Automated Topology Builder Version 3.0

Stroet, Martin; Caron, Bertrand; Visscher, Koen M.; Geerke, Daan P.; Malde, Alpeshkumar K.; Mark, Alan E.

published in
Journal of Chemical Theory and Computation
2018

DOI (link to publisher)
10.1021/acs.jctc.8b00768

document version
Publisher's PDF, also known as Version of record

JCTC Journal of Chemical Theory and Computation

Article

Cite This: J. Chem. Theory Comput. 2018, 14, 5834−5845

pubs.acs.org/JCTC

Automated Topology Builder Version 3.0: Prediction of Solvation Free Enthalpies in Water and Hexane

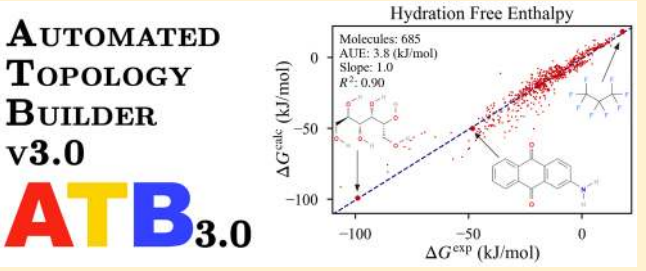
Martin Stroet,[†] Bertrand Caron,[†] Koen M. Visscher,[‡] Daan P. Geerke,[‡] Alpeshkumar K. Malde,[†] and Alan E. Mark*,[†]

[†]School of Chemistry & Molecular Biosciences, University of Queensland, St Lucia, Queensland 4072, Australia

[‡]AIMMS Division of Molecular Toxicology, Department of Chemistry and Pharmaceutical Sciences, Faculty of Science, Vrije Universiteit Amsterdam, De Boelelaan 1108, 1081 HZ Amsterdam, The Netherlands

Supporting Information

ABSTRACT: The ability of atomic interaction parameters generated using the Automated Topology Builder and Repository version 3.0 (ATB3.0) to predict experimental hydration free enthalpies (ΔG^{water}) and solvation free enthalpies in the apolar solvent hexane (ΔG^{hexane}) is presented. For a validation set of 685 molecules the average unsigned error (AUE) between ΔG^{water} values calculated using the ATB3.0 and experiment is 3.8 kJ·mol^{-1}. The slope of the line of best fit is 1.00, the intercept −1.0 kJ·mol^{-1}, and the R^2 0.90. For the more restricted set of 239 molecules used to validate OPLS3 (*J. Chem. Theory Comput.* **2016**, *12*, 281−296, DOI: 10.1021/acs.jctc.5b00864) the AUE using the ATB3.0 is just 2.7 kJ·mol^{-1} and the R^2 0.93. A roadmap for further improvement of the ATB parameters is presented together with a discussion of the challenges of validating force fields against the available experimental data.



INTRODUCTION

The ability to model molecular systems using classical molecular dynamics (MD) is critically dependent on the accuracy of the potential energy function (force field) used to represent atomic interactions. Extensively parametrized fixed-charge force fields such as AMBER,[1] CHARMM,[2] OPLS,[3] and GROMOS[4,5] are available for modeling common biomolecules such as amino acids, nucleic acids, lipids, and certain sugars. However, obtaining reliable parameters for heterogeneous classes of compounds such as substrates, inhibitors, co-factors, and drug-like molecules compatible with these force fields remains a major challenge.[6−8] While the functional forms used to describe interatomic interactions in the force fields listed above differ only superficially and all are fitted and validated against similar experimental and/or computational data, the philosophy underlying how they are optimized can vary significantly. Some, such as GROMOS and OPLS, have historically focused on fitting to thermodynamic properties such as the heats of vaporization, liquid densities, and the solvation properties of small molecules. While others, such as AMBER and CHARMM, have placed more emphasis on the reproduction of structural properties. The parameter development strategies also differ between force fields. For example, OPLS uses highly refined approaches to assign unique bonded and nonbonded parameters to atoms based on their local chemical environment[9] leading to many thousands of individual parameters, while GROMOS considers molecules as a collection of functional groups and allocates only a limited range of parameters. The argument for the former is that it allows specific details associated with individual molecules to be described while the aim of the later is to enhance transferability and mitigate the possibility of overfitting.

Irrespective of the philosophy used, the parametrization and testing of novel molecules can be laborious. To facilitate this process, a number of automated methods have been implemented as either standalone applications, e.g., Antechamber,[10,11] or as Web servers, e.g., PRODRG,[12] RED,[13] YASARA AutoSMILES,[14] SwissParam,[15] ParamChem,[16−18] GAAMP,[19,20] LigParGen,[21] and, the focus of this study, the Automated Topology Builder (ATB).[22] The ATB is a Web server that provides all-atom and united-atom interaction parameters as topology and building-block files in formats compatible with a range of molecular simulation and X-ray refinement packages: GROMOS,[23] GROMACS,[24,25] LAMMPS,[26] CNS,[27] Phenix,[28] and Refmac5,[29] as well as APBS.[30] A major difference between the ATB and other automated builders is that the parametrization procedures can be progressively refined based on an internal database of structural and quantum mechanical (QM) data. This internal database has been built from molecules submitted for parametrization by general users utilizing computational facilities provided by the University of Queensland and the Australian National Cyber Infrastructure and by researchers at the Lawrence Livermore National Laboratory using its computational facilities. The publicly accessible ATB reposi-

Received: July 25, 2018
Published: October 5, 2018

tory currently contains in excess of 250,000 molecules, including most of the molecules deposited in the PDB as ligands together with all molecules in the ChEMBL database containing 35 atoms or less. Access to the site and submission pipeline is made freely available for academic (noncommercial) use (atb.uq.edu.au).

While the ATB and each of the automated protocols listed above are widely used, the degree to which the interaction parameters they produce are reliable and accurate remains an open question.[22] The validation of molecular force fields in general is highly challenging.[31] This is because the amount of experimental data that can be used for validation is limited compared to the potential number of degrees of freedom in the model, and the fact that multibody effects along with the average contribution of the omitted electronic degrees of freedom are incorporated implicitly. In the case of small organics used in medicinal chemistry, much emphasis has been placed on the ability to reproduce the free enthalpy of solvation, i.e., the free enthalpy of transferring a molecule from the gas phase into a solvent such as water. This is because the free enthalpy of solvation can be directly measured experimentally for many molecules and is widely viewed as a robust method for validating solute−solvent interactions which are critical in processes such as the binding of a ligand to a protein or the partitioning of a molecule into a biological membrane, both of which involve the desolvation from water and the resolvation in a less polar environment. Indeed, studies performed by others suggest that improvements in the prediction of solvation data are correlated with improvements in the ability to predict ligand−protein binding affinities.[9,32]

Despite the fact solvation data is widely used to validate force fields, it is not without its challenges. Although free enthalpies of solvation can in principle be measured with high accuracy, the range of experimental data in the literature is limited and often dates from over a half-century ago. In addition, the process of collation, conversions between units, and the prioritization of specific sources introduces additional uncertainty in the values assigned to particular molecules. One way the uncertainty in the historical data can be mitigated is by validating against as broad a set of experimental data as possible and by considering the sensitivity of the results to subsets of the data. This has been widely recognized, and over the past decade there has been a trend toward progressively larger validation sets.[9,20,33] For example, in 2011 ATB1.0 was validated by computing the hydration free enthalpy of 190 compounds, 90 of which were biologically relevant small organic molecules and 100 were drug-like molecules taken from the CUP8/SAMPL0,[34] SAMPL1,[35] and SAMPL2[36] data sets representing a diverse range of functional groups. ATB1.0 used a semiautomated protocol and was based directly on the GROMOS53A6 united-atom force field.[4] As expected ATB1.0 performed well for molecules comprised of functional groups found within proteins but less well for molecules containing halogens and functional groups not commonly found in proteins and not part of the GROMOS force field including nitro, ether, and *N*-alkyl groups among others. In 2014 the ATB2.0 was validated against hydration free enthalpies for a set of 214 drug-like molecules, including 47 molecules that formed part of the SAMPL4 challenge. ATB2.0 used a fully automated parametrization and validation protocol. This meant that hydration free enthalpies could be obtained without manual intervention following the submission of a molecule to the ATB.[37] ATB2.0 was based on the GROMOS 54A7 united-atom force field,[4,5] and while it contained refinements in regard to the procedures used to assign bonded and Lennard-Jones (LJ) parameters, no attempt was made to refine the underlying force field, e.g., to refine the Lennard-Jones terms assigned to particular functional groups. The average unsigned error (AUE) compared to experiment using ATB 2.0 topologies for the set of 214 molecules considered was 6.7 kJ·mol^{-1} and the root-mean-square error (RMSE) was 9.5 kJ·mol^{-1}. However, for molecules containing functional groups that form part of the GROMOS 54A7 united-atom force field, the AUE was 3.4 kJ·mol^{-1} and the RMSE was 4.0 kJ·mol^{-1}. This suggested that small refinements to the parameters used to represent functional groups not specifically parametrized in the GROMOS 54A7 force field could lead to significant improvements in the ability to reproduce solvation free enthalpies. It also led to the development of improved refinement strategies as described in Stroet et al.[38]

In this work we present results from the first major attempt to refine the parameters used within the ATB. The ATB3.0 parameters have been validated by comparing the calculated and experimental solvation free enthalpies for a set of 685 compounds in water and 218 compounds in hexane. It is shown that the ATB3.0 parameters are predictive of solvation free enthalpies in both water and hexane. Comparisons using more restricted validation sets proposed by others suggest that the performance of the ATB3.0 parameters with respect to the prediction of hydration and apolar solvation is highly competitive with respect to other similar force fields for which published data are available. We also highlight cases where further improvements in the performance of the parameters should be possible.

