

A method for accurately predicting solvation structure.

Timothy T. Duignan,^{*,†} Christopher J Mundy,^{‡,¶} Gregory K. Schenter,[‡] and X.
S. Zhao^{*,†}

*†School of Chemical Engineering, The University of Queensland, St Lucia, Brisbane 4072,
Australia*

*‡Physical Science Division, Pacific Northwest National Laboratory, P.O. Box 999, Richland,
Washington 99352, USA*

*¶Affiliate Professor, Department of Chemical Engineering, University of Washington,
Seattle, Washington, USA*

E-mail: t.duignan@uq.edu.au; george.zhao@uq.edu.au

Abstract

Accurately predicting the molecular structure of solutions is a fundamental scientific challenge. Using quantum mechanical density functional theory (DFT) to make these predictions is hindered by significant variation depending on which DFT functional is used. Here, we present a simple metric that can determine the reliability of a DFT functional for predicting solvation structure. We then show that including a simple interaction term to correct this metric leads to quantitative agreement with experimental measurements of liquid structure. We demonstrate the utility of this method by using it to accurately describe the hydration structure around the Na^+ and K^+ ions as well as the structural properties of pure water with a computationally cheap functional.

Introduction

The inability to make accurate predictions of the molecular structure of liquids solutions limits our understanding of many important phenomena. For example, understanding the transport of Na^+ and K^+ through the membrane of a biological cell or through the electrolyte solution of a battery depends on having a reliable and accurate picture of how solvent is structured around these ions.^{1,2}

The most direct approach to making these predictions is through the use of density functional theory with molecular dynamics (DFT-MD) simulation. However, this approach is currently limited by the fact that the predictions can often vary significantly depending on which DFT functional is used.³ For example, we have shown that revised Perdew, Burke, and Ernzerhof functional with Grimme dispersion correction (revPBE-D3)⁴⁻⁶ can reproduce water structure reasonably well but fails when used to predict the structure around cations such as Na^+ and K^+ .⁷⁻⁹ In contrast, the strongly constrained and appropriately normed (SCAN) functional,^{10,11} which is a higher level meta generalised gradient approximation (GGA) functional, works well for Na^+ and K^+ structure but is somewhat less accurate than

revPBE-D3 for pure water structure.⁹ Using even higher level DFT functionals that incorporate explicit Hartree-Fock or electron correlation can potentially mitigate this problem somewhat but these approaches are so computationally expensive that they can only be applied by a very small group of researchers to a tiny set of problems.^{9,12-17} Even meta-GGA functionals such as SCAN are significantly slower than GGAs, limiting their usefulness to study practically important systems.⁹ This functional dependence is consistent with the fact that these functionals can show significant errors in the 2-body ion-water and water-water interactions computed at higher levels of theory.^{18,19} The good experimental agreement for bulk liquid structural properties can be the result of cancellation of errors. For example, between two and three body interactions²⁰ and errors in the interaction and the neglect of quantum nuclear effects in the case of revPBE-D3.¹⁶ Additionally many DFT functionals, particularly GGAs can have issues with reproducing charged group interactions.²¹ There is a great deal of ongoing research to address the various shortcomings of DFT.²²⁻²⁴ However, it is unclear whether it will be possible to find a functional that is accurate enough reproduce liquid structure while still having low enough computational demands to be practically useful.

A standard metric used to evaluate a force field or functional is the calculation of the dimer binding energies using the minimum energy structure compared with higher level methods.^{19,25,26} This is not ideal for assessing the ability of a functional to reproduce bulk condensed phase structure for several reasons. Firstly, at room temperature in the liquid phase the structures will not stay in a single minimum energy structure.²⁷ Secondly, the balance of contributions such as dispersion and induction can be quite different in vacuum compared with the condensed phase.^{28,29} Thirdly, It is not clear what basis sets to use for the dimer calculations as DFT-MD simulations are often performed with basis sets optimised for the condensed phase.³⁰ Fourthly, many body effects beyond the dimer interaction are known to be important.³¹ Finally, it is the forces that determine the structural properties rather than absolute binding energies. Therefore the ability of a functional to reproduce the energy

of totally dissociating the molecules of a liquid is not relevant if the goal is to reproduce the structural details of the solvent molecules in the condensed phase.^{9,32}

An impressive body of work in recent years has demonstrated that many-body models can correctly reproduce many-body effects at all orders and reproduce structural and other properties of water and ions in water.³³⁻³⁵ These models are fitted to reproduce energies of clusters larger than the simple dimer and at a range of distorted geometries away from the energy minimum. We feel that it is valuable to better understand the relation between this approach and DFT.

Here, we propose a new metric to determine the accuracy of a DFT functional for reproducing liquid structure, which does not suffer from the limitations outlined above. We show that it can accurately predict the ability of a DFT functional to reproduce liquid structure and provides a straightforward method for correcting the error in the liquid structure.

Results and Discussion

Following on from our work in Ref. 9 we use clusters of 8 water molecules surrounding a central molecule in order to fully sample the entire first hydration shell. These clusters are extracted from simulation in order to have a representative sample of structures. We apply this method to pure water and to cations in water so the central molecule in these clusters is either a water molecule or a Na^+ or K^+ ion as shown in Figure 1.

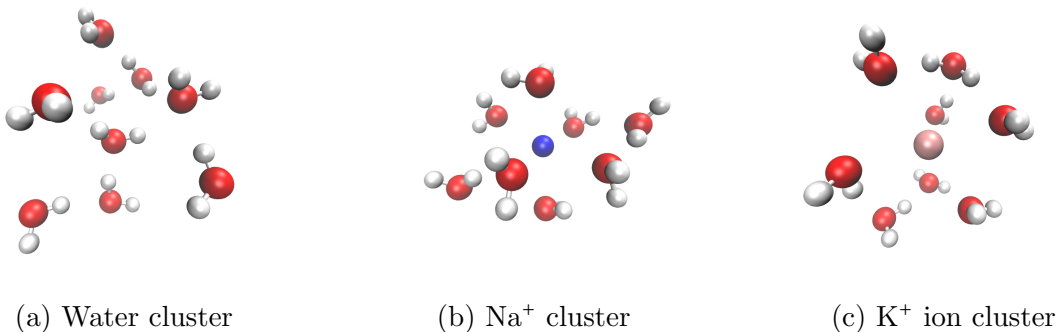


Figure 1: Cluster structures.

We compute the forces on the water molecules in these clusters with the revPBE-D3 functional and compare with Moller–Plesset perturbation theory level (MP2) as a reference. MP2 should be a reliable benchmark as it reproduces the dimer binding energy of water at the CCSD(T) level³⁶ and reproduces cation-water binding energies well.²⁹ Simulation of water with MP2 has also shown reasonable agreement with experimental water structure although approximations are required to reduce computational cost.^{12,14,36} However, MP2 can have significant errors for some cases and cannot always be used as a reliable benchmark.³⁷ In these cases a more sophisticated approach like CCSD(T) would be preferable.

