Towards accurate solvation dynamics of divalent cations in water using the polarizable amoeba force field: From energetics to structure

Jean-Philip Piquemal,a) Lalith Perera,a) and G. Andrés Cisneros
Laboratory of Structural Biology, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina 27709

Pengy Ren
Department of Biomedical Engineering, The University of Texas at Austin, Austin, Texas 78712-0238

Lee G. Pedersen and Thomas A. Dardenb)
Laboratory of Structural Biology, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina 27709

(Received 13 April 2006; accepted 28 June 2006; published online 4 August 2006)

Molecular dynamics simulations were performed using a modified amoeba force field to determine hydration and dynamical properties of the divalent cations Ca$^{2+}$ and Mg$^{2+}$. The extension of amoeba to divalent cations required the introduction of a cation specific parametrization. To accomplish this, the Tholé polarization damping model parametrization was modified based on the \textit{ab initio} polarization energy computed by a constrained space orbital variation energy decomposition scheme. Excellent agreement has been found with condensed phase experimental results using parameters derived from gas phase \textit{ab initio} calculations. Additionally, we have observed that the coordination of the calcium cation is influenced by the size of the periodic water box, a recurrent issue in first principles molecular dynamics studies. © 2006 American Institute of Physics. [DOI: 10.1063/1.2234774]

I. INTRODUCTION

Calcium and magnesium cations play a critical role in many biological systems. In most cases, these divalent cations are specific in their ability to bind to proteins and thereby confer appropriate biological function or activity. For example, the presence of calcium cations is important in blood clotting, signal transduction, and cell division. Specifically, in the case of blood clotting, it has been experimentally observed that the presence of calcium is required for clot formation.1 Indeed, calcium cations directly participate in the binding and folding of the \textit{γ}-carboxyglutamic acid (GLA) rich domain that is common to the vitamin-K-dependent proteins present in the blood coagulation cascade. Blood plasma does not coagulate in the sole presence of magnesium ions,1 an effect attributed to the concomitant lack of binding of the GLA residue to negatively charged phospholipids when only magnesium ions are present. More precisely, recent x-ray crystal structures with mixed ions show that the \textit{N}-terminus loop segment that is thought to be the key determinant of the binding of GLA domain to membranes has a disordered structure when magnesium ions are present. In the presence of calcium ions, GLA domains have been crystallized and a strong GLA-calcium network is observed with varying degree of calcium coordination.1,2

The ability of calcium to coordinate water molecules with flexible coordination is thought to be important for its function. In recent Car-Parrinello (CP) simulation studies the hydration shells and the preferred coordination numbers for calcium and magnesium cations were reported.3–6 However, if one wishes to extend these dynamical studies to proteins involved in blood coagulation, it is apparent that force field simulation methods must be employed.

Divalent cations exhibit strong many body effects in the total energy.7,8 For that reason, polarizable force fields are potentially better able to accurately represent the potential energy surface. Recently, many polarizable force fields have been designed (see Refs. 8 and 9 for extensive review of the field) but only a few of them such as sum of interactions between fragments \textit{ab initio},10–12 (SIBFA) or nonempirical modeling13 (NE\textit{MO}) are parametrized for divalent cations. Performing long molecular dynamics simulations on large systems using periodic boundary conditions with these methods remains a daunting challenge. In this contribution, we introduce a new implementation of the amoeba force field14,15 in AMBER 9.0.17 The key feature is the reimplementation of the original Tinker14–16 subroutines using an improved particle mesh Ewald18–20 (PME) algorithm for electrostatic and polarization energies. The implementation, which uses distributed multipoles (up to quadrupoles),20 provides a significant speedup necessary for the performance of realistic simulations. Additionally, we explore a parametrization of the amoeba force field based on \textit{ab initio} energy decomposition using the constrained space orbital variation (CSOV) approach.21 The idea is to employ \textit{ab initio} results obtained from energy decomposition for a cation interacting with a single water molecule in the gas phase. In this way, it
is then possible to accurately and specifically refine parameters for the Tholé model \(^\text{11,22}\) of polarization damping at short range. Properties calculated using the modified amoeba force field are compared to current experimental values, as well as state of the art Car-Parrinello first principles simulations. We will also discuss artifacts arising from the use of small water boxes in simulating the periodic cell.

