THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

Toward Automated Benchmarking of Atomistic Force Fields: Neat Liquid Densities and Static Dielectric Constants from the ThermoML Data Archive

Kyle A. Beauchamp,[*,†] Julie M. Behr,[‡] Ariën S. Rustenburg,[¶] Christopher I. Bayly,[§] Kenneth Kroenlein,[∥] and John D. Chodera[*,†]

[†]Computational Biology Program, Sloan Kettering Institute, Memorial Sloan Kettering Cancer Center, New York, New York 10065, United States

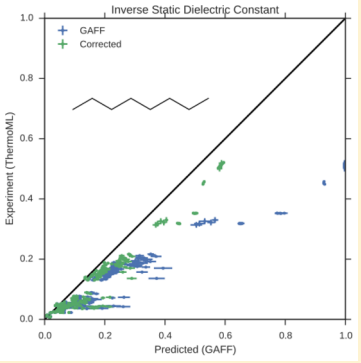
[‡]Tri-Institutional Program in Computational Biology and Medicine and [¶]Graduate Program in Physiology, Biophysics, and Systems Biology, Weill Cornell Medical College, New York, New York 10065, United States

[§]OpenEye Scientific Software Inc., Santa Fe, New Mexico 87508, United States

[∥]Thermodynamics Research Center, NIST, Boulder, Colorado 80305, United States

Supporting Information

ABSTRACT: Atomistic molecular simulations are a powerful way to make quantitative predictions, but the accuracy of these predictions depends entirely on the quality of the force field employed. Although experimental measurements of fundamental physical properties offer a straightforward approach for evaluating force field quality, the bulk of this information has been tied up in formats that are not machine-readable. Compiling benchmark data sets of physical properties from non-machine-readable sources requires substantial human effort and is prone to the accumulation of human errors, hindering the development of reproducible benchmarks of force-field accuracy. Here, we examine the feasibility of benchmarking atomistic force fields against the NIST ThermoML data archive of physicochemical measurements, which aggregates thousands of experimental measurements in a portable, machine-readable, self-annotating IUPAC-standard format. As a proof of concept, we present a detailed benchmark of the generalized Amber small-molecule force field (GAFF) using the AM1-BCC charge model against experimental measurements (specifically, bulk liquid densities and static dielectric constants at ambient pressure) automatically extracted from the archive and discuss the extent of data available for use in larger scale (or continuously performed) benchmarks. The results of even this limited initial benchmark highlight a general problem with fixed-charge force fields in the representation low-dielectric environments, such as those seen in binding cavities or biological membranes.



1. INTRODUCTION

Recent advances in hardware and software for molecular dynamics simulation now permit routine access to atomistic simulations at the 100 ns time scale and beyond.[1] Leveraging these advances in combination with consumer GPU clusters, distributed computing, or custom hardware has brought microsecond and millisecond simulation time scales within reach of many laboratories. These dramatic advances in sampling, however, have revealed deficiencies in force fields as a critical barrier to enabling truly predictive simulations of physical properties of biomolecular systems.

Protein and water force fields have been the subject of numerous benchmarks[2−4] and enhancements,[5−7] with key outcomes including the ability to fold fast-folding proteins,[8−10] improved fidelity of water thermodynamic properties,[11] and improved prediction of NMR observables. Although small-molecule force fields have also been the subject of benchmarks[12−14] and improvements,[15] such work has typically focused on small perturbations to specific functional groups. For example, a recent study found that modified hydroxyl nonbonded parameters led to the improved prediction of static dielectric constants and hydration free energies.[15] There are also outstanding questions of generalizability of these targeted perturbations; it is uncertain whether changes to the parameters for a specific chemical moiety will be compatible with seemingly unrelated improvements to other groups. Addressing these questions requires establishing community agreement upon shared benchmarks that can be easily replicated among laboratories, used to test proposed force-field enhancements, and expanded as the body of experimental data grows.

A key barrier to establishing reproducible and extensible force-field accuracy benchmarks is that many experimental data sets are heterogeneous, paywalled, and unavailable in machine-readable formats (although notable counterexamples exist, e.g.,

Received: July 13, 2015
Revised: August 24, 2015
Published: September 4, 2015

DOI: 10.1021/acs.jpcb.5b06703
J. Phys. Chem. B 2015, 119, 12912−12920

the Protein Data Bank,[16] FreeSolv,[17] and the Biological Magnetic Resonance Bank[18]). Although this inconvenience is relatively minor for benchmarking force-field accuracy for a single target system (e.g., water), it becomes prohibitive for studies spanning relevant chemical spaces, such as force fields intended to describe a large variety of druglike small organic molecules.

In addition to inconvenience, the number and kind of human-induced errors that can corrupt hand-compiled benchmarks are legion. A United States Geological Survey (USGS) case study examining the reporting and use of literature values of the aqueous solubility (S_w) and octanol−water partition coefficients (K_{ow}) for DDT and its persistent metabolite DDE provides remarkable insight into a variety of common errors.[19] Secondary sources are often cited as primary sources, a phenomenon that occurred up to five levels deep in the case of DDT and DDE; citations for data are often incorrect, misattributed to unrelated publications, or omitted altogether; and numerical data can be mistranscribed, transposed, or incorrectly converted among unit systems.[19] In the case of DDT and DDE, these errors occur to such a degree that the authors note that "strings of erroneous data compose as much as 41−73 percent of the total data".[19] Given the number and importance of these measurements, the quality of physicochemical data sets of lesser importance may be suspect.

To ameliorate problems of data archival, the National Institute of Standards and Technology (NIST) Thermodynamics Research Center (TRC) has developed an IUPAC-standard XML-based format, ThermoML[20−22], for storing physicochemical measurements, uncertainties, and metadata. Manuscripts containing new experimental measurements submitted to several journals (the Journal of Chemical Engineering Data, the Journal of Chemical Thermodynamics, Fluid Phase Equilibria, Thermochimica Acta, and the International Journal of Thermophysics) are guided through a data-archiving process that involves sanity checks, conversion to a standard machine-readable format, and archival at the TRC (http://trc.nist.gov/ThermoML.html).

Here, we examine the ThermoML archive as a potential source for a reproducible, extensible accuracy benchmark of biomolecular force fields. As a proof of concept, we concentrate on two important physical property measurements easily computable in many simulation codes, neat liquid density and static dielectric constant, with the goal of developing a standard benchmark for validating these properties in fixed-charge force fields of druglike molecules and biopolymer residue analogues. These two properties provide sensitive tests of force field accuracy that are nonetheless straightforward to calculate. Using these data, we evaluate the generalized Amber small-molecule force field (GAFF)[23,24] with the AM1-BCC charge model[25,26] and identify systematic biases to aid further force-field refinement.

2. METHODS

2.1. ThermoML Archive Retrieval and Processing. A snapshot of the ThermoML Archive was obtained from the NIST TRC on April 8, 2015. To explore the content of this archive, we created a Python (version 2.7.9) tool, ThermoPyL (https://github.com/choderalab/ThermoPyL), that formats the XML content into a spreadsheet-like format accessible via the Pandas (version 0.15.2) library. This tool also contains a preliminary version of scripts for maintaining an up-to-date version of the ThermoML archive. First, we obtained the XML schema (http://media.iupac.org/namespaces/ThermoML/ThermoML.xsd) defining the layout of the data. This schema was converted into a Python object via PyXB 1.2.4 (http://pyxb.sourceforge.net/). Finally, this schema was used to extract the data into Pandas[27] dataframes, and successive data filters described in section 3.1 were applied to explore the composition of the data.

