JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

pubs.acs.org/jcim

Article

Accurate Simulations of Water and Aqueous Solutions through Fine-Tuned Dispersion-Corrected Density Functional Theory and Machine-Learning Interatomic Potentials

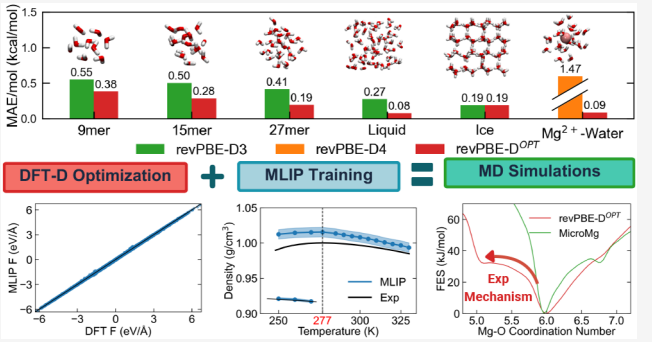
Alfonso Ferretti, Giacomo Melani, Luca Benedetti, Robert A. Sorodoc, Alessando Fortunelli,* and Giuseppe Brancato*

Cite This: J. Chem. Inf. Model. 2025, 65, 12437−12447

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Dispersion-corrected density functional theory (DFT-D) is widely employed to model large molecular systems at an affordable computational cost and to develop machine-learning interatomic potentials (MLIPs), enabling reliable molecular dynamics (MD) simulations of condensed-phase systems. Yet, given a molecular system, the choice of a specific DFT-D model that can achieve the necessary accuracy over an extended range of physicochemical properties and conditions is generally not trivial. Here, we report an effective computational strategy for enhancing the accuracy of standard DFT-D models toward high-level quantum mechanical data and for developing MLIPs preserving the same high fidelity. Taking water as a paradigmatic example, we derive a novel MLIP and demonstrate that its use allows us to accurately predict a wide range of properties in diverse forms, from small clusters to bulk liquid and ice, such as radial distribution functions, fusion/vaporization enthalpies, diffusion constants, and density isobars, capturing remarkably well its peculiar and anomalous behavior, often elusive even to standard first-principle MD simulations. Furthermore, we show how the same computational strategy can be readily extended to treat aqueous solutions. Considering $MgCl_2$ in water as a test case, we develop a MLIP and use it to predict the metal ion hydration structure and the water exchange dynamics exhibiting a significantly improved agreement with experiments with respect to both standard DFT-D and classical force fields.



INTRODUCTION

Machine learning interatomic potentials (MLIPs) based on neural networks are becoming the *de facto* approach to reliably simulating condensed-phase molecular systems at an affordable computational cost.[1−3] However, the quality and performances of MLIPs are mostly determined by two key factors: (1) the ability to correctly infer short-range interactions that implicitly include long-range interactions such as dispersion and electrostatics (since long-range interactions are still difficult to incorporate into MLIPs)[4] and (2) the generation of large and accurate training databases relative to condensed-phase molecular structures,[5] especially those representing complex and disordered liquid phases. In this regard, reliable training data for MLIPs can be conveniently obtained through density functional theory (DFT) methods,[6−8] which generally show a good balance between accuracy and efficiency. Indeed, DFT offers a computationally more favorable scaling with the number of electrons than gold standard quantum mechanical methods (i.e., explicitly correlated coupled cluster or quantum Monte Carlo methods), although approximate exchange-correlation functionals may compromise accuracy, e.g., underestimating van der Waals (VdW) dispersion interactions.[9] For a given molecular system, it is then usually not trivial to choose a density functional that achieves the needed accuracy over a wide range of physicochemical properties and conditions (from small isolated clusters to the liquid and solid phases, and from local structural properties to basic thermodynamics, such as bulk density, diffusion, enthalpy of melting/vaporization, vibrational density of states, etc.). This is clearly illustrated by the case of water, that is still debated, despite its intrinsic interest justifying the tremendous efforts devoted to its computational description and despite the improvements demonstrated recently by the latest DFT approximations.[8,10−15] In this context, recent developments in density-corrected SCAN functionals, belonging to the third (meta-GGA) rung of the Jacob's ladder of DFT approximations, are

Received: September 3, 2025
Revised: November 3, 2025
Accepted: November 5, 2025
Published: November 12, 2025

particularly noteworthy, such as the works of Paesani and coworkers[16−18] in which applications to pure water systems led to the accurate description of several water properties when combined with a many-body expansion potential.[13] Nevertheless, weak dispersion forces are generally not well represented by these functionals and further corrections have been proposed.[14] Among others, two recently proposed approaches are the experiment directed simulations[19] (EDS) or the "Δ-learning"[20] used to promote DFT to Coupled-Cluster accuracy. In this context, the addition of semiempirical dispersion corrections to standard Kohn−Sham DFT (DFT-D) models[21] represents one of the most widely used and well-assessed in silico approaches for treating complex molecular systems because of the demonstrated capability to recover the missing VdW interactions affecting popular exchange-correlation functionals.[22,23]

Recently, we proposed the idea of fine-tuning Grimme's DFT-D models toward specific molecular systems, so as to elevate the accuracy of the computed noncovalent interactions up to high-level reference calculations (i.e., achieving a mean absolute error of 0.1 kcal/mol per molecule),[24,25] with the final goal of improving the description of condensed-phase properties. It should be noted, however, that this approach is physically sound and computationally feasible as far as the focus is on standard equilibrium and dynamical properties, while more subtle electronic properties may need different strategies.[26−28] Accordingly, the DFT-D models were reoptimized by tweaking the empirical parameters of the potential (i.e., scaling factors) to match benchmark interaction energies issuing from hundreds of nonequilibrium molecular configurations. This was achieved by factoring out from the reference calculations the energy deviations of the one-body contributions (i.e., monomer energy deviations), thus refining the dispersion corrections more effectively. Test calculations carried out on acetonitrile[24] and water[25] proved this approach successful in obtaining highly accurate DFT-D models with respect to reference diffusion Monte Carlo (DMC),[29] making it particularly well-suited for large-scale applications, such as *ab initio* molecular dynamics (AIMD) simulations or data generation for MLIPs.

Here, based on our previous work, we applied our DFT-D optimization protocol to water and an aqueous solution of $MgCl_2$. Then, we exploited the high-fidelity data generated by the optimized DFT to build effective MLIPs based on the message-passing atomic cluster expansion (MACE) architecture.[30] Remarkably, the obtained MLIPs are capable of representing liquid water and ice at the level of the best state-of-the-art potentials. Moreover, we demonstrate that the water MLIP is transferable to salt solutions, showing that the $MgCl_2$/water MLIP we derive starting from the pure water MLIP is able to correctly represent both the hydration structure and the water exchange mechanisms in the first solvation shell of the magnesium ion in better agreement with experiment than alternative approaches.

METHODS

DFT-D Optimization. The *rev*PBE-D3 model was reoptimized toward the reference DMC data set of Alfè et al.[12] on water interaction energies obtained from liquid-phase configurations containing either 32 or 64 water molecules, enforcing periodic boundary conditions. Following the procedure originally presented in refs 24,25 we obtained an optimal parametrization that minimized the mean absolute error (MAE) in the interaction energy relative to diffusion Monte Carlo (DMC) calculations by excluding the one-body energy contribution issuing from individual water molecules using the Partridge and Schwenke potential.[31] All DFT calculations were carried out using a plane-wave basis set and soft pseudopotentials[32] as implemented in the Quantum Espresso package.[33] No significant differences were observed in comparison with all-electron quantum mechanical calculations employing the extended aug-cc-pVTZ[34] basis set (see Section S1 for further details). To demonstrate the modularity of our approach, after optimizing the *rev*PBE-D3 parameters for the water−water interactions, we further reoptimized also the Mg^{2+}-water interaction as provided by the *rev*PBE-D4 model. In this case, 40 first-hydration-shell configurations (i.e., $Mg(H_2O)_6^{2+}$) and eight second-shell configurations (i.e., $Mg(H_2O)_6(H_2O)_{13-17}$) were extracted from classical MD trajectories and evaluated at the DLPNO−CCSD(T)-F12 level using the cc-PVQZ basis set as implemented in ORCA.[35] For each configuration, we computed (i) the total cluster energy and (ii) the energy of the water fragment frozen in the same geometry. The interaction energy for method X (X = DFT or DLPNO−CCSD(T)) was obtained as follows

$$\Delta E_{\mathrm{int,X}} = E_{\mathrm{tot,X}} - E_{\mathrm{wat,X}} - E_{\mathrm{Mg^{2+},X}} \quad (1)$$

Energy deviations (i.e., $\Delta E_{\mathrm{int,DFT}} - \Delta E_{\mathrm{int,CC}}$) were minimized by tuning the S_8 coefficient of the D4 potential to reproduce the metal ion−water interaction energies, while keeping the same optimized parametrization for the water−water description (see Section S1 for further details).

