ACS PHYSICAL CHEMISTRY Au
AN OPEN ACCESS JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/physchemau

Article

Performing Molecular Dynamics Simulations and Computing Hydration Free Energies on the B3LYP-D3(BJ) Potential Energy Surface with Adaptive Force Matching: A Benchmark Study with Seven Alcohols and One Amine

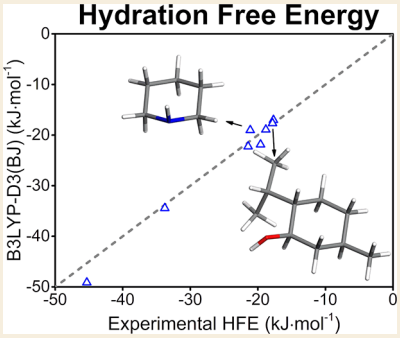
Dong Zheng and Feng Wang*

Cite This: *ACS Phys. Chem Au* 2021, 1, 14–24

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The potential energy surfaces at the B3LYP-D3(BJ) level for eight solutes in dilute aqueous solutions were mapped into simple pairwise additive force field expressions using the adaptive force matching (AFM) method. The quality of the fits was validated by computing the hydration free energy (HFE), enthalpy of hydration, and diffusion constant for each solute. By force matching B3LYP-D3(BJ), the predictions from the models agree with the closest experimental HFE and enthalpy of hydration within chemical accuracy. The diffusion constants from the models are also in good agreement with experimental references. The good agreement provides confidence on the quality of B3LYP-D3(BJ) in producing potential energy surfaces for thermodynamic property calculations through AFM for the molecules studied. Accurate computational predictions could potentially provide validations to experimental measurements in cases where experimental measurements from different sources do not agree.



KEYWORDS: *Hydration free energy, Adaptive force matching, Molecular dynamics, Density functional theory*

I. INTRODUCTION

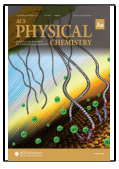
While advances in electronic structure theory have enabled the computation of highly accurate energies, conformational space sampling with electronic structure methods is often challenging as a result of the high computational cost associated with such methods. Electronic structure based property computations are frequently performed with minimum energy conformations instead of averaging over proper ensembles. This limitation could lead to problems for certain properties. An example demonstrating such a limitation is liquid water. Geometry optimizations would lead to conformations that maximize the number of hydrogen bonds, leading to ice-like conformations as local minima. If properties were only computed on such ice-like local minima, important properties of the liquid could be missed, since spontaneous breaking of hydrogen bonds happens on a regular basis in the liquid phase.[1−3]

One way to get ensemble properties based on electronic structure methods is to performed so-called *ab initio* molecular dynamics (MD)[4−6] simulation. Depending on how much electron density is allowed to deviate from the Born−Oppenheimer solution and what type of electronic structure method to use,[7−10] *ab initio* MD encompasses a family of approaches. Although it is debated whether density functional theory (DFT) should be referred to as *ab initio*, DFT based MD is generally considered as one example of *ab initio* MD. In fact, most *ab initio* MD is based on DFT due to its ability to achieve good accuracy at a relatively low computational cost.

Ab initio MD has many strengths, such as an accurate description of many body effects and the easiness to model reactivity. However, despite recent progress with linear scaling algorithms[11,12] and the rapid increase of computational power, performing *ab initio* MD with thousands of atoms at nanosecond time scale is challenging.

One alterative way to perform MD on an *ab initio* potential energy surface (PES) is through fitting such a PES to a force field model. Such a fitting can be accomplished with either energy matching or force matching.[13−20] Electronic structure based force field development has a long history,[21,22] although such approaches have rarely been considered as a way to perform *ab initio* MD. This is at least partly due to the need to add empirical parameters to early *ab initio* based potentials. Without empirical adjustments, such potentials tend to provide rather poor agreement with experiments. Once empirical parameters are introduced, the goal of such a potential will no longer be to reproduce the *ab initio* PES. If a potential has reproducing the *ab initio* PES as its sole objective, we feel

Received: April 29, 2021
Published: July 21, 2021



https://doi.org/10.1021/acsphyschemau.1c00006
ACS Phys. Chem Au 2021, 1, 14–24

appropriate to consider such an approach as an indirect way to perform *ab initio* MD.

For MD on *ab initio* PES, some noteworthy approaches[23−26] in recent years include advanced potentials that are fit to energy decompositions based on symmetry adapted perturbation theory (SAPT).[27] Another noteworthy direction is to fit the potential under the framework of force matching.[13−15] As an force matching based approach, the adaptive force matching (AFM) method has showed some good success.[28,29] It has been shown that AFM derived potentials give good predictions of hydration free energies (HFEs) for simple salts[30] and small neutral solutes.[31,32] With PBE/D3 as a reference, an AFM based potential predicted the conformation distribution for hydrated alanine in excellent agreement with experimental NMR scalar coupling constants.[33]

With AFM, an energy expression is fitted to condensed phase reference forces computed with an electronic structure method. Many-body effects are typically captured implicitly in favor of minimizing the computational cost by utilizing a simple pairwise additive energy expression. A pairwise potential fitted through AFM will not be as accurate as a direct *ab initio* MD. However, with a simple molecular mechanics energy expression, it is much easier to address systems where finite size effects are significant and to compute properties where long time fluctuation needs be averaged out. Another advantage of a molecular mechanistic potential is the ability to do alchemical transformations.[34,35] Such an ability allows the construction of efficient thermodynamic pathways to compute free energy differences.

In this work, we demonstrate the use of B3LYP to create force field models with AFM and validate such force fields by computing the HFEs for eight different molecules, ethanol, isobutanol, 2-butanol, 1,2-butanediol, 1,4-butanediol, 1-hexanol, menthol, and piperidine. For some of these molecules, force fields based on MP2 and local MP2 (LMP2) have been developed.[32] However, LMP2 with projected atomic orbitals (PAO)[36] has been shown to be provide unsatisfactory accuracy for alcohols.[32] Seven of the eight molecules in this study are alcohols. An exchange-correlation functional with dispersion-correction is selected to match the CCSD(T)[37−40] PES for ethanol.

Our work showed that models for simple alcohols can be reliably created with B3LYP-D3(BJ), which overcomes the deficiency shown previously with LMP2.[32] Identifying a good DFT functional for computing thermodynamic properties is valuable, as such a functional would allow larger molecules to be studied in the framework of AFM. Of the eight molecules studied, piperidine is an amine. Experimental determination of amine HFE is complicated by the need to account for hydrolysis. Menthol is a widely used drug and is the largest molecule of this group.

Our approach bears some resemblance to quantum mechanics/molecular mechanics (QM/MM) based free energy calculations.[41−43] However, in typical QM/MM based HFE calculations, only the solute is treated with QM. The fitting step in our approach treats both the solute and water with QM. Thus, the model captures the PES of QM solute in QM solvent. A typical free energy perturbation[44−49] (FEP) based QM/MM HFE calculation performs sampling only with a force field that is not iteratively improved. If the sampling is less than satisfactory, the Zwanzig based FEP[50] will be slow to converge. AFM relies on a water model that is fitted to electronic structure calculations at an appropriate quality. The solute−water cross terms are iteratively improved during the AFM iterations based on QM data. Our approach thus enables the sampling of the hydration structure also at the quality of the reference method.

The paper is organized in four sections, after the introduction, a brief description of the fitting protocol is provided in Section II. The computational details for property calculations are provided in Section III, the results and discussion in Section IV, and a summary and conclusion in Section V.

II. SPECIFIC DETAILS OF THE FITTING PROTOCOL FOR THE DEVELOPMENT OF THE SOLUTE−WATER MODELS

In this work, only the solute−water interactions will be fitted with AFM. The BLYPSP-4F water model, which was also developed with AFM,[29] will be used to model hydration water. This model was fitted to a coupled-cluster quality PES for liquid water obtained using the DFT with supplemental potential approach.[51,52] It has been shown that the BLYPSP-4F model gives good water properties, such as diffusion constant, surface tension, dielectric constant, heat of vaporization, etc.[29,53,54] The atom typing for the solutes is shown in Figure 1. The atom types are constructed with the name of the element followed by a number. In principle, we prefer to give each symmetry unique atom a different atom type. For 1-hexanol and menthol, atoms with very similar chemical environments are allowed to share the same atom type. For example, in 1-hexanol, the α and β carbons of the hydroxyl have different types, and subsequent secondary carbons share

