JCTC Journal of Chemical Theory and Computation

Spontaneous Formation of KCl Aggregates in Biomolecular Simulations: A Force Field Issue?

Pascal Auffinger,*,[†] Thomas E. Cheatham III,[‡] and Andrea C. Vaiana[†,§]

Architecture et Réactivité de l'ARN, Université Louis Pasteur de Strasbourg, CNRS, IBMC, 15 rue René Descartes, 67084 Strasbourg, France, and Department of Medical Chemistry, Pharmaceutical Chemistry and Pharmaceutics and Bioengineering, University of Utah, Salt Lake City, Utah 84112

Abstract: Realistic all-atom simulation of biological systems requires accurate modeling of both the biomolecules and their ionic environment. Recently, ion nucleation phenomena leading to the rapid growth of KCl or NaCl clusters in the vicinity of biomolecular systems have been reported. To better understand this phenomenon, molecular dynamics simulations of KCl aqueous solutions at three (1.0, 0.25, and 0.10 M) concentrations were performed. Two popular water models (TIP3P and SPC/E) and two Lennard-Jones parameter sets (AMBER and Dang) were combined to produce a total of 80 ns of molecular dynamics trajectories. Results suggest that the use of the Dang cation Lennard-Jones parameters instead of those adopted by the AMBER force-field produces a more accurate description of the ionic solution. In the later case, formation of salt aggregates is probably indicative of an artifact resulting from misbalanced force-field parameters. Because similar results were obtained with two different water parameter sets, the simulations exclude a water model dependency in the formation of anomalous ionic clusters. Overall, the results strongly suggest that for accurate modeling of ions in biomolecular systems, great care should be taken in choosing balanced ionic parameters even when using the most popular force-fields. These results invite a reexamination of older data obtained using available force-fields and a thorough check of the quality of current parameters sets by performing simulations at finite (>0.25 M) instead of minimal salt conditions.

Introduction

Biomolecular systems are surrounded by solvent particles (including water molecules, cations, and anions), and this environment modulates to a significant degree the physico-chemical properties of these systems.[1] Theoretical methods such as molecular dynamics (MD) simulations are often used to gain microscopic insight into the complex interplay of interactions between biomolecular species and solvent particles. These methods use empirical force fields specifically

developed and validated by extensive use of experimental and high-level ab initio computational data. Given the importance of the various ionic species surrounding biomolecular systems, a significant effort has been put into fine-tuning various sets of Lennard-Jones (LJ) parameters for monovalent ions such as Na^+, K^+, and Cl (for example, see refs 2−6). These parameters have subsequently been included in major force-fields. Recently, some of these parameters, in conjunction with a choice of water models, have been evaluated by comparison of a large array of calculated structural and thermodynamic properties.[7] The authors note that the use of different parametrizations leads to a *large* dispersion of calculated properties resulting mainly from incomplete experimental knowledge of the structural features of ionic aqueous solutions at finite molarity. Therefore, they

* Corresponding author phone: (33) 388 41 70 49; fax: (33) 388 60 22 18; e-mail: p.auffinger@ibmc.u-strasbg.fr.

[†] Université Louis Pasteur de Strasbourg.

[‡] University of Utah.

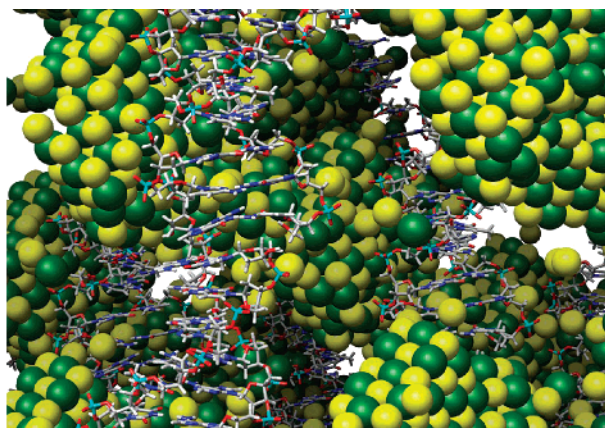[§] Present address: Los Alamos National Laboratory, MS K710, Los Alamos, NM 87545.

Table 1. Characteristic Simulation Parameters

	Amber_ TIP3P	Amber_ SPC/E	Dang_ TIP3P	Dang_ SPC/E
~1.0 M 145 KCl pairs 7568 H$_2$O	5 ns	5 ns	5 ns	5 ns
~0.25 M 36 KCl pairs 7785 H$_2$O	5 ns	5 ns	5 ns	5 ns
~0.10 M 15 KCl pairs 7827 H$_2$O	10 ns	10 ns	10 ns	10 ns

Figure 1. Spontaneous formation of NaCl aggregates in a simulation of a d(CGCGAATTCGCG)$_2$ duplex in 4 M salt solution using the AMBER adopted ion parameters and the TIP3P model. Shown are the unit cell (omitting the water) and images ±1 unit cell in each direction. The sodium and chloride ions are yellow and green, respectively.

Table 2. Lennard-Jones Parameters (r^* and ϵ) and Partial Charges (q) for the Water and Ion Models[a]

	model	q^b	r^* (Å)^c	ϵ (kcal/mol) ^{c}
water^d	TIP3P	−0.8340	1.7683	0.1520
	SPC/E	−0.8476	1.7766	0.1553
K$^+$	Amber	+1	2.6580	0.000328
	Dang	+1	1.8687	0.100000
Cl$^-$	Amber	−1	2.4700	0.10
	Dang	−1	2.4700	0.10

[a] Note That the AMBER and Dang Parameters for the Cl$^-$ Ion Are Identical. [b] Partial charge for the oxygen atom of the water model and the monovalent ions. [c] r^* corresponds to the position of the Lennard-Jones minimum, and ϵ corresponds to the depth of this minimum. [d] For the TIP3P and SPC/E models, the OW−HW and HW−HW distances are constrained to 0.9572 and 1.5136 Å and to 1.0000 and 1.6330 Å, respectively.

did not reach a conclusive ranking of the investigated models. Indeed, finding criteria that allow one to select the most appropriate models and to unambiguously discard defective parameter sets are not straightforward unless some clear-cut artifacts can be identified (see for example, early studies that demonstrated the limits of truncation methods in the calculation of electrostatic interactions).[8] Similar artifacts are likely present in several recently published studies that in some cases reported the formation of salt aggregates in the vicinity of biomolecular systems.[9−13] For example, in a series of simulations exploring "A to B" DNA transitions in >1.0 M NaCl salt solution with AMBER force fields, clear formation of NaCl aggregates were observed (Figure 1; for computational details, see the Supporting Information). In fact, spontaneous and systematic formation of salt aggregates at concentrations around and below 1.0 M is not expected in NaCl and KCl electrolytes (the experimental solubility limits at 20 °C for KCl and NaCl are around 3.2[14,15] and 5.4 mol/L,[14] respectively). Interestingly, all these biomolecular simulations make use of the AMBER force-fields.[16]

To identify the parameters that may be involved in this atypical aggregation process, we performed MD simulations of model systems of aqueous KCl solutions at three different concentrations (1.0, 0.25, and 0.10 M) using two popular water models (TIP3P and SPC/E), as well as two Lennard-Jones (LJ) parameter sets for the K$^+$ cation. One of these parameter sets (Åqvist)[2] is widely used as a part of the parm99 force-field (and all earlier versions) delivered with the AMBER package.[16] The other is derived from the Dang and Kollman's work[17] and has been extensively tested in our MD investigations on nucleic acid fragments,[1,18−21] as well as in studies from other groups.