METHODS

Parametrization Algorithm. The general protocol used by the ATB to assign interaction parameters for small molecules has been described in detail previously.[22,37,39] Briefly, data derived from quantum mechanical (QM) calculations are combined with a rule-based approach aimed at generating interaction parameters compatible with a given force field. Currently, the bonded parameters and 6−12 Lennard-Jones parameters are based around the GROMOS 54A7 force field.[4,5] However, this parameter set has been augmented with additional parameters specific to the ATB to enable a broad range of molecules to be described. For example, where a bonded interaction cannot be described appropriately using terms within the existing GROMOS 54A7 force field, additional parameters are generated dynamically based on an analysis of the QM Hessian. In addition, a number of Lennard-Jones terms have been optimized based on the experimental properties of a series of pure liquids properties and solvation free enthalpy data. Atomic charges are obtained by performing a least-squares fit of Coulomb point charges located at the center of each atom to electrostatic potential (ESP) surfaces as per the Merz and Kollman fitting scheme[40] using the program FieldFit.[41] ESP surfaces were calculated with a point density of 1.0 bohr^{-2} from the optimized geometry at the B3LYP/6-31G* level of theory in conjunction with the PCM[42] implicit solvation model for water. Fitting artifacts associated with assigning ESP charges based on the use of a single configuration were mitigated by averaging the resulting charges for chemically equivalent atoms.[43] Chemically equivalent atoms within a given molecule—arising as a result of molecular symmetry or rapid exchange—were

DOI: 10.1021/acs.jctc.8b00768
J. Chem. Theory Comput. 2018, 14, 5834−5845

identified from the automorphic nodes of the corresponding molecular graph calculated using the package nauty.[44]

Solvation Free Enthalpy Calculations. Solvation free enthalpies (ΔG^{solv}) were calculated using thermodynamic integration, in which the difference in free enthalpy between two states of a system A and B is expressed as

$$\Delta G_{AB} = \int_{\lambda_A}^{\lambda_B} \left\langle \frac{\partial V(\mathbf{r})}{\partial \lambda} \right\rangle_{\lambda} \mathrm{d}\lambda \quad (1)$$

where $V(\mathbf{r})$ is the potential energy of the system as a function of the coordinate vector $\mathbf{r}$ and λ is a parameter that couples the two states A and B.[45] In this case the coupling parameter λ was used to scale the inter- and intramolecular nonbonded interactions of the solute from full interaction to no interaction in the presence of solvent and in vacuum, the difference between these values yielding ΔG^{solv}. To avoid sampling singularities in the potential energy function and in the derivative with respect to λ (as well as to reduce numerical instabilities during the simulations), the nonbonded interactions were scaled using the λ-dependent soft-core interaction function of Beutler et al.[46] with $\alpha_{LJ} = 0.5$ and $\alpha_{electrostatic} = 0.5$ nm^2. Note that when using the λ-dependent soft-core interaction function as implemented in GROMOS (eqs 7 and 8 in Beutler et al.[46]), there is no requirement or advantage in performing the removal of the charge and Lennard-Jones interactions in separate steps as is sometimes required by other codes.

All thermodynamic integration calculations were performed in a fully automated manner. The statistical and systematic error arising from the calculations of the ensemble averages and numerical integration were analyzed dynamically to optimize computational efficiency. This also allowed the uncertainty in the calculated value of ΔG^{solv} to be iteratively refined to a specified precision by the extension of the MD simulation at a given λ-point or the addition of extra λ-points according to the largest sources of uncertainty.

The standard error in the ensemble average was estimated using the two-sample Kolmogorov−Smirnov statistic to compare the distribution of fluctuations in progressively larger portions of a time series. Full details are provided as Supporting Information. The code used to perform this analysis is available via GitHub.[47] This approach enables the effective equilibration time (i.e., the portion of a time series which should be omitted from ensemble average calculations due to relaxation from initial conditions) to be determined automatically and also provides a heuristic for the robustness of the convergence which indicates the reliability of the estimate of the error in the ensemble with respect to a predefined target uncertainty.

The error in the numerical integration of eq 1 involves two terms: (1) the uncertainty in the ensemble average ($\langle \partial V/\partial \lambda \rangle_{\lambda}$) value, and (2) truncation errors associated with the approach used to interpolate between the series of discrete λ-points, in this case the trapezoidal method. The contribution to the integration error from the uncertainty in the ensemble average at each λ-point was determined analytically for the trapezoidal method by applying the general Gaussian error propagation formula,[48] while the truncation error was estimated using a Taylor series expansion. Full details are provided as Supporting Information and the code used in this analysis is available via GitHub.[49] Using the approach outlined above, the individual contributions to the overall uncertainty can be calculated explicitly. It is therefore possible to efficiently reduce the total error by systematically targeting the largest individual components of the error.

The calculations of ΔG^{solv} were initialized by performing simulations of the solute in the solvent at 11 equally spaced λ-points ($\langle \partial V/\partial \lambda \rangle_{\lambda,solvent}$). The calculations in vacuum were performed at the same λ-points as used for the solvent calculations. These were initialized using 20 configurations of the solute extracted from the trajectories of the solute in solvent at regular intervals. The convergence of the ensemble averages was monitored using the method outlined above. The final ΔG^{solv} values were obtained by using the trapezoidal method to numerically integrate the curve given by $\langle \partial V/\partial \lambda \rangle_{\lambda,vac} - \langle \partial V/\partial \lambda \rangle_{\lambda,solvent}$. If the uncertainty using the initial 11 λ-points was greater than the target error of 1.5 kJ·mol^{-1}, additional λ-points were added or the simulations extended until the target error had been met.

Molecular Dynamics Simulations. All solvation free enthalpy calculations were performed on all-atom representations as generated by the ATB3.0 using the GROMOS11 (version 1.2.4) simulation package.[23] The simulations in water were started from structures optimized with GAMESS-US[50] at the B3LYP/6-31G* level of theory in conjunction with the PCM[42] implicit solvation model for water. To generate the solvated systems, a given molecule was placed at the center of a cubic periodic box. The size of the box was chosen such that the minimum distance between the solute and the box wall was 1.45 and 2.0 nm for solvation in water and hexane, respectively. The solute was then solvated using an equilibrated box of SPC[51] water or all-atom hexane (ATB3.0). The ATB3.0 all-atom hexane model reproduces the heat of vaporization, excess free energy, and hydration free in SPC water to within 1 kJ·mol^{-1} of the experimental values. The density of liquid hexane at 298.15 K is within 2% of the experimental value. Full details are provided as Supporting Information (Table S1). The solvated systems were energy minimized using a steepest decent algorithm. Initial velocities were taken from a Maxwell−Boltzmann distribution at 298 K. Bond lengths were constrained using SHAKE[52] with a geometric tolerance of 10^{-4}. The equations of motion were integrated using a time step of 2 fs. All simulations were performed at constant temperature (298 K) and pressure (1 atm) using a Berendsen thermostat and barostat.[53] The coupling times were 0.1 and 0.5 ps, respectively. The isothermal compressibility was 4.575×10^{-4} (kJ/mol/nm^3)$^{-1}$. Note, as the interactions between the solute and the rest of the system are scaled to zero, the solute can become thermally decoupled from the system. To prevent this, and to maintain equipartition, the solute was stochastically coupled to a reference temperature of 298 K using an atomic friction coefficient of 1 ps^{-1}. Nonbonded interactions were calculated using a single-range cutoff at 1.4 nm with the pair list updated every 5 steps. A reaction field correction[54] was applied to correct for the truncation of electrostatic interactions beyond the long-range cutoff using a relative dielectric permittivity of 61.[55]

The vacuum systems were generated by deleting the solvent molecules within the simulation box. Pressure coupling was not applied and the temperature was maintained using stochastic dynamics with a reference temperature of 298 K and an atomic friction coefficient of 91 ps^{-1}.[56]

Experimental Data. The experimental solvation free enthalpy data used in this study were obtained by aggregating four published data sets. The data sets used were as follows:

DOI: 10.1021/acs.jctc.8b00768
J. Chem. Theory Comput. 2018, 14, 5834−5845

the Minnesota Solvation Database—version 2012, which contains 3037 solvation energies and transfer free energies for 790 unique molecules in 92 solvents;[57] FreeSolv, a data set of 642 hydration free energies of small, neutral molecules;[58,59] the Schrodinger OPLS solvation free enthalpy validation data set published by Shivakumar et al.,[60] which contains hydration free enthalpies of 239 small molecules; and the 292 hydration free enthalpies collated in Gerber.[61] Note that all four data sets are secondary sources, and while many of the molecules within these databases are identical, the reported experimental values differed in some cases. The resulting set contained 773 neutral molecules with experimental ΔG^{water} values and 59 with experimental ΔG^{hexane} values. Since the number of molecules with experimental ΔG^{hexane} values was limited, this data set was augmented by exploiting the fact there is a very strong correlation between the solvation free enthalpy of solutes in a range of linear and cyclic alkanes. This enabled us to extend the available apolar solvation data to 218 molecules. See Stroet et al.[38] for details and for the validation of this approach. All experimental reference data that have been collated to date for molecules used in this work are provided as Supporting Information and is also available via GitHub.[62]