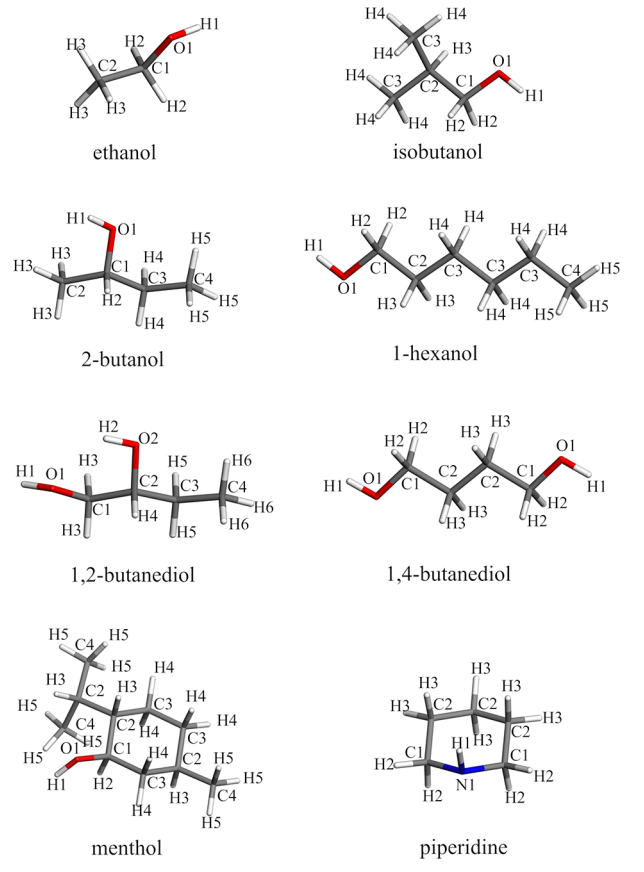


Figure 1. Atom type definitions for the solutes investigated.

https://doi.org/10.1021/acsphyschemau.1c00006
ACS Phys. Chem Au 2021, 1, 14−24

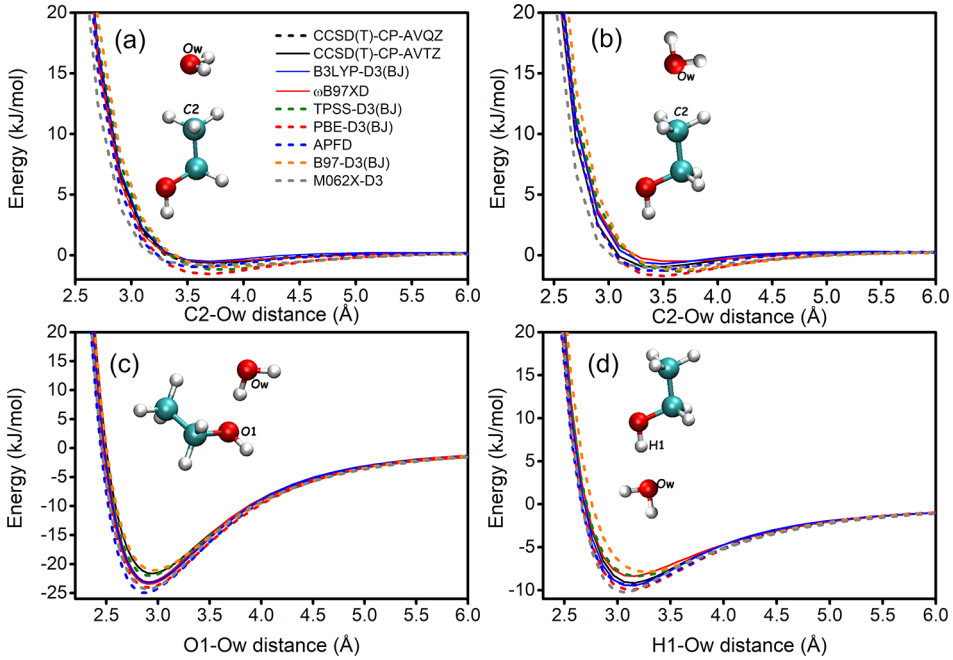


Figure 2. Potential energy scans for the ethanol−water dimer with four different orientations around the methyl and hydroxyl group of ethanol. The scans were performed with Gaussian 16.[129] All DFT calculations were performed with the aug-cc-pVTZ basis set. The CCSD(T) calculations were performed with counter-poise correction, aug-cc-pVTZ(AVTZ), and aug-cc-pVQZ(AVQZ). Note that the aug-cc-pVQZ CCSD(T) calculation is only performed for orientation A.

the same atom type. We also note that the parameters are specific to each molecule. In other words, the same atom type in different molecules will not have the same parameter.

A typical AFM iteration contains three steps: the MD step, the QM/MM step, and the force matching (FM) step. The MD step samples conformations to be used for FM, and the QM/MM step performs single point force calculations using the conformations sampled in the MD step. We note that although forces are computed, they are only used for fitting a force field and are not used to propagate the equations of motion. The FM step refits the force field to best reproduce the QM forces obtained in the QM/MM step.

In this study, the MD step was performed with a single solute in a box of 343 waters at 298 and 328 K as done previously for other solute models. Although we design the force field to work under the ambient temperature, simulations at the slightly higher temperature improve the sampling of the repulsive wall of the potential and are expected to improve the convergence of AFM, especially when the training set is small.

In the QM/MM step, the system is divided into a QM region and an MM region. The MM region is represented by partial charges. The QM region is further divided into a central region and a buffer region. Only forces on the central region atoms are fitted. The forces on the buffer region atoms are discarded to remove potential boundary effects when electrons in the QM region spill over to partial charges in the MM region. The protocol for selecting the QM/MM region is the same as that used in our previous work for developing MP2 and LMP2 based solute models and is thus only provided in the Supporting Information.

In the QM/MM step, DFT forces are computed with the B3LYP exchange correlation functional.[55−57] The aug-cc-pVTZ basis set was used for heavy atoms and the cc-pVTZ basis set for hydrogens. The removal of diffuse augmented functions on hydrogens reduces basis set linear dependency. The long-range dispersion is modeled by Grimme's D3 approach[58] with Becke−Johnson (BJ) damping.[59,60]

Figure 2 shows potential energy surface scans of ethanol−water dimers around the methyl and hydroxyl groups of ethanol. In orientation A, where a water approaches the methyl group of ethanol on its side, reference calculations were performed with CCSD(T) using both the aug-cc-pVTZ[61] and aug-cc-pVQZ basis sets[62] with the counterpoise correction (CP).[63] At least for this orientation, the difference between aug-cc-pVTZ and aug-cc-pVQZ is very small, probably as a result of the counterpoise correction. For other orientations, only CCSD(T) with aug-cc-pVTZ and counterpoise correction was used as reference. Of all the functionals tested, B3LYP-D3(BJ) seems to provide the best agreement with CCSD(T)-CP with ωB97XD[64] being a close second. We note that no counter-poise correction was used for any of the DFT calculations. Based on this finding, B3LYP-D3(BJ) is used to provide reference forces for this study.

A three step fitting procedure is performed as in our previous study.[32] Only the second and third step fittings were done iteratively with AFM. The objective of the first step is to determine the dispersion parameters, which will be held constant in the subsequent fits using AFM.

Proper damping of dispersion is important especially when the short-range repulsion is modeled with an exponential term that does not approach infinity at short-range. In our models, dispersion will be modeled with the short-range damped form,[33]

$$V_{disp}(r_{ij}) = \frac{C_6}{r_0^6 + r_{ij}^6} \quad (1)$$

where the r_0 is 0.6 times the sum of the atomic van der Waals radius taken from the work of Tkatchenko.[65] The dispersion terms are only placed between heavy atoms. The intermolecular dispersion is fitted to the D3(BJ) dispersion energy

https://doi.org/10.1021/acsphyschemau.1c00006
ACS Phys. Chem Au 2021, 1, 14−24

gradient between a solute and a water. The intramolecular dispersion is fitted with two isolated solutes without water. The two solute molecules were sampled with an MD simulation in an aqueous solution with Optimized Potentials for Liquid Simulations All Atom (OPLS-AA).[66,67] All water molecules are discarded for the fitting of intramolecular dispersion.

The first fitting step only fits the C_6 parameters in eq 1 to D3(BJ) dispersion. After the dispersion parameters are determined, other intermolecular parameters, such as partial charges and short-range repulsion terms, are fitted in AFM iterations to reproduce total forces and total torques of each molecule using the B3LYP-D3(BJ) reference. The short-range repulsion is modeled using a simple exponential form,

$$U_{rep}(r_{ij}) = A \exp(-r_{ij}) \tag{2}$$

where r_{ij} is the distance between atoms i on the solute and atom j on a water. A repulsion term is placed between every heavy atom of the solute and the water oxygen. In addition, every pair of atoms with opposite charges will also have a repulsion term between them.

After the intermolecular parameters are determined, the intramolecular parameters were fitted to reproduce B3LYP-D3(BJ) atomic forces. The intramolecular Coulombic and dispersion terms had been fitted in the previous steps. The intramolecular parameters to be determined in this step include bond, angle, torsional, and short-range nonbonded repulsion terms. The intramolecular nonbonded interactions are placed between all atom pairs separated by three or more covalent bonds. While the α for intermolecular repulsion is fitted, the α for intramolecular repulsion is fixed to 3.6 $Å^{-1}$. Fixing α for intramolecular repulsion will lead to a small reduction in accuracy but significantly reduces the number of nonlinear parameters to be determined. Our results do not seem to indicate any problem associated with using a fixed value for the intramolecular α parameters.

