UC Davis

UC Davis Previously Published Works

Title

Tuning Potential Functions to Host—Guest Binding Data

Permalink

https://escholarship.org/uc/item/2ts1v38g

Journal

Journal of Chemical Theory and Computation, 20(1)

ISSN

1549-9618

Authors

Setiadi, Jeffry

Boothroyd, Simon

Slochower, David R

et al.

Publication Date

2024-01-09

DOI

10.1021/acs.jctc.3c01050

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

J Chem Theory Comput. Author manuscript; available in PMC 2025 January 09.

Published in final edited form as:

J Chem Theory Comput. 2024 January 09; 20(1): 239–252. doi:10.1021/acs.jctc.3c01050.

Tuning Potential Functions to Host–Guest Binding Data

Jeffry Setiadi[†], **Simon Boothroyd**[‡,¶], **David R. Slochower**[§], **David L. Dotson**[‖,⊥], **Matthew W. Thompson**[⊥], **Jeffrey R. Wagner**[⊥], **Lee-Ping Wang**[#], **Michael K. Gilson**[†]

[†]Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, 9255 Pharmacy Lane, La Jolla, California 92093, United States

[‡]Boothroyd Scientific Consulting Ltd., London WC2H 9JQ, United Kingdom

[¶]Psivant Therapeutics, Boston, Massachusetts 02210, United States

[§]Vertex Pharmaceuticals, San Diego, California 92121, United States

[‖]Datryllic LLC, Phoenix, Arizona 85003, United States

[⊥]The Open Force Field Consortium, Open Molecular Software Foundation, Davis, California 95616, United States

[#]Chemistry Department, University of California Davis, Davis, California 95616, United States

Abstract

Software to more rapidly and accurately predict protein–ligand binding affinities is of high interest for early-stage drug discovery, and physics-based methods are among the most widely used technologies for this purpose. The accuracy of these methods depends critically on the accuracy of the potential functions they use. Potential functions are typically trained against a combination of quantum chemical and experimental data. However, although binding affinities are among the most important quantities to predict, experimental binding affinities have not to date been integrated into the experimental dataset used to train potential functions. In recent years, the use of host–guest complexes as simple and tractable models of binding thermodynamics has gained popularity due to their small size and simplicity, relative to protein–ligand systems. Host–guest complexes can also avoid ambiguities that arise in protein–ligand systems, such as uncertain protonation states. Thus, experimental host–guest binding data are an appealing additional data type to integrate into the experimental dataset used to optimize potential functions. Here, we report the extension of the Open Force Field Evaluator framework to enable the systematic calculation of host–guest binding free energies and their gradients with respect to force field parameters, coupled with the curation of 126 host–guest complexes with available experimental binding free energies. As an initial application of this novel infrastructure, we optimized generalized

mgilson@health.ucsd.edu .

Supporting Information Available

Figure S1, conformational restraints applied on host molecules; Figure S2–S4, the host–guest test set complexes; Figure S5, the evolution of parameters from ForceBalance optimization; Table S1–S2, the host–guest binding free energies for the training set and test set; Table S3–S4, statistics of host–guest binding free energies for the training set and test set; Figure S6 – plots of electrostatic potentials in CB7 along the axes of βCD; Figure S7–S11, protein–ligand data set; Table S5–S8, protein–ligand binding affinities; Figure S12, the plot of protein–ligand binding with ff99SB-ILDN force field; Table S9–S10, statistics of protein–ligand binding free energies; Table S11, hydration free energies; Table S12, statistics of hydration free energies; Figure S13, small molecular hydration free energy data set; Figure S14, chemical structures of outliers in hydration free energy calculations

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Born (GB) cavity radii for the OBC2 GB implicit solvent model against experimental data for 36 host–guest systems. This refitting led to a dramatic improvement in accuracy for both the training set and a separate test set with 90 additional host–guest systems. The optimized radii also showed encouraging transferability from host–guest systems to 59 protein-ligand systems. However, the new radii are significantly smaller than the baseline radii and lead to excessively favorable hydration free energies (HFE). Thus, users of the OBC2 GB model currently may choose between GB cavity radii that yield more accurate binding affinities or GB cavity radii that yield more accurate HFEs. We suspect that achieving good accuracy on both will require more far-reaching adjustments to the GB model. We note that binding free energy calculations using the OBC2 model in OpenMM gain about a 10x speedup relative to corresponding explicit solvent calculations, suggesting a future role for implicit solvent absolute binding free energy (ABFE) calculations in virtual compound screening. This study proves the principle of using host–guest systems to train potential functions that are transferrable to protein–ligand systems, and provides an infrastructure that enables a range of applications.