RESULTS AND DISCUSSION

Comparison Against All Available Data. The ATB3.0 all-atom interaction parameters were used to calculate the solvation free enthalpies in water and hexane for all molecules for which corresponding experimental data had been identified. This included a total of 773 molecules for which experimental data in water (ΔG^{water}) were available and 218 molecules for which experimental data were available in either hexane (ΔG^{hexane}) or a surrogate apolar solvent (pentane, cyclohexane, heptane, octane, nonane, decane, undecane, dodecane, pentadecane, and hexadecane).[38] The calculations were performed using the automated thermodynamic integration protocol described above. A complete list of all molecules, their ATB repository identifiers, and the calculated values of (ΔG^{water}) and (ΔG^{hexane}) together with the collated experimental data is provided as Supporting Information. A tar file containing the optimized geometries used as starting configurations in the calculations as well as full topology files for all molecules examined are also provided as Supporting Information, so that these results can be readily reproduced by others. The starting configurations and topology files are also available for download from the ATB website[22] (https://atb.uq.edu.au/). In addition, the website contains plots of $\langle \partial V/\partial \lambda \rangle$ versus λ allowing the precise λ-values used and the nature of the convergence of the thermodynamic integration calculations to be examined. A comparison between the calculated and experimental values for ΔG^{water} yielded an average unsigned error of 4.9 kJ·mol^{-1} and a root-mean-square error of 8.1 kJ·mol^{-1}. For ΔG^{hexane} the AUE was 3.9 kJ·mol^{-1} and the RMSE was 5.7 kJ·mol^{-1}. The values in the case of ΔG^{water} were large compared to the expected uncertainty, so the reliability of the experimental data was examined further.

ATB3.0 Water Validation Set. Of the 773 molecules for which experimental ΔG^{water} values are available, the reported experimental uncertainty in the case of 88 molecules is greater than 4.18 kJ·mol^{-1}. This is more than twice the statistical uncertainty in the calculated values (1.5 kJ·mol^{-1}). Removing these molecules from the test set resulted in an AUE between calculated and experimental ΔG^{water} of 3.8 kJ·mol^{-1} and an RMSE of 5.5 kJ·mol^{-1}. The resulting set of 685 molecules will be referred to as the ATB3.0 water validation set. A plot of the calculated versus experimental ΔG^{water} data for the full set of 773 molecules showing the experimental uncertainty and the distribution of the 88 molecules with an experimental uncertainty of greater than 4.18 kJ·mol^{-1} is provided as Supporting Information (Figure S1).

The fact that the removal of approximately 10% of the ΔG^{water} data (that with high experimental uncertainty) resulted in a 30% drop in the RMSE (8.1 to 5.5 kJ·mol^{-1}) highlights that uncertainty in the available experimental data remains a major challenge in the validation of molecular force fields. Indeed, a number of inconsistencies in the reported experimental data were also identified as part of this work. For example, the experimental ΔG^{water} for 1,2-dimethoxyethane reported by Rizzo et al.[63] (−20.3 kJ·mol^{-1}) differs significantly from that reported by Shivakumar et al.[64] (−16.1 kJ·mol^{-1}). The value for ΔG^{water} calculated for 1,2-dimethoxyethane using the ATB3.0 parameters was −11.0 ± 1.5 kJ·mol^{-1}; however, it is not clear whether either of the proposed experimental values are reliable. In the case of hexafluoropropylene an even greater discrepancy was found. The value for ΔG^{water} reported by Wilhelm et al.[65] (12.3 kJ·mol^{-1}) differs by 28 kJ·mol^{-1} from that reported by Boulanger et al.[20] (−15.7 kJ·mol^{-1}). The value calculated using ATB3.0 was 13.4 ± 1.5 kJ·mol^{-1}. This, combined with the fact that Wilhelm et al.[65] report primary data, strongly suggests that the value reported by Boulanger et al.[20] is unreliable. There are also three significantly different values for ΔG^{water} of methoxybenzene (anisole) reported: −4.34 kJ·mol^{-1} (Cabani et al.[66]), −10.25 kJ·mol^{-1} (Rizzo et al.[63]), and −15.61 kJ·mol^{-1} (Shivakumar et al.[64]). The value calculated by ATB3.0 is −13.4 kJ·mol^{-1}. It is unclear which value, if any, is reliable.

ATB3.0 Apolar Validation Set. In all cases the reported experimental uncertainty for the ΔG of solvation in the apolar solvents considered was less than or equal to 4.18 kJ·mol^{-1}. As a consequence, no molecules were eliminated from this part of the study. Nevertheless, a range of inconsistencies in the published experimental data were evident. One illustrative case is the unexpected difference between the solvation free enthalpy of *p*-hydroxybenzaldehyde in hexane and cyclohexane reported in the Minnesota Solvation Database—version 2012.[57] As noted above, the solvation free enthalpies of many compounds are highly correlated in a range of alkane solvents.[38] For example, the variation in the experimental solvation free enthalpy of phenol in hexane (−23.0 kJ·mol^{-1}) and seven alternative alkane solvents (pentane, cyclohexane, heptane, octane, nonane, decane, and hexadecane) is only 1.5 kJ·mol^{-1}. Similarly, the variation between benzaldehyde in hexane (−23.1 kJ·mol^{-1}) and three alternative alkane solvents (cyclohexane, heptane, and hexadecane) is only 0.75 kJ·mol^{-1}. However, the reported values for *p*-hydroxybenzaldehyde, which contains a combination of the functional groups in phenol and benzaldehyde, in hexane and cyclohexane are −38.4 kJ·mol^{-1} and −30.1 kJ·mol^{-1}, respectively, a difference of 8.3 kJ·mol^{-1}. The value of ΔG^{hexane} for *p*-hydroxybenzaldehyde calculated using ATB3.0 is −26.0 ± 1.5 kJ·mol^{-1}. Because the ATB3.0 parameters predict the experimental solvation free enthalpies of a range of closely related molecules including phenol, benzaldehyde, and *m*-hydroxybenzaldehyde to within 3 kJ·mol^{-1}, we suspect the value for *p*-hydroxybenzaldehyde in cyclohexane is more representative of the true ΔG of solvation of an isolated molecule of *p*-hydroxybenzaldehyde in an alkane solvent than the value listed for hexane.

DOI: 10.1021/acs.jctc.8b00768
J. Chem. Theory Comput. 2018, 14, 5834−5845

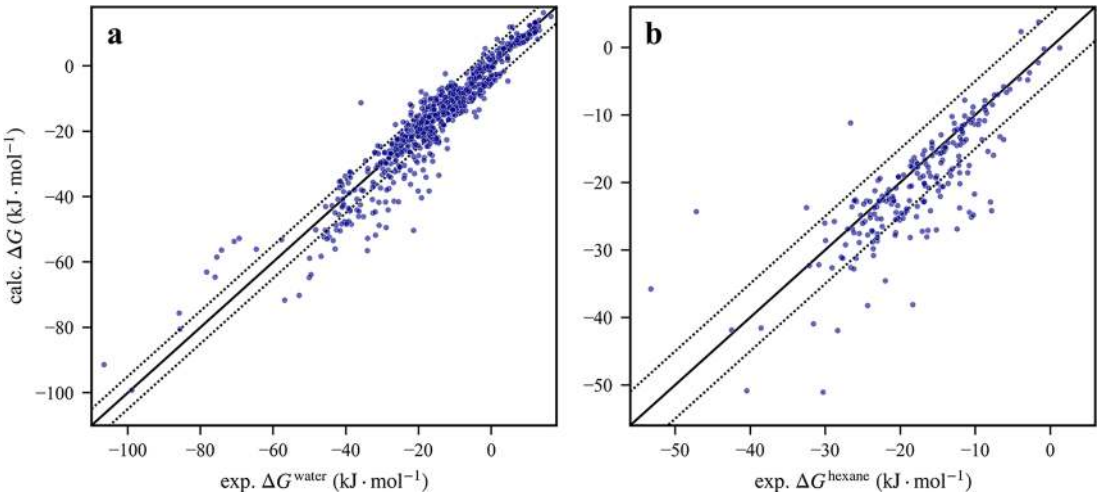


Figure 1. Calculated (calc.) versus experimental (exp.) solvation free enthalpies (ΔG^{water} and ΔG^{hexane}) for molecules parametrized using ATB3.0: (a) 685 molecules in water; (b) 218 molecules in hexane. The solid line corresponds to a one-to-one agreement between calculated and experimental values. The dotted lines correspond to ±5.0 kJ·mol^{-1} from one-to-one agreement.