Machine Learning Interatomic Potentials. We used the functionals so optimized to develop a reference database for water MLIP training using the MACE[30] architecture (see Sections S2 and S3 for further details). The construction of the fitting database is critical in determining MLIP accuracy.[36,37] A flowchart of the main steps of the computational protocol is illustrated in Figure 1. Initially, we generated 10 200 representative liquid-phase structures through classical MD simulations using the TIP4P/2005[38] water model. To build up the training set, energies and atomic forces were evaluated at the *rev*PBE-D3^{OPT} level and used to train the first version of the MLIP, MLIP$^{(1)}$. In an active learning strategy,[39] the obtained MLIP$^{(1)}$ model was then used to sample additional liquid water structures and enrich the original training set, from which we developed an intermediate MLIP model, MLIP$^{(2)}$. In a second step, we used MLIP$^{(2)}$ to generate further liquid configurations from classical MD and path-integral MD (PI-MD) simulations to ensure well-balanced performances when including nuclear quantum effects. Upon further training, the final MLIP model (i.e., the third version MLIP$^{(3)}$ or MLIP-*rev*PBE-D3^{OPT}) was obtained. MLIP-*rev*PBE-D3^{OPT} was initially validated on a large number of *rev*PBE-D3^{OPT} interaction energies obtained from liquid water and ice structures not included in the training set (Figure S4) and energies and forces from Born−Oppenheimer AIMD simulation of liquid water (Figure S8), achieving excellent results in terms of both interaction energies (RMSE of 0.1 meV/atom) and atomic forces (RMSE of 8 meV/Å). An analogous active learning protocol was applied to develop an effective MLIP for describing Mg^{2+} in water. A system containing $MgCl_2$ in 125 water molecules was initially simulated using classical MD, then selected configurations were re-evaluated at the *rev*PBE−D^{OPT} level and used to train an initial MLIP Mg^{2+}−water potential. The first MLIP model was used to carry out

https://doi.org/10.1021/acs.jcim.5c02079
J. Chem. Inf. Model. 2025, 65, 12437−12447

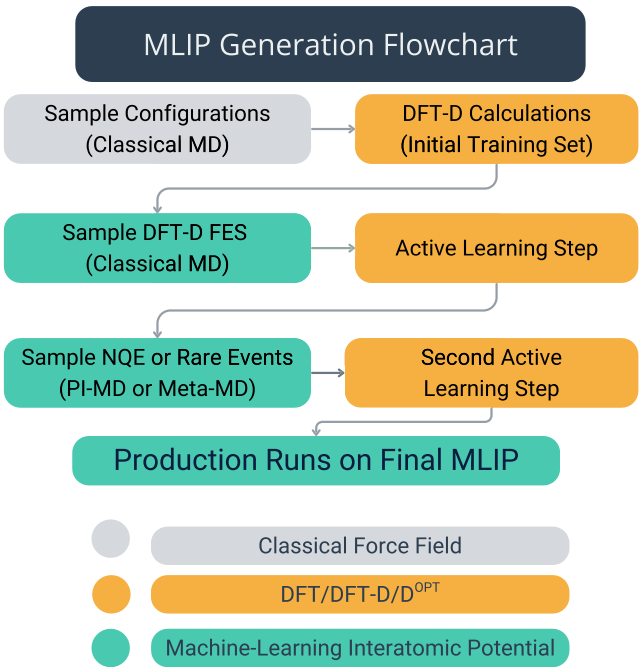


Figure 1. Flowchart illustrating the main steps followed during the MLIP generation. First, uncorrelated structures of liquid water or aqueous solutions were generated through classical MD simulations, evaluating interaction energy and atomic forces for each one at the *rev*PBE-D^{OPT} level. The initial MLIP was then used to carry out active learning in two further steps, the first one to correctly sample the free energy surface relative to the DFT model used and the second to sample nuclear quantum effect, for the case of water, or rare events (i.e., water exchange the first solvation shell) for the case of Mg^{2+} aqueous solutions.

additional MD and well-tempered metadynamics[40,41] simulations in which the ion−water coordination number served as a collective variable (see Section 2.5 below), efficiently exploring the water coordination space. The resulting enlarged data set was employed for the final training, yielding a MLIP that reproduced *rev*PBE−D^{OPT} energies and forces with the same accuracy achieved for pure water (RMSE energy per atom of 0.2 meV/atom, RMSE forces of 9 meV/Å). The same configurations (i.e., $MgCl_2$ in water) were also used to develop a MLIP model based upon standard DFT-D3(0) for water−water interaction and standard DFT-D4 for Mg^{2+}-water interactions for the sake of comparison, referred to as MLIP-*rev*PBE-D4. We note in passing that the parametrization procedure we exploit here could be applied not only to MLIPs, but also to classical force fields, some of which have been proven capable of reproducing the properties of water accurately.[42]

Molecular Dynamics Simulations. Production Molecular dynamics (MD) simulations were performed using machine learning interatomic potentials (MLIPs) based on the MACE framework.[30] Classical simulations of bulk liquid water contained $N = 510$ molecules in a cubic box. The equations of motion were integrated with LAMMPS[43] using a 0.5[44] fs time step. The starting system was first equilibrated in the *NVT* ensemble for 1 ns using a Langevin thermostat. Structural observables—radial distribution functions (RDFs), hydrogen−bond angle distributions, and the tetrahedral order parameter q—were extracted from separate *NVT* runs, while dynamic quantities (diffusion coefficients, vibrational density of states, IR spectra) were evaluated from *NVE* trajectories (see SI, Section S4 for more details). Density−temperature curves (Figure 4) were obtained from *NpT* trajectories of 2.5 ns under a Langevin thermostat and barostat. Equilibrium properties were averaged over the final ns.

Quantum nuclear effects were included through either path−integral MD (PI-MD) with the PIGLET scheme[45] or Thermosttated Ring-Polymer Molecular Dynamics[46] (TRPMD) as implemented in i-PI[47] and using ASE[48] implementation of MACE to evaluate energy and forces. The systems consisted of 64 water molecules represented by 56 beads using a time step of 0.25 fs. A PIGLET generalized Langevin thermostat was applied to the centroid and internal modes to replicate *NVT* conditions, while *NpT* simulations used an additional isotropic Langevin barostat.

Thermodynamic and Dynamical Properties. The vaporization and fusion enthalpies were evaluated from the following expressions

$$\Delta H_{\mathrm{fus}} = H_l - H_s = (\langle U_l \rangle - \langle U_s \rangle) + PV \tag{2}$$

$$\Delta H_{\mathrm{vap}} = H_g - H_l = (\langle U_g \rangle - \langle U_l \rangle) + RT \tag{3}$$

where $\langle U_x \rangle$ denotes the ensemble-average internal energy of the system, with s the solid, l the liquid, and g the gas-phase. Energy averages were obtained from 200 ps *NVT* simulations, starting from initial configurations previously equilibrated under *NpT* conditions. The *PV* and *RT* terms were set from the corresponding control values of the relative *NVT* and *NpT* simulations. Isobaric heat capacity, c_p and thermal expansion coefficient α_P defined as

$$c_p = \left(\frac{\partial H}{\partial T}\right)_p \tag{4}$$

$$\alpha_p = -\frac{1}{V}\left(\frac{\partial V}{\partial T}\right)_p = -\frac{1}{\rho}\left(\frac{\partial \rho}{\partial T}\right)_p \tag{5}$$

were evaluated following the procedures described in Reddy et al.[16] Diffusion coefficients were calculated at 280, 300, and 320 K by linear regression of the mean square displacement (MSD) slopes, excluding the initial and final picoseconds to focus solely on the diffusive regime.

$$D_L = \lim_{t\to\infty} \frac{1}{6t}\langle |\mathbf{r}(t) - \mathbf{r}(0)|^2 \rangle \tag{6}$$

For classical MD simulations at 300 K, diffusion coefficients were averaged over 10 simulation replicas of 180 ps each. At 280 and 320 K, averages were obtained from 10 replicas of 10 ps each. The TRPMD simulations used five replicas of 8 ps per temperature point. Errors were estimated via bootstrap resampling, with corrections applied to account for finite-size effects of the simulation cell box.[49]

$$D_\infty = D_L + \frac{k_B T \xi}{6\pi\eta L} \tag{7}$$

The vibrational density of states (vDOS) was obtained from *NVE* trajectories for classical dynamics and using TRPMD to include nuclear quantum effects. We ran 13 (30 ps) independent simulations for TRPMD and 10 (75 ps) for classical MD simulations. The vDOS was computed by applying a Hann window to the velocity autocorrelation function, followed by a Fourier transform. The resulting spectra were corrected for quantum effects and further

https://doi.org/10.1021/acs.jcim.5c02079
J. Chem. Inf. Model. 2025, 65, 12437−12447