1 Introduction

Molecular dynamics (MD) simulations with empirical force fields are a valuable tool in the field of computational chemistry and biophysics. MD simulations allow us to probe protein–ligand systems at the molecular level and estimate physical quantities of interest at a lower cost than experimentation and more quickly and with better conformational sampling than quantum mechanical methods that explicitly treat electronic degrees of freedom. A quantity of particular interest in computer-aided drug design (CADD) is the affinity, or binding free energy, of a targeted protein with a candidate small molecule ligand. This quantity can be computed with MD simulations,[1-6] but these calculations remain time-consuming and error-prone. The errors can arise from a number of sources, including uncertainties in the protonation states of the protein and ligand, both before and after binding; the difficulty of establishing a complete and suitable initial protein structure with a sufficiently accurate ligand pose; the challenge of achieving adequate numerical convergence of slow protein motions; and approximations made by the force field in modeling the potential energy surface. This array of error sources makes it difficult to isolate or attribute discrepancies from the experiment to any one source. Aqueous host–guest systems embody much the same physical chemistry, such as conformational changes upon binding, hydrogen bonding between host and guest, and hydrophobic interactions. However, they are much smaller and simpler than proteins, enabling faster convergence and posing fewer ambiguities in protonation states and structures. For these reasons, host–guest systems have been used to test the accuracy of binding affinity calculations in the SAMPL blinded prediction challenges[7-11] and to evaluate empirical force fields in retrospective studies.[12-14]

Indeed, the force field is a critical determinant of the accuracy of a binding free energy calculation. Empirical force fields are typically tuned to fit quantum mechanical (QM) chemical data and/or liquid-state physical properties. There are also recent efforts to fit empirical force fields to small molecule crystal structures.[15] However, curating the experimental data and organizing them into a uniform format for use in testing and tuning force fields can be laborious. In addition, it is a non-trivial task to optimize

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

force field parameters to many physical properties simultaneously. The Open Force Field (OpenFF) initiative has addressed this challenge by creating OpenFF-Evaluator,[16] a software framework that provides an automated and highly scalable interface to estimate multiple physical properties, such as densities, dielectric constants, enthalpies, and hydration free energies (HFE). In addition, OpenFF has curated experimental data for these physical properties from the National Institute of Standards and Technology (NIST) ThermoML[17] and FreeSolv[18,19] databases. The framework is also designed to integrate well with ForceBalance,[20-23] which allows the optimization of force field parameters against the physical property data. However, even though a key intended application of OpenFF force fields is the calculation of protein–ligand affinities in support of structure-based drug discovery, binding data have not so far been integrated into the standard force field training set.

There is some precedent for using host–guest binding thermodynamics to tune force field parameters. For example, experimental binding affinities of octa-acid complexes were used to reoptimize the oxygen Lennard-Jones (LJ) parameters of the TIP3P water model, resulting in a new water model called Bind3P.[24] Bind3P gave a consistent drop in the mean signed error in a host–guest test set and improved HFEs of small organic molecules, while still providing accurate results for the properties of pure water. In another study, selected solute LJ parameters were optimized against experimental cucurbit[*n*]uril-guest binding enthalpies, leading to improved test-set accuracy.[25] However, although these trial studies yielded favorable results, they did not offer a way of systematizing and automating the process of training force fields against host–guest binding thermodynamics. Here, we implement this concept by extending the OpenFF-Evaluator framework to estimate host–guest binding free energies and their gradients with respect to force field parameters, and integrating this capability with ForceBalance.

As an initial case study, we use this new capability to optimize the parameters of a generalized Born (GB)[26] implicit solvent model. There are several advantages to working at this stage with an implicit solvent model rather than optimizing, e.g., LJ parameters in the context of explicit solvent. Firstly, using continuum electrostatics to approximate the interactions of atoms with water reduces the system size from thousands to hundreds of atoms, affording a ~10x computational speedup, which allows us to include more host–guest systems in our training set. Secondly, prior computational studies of cyclodextrin–guest complexes have proven the applicability of host–guest systems to evaluate implicit solvent models.[13,27] Thirdly, fewer parameters are involved in optimizing implicit solvent models than e.g., nonbonded force field parameters, due to the small number of GB atom types, so we have a favorable parameter-to-training data ratio. Finally, we note that the greater speed of absolute binding free energy calculations with a GB model means that an improved GB model that can reproduce binding affinities accurately would be invaluable for high-throughput virtual screening of candidate ligands for targeted proteins. We focus on the OBC2[28] form of the GB model, which was originally optimized to fit energies derived from finite difference solutions of the Poisson-Boltzmann (PB) equation.[26] Though there are more advanced models, such as the GBneck[29,30] and R6 effective radii[31] approaches, and the AGBNP2 model, which adds further interaction terms,[32,33] we selected the OBC2 model due to its simplicity, speed, and ready availability in the OpenMM simulation package.[34]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Here we report the optimization of GB cavity radii used in the OBC2 implicit solvent model against a collection of host–guest binding data and the evaluation of the resulting adjusted GB model against a larger test set of benchmarks comprising additional host–guest binding data, protein–ligand binding data, and small molecule HFEs. We find that the new GB cavity radii decrease the accuracy of the HFEs but provide excellent transferability to the test set binding free energies. This pattern of results traces to marked reductions in certain GB cavity radii by the optimization process. Our extended OpenFF-Evaluator infrastructure and the present case study provides a novel foundation for future work integrating host–guest binding data into the experimental data sets used to tune empirical force fields. We believe this approach will generate force fields that provide increasingly accurate protein–ligand binding affinity predictions.