Overall, the fitting protocols are essentially the same as we did previously for several molecules with MP2 and LMP2 as reference methods. The only major difference is to use B3LYP-D3(BJ) to provide reference forces. To be consistent with the B3LYP-D3(BJ) reference, the dispersion in this work was fitted to D3(BJ) dispersion rather than to the E2 dispersion from SAPT.[32]

III. COMPUTATIONAL DETAILS FOR PROPERTY CALCULATIONS

The HFEs were computed with an alchemical path using the Bennett acceptance ratio (BAR) method[68,69] as implemented in Gromacs. It is worth noting that by default Gromacs will automatically remove gas phase contributions of the solute with the couple-intramol keyword set to no. This option has severe limitations when used with tabulated potentials. Our force fields used different tabulated potentials for each pair and are not compatible with the default implementation in Gromacs. We perform alchemical integrations in both the solution and the gas phase with the couple-intramol keyword set to yes. The HFE is computed as the difference between the two BAR free energies.

A soft-core potential is used to avoid numerical issues when particles overlap.[70] The soft-core parameters were chosen with α and σ being 1 and 0.3 nm, respectively. A total of 21 alchemical windows were used with 11 windows for removing Coulombic interactions followed by 10 windows for removing the short-range nonbonded interactions. For each window, a total of 5 ns of simulation was performed with stochastic Langevin dynamics in the NPT ensemble at 298 K and 1 bar with a 0.5 fs time step. For solution phase simulations, the pressure was enforced with the Parrinello−Rahman[71,72] barostat with a relaxation constant of 5 ps. The gas phase simulation was performed under constant volume at the mean box size of the solution.

The HFEs of the Generalized AMBER Force field (GAFF)[73,74] and OPLS-AA[66,67] were also computed in TIP4P water.[32] The GAFF simulations used the restrained electrostatic potential (RESP) charges computed with Hartree−Fock/6-31G*.[75,76] The measurement of the HFEs for the GAFF/RESP and OPLS-AA models followed the same BAR based procedure with the same simulation length except with a 1 fs time step. Although the couple-intramol keyword does function correctly for GAFF and OPLS-AA, the HFEs were computed still with the finite difference and couple-intramol set to yes. For rigid molecules, removing intramolecular coupling automatically by Gromacs might be more convenient. For flexible molecules, the solute might sample different conformations in the gas phase and the liquid phase.[18,77] Thus, using the finite difference should be more accurate, although it increases statistical uncertainty.

The heat of hydration is computed with the finite difference method according to the formula,

$$\Delta H_{solution} = \langle E\rangle_{sol} - (\langle E\rangle_{water} + \langle E\rangle_{solute}) + PV_{sol} - P\langle V\rangle_{water} - RT \tag{3}$$

where $\langle E\rangle$ is the average internal energy of the solution (sol), water, and solute and $\langle V\rangle$ is the average volume. The ideal gas law is used to estimate the gas phase PV term. For each solute, the internal energy E was measured from a 200 ns simulation in solution and 100 ns simulation in the gas phase.

Diffusion constants were measured from the root-mean-square displacement of the solute from 10 independent 5 ns MD simulations performed in the canonical ensemble at 298 K. The temperature is also controlled with the Nosé−Hoover thermostat with a relatively long 5 ps relaxation time. The heat of hydration and the diffusion constants were measured only for AFM based models.

IV. RESULTS AND DISCUSSION

The parameters for all the force fields developed are summarized in Tables S1−S8 of the Supporting Information. Gromacs input files for all the models are provided at http://wanglab.uark.edu/Models.

With the B3LYP-D3(BJ) based AFM model, the HFE of ethanol is −22.24 kJ/mol. The experimental values range from −19.61 kJ/mol to −21.44 kJ/mol. The −22.24 kJ/mol estimate is thus within 1 kJ/mol of the most negative experimental value. Two MP2 based AFM models have been developed previously with slightly different placements of short-range repulsion sites. One study with the repulsion sites optimized give an HFE of −20.45 kJ/mol.[31] The other with a simple approach similar to this work predicted an HFE of −19.65 kJ/mol.[32] Thus, the B3LYP-D3(BJ) HFE is slightly more negative than the MP2 one by about 2 kJ/mol. The difference between the B3LYP-D3(BJ) HFE and the MP2 HFE is not greater than the difference between different experiments.

https://doi.org/10.1021/acsphyschemau.1c00006
ACS Phys. Chem Au 2021, 1, 14−24

Table 1. HFEs of the Molecules and Various Experimental References[a]

solute	B3LYP-D3(BJ)	OPLS-AA	GAFF/RESP	experiment
ethanol	−22.24 ± 0.12	−19.80 ± 0.11	−19.32 ± 0.06	−19.61,[102] −20.02,[103] −20.83,[104] −20.96,[105] −21.21,[106] −21.44[107]
2-butanol	−21.85 ± 0.26	−21.34 ± 0.26	−19.16 ± 0.19	−15.02,[108] −19.32,[109] −19.61[105]
isobutanol	−18.91 ± 0.21	−22.01 ± 0.27	−22.46 ± 0.94	−14.62,[108] −15.62,[110] −16.91,[111] −18.82[112]
1,2-butanediol	−34.44 ± 0.13	−34.72 ± 0.74	−39.83 ± 0.25	<−33.78,[103] −38.34[97]
1,4-butanediol	−49.14 ± 0.12	−50.75 ± 0.63	−47.19 ± 0.17	<−36.21,[103] −41.52,[109] −45.31[97]
1-hexanol	−17.01 ± 0.27	−25.18 ± 0.22	−22.09 ± 0.20	−16.07,[108] −17.70,[113] −17.98,[114] −18.27,[115] −18.45,[116] −19.32[103]
piperidine	−19.02 ± 0.10	−14.34 ± 0.24	−20.03 ± 0.20	−3.48,[90] −21.13,[89] −21.93[88]
menthol	−17.65 ± 0.32	−14.78 ± 0.32	−26.90 ± 0.37	−14.80,[85,86] −17.83[85,87]

[a]The B3LYP-D3(BJ) values were determined with the AFM potentials reported in this work. Both OPLS-AA and GAFF/RESP HFEs were computed in this work in TIP4P water. All values are in kJ/mol.

We note that previously the MP2 and B3LYP HFEs have been computed in the context of QM/MM FEP and found to be close to each other for the same water model.[49] On the other hand, different water models could have a large influence on the HFE of the solute computed with QM/MM. With AFM, both solute and water were included in the QM region; thus, the solute−water cross terms are fitted against QM. This is advantageous as the HFE has a significant contribution from the cavitation energy in water. Hydrogen bonds between water molecules have to be broken to accommodate the solute, and the cavitation energy depends on the size of the solute. Fitting the solute−water interactions at the QM level will ensure a good description of the size of the cavity. Along with the use of a high-quality water model, a reliable estimate of the cavitation energy is obtained. With AFM based cross terms, it has been shown that the solute HFE has little dependence on the choice of model for water−water interactions.[78]

For 2-butanol and isobutanol, experimental values from different groups span a range of 4 kJ/mol. The B3LYP-D3(BJ) based AFM models gave HFEs of −21.85 kJ/mol for 2-butanol and −18.91 kJ/mol for isobutanol. The 2-butanol HFE is more negative than the most negative experimental value by 2 kJ/mol, and the isobutanol HFE is in good agreement with the most negative experimental value of −18.82 kJ/mol. We note that the chemical accuracy is generally considered to be 1 kcal/mol, and the 2 kJ/mol difference for 2-butanol is thus within the chemical accuracy.

It is interesting that the isobutanol seems to have a smaller absolute HFE than 2-butanol based on experimental values, although the trend is not entirely clear considering the large variance in experimental HFEs. The B3LYP-D3(BJ) based AFM models predict a trend similar to experiments with the isobutanol absolute HFE smaller by 3 kJ/mol, while both OPLS-AA and GAFF/RESP predict isobutanol to have a larger absolute HFE.

For 1-hexanol, the experimental HFEs span a range of 3 kJ/mol. The B3LYP-D3(BJ) value obtained through AFM is −17.01 kJ/mol in close agreement with the more positive experimental values of −16.07 kJ/mol or −17.70 kJ/mol. The OPLS-AA HFE is too negative even when compared to the most negative references.