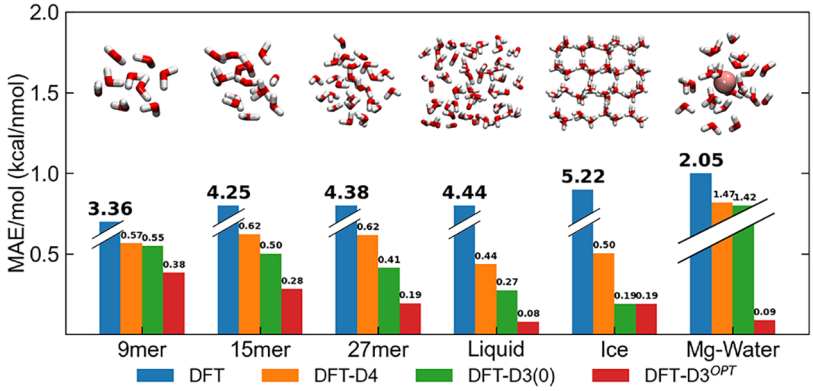


Figure 2. Mean absolute error (MAE) per molecule of the interaction energy computed using the standard *rev*PBE-D4, *rev*PBE-D3, and reoptimized *rev*PBE-D3^{OPT} (or D4 for the Mg-Water interaction) model with respect to DMC on isolated water clusters of growing size (i.e., 9-mer, 15-mer, and 27-mer),[64] liquid-phase structures (including 32 and 64 water molecules),[12] ice polymorph structures from the DMC-ICE13[11] data set. The references for the Mg-Water systems were evaluated at the DLPNO−CCSD(T)-F12 with the cc-PVQZ basis set

adjusted according to the procedure described in Melani et al.[50]

$$\mathrm{vDOS}(\omega) = \left(\frac{\hbar\omega}{k_B T}\cdot\frac{1}{1-e^{-\hbar\omega/k_B T}}\right)\cdot\int_0^{T_{max}} C_{vv}(t)\cdot \sin^2\left(\frac{\pi t}{T_{max}}\right)\cdot\cos(\omega t)\mathrm{d}t \quad (8)$$

where T_{max} is the total length of the time window and N is the total number of atoms in the system. C_{vv} is the velocity−velocity autocorrelation function

$$C_{vv}(t) = \frac{1}{N}\sum_{i=1}^{N}\langle \mathbf{v}_i(0)\cdot\mathbf{v}_i(t)\rangle \quad (9)$$

Finally, the spectra were averaged for different runs.

Free Energy of Ion Coordination. To evaluate the free energy change associated with the first−shell water coordination, we adopted the water coordination number s, as described in refs 51,52, as a continuous collective variable defined as follows

$$s = \sum_{i=1}^{N}\left[1 - \frac{1}{1+\exp[a(r_i - r_0)]}\right] \quad (10)$$

where r_i is the ion−oxygen distance of the i-th solvent molecule, N is the total number of water molecules in the simulation box, r_0 denotes the ion−oxygen cutoff distance, and a (=4.0 Å^{-1}) controls the steepness of the Fermi−type switching function. The cutoff r_0 was set on the basis of the first minimum of the ion−oxygen radial distribution function (in this case, r_0 = 3.0 Å). An extensive sampling with well-tempered metadynamics[40,41] (computational details of the meta-MD simulations can be found in the Supporting Information in Section S4) was carried out using LAMMPS in combination with the PLUMED library (v2.9.0).[53,54]

RESULTS AND DISCUSSION

DFT-D Optimization. In the present work, we adopted the *rev*PBE GGA exchange-correlation functional[55] and performed a careful reoptimization of the D3(0)[56] potential since the standard *rev*PBE-D3 reported already quite satisfactory results on various properties of water under different physical conditions,[57−60] although the good performance of *rev*PBE-D3 for liquid water could be due to a favorable compensation between functional-driven and density-driven errors.[61] Note that *rev*PBE-D3 resulted the best performing DFT approximation toward the DMC-ICE13 data set.[11] Besides, *rev*PBE-D3 is better suited for AIMD simulations or large-scale calculations than recently proposed hybrid or meta-GGA functionals, due to its computational efficiency. Dispersion corrections were optimized against DMC benchmark data.[12] A MACE MLIP was then carefully trained against *rev*PBE data via a judicious choice of the training database (see Section 2.2 and Figure 1), and then assessed for its ability to predict key structural, thermodynamic, and dynamic properties of water, ice, and aqueous solutions of Mg^{2+}, providing a remarkable agreement with available experimental data, as discussed in the following.

The quality of the refined *rev*PBE-D3 model (hereafter referred to as *rev*PBE-D3^{OPT}) is illustrated in Figure 2 where relative errors in the interaction energy per molecule (MAE/mol) were compared to the results issued from the standard *rev*PBE, *rev*PBE-D3/D4[62] and the hybrid *rev*PBE0[63]-D4 on various systems, including: (i) water clusters of growing size (i.e., 9-mer, 15-mer, and 27-mer),[64] (ii) liquid water configurations (used in the optimization step),[12] and (iii) different ice polymorph structures from the ICE13 data set.[11] Our optimized model consistently improved the estimated interaction energy across all examined systems, from small water clusters to multiple liquid and ice structures, reaching a remarkably low MAE/mol of less than 0.1 kcal/mol for the liquid phase (i.e., within DMC intrinsic statistical error). Note that energy deviations remained significantly small throughout the individual structures of the benchmark data sets, showing a noticeable improvement compared to the standard D3 and D4 parametrization (Figures S3−S5). Similar results were obtained on test calculations on the optimized neutral water clusters of the WATER27 data set[65] (MAE/mol: 0.11 kcal/mol, while D3 and D4 reported 0.21 and 0.20 kcal/mol, respectively). Besides, the present optimized model provided the same very good agreement as the standard *rev*PBE-D3 functional with the recent DMC-ICE13 ice polymorph energies,[11] showing an error of only 0.19 kcal/mol per molecule. The optimization of the magnesium-water interaction is also reported in Figure 2 showing a significant increase in accuracy with respect to the default values. We have optimized the empirical D4 correction for the interaction

https://doi.org/10.1021/acs.jcim.5c02079
J. Chem. Inf. Model. 2025, 65, 12437−12447

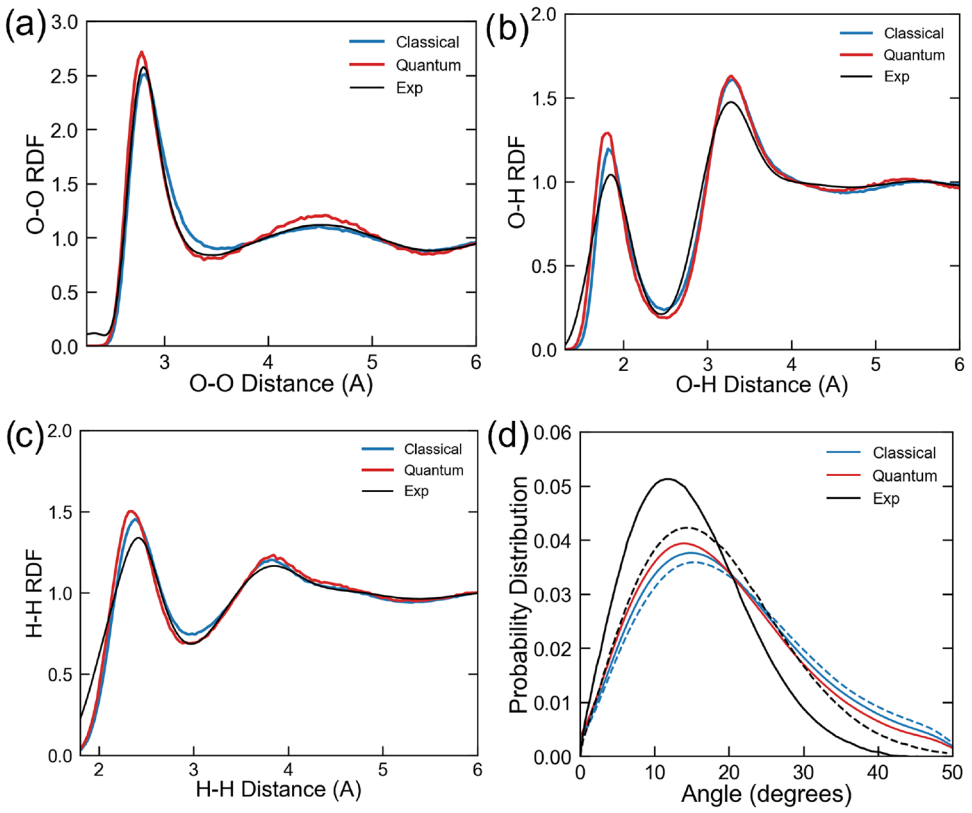


Figure 3. Experimental and simulated (a) O−O, (b) O−H, and (c) H−H RDFs obtained from classical and path-integral MD simulations using the MLIP-*rev*PBE-D3^{OPT} model. O−O Experiment was replotted from Skinner et al.,[67] O−H and H−H were replotted from Soper et al.[68] (d) Experimental and simulated H-bond angle distributions obtained at room temperature (solid line) and T = 323 K (dashed line). Experiments were extracted and replotted/reproduced from Modig et al.[69]

between Mg^{2+} and water changing the S_8 parameter to minimize the error, with respect to the reference data, in the interaction energy reported in eq 1. This was done using only two-shell structures as a reference because they are more representative of the environment observed in molecular dynamics. While the standard parametrizations of the D corrections have proven to be sufficiently accurate for water−water interactions, this is not true for the Mg^{2+}-water interaction. In Figure 2, we show how standard parametrizations do not mitigate the initial error of the DFT (MAE/mol(DFT): 2.05 kcal/mol MAE/mol(DFT-D4): 1.42 kcal/mol). To obtain good agreement (MAE/mol(DFT-D^{OPT}): 0.09 kcal/mol) with the DLPNO−CCSD(T)-F12 reference calculations, we increased the value of the S_8 parameter from 1.75 to 9.01. In Figure S5 we report the deviations with respect to DLPNO−CCSD(T)-F12 for all the single and double shell cluster structures. The MAE on the single-shell $Mg(H_2O)_6^{2+}$ clusters rises to 1.21 kcal/mol, indicating a slight overbinding caused by tuning on the larger set, yet the optimized functional still outperforms plain *rev*PBE (MAE/mol: 2.98 kcal/mol) and the default D4 variant (MAE/mol: 1.79 kcal/mol) by a comfortable margin. A recent study[66] demonstrates that DC-SCAN is capable of representing ion−water clusters with near-chemical accuracy, albeit at significantly higher computational cost (i.e., meta-GGA on HF densities). In this work, the ad hoc optimization of D4 yielded a system-specific *rev*PBE-D4 model that attains negligible errors in the ion−solvent interaction energy, comparable to those reported by Palos et al.,[66] at lower cost, thus enabling longer time scales and larger systems. Accordingly, our protocol is advantageous when high-throughput or large-scale simulations are needed, where the one-time optimization cost is amortized.