Experimental Limitations. The examples of discrepancies in the experimental data given above have been included to illustrate the magnitude of the uncertainty present. They are not a complete list of all the problems identified. As noted in the Introduction, much of the data that is being used for parametrization and validation dates from over a half-century ago. This does not mean the data are flawed. However, the process of collation, conversions between units and the prioritization of specific sources over others means that the accuracy of any particular value is uncertain. For example, the data collated by Cabani et al.[66] forms part of many modern collections of experimental data. Cabani et al.[66] used the data to develop an empirical model for estimating thermodynamic properties of molecules based on group contributions. In their article multiple potential sources for the hydration free enthalpies of particular compounds are cited. However, only those values used to parametrize their model were recorded. No errors or uncertainties are indicated, and most importantly, no indication was provided for why a particular value was chosen over the alternatives or even why a particular source was preferred in a given case. This is important as the preference given to particular sources by Cabani et al.[66] is inconsistent. For example, in the case of 1-propene the hydration free enthalpy from Hine and Mookerjee[67] was selected and that of Wilhelm et al.[65] rejected while for 2-methylpropene the value of Wilhelm et al.[65] was selected and that of Hine and Mookerjee[67] rejected. A similar selection of data is likely to be a feature of all collations including Gerber[61] and Abraham et al.,[68] which have been incorporated into modern collections such as the Minnesota Solvation Database and FreeSolv. It is also important to note that the estimates of the errors now associated with these data have often been added subsequently.

Even in more recent studies it is unclear how the specific molecules included in a given fitting or validation test set were selected or why one particular experimental value is used as opposed to another. While we have identified inconsistencies in a number of cases (e.g., duplicate entries, ambiguous molecule identifiers, and significant differences between the values used and the value in the source that was cited), no attempt was made to systematically curate all the data and it is highly likely that other inconsistencies remain. This highlights the need to parametrize and validate interaction parameters with respect to large groups of related molecules in order to reduce the dependence on individual experimental values.

Finally, given the discrepancies in the reported experimental ΔG^{water} values and the fact that subsequent attempts to quantify the experimental errors are also inconsistent, the most appropriate method to choose between alternative values is unclear. Due to the ambiguity in the experimental data, we have chosen to compare to the reported experimental value which deviates least from that calculated with ATB3.0. However, all values that have been considered along with their source are included in the Supporting Information.

Overall Performance of ATB3.0. Figure 1 shows the relationship between the calculated and experimental values of ΔG^{water} and ΔG^{hexane} for 685 molecules and 218 molecules, respectively. As can be seen in Figure 1a, ATB3.0 is able to predict ΔG^{water} values for a very diverse set of molecules with values ranging from −107 to 18 kJ·mol^{-1}. The least-squares linear fit between the calculated and experimental ΔG^{water} values for the data shown in Figure 1a has a slope of 1.00, an intercept of −1.0 kJ·mol^{-1}, and an R^2 of 0.90. While the overall trend is very well reproduced, the ATB3.0 parameters, on average, underestimate ΔG^{water} by 1 kJ·mol^{-1} meaning that the average interaction of the molecules with SPC water is slightly too favorable. It is also evident that there is increasing scatter with increasing affinity for water. This may reflect challenges associated with the experimental measurements. Both observations also suggest further refinement of the parameters is possible.

In many common applications, such as the determination of binding affinity in computational drug design or predicting how compounds partition within biological membranes, the property of interest is dependent on the difference in free enthalpy between a molecule in water and the same molecule in a weakly polar or apolar environment. The ability of the ATB3.0 parameters to reproduce the aggregated experimental data in an apolar solvent (ΔG^{hexane}) is shown in Figure 1b. In this case the AUE is 3.9 kJ·mol^{-1}, RMSE is 5.7 kJ·mol^{-1}, and the line of best fit has a slope of 0.87, an intercept of −4.8 kJ·mol^{-1} and an R^2 of 0.65. The fact that the ΔG^{water} values are better reproduced than ΔG^{hexane} values in part reflects the focus of previous parametrization efforts and the fact that

DOI: 10.1021/acs.jctc.8b00768
J. Chem. Theory Comput. 2018, 14, 5834−5845

values calculated in hexane are compared to experimental data from a range of apolar solvents. The limited amount of experimental data in apolar solvents also means the ΔG^{hexane} results are significantly impacted by a small number of outliers.

Analysis of Specific Functional Groups. To determine whether the observed differences between the calculated and experimental values were associated with particular functional groups, the results were grouped and analyzed. Figure 2 shows

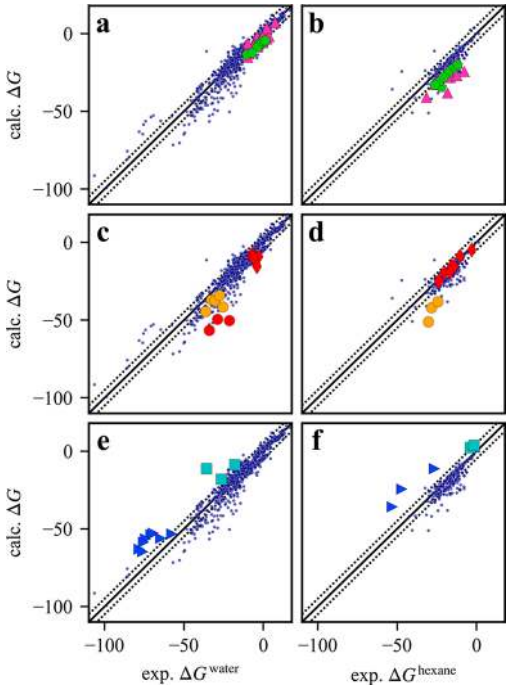


Figure 2. Examples of functional groups that lead to a systematic deviation between the calculated (calc.) and experimental (exp.) solvation free enthalpy (ΔG) in water (left) and hexane (right). Molecules containing bromo (purple triangles) and iodo (green pentagons) groups are highlighted in panels a and b; thiol and sulfide (red diamonds), phosphate (orange circles) and thiophosphate groups (red circles), in panels c and d; urea moieties (blue triangles) and molecules without carbon atoms (cyan squares), in panels e and f. In each case all other molecules are also indicated (blue dots). The solid line corresponds to a one-to-one agreement between the calculated and experimental values. The dotted lines correspond to ±5.0 kJ·mol^{-1} from a one-to-one agreement.

the results for ΔG^{water} (left) and ΔG^{hexane} (right) with particular subsets of related molecules which showed large deviations from the experimental values highlighted. In panels a and b, molecules containing bromine are indicated by purple triangles and those containing iodine by green pentagons. While the values for many of the bromine containing molecules in water are roughly in line with experiment, the values in hexane are systematically too favorable. The molecules containing iodine are systematically too favorable in both water and hexane. In panels c and d, thiol and sulfide containing groups are indicated by red diamonds, phosphates by orange circles, and thiophosphates by red circles. Clearly, the systematic error in ΔG^{water} for molecules containing thiophosphate groups (red circles) is to a first approximation additive, the deviation reflecting the sum of the error in thiols/sulfides and phosphates individually. This suggests that refining the thiol, sulfide, and phosphate parameters independently would also reduce the error in molecules containing a thiophosphate group. In panels e and f, molecules containing the urea moiety are indicated by blue triangles and molecules without carbon atoms (e.g., hydrogen peroxide and ammonia) by cyan squares. The solvation free enthalpies of molecules containing urea moieties are systematically overestimated in both water and hexane. In this case although the individual values do not match experiment, the compensation of errors would mean partition properties could still be reproduced. Molecules which do not contain carbon atoms such as hydrogen peroxide and ammonia are not well described by the level of theory used and it is not surprising that these molecules are not well represented by ATB3.0. It is also important to note in regard to Figure 2 that the 6−12 Lennard-Jones parameters used by ATB3.0 for sulfur and phosphorus as well as the parameters used to describe urea groups are taken directly from the GROMOS 54A7 force field and have not been refined. The GROMOS 54A7 bromine Lennard-Jones parameters were optimized to better reproduce ΔG^{water} for a set of 19 reference molecules.[69] The Lennard-Jones parameters for iodine were based on those used in the OPLS2005 force field but scaled to make them more compatible with other GROMOS 54A7 parameters. This suggests that refinement of these parameters would lead to a significant improvement in the overall performance of ATB3.0. Indeed, in the case of the ΔG^{hexane} results, if molecules containing bromo, iodo, and phosphate groups are excluded, the AUE for the remaining 189 molecules drops from 3.9 to 2.9 kJ·mol^{-1} and the RMSE from 5.7 to 4.2 kJ·mol^{-1}.

Comparing the Performance of the ATB3.0, OPLS3, GAFF/AM1-BCC, GAAMP, and LigParGen. As noted in the Introduction a number of automated parametrization schemes, both academic and commercial, have been developed.[9−11,19,20] Direct comparison between the parameters generated by these schemes is often challenging. Although these schemes are often validated against solvation free enthalpies, in particular hydration free enthalpies, the values are often calculated using different solvent models, different simulation packages, and different sets of molecules. In addition, the starting geometries of the molecules, the extent to which the calculations have converged, and even the parameter sets and topologies are not always made public.

