Many-Body Effects Determine the Local Hydration Structure of Cs⁺ in Solution

Debbie Zhuang,† Marc Riera,† Gregory K. Schenter,‡ John L. Fulton,*†§ and Francesco Paesani*,†¶§

†Department of Chemistry and Biochemistry, University of California, San Diego
    La Jolla, California 92093, United States

‡Physical and Computational Sciences Directorate, Pacific Northwest National Laboratory
    Richland, WA 99352, United States

¶Materials Science and Engineering, University of California, San Diego
    La Jolla, California 92093, United States

§San Diego Supercomputer Center, University of California, San Diego
    La Jolla, California 92093, United States

E-mail: john.fulton@pnnl.gov; fpaesani@ucsd.edu
Abstract

A systematic analysis of the hydration structure of Cs\(^+\) ions in solution is derived from simulations carried out using a series of molecular models built upon a hierarchy of approximate representations of many-body effects in ion–water interactions. It is found that a pairwise-additive model, commonly used in biomolecular simulations, provides poor agreement with experimental X-ray spectra, indicating an incorrect description of the underlying hydration structure. Although the agreement with experiment improves in simulations with a polarizable model, the predicted hydration structure is found to lack the correct sequence of water shells. Progressive inclusion of explicit many-body effects in the representation of Cs\(^+\)–water interactions as well as account for nuclear quantum effects is shown to be necessary for quantitatively reproducing the experimental X-ray spectra. Besides emphasizing the importance of many-body effects, these results suggest that molecular models rigorously derived from many-body expansions hold promise for realistic simulations of aqueous solutions.

Graphical TOC Entry
Determining the driving forces and molecular mechanisms that regulate the hydration properties of alkali-metal ions is fundamental for a microscopic understanding of several processes taking place in aqueous clusters, solutions, and interfaces, which, in turn, have major implications for various fields of science and engineering. For example, lithium salts are widely used in rechargeable batteries,\textsuperscript{1,2} and are effective agents in the treatment of manic-depressive illness.\textsuperscript{3} Sodium and potassium ions play important roles in the stabilization of biomolecules,\textsuperscript{4–6} intracellular signal transduction,\textsuperscript{7,8} and enzyme and nucleic acid catalysis.\textsuperscript{9–12} On the other hand, the heavier alkali-metal ions (i.e., rubidium and cesium) are not as ubiquitous in either the environment or in living systems. However, the accident at the Fukushima Daiichi reactor in 2011 has drawn attention to the importance of a molecular-level understanding of adsorption and desorption processes of soluble radionuclides, such as $^{137}$Cs, for the development of efficient technologies for the treatment of nuclear waste.\textsuperscript{13}

In this context, deriving a molecular picture of the hydration properties of Cs$^+$ ions presents significant challenges to both experiment and theoretical modeling due to the intricate interplay between Cs$^+$–water and water–water interactions, which, being of similar magnitude, are difficult to disentangle. As a result, large variability is found in the values reported in the literature for both Cs$^+$ coordination number and spatial extent of the hydration shells around Cs$^+$ ions in solution. From large-angle X-ray (LAXS) and double difference infrared (DDIR) spectroscopic experiments performed on a 2.0 M cesium iodide solution, the mean distance between Cs$^+$ and the oxygen atoms of the water molecules (Cs$^+$–O) within the first hydration shell was estimated to be 3.07 Å, corresponding to a coordination number of 8.\textsuperscript{14} Anomalous X-ray diffraction patterns measured for a 3 molal cesium iodide solution were used to determine a coordination number of 7.9, assuming a Cs$^+$–O distance of 3.0 Å.\textsuperscript{15} Values in the range of 7.7 – 8 for the coordination number and 3.1 Å – 3.2 Å for and Cs$^+$–O distance were reported from subsequent neutron and X-ray diffraction measurements.\textsuperscript{16,17} Various theoretical and computational approaches, including \textit{ab initio} molecular dynamics (AIMD), quantum mechanics/molecular mechanics (QM/MM), and classical molecular dy-
namics (MD), have also been used to investigate the hydration structure of Cs$^+$, resulting in predictions for the coordination number and Cs$^+$–O distance in the range of 7 – 10 and 3.0 – 3.3 Å, respectively.\textsuperscript{18–23}