2.2. Simulation. To enable automated accuracy benchmarking of physicochemical properties of neat liquids, such as mass density and static dielectric constant, we developed a semiautomated pipeline for preparing simulations, running them on a standard computer cluster using a portable simulation package, and analyzing the resulting data. All of the code for this procedure is available at https://github.com/choderalab/LiquidBenchmark. Below, we describe the operation of the various stages of this pipeline and their application to the benchmark reported here.

2.2.1. Preparation. Chemical names were parsed from the ThermoML extract and converted to both CAS and SMILES strings using CIRpy (https://github.com/mcs07/CIRpy). SMILES strings were converted into molecular structures using the OpenEye Python Toolkit version 2015-2-3,[28] as wrapped in openmoltools (http://github.com/choderalab/openmoltools).

Simulation boxes containing 1000 molecules were constructed using PackMol version 14-225[29] wrapped in the Python automation library openmoltools. To ensure stable automated equilibration, we chose PackMol box volumes to accommodate twice the volume of the enclosed atoms, with atomic radii estimated as 1.06 and 1.53 Å for hydrogens and nonhydrogens, respectively.

For this illustrative benchmark, we utilized the generalized Amber small-molecule force field (GAFF)[23,24] with the AM1-BCC charge model,[25,26] which we shall refer to as the GAFF/AM1-BCC force field.

Canonical AM1-BCC[25,26,30] charges were generated with the OpenEye Python Toolkit version 2015-2-3[28] using the `oequacpac.OEAssignPartialCharges` module with the `OECharges_AM1BCCSym` option, which utilizes a conformational expansion procedure (using `oeomega.OEOmega`[31]) prior to charge-fitting to minimize artifacts from intramolecular contacts. The `OEOmega` selected conformer was then processed using `antechamber` (with `parmchk2`) and `tleap` in AmberTools 14[32] to produce Amber-format `prmtop` and `inpcrd` files, which were then read into OpenMM to perform molecular simulations using the `simtk.openmm.app` module.

The simulations reported here used libraries openmoltools 0.6.4, OpenMM 6.3,[33] and MDTraj 1.3.[34] Exact commands to install various dependencies can be found in section S1.1 in the Supporting Information.

2.2.2. Equilibration and Production. All simulations were performed using OpenMM 6.3.[33] Simulation boxes were first minimized using the L-BFGS algorithm[35] using the `LocalEnergyMinimizer` default parameters and subsequently equilibrated for 10^7 steps with an equilibration time-step of 0.4 fs and a collision rate of 5 ps^{-1}. Production simulations employed a Langevin Leapfrog integrator[36] with a collision rate of 1 ps^{-1} and a 1 fs time-step, as we found that time-steps of 2 fs or greater led to a significant time-step dependence in computed equilibrium densities (Figure S1).

Equilibration and production simulations utilized a Metropolis Monte Carlo barostat with a control pressure of 1 atm

DOI: 10.1021/acs.jpcb.5b06703
J. Phys. Chem. B 2015, 119, 12912−12920

(101.325 kPa), utilizing molecular scaling and automated step-size adjustment during equilibration, with volume moves attempted every 25 steps. The particle mesh Ewald (PME) method with conducting boundary conditions[37] was used with a long-range cutoff of 0.95 nm, and a long-range isotropic dispersion correction was employed to correct for the truncation of Lennard-Jones interactions outside of the 0.95 nm cutoff. PME grid and spline parameters were automatically selected using the default settings in OpenMM 6.3 for the CUDA platform,[33] which was operated in default mixed-precision mode. Instantaneous densities were stored every 250 fs, and trajectory snapshots were stored every 5 ps.

Production simulations were continued until automatic analysis showed standard errors in densities were less than 2 $\times 10^{-4}$ g/cm^3. Automatic analysis of the production simulation data was run every 1 ns of simulation time, and utilized the `detectEquilibration` method[38] in the `timeseries` module of pymbar 2.1[39] to automatically discard the initial portion of the production simulation containing strong far-from-equilibrium configurations. This procedure selects the equilibration end point T_{eq} by maximizing the number of effectively uncorrelated samples in the remainder of the production simulation, $N_{eff} = (T - T_{equil})/g$, with T being the total simulation length. The statistical inefficiency g was determined by autocorrelation analysis using the fast-adaptive statistical inefficiency computation method implemented in the `timeseries.computeStatisticalInefficiency` method of pymbar 2.1 (with the algorithm is described in ref 40). This approach is essentially the same as the fixed-width procedure described by eq 7.12 of ref 41, with n^* equal to 4000 and the sequential testing correction (n^{-1} term) ignored due to the large value of n.

Statistical errors in estimated average density $\langle\rho\rangle$ were computed by the Markov chain standard error (MCSE)

$$\delta\rho \approx \frac{\mathrm{std}(\rho)}{\sqrt{N_{\mathrm{eff}}}} \tag{1}$$

where std(ρ) is the sample standard deviation of the density, and N_{eff} is the number of effectively uncorrelated samples.

Using this adaptive protocol, we found starting trajectory lengths of 12000 (8000, 16000) density frames (250 fs each), discarded regions of 28 (0, 460) density frames, and statistical inefficiencies of 20 (15, 28) density frames; the reported numbers indicate (median, (25% quartile, 75% quartile)).

2.3. Timings. The wall time required for a given simulation depends on the number of atoms in the simulation system (3000−29000 atoms), the GPU used (GTX 680 or GTX Titan), and the time required for automated termination. For butyl acrylate (21000 atoms) on a GTX Titan, the wall-clock performance is approximately 80 ns/day. Using a value of 80 ns/day with approximately 3 ns of production simulation corresponds to 1 h for the production segment of the simulation and 3 h for the fixed equilibration portion of 10^7 steps.

2.3.1. Data Analysis and Statistical Error Estimation. Trajectory analysis was performed using OpenMM 6.3[33] and MDTraj 1.3.[34]

Mass Density. Mass density ρ was computed via the relation

$$\rho = \left\langle \frac{M}{V} \right\rangle \tag{2}$$

where M is the total mass of all particles in the system, and V is the instantaneous volume of the simulation box.

Static Dielectric Constants. Static dielectric constants were calculated using the dipole fluctuation approach appropriate for PME with conducting ("tinfoil") boundary conditions,[11,42] with the total system box dipole μ computed from trajectory snapshots using MDTraj 1.3.[34]

$$\epsilon = 1 + \beta\frac{4\pi}{3}\cdot\frac{\langle\mu\cdot\mu\rangle - \langle\mu\rangle\cdot\langle\mu\rangle}{\langle V\rangle} \tag{3}$$

where $\beta \equiv 1/(k_B\cdot T)$ is the inverse temperature.

Computation of Expectations. Expectations were estimated by computing sample means over the production simulation after discarding the initial far-from-equilibrium portion to equilibration (as described in the automatic termination criteria above).