II. COMPUTATIONAL DETAILS

A. Gas phase ab initio calculations

MP2 (full) intermolecular interaction energy-distance scan calculations have been performed using GAUSSIAN 03. \(^\text{23}\) The aug-cc-pVTZ basis set \(^\text{24}\) was employed for water and cc-pVDZ basis set \(^\text{24}\) for the Ca\(^{2+}\) and Mg\(^{2+}\) ions. CSOV polarization energy calculations have been performed using a modified version of HONDO95.3 (Ref. 25) with the Hartree-Fock and B3LYP (Refs. 26 and 27) methods using the above basis sets. The water geometry was selected to be as in the \(/\text{H}_2\text{O}\) \(^\text{85}\) Fock and B3LYP polarization energy calculations have been performed using a cc-pVDZ basis set \(^\text{24}\) for the Ca\(^{2+}\) and Mg\(^{2+}\) ions. CSOV polarization energies were computed using GAUSSIAN 03 at the MP2/CC-pVDZ level.

B. Efficient PME calculations for multipole-based electrostatic and polarization energies

In the present AMBER 9.0 implementation of the amoeba force field, a Cartesian tensor implementation of the multipole-based (up to quadrupoles) PME has been developed following Sagui et al. \(^\text{30}\) The direct sum which uses the McMurchie-Davidson recursion \(^\text{28}\) has been modified in order to treat polarization by a Tholé damping model. Subsequently, the PME algorithm allows a damping of the permanent-multipole induced-dipoles and induced-dipole induced-dipole interactions at short range (i.e., in the direct sum). The procedure is self-consistent: the first step generates the induced dipoles due to the permanent fields (charges, dipoles, and quadrupoles) then the procedure includes fields due to induced dipoles and iterates until self-consistency (a convergence criteria of \(10^{-8}\) D for the sum of the induced dipoles was used). The iteration step uses a successive over-relaxation (SOR) method \(^\text{29}\) in order to accelerate the convergence,

\[
\mu_{n+1} = \alpha F_n,
\]

\[
\mu_{n+1} = \gamma \mu_n + (1 - \gamma) \mu'_{n+1},
\]

where \(\mu\) is the induced dipole, \(\alpha\) the scalar polarizability, and \(F\) the field due to the permanent and induced dipole moments. \(\gamma\) is a parameter set to 0.7. It is worth noting that the implementation is general and can readily accommodate protein simulations. \(^\text{30}\)

C. Parametrization of the Tholé damping model using the CSOV approach

The purpose of our approach is to refine the amoeba polarization model so that it will accurately accommodate divalent cations. The Tholé method is pairwise additive and with a modification for the multipole charge distribution at short range. The original amoeba damped charge distribution \(^\text{14}\) has the form

\[
\rho = \frac{3a}{4\pi} \exp(-au^2),
\]

where \(u=R_{ij}/(\alpha_i \alpha_j)^{1/6}\) is the effective distance as a function of atomic polarizabilities of sites \(i(\alpha_i)\) and \(j(\alpha_j)\). The dimensionless parameter \(a\) controls the strength of the damping.

In the original amoeba model, a universal Tholé parameter \(a\) has been introduced. That way, in its original formulation, the ability of the amoeba force field to discriminate cations is limited to a specific parametrization of the van der Waals interactions. For that reason, two chemically different cations bearing the same charge but presenting a different electronic structure would exhibit the same electrostatic and polarization interactions (owing to the very small cation polarizability), showing none of the differences observed at the \(ab\) \textit{initio} level. However, Masia et al. \(^\text{31}\) recently proposed to introduce a specific Tholé parameter for a given water-cation complex to account for the specific properties of water-cation interaction. Following this idea, in the present study, we decouple the screening parameter used for the water-water interaction in the original water model and use instead cation specific damping parameters for different cation-water interactions. Moreover, we propose to extend the parametrization approach \(^\text{31}\) of the original work of Masia et al. by adding energetic criteria based on the \(ab\) \textit{initio} polarization energy.