In this paper, we show that the monovalent cation parameters[2] that are part of the AMBER force-field are involved in the observed aggregation phenomena and that the chosen water model has no impact on the manifestation of this artifact. Hence, we suggest that current AMBER-adopted Åqvist parameters should no longer be used for

simulation of ionic solutions because they may affect to an unknown degree the physicochemical properties of the investigated system. Instead, we observed that the K$^+$ parameters of Dang et al.[17] prevent the formation of salt aggregates. Hence, those or similar parameters should be more thoroughly tested and, if considered appropriate, used in replacement of the ones integrated in AMBER that are clearly imbalanced and not adapted for conducting long MD simulations.

Computational Methods

Twelve molecular dynamics (MD) simulations of aqueous KCl solutions at different ionic strength (1.0, 0.25, and 0.10 M) totaling 80 ns, each on a 5−10 ns scale, were carried out (Table 1). Two water models, TIP3P and SPC/E,[22] as well as two parameter sets for the K$^+$ cation, were used (Table 2). The first set, which contains K$^+$ parameters adapted from the work of Åqvist,[2] is extracted from the AMBER force-field.[23] The second set has been optimized for the SPC/E water model and is extracted from a work of Dang and Kollman.[24] The parameters for the Cl$^-$ anions, which have been used along with the SPC/E water model, are derived from the work of Smith and Dang.[3] Interestingly, these chloride parameters are implemented in the AMBER force field, although they have been adjusted to match the SPC/E (and not the TIP3P) water model.[3] The simulations performed here are named after the type of K$^+$ parameters (AMBER or Dang) and water models (TIP3P or SPC/E) that were used (see Table 1). Note that in the following, AMBER parameters refer to the Åqvist monovalent cation parameters adopted by the all AMBER force-field versions.
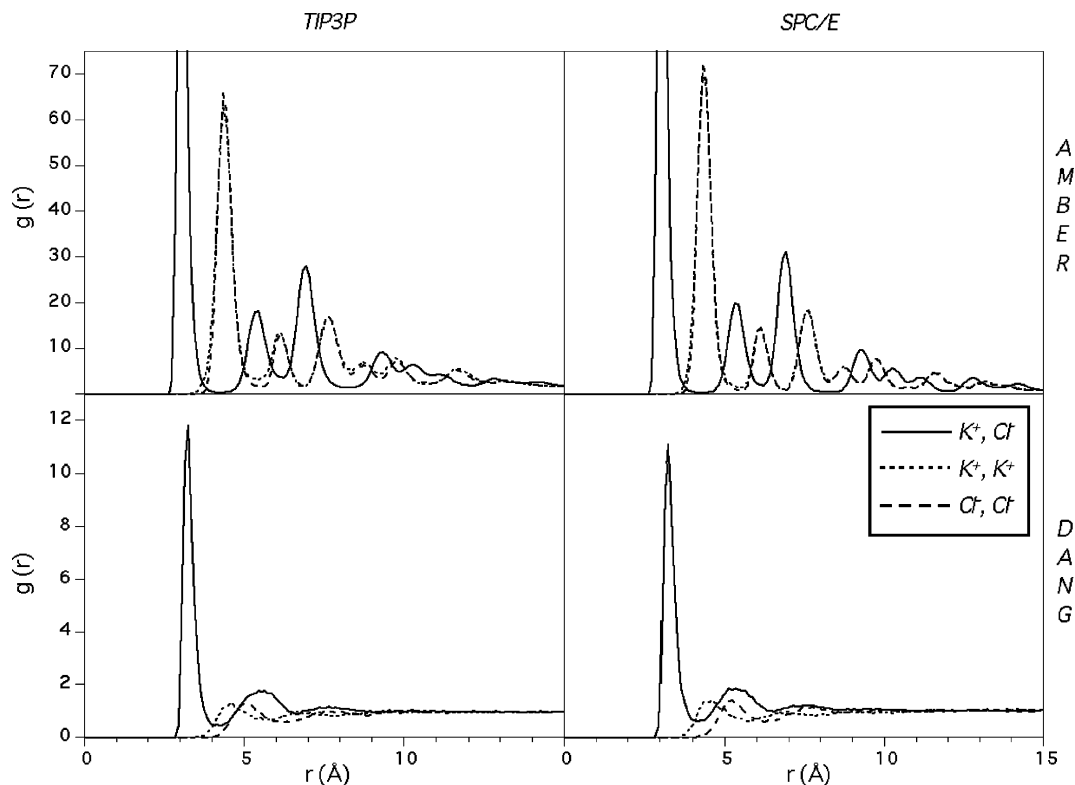
Ionic Aggregation in Biomolecular Simulations

J. Chem. Theory Comput., Vol. 3, No. 5, 2007 **1853**

Figure 2. Ion−ion radial distribution functions calculated over the last 200 ps of the 5 ns long MD trajectories of 1.0 M KCl aqueous solutions generated by using the AMBER (top) and Dang (bottom) parameters for the K^+ cation and the TIP3P (left) and SPC/E (right) water parameters.

All simulations were run at constant temperature (298 K) and pressure (1 atm = 101.325 Pa) by using the PMEMD module of the AMBER 8.0 simulation package.[16,25] This module treats the long-range electrostatic interactions[26] with the particle mesh Ewald (PME) summation method. The chosen charge grid spacing is close to 1.0 Å, and a cubic interpolation scheme was used. A cutoff of 9 Å for the Lennard-Jones interaction and the Berendsen temperature-coupling scheme with a time constant of 0.4 ps were used. The trajectories were run with a 2 fs time step. The lengths of the simulations conducted with different parameter sets and at various ionic concentrations are given in Table 1. Note that the simulation times were doubled (from 5 to 10 ns) for the 0.10 M electrolyte solution to ensure a better sampling of the configurational space. The ions were initially placed such that no two ions were closer than 8, 12, or 16 Å in the 1.0, 0.25, and 0.10 M setups, respectively. In this manner, the ions are at the beginning of the simulations uniformly distributed in the simulation box (a cube with a ~62 Å edge). The number of water molecules per ion is ~26, ~108, and ~260 at the 1.0, 0.25, and 0.10 M concentrations, respectively.

