Accurate Modeling of Bromide and Iodide Hydration with Data-Driven Many-Body Potentials

Alessandro Caruso,[*,†] Xuanyu Zhu,[†] 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, Washington 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: acaruso@ucsd.edu; fpaesani@ucsd.edu

Abstract

Ion–water interactions play a central role in determining the properties of aqueous systems in a wide range of environments. However, a quantitative understanding of how the hydration properties of ions evolve from small aqueous clusters to bulk solutions and interfaces remains elusive. Here, we introduce the second generation of data-driven many-body energy (MB-nrg) potential energy functions (PEFs) representing bromide–water and iodide–water interactions. The MB-nrg PEFs use permutationally invariant polynomials to reproduce two-body and three-body energies calculated at the coupled cluster level of theory, and implicitly represent all higher-body energies using classical many-body polarization. A systematic analysis of the hydration structure of small $Br^-(H_2O)_n$ and $I^-(H_2O)_n$ clusters demonstrates that the MB-nrg PEFs predict interaction energies in quantitative agreement with "gold standard" coupled cluster reference values. Importantly, when used in molecular dynamics simulations carried out in the isothermal-isobaric ensemble for single bromide and iodide ions in liquid water, the MB-nrg PEFs predict extended X-ray absorption fine structure (EXAFS) spectra that accurately reproduce the experimental spectra, which thus allows for characterizing the hydration structure of the two ions with high level of confidence.

Introduction

Halide ions are among the most studied electrolytes, both in experiments and in simulations, due to their role in various natural and industrial processes.[1–4] They are frequently encountered in biological systems[5] and found to play important roles in electrochemistry,[6] and environmental chemistry.[7–9] Moreover, due to their spherical symmetry and short-lived interactions with water molecules, halide ions are frequently used as benchmarks in fundamental studies of hydration phenomena of negative ions.[10]

Developing a molecular-level understanding of the properties of hydrated ions from small gas-phase clusters to aqueous solutions poses several challenges to both experiment and simulation.[10–22] Specific ion effects on the local structure of the hydrogen-bond (H-bond) network of water have been the focus of extensive investigations.[10,23–32] It is established that the presence of ionic species in aqueous environments gives rise to structural rearrangements of the water H-bond network. However, the extent of these rearrangements remains matter of debate. Until recently, the common view has been to classify ions as "structure makers" or "structure breakers" depending on whether they strengthen or disrupt the H-bond network of the surrounding water molecules. According to this classification, "structure makers" correspond to ions with high charge density while "structure breakers" are ions with small and diffuse charge density. This classification, inspired by Hofmeister's series on protein stability,[33] has often appeared to be too simplistic and has been repeatedly challenged by experimental measurements.[16,29,30,34,35]

Several theoretical and computational studies of the hydration properties of bromide and iodide ions based on either force fields (FFs) or density functional theory (DFT) have been reported in the literature, starting from pioneering simulations showing enhanced propensity of these ions for the water surface.[23–26,36] A polarizable FF derived from first principles[37] was used to characterize the hydration structure and calculate the extended X-ray absorption fine structure (EXAFS) spectrum of bromide in solution which was found to be in qualitative agreement with the corresponding experimental data.[38] Molecular dynamics (MD) simulations carried out with the AMOEBA polarizable force field found that the interactions among water molecules in the first solvation shell

around a bromide ion are similar to those in pure bulk water.[39] A systematic analysis of the hydration properties in ion–water clusters, including $Br^-(H_2O)_n$ clusters, carried out at the DFT level found that the differences in dipole moments between molecules residing inside and outside of the first hydration shell of the ion become smaller as the cluster size increases, which was interpreted as evidence in support of the use of nonpolarizable FFs in MD simulations.[40] Car-Parrinello MD simulations of bromide in water carried out with the BLYP functional were used to calculate the EXAFS spectrum that was found to be in better agreement with the experimental data than the analogous spectrum calculated using the TIP3P and OPLS models.[41] The BLYP functional was also used in DFT and quantum mechanics/molecular mechanics (QM/MM) simulations of iodide in water.[42] Somewhat surprisingly, it was found that the EXAFS spectrum calculated from the QM/MM simulations was in better agreement with the experimental data than the analogous spectrum calculated from the DFT simulations. DFT simulations carried out with the BLYP-D functional were used to calculate the EXAFS spectra of iodide in water which were found to be in qualitative agreement with the experimental data, displaying some differences in both phases and amplitudes.[43] More recently, polarizable FFs based on the BK3 water model were used to determine the hydration structure of both halide and alkali-metal ions.[44] A subsequent study found that these BK3-based FFs overpredict water structuring around the ions in solution and concluded that the ion–water interactions are not adequately represented by these FFs.[45] MD simulations carried out with the ONIOM-XS approach found that bromide–water interactions in solution are weak and give rise to a loosely bound first shell.[46]

In previous studies,[47–50] we introduced the many-body energy (MB-nrg) theoretical/computational framework for data-driven many-body potential energy functions (PEFs) that are rigorously derived from the many-body expansion (MBE) of the energy calculated using coupled cluster theory, including single, double, and perturbative triple excitations, i.e., CCSD(T), which is currently considered as the "gold standard" for molecular interactions.[51] When used to model halide–water interactions, the MB-nrg PEFs were shown to provide high accuracy, quantitatively reproducing the energetics of small $X^-(H_2O)_N$ clusters, with X = F, Cl, Br, and I,[52,53] as well as the vibrational

spectra and tunneling splitting of halide monohydrate[54] and dihydrate complexes.[32,53] More recently, we demonstrated that the original chloride–water MB-nrg PEF[47] could be improved by refining the training set of the 2-body energies via active learning[55] and including an explicit 3-body energy term.[56] The new chloride–water MB-nrg PEF was shown to achieve CCSD(T) accuracy in representing the interaction energies of $Cl^-(H_2O)_n$ clusters and, when used in MD simulations, it predicted the correct hydration structure of chloride in water as demonstrated by the quantitative agreement between the experimental and calculated extended EXAFS spectra.[56]

In this study, we continue our analysis of many-body effects in ion hydration by introducing second-generation bromide–water and iodide–water MB-nrg PEFs. The article is structured as follows: in the "Methods" section, we present the new MB-nrg PEFs that are developed from expanded training sets for 2-body and 3-body energies generated using the active learning scheme of Ref. 55. We assess the overall accuracy of the new MB-nrg PEFs in the "Results" section by analyzing their ability to reproduce the energetics of small $Br^-(H_2O)_n$ and $I^-(H_2O)_n$ clusters as well as the hydration structure of each ion in solution. In the "Conclusion", we summarize our work and provide an outlook for potential future applications of the MB-nrg PEFs.

Methods

MB-nrg PEFs

The MB-nrg PEFs for bromide and iodide in water were developed following Ref. 56. Within the MBE, the energy of a system is obtained as the sum of individual n-body energy terms, $\varepsilon^{\mathrm{nB}}$, according to

$$E_N(1,..,N) = \sum_{i=1}^{N} \varepsilon^{1\mathrm{B}}(i) + \sum_{i<j}^{N} \varepsilon^{2\mathrm{B}}(i,j) + \sum_{i<j<k}^{N} \varepsilon^{3\mathrm{B}}(i,j,k) + ... + \varepsilon^{\mathrm{NB}}(1,..,N), \qquad (1)$$

Since the MB-nrg theoretical/computational framework has already been described in the literature,[47–50,57] we will only describe here the salient features and provide details specific to the

bromide–water and iodide–water PEFs. Briefly, the MB-nrg PEFs use a combination of short-range permutationally invariant polynomials (PIPs)[58] trained on electronic structure data and physics-based functions to represent the 1B, 2B, and 3B terms of the MBE in Eq. 1,[57,59] while all other nB terms with $n > 3$ are represented by an implicit many-body term derived from the Thole model of classical polarization.[60]

Within the MB-nrg theoretical/computational framework, Eq. 1 is thus expressed as

$$E_N = \sum_{i=1}^{N} \varepsilon^{1\mathrm{B}}(i) + \sum_{i>j}^{N} \varepsilon^{2\mathrm{B}}(i,j) + \sum_{i>j>k}^{N} \varepsilon^{3\mathrm{B}}(i,j,k) + V_{\mathrm{pol}}^{>3\mathrm{B}}, \tag{2}$$

where $\varepsilon^{1\mathrm{B}}(i)$ is the distortion energy of the ith monomer in the system, and all other nB terms are defined recursively as

$$\begin{aligned} \varepsilon^{\mathrm{nB}}(1,\ldots,n) &= E_n(1,\ldots,n) - \sum_i \varepsilon^{1\mathrm{B}}(i) - \sum_{i<j} \varepsilon^{2\mathrm{B}}(i,j) - \ldots \\ &- \sum_{i<j<k} \varepsilon^{3\mathrm{B}}(i,j,k) - \ldots - \varepsilon^{(\mathrm{n}-1)\mathrm{B}}(1,\ldots,n-1) \end{aligned} \tag{3}$$

Here, $\varepsilon^{\mathrm{nB}}$ is the n-body energy, and $\mathrm{E}_n(1,\ldots,n)$ is the energy of a subsystem containing n monomers. Since the halide ions are monoatomic species, the 1-body term of the halide–water MB-nrg PEFs contains contributions only from the intramolecular distortions of the water molecules, which are described by the MB-pol PEF[61–63] through the model developed by Partridge and Schwenke.[64]

In Eq. 2, the 2-body energy, $\varepsilon^{2\mathrm{B}}$, takes the form

$$\varepsilon^{2\mathrm{B}} = V_{\mathrm{sr}}^{2\mathrm{B}} + V_{\mathrm{elec}}^{2\mathrm{B}} + V_{\mathrm{disp}}^{2\mathrm{B}} + V_{\mathrm{pol}}^{2\mathrm{B}} \tag{4}$$

where $V_{\mathrm{sr}}^{2\mathrm{B}}$ is a short-range term expressed by a PIP that is fitted to reproduce CCSD(T) 2-body energies and switched off when the distance between the halide ion (X) and the oxygen atom (O) of the water molecule within the dimer is larger than a predefined cutoff,

$$V_{\mathrm{sr}}^{2\mathrm{B}} = s_2\left(R_{\mathrm{X^-O}}\right) \cdot V_{\mathrm{PIP}}^{2\mathrm{B}} \tag{5}$$

The switching function, $s_2\left(R_{\mathrm{X^-O}}\right)$, is given by[56]

$$s_2(\mathrm{R_{X^-O}}) = \begin{cases} 1, & \text{if } t_2(R_{\mathrm{X^-O}}) < 0 \\ cos^2\left[t_2(R_{\mathrm{X^-O}})\pi/2\right], & \text{if } 0 \le t_2(R_{\mathrm{X^-O}}) < 1 \\ 0, & \text{if } 1 \le t_2(R_{\mathrm{X^-O}}) \end{cases} \tag{6}$$

with

$$t_2(R_{\mathrm{X^-O}}) = \frac{R_{\mathrm{X^-O}} - R_{\mathrm{in}}^{\mathrm{2B}}}{R_{\mathrm{out}}^{\mathrm{2B}} - R_{\mathrm{in}}^{\mathrm{2B}}}. \tag{7}$$

$R_{\mathrm{in}}^{\mathrm{2B}}$ and $R_{\mathrm{out}}^{\mathrm{2B}}$ are the inner and outer cutoffs of $s_2\left(R_{\mathrm{X^-O}}\right)$ which are chosen in order to ensure a smooth and continuous variation of $\varepsilon^{\mathrm{2B}}$ in the switching region. ($R_{\mathrm{in}}^{\mathrm{2B}}$, $R_{\mathrm{out}}^{\mathrm{2B}}$) = (5.9 Å, 7.9 Å) and (6.2 Å, 8.2 Å) for the bromide–water and iodide–water MB-nrg PEFs, respectively.

