Data-Driven Many-Body Models Enable a Quantitative Description of Chloride Hydration from Clusters to Bulk

Alessandro Caruso[1, a)] and Francesco Paesani[1, 2, 3, b)]

[1)]*Department of Chemistry and Biochemistry, University of California San Diego, La Jolla, California 92093, United States*

[2)]*Materials Science and Engineering, University of California San Diego, La Jolla, California 92093, United States*

[3)]*San Diego Supercomputer Center, University of California San Diego, La Jolla, California 92093, United States*

We present a new data-driven potential energy function (PEF) describing chloride–water interactions which is developed within the many-body-energy (MB-nrg) theoretical framework. Besides quantitatively reproducing low-order many-body energy contributions, the new MB-nrg PEF is able to correctly predict the interaction energies of small chloride–water clusters calculated at the coupled cluster level of theory. Importantly, classical and quantum molecular dynamics simulations of a single chloride ion in water demonstrate that the new MB-nrg PEF predicts X-ray spectra in close agreement with the experimental results. Comparisons with an popular empirical model and a polarizable PEF emphasize the importance of an accurate representation of short-range many-body effect while demonstrating that pairwise additive representations of chloride–water and water–water interactions are inadequate for correctly representing the hydration structure of chloride in both gas-phase clusters and solution. We believe that the analyses presented in this study provide additional evidence for the accuracy and predictive ability of the MB-nrg PEFs which can then enable more realistic simulations of ionic aqueous systems in different environments.

[a)]Electronic mail: acaruso@ucsd.edu

[b)]Electronic mail: fpaesani@ucsd.edu

I. INTRODUCTION

A molecular-level characterization of the hydration properties of charged species from small clusters to bulk solutions and interfaces is key to understanding many physicochemical processes. Notably, hydrated ions are found as stabilizing species for biomolecules[1–3] and are involved in catalytic and transport processes.[4–10] It is hard to overstate the importance of ions in electrochemical processes, as charged species are a necessary component of electrolytic and galvanic cells.[11] Hydrated ions have also been shown to take part in the growth process of cloud condensation nuclei.[12–15]

While ions are ubiquitous in natural and industrial processes, a predictive understanding of the driving forces that determine the molecular properties of ions in aqueous solutions is still missing. For instance, it is known that, depending on their intrinsic electronic structure, ions can either strengthen or weaken the structure of the surrounding water hydrogen-bonding (H-bonding) network. To describe this effect, Hofmeister's original classification of ions according to their ability to modulate protein solubility[16] subsequently led to the classification of ions as "structure makers" and "structure breakers".[17] Broadly speaking, "structure makers" are small and highly charged ions that are strongly hydrated and are, therefore, believed to strengthen the H-bonds between the surrounding water molecules. On the other hand, "structure breakers" are large and usually monovalent ions that, because of their size, disrupt the structure of the water H-bonding network. Although appealing due its simplicity, this classification has been challenged by measurements performed with various spectroscopic techniques.[18–22]

One of the most striking examples of the uncertainties in the current understanding of specific ion effects is perhaps represented by the ongoing debate around the distribution of ions at the air/water interface. Surface tension values larger than those for pure water were measured for salt solutions by Heydweiller more than one hundred years ago.[23] Importantly, while the surface tension was found to be independent of the nature of the cations, it varied significantly depending on the type of anion present in solution. Moreover, the effects of the anions on the surface tension were found to follow an inverse Hofmeister series. First Wagner,[24] and later Onsager and Samaras[25] proposed that image charges at the air/water interface are responsible for the local depletion of ions in the interfacial region which, in turn, leads to the observed variation of the surface tension in salt solutions. However, subsequent experiments found that the electrostatic potential difference across the air/water interface measured for solutions of halide salts (with the exception

of fluoride salts) is more negative than that for pure water, thus suggesting that larger halide ions have higher propensity for the interface than cations, in contradiction with predictions derived from Wagner, and Onsager and Samaras theories.[26] Since then, various experimental approaches have been used to characterize the physical mechanisms that determine the distribution of different ions at the water surface, sometimes with conflicting results.[27–29]

Pioneering molecular dynamics (MD) simulations carried out for halide–water clusters using polarizable force fields (FFs) predicted that all halide ions, except fluoride, are preferentially located at the surface of the clusters.[30–34] In contrast, MD simulations carried out for $Cl^-(H_2O)_n$ clusters with nonpolarizable FFs found that the chloride ion is always located in the interior of the clusters.[35,36] The different results obtained from MD simulations with polarizable and nonpolarizable FFs were interpreted as an indication of the importance of many-body effects in the hydration of halide ions.

Higher propensity for the water surface relative to the bulk was found for larger halide ions from MD simulations of concentrated salt solutions as well as from calculations of single-ion potentials of mean force (PMFs) carried out using polarizable FFs.[37,38] In particular, the surface propensity was found to increase from Cl^- to I^-, with F^- being repelled from the interface. Similar results were later obtained with various polarizable models.[39–41] More recent *ab initio* MD simulations based on density functional theory (DFT) found a much lower propensity of the iodide ion, for the air/water interface, compared to predictions obtained with polarizble FFs.[42] However, these MD-DFT simulations were carried out with the dispersion-corrected BLYP functional which has been shown to suffer from some intrinsic limitations when applied to the modeling of liquid water.[43–45] A lower surface propensity than that calculated for larger halide ions using MD simulations with polarizable FFs has also been predicted by an extended dielectric continuum (DC) theory which takes into account both the dimension and the polarizability of the ions.[46,47] An alternative model emphasizing the impact that ions may have on surface fluctuations has also been proposed to explain the experimental observations of selective ion adsorption at the air/water interface.[48]

Despite much recent effort in characterizing the molecular driving forces that contribute to modulating both structural and thermodynamic properties of ions in solution, it has become increasingly apparent that the development of a unified, molecular theory of ion hydration requires a quantitative description of the interplay between ion–water and water–water interactions, which has so far remained elusive. In this context, the last decade has witnessed the emergence of many-body potential energy functions (PEFs) which, rigorously derived from the many-body expansion

(MBE) of the energy, hold great promise for predictive MD simulations of aqueous systems, from small clusters in the gas phase to bulk solutions and interfaces.[49–51] In particular, building upon the MB-pol PEF for water,[52–54] we developed two families of many-body PEFs, the TTM-nrg[55,56] and MB-nrg[57,58] PEFs, which have been shown to accurately reproduce both structural and thermodynamics properties of $X^-(H_2O)_n$ (X = F, Cl, Br, I) and $M^+(H_2O)_n$ (M = Li, Na, K, Rb, Cs) clusters, including quantum-mechanical effects in H-bonding rearrangements and isomeric equilibria.[59–63]

In this study, we present a systematic analysis of many-body effects in the hydration properties of Cl^- through detailed comparisons of different models of chloride–water interactions, from simple point-charge FFs to polarizable FFs, and explicit many-body PEFs. For this purpose, we introduce an extended MB-nrg PEF that, building upon the results of Ref. 57, includes an explicit 3-body (3B) term as well as a refined 2-body term (2B) derived from an expanded training set of dimer configurations generated using a recently developed active learning scheme for many-body PEFs.[64]

II. METHODS

The MBE expresses the energy, E_N, of a system containing N (atomic or molecular) monomers as the sum of individual n-body energies, $\varepsilon^{\mathrm{nB}}$, where $n \leq N$,[65]

$$E_N(r_1,..,r_N) = \sum_{i=1}^{N} \varepsilon^{\mathrm{1B}}(r_i) + \sum_{i<j}^{N} \varepsilon^{\mathrm{2B}}(r_i,r_j) + \sum_{i<j<k}^{N} \varepsilon^{\mathrm{3B}}(r_i,r_j,r_k) + ... + \varepsilon^{\mathrm{NB}}(r_1,..,r_N) \qquad (1)$$

Here, r_i collectively represents the coordinates of all atoms in monomer i, $\varepsilon^{\mathrm{1B}}$ represents the distortion energy of an isolated (molecular) monomer, and the n-body energies $\varepsilon^{\mathrm{nB}}$ are defined as

$$\varepsilon^{\mathrm{nB}} = E_n(r_1,...,r_n) - \sum_{i=1}^{N} \varepsilon^{\mathrm{1B}}(r_i) - \sum_{i<j}^{N} \varepsilon^{\mathrm{2B}}(r_i,r_j) - \ldots - \sum_{i<j<k<\ldots}^{N} \varepsilon^{\mathrm{(n\text{-}1)B}}(r_i,r_j,r_k,\ldots) \qquad (2)$$

Since the MBE converges quickly for systems with localized electron densities and large band gaps, Eq. 1 can be used as a effective theoretical/computational framework for developing many-body PEFs where each individual term of the MBE is independently fitted to the corresponding electronic structure data.[66,67]

As discussed in detail in the original studies, both TTM-nrg[55,56] and MB-nrg[57,58] PEFs are derived from Eq. 1, and use the MB-pol PEF for representing water–water interactions.[52–54] MB-pol has been shown by us and others to correctly reproduce the properties of water,[66,68] from small clusters in the gas phase[69–80] to liquid water,[81–87] the air/water interface,[88–92], and ice.[93–96]

In the TTM-nrg PEFs all ion–water many-body contributions, i.e., ε^{2B} to ε^{NB}, are described by an implicit NB term represented by classical polarization.[55,56] The MB-nrg PEFs instead approximate Eq. 1 with the sum of explicit low-order terms, generally up to the 3B term, along with the same implicit many-body term used by the TTM-nrg PEFs to represent all higher-body interactions,[57,58]

$$E_N = \sum_{i=1}^{N} \varepsilon_i^{1B} + \sum_{i>j}^{N} \varepsilon_{i,j}^{2B} + \sum_{i>j>k}^{N} \varepsilon_{i,j,k}^{3B} + V_{pol} \tag{3}$$

Each term of Eq. 3 is fitted to reproduce the corresponding nB reference energies that, as discussed below, are calculated at the explicitly correlated coupled cluster level of theory including single, double, and perturbative triple excitations, i.e., CCSD(T)-F12b.[97,98] Since the theoretical/computational framework behind the MB-pol[52,53] and MB-nrg[57,58] PEFs is described in the original references, we will only discuss here specific details related to the development of the present chloride–water MB-nrg PEF.