MLIP Simulations of Bulk Water and Ice. The MLIP-*rev*PBE-D3^{OPT} model was then used to carry out liquid water and ice simulations under different conditions, performing both classical and quantum MD simulations. A system containing 510 water molecules was set up to reproduce the liquid-phase properties, while different systems were generated to represent the ice polymorph structures (see Section S4 for further details). As depicted in Figure 3, the computed O−O, O−H, and H−H RDFs of liquid water at normal conditions were all very close to the experimental counterparts.[67,68] The obtained RDFs also matched very closely previous first-principle MD simulations based on the standard *rev*PBE-D3 functional,[58,59] thus indicating that local structural properties were not much affected by the present D3 reoptimization. Moreover, the introduction of nuclear quantum effects was more noticeable in the first peak of the O−O RDF, which showed a slightly enhanced height and a better agreement with experiments, in line with the observation reported in ref 59. Figure 3d depicts the H-bond angle distribution, showing again a good agreement between simulation and experiments. In particular, at room temperature classical and PI MD simulations showed a distribution peaked at 15.0 and 13.9°, respectively (exp., 12.1°), while at higher temperature (T = 323 K) both distributions shifted slightly toward larger angles.

What is most notable, *NpT* simulations exploiting the MLIP-*rev*PBE-D3^{OPT} model reported an accurate density−temperature profile of both liquid water and ice. In Figure 4a, the

https://doi.org/10.1021/acs.jcim.5c02079
J. Chem. Inf. Model. 2025, 65, 12437−12447

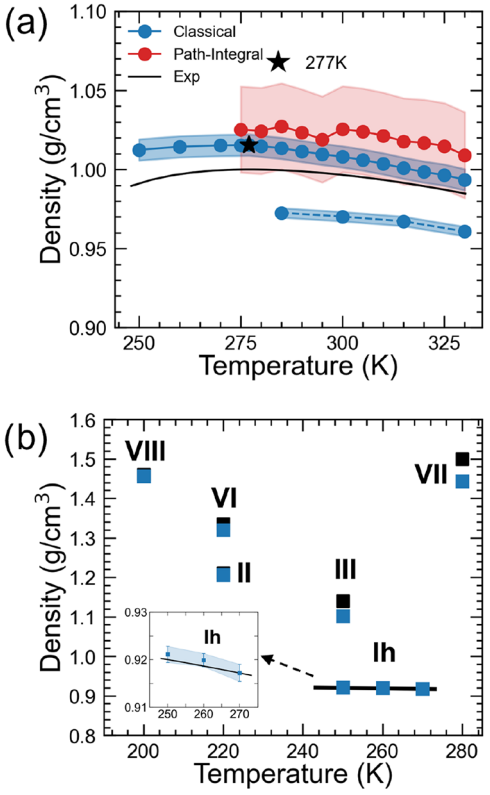


Figure 4. (a) Experimental and simulated density isobars of liquid water evaluated from classical and path-integral MD simulations using the MLIP-*rev*PBE-D3^{OPT} model. A star indicates the density maximum from classical MD simulations. Results from the unmodified *rev*PBE-D3 model are also reported (dashed line). Experimental densities were replotted from NIST.[70] See also Figure S9 of the Supporting Infromation for a comparison with other simulations. (b): Experimental and simulated (classical MD) densities for several ice polymorphic phases evaluated at ambient pressure and stability temperature. Inset, zoomed view of the Ih ice phase density at different temperatures. Experimental densities of ice III, VII, and VIII were replotted from ref 71, ice II and ice VI from ref 72, and ice Ih from NIST.[70]

density prediction of liquid water obtained from classical MD in the temperature range 250−330 K closely followed the experimental density ($\Delta\rho$ < 0.015 g/cm^3 between 275 and 330 K). Besides, the computed density correctly displayed a maximum at 277 K (sim. 1.015 g/cm^3, exp. 1.0 g/cm^3), while at room temperature (300 K) the estimated density was 1.008 g/cm^3 (exp. 0.997 g/cm^3, $\Delta\rho$ = 0.01 g/cm^3). Below 273 K, in the supercooled water regime, we noticed somewhat larger discrepancies from experiments, though they remained rather limited ($\Delta\rho$ < 0.02 g/cm^3). The PI-MD simulations predicted a slightly larger density than classical MD over the same temperature range (275−330 K), although the difference is not significant after taking into consideration the increased statistical error (σ = 0.03 g/cm^3), hence again satisfactorily close to the experimental densities. It is worth noting that the reoptimized *rev*PBE-D3^{OPT} model outperformed previous first-principle and MLIP simulation studies based on various DFT functionals.[57] For comparison, two of the most recent water simulation studies, one based on a neural network trained on standard *rev*PBE0-D3 functional[7] and another one on a many-body potential derived from the density-corrected SCAN model,[13] underestimated the density−temperature dependence (Figure S9), while showing an overall reliable prediction for liquid and solid water. Similarly, the AIMD simulation by Pestana et al., employing the *rev*PBE-D3 approximation, provided a density of about 0.97 g/cm^3 at ambient conditions.[58] Indeed, simulations carried out in this work using a MLIP trained on the unmodified *rev*PBE-D3 functional predict a density profile versus temperature well in line with the ab initio results of ref 58 (Figure 4a), confirming that the original D3 is somewhat underbinding. On the contrary, our reoptimized model showed a nearly perfect match with the density isobars (and the related thermal expansion coefficient, see Figure 5d below) issued from the MB-pol[16,18] many-body

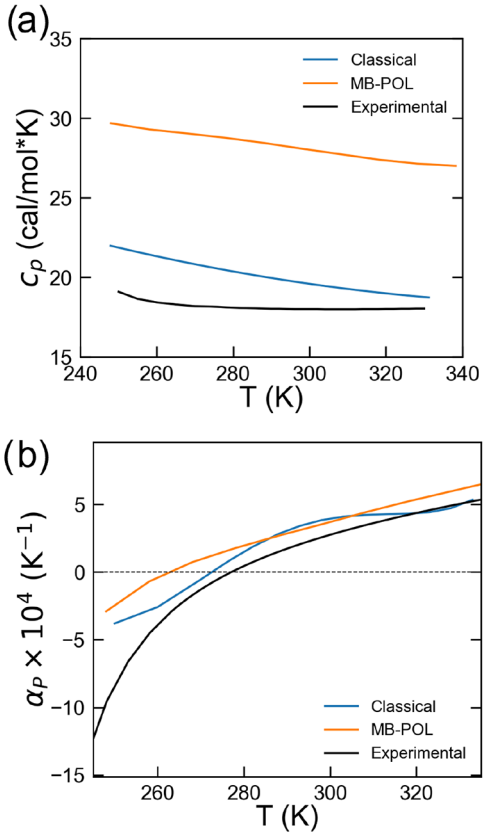


Figure 5. (a): Experimental and simulated isobaric specific heat capacity from classical MLIP-*rev*PBE-D3^{OPT} and MB-Pol extracted and replotted from.[16] Experimental values were replotted from Kell.[73] (b): Experimental and simulated thermal expansion coefficient isobars from classical MLIP-*rev*PBE-D3^{OPT} and MB-Pol extracted and replotted from.[16] Experimental values were replotted from Kell.[73]

potential (Figure S9), which represents a gold standard for modeling physical properties of water. It is worth noting that the slight increase in liquid density issuing from our model with respect to standard *rev*PBE-D3 reflects well the improved water interaction energy of the former, as shown in Figure 4a. Furthermore, in a previous work, Montero De Hijes et al.[57] showed how the density isobar and the melting temperature of water are highly sensitive to the damping function of the D3 term. Their heuristic revPBE0−D3/mix model (i.e., a mix of the zero-damping and Becke-Johnson damping functions) reproduced these properties well, but it was flagged as potentially nontransferable.

Furthermore, we carried out *NpT* simulations on six different ice phases (i.e., Ih, II, III, VI, VII, and VIII) at

https://doi.org/10.1021/acs.jcim.5c02079
J. Chem. Inf. Model. 2025, 65, 12437−12447

ambient pressure and relative stability temperature. As shown in Figure 4b, the predicted densities were again in excellent agreement with available experimental data. In particular, the ice phase Ih, II, VI, and VIII exhibited negligible density deviations ($\leq$0.01 g/cm^3), while III and VII reported larger but limited discrepancies, below 0.06 g/cm^3. Moreover, MLIP simulations also captured minor physical changes, such as the slight density reduction in the temperature range 250−270 K of ice phase Ih (inset of Figure 4b).