2 Methods

2.1 Optimization of GB Cavity Radii Against Host–Guest Binding Data

The standard OpenFF-Evaluator framework can estimate several condensed-phase physical properties, including liquid densities, dielectric constants, enthalpies of mixing, and hydration free energies. Here, we extend the OpenFF-Evaluator framework to include a workflow that estimates host–guest absolute binding free energies using either explicit solvent or generalized Born surface area (GBSA) implicit solvent. We have also modified the hydration free energy workflow to work with implicit solvents. The extension to the framework is illustrated in Figure 1. Currently, the OpenFF-Toolkit can assign small molecule parameters for the HCT[35] and the OBC[28] implicit solvent models. Here, we chose to focus on the OBC2 variant of the OBC model for the implicit solvent, but other models can be implemented using the OpenFF-SMIRNOFF plugin (https://github.com/openforcefield/smirnoff-plugins).

The hydration free energy from the OBC model is given by the following equations

$$\Delta G_{\mathrm{GB}}^{\mathrm{self}} = -\frac{1}{2}\left(\frac{1}{\epsilon_{\mathrm{solute}}} - \frac{1}{\epsilon_{\mathrm{solvent}}}\right)\sum_i \frac{q_i^2}{R_i}, \tag{1a}$$

$$\Delta G_{\mathrm{GB}}^{\mathrm{cross}} = -\frac{1}{2}\left(\frac{1}{\epsilon_{\mathrm{solute}}} - \frac{1}{\epsilon_{\mathrm{solvent}}}\right)\sum_{ij} \frac{q_i q_j}{f_{\mathrm{GB}}(d_{ij}, R_i, R_j)}, \tag{1b}$$

$$\Delta G_{\mathrm{SA}} = E_{\mathrm{SA}}.4\pi\sum_i (r_i + r_{\mathrm{solvent}})^2\left(\frac{r_i}{R_i}\right)^6. \tag{1c}$$

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Here, i and j are atom indices, ϵ is the dielectric constant, q_i is the partial charge, d_{ij} is the distance between atoms i and j, r_i is the GB cavity radius for atom i, r_{solvent} is the radius of water in the surface-area term, R_i is the effective GB radius, f_{GB} is the effective interaction distance and E_{SA} is the surface area (SA) energy scale (2.26 kJ/mol/nm^2). The SA term in Eq (1c) is a modified version of Eq (2) from Schaefer et al.[36] that is implemented in OpenMM, which is different from the SA terms implemented in the AMBER program.[37,38] The effective interaction distance is given by[39]

$$f_{\text{GB}} = \left[d_{ij}^2 + R_i R_j \exp\left(-\frac{d_{ij}^2}{4 R_i R_j} \right) \right]^{1/2}. \tag{2}$$

In the OBC model, the GB cavity radius r_i is transformed into an effective radius R_i, which depends on the configuration of the system via the equation

$$R_i = \frac{1}{\rho_i^{-1} - r_i^{-1} \tanh(\alpha \Psi_i - \beta \Psi_i^2 + \gamma \Psi_i^3)}, \tag{3}$$

where $\rho_i = r_i - 0.09$ Å is the shifted radius and $\alpha = 1.0$, $\beta = 0.8$, and $\gamma = 4.85$ for OBC2. The variable Ψ_i is an integral of van der Waals (vdW) spheres of all atoms multiplied by the shifted GB radius

$$\Psi_i = \frac{\rho_i}{4\pi} \int_{\text{vdW}} \theta(|r| - \rho_i) \frac{1}{|r|^4} d^3 r, \tag{4}$$

where the step function $\theta(r)$ excludes the volume of atom i from the integration. We optimized the GB cavity radii, r_i, using mBondi2 radii[28] as the starting point. These are based on the Bondi[40] radii but provide an increased radius for hydrogen bound to nitrogen from 1.2 Å to 1.3 Å (see Table 1).

The baseline OpenFF-Evaluator framework can also estimate the gradients of many physical properties with respect to force field parameters. Here, we also implement the gradient calculations for host–guest absolute binding free energy (ABFE) into the OpenFF-Evaluator framework using the following relation[25]

$$\frac{\partial \Delta G_{\text{d}}}{\partial \theta} = \left\langle \frac{\partial U}{\partial \theta} \right\rangle_{\text{bound}} - \left\langle \frac{\partial U}{\partial \theta} \right\rangle_{\text{free}}, \tag{5}$$

where U is the total potential energy, θ is the force field parameter, and $\langle\rangle_{\text{bound}}$ and $\langle\rangle_{\text{free}}$ indicate ensemble averages of the host–guest system without any restraining potentials included in the bound and unbound (i.e., free) states, respectively. These two states

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

correspond to the first and last windows of the attach–pull–release (APR) calculations (see Method 2.2 for details). The restraining potentials in the bound and free states do not alter the internal coordinates of the host–guest molecules and, thus, do not contribute to the total potential energy. The restraining potentials in the free state keep the host and guest molecules far apart; thus, we simulate the free host and free guest in the same system. This approach reduces the number of simulations from four to two, as was originally defined by Yin et al.[25] In practice, the gradient $\partial U / \partial \theta$ is estimated using the central difference method by re-evaluating the potential energy U with a perturbed θ

$$\frac{dU}{d\theta} \approx \frac{U(\theta + h) - U(\theta - h)}{2h}. \tag{6}$$

We used a step size of $h = \theta \times 10^{-4}$ Å to perturb the GB cavity radii. For host–guest complexes with multiple binding poses N_b, as in the case of cyclodextrins, the binding free energy gradient is given by the following equation

$$\frac{\partial \Delta G_{\mathrm{b}}}{\partial \theta} = \frac{\sum_{i=1}^{N_b} \frac{\partial \Delta G_i}{\partial \theta} \exp(-\beta \Delta G_i)}{\sum_{i=1}^{N_b} \exp(-\beta \Delta G_i)}. \tag{7}$$