A. 2-body energies

Following Refs. 57 and 58, ε^{2B} in Eq. 3 is represented by three terms:

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

Here, V_{sr}^{2B} describes quantum-mechanical short-range 2B interactions (e.g., Pauli repulsion, and charge transfer and penetration) that arise from the overlap of the electron densities of the chloride ion and a water molecule, which cannot be represented in terms of classical expressions.[50,51] In the MB-nrg PEF, V_{sr}^{2B} is represented by a permutationally invariant polynomial, V_{PIP}^{2B}, that is dampened to zero at long range by a switching function, $s_2(R_{Cl^-O})$, of the distance R_{Cl^-O} between the chloride ion (Cl^-) and the oxygen atom (O) of the water molecule within a Cl^-–H_2O dimer,

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

where

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

and

$$t_2(R_{Cl^-O}) = \frac{R_{Cl^-O} - R_i^{2B}}{R_{out}^{2B} - R_{in}^{2B}} \tag{7}$$

Here, $R_{\text{in}}^{\text{2B}} = 5.8$ Å and $R_o^{2B} = 7.8$ Å are the predefined inner and outer cutoff distances of the switching function. These cutoff distances, which differ slightly from those used in Ref. 57, were found to guarantee a smooth and continuous representation of ε^{2B}. In particular, both $R_{\text{in}}^{\text{2B}}$ and the range of the switching function were increased compared to the original values[57] since the gain in stability provided by a more slowly varying switching function was found to overcome the higher computational cost associated with a larger number of 2B interactions to compute. As in Ref. 57, $V_{\text{PIP}}^{\text{2B}}$ is a function of all pairwise distances among the physical atoms (H, O, and Cl^-) and the lone-pair sites of the MB-pol water molecule (L_1 and L_2)[52] within a Cl^-–H_2O dimer. $V_{\text{PIP}}^{\text{2B}}$ 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. By construction, $V_{\text{PIP}}^{\text{2B}}$ thus contains 496 linear fitting parameters (c_i) and 9 nonlinear fitting parameters.[57]

As in Ref. 57, V_{elec} in Eq. 4 represents permanent electrostatics between the negative (-1e) charge of the chloride ion and the MB-pol geometry-dependent point charges of the water molecule which reproduce the *ab initio* dipole moment of an isolated water molecule.[99]

The last term in Eq. 4, V_{disp}, describes the 2B dispersion energy:

$$V_{\text{disp}} = -f(\delta_{\text{Cl}^-\text{O}})\frac{C_{6,\text{Cl}^-\text{O}}}{R^6_{\text{Cl}^-\text{O}}} - f(\delta_{\text{Cl}^-\text{H}_1})\frac{C_{6,\text{Cl}^-\text{H}_1}}{R^6_{\text{Cl}^-\text{H}_1}} - f(\delta_{\text{Cl}^-\text{H}_2})\frac{C_{6,\text{Cl}^-\text{H}_2}}{R^6_{\text{Cl}^-\text{H}_2}} \tag{8}$$

where $R_{\text{Cl}^-\text{O}}$, $R_{\text{Cl}^-\text{H}_1}$, and $R_{\text{Cl}^-\text{H}_2}$ are the distances between the Cl^- ion and the O, and the two H atoms of the water molecule within the dimer, and $f(\delta)$ and C_6 are the corresponding Tang-Toennies damping functions[100] and dispersion coefficients.

B. 3-body energies

Building upon the same theoretical framework used in the development of MB-pol and the Cs^+–H_2O MB-nrg PEF, ε^{3B} in Eq. 3 is represented by a 3B short-range term that effectively takes into account 3B energy contributions of quantum-mechanical origin arising from the overlap of the electronic densities of the chloride ion and two water (a and b) molecules at a time as well as short-range 3B dispersion energy contributions,

$$\varepsilon^{\text{3B}} = [s_3(R_{\text{Cl}^-\text{O}_a})s_3(R_{\text{Cl}^-\text{O}_b}) + s_3(R_{\text{Cl}^-\text{O}_a})s_3(R_{\text{O}_a\text{O}_b}) + s_3(R_{\text{Cl}^-\text{O}_b})s_3(R_{\text{O}_a\text{O}_b})] \cdot V_{\text{PIP}}^{\text{3B}} \tag{9}$$

Here, s_3 is a 3-body switching function given by

$$s_3(R_{kl}) = \begin{cases} 1, & \text{if } t_3(R_{kl}) < 0 \\ cos^2\left[t_3(R_{kl})\pi/2\right], & \text{if } 0 \leq t_3(R_{kl}) < 1 \\ 0, & \text{if } 1 \leq t_3(R_{kl}) \end{cases} \tag{10}$$

where

$$t_3(R_{kl}) = \frac{R_{kl} - R_i^{3B}}{R_{\text{out}}^{3B} - R_{in}^{3B}} \tag{11}$$

In Eqs. 9 and 10, R_{kl} is the distance between any (k, l) pair of Cl^- and O atoms within a $Cl^-(H_2O)_2$ trimer, R_{in}^{3B} and R_{out}^{3B} are the inner and outer cutoff distances. As an optimal compromise between accuracy and computational cost, 3B effects are only included within the first hydration shell of the chloride ion which is achieved by setting R_{in}^{3B} and R_{out}^{3B} to 2.5 Åand 4.5 Å, respectively. Although the following analyses demonstrate that the 2B and 3B cutoff ranges adopted in the present study allow for the accurate representation of the hydration structure of a chloride ion both in gas-phase clusters and in solution, it should be noted that the MB-nrg framework gives the user complete freedom in the choice of the inner and outer cutoffs.

V_{PIP}^{3B} is a function of all 41 pairwise distances between the physical atoms (H, O, and Cl^-) and the lone-pair sites of the two water molecules (L_1 and L_2) within a $Cl^-(H_2O)_2$ trimer. V_{PIP}^{3B} contains 1575 symmetrized monomials, ξ_i: 39 second-degree monomials, 613 third-degree monomials, and 923 fourth-degree monomials. Therefore, V_{PIP}^{3B} contains 1575 linear fitting parameters and 13 nonlinear fitting parameters.

Specific details about V_{PIP}^{2B} and V_{PIP}^{3B}, along with the definition of all monomials are given in the Supplementary Material.

C. Reference energies

The reference dimer configurations used in the parameterization of the 2B energy term, ε^{2B}, were selected using the active learning approach described in Ref. 64, starting from a pool of 150 000 configurations generated by sampling a spherical grid between 2–8 Å from the chloride ion as well as the normal modes of the Cl^-–H_2O dimer. 9059 and 854 dimer configurations were used in the training and test sets, respectively. The reference 2B energies were calculated at the CCSD(T)-F12b level of theory[97,98] in the complete basis set (CBS) limit that was achieved via a two-point extrapolation.[101,102] The CCSD(T)-F12b calculations were performed with the

augmented correlation-consistent polarized valence triple-/quadruple-ζ (aug-cc-pV[T/Q]Z) basis sets.[103–106]

A Cl^-–H_2O MB-nrg PEF, without the explicit 3B term, ε^{3B}, was initially developed and used in MD simulations with a single Cl^- ion in water which were carried out at ambient conditions in the isobaric-isothermal (NPT) ensemble to generate the 3B pool. 13140 and 1240 $Cl^-(H_2O)_2$ trimer configurations were extracted from the MD trajectories and included in the training and test sets, respectively. The 3B energies were calculated at the CCSD(T)-F12b level of theory[97,98] using the aug-cc-pVTZ basis set.[103–106]

All CCSD(T)-F12b electronic structure calculations were carried out using MOLPRO (version 2020.1).[107]

D. Fitting procedure

We followed the same fitting procedure used in the development of MB-pol[52,53] and other MB-nrg PEFs.[57,58,108–110] Specifically, the linear parameters were optimized through singular value decomposition, while the non-linear parameters were optimized using the simplex algorithm. The following regularized weighted sum of squared deviations was calculated and minimized:

$$\chi^2 = \sum_{n\in\mathscr{S}} w_n[\varepsilon_{\text{model}}(n) - \varepsilon_{\text{ref}}(n)]^2 + \Gamma^2 \sum_{l=1}^{L} c_l^2 \tag{12}$$

Here, $\mathscr{S}$ represents the training set and L is the number of linear parameters. The weights w_n are introduced in Eq. 12 to emphasize configurations with lower binding energies according to[111]

$$w(E_i) = \left[\frac{\Delta E}{E_i - E_{\text{min}} + \Delta E}\right]^2 \tag{13}$$

Here, E_{min} is the lowest binding energy in the training set and ΔE is a parameter that defines the range of favorably weighted energies. $\Delta E = 35$ kcal/mol and $\Delta E = 47.5$ kcal/mol were used for the 2B and 3B energies, respectively. The regularization parameter Γ was set to 0.0005 to reduce the variation of the linear parameters while preserving the overall accuracy.

E. MD simulations and analysis

All MD simulations were carried out in the NPT ensemble for a box containing a single chloride ion and 277 water molecules at 298.15 K and 1.0 atm, corresponding to a $\sim$0.2 M solution. The

velocity-Verlet algorithm[112] was used to propagate the equations of motion in the MD simulations with the TTM-nrg and MB-nrg PEFs. A timestep δt of 0.2 fs was used, which guarantees a correct sampling of the molecular degrees of freedom as well as the TTM-nrg and MB-nrg induced dipole moments that were propagated according to the always stable predictor–corrector algorithm.[113] Nosé-Hoover chains with 4 thermostats attached to each degree of freedom were used to control the temperature while the pressure was controlled using the algorithm described in Ref. 114. The path-integral molecular dynamics (PIMD) simulations with the MB-nrg PEF were carried out using the normal-mode representation, with each atom being described by a ring-polymer with 32 beads.[115] All MD simulations with the TTM-nrg and MB-nrg PEFs were carried out with an in-house version of DL_POLY 2.0[116].