Statistical Uncertainties. For density uncertainties, the Markov chain standard error (MCSE) was estimated using eq 1. For dielectric uncertainties, to avoid the complexities of computing and propagating correlated errors in eq 3, we employed a bootstrap procedure: the portion of the production simulation not discarded to equilibration was used as input to a circular-block bootstrapping procedure,[43] with block sizes automatically selected to maximize the error.[44]

2.3.2. Code Availability. All code to perform data analysis and create figures for this work, as well as all intermediate data (except configurational trajectories, due to their large size), is available at https://github.com/choderalab/LiquidBenchmark.

3. RESULTS

3.1. Extracting Neat Liquid Measurements from the NIST TRC ThermoML Archive. As described in section 2.1, we retrieved a copy of the ThermoML Archive on Apr 8, 2015 and performed a number of sequential filtering steps to produce a ThermoML extract relevant for benchmarking force fields describing small organic molecules. As our aim is to explore neat liquid data with functional groups relevant to biopolymers and druglike molecules, we applied the following ordered filters, starting with all data containing density or static dielectric constants:

1. The measured sample contains only a single component (pure substances)
2. The molecule contains only druglike elements (defined here as H, N, C, O, S, P, F, Cl, and Br)
3. The molecule has ≤10 nonhydrogen atoms
4. The measurement was performed in a biophysically relevant temperature range ($270 \le T$ [K] ≤ 330)
5. The measurement was performed at ambient pressure ($100 \le P$ [kPa] ≤ 102)
6. Only measurements in the liquid phase were retained
7. The temperature and pressure were rounded to nearby values (as described below), averaging all measurements within each group of like conditions
8. Only conditions (molecule, temperature, and pressure) for which *both* density and dielectric constants were available were retained

The temperature-and-pressure rounding step was motivated by common data reporting variations; for example, an experiment performed at the freezing temperature of water and ambient pressure might be entered as either 101.325 or 100 kPa, with a temperature of either 273 or 273.15 K. Therefore, all pressures within the range of $100 \le P \le 102$ (in kPa) were rounded to

DOI: 10.1021/acs.jpcb.5b06703
J. Phys. Chem. B 2015, 119, 12912−12920

exactly 1 atm (101.325 kPa). Temperatures were rounded to one decimal place in K.

The application of these filters (Table 1) leaves 246 conditions, where a *condition* here indicates a molecule, temperature, and pressure tuple for which both density and dielectric data are available. The functional groups present in the resulting data set are summarized in Table 2; see section 2.1 for further description of the software pipeline used.

Table 1. Successive Filtration of the ThermoML Archive[a]

	number of measurements remaining	
filter step	mass density	static dielectric
1. single component	136 212	1651
2. druglike elements	125 953	1651
3. heavy atoms	71 595	1569
4. temperature	38 821	964
5. pressure	14 103	461
6. liquid state	14 033	461
7. aggregate T, P	3592	432
8. density and dielectric	246	246

[a]A set of successive filters were applied to all measurements in the ThermoML archive that contained either mass density or static dielectric constant measurements. Each column reports the number of measurements remaining after successive application of the corresponding filtration step. The 246 final measurements correspond to 45 unique molecules measured at several temperature conditions.

3.2. Benchmarking GAFF/AM1-BCC against the ThermoML Archive. *3.2.1. Mass Density.* Mass densities of bulk liquids have been widely used for parametrizing and testing force fields, particularly the Lennard-Jones parameters representing dispersive and repulsive interactions.[46,47] We therefore used the present ThermoML extract as a benchmark of the GAFF/AM1-BCC force field (Figure 1).

Table 2. Functional Groups Present in Filtered Dataset[a]

functional group	occurrences
1,2-aminoalcohol	4
1,2-diol	3
alkene	3
aromatic compound	1
carbonic acid diester	2
carboxylic acid ester	4
dialkyl ether	7
heterocyclic compound	3
ketone	3
lactone	1
primary alcohol	19
primary aliphatic amine (alkylamine)	2
primary amine	2
secondary alcohol	4
secondary aliphatic amine (dialkylamine)	2
secondary aliphatic−aromatic amine (alkylarylamine)	1
secondary amine	3
sulfone	1
sulfoxide	1
tertiary aliphatic amine (trialkylamine)	3
tertiary amine	3

[a]The filtered ThermoML dataset contained 246 distinct (molecule, temperature, pressure) conditions, spanning 45 unique compounds. The functional groups represented in these compounds (as identified by the program `checkmol` v0.5)[45] is summarized here.

Overall Accuracy. Overall, the densities show reasonable accuracy, with a root-mean square (RMS) relative error over all measurements of (3.0 ± 0.1)%, especially encouraging given that this force field was not designed with the intention of modeling bulk liquid properties of organic molecules.[23,24] This is reasonably consistent with previous studies reporting a relative error of 4% on a different benchmark set.[12]

Temperature Dependence. For a given compound, the signs of the errors typically do not change at different temperatures (Figures 1 and S4). Furthermore, the magnitudes of the error also remain largely constant (vertical lines in Figure 1B), although several exceptions do occur. It is possible that these systematic density offsets indicate correctable biases in force field parameters.

Outliers. The largest density errors occur for a number of oxygen-containing compounds: 1,4-dioxane; 2,5,8-trioxanonane; 2-aminoethanol; dimethyl carbonate; formamide; and water (Figure S4). The absolute error on these poor predictions is on the order of 0.05 g/cm^3, which is substantially higher than the measurement error (≤0.008 g/cm^3; see Figure S2).

We note that our benchmark includes a GAFF/AM1-BCC model for water due to our desire to automate benchmarks against a force field capable of modeling a large variety of small molecular liquids. Water, an incredibly important solvent in biomolecular systems, is generally treated with a special-purpose model (such as TIP3P[46] or TIP4P-Ew[11]) parametrized to fit a large quantity of thermophysical data. As expected, the GAFF/AM1-BCC model performs poorly in reproducing liquid densities for this very special solvent. We conclude that it remains highly advisible that the field continue to use specialized water models when possible.

3.2.2. Static Dielectric Constant. Overall Accuracy. As a measure of the dielectric response, the static dielectric constant of neat liquids provides a critical benchmark of the accuracy of electrostatic treatment in force field models. Discussing the accuracy in terms of the ability of GAFF/AM1-BCC to reproduce the static dielectric constant ϵ is not necessarily meaningful because of the way that the solvent dielectric ϵ enters into the Coulomb potential between two point charges separated by a distance r,

$$U(r) = \frac{q_1 q_2}{\epsilon r} \propto \frac{1}{\epsilon} \tag{4}$$

It is evident that $1/\epsilon$ is a much more meaningful quantity to compare than ϵ directly, as a 5% error in $1/\epsilon$ will cause a 5% error in the Coulomb potential between two point charges (assuming a uniform dielectric), while a 5% error in ϵ will have a much more complex ϵ-dependent effect on the Coulomb potential. We therefore compare simulations against measurements in our ThermoML extract on the $1/\epsilon$ scale in Figure 2.