Results

Ion−Ion Radial Distribution Functions at High Salt Concentration. The ion−ion radial distribution functions (RDF) derived from the *Amber_TIP3P_1.0M* and *Amber_SPC/E_1.0M* simulations calculated over the last 200 ps of the 5 ns long trajectories display comparable regular patterns indicative of a highly ordered ionic structure (Figure 2). A visual examination of the MD trajectories allows the

clear identification of a rapid aggregation process that leads to the formation of very large KCl clusters: after 5 ns of simulation, the largest of these clusters comprises more than 100 ions (or 50 KCl ion pairs), while a total of less than 10 ions remain unpaired (Figure 3 and Supporting Information). These clustered ions are arranged in a three-dimensional face-centered cubic lattice, typical of KCl, NaCl, or LiF[27] crystals. In these clusters, each ion is, on the average, surrounded by ~4 ions of opposite charge and by ~6 ions carrying the same charge (Table 3). Another characteristic of this crystal-like ionic arrangement resides in comparable cation−cation and anion−anion radial distribution functions (Figure 2). Indeed, one does not expected to find such ordered structures in a liquid phase.

A glimpse into the dynamics of formation of these "nanocrystals" is given by the K−Cl radial distribution functions calculated over four different 500 ps time intervals for the *Amber_TIP3P_1.0M* simulation (Figure 4). They indicate that, after 5 ns, the ionic aggregates are still growing, suggesting that no equilibrium was reached at this point. Hence, it can be inferred that for sufficiently long simulation times, all 290 ions present in the simulation box will aggregate and form a single nanocrystal. The profile of the RDF calculated over the first 200 ps is also remarkable in that it clearly indicates that the artifactual formation of ionic aggregates is difficult to observe in subnanosecond MD simulations.

On the contrary, no salt aggregation is observed in the *Dang_TIP3P_1.0M* and *Dang_SPC/E_1.0M* simulations. Here, the calculated ion−ion RDF's are close to what is expected for a simulation of a dissociated salt solution
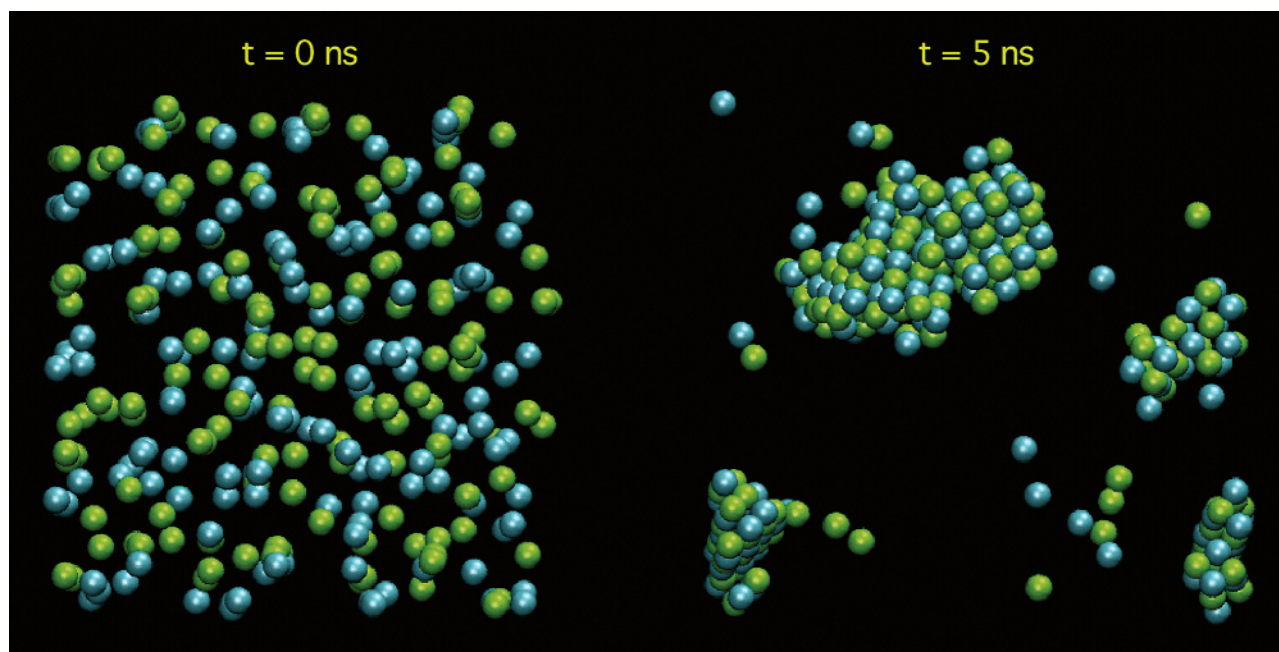
Figure 3. Initial (left) and final (right) configuration of the *Amber_TIP3P_1.0M* simulation. The K^+ and Cl^- ions are shown in green and cyan, respectively. For clarity, water molecules are not shown.

Table 3. First Maxima (r_{max}) of the Ion−Ion Radial Distribution Functions (See Figure 2) and Average Number of Ions (n) Present in Their First Coordination Shell

		K−K		K−Cl		Cl−Cl	
		r_{max} (Å)	n	r_{max} (Å)	n	r_{max} (Å)	n
1.0 M[a]	Amber_TIP3P	4.3	5.8	3.0	3.8	4.3	5.6
	Amber_SPC/E	4.3	5.9	3.0	3.9	4.3	5.7
	Dang_TIP3P	4.5	0.3	3.2	0.4	5.0	0.4
	Dang_SPC/E	4.6	0.3	3.2	0.4	5.2	0.4
0.25 M[a]	Amber_TIP3P	4.3	1.7	3.0	1.6	4.3	1.6
	Amber_SPC/E	4.2	1.1	3.0	1.3	4.3	1.1
	Dang_TIP3P	4.5	<0.1	3.2	0.1	5.3	<0.1
	Dang_SPC/E	4.5	<0.1	3.2	0.1	5.2	<0.1
0.10 M[b]	Amber_TIP3P	4.2	0.3	3.0	0.6	4.3	0.2
	Amber_SPC/E	4.3	0.1	3.0	0.5	4.8	0.1
	Dang_TIP3P	4.2	<0.1	3.2	0.1	5.1	<0.1
	Dang_SPC/E	4.5	<0.1	3.2	0.1	5.2	<0.1

[a] These values (1.0 and 0.25 M) have been calculated by using the last 500 ps of the 5 ns trajectories. [b] These values (0.10 M) have been calculated by using the last ns of the 10 ns trajectories.