The development of explicit many-body potential energy functions (PEFs), rigorously derived from the corresponding many-body expansions of the underlying interaction energies, represents a significant step toward predictive computer simulations of aqueous systems at the molecular level.\textsuperscript{24–39} By combining the MB-pol PEF,\textsuperscript{32–34} which correctly predicts the properties of water across different phases,\textsuperscript{40} with the TTM-nrg and MB-nrg PEFs,\textsuperscript{38,39} which accurately describe molecular interactions between halide and alkali-metal ions with water, theoretical studies have been carried out to characterize the tunneling dynamics in $X^–(H_2O)$ and $X^–(D_2O)$ dimers, with $X = F, Cl, Br, \text{and I}$,\textsuperscript{41} as well as quantum isomeric equilibria in small $M^+(H_2O)_n$ clusters, with $M = Li, Na, K, Rb, \text{and Cs}$.\textsuperscript{42}

In this study, a series of molecular models built upon a hierarchy of approximate representations of Cs$^+$–water interactions is used in MD and path-integral molecular dynamics (PIMD) simulations\textsuperscript{43} to investigate many-body effects in the hydration structure of Cs$^+$ ions. Specifically, the analyses presented in the following include the empirical point-charge force field obtained by combining the TIP4P-Ew model for water\textsuperscript{44} with the corresponding Cs$^+$–water parameterization introduced in Ref. 45, hereafter referred to as TIP4P-Ew, the polarizable TTM-nrg PEF, with an implicit representation of N-body (NB) effects based on classical induction,\textsuperscript{37} and the many-body MB-nrg PEF, with either explicit inclusion of two-body (2B) interactions in addition to the same classical NB term adopted by TTM-nrg, hereafter referred to as (2B+NB)-MB-nrg, or explicit inclusion of both two-body (2B) and three-body (3B) interactions in addition to the TTM-nrg classical NB term, hereafter referred to as (2B+3B+NB)-MB-nrg.\textsuperscript{39} The accuracy of the four different models in reproducing the hydration structure of Cs$^+$ in diluted solutions as well as the role played by nuclear quantum effects are then assessed through systematic comparisons between measured and simulated L$_1$-edge and L$_3$-edge extended X-ray absorption fine structure (EXAFS) spectra.
Specific details about both simulations and experimental measurements are reported in the Supporting Information.

To first assess the ability of the TIP4P-Ew, TTM-nrg, (2B+NB)-MB-nrg, and (2B+3B+NB)-MB-nrg models to describe Cs$^+$–water interactions, Fig. 1 shows the correlation plots between 2B and 3B energies calculated with each model and the corresponding reference values obtained at the coupled cluster level of theory with single, double, and iterative triple excitations, CCSD(T). In this analysis, 2B and 3B energies are calculated for 380 distinct Cs$^+$(H$_2$O) and Cs$^+$(H$_2$O)$_2$ configurations with distorted water geometries, respectively. In the case of the (rigid) TIP4P-Ew model, the CCSD(T) reference energies are calculated for the same dimer and trimer configurations used for the other models after rescaling both OH bonds and HOH angles of the water molecules to match the TIP4P-Ew parameterization. Specific details about the CCSD(T) calculations are reported in the Supporting Information.

Among the four models, TIP4P-Ew predicts 2B energies that deviate significantly from the CCSD(T) data (panel a), resulting in a root-mean-square error (RMSE) of 10.55 kcal/mol.

![Figure 1: Correlation plots for 2B (top panels) and 3B (bottom panels) Cs$^+$–water interaction energies for the TIP4P-Ew, TTM-nrg, (2B+NB)-MB-nrg and (2B+3B+NB)-MB-nrg models. On the x-axes are the CCSD(T) reference data while on the y-axes are the corresponding values calculated with each model.](image-url)
This large RMSE can be explained by considering that, as in all empirical pairwise-additive force fields, the 2B term of TIP4P-Ew does not strictly represent 2B interactions but effectively also accounts for higher-order interactions that, by construction, are not explicitly included in the model (e.g., 3B interactions in panel b). As shown by the correlation plots for the TTM-nrg model (panels c and d), implicit inclusion of many-body effects through classical polarization significantly improves the agreement with the CCSD(T) data for both 2B and 3B energies, resulting in RMSEs of 1.23 kcal/mol and 0.14 kcal/mol, respectively. Explicit representations of 2B interactions, which is accomplished in the MB-nrg PEFs (panels e and g) by introducing permutationally invariant polynomials (PIPs) that effectively represent non-classical contributions to molecular interactions (e.g., charge transfer and penetration, and Pauli repulsion),\textsuperscript{38,39} leads to further reduction of the 2B RMSEs to 0.05 kcal/mol. Finally, the correlation plots shown in panels d, f and h demonstrate that purely classical representations of 3B interactions adopted by the TTM-nrg and (2B+NB)-MB-nrg models are not sufficient to quantitatively reproduce the CCSD(T) data, resulting in RMSEs of 0.17 kcal/mol. Significantly higher accuracy in the description of 3B energies, with an associated RMSE of 0.02 kcal/mol, is exhibited by the (2B+3B+NB)-MB-nrg model, which supplements the classical description of 3B Cs\textsuperscript{+}–water interactions adopted by the TTM-nrg and (2B+NB)-MB-nrg models with explicit PIPs representing non-classical 3B contributions.

