Title
Simple and Accurate One-Body Energy and Dipole Moment Surfaces for Water and Beyond

Journal
The Journal of Physical Chemistry Letters, 15(26)

ISSN
1948-7185

Authors
Sami, Selim
LaCour, R Allen
Heindel, Joseph P
et al.

Publication Date
2024-07-04

DOI
10.1021/acs.jpclett.4c00587

Peer reviewed

THE JOURNAL OF PHYSICAL CHEMISTRY LETTERS
A JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/JPCL
Letter

Simple and Accurate One-Body Energy and Dipole Moment Surfaces for Water and Beyond

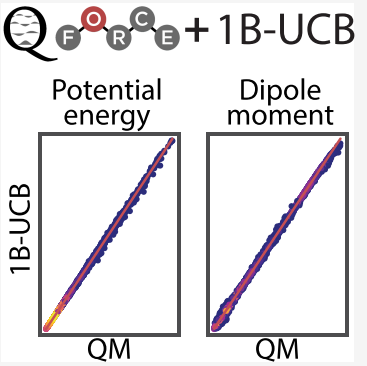
Selim Sami,* R. Allen LaCour, Joseph P. Heindel, and Teresa Head-Gordon*

Cite This: J. Phys. Chem. Lett. 2024, 15, 6712−6721

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Water is often the testing ground for new, advanced force fields. While advanced functional forms for intermolecular interactions have been integral to the development of accurate water models, less attention has been paid to a transferable model for intramolecular valence terms. In this work, we present a one-body energy and dipole moment surface model, named 1B-UCB, that is simple yet accurate and can be feasibly adapted for both standard and advanced potentials. 1B-UCB for water is comparable in accuracy to those with much more complex functional forms, despite having drastically fewer parameters. The parametrization protocol has been implemented as part of the Q-Force automated workflow and requires only a quantum mechanical Hessian calculation as reference data, hence allowing it to be easily extended to a variety of molecular systems beyond water, which we demonstrate on a selection of small molecules with different symmetries.

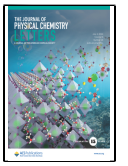


Molecular dynamics simulations using force fields (FFs) are fundamental for improving our understanding of biological and chemical systems and developing new materials. As intermolecular interactions often play an important role in condensed phase phenomena, they are typically the main focus of FF development. To reach the highest levels of accuracy for intermolecular interactions, new advanced FFs can now account for quantum mechanical (QM) effects, including polarization, charge transfer, and charge penetration.[1,2] Water is typically the testing ground for new models of the potential energy surface, and many successful advanced FFs for water have been developed, such as MB-Pol,[3−5] q-AQUA,[6−8] TTM,[9,10] AMOEBA,[11−14] HIPPO,[15,16] FFlux,[17] and MB-UCB.[18] Some of these water potentials have reached a high level of accuracy, such that they work well over a wide range of the phase diagram of water.

Often a similarly advanced functional form is used to describe water intramolecular interactions, such as the Partridge and Schwenke (PS) one-body (1B) potential,[19] which is used by MB-Pol, q-AQUA, and TTM models. These advanced water potentials have reached a high level of accuracy for spectroscopic properties as a result of better treatment of the 1B potential, yielding, for example, more accurate infrared intensities in line with experiments. However, the PS energy and dipole moment surface (DMS) involves hundreds of parameters fit using tens of thousands of data points, and thus, it is challenging to extend such approaches to larger organic molecule or material systems.[20]

On the other end of the complexity spectrum are the harmonic bonds and angles that are the workhorse of many classical force fields.[21−25] These simple 1B potentials offer ease of parametrization for a large range of molecules while having a low computational cost. However, they often neglect coupling between different bonding terms as well as anharmonicity, resulting in large errors in the spectroscopic properties. Physically motivated functional forms to account for anharmonicity, such as Morse potentials for bonds and cosine angles for bending, and the coupling between these terms have long been known,[20] but their application to commonly used FFs have been lacking. Several FFs, such as AMOEBA and MB-UCB, have introduced some additional terms to overcome these problems, such as the Urey−Bradley potential, which introduces a harmonic coupling between bonds and angles as well as using higher order polynomials for the bonding potentials to account for anharmonicity. However, even when these terms are included in the FF, they tend to exhibit problems in the 1B potential. For example, the anharmonicity terms in the AMOEBA and MB-UCB water model are taken from the MM3 model for hydrocarbons[26] and were not reoptimized, and the harmonic force constants are then fit to anharmonic experimental frequencies; thus, the anharmonicity is double-counted. q-TIP4P/F[27] uses a Morse potential for stretches for water but lacks anharmonic angles and coupling terms while also empirically fitting its 1B potential to condensed phase properties.

Received: February 23, 2024
Revised: June 7, 2024
Accepted: June 13, 2024
Published: June 20, 2024



https://doi.org/10.1021/acs.jpclett.4c00587
J. Phys. Chem. Lett. 2024, 15, 6712−6721

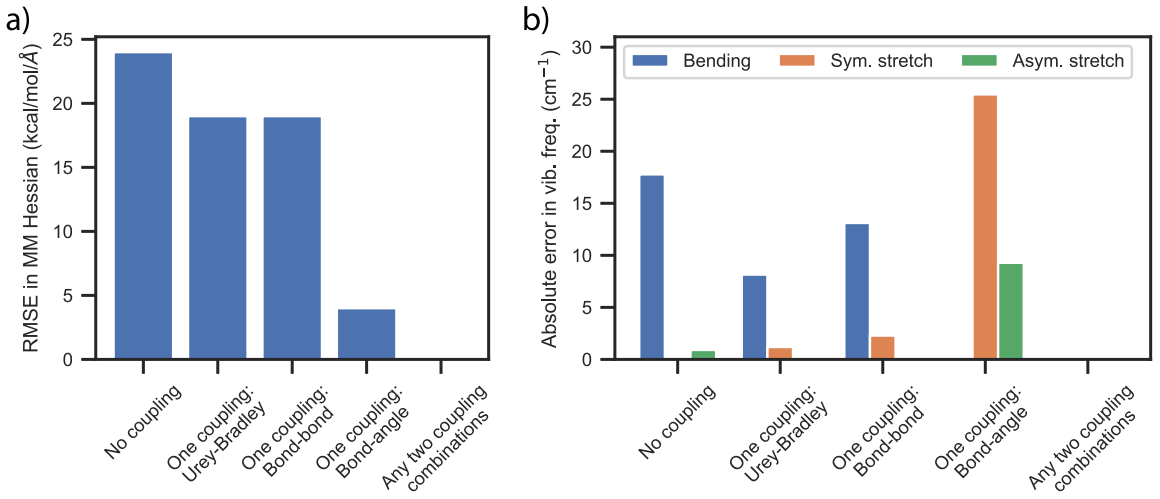


Figure 1. Accuracy of different coupling models for the 1B MM potential. (a) RMSE to CCSD(T) Hessian and (b) absolute error in vibrational frequencies for no coupling term, various single coupling terms, and any combination of two coupling terms. Using any two coupling terms results in a RMSE smaller than 0.01 kcal mol^{-1} Å^{-1}; thus, these are plotted together. We found that fitting the bonded terms to both the modes and frequencies created errors on the QM Hessian terms as large as 50% compared to fitting directly to Hessian. For panel b, errors smaller than 1 cm^{-1} are set to zero with the assumption that this is within numerical error of our Hessian calculation.