For 1,2-butanediol, B3LYP-D3(BJ) gave a HFE of −34.44 kJ/mol, which is in good agreement with the more positive experimental reference of −33.78 kJ/mol or lower. For 1,4-butanediol, the predicted HFE of −49.14 kJ/mol is about 3.8 kJ/mol more negative than the most negative experimental reference of −45.31 kJ/mol. This is the worst agreement among all the solutes studied. However, even for this molecule, the agreement is within the commonly accepted chemical accuracy of 4 kJ/mol. We note that 1,2-butanediol and 1,4-butanediol are challenging molecules for HFE computations since the solute can form intramolecular hydrogen bonds.[79−81] Considering a possible correlation of the hydration structure with intramolecular hydrogen bonds, the proper sampling of solute conformations could be important. Previously, we validated that the 5 ns alchemical windows are sufficient to converge the HFE for 1,4-butanediol for an LMP2 based AFM model. In the case of 1,4-butanediol, for some conformations, a water can form hydrogen bonds with both hydroxyl groups simultaneously. This would be a challenging molecule to study if only the solute is treated quantum mechanically. AFM has the advantage that both solute and nearby solvent molecules are treated quantum mechanically in the QM/MM calculations.

The largest alcohol studied is menthol. Menthol has three chiral atoms and eight stereoisomers.[82] Two groups of four with different physical properties are formed. The form investigated in this work is the most occurring stereoisomer in nature: (1R,2S,5R), also known as l-menthol. Some clarification in our experimental references are required. While Mobley et al. reported the HFE of menthol to be −13.39 kJ/mol,[83] the value cannot be reproduced from the Guthrie database,[84] which was the source of the experimental reference for the Mobley SAMPL4 study. For l-menthol, the database lists one source for the experimental vapor pressure from Chickos et al.[85] and two solubility measurements, one by Weidenhamer et al.[86] and another by Ajisaka et al.[87] The Chickos study provided a fit to the Clausius−Clapeyron equation. From the reported parameters for the Clapeyron equation, the vapor pressure at 298 K is 7.32×10^{-5} atm, which is used as our experimental vapor pressure rather than the 7.90×10^{-5} atm in the curated database of Guthrie. We note the measured and the Clapeyron based vapor pressures lead to a very small 0.3 kJ/mol difference in the experimental HFE estimate. With the Clapeyron equation based vapor pressure, the Ajisaka et al.[87] solubility value would give a menthol HFE of −17.83 kJ/mol, and the Weidenhamer solubility would lead to a menthol HFE of −14.80 kJ/mol. The B3LYP-D3(BJ) based AFM model predicts a HFE of −17.65 kJ/mol, which is in good agreement with the estimate based the Ajisaka et al.[87] solubility. The −13.39 kJ/mol value reported by Mobley et al. is not included in Table 1 due to our inability to confirm its origin.

Piperidine is the only molecule investigated that is an amine. Experimental measurements of amine HFEs are complicated by the need to account for hydrolysis to ensure the measured thermodynamic data corresponds to the neutral molecule. An amine may change its protonation state in water, and such a

https://doi.org/10.1021/acsphyschemau.1c00006
ACS Phys. Chem Au 2021, 1, 14−24

process is frequently referred to as hydrolysis. Although hydrolysis is small under the condition of measurements by Bernauer et al.,[88] it is was nonetheless accounted for in deriving their experimental value. Whereas the experimental HFEs of Bernauer et al.[88] and Amoore et al.[89] agree, that of Cabani[90] et al. is quite different. The B3LYP-D3(BJ) HFE is in very good agreement with the experimental values by Bernauer or Amoore.

Table 1 also compared B3LYP-D3(BJ) HFEs to those computed with the GAFF/RESP and OPLS-AA models in TIP4P water. Several other small molecule force fields, such as GROMOS[91,92] and CHARMM General Force Fields (CGenFF),[93,94] also have parameters for these molecules but are not studied in this work. Both GAFF/RESP and OPLS-AA performed quite well overall.[95] OPLS-AA produced larger errors for piperidine and 1-hexanol. The HFEs based on GAFF/RESP are not very satisfactory only for menthol.

To provide a quantitative judgement of the performance of our B3LYP-D3(BJ) based AFM models, the root-mean-square error (RMSE) is computed using the median of experimental values as a reference. The use of the median removes strong influences from experimental values that do not agree well with the rest. The HFE RMSE is 3.2 kJ/mol for the AFM models, 5.3 kJ/mol for OPLS-AA, and 5.2 kJ/mol for GAFF/RESP. The B3LYP-D3(BJ) based AFM models are thus in better agreement with experiments compared to the other two models studied. It is worth noting that no experimental information is used to create these AFM models. All the parameters were only fit to best reproduce the B3LYP-D3(BJ) energy gradients.

Table 2 reports the enthalpy of hydration for the eight solutes computed with the finite difference method mentioned in Section III using the B3LYP-D3(BJ) based AFM models. Experimental values were found only for seven of the eight solutes. Overall the agreement is excellent. The computed estimate is consistently within 2 kJ/mol of the closest experimental reference. The most surprising result is the close agreement with experimental heat of hydration for piperidine since the experimental value was taken from the work of Cabani,[90] which gives a very different HFE. The Cabani measurements of the enthalpy of hydration and HFE are independent from each other. The Cabani enthalpy measurement carefully considered hydrolysis by extrapolating in HCl, NaOH, and water solutions and checked measured value against that of Sacconi.[96] This suggests that the Cabani enthalpy measurement is reliable.

Table 2. Enthalpies of Hydration at 298 K Computed with the AFM Force Fields Fitted to B3LYP-D3(BJ)[a]

solute	B3LYP-D3(BJ)	experiment
ethanol	−52.03 ± 0.23	−52.40,[117] −50.42,[118] −50.6,[119] −52.62,[120] −52.65[121]
2-butanol	−62.47 ± 0.14	−62.7,[119] −62.72,[117] −62.83[121]
isobutanol	−58.27 ± 0.48	−60.2,[119] −60.15,[117] −60.11[121]
1,2-butanediol	−72.19 ± 0.26	−82.1,[97] −73.8[99,100]
1,4-butanediol	−89.10 ± 0.45	−89.6[97]
1-hexanol	−66.28 ± 0.25	−66.20,[117] −67.4,[119] −68.02,[120] −68.17[122]
piperidine	−65.87 ± 0.41	−65.4,[90] −63.9,[96] −66.41[117]
menthol	−76.99 ± 0.52	

[a]The various experimental references are listed, where available. All values are in kJ/mol.

The enthalpy of hydration of 1,2-butanediol is worthy of additional discussion. The −82.1 kJ/mol reference was taken from the Compernolle and Müller compilation of experimental data from various sources[97], and the −73.8 kJ/mol value is computed by us based on published experimental data. For the enthalpy of hydration, the Compernolle and Müller work used Hess's Law and computed the enthalpy of hydration as

$$\Delta H_{hyd} = \Delta H_{sol}^{\infty} - \Delta H_{vap} \tag{4}$$

Equation 4 assumes the vapor first condenses into liquid and then dissolves at infinite dissolution in the solvent. ΔH_{sol} is the enthalpy of the solution at low concentration of the pure solute in the liquid state measured by Lopes Jesus,[98] and the ΔH_{vap} is the heat of vaporization measured by Verevkin.[99]

Following the discussion by Compernolle and Müller, another experimental estimate of ΔH_{sol} can be obtained with the van Hoff equation from the temperature dependent activity coefficients at infinite dilution, γ_s^{∞}, measured by Suleiman and Eckert,[100]

$$\Delta H_{sol} = R\frac{\mathrm{d}\ln\gamma_s^{\infty}}{\mathrm{d}(1/T)} \tag{5}$$

The fit shows excellent linearity and is reported in the Supporting Information. Linear regression of eq 5 based on the Suleiman and Eckert data gives a ΔH_{sol} of −3.4 kJ/mol. When combined with the same ΔH_{vap} measured by Verevkin, an experimental value of −73.8 kJ/mol can be reached for the enthalpy of hydration for 1,2-butanediol. It is worth mentioning that multiple estimates of ΔH_{vap} are available with the span of 3 kJ/mol.[99] The less positive ΔH_{vap} estimate would give an enthalpy of hydration estimate of −72.8 kJ/mol in even better agreement with our computed value.

It is interesting to note that the difference in HFEs between the two diols, 1,2-butanediol and 1,4-butanediol, is 14.7 kJ/mol and the difference in enthalpy of hydration is 17.2 kJ/mol. 1,2-Butanediol is stabilized less by the solvent than 1,4-butanediol for both quantities. It is easier for 1,2-butanediol to form an intramolecular hydrogen bond since such a hydrogen bond will result in a five-member ring. In the gas phase, our simulation indicates that 1,2-butanediol will form an intramolecular hydrogen bond in 75% of the cases and 1,4-butanediol will form an intramolecular hydrogen bond in 45% of the configurations. We note that intramolecular hydrogen bonds will not affect the solution phase stability appreciably. There would be little difference in the total number of hydrogen bonds in an aqueous solution as the hydroxyl groups can also form hydrogen bonds with water. The 30% extra hydrogen bond stabilizes 1,2-butanediol more in the gas phase. This stabilization would lead to less favorable enthalpy of solution when compared to that of 1,4-butanediol. The observed 2.8 kJ/mol less enthalpic cost associated with the dehydration of 1,2-butanediol is consistent with 30% extra hydrogen bonds in the gas phase, which corresponds to a hydrogen bond strength of 9.3 kJ/mol (2.2 kcal/mol).