The error in the radial component of the total force on each water molecule as a function of distance from the central molecule are calculated and shown in Figure 2 for the three cases. The oxygen–oxygen and cation–oxygen radial distribution functions (RDFs) from the simulation at the DFT-MD level of theory are shown below in Figure 3. We compare the theoretical calculations with the O – O RDFs from Skinner et al.³⁸ (black dashed line) and from Brookes & Head-Gordon (black circles).³⁹ For Na⁺ we compared with the rescaled Na⁺-O peak extracted from XRD of NaCl at 6M⁸ and for K⁺ we compare with the experimentally determined peak position.^{40–42} We also include simulation results with the SCAN functional, which we have previously shown provides an excellent description of the hydration structure around Na⁺ and K⁺ in comparison with X-ray adsorption fine structure (EXAFS) measurements.⁹

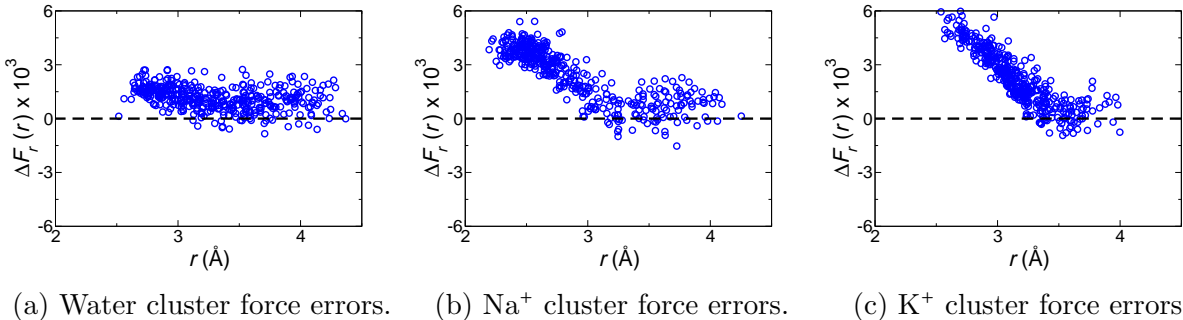


Figure 2: Error in the radial component of the force on water molecules as a function of distance from the central molecule with revPBE-D3. Forces are in atomic units times 10^3 .

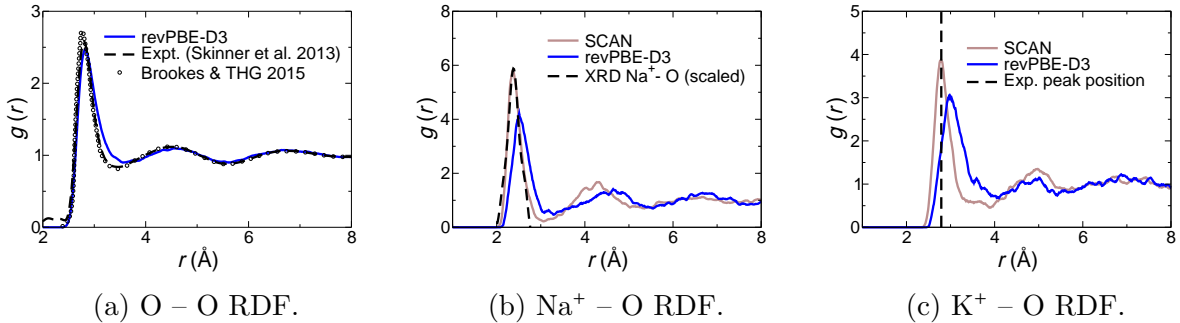


Figure 3: RDFs with revPBE-D3.

Comparing Figures 2 and 3 shows that the error plots provide an explanation for the errors observed in the simulations. Water is slightly too repelled on average from a central water molecule meaning the peak position is slightly too large in the water-water RDF. The water molecules in the first solvation shell of Na^+ and K^+ are repelled significantly too strongly explaining why the peak position is too large for the ion-oxygen RDFs. This method can therefore be used to predict the performance of a DFT functional without having to perform expensive simulations.

Pairwise interaction correction

A clear radial dependence in the error as a function of distance is observable for Na^+ and K^+ in Figure 2. This immediately suggests a simple fix to correct the revPBE-D3 functional. We can add a pairwise interaction term between the water and ions in order to remove this error. This philosophy is analogous to Grimme’s method for correcting DFT functionals for the lack of dispersion interactions,⁶ although in this case we are not necessarily correcting only the dispersion interaction specifically. This approach can achieve better accuracy than Grimme’s method because we are adjusting the parameters to fit high level quantum mechanical calculations on the specific system of interest itself rather than using parameters that need to be generalisable to any molecule.

There is only a weak radial dependence in the error for water. This is because any radial

dependence in the error of the water-water interaction is obscured by the interaction with water molecules in all directions. It is therefore more difficult to extract a clear correction term to improve the water-water interactions. However, revPBE-D3 already includes the Grimme C_6/r^6 dispersion correction term. The parameters for this correction are generally applicable but have not been specifically adjusted for water with this functional and basis set combination in the condensed phase. It is therefore reasonable to use the same function to correct the water-water interactions where the C_6 parameter is adjusted to minimise the new error metric provided here. We do not include short-range damping as it has only a minimal effect in this case. Dipole induction and orientational interactions also both have an inverse power sixth law dependence further motivating this choice of functional from.⁴³ We examine the dependence on this choice for the functional form below. Although this is a pairwise functional it is not simply correcting for the errors in the two-body interaction. Because it is fit to the error in the forces on small clusters it will also be effectively correcting for errors in the many-body interactions.

The C_6/r^6 function to correct the water water interactions acts between every oxygen pair. The C_6 parameter was determined to be -22.1 in a.u. For the ion-water correction interaction a quadratic form was used with damping by the Fermi function to ensure that it was monotonic and went to zero for large separations. The functional form is:

$$U(r) = A(r - b_1)^2 \left(1 - \frac{1}{1 + \exp\left(-\frac{r-b_2}{b_3}\right)} \right) \quad (1)$$

Taking the derivative of this gives the functional form for the force which is what we fit to the error given in Figure 2. In order to fit to the error plots we take the running average of the errors in order to smooth them. The parameters for the Na^+ function are: $A = -0.00053$, $b_1 = 6.905$, $b_2 = 5.19$ and $b_3 = 0.33$. For K^+ they are $A = -0.0013$, $b_1 = 6.63$, $b_2 = 6.0$ and $b_3 = 0.3$. All parameters are given in a.u. and were manually adjusted to minimise the errors. In Figure 4 we show the negative of the forces that result from these pairwise interactions compared with the error plots. Note that the error plots for Na^+ and K^+ are slightly different

from the ones given in Figure 2 as they include the correction from the C_6/r^6 interaction between water molecules.