Likewise, the 1B DMS is also quite poor for models with charges that are fixed with respect to geometrical distortions, resulting in unphysical trends, such as a decrease in the bond angle of water when moving from the gas to condensed phase. As mentioned earlier, PS DMS is exceptionally accurate yet requires hundreds of parameters. Similarly, DMS of water by Lodi and Tennyson[28] has comparable if not higher accuracy and complexity than PS. This was used by Bowman and co-workers for the DMS for multiple properties of liquid water and ice, including the infrared (IR) spectrum. An alternative approach is predicting not just the DMS but the whole geometry-dependent atomic multipole moments based on machine-learning approaches, as done in the FFlux model.[29,30] On the simpler side of things, the three-parameter charge flux approach for water has been developed by Dinur et al.[31] and since been used by AMOEBA+,[32] HIPPO,[15] and TTM3-F.[10] In terms of going beyond water, AMOEBA+ has recently made significant progress in this regard, parametrizing charge flux for small molecules with a variety of functional groups based on bond and angle scans. However, this approach neglects parts of the DMS, namely, the coupling between bond and angle distortions, and is also unwieldy for larger molecules.

In this work, we first design 1B energy and dipole moment surfaces for water using QM Hessian, which we believe is a more generalizable and transferable approach. First, we demonstrate that, by judicious choice of functional form and parametrization strategy, we can obtain a 1B energy potential that reproduces QM Hessian as well as energies in the anharmonic region. Next, we demonstrate that an accurate DMS can be obtained by parametrizing a charge flux model using QM Hessian and the associated dipole derivatives. Additionally, we show that a novel alternative method is to use virtual sites to obtain an accurate DMS, which may be especially useful for models that possess such sites. Finally, we demonstrate on four other molecules with different symmetries that our simple yet accurate 1B model, 1B-UCB, is conveniently extensible to molecules beyond water as a result of the automated parametrization strategy.

Our first goal is to construct an energy surface that reproduces the QM 1B energy surface. We note that we fit two parameter sets. The first will be fit to energies computed using ωB97X-V[33] with the def2-QZVPPD basis set,[34] while the other will be fit to energies computed using CCSD(T) with aug-cc-pV5Z[35] in the other. These two models will be termed density functional theory (DFT)- and coupled cluster (CC)-based molecular mechanics (MM) models, respectively. We fit to two levels of QM because, while CCSD(T) energies are relatively inexpensive for the water monomer, DFT calculations will be much more affordable for larger molecules, and thus, it is important to compare their relative accuracy. We also use DFT to evaluate the larger water validation data sets that are not feasible for CC calculations. We justify the use of ωB97X-V[33] because it has been shown to be the best hybrid generalized gradient approximation (GGA) functional compared to CCSD(T) and beyond.[36,37] QM calculations were performed with Q-Chem 6.1.[38] Our models were implemented in and evaluated within the OpenMM software package.[39]

We begin by constructing a model that reproduces QM Hessian, which describes the energy surface near its minimum. While the PS energy surface does reproduce Hessian, harmonic bond and angle potentials cannot because they neglect coupling between different types of bonds and angles. While coupling is most often introduced through the Urey–Bradley expression, which is an additional harmonic bond potential between two water hydrogens, alternative schemes could include bond–bond

$$V_{\text{bond–bond}}(r_1, r_2) = k_{\text{bb}}(r_1 - r_1^0)(r_2 - r_2^0) \quad (1)$$

and bond–angle

$$V_{\text{bond–angle}}(r, \theta) = k_{\text{ba}}(r - r^0)(\cos\theta - \cos\theta^0) \quad (2)$$

coupling terms. With the aim of minimizing the number of FF terms while fully reproducing QM Hessian, we examined the performance of every FF that could be constructed by combining the standard harmonic bond (for the OH stretches) and angle potentials (for the HOH angle) with zero, one, or two coupling terms. We parametrized the force constants by fitting MM Hessian to QM Hessian using the Q-Force toolkit,[40] which seeks to minimize the difference between the

https://doi.org/10.1021/acs.jpclett.4c00587
J. Phys. Chem. Lett. 2024, 15, 6712−6721

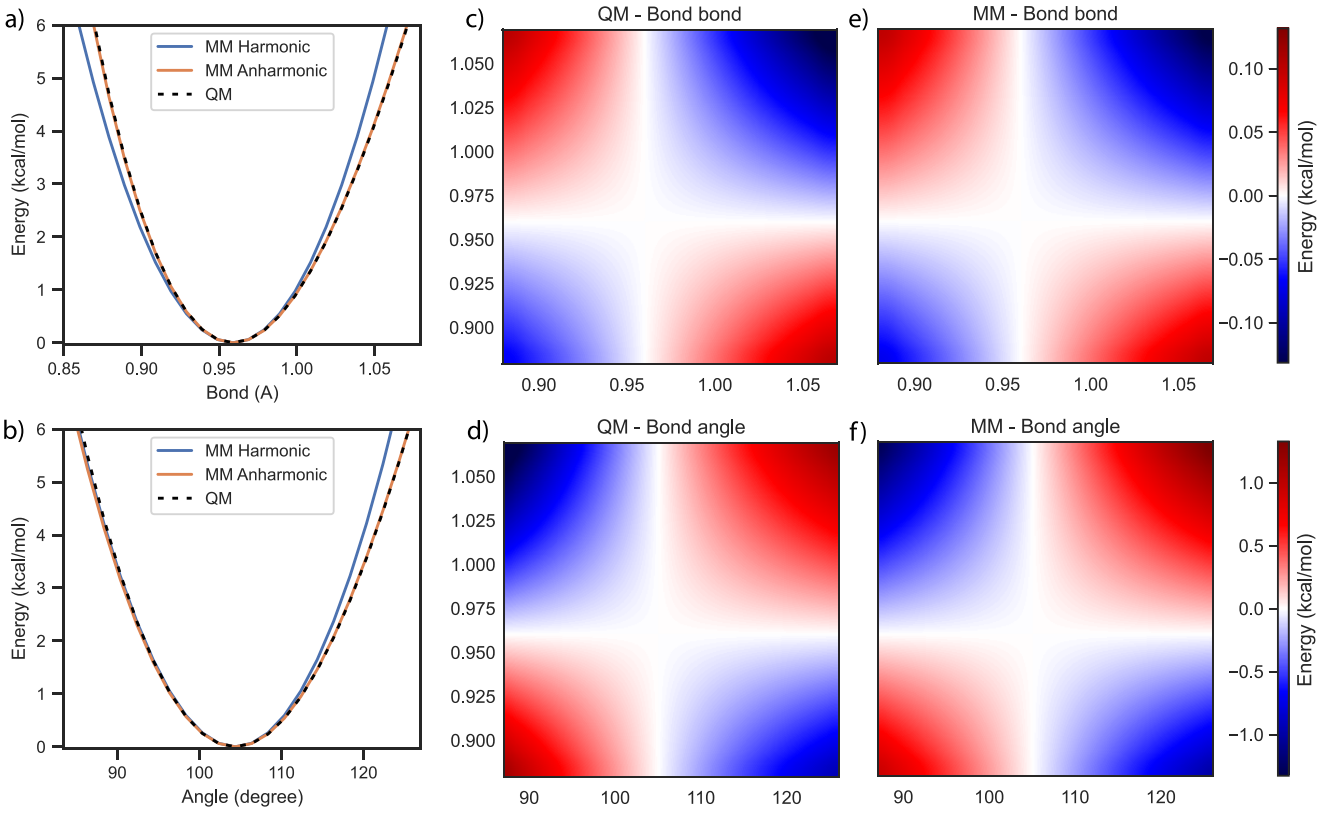


Figure 2. Validation of the 1B MM model against the bond and angle CCSD(T) scan data. 1D scan of the (a) bonds and (b) angles for energies below 10 kbT. We obtained MAEs of 0.014 and 0.038 kcal/mol for the bond and angle energies, respectively. The maximum errors at the largest displacements were 0.06 and 0.26 kcal/mol for bond and angle energies, respectively. 2D scan of (c) QM bond−bond, (d) MM bond−bond, (e) QM bond−angle, and (f) MM bond−angle coupling energies (kcal/mol). We obtained MAEs of 0.0005 and 0.015 kcal/mol for the bond−bond and bond−angle couplings, respectively. The maximum errors at 10 kbT displacements were 0.003 and 0.11 kcal/mol for the bond−bond and bond−angle couplings, respectively.