Systematic Underestimation of the Dielectric Constants of Nonpolar Liquids. Overall, we find the dielectric constants to be qualitatively reasonable but with clear deviations from experiment, particularly for nonpolar liquids. This is not surprising given the complete neglect of electronic polarization that will be the dominant contribution for such liquids. In particular, GAFF/AM1-BCC systematically underestimates the dielectric constants for nonpolar liquids, with the predictions of $\epsilon \approx 1.0$ being substantially smaller than the measured value of $\epsilon \approx 2$. Because this deviation likely stems from the lack of an explicit treatment of electronic polarization, we used a simple

DOI: 10.1021/acs.jpcb.5b06703
J. Phys. Chem. B 2015, 119, 12912−12920

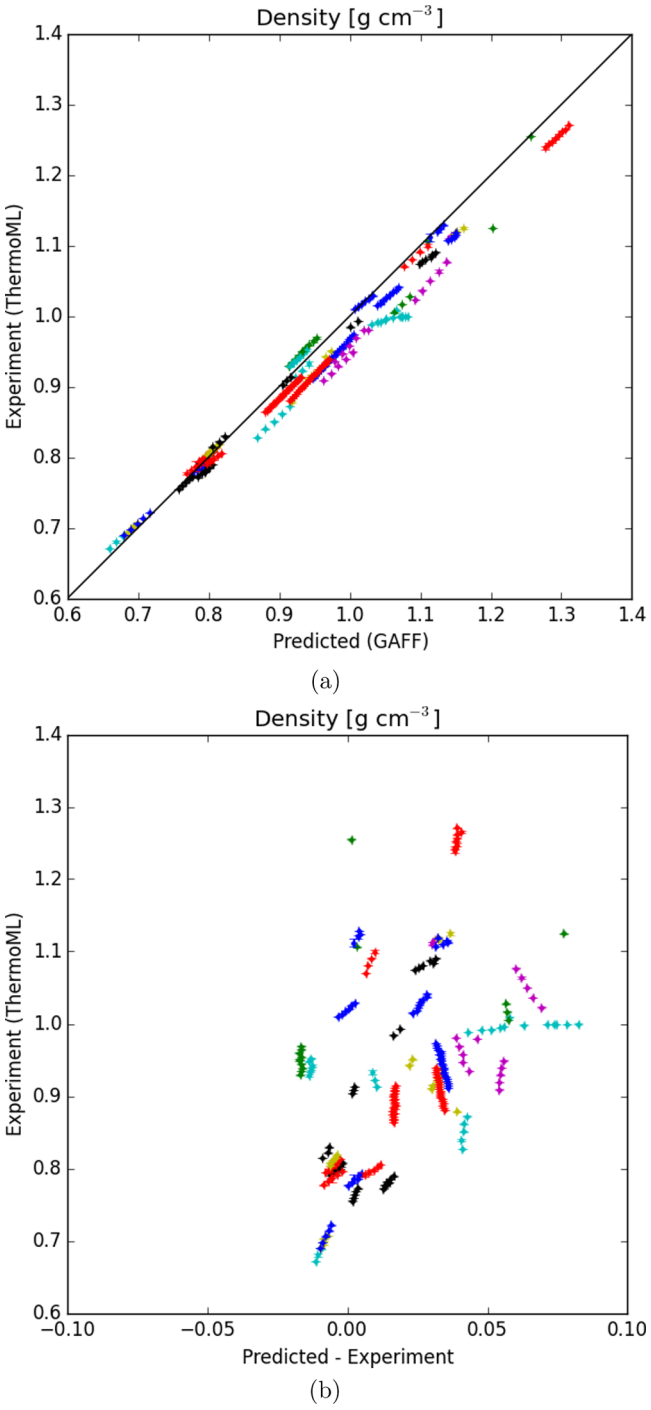


Figure 1. Comparison of liquid densities between experiment and simulation. (a) Liquid-density measurements extracted from ThermoML are compared against densities predicted using the GAFF/AM1-BCC small-molecule fixed-charge force field. Color groupings represent identical chemical species, although the color map repeats itself due to the large number (45) of unique compounds. Plots of density vs temperature grouped by chemical species are available in Figure S4. Simulation error bars represent one standard error of the mean, with the number of effective (uncorrelated) samples estimated using pymbar. Experimental error bars indicate the standard deviation between independently reported measurements, when available, or author-reported standard deviations in ThermoML entries; for some measurements, neither uncertainty estimate is available. See Figure S2 for further discussion of error. (b) The same plot with the residual (predicted minus experimental values) on the x-axis. Note that the error bars are all smaller than the symbols.

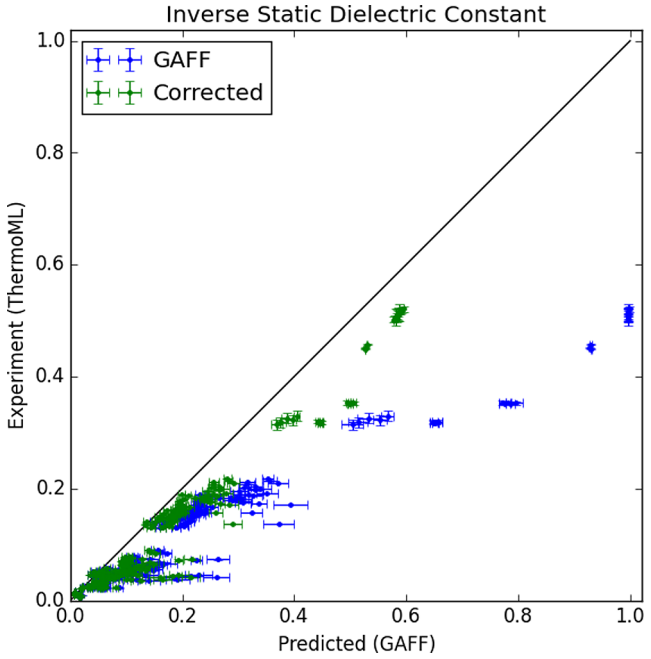


Figure 2. Measured (ThermoML) vs predicted (GAFF/AM1-BCC) inverse static dielectrics. Simulation error bars represent one standard error of the mean. Experimental error bars indicate the larger of standard deviation between independently reported measurements and the authors' reported standard deviations; for some measurements, neither uncertainty estimate is available. See Figure S3 for further discussion of error. See section 3.2.2 for an explanation of why the inverse dielectric constant (rather than the dielectric constant) is plotted. For nonpolar liquids, it is clear that the force field predicts electrostatic interactions that are substantially biased by missing polarizability. Plots of dielectric constant vs temperature grouped by chemical species are available in Figure S5.

empirical polarization model that computes the molecular electronic polarizability α as a sum of elemental atomic polarizability contributions.[48]

From the computed molecular electronic polarizability α, an additive correction to the simulation-derived static dielectric constant accounting for the missing electronic polarizability can be computed:[11]

$$\Delta\epsilon = 4\pi N\frac{\alpha}{\langle V\rangle} \quad (5)$$

A similar polarization correction was used in the development of the TIP4P-Ew water model, where it had a minor effect[11] because almost all of the high static dielectric constants for water comes from the configurational response of its strong dipole. However, the missing polarizability is a dominant contribution to the static dielectric constant of nonpolar organic molecules; in the case of water, the empirical atomic polarizability model predicts a dielectric correction (to ϵ) of 0.52, while 0.79 was used for the TIP4P-Ew model. Averaging this dielectric constant over all liquids in the present work leads to average polarizability corrections (to ϵ) of 0.74 ± 0.08. Taking the data set as a whole, we find that the relative error in uncorrected dielectric (to ϵ) is on the order of −0.34 ± 0.02, as compared to −0.25 ± 0.02 for the corrected dielectric.