(Figure 2). They display first and second peaks revealing the transient formation of contact and water-mediated ion pairs. The RDF's calculated at different time intervals are almost indistinguishable suggesting that the simulations have rapidly converged with respect to the distribution of ions in the simulation box (Figure 4). No formation of ion aggregates can be observed (Supporting Information). Interestingly, the average ion−ion contact distances are larger here than in the simulations where salt aggregates were observed. The average K^+ to K^+, K^+ to Cl^-, and Cl^- to Cl^- distances are increased by ∼0.2, ∼0.2, and ∼0.8 Å, respectively, with only ∼0.4 ions of the same and of opposite charge present in their first coordination shell (Table 3). These numbers suggest that, with the Dang parameters, almost no K−K and Cl−Cl contact ion pairs are formed even at the high 1.0 M salt
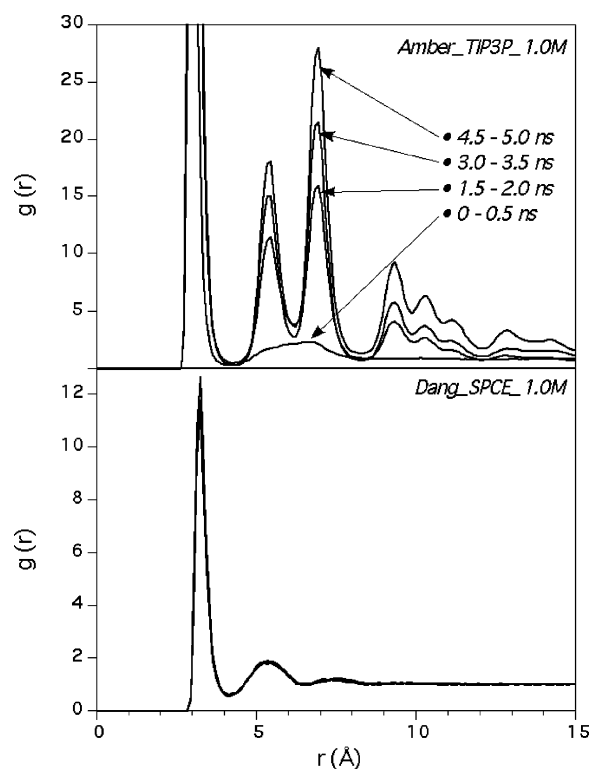

Figure 4. K−Cl radial distribution functions calculated over four 500 ps time intervals (see top panel) of the 5 ns long *Amber_TIP3P_1.0M* (top) and *Dang_SPC/E_1.0M* (bottom) trajectories.

concentration. The results appear to be strongly dependent on the type of parameters used for the K^+ ions but rather insensitive to the water model chosen. The dependence on the Cl^- parameters has not been explicitly evaluated here because we believe that these parameters were derived in a more consistent manner and that they work well in simulation of both salt solution and biomolecules and because there is
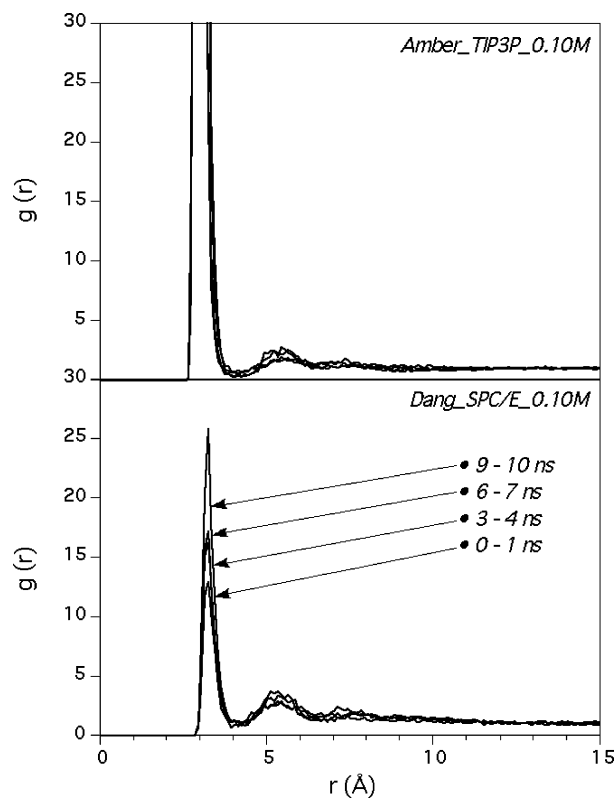
Ionic Aggregation in Biomolecular Simulations

J. Chem. Theory Comput., Vol. 3, No. 5, 2007 **1855**

Figure 5. K−Cl radial distribution functions calculated over four 1 ns time intervals (marked in the bottom panel) of the *Amber_TIP3P_0.10M* (top) and *Dang_SPC/E_0.01M* (bottom) 10 ns long MD trajectories.

fairly good consensus on the use of the Smith and Dang parameters.[3]

Medium to Low Ionic Concentrations. At 1.0 M, the interpretation of the collected data is unambiguous. At 0.25 M, the formation of KCl aggregates, though less dramatic, is still clearly observable on the 5 ns time scale. However, at the low salt concentration (0.10 M) ion aggregation is not observed. Small clusters composed of up to five ions form and disaggregate relatively rapidly (Supporting Information). Indeed, at 0.10 M, it is more difficult to identify the formation of aggregates from a visual inspection of the trajectories because only 15 ion pairs are present in the simulation cell. Yet, the ion−ion RDF's calculated over four different 1 ns time intervals reveal a clear dependence on the type of K$^+$ parameters that were used (Figure 5). This is most clearly revealed by the height of the first peaks and associated coordination numbers. For example, the numbers of ions of opposite charge surrounding a given ion are five times larger when the AMBER rather than the *Dang* parameters are used (∼0.5 instead of ∼0.1), revealing an increased occurrence of KCl contact pairs (Table 3). The time evolution of the RDF's calculated over four different 1 ns time intervals for the *Dang_SPC/E_0.10M* trajectories (Figure 5) suggest that these simulations have not converged over the 10 ns time scale. Yet, for such diluted solutions, a rather slow convergence rate is expected. On the other hand, these results could be indicative of a phase transition associated with the AMBER parameters with a critical concentration between 0.10 (no aggregation) and 0.25 M (aggregation).

Table 4. First Maxima (r_{max}) of the Ion−Water Radial Distribution Functions and First Coordination Shell Ion Hydration Number (n) Calculated for the Four 0.10 M Simulations[a]

	K−O$_w$		Cl−O$_w$	
	r_{max} (Å)	n	r_{max} (Å)	n
Amber_TIP3P_ (0.10M/1.0M)	2.73/2.74	6.2/2.2	3.23/3.24	7.0/2.4
Amber_SPC/E_ (0.10M/1.0M)	2.75/2.75	5.6/2.1	3.23/3.23	6.1/2.2
Dang_TIP3P_ (0.10M/1.0M)	2.82/282	7.2/6.9	3.24/3.24	7.6/7.1
Dang_SPC/E_ (0.10M/1.0M)	2.83/2.82	7.1/6.8	3.24/3.22	7.1/6.8
experimental	2.8	6.0−8.0	3.2	6.0−8.0

[a] Experimental values are derived from refs 28 and 29.