The equation above is a weighted sum of the gradients over all binding poses with the Boltzmann probability of the binding affinities ΔG_{b} as the weights. The force field parameter θ optimization is handled with ForceBalance[20-23] (https://github.com/leeping/forcebalance), and we used a regularized least-squares objective function

$$L(\theta) = \sum_{n}^{N} \frac{w_n}{M_n} \sum_{m}^{M_n} \left(\frac{y_m^{ref} - y_m(\vec{\theta})}{d_n} \right)^2 + \sum_{i \in \theta} \frac{|\Delta\theta_i|^2}{\sigma_i}, \tag{8}$$

where N is the number of types of physical properties, M_n is the number of systems for each physical property, w_n is a weight associated with each property, d_n is a scaling factor with the same unit as the property, y_m^{ref} is the experimental value, y_m is the estimated value from simulation, $\vec{\theta}$ is a vector of parameters being optimized and σ_i is the prior width hyperparameter. In this study, because there is only one property being optimized, $N = 1$ and there is only a single weight $w = 1$. The scaling factor is set as $d_n = 1$ kcal / mol, and chosen based on an order-of-magnitude estimate of the experimental measurement error. Therefore, the first term in Eq. 8 is simply the mean squared error of the binding free energy in units of $(\mathrm{kcal/mol})^2$.

ForceBalance uses a Gauss-Newton approach to optimize the parameters by using the property gradients from Eq (7) to build the gradient and approximate the Hessian of the objective function. Here, we optimize θ to fit host–guest experimental binding data. The

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

objective function in Eq (8) includes an L2 regularization penalty, and we use a prior width σ_i of 0.5 Å for all radii.

We implemented the workflow for host–guest ABFE calculations in implicit solvent and its gradient with respect to θ in a forked version of OpenFF-Evaluator (https://github.com/jeff231li/openff-evaluator). A GitHub repository is available online detailing the installation of the software, with its dependencies, and a short tutorial on running an optimization (https://github.com/jeff231li/HG-GBSA-Optimization).

2.2 Calculation of Host–Guest Absolute Binding Free Energies

We estimate the ABFE, ΔG_b, for host–guest complexes using the APR method[41,42] as implemented in version 1.2.0 of the Python package pAPRika (https://github.com/GilsonLabUCSD/pAPRika), which is integrated into OpenFF-Evaluator. The protocol for the host–guest calculations is similar to that of an earlier study of cyclodextrin complexes.[12,14] Briefly, we apply Boresch-style restraints[2] on the host and guest molecules to restrain their translational and orientational motion. These restraints consist of bond, angle, and dihedral restraints that orient the host–guest complex to pull the guest molecule in the z-axis. To aid in defining the Boresch restraints, we include three noninteracting dummy atoms in each host–guest complex. Note that the Boresch restraints applied to the host molecule do not contribute to the binding free energy. We use a spring constant of 5 kcal/mol/Å^2 for bond restraint and 100 kcal/mol/rad^2 for the angle and dihedral restraints. To improve convergence in the pull phase calculations, we also apply conformational restraints on the host molecules, which is particularly helpful for larger guest molecules that tightly bind to the host cavity. For cyclodextrins, we apply two dihedral restraints between neighboring glucose units $O5_n - C1_n - O1_n - C4_{n+1}$ and $C1_n - O1_n - C4_{n+1} - C5_{n+1}$ with a weak spring constant of 6 kcal/mol/rad^2. For cucurbit[n]urils, we use several distance jack restraints placed between amide nitrogen across the portal region to enlarge the opening. The jack distance is 13.5 Å and 15.0 Å for CB7 and CB8, respectively, with a spring constant of 15 kcal/mol/Å^2. For octa-acids, we apply jack restraints with a distance of 15 Å on opposite atoms across the portal with a spring constant of 25 kcal/mol/Å^2. Figure S1 shows the different conformational restraints applied to cyclodextrins, cucurbit[n]urils, and octa-acids. The attach and release phases include the free energy costs of applying and releasing the conformational restraints on the host molecule. The initial structures and definition of the anchor atoms and restraints are available in the *Taproom* repository (see Method 2.5 for details). During the attachment phase, we scale the force constant on the restraints across 15 independent windows [0.0, 0.004, 0.008, 0.016, 0.024, 0.04, 0.055, 0.0865, 0.118, 0.181, 0.244, 0.37, 0.496, 0.748, 1.0]. We then pull the guest molecule out of the host cavity by 18 Å over 46 equally-spaced windows. The release phase is separated into two contributions; (1) release of the host conformational restraints and (2) release of the guest restraints in bulk (i.e., at r = 18 Å). The former is calculated over the same 15 windows as in the attach phase but in reverse and without the guest molecule. The latter is calculated analytically, including the standard-state correction at 1 M. For cyclodextrin complexes, we perform the APR calculations for two binding poses due to the asymmetry of the host molecules. The binding poses are defined by the orientation of the guest molecule with respect to the