4. DISCUSSION

4.1. Mass Densities. Our simulations have indicated the presence of systematic density biases with magnitudes larger

DOI: 10.1021/acs.jpcb.5b06703
J. Phys. Chem. B 2015, 119, 12912−12920

than the measurement error. Correcting these errors may be a low-hanging fruit for future force-field refinements. As an example of the feasibility of improved accuracy in densities, a recent three-point water model was able to recapitulate water density with errors of less than 0.005 g/cm^3 over the temperature range [280 to 320 K].[49] This improved accuracy in density prediction was obtained alongside accurate predictions of other experimental observables, including the static dielectric constant. We suspect that such accuracy might be obtainable for GAFF-like force fields across some portion of chemical space. A key challenge for the field is to demarcate the fundamental limit of fixed-charge force fields for predicting orthogonal classes of experimental observables. For example, is it possible to achieve a relative density error of 10^{-4} without sacrificing accuracy of other properties, such as enthalpies of vaporization? In our opinion, the best way to answer such questions is to systematically build force fields with the goal of predicting various properties to within their known experimental uncertainties, similar to what has been done for water.[11,49]

4.2. Dielectric Constants in Force Field Parametrization. A key feature of the static dielectric constant for a liquid is that, for force-field purposes, it consists of two very different contributions, distinguished by the dependence on the fixed charges of the force field and the dynamic motion of the molecule. One component, the electronic polarizability (which can be separately quantified through the high-frequency dielectric constant), arises from the almost-instantaneous electronic polarization in response to the external electric field. Electronic polarizability contributes a small component, generally around $\epsilon = 2$, which can be dominant for nonpolar liquids but is completely neglected by the nonpolarizable force fields in common use for biomolecular simulations.

The other component arises from the dynamical response of the molecule through reorientation or conformational relaxation via nuclear motion. For small polar liquids that lack significant internal degrees of freedom, such as water, reorientation of various molecular multipoles in response to the external electric field contributes the majority of the static dielectric constant. As a result, for polar liquids, we expect the parametrized atomic charges to play a major role in the static dielectric.

Recent force field development has seen a resurgence of papers fitting dielectric constants during force field parametrization.[15,49] However, a number of authors have pointed out potential challenges in constructing self-consistent fixed-charge force fields.[50,51]

Interestingly, recent work by Dill and co-workers[50] observed that, for CCl_4, reasonable choices of point charges are incapable of recapitulating the observed dielectric of $\epsilon = 2.2$, instead producing dielectric constants in the range of $1.0 \leq \epsilon \leq 1.05$. This behavior is quite general: fixed point charge force fields will predict $\epsilon \approx 1$ for many nonpolar or symmetric molecules, but the measured dielectric constants are instead $\epsilon \approx 2$ (Figure 3). Although this behavior is well-known and results from missing physics of polarizability, we suspect it may have several profound consequences, which we discuss below.

Suppose, for example, that one attempts to fit force-field parameters to match the static dielectric constants of CCl_4, $CHCl_3$, CH_2Cl_2, and CH_3Cl. In moving from the tetrahedrally symmetric CCl_4 to the asymmetric $CHCl_3$, it suddenly becomes possible to achieve the observed dielectric constant of 4.8 by an appropriate choice of point charges. However, the model for

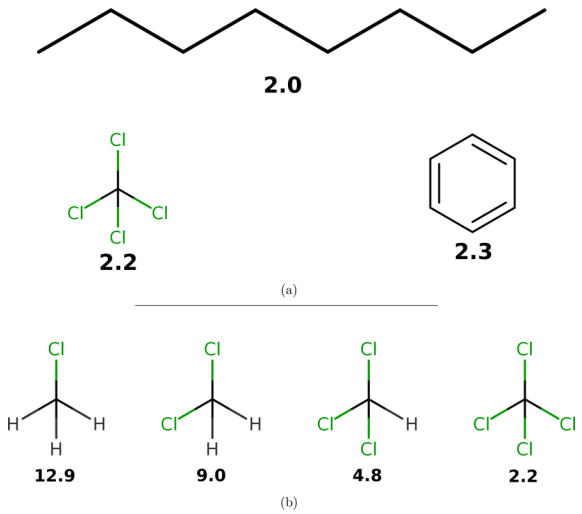


Figure 3. Typical experimental static dielectric constants of some nonpolar compounds. (a) Measured static dielectric constants of various nonpolar or symmetric molecules.[52,53] Fixed-charge force fields give $\epsilon \approx 1$ for each species; for example, we calculated $\epsilon = 1.0030 \pm 0.0002$ for octane. (b) A congeneric series of chloro-substituted methanes have static dielectric constants between 2 and 13. Reported dielectric constants are at near-ambient temperatures.

$CHCl_3$ uses fixed point charges to account for *both* the permanent dipole moment and the electronic polarizability, whereas the CCl_4 model contains no treatment of polarizability. We hypothesize that this inconsistency in parametrization may lead to discontinuous jumps in physical properties in related molecular series, where symmetric molecules (e.g., benzene and CCl_4) have qualitatively different properties than closely related asymmetric molecules (e.g., toluene and $CHCl_3$).

How important is this effect? We expect it to be important wherever we encounter the transfer of a polar molecule (such as a peptide, native ligand, or pharmaceutical small molecule) from a polar environment (such as the cytosol, interstitial fluid, or blood) into a nonpolar environment (such as a biological membrane or nonpolar binding site of an enzyme or receptor). Thus, we expect this to be implicated in biological processes ranging from ligand binding to absorption and distribution within the body. To understand this conceptually, consider the transfer of a polar small-molecule transfer from the nonpolar interior of a lipid bilayer to the aqueous and, hence, very polar cytosol.

As a real-world example, we imagine that the missing atomic polarizability could be important in determining accurate transfer free energies involving low-dielectric solvents, such as the small-molecule transfer free energy from octanol or cyclohexane to water. The Onsager model for solvation of a dipole μ of radius a gives us a way to estimate the magnitude of error introduced by making an error of $\Delta\epsilon$ in the static dielectric constant of a solvent. The free energy of dipole solvation is given by this model as

$$\Delta G = -\frac{\mu^2}{a^3}\cdot\frac{\epsilon - 1}{2\epsilon + 1} \tag{6}$$

such that, for an error of $\Delta\epsilon$ departing from the true static dielectric constant ϵ, we find the error in solvation is

$$\Delta\Delta G = -\frac{\mu^2}{a^3}\left[\frac{(\epsilon + \Delta\epsilon) - 1}{2(\epsilon + \Delta\epsilon) + 1} - \frac{\epsilon - 1}{2\epsilon + 1}\right] \tag{7}$$

DOI: 10.1021/acs.jpcb.5b06703
J. Phys. Chem. B 2015, 119, 12912−12920

For example, the solvation of water (a = 1.93 Å, μ = 2.2 D) in a low-dielectric medium such as tetrachloromethane or benzene ($\epsilon \sim 2.2$, but $\Delta\epsilon = -1.2$) gives an error of $\Delta\Delta G \sim -8$ kJ/mol (−2 kcal/mol).