The diffusion constants of the solutes are reported in Table 3. Very good agreement can be observed between experimental values and predictions based on B3LYP-D3(BJ) through AFM. For the six molecules where experimental references are available, the RMSE is 0.10×10^{-5} cm^2/ps, the largest error being 2-butanol at 0.14×10^{-5} cm^2/ps. This indicates that

https://doi.org/10.1021/acsphyschemau.1c00006
ACS Phys. Chem Au 2021, 1, 14−24

Table 3. Diffusion Coefficients in Dilute Aqueous Solution Based on the AFM Models Fitted to B3LYP-D3(BJ)[a]

solute	B3LYP-D3(BJ)	experiment
ethanol	1.10 ± 0.11	1.24,[123,124] 1.23,[125] 1.22[126,127]
2-butanol	0.80 ± 0.06	0.94[125]
isobutanol	0.86 ± 0.08	0.95[125]
1,2-butanediol	0.87 ± 0.06	0.93[128]
1,4-butanediol	0.92 ± 0.06	0.91[128]
1-hexanol	0.90 ± 0.09	0.83[125]
piperidine	0.76 ± 0.04	
menthol	0.68 ± 0.05	

[a]Error bars are the standard error of the mean. All values are in 10^{-5} cm^2/s.

AFM models based on B3LYP-D3(BJ) are able to predict good diffusivity for simple alcohols and also for the amine. The good agreements definitely benefitted from the good diffusivity of the BLYPSP-4F model, and one could argue that diffusion constants are not particularly challenging to predict given their narrow range for these compounds.

It has been seen with our MP2 and LMP2 based models that the AFM model diffusion constants are in excellent agreement with experiments even when the agreement in HFE is less than satisfactory. We thus anticipate that for the two molecules where experimental references are unavailable, the prediction from our force fields is likely to be the best available estimate.

V. SUMMARY AND CONCLUSION

Simple pairwise additive force fields for seven alcohols and one amine, piperidine, are developed to reproduce the B3LYP-D3(BJ) PES of these molecules in dilute aqueous solutions. With counterpoise corrected CCSD(T) as a reference, B3LYP-D3(BJ) is found to produce a good description of the ethanol−water dimer PES. The force fields were developed with the AFM method by fitting B3LYP-D3(BJ) gradients from QM/MM calculations designed to capture many-body effects in the condensed phase. The adequacy of simple energy expressions to fit DFT can only be justified with the realization that the mapping of the PES is only for a specific solute under a particular thermodynamic condition of interest. In this narrow window of applicability, our results indicate that the mapping of B3LYP-D3(BJ) PES with AFM performed satisfactorily for investigating many ensemble properties that would be challenging to study directly with DFT.

With the B3LYP-D3(BJ) PES, the HFE, enthalpy of hydration, and diffusion constants are computed and compared to experimental values to validate the performance of the models developed by AFM. In all the cases, the force fields developed were able to accurately reproduce HFEs of the solutes, outperforming both GAFF/RESP and OPLS-AA. Comparison with experiments are complicated by the large differences between experimental values from different measurements. The B3LYP-D3(BJ) based predictions are always within chemical accuracy from the closest experimental value and produce an RMSE of 3.2 kJ/mol when the medians of the experimental values were used as a reference. Similar agreements of chemical accuracy or better are also observed for enthalpies of hydration for all the molecules. For 1,2-butanediol, the computed value is different from the compiled experimental value from Compernolle and Müller. It is, however, in good agreement with our estimate based on the experimental activity coefficients of Suleiman and Eckert.

Diffusion constants predicted by the B3LYP-D3(BJ) based models in the BLYPSP-4F water are also in excellent agreement with experiments for all the solutes where experimental diffusivities are available. Our previous studies based on MP2 and LMP2 showed that the AFM based diffusion constants were in good agreement with experiments even when the agreement for HFE was less than satisfactory.[32] Thus, there is good confidence that the AFM based diffusion constants are reliable even for the molecules, where experimental references are not available.

We believe the good agreements for the solute properties indicate that the models from AFM are faithfully mapping the underlying B3LYP-D3(BJ) PES. Mapping B3LYP-D3(BJ) with AFM provides a powerful method to obtain DFT based thermodynamic properties for small molecules. While the typical QM/MM based approach for prediction of HFE only models the solute with QM, one clear advantage of an AFM based approach is that the training is performed with both the solute and the nearby solvent molecules modeled with QM. In addition, AFM produces a customized model for each solute that is able to study many thermodynamic properties. It is thus not limited to the prediction of HFE.

This study investigated seven alcohols and one amine as a demonstration. It is worth noting that although only a few properties were computed in this work, our force fields were designed to capture the B3LYP-D3(BJ) PES rather than reproducing any particular property. Thus, the models are expected to capture other properties of the dilute solution. We note that the pairwise additive potentials can only capture the B3LYP-D3(BJ) PES in an average sense, since detailed many-body effects are not modeled explicitly. For some properties, the simplistic energy expressions used would cause important physics to be missed, and thus the prediction may not be accurate. Also, some properties may require explicit considerations of nuclear quantum effects.[101] While we are confident that there are other properties where AFM based models would provide accurate predictions, the actual performance for properties not studied in this work would need to be further investigated and validated.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsphyschemau.1c00006.

QM/MM region selection protocol, summary of the parameters of the eight molecules, and calculation of the enthalpy of hydration for 1,2-butanediol (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Feng Wang − *Department of Chemistry and Biochemistry, University of Arkansas, Fayetteville, Arkansas 72701, United States*; orcid.org/0000-0002-2740-3534; Email: fengwang@uark.edu

Author

Dong Zheng − *Department of Chemistry and Biochemistry, University of Arkansas, Fayetteville, Arkansas 72701, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acsphyschemau.1c00006

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the National Institutes of Health under Grant Nos. 1R01GM120578 and 2P20GM103429. Computational resources were provided by the Arkansas High Performance Computing Center through Grant MRI-R2 No. 09559124 provided by the NSF; partial funding for the computational resource is provided by the Arkansas Bioscience Institute.