For comparison, MD simulations were also carried out using the Cl^-–H_2O empirical parameterization compatible with the TIP4P/Ew water model[117] which was introduced in Ref. 118 These simulations were carried out with AMBER[119] using a global Langevin thermostat and a Monte Carlo barostat to control the temperature and the pressure, respectively.

The FEFF software was used to calculate the EXAFS signals.[120–122] Following Ref. 123, all FEFF calculations were performed using clusters containing the chloride ion and its 33 closest water molecules which were extracted from the corresponding MD and PIMD trajectories at intervals of 0.5 ps.

III. RESULTS

Fig. 1 shows the correlation plots between the reference CCSD(T)-F12b/CBS 2B energies and the corresponding MB-nrg values for both training (panel a) and test (panel b) sets. Root-mean-square errors (RMSEs) of 0.2387 kcal/mol and 0.2027 kcal/mol for the training and test sets, respectively, demonstrate that the present MB-nrg PEF is able to describe the Cl^-–H_2O 2B energies with coupled cluster accuracy over a wide range of values, without overfitting.

Fig. 1c shows one-dimensional potential energy radial scans for different (θ, ϕ) orientations of Cl^- relative to H_2O (see Fig. 1d for the definition of the coordinate system). Independently of the relative orientation, the MB-nrg PEF quantitatively reproduces the CCSD(T)/CBS values at all Cl^-–H_2O separations, which provides further evidence for the overall high accuracy of the present MB-nrg PEF at the 2B level.

After assessing the accuracy of the 2B term of the MB-nrg PEF, Fig. 2 shows the correlation

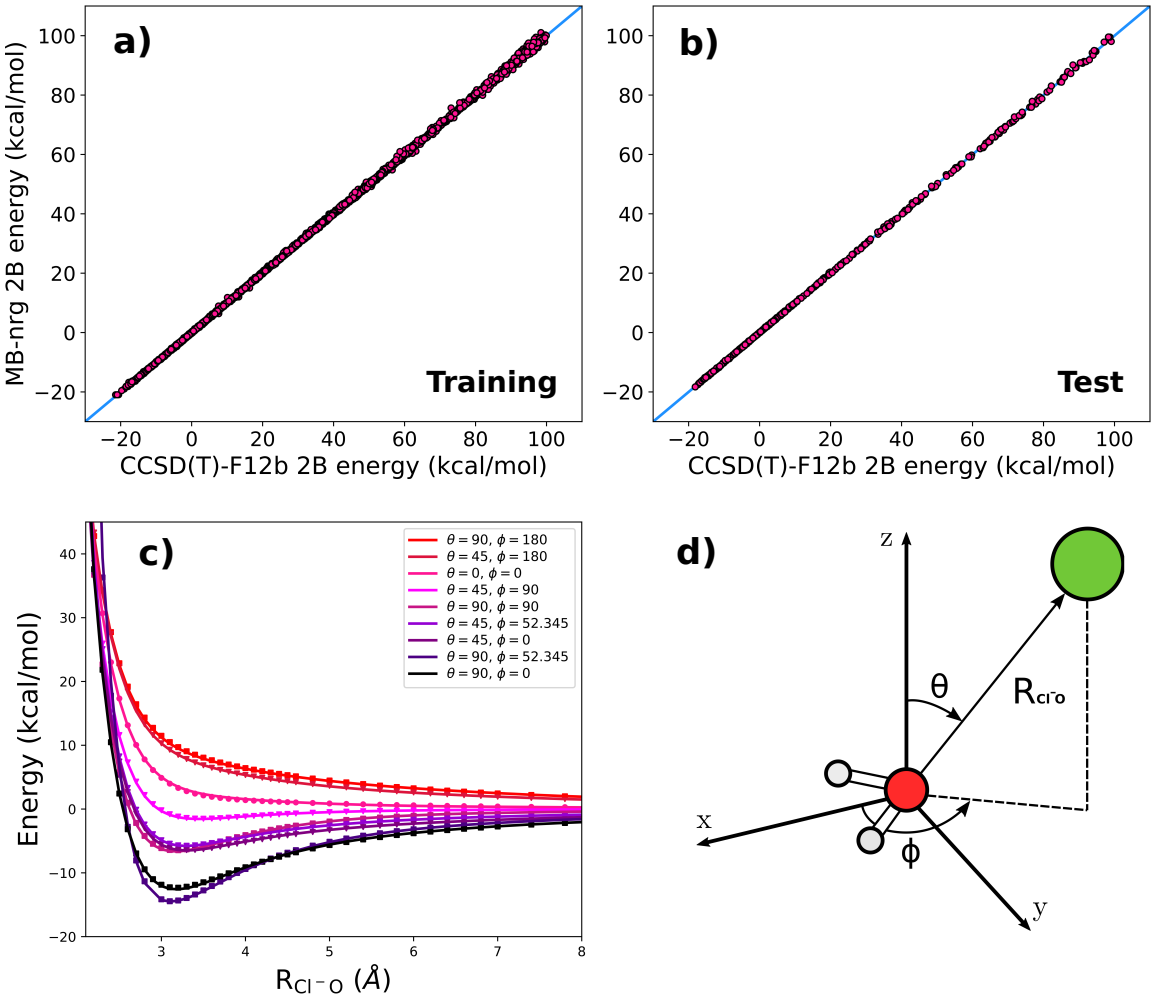

FIG. 1. Top panels: 2B energy correlation plots between the CCSD(T)-F12b/CBS reference values (x-axis) and corresponding MB-nrg values (y-axis) for the training (a) and test (b) sets. Bottom panels: Cl^-–H_2O potential energy scans (c) along the Cl^-–O distance (R_{Cl^-O}) for different orientations θ and ϕ (d). The symbols corresponds to the CCSD(T)-F12b/CBS reference energies, while the corresponding MB-nrg values are shown as solid lines.

plots between the CCSD(T)-F12b 3B reference energies and the corresponding MB-nrg values for both training (panel a) and test (panel b) sets. Also in this case, the present Cl^-–H_2O MB-nrg PEF is able to quantitatively reproduce the CCSD(T)-F12b values over the a wide range of 3B energies, with RMSEs of 0.0655 kcal/mol and 0.0506 kcal/mol for the training and test sets, respectively. As for the 2B energies, this analysis indicates that the high accuracy achieved by the MB-nrg PEF does not result from overfitting. MB-nrg 2B and 3B energy error plots are reported in the Supplementary Material.

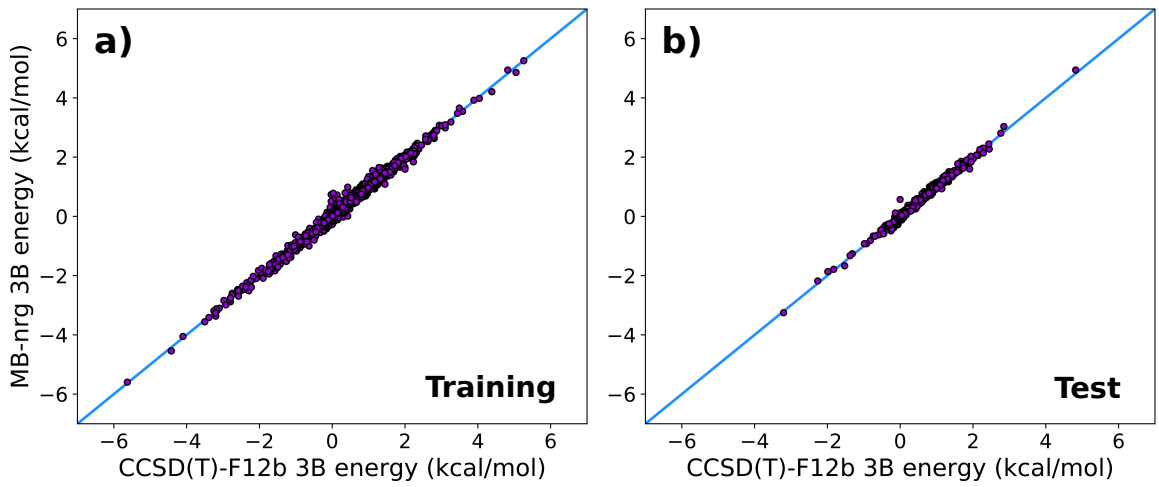

FIG. 2. 3B energy correlation plots between the CCSD(T)-F12b reference values (x-axis) and corresponding MB-nrg values (y-axis) for the training (a) and test (b) sets.

Fig. 3 shows a comparison between the interaction energies calculated for the low-energy isomers of the $Cl^-(H_2O)_n$ clusters ($n = 1-4$) using the empirical TIP4P/Ew-based model of Ref. 118, the TTM-nrg PEF of Ref. 55, and the present MB-nrg PEF. To investigate the role played by short-range many-body effects, in this and following analyses, we consider two versions of the MB-nrg PEF: the (2B+NB)-MB-nrg PEF that includes the explicit 2B term of Eq. 4 in addition to the classical NB polarization term of Eq. 3, and the (2B+3B+NB)-MB-nrg PEF that includes both explicit 2B (Eq. 4) and 3B (Eq. 9) terms in addition to the classical NB polarization term of Eq. 3. The results obtained with these four different representations of the Cl^-–H_2O interactions are compared with the the corresponding CCSD(T)-F12b reference values.[49,60] As expected, being an empirical pairwise additive model developed for bulk simulations, the TIP4P/Ew-based model is unable to correctly reproduce the energetics of $Cl^-(H_2O)_n$ clusters, independently of the size and H-bonding arrangements. By including NB effects through a classical polarization term, the TTM-nrg PEF clearly provides a more accurate representations of all clusters, which somewhat deteriorates for structures with more cooperative H-bonding arrangements, as occurs in the isomer 4f of $Cl^-(H_2O)_4$. The agreement with the CCSD(T)-F12b results effectively becomes quantitative with the inclusion of the explicit 2B term in the (2B+NB)-MB-nrg PEF, with the exception of the isomer 4f of $Cl^-(H_2O)_4$, for which only the (2B+3B+NB)-MB-nrg PEF is able to accurately reproduce the reference interaction energy. While emphasizing the role of many-body effects in chloride–water interactions, this analysis also highlights the limitations of a representation of