Implications for Transfer Free Energies. As another example, consider the transfer of small druglike molecules from a nonpolar solvent (such as cyclohexane) to water, a property often measured to indicate the expected degree of lipophilicity of a compound. To estimate the magnitude of error expected, for each molecule in the FreeSolv database (as of February 20, 2015),[17] we estimated the expected error in computed transfer free energies should GAFF/AM1-BCC be used to model the nonpolar solvent cyclohexane using the Onsager model (eq 7). We took the cavity radius a to be half the maximum interatomic distance and calculated $\mu = \sum_i q_i r_i$ using the provided mol2 coordinates and AM1-BCC charges. This calculation predicts a mean error of (−3.8 ± 0.3) kJ/mol [(−0.91 ± 0.07) kcal/mol] for the 643 molecules (where the standard error is computed from bootstrapping over FreeSolv compound measurements), suggesting that the missing atomic polarizabilty unrepresentable by fixed point charge force fields could contribute substantially to errors in predicted transfer and solvation properties of druglike molecules. In other words, the use of a fixed-charge physics may lead to errors of 3.8 kJ/mol in cyclohexane transfer free energies. We conjecture that these missing physics will be important in the upcoming (2015) SAMPL challenge,[54] which will examine transfer free energies in several low-dielectric media.

Utility in Parametrization. Given their ease of measurement and direct connection to long-range electrostatic interactions, static dielectric constants have high potential utility as primary data for force-field parametrization efforts. Although this will require the use of force fields with explicit treatment of atomic polarizability, the inconsistency of fixed-charge models in low-dielectric media is sufficiently alarming to motivate further study of polarizable force fields. In particular, continuum methods,[55−57] point dipole methods,[58,59] and Drude methods[60,61] have been maturing rapidly. Finding the optimal balance of accuracy and performance remains an open question; however, the use of experimentally parametrized direct polarization methods[62] may provide polarizability physics at a cost not much greater than that of fixed charge force fields. This is not to say that we suggest an immediate transition to polarizable force fields; the efficiency benefits and pervasiveness of fixed-charge models are important. Furthermore, empirical corrections such as overpolarized charges[50] and others[51] may provide the ability to model some low-dielectric behaviors at the cost of some transferability.

4.3. ThermoML as a Data Source. The present work has focused on the neat liquid density and dielectric measurements present in the ThermoML Archive[20,21,63] as a target for molecular-dynamics force-field validation. While liquid mass densities and static dielectric constants have already been widely used in force field work, several aspects of ThermoML make it a unique resource for the force-field community. First, the aggregation, support, and dissemination of ThermoML data sets through the ThermoML archive is supported by NIST, whose mission makes these tasks a long-term priority. Second, the ThermoML archive is actively growing through partnerships with several journals, and the new experimental measurements published in these journals are critically examined by the TRC and included in the archive. Finally, the files in the ThermoML Archive are portable and machine-readable via a formal XML schema, allowing facile access to hundreds of thousands of measurements. Numerous additional physical properties contained in ThermoML (including activity coefficients, diffusion constants, boiling-point temperatures, critical pressures and densities, coefficients of expansion, speed-of-sound measurements, viscosities, excess molar enthalpies, heat capacities, and volumes) for neat phases and mixtures represent a rich data set of high utility for force field validation and parametrization.

5. CONCLUSIONS

High-quality, machine-readable data sets of physicochemical measurements will be required for the construction of next-generation small-molecule force fields. Here we have discussed the NIST ThermoML archive as a growing source of physicochemical measurements that may be useful for the force field community. From the NIST ThermoML archive, we selected a data set of 246 ambient, neat liquid systems for which both densitt and static dielectric constant data are available. Using this data set, we benchmarked GAFF/AM1-BCC, one of the most popular small-molecule force fields and found systematic biases in densities and particularly in static dielectric constants. Element-based empirical polarizabilty models are able to account for much of the systematic differences between the GAFF/AM1-BCC model and the experiment. Non-polarizable force fields may show unacceptable biases when predicting transfer and binding properties of nonpolar environments, such as binding cavities or membranes.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcb.5b06703.

Figures showing the time-step dependence of density, error analyses for the ThermoML dataset, and the temperature dependence of density and the static dielectric constant; and commands to install dependencies. (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*kyle.beauchamp@choderalab.org.
*john.chodera@choderalab.org.

Notes

Source code for ThermoPyL, the ThermoML Archive retrieval and processing tools described in Section 2.1, can be found at https://github.com/choderalab/ThermoPyL. All source code for the automated ThermoML benchmark described in this paper can be found at https://github.com/choderalab/LiquidBenchmark.

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Patrick B. Grinaway (MSKCC), Vijay S. Pande (Stanford University), Lee-Ping Wang (Stanford University), Peter Eastman (Stanford University), Robert McGibbon (Stanford University), Jason Swails (Rutgers University), David L. Mobley (University of California, Irvine), Michael R. Shirts (University of Virginia), William C. Swope (IBM), Julia E. Rice (IBM), Hans Horn (IBM), Jed W. Pitera (IBM), and members of the Chodera lab for helpful discussions. Support for J.M.B. was provided by the Tri-Institutional

DOI: 10.1021/acs.jpcb.5b06703
J. Phys. Chem. B 2015, 119, 12912−12920

Training Program in Computational Biology and Medicine (via NIH training grant 1T32GM083937). K.A.B. was supported in part by Starr Foundation grant I8-A8-058. J.D.C. and K.A.B. acknowledge partial support from NIH grant P30 CA008748. K.A.B., J.M.B., A.S.R., and J.D.C. acknowledge the generous support of this research by the Sloan Kettering Institute. The authors gratefully acknowledge OpenEye Scientific for generously providing their toolkit for use in noncommercial projects that generate results for the public domain. This contribution of the National Institute of Standards and Technology (NIST) is not subject to copyright in the United States. Products or companies named here are cited only in the interest of complete technical description, and neither constitute nor imply endorsement by NIST or by the U.S. government. Other products may be found to serve as well.

REFERENCES

(1) Salomon-Ferrer, R.; Gotz, A. W.; Poole, D.; Le Grand, S.; Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *J. Chem. Theory Comput.* **2013**, *9*, 3878−3888.

(2) Lindorff-Larsen, K.; Maragakis, P.; Piana, S.; Eastwood, M.; Dror, R.; Shaw, D. Systematic Validation of Protein Force Fields against Experimental Data. *PLoS One* **2012**, *7*, e32131.

(3) Beauchamp, K.; Lin, Y.; Das, R.; Pande, V. Are Protein Force Fields Getting Better? A Systematic Benchmark on 524 Diverse NMR Measurements. *J. Chem. Theory Comput.* **2012**, *8*, 1409−1414.

(4) Best, R.; Buchete, N.; Hummer, G. Are Current Molecular Dynamics Force Fields too Helical? *Biophys. J.* **2008**, *95*, L07−L09.

(5) Li, D.-W.; Bruschweiler, R. Iterative Optimization of Molecular Mechanics Force Fields from NMR Data of Full Length Proteins. *J. Chem. Theory Comput.* **2011**, *7*, 1773−1782.

(6) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E.; Mittal, J.; Feig, M.; MacKerell, A. D. Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone φ, ψ and Side-Chain χ1 and χ2 Dihedral Angles. *J. Chem. Theory Comput.* **2012**, *8*, 3257−3273.