In addition to the above structural properties, the MLIP-*rev*PBE-D3^{OPT} model ensured very satisfactory results when tested on thermodynamic quantities, namely the enthalpy of melting (ΔH^{fus}) and vaporization (ΔH^{vap}) obtained according to eqs 2 and 3. The predicted ΔH^{fus} was 1.67 kcal/mol (exp. 1.44 kcal/mol), and ΔH^{vap} was 10.63 kcal/mol (exp. 10.49 kcal/mol), in remarkable agreement with experiments.

The specific heat at constant pressure, c_p, and the isobaric thermal-expansion coefficient, α_P, were also extracted from our *NPT* simulations according to eqs 4 and 5. Figure 5a shows $c_p(T)$ from the MLIP-*rev*PBE-D3^{OPT} and the many-body MB-Pol[16] model alongside experiments. Our model somewhat overestimated c_p at low temperature (<290 K) while approaching experiments at room temperature and beyond. This excess heat capacity at low T likely reflected anharmonic motions that are too facile in the absence of nuclear quantum effects. Noteworthy, our model outperformed MB-Pol in the same investigated temperature range (Figure 5a). Moreover, Figure 5b reports the isobaric thermal-expansion coefficient α_P, plotted as $\alpha_P \times 10^4$ K^{-1} for convenience. The MLIP-*rev*PBE-D3^{OPT} reproduced well the anomalous negative thermal expansion below the density maximum, while crossing zero at 277 K. These results confirmed the accurate description of the water density in a rather extended temperature range, as compared to experiments.

Finally, we performed *NVE* and TRPMD simulations from which we evaluated dynamical properties such as the self-diffusion coefficient and the vibrational density of states (vDOS) of liquid water. In Figure 6a, the self-diffusion coefficient, as obtained according to eq 6 and corrected for infinite size according to eq 7 of water structures using both classical and path-integral MD, is depicted as a function of temperature. Results show how the inclusion of quantum effects significantly improved the description of water mobility. The diffusion coefficient was overestimated at the classical level (at 300 K, 3.05·10^{-9} m^2 s^{-1}) but well-reproduced at the path-integral MD level (2.27·10^{-9} m^2 s^{-1}, exp: 2.41·10^{-9} m^2 s^{-1}). In Figure 6b, we compare the vDOS obtained according to eqs 8 and 9 with the experimental IR power spectrum of water. The vDOS was evaluated according to the procedure described in.[50] In classical MD, the O−H stretching frequency was slightly blue-shifted relative to experimental data, whereas PI-MD resulted in a red-shift, showing a discrepancy of ≈1.5% on peak position. These trends were consistent with previous AIMD studies employing the *rev*PBE-D3 functional.[58,59] In particular, quantum nuclear effects red-shifted the O−H stretching peak position by about 250 cm^{-1}, similar to the AIMD results obtained in ref 59. Concerning the observed deviation from experiments, this is likely due to a still excessive "softness" of the O−H bond in *rev*PBE-D3^{OPT} with respect to the exact or hybrid-functional.

Mg^{2+} in Aqueous Solution. The MLIP-*rev*PBE-D^{OPT} model was further tested to perform MD simulations of Mg^{2+} in water. In Figure 7a, the Mg−O RDF obtained from

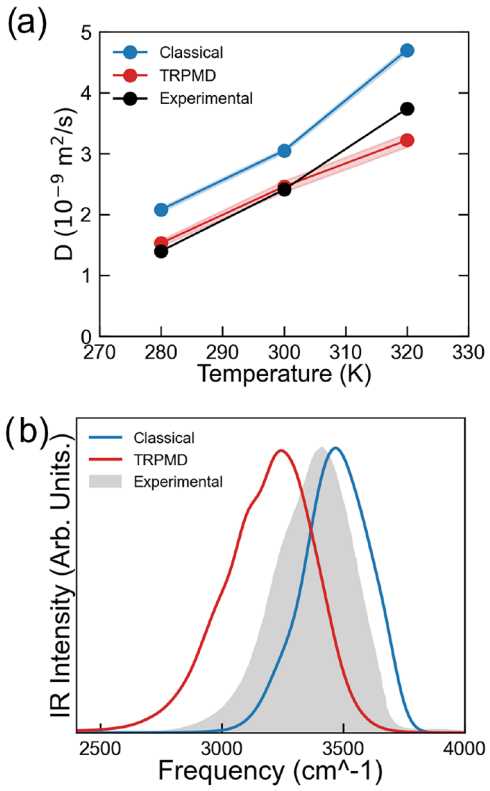


Figure 6. (a): Experimental and simulated Diffusion Coefficient isobars for water from classical and PI-MD simulation using MLIP-*rev*PBE-D3^{OPT}. Experimental data were replotted from Prof. Dietrich's Web site https://dtrx.de/od/diff/. (b): Experimental IR Spectrum and simulated vDOS rescaled to experimental density for comparison from classical and PI MD simulation employing MLIP-*rev*PBE-D3^{OPT}. Experimental values were extracted and replotted from Bertie et al.[74]

our MLIP simulation is depicted along with results issuing from other popular classical force fields, namely the MicroMg[75] and the 12−6−4 from Li and Merz,[76] as well as a MLIP derived from standard dispersion corrections (i.e., MLIP-*rev*PBE-D). The MLIP-*rev*PBE-D^{OPT} predicted well the position of the first RDF peak at 2.07 Å, matching experiments within the statistical error (2.09 ± 0.04 Å).[77] Note that our MLIP, similarly to other recent ML models,[78,79] showed a broader first peak in comparison to classical force fields, thus indicating a more flexible first solvation shell. The beneficial effect of the optimization first became apparent in the description of the second solvation shell, since MLIP-*rev*PBE-D^{OPT} reported the second Mg−O peak appreciably closer to experiments than both the unoptimized DFT-D4 model and the classical force fields. The optimization also improved the predicted density with respect to the other tested models, as reported in Table 1. Upon reoptimization, the obtained MLIP-*rev*PBE-D^{OPT} displayed a density of 0.99 ± 0.01 g/cm^3, in close agreement with experiments (1.005 g/cm^3 from Al Ghafri et al.[80]), while the same property was somewhat underestimated by the default parametrization. In a previous work, Kostal et al.[81] observed that, in the case of monovalent cations, the use of D3 can worsen the description of structural properties due to overbinding of the pristine *rev*PBE functional. Here, we showed how, in the case of Mg^{2+}, the use of a reliable DFT approach in combination with Dx reoptimization allowed us to recover the correct structural, thermodynamic, and kinetic behavior.

https://doi.org/10.1021/acs.jcim.5c02079
J. Chem. Inf. Model. 2025, 65, 12437−12447

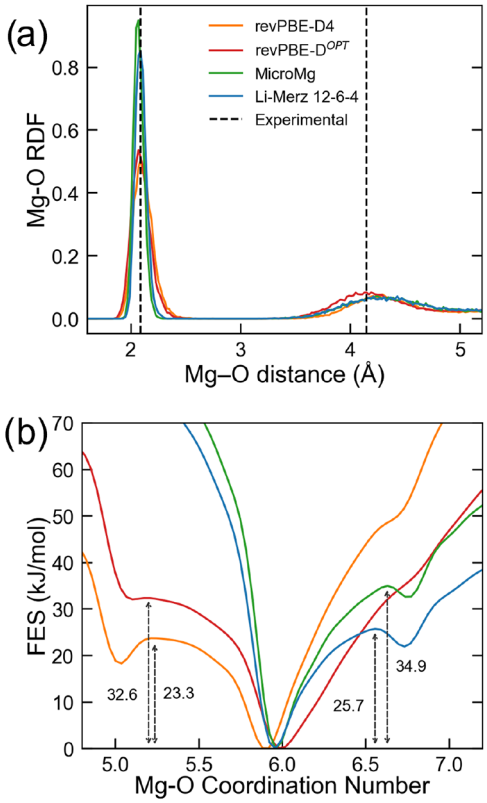


Figure 7. (a): Experimental and simulated Mg−O RDF as obtained from MLIP-*rev*PBE-D4, D^{OPT} and from MicroMg[75] and Li-Merz[76] 12−6−4 Classical Force Fields. Experimental reference (black dashed line 2.09 ± 0.4), was replotted from Marcus.[77] (b): Free Energy Surface of Mg^{2+}-Water coordination as obtained from MLIP-*rev*PBE-D4, D^{OPT} and from MicroMg and Li-Merz 12−6−4 Classical Force Fields.