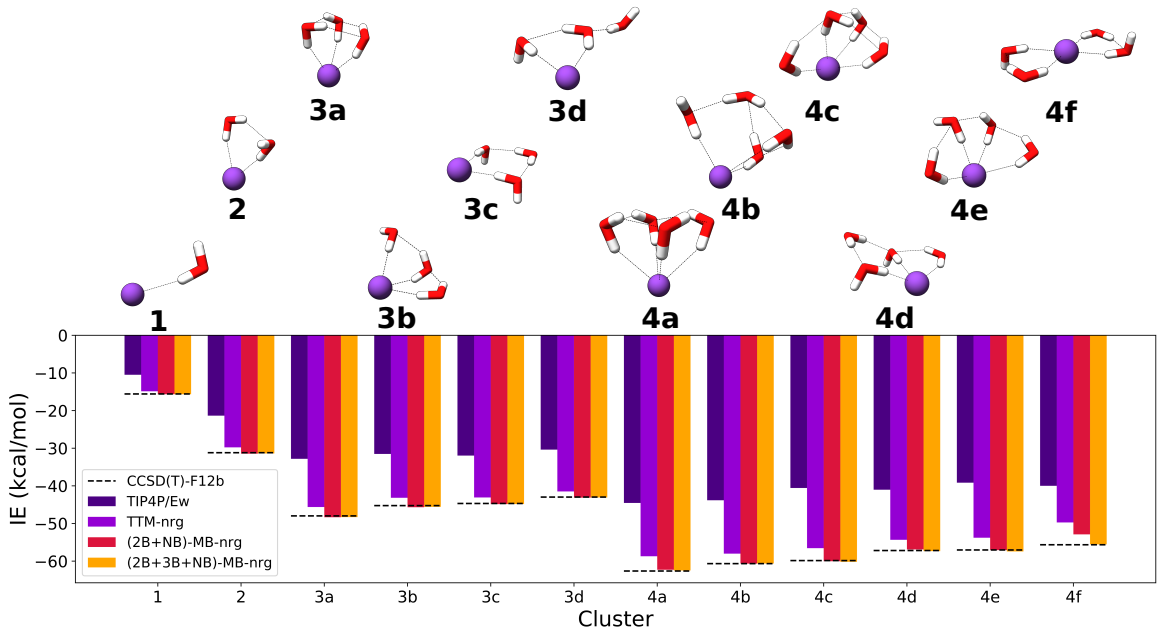

FIG. 3. Comparison between the interaction energies calculated for the low-energy isomers of $Cl^-(H_2O)_n$ clusters (with $n = 1-4$) using the empirical TIP4P/Ew-based model of Ref. 118, the TTM-nrg PEF of Ref. 55, and the present (2B+NB)-MB-nrg and (2B+3B+NB)-MB-nrg PEFs. For each cluster, the CCSD(T)-F12b reference values[49,60] are shown as horizontal dashed lines.

many-body effects which is entirely based on classical polarization, as in the TTM-nrg PEF and other common polarizable models, and further demonstrates the importance of correctly describing short-range low-order (i.e., 2B and 3B) interactions.[49] In this context, it should be noted that the analyses reported in Ref. 49 demonstrate that all nB interactions with $n > 3$ in $Cl^-(H_2O)_n$ clusters are correctly described by the classical polarization term, $V_{\rm pol}$ employed by the TTM-nrg and MB-nrg PEFs.

Although the analyses presented in Figs. 1-3 provide a quantitative assessment of the ability of the present MB-nrg PEF to reproduce, at the fundamental level, many-body interactions in $Cl^-(H_2O)_n$ clusters for which calculations at the coupled cluster level of theory are feasible, the MB-nrg PEFs also enable computer simulations of condensed-phase systems for which CCSD(T)-level calculations are currently out of reach. In absence of high-quality *ab initio* reference data for bulk simulations, the following analysis uses available EXAFS data to assess the reliability of the MB-nrg PEF in predicting the hydration structure of Cl^-. As for the analysis in Fig. 3, comparisons are made with results obtained from MD simulations carried out with the TIP4P/Ew-

based model, the TTM-nrg PEF, and the two (2B+NB)-MB-nrg and (2B+3B+NB)-MB-nrg PEFs. In addition, results from PIMD simulations carried out with the (2B+3B+NB)-MB-nrg PEF are used to quantify the role played by nuclear quantum effects in determining the local hydration structure of Cl^-.

Fig. 4a shows the Cl^-–O radial distribution functions (RDFs) calculated from the five sets of NPT simulations that were carried out in this study. Within 8 Å from the Cl^- ion, the TIP4P/Ew-based model predicts highly structured hydration shells located at 3.15 Å, 4.80 Å, and 7.15 Å. The inclusion of many-body effects through classical polarization in the TTM-nrg PEF leads to significant disruption of the first hydration shell compared to the TIP4P/Ew RDF, which is accompanied

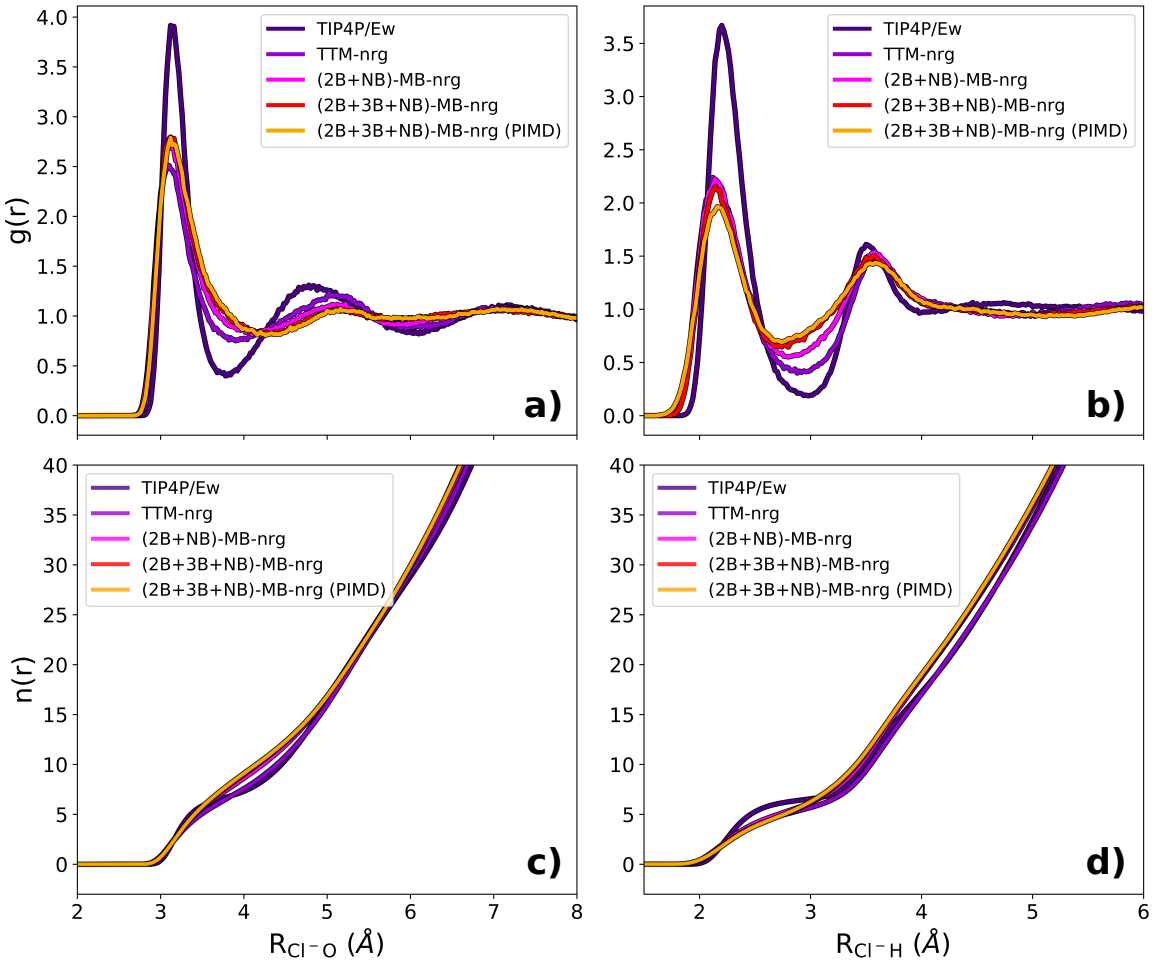

FIG. 4. Top panels: Cl-O (a) and Cl-H (b) radial distribution functions (RDFs) calculated from MD simulations carried out with the TIP4P/Ew-based model, and TTM-nrg, (2B+NB)-MB-nrg, and (2B+3B+NB)-MB-nrg PEFs as well as from PIMD simulations with the (2B+3B+NB)-MB-nrg PEF. Bottom panels: Corresponding Cl-O (c) and Cl-H (d) cumulative distribution functions (CDFs).

by a shift of the second hydration shell to larger distances (5.13 Å). This effect becomes more pronounced with the inclusion of explicit 2B and 3B contributions as shown by the RDFs calculated with the (2B+NB)-MB-nrg and (2B+3B+NB)-MB-nrg PEFs, respectively. In particular, the (2B+3B+NB)-MB-nrg PEF predicts more diffuse first and second hydration shells which indicates a significantly less structured spatial arrangement of the water molecules around the chloride ion compared to that predicted by the TIP4P/Ew-based model. Similar trends are found in the Cl^-–H RDFs shown in Fig. 4b, with the TIP4P/Ew-based model and (2B+3B+NB)-MB-nrg PEF predicting the most and least structured first Cl^-–H shells, respectively. Importantly, while the TIP4P/Ew-based model predicts a well-defined and more compact second Cl^-–H shell along with a still distinguishable third Cl^-–H shell, all many-body PEFs predict more diffuse distributions of the water hydrogen atoms of the water molecules that more closely hydrate the Cl^- ion. It should be noted that MD and PIMD simulations carried out with the (2B+3B+NB)-MB-nrg PEF effectively provide the same progression of hydration shells, with minor differences only visible in the first peak of the Cl^--H RDF, which suggests that nuclear quantum effects play a negligible role in determining the hydration structure of Cl^-.