Ion−Water Coordination Numbers. The ion−water coordination numbers have been determined for the four simulations conducted at 0.10 M KCl (Table 4). The calculated K−O$_w$ coordination distances, although slightly different (AMBER ≈ 2.74 Å; *Dang* ≈ 2.83 Å), are close to the experimental consensus value of 2.8 Å.[28,29] Not surprisingly, the calculated Cl−O$_w$ coordination distances are identical in all simulations because the same parameters for Cl$^-$ were used in all of them. The calculated 3.23 Å value is close to the experimental consensus value of 3.2 Å.[28,29] For both ions, the number of water molecules located in the first shell is larger when the *Dang* parameters are chosen, reflecting again the fact that the parameters adopted by AMBER favor the formation of ion pairs and aggregates. Interestingly, at higher concentrations the ion hydration number decreases as expected.

Discussion and Conclusion

AMBER Lennard-Jones Parameters for K$^+$ Favor the Rapid Formation of KCl Aggregates. The formation of salt aggregates in MD simulations of biomolecular systems has already been described in a few studies using KCl or NaCl salts.[9−13] However, these studies did not identify the origin of this phenomenon. Our investigation reveals that the Lennard-Jones parameters for the K$^+$ cation extracted from the AMBER force field[23] and derived from an early parametrization study[2] are likely at the origin of a rapid, irreversible, and unnatural formation of KCl aggregates at high (1.0 M), as well as near physiological (0.25 M), salt concentration. In addition, the simulations clearly show that the observed aggregation behavior is not dependent on the properties of two of the most widely used rigid water models (TIP3P or SPC/E).

To the best of our knowledge, no biomolecular simulations based on "non-AMBER" force-fields have reported such artifactual behavior. Feig and Pettitt,[30] who investigated the distribution of sodium and chlorine ions around DNA duplexes by comparing the AMBER and CHARMM force fields, used parameters developed by Roux[4] for the Na$^+$ and Cl$^-$ ions, and did not report any strange behavior in the ionic atmosphere. Similarly, K$^+$ parameters extracted from a study by Dang and Kollman[24] did not lead, in this and earlier

nanosecond long simulations from our group,[12,18,20,21,31,32] to detectable aggregation artifacts, while the use of the AMBER parameters resulted in rapid aggregation of K^+ and Cl^- particles.[12]

A large ensemble of MD studies of aqueous ionic solutions using various parameter sets and particle mesh Ewald (PME) summation methods for the treatment of long-range electrostatic interactions have been published, including simulations of LiF,[27] LiCl,[33−35] NaCl,[7,36−45] KCl,[15,40,45,46] RbCl,[35,45] CsCl,[45] NaBr, KBr, RbBr, CsBr,[45] and CsI.[35] Polarizable force-fields have also been used in other force-fields.[3,44,47−50] Among all these simulations, the use of the Smith and Dang[3] parameters is quite popular (at least for NaCl salts). As reported here, no ion aggregation has been reported in simulations using the Dang parameters even under high and supersaturated salt conditions.[38,42,43] AMBER parameters are rarely used in simulations of ionic aqueous solutions,[7,45] while they are recurrently used in MD simulations of biomolecular systems.[9−13] In one study, a comparison of calculated properties of a ~1.0 M NaCl aqueous solution generated by using six different parameter sets revealed some level of aggregation for various force-fields including AMBER and GROMOS, while as expected, the Dang parameters did not lead to any detectable formation of ion clusters during 2 ns MD simulations.[7] In another study,[45] the GROMACS program[51] was used to simulate NaCl to CsCl and NaBr to CsBr aqueous salts at various concentrations ranging from 0.10 to 1.0 M. The Åqvist parameters were used for the Na^+, K^+, Rb^+, and Cs^+ cations and different parameters were used for the Cl^{-52} and Br^{-53} anions. The authors reported the formation of ion clusters for all salts at 1.0 M, but not at 0.10 M, in agreement with our own data. However, these clusters that comprise approximately one-third of all ions present in solution appear to be in rapid equilibrium with dissociated ions. Since the formation of ion aggregates was apparently not as dramatic as the one we observed in simulations conducted with the AMBER program and force-fields, these cluster formations were considered as representative of a nonideal behavior observed at the higher ionic concentrations.

Ionic aggregation was also observed in MD simulations of LiF,[27] LiCl,[33] and NaCl[36,37] at 1.0 M concentrations and above. The authors of these studies used self-developed[15,27,33] or GROMOS-adapted parameters[36,37] for the ions. For a 1.0 M solution of LiF, a phase separation was observed. The resulting data indicated that all ions had formed a large and unique cluster geometrically described as a face-centered cubic lattice, the same crystalline structure as that exhibited by LiF, NaCl, or KCl. Smaller clusters were observed in NaCl simulations, mainly, because simulation times below 0.5 ns limited the full formation of aggregates.[36,37] With self-developed parameters, ionic association in 1.0 M of KCl was still observed but was considered to be weak.[15]

More generally, from an experimental point of view, it can be stated that in dilute electrolyte solutions the tendency to aggregate is counterbalanced by thermal fluctuations. Above the saturation point, however, the number of water molecules per ion pair becomes too small to prevent initial ion nucleation followed in most cases by crystallization.

From a theoretical point of view, instead, it appears that the interatomic potentials must be correctly balanced to reproduce these subtle equilibria. Any imbalance would lead to observable microscopic catastrophes such as physically improbable aggregation processes. Correct parametrization of three component systems (water, cation, anion) is certainly not straightforward because it involves the fine-tuning of ten different intermolecular potentials. The AMBER potentials by Åqvist[2] were obtained by fitting to experimental free energies of ion hydration, whereas those by Dang were constructed by fitting to gas-phase binding enthalpy data. A recent study devoted to the calculation of ion−ion potential of mean forces also addressed the respective qualities of the Åqvist (AMBER) and Dang models.[54] According to this author, "the Na^+ Lennard-Jones (LJ) parameters of Dang and Åqvist differ considerably with respect to each other. Thus, if only one experimental property is used to determine the LJ parameters, the determined LJ parameters become not necessarily unique. Hence, LJ parameters should ideally be optimized with respect to independent experimental properties to narrow down the ambiguity in the assessment of their values".