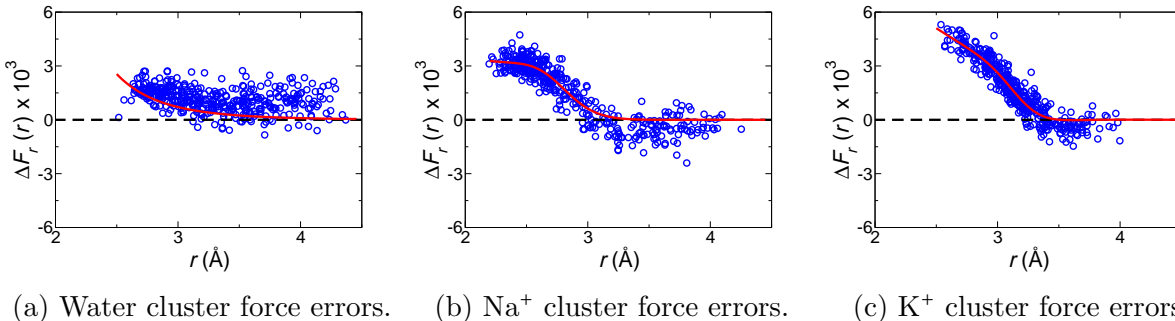
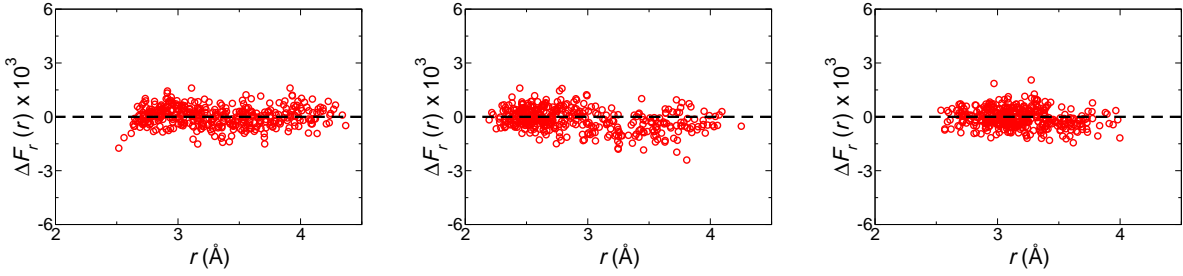


Figure 4: Error in the radial component of the force on water molecules as a function of distance from the central molecule with revPBE-D3 compared with the negative of the pairwise forces used to correct them.

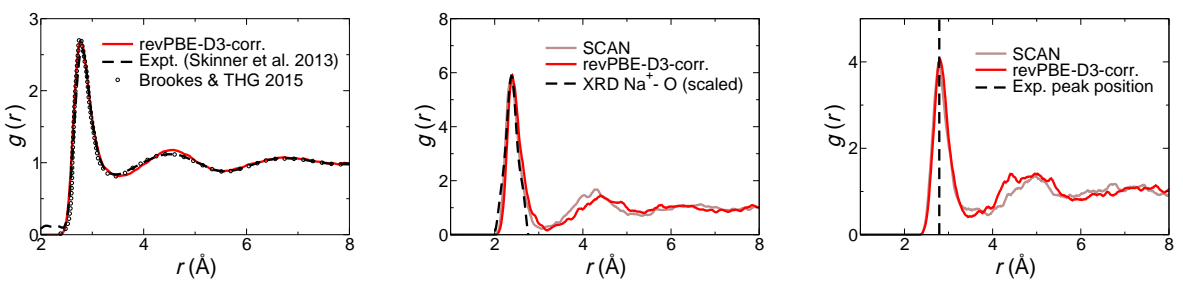
Combining these pairwise correction terms with the revPBE-D3 functional we recalculate the errors in the forces for these clusters as shown in Figure 5, which demonstrates that the average error is much closer to zero. In addition, the radial dependence of the error in the force is also significantly reduced. The error is slightly negative for the water molecules in the second hydration shell around the cations. For simplicity, we do not correct for this and use a simple monotonic energy function instead as the forces at these distances are quite sensitive to the water–water interaction and so these residual negative errors are likely an artefact of the choice of the pair wise function for the correction to the water–water interaction as well as possible failures of many body effects associated with the DFT functional. However, these errors are relatively minor and a more exhaustive search for pairwise functions or the use of many body interaction potentials should improve the agreement even more.

Figure 6 shows the RDFs that result from simulating with these correction forcefields added to the DFT forces. The resulting agreement for the RDF for water is significantly improved in almost total agreement with experiment. This agreement is comparable to water simulated with much higher levels of theory with significantly lower computational demands^{12,17} or to methods extensively parameterised to match cluster data.^{44,45}



(a) Water cluster force errors. (b) Na^+ cluster force errors. (c) K^+ cluster force errors.

Figure 5: Error in the radial component of the force on water molecules as a function of distance from the central molecule with corrected revPBE-D3.



(a) Corrected O – O RDF. (b) Corrected Na^+ – O RDF. (c) Corrected K^+ – O RDF.

Figure 6: RDFs with corrected revPBE-D3.

Additionally, the correct peak position is reproduced for the Na^+ and K^+ cases. These RDFs also match very well with the results of the higher level SCAN function, which we have previously shown agrees very well with experimental EXAFS results.⁹ This is particularly significant considering the quite poor performance of uncorrected revPBE-D3 to reproduce the hydration structure around these ions.

The approach outlined here has some similarities to traditional force matching approaches.^{46–48} Force matching normally works by assuming a functional form for the full potential energy surface with many parameters that are adjusted to minimise the mean error in the forces. This approach has shown promising results in application to ions in water.^{49–51} The key advantages of our new approach is that it uses a considered projection of the errors in the forces to develop a simple visualisable and intuitive metric of the error in a DFT functional, which can identify the key systematic error resulting in the failure to reproduce experimental liquid structure and provides a straightforward path to directly corrected for that error. Alternatively, correction potentials can be fitted to minimise the error in the experimental properties such as the RDF.^{52,53} This is obviously restricted to cases where accurate experimental data is available.