MM and QM Hessian elements. We note that bonded terms are frequently parametrized to match either the vibrational frequencies or a combination of vibrational frequencies and modes, but these typically result in very inaccurate Hessians. Instead parametrizing a FF that fully reproduces QM Hessian, as we do, ensures that vibrational modes and frequencies will also be reproduced.

As seen in Figure 1a, the QM Hessian pattern cannot be reproduced well without any coupling. Introducing one coupling term is still insufficient, although introducing bond−angle coupling performs better than alternative single coupling terms. In contrast, any combination of two coupling terms can fully reproduce QM Hessian, with the error in any Hessian element below 0.01 kcal mol^{-1} Å^{-1}. We choose the bond−bond (eq 1) and bond−angle (eq 2) coupling terms for our final FF because we find that using one term that exclusively couples the two bonds and one that couples the bonds with the angle is more intuitive, as the Urey−Bradley term does some degree of both. We also find that combining the bond−bond and bond−angle coupling terms performs slightly better in the anharmonic region than any combination involving the Urey−Bradley term. When we look at the absolute errors in individual vibrational modes (Figure 1b), without bond−angle coupling, the bending mode makes significant errors, while having only a bond−angle coupling term results in large errors in stretch modes.

While harmonic potentials suffice to reproduce QM Hessian, they will obviously perform poorly outside the harmonic region. One solution is to scan the energy surface of the water monomer and fit the higher order terms of a polynomial to match the anharmonicity. However, this requires one to compute significantly more QM data and add several more parameters; while not an issue for water specifically, this would complicate extending the parametrization strategy to larger molecules. To circumvent this issue, we make use of the physically motivated functional forms given by Dateo et al.,[20] whose parameters require less QM data to fit. For the angle term, this is a cosine angle expression

$$V_{\text{angle}}(\theta) = \frac{k_a}{2}(\cos\theta - \cos\theta^0)^2 \tag{3}$$

which has a single force constant k_a that can be fit to Hessian. For the bonding term, this is the Morse potential

$$V_{\text{Morse}}(r) = D[1 - \exp(-\beta(r - r^0)]^2 \tag{4}$$

where

$$\beta = \sqrt{\frac{k_b}{2D}} \tag{5}$$

with k_b and D being the harmonic force constant and the dissociation energy, respectively. Conveniently, k_b can be fit to QM Hessian, and thus, only D has to be determined, for which either experimental value can be taken or it can be computed by a single QM calculation. For the dissociation energy of hydrogen in water, both experimental (118.8 kcal/mol) and

https://doi.org/10.1021/acs.jpclett.4c00587
J. Phys. Chem. Lett. 2024, 15, 6712−6721

theoretical (DFT, 123.0 kcal/mol; CC, 125.3 kcal/mol) values were tested. We found minimal difference between using either experimental or theoretical values in energy regimes that would be relevant for non-reactive FFs. This is encouraging because it likely enables using generic D values for specific atom pairs or functional groups that do not have to be optimized for each molecule. Furthermore, it can be desirable sometimes to prohibit any bond dissociation, which is technically possible for the Morse potential. Conveniently, for these cases, both Morse and cosine angle potentials can be converted to a Taylor expansion without requiring any additional parameters.

We first validate our energy surface by comparing our model-predicted energies to QM energies for various bond and angle deformations. In panels a and b of Figure 2, we compare one-dimensional (1D) scans of the bond and angle energies while keeping the other degrees of freedom fixed. While the harmonic potential starts deviating significantly at larger displacements, both anharmonic potentials perform very well, even at large displacements of 10 kbT (~6 kcal/mol at room temperature). In panels c–f of Figure 2, we compare two-dimensional (2D) scans of the bond–bond and bond–angle coupling energies. We see that bond–angle coupling energies are about 1 order of magnitude larger than the bond–bond coupling energies, which may explain the effectiveness of the bond–angle coupling shown in Figure 1. When all three degrees of freedom are compared against the benchmark reference for the 1B energies, the mean absolute error (MAE) and maximum error at the 10 kbT displacements are 0.025 and 0.29 kcal/mol, respectively. The FF parameters for the 1B energy surface are provided in Table 1; the Taylor expansion equations and parameters are provided in the Supporting Information.

Table 1. 1B-UCB Potential Energy Surface Parameters for the Two Models[a]

parameter	ωB97x-V/def2-qzvppd	CCSD(T)/aug-cc-pV5Z
r^0 (Å)	0.959274	0.958413
k_b (kJ mol^{-1} Å^{-2})	5151.75	5098.15
D (kJ/mol)	514.757	524.265
θ^0 (deg)	105.0387	104.4234
k_a (kJ/mol)	445.977	452.183
k_{bb} (kJ mol^{-1} Å^{-2})	−45.4801	−61.1423
k_{ba} (kJ mol^{-1} Å^{-1})	−153.552	−159.886

[a]The 1B potential is parameterized against wB97X-V/def2-QZVPPD and CCSD(T)/aug-cc-pV5Z Hessians.

In Figure 3a, we compare our predicted sum of 1B energies of water clusters containing 7, 10, or 16 waters to DFT, CCSD(T), and PS using the water cluster geometries taken from Herman and Xantheas.[41] We see that our 1B model, fit with either DFT or CCSD(T), accurately reproduces its QM reference, with per molecule errors of less than 0.0025 kcal/mol. The 1B energies predicted by PS also compare quite well but with seemingly larger differences compared to the CCSD(T)/aug-cc-pV5Z energies; this is likely because it was fit to additional empirical corrections beyond CCSD(T)/aug-cc-pV5Z energies.

Besides these three specific water clusters, 1B energies were also calculated for 1200 structures, each of 2-, 3-, 4-, 5-mer clusters, that were taken Wang et al.;[11] we only compute the DFT energies because the many CCSD(T) calculations are too expensive computationally. In Figure 3b, we show that our DFT-fitted model is quite accurate, while in Figure 3c, we compared the DFT energies to our model fit to CC Hessian. Considering the close match between the CC results and our CC-based model in Figure 3a, this comparison serves more as a validation of the DFT functional than our model. Thus, we conclude that the DFT error is less than 0.06 kcal/mol per water molecule, showing that ωB97X-V/def2-QZVPPD performs exceptionally well as a reference when CCSD(T)/aug-cc-pV5Z is no longer viable.