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

primary/secondary face of the cyclodextrin, and we combine the binding free energies of the two poses with[42,43]

$$\Delta G_{\mathrm{b}} = -RT \ln[\exp(-\beta \Delta G_{\mathrm{p}}) + \exp(-\beta \Delta G_{\mathrm{s}})]\,. \tag{9}$$

A brief derivation of this equation is also provided in Section 5. We perform an energy minimization for each APR window followed by 1 ns of equilibration and 30 ns of the production run. The free energy across the three phases was computed with the thermodynamic integration (TI) method, and blocking analysis was used to estimate the uncertainty in work done by each restraint in each window. The final SEM in ΔG_{b} is obtained from bootstrap resampling (2000 iterations).

2.3 Calculation of Protein–Ligand Absolute Binding Free Energies

For the protein–ligand systems, we chose the alchemical double–decoupling method (DDM)[1] instead of APR because it is often more challenging to define a suitable pulling pathway for a protein–ligand system. The protein–ligand ABFE calculations were performed outside of the OpenFF-Evaluator framework, using pAPRika and other Python tools, because a workflow to handle protein–ligand ABFE is not yet implemented in the framework. The alchemical calculation protocol is similar to those of Aldeghi et al.[44,45] and Alibay et al.[6] First, we annihilate the partial charges on the ligand in the bulk and perform the opposite in the binding site. The GBSA interactions are annihilated simultaneously with the partial charges, and the nonbonded potentials were alchemically modified with *openmmtools* version 0.21.5 (https://github.com/choderalab/openmmtools). To accomplish this, *openmmtools* introduces the following alchemical attenuation function to the GBSA potential

$$s_i(\lambda, \eta_i) = \lambda \eta_i + (1 - \eta_i), \tag{10a}$$

$$s_{ij}(\lambda, \eta_i, \eta_j) = s_i(\lambda, \eta_i) \cdot s_j(\lambda, \eta_j), \tag{10b}$$

where i and j are atom indices, λ is the alchemical coupling parameter, and η is an indicator function that is equal to 1 for atoms in the alchemical region and 0 otherwise. The attenuation function above ensures that only interactions involving the ligand atoms are perturbed. The GBSA potential within the protein is left unperturbed. Inserting the attenuation function of Eq (10) to the GBSA potential in Eqs (1a)-(1c) gives the alchemically-modified potential

$$\Delta G_{\mathrm{GB}}^{\mathrm{self}}(\lambda, \eta) = -\frac{1}{2}\left(\frac{1}{\epsilon_{\mathrm{solute}}} - \frac{1}{\epsilon_{\mathrm{solvent}}}\right) \sum_i s_i(\lambda, \eta_i) \frac{q_i^2}{R_i},$$

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(11a)

$$\Delta G_{\mathrm{GB}}^{\mathrm{cross}}(\lambda,\eta) = -\frac{1}{2}\left(\frac{1}{\epsilon_{\mathrm{solute}}}-\frac{1}{\epsilon_{\mathrm{solvent}}}\right)\sum_{i,j} s_{ij}(\lambda,\eta_i,\eta_j)\frac{q_i q_j}{f_{\mathrm{GB}}(d_{ij},R_i,R_j)},$$

(11b)

$$\Delta G_{\mathrm{SA}}(\lambda,\eta) = E_{\mathrm{SA}}.4\pi\sum_i s_i(\lambda,\eta_i)(r_i+r_{\mathrm{solvent}})^2\left(\frac{r_i}{R_i}\right)^6.$$

(11c)

We perform the annihilation of the electrostatic interactions of the guest in the bulk and binding site in separate systems over 11 equally-spaced λ values. The vdW interactions between the ligand and the protein are decoupled with the soft-core potential from Pham and Shirts[46] across 21 equally-spaced λ values. We applied Boresch-style restraints on the ligand in reference to the protein to restrict its translational and orientational motion during the alchemical transformation. To simplify the procedure, we used the *MDRestraintsGenerator* (https://doi.org/10.5281/zenodo.4570555) Python program to determine the anchor atoms automatically from 5 ns of unrestrained MD simulation. The free energy cost of applying the Boresch restraints is split across 15 windows, similar to the host–guest attach phase calculations. The free energy cost of releasing the restraints with the standard-state correction at 1 M is done analytically. We use a force constant of 5 kcal/mol/Å^2 and 100 kcal/mol/rad^2 for the distance and angle restraints, respectively. For each ABFE window, we performed an energy minimization followed by 2 ns and 20 ns of equilibration and production run, respectively. The free energy is recovered using the TI method, and the error in ΔG_{b} was estimated with bootstrap resampling similar to the host–guest calculations.