Table 1. Experimental and Simulated Exchange Kinetics in the First Solvation Shell and Aqueous Solution Density

model	ρ [g cm^{-3}][a]	$\Delta A^{\ddagger}$ [kJ mol^{-1}]	k_X [s^{-1}]
MicroMg (ref)	0.99	34.9	8.0×10^5
*rev*PBE−D4	0.95	23.3	8.4×10^7
*rev*PBE−D^{OPT}	0.99	32.6	2.0×10^6
Exp	1.005[b]		6.7×10^5[c]

[a]Density of a 0.109 M aqueous solution of $MgCl_2$. [b]Experimental density was taken from Al Ghafri et al.[80] [c]Experimental rate was taken from Bleuzen et al.[82]

Switching to dynamic properties, we focused on the free-energy barrier for water exchange in the first solvation shell. To this purpose, we evaluated the free energy change as a function of the ion−water coordination number (as defined in eq 10) as obtained from purposely performed meta-MD simulations according to the computational protocol described in refs 51,52 (see Free Energy of Ion Coordination section for further details), and we report the results of these simulations in Figure 7b, where we compare DFT and force-field predictions for this quantity. As apparent from Figure 7b, all meta-MD simulations predicted the 6-fold water coordination as the most stable for Mg^{2+}. However, importantly, only the DFT-based potentials correctly indicated a dissociative water exchange mechanism with a 5-fold water configuration as the second most favorable state (a shallow local minimum in the free-energy). A dissociative water exchange mechanism is indeed consistent with the findings reported in the NMR experimental study by Bleuzen et al.[82] In contrast, both classical force-field simulations predicted an erroneous associative mechanism through a hepta-coordinated state, possibly due to the lack of many-body interactions. It is also worthwhile to observe that, only using a MLIP parametrized against a DFT functional in which the dispersion terms have been reoptimized, i.e., only using MLIP-*rev*PBE-D^{OPT}, the simulations were able to reproduce fairly well also the predicted energy barrier for water exchange and the experimental rate constant for the process (Table 1). Here, we assume as the reference energy barrier the one obtained from the MicroMg[75] parametrization (34.9 kJ/mol), a model purposely developed to match the experimental water exchange rate. The rate constants were estimated by considering an exponential scaling factor with respect to the MicroMg model ($k_{X/kMicroMg} = \exp(\Delta\Delta A^{\ddagger}/kT)$). Noteworthy, the default *rev*PBE−D4 model underestimated the barrier by about 12 kJ/mol, hence overshooting the predicted rate constant by 2 orders of magnitude (Figure 7b and Table 1). We thus conclude that a physically sound re−parametrization of the D4 term is necessary and sufficient to restore a realistic description of the first−shell solvation structure around the magnesium ion and its water exchange kinetics, simultaneously reproducing the experimentally established dissociative mechanism. Note that our protocol is not only computationally effective but also easily transferable to other mixture and/or aqueous solutions in comparison with state-of-the-art gold standard approaches, e.g., MB-pol.

CONCLUSIONS

In summary, the key features of our proposed in silico strategy are (1) fine-tuning a computationally efficient DFT-D (GGA) functional enhanced by dispersion corrections (MAE/mol <0.1 kcal/mol) toward high-fidelity reference data; (2) building a proper training database to generate an effective MLIP able to reproduce interaction energies and atomic forces of the optimized DFT-D model with great accuracy (RMSE of 0.1 meV/atom and 8 meV/Å); (3) validating the predictive capability of the trained MLIP over a wide ensemble of physicochemical properties of diverse systems, from nano-clusters to bulk liquid and solid phases.

In particular, our results demonstrate, first, that system-specific optimization of dispersion parameters can be successfully extended to treat neat water and aqueous solutions. By then incorporating the resulting high-quality energy and gradient evaluations into a data-driven neural network potential, the MLIP-*rev*PBE-$D3^{OPT}$ model, accurate MD simulations of complex condensed-phase systems can be carried out at an affordable computational cost. Notably, the so obtained MLIP-*rev*PBE-$D3^{OPT}$ model is able to reproduce the energetics of the dispersion-corrected DFT approach for a diverse ensemble of systems, from small clusters to bulk liquid and solid phases. It also successfully predicted a wide range of experimental equilibrium and dynamic quantities such as radial distribution functions, self-diffusion coefficients, vibrational density of states, enthalpy of melting and enthalpy of vaporization, and, impressively, the mass density of water in several phases, ranging from the liquid phase (matching the observed maximum at 277 K) to various forms of ice as evaluated at different temperatures. The protocol extends naturally to interfacial systems (e.g., solid−liquid, liquid−vapor) by harvesting snapshots from coexistence or slab

https://doi.org/10.1021/acs.jcim.5c02079
J. Chem. Inf. Model. 2025, 65, 12437−12447

simulations at the target *T-P* conditions and inserting them (via the same active-learning cycle used for PIMD and Mg^{2+} metadynamics) into the fitting database.

Finally, in view of applications to, e.g., saline aqueous solutions that do require an independent optimization of solvent−solvent and solute−solvent interactions, we have tested the case of Mg^{2+} in water. Transfering the accurate description of water−water interactions into a purposedly derived MLIP for the accurate description of the metal ion−water interaction, as provided by the present approach, led to a reliable representation of both the Mg^{2+} microsolvation structure and its water exchange dynamics, which were otherwise misrepresented by classical force fields and standard (unoptimized) DFT-D models. We believe that these findings have proved the fidelity, transferability and modularity of the obtained MLIP models, thus supporting the use of the present approach for developing realistic MLIPs for a large variety of molecular systems.

■ ASSOCIATED CONTENT

Data Availability Statement

All input files, trained models, and analysis scripts needed to reproduce the simulations and results are openly available at https://github.com/SNS-Brancato-Lab/Water_MG_MLIP.

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jcim.5c02079.

Additional details on the DFT setup, MLIP validation, performance and database sampling, and the protocols used for MD and metadynamics (Meta-MD) simulations (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Alessando Fortunelli − *Consiglio Nazionale delle Ricerche, CNR-ICCOM, 56124 Pisa, Italy;* orcid.org/0000-0001-5337-4450; Email: alessandro.fortunelli@cnr.it

Giuseppe Brancato − *Scuola Normale Superiore and CSGI, I-56127 Pisa, Italy; Istituto Nazionale di Fisica Nucleare (INFN), Sezione di Pisa, I-56127 Pisa, Italy;* orcid.org/0000-0001-8059-2517; Email: giuseppe.brancato@sns.it

Authors

Alfonso Ferretti − *Scuola Normale Superiore, I-56127 Pisa, Italy; Istituto Nazionale di Fisica Nucleare (INFN), Sezione di Pisa, I-56127 Pisa, Italy;* orcid.org/0000-0003-1567-0616

Giacomo Melani − *Consiglio Nazionale delle Ricerche, CNR-ICCOM, 56124 Pisa, Italy;* orcid.org/0000-0003-1632-8935

Luca Benedetti − *Scuola Normale Superiore, I-56127 Pisa, Italy; Istituto Nazionale di Fisica Nucleare (INFN), Sezione di Pisa, I-56127 Pisa, Italy*

Robert A. Sorodoc − *NEST Istituto Nanoscienze-CNR and Scuola Normale Superiore, 56127 Pisa, Italy;* orcid.org/0009-0008-8851-8459

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jcim.5c02079

Author Contributions

A.Fer. developed the DFT-D model, performed the ab initio, MD and ML simulations, and performed the corresponding analyses; G.M. contributed to the method development and to the VDOS analysis; L.B. contributed to the classical MD simulations and R.A.S. to the DFT-D optimization of Mg in water; AFer., G.M., AFor. and G.B. developed the ML model., A.Fer., G.M., A.For. and G.B. contributed to the rationalization of the results and to the computational design; G.B. supervised the work; A.Fer., G.M., L.B., A.For., and G.B. jointly wrote the manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

G.B. acknowledges financial support under the National Recovery and Resilience Plan (NRRP), Mission 4, Component 2, Investment 1.1, Call for tender No. 1409 published on 14.9.2022 by the Italian Ministry of University and Research (MUR), funded by the European Union—NextGenerationEU—Project Title Efficient Sequestration of Metal Ions from Aqueous Systems for Green and Sustainable Applications—AquaGreen—CUP E53D23015550001—Grant Assignment Decree No. 1409 adopted on 14/09/2022 by the Italian Ministry of Ministry of University and Research (MUR). G.M. and A.F. acknowledge financial support from ICSC/EU (grant number CN00000013), and from the Italian Ministry of Environment and Energy Security/POR H2 AdP MMES/ENEA project. We gratefully acknowledge the computational resources and technical support of the Center for High Performance Computing (CHPC) at SNS and the CINECA. Networking within the COST Action CA21101 "Confined molecular systems: from a new generation of materials to the stars" (COSY) supported by COST (European Cooperation in Science and Technology) is also acknowledged.

■ REFERENCES

(1) Deringer, V. L.; Caro, M. A.; Csányi, G. Machine Learning Interatomic Potentials as Emerging Tools for Materials Science. *Adv. Mater.* **2019**, *31*, No. 1902765.

(2) Behler, J. Four Generations of High-Dimensional Neural Network Potentials. *Chem. Rev.* **2021**, *121* (16), 10037−10072.

(3) Omranpour, A.; De Hijes, P. M.; Behler, J.; Dellago, C. Perspective: Atomistic simulations of water and aqueous systems with machine learning potentials. *J. Chem. Phys.* **2024**, *160*, No. 170901.

(4) Anstine, D. M.; Isayev, O. Machine Learning Interatomic Potentials and Long-Range Physics. *J. Phys. Chem. A* **2023**, *127*, 2417−2431.

(5) Kulichenko, M.; Nebgen, B.; Lubbers, N.; Smith, J. S.; Barros, K.; Allen, A. E. A.; Habib, A.; Shinkle, E.; Fedik, N.; Li, Y. W.; Messerly, R. A.; Tretiak, S. Data Generation for Machine Learning Interatomic Potentials and Beyond. *Chem. Rev.* **2024**, *124*, 13681−13714.