For further comparison, we also calculated harmonic and anharmonic frequencies from our MM model to compare against those of the PS model and the CCSD(T)/aug-cc-pV5Z benchmark. We denote poly5z as a high-level fit to the CCSD(T)/aug-cc-pV5Z energy surface using 12th-order polynomials, similar to our previous work[42] but without a basis set extrapolation. In Table 2, we show that all models reproduce the harmonic frequencies quite accurately. For the experimental anharmonic frequencies, the PS model has an exact match because it was explicitly fit to these data empirically. Therefore, the poly5z models provide a better comparison for our MM model in terms of its accuracy relative to the CCSD(T) surface. We see that the poly5z and MM models are both accurate for anharmonic frequencies, with errors ranging from 5 to 20 wavenumbers. Our MM model performs very similarly to the poly5z model, with the greatest difference being the 11 wavenumber difference in the bending mode. The greater error in the bending mode likely stems from the Morse potential predicting slightly more accurate bond energies than the cosine angle potential predicting angle energies, as shown in Figure 2. With our DFT-fitted MM model, the harmonic and anharmonic frequencies are all blue-shifted by 10–20 wavenumbers while retaining the correct trends, as typically experienced with hybrid functionals.[43]

We next turn to parametrizing a 1B DMS. We note that, without explicit charge flow, any point charge or multipole model will overestimate the change in the dipole moment with bond length or angle. A well-known consequence of this is that the average bond angle of water in the condensed phase will be lower than that in the gas phase, which contradicts the experiment. While the PS model circumvents this problem using a complex functional form that requires many data points to fit, a common approach is to use a simple charge flux model,[31] where some amount of charge moves between the atoms based on the deviation from equilibrium bond lengths and angles. In this case, the 1B DMS can be expressed as

$$
\begin{aligned}
dq_{H_1} &= j_b(r_{OH_1} - r^0_{OH}) + j_\theta(\theta - \theta_0) + j_{bb}(r_{OH_2} - r^0_{OH}) \\
dq_{H_2} &= j_b(r_{OH_2} - r^0_{OH}) + j_\theta(\theta - \theta_0) + j_{bb}(r_{OH_1} - r^0_{OH}) \\
dq_O &= -(dq_{H_1} + dq_{H_2})
\end{aligned}
\tag{6}
$$

where j_b is the bond flux parameter, j_θ is the angle flux parameter, and j_{bb} is the bond–bond coupling flux parameter. Again, these three parameters are typically fit by scanning all relevant bonds and angles, which is feasible for water but will become more difficult for larger molecules. Instead, we used the dipole derivative matrix obtained from a QM Hessian calculation. The first advantage to this approach is that it requires no extra calculations to parametrize the DMS. An additional and significant benefit is that fitting to the QM dipole derivatives is a more rigorous approach than fitting to

https://doi.org/10.1021/acs.jpclett.4c00587
J. Phys. Chem. Lett. 2024, 15, 6712–6721

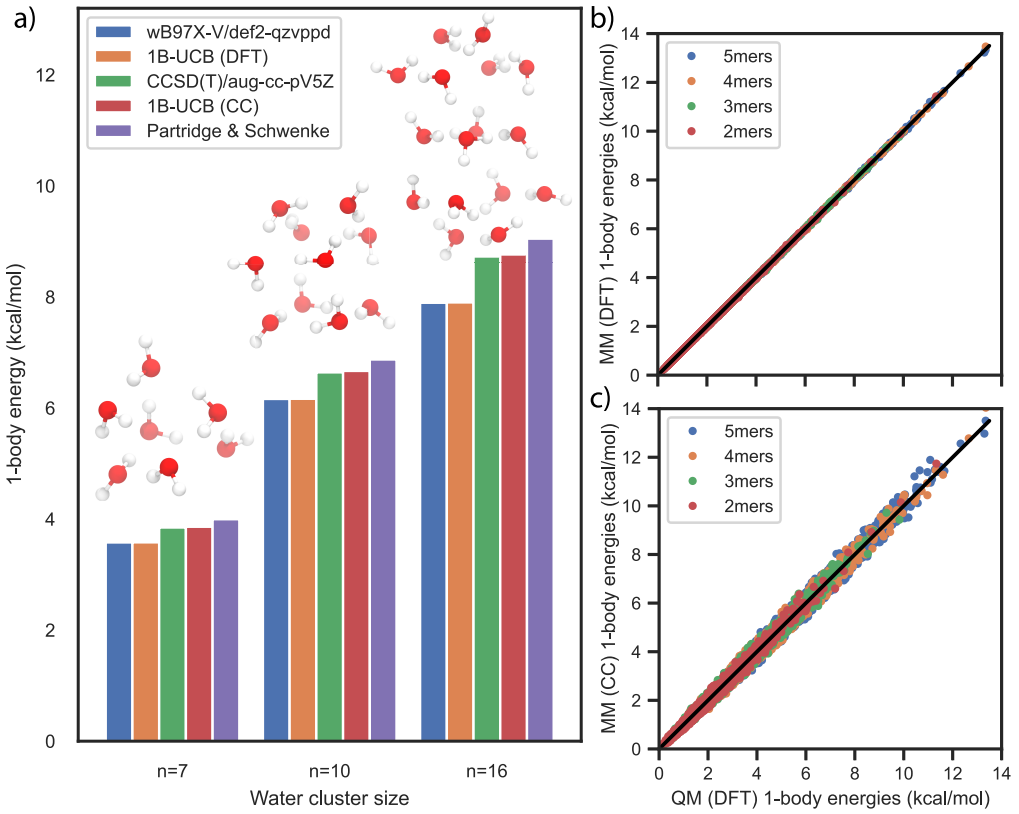


Figure 3. Comparison of 1B potential energies between QM and MM models. (a) 7-, 10-, and 16-mer water cluster 1B energies for DFT, CCSD(T) and PS references and our model on MP2/aug-cc-pVTZ optimized structures taken from Herman and Xantheas.[41] Our simple model has errors of 0.014, 0.026, and 0.036 kcal/mol for CC and 0.002, 0.003, and 0.006 kcal/mol for DFT for the 7-, 10-, and 16-mer water clusters, respectively. (b) Energies of our 1B model derived from DFT plotted against ωB97X-V/def2-qzvppd energies for 1200 snapshots taken from each cluster size, from 2- to 5-mers. We find MAEs of 0.007, 0.009, 0.011, and 0.013 kcal/mol and maximum errors of 0.115, 0.122, 0.113, and 0.115 kcal/mol for the 2-, 3-, 4-, and 5-mers, respectively. (c) Energies of our 1B model derived from CCSD(T) plotted against ωB97X-V/def2-qzvppd energies for 1200 snapshots taken from each cluster size, with MAEs of 0.11067, 0.1456, 0.17435, and 0.19496 kcal/mol for 2-, 3-, 4-, and 5-mers, respectively. The small cluster structures are taken from Wang et al.[11]

Table 2. Harmonic and Anharmonic Frequencies for Various QM and MM Models[a]

mode	experimental	CCSD(T)	ωB97x-V	PS	poly5z	1B-UCB (CC)	1B-UCB (DFT)
			Harmonic Frequencies (cm^{-1})				
bending	n/a	1650	1636	1650	1650	1650	1636
symmetric stretching	n/a	3835	3860	3833	3835	3835	3860
asymmetric stretching	n/a	3945	3961	3945	3945	3945	3961
			Anharmonic Frequencies (cm^{-1})				
bending	1595	n/a	n/a	1595	1583	1573	1557
symmetric stretching	3657	n/a	n/a	3657	3659	3654	3674
asymmetric stretching	3756	n/a	n/a	3756	3743	3740	3751

[a]We computed the anharmonic frequencies using a variational calculation, whose details we include in the Supporting Information and available in our recent paper on Raman spectroscopy.[42]

only a 1D bond and angle scan, as the dipole derivative matrix contains more information, such as the coupling between bonds and angles.