2.4 Calculation of Hydration Free Energies

We compute ΔG_{solv} alchemically using the Yank program (https://doi.org/10.5281/zenodo.3534289) through the OpenFF-Evaluator framework. The GBSA potential was annihilated simultaneously with the partial charges, as in the protein–ligand calculations. Here, the electrostatic interactions are annihilated in both the solution and vacuum environments. For the vdW interactions, we decouple the molecule's interaction only in solution, thus removing the need to perform the calculation in vacuum. For both electrostatics and vdW interactions, we allow Yank to automatically determine the number of λ points with the *trailblaze* algorithm from Rizzi et al.[47] (unpublished). We run Yank with Hamiltonian replica exchange over 2000 iterations with a time step of 2 fs. The analysis was done using Yank's analysis framework, which employs the multistate Bennett acceptance ratio (MBAR) method[48] to recover the free energy from the MD trajectories.

2.5 Training and Test Sets

We curated 126 aqueous host–guest complexes involving cyclodextrin, cucurbit[*n*]uril, and octa-acid hosts. For complexes with cyclodextrins, there are 44 host–guest complexes

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

with either α-cyclodextrin (αCD) or β-cyclodextrin (βCD) hosts. The majority of these complexes were obtained from Rekharsky et al.,[49] and one complex was curated from SAMPL7[10] (βCD with R-rimantadine[50]). These cyclodextrin host–guest complexes have been studied computationally[12,14] and include a series of small molecules containing ammonium, carboxylate, and cyclic alcohol groups. For cucurbit[*n*]urils, we selected 23 complexes with cucurbit[7]uril (CB7) and 22 complexes with cucurbit[8]uril (CB8). These were curated from SAMPL challenges 3,[51] 4,[7] 6,[9] and 8[11] and several more CB7 complexes were curated from Moghaddam et al.[52] and Kim et al.[53] We also selected two variants of the octa-acid host molecule, octa-acid[54] (OA) with 23 complexes and tetra-endo-methyl octa-acid[55] (TEMOA) with 14 complexes. The octa-acid complexes were curated from SAMPL challenges 4,[7] 5,[8] and 6.[9] All carboxylic acid and amine groups were treated as ionized because their expected pK_a were far from the neutral pH at which the binding measurements were done. We have made the structures of these 126 host–guest complexes available in the *Taproom* repository (https://github.com/slochower/host-guest-benchmarks). In addition to the structures, metadata files are also included that are read by OpenFF-Evaluator and contain the information necessary to perform APR calculations on these complexes. We hand-selected 36 complexes for the training set, six from each host molecule, leaving 90 complexes for the test set. The host–guest complexes used in the training set are shown in Figure 2, and those in the test set are shown in Figures S2-S4.

For the protein–ligand benchmark, we use 59 protein–ligand systems previously investigated by Alibay et al.[6] because of the wide range of ΔG_b (−2.7 to −12.6 kcal/mol) and the availability of the bound structures (https://doi.org/10.5281/zenodo.5913469). These systems are based on published fragment-based drug designs (FBDD) for four protein targets: PWWP1, HSP90, MCL-1, and Cyclophilin D.[56-59] The protonation states of all ligands were kept the same as the input files in the GitHub repository (https://github.com/bigginlab/fragment-opt-abfe-benchmark).

For the hydration free energy tests, we focused on non-ionized molecules, thus avoiding ambiguity in interpreting the experimental data. The data are drawn from the FreeSolv database,[18,19] which has been tried and tested in previous studies.[60] The FreeSolv database currently contains 643 neutral small molecules. We selected 100 molecules containing only hydrogen, carbon, nitrogen, and oxygen atoms. With the chosen molecules, the ΔG_{solv} ranges from +3 kcal/mol to −25 kcal/mol. The chemical structures are shown in Figure S13.

2.6 Simulation Details and Force Fields Used

We performed all MD simulations with the OpenMM MD engine version 7.5.1.[34] Except for the protein–ligand binding calculations, the OpenMM simulations were run through the OpenFF-Evaluator framework.[16] For the hydration free energy and host–guest binding affinity calculations, we use the Sage (OpenFF 2.0.0[61]) force field from OpenFF with AM1-BCC[62,63] partial charges and ELF10 (electrostatically least-interacting functional) conformer selection. Partial charges were assigned to the small molecules with the OpenFF-Toolkit using OpenEye OEChem (version 2022) as its backend. We ran four sets of calculations for each protein–ligand system, the result of testing two AMBER force fields for the protein, i.e., ff99SB-ILDN[64,65] and ff14SB,[66] and two small molecule force fields for

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

the ligands, i.e., Sage and GAFF2.1. With GAFF2.1, we use AM1-BCC partial changes, and the AMBER parameters were applied to the protein and ligand using the AmberTools[67] program suite. ParmEd version 3.4.3 (https://github.com/ParmEd/ParmEd) was used to combine the ligand parameterized with Sage and the protein with the AMBER force field. In all simulations, the temperature was maintained at a constant value of 298.15 K with the Langevin thermostat and a 2 fs integration time step.