V_{elec} in Eq. 4 represents electrostatic interactions between the negative (-1 e) charge of the halide ions and the geometry-dependent point charges of the water molecule which are obtained by fitting the *ab initio* dipole moment of an isolated water molecule calculated in Ref. 64.

$V_{\mathrm{disp}}^{\mathrm{2B}}$ in Eq. 4 represents the 2-body dispersion energy and is expressed as

$$V_{\mathrm{disp}}^{\mathrm{2B}} = -f(\delta_{\mathrm{X^-O}})\frac{C_{6,\mathrm{X^-O}}}{R_{\mathrm{X^-O}}^6} - f(\delta_{\mathrm{X^-H_1}})\frac{C_{6,\mathrm{X^-H_1}}}{R_{\mathrm{X^-H_1}}^6} - f(\delta_{\mathrm{X^-H_2}})\frac{C_{6,\mathrm{X^-H_2}}}{R_{\mathrm{X^-H_2}}^6} \tag{8}$$

where $R_{\mathrm{X^-O}}$, $R_{\mathrm{X^-H_1}}$, and $R_{\mathrm{X^-H_2}}$ are the distances between the ion ($\mathrm{X^-}$), and the oxygen (O) and the two hydrogen (H) atoms of the water molecule within a $\mathrm{X^-}$–$\mathrm{H_2O}$ dimer, and $f(\delta)$ and C_6 are the corresponding Tang-Toennies damping functions[65] and dispersion coefficients determined in Ref. 47.

The 3-body energy in Eq. 2, $\varepsilon^{\mathrm{3B}}$, is given by

$$\varepsilon^{\mathrm{3B}} = V_{\mathrm{sr}}^{\mathrm{3B}} + V_{\mathrm{pol}}^{\mathrm{3B}} \tag{9}$$

As in $\varepsilon^{\mathrm{2B}}$ of Eq. 4, $V_{\mathrm{sr}}^{\mathrm{3B}}$ is a short-range term expressed by a PIP that is fitted to reproduce CCSD(T)

3-body energies and switched off when two or more of the X-O and O-O distances are larger than a predefined cutoff value according to:

$$\begin{aligned} V_{\rm sr}^{\rm 3B} = [&s_3(R_{\rm X^-O_a})s_3(R_{\rm X^-O_b}) \\ &+ s_3(R_{\rm X^-O_a})s_3(R_{\rm O_aO_b}) \\ &+ s_3(R_{\rm X^-O_b})s_3(R_{\rm O_aO_b}) \\ &- 2\cdot s_3(R_{\rm X^-O_a})s_3(R_{\rm X^-O_b})s_3(R_{\rm O_aO_b})]\cdot V_{\rm PIP}^{\rm 3B} \end{aligned} \tag{10}$$

A combination of switching functions $s_3(R_{kl})$ acting on each (X,O) and (O,O) pair is used to guarantee that $\varepsilon^{\rm 3B}$ transitions smoothly between $(V_{\rm PIP}^{\rm 3B}+V_{\rm pol}^{\rm 3B})$ at short range and $V_{\rm pol}^{\rm 3B}$ at long range, with $s_3(R_{kl})$ given by

$$s_3(R_{kl}) = \begin{cases} 1, & \text{if } t_3(R_{kl}) < 0 \\ cos^2\left[t_3(R_{kl})\pi/2\right], & \text{if } 0 \le t_3(R_{kl}) < 1 \\ 0, & \text{if } 1 \le t_3(R_{kl}) \end{cases} \tag{11}$$

Here, the variable $t_3(R_{kl})$ depends on the inner ($R_{\rm in}^{\rm 3B}$) and outer ($R_{\rm out}^{\rm 3B}$) cutoff values according to:

$$t_3(R_{kl}) = \frac{R_{kl}-R_i^{\rm 3B}}{R_{\rm out}^{\rm 3B}-R_{in}^{\rm 3B}} \tag{12}$$

As discussed in Ref. 56, the MB-nrg framework provides the user with complete freedom for the choice of the inner and outer cutoffs. To account for the relatively large size and polarizability of the two halide ions, ($R_{\rm in}^{\rm 3B}$ $R_{\rm out}^{\rm 3B}$) were set equal to (3.9 Å, 5.9 Å) and (5.5 Å, 6.0 Å) for bromide and iodide, respectively.

Finally, $V_{\rm pol}^{\rm 2B}$ in Eq. 4, $V_{\rm pol}^{\rm 3B}$ in Eq. 9, and $V_{\rm pol}^{>\rm 3B}$ in Eq. 2 are implicitly included in a classical N-body polarization term, $V_{\rm pol}^{\rm NB}$, derived from the Thole model.[60] The effective atomic polarizabilities, $\alpha^{\rm eff}$, for the bromide and iodide ions in water were determined from exchange-dipole moment (XDM)[66,67] calculations carried out with Gaussian 16[68] and postg[69,70] for $Br^-(H_2O)_n$

and $I^-(H_2O)_n$ clusters. Specifically, clusters of increasingly larger radius were extracted from MD simulations for systems containing a single ion and 277 water molecules carried out in the isothermal-isobaric (NPT) ensemble at 298 K and 1 atm using the TTM-nrg PEFs.[71] For each radius, 20 clusters were randomly selected and the corresponding effective atomic polarizabilities were determined as[72]

$$\alpha^{\text{eff}} = \alpha^{\text{free}} \left(\frac{V^{\text{eff}}}{V^{\text{free}}} \right)^{4/3} \tag{13}$$

Here V^{eff} and V^{free} are the effective and free volumes of the halide ions, respectively, obtained from the XDM calculations. Fig. 1 shows the average values of α^{eff} for the bromide and ions as a function of the cluster size. The bulk values of α^{eff} were determined from the asymptotic limit of the two curves, which results in the values of 3.7819 Å^3 and 5.9563 Å^3 for Br^- and I^-, respectively.

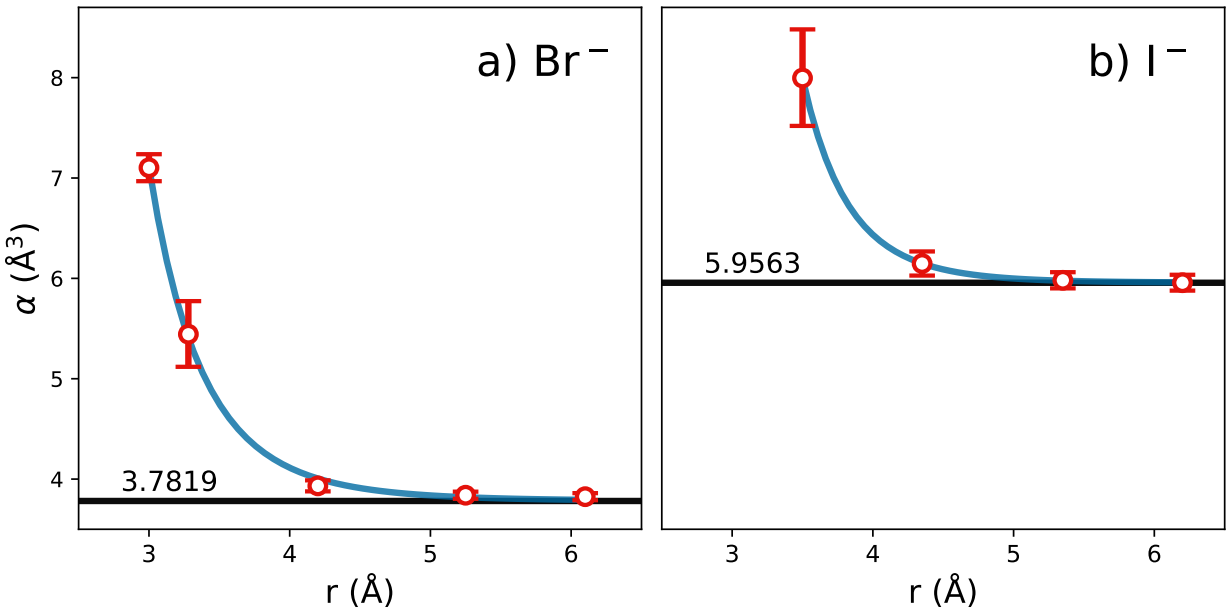

Figure 1: Variation of bromide (panel a) and iodide (panel b) polarizabilities calculated with XDM as a function of the radius (r) of the corresponding $Br^-(H_2O)_n$ and $Br^-(H_2O)_n$ clusters. The error bars are determined as 95% confidence interval.

Pemutationally invariant polynomials

Both $V_{\mathrm{PIP}}^{\mathrm{2B}}$ and $V_{\mathrm{PIP}}^{\mathrm{3B}}$ are functions of the pairwise distances between the ion (X), the hydrogen and oxygen atoms (H and O), and the lone-pair sites of the MB-pol water molecules (L_1 and L_2).[61] $V_{\mathrm{PIP}}^{\mathrm{2B}}$ contains 496 symmetrized monomials (ξ_i): 3 first-degree monomials, 15 second-degree monomials, 49 third-degree monomials, 130 fourth-degree monomials, and 299 fifth degree monomials. $V_{\mathrm{PIP}}^{\mathrm{2B}}$ thus contains 496 linear fitting parameters (c_i) and 9 nonlinear fitting parameters.[47,56] $V_{\mathrm{PIP}}^{\mathrm{3B}}$ contains 1575 symmetrized monomials, ξ_i: 39 second-degree monomials, 613 third-degree monomials, and 923 fourth-degree monomials. Therefore, $V_{\mathrm{PIP}}^{\mathrm{3B}}$ contains 1575 linear fitting parameters and 13 nonlinear fitting parameters.

Fitting procedure

As in MB-pol[61,62] and other MB-nrg PEFs,[47–50,56,73] the linear parameters of the 2- and 3-body PIPs were fitted through singular value decomposition while the simplex algorithm was used for the non-linear parameters.

