OS (DMP$^-$)	3.86152	3.84758	0.01394

[a] AMBER-99 parameters are generated with the Lorentz−Berthelot mixing rules (additive σ mean, geometric ϵ Mean). OPLS parameters are generated using geometric means for both σ and ϵ. The most-significant differences are shown in boldface.
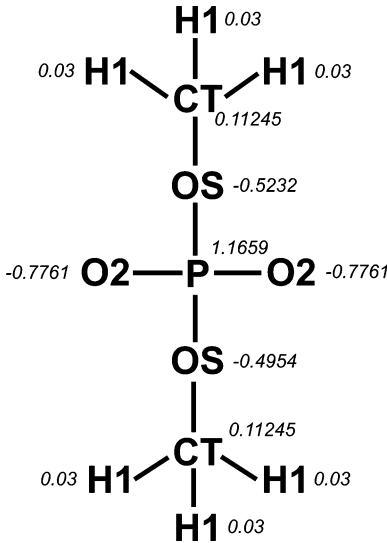

Figure 2. Schematic of the DMP$^-$ anion. Also shown are the AMBER-99 parameters used to model the anion. Charges (italics) and atom types (boldface) are taken verbatim from the parameters of Cornell et al.[22] for the nucleic acid backbone with the addition of capping methyl groups.

between ions in the bulk and between ions and water molecules we used the Åqvist/OPLS parameters. At first glance, this might seem like a curious strategy. However, the dominant interactions between bulk solution ions and the DMP$^-$ anion are likely to involve the phosphate oxygen atoms. The original AMBER-99 σ_{ij} parameters used to model interactions between phosphate
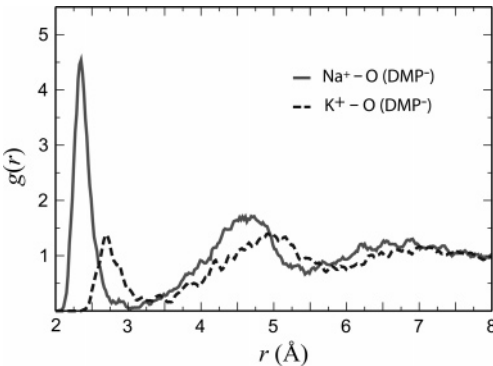

Figure 3. RDFs of Na$^+$ and K$^+$ around the negatively charged oxygen atoms of the dimethylphosphate anion (DMP$^-$), a phosphate backbone mimic. Simulations used the Åqvist/OPLS values for ion−ion and ion−water L-J interactions in conjunction with default AMBER-99 values for the ion−DMP$^-$ interactions.

TABLE 5: Comparison of Simulated Cation−DMP Anion Contact Distances[a]

ion	simulated cation−DMP$^-$ oxygen distance (Å)[b]	crystallographic cation−DMP$^-$ oxygen distance (Å)[23]
Na$^+$	2.35 ± 0.10	2.21/2.41
K$^+$	2.70 ± 0.15	2.43/2.80

[a] X-ray crystal values correspond to the two regions of highest electron density reported by Schneider and Kabelá.[23] [b] Error bars represent standard errors.

oxygen atoms (AMBER atom types O2 and OS in Figure 2) and monovalent cations differ from their OPLS counterparts only in the second decimal place and this comparison is shown in the last three columns of Table 4. Therefore, we propose that the strategy outlined above is likely to be a sufficient, minimalist modification.