Having established the accuracy of the four models in reproducing the lower-order, and more relevant, many-body effects in Cs\textsuperscript{+}–water interactions, Fig. 2 analyzes how the differences found in the correlation plots of Fig. 1 impact the hydration structure of Cs\textsuperscript{+} in solution. The comparison between the Cs\textsuperscript{+}–O radial distribution functions (RDFs), $g_{\text{Cs}^+\text{–O}}$, calculated from MD and PIMD simulations with the four models (Fig. 2a) clearly shows that TIP4P-Ew predicts a more structured distribution of water molecules around Cs\textsuperscript{+}, with a sequence of well-defined peaks located at $\sim$3.0 Å, $\sim$5.0 Å, $\sim$7.0 Å, and $\sim$9.0 Å. Inclusion of an implicit description of many-body effects through a classical polarization term as implemented in the TTM-nrg model effectively leads to the collapse of the hydration shell located
at \( \sim 5.0 \) \( \text{Å} \) in the TIP4P-Ew RDF. This is accompanied by the broadening of the first peak, corresponding to an expansion of the 1\(^{\text{st}}\) hydration shell, whose position consequently shifts to relatively larger \( \text{Cs}^{+} - \text{O} \) distances. As a result of this structural reorganization of the water molecules around \( \text{Cs}^{+} \), the second and third peaks (i.e., 2\(^{\text{nd}}\) and 3\(^{\text{rd}}\) hydration shells) in the TTM-nrg RDF effectively correspond to the third and fourth peaks (i.e., 3\(^{\text{rd}}\) and 4\(^{\text{th}}\) hydration shells) in the TIP4P-Ew RDF.

Qualitatively different hydration structures are predicted by the (2B+NB)-MB-nrg and (2B+3B+NB)-MB-nrg models, which progressively include explicit representations of 2B and

![Figure 2: a) Radial distribution functions (RDFs) describing the spatial correlation between the \( \text{Cs}^{+} \) ion and the oxygen (O) atoms of the water molecules calculated from MD simulations with the TIP4P-Ew, TTM-nrg, (2B+NB)-MB-nrg and (2B+3B+NB)-MB-nrg models as well as from PIMD simulations with the (2B+3B+NB)-MB-nrg model. b) Corresponding coordination numbers calculated as a function of the \( \text{Cs}^{+} - \text{O} \) distance. c) Two-dimensional plots comparing the density profiles calculated from MD simulations with the TIP4P-Ew, TTM-nrg, (2B+NB)-MB-nrg and (2B+3B+NB)-MB-nrg models on the left of each panel with the corresponding results obtained from PIMD simulations with the (2B+3B+NB)-MB-nrg model.](image-url)
3B contributions to Cs\(^+\)–water interactions. Compared to the TIP4P-Ew RDF, the RDFs calculated with both MB-nrg models display a lower but broader first peak that extends up to \(\sim 4.5\ \text{Å}\), as well as a second, broader peak extending from \(\sim 4.5\ \text{Å}\) to \(\sim 8.0\ \text{Å}\). Contrary to TTM-nrg predictions, the MB-nrg RDFs display a well-defined sequence of hydration shells, with a sharper first peak at \(\sim 3.15\ \text{Å}\) and a second, broader peak at \(\sim 6.2\ \text{Å}\). Although comparisons between RDFs calculated from simulations with (2B+NB)-MB-nrg and (2B+3B+NB)-MB-nrg suggests that the inclusion of an explicit 3B term in the (2B+3B+NB)-MB-nrg model only leads to minor changes to the overall hydration structure of Cs\(^+\), it nevertheless contributes to sharpening the interstitial region between the first and second peak, providing further evidence for the importance of non-classical 3B effects in Cs\(^+\)–water interactions.

Finally, the comparison between the RDFs obtained from MD and PIMD simulations with the (2B+3B+NB)-MB-nrg model indicates that nuclear quantum effects play a minimal role in determining the hydration structure of Cs\(^+\) in solution.