Obviously using pairwise interactions to correct this error metric cannot guarantee the final RDFs will be correct, particularly the O – O RDF, as there will be other errors this metric cannot detect. However, it appears that in this case it is detecting the main error of the revPBE-D3 functional. It should also be possible to use a modification of this approach to examine properties such as torques by looking at the non-radial components of the forces in order to ensure that angular distributions are also reproduced correctly. We can hopefully apply a modified version of this method to examine more complex cases such as solute distributions at surfaces and solute pairing in solutions. An additional issue is that ensuring the short-range forces are correct does not mean that long range slowly varying interactions or the and therefore the total binding energies will be correctly reproduced. For example, simulations with the isothermal isobaric ensemble show the density is significantly overesti-

mated with the addition of the C_6/r^6 form to correct the water-water interaction. This is consistent with the results of Local Molecular field theory,^{54,55} which demonstrates that it is possible to reproduce the correction hydration structure by only considering short-range strongly varying forces. But that longer range more slowly varying forces are important for determining bulk thermodynamic properties such as density. Properties that depend on these quantities will need additional fitting or other correction methods to be accurately reproduced. For instance, the error in the density can be easily corrected by adding a weak long-range potential that does not perturb the local forces significantly.

The use of a C_6/r^6 form for the water-water interaction is not rigorously justifiable. A more sophisticated method that also fit a function to simultaneously correct both the forces and cluster energies would be one option to improve this fitting.⁵⁶ Additionally, ideally the O-H and O-O interactions should be corrected separately to account for the fact that GGAs can over-stabilise hydrogen bonding in particular.⁵⁷ However, we examine the dependence on this choice by also fitting a C_3/r^3 function to correct the water-water error and then simulate the O-O RDF with this correction. As Figure 7 shows the C_3/r^3 functional form also removes the error in the water-water cluster forces adequately and results in excellent agreement with the RDF of pure water showing that this method does not seem to rely sensitively on the specific function chosen to correct for the error. The C3 parameter is set to -0.19 a.u.

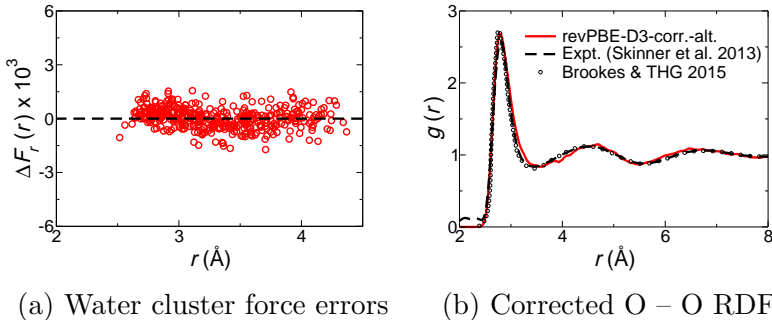


Figure 7: Errors and RDFs for water using the C_3/r^3 functional form to correct the errors in revPBE-D3.

Figure 8 compares the errors on clusters of 8 water molecules with the errors computed for a single dimer. The dimer structures are determined by taking the minimum energy structures and shifting the molecules positions. This shows that the dimer forces can not adequately explain the errors in the forces on the clusters and therefore will not be able to reliably correct the errors in the condensed phase. It is clear that the water-water dimer is too repulsive and the cation-water interactions have a long range over attraction that is not observed in the clusters likely due to electrostatic screening effects of the other waters. Some non-monotonic deviations at smaller separation are also observed. The error in the water-water dimer can be fit with the same the function as used for the cation-water case with parameters of $A = -0.0006$, $b_1 = 6.83$, $b_2 = 6.5$ and $b_3 = 0.2$. Figure 8 shows that the RDF of water simulated including this correction shows very poor experimental agreement. This demonstrates the limited usefulness of examining only the errors in the dimer interaction and shows the need for calculations on larger clusters.

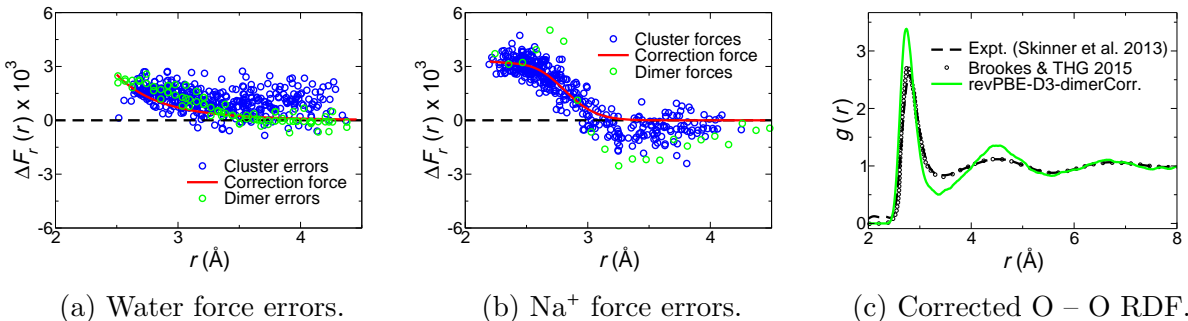


Figure 8: (a) and (b) show the errors in the radial component of the force on water molecules as a function of distance for clusters (blue) and dimers (green). (c) shows the O – O RDF that results from using a correction potential fitted to the error in the dimer force.

Application to chloride

Next we apply this method to the chloride anion to test its generalisability. Figure ?? shows this error metric for water molecules around the chloride anion. It is clear that the deviation from zero is dramatically smaller than for cations or for pure water. Including the C_6/r^6

water-water correction slightly reduces the overall error bringing it even closer to zero. This is consistent with the chloride–oxygen RDF calculated without any corrections, which shows good agreement with the experimentally determined peak position.^{8,58}

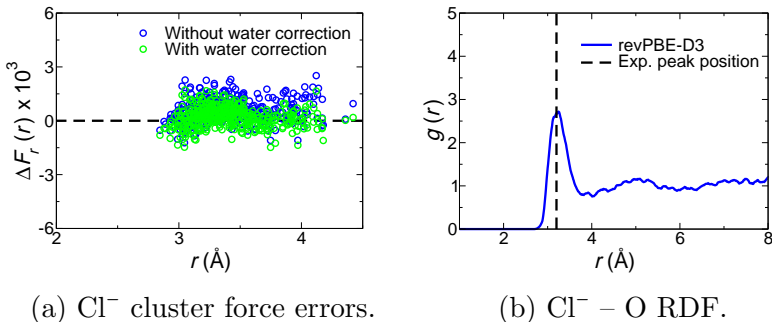
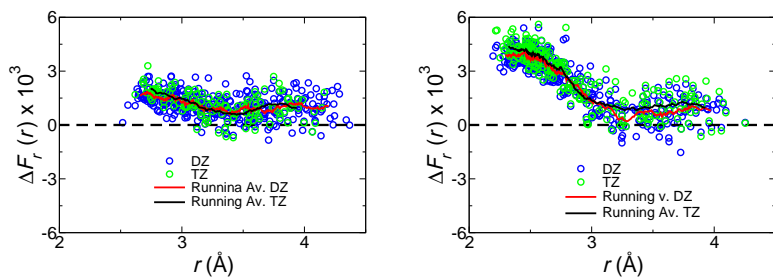


Figure 9: Error in the radial component of the force on water molecules as a function of distance from the Cl^- anion with revPBE-D3. And the Cl^- – O RDF simulated using revPBE-D3.