Site−site RDFs for cations around the oxygen atoms of the DMP$^-$ anion (O2 atoms in Figure 2) are shown in Figure 3. Locations of the first peaks in cation−DMP$^-$ oxygen atom RDFs may be compared to values tabulated from analyses of high-resolution crystal structures,[23] which have been carried out for Na$^+$ and K$^+$ around the phosphate oxygen atoms (equivalent to the O2 atoms shown in Figure 2) of the DMP$^-$ anion. These comparisons are shown in Table 5. We find that the observed distances for the first peak lies within the bounds prescribed by analysis of small-molecule crystallographic data.[23] Because of the absence of high-resolution crystal structures containing Cs$^+$, we did not simulate CsCl with DMP$^-$. Encouraged by our preliminary result, we suggest that our strategy may be generalizable to simulations of monovalent ions around nucleic acids. Recent simulations of double-stranded DNA in NaCl and KCl showed ion clustering and anion accumulation around the phosphate backbone.[6] Although these results are intriguing, it is possible that the erroneous adaptation of Åqvist's cation

parameters play some role in the observed behavior. If our prescription is accepted as a reasonable strategy, then it would be useful to revisit these types of simulations.

The strategy outlined above is the minimalist fix required to facilitate continued usage of AMBER-99 nucleic acid parameters. It is important to note that the ion parameters in AMBER-99 are identical to those in other versions of the AMBER forcefield including, to the best of our knowledge, all versions of AMBER subsequent to AMBER-94.[22] To facilitate usage of our prescription, we have provided a detailed tabulated inventory of the σ_{ij} and ϵ_{ij} values for all relevant cross types involving ions and AMBER-99 phosphate backbone parameters (Table 4). For all other ion–nucleic acid interactions, we are recommending the continued use of default AMBER-99 parameters as a minimalist perturbation of the AMBER forcefield. Parameter modifications for interactions such as those between negatively charged carboxyl groups on proteins and bulk cations will require additional tests such as those carried out here. Should such tests reveal inaccuracies or inconsistencies, then changes along the lines of our prescription are warranted.

Recently, Auffinger et al.[24] studied ion aggregation of KCl using different combinations of ion forcefields and water models. In their study, the formation of large aggregates was unique to simulations based on the AMBER ion parameters. They conclude that spontaneous ion aggregation in KCl is an artifact of errors in the AMBER ion parameters. Auffinger et al. recommend that the adapted Åqvist ion parameters in AMBER be replaced with the parameters of Dang and Smith.[20] Here, we have suggested a different approach based on two criteria: First, in previous work[10] we showed that nonidealities in strong 1:1 electrolytes obtained using Åqvist's parameters are in agreement with those obtained from conductometry.[11] Second, we have tested the accuracy of our approach by comparing computed ion densities around the DMP^- anion to those obtained from analysis of crystallographic data. We suggest that similar tests need to be carried out for the approach proposed by Auffinger et al.

Methods

All simulations were carried out using version 3.3.1 of the GROMACS molecular dynamic package[25] and the AMBER-99 port of Sorin and Pande[26] (which we cross-checked against the parm99.dat parameter file distributed by the AMBER developers).[7] The isothermal–isobaric ensemble (1 bar, 298 K) was simulated using the weak coupling algorithms of Berendsen and co-workers.[27] The time constants used for manostat and thermostat coupling were 1 and 0.2 ps, respectively. Four simulations of 10-ns length (after an equilibration of 2 ns) were performed for each set of alkali–chloride parameters, each with randomized starting coordinates and velocities. Because the "unfixed" AMBER-99 simulations of CsCl and KCl exhibited rapid lattice formation on the nanosecond time scale at both 1 *m* and 400 m*m* concentrations, pairing constants were extracted from 250 m*m* simulations containing 125 ion pairs and 27750 TIP3P waters in cubic boxes approximately 95 Å on a side. All other simulations (including "fixed" AMBER-99) were simulated at an electrolyte concentration of 1 *m*, with 204 ions of each type and 11 341 TIP3P water molecules in cubic boxes approximately 71 Å on a side. Water bond lengths and angles

were constrained using the SETTLE algorithm.[28] DMP anion simulations contained one additional DMP^- molecule (Figure 2) and one additional neutralizing cation and were otherwise identical to the "fixed" AMBER-99 simulations. Large box sizes were employed to minimize finite size effects caused by periodic boundaries conditions. Long-range electrostatics were treated with the particle-mesh Ewald algorithm.[29] A cutoff of 10 Å was employed for real-space Coulomb interactions, van der Waals interactions, and neighbor list generation. Equations of motion were integrated using the leapfrog algorithm[30] with a time step of 2 fs. Detailed protocols for the pairing/clustering measurements and their associated statistical error bars are provided in previous work.[10]

References and Notes

(1) Record, M. T., Jr.; Anderson, C. F.; Lohman, T. M. *Q. Rev. Biophys.* **1978**, *11*, 103.