(7) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J.; Dror, R.; Shaw, D. Improved Side-Chain Torsion Potentials for the Amber ff99SB Protein Force Field. *Proteins: Struct., Funct., Genet.* **2010**, *78*, 1950−1958.

(8) Lindorff-Larsen, K.; Piana, S.; Dror, R.; Shaw, D. How Fast-Folding Proteins Fold. *Science* **2011**, *334*, 517−520.

(9) Ensign, D.; Kasson, P.; Pande, V. Heterogeneity Even at the Speed Limit of Folding: Large-scale Molecular Dynamics Study of a Fast-folding Variant of the Villin Headpiece. *J. Mol. Biol.* **2007**, *374*, 806−816.

(10) Voelz, V.; Bowman, G.; Beauchamp, K.; Pande, V. Molecular Simulation of ab Initio Protein Folding for a Millisecond Folder NTL9 (1- 39). *J. Am. Chem. Soc.* **2010**, *132*, 1526−1528.

(11) Horn, H.; Swope, W.; Pitera, J.; Madura, J.; Dick, T.; Hura, G.; Head-Gordon, T. Development of an Improved Four-site Water Model for Biomolecular Simulations: TIP4P-Ew. *J. Chem. Phys.* **2004**, *120*, 9665−9678.

(12) Caleman, C.; van Maaren, P. J.; Hong, M.; Hub, J. S.; Costa, L. T.; van der Spoel, D. Force Field Benchmark of Organic Liquids: Density, Enthalpy of Vaporization, Heat Capacities, Surface Tension, Isothermal Compressibility, Volumetric Expansion Coefficient, and Dielectric Constant. *J. Chem. Theory Comput.* **2012**, *8*, 61−74.

(13) Fischer, N. M.; van Maaren, P. J.; Ditz, J. C.; Yildirim, A.; van der Spoel, D. Properties of Organic Liquids when Simulated with Long-Range Lennard-Jones interactions. *J. Chem. Theory Comput.* **2015**, *11*, 2938−2944.

(14) Zhang, J.; Tuguldur, B.; van der Spoel, D. Force Field Benchmark of Organic Liquids II: Gibbs Energy of Solvation. *J. Chem. Inf. Model.* **2015**, *55*, 1192−1201.

(15) Fennell, C. J.; Wymer, K. L.; Mobley, D. L. A Fixed-Charge Model for Alcohol Polarization in the Condensed Phase, and Its Role in Small Molecule Hydration. *J. Phys. Chem. B* **2014**, *118*, 6438−6446.

(16) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235−242.

(17) Mobley, D. L. Experimental and Calculated Small Molecule Hydration Free Energies. http://www.escholarship.org/uc/item/6sd403pz, UC Irvine Department of Pharmaceutical Sciences (accessed Feb 20, 2015).

(18) Ulrich, E.; Akutsu, H.; Doreleijers, J.; Harano, Y.; Ioannidis, Y.; Lin, J.; Livny, M.; Mading, S.; Maziuk, D.; Miller, Z.; et al. BioMagResBank. *Nucleic Acids Res.* **2008**, *36*, D402−D408.

(19) Pontolillo, J.; Eganhouse, R. P. *The Search for Reliable Aqueous Solubility (S_w) and Octanol-Water Partition Coefficient (K_{ow}) Data for Hydrophobic Organic Compounds: DDT and DDE as a Case Study*; Water-Resources Investigations Report 01-4201; U.S. Geological Survey: Reston, VA, 2001.

(20) Frenkel, M.; Chirico, R. D.; Diky, V. V.; Dong, Q.; Frenkel, S.; Franchois, P. R.; Embry, D. L.; Teague, T. L.; Marsh, K. N.; Wilhoit, R. C. ThermoML an XML-based Approach for Storage and Exchange of Experimental and Critically Evaluated Thermophysical and Thermochemical Property Data. 1. Experimental data. *J. Chem. Eng. Data* **2003**, *48*, 2−13.

(21) Frenkel, M.; Chiroco, R. D.; Diky, V.; Dong, Q.; Marsh, K. N.; Dymond, J. H.; Wakeham, W. A.; Stein, S. E.; Königsberger, E.; Goodwin, A. R. XML-based IUPAC Standard for experimental, Predicted, and Critically Evaluated Thermodynamic Property Data Storage and Capture (ThermoML) (IUPAC Recommendations 2006). *Pure Appl. Chem.* **2006**, *78*, 541−612.

(22) Chirico, R. D.; Frenkel, M.; Magee, J. W.; Diky, V.; Muzny, C. D.; Kazakov, A. F.; Kroenlein, K.; Abdulagatov, I.; Hardin, G. R.; Acree, W. E., Jr; et al. Improvement of Quality in Publication of Experimental Thermophysical Property Data: Challenges, Assessment Tools, Global Implementation, and Online Support. *J. Chem. Eng. Data* **2013**, *58*, 2699−2716.

(23) 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.

(24) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic Atom Type and Bond Type Perception in Molecular Mechanics Calculations. *J. Mol. Graphics Modell.* **2006**, *25*, 247−260.

(25) Jakalian, A.; Bush, B. L.; Jack, D. B.; Bayly, C. I. Fast, Efficient Generation of High-quality Atomic Charges. AM1-BCC Model: I. Method. *J. Comput. Chem.* **2000**, *21*, 132−146.

(26) Jakalian, A.; Jack, D. B.; Bayly, C. I. Fast, Efficient Generation of High-quality Atomic Charges. AM1-BCC Model: II. Parameterization and validation. *J. Comput. Chem.* **2002**, *23*, 1623−1641.

(27) McKinney, W. Data Structures for Statistical Computing in Python. In *Proceedings of the 9th Python in Science Conference*, Austin, Texas, 2010; pp 51−56.

(28) *OpenEye Toolkit*, version 2014; OpenEye Scientific: Santa Fe, NM, 2014; http://www.eyesopen.com.

(29) Martínez, L.; Andrade, R.; Birgin, E. G.; Martínez, J. M. Packmol: A Package for Building Initial Configurations for Molecular Dynamics Simulations. *J. Comput. Chem.* **2009**, *30*, 2157−2164.

(30) Velez-Vega, C.; McKay, D. J.; Aravamuthan, V.; Pearlstein, R.; Duca, J. S. Time-Averaged Distributions of Solute and Solvent Motions: Exploring Proton Wires of GFP and PfM2DH. *J. Chem. Inf. Model.* **2014**, *54*, 3344−3361.

(31) Hawkins, P. C.; Nicholls, A. Conformer Generation with OMEGA: Learning from the Data Set and the Analysis of Failures. *J. Chem. Inf. Model.* **2012**, *52*, 2919−2936.

(32) Case, D.; Babin, V.; Berryman, J.; Betz, R.; Cai, Q.; Cerutti, D.; Cheatham, T.III; Darden, T.; Duke, R.; Gohlke, H. et al. *AMBER 14*; University of California: San Francisco, CA, 2014.