To compare the performance of the ATB3.0 to alternative automated parametrization methods, we have not attempted to calculate the hydration free enthalpies using parameters generated by other schemes based on their associated force fields for all 685 molecules in the ATB3.0 validation set. Rather, we have analyzed the ATB3.0 ΔG^{water} results for the data set used by Shivakumar et al.[60] to validate OPLS3,[9] the data set used by Boulanger et al.[20] to validate GAAMP, that of Dodda et al.[70] used to validate the 1.14*CM1A and 1.14*CM1A-LBCC charges used in LigParGen, and that of Mobley et al.[58,59] to test GAFF/AM1-BCC. These are shown in Figure 3a−d, respectively. A summary of the performance statistics for ATB3.0 is listed in Table 1 alongside equivalent statistics for OPLS3, GAFF/AM1-BCC, GAAMP, 1.14*CM1A, and 1.14*CM1A-LBCC where available. The data set proposed by Shivakumar et al.[60] contains 239 molecules, that of Boulanger et al.[20] contains 419 unique molecules, and that of Dodda et al.[70] consists of 425 unique molecules; all are subsets of the ATB3.0 water validation set. The FreeSolv data set of Mobley et al.[58,59] contains 642 molecules; however, 59 of these have a reported uncertainty of 4.18 kJ·mol^{-1} or greater and were thus excluded from the

DOI: 10.1021/acs.jctc.8b00768
J. Chem. Theory Comput. 2018, 14, 5834−5845

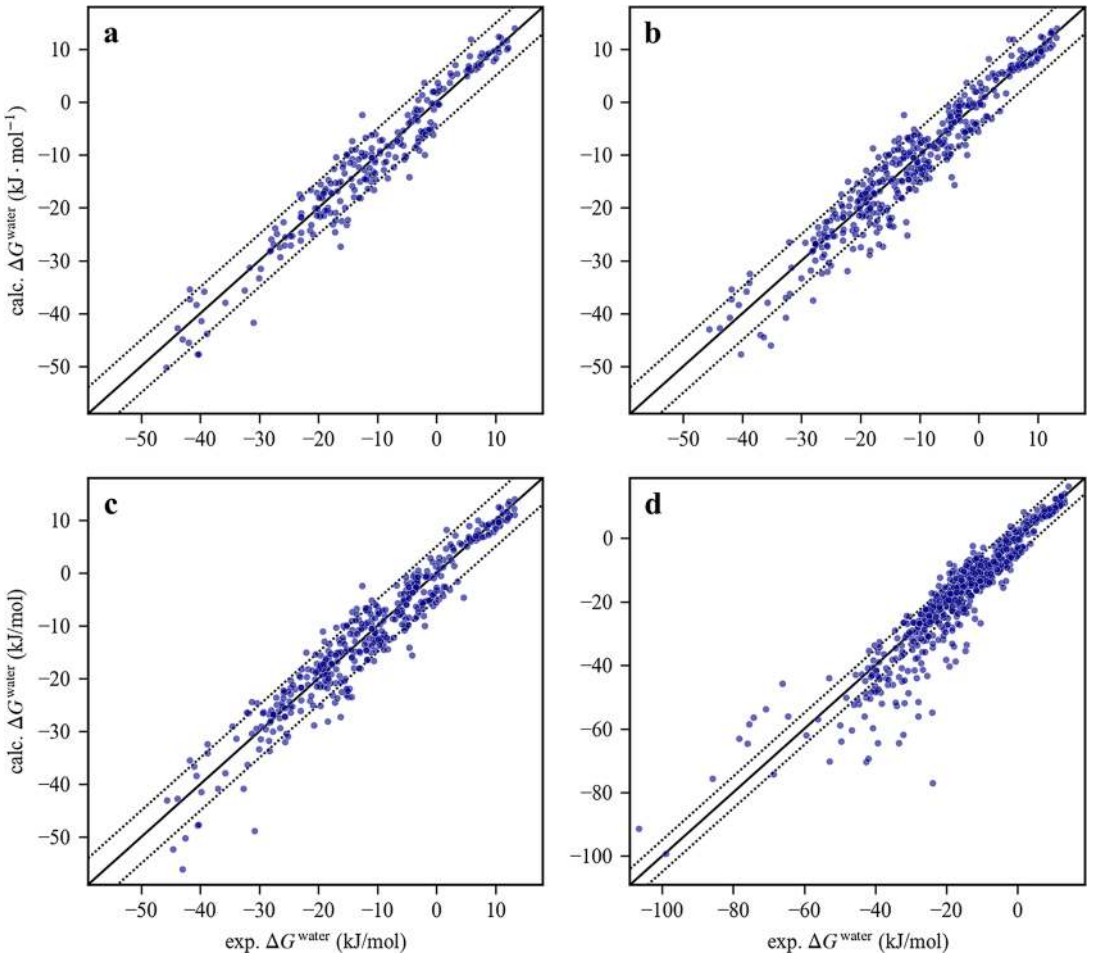


Figure 3. Comparison between the ATB3.0 calculated (calc.) and experimental (exp.) hydration free enthalpy (ΔG^{water}) values for the data sets proposed by Shivakumar et al.[60] (a), Boulanger et al.[20] (b), Dodda et al.[70] (c), and Mobley et al.[58,59] (d). The solid line corresponds to a one-to-one agreement between the calculated and experimental values. The dotted lines correspond to ±5.0 kJ·mol^{-1} from a one-to-one agreement.

Table 1. Comparison between the ATB3.0 and Other Automated Parametrization Schemes for the Prediction of Hydration Free Enthalpies[a]

data set	force field	N	AUE (kJ·mol^{-1})	RMSE (kJ·mol^{-1})	AE (kJ·mol^{-1})	slope	intercept (kJ·mol^{-1})	R^2
Shivakumar et al.[60]	ATB3.0	239	2.7	3.4	−0.4	1.01	−0.3	0.93
	OPLS3[9]	239	3.0	3.6	-	-	-	-
	GAFF/AM1-BCC[64]	239	4.9	5.8	-	0.97	3.3	0.87
Boulanger et al.[20]	ATB3.0	419[b]	2.9	3.7	−0.3	1.00	−0.3	0.91
	GAAMP[20]	426	6.2	7.3	5.9	0.90	4.9	0.88
	GAAMP scaled[20]	426	3.3	4.5	0.5	0.88	−0.8	0.86
	GAFF/AM1-BCC[20c]	426	4.3	5.3	3.1(−6.3)	0.98(0.90)	2.8(−3.9)	0.88
Dodda et al.[70]	ATB3.0	425[b]	2.9	3.7	−0.3	0.98	−0.5	0.92
	1.14*CM1A[70d]	426	5.3	6.6	−1.2	0.76	−3.6	0.81
	1.14*CM1A-LBCC[70d]	426	2.5	3.3	0.0	0.90	−1.2	0.94
Mobley et al.[58,59]	ATB3.0	642	4.2	6.6	−1.6	1.02	−1.2	0.87
	GAFF/AM1-BCC[58,59]	642	4.7	6.4	1.3	1.01	1.6	0.87
ATB validation set	ATB3.0	685	3.8	5.5	−1.0	1.00	−1.0	0.90

[a]Number of molecules (N), average unsigned error (AUE), root-mean-square error (RMSE), average error (AE), least-squares linear regression (slope, intercept), and coefficient of determination (R^2). [b]Duplicate and ambiguous entries removed; details in Supporting Information. [c]AE, slope, and intercept values recalculated using supporting data provided by Boulanger et al.;[20] reported values shown in brackets. [d]Summary statistics calculated with experimental reference values on the y-axis.

ATB3.0 water validation set. As is clear from Figure 3, the range of ΔG^{water} values and the diversity of chemical groups in the Shivakumar et al.,[60] Boulanger et al.,[20] and Dodda et al.[70] data sets is significantly less than that of FreeSolv[58] and the ATB3.0 water validation data set. It is unclear why only specific fractions of the available experimental data were used in these validation studies. Note that while both Boulanger et al.[20] and Dodda et al.[70] report data sets of 426 molecules, they do not

DOI: 10.1021/acs.jctc.8b00768
J. Chem. Theory Comput. 2018, 14, 5834−5845

correspond to exactly the same set of molecules and even where the same molecule is used in some cases the quoted experimental values differ. Boulanger et al.[20] do not explicitly cite the source of their data. Dodda et al.[70] cites FreeSolv.[33] However, the data set of Dodda et al.[70] contains 5 molecules not found in FreeSolv[33] or the more recent version of FreeSolv.[58,59] Of those 420 molecules found in FreeSolv,[58,59] 15 values differ by more than 0.5 kJ·mol^{-1} with the largest being 2-methylpropyl 2-methylpropanoate which differs by 9.2 kJ·mol^{-1}.

As can be seen from Table 1, the performance of ATB3.0 in terms of AUE and RMSE is equivalent to that of OPLS3[9] using the Shivakumar et al.[60] data set (AUE = 2.7 for ATB3.0 versus 3.0 kJ·mol^{-1} for OPLS3; RMSE = 3.4 versus 3.6 kJ·mol^{-1}). The ATB3.0 parameters are highly predictive with the line of best fit having a slope of 1.01, an intercept of −0.3 kJ·mol^{-1}, and an R^2 of 0.93. The results for GAFF/AM1-BCC[64] on this particular subset are significantly worse for all metrics except for the slope which is 0.97.