(2) Young, M. A.; Jayaram, B.; Beveridge, D. L. *J. Am. Chem. Soc.* **1997**, *119*, 59.

(3) Tan, Z.-J.; Chen, S.-J. *J. Chem. Phys.* **2005**, *122*, 044903.

(4) Ponomarev, S. Y.; Thayer, K. M.; Beveridge, D. L. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 14771.

(5) Cheatham, T. E., III; Kollman, P. A. *Structure* **1997**, *5*, 1297.

(6) Savelyev, A.; Papoian, G. A. *J. Am. Chem. Soc.* **2006**, *128*, 14506.

(7) Wang, J.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049.

(8) Åqvist, J. *J. Phys. Chem.* **1990**, *94*, 8021.

(9) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225.

(10) Chen, A. A.; Pappu, R. V. *J. Phys. Chem. B* **2007**, *111*, 6469.

(11) Fuoss, R. M. *Proc. Natl. Acad. Sci. U.S.A.* **1980**, *77*, 34.

(12) Srinivasan, J.; Cheatham, T. E., III; Cieplak, P.; Kollman, P. A.; Case, D. A. *J. Am. Chem. Soc.* **1998**, *120*, 9401.

(13) Bevan, D. R.; Li, L.; Pedersen, L. G.; Darden, T. A. *Biophys. J.* **2000**, *78*, 668.

(14) Cheatham, T. E., III; Srinivasan, J.; Case, D. A.; Kollman, P. A. *J. Biomol. Struct. Dyn.* **1998**, *16*, 265.

(15) Williams, D. J.; Hall, K. B. *Biophys. J.* **1999**, *76*, 3192.

(16) Showalter, S. A.; Baker, N. A.; Tang, C.; Hall, K. B. *J. Biol. NMR* **2005**, *32*, 179.

(17) Shannon, R. D. *Acta Crystallogr., Sect. A: Found. Crystallogr.* **1976**, *32*, 751.

(18) Neilson, G. W.; Adya, A. K. *Annu. Rep. Prog. Chem., Sect. C: Phys. Chem.* **1997**, *93*, 101.

(19) van der Waals Parameters for Cations in AMBER Format. http://amber.scripps.edu/Questions/vdw.html (accessed July 30, 2007).

(20) Smith, D. E.; Dang, L. X. *J. Chem. Phys.* **1994**, *100*, 3757.

(21) Chandrasekhar, J.; Spellmeyer, D. C.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1984**, *106*, 903.

(22) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.

(23) Schneider, B.; Kabelá, M. *J. Am. Chem. Soc.* **1998**, *120*, 161.

(24) Auffinger, P.; Cheatham, T. E., III; Vaiana, A. C. Series 2007. http://pubs3.acs.org/acs/journals/doilookup?in_doi=10.1021/ct700143s.

(25) van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. *J. Comput. Chem.* **2005**, *26*, 1701.

(26) Sorin, E. J.; Pande, V. S. *Biophys. J.* **2005**, *88*, 2472.

(27) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684.

(28) Miyamoto, S.; Kollman, P. A. *J. Comput. Chem.* **1992**, *13*, 952.

(29) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089.

(30) Hockney, R. W.; Eastwood, J. W. *Computer Simulations Using Particles*; McGraw-Hill: New York, 1981.

(31) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Phys. Chem.* **1983**, *79*, 926.