MP2 basis set dependence

One important question is whether the basis sets used for MP2 are sufficiently accurate to use as a benchmark for the errors in the forces as the double zeta basis sets used for these calculations may not be at the complete basis set limit which is normally closer to the quadruple zeta. To examine whether this is an issue or not Figure 10 compares the error calculated using the double zeta basis sets with triple zeta basis sets to show that there is no significant sensitivity on the basis sets used for the MP2 calculation for this metric. This is significant as triple zeta basis sets are much more expensive than double zeta. Double zeta are not at the complete basis set limit but luckily we do not need such high accuracy for this metric whereas we would if we were trying to reproduce absolute binding energies. We did not compute this quantity for potassium due to the higher computational resources required for triple-zeta basis sets.

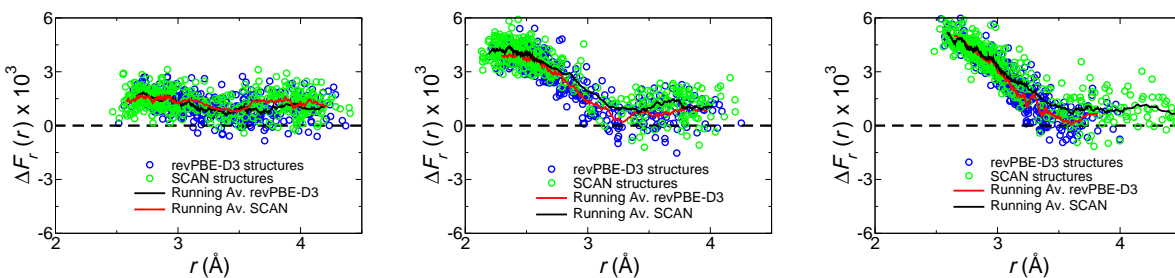


(a) Water cluster force errors. (b) Na^+ cluster force errors.

Figure 10: Comparison of double-zeta and triple-zeta basis sets for MP2 when calculating the error in the radial component of the force on water molecules as a function of distance from the central molecule with revPBE-D3.

Cluster structure dependence

Next we examine whether the cluster error plots depend on the underlying simulation from which the structures are drawn. To examine this we take simulations of the sodium and potassium ions performed with the SCAN functional, which we have presented in previous work. Figure 11 shows that there is no significant difference. This is important as we have shown that SCAN provides a much better description of the hydration structure of sodium and potassium but a worse description of water structure compared with revPBE-D3. This means that the error plots shown are not an artefact of the fact that we are drawing from an incorrect distribution.



(a) Water cluster force errors. (b) Na^+ cluster force errors. (c) K^+ cluster force errors.

Figure 11: Error in the radial component of the force on water molecules as a function of distance from the central molecule with revPBE-D3. Clusters are extracted from simulations using either the SCAN functional or the revPBE-D3 functional.

Cluster size dependence

Next we show that cluster error plots do not appear to depend sensitively on the number of water molecules used in the cluster. Figure 12 shows the results for potassium using clusters of 8 water molecules and 16 water molecules respectively. We see that the error is basically unchanged with the larger clusters. For water we also performed a calculation with clusters containing only two water molecules. This is shown in Figure 12 and demonstrates that the error in the force with revPBE-D3 water is consistent with the nine water molecule clusters although only a much smaller radial range is scanned. The running averages are also shown to demonstrate the similarity.

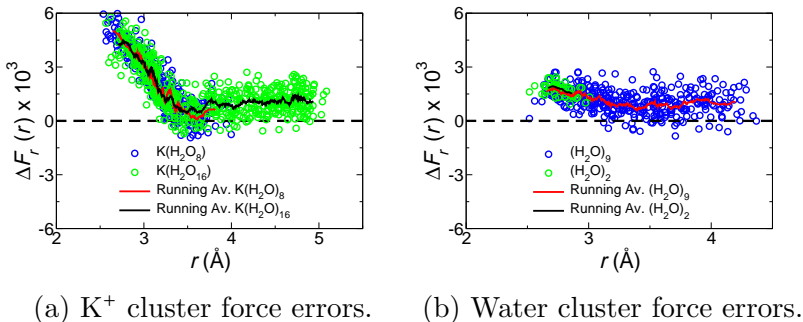


Figure 12: Error in the radial component of the force on water molecules as a function of distance from the central molecule with revPBE-D3. For potassium clusters or either 8 or 16 water molecules are extracted from simulations. For pure water clusters clusters of 2 and nine are examined.

Application to SCAN

Finally we compute this error metric for the SCAN functional. This is shown in Figure 13 We have shown SCAN accurately reproduces hydration structure of sodium and potassium but does not reproduce the hydration structure of water as well as revPBE-D3 showing a somewhat over-structured oxygen-oxygen RDF. The error plots computed are consistent with this interpretation. There is generally too much attraction between the water molecules, which would cause the peak position to be too small for the water-water rdf. The error for sodium

and potassium does not show any significant radial dependence and is just slightly too attractive but this is attributable to the error in the water-water interactions. This implies that the cation-water interactions are reasonably well reproduced with SCAN and explains why it reproduces the hydration structure around these ions well. The better agreement of SCAN is to be expected considering that it has been developed to satisfy 17 known constraints that a general exchange-correlation functional should satisfy. It can accurately reproduce binding energies and structures of a variety of molecules without empirical dispersion corrections.^{10,11} The RDFs with SCAN are shown below in Figure 14 for comparison.

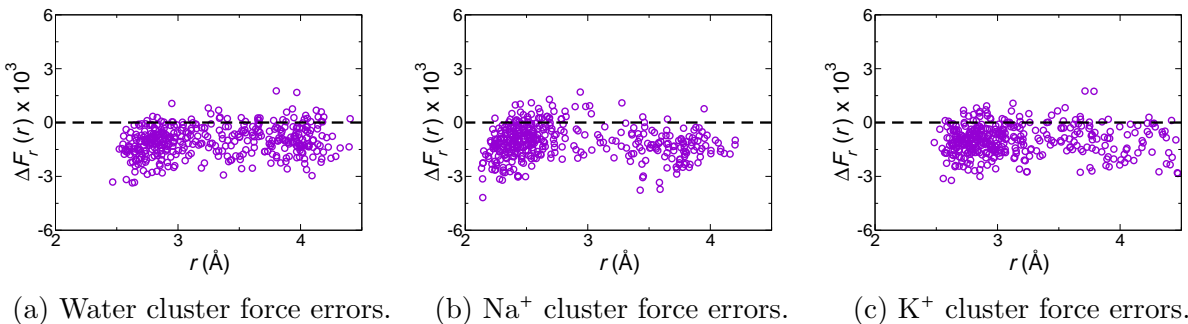


Figure 13: Error in the radial component of the force on water molecules as a function of distance from the central molecule with SCAN.