Are Long Simulation Times Needed to Detect Artifacts? Short simulation times may lead to insufficient equilibration of the ionic atmosphere surrounding biomolecules. To achieve a "significant" level of equilibration, simulation times of tens of nanoseconds[55,56] and up to ~500 ns[10] were suggested for the monovalent cation distribution within DNA grooves to converge. Indeed, short simulation times may significantly complicate the detection of ionic aggregation, as well as other potential artifacts that may only manifest themselves on the longer timescales because of accumulation of errors during the MD runs.[57] Yet, convergence times strongly depend on the type of properties and system investigated. For example, convergence of the ion−ion radial distribution functions is achieved in less than 1 ns for the *Dang_1.0M* simulations (Figure 4), while convergence of the same properties for the *Dang_0.10M* is not achieved after 10 ns (Figure 5). Similarly, ion aggregation is observable already after 0.5 ns for the *Amber_1.0M* simulations, while it is very difficult to detect this phenomenon in simulations conducted at low concentration (see Movies S1 and S2). Indeed, the fastest equilibration times are probably obtained for the most homogeneous systems, (i.e., highly concentrated ionic solutions or pure water systems). On the other hand, equilibration is difficult to achieve for highly diluted electrolytes.[41,58] An extreme case of dilute solutions is related to "minimal salt conditions" and will be discussed in the following section.

Minimal Salt Strategies: Implications for Biomolecular Simulations. Salt effects should be taken into account with the greatest possible accuracy in MD simulations of biomolecular systems. This is especially true for highly charged nucleic acids. However, MD simulations of nucleic acid systems taking into account a complete representation of the salt environment are relatively infrequent (especially among AMBER users) because it is generally believed that the Lennard-Jones parameters for monovalent cations are more reliable than those for the highly polarizable chloride anion

Ionic Aggregation in Biomolecular Simulations

J. Chem. Theory Comput., Vol. 3, No. 5, 2007 **1857**

(see ref 59 and associated Supporting Information). Hence, a minimal salt strategy, in which only charge neutralizing cations are taken into account, was used by many groups to prevent the occurrence of anion related artifacts (see ref 1). Although such a strategy seems at first sight reasonable, its use was not based on a precise evaluation of the reliability of available parameter sets. The present investigation suggest that, at least for AMBER users and probably also for users of other force-fields (see ref 7), a misbalance in the ionic Lennard-Jones parameters is at the origin of the serious ionic aggregation problems described above that might affect to an unknown degree the quality of the generated MD trajectories. This misbalance might have its roots in the Lennard-Jones parameters for monovalent cations (Na^+, K^+, ...) as suggested by our data. Consequently, the community is strongly encouraged to reevaluate all the data collected using MD simulations of biomolecular systems in monovalent salt solution, especially data using the default parameters supplied with AMBER. Of particular concern are data related to the interaction of monovalent cations with nucleic acid bases (and other biomolecular fragments) because the $M^+ \cdot \cdot \cdot O/N$ interactions are certainly affected to an unknown degree by the use of misbalanced ionic parameters. On the other hand, recent MD simulations successfully reproduced the nucleic acid anion binding sites observed in crystallographic data,[19] suggesting that the Dang Cl^- parameters, although certainly far from being perfect, can be used in MD simulations to reproduce salient features of the ionic atmosphere surrounding biomolecules.

Possible Application to Nucleation Studies. Interestingly, alteration of Lennard-Jones parameters has been used to initiate a nucleation process for NaCl that was subsequently investigated by using the path sampling method developed by Chandler and co-workers.[60] The authors modified the ion−water interactions to obtain an artificial system that crystallizes in a few tens of picoseconds. Hence, nucleation could be studied from simulations generated by using AMBER parameters. Although such trajectories do not correspond to realistic models of ionic solutions, this approach may still be used for getting insight into nucleation phenomena. This is especially true in view of the fact that the ion clusters seem to adopt the same ion ordering as in the crystalline state. Furthermore, phase transition points, such as those occurring at concentrations between 0.10 and 0.25 M in KCl aqueous solutions could be characterized from such MD simulations.[61] Indeed, dissolution of NaCl clusters or nucleation at an NaCl/water interface have been investigated by using the AMBER force-field parameters.[62,63] However, this is not the scope of the present investigation.

Which Monovalent Cation Parameter Sets Should be Used?. Unfortunately, no straightforward answer to this question can be provided. Current parameter sets are far from perfect and suffer from various drawbacks. In 1996, for instance, Lyubartsev and Laaksonen noted "reported RDF (Radial Distribution Function) or PMF (Potential of Mean Force) results appear to be quite often in contradiction to each other and show an apparent dependence on the used model. A general picture arises in several works: the anion-cation potential of mean force has usually two minima, first corresponding to the contact ion pairs (CIP configuration) and the second corresponding to the solvent separated ion pairs (SSIP configuration). However, the intensities, i.e., the relative importance of these minima, vary largely from work to work."[38] Patra and Karttunen[7] reached the same conclusion by analyzing MD simulations of aqueous NaCl obtained by using six different ion parameter sets and four different water models and concluded that the observed uncertainty in calculated data reflects our current fragmentary experimental knowledge of the structural properties of ionic solutions at finite molarity.

It is not the scope of this study to develop new parameter sets. Nevertheless, on the basis of our data, it can be suggested that it would be worth abandoning ionic models that display any propensity toward anomalous aggregation (for instance AMBER; see ref 7) in favor of those leading to an "appropriate level" of dissociation (Dang), as spontaneous ion aggregation is not expected for molar aqueous solutions of KCl or NaCl salts. Control simulations with other alkali cation models (Li^+, Na^+, Rb^+, or Cs^+) included in the AMBER force-field were not performed. But we suspect that these parameters suffer from the same flaws because they have been parametrized in a similar manner.[45] A reevaluation of the performances of all available parameters in the context of three component electrolyte solutions should be undertaken with a special emphasis on this aggregation issue in the framework of a recent proposal devoted to create a set of descriptive parameters and measures allowing us to judge the "quality" and reliability of MD simulations.[64]

In conclusion, the combination of more precise experimental and theoretical studies will lead to a generation of force-fields free from such imbalanced interatomic potential terms. In this respect, polarizable force-fields will certainly be key players in allowing the generation of more accurate and informative biomolecular simulations.[44,47,49,50,65] Finally, the issues discussed above are not limited to nucleic acid systems but are also relevant for all other biomolecular systems including proteins, membranes, and ion channels,[66] for which the electrolytic environment plays a determining role.

Supporting Information Available: Computational details related to Figure 1 and two movies showing MD trajectories of 1.0 and 0.10 M KCl aqueous electrolytes generated with the Åqvist (AMBER) and Dang parameters. This material is available free of charge via the Internet at http://pubs.acs.org.

1858 *J. Chem. Theory Comput., Vol. 3, No. 5, 2007*

Auffinger et al.

References

(1) Auffinger, P.; Hashem, Y., Nucleic acid solvation: From outside to insight. *Curr. Opin. Struct. Biol.* **2007**, in press.

(2) Åqvist, J. Ion−water interaction potentials derived from free energy perturbation simulations. *J. Phys. Chem.* **1990**, *94*, 8021−8024.