Figs. 4c-d show the corresponding Cl^-–O and Cl^-–H cumulative distribution functions (CDFs). The structural features highlighted in the analysis of the Cl^-–O RDFs clearly have a direct effect on the coordination number. In particular, due to the underlying more ordered hydration structure, the TIP4P/Ew-based model predicts that 7 water molecules constitute the first shell. A similar evolution of the Cl^-–O CDF is obtained with the TTM-nrg PEF model although, due to a more diffuse arrangement of the water molecules, it is difficult to precisely determine the coordination number within the first hydration shell of Cl^-, with $6 < n_{1\mathrm{st}} < 8$. As it could be inferred from the analyses of the RDFs, a broader first hydration shell, containing ~ 11 water molecules, is predicted by both MB-nrg PEFs.

To further explore the structural properties of the hydrated chloride ion, we calculated the incremental radial distribution functions (iRDFs), which describe individual contributions to the total Cl^-–O RDF associated with each water molecule i as a function of its distance ($R_{Cl^-O_i}$) Cl^-, and the radial-angular distribution functions (RADFs) in the first hydration shell region. The iRDFs shown in Fig. 5 indicate that the TIP4P/Ew-based model predicts a clear separation between the first and second hydration shells which is located between the seventh and eighth water molecule. As indicated by the CDFs in Fig. 4c, the separation between first and second hydration shells becomes increasingly less distinguishable and shifts to larger distances as many-

body contributions to the Cl^-–H_2O interactions are progressively included, from the TTM-nrg to the (2B+3B+NB)-MB-nrg PEFs. It should also be noted that the TIP4P/Ew-based model predicts significantly narrower iRDFs for water molecules located in the first hydration shell of Cl^- ($i = 1-6$) as well as much broader distributions for the 7th and 8th water molecules located in the transition region between the first and second hydration shells compared to those obtained with the many-body PEFs. This analysis thus provides further evidence for the TTM-nrg and MB-nrg PEFs predicting less tightly bound water molecules in the first hydration shell of Cl^- than the TIP4P/Ew-based model.

The analysis of the RADFs shown Fig. 6 provides direct insights into the average distribution of water molecules within the first hydration shell of Cl^-. While predicting a tighter first hydration shell along the Cl^--O distance, as already inferred from the analyses of both RDFs and iRDFs, the TIP4P/Ew-based model is also characterized by a slightly broader distribution along the angular coordinate than both TTM-nrg and MB-nrg PEFs, which results in a relatively higher intensity between 55^o and 75^o. Particularly evident is the lack of the feature at 130° that becomes more

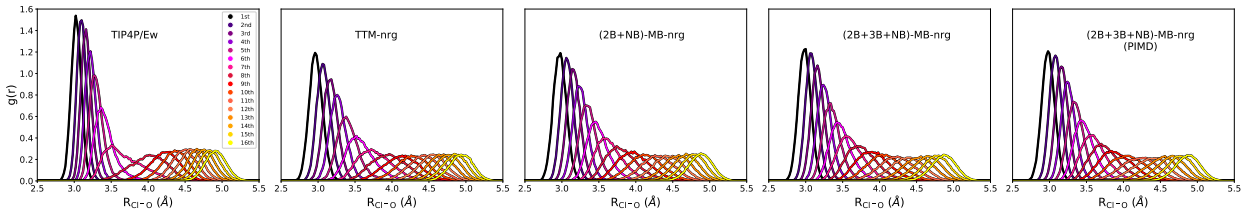

FIG. 5. Incremental radial distribution functions (iRDFs) calculated from MD simulations with the TIP4P/Ew-based model, and TTM-nrg, (2B+NB)-MB-nrg, and (2B+3B+NB)-MB-nrg PEFs as well as from PIMD simulations with the (2B+3B+NB)-MB-nrg PEF.

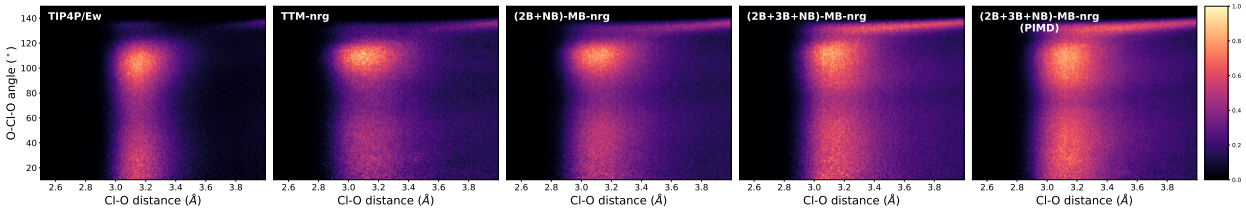

FIG. 6. Radial-angular distribution functions (RADFs) of the first hydration shell calculated from MD simulations with the TIP4P/Ew-based model, and TTM-nrg, (2B+NB)-MB-nrg, and (2B+3B+NB)MB-nrg PEFs as well as from PIMD simulations with the (2B+3B+NB)-MB-nrg PEF. The Cl-O distance (in Å) is shown on the *x*-axis, and the O-Cl-O angle (in degrees) is shown in the *y*-axis.

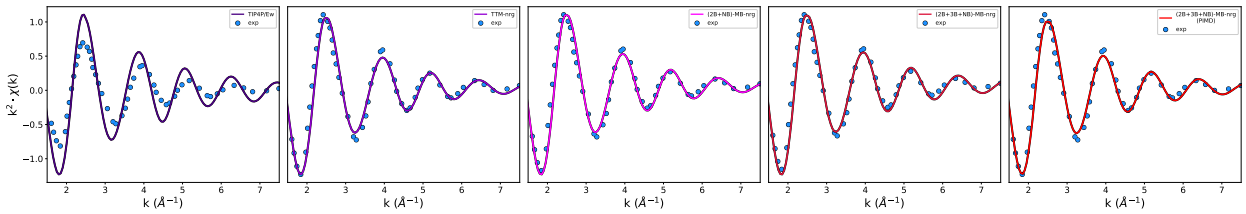

FIG. 7. K-edge EXAFS spectra, $k^2\chi(k)$, calculated from MD simulations with the TIP4P/Ew-based model, and TTM-nrg, (2B+NB)-MB-nrg, and (2B+3B+NB)-MB-nrg PEFs as well as from PIMD simulations with the (2B+3B+NB)-MB-nrg PEF. The experimental data from Ref. 124 are shown as blue circles.

pronounced as many-body contributions are progressively included in the description of the Cl^-–H_2O interactions. Due to the angular and radial diffuseness of the first hydration shell, it is difficult to extract contributions from individual molecules to the RADFs, and the average distribution of oxygen atoms around the chloride ion cannot be easily inferred. However, the different features exhibited by the different models show that short-range low-order interactions directly influence the geometry of the hydration complex. It should be noted that the RADF calculated from PIMD simulations with the (2B+3B+NB)-MB-nrg PEF is effectively indistinguishable from that obtained from the corresponding MD simulations, which provides further support for nuclear quantum effects playing a negligible role in determining the hydration structure of Cl^-.

To determine which representation of the Cl^-–H_2O interactions provides the most realistic description of the hydration structure of Cl^- in solution, Fig. 7 shows comparisons of the K-edge EXAFS spectra calculated with the TIP4P/Ew-based model, and TTM-nrg and MB-nrg PEFs with the corresponding experimental data from Ref. 124. This analysis shows that the TIP4P/Ew-based model is unable to correctly reproduce the amplitude of the EXAFS spectra and slightly underestimates the period of the oscillations, with the calculated peaks appearing at relatively smaller k values. On the other hand, the agreement with the experimental data systematically improves as many-body effects are progressively included, with the MD and PIMD simulations carried out with the (2B+3B+NB)-MB-nrg PEFs providing nearly quantitative agreement with the experimental EXAFS spectrum. The comparisons shown in Fig. 5 also indicate that 3B interactions have minimal impact on the simulated EXAFS spectra, while the inclusion of nuclear quantum effects in the PIMD simulations improves the agreement with the experimental data at larger k but worsens it for k values between 2 and 4 Å^{-1}.

IV. CONCLUSIONS

In this study, we have introduced a new many-body MB-nrg PEF describing chloride–water interactions which includes explicit 2B and 3B terms derived from CCSD(T)-F12b data along with an implicit NB term based on classical polarization. Although the new MB-nrg PEF is trained only on $Cl^-(H_2O)$ dimer and $Cl^-(H_2O)_2$ trimer configurations, it is able to quantitatively reproduce the interaction energies of small $Cl^-(H_2O)_n$ clusters, with $n = 1-4$. Importantly, when used in MD and PIMD simulations of a single Cl^- ion in water, we have demonstrated that the new MB-nrg PEF enables the calculation of EXAFS spectra that are in close agreement with the available experimental data, correctly reproducing both the amplitude and phase of the EXAFS oscillations.

Comparisons with the results obtained with a popular empirical force field[118] based on the TIP4P/Ew model of water[117] show that pairwise additive representations of chloride–water and water–water interactions are inadequate for representing chloride hydration structure in both gas-phase clusters and solution, underestimating the strength of the interactions in the first case while predicting an overly tight first hydration shell in the second case. On the other hand, comparisons with the results obtained with the TTM-nrg PEF[55] emphasize the importance of many-body effects in determining the hydration structure of Cl^- but, at the same time, highlight the limitations associated with a representation of these effects entirely based on classical many-body polarization.

We believe that the analyses presented here, along with results reported in previous studies,[57–63,123,125] provide further evidence that, as the MB-pol PEF has enabled an accurate description of the properties of water across different phases,[66] the MB-nrg PEFs for ion–water interactions can enable more realistic simulations of ionic aqueous systems from gas-phase clusters to bulk solutions and interfaces.

V. SUPPLEMENTARY MATERIAL

The supplementary material includes tables with the complete lists of both distances and variables of the 2B and 3B permutationally invariant polynomials, and the energy correlation and error plots for MB-nrg 2B and 3B.