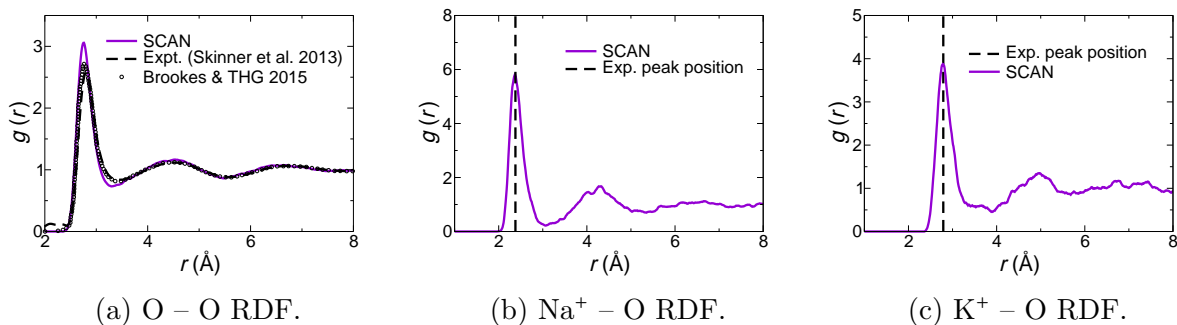


Figure 14: RDFs with SCAN.

Conclusions

We have described an improved metric for the error in a DFT functional. The significance of this metric is that it provides a test that can determine which DFT functional should be used for a given system and how reliable the liquid structure predictions are likely to be with that functional. This test requires only modest computational resources and we therefore advocate for its use whenever a new method is used to predict solvation structure. Additionally, this metric provides a straightforward method to quickly and cheaply correct even severe errors in the DFT functional resulting in significantly better experimental agreement. This result also means we should be able to reliably use cheap GGA functionals to directly calculate important properties of solutes in solution such as solvation free energies because once the hydration structures are properly reproduced similar simple methods can be applied to determine the absolute binding energies.²⁹

Simulation details

For the revPBE-D3 functional,^{4,5} Born-Oppenheimer *ab initio* molecular dynamics simulations within the constant volume NVT (at 300 K) ensemble using periodic boundary conditions are performed within the CP2K simulation suite (<http://www.cp2k.org>) containing the *QuickStep* module for the DFT calculations.³² The D3 dispersion correction due to Grimme⁶ was used for revPBE. A 0.5 fs time step was used. We used a double ζ basis set that has been optimized for the condensed phase³⁰ in conjunction with GTH pseudopotentials⁵⁹ using a 400 Ry cutoff for the auxiliary plane wave basis for the revPBE-D3 simulations and a 1200 Ry cutoff for the SCAN¹⁰ simulations.^{60,61} A Nosé-Hoover thermostat was attached to every degree of freedom to ensure equilibration.⁶² The energies were accumulated for ≈ 12 ps after 3 ps of equilibration. The sodium and potassium simulations for revPBE-D3 consisted of one sodium ion in a box of 96 water molecules of fixed dimensions of 14.3^3 \AA^3 giving a density of 1 g cm^{-3} . The same settings were used for the corrected revPBE-D3 simulations

with the exception that the multiple force evaluation option was used to combine the DFT forces with the pairwise forces computed using the FIST method. The SCAN simulation results used essentially the same set up and were taken from Ref. 9.⁹ The chloride anion was simulated in a 12.46^3 \AA^3 box with 64 water molecules giving a density of 1 g cm^{-3} for 17 ps after 3 ps of equilibration a newly optimised pseudopotential for chloride with the revPBE-D3 functional was used.

ORCA⁶³ was used to calculate the cluster forces at the MP2 level of theory. Clusters of 8 water molecules were used in the cluster correction calculation with 50 frames extracted from the 15 ps trajectory with the revPBE-D3 functional. The aug-cc-pVDZ basis set was used for the oxygen and hydrogen atoms.⁶⁴ Similarly, the cc-pCVDZ basis set was used for the sodium ion⁶⁵ and the cc-pwCVDZ basis set for the potassium ion.⁶⁶ Frozen cores were used for the MP2 calculations. For the revPBE-D3 and SCAN cluster energy calculations CP2K was used with the periodicity none option and a larger cell size to remove any box size dependence. Otherwise, the same parameters, basis sets *etc.* as the simulation were used.

Acknowledgements

XSZ and TTD acknowledge the Australian Research Council (ARC) funding via project number FL170100101 and DE200100794. GKS and CJM were supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences, and Biosciences. PNNL is a multiprogram national laboratory operated by Battelle for the U.S. Department of Energy. The SCAN calculations used resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility operated under Contract No. DE-AC02-05CH11231. Calculations were also performed on PNNL's Institutional Computing resource. This research was undertaken with the assistance of resources and services from the National Computational Infrastructure (NCI), which is supported by the Australian

Government and with the assistance of resources from QCIF (<http://www.qcif.edu.au>) and The University of Queensland's Research Computing Centre (RCC).

References

- (1) Doyle, D. A.; Cabral, J. M.; Pfuetzner, R. A.; Kuo, A.; MacKinnon, R.; Cabral, J. M.; Gulbis, J. M.; Cohen, S. L.; Chait, B. T.; MacKinnon, R. The structure of the potassium channel: Molecular basis of K⁺ conduction and selectivity. *Science* **1998**, *280*, 69–77.
- (2) Åvall, G.; Mindemark, J.; Brandell, D.; Johansson, P. Sodium-ion battery electrolytes: modeling and simulations. *Adv. Energy Mater.* **2018**, *8*, 1703036.
- (3) Gillan, M. J.; Alfè, D.; Michaelides, A. Perspective: how good is DFT for water? *J. Chem. Phys.* **2016**, *144*, 130901.
- (4) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
- (5) Zhang, Y.; Yang, W. Comment on Generalized gradient approximation made simple. *Phys. Rev. Lett.* **1998**, *80*, 890–890.
- (6) 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*, 154104.
- (7) Galib, M.; Duignan, T. T.; Misteli, Y.; Baer, M. D.; Schenter, G. K.; Hutter, J.; Mundy, C. J. Mass density fluctuations in quantum and classical descriptions of liquid water. *J. Chem. Phys.* **2017**, *146*, 244501.
- (8) Galib, M.; Baer, M. D.; Skinner, L. B.; Mundy, C. J.; Huthwelker, T.; Schenter, G. K.; Benmore, C. J.; Govind, N.; Fulton, J. L. Revisiting the hydration structure of aqueous Na⁺. *J. Chem. Phys.* **2017**, *146*, 084504.