(3) Smith, D. E.; Dang, L. X. Computer simulations of NaCl association in polarizable water. *J. Chem. Phys.* **1994**, *100*, 3757−3765.

(4) Roux, B.; Prod'hom, B.; Karplus, M. Ion transport in the gramicidin channel: Molecular dynamics study of single and double occupancy. *Biophys. J.* **1995**, *68*, 876−892.

(5) Peng, Z.; Ewig, C. S.; Hwang, M. J.; Waldman, M.; Hagler, A. T. Derivation of class II force fields. 4. van der Waals parameters of alkali metal cations and halide anions. *J. Phys. Chem. A* **1997**, *101*, 7243−7252.

(6) Jensen, K. P.; Jorgensen, W. L. Halide, Ammonium, and alkali metal ion parameters for modeling aqueous solutions. *J. Chem. Theory Comput.* **2006**, *2*, 1499−1509.

(7) Patra, M.; Karttunen, M. Systematic comparison of force fields for microscopic simulations of NaCl in aqueous solutions: Diffusion, free energy of hydration, and structural properties. *J. Comput. Chem.* **2004**, *25*, 678−689.

(8) Auffinger, P.; Beveridge, D. L. A simple test for evaluating the truncation effects in simulation of systems involving charged groups. *Chem. Phys. Lett.* **1995**, *234*, 413−415.

(9) Mazur, A. K. Titration in silico of reversible B ↔ A transitions in DNA. *J. Am. Chem. Soc.* **2003**, *125*, 7849−7859.

(10) Cheatham, T. E. Simulations and modeling of nucleic acid stucture, dynamics and interactions. *Curr. Opin. Struct. Biol.* **2004**, *14*, 360−367.

(11) Savelyev, A.; Papoian, G. A. Electrostatic, steric, and hydration interactions favor Na(+) condensation around DNA compared with K(+). *J. Am. Chem. Soc.* **2006**, *128*, 14506−14518.

(12) Vaiana, A. C.; Westhof, E.; Auffinger, P. A molecular dynamics simulation study of an aminoglycoside/A-site RNA complex: Conformational and hydration patterns. *Biochimie* **2006**, *88*, 1061−1073.

(13) Savelyev, A.; Papoian, G. A. Inter-DNA electrostatics from explicit solvent molecular dynamics simulations. *J. Am. Chem. Soc.* **2007**, *129*, 6060−6061.

(14) Stephen, H.; Stephen, T. *Solubilities of Inorganic and Organic Compounds*; Macmillan: New York, 1963.

(15) Vieira, D. S.; Degrève, L. Molecular simulation of a concentrated aqueous KCl solution. *J. Mol. Struct.* **2002**, *580*, 127−135.

(16) Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber biomolecular simulation programs. *J. Comput. Chem.* **2005**, *26*, 1668−88.

(17) Dang, L. X.; Kollman, P. A. Free energy of association of the K+/18-crown-6 complex in water: a new molecular dynamics study. *J. Phys. Chem.* **1995**, *99*, 55−58.

(18) Auffinger, P.; Westhof, E. Water and ion binding around RNA and DNA (C,G)-oligomers. *J. Mol. Biol.* **2000**, *300*, 1113−1131.

(19) Auffinger, P.; Bielecki, L.; Westhof, E. Anion binding to nucleic acids. *Structure* **2004**, *12*, 379−388.

(20) Auffinger, P.; Westhof, E. Water and ion binding around r(UpA)$_{12}$ and d(TpA)$_{12}$ oligomers−Comparison with RNA and DNA (CpG)$_{12}$ duplexes. *J. Mol. Biol.* **2001**, *305*, 1057−1072.

(21) Auffinger, P.; Bielecki, L.; Westhof, E. Symmetric K+ and Mg2+ ion binding sites in the 5S rRNA loop E inferred from molecular dynamics simulations. *J. Mol. Biol.* **2004**, *335*, 555−571.

(22) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. The missing term in effective pair potential. *J. Phys. Chem.* **1987**, *97*, 6269−6271.

(23) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwel, 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) Dang, L. X. Mechanism and thermodynamics of ion selectivity in aqueous solutions of 18-crown-6 ether: A molecular dynamics study. *J. Am. Chem. Soc.* **1995**, *117*, 6954−6960.

(25) Case, D.; Darden, T. A.; Cheatham, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Merz, K. M.; Wang, B.; Pearlman, D. A.; Crowley, M.; Brozell, S.; Tsui, V.; Gohlke, H.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Schafmeister, C.; Caldwell, J. W.; Ross, W. S.; Kollman, P. A. *AMBER 8*; University of California: San Francisco, CA, 2004.

(26) Darden, T.; Perera, L.; Li, P.; Pedersen, L. New tricks for modelers from the crystallography toolkit: The particle mesh Ewald algorithm and its use in nucleic acid simulations. *Structure* **1999**, *7*, R55−R60.

(27) Degrève, L.; Borin, A. C.; Mazzé, F. M.; Rodrigues, A. L. G. Molecular simulation of a phase separation in a non-primitive electrolyte solution. *Chem. Phys.* **2001**, *265*, 193−205.

(28) Marcus, Y. Ionic radius in aqueous solutions. *Chem. Rev.* **1988**, *88*, 1475−1498.

(29) Ohtaki, H. Ionic solvation in aqueous and nonaqueous solutions. *Monatsh. Chem.* **2001**, *132*, 1237−1268.

(30) Feig, M.; Pettitt, B. M. Sodium and chlorine ions as part of the DNA solvation shell. *Biophys. J.* **1999**, *77*, 1769−1781.

(31) Auffinger, P.; Bielecki, L.; Westhof, E. The Mg2+ binding sites of the 5S rRNA loop E motif as investigated by molecular dynamics simulations. *Chem. Biol.* **2003**, *10*, 551−561.

(32) Auffinger, P.; Westhof, E. RNA hydration: Three nanoseconds of multiple molecular dynamics simulations of the solvated tRNA^{Asp} anticodon hairpin. *J. Mol. Biol.* **1997**, *269*, 326−341.

(33) Degrève, L.; Mazzé, F. M. Molecular simulation of LiCl aqueous solutions. *Mol. Phys.* **2003**, *101*, 1443−1453.

(34) Egorov, A. V.; Komolkin, A. V.; Chizhik, V. I.; Yushmanov, P. V.; Lyubartsev, A. P.; Laaksonen, A. Temperature and concentration effects on Li+-ion hydration. A molecular dynamics simulation study. *J. Phys. Chem. B* **2003**, *107*, 3234−3242.

(35) Du, H.; Rasaiah, J. C.; Miller, J. D. Structural and dynamic properties of concentrated alkali halide solutions: A molecular dynamics simulation study. *J. Phys. Chem. B* **2007**, *111*, 209−217.

(36) Degrève, L.; da Silva, F. L. B. Large ionic clusters in concentrated aqueous NaCl solution. *J. Chem. Phys.* **1999**, *111*, 5150−5156.