To do so, we chose to use the CSOV decomposition scheme which allows us to split the interaction energy into components having physical interpretations: i.e., Coulomb, exchange-repulsion, polarization, and charge transfer energies (no dispersion at the levels treated by CSOV). The CSOV approach can further separate the induction energy into two parts: the polarization energy approximated using polarization models such as Applequist-Tholé, and the charge transfer energy, a nonclassical term arising in the wave function (donation, back donation, and incompleteness of the basis set).

\[
\Delta E_{\text{tot}} = E_{\text{Coulomb}} + E_{\text{exch-repulsion}} + E_{\text{pol}} + E_{\text{ct}}
\]

\[
= E_{\text{frozen core}} + E_{\text{induction}}.
\]

CSOV is not available at the MP2 level but we can consider that the MP2 polarization energy value, which includes dynamic electronic correlation, should be intermediate between the Hartree-Fock (uncorrelated and therefore underestimated) and B3LYP [correlated but overestimated due to the self interaction artifacts present in density functional theory \(^\text{32}\) (DFT)] values. Since this approach can be coupled to the original dipole criteria \(^\text{31}\) it is thus possible to fit the Tholé parameter so as to reproduce the behavior of the polarization energy of an amoeba water interaction with any cation.

Another crucial point is the importance of the charge transfer contribution. In the case of a magnesium ion, this is negligible according to CSOV (only around 2% of the total interaction energy around the first shell coordination distance). Likewise, the water-calcium complex exhibits a slightly larger charge transfer component (around 3% of the
total energy) compared to Mg$^{2+}$. In the amoeba energy scheme, this quantity will have to be included in the so-called “van der Waals” term (including short range electrostatic, penetration, exchange-repulsion, and dispersion energies) in order not to perturb the dominant many body polarization term. Finally, the parameters $R_e$ and $e$ of the Halgren function$^{33}$ representing the van der Waals interaction are fitted to the remainder of the total interaction energy minus the sum of coulomb and polarization amoeba energies,

$$E_{\text{Halgren}} = \Delta E_{\text{int}} (\text{MP2BSSE corrected}) - E_{\text{Coulomb}}(\text{amoeba multipoles}) - E_{\text{polarization}}(\text{amoeba}).$$

Concerning the van der Waals interactions, it is important to point out that a long range periodic correction is now included in the reciprocal sum of the PME approach using an analytic continuum (see Sec. 2.8 of Ref. 34 for more details about the analytic continuum expression).

D. Simulation protocol

For each cation, three sets of trajectory calculations were performed: (a) with 60 water molecules, (b) with 216 water molecules, and (c) with 516 water molecules. In all cases, the flexible amoeba water model$^{14}$ has been used. The PME algorithm uses a 0.8 Å grid, a spline order of 5 and 0.45 Ewald coefficient. The cutoffs are about 6.7 Å for the direct space Coulomb and 8.0 Å for the van der Waals. All constant pressure production simulations were carried out at 300 K with 1 fs time step for a total of 5 ns or more for each trajectory (after a nanosecond of equilibration period). Configurations recorded at every 1 ps were used for the analysis. Residence times were computed following Eq. (9) of Ref. 35. As proposed, we introduced a parameter $\zeta$ to take into account the molecules which leave the first coordination shell only temporarily. The values of $\zeta$ were taken to be 0 and 2 ps in order to include or exclude these effects. Additionally, we verified that the choice of a convergence criteria on the induced dipoles within the simulation did not affect the final result by making sure that the radial distribution functions and dynamical properties were identical.