VI. ACKNOWLEDGEMENT

This research was supported by the National Science Foundation through Grant No. CHE-1453204. 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; the National Energy Research Scientific Computing Center (NERSC), which is supported by the Office of Science of the U.S. Department of Energy under Contract DE-AC02-05CH11231; and the Triton Shared Computing Cluster (TSCC) at the San Diego Supercomputer Center.

VII. DATA AVAILABILITY

Any data generated and analyzed in this study are available from the authors upon request. The TTM-nrg and MB-nrg PEFs used in this study are available in our open-access MBX software that can be downloaded from http://paesanigroup.ucsd.edu/software/mbx.html.

REFERENCES

[1]S. Nahar and H. Tajmir-Riahi, "Do metal ions alter the protein secondary structure of a light-harvesting complex of thylakoid membranes?" J. Inorg. Biochem. **58**, 223–234 (1995).

[2]S. A. Woodson, "Metal ions and rna folding: A highly charged topic with a dynamic future," Curr. Opin. Chem. Biol. **9**, 104–109 (2005).

[3]D. E. Draper, "Rna folding: Thermodynamic and molecular descriptions of the roles of ions," Biophys. J. **95**, 5489–5495 (2008).

[4]R. A. Sneen, "Substitution at a saturated carbon atom. XVII. Organic ion pairs as intermediates in nucleophilic substitution and elimination reactions," Acc. Chem. Res. **6**, 46–53 (1973).

[5]M. Pregel, E. Dunn, R. Nagelkerke, G. Thatcher, and E. Buncel, "Alkali-metal ion catalysis and inhibition in nucleophilic displacement reaction of phosphorus-, sulfur-, and carbon-based esters," Chem. Soc. Rev. **24**, 449–455 (1995).

[6]N. Sträter, W. N. Lipscomb, T. Klabunde, and B. Krebs, "Two-metal ion catalysis in enzymatic acyl- and phosphoryl-transfer reactions," Angew. Chem. Int. Ed. **35**, 2024–2055 (1996).

[7]R. Hanna and J. A. Doudna, "Metal ions in ribozyme folding and catalysis," Curr. Opin. Chem. Biol. **4**, 166–170 (2000).

[8]M. Mucha, T. Frigato, L. M. Levering, H. C. Allen, D. J. Tobias, L. X. Dang, and P. Jungwirth, "Unified molecular picture of the surfaces of aqueous acid, base, and salt solutions," J. Phys. Chem. B **109**, 7617–7623 (2005).

[9]M. R. Stahley and S. A. Strobel, "Rna splicing: Group I intron crystal structures reveal the basis of splice site selection and metal ion catalysis," Curr. Opin. Struct. Biol. **16**, 319–326 (2006).

[10]R. K. Sigel and A. M. Pyle, "Alternative roles for metal ions in enzyme catalysis and the implications for ribozyme chemistry," Chem. Rev. **107**, 97–113 (2007).

[11]M. Winter and R. J. Brodd, "What are batteries, fuel cells, and supercapacitors?" Chem. Rev. **104**, 4245–4270 (2004).

[12]R. G. Harrison and H. Tammet, "Ions in the terrestrial atmosphere and other solar system atmospheres," Space Sci. Rev. **137**, 107–118 (2008).

[13]D. J. Tobias, A. C. Stern, M. D. Baer, Y. Levin, and C. J. Mundy, "Simulation and theory of ions at atmospherically relevant aqueous liquid-air interfaces," Annu. Rev. Phys. Chem. **64**, 339–359 (2013).

[14]N. S. Shuman, D. E. Hunton, and A. A. Viggiano, "Ambient and modified atmospheric ion chemistry: From top to bottom," Chem. Rev. **115**, 4542–4570 (2015).

[15]K. Lehtipalo, L. Rondo, J. Kontkanen, S. Schobesberger, T. Jokinen, N. Sarnela, A. Kürten, S. Ehrhart, A. Franchin, T. Nieminen, F. Riccobono, M. Sipilä, T. Yli-Juuti, J. Duplissy, A. Adamov, L. Ahlm, J. a. Almeida, A. Amorim, F. Bianchi, M. Breitenlechner, J. Dommen, A. J. Downard, E. M. Dunne, R. C. Flagan, R. Guida, J. Hakala, A. Hansel, W. Jud, J. Kangasluoma, V.-M. Kerminen, H. Keskinen, J. Kim, J. Kirkby, A. Kupc, O. Kupiainen-Määttä, A. Laaksonen, M. J. Lawler, M. Leiminger, S. Mathot, T. Olenius, I. K. Ortega, A. Onnela, T. Petäjä, A. Praplan, M. P. Rissanen, T. Ruuskanen, F. D. Santos, S. Schallhart, R. Schnitzhofer, M. Simon, J. N. Smith, G. Tröstl, Jasmin Tsagkogeorgas, A. Tomé, P. Vaattovaara, H. Vehkamäki, A. E. Vrtala, P. E. Wagner, C. Williamson, D. Wimmer, P. M. Winkler, A. Virtanen, N. M. Donahue, K. S. Carslaw, U. Baltensperger, I. Riipinen, J. Curtius, D. R. Worsnop, and M. Kulmala, "The effect of acid–base clustering and ions on the growth of atmospheric nano-particles," Nat. Commun. **7**, 1–9 (2016).

[16]F. Hofmeister, "Zur lehre von der wirkung der salze," Naunyn-Schmiedeberg's Arch. Pharmacol. **24**, 247–260 (1888).

[17]K. D. Collins and M. W. Washabaugh, "The Hofmeister effect and the behaviour of water at interfaces," Q. Rev. Biophys. **18**, 323–422 (1985).

[18]C. D. Cappa, J. D. Smith, K. R. Wilson, B. M. Messer, M. K. Gilles, R. C. Cohen, and R. J. Saykally, "Effects of alkali metal halide salts on the hydrogen bond network of liquid water," J. Phys. Chem. B **109**, 7046–7052 (2005).

[19]J. D. Smith, R. J. Saykally, and P. L. Geissler, "The effects of dissolved halide anions on hydrogen bonding in liquid water," J. Am. Chem. Soc. **129**, 13847–13856 (2007).

[20]D. A. Schmidt, O. Birer, S. Funkner, B. P. Born, R. Gnanasekaran, G. W. Schwaab, D. M. Leitner, and M. Havenith, "Rattling in the cage: Ions as probes of sub-picosecond water network dynamics," J. Am. Chem. Soc. **131**, 18512–18517 (2009).

[21]S. Funkner, G. Niehues, D. A. Schmidt, M. Heyden, G. Schwaab, K. M. Callahan, D. J. Tobias, and M. Havenith, "Watching the low-frequency motions in aqueous salt solutions: The terahertz vibrational signatures of hydrated ions," J. Am. Chem. Soc. **134**, 1030–1035 (2012).

[22]I. Waluyo, D. Nordlund, U. Bergmann, D. Schlesinger, L. G. Pettersson, and A. Nilsson, "A different view of structure-making and structure-breaking in alkali halide aqueous solutions through X-ray absorption spectroscopy," J. Chem. Phys. **140**, 244506 (2014).

[23]A. Heydweiller, "Über physikalische eigenschaften von lösungen in ihrem zusammenhang. ii. oberflächenspannung und elektrisches leitvermögen wässeriger salzlösungen," Ann. Phys. **4**, 145–185 (1910).

[24]C. Wagner, "Die oberflächenspannung verdünnter elektrolytlösungen," Phys. Z. **25**, 474 (1924).

[25]L. Onsager and N. N. Samaras, "The surface tension of debye-hückel electrolytes," J. Chem. Phys. **2**, 528–536 (1934).

[26]A. Frumkin, "Phasengrenzkräfte und adsorption an der trennungsfläche luft: Lösung anorganischer elektrolyte," Z. Phys. Chem. **109**, 34–48 (1924).

[27]D. Liu, G. Ma, L. M. Levering, and H. C. Allen, "Vibrational spectroscopy of aqueous sodium halide solutions and air-liquid interfaces: Observation of increased interfacial depth," J. Phys. Chem. B **108**, 2252–2260 (2004).

[28]E. A. Raymond and G. L. Richmond, "Probing the molecular structure and bonding of the surface of aqueous salt solutions," J. Phys. Chem. B **108**, 5051–5059 (2004).

[29]P. B. Petersen and R. J. Saykally, "Confirmation of enhanced anion concentration at the liquid water surface," Chem. Phys. Lett. **397**, 51–55 (2004).

[30]L. Perera and M. L. Berkowitz, "Many-body effects in molecular dynamics simulations of $Na^+(H_2O)_n$ and $Cl^-(H_2O)_n$ clusters," J. Chem. Phys. **95**, 1954–1963 (1991).

[31]L. Perera and M. L. Berkowitz, "Structure and dynamics of $Cl^-(H_2O)_{20}$ clusters: The effect of the polarizability and the charge of the ion," J. Chem. Phys. **96**, 8288–8294 (1992).

[32]L. Perera and M. L. Berkowitz, "Stabilization energies of Cl^-, Br^-, and I^- ions in water clusters," J. Chem. Phys. **99**, 4222–4224 (1993).

[33]L. Perera and M. L. Berkowitz, "Structures of $Cl^-(H_2O)_n$ and $F^-(H_2O)_n$ (n= 2, 3,..., 15) clusters. molecular dynamics computer simulations," J. Chem. Phys. **100**, 3085–3093 (1994).

[34]L. X. Dang and D. E. Smith, "Molecular dynamics simulations of aqueous ionic clusters using polarizable water," J. Chem. Phys. **99**, 6950–6956 (1993).

[35]S. J. Stuart and B. Berne, "Effects of polarizability on the hydration of the chloride ion," J. Phys. Chem. **100**, 11934–11943 (1996).

[36]S. J. Stuart and B. Berne, "Surface curvature effects in the aqueous ionic solvation of the chloride ion," J. Phys. Chem. A **103**, 10300–10307 (1999).