■ REFERENCES

(1) Laage, D.; Elsaesser, T.; Hynes, J. T. Water Dynamics in the Hydration Shells of Biomolecules. *Chem. Rev.* **2017**, *117*, 10694−10725.
(2) Ross, G. A.; Bodnarchuk, M. S.; Essex, J. W. Water Sites, Networks, and Free Energies with Grand Canonical Monte Carlo. *J. Am. Chem. Soc.* **2015**, *137*, 14930−14943.
(3) Ebbinghaus, S.; Kim, S. J.; Heyden, M.; Yu, X.; Heugen, U.; Gruebele, M.; Leitner, D. M.; Havenith, M. An Extended Dynamical Hydration Shell around Proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 20749−20752.
(4) Iftimie, R.; Minary, P.; Tuckerman, M. E. Ab Initio Molecular Dynamics: Concepts, Recent Developments, and Future Trends. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 6654−6659.
(5) Paquet, E.; Viktor, H. L. Computational Methods for Ab Initio Molecular Dynamics. *Adv. Chem.* **2018**, *2018*, 1−14.
(6) Marx, D.; Hutter, J. *Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods*; Cambridge University Press: Cambridge, 2009.
(7) Scuseria, G. E. Linear Scaling Density Functional Calculations with Gaussian Orbitals. *J. Phys. Chem. A* **1999**, *103*, 4782−4790.
(8) Iyengar, S. S.; Schlegel, H. B.; Millam, J. M.; Voth, G. A.; Scuseria, G. E.; Frisch, M. J. Ab Initio Molecular Dynamics: Propagating the Density Matrix with Gaussian Orbitals. II. Generalizations Based on Mass-Weighting, Idempotency, Energy Conservation and Choice of Initial Conditions. *J. Chem. Phys.* **2001**, *115*, 10291−10302.
(9) Schlegel, H. B.; Millam, J. M.; Iyengar, S. S.; Voth, G. A.; Daniels, A. D.; Scuseria, G. E.; Frisch, M. J. Ab Initio Molecular Dynamics: Propagating the Density Matrix with Gaussian Orbitals. *J. Chem. Phys.* **2001**, *114*, 9758−9763.
(10) Schlegel, H. B.; Iyengar, S. S.; Li, X.; Millam, J. M.; Voth, G. A.; Scuseria, G. E.; Frisch, M. J. Ab Initio Molecular Dynamics: Propagating the Density Matrix with Gaussian Orbitals. III. Comparison with Born−Oppenheimer Dynamics. *J. Chem. Phys.* **2002**, *117*, 8694−8704.
(11) Artacho, E.; Sánchez-Portal, D.; Ordejón, P.; García, A.; Soler, J. M. Linear-Scaling Ab-Initio Calculations for Large and Complex Systems. *Phys. Status Solidi B* **1999**, *215*, 809−817.
(12) Mohr, S.; Ratcliff, L. E.; Genovese, L.; Caliste, D.; Boulanger, P.; Goedecker, S.; Deutsch, T. Accurate and Efficient Linear Scaling DFT Calculations with Universal Applicability. *Phys. Chem. Chem. Phys.* **2015**, *17*, 31360−31370.
(13) 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.
(14) Izvekov, S.; Parrinello, M.; Burnham, 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.
(15) Wang, L. P.; Chen, J.; Van Voorhis, T. Systematic Parametrization of Polarizable Force Fields from Quantum Chemistry Data. *J. Chem. Theory Comput.* **2013**, *9*, 452−460.
(16) Ercolessi, F.; Adams, J. B. Interatomic Potentials from First-Principles Calculations: The Force-Matching Method. *Europhys. Lett.* **1994**, *26*, 583−588.
(17) Kvaal, S. Ab Initio Quantum Dynamics Using Coupled-Cluster. *J. Chem. Phys.* **2012**, *136*, 194109.
(18) Hudson, P. S.; Boresch, S.; Rogers, D. M.; Woodcock, H. L. Accelerating QM/MM Free Energy Computations Via Intramolecular Force Matching. *J. Chem. Theory Comput.* **2018**, *14*, 6327−6335.
(19) Giese, T. J.; York, D. M. Development of a Robust Indirect Approach for MM –> QM Free Energy Calculations That Combines Force-Matched Reference Potential and Bennett's Acceptance Ratio Methods. *J. Chem. Theory Comput.* **2019**, *15*, 5543−5562.
(20) Brunken, C.; Reiher, M. Self-Parametrizing System-Focused Atomistic Models. *J. Chem. Theory Comput.* **2020**, *16*, 1646−1665.
(21) Bounds, D. G. On the MCY Potential for Liquid Water. *Chem. Phys. Lett.* **1983**, *96*, 604−606.
(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) Metz, M. P.; Piszczatowski, K.; Szalewicz, K. Automatic Generation of Intermolecular Potential Energy Surfaces. *J. Chem. Theory Comput.* **2016**, *12*, 5895−5919.
(24) McDaniel, J. G.; Schmidt, J. R. Physically-Motivated Force Fields from Symmetry-Adapted Perturbation Theory. *J. Phys. Chem. A* **2013**, *117*, 2053−2066.
(25) McDaniel, J. G.; Schmidt, J. R. Next-Generation Force Fields from Symmetry-Adapted Perturbation Theory. *Annu. Rev. Phys. Chem.* **2016**, *67*, 467−488.
(26) Mas, E. M.; Szalewicz, K.; Bukowski, R.; Jeziorski, B. Pair Potential for Water from Symmetry-Adapted Perturbation Theory. *J. Chem. Phys.* **1997**, *107*, 4207−4218.
(27) Jeziorski, B.; Moszynski, R.; Szalewicz, K. Perturbation Theory Approach to Intermolecular Potential Energy Surfaces of van der Waals Complexes. *Chem. Rev.* **1994**, *94*, 1887−1930.
(28) Yuan, Y.; Ma, Z.; Wang, F. Leveraging Local MP2 to Reduce Basis Set Superposition Errors: An Efficient First-Principles Based Force-Field for Carbon Dioxide. *J. Chem. Phys.* **2019**, *151*, 184501.
(29) Wang, F.; Akin-Ojo, O.; Pinnick, E.; Song, Y. Approaching Post-Hartree−Fock Quality Potential Energy Surfaces with Simple Pair-Wise Expressions: Parameterising Point-Charge-Based Force Fields for Liquid Water Using the Adaptive Force Matching Method. *Mol. Simul.* **2011**, *37*, 591−605.
(30) 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**, *121*, 6637−6645.
(31) Rogers, T. R.; Wang, F. Accurate MP2-Based Force Fields Predict Hydration Free Energies for Simple Alkanes and Alcohols in Good Agreement with Experiments. *J. Chem. Phys.* **2020**, *153*, 244505.
(32) Zheng, D.; Yuan, Y.; Wang, F. Determining the Hydration Free Energies of Selected Small Molecules with MP2 and Local MP2 through Adaptive Force Matching. *J. Chem. Phys.* **2021**, *154*, 104113.
(33) Yuan, Y.; Ma, Z.; Wang, F. Development and Validation of a DFT-Based Force Field for a Hydrated Homoalanine Polypeptide. *J. Phys. Chem. B* **2021**, *125*, 1568−1581.
(34) Mezei, M.; Beveridge, D. L. Free Energy Simulations. *Ann. N. Y. Acad. Sci.* **1986**, *482*, 1−23.
(35) Squire, D. R.; Hoover, W. G. Monte Carlo Simulation of Vacancies in Rare-Gas Crystals. *J. Chem. Phys.* **1969**, *50*, 701−706.
(36) Pulay, P. Localizability of Dynamic Electron Correlation. *Chem. Phys. Lett.* **1983**, *100*, 151−154.
(37) Purvis, G. D., III; Bartlett, R. J. A Full Coupled-Cluster Singles and Doubles Model: The Inclusion of Disconnected Triples. *J. Chem. Phys.* **1982**, *76*, 1910−1918.
(38) Scuseria, G. E.; Janssen, C. L.; Schaefer, H. F., III An Efficient Reformulation of the Closed-Shell Coupled Cluster Single and Double Excitation (CCSD) Equations. *J. Chem. Phys.* **1988**, *89*, 7382−7387.