Fig 2b shows that the differences in the RDFs obtained from simulations with the four models directly translate into different distributions of water molecules around the Cs\(^+\) ion. In particular, while MD simulations with the TIP4P-Ew model predict that \(\sim 8\) molecules are within the 1\(^{\text{st}}\) hydration shell of Cs\(^+\), the lack of a well-defined 2\(^{\text{nd}}\) hydration shell makes it difficult to unambiguously determine a coordination number from the corresponding TTM-nrg simulations. The wider 1\(^{\text{st}}\) hydration shells predicted by simulations with both MB-nrg models result in relatively larger coordination numbers (\(\sim 10\)), with negligible differences due to nuclear quantum effects. The more compact nature of the hydration structure predicted by the TIP4P-Ew model is visually apparent in the two-dimensional density plots shown in the first panel of Fig. 2c. Fig. 2c also shows that the systematic inclusion of more accurate representations of many-body effects, from the TTM-nrg to the (2B+3B+NB)-MB-nrg model, leads to the progressive reshaping of the hydration shells, particularly at short Cs\(^+\)–water distances. Overall, the comparisons in Fig. 2 demonstrate that many-body effects have significant impact on the local Cs\(^+\) hydration structure, while pairwise-additive representa-
tions, which approximate Cs$^+$–water interactions in a mean-field sense, become increasingly more accurate at larger Cs$^+$–water distances ($R_{Cs^+-O} > 7.0$ Å).

Further insights into the local Cs$^+$ hydration structure are obtained from decomposing the total Cs$^+$–O RDFs shown in Fig. 2a into individual contributions associated with each water molecule $i$ as a function of its distance ($R_{Cs^+-O_i}$) from the Cs$^+$ ion. These incremental radial distribution functions ($i$-RDFs) calculated from MD simulations with TIP4P-Ew, TTM-nrg, (2B+NB)-MB-nrg, and (2B+3B+NB)-MB-nrg as well as from PIMD simulations with (2B+3B+NB)-MB-nrg are shown in panels a-e of Fig. 3, respectively. While the TIP4P-Ew $i$-RDFs show a clear separation between water molecules belonging to the 1st and 2nd hydration shells, with the 9th molecule located in the interstitial region between the two shells, all other models predict a more gradual transition between the first two hydration shells. This transition is extremely hard to detect in the TTM-nrg $i$-RDFs and becomes more distinct as both many-body and nuclear quantum effects are explicitly taken into account in the PIMD simulations with the (2B+3B+NB)-MB-nrg model.

The different evolution of the hydration shells predicted by the four models becomes more evident from the analysis of the average distances, $\langle R_{Cs^+-O_i} \rangle$, and associated variances, $\sigma_i^2$,
calculated in Fig. 3f-j for the individual \(i\)-RDFs. MD simulations with the TIP4P-Ew model predict narrower distributions for water molecules closer to the Cs\(^+\) ion (\(\langle R_{\text{Cs}^+-O_i} \rangle \leq 3.4 \, \text{Å}\)), indicating relatively stronger Cs\(^+\)–water interactions, as well as the broadest distributions for water molecules (8\(^{th}\) to 11\(^{th}\)) located at the boundary of the 2\(^{nd}\) and 3\(^{rd}\) hydration shells, with \(\sigma_i^2\) becoming effectively constant for the \(i\)-RDFs corresponding to water molecules with \(i \geq 16\). The distinction between water molecules located between 3.5 Å and 4.5 Å and those at both smaller and larger distances from the Cs\(^+\) ion becomes less marked when many-body effects are implicitly included in simulations with the TTM-nrg model. Fig. 3 shows that explicit account for many-body Cs\(^+\)–water interactions and nuclear quantum effects in simulations with the MB-nrg models progressively leads to sharper distributions of \(\sigma_i^2\) as a function of \(\langle R_{\text{Cs}^+-O_i} \rangle\). Although these distributions are qualitatively similar to that obtained from MD simulations with the TIP4P-Ew model, they are centered at larger \(\langle R_{\text{Cs}^+-O_i} \rangle\) values, with their maxima shifted by \(\sim 0.5 \, \text{Å}\) relative to the TIP4P-Ew distribution, and more uniform, with the difference in \(\sigma_i^2\) between the narrowest and broadest \(i\)-RDFs being \(\sim 0.5 \, \text{Å}^2\) compared to \(\sim 0.75 \, \text{Å}^2\) for the corresponding TIP4P-Ew \(i\)-RDFs.