We also developed a novel way to reproduce dipole derivatives using virtual sites (VS). In the four- or six-site water models, such as the TIPnP[23] or CHARMM[21] series, there is an in-plane VS constructed from oxygen and its two neighbors

$$\overrightarrow{r_{VS}} = \overrightarrow{r_O} + w(\overrightarrow{r_{OH_1}} + \overrightarrow{r_{OH_2}}) \tag{7}$$

where w is the weight parameter determining the position of the VS. We modify the VS model by adding three charge flux terms that determine the position of the VS based on the deviations from the equilibrium bond lengths and angles.

$$\begin{aligned}\overrightarrow{r_{VS}} = \overrightarrow{r_O} + w(\overrightarrow{r_{OH_1}} + \overrightarrow{r_{OH_2}}) + j_b\overrightarrow{r_{OH_1}}(r_{OH_1} - r^0_{OH}) \\ + j_b\overrightarrow{r_{OH_2}}(r_{OH_2} - r^0_{OH}) + j_\theta(\overrightarrow{r_{OH_1}} + \overrightarrow{r_{OH_2}})(\theta - \theta_0) \\ + j_{bb}\overrightarrow{r_{OH_1}}(r_{OH_2} - r^0_{OH}) + j_{bb}\overrightarrow{r_{cb}}(r_{OH_1} - r^0_{OH})\end{aligned} \tag{8}$$

Because the VS has a partial charge, deviations from the equilibrium geometry lead to charge flux through the movement of the VS.

https://doi.org/10.1021/acs.jpclett.4c00587
J. Phys. Chem. Lett. 2024, 15, 6712−6721

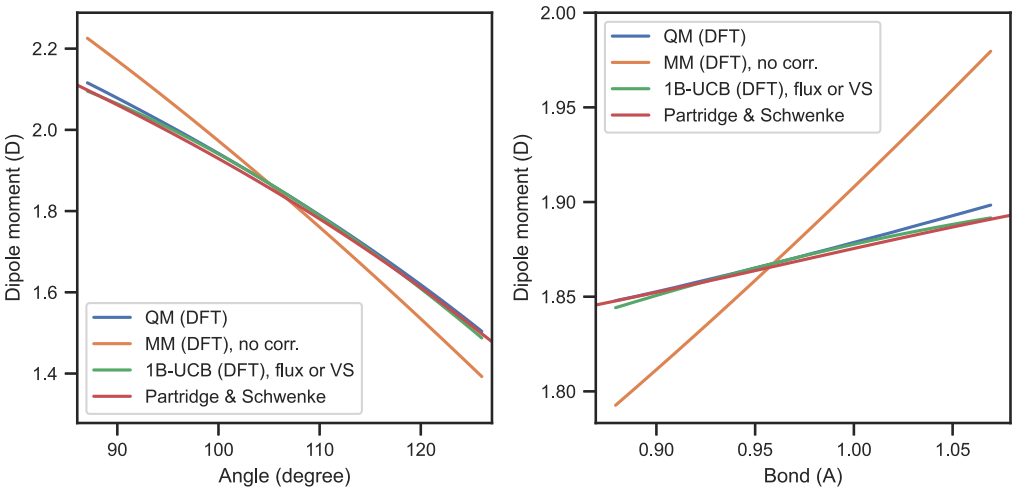


Figure 4. Change in the dipole moment over angle and bond distortions. (a) Angle and (b) bond lengths are scanned for QM (DFT), MM (DFT) with and without one of the DMS corrections, and PS model. Both the CF and VS approaches give almost identical results.

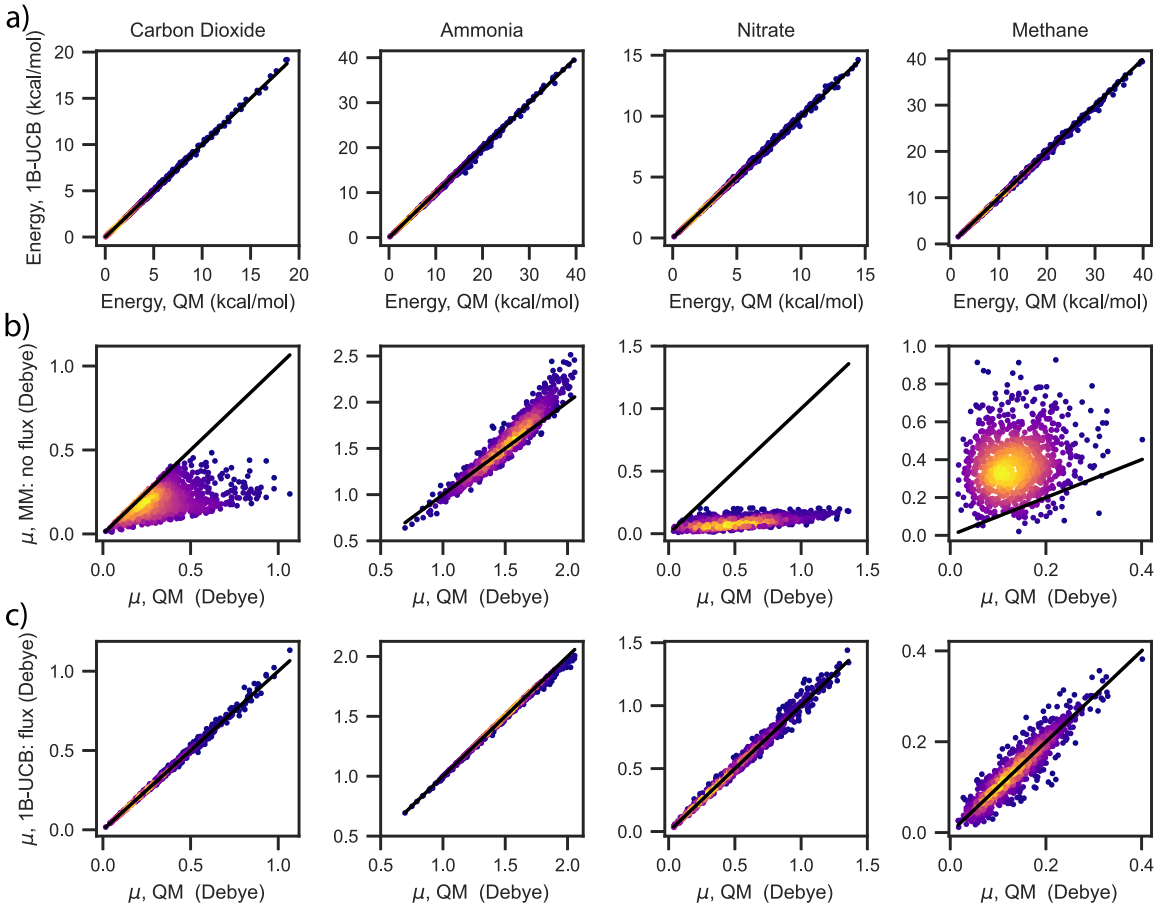


Figure 5. Comparison of 1B between QM and MM models for carbon dioxide, ammonia, nitrate, and methane molecules. (a) Potential energies and total dipole moments (b) without and (c) with charge flux. The QM method is ωB97x-V/def2-qzvppd.

However, unlike the 1B energy surface, the 1B DMS j parameters will depend upon the non-bonded parts of the FF, namely, the point charges, atomic dipoles, and virtual site position. Therefore, they must be reparameterized for the specific model (e.g., they depend upon the partial charges assigned to the H and O of a water model), which is a simple task achieved by fitting the MM dipole derivative matrix to the QM dipole derivative matrix provided in the Supporting Information. In this work, for both approaches, we have chosen to use point charges (and a w parameter in the VS model) to exactly reproduce the dipole moment of the optimized structure, as this is typically the case for both four- or six-site water models.