(6) Morawietz, T.; Behler, J. A Density-Functional Theory-Based Neural Network Potential for Water Clusters Including van der Waals Corrections. *J. Phys. Chem. A* **2013**, *117*, 7356−7366.

(7) Cheng, B.; Engel, E. A.; Behler, J.; Dellago, C.; Ceriotti, M. Ab initio thermodynamics of liquid and solid water. *Proc. Natl. Acad. Sci. U.S.A.* **2019**, *116*, 1110−1115.

(8) Zhang, C.; Tang, F.; Chen, M.; Xu, J.; Zhang, L.; Qiu, D. Y.; Perdew, J. P.; Klein, M. L.; Wu, X. Modeling Liquid Water by Climbing up Jacob's Ladder in Density Functional Theory Facilitated by Using Deep Neural Network Potentials. *J. Phys. Chem. B* **2021**, *125*, 11444−11456.

(9) Wang, J.; Román-Pérez, G.; Soler, J. M.; Artacho, E.; Fernández-Serra, M.-V. Density, structure, and dynamics of water: The effect of van der Waals interactions. *J. Chem. Phys.* **2011**, *134*, No. 024516.

(10) Brandenburg, J. G.; Maas, T.; Grimme, S. Benchmarking DFT and semiempirical methods on structures and lattice energies for ten ice polymorphs. *J. Chem. Phys.* **2015**, *142*, No. 124104.
(11) Della Pia, F.; Zen, A.; Alfè, D.; Michaelides, A. DMC-ICE13: Ambient and high pressure polymorphs of ice from diffusion Monte Carlo and density functional theory. *J. Chem. Phys.* **2022**, *157*, No. 134701.
(12) Alfè, D.; Bartók, A. P.; Csányi, G.; Gillan, M. J. Communication: Energy benchmarking with quantum Monte Carlo for water nano-droplets and bulk liquid water. *J. Chem. Phys.* **2013**, *138*, No. 221102.
(13) Dasgupta, S.; Lambros, E.; Perdew, J. P.; Paesani, F. Elevating density functional theory to chemical accuracy for water simulations through a density-corrected many-body formalism. *Nat. Commun.* **2021**, *12*, No. 6359.
(14) Song, S.; Vuckovic, S.; Kim, Y.; Yu, H.; Sim, E.; Burke, K. Extending density functional theory with near chemical accuracy beyond pure water. *Nat. Commun.* **2023**, *14*, No. 799.
(15) Villard, J.; Bircher, M. P.; Rothlisberger, U. Structure and dynamics of liquid water from ab initio simulations: adding Minnesota density functionals to Jacob's ladder. *Chem. Sci.* **2024**, *15*, 4434−4451.
(16) Reddy, S. K.; Straight, S. C.; Bajaj, P.; Huy Pham, C.; Riera, M.; Moberg, D. R.; Morales, M. A.; Knight, C.; Götz, A. W.; Paesani, F. On the accuracy of the MB-pol many-body potential for water: Interaction energies, vibrational frequencies, and classical thermodynamic and dynamical properties from clusters to liquid water and ice. *J. Chem. Phys.* **2016**, *145* (19), No. 194504.
(17) Dasgupta, S.; Shahi, C.; Bhetwal, P.; Perdew, J. P.; Paesani, F. How Good Is the Density-Corrected SCAN Functional for Neutral and Ionic Aqueous Systems, and What Is So Right about the Hartree-Fock Density? *J. Chem. Theory Comput.* **2022**, *18*, 4745−4761.
(18) Zhu, X.; Riera, M.; Bull-Vulpe, E. F.; Paesani, F. MB-pol(2023): Sub-chemical Accuracy for Water Simulations from the Gas to the Liquid Phase. *J. Chem. Theory Comput.* **2023**, *19*, 3551−3566.
(19) Calio, P. B.; Hocky, G. M.; Voth, G. A. Minimal Experimental Bias on the Hydrogen Bond Greatly Improves Ab Initio Molecular Dynamics Simulations of Water. *J. Chem. Theory Comput.* **2020**, *16*, 5675−5684.
(20) Daru, J.; Forbert, H.; Behler, J.; Marx, D. Coupled Cluster Molecular Dynamics of Condensed Phase Systems Enabled by Machine Learning Potentials: Liquid Water Benchmark. *Phys. Rev. Lett.* **2022**, *129*, No. 226001.
(21) Grimme, S. Accurate description of van der Waals complexes by density functional theory including empirical corrections. *J. Comput. Chem.* **2004**, *25*, 1463−1473.
(22) Gillan, M. J.; Alfè, D.; Michaelides, A. Perspective: How good is DFT for water? *J. Chem. Phys.* **2016**, *144*, No. 130901.
(23) Grimme, S.; Hansen, A.; Brandenburg, J. G.; Bannwarth, C. Dispersion-Corrected Mean-Field Electronic Structure Methods. *Chem. Rev.* **2016**, *116*, 5105−5154.
(24) Barbosa, N.; Pagliai, M.; Sinha, S.; Barone, V.; Alfè, D.; Brancato, G. Enhancing the Accuracy of Ab Initio Molecular Dynamics by Fine Tuning of Effective Two-Body Interactions: Acetonitrile as a Test Case. *J. Phys. Chem. A* **2021**, *125*, 10475−10484.
(25) Ferretti, A.; Canal, L.; Sorodoc, R. A.; Sinha, S.; Brancato, G. Fine Tuning the Intermolecular Interactions of Water Clusters Using the Dispersion-Corrected Density Functional Theory. *Molecules* **2023**, *28* (9), No. 3834.
(26) Brancato, G.; Rega, N.; Barone, V.; et al. Accurate Density Functional Calculations of Near-Edge X-Ray and Optical Absorption Spectra of Liquid Water Using Nonperiodic Boundary Conditions: The Role of Self-Interaction and Long-Range Effects. *Phys. Rev. Lett.* **2008**, *100*, No. 107401, DOI: 10.1103/PhysRevLett.100.107401.
(27) Fransson, T.; Pettersson, L. G. M. TDDFT and the x-ray absorption spectrum of liquid water: Finding the "best" functional. *J. Chem. Phys.* **2024**, *160*, No. 234105.
(28) Nguyen, N. L.; Ma, H.; Govoni, M.; et al. Finite-Field Approach to Solving the Bethe-Salpeter Equation. *Phys. Rev. Lett.* **2019**, *122*, No. 237402, DOI: 10.1103/PhysRevLett.122.237402.
(29) Ambrosetti, A.; Alfè, D.; DiStasio, R. A. J.; Tkatchenko, A. Hard Numbers for Large Molecules: Toward Exact Energetics for Supramolecular Systems. *J. Phys. Chem. Lett.* **2014**, *5*, 849−855.
(30) Batatia, I.; Kovács, D. P.; Simm, G. N. C.; Ortner, C.; Csányi, G. NMACE: Higher Order Equivariant Message Passing Neural Networks for Fast and Accurate Force Fields. 2023, arXiv:2206.07697. arXiv.org e-Printarchive. http://arxiv.org/abs/2206.07697.
(31) 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.
(32) Dal Corso, A. Pseudopotentials periodic table: From H to Pu. *Comput. Mater. Sci.* **2014**, *95*, 337−350.
(33) Giannozzi, P.; Baroni, S.; Bonini, N.; et al. QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials. *J. Phys.: Condens. Matter* **2009**, *21*, No. 395502.
(34) Dunning, T. H., Jr. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007−1023.
(35) Neese, F. The ORCA program system. *WIREs Comput. Mol. Sci.* **2012**, *2*, 73−78.
(36) Roongcharoen, T.; Giorgio, C.; Sementa, L.; Melani, G.; Fortunelli, A. Machine-Learning-Accelerated DFT Conformal Sampling of Catalytic Processes. *J. Chem. Theory Comput.* **2024**, *20*, 9580−9591.
(37) Kuryla, D.; Csányi, G.; van Duin, A. C. T.; Michaelides, A. Efficient exploration of reaction pathways using reaction databases and active learning. *J. Chem. Phys.* **2025**, *162*, No. 114122.
(38) Abascal, J. L. F.; Vega, C. A general purpose model for the condensed phases of water: TIP4P/2005. *J. Chem. Phys.* **2005**, *123*, No. 234505.
(39) Deringer, V. L.; Bartók, A. P.; Bernstein, N.; Wilkins, D. M.; Ceriotti, M.; Csányi, G. Gaussian Process Regression for Materials and Molecules. *Chem. Rev.* **2021**, *121*, 10073−10141.
(40) Laio, A.; Parrinello, M. Escaping free-energy minima. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 12562−12566.
(41) Barducci, A.; Bussi, G.; Parrinello, M. Well-Tempered Metadynamics: A Smoothly Converging and Tunable Free-Energy Method. *Phys. Rev. Lett.* **2008**, *100*, No. 020603.
(42) Naserifar, S.; Goddard, W. A. Liquid water is a dynamic polydisperse branched polymer. *Proc. Natl. Acad. Sci. U.S.A.* **2019**, *116*, 1998−2003.
(43) 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.; Shan, R.; Stevens, M. J.; Tranchida, J.; Trott, C.; Plimpton, S. J. LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. *Comput. Phys. Commun.* **2022**, *271*, No. 108171.
(44) Asthagiri, D. N.; Beck, T. L. MD Simulation of Water Using a Rigid Body Description Requires a Small Time Step to Ensure Equipartition. *J. Chem. Theory Comput.* **2024**, *20*, 368−374.
(45) Uhl, F.; Marx, D.; Ceriotti, M. Accelerated path integral methods for atomistic simulations at ultra-low temperatures. *J. Chem. Phys.* **2016**, *145*, No. 054101.
(46) Rossi, M.; Ceriotti, M.; Manolopoulos, D. E. How to remove the spurious resonances from ring polymer molecular dynamics. *J. Chem. Phys.* **2014**, *140*, No. 234116.
(47) Kapil, V.; Rossi, M.; Marsalek, O.; et al. i-PI 2.0: A universal force engine for advanced molecular simulations. *Comput. Phys. Commun.* **2019**, *236*, 214−223.
(48) Larsen, A. H.; Mortensen, J. J.; Blomqvist, J.; et al. The atomic simulation environment—a Python library for working with atoms. *J. Phys.: Condens. Matter* **2017**, *29*, No. 273002.