The regularized weighted sum of squared deviations, χ^2, was minimized over the training set $\mathscr{S}$, while the L linear parameters were regularized with $\Gamma = 0.0005$:

$$\chi^2 = \sum_{n\in\mathscr{S}} w_n[\varepsilon_{\mathrm{model}}(n) - \varepsilon_{\mathrm{ref}}(n)]^2 + \Gamma^2 \sum_{l=1}^{L} c_l^2 \tag{14}$$

Here, w_n are weights to emphasize low binding energy configurations according to

$$w(E_i) = \left[\frac{\Delta E}{E_i - E_{\mathrm{min}} + \Delta E}\right]^2. \tag{15}$$

where E_{min} is the lowest binding energy in $\mathscr{S}$ and ΔE is the range of favorable configurations which was set to 30 kcal/mol and 42.5 kcal/mol for the 2-body and 3-body energies, respectively. The fitting process was performed using a development version of the MB-Fit software.[57,74]

Reference energies

The training sets for the 2-body energies were generated using the method described in Refs. 55 and 56. A pool of $\sim$ 150000 configurations was generated from different sources: a spherical grid with the water molecule in a $\sim 2-8$ Å shell from the ion, normal modes of the ion–water dimer, and MD simulations carried out in periodic boundary conditions using the TTM-nrg PEFs[71] for a box containing a single ion and 277 water molecules (see Section "Molecular dynamics simulations". The reference 2-body energies were obtained at the CCSD(T)-F12b level of theory[75,76] in the complete basis set (CBS) limit that was achieved via a two-point extrapolation[77,78] between 2-body energies calculated with the augmented correlation-consistent polarized valence triple- (aug-cc-pVTZ) and quadruple-ζ (aug-cc-pVQZ) basis sets.[79–82] The final 2-body training sets consist of 17057 bromide–water and 15810 iodide–water dimers, while the corresponding test sets consist of 1795 and 1668 dimer configurations, respectively.

3-body training sets consisting of 33830 $Br^-(H_2O)_2$ and 33985 $I^-(H_2O)_3$ trimers were also generated from MD simulations carried with the TTM-nrg PEFs.[71] The corresponding tests consist of 3576 and 3621 trimer configurations, respectively. The 3-body energies were calculated at the CCSD(T)-F12b level of theory[75,76] using the aug-cc-pVTZ basis set.[79–82] All CCSD(T)-F12b calculations were carried out using MOLPRO (version 2020.1).[83]

Molecular dynamics simulations

All the MD simulations were carried out in the NPT ensemble at 298.15 K and 1.0 atm for an orthorhombic box containing a single ion and 277 water molecules, which corresponds to a concentration of $\sim$0.2 M concentration. The velocity-Verlet algorithm was used to propagate the equations of motion with a time step of 0.5 fs according to Ref. 84. The temperature and pressure were maintained using a global Nosé–Hoover chain of 3 thermostats with a relaxation time of 0.05 ps, and a global Nosé–Hoover barostat with a relaxation time of 0.5 ps which was thermosttated by a chain of three thermostats. The NPT simulations consisted of 0.1 ns of equilibration followed by 1 ns of production. Short-range interactions were evaluated with a real-space cutoff of 9 Å,

while long-range interactions (including electrostatic, dispersion, and polarization contributions) were calculated in reciprocal space using a particle–particle particle–mesh solver.[85] All MD simulations were carried out using the Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS)[86] package interfaced with the MBX software for many-body PEFs.[87]

Extended X-ray absorption spectroscopy

Two different EXAFS analysis methods were used to compare the experimentally measured structures with the structures obtained from MD simulations. The first method is exact and consists in generating an ensemble average of EXAFS spectra calculated for a set of snapshots extracted from the simulated trajectory (MD-EXAFS). The MD-EXAFS method has previously been described in Ref. 88. From the equilibrated portion of the MD trajectory, 2000 equally spaced frames are selected and the Cartesian coordinates of the halide ion, and oxygen and hydrogen atoms of the water molecules are retrieved. Each set is used as input to the EXAFS scattering code, FEFF9,[89–91] in order to generate the ensemble average $\chi(k)$ spectra. As in Refs. 92 and 56, all FEFF calculations were performed using clusters containing the halide ion and its 33 closest water molecules, extracted from the corresponding NPT trajectories.

The second method involves fitting a small set of theoretical standards to measured or calculated spectra. The EXAFS analysis software Artemis was used for this approach.[93] The method adopts a Gaussian model and thus approximates the position of each atom with a normal distribution centered on its average value. The limitations of this approach have been discussed with respect to its application to disordered systems.[94,95] Fitting to the theoretical standards requires some a priori insight into the chemical makeup of the system in order to judiciously select a set of the most important nearby neighbor atoms for which 3 or 4 scattering paths are created. Then using a least-squares fitting procedure, refinements to coordination numbers, distances, and disorder are used to provide a best-fit to the measured or simulated spectrum.

Results

To assess the accuracy of the 2-body terms of the MB-nrg PEFs, Fig. 2 shows correlation plots between the CCSD(T)-F12b/CBS reference 2-body energies and the corresponding MB-nrg values calculated for the bromide–water and iodide–water dimers in the test sets. The associated root mean square errors (RMSEs) for both training and test sets are reported in Table 1. Both MB-nrg PEFs achieve CCSD(T)/CBS accuracy over the entire energy range from -25 kcal/mol to 105 kcal/mol. The ability of the MB-nrg PEFs to provide a high-fidelity representation of the CCSD(T)/CBS dimer multidimensional energy landscape is further demonstrated in Fig. 3 that shows comparisons between CCSD(T)/CBS and MB-nrg one-dimensional potential energy radial scans calculated for various orientations (θ, ϕ) of each ion relative to the water molecule within a ion–water dimer.

Fig. 4 shows correlation plots between the CCSD(T)-F12b/CBS reference 3-body energies and the corresponding MB-nrg values calculated for the bromide–water and iodide–water trimers in the

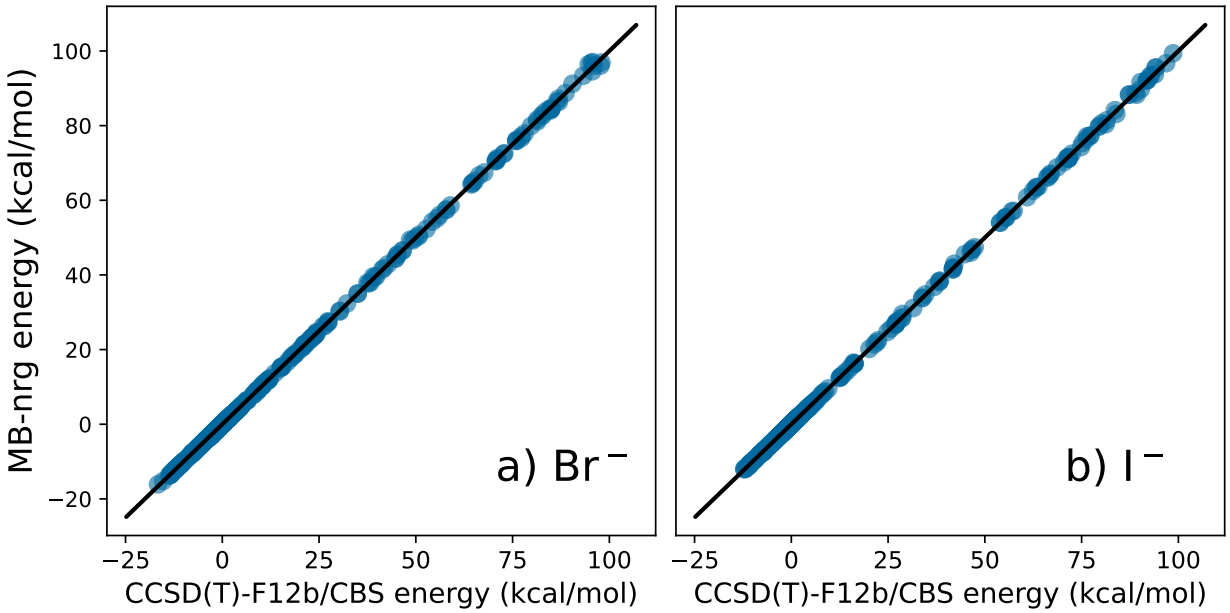

Figure 2: 2-body energy correlation plots between the CCSD(T)-F12b/CBS reference values (x axis) and corresponding MB-nrg values (y axis) for the bromide–water (panel a) and iodide–water (panel b) test sets.

Table 1: Root mean square errors (RMSEs) associated with bromide–water and iodide–water 2-body and 3-body energies calculated with the MB-nrg PEFs relative to the corresponding CCSD(T)-F12b/CBS reference values of the training and test sets.

MB-nrg PEF	2-body RMSE (kcal/mol)		3-body RMSE (kcal/mol)	
	Training	Test	Training	Test
bromide–water	0.1947	0.1871	0.0404	0.0470
iodide–water	0.2010	0.2287	0.0513	0.0677

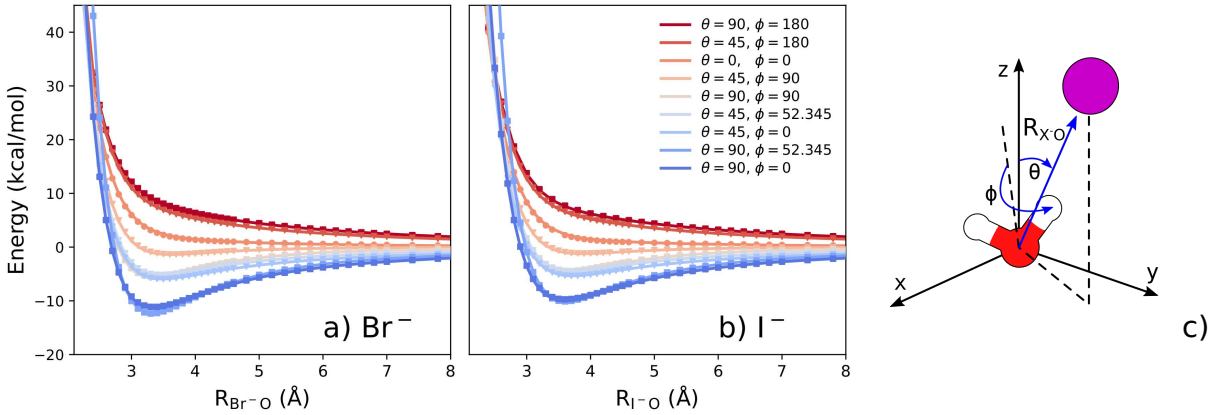

Figure 3: Interaction energy scans along the X^-–O distance (R_{X^-O}) for selected orientations (θ, ϕ) of the halide ion relative to the water molecule in a $X^-(H_2O)$ dimer, with X = Br (panel a) and I (panel b). R_{X^-O}, θ, and ϕ are defined in panel c). The symbols correspond to the CCSD(T)-F12b/CBS reference interaction energies, while the corresponding MB-nrg values are shown as solid lines.

test sets, while the RMSEs calculated for both training and test sets are reported in Table 1. The comparisons shown in Figs. 2 and 4 demonstrate that the both bromide–water and iodide–water MB-nrg PEFs are able to quantitatively reproduce the corresponding CCSD(T)/CBS 2-body and 3-body energies, without overfitting.