III. RESULTS AND DISCUSSION

A. Parametrization

As it can be seen from Figs. 1 and 2, both the modified amoeba polarization and total interaction energies appear to be in a good agreement with the $ab$ initio polarization and the supermolecular intermolecular interaction energy. For both cations, the quantum behavior is preserved and the modified amoeba polarization energy is intermediate between Hartree-Fock and B3LYP values in the range of interest. It is important to point out that the reported damping parameter extracted from a dipole observation method$^{31}$ is very close to our best fit obtained from CSOV decomposition in the case of calcium ($a=0.088$). However, for magnesium, the approach of Masia et al. ($a=0.05$) performs well only at medium and long range but it does not correctly represent the short range behavior of the $ab$ initio polarization energy,$^{36}$ compared to our new parameter $a=0.076$. It is important to point out that without the modification in Tholé damping parameters, no selectivity for a specific cation can be included into the Tholé polarization energy. In the original amoeba parametrization scheme which uses a universal Tholé parameter, any divalent (or trivalent) cation would present the same polarization energy in the presence of water. This obviously is not physical, as shown in the quantum results, thus limiting the validity/applicability of the prior many body approach,$^{13,14}$ to treat divalent cations. As explained earlier, the Halgren function has been adjusted to reproduce the $ab$ initio interaction energy. The two parameters ($R_e$ and $e$) are 3.22 Å and 0.96 kcal/mol for water-Ca$^{2+}$ and 2.044 Å and 0.69 kcal/mol for water-Mg$^{2+}$. We observed that this function slightly underestimates the interaction strength in the range between the first and second water shell. Nevertheless, the errors in the Halgren function remain relatively small and should not seriously influence the dynamics. The atomic polarizabilities of Ca$^{2+}$ and Mg$^{2+}$ are 0.376 and 0.107 Å$^3$, respectively.
At least two bulklike water shells are observed in the central simulation box while this number increases to 3 (or more) in the case of a large 516 water box.

C. Calcium ion

As with magnesium, good agreement is observed for structural and dynamical properties between the most recent experiment\(^3,39,40\) and the present work for calcium hydration (see Table I and Fig. 3). Concerning the radial distributions of water around the cations, the peak positions of the Ca\(^{2+}\)-O and Ca\(^{2+}\)H (i.e., 2.42–2.46 Å and 3.01–3.05 Å, respectively), are fully in line with experiment\(^3,39,40\). For Ca\(^{2+}\), a significant exchange of the water molecules in the first shell is observed. Due to this exchange, the average coordination of the cation oscillated between seven and eight water molecules. However, eight coordination is preferred over the seven coordination for the larger box sizes (216 or 516 water). In the case of small water boxes, this preference disappears indicating a dependency of the dynamical properties on the periodic simulation cell size. This tendency towards eight coordination is in agreement with the experimental observations\(^3,39,40\) which predict that the coordination increases at infinite dilution (starting from 7 at medium concentrations and increasing towards 8), the condition present in the simulations.

As can be seen from Figs. 3(c) and 3(d), the radial distribution function \([g(r)]\) of water molecules around Ca\(^{2+}\) is strongly influenced by the choice of the box size. Larger box sizes appear to favor a larger coordination number. This same effect is observed on the average first shell residence time, with about 30% increase in the larger simulation boxes. As in the magnesium case, our systems with 216 and 512 water molecules exhibit at least two structured water layers around the cation with two or more bulklike water layers. As the dynamical properties are conserved between the 216 and 516 water boxes (see Table I and Figure 3), a 216 water molecule simulation cell appears sufficient to perform a satisfactory simulation.

Thus, a note of caution is warranted here for interpreting results obtained from some first principle Car-Parrinello simulations which are presently capable of dealing with only small water boxes and small simulation times.

Indeed, our multiple nanosecond simulations have exposed potential artifacts for systems with a relatively small number of water molecules. For example, we find that the coordination of calcium, as well as the residence time, is dependent on the size of the periodic water box. This effect, however, is not present in the magnesium case, due to the absence of exchanges in the first shell. It appears that too small a water box will tend to decrease the average coordination of Ca\(^{2+}\). Thus, it is almost certain that for small box sizes, as have been used, for example, in first principle Car-Parrinello molecular dynamics computations, the box size effect may lead to an underestimation of the coordination in Ca\(^{2+}\). Indeed, concerning the ion-oxygen radial distribution, we have observed a bulklike behavior of the water molecules beyond the cation’s second solvation shell (beyond 5 Å for Mg\(^{2+}\) and beyond 5.3 Å for Ca\(^{2+}\)). In the case of the small