Finally, the ability of the different models to realistically describe the hydration structure of Cs\(^+\) in solution is determined from comparisons between calculated and measured L\(_1\)-edge and L\(_3\)-edge EXAFS spectra shown in Fig. 4a-e and Fig. 4f-j, respectively. Two different methods are used to calculate the EXAFS spectra from the atomic positions of molecular configurations extracted from either MD or PIMD trajectories. The first method (solid line) uses more accurate representations of the Cs\(^+\) and O muffin tin potentials within the FEFF9 EXAFS calculation by selecting only the Cs\(^+\) and O positions (i.e., H atoms are not included in the calculations), while the second method (dashed lines) employs the default settings that also includes the H positions (see Supporting Information for specific details). The TIP4P-Ew model provides the poorest performance in reproducing both L\(_1\)-edge and L\(_3\)-edge spectra, predicting more pronounced oscillations whose amplitudes are too high and that are slightly out of phase with the experiment at mid- to high-\(k\) values. This reflects
Figure 4: Comparisons between L₁-edge (panels a-e) and L₃-edge (panels f-j) EXAFS spectra, \( k^2\chi(k) \), calculated from MD simulations with the TIP4P-Ew, TTM-nrg, (2B+NB)-MB-nrg and (2B+3B+NB)-MB-nrg models as well as from PIMD simulations with the (2B+3B+NB)-MB-nrg model with the corresponding experimental data. See main text for details. Two sets of theoretical spectra are shown, corresponding to FEFF9 calculations carried out on molecular configurations that include only Cs⁺ and O atoms (solid lines) and all atoms (dashed lines). See main text for details.

the narrow width of the first peak in the Cs⁺–O RDF and the shorter Cs⁺–O distance predicted by the TIP4P-Ew model (Fig. 2a). The inclusion of an implicit representation of many-body effects through classical polarization as implemented in the TTM-nrg model results in EXAFS spectra that more closely follow the oscillations in the corresponding experimental traces, although the calculated amplitudes of the oscillations are somewhat lower than in the measurements. The low amplitude and the frequency mismatch for the TTM-nrg model are progressively improved as 2B and 3B effects in the underlying Cs⁺–water interactions are taken explicitly into account in the simulations with the (2B+NB)-MB-nrg and (2B+3B+NB)-MB-nrg models, which lead to excellent agreement with the experimental data for both L₁-edge and L₃-edge spectra. As expected from the analysis of the RDFs shown in Fig. 2, nuclear quantum effects appear to have only minimal impact on the EXAFS spectra.

In conclusion, the role played by many-body effects in the hydration structure of Cs⁺ has been investigated through a systematic analysis of predictions obtained from MD and PIMD simulations carried out with four different molecular models built upon a hierarchy of approx-
imate representations of Cs$^+$–water interactions. As expected from the simpler functional form, the empirical, pairwise-additive TIP4P-Ew model provides the poorest agreement with the available EXAFS spectra, which suggests a physically incorrect description of the underlying hydration structure. Although inclusion of many-body effects through an implicit representation based on classical polarization as implemented in the TTM-nrg model improves the agreement with experiment, the predicted Cs$^+$ hydration structure appears to lack the correct sequence of well-defined water shells. Systematic inclusion of explicit many-body Cs$^+$–water interactions in the MB-nrg models as well as account of nuclear quantum effects in PIMD simulations progressively leads to nearly quantitative agreement with the experimental EXAFS spectra. While emphasizing the importance of a physically correct representation of many-body contributions to Cs$^+$–water interactions at both short and long ranges, the present analysis also suggests that PEFs rigorously built upon many-body expansions hold great promise for “realistic” simulations of aqueous solutions.

Acknowledgements

This research was supported by the National Science Foundation through grant CHE-1453204 and used resources of the Extreme Science and Engineering Discovery Environment (XSEDE),\textsuperscript{47} which is supported by the National Science Foundation through grant ACI-1053575, under allocation TG-CHE110009. M.R.R. was supported by a Software Fellowship from the Molecular Sciences Software Institute, which is funded by the National Science Foundation through grant ACI-1547580. Work by J.L.F. and G.K.S. was supported by the U.S. Department of Energy (DOE), Office of Science, Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences & Biosciences. Pacific Northwest National Laboratory (PNNL) is a multiprogram national laboratory operated for DOE by Battelle. The Cs$^+$ L$_3$ XAFS measurements were performed at the PHOENIX beamline of the Swiss Light Source, Paul Scherrer Institute, Villigen, Switzerland. The Cs$^+$ L$_1$ measurements at Advanced Photon
Source, Sector 20 were supported by the U.S. Department of Energy and the Canadian Light Source.

References


(7) Endo, M. Calcium Ion as a Second Messenger with Special Reference to Excitation-Contraction Coupling. J. Pharm. Sci. 2006, 100, 519–524.