While the accuracy exhibited by the 2-body and 3-body terms of the bromide–water and iodide–water PEFs is certainly remarkable, it is also somewhat expected since these terms are explicitly fitted to reproduce the corresponding CCSD(T)-F12b/CBS reference energies. In this context, one of the most arduous challenges for data-driven PEFs is achieving full transferability across phases and/or thermodynamic state points different from those represented in the training sets. To address this challenge, we first analyze the ability of the MB-nrg PEFs to correctly reproduce the CCSD(T)-F12b/CBS interaction energies of small $X^-(H_2O)_n$ clusters shown in Fig. 5. It is im-

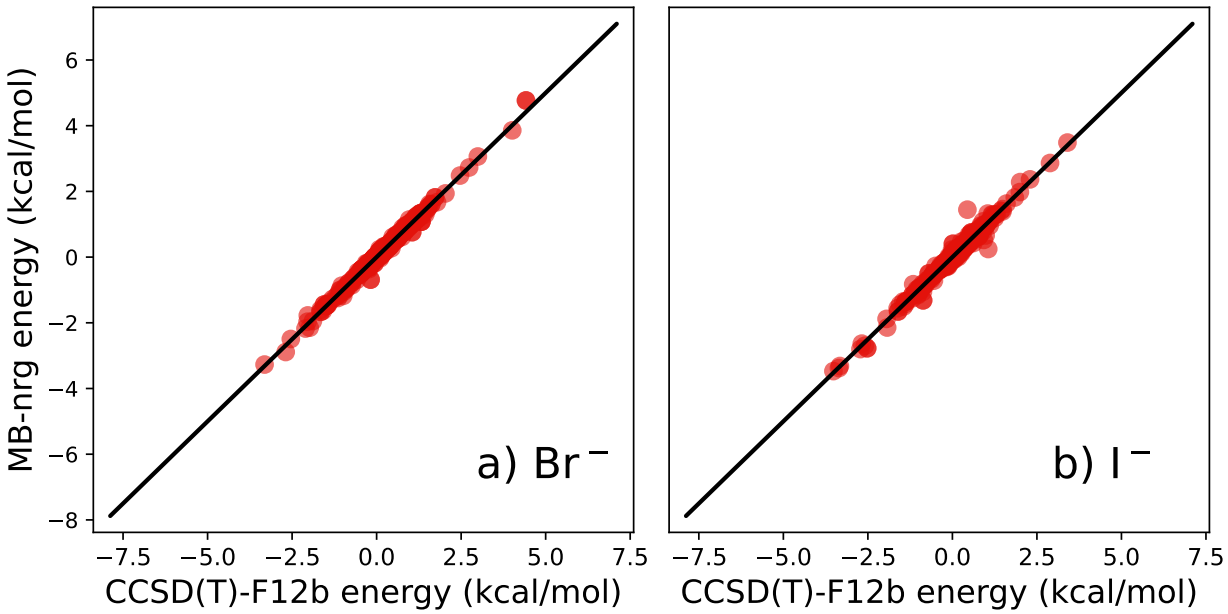

Figure 4: 3-body energy correlation plots between the CCSD(T)-F12b/CBS reference values (x axis) and corresponding MB-nrg values (y axis) for the bromide–water (panel a) and iodide–water (panel b) test sets.

portant to emphasize that the calculations carried out for systems containing a single ion and more than two water molecules correspond to actual predictions since, by construction, the MB-nrg PEFs were only trained up to the 3-body energies in the trimers and all higher many-body terms are represented by classical polarization. Fig. 6 shows comparisons between the CCSD(T)-F12b/CBS reference interaction energies[52,53] and the corresponding values calculated with the MB-nrg PEFs for 15 different $X^-(H_2O)_n$ clusters (with $n = 1-4$), with X = Br and I. Besides the full MB-nrg PEFs, hereafter referred to as (2B+3B+NB)-MB-nrg PEFs, which include explicit PIP-based representations of both 2-body and 3-body energies (Eqs. 6-14), Fig. 6 also shows results obtained with MB-nrg PEFs that include PIP terms only for the 2-body energies and are hereafter referred to as (2B+NB)-MB-nrg PEFs. The comparisons in Fig. 6, therefore, allow for assessing not only the overall accuracy of the full MB-nrg PEFs but also the relative importance of 2-body, 3-body, and higher n-body contributions to the interactions energies of larger bromide–water and iodide–water systems. The (2B+3B+NB)-MB-nrg PEFs quantitatively reproduce the CCSD(T)-F12b/CBS reference energies of both $Br^-(H_2O)_n$ and $I^-(H_2O)_n$, independently of the cluster size and structure.

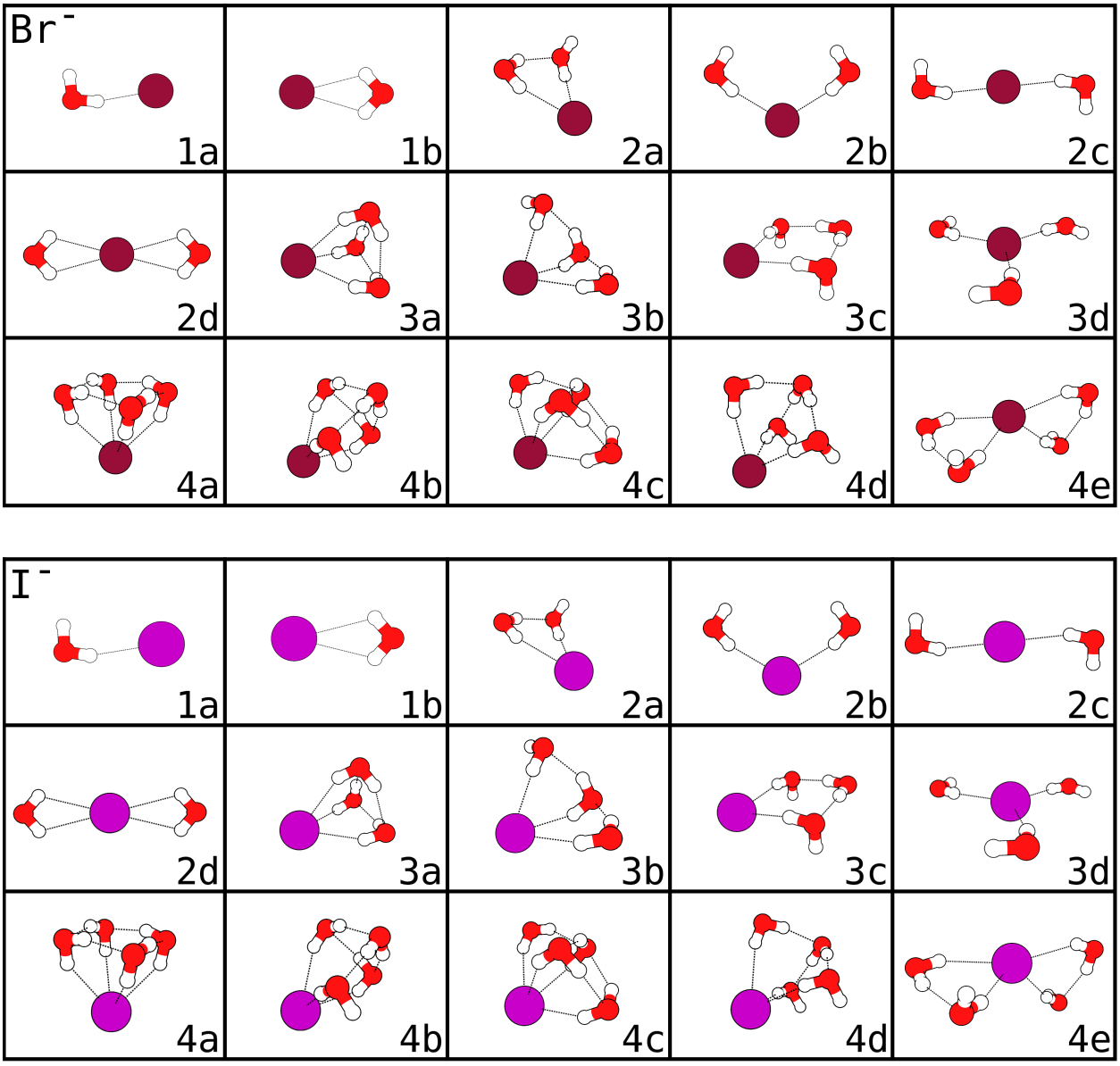

Figure 5: Low-lying energy isomers of $Br^-(H_2O)_n$ (top panels) and $I^-(H_2O)_n$ (bottom panels) clusters ($n = 1-4$).

Importantly, the performance of the (2B+3B+NB)-MB-nrg and (2B+NB)-MB-nrg PEFs is very similar, with only small differences found for the isomers with relatively higher interaction energy (e.g., isomers 3d and 4e), which indicates that n-body energy contributions with $n > 2$ are primarily due to classical electrostatic interactions.

The last challenge that remains to be addressed in order to assess the transferability of the MB-nrg PEFs is to determine if the high accuracy displayed in predicting the interaction energies of

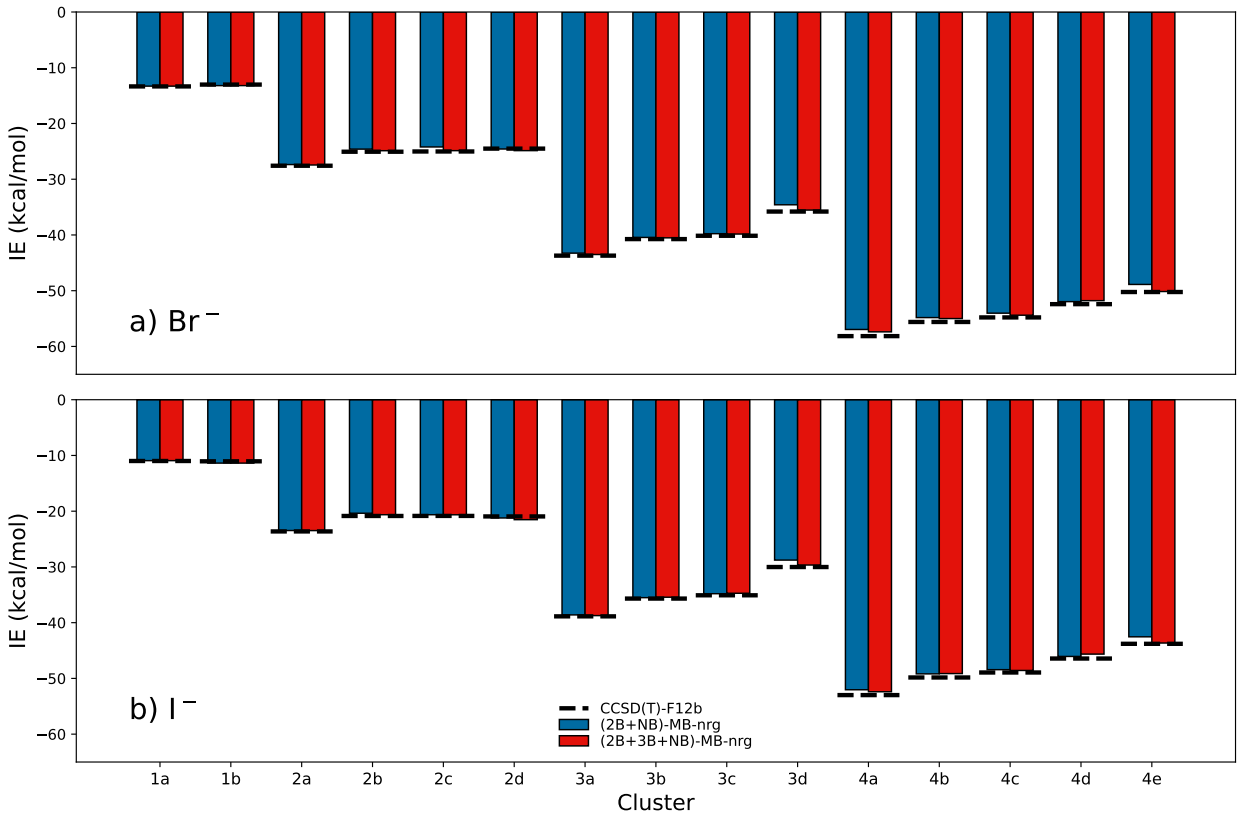

Figure 6: Comparison between the interaction energies calculated for the low-energy isomers of $Br^-(H_2O)_n$ (panel a) and $I^-(H_2O)_n$ (panel b) clusters ($n = 1-4$) using the (2B+NB)-MB-nrg and (2B+3B+NB)-MB-nrg PEFs. For each cluster, the CCSD(T)-F12b reference values[52,53] are shown as horizontal dashed lines.

small clusters in the gas phase translates into realistic descriptions of the hydration structure of the bromide and iodide ions in solution at finite temperature.