- (9) Duignan, T.; Schenter, G. K.; Fulton, J.; Huthwelker, T.; Balasubramanian, M.; Galib, M.; Baer, M. D.; Wilhelm, J.; Hutter, J.; Ben, M. D. et al. Quantifying the hydration structure of sodium and potassium ions: taking additional steps on Jacob’s Ladder. *chemrxiv* **2018**, doi: 10.26434/chemrxiv.7466426.v2.
- (10) Sun, J.; Ruzsinszky, A.; Perdew, J. Strongly constrained and appropriately normed semilocal density functional. *Phys. Rev. Lett.* **2015**, *115*, 036402.
- (11) Sun, J.; Remsing, R. C.; Zhang, Y.; Sun, Z.; Ruzsinszky, A.; Peng, H.; Yang, Z.; Paul, A.; Waghmare, U.; Wu, X. et al. Accurate first-principles structures and energies of diversely bonded systems from an efficient density functional. *Nat. Chem.* **2016**, *8*, 831–836.
- (12) Del Ben, M.; Schönherr, M.; Hutter, J.; VandeVondele, J. Bulk liquid water at ambient temperature and pressure from MP2 theory. *J. Phys. Chem. Lett.* **2013**, *4*, 3753–3759.
- (13) DiStasio, R. A.; Santra, B.; Li, Z.; Wu, X.; Car, R. The individual and collective effects of exact exchange and dispersion interactions on the *ab initio* structure of liquid water. *J. Chem. Phys.* **2014**, *141*, 084502.
- (14) Del Ben, M.; Hutter, J.; VandeVondele, J. Probing the structural and dynamical properties of liquid water with models including non-local electron correlation. *J. Chem. Phys.* **2015**, *143*, 054506.
- (15) Del Ben, M.; Schütt, O.; Wentz, T.; Messmer, P.; Hutter, J.; VandeVondele, J. Enabling simulation at the fifth rung of DFT: large scale RPA calculations with excellent time to solution. *Comput. Phys. Commun.* **2015**, *187*, 120–129.
- (16) 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.

- (17) Pestana, L. R.; Marsalek, O.; Markland, T. E.; Head-Gordon, T. The quest for accurate liquid water properties from first principles. *J. Phys. Chem. Lett.* **2018**, *9*, 5009–5016.
- (18) Bajaj, P.; Götz, A. W.; Paesani, F. Toward Chemical Accuracy in the Description of IonWater Interactions through Many-Body Representations. I. HalideWater Dimer Potential Energy Surfaces. *J. Chem. Theory Comput.* **2016**, *12*, 2698–2705.
- (19) Riera, M.; Mardirossian, N.; Bajaj, P.; Götz, A. W.; Paesani, F. Toward chemical accuracy in the description of ion–water interactions through many-body representations. Alkali-water dimer potential energy surfaces. *J. Chem. Phys.* **2017**, *147*, 161715.
- (20) Riera, M.; Lambros, E.; Nguyen, T. T.; Götz, A. W.; Paesani, F. Low-order many-body interactions determine the local structure of liquid water. *Chem. Sci.* **2019**,
- (21) Otero-de-la Roza, A.; Johnson, E. R. Analysis of Density-Functional Errors for Noncovalent Interactions between Charged Molecules. *J. Phys. Chem. A* **2020**, *124*, 353–361.
- (22) Li, C.; Zheng, X.; Cohen, A. J.; Mori-sánchez, P.; Yang, W. Local Scaling Correction for Reducing Delocalization Error in Density Functional Approximations. *Phys. Rev. Lett.* **2015**, *114*, 053001.
- (23) Yang, Z.-h.; Pederson, M. R.; Perdew, J. P. Full self-consistency in the Fermi-orbital self-interaction correction. *Phys. Rev. A* **2017**, *95*, 052505.
- (24) Sharkas, K.; Wagle, K.; Santra, B.; Akter, S.; Zope, R. R.; Baruah, T.; Jackson, K. A.; Perdew, J. P.; Peralta, J. E. Self-interaction error overbinds water clusters but cancels in structural energy differences. *Proc. Natl. Acad. Sci.* **2020**, *117*, 11283–11288.
- (25) Yu, H.; Whitfield, T. W.; Harder, E.; Lamoureux, G.; Vorobyov, I.; Anisimov, V. M.; Mackerell, A. D.; Roux, B. Simulating monovalent and divalent ions in aqueous solution using a Drude polarizable force field. *J. Chem. Theory Comput.* **2010**, *6*, 774–786.