For the larger Boulanger et al.[20] data set, there is a slight loss of performance using ATB3.0 compared to the Shivakumar et al.[60] data set. AUE increases from 2.7 to 2.9 kJ·mol^{-1}, and RMSE increases from 3.3 to 3.7 kJ·mol^{-1}. The line of best fit has a slope of 1.00, an intercept of −0.3 kJ·mol^{-1}, and an R^2 of 0.91. A loss in performance with a larger test set could reflect overfitted force field parameters, a higher degree of uncertainty in the additional experimental data, or the inclusion of a larger proportion of molecules containing functional groups for which the ATB3.0 performs less well. In this case the loss in performance can be attributed to just 5 of the additional 180 molecules: hydrazine, methylimidazole, 1-methypyrrole, triethylamine, and butanethiol, which suggests that the loss in performance is not due to overfitting. The performance of GAAMP,[20] GAAMP scaled,[20] and GAFF/AM1-BCC[20] on this particular subset are all significantly worse than that of ATB3.0. Note that the metrics reported by Boulanger et al.[20] for the GAFF/AM1-BCC results (shown in brackets in Table 1) could not be reproduced when calculated using the supplementary data that the authors provided and that the recalculated results are significantly better than the metrics reported in the main body of the publication itself.

The Dodda et al.[70] data set used to validate the 1.14*CM1A and 1.14*CM1A-LBCC charge models used in LigParGen has a high degree of overlap with that of Boulanger et al.[20] but differs in a number of respects. Of the 425 and 419 unique molecules in the Dodda et al.[70] and Boulanger et al.[20] data sets, respectively, only 356 are found in both sets. Of these the experimental hydration free enthalpies differ by more than 0.2 kJ·mol^{-1} in 10 cases. For the Dodda et al.[70] data set the AUE of ATB3.0 is 2.9 kJ·mol^{-1} and the RMSE is 3.7 kJ·mol^{-1}. The line of best fit has a slope of 0.98, an intercept of −0.5 kJ·mol^{-1}, and an R^2 of 0.92. The results for ATB3.0 are significantly better than those obtained using the 1.14*CM1A charges on all criteria. The AUE and RMSE using 1.14*CM1A-LBCC, which contains 19 specially fitted bond correction terms, is lower than that obtained using the ATB3.0 by 0.4 kJ·mol^{-1} in both cases. However, the slope and intercept of the line of best fit reported by Dodda et al.[70] for the 1.14*CM1A-LBCC charge model are 0.90 and −1.2 kJ·mol^{-1}, respectively. Note, the molecules chosen by Dodda et al.[70] only range in ΔG^{water} values from −46 to 14 kJ·mol^{-1}. Given the error in the slope, the predictive ability over a more diverse range of compounds is uncertain.

The results obtained using the FreeSolv data set of Mobley et al.[58,59] (Table 1) show that the slope of best fit, the intercept, and the R^2 obtained using GAFF/AM1-BCC and ATB3.0 are almost identical. ATB3.0 has a lower AUE but a higher RMSE; the higher RMSE can be attributed to just 2 molecules. This is not entirely surprising—despite the fact that ATB3.0 performs significantly better than GAFF/AM1-BCC on the other data sets—given that FreeSolv contains a number of molecules for which the uncertainty in the experimental data is reported to be greater than 4.18 kJ·mol^{-1}, indicating that these values may be unreliable. Note that these molecules with high reported uncertainty were omitted from the ATB3.0 validation set and were not present in any of the alternative data sets that have been considered. The ATB3.0 results for the omitted molecules are highlighted in Figure S1.

No comparison between the performance of the ATB3.0 and either OPL3, GAFF/AM1-BCC, or GAAMP for apolar solvation is presented since apolar solvation data are not currently available for these alternative parameter sets.

Key Parameters and Parametrization Strategy. A key element of ATB3.0 is the choice of a simple underlying model and parametrization strategy. Much has been written by others on the choice of charge fitting model, the fitting of dihedral parameters, the choice of water model, and the treatment of long-range electrostatic interactions in the calculation of hydration free enthalpies.[20,70−72] Indeed, all of these elements are potentially important. However, given that the optimization of atomistic force fields with respect to the available experimental data is an underdetermined problem, there is a danger that the attempt to optimize each specific element of the calculation will lead to a proliferation of parameters and potentially an overfitted model. In this initial round of optimization we have deliberately attempted to restrict the number of parameters used by the ATB to avoid overfitting. Rather than attempting to optimize parameters for specific cases, we have developed generalized rules that can be applied across all molecules. This ensures that the parameters proposed are transferable over as broad a region of chemical space as possible. The results presented in this work involve 118 bond types, 73 angle types, and only 20 dihedral types. For comparison, OPLS3, which achieves similar fidelity in terms of the prediction of hydration free energies, includes 1187 bond stretching types, 15236 bond bending types, and tens of thousands of individually parametrized dihedrals. While these numbers are for the complete force field and not just the specific molecules considered in this work, OPLS3 does contain a much greater number of parameters than ATB3.0 or any other comparable force field, a fact repeatedly highlighted by its developers.[9] By default, the bonded parameters in this work were taken from the GROMOS 54A7 biomolecular force field which are primarily derived based on the element type and hybridization state of the atoms involved. Where this was not possible, additional bond and angle parameters are generated dynamically based on an analysis of the QM Hessian. In particular, the number of dihedral terms used in this work is limited and very small compared to that used in other comparable force fields. However, it is important to note in this regard that rather than simply introducing a fixed scaling factor, the GROMOS force field includes special Lennard-Jones terms to describe 1−4 interactions and that these terms act in concert with the dihedral terms. While this approach does have limitations, for the 250,000 molecules currently in the ATB repository the root-mean-square positional deviation

DOI: 10.1021/acs.jctc.8b00768
J. Chem. Theory Comput. 2018, 14, 5834−5845

(RMSD) between the QM optimized geometry and the geometry of the molecule after minimization in vacuum is less than 0.03 nm in over 95% of cases. Nevertheless, the results of others suggest that it should be possible to further improve the performance of the ATB by expanding the number of dihedral terms.[9,73]

The GROMOS 54A7 united-atom force field contains 54 sets of C6 and C12 Lennard-Jones parameters (atom types). However, only 23 atom types were used to represent the 773 molecules considered in this work. The Lennard-Jones parameters included 2 atom types for hydrogen atoms, 4 atom types for carbon, 5 atom types for oxygen, 5 for nitrogen atoms, 2 for sulfur, 1 for silicon and phosphorus combined, and 1 each for chlorine, bromine, and iodine. For completeness it should be noted that the GROMOS force field allows for alternative C12 terms to be specified for particular pairs of interactions. Alternative C12 parameters were used for 3 of the 5 oxygen types and all of the 5 nitrogen types, effectively yielding 31 unique sets of Lennard-Jones parameters. For comparison, the developers of GAAMP highlight the fact they used a total of 52 empirical Lennard-Jones parameters to describe their set of 419 unique molecules noting that this was a relatively small number.[20] OPLS3 contains a total of 124 atom types with different van der Waals (Lennard-Jones) parameter sets. Again, this is for the complete force field which covers a broader range of chemical space than considered here. The Lennard-Jones parameters used in the ATB3.0 force field were based on those in the GROMOS 54A7 parameter set which have been fitted to reproduce the densities and heats of vaporization for a range of simple liquids at room temperature and atmospheric pressure. However, a number of modifications and extensions were introduced. The Lennard-Jones parameters for chlorine have been systematically optimized as described previously.[38] The parameters for bromine were optimized against hydration free enthalpies. The charges assigned by the ATB3.0 to oxygen atoms are systematically more negative than those in the GROMOS 54A7 protein force field. To compensate for the increase in Coulombic interactions, the oxygen Lennard-Jones parameters used for carbonyl atoms in aldehyde, ketone, ester, and carboxylic acid groups were refined accordingly, as were the parameters for oxygen atoms in hydroxy groups. Conversely, the charges assigned by the ATB3.0 to aromatic hydrogen and carbon atoms are systematically lower than those in the GROMOS 54A7 force field. The Lennard-Jones parameters for aromatic carbon atoms were adjusted to compensate. As noted by Pechlaner et al.,[74] the original GROMOS 54A7 parameters overestimate the free enthalpy of hydration for amines.[75] In the ATB3.0 separate parameters are used to describe primary, secondary, and tertiary amines. Finally, the GROMOS 54A7 force field has a single set of parameters for a bare carbon used in an all-atom representation irrespective of its partial charge. While the partial charge on most aliphatic carbons is either slightly negative or close to zero, the partial charge on some carbon atoms associated with functional groups such as alcohols, aldehydes, ketones, and esters are significantly greater than zero. Such positively charged carbons were assigned an alternative set of Lennard-Jones parameters.

The charge assignment model used in this work is also relatively simple. The point charges located at the center of each atom were obtained using the Merz and Kollman fitting scheme.[40] The primary variation from previously published protocols was an increase in the density of points used to represent the ESP surfaces. To avoid overpolarization and ensure that the charges were in the same range as those commonly used in the GROMOS force field and thus compatible with the existing Lennard-Jones parameters, the ESP surfaces were calculated in conjunction with the PCM[42] implicit solvation model of water. Fitting artifacts resulting from the use of a single conformation were mitigated by identifying equivalent atoms that arise as a result of molecular symmetry or rapid exchange using the nauty package and averaging the charges.[44] In contrast to force fields such as 1.14*CM1A-LBCC and GAFF/AM1-BCC, the introduction of fitted charge scaling parameters and localized terms such as bond charge corrections (BCCs)[70] have been avoided.

A copy of the ATB3.0 augmented GROMOS 54A7 interaction parameter file used in this work is provided as Supporting Information. Note, this file contains both the original GROMOS parameters as well as the parameters used by molecules parametrized with ATB3.0 so that ATB molecules can be combined directly with the existing elements of the GROMOS biomolecular force field.