B. Magnesium ion

As shown in Table I and in Figs. 3(a) and 3(b), the results from the modified amoeba simulation are in good agreement with those of both experiments (see Ref. 3 and references therein), Car-Parrinello\(^3\) and quantum mechanical/molecular mechanical (QM/MM) simulations.\(^37\) The first Mg\(^{2+}\)-O peak position is accurately predicted [see Fig. 3(a)] at 2.08 Å (2.09 Å experimentally),\(^3\), and no exchange of the six water molecules in the first shell was observed during the simulation time. These results are in line with the long exchange times observed experimentally (microsecond scale).\(^38\)

Since no exchange in the first shell is observed, one should not expect any dependency of the dynamical properties on the periodic simulation box size. Also, we noticed that the bulklike water molecules are observed beyond the second shell of water molecules around the cation in the 216 and 516 water boxes. The number of bulklike water depends on the size of the box. Indeed, in the case of 216 water box at...
box (60 water molecules) with a box length of 12 Å, the maximum ion-water distance is 6 Å which corresponds only to the two structured layers (see Fig. 4). This situation leads to artifacts since the two structured layers can interact directly with their periodic images in the absence of significant bulklike water separating them. Moreover, it is likely that some discrepancies between experiment and certain CP results may be linked to the incomplete relaxation of the water molecules. Indeed, we observed that amoeba flexible water molecules, which we employ here, require a significantly larger relaxation time compared to regular rigid classical models. The smaller relaxation times of rigid water models may account for the observation of a larger coordination number appearing in CP when rigid water molecules were used. Moreover, it seems that two additional factors are important for the CP simulations, the choices of the DFT functional and of the fictitious electronic masses used for the extended Lagrangian in CP calculations. Indeed, the Car-Parrinello molecular dynamics (CPMD) simulations of Naor et al., using BLYP and a mass of 890 a.u., are in relative agreement with experiment despite a small simulation box. This choice differs from the one of Lightstone et al. who preferred the PBE functional and a mass of 1100 a.u., leading to the underestimation of the coordination of Ca\(^{2+}\). The magnitude of these combined effects is for the moment unknown for Ca\(^{2+}\) and Mg\(^{2+}\) Car-Parrinello computations and will require careful studies since the recommended fictitious mass is about 400 a.u. for liquid water simulations.

The computed residence times appear smaller than the experimental values that have been postulated to be in the \(10^{-7} - 10^{-9}\) s range but are in better agreement than results from CP and QM/MM simulations. A key point in understanding these discrepancies could be the recent experimental description of the important coordination of counterions such as Cl\(^{-}\) to the first shell water molecules at the concentrations used in the experiment. The counterion effects are probably increasing the residence times. In other words, the water molecules are more tightly bound to the counterions, which leads to longer residence times.

### Table I: Summary of structural and dynamical properties for Ca\(^{2+}\) and Mg\(^{2+}\) simulations.

<table>
<thead>
<tr>
<th></th>
<th>(M = \text{Ca}^{2+})</th>
<th>(M = \text{Mg}^{2+})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water coordination</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experimental(^a),(^b)</td>
<td>7.0–8.0</td>
<td>6.0</td>
</tr>
<tr>
<td>Car-Parrinello(^c)</td>
<td>6.2–7.2, 8–9</td>
<td>6.0</td>
</tr>
<tr>
<td>This work</td>
<td>7.7 (216 and 516 waters), 7.2 (60 waters)</td>
<td></td>
</tr>
<tr>
<td>First (M(II)–O) peak (Å)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experimental(^d)</td>
<td>2.41–2.44; 2.437; 2.46</td>
<td>2.09</td>
</tr>
<tr>
<td>Car-Parrinello(^e)</td>
<td>2.43–2.44</td>
<td>2.13</td>
</tr>
<tr>
<td>This work</td>
<td>2.42–2.46</td>
<td>2.08</td>
</tr>
<tr>
<td>First (M(II)–H) peak (Å)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experimental(^d)</td>
<td>2.97–3.07</td>
<td>…</td>
</tr>
<tr>
<td>Car-Parrinello(^d)</td>
<td>3.03–3.07</td>
<td>…</td>
</tr>
<tr>
<td>This work</td>
<td>3.01–3.05</td>
<td>2.63–2.73</td>
</tr>
<tr>
<td>Average first shell water tilt angle</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experimental(^d)</td>
<td>34.0–38.0; 39.7</td>
<td>…</td>
</tr>
<tr>
<td>Car-Parrinello(^d)</td>
<td>40.1</td>
<td>…</td>
</tr>
<tr>
<td>This work</td>
<td>40.2–40.7</td>
<td>39.7–40.6</td>
</tr>
<tr>
<td>First shell water residence time (ps) for (n) water molecules</td>
<td>(N)</td>
<td>(t' = 0) ps</td>
</tr>
<tr>
<td>60</td>
<td>131</td>
<td>154</td>
</tr>
<tr>
<td>216</td>
<td>158</td>
<td>182</td>
</tr>
<tr>
<td>516</td>
<td>160</td>
<td>203</td>
</tr>
<tr>
<td>Percentage coordination with (n) water molecules</td>
<td>(n = 60)</td>
<td>(n = 216)</td>
</tr>
<tr>
<td>6</td>
<td>0.84</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>46.16</td>
<td>33.16</td>
</tr>
<tr>
<td>8</td>
<td>52.72</td>
<td>66.32</td>
</tr>
<tr>
<td>9</td>
<td>0.28</td>
<td>0.52</td>
</tr>
</tbody>
</table>