In this regard, Fig. 7 shows the radial distribution functions (RDFs) and corresponding coordination numbers calculated from NPT simulations carried out with both (2B+NB)-MB-nrg and (2B+3B+NB)-MB-nrg PEFs at 298 K and 1 atm for a box containing a single ion and 277 water molecules in periodic boundary conditions. The RDFs describing spatial 2-body correlations between the Br^- ion and the oxygen (O) atoms of the water molecules exhibit a well-define hydration sttructure, with a prominent, first-shell peak at $\sim$3.4 Å. Both (2B+NB)-MB-nrg and (2B+3B+NB)-MB-nrg PEFs effectively predict the same position and shape for this first peak. Some minor differences exist between the two MB-nrg PEFs in the region of the second, broader peak cor-

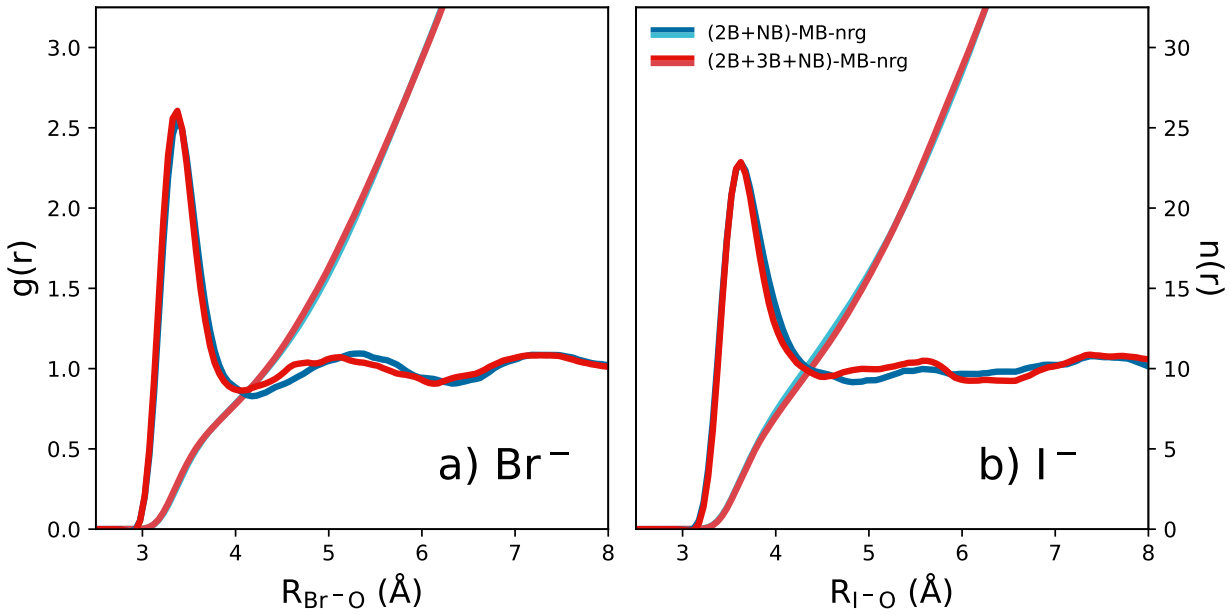

Figure 7: Bromide-oxygen (panel a) and iodide-oxygen (panel b) radial distribution functions, $g(r)$, and corresponding coordination numbers, $n(r)$, calculated from NPT simulations carried out at 298 K and 1 atm with the (2B+NB)-MB-nrg (blue) and (2B+3B+NB)-MB-nrg (red) PEFs.

responding to the second hydration shell that extends between ∼4.2 Å and ∼6.2 Å. Specifically, the (2B+3B+NB)-MB-nrg PEF predicts a small shift of the second hydration shell towards shorter distances, signalling relatively stronger bromide–water interactions. The variation of the water coordination number as a function of the distance from the bromide ion shows an inflection point at ∼4.2 Å, indicating that the first hydration shell contains ∼ 7-8 water molecules.

The iodide–oxygen RDFs predicted by both (2B+NB)-MB-nrg and (2B+3B+NB)-MB-nrg PEFs shows a first peak similar to that observed in the corresponding bromide–oxygen RDFs, but at slightly larger distances due to the larger size of the iodide ion. However, the evolution of the subsequent hydration shells around iodide is appreciably different. In particular, both (2B+NB)-MB-nrg and (2B+3B+NB)-MB-nrg PEFs predict a shallow hydration structure beyond 4.5 Å, with the presence of a second and third hydration shell only barely visible in the (2B+3B+NB)-MB-nrg PEF. In this regard, it should be noted that the presence of the "kink" at ∼5.5 Å in the iodide-oxygen RDF calculated with the (2B+3B+NB)-MB-nrg PEF might also be due to a less-than-perfect transition from the data-driven component (i.e., PIPs + polarization) to the purely classical component (i.e.,

polarization) of the 3-body term in Eq. 9. The role of the switching functions in the 2-body (Eq. 5) and 3-body (Eq. 10) terms of the MB-nrg potentials will be the subject of a forthcoming study.

As expected from the shape of the iodide-oxygen RDFs calculated with the (2B+NB)-MB-nrg and (2B+3B+NB)-MB-nrg PEFs, the water coordination number calculated with both PEFs effectively shows a monotonic increase as a function of the distance from the iodide ion, with only a weak inflection point at $\sim$4.5 Å which precludes a precise determination of the first-shell coordination number. The differences between the bromide–oxygen and iodide–oxygen RDFs primarily arise from the competition between 2-body halide–water and water–water interactions that are further modulated by 3-body iodide–water–water interactions. As shown in Refs. 32, the strength of 2-body halide–water interactions decreases with the size of the halide ions and becomes comparable to the strength of 2-body water–water interactions in the case of the iodide ion. The competition between iodide–water and water–water interactions thus results in a shallower iodide–oxygen RDF beyond the first hydration shell. When compared to the theoretical models of Ref. 43, the MB-nrg model predicts a nearly identical first peak. However, DFT-based MD and the Dang and Chang (D/C) model predict less interstitial water between the first and the second solvation shell, as reported in the Supporting Information.

While the analysis of the RDFs discussed above allows for gaining insights into the hydra-

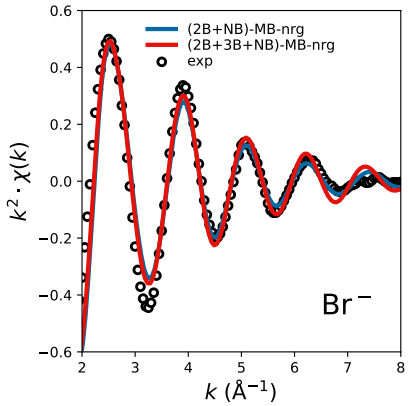

Figure 8: K-edge EXAFS spectrum, $k^2\chi(k)$, of bromide in water calculated from NPT simulations carried out at 298 K and 1 atm with the (2B+NB)-MB-nrg and (2B+3B+NB)-MB-nrg PEFs. The experimental EXAFS spectrum from Ref. 96 is shown as black circles.

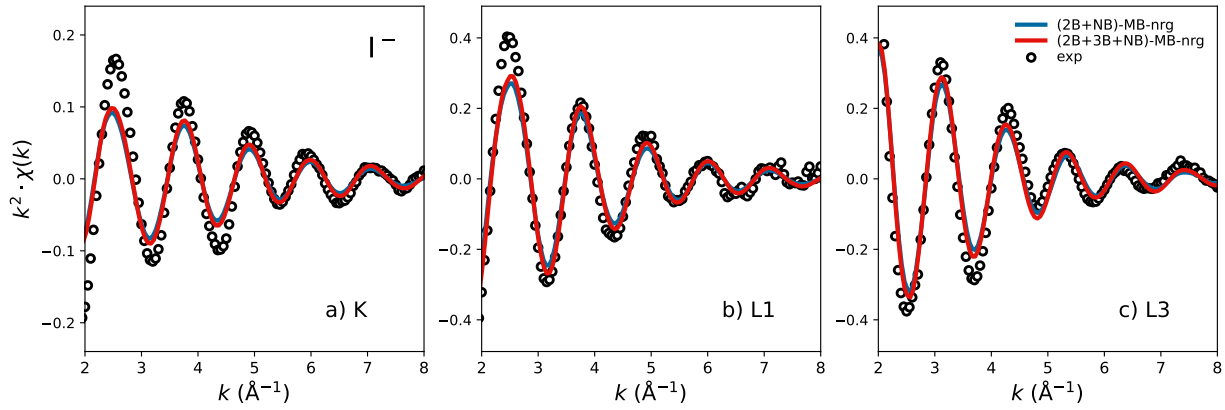

Figure 9: K-edge (panel a), L1-edge (panel b), and L3-edge (panel c) EXAFS spectra, $k^2\chi(k)$, of iodide in water calculated from NPT simulations carried out at 298 K and 1 atm with the (2B+NB)-MB-nrg and (2B+3B+NB)-MB-nrg PEFs. The corresponding experimental EXAFS spectra from Ref. 43 are shown as black circles.

tion structure of bromide and iodide ions in solution, it does not provide any evidence for the accuracy and realism of the RDFs predicted by the MB-nrg PEFs. To address this last challenge, in Figs. 8 and 9, we show comparisons between the experimental and simulated EXAFS spectra, calculated with the MD-EXAFS methodology introduced in the Methods section, for bromide and iodide in water, respectively. The K-edge spectra calculated with the (2B+NB)-MB-nrg and (2B+3B+NB)-MB-nrg PEFs for bromide in water are effectively indistinguishable from each other and in quantitative agreement with the experimental spectrum, with only small deviations for k in the 3-3.5 Å^{-1} range. Similar agreement between the experimental and MB-nrg results was reported for the K-edge spectrum of chloride in water.[56]

In Fig. 9, we report results relative to three different absorption edges for iodide (K-, L1-, and L3-edges). With regard to the EXAFS single scattering paths, there is a simple 90° phase shift in the $\chi(k)$ oscillations from the K- or L1- edge spectra (1s and 2s initial states, respectively) with respect to those observed for the L3-edge (2p initial state). However, the symmetry selection rules dictate that the K-, L1-, and L3-edge spectra provide independent, non-redundant, measurements of the local symmetry with respect to multiple scattering paths.[97]

All three EXAFS spectra calculated with the iodide–water MB-nrg PEFs are also in good agreement with the experimental data, with the agreement improving as k increases. However, while the

phase of each spectrum is quantitatively reproduced by the MB-nrg PEFs over the entire range of k values, some discrepancies in the amplitudes of the oscillations exist at small values of k, especially in the case of the K-edge spectrum. Interestingly, the differences seen in the second shells of the RDFs predicted by the (2B+NB)-MB-nrg and (2B+3B+NB)-MB-nrg PEFs appear to only have minimal effects on the corresponding EXAFS spectra. This provides further evidence for the local nature of the EXAFS measurements that are primarily sensitive to the first hydration shell. The amplitude discrepancies for the K- and L1-edge spectra (1s and 2s initial states, respectively), especially in the region around 2.5 Å^{-1} and 4.2 Å^{-1}, are due to multi-electron excitations that are part of the atomic background function present in the experimental spectra.[43,98] Importantly, the multi-electron excitations for the L3-edge (2p initial state) have different distributions and intensities of transitions. The complete set of K-, L1- and L3-edge atomic-background multi-electron excitation features are also observed in the residuals for the model fit to the experimental spectra that is shown in the $\chi(k)$ spectra of Fig. S6a in the Supporting Information. In the light of these moderate distortions of the experimental EXAFS spectra, the MB-nrg provides excellent overall agreement.