[37]P. Jungwirth and D. J. Tobias, "Molecular structure of salt solutions: A new view of the interface with implications for heterogeneous atmospheric chemistry," J. Phys. Chem. B **105**, 10468–10472 (2001).

[38]P. Jungwirth and D. J. Tobias, "Ions at the air/water interface," J. Phys. Chem. B **106**, 6361–6373 (2002).

[39]L. X. Dang, "Computational study of ion binding to the liquid interface of water," J. Phys. Chem. B **106**, 10388–10394 (2002).

[40]L. X. Dang and T.-M. Chang, "Molecular mechanism of ion binding to the liquid/vapor interface of water," J. Phys. Chem. B **106**, 235–238 (2002).

[41]C. D. Wick, I.-F. W. Kuo, C. J. Mundy, and L. X. Dang, "The effect of polarizability for understanding the molecular structure of aqueous interfaces," J. Chem. Theory. Comput. **3**, 2002–2010 (2007).

[42]M. D. Baer and C. J. Mundy, "Toward an understanding of the specific ion effect using density functional theory," J. Phys. Chem. Lett. **2**, 1088–1093 (2011).

[43]M. J. Gillan, D. Alfè, and A. Michaelides, "Perspective: How good is DFT for water?" J. Chem. Phys. **144**, 130901 (2016).

[44]G. A. Cisneros, K. T. Wikfeldt, L. Ojamäe, J. Lu, Y. Xu, H. Torabifard, A. P. Bartók, G. Csányi, V. Molinero, and F. Paesani, "Modeling molecular interactions in water: From pairwise to many-body potential energy functions," Chem. Rev. **116**, 7501–7528 (2016).

[45]M. Galib, T. T. Duignan, Y. Misteli, M. D. Baer, G. K. Schenter, J. Hutter, and C. J. Mundy, "Mass density fluctuations in quantum and classical descriptions of liquid water," J. Chem. Phys. **146**, 244501 (2017).

[46]Y. Levin, A. P. Dos Santos, and A. Diehl, "Ions at the air-water interface: An end to a hundred-year-old mystery?" Phys. Rev. Lett. **103**, 257802 (2009).

[47]Y. Levin, "Polarizable ions at interfaces," Phys. Rev. Lett. **102**, 147803 (2009).

[48]D. E. Otten, P. R. Shaffer, P. L. Geissler, and R. J. Saykally, "Elucidating the mechanism of selective ion adsorption to the liquid water surface," Proc. Natl. Acad. Sci. U.S.A. **109**, 701–705 (2012).

[49]F. Paesani, P. Bajaj, and M. Riera, "Chemical accuracy in modeling halide ion hydration from many-body representations," Adv. Phys. X **4**, 1631212 (2019).

[50]B. B. Bizzarro, C. K. Egan, and F. Paesani, "Nature of halide–water interactions: Insights from many-body representations and density functional theory," J. Chem. Theory. Comput. **15**, 2983–2995 (2019).

[51]C. K. Egan, B. B. Bizzarro, M. Riera, and F. Paesani, "Nature of alkali ion–water interactions: Insights from many-body representations and density functional theory. II," J. Chem. Theory. Comput. **16**, 3055–3072 (2020).

[52]V. Babin, C. Leforestier, and F. Paesani, "Development of a "first principles" water potential with flexible monomers: Dimer potential energy surface, VRT spectrum, and second virial coefficient," J. Chem. Theory. Comput. **9**, 5395–5403 (2013).

[53]V. Babin, G. R. Medders, and F. Paesani, "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. **10**, 1599–1607 (2014).

[54]G. R. Medders, V. Babin, and F. Paesani, "Development of a "first principles" water potential with flexible monomers. III. liquid phase properties," J. Chem. Theory. Comput. **10**, 2906–2910 (2014).

[55]D. J. Arismendi-Arrieta, M. Riera, P. Bajaj, R. Prosmiti, and F. Paesani, "i-TTM model for ab initio-based ion–water interaction potentials. 1. Halide–water potential energy functions," J. Phys. Chem. B **120**, 1822–1832 (2016).

[56]M. Riera, A. W. Götz, and F. Paesani, "The i-TTM model for ab initio-based ion–water interaction potentials. II. Alkali metal ion–water potential energy functions," Phys. Chem. Chem. Phys. **18**, 30334–30343 (2016).

[57]P. Bajaj, A. W. Götz, and F. Paesani, "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. **12**, 2698–2705 (2016).

[58]M. Riera, N. Mardirossian, P. Bajaj, A. W. Götz, and F. Paesani, "Toward chemical accuracy in the description of ion–water interactions through many-body representations. Alkali-water dimer potential energy surfaces," J. Chem. Phys. **147**, 161715 (2017).

[59]M. Riera, S. E. Brown, and F. Paesani, "Isomeric equilibria, nuclear quantum effects, and vibrational spectra of $M^+(H_2O)_{n=1-3}$ clusters, with M = Li, Na, K, Rb, and Cs, through many-body representations," J. Phys. Chem. A **122**, 5811–5821 (2018).

[60]P. Bajaj, M. Riera, J. K. Lin, Y. E. Mendoza Montijo, J. Gazca, and F. Paesani, "Halide ion microhydration: Structure, energetics, and spectroscopy of small halide–water clusters," J. Phys. Chem. A **123**, 2843–2852 (2019).

[61]P. Bajaj, J. O. Richardson, and F. Paesani, "Ion-mediated hydrogen-bond rearrangement through tunnelling in the iodide–dihydrate complex," Nat. Chem. **11**, 367–374 (2019).

[62]P. Bajaj, D. Zhuang, and F. Paesani, "Specific ion effects on hydrogen-bond rearrangements in the halide–dihydrate complexes," J. Phys. Chem. Lett. **10**, 2823–2828 (2019).

[63]M. Riera, J. J. Talbot, R. P. Steele, and F. Paesani, "Infrared signatures of isomer selectivity and symmetry breaking in the $Cs^+(H_2O)_3$ complex using many-body potential energy functions," J. Chem. Phys. **153**, 044306 (2020).

[64]Y. Zhai, A. Caruso, S. Gao, and F. Paesani, "Active learning of many-body configuration space: Application to the Cs^+–water MB-nrg potential energy function as a case study," J. Chem. Phys. **152**, 144103 (2020).

[65]D. Hankins, J. Moskowitz, and F. Stillinger, "Water molecule interactions," J. Chem. Phys. **53**, 4544–4554 (1970).

[66]F. Paesani, "Getting the right answers for the right reasons: Toward predictive molecular simulations of water with many-body potential energy functions," Acc. Chem. Res. **49**, 1844–1851 (2016).

[67]E. Lambros, S. Dasgupta, E. Palos, S. Swee, J. Hu, and F. Paesani, "General many-body framework for data-driven potentials with arbitrary quantum mechanical accuracy: Water as a case study," ChemRxiv , https://doi.org/10.26434/chemrxiv.14710815.v1 (2021).

[68]S. K. Reddy, S. C. Straight, P. Bajaj, C. Huy Pham, M. Riera, D. R. Moberg, M. A. Morales, C. Knight, A. W. Götz, and F. Paesani, "On the accuracy of the MB-pol many-body poten-

tial for water: Interaction energies, vibrational frequencies, and classical thermodynamic and dynamical properties from clusters to liquid water and ice," J. Chem. Phys. **145**, 194504 (2016).

[69]J. O. Richardson, C. Pérez, S. Lobsiger, A. A. Reid, B. Temelso, G. C. Shields, Z. Kisiel, D. J. Wales, B. H. Pate, and S. C. Althorpe, "Concerted hydrogen-bond breaking by quantum tunneling in the water hexamer prism," Science **351**, 1310–1313 (2016).

[70]W. T. Cole, J. D. Farrell, D. J. Wales, and R. J. Saykally, "Structure and torsional dynamics of the water octamer from thz laser spectroscopy near 215 μm," Science **352**, 1194–1197 (2016).

[71]J. D. Mallory and V. A. Mandelshtam, "Diffusion Monte Carlo studies of MB-pol $(H_2O)_{2-6}$ and $(D_2O)_{2-6}$ clusters: Structures and binding energies," J. Chem. Phys. **145**, 064308 (2016).

[72]P. E. Videla, P. J. Rossky, and D. Laria, "Communication: Isotopic effects on tunneling motions in the water trimer," J. Chem. Phys. **144**, 061101 (2016).

[73]S. E. Brown, A. W. Götz, X. Cheng, R. P. Steele, V. A. Mandelshtam, and F. Paesani, "Monitoring water clusters "melt" through vibrational spectroscopy," J. Am. Chem. Soc. **139**, 7082–7088 (2017).

[74]C. L. Vaillant and M. T. Cvitaš, "Rotation-tunneling spectrum of the water dimer from instanton theory," Phys. Chem. Chem. Phys. **20**, 26809–26813 (2018).

[75]C. Vaillant, D. Wales, and S. Althorpe, "Tunneling splittings from path-integral molecular dynamics using a Langevin thermostat," J. Chem. Phys. **148**, 234102 (2018).

[76]M. Schmidt and P.-N. Roy, "Path integral molecular dynamic simulation of flexible molecular systems in their ground state: Application to the water dimer," J. Chem. Phys. **148**, 124116 (2018).

[77]K. P. Bishop and P.-N. Roy, "Quantum mechanical free energy profiles with post-quantization restraints: Binding free energy of the water dimer over a broad range of temperatures," J. Chem. Phys. **148**, 102303 (2018).

[78]P. E. Videla, P. J. Rossky, and D. Laria, "Isotopic equilibria in aqueous clusters at low temperatures: Insights from the MB-pol many-body potential," J. Chem. Phys. **148**, 084303 (2018).

[79]N. R. Samala and N. Agmon, "Temperature dependence of intramolecular vibrational bands in small water clusters," J. Phys. Chem. B **123**, 9428–9442 (2019).

[80]M. T. Cvitaš and J. O. Richardson, "Quantum tunnelling pathways of the water pentamer," Phys. Chem. Chem. Phys. **22**, 1035–1044 (2020).

[81]G. R. Medders and F. Paesani, "Infrared and raman spectroscopy of liquid water through "first-principles" many-body molecular dynamics," J. Chem. Theory Comput. **11**, 1145–1154 (2015).