3 Results and Discussion

3.1 Standard GB Cavity Radii Overestimate Host–Guest Affinities

As shown in Figure 3A and detailed in Table S1, ABFE calculations with the original set of GB cavity radii markedly overestimate the affinities of our 36 training-set host–guest systems. The RMSE for the whole training set is 21.0 kcal/mol, with a coefficient of determination R^2 of 0.7. The least accurate results are for the cucurbit[n]uril complexes, for which the computed binding free energies range as negative as −68.5 kcal/mol, and the worst results are for those involving cucurbit[7]uril (CB7), with an RMSE of 41.2 kcal/mol (Table S4). We confirmed the robustness of the problematic CB7 results – i.e., the gross overestimation of their binding – across force fields and software implementations by redoing the runs in AMBER using the GAFF2 force field parameters and the default GB cavity radii. The binding affinities are still grossly overestimated, with a mean signed error of −35 kcal/mol compared with −40 kcal/mol for Sage in OpenMM; see Table S2 for details. The results are less problematic for the cyclodextrins and octa-acid systems, which give RMSEs of 3.7 and 5.8 kcal/mol, respectively. The mean signed error (MSE) for cyclodextrins, octa-acids, and cucurbit[*n*]urils are −3.3, −3.5, and −34.9 kcal/mol, respectively. Thus, on average, ABFE calculations with the baseline OBC2 implicit solvent model overestimate the binding free energies of the cucurbit[*n*]uril complexes ~10 times more than the cyclodextrin and octa-acid complexes.

The overestimation of the cucurbit[*n*]uril affinities observed here traces to excessively favorable net electrostatic interactions between cationic guests and the cucurbit[*n*]urils' rings of electronegative carbonyl atoms. As shown in Figure S6, the electrostatic potentials generated by CB7 *in vacuo* are much stronger than those generated by the cyclodextrins. Thus, the overall host-guest electrostatic interaction, which is the sum of the *in vacuo* electrostatic energy and the GB solvation free energy, is highly sensitive to the details of the GB model, and there is clearly an imbalance between the direct, Coulombic, *in vacuo* interaction and the solvent-mediated interaction provided by the GB model. It is worth noting that calculations of cucurbit[*n*]uril-guest binding in explicit solvent[42] also show an overestimation of the binding affinities, though not as severe as the results with OBC2 implicit solvent presented here. A previous study by Zhang et al.[13] revealed that the GB implicit solvent model underestimates the screening of Coulomb interactions in cyclodextrins by a factor of three compared to explicit solvent. In summary, the excessively negative ΔG_b values observed here for cucurbit[*n*]urils reflect underscreening of strong Coulomb interactions by the baseline OBC2 implicit solvent model.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

3.2 Optimized GB Cavity Radii Improve Host–Guest Accuracy for Training and Test Sets

Optimization of the five GB cavity radii (carbon, nitrogen, oxygen, general hydrogen, and hydrogen bonded to nitrogen) led to a marked reduction in error for the training set (Figure 3B, Table S4) with the overall RMSE falling from 21.0 kcal/mol to 2.9 kcal/mol. Improvements were also seen for each separate host class, falling to 2.5 kcal/mol, 1.4 kcal/mol, and 4.4 kcal/mol for the cucurbit[*n*]urils, cyclodextrins, and octa-acids, respectively. However, the coefficient of determination R^2 for the whole training set decreased slightly from 0.7 to 0.6 (Figure 3A). The individual analysis reveals that cucurbit[*n*]urils gained an improvement in R^2 from 0.2 to 0.4, while cyclodextrin complexes decreased from 0.5 to 0.3 (Table S4). For the octa-acids, the R^2 stayed the same at around 0.2. The MSE for each host class also improved significantly to about −1.0 kcal/mol for cucurbit[n]urils and octa-acids and −0.1 kcal/mol for cyclodextrins. The evolution of the ForceBalance optimization is summarized in Figure S5. The objective function, defined in Eq (8), starts at about 450 and drops to 140 after one iteration (Figure S5A). After nine iterations, the objective function fluctuates around 8–10, and the convergence criterion is reached at iteration 22.

We also track the GB cavity radii's evolution over the course of the optimization (Figures S5B-F); the initial and final values are listed in Table 1. We will refer to the optimized radii as HG-optimized from here on. For all atom types except oxygen, the GB cavity radius decreased from its original value, while that of oxygen increased by about 0.13 Å. The final radii for hydrogen and nitrogen are unphysically small, with values of 0.71 Å and 0.53 Å, respectively. As mentioned in the previous section, the observed overestimation in ΔG_b is due to the weak screening of Coulomb interactions using OBC2 implicit solvent with the original radii values. To correct this, ForceBalance shrinks the radius of hydrogen and nitrogen to make them more hydrated, effectively increasing the screening of Coulomb interactions, which decreases the magnitude of ΔG_b. Note that the results of this optimization are influenced by the value of the hyperparameter σ_i in Eq 8, with smaller values leading to smaller changes in the radii from their initial values. The sensitivity of the present results to σ_i is considered in Section S4 of the SI.