CONCLUSIONS

In this work atomic interaction parameters generated using version 3.0 of ATB were used to compare the calculated and experimental free enthalpy of solvation in water and hexane for 773 and 218 molecules, respectively. Based on the results of these calculations and an analysis of the reported uncertainties in the experimental data, we have proposed a set of 685 molecules for which the uncertainty in the experimental hydration data is $\leq$4.18 kJ·mol^{-1} as a validation set. For these 685 molecules the slope of the line of best fit between the values calculated with ATB3.0 and experiment is 1.00; the intercept, −1.0 kJ·mol^{-1}; and R^2, 0.90. The AUE = 3.8 kJ·mol^{-1}, and RMSE = 5.5 kJ·mol^{-1}. This shows that the ATB3.0 parameters are predictive for a wide range of compounds. Furthermore, comparisons using more restricted validation sets proposed by others suggest that the performance of the ATB3.0 parameters with respect to the prediction of hydration and apolar solvation is better than or comparable to that of similar force fields including OPLS3, GAAMP, LigParGen, and GAFF/AM1-BCC[20] based on published data.

The results described in this work represent the first stage in an ongoing program to refine systematically the parameters distributed as part of ATB. The parameters for molecules in the ATB repository are continuously updated and are made freely available for academic (noncommercial) use. The parameters are distributed in a number of formats and can be directly used in conjunction with a variety of simulation packages including the current versions of GROMOS, GROMACS and LAMMPS. A converter is also available to allow systems generated using GROMOS to be simulated using AMBER. This will be described in more detail elsewhere.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.8b00768.

Additional computational details, hydration free enthalpies, ATB all-atom alkane model results, duplicate and ambiguous data, and hydration and solvation free energies (PDF)

DOI: 10.1021/acs.jctc.8b00768
J. Chem. Theory Comput. 2018, 14, 5834−5845

Optimized geometries used as starting configurations and full topology files (ZIP)

AUTHOR INFORMATION

Corresponding Author

*Tel.: +61 7 336 54180. Fax: +61 7 336 53872. E-mail: a.e.mark@uq.edu.au.

ORCID

Martin Stroet: 0000-0002-9570-2376
Bertrand Caron: 0000-0003-2305-1452
Daan P. Geerke: 0000-0002-5262-6166
Alpeshkumar K. Malde: 0000-0002-8181-1619
Alan E. Mark: 0000-0001-5880-4798

Funding

This work was funded from the Australian Grants Commission (Grants DP150101097 and DP180101421) with the assistance of high-performance computing resources provided through the National Computational Merit Allocation Scheme supported by the Australian Government (Project m72).

Notes

The authors declare no competing financial interest.

The research data (and/or materials) supporting this publication can be accessed and are made freely available for academic (noncommercial) use at http://atb.uq.edu.au.

ACKNOWLEDGMENTS

We thank Dr. Brian Bennion, Dr. Felice Lightstone, and Ms. Svetlana Gelpi-Dominguez of the Lawrence Livermore National Laboratory (LLNL) for their assistance in expanding the ATB public database by performing QM calculations on molecules within ChEMBL using LLNL computing resources.

ABBREVIATIONS

AE,average error; ATB,automated topology builder; AUE,average unsigned error; ESP,electrostatic potential; MD,molecular dynamics; QM,quantum mechanical; RMSE,root-mean-square error; TI,thermodynamic integration

REFERENCES

(1) 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.

(2) MacKerell, A. D., Jr; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.

(3) 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.

(4) 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.

(5) 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.

(6) Verlinde, C. L. M. J.; Hol, W. G. J. Structure-based drug design: progress, results and challenges. *Structure* **1994**, *2*, 577−587.

(7) Tollenaere, J. P. The role of structure-based ligand design and molecular modelling in drug discovery. *Pharm. World Sci.* **1996**, *18*, 56−62.

(8) Ooms, F. Molecular modeling and computer aided drug design. Examples of their applications in medicinal chemistry. *Curr. Med. Chem.* **2000**, *7*, 141−158.

(9) Harder, E.; Damm, W.; Maple, J.; Wu, C.; Reboul, M.; Xiang, J. Y.; Wang, L.; Lupyan, D.; Dahlgren, M. K.; Knight, J. L.; Kaus, J. W.; Cerutti, D. S.; Krilov, G.; Jorgensen, W. L.; Abel, R.; Friesner, R. A. OPLS3: A force field providing broad coverage of drug-like small molecules and proteins. *J. Chem. Theory Comput.* **2016**, *12*, 281−296.

(10) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graphics Modell.* **2006**, *25*, 247−260.

(11) 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.

(12) Schuttelkopf, A. W.; van Aalten, D. M. F. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 1355−1363.

(13) Vanquelef, E.; Simon, S.; Marquant, G.; Garcia, E.; Klimerak, G.; Delepine, J. C.; Cieplak, P.; Dupradeau, F.-Y. RED Server: a web service for deriving RESP and ESP charges and building force field libraries for new molecules and molecular fragments. *Nucleic Acids Res.* **2011**, *39*, W511−W517.

(14) Krieger, E.; Koraimann, G.; Vriend, G. Increasing the precision of comparative models with YASARA NOVA—a self-parameterizing force field. *Proteins: Struct., Funct., Genet.* **2002**, *47*, 393−402.

(15) Zoete, V.; Cuendet, M. A.; Grosdidier, A.; Michielin, O. SwissParam: A fast force field generation tool for small organic molecules. *J. Comput. Chem.* **2011**, *32*, 2359−2368.

(16) 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.* **2010**, *31*, 671−690.

(17) Vanommeslaeghe, K.; MacKerell, A. D., Jr Automation of the CHARMM General Force Field (CGenFF) I: Bond perception and atom typing. *J. Chem. Inf. Model.* **2012**, *52*, 3144−3154.

(18) Vanommeslaeghe, K.; Raman, E. P.; MacKerell, A. D., Jr Automation of the CHARMM General Force Field (CGenFF) II: Assignment of Bonded parameters and partial atomic charges. *J. Chem. Inf. Model.* **2012**, *52*, 3155−3168.

(19) Huang, L.; Roux, B. Automated force field parameterization for nonpolarizable and polarizable atomic models based on ab Initio target data. *J. Chem. Theory Comput.* **2013**, *9*, 3543−3556.

(20) Boulanger, E.; Huang, L.; Rupakheti, C.; MacKerell, A. D., Jr.; Roux, B. Optimized Lennard-Jones parameters for druglike small molecules. *J. Chem. Theory Comput.* **2018**, *14*, 3121−3131.

(21) Dodda, L. S.; Cabeza de Vaca, I.; Tirado-Rives, J.; Jorgensen, W. L. LigParGen web server: an automatic OPLS-AA parameter generator for organic ligands. *Nucleic Acids Res.* **2017**, *45*, W331−W336.

(22) Malde, A. K.; Zuo, L.; Breeze, M.; Stroet, M.; Poger, D.; Nair, P. C.; Oostenbrink, C.; Mark, A. E. An automated force field topology builder (ATB) and repository: Version 1.0. *J. Chem. Theory Comput.* **2011**, *7*, 4026−4037.

(23) Schmid, N.; Christ, C. D.; Christen, M.; Eichenberger, A. P.; van Gunsteren, W. F. Architecture, implementation and parallelisation of the GROMOS software for biomolecular simulation. *Comput. Phys. Commun.* **2012**, *183*, 890−903.

(24) van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. GROMACS: Fast, flexible, and free. *J. Comput. Chem.* **2005**, *26*, 1701−1718.