[82]S. C. Straight and F. Paesani, "Exploring electrostatic effects on the hydrogen bond network of liquid water through many-body molecular dynamics," J. Phys. Chem. B **120**, 8539–8546 (2016).

[83]S. K. Reddy, D. R. Moberg, S. C. Straight, and F. Paesani, "Temperature-dependent vibrational spectra and structure of liquid water from classical and quantum simulations with the MB-pol potential energy function," J. Chem. Phys. **147**, 244504 (2017).

[84]K. M. Hunter, F. A. Shakib, and F. Paesani, "Disentangling coupling effects in the infrared spectra of liquid water," J. Phys. Chem. B **122**, 10754–10761 (2018).

[85]Z. Sun, L. Zheng, M. Chen, M. L. Klein, F. Paesani, and X. Wu, "Electron-hole theory of the effect of quantum nuclei on the X-ray absorption spectra of liquid water," Phys. Rev. Lett. **121**, 137401 (2018).

[86]A. P. Gaiduk, T. A. Pham, M. Govoni, F. Paesani, and G. Galli, "Electron affinity of liquid water," Nat. Commun. **9**, 1–6 (2018).

[87]V. Cruzeiro, A. Wildman, X. Li, and F. Paesani, "Relationship between hydrogen-bonding motifs and the $1b_1$ splitting in the x-ray emission spectrum of liquid water," J. Phys. Chem. Lett. **12**, 3996–4002 (2021).

[88]G. R. Medders and F. Paesani, "Dissecting the molecular structure of the air/water interface from quantum simulations of the sum-frequency generation spectrum," J. Am. Chem. Soc. **138**, 3912–3919 (2016).

[89]D. R. Moberg, S. C. Straight, and F. Paesani, "Temperature dependence of the air/water interface revealed by polarization sensitive sum-frequency generation spectroscopy," J. Phys. Chem. B **122**, 4356–4365 (2018).

[90]S. Sun, F. Tang, S. Imoto, D. R. Moberg, T. Ohto, F. Paesani, M. Bonn, E. H. Backus, and Y. Nagata, "Orientational distribution of free OH groups of interfacial water is exponential," Phys. Rev. Lett. **121**, 246101 (2018).

[91]S. Sengupta, D. R. Moberg, F. Paesani, and E. Tyrode, "Neat water–vapor interface: Proton continuum and the nonresonant background," J. Phys. Chem. Lett. **9**, 6744–6749 (2018).

[92]M. C. Muniz, T. E. Gartner III, M. Riera, C. Knight, S. Yue, F. Paesani, and A. Z. Panagiotopoulos, "Vapor-liquid equilibrium of water with the MB-pol many-body potential," J. Chem. Phys. **154**, 211103 (2021).

[93]C. H. Pham, S. K. Reddy, K. Chen, C. Knight, and F. Paesani, "Many-body interactions in ice," J. Chem. Theory Comput. **13**, 1778–1784 (2017).

[94]D. R. Moberg, S. C. Straight, C. Knight, and F. Paesani, "Molecular origin of the vibrational structure of ice I_h," J. Phys. Chem. Lett. **8**, 2579–2583 (2017).

[95]D. R. Moberg, P. J. Sharp, and F. Paesani, "Molecular-level interpretation of vibrational spectra of ordered ice phases," J. Phys. Chem. B **122**, 10572–10581 (2018).

[96]D. R. Moberg, D. Becker, C. W. Dierking, F. Zurheide, B. Bandow, U. Buck, A. Hudait, V. Molinero, F. Paesani, and T. Zeuch, "The end of ice I," Proc. Natl. Acad. Sci. U.S.A. **116**, 24413–24419 (2019).

[97]T. Adler, G. Knizia, and H. Werner, "A simple and efficient CCSD(T)-F12 approximation," J. Chem. Phys. **127**, 221106–221106 (2007).

[98]G. Knizia, T. B. Adler, and H.-J. Werner, "Simplified CCSD(T)-F12 methods: Theory and benchmarks," J. Chem. Phys. **130**, 054104 (2009).

[99]H. Partridge and D. W. Schwenke, "The determination of an accurate isotope dependent potential energy surface for water from extensive ab initio calculations and experimental data," J. Chem. Phys. **106**, 4618–4639 (1997).

[100]K. Tang and J. P. Toennies, "An improved simple model for the van der waals potential based on universal damping functions for the dispersion coefficients," J. Chem. Phys. **80**, 3726–3741 (1984).

[101]J. G. Hill, K. A. Peterson, G. Knizia, and H.-J. Werner, "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. **131**, 194105 (2009).

[102]U. Góra, R. Podeszwa, W. Cencek, and K. Szalewicz, "Interaction energies of large clusters from many-body expansion," J. Chem. Phys. **135**, 224102 (2011).

[103]B. P. Pritchard, D. Altarawy, B. Didier, T. D. Gibson, and T. L. Windus, "New basis set exchange: An open, up-to-date resource for the molecular sciences community," J. Chem. Inf. Model. **59**, 4814–4820 (2019).

[104]T. H. Dunning, "Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen," J. Chem. Phys. **90**, 1007–1023 (1989).

[105]R. A. Kendall, T. H. Dunning, and R. J. Harrison, "Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions," J. Chem. Phys. **96**, 6796–6806 (1992).

[106]D. E. Woon and T. H. Dunning, "Gaussian basis sets for use in correlated molecular calculations. III. The atoms aluminum through argon," J. Chem. Phys. **98**, 1358–1371 (1993).

[107]H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, and M. Schütz, "Molpro: A general-purpose quantum chemistry program package," Wiley Interdiscip. Rev. Comput. Mol. Sci. **2**, 242–253 (2012).

[108]M. Riera, E. P. Yeh, and F. Paesani, "Data-driven many-body models for molecular fluids: CO_2/H_2O mixtures as a case study," J. Chem. Theory. Comput. **16**, 2246–2257 (2020).

[109]M. Riera, A. Hirales, R. Ghosh, and F. Paesani, "Data-driven many-body models with chemical accuracy for CH_4/H_2O mixtures," J. Phys. Chem. B **124**, 11207–11221 (2020).

[110]V. W. D. Cruzeiro, E. Lambros, M. Riera, R. Roy, F. Paesani, and A. W. Götz, "Highly accurate many-body potentials for simulations of N_2O_5 in water: Benchmarks, development, and validation," J. Chem. Theory. Comput. , https://doi.org/10.1021/acs.jctc.1c00069 (2020).

[111]A. N. Tihonov, "Solution of incorrectly formulated problems and the regularization method," Soviet Math. **4**, 1035–1038 (1963).

[112]M. Tuckerman, B. J. Berne, and G. J. Martyna, "Reversible multiple time scale molecular dynamics," J. Chem. Phys. **97**, 1990–2001 (1992).

[113]J. Kolafa, "Time-reversible always stable predictor–corrector method for molecular dynamics of polarizable molecules," J. Comp. Chem. **25**, 335–342 (2004).

[114]G. J. Martyna, M. L. Klein, and M. Tuckerman, "Nosé–Hoover chains: The canonical ensemble via continuous dynamics," J. Chem. Phys. **97**, 2635–2643 (1992).

[115]B. J. Berne and D. Thirumalai, "On the simulation of quantum systems: Path integral methods," Annu. Rev. Phys. Chem. **37**, 401–424 (1986).

[116]W. Smith and T. Forester, "DL_POLY_2. 0: A general-purpose parallel molecular dynamics simulation package," J. Mol. Graph. **14**, 136–141 (1996).

[117]H. W. Horn, W. C. Swope, J. W. Pitera, J. D. Madura, T. J. Dick, G. L. Hura, and T. Head-Gordon, "Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew," J. Chem. Phys. **120**, 9665–9678 (2004).

[118]I. S. Joung and T. E. Cheatham III, "Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations," J. Phys. Chem. B **112**, 9020–9041 (2008).

[119]D. Case, H. Aktulga, K. Belfon, I. Ben-Shalom, S. Brozell, D. Cerutti, T. Cheatham III, V. Cruzeiro, T. Darden, R. Duke, G. Giambasu, M. Gilson, H. Gohlke, A. Goetz, R. Harris, S. Izadi, S. Izmailov, C. Jin, K. Kasavajhala, M. Kaymak, E. King, A. Kovalenko, T. Kurtzman, T. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, M. Machado, V. Man,

M. Manathunga, K. Merz, Y. Miao, O. Mikhailovskii, G. Monard, H. Nguyen, K. O'Hearn, A. Onufriev, F. Pan, S. Pantano, R. Qi, A. Rahnamoun, D. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, J. Shen, C. Simmerling, N. Skrynnikov, J. Smith, J. Swails, R. Walker, J. Wang, H. Wei, R. Wolf, X. Wu, Y. Xue, D. York, S. Zhao, and P. Kollman, "Amber 2021," `https://ambermd.org` (2021).

[120]J. J. Rehr, J. J. Kas, F. D. Vila, M. P. Prange, and K. Jorissen, "Parameter-free calculations of X-ray spectra with FEFF9," Phys. Chem. Chem. Phys. **12**, 5503–5513 (2010).

[121]J. J. Rehr, J. J. Kas, M. P. Prange, A. P. Sorini, Y. Takimoto, and F. Vila, "Ab initio theory and calculations of X-ray spectra," C. R. Phys. **10**, 548–559 (2009).

[122]J. J. Rehr and R. C. Albers, "Theoretical approaches to X-ray absorption fine structure," Rev. Mod. Phys. **72**, 621 (2000).

[123]D. Zhuang, M. Riera, G. K. Schenter, J. L. Fulton, and F. Paesani, "Many-body effects determine the local hydration structure of Cs^+ in solution," J. Phys. Chem. Lett. **10**, 406–412 (2019).

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

[125]P. Bajaj, X.-G. Wang, T. CarringtonJr, and F. Paesani, "Vibrational spectra of halide-water dimers: Insights on ion hydration from full-dimensional quantum calculations on many-body potential energy surfaces," J. Chem. Phys. **148**, 102321 (2017).