In Figure 4, the change in the dipole moment with respect to the angle and bond displacements is given for the QM reference, PS model, MM with no charge flux correction, and

https://doi.org/10.1021/acs.jpclett.4c00587
J. Phys. Chem. Lett. 2024, 15, 6712−6721

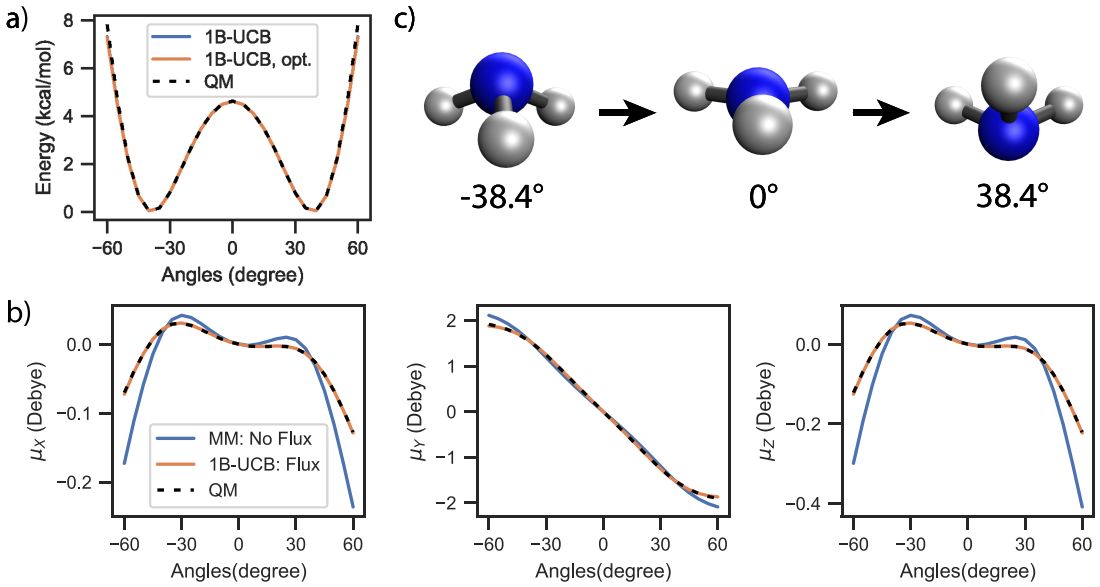


Figure 6. Ammonia pyramidal inversion for QM, MM, and 1B-UCB model. (a) Energies and (b) each dipole moment component. (c) Geometry at the minima and saddle point. *Y* axis is orthogonal to the plane of three hydrogens. *X* and *Z* are parallel to the plane of the three hydrogens, with *Z* pointing toward one of the hydrogens and *X* being orthogonal to *Z*.

MM with either the charge flux (eq 6) or VS correction (eqs 7 and 8). We see that the dipole moment is evidently much more sensitive to the angle than bond displacements. Without any charge flux correction, changes to the dipole moment are significantly overestimated [root mean square error (RMSE): angles, 0.43 D; bonds, 0.19 D], whereas with either charge flux or VS correction, there is virtually an exact match at small displacement, with small deviations at larger displacements (RMSE: angles, 0.056 D; bonds, 0.012 D). The PS dipole moments are also very similar to the QM benchmark, but our MM model is much simpler and more generalizable.

To demonstrate the applicability of this protocol and model to other molecules, we have parametrized four other molecules with different symmetries: carbon dioxide (linear), ammonia (trigonal pyramid), nitrate (trigonal planar), and methane (tetrahedral). A few other considerations were needed for these molecules, which are now added to the protocol:

- For linear three-atom molecules (carbon dioxide), bond−angle coupling terms were unnecessary. Additionally, a harmonic angle potential was used instead of a cosine angle as a result of the equilibrium angle of 180°.
- For four- and five-atom molecules, an angle−angle coupling term was needed.

$$V_{\text{angle-angle}}(\theta_1, \theta_2) = k_{\text{aa}}(\cos\theta_1 - \cos\theta_1^0)(\cos\theta_2 - \cos\theta_2^0) \quad (9)$$

- For trigonal planar molecules (nitrate), an improper dihedral term was needed, and a harmonic potential proved suitable for this purpose.

$$V_{\text{improper}}(\phi_1) = k_{\text{improper}}(\phi - \phi^0)^2 \quad (10)$$

As scanning the degrees of freedom, especially angles, becomes more challenging with four- and five-atom molecules, for the validation data set, we have instead opted to perform Wigner sampling using the Wigner module in the SHARC program[44] to generate 1000 distorted structures for each molecule. In Figure 5a, the QM and 1B-UCB energies are shown for carbon dioxide, ammonia, nitrate, and methane. A good match between QM and 1B-UCB energies can be seen for all molecules, with MAEs of 0.04, 0.15, 0.06, and 0.19 kcal/mol, respectively. Note that the higher MAEs of ammonia and methane are not because these models are less accurate but because their light hydrogen atoms cause Wigner sampling to distort them to higher energy structures. In Figure 5b, QM and MM total dipole moments without any dipole corrections are shown. It can be seen that, without any correction, the dipole moments are highly inaccurate, with MAEs of 0.071, 0.099, 0.20, and 0.17 D, respectively. With the use of the charge flux approach, as shown in Figure 5c, the accuracy significantly improves, with the MAEs reducing by about an order of magnitude to values of 0.003, 0.012, 0.013, and 0.019 D, respectively.

Another demonstration of the success of Hessian fitting outside the near-equilibrium distortions is ammonia pyramidal inversion (Figure 6c). As shown in Figure 6a, both the barrier height and the forces throughout the scan (shown by obtaining the same geometry and energy after an MM geometry optimization) are captured by fitting only to Hessian, with MAEs in energies of 0.068 and 0.072 kcal/mol with and without optimization, respectively. Similarly, as shown in Figure 6b, each component of the dipole moment is significantly improved by the charge flux approach and closely matches the QM dipole moments, with MAEs of 0.0004, 0.0077, and 0.0007 D for *X*, *Y*, and *Z* directions (defined in the figure caption), respectively.

In summary, we developed the 1B-UCB model, defined by a well-defined protocol to generate an accurate 1B energy and DMS for water and other small molecules. By keeping the model relatively simple (a total of 10 parameters for water: 2 for equilibrium bond lengths and angles, 4 for the energy surface, and 3 for the DMS) and through deliberate choices of functional forms, we have demonstrated a simple parametrization strategy, where the bonded parameters are obtained by fitting MM Hessian to QM Hessian and the

https://doi.org/10.1021/acs.jpclett.4c00587
J. Phys. Chem. Lett. 2024, 15, 6712−6721

DMS is obtained by fitting the MM dipole derivatives to the QM dipole derivatives, both of which are readily available from a single QM Hessian calculation. We have used two different approaches for the DMS: the charge flux method and a novel approach taking advantage of the virtual interaction site common to many molecular mechanics models. Both approaches perform equally well; therefore, the VS approach should be appropriate for models that already have a VS, while the charge flux approach is appropriate for those that do not.

As a favorite testing ground for new force fields, the 1B-UCB model for water performs competitively with much more complex 1B water models in reproducing reference QM calculations. However, while complexity is rarely held back when it comes to advanced water models, the simplicity of the 1B-UCB model, combined with the automated parametrization protocols using the Q-Force toolkit,[40] makes it much more feasible to parametrize larger numbers and sizes of molecules effortlessly. To demonstrate this point, we have parametrized, using these automated protocols, four additional molecules with different symmetries and shown that the accuracy of both energies and dipoles are consistent among these molecules. The key advantage of our approach is combining accuracy, simplicity, and automation, and we are currently working to leverage this advantage to biologically relevant molecules while tackling the additional challenges of torsions and intramolecular non-bonded interactions.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpclett.4c00587.