(37) Degrève, L.; Da, Silva, F. L. B. Detailed microscopic study of 1 M aqueous NaCl, solution by computer simulations. *J. Mol. Liquids* **2000**, *87*, 217−232.

(38) Lyubartsev, A. P.; Laaksonen, A. Concentration effects in aqueous NaCl solutions. A molecular dynamics simulation. *J. Phys. Chem.* **1996**, *100*, 16410−16418.

(39) Koneshan, S.; Rasaiah, J. C. Computer simulation studies of aqueous sodium chloride solutions at 298 K and 693 K. *J. Chem. Phys.* **2000**, *113*, 8125−8137.

(40) Chowdhuri, S.; Chandra, A. Molecular dynamics simulations of aqueous NaCl and KCl solutions: effects of ion concentration on the single-particle, pair, and collective dynamical properties of ions an water molecules. *J. Chem. Phys.* **2001**, *115*, 3732−3741.

(41) Lyubartsev, A. P.; Marcelja, S. Evaluation of effective ion−ion potentials in aqueous electrolytes. *Phys. Rev. E* **2002**, *65*, 041202.

(42) Uchida, H.; Matsuoka, M. Molecular dynamics simulation of solution structure and dynamics of aqueous sodium chloride solutions from dilute to supersaturated concentration. *Fluid Phase Equilib.* **2004**, *219*, 49−54.

(43) Bouazizi, S.; Nasr, S.; Jaidane, N.; Bellissent-Funel, M. C. Local order in aqueous NaCl solutions and pure water: X-ray scattering and molecular dynamics simulations study. *J. Phys. Chem. B* **2006**, *110*, 23515−23523.

(44) Wick, C. D.; Dang, L. X. Computational observation of enhanced solvation of the hydroxyl radical with increased NaCl concentration. *J. Phys. Chem. B* **2006**, *110*, 8917−8920.

(45) Chen, A. A.; Pappu, R. V. Quantitative characterization of ion pairing and cluster formation in strong 1:1 electrolytes. *J. Phys. Chem. B* **2007**, *111*, 6469−6478.

(46) Chang, T.-M.; Dang, L. X. Detailed study of potassium solvation using molecular dynamics techniques. *J. Phys. Chem. B* **1999**, *103*, 4714−4720.

(47) Grossfield, A.; Ren, P.; Ponder, J. W. Ion solvation thermodynamics from simulation with a polarizable force field. *J. Am. Chem. Soc.* **2003**, *125*, 15671−15682.

(48) Cavallari, M.; Cavazzoni, C.; Ferrario, M. Structure of NaCl and KCl concentrated aqueous solutions by ab initio molecular dynamics. *Mol. Phys.* **2004**, *102*, 959−966.

(49) Dang, L. X.; Schenter, G. K.; Glezakou, V. A.; Fulton, J. L. Molecular simulation analysis and X-ray absorption measurement of Ca^{2+}, K^{+}, and Cl^{-} ions in solution. *J. Phys. Chem. B* **2006**, *110*, 23644−23654.

(50) Lamoureux, G.; Roux, B. Absolute hydration free energy scale for alkali and halide ions established from simulations with a polarizable force field. *J. Phys. Chem. B* **2006**, *110*, 3308−3322.

(51) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. GROMACS: Fast, flexible, and free. *J. Comput. Chem.* **2005**, *26*, 1701−1718.

(52) Chandrasekhar, J.; Spellmeyer, D. C.; Jorgensen, W. L. Energy component analysis for dilute aqueous solutions of Li+, Na+, F-, and Cl-, ions. *J. Am. Chem. Soc.* **1984**, *106*, 903−910.

(53) Lybrand, T. P.; Ghosh, I.; McCammon, J. A. Hydration of chloride and bromide anions: Determination of relative free energy by computer simulation. *J. Am. Chem. Soc.* **1985**, *107*, 7793−7794.

(54) Gavryushov, S.; Linse, P. Effective interaction potentials for alkali and alkaline earth metal ions in SPC/E water and prediction of mean ion activity coefficients. *J. Phys. Chem. B* **2006**, *110*, 10878−10887.

(55) Ponomarev, S. Y.; Thayer, K. M.; Beveridge, D. L. Ion motions in molecular dynamics simulations on DNA. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 14771−14775.

(56) Varnai, P.; Zakrzewska, K. DNA and its counterions: A molecular dynamics study. *Nucleic Acids Res.* **2004**, *32*, 4269−4280.

(57) Auffinger, P.; Vaiana, A. C., Molecular dynamics simulations of RNA systems. In *Handbook of RNA Biochemistry*; Westhof, X. X., Bindereif, X. X., Schön, X. X., Hartmann, X. X., Eds.; Willey-VCH: Manheim, Germany, 2005; pp 560−576.

(58) Donnini, S.; Mark, A. E.; Juffer, A. H.; Villa, A. Incorporating the effect of ionic strength in free energy calculations using explicit ions. *J. Comput. Chem.* **2005**, *26*, 115−122.

(59) Krasovska, M. V.; Sefcikova, J.; Reblova, K.; Schneider, B.; Walter, N. G.; Sponer, J. Cations and hydration in catalytic RNA: Molecular dynamics of the hepatitis delta virus ribozyme. *Biophys. J.* **2006**, *91*, 626−638.

(60) Zahn, D. Atomistic mechanism of NaCl nucleation from an aqueous solution. *Phys. Rev. Lett.* **2004**, *92*, 040801.

(61) Debenedetti, P. G. Thermodynamics: When a phase is born. *Nature* **2006**, *441*, 168−169.

(62) Yang, Y.; Meng, S.; Xu, L. F.; Wang, E. G.; Gao, S. Dissolution dynamics of NaCl nanocrystal in liquid water. *Phys. Rev. E* **2005**, *72*, 012602.

(63) Yang, Y.; Meng, S. Atomistic nature of NaCl nucleation at the solid−liquid interface. *J. Chem. Phys.* **2007**, *126*, 044708.

(64) Murdock, S. E.; Tai, K.; M. H., N.; Johnston, S.; Wu, B.; Fangohr, H.; Laughton, C. A.; Essex, J. W.; Sansom, M. S. P. Quality assurance for biomolecular simulations. *J. Chem. Theory Comput.* **2006**, *2*, 1477−1481.

(65) Glezakou, V.; Chen, Y.; Fulton, J. L.; Schenter, G. K.; Dang, L. X. Electronic structure, statistical mechanical simulaitons, and EXAFS spectroscopy of aqueous potassium. *Theor. Chem. Acc.* **2006**, *115*, 86−99.

(66) Noskov, S. Y.; Berneche, S.; Roux, B. Control of ion selectivity in potassium channels by electrostatic and dynamic properties of carbonyl ligands. *Nature* **2004**, *431*, 830−4.