For comparison, EXAFS spectra calculated only including water molecules within the the first hydration shell of the halide ions are shown in Figs. S2 and S3 in the Supporting Information. Although very similar to those calculated by including 33 water molecules, i.e., including water molecules residing beyond the first hydration shell, the EXAFS spectra calculated with only water molecules within the first solvation shell show small differences in the amplitudes as well as at large k values for both halide ions. This suggests that, when possible, including a larger number of water molecules beyond the first hydration shell in the FEFF calculations may lead to better converged EXAFS spectra. For comparisons, the EXAFS spectra calculated for I^- using DFT-based simulations as well as simulations with the D/C polarizable model reported in Ref. 43 are shown in Fig. S4.

It is important to emphasize that all the structural details of the RDFs are encoded into the various regions of the calculated EXAFS spectra. For instance, the EXAFS spectra contain infor-

Table 2: Top: Comparison of the structural parameters from K-edge fits to experimental EXAFS gathered from Ref. 96 and MD-EXAFS from this work for the aqueous Br^- first-shell structure. Bottom: Comparison of the structural parameters from simultaneous K-, L1-, and L3-edge fits to experimental EXAFS and MD-EXAFS gathered from Ref. 43 with MD-EXAFS fits from this work for the aqueous I^- first-shell structure. N is the coordination number, R refers to the measured X-H and X-O distances, σ^2 is the Debye-Waller factor, $\phi_{\text{X-H-O}}$ is the X-H-O angle, and the subscript XHO refers to three-leg I-H-O paths. The goodness of the fit, $\mathscr{R}$, is calculated as the sum of the errors squared scaled by the magnitude of the data. See main text for details.

System	Structure						
	Scatterer	N*	R (Å)	$\sigma^2 \times 10^3$ (Å^2)	$\sigma^2_{\text{BrHO}} \times 10^3$ (Å^2)	$\phi_{\text{Br-H-O}}$	$\mathscr{R}$
0.5m RbBr	O_{H_2O}	6	3.274(024)	13.0(4.2)	12.4(2.0)	158°	0.020
	H_{H_2O}	6	2.357(049)	21.6(9.6)			
MB-nrg MD	O_{H_2O}	6	3.392(017)	13.2(2.4)	18.4(4.8)	154°	0.016
	H_{H_2O}	6	2.501(037)	27.1(11.5)			
System	Structure						
	Scatterer	N	R (Å)	$\sigma^2 \times 10^3$ (Å^2)	$\sigma^2_{\text{IHO}} \times 10^3$ (Å^2)	$\phi_{\text{I-H-O}}$	$\mathscr{R}$
0.4m NaI	O_{H_2O}	6.3(0.9)	3.498(025)	17.1(4.7)	20.9(6.8)	148°	0.046
	H_{H_2O}	6.3(0.9)	2.649(028)	35.6(4.5)			
DFT MD	O_{H_2O}	6.0(0.4)	3.584(007)	14.4(1.1)	15.9(1.6)	156°	0.008
	H_{H_2O}	6.0(0.4)	2.686(026)	36.6(5.1)			
D/C MD	O_{H_2O}	5.0(0.4)	3.526(007)	12.7(1.0)	21.4(2.7)	163°	0.013
	H_{H_2O}	5.0(0.4)	2.597(017)	22.7(2.7)			
MB-nrg MD	O_{H_2O}	5.6(0.5)	3.569(010)	16.6(1.4)	19.1(1.0)	147°	0.011
	H_{H_2O}	5.6(0.5)	2.723(029)	34.9(5.4)			

* Fixed parameter.

mation about the rather broad and asymmetric first peak in the calculated RDFs for which both halide ions show a relatively large amount of disorder. Furthermore, the EXAFS spectra are sensitive to interstitial water molecules that reside in the region from 4 to 4.5 Å. At larger distances, however, the lack of any or the presence of very weak structures makes detection more difficult. The signals $\chi(k)$ also contain angular correlation components that are a feature of photoelectron multiple scattering paths.[43]

In order to explore these characteristics of the EXAFS spectra, in Table 2 we report the structural parameters obtained from fitting the K-edge of bromide and simultaneously fitting the K-, L1-, and L3-edge of iodide to the theoretical standards, using the second methodology discussed in the Methods section. Specifically, listed in Table 2 are the coordination number N, the measured

I-H and I-O distances R, the Debye-Waller factor σ^2, and the X-H-O angle $\phi_{\text{X-H-O}}$. The resulting fits are shown in Figs. S5 and S6 of the Supporting Information. For both bromide and iodide (X = Br, I), the X-O, X-H, and X-H-O scattering paths were selected. The bromide K-edge EXAFS data were weighted by k^2, and fitted over the range $1.8 < k < 10.0$ Å^{-1}. Both the real and imaginary parts of $\chi(R)$ are in the region of $1.25 < R < 4.5$ Å. A value $S_0^2 = 0.91$ of the core hole factor was used. Given the small number of independent points collected, the coordination number was kept fixed at 6 for bromide; for the same reason, the fitted parameters generally show a larger error when compared to iodide. In the case of iodide, the K-, L1- and L3-edge experimental data were weighted by k^3, and fitted simultaneously over the ranges $1.9 < k < 8.8$, $1.9 < k < 8.6$, and $1.9 < k < 7.8$ Å^{-1}, respectively. In all cases, both the real and imaginary parts of the resulting $\chi(R)$ are in the region of $1.25 < R < 4.5$ Å. A value $S_0^2 = 1.0$ of the core hole factor was used. It is important to note that, since the exact same fitting model was applied to the experimental, DFT-based, D/C model, and MB-nrg model spectra, their differences can be quantitatively compared. As shown in Table 2, the I-O and I-H distance from the MB-nrg model fit are within 0.07 Å of the experimental values. The Debye Waller factors for I-O, I-H, and I-H-O from MB-nrg are identical to those of the experimental values within fitting errors; this is especially striking when compared to other reported MD-EXAFS results. Similar agreement is observed for the bromide-water structure. Moreover, the MB-nrg iodide-water model is capable of faithfully reproducing the I-H-O angle. The small differences observed in the coordination number are within the error recorded for the experimental fit.

Conclusions

In this work, we have introduced second-generation bromide–water and iodide–water MB-nrg PEFs. Within the MB-nrg theoretical/computational framework, the two MB-nrg PEFs are derived from the MBE of the energy and includes explicit data-driven 2-body and 3-body energy terms along with an implicit term describing all n-body energy contributions with $n > 3$. The 2-body and

3-body terms are represented by PIP optimized to reproduce the corresponding CCSD(T)-F12b reference data, while the implicit term is represented by classical many-body polarization.

The two MB-nrg PEFs are able to quantitatively reproduce the interaction energies of $Br^-(H_2O)_n$ and $I^-(H_2O)_n$, effectively achieving CCSD(T)-F12b accuracy in all cases examined in this study. A systematic analysis of the interaction energies show that 3-body energy contributions are primarily due to classical polarization. However, the inclusion of an explicit, short-range 3-body term is shown to be important for retrieving specific features of the hydration shells around the two ions in solution. The second solvation shell is found to be particularly sensitive to 3-body interactions: while in the case of bromide the shell slightly shifts to shorter distances when the short-range PIP is included in the MB-nrg PEF, neglecting short-range iodide-water-water interactions leads to a shallow and nearly featureless iodide–oxygen radial distribution function beyond the first hydration shell.

The structural features that characterize the hydration structure of bromide and iodide in solution predicted by the MB-nrg PEFs are confirmed by the agreement between the experimental and simulated EXAFS spectra. It should, however, be noted that while the phases of the EXAFS spectra simulated with the MB-nrg PEFs correctly reproduce the experimental values, small variations in the amplitudes, especially in the case of the K-edge of iodide, exist, which may be due to inaccuracies in the MB-nrg PEFs and/or multielectron scattering effects that are not accounted for in the simulated spectra.

We believe that the results presented in this study further demonstrate the ability of the MB-nrg PEFs to correctly predict the physics of hydrated halide ions, providing a quantitative representation of halide–water interactions from the gas to the condensed phase and enabling affordable MD simulations of ionic aqueous solutions with CCSD(T) accuracy.

Supporting Information

Additional analyses of the EXAFS spectra and comparisons with previous simulations of iodide in solution carried out using a polarizable force field as well as a DFT model.

Acknowledgement

We thank Greg Schenter for stimulating discussions about the calculation and interpretation of the EXAFS spectra. Work by A.C., X.Z, and F.P. was supported by the National Science Foundation through Grant No. CHE-1453204. Work by J.L.F was supported under project 16248, funded by the U.S. Department of Energy (DOE), Office of Science, Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences, and Biosciences. The simulations used resources of the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by the National Science Foundation through Grant No. ACI-1053575, and the Triton Shared Computing Cluster (TSCC) at the San Diego Supercomputer Center.

Data availability

Any data generated and analyzed in this study are available from the authors upon request. The MB-nrg PEFs used in this study are available in our open-access MBX software that can be downloaded from https://github.com/paesanilab/MBX.

References

(1) Marcus, Y. *Ion Solvation*; Wiley: Chichester, 1985.

(2) El-Dessouky, H. T.; Ettouney, H. M. *Fundamentals of Salt Water Desalination*; Elsevier: Amsterdam, 2002.

(3) Grebel, J. E.; Pignatello, J. J.; Mitch, W. A. Effect of Halide Ions and Carbonates on Organic Contaminant Degradation by Hydroxyl Radical-Based Advanced Oxidation Processes in Saline Waters. *Environ. Sci. Technol.* **2010**, *44*, 6822–6828.

(4) Ghosh, S.; Manna, L. The Many "Facets" of Halide Ions in the Chemistry of Colloidal Inorganic Nanocrystals. *Chem. Rev.* **2018**, *118*, 7804–7864.

(5) Duan, D.; Winter, C.; Cowley, S.; Hume, J. R.; Horowitz, B. Molecular Identification of a Volume-Regulated Chloride Channel. *Nature* **1997**, *390*, 417–421.

(6) Winter, M.; Brodd, R. J. What Are Batteries, Fuel Cells, and Supercapacitors? *Chem. Rev.* **2004**, *104*, 4245–4270.

(7) Harrison, R. G.; Tammet, H. Ions in the Terrestrial Atmosphere and Other Solar System Atmospheres. *Space Sci. Rev.* **2008**, *137*, 107–118.