(39) Scuseria, G. E.; Schaefer, H. F., III Is Coupled Cluster Singles and Doubles (CCSD) More Computationally Intensive Than Quadratic Configuration Interaction (QCISD)? *J. Chem. Phys.* **1989**, *90*, 3700−3703.
(40) Christiansen, O.; Koch, H.; Jo/rgensen, P. Perturbative Triple Excitation Corrections to Coupled Cluster Singles and Doubles Excitation Energies. *J. Chem. Phys.* **1996**, *105*, 1451−1459.
(41) Yang, W.; Cui, Q.; Min, D.; Li, H. QM/MM Alchemical Free Energy Simulations: Challenges and Recent Developments. In *Annual Reports in Computational Chemistry*; Wheeler, R. A., Ed.; Elsevier: 2010; Vol. *6*, Chapter 4, pp 51−62.
(42) König, G.; Hudson, P. S.; Boresch, S.; Woodcock, H. L. Multiscale Free Energy Simulations: An Efficient Method for Connecting Classical MD Simulations to QM or QM/MM Free Energies Using Non-Boltzmann Bennett Reweighting Schemes. *J. Chem. Theory Comput.* **2014**, *10*, 1406−1419.
(43) Jia, X.; Wang, M.; Shao, Y.; König, G.; Brooks, B. R.; Zhang, J. Z. H.; Mei, Y. Calculations of Solvation Free Energy through Energy Reweighting from Molecular Mechanics to Quantum Mechanics. *J. Chem. Theory Comput.* **2016**, *12*, 499−511.
(44) König, G.; Mei, Y.; Pickard, F. C.; Simmonett, A. C.; Miller, B. T.; Herbert, J. M.; Woodcock, H. L.; Brooks, B. R.; Shao, Y. Computation of Hydration Free Energies Using the Multiple Environment Single System Quantum Mechanical/Molecular Mechanical Method. *J. Chem. Theory Comput.* **2016**, *12*, 332−344.
(45) Giese, T. J.; York, D. M. Development of a Robust Indirect Approach for MM → QM Free Energy Calculations That Combines Force-Matched Reference Potential and Bennett's Acceptance Ratio Methods. *J. Chem. Theory Comput.* **2019**, *15*, 5543−5562.
(46) Genheden, S.; Cabedo Martinez, A. I.; Criddle, M. P.; Essex, J. W. Extensive All-Atom Monte Carlo Sampling and QM/MM Corrections in the SAMPL4 Hydration Free Energy Challenge. *J. Comput.-Aided Mol. Des.* **2014**, *28*, 187−200.
(47) Sodt, A. J.; Mei, Y.; König, G.; Tao, P.; Steele, R. P.; Brooks, B. R.; Shao, Y. Multiple Environment Single System Quantum Mechanical/Molecular Mechanical (MESS-QM/MM) Calculations. 1. Estimation of Polarization Energies. *J. Phys. Chem. A* **2015**, *119*, 1511−1523.
(48) König, G.; Pickard, F. C.; Huang, J.; Simmonett, A. C.; Tofoleanu, F.; Lee, J.; Dral, P. O.; Prasad, S.; Jones, M.; Shao, Y.; Thiel, W.; Brooks, B. R. Calculating Distribution Coefficients Based on Multi-Scale Free Energy Simulations: An Evaluation of MM and QM/MM Explicit Solvent Simulations of Water-Cyclohexane Transfer in the SAMPL5 Challenge. *J. Comput.-Aided Mol. Des.* **2016**, *30*, 989−1006.
(49) König, G.; Pickard, F. C.; Huang, J.; Thiel, W.; MacKerell, A. D.; Brooks, B. R.; York, D. M. A Comparison of QM/MM Simulations with and without the Drude Oscillator Model Based on Hydration Free Energies of Simple Solutes. *Molecules* **2018**, *23*, 2695.
(50) Zwanzig, R. W. High-Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases. *J. Chem. Phys.* **1954**, *22*, 1420−1426.
(51) Song, Y.; Wang, F. Accurate Ranking of CH4·(H2O)20 Clusters with the Density Functional Theory Supplemental Potential Approach. *Theor. Chem. Acc.* **2013**, *132*, 1324.
(52) Song, Y.; Akin-Ojo, O.; Wang, F. Correcting for Dispersion Interaction and Beyond in Density Functional Theory through Force Matching. *J. Chem. Phys.* **2010**, *133*, 174115.
(53) Hu, H.; Wang, F. The Liquid-Vapor Equilibria of TIP4P/2005 and BLYPSP-4F Water Models Determined through Direct Simulations of the Liquid-Vapor Interface. *J. Chem. Phys.* **2015**, *142*, 214507.
(54) Hu, H.; Ma, Z.; Wang, F. On the Transferability of Three Water Models Developed by Adaptive Force Matching. In *Annual Reports in Computational Chemistry*; Wheeler, R. A., Ed.; Elsevier: 2014; Vol. *10*, Chapter 2, pp 25−43.
(55) Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(56) Becke, A. D. A New Mixing of Hartree−Fock and Local Density-Functional Theories. *J. Chem. Phys.* **1993**, *98*, 1372−1377.
(57) Hertwig, R. H.; Koch, W. On the Parameterization of the Local Correlation Functional. What Is Becke-3-LYP? *Chem. Phys. Lett.* **1997**, *268*, 345−351.
(58) 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.
(59) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *J. Comput. Chem.* **2011**, *32*, 1456−65.
(60) Smith, D. G.; Burns, L. A.; Patkowski, K.; Sherrill, C. D. Revised Damping Parameters for the D3 Dispersion Correction to Density Functional Theory. *J. Phys. Chem. Lett.* **2016**, *7*, 2197−2203.
(61) 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.
(62) Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. Electron Affinities of the First-Row Atoms Revisited. Systematic Basis Sets and Wave Functions. *J. Chem. Phys.* **1992**, *96*, 6796−6806.
(63) Boys, S. F.; Bernardi, F. The Calculation of Small Molecular Interactions by the Differences of Separate Total Energies. Some Procedures with Reduced Errors. *Mol. Phys.* **1970**, *19*, 553−566.
(64) Chai, J.-D.; Head-Gordon, M. Long-Range Corrected Hybrid Density Functionals with Damped Atom−Atom Dispersion Corrections. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615−6620.
(65) Anatole von Lilienfeld, O.; Tkatchenko, A. Two- and Three-Body Interatomic Dispersion Energy Contributions to Binding in Molecules and Solids. *J. Chem. Phys.* **2010**, *132*, 234109.
(66) 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.
(67) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. Evaluation and Reparametrization of the OPLS-AA Force Field for Proteins Via Comparison with Accurate Quantum Chemical Calculations on Peptides. *J. Phys. Chem. B* **2001**, *105*, 6474−6487.
(68) Bennett, C. H. Efficient Estimation of Free Energy Differences from Monte Carlo Data. *J. Comput. Phys.* **1976**, *22*, 245−268.
(69) Shirts, M. R.; Chodera, J. D. Statistically Optimal Analysis of Samples from Multiple Equilibrium States. *J. Chem. Phys.* **2008**, *129*, 124105.
(70) Beutler, T. C.; Mark, A. E.; van Schaik, R. C.; Gerber, P. R.; van Gunsteren, W. F. Avoiding Singularities and Numerical Instabilities in Free Energy Calculations Based on Molecular Simulations. *Chem. Phys. Lett.* **1994**, *222*, 529−539.
(71) Parrinello, M.; Rahman, A. Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method. *J. Appl. Phys.* **1981**, *52*, 7182−7190.
(72) Nosé, S.; Klein, M. L. Constant Pressure Molecular Dynamics for Molecular Systems. *Mol. Phys.* **1983**, *50*, 1055−1076.
(73) 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.
(74) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(75) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Kollman, P. A. Application of RESP Charges to Calculate Conformational Energies, Hydrogen Bond Energies, and Free Energies of Solvation. *J. Am. Chem. Soc.* **1993**, *115*, 9620−9631.
(76) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. *J. Phys. Chem.* **1993**, *97*, 10269−10280.