Dipole moment surface parameters and parametrization strategy and the Taylor expansion of the bonded potentials (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Selim Sami − *Kenneth S. Pitzer Theory Center and Department of Chemistry, University of California, Berkeley, Berkeley, California 94720, United States;* orcid.org/0000-0002-4484-0322; Email: s.sami@berkeley.edu

Teresa Head-Gordon − *Kenneth S. Pitzer Theory Center and Department of Chemistry, University of California, Berkeley, Berkeley, California 94720, United States; Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States; Departments of Bioengineering and Chemical and Biomolecular Engineering, University of California, Berkeley, Berkeley, California 94720, United States;* orcid.org/0000-0003-0025-8987; Email: thg@berkeley.edu

Authors

R. Allen LaCour − *Kenneth S. Pitzer Theory Center and Department of Chemistry, University of California, Berkeley, Berkeley, California 94720, United States; Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States*

Joseph P. Heindel − *Kenneth S. Pitzer Theory Center and Department of Chemistry, University of California, Berkeley, Berkeley, California 94720, United States; Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States;* orcid.org/0000-0002-9748-1730

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpclett.4c00587

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Selim Sami thanks the Dutch Research Council (NWO) Rubicon Grant (019.212EN.004) for fellowship support. Teresa Head-Gordon acknowledges support from the U.S. National Science Foundation through Grant CHE-2313791. Computational resources were provided by the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility operated under Contract DE-AC02-05CH11231.

■ REFERENCES

(1) Demerdash, O.; Yap, E. H.; Head-Gordon, T. Advanced Potential Energy Surfaces for Condensed Phase Simulation. *Annu. Rev. Phys. Chem.* **2014**, *65*, 149−174.

(2) Demerdash, O.; Wang, L.-P.; Head-Gordon, T. Advanced models for water simulations. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2018**, *8*, e1355.

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

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

(5) 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.

(6) Yu, Q.; Qu, C.; Houston, P. L.; Conte, R.; Nandi, A.; Bowman, J. M. q-AQUA: A Many-Body CCSD(T) Water Potential, Including Four-Body Interactions, Demonstrates the Quantum Nature of Water from Clusters to the Liquid Phase. J. Phys. *Chem. Lett.* **2022**, *13*, 5068−5074.

(7) Qu, C.; Yu, Q.; Houston, P. L.; Conte, R.; Nandi, A.; Bowman, J. M. Interfacing q-AQUA with a Polarizable Force Field: The Best of Both Worlds. *J. Chem. Theory Comput.* **2023**, *19*, 3446−3459.

(8) Yu, Q.; Qu, C.; Houston, P. L.; Nandi, A.; Pandey, P.; Conte, R.; Bowman, J. M. A Status Report on Gold Standard Machine-Learned Potentials for Water. J. Phys. *Chem. Lett.* **2023**, *14*, 8077−8087.

(9) Fanourgakis, G. S.; Xantheas, S. S. The Flexible, Polarizable, Thole-Type Interaction Potential for Water (TTM2-F) Revisited. *J. Phys. Chem. A* **2006**, *110*, 4100−4106.

(10) Fanourgakis, G. S.; Xantheas, S. S. Development of transferable interaction potentials for water. V. Extension of the flexible, polarizable, Thole-type model potential (TTM3-F, v. 3.0) to describe the vibrational spectra of water clusters and liquid water. *J. Chem. Phys.* **2008**, *128*, 074506.

(11) Wang, L.-P.; Head-Gordon, T.; Ponder, J. W.; Ren, P.; Chodera, J. D.; Eastman, P. K.; Martinez, T. J.; Pande, V. S. Systematic Improvement of a Classical Molecular Model of Water. *J. Phys. Chem. B* **2013**, *117*, 9956−9972.

(12) Laury, M. L.; Wang, L.-P.; Pande, V. S.; Head-Gordon, T.; Ponder, J. W. Revised Parameters for the AMOEBA Polarizable Atomic Multipole Water Model. *J. Phys. Chem. B* **2015**, *119*, 9423−9437.

(13) Liu, C.; Piquemal, J.-P.; Ren, P. Implementation of Geometry-Dependent Charge Flux into the Polarizable AMOEBA+ Potential. J. Phys. *Chem. Lett.* **2020**, *11*, 419−426.