(8) Tobias, D. J.; Stern, A. C.; Baer, M. D.; Levin, Y.; Mundy, C. J. Simulation and Theory of Ions at Atmospherically Relevant Aqueous Liquid-Air Interfaces. *Annu. Rev. Phys. Chem.* **2013**, *64*, 339–359.

(9) Shuman, N. S.; Hunton, D. E.; Viggiano, A. A. Ambient and Modified Atmospheric Ion Chemistry: From Top to Bottom. *Chem. Rev.* **2015**, *115*, 4542–4570.

(10) Robertson, W. H.; Johnson, M. A. Molecular Aspects of Halide Ion Hydration: The Cluster Approach. *Annu. Rev. Phys. Chem.* **2003**, *54*, 173.

(11) Ohtaki, H.; Radnai, T. Structure and Dynamics of Hydrated Ions. *Chem. Rev.* **1993**, *93*, 1157–1204.

(12) Liu, D.; Ma, G.; Levering, L. M.; Allen, H. C. Vibrational Spectroscopy of Aqueous Sodium Halide Solutions and Air-Liquid Interfaces: Observation of Increased Interfacial Depth. *J. Phys. Chem. B* **2004**, *108*, 2252–2260.

(13) Jungwirth, P.; Tobias, D. J. Specific Ion Effects at the Air/Water Interface. *Chem. Rev.* **2006**, *106*, 1259–1281.

(14) Petersen, P. B.; Saykally, R. J. On the Nature of Ions at the Liquid Water Surface. *Annu. Rev. Phys. Chem.* **2006**, *57*, 333–364.

(15) Fennell, C. J.; Bizjak, A.; Vlachy, V.; Dill, K. A. Ion Pairing in Molecular Simulations of Aqueous Alkali Halide Solutions. *J. Phys. Chem. B* **2009**, *113*, 6782–6791.

(16) Funkner, S.; Niehues, G.; Schmidt, D. A.; Heyden, M.; Schwaab, G.; Callahan, K. M.; Tobias, D. J.; Havenith, M. Watching the Low-Frequency Motions in Aqueous Salt Solutions: The Terahertz Vibrational Signatures of Hydrated ions. *J. Am. Chem. Soc.* **2012**, *134*, 1030–1035.

(17) Wolk, A. B.; Leavitt, C. M.; Garand, E.; Johnson, M. A. Cryogenic Ion Chemistry and Spectroscopy. *Acc. Chem. Res.* **2014**, *47*, 202–210.

(18) Piatkowski, L.; Zhang, Z.; Backus, E. H.; Bakker, H. J.; Bonn, M. Extreme Surface Propensity of Halide Ions in Water. *Nat. Commun.* **2014**, *5*, 1–7.

(19) Van Der Vegt, N. F.; Haldrup, K.; Roke, S.; Zheng, J.; Lund, M.; Bakker, H. J. Water-Mediated Ion Pairing: Occurrence and Relevance. *Chem. Reviews* **2016**, *116*, 7626–7641.

(20) Chen, Y.; Okur, H. I.; Gomopoulos, N.; Macias-Romero, C.; Cremer, P. S.; Petersen, P. B.; Tocci, G.; Wilkins, D. M.; Liang, C.; Ceriotti, M. et al. Electrolytes Induce Long-Range Orientational Order and Free Energy Changes in the H-bond Network of Bulk Water. *Sci. Adv.* **2016**, *2*, e1501891.

(21) Duignan, T. T.; Kathmann, S. M.; Schenter, G. K.; Mundy, C. J. Toward a First-Principles Framework for Predicting Collective Properties of Electrolytes. *Acc. Chem. Res.* **2021**, *54*, 2833–2843.

(22) Weitzner, S. E.; Pham, T. A.; Orme, C. A.; Qiu, S. R.; Wood, B. C. Beyond Thermodynamics: Assessing the Dynamical Softness of Hydrated Ions from First Principles. *J. Phys. Chem. Lett.* **2021**, *12*, 11980–11986.

(23) Perera, L.; Berkowitz, M. L. Many-Body Effects in Molecular Dynamics Simulations of $Na^{+}(H_2O)_n$ and $Cl^{-}(H_2O)_n$ Clusters. *J. Chem. Phys.* **1991**, *95*, 1954–1963.

(24) Perera, L.; Berkowitz, M. L. Structure and Dynamics of $Cl^{-}(H_2O)_{20}$ Clusters: The Effect of the Polarizability and the Charge of the Ion. *J. Chem. Phys.* **1992**, *96*, 8288–8294.

(25) Dang, L. X.; Smith, D. E. Molecular Dynamics Simulations of Aqueous Ionic Clusters Using Polarizable Water. *J. Chem. Phys.* **1993**, *99*, 6950–6956.

(26) Perera, L.; Berkowitz, M. L. Structures of $Cl^{-}(H_2O)_n$ and $F^{-}(H_2O)_n$ (n= 2, 3,..., 15) Clusters. Molecular Dynamics Computer Simulations. *J. Chem. Phys.* **1994**, *100*, 3085–3093.

(27) Combariza, J. E.; Kestner, N. R.; Jortner, J. Energy-Structure Relationships for Microscopic Solvation of Anions in Water Clusters. *J. Chem. Phys.* **1994**, *100*, 2851–2864.

(28) Xantheas, S. S. Quantitative Description of Hydrogen Bonding in Chloride-Water Clusters. *J. Phys. Chem.* **1996**, *100*, 9703–9713.

(29) Cappa, C. D.; Smith, J. D.; Wilson, K. R.; Messer, B. M.; Gilles, M. K.; Cohen, R. C.; Saykally, R. J. Effects of Alkali Metal Halide Salts on the Hydrogen Bond Network of Liquid Water. *J. Phys. Chem. B* **2005**, *109*, 7046–7052.

(30) Schmidt, D. A.; Birer, O.; Funkner, S.; Born, B. P.; Gnanasekaran, R.; Schwaab, G. W.; Leitner, D. M.; Havenith, M. Rattling in the Cage: Ions as Probes of Sub-Picosecond Water Network Dynamics. *J. Am. Chem. Soc.* **2009**, *131*, 18512–18517.

(31) Shalit, A.; Ahmed, S.; Savolainen, J.; Hamm, P. Terahertz Echoes Reveal the Inhomogeneity of Aqueous Salt Solutions. *Nat. Chem.* **2017**, *9*, 273–278.

(32) Bajaj, P.; Zhuang, D.; Paesani, F. Specific Ion Effects on Hydrogen-Bond Rearrangements in the Halide–Dihydrate Complexes. *J. Phys. Chem. Lett.* **2019**, *10*, 2823–2828.

(33) Hofmeister, F. Zur Lehre von der Wirkung der Salze. *Naunyn-Schmiedeberg's Arch. Pharmacol.* **1888**, *24*, 247–260.

(34) Smith, J. D.; Saykally, R. J.; Geissler, P. L. The Effects of Dissolved Halide Anions on Hydrogen Bonding in Liquid Water. *J. Am. Chem. Soc.* **2007**, *129*, 13847–13856.

(35) Waluyo, I.; Nordlund, D.; Bergmann, U.; Schlesinger, D.; Pettersson, L. G.; Nilsson, A. A Different View of Structure-Making and Structure-Breaking in Alkali Halide Aqueous Solutions through X-ray Absorption Spectroscopy. *J. Chem. Phys.* **2014**, *140*, 244506.

(36) Perera, L.; Berkowitz, M. L. Stabilization Energies of Cl^-, Br^-, and I^- Ions in Water Clusters. *J. Chem. Phys.* **1993**, *99*, 4222–4224.

(37) Ayala, R.; Martínez, J. M.; Pappalardo, R. R.; Sánchez Marcos, E. On the Halide Hydration Study: Development of First-Principles Halide Ion-Water Interaction Potential Based on a Polarizable Model. *J. Chem. Phys.* **2003**, *119*, 9538–9548.

(38) Merkling, P. J.; Ayala, R.; Martínez, J. M.; Pappalardo, R. R.; Sánchez Marcos, E. Interplay of Computer Simulations and X-ray Absorption Spectra in the Study of the Bromide Hydration Structure. *J. Chem. Phys.* **2003**, *119*, 6647–6654.

(39) Grossfield, A. Dependence of Ion Hydration on the Sign of the Ion's Charge. *J. Chem. Phys.* **2005**, *122*, 024506.

(40) Krekeler, C.; Hess, B.; Delle Site, L. Density Functional Study of Ion Hydration for the Alkali Metal Ions (Li^+, Na^+, K^+) and the Halide Ions (F^-, Br^-, Cl^-). *J. Chem. Phys.* **2006**, *125*, 054305.

(41) D'Angelo, P.; Migliorati, V.; Guidoni, L. Hydration Properties of the Bromide Aqua Ion:

The Interplay of First Principle and Classical Molecular Dynamics, and X-ray Absorption Spectroscopy. *Inorg. Chem.* **2010**, *49*, 4224–4231.

(42) Pham, V.; Tavernelli, I.; Milne, C.; van der Veen, R.; D'Angelo, P.; Bressler, C.; Chergui, M. The Solvent Shell Structure of Aqueous Iodide: X-ray Absorption Spectroscopy and Classical, Hybrid QM/MM and Full Quantum Molecular Dynamics Simulations. *Chem. Phys.* **2010**, *371*, 24–29.

(43) Fulton, J. L.; Schenter, G. K.; Baer, M. D.; Mundy, C. J.; Dang, L. X.; Balasubramanian, M. Probing the Hydration Structure of Polarizable Halides: A Multiedge XAFS and Molecular Dynamics Study of the Iodide Anion. *J. Phys. Chem. B* **2010**, *114*, 12926–12937.

(44) Kiss, P. T.; Baranyai, A. A New Polarizable Force Field for Alkali and Halide Ions. *J. Chem. Phys.* **2014**, *141*, 114501.

(45) Yue, S.; Panagiotopoulos, A. Z. Dynamic Properties of Aqueous Electrolyte Solutions from Non-Polarisable, Polarisable, and Scaled-Charge Models. *Mol. Phys.* **2019**, *117*, 3538–3549.

(46) Sripa, P.; Tongraar, A. The "Surface"(S) State of the Br^- Hydration: An ONIOM-XS MD Simulation Study. *Chem. Phys. Lett.* **2020**, *738*, 136853.

(47) Bajaj, P.; Götz, A. W.; Paesani, F. Toward Chemical Accuracy in the Description of Ion–Water Interactions through Many-Body Representations. I. Halide–Water Dimer Potential Energy Surfaces. *J. Chem. Theory. Comput.* **2016**, *12*, 2698–2705.

(48) Riera, M.; Mardirossian, N.; Bajaj, P.; Götz, A. W.; Paesani, F. Toward Chemical Accuracy in the Description of Ion–Water Interactions through Many-Body Representations. Alkali-Water Dimer Potential Energy Surfaces. *J. Chem. Phys.* **2017**, *147*, 161715.

(49) Riera, M.; Yeh, E. P.; Paesani, F. Data-Driven Many-Body Models for Molecular Fluids: CO_2/H_2O Mixtures as a Case Study. *J. Chem. Theory. Comput.* **2020**, *16*, 2246–2257.