https://doi.org/10.1021/acsphyschemau.1c00006
ACS Phys. Chem Au 2021, 1, 14−24

(77) Hudson, P. S.; Woodcock, H. L.; Boresch, S. Use of Interaction Energies in QM/MM Free Energy Simulations. *J. Chem. Theory Comput.* **2019**, *15*, 4632−4645.
(78) Li, J.; Wang, F. The Effects of Replacing the Water Model While Decoupling Water-Water and Water-Solute Interactions on Computed Properties of Simple Salts. *J. Chem. Phys.* **2016**, *145*, 044501.
(79) Yunger, L. M.; Cramer, R. D. Measurement and Correlation of Partition Coefficients of Polar Amino Acids. *Mol. Pharmacol.* **1981**, *20*, 602−608.
(80) König, G.; Boresch, S. Hydration Free Energies of Amino Acids: Why Side Chain Analog Data Are Not Enough. *J. Phys. Chem. B* **2009**, *113*, 8967−8974.
(81) Roseman, M. A. Hydrophilicity of Polar Amino Acid Side-Chains Is Markedly Reduced by Flanking Peptide Bonds. *J. Mol. Biol.* **1988**, *200*, 513−522.
(82) Kamatou, G. P.; Vermaak, I.; Viljoen, A. M.; Lawrence, B. M. Menthol: A Simple Monoterpene with Remarkable Biological Properties. *Phytochemistry* **2013**, *96*, 15−25.
(83) Mobley, D. L.; Wymer, K. L.; Lim, N. M.; Guthrie, J. P. Blind Prediction of Solvation Free Energies from the SAMPL4 Challenge. *J. Comput.-Aided Mol. Des.* **2014**, *28*, 135−150.
(84) Guthrie, J. P.; Mobley, D. L. *The Guthrie Hydration Free Energy Database of Experimental Small Molecule Hydration Free Energies*. Department of Pharmaceutical Sciences, UC Irvine, 2018. https://escholarship.org/uc/item/53n2h10t (accessed July 8, 2021).
(85) Chickos, J. S.; Garin, D. L.; Hitt, M.; Schilling, G. Some Solid State Properties of Enantiomers and Their Racemates. *Tetrahedron* **1981**, *37*, 2255−2259.
(86) Weidenhamer, J. D.; Macias, F. A.; Fischer, N. H.; Williamson, G. B. Just How Insoluble Are Monoterpenes? *J. Chem. Ecol.* **1993**, *19*, 1799−1807.
(87) Ajisaka, N.; Hara, K.; Mikuni, K.; Hara, K.; Hashimoto, H. Effects of Branched Cyclodextrins on the Solubility and Stability of Terpenes. *Biosci., Biotechnol., Biochem.* **2000**, *64*, 731−734.
(88) Bernauer, M.; Dohnal, V. Temperature Dependences of Limiting Activity Coefficients and Henry's Law Constants for N-methylpyrrolidone, Pyridine, and Piperidine in Water. *Fluid Phase Equilib.* **2009**, *282*, 100−107.
(89) Amoore, J. E.; Buttery, R. G. Partition Coefficient and Comparative Olfactometry. *Chem. Senses* **1978**, *3*, 57−71.
(90) Cabani, S.; Conti, G.; Lepori, L. Thermodynamic Study on Aqueous Dilute Solutions of Organic Compounds. Part 1.—Cyclic Amines. *Trans. Faraday Soc.* **1971**, *67*, 1933−1942.
(91) Oostenbrink, C.; Villa, A.; Mark, A. E.; van Gunsteren, W. F. A Biomolecular Force Field Based on the Free Enthalpy of Hydration and Solvation: The Gromos Force-Field Parameter Sets 53A5 and 53A6. *J. Comput. Chem.* **2004**, *25*, 1656−1676.
(92) Schmid, N.; Eichenberger, A. P.; Choutko, A.; Riniker, S.; Winger, M.; Mark, A. E.; van Gunsteren, W. F. Definition and Testing of the Gromos Force-Field Versions 54A7 and 54B7. *Eur. Biophys. J.* **2011**, *40*, 843−856.
(93) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; Mackerell, A. D., Jr. CHARMM General Force Field: A Force Field for Drug-Like Molecules Compatible with the CHARMM All-Atom Additive Biological Force Fields. *J. Comput. Chem.* **2009**, *31*, 671−690.
(94) Yu, W.; He, X.; Vanommeslaeghe, K.; MacKerell, A. D., Jr Extension of the CHARMM General Force Field to Sulfonyl-Containing Compounds and Its Utility in Biomolecular Simulations. *J. Comput. Chem.* **2012**, *33*, 2451−2468.
(95) Mobley, D. L.; Dumont, É.; Chodera, J. D.; Dill, K. A. Comparison of Charge Models for Fixed-Charge Force Fields: Small-Molecule Hydration Free Energies in Explicit Solvent. *J. Phys. Chem. B* **2007**, *111*, 2242−2254.
(96) Sacconi, L.; Paoletti, P.; Ciampolini, M. Thermochemical Studies. I. Thermodynamic Functions of Solutions of Pyridine Bases in Water. *J. Am. Chem. Soc.* **1960**, *82*, 3828−3831.
(97) Compernolle, S.; Müller, J. F. Henry's Law Constants of Polyols. *Atmos. Chem. Phys.* **2014**, *14*, 12815−12837.
(98) Lopes Jesus, A. J.; Ermelinda Eusébio, M.; Redinha, J. S.; Leitão, M. L. P. Enthalpy of Solvation of Butanediols in Different Solvents. *Thermochim. Acta* **2000**, *344*, 3−8.
(99) Verevkin, S. P. Determination of Vapor Pressures and Enthalpies of Vaporization of 1,2-Alkanediols. *Fluid Phase Equilib.* **2004**, *224*, 23−29.
(100) Suleiman, D.; Eckert, C. A. Limiting Activity Coefficients of Diols in Water by a Dew Point Technique. *J. Chem. Eng. Data* **1994**, *39*, 692−696.
(101) Yuan, Y.; Li, J.; Li, X. Z.; Wang, F. The Strengths and Limitations of Effective Centroid Force Models Explored by Studying Isotopic Effects in Liquid Water. *J. Chem. Phys.* **2018**, *148*, 184102.
(102) Ueberfeld, J.; Zbinden, H.; Gisin, N.; Pellaux, J.-P. Determination of Henry's Constant Using a Photoacoustic Sensor. *J. Chem. Thermodyn.* **2001**, *33*, 755−764.
(103) Altschuh, J.; Brüggemann, R.; Santl, H.; Eichinger, G.; Piringer, O. G. Henry's Law Constants for a Diverse Set of Organic Chemicals: Experimental Determination and Comparison of Estimation Methods. *Chemosphere* **1999**, *39*, 1871−1887.
(104) Vitenberg, A. G.; Dobryakov, Y. G. Gas-Chromatographic Determination of the Distribution Ratios of Volatile Substances in Gas-Liquid Systems. *Russ. Russ. J. Appl. Chem.* **2008**, *81*, 339−359.
(105) Snider, J. R.; Dawson, G. Tropospheric Light Alcohols, Carbonyls, and Acetonitrile: Concentrations in the Southwestern United States and Henry's Law Data. *J. Geophys. Res.* **1985**, *90*, 3797−3805.
(106) Burnett, M. G. Determination of Partition Coefficients at Infinite Dilution by the Gas Chromatographic Analysis of the Vapor above Dilute Solutions. *Anal. Chem.* **1963**, *35*, 1567−1570.
(107) Rohrschneider, L. Solvent Characterization by Gas-Liquid Partition Coefficients of Selected Solutes. *Anal. Chem.* **1973**, *45*, 1241−1247.
(108) Butler, J.; Ramchandani, C.; Thomson, D. 58. The Solubility of Non-electrolytes. Part I. The Free Energy of Hydration of Some Aliphatic Alcohols. *J. Chem. Soc.* **1935**, 280−285.
(109) Sander, R. Compilation of Henry's Law Constants (Version 4.0) for Water as Solvent. *Atmos. Chem. Phys.* **2015**, *15*, 4399−4981.
(110) Kim, Y. H.; Kim, K. H. Recent Advances in Thermal Desorption-Gas Chromatography-Mass Spectrometery Method to Eliminate the Matrix Effect between Air and Water Samples: Application to the Accurate Determination of Henry's Law Constant. *J. Chromatogr. A* **2014**, *1342*, 78−85.
(111) Shiu, W.-Y.; Mackay, D. Henry's Law Constants of Selected Aromatic Hydrocarbons, Alcohols, and Ketones. *J. Chem. Eng. Data* **1997**, *42*, 27−30.
(112) Rytting, J. H.; Huston, L. P.; Higuchi, T. Thermodynamic Group Contributions for Hydroxyl, Amino, and Methylene Groups. *J. Pharm. Sci.* **1978**, *67*, 615−618.
(113) Lei, Y. D.; Shunthirasingham, C.; Wania, F. Comparison of Headspace and Gas-Stripping Techniques for Measuring the Air−Water Partitioning of Normal Alkanols (C4 to C10): Effect of Temperature, Chain Length, and Adsorption to the Water Surface. *J. Chem. Eng. Data* **2007**, *52*, 168−179.
(114) Shunthirasingham, C.; Cao, X.; Lei, Y. D.; Wania, F. Large Bubbles Reduce the Surface Sorption Artifact of the Inert Gas Stripping Method. *J. Chem. Eng. Data* **2013**, *58*, 792−797.
(115) Li, J.; Carr, P. W. Measurement of Water-Hexadecane Partition Coefficients by Headspace Gas Chromatography and Calculation of Limiting Activity Coefficients in Water. *Anal. Chem.* **1993**, *65*, 1443−1450.
(116) Rytting, J. H.; Huston, L. P.; Higuchi, T. Thermodynamic Group Contributions for Hydroxyl, Amino, and Methylene Groups. *J. Pharm. Sci.* **1978**, *67*, 615−618.
(117) Cabani, S.; Gianni, P.; Mollica, V.; Lepori, L. Group Contributions to the Thermodynamic Properties of Non-Ionic Organic Solutes in Dilute Aqueous Solution. *J. Solution Chem.* **1981**, *10*, 563−595.

(118) Ben-Naim, A.; Marcus, Y. Solvation Thermodynamics of Nonionic Solutes. *J. Chem. Phys.* **1984**, *81*, 2016−2027.
(119) Haynes, W. M. *CRC Handbook of Chemistry and Physics*; CRC Press: 2014.
(120) Hallén, D.; Nilsson, S. O.; Rothschild, W.; Wadsö, I. Enthalpies and Heat Capacities for n-Alkan-1-ols in H2O and D2O. *J. Chem. Thermodyn.* **1986**, *18*, 429−442.
(121) Rouw, A. C.; Somsen, G. The Solvation of Some Alcohols in Binary Solvents: Enthalpies of Solution and Enthalpies of Transfer. *J. Chem. Thermodyn.* **1981**, *13*, 67−76.
(122) Hill, D.; White, L. The Enthalpies of Solution of Hexa-1-ol and Heptan-1-ol in Water. *Aust. J. Chem.* **1974**, *27*, 1905−1916.
(123) Hammond, B. R.; Stokes, R. H. Diffusion in Binary Liquid Mixtures. Part 1.—Diffusion Coefficients in the System Ethanol + Water at 25°. *Trans. Faraday Soc.* **1953**, *49*, 890−895.
(124) Pratt, K. C.; Wakeham, W. A. The Mutual Diffusion Coefficient of Ethanol-Water Mixtures: Determination by a Rapid, New Method. *Proc. R. Soc. London. A* **1974**, *336*, 393−406.
(125) Hao, L.; Leaist, D. G. Binary Mutual Diffusion Coefficients of Aqueous Alcohols. Methanol to 1-Heptanol. *J. Chem. Eng. Data* **1996**, *41*, 210−213.
(126) Easteal, A. J.; Woolf, L. A. Pressure and Temperature Dependence of Tracer Diffusion Coefficients of Methanol, Ethanol, Acetonitrile, and Formamide in Water. *J. Phys. Chem.* **1985**, *89*, 1066−1069.
(127) Dullien, F. A. L.; Shemilt, L. W. Diffusion Coefficients for the Liquid System: Ethanol-Water. *Can. J. Chem. Eng.* **1961**, *39*, 242−247.
(128) Tominaga, T.; Matsumoto, S. Diffusion of Polar and Nonpolar Molecules in Water and Ethanol. *Bull. Chem. Soc. Jpn.* **1990**, *63*, 533−537.
(129) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; et al. *Gaussian 16*, Revision A.01; Gaussian Inc.: 2016.

https://doi.org/10.1021/acsphyschemau.1c00006
ACS Phys. Chem Au 2021, 1, 14−24