- (26) Mao, Y.; Demerdash, O.; Head-Gordon, M.; Head-Gordon, T. Assessing Ion–Water interactions in the AMOEBA force field using energy decomposition analysis of electronic structure calculations. *J. Chem. Theory Comput.* **2016**, *12*, 5422–5437.
- (27) Bryantsev, V. S.; Diallo, M. S.; Goddard, W. A. Calculation of solvation free energies of charged solutes using mixed cluster/continuum models. *J. Phys. Chem. B* **2008**, *112*, 9709–9719.
- (28) Pollard, T. P.; Beck, T. L. Toward a quantitative theory of Hofmeister effects: from quantum effects to thermodynamics. *Curr. Opin. Colloid Interface Sci.* **2016**, *23*, 110–118.
- (29) Duignan, T. T.; Baer, M. D.; Schenter, G. K.; Mundy, C. J. Real single ion solvation free energies with quantum mechanical simulation. *Chem. Sci.* **2017**, *8*, 6131 – 6140.
- (30) Vandevondele, J.; Hutter, J. Gaussian basis sets for accurate calculations on molecular systems in gas and condensed phases. *J. Chem. Phys.* **2007**, *127*, 114105.
- (31) Zhuang, D.; Riera, M.; Schenter, G. K.; Fulton, J. L.; Paesani, F. Many-body effects determine the local hydration structure of Cs⁺ in solution. *J. Phys. Chem. Lett.* **2019**, *10*, 406–412.
- (32) Vandevondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. Quickstep: Fast and accurate density functional calculations using a mixed gaussian and plane waves approach. *Comput. Phys. Commun.* **2005**, *167*, 103–128.
- (33) Medders, G. R.; Götz, A. W.; Morales, M. A.; Bajaj, P.; Paesani, F. On the representation of many-body interactions in water. *J. Chem. Phys.* **2015**, *143*.
- (34) Reddy, S. K.; Straight, S. C.; Bajaj, P.; Pham, C. H.; 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. **2016**, 1–37.
- (35) Paesani, F.; Bajaj, P.; Riera, M. Chemical accuracy in modeling halide ion hydration from many-body representations. *Adv. Phys. X* **2019**, *4*, 1631212.
- (36) Willow, S. Y.; Salim, M. A.; Kim, K. S.; Hirata, S. Ab initio molecular dynamics of liquid water using embedded-fragment second-order many-body perturbation theory towards its accurate property prediction. *Sci. Rep.* **2015**, *5*, 14358.
- (37) Cybulski, S. M.; Lytle, M. L. The origin of deficiency of the supermolecule second-order Møller-Plesset approach for evaluating interaction energies. *J. Chem. Phys.* **2007**, *127*.
- (38) Skinner, L. B.; Huang, C.; Schlesinger, D.; Pettersson, L. G. M.; Nilsson, A.; Benmore, C. J. Benchmark oxygen-oxygen pair-distribution function of ambient water from x-ray diffraction measurements with a wide Q-range. *J. Chem. Phys.* **2013**, *138*, 074506.
- (39) Brookes, D. H.; Head-Gordon, T. Family of oxygen-oxygen radial distribution functions for water. *J. Phys. Chem. Lett.* **2015**, *6*, 2938–2943.
- (40) Ohtaki, H.; Radnai, T. Structure and dynamics of hydrated ions. *Chem. Rev.* **1993**, *93*, 1157–1204.
- (41) Marcus, Y. Effect of ions on the structure of water: structure making and breaking. *Chem. Rev.* **2009**, *109*, 1346–1370.
- (42) Glezakou, V. A.; Chen, Y.; Fulton, J. L.; Schenter, G. K.; Dang, L. X. Electronic structure, statistical mechanical simulations, and EXAFS spectroscopy of aqueous potassium. *Theor. Chem. Acc.* **2006**, *115*, 86–99.
- (43) Israelachvili, J. N. *Intermolecular and Surface Forces*, 3rd ed.; Academic Press: San Diego, 2011.

- (44) Medders, G. R.; Babin, V.; Paesani, F. Development of a “first-principles” water potential with flexible monomers. III. Liquid phase properties. *J. Chem. Theory Comput.* **2014**, *10*, 2906–2910.
- (45) Das, A. K.; Urban, L.; Leven, I.; Loipersberger, M.; Aldossary, A.; Head-Gordon, M.; Head-Gordon, T. Development of an advanced force field for water using variational energy decomposition analysis. *J. Chem. Theory Comput.* **2019**, doi: 10.1021/acs.jctc.9b00478.
- (46) Ercolesi, F.; Adams, J. B. Interatomic potentials from first-principles calculations: The force-matching method. *Europhys. Lett.* **1994**, *26*, 583–588.
- (47) Izvekov, S.; Parrinello, M.; Bumham, C. J.; Voth, G. A. Effective force fields for condensed phase systems from *ab initio* molecular dynamics simulation: A new method for force-matching. *J. Chem. Phys.* **2004**, *120*, 10896–10913.
- (48) Masia, M.; Guàrdia, E.; Nicolini, P. The force matching approach to multiscale simulations: Merits, shortcomings, and future perspectives. *Int. J. Quantum Chem.* **2014**, *114*, 1036–1040.
- (49) Spångberg, D.; Guàrdia, E.; Masia, M. Aqueous halide potentials from force matching of Car-Parrinello data. *Comput. Theor. Chem.* **2012**, *982*, 58–65.
- (50) Li, J.; Wang, F. Pairwise-additive force fields for selected aqueous monovalent ions from adaptive force matching. *J. Chem. Phys.* **2015**, *143*, 194505.
- (51) Li, J.; Wang, F. Accurate Prediction of the Hydration Free Energies of 20 Salts through Adaptive Force Matching and the Proper Comparison with Experimental References. *J. Phys. Chem. B* **2017**, acs.jpcc.7b04618.
- (52) White, A. D.; Knight, C.; Hocky, G. M.; Voth, G. A. Communication: Improved *ab*

- initio* molecular dynamics by minimally biasing with experimental data. *J. Chem. Phys.* **2017**, *146*, 041102.
- (53) Amirkulova, D. B.; White, A. D. Recent advances in maximum entropy biasing techniques for molecular dynamics. *Mol. Simul.* **2019**, *45*, 1285–1294.
- (54) Rodgers, J. M.; Weeks, J. D. Accurate thermodynamics for short-ranged truncations of Coulomb interactions in site-site molecular models. *J. Chem. Phys.* **2009**, *131*.
- (55) Remsing, R. C.; Rodgers, J. M.; Weeks, J. D. Deconstructing Classical Water Models at Interfaces and in Bulk. *J. Stat. Phys.* **2011**, *145*, 313–334.
- (56) Wang, L. P.; Van Voorhis, T. Communication: Hybrid ensembles for improved force matching. *J. Chem. Phys.* **2010**, *133*.
- (57) Calio, P. B.; Hocky, G. M.; Voth, G. A. A Minimal Experimental Bias on the Hydrogen Bond Greatly Improves *ab initio* Molecular Dynamics Simulations of Water. *arXiv* **2020**, arXiv:2006.01929.
- (58) Duignan, T. T.; Parsons, D. F.; Ninham, B. W. A Continuum Solvent Model of the Multipolar Dispersion Solvation Energy. *J. Phys. Chem. B* **2013**, *117*, 9412–9420.
- (59) Goedecker, S.; Teter, M.; Hutter, J. Separable dual-space Gaussian pseudopotentials. *Phys. Rev. B* **1996**, *54*, 1703–1710.
- (60) Miceli, G.; Hutter, J.; Pasquarello, A. Liquid water through density-functional molecular dynamics: Plane-wave vs atomic-orbital basis sets. *J. Chem. Theory Comput.* **2016**, *12*, 3456–3462.
- (61) Yao, Y.; Kanai, Y. Free energy profile of NaCl in water: First-principles molecular dynamics with SCAN and ω B97X-V exchange-correlation functionals. *J. Chem. Theory Comput.* **2018**, *14*, 884–893.

- (62) Martyna, G. J.; Klein, M. L.; Tuckerman, M. Nosé–Hoover chains: the canonical ensemble via continuous dynamics. *J. Chem. Phys.* **1992**, *97*, 2635–2643.
- (63) Neese, F. The ORCA program system. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 73–78.
- (64) Dunning, Jr., T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023.
- (65) Woon, D. E.; Dunning, Jr., T. H. Gaussian basis sets for use in correlated molecular calculations. IV. Calculation of static electrical response properties. *J. Chem. Phys.* **1994**, *100*, 2975–2988.
- (66) Peterson, K. Private communication. 2017.

Graphical TOC Entry