(50) Riera, M.; Hirales, A.; Ghosh, R.; Paesani, F. Data-Driven Many-Body Models with Chemical Accuracy for CH_4/H_2O Mixtures. *J. Phys. Chem. B* **2020**, *124*, 11207–11221.

(51) Rezac, J.; Hobza, P. Benchmark Calculations of Interaction Energies in Noncovalent Complexes and their Applications. *Chem. Rev.* **2016**, *116*, 5038–5071.

(52) Paesani, F.; Bajaj, P.; Riera, M. Chemical Accuracy in Modeling Halide Ion Hydration from Many-Body Representations. *Adv. Phys. X* **2019**, *4*, 1631212.

(53) Bajaj, P.; Riera, M.; Lin, J. K.; Mendoza Montijo, Y. E.; Gazca, J.; Paesani, F. Halide Ion Microhydration: Structure, Energetics, and Spectroscopy of Small Halide–Water Clusters. *J. Phys. Chem. A* **2019**, *123*, 2843–2852.

(54) Bajaj, P.; Wang, X.-G.; CarringtonJr, T.; Paesani, F. Vibrational spectra of halide-water dimers: Insights on Ion Hydration from Full-Dimensional Quantum Calculations on Many-Body Potential Energy Surfaces. *J. Chem. Phys.* **2017**, *148*, 102321.

(55) Zhai, Y.; Caruso, A.; Gao, S.; Paesani, F. Active Learning of Many-Body Configuration Space: Application to the Cs^+–Water MB-nrg Potential Energy Function as a Case Study. *J. Chem. Phys.* **2020**, *152*, 144103.

(56) Caruso, A.; Paesani, F. Data-Driven Many-Body Models Enable a Quantitative Description of Chloride Hydration from Clusters to Bulk. *J. Chem. Phys.* **2021**, *155*, 064502.

(57) Bull-Vulpe, E. F.; Riera, M.; Götz, A. W.; Paesani, F. MB-Fit: Software Infrastructure for Data-Driven Many-Body Potential Energy Functions. *J. Chem. Phys.* **2021**, *155*, 124801.

(58) Braams, B. J.; Bowman, J. M. Permutationally Invariant Potential Energy Surfaces in High Dimensionality. *Int. Rev. Phys. Chem.* **2009**, *28*, 577–606.

(59) Lambros, E.; Dasgupta, S.; Palos, E.; Swee, S.; Hu, J.; Paesani, F. General Many-Body Framework for Data-Driven Potentials with Arbitrary Quantum Mechanical Accuracy: Water as a Case Study. *J. Chem. Theory Comput.* **2021**, *17*, 5635–5650.

(60) Thole, B. T. Molecular Polarizabilities Calculated with a Modified Dipole Interaction. *Chem. Phys.* **1981**, *59*, 341–350.

(61) Babin, V.; Leforestier, C.; Paesani, F. Development of a "First Principles" Water Potential with Flexible Monomers: Dimer Potential Energy Surface, VRT Spectrum, and Second Virial Coefficient. *J. Chem. Theory. Comput.* **2013**, *9*, 5395–5403.

(62) Babin, V.; Medders, G. R.; Paesani, F. Development of a "First Principles" Water Potential with Flexible Monomers. II: Trimer Potential Energy Surface, Third Virial Coefficient, and Small Clusters. *J. Chem. Theory. Comput.* **2014**, *10*, 1599–1607.

(63) Medders, G. R.; Babin, V.; Paesani, F. Development of a "First Principles" Water Potential with Flexible Monomers. III. Liquid Phase Properties. *J. Chem. Theory. Comput.* **2014**, *10*, 2906–2910.

(64) Partridge, H.; Schwenke, D. W. The Determination of an Accurate Isotope Dependent Potential Energy Surface for Water From Extensive Ab Initio Calculations and Experimental Data. *J. Chem. Phys.* **1997**, *106*, 4618–4639.

(65) Tang, K.; Toennies, J. P. An Improved Simple Model for the van der Waals Potential Based on Universal Damping Functions for the Dispersion Coefficients. *J. Chem. Phys.* **1984**, *80*, 3726–3741.

(66) Becke, A. D.; Johnson, E. R. Exchange-Hole Dipole Moment and the Dispersion Interaction Revisited. *J. Chem. Phys.* **2007**, *127*, 154108.

(67) Becke, A. D.; Johnson, E. R. A Unified Density-Functional Treatment of Dynamical, Nondynamical, and Dispersion Correlations. *J. Chem. Phys.* **2007**, *127*, 124108.

(68) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H. et al. Gaussian 16 Revision C.01. 2016; Gaussian Inc. Wallingford CT.

(69) Otero-De-La-Roza, A.; Johnson, E. R. Non-Covalent Interactions and Thermochemistry Using XDM-Corrected Hybrid and Range-Separated Hybrid Density Functionals. *J. Chem. Phys.* **2013**, *138*, 204109.

(70) Kannemann, F. O.; Becke, A. D. Van der Waals Interactions in Density-Functional Theory: Intermolecular Complexes. *J. Chem. Theory Comput.* **2010**, *6*, 1081–1088.

(71) Arismendi-Arrieta, D. J.; Riera, M.; Bajaj, P.; Prosmiti, R.; Paesani, F. i-TTM Model for Ab Initio-Based Ion–Water Interaction Potentials. 1. Halide–Water Potential Energy Functions. *J. Phys. Chem. B* **2016**, *120*, 1822–1832.

(72) Gobre, V. V. *Efficient Modelling of Linear Electronic Polarization in Materials Using Atomic Response Functions*; Technische Universitaet Berlin (Germany), 2016.

(73) Cruzeiro, V. W. D.; Lambros, E.; Riera, M.; Roy, R.; Paesani, F.; Götz, A. W. Highly Accurate Many-Body Potentials for Simulations of N_2O_5 in Water: Benchmarks, Development, and Validation. *J. Chem. Theory Comput.* **2021**, *17*, 3931–3945.

(74) GitHub, MB-Fit: Software Infrastructure for Data-Driven Many-Body Potential Energy Functions. `https://github.com/paesanilab/MB-Fit`.

(75) Adler, T.; Knizia, G.; Werner, H. A Simple and Efficient CCSD(T)-F12 Approximation. *J. Chem. Phys.* **2007**, *127*, 221106–221106.

(76) Knizia, G.; Adler, T. B.; Werner, H.-J. Simplified CCSD(T)-F12 Methods: Theory and Benchmarks. *J. Chem. Phys.* **2009**, *130*, 054104.

(77) Hill, J. G.; Peterson, K. A.; Knizia, G.; Werner, H.-J. Extrapolating MP2 and CCSD Explicitly Correlated Correlation Energies to the Complete Basis Set Limit With First and Second Row Correlation Consistent Basis Sets. *J. Chem. Phys.* **2009**, *131*, 194105.

(78) Góra, U.; Podeszwa, R.; Cencek, W.; Szalewicz, K. Interaction Energies of Large Clusters From Many-Body Expansion. *J. Chem. Phys.* **2011**, *135*, 224102.

(79) Pritchard, B. P.; Altarawy, D.; Didier, B.; Gibson, T. D.; Windus, T. L. New Basis Set Exchange: An Open, Up-to-Date Resource for the Molecular Sciences Community. *J. Chem. Inf. Model.* **2019**, *59*, 4814–4820.

(80) Dunning, T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms Boron through Neon and Hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

(81) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. Electron Affinities of the First-Row Atoms Revisited. Systematic Basis Sets and Wave Functions. *J. Chem. Phys.* **1992**, *96*, 6796–6806.

(82) Woon, D. E.; Dunning, T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. III. The Atoms Aluminum through Argon. *J. Chem. Phys.* **1993**, *98*, 1358–1371.

(83) Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M. Molpro: A General-Purpose Quantum Chemistry Program Package. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 242–253.

(84) Shinoda, W.; Shiga, M.; Mikami, M. Rapid Estimation of Elastic Constants by Molecular Dynamics Simulation under Constant Stress. *Phys. Rev. B* **2004**, *69*, 134103.

(85) Hockney, R. W.; Eastwood, J. W. *Computer Simulation Using Particles*; CRC Press: New York, 2021.

(86) Thompson, A. P.; Aktulga, H. M.; Berger, R.; Bolintineanu, D. S.; Brown, W. M.; Crozier, P. S.; in't Veld, P. J.; Kohlmeyer, A.; Moore, S. G.; Nguyen, T. D. et al. LAMMPS - A Flexible Simulation Tool for Particle-Based Materials Modeling at the Atomic, Meso, and Continuum Scales. *Comput. Phys. Commun.* **2022**, *271*, 108171.

(87) MBX: A Many-Body Energy and Force Calculator. `https://paesanigroup.ucsd.edu/software/mbx.html`.

(88) Schenter, G. K.; Fulton, J. L. *XAFS Techniques for Catalysts, Nanomaterials, and Surfaces*; Springer, 2017; pp 251–270.

(89) Rehr, J. J.; Kas, J. J.; Vila, F. D.; Prange, M. P.; Jorissen, K. Parameter-Free Calculations of X-ray Spectra with FEFF9. *Phys. Chem. Chem. Phys.* **2010**, *12*, 5503–5513.

(90) Rehr, J. J.; Kas, J. J.; Prange, M. P.; Sorini, A. P.; Takimoto, Y.; Vila, F. Ab Initio Theory and Calculations of X-ray Spectra. *C. R. Phys.* **2009**, *10*, 548–559.

(91) Rehr, J. J.; Albers, R. C. Theoretical Approaches to X-ray Absorption Fine Structure. *Rev. Mod. Phys.* **2000**, *72*, 621.

(92) Zhuang, D.; Riera, M.; Schenter, G. K.; Fulton, J. L.; Paesani, F. Many-Body Effects Determine the Local Hydration Structure of Cs^+ in Solution. *J. Phys. Chem. Lett.* **2019**, *10*, 406–412.

(93) Ravel, B.; Newville, M. ATHENA, ARTEMIS, HEPHAESTUS: Data Analysis for X-ray Absorption Spectroscopy using IFEFFIT. *J. Synchrotron Radiat.* **2005**, *12*, 537–541.

(94) Glezakou, V.-A.; Chen, Y.; Fulton, J. L.; Schenter, G. K.; Dang, L. X. Electronic Structure, Statistical Mechanical Simulations, and EXAFS Spectroscopy of Aqueous Potassium. *Theor. Chem. Acc.* **2006**, *115*, 86–99.

(95) Filipponi, A. EXAFS for Liquids. *J. Phys. Condens. Matter* **2001**, *13*, R23.

(96) Pham, V.-T.; Fulton, J. L. Ion-Pairing in Aqueous $CaCl_2$ and RbBr Solutions: Simultaneous Structural Refinement of XAFS and XRD Data. *J. Chem. Phys.* **2013**, *138*, 044201.

(97) Fulton, J. L.; Kathmann, S. M.; Schenter, G. K.; Balasubramanian, M. Hydrated Structure of Ag(I) Ion from Symmetry-Dependent, K-and L-edge XAFS Multiple Scattering and Molecular Dynamics Simulations. *J. Phys. Chem. A* **2009**, *113*, 13976–13984.

(98) D'Angelo, P.; Zitolo, A.; Migliorati, V.; Pavel, N. V. Measurement of X-ray Multielectron Photoexcitations at the I^- K edge. *Phys. Rev. B* **2008**, *78*, 144105.

TOC graphic