(49) Yeh, I.-C.; Hummer, G. System-Size Dependence of Diffusion Coefficients and Viscosities from Molecular Dynamics Simulations Periodic Boundary Conditions. *J. Phys. Chem. B* **2004**, *108*, 15873−15879.
(50) Melani, G.; Nagata, Y.; Wirth, J.; Saalfrank, P. Vibrational spectroscopy of hydroxylated *alpha*-Al2O3(0001) surfaces with and without water: An ab initio molecular dynamics study. *J. Chem. Phys.* **2018**, *149*, No. 014707.
(51) Brancato, G.; Barone, V. Free Energy Landscapes of Ion Coordination in Aqueous Solution. *J. Phys. Chem. B* **2011**, *115*, 12875−12878.
(52) Sagresti, L.; Peri, L.; Ceccarelli, G.; Brancato, G. Stochastic Model of Solvent Exchange in the First Coordination Shell of Aqua Ions. *J. Chem. Theory Comput.* **2022**, *18*, 3164−3173.
(53) Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. PLUMED 2: New feathers for an old bird. *Comput. Phys. Commun.* **2014**, *185*, 604−613.
(54) Bonomi, M.; Bussi, G.; Camilloni, C.; et al. Promoting transparency and reproducibility in enhanced molecular simulations. *Nat. Methods* **2019**, *16*, 670−673.
(55) Zhang, Y.; Yang, W. Comment on "Generalized Gradient Approximation Made Simple". *Phys. Rev. Lett.* **1998**, *80*, 890.
(56) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, No. 154104.
(57) Montero de Hijes, P.; Dellago, C.; Jinnouchi, R.; Kresse, G. Density isobar of water and melting temperature of ice: Assessing common density functionals. *J. Chem. Phys.* **2024**, *161*, No. 131102.
(58) Pestana, L. R.; Mardirossian, N.; Head-Gordon, M.; Head-Gordon, T. Ab initio molecular dynamics simulations of liquid water using high quality meta-GGA functionals. *Chem. Sci.* **2017**, *8*, 3554−3565.
(59) Marsalek, O.; Markland, T. E. Quantum Dynamics and Spectroscopy of Ab Initio Liquid Water: The Interplay of Nuclear and Electronic Quantum Effects. *J. Phys. Chem. Lett.* **2017**, *8*, 1545−1551.
(60) Lausch, K. N.; El Haouari, R.; Trzewik, D.; Behler, J. Impact of the damping function in dispersion-corrected density functional theory on the properties of liquid water. *J. Chem. Phys.* **2025**, *163*, No. 034101.
(61) Palos, E.; Lambros, E.; Swee, S.; Hu, J.; Dasgupta, S.; Paesani, F. Assessing the Interplay between Functional-Driven and Density-Driven Errors in DFT Models of Water. *J. Chem. Theory Comput.* **2022**, *18*, 3410−3426.
(62) Caldeweyher, E.; Ehlert, S.; Hansen, A.; Neugebauer, H.; Spicher, S.; Bannwarth, C.; Grimme, S. A generally applicable atomic-charge dependent London dispersion correction. *J. Chem. Phys.* **2019**, *150*, No. 154122.
(63) Adamo, C.; Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **1999**, *110*, 6158−6170.
(64) Alfè, D.; Bartók, A. P.; Csányi, G.; Gillan, M. J. Analyzing the errors of DFT approximations for compressed water systems. *J. Chem. Phys.* **2014**, *141*, No. 014104.
(65) Bryantsev, V. S.; Diallo, M. S.; van Duin, A. C. T.; Goddard, W. A. I. Evaluation of B3LYP, X3LYP, and M06-Class Density Functionals for Predicting the Binding Energies of Neutral, Protonated, and Deprotonated Water Clusters. *J. Chem. Theory Comput.* **2009**, *5*, 1016−1026.
(66) Palos, E.; Caruso, A.; Paesani, F. Consistent density functional theory-based description of ion hydration through density-corrected many-body representations. *J. Chem. Phys.* **2023**, *159*, No. 181101.
(67) Skinner, L. B.; Galib, M.; Fulton, J. L.; Mundy, C. J.; Parise, J. B.; Pham, V.-T.; Schenter, G. K.; Benmore, C. J. The structure of liquid water up to 360 MPa from x-ray diffraction measurements using a high Q-range and from molecular simulation. *J. Chem. Phys.* **2016**, *144*, No. 134504.
(68) Soper, A. K. The radial distribution functions of water and ice from 220 to 673 K and at pressures up to 400 MPa. *Chem. Phys.* **2000**, *258*, 121−137.
(69) Modig, K.; Pfrommer, B. G.; Halle, B. Temperature-Dependent Hydrogen-Bond Geometry in Liquid Water. *Phys. Rev. Lett.* **2003**, *90*, No. 075502.
(70) Harvey, A. H. Properties of Ice and Supercooled Water. In *NIST*; Last Modified: 2020-09-28T11:09−04:00 Publisher: Allan H. Harvey, 2019.
(71) Chaplin, M. The Ice Phases of Water. https://water.lsbu.ac.uk/water/ice_phases.html#ba.
(72) Loerting, T.; Bauer, M.; Kohl, I.; Watschinger, K.; Winkel, K.; Mayer, E. Cryoflotation: Densities of Amorphous and Crystalline Ices. *J. Phys. Chem. B* **2011**, *115*, 14167−14175.
(73) Kell, G. S. Density, thermal expansivity, and compressibility of liquid water from 0.deg. to 150.deg.. Correlations and tables for atmospheric pressure and saturation reviewed and expressed on 1968 temperature scale. *J. Chem. Eng. Data* **1975**, *20*, 97−105.
(74) Bertie, J. E.; Lan, Z. Infrared Intensities of Liquids XX: The Intensity of the OH Stretching Band of Liquid Water Revisited, and the Best Current Values of the Optical Constants of H2O(l) at 25°C between 15,000 and 1 cm-1. *Appl. Spectrosc.* **1996**, *50*, 1047−1057.
(75) Grotz, K. K.; Cruz-León, S.; Schwierz, N. Optimized Magnesium Force Field Parameters for Biomolecular Simulations Accurate Solvation, Ion-Binding, and Water-Exchange Properties. *J. Chem. Theory Comput.* **2021**, *17*, 2530−2540.
(76) Li, P.; Merz, K. M. J. Taking into Account the Ion-Induced Dipole Interaction in the Nonbonded Model of Ions. *J. Chem. Theory Comput.* **2014**, *10*, 289−297.
(77) Marcus, Y. Ionic radii in aqueous solutions. *Chem. Rev.* **1988**, *88*, 1475−1498.
(78) Juraskova, V.; Tusha, G.; Zhang, H.; Schäfer, V.; Duarte, L. Modelling ligand exchange in metal complexes with machine learning potentials. *Faraday Discuss.* **2025**, *256*, 156−176.
(79) Li, W.; Sun, Z. Second Hydration Shell of Mg2+: Competition between Ion−Water Interaction and Hydrogen Bonding Interaction. *J. Phys. Chem. Lett.* **2025**, *16*, 503−509.
(80) Al Ghafri, S.; Maitland, G. C.; Trusler, J. P. M. Densities of Aqueous MgCl2(aq), CaCl2(aq), KI(aq), NaCl(aq), KCl(aq), AlCl3(aq), and (0.964 NaCl + 0.136 KCl)(aq) at Temperatures Between (283 and 472) K, Pressures up to 68.5 MPa, and Molalities up to 6 mol·kg−1. *J. Chem. Eng. Data* **2012**, *57*, 1288−1304.
(81) Kostal, V.; Mason, P. E.; Martinez-Seara, H.; Jungwirth, P. Common Cations Are Not Polarizable: Effects of Dispersion Correction on Hydration Structures from Ab Initio Molecular Dynamics. *J. Phys. Chem. Lett.* **2023**, *14*, 4403−4408.
(82) Bleuzen, A.; Pittet, P.-A.; Helm, L.; Merbach, A. E. Water exchange on magnesium(II) in aqueous solution: a variable temperature and pressure 17O NMR study. *Magn. Reson. Chem.* **1997**, *35*, 765−773.

https://doi.org/10.1021/acs.jcim.5c02079
J. Chem. Inf. Model. 2025, 65, 12437−12447