(14) Liu, C.; Piquemal, J.-P.; Ren, P. AMOEBA+ Classical Potential for Modeling Molecular Interactions. *J. Chem. Theory Comput.* **2019**, *15*, 4122−4139.
(15) Rackers, J. A.; Silva, R. R.; Wang, Z.; Ponder, J. W. Polarizable Water Potential Derived from a Model Electron Density. *J. Chem. Theory Comput.* **2021**, *17*, 7056−7084.
(16) Chung, M. K. J.; Wang, Z.; Rackers, J. A.; Ponder, J. W. Classical Exchange Polarization: An Anisotropic Variable Polarizability Model. *J. Phys. Chem. B* **2022**, *126*, 7579−7594.
(17) Symons, B. C. B.; Popelier, P. L. A. Application of Quantum Chemical Topology Force Field FFLUX to Condensed Matter Simulations: Liquid Water. *J. Chem. Theory Comput.* **2022**, *18*, 5577−5588.
(18) 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**, *15*, 5001−5013.
(19) 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.
(20) Dateo, C. E.; Lee, T. J.; Schwenke, D. W. An accurate quartic force field and vibrational frequencies for HNO and DNO. *J. Chem. Phys.* **1994**, *101*, 5853−5859.
(21) Mackerell, A. D., Jr. Empirical force fields for biological macromolecules: Overview and issues. *J. Comput. Chem.* **2004**, *25*, 1584−1604.
(22) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **1983**, *4*, 187−217.
(23) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225−11236.
(24) Scott, W. R. P.; Hünenberger, P. H.; Tironi, I. G.; Mark, A. E.; Billeter, S. R.; Fennen, J.; Torda, A. E.; Huber, T.; Krüger, P.; van Gunsteren, W. F. The GROMOS Biomolecular Simulation Program Package. *J. Phys. Chem. A* **1999**, *103*, 3596−3607.
(25) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25*, 1157−1174.
(26) Allinger, N. L.; Yuh, Y. H.; Lii, J. H. Molecular mechanics. The MM3 force field for hydrocarbons. 1. *J. Am. Chem. Soc.* **1989**, *111*, 8551−8566.
(27) Habershon, S.; Markland, T. E.; Manolopoulos, D. E. Competing quantum effects in the dynamics of a flexible water model. *J. Chem. Phys.* **2009**, *131*, 024501.
(28) Lodi, L.; Tennyson, J.; Polyansky, O. L. A global, high accuracy *ab initio* dipole moment surface for the electronic ground state of the water molecule. *J. Chem. Phys.* **2011**, *135*, 034113.
(29) Davie, S. J.; Di Pasquale, N.; Popelier, P. L. A. Incorporation of local structure into kriging models for the prediction of atomistic properties in the water decamer. *J. Comput. Chem.* **2016**, *37*, 2409−2422.
(30) Konovalov, A.; Symons, B. C.; Popelier, P. L. On the many-body nature of intramolecular forces in FFLUX and its implications. *J. Comput. Chem.* **2021**, *42*, 107−116.
(31) Dinur, U.; Hagler, A. T. Geometry-dependent atomic charges: Methodology and application to alkanes, aldehydes, ketones, and amides. *J. Comput. Chem.* **1995**, *16*, 154−170.
(32) Yang, X.; Liu, C.; Walker, B. D.; Ren, P. Accurate description of molecular dipole surface with charge flux implemented for molecular mechanics. *J. Chem. Phys.* **2020**, *153*, 064103.
(33) Mardirossian, N.; Head-Gordon, M. ωB97X-V: A 10-parameter, range-separated hybrid, generalized gradient approximation density functional with nonlocal correlation, designed by a survival-of-the-fittest strategy. *Phys. Chem. Chem. Phys.* **2014**, *16*, 9904−9924.
(34) Weigend, F. Accurate Coulomb-fitting basis sets for H to Rn. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1057−1065.
(35) 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.
(36) Mardirossian, N.; Head-Gordon, M. Thirty years of density functional theory in computational chemistry: An overview and extensive assessment of 200 density functionals. *Mol. Phys.* **2017**, *115*, 2315−2372.
(37) Goerigk, L.; Hansen, A.; Bauer, C.; Ehrlich, S.; Najibi, A.; Grimme, S. A look at the density functional theory zoo with the advanced GMTKN55 database for general main group thermochemistry, kinetics and noncovalent interactions. *Phys. Chem. Chem. Phys.* **2017**, *19*, 32184−32215.
(38) Epifanovsky, E.; Gilbert, A. T. B.; Feng, X.; Lee, J.; Mao, Y.; Mardirossian, N.; Pokhilko, P.; White, A. F.; Coons, M. P.; Dempwolff, A. L.; Gan, Z.; Hait, D.; Horn, P. R.; Jacobson, L. D.; Kaliman, I.; Kussmann, J.; Lange, A. W.; Lao, K. U.; Levine, D. S.; Liu, J.; McKenzie, S. C.; Morrison, A. F.; Nanda, K. D.; Plasser, F.; Rehn, D. R.; Vidal, M. L.; You, Z.-Q.; Zhu, Y.; Alam, B.; Albrecht, B. J.; Aldossary, A.; Alguire, E.; Andersen, J. H.; Athavale, V.; Barton, D.; Begam, K.; Behn, A.; Bellonzi, N.; Bernard, Y. A.; Berquist, E. J.; Burton, H. G. A.; Carreras, A.; Carter-Fenk, K.; Chakraborty, R.; Chien, A. D.; Closser, K. D.; Cofer-Shabica, V.; Dasgupta, S.; de Wergifosse, M.; Deng, J.; Diedenhofen, M.; Do, H.; Ehlert, S.; Fang, P.-T.; Fatehi, S.; Feng, Q.; Friedhoff, T.; Gayvert, J.; Ge, Q.; Gidofalvi, G.; Goldey, M.; Gomes, J.; González-Espinoza, C. E.; Gulania, S.; Gunina, A. O.; Hanson-Heine, M. W. D.; Harbach, P. H. P.; Hauser, A.; Herbst, M. F.; Hernández Vera, M.; Hodecker, M.; Holden, Z. C.; Houck, S.; Huang, X.; Hui, K.; Huynh, B. C.; Ivanov, M.; Jász, A.; Ji, H.; Jiang, H.; Kaduk, B.; Kähler, S.; Khistyaev, K.; Kim, J.; Kis, G.; Klunzinger, P.; Koczor-Benda, Z.; Koh, J. H.; Kosenkov, D.; Koulias, L.; Kowalczyk, T.; Krauter, C. M.; Kue, K.; Kunitsa, A.; Kus, T.; Ladjánszki, I.; Landau, A.; Lawler, K. V.; Lefrancois, D.; Lehtola, S.; Li, R. R.; Li, Y.-P.; Liang, J.; Liebenthal, M.; Lin, H.-H.; Lin, Y.-S.; Liu, F.; Liu, K.-Y.; Loipersberger, M.; Luenser, A.; Manjanath, A.; Manohar, P.; Mansoor, E.; Manzer, S. F.; Mao, S.-P.; Marenich, A. V.; Markovich, T.; Mason, S.; Maurer, S. A.; McLaughlin, P. F.; Menger, M. F. S. J.; Mewes, J.-M.; Mewes, S. A.; Morgante, P.; Mullinax, J. W.; Oosterbaan, K. J.; Paran, G.; Paul, A. C.; Paul, S. K.; Pavošević, F.; Pei, Z.; Prager, S.; Proynov, E. I.; Rák, A.; Ramos-Cordoba, E.; Rana, B.; Rask, A. E.; Rettig, A.; Richard, R. M.; Rob, F.; Rossomme, E.; Scheele, T.; Scheurer, M.; Schneider, M.; Sergueev, N.; Sharada, S. M.; Skomorowski, W.; Small, D. W.; Stein, C. J.; Su, Y.-C.; Sundstrom, E. J.; Tao, Z.; Thirman, J.; Tornai, G. J.; Tsuchimochi, T.; Tubman, N. M.; Veccham, S. P.; Vydrov, O.; Wenzel, J.; Witte, J.; Yamada, A.; Yao, K.; Yeganeh, S.; Yost, S. R.; Zech, A.; Zhang, I. Y.; Zhang, X.; Zhang, Y.; Zuev, D.; Aspuru-Guzik, A.; Bell, A. T.; Besley, N. A.; Bravaya, K. B.; Brooks, B. R.; Casanova, D.; Chai, J.-D.; Coriani, S.; Cramer, C. J.; Cserey, G.; DePrince, A. E., III; DiStasio, R. A., Jr.; Dreuw, A.; Dunietz, B. D.; Furlani, T. R.; Goddard, W. A., III; Hammes-Schiffer, S.; Head-Gordon, T.; Hehre, W. J.; Hsu, C.-P.; Jagau, T.-C.; Jung, Y.; Klamt, A.; Kong, J.; Lambrecht, D. S.; Liang, W.; Mayhall, N. J.; McCurdy, C. W.; Neaton, J. B.; Ochsenfeld, C.; Parkhill, J. A.; Peverati, R.; Rassolov, V. A.; Shao, Y.; Slipchenko, L. V.; Stauch, T.; Steele, R. P.; Subotnik, J. E.; Thom, A. J. W.; Tkatchenko, A.; Truhlar, D. G.; Van Voorhis, T.; Wesolowski, T. A.; Whaley, K. B.; Woodcock, H. L.; Zimmerman, P. M.; Faraji, S.; Gill, P. M. W.; Head-Gordon, M.; Herbert, J. M.; Krylov, A. I. Software for the frontiers of quantum chemistry: An overview of developments in the Q-Chem 5 package. *J. Chem. Phys.* **2021**, *155*, 084801.
(39) Eastman, P.; Swails, J.; Chodera, J. D.; McGibbon, R. T.; Zhao, Y.; Beauchamp, K. A.; Wang, L.-P.; Simmonett, A. C.; Harrigan, M. P.; Stern, C. D.; Wiewiora, R. P.; Brooks, B. R.; Pande, V. S. OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. *PLoS Comput. Biol.* **2017**, *13*, e1005659.

https://doi.org/10.1021/acs.jpclett.4c00587
J. Phys. Chem. Lett. 2024, 15, 6712−6721

(40) Sami, S.; Menger, M. F. S. J.; Faraji, S.; Broer, R.; Havenith, R. W. A. Q-Force: Quantum Mechanically Augmented Molecular Force Fields. *J. Chem. Theory Comput.* **2021**, *17*, 4946.
(41) Herman, K. M.; Xantheas, S. S. An extensive assessment of the performance of pairwise and many-body interaction potentials in reproducing *ab initio* benchmark binding energies for water clusters n = 2−25. *Phys. Chem. Chem. Phys.* **2023**, *25*, 7120−7143.
(42) LaCour, R. A.; Heindel, J. P.; Head-Gordon, T. Predicting the Raman Spectra of Liquid Water with a Monomer-Field Model. *J. Phys. Chem. Lett.* **2023**, *14*, 11742−11749.
(43) Laury, M. L.; Boesch, S. E.; Haken, I.; Sinha, P.; Wheeler, R. A.; Wilson, A. K. Harmonic vibrational frequencies: Scale factors for pure, hybrid, hybrid meta, and double-hybrid functionals in conjunction with correlation consistent basis sets. *J. Comput. Chem.* **2011**, *32*, 2339−2347.
(44) Richter, M.; Marquetand, P.; González-Vázquez, J.; Sola, I.; González, L. SHARC: Ab Initio Molecular Dynamics with Surface Hopping in the Adiabatic Representation Including Arbitrary Couplings. *J. Chem. Theory Comput.* **2011**, *7*, 1253−1258.

https://doi.org/10.1021/acs.jpclett.4c00587
J. Phys. Chem. Lett. 2024, 15, 6712−6721