\(^a\)References 39 and 40.  
\(^b\)Reference 3 and references therein.  
\(^c\)References 3–6.  
\(^d\)References 3 and 39.  
\(^e\)Reference 6.  
\(^f\)References 40 and 52.
words, we did not use explicit counterions (we have instead a neutralizing background) in our simulations. This may account for the somewhat shorter simulation exchange times.

IV. CONCLUSIONS

Structural and dynamical properties of ion-water interactions for calcium and magnesium divalent cations were captured through multiple nanosecond simulations using a modified amoeba polarizable force field. It is encouraging that the parameters obtained from quantum calculations performed in vacuum for cation water pairs (note that we have used a high level \textit{ab initio} decomposition scheme to correctly obtain the contributions to the total energy) yield very good results when used in bulk simulations \textit{without condensed phase adjustment} (contrary to the approach generally used for the parametrization of classical force fields). This should lead to a more robust force field which is also able to treat nonhomogenous environments with difficult local interactions.\textsuperscript{49}

The benefits of tailoring specific atom-cation pair Tholé parameters are obvious and should be preferred over the fixed universal Tholé parameter used in the original amoeba scheme. The parametrization by means of CSOV calculation appears to be effective in enabling the decomposition of many body effects around the cation. Limitations in fitting the van der Waals interactions to the Halgren functional form for ions with substantial charge transfer may require additional modifications; however, for the cases considered here, the current approach is effective.

We have observed that the structural and dynamical properties of the calcium cation is influenced by the size of the periodic water box, a recurrent issue in first principles molecular dynamics studies with strong consequences on the dynamical properties.

This work validates the use of PME as an important tool in performing long simulations with polarizable force fields. As the methodology requires more attention, details will be given in an incoming paper\textsuperscript{50} generalizing the approach. Regarding the timings, a nanosecond simulation using a 216 water molecule box takes 34 h on a single 1.8 Ghz AMD Opteron processor. For a system of this size, the PME algorithm appears already ten times faster than the regular Ewald summation \((n^2\) scaling). This timing increases to 94 h when a large 516 box is considered following the \(n(\log(n))\) dependency of PME. On average, amoeba appears 12 times slower than the classical AMBER force field in AMBER 9.0. These timings have to be compared with the 20–50 ps simulations currently performed at the CP level on massively parallel machines using months of computer time.

For these reasons, we think that the possibility of the use of large simulations cells in the framework of the AMBER 9.0 amoeba implementation could be a reasonable complement to first principle approaches. Future work will focus on the importance of charge transfer in the solvation of metallic cations, on hydration free energy computations\textsuperscript{51} as well...
as in the implementation of the PME approach into more realistic SIBFA-type potentials.

ACKNOWLEDGMENTS

This research was supported by the Intramural Research program of the NIH and NIEHS. The authors wish to thank Robert Bass (NIEHS) and Jay Ponder (Washington University at St. Louis) for their support. Two of the authors (J.-P.P. and L.P.) have contributed equally to this work.