DOI: 10.1021/acs.jctc.8b00768
J. Chem. Theory Comput. 2018, 14, 5834−5845

(25) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1−2*, 19−25.
(26) Plimpton, S. Fast parallel algorithms for short-range molecular dynamics. *J. Comput. Phys.* **1995**, *117*, 1−19.
(27) Brünger, A. T.; Adams, P. D.; Clore, G. M.; DeLano, W. L.; Gros, P.; Grosse-Kunstleve, R. W.; Jiang, J. S.; Kuszewski, J.; Nilges, M.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Warren, G. L. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1998**, *54*, 905−921.
(28) Adams, P. D.; Afonine, P. V.; Bunkoczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L.-W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 213−221.
(29) Murshudov, G. N.; Skubák, P.; Lebedev, A. A.; Pannu, N. S.; Steiner, R. A.; Nicholls, R. A.; Winn, M. D.; Long, F.; Vagin, A. A. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *67*, 355−367.
(30) Baker, N. A.; Sept, D.; Joseph, S.; Holst, M. J.; McCammon, J. A. Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 10037−10041.
(31) van Gunsteren, W. F.; Daura, X.; Hansen, N.; Mark, A. E.; Oostenbrink, C.; Riniker, S.; Smith, L. Validation of molecular simulation: An overview of issues. *Angew. Chem., Int. Ed.* **2018**, *57*, 884−902.
(32) Shivakumar, D.; Harder, E.; Damm, W.; Friesner, R. A.; Sherman, W. Improving the prediction of absolute solvation free energies using the next generation OPLS force field. *J. Chem. Theory Comput.* **2012**, *8*, 2553−2558.
(33) Mobley, D. L.; Guthrie, J. P. FreeSolv: a database of experimental and calculated hydration free energies, with input files. *J. Comput.-Aided Mol. Des.* **2014**, *28*, 711−720.
(34) Nicholls, A.; Mobley, D. L.; Guthrie, J. P.; Chodera, J. D.; Bayly, C. I.; Cooper, M. D.; Pande, V. S. Predicting small-molecule solvation free energies: an informal blind test for computational chemistry. *J. Med. Chem.* **2008**, *51*, 769−79.
(35) Guthrie, J. P. A blind challenge for computational solvation free energies: introduction and overview. *J. Phys. Chem. B* **2009**, *113*, 4501−4507.
(36) Geballe, M. T.; Skillman, A. G.; Nicholls, A.; Guthrie, J. P.; Taylor, P. J. The SAMPL2 blind prediction challenge: introduction and overview. *J. Comput.-Aided Mol. Des.* **2010**, *24*, 259−279.
(37) Koziara, K. B.; Stroet, M.; Malde, A. K.; Mark, A. E. Testing and validation of the automated topology builder (ATB) version 2.0: Prediction of hydration free enthalpies. *J. Comput.-Aided Mol. Des.* **2014**, *28*, 221−233.
(38) Stroet, M.; Koziara, K. B.; Malde, A. K.; Mark, A. E. Optimization of empirical force fields by parameter space mapping: A single-step perturbation approach. *J. Chem. Theory Comput.* **2017**, *13*, 6201−6212.
(39) Canzar, S.; El-Kebir, M.; Pool, R.; Elbassioni, K.; Malde, A. K.; Mark, A. E.; Geerke, D. P.; Stougie, L.; Klau, G. W. Charge group partitioning in biomolecular simulation. In *Research in computational molecular biology*, Chor, B., Ed.; Springer: Berlin, Heidelberg, 2012; Vol. *7262*, pp 29−43, DOI: 10.1007/978-3-642-29627-7_3.
(40) Singh, U. C.; Kollman, P. A. An approach to computing electrostatic charges for molecules. *J. Comput. Chem.* **1984**, *5*, 129−145.
(41) Visscher, K. M.; Geerke, D. P. FieldFit. github.com/GeerkeLab/fieldFit.
(42) Miertuš, S.; Scrocco, E.; Tomasi, J. Electrostatic interaction of a solute with a continuum. A direct utilizaion of ab initio molecular potentials for the prevision of solvent effects. *Chem. Phys.* **1981**, *55*, 117−129.
(43) Malde, A. K.; Stroet, M.; Caron, B.; Visscher, K. M.; Mark, A. E. Predicting the prevalence of alternative Warfarin tautomers in solution. *J. Chem. Theory Comput.* **2018**, *14*, 4405.
(44) McKay, B. D.; Piperno, A. Practical graph isomorphism, II. *J. Symb. Comput.* **2014**, *60*, 94−112.
(45) van Gunsteren, W. F.; Weiner, P. K.; Wilkinson, A. J. *Computer simulation of biomolecular systems: Theoretical and experimental applications*. Springer Science & Business Media: 1997; Vol. 3.
(46) 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.
(47) Stroet, M. *ks_convergence_analysis*, Version 1.1; 2018, DOI: 10.5281/zenodo.1313260, supplement to https://github.com/martinstroet/ks_convergence_analysis.
(48) Ku, H. H. Notes on the use of propagation of error formulas. *J. Res. Natl. Bur. Stand., Sect. C* **1966**, *70C*, 263−273.
(49) Stroet, M. *trapz_errors*, Version 1.1, 2018, DOI: 10.5281/zenodo.1313404, supplement to https://github.com/martinstroet/trapz_errors.
(50) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. General atomic and molecular electronic structure system. *J. Comput. Chem.* **1993**, *14*, 1347−1363.
(51) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. Interaction models for water in relation to protein hydration. In *Intermolecular forces: Proceedings of the fourteenth Jerusalem symposium on quantum chemistry and biochemistry held in Jerusalem, Israel, April 13−16, 1981*; Pullman, B., Ed.; Springer: Dordrecht, The Netherlands, 1981; pp 331−342.
(52) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **1977**, *23*, 327−341.
(53) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(54) Tironi, I. G.; Sperb, R.; Smith, P. E.; van Gunsteren, W. F. A generalized reaction field method for molecular dynamics simulations. *J. Chem. Phys.* **1995**, *102*, 5451−5459.
(55) Heinz, T. N.; van Gunsteren, W. F.; Hünenberger, P. H. Comparison of four methods to compute the dielectric permittivity of liquids from molecular dynamics simulations. *J. Chem. Phys.* **2001**, *115*, 1125−1136.
(56) Yun-yu, S.; Lu, W.; van Gunsteren, W. F. On the approximation of solvent effects on the conformation and dynamics of cyclosporin A by stochastic dynamics simulation techniques. *Mol. Simul.* **1988**, *1*, 369−383.
(57) Marenich, A. V.; Kelly, C. P.; Thompson, J. D.; Hawkins, G. D.; Chambers, C. C.; Giesen, D. J.; Winget, P.; Cramer, C. J.; Truhlar, D. G. *Minnesota Solvation Database − Version 2012*; University of Minnesota: Minneapolis, MN, USA, 2012.
(58) Mobley, D. L.; Shirts, M.; Lim, N.; Chodera, J.; Beauchamp, K.; Lee-Ping *MobleyLab/FreeSolv*, Version 0.52; 2018, DOI: 10.5281/zenodo.1161245 supplement to https://github.com/MobleyLab/FreeSolv/tree/v0.52.
(59) Duarte Ramos Matos, G.; Kyu, D. Y.; Loeffler, H. H.; Chodera, J. D.; Shirts, M. R.; Mobley, D. L. Approaches for calculating solvation free energies and enthalpies demonstrated with an update of the FreeSolv database. *J. Chem. Eng. Data* **2017**, *62*, 1559−1569.
(60) Shivakumar, D.; Deng, Y.; Roux, B. Computations of absolute solvation free energies of small molecules using explicit and implicit solvent model. *J. Chem. Theory Comput.* **2009**, *5*, 919−930.
(61) Gerber, P. R. Charge distribution from a simple molecular orbital type calculation and non-bonding interaction terms in the force field MAB. *J. Comput.-Aided Mol. Des.* **1998**, *12*, 37−51.
(62) Stroet, M.; Caron, B.; Visscher, K. M.; Geerke, D. P.; Malde, A. K.; Mark, A. E. *ATB-UQ/ATB3.0_validation_data*, Version 1.1; 2018,

DOI: 10.1021/acs.jctc.8b00768
J. Chem. Theory Comput. 2018, 14, 5834−5845

DOI: 10.5281/zenodo.1412161 supplement to https://github.com/ATB-UQ/ATB3.0_validation_data.
(63) Rizzo, R. C.; Aynechi, T.; Case, D. A.; Kuntz, I. D. Estimation of absolute free energies of hydration using continuum methods: Accuracy of partial charge models and optimization of nonpolar contributions. *J. Chem. Theory Comput.* **2006**, *2*, 128−139.
(64) Shivakumar, D.; Williams, J.; Wu, Y.; Damm, W.; Shelley, J.; Sherman, W. Prediction of absolute solvation free energies using molecular dynamics free energy perturbation and the OPLS force field. *J. Chem. Theory Comput.* **2010**, *6*, 1509−1519.
(65) Wilhelm, E.; Battino, R.; Wilcock, R. J. Low-pressure solubility of gases in liquid water. *Chem. Rev.* **1977**, *77*, 219−262.
(66) 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.
(67) Hine, J.; Mookerjee, P. K. Structural effects on rates and equilibriums. XIX. Intrinsic hydrophilic character of organic compounds. Correlations in terms of structural contributions. *J. Org. Chem.* **1975**, *40*, 292−298.
(68) Abraham, M. H.; Whiting, G. S.; Fuchs, R.; Chambers, E. J. Thermodynamics of solute transfer from water to hexadecane. *J. Chem. Soc., Perkin Trans. 2* **1990**, 291−300.
(69) Koziara, K. B. Development and validation of the force field parameters for drug-like molecules and their applications in structure-based drug design. Ph.D. Thesis; University of Queensland, St Lucia, Australia, 2016; DOI: 10.14264/uql.2016.1075.
(70) Dodda, L. S.; Vilseck, J. Z.; Tirado-Rives, J.; Jorgensen, W. L. 1.14*CM1A-LBCC: Localized bond-charge corrected CM1A charges for condensed-phase simulations. *J. Phys. Chem. B* **2017**, *121*, 3864−3870.
(71) 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.
(72) MacKerell, A. D., Jr.; Feig, M.; Brooks, C. L., III Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J. Comput. Chem.* **2004**, *25*, 1400−1415.
(73) DuBay, K. H.; Hall, M. L.; Hughes, T. F.; Wu, C.; Reichman, D. R.; Friesner, R. A. Accurate force field development for modeling conjugated polymers. *J. Chem. Theory Comput.* **2012**, *8*, 4556−4569.
(74) Pechlaner, M.; Reif, M. M.; Oostenbrink, C. Reparametrisation of united-atom amine solvation in the GROMOS force field. *Mol. Phys.* **2017**, *115*, 1144−1154.
(75) Oostenbrink, C.; Juchli, D.; van Gunsteren, W. F. Amine hydration: A united-atom force-field solution. *ChemPhysChem* **2005**, *6*, 1800−1804.

DOI: 10.1021/acs.jctc.8b00768
J. Chem. Theory Comput. 2018, 14, 5834−5845