(33) Eastman, P.; Friedrichs, M. S.; Chodera, J. D.; Radmer, R. J.; Bruns, C. M.; Ku, J. P.; Beauchamp, K. A.; Lane, T. J.; Wang, L.-P.; Shukla, D.; et al. OpenMM 4: A Reusable, Extensible, Hardware

DOI: 10.1021/acs.jpcb.5b06703
J. Phys. Chem. B 2015, 119, 12912−12920

Independent Library for High Performance Molecular Simulation. *J. Chem. Theory Comput.* **2012**, *9*, 461−469.
(34) McGibbon, R. T.; Beauchamp, K. A.; Schwantes, C. R.; Wang, L.-P.; Hernández, C. X.; Harrigan, M. P.; Lane, T. J.; Swails, J. M.; Pande, V. S. MDTraj: a Modern, Open Library for the Analysis of Molecular Dynamics Trajectories. 2014; http://biorxiv.org/content/early/2014/09/09/008896.short, accessed Aug. 27, 2015.
(35) Liu, D. C.; Nocedal, J. On the Limited Memory BFGS Method for Large Scale Optimization. *Mathematical programming* **1989**, *45*, 503−528.
(36) Izaguirre, J. A.; Sweet, C. R.; Pande, V. S. Multiscale Dynamics of Macromolecules using Normal Mode Langevin. In *Pacific Symposium on Biocomputing*, Waimei, HI, 2010.
(37) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N Log (N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(38) Chodera, J. D. A Simple Method for Automated Equilibration Detection in Molecular Simulations. *bioRxiv* [Online] **2015**, http://biorxiv.org/content/early/2015/07/04/021659.
(39) Shirts, M. R.; Chodera, J. D. Statistically Optimal Analysis of Samples from Multiple Equilibrium States. *J. Chem. Phys.* **2008**, *129*, 124105.
(40) Chodera, J. D.; Singhal, N.; Pande, V. S.; Dill, K. A.; Swope, W. C. Automatic Discovery of Metastable States for the Construction of Markov models of Macromolecular Conformational Dynamics. *J. Chem. Phys.* **2007**, *126*, 155101.
(41) Brooks, S.; Gelman, A.; Jones, G.; Meng, X.-L. *Handbook of Markov Chain Monte Carlo*; CRC Press: Boca Raton, FL, 2011.
(42) Neumann, M. Dipole Moment Fluctuation Formulas in Computer Simulations of Polar Systems. *Mol. Phys.* **1983**, *50*, 841−858.
(43) Sheppard, K. ARCH Toolbox for Python. http://dx.doi.org/10.5281/zenodo.15681; GitHub repository: https://github.com/bashtage/arch.
(44) Flyvbjerg, H.; Petersen, H. G. Error Estimates on Averages of Correlated Data. *J. Chem. Phys.* **1989**, *91*, 461.
(45) Haider, N. Functionality Pattern Matching as an Efficient Complementary Structure/Reaction Search Tool: an Open-source Approach. *Molecules* **2010**, *15*, 5079−5092.
(46) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926−935.
(47) Jorgensen, W. L.; Madura, J. D.; Swenson, C. J. Optimized Intermolecular Potential Functions for Liquid Hydrocarbons. *J. Am. Chem. Soc.* **1984**, *106*, 6638−6646.
(48) Bosque, R.; Sales, J. Polarizabilities of Solvents from the Chemical Composition. *J. Chem. Inf. Model.* **2002**, *42*, 1154−1163.
(49) Wang, L.-P.; Martínez, T. J.; Pande, V. S. Building Force Fields-an Automatic, Systematic and Reproducible Approach. *J. Phys. Chem. Lett.* **2014**, *5*, 1885−1891.
(50) Fennell, C. J.; Li, L.; Dill, K. A. Simple Liquid Models with Corrected Dielectric Constants. *J. Phys. Chem. B* **2012**, *116*, 6936−6944.
(51) Leontyev, I. V.; Stuchebrukhov, A. A. Polarizable Molecular Interactions in Condensed Phase and their Equivalent Nonpolarizable Models. *J. Chem. Phys.* **2014**, *141*, 014103.
(52) D'Aprano, A.; Donato, I. D. Dielectric Polarization and Polarizability of 1-pentanol+ n-octane Mixtures from Static Dielectric Constant and Refractive Index Data at 0, 25 and 45 C. *J. Solution Chem.* **1990**, *19*, 883−892.
(53) Haynes, W. M. *CRC Handbook of Chemistry and Physics*; CRC Press: Boca Raton, FL, 2011.
(54) Newman, J.; Fazio, V. J.; Caradoc-Davies, T. T.; Branson, K.; Peat, T. S. Practical Aspects of the SAMPL Challenge: Providing an Extensive Experimental Data Set for the Modeling Community. *J. Biomol. Screening* **2009**, *14*, 1245−1250.
(55) Truchon, J.-F.; Nicholl's, A.; Grant, J. A.; Iftimie, R. I.; Roux, B.; Bayly, C. I. Using Electronic Polarization from the Internal Continuum (EPIC) for Intermolecular Interactions. *J. Comput. Chem.* **2010**, *31*, 811−824.
(56) Truchon, J.-F.; Nicholls, A.; Roux, B.; Iftimie, R. I.; Bayly, C. I. Integrated Continuum Dielectric Approaches to Treat Molecular Polarizability and the Condensed Phase: Refractive Index and Implicit Solvation. *J. Chem. Theory Comput.* **2009**, *5*, 1785−1802.
(57) Truchon, J.-F.; Nicholls, A.; Iftimie, R. I.; Roux, B.; Bayly, C. I. Accurate Molecular Polarizabilities Based on Continuum Electrostatics. *J. Chem. Theory Comput.* **2008**, *4*, 1480−1493.
(58) Ponder, J.; Wu, C.; Ren, P.; Pande, V.; Chodera, J.; Schnieders, M.; Haque, I.; Mobley, D.; Lambrecht, D.; DiStasio, R., Jr; et al. Current Status of the AMOEBA Polarizable Force Field. *J. Phys. Chem. B* **2010**, *114*, 2549−2564.
(59) Ren, P.; Ponder, J. W. Temperature and Pressure Dependence of the AMOEBA Water Model. *J. Phys. Chem. B* **2004**, *108*, 13427−13437.
(60) Lamoureux, G.; Roux, B. Modeling Induced Polarization with Classical Drude Oscillators: Theory and Molecular Dynamics Simulation Algorithm. *J. Chem. Phys.* **2003**, *119*, 3025−3039.
(61) Anisimov, V. M.; Lamoureux, G.; Vorobyov, I. V.; Huang, N.; Roux, B.; MacKerell, A. D. Determination of Electrostatic Parameters for a Polarizable Force Field Based on the Classical Drude Oscillator. *J. Chem. Theory Comput.* **2005**, *1*, 153−168.
(62) Wang, L.-P.; Head-Gordon, T. L.; Ponder, J. W.; Ren, P.; Chodera, J. D.; Eastman, P. K.; Martínez, T. J.; Pande, V. S. Systematic Improvement of a Classical Molecular Model of Water. *J. Phys. Chem. B* **2013**, *117*, 9956−9972.
(63) Chirico, R. D.; Frenkel, M.; Diky, V. V.; Marsh, K. N.; Wilhoit, R. C. ThermoML An XML-Based Approach for Storage and Exchange of Experimental and Critically Evaluated Thermophysical and Thermochemical Property Data. 2. Uncertainties. *J. Chem. Eng. Data* **2003**, *48*, 1344−1359.

DOI: 10.1021/acs.jpcb.5b06703
J. Phys. Chem. B 2015, 119, 12912−12920