The trained GB parameters also provide a large improvement in accuracy for the host–guest test set, which, with 90 systems, is about three times the size of the training set (Figure 4 and Table S5). Just as for the training set, the binding free energies are overestimated with the original mBondi2 radii, especially for the cucurbit[*n*]uril complexes, and the new radii correct this bias. Going from the original radii to the trained set reduces the overall RMSE from 19.5 to 2.1 kcal/mol and the R^2 value from 0.5 to 0.8. The cucurbit[n]uril complexes improved the most, with RMSE falling from 32.1 to 2.8 kcal/mol. Similarly, cyclodextrin complexes improved from 3.4 to 1.2 kcal/mol, while the RMSE stayed roughly the same for octa-acids. We see a big improvement in the R^2 value for cucurbit[*n*]urils and octa-acids, but the value fell slightly from 0.6 to 0.5 for cyclodextrins.

Interestingly, the HG-optimized radii actually gave better accuracy for the test set than for the training set, with RMSE lower by 1.0 kcal/mol and R^2 increased by about 0.3. Closer examination reveals that the RMSE of the cyclodextrins and cucurbit[*n*]urils are similar between the training and tests, but the RMSE of the octa-acids falls from 4.4 kcal/mol in

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

the training set to 1.8 kcal/mol in the test set. We observed that the octa-acid systems in the training set that are overestimated (more negative) by at least 4 kcal/mol when using the HG-optimized radii (Figure 3B and Table S1) have guest molecules with cationic nitrogens (Figure 2). We suspect that the reduction in the nitrogen GB radius, which greatly improves $\Delta^{\wedge}$ for cucurbit[*n*]urils, slightly worsens the accuracy for octa-acids with ammonium guests, and that this may explain the fact that the overall correlation with experiment is better for the test set than the training set.

3.3 GB Cavity Radii Tuned for Host–Guest Binding Reduce the RMSE for Protein–Ligand Binding Calculations

We hope that training a potential function against host–guest binding data will lead to improved accuracy for protein–ligand binding free energies. We therefore ran ABFE calculations for 59 protein–ligand systems with the original mBondi2 radii and with our HG-optimized radii. We performed these calculations with two AMBER force fields for the protein and two small molecule force fields for the ligands. The results for protein force field ff14SB are illustrated in Figure 5. As already seen for the host–guest systems, ABFE calculations using OBC2 with its standard mBondi2 radii lead to marked overbinding (i.e., ΔG_b, too negative). The HSP90 complexes are overestimated the most, with an RMSE of 10 kcal/mol. We suspect this may be related to the importance of an explicit treatment of key water molecules in the binding site.[6,68] However, further investigation is needed to assess this. Very encouragingly, the new GB radii optimized against host–guest binding data, generate markedly reduced errors for the protein-ligand systems, with the RMSE falling from 7.4 to 2.4 kcal/mol for Sage and from 6.4 to 2.0 kcal/mol for GAFF2. However, the value of R^2 falls by about a factor of 2 using the optimized radii. Thus, the HG-optimized radii transfer reasonably well to protein–ligand systems.

Similar trends are obtained with protein force field ff99SB-ILDN (Figure S12), and the same trends are also observed for each of the four proteins individually (Tables S10 and S11). Interestingly, our results for ff99SB-ILDN and GAFF2 with the HG-optimized radii give a slightly better RMSE than the explicit solvent results from Alibay et al.;[6] 2.8 kcal/mol in explicit solvent and 2.5 kcal/mol in implicit solvent with the HG-optimized radii. However, the results in explicit solvent provide a higher correlation with experimental data than in implicit solvent, R^2 of 0.8 compared to 0.1.

We also examined the consequences of the new GB cavity radii for simple protein simulations by performing 100 ns of unrestrained MD for each protein, using ff14SB, without any bound ligand, in implicit solvent with the mBondi2 and HG-optimized radii, as well with the explicit TIP3P water model. As shown by the RMSD plots in Figure 6, all four proteins are stable when simulated with explicit solvent, with the RMSDs fluctuating around 1–2 Å. The proteins are also stable when simulated in OBC2 implicit solvent with the original mBondi2 radii, with RMSDs similar to TIP3P except for HSP90, where we observed excess flexibility in residues 100–124. However, implicit solvent simulations with the HG-optimized radii tended to generate greater flexibility, with RMSD values of about 2–4 Å. PWWP1 remains the most stable, with an average RMSD close to the TIP3P results. Inspection of the trajectories reveals that the larger RMSDs with HG-optimized radii are

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

due to increased flexibility in some alpha helices. This makes sense, as the HG-optimized shrinks the GB cavity radii for hydrogen and nitrogen atoms, resulting in weaker backbone hydrogen bonds along the helix. Superposed representative structures of the four proteins, comparing their conformations in TIP3P and OBC2(HG-optimized) solvent, are shown in Figure 7. Despite the increased flexibility in some protein regions, all four proteins maintained their tertiary structure with the HG-optimized radii. We also note that the ligands were stable in their binding pockets during the ABFE calculations.

3.4 HG-Optimized GB Cavity Radii Lead to Excessively Favorable Hydration Free Energies

An ideal set of GB cavity radii would provide not only accurate binding free energies but also accurate HFEs. We tested this by using both the original (mBondi2) and HG-optimized radii to estimate ΔG_{solv} for 100 neutral small molecules from the FreeSolv database (see Methods 2.4), using the Sage small molecule force field. As shown in Figure 8, the original mBondi2 radii give good accuracy, with an RMSE of 2.1 kcal/mol and R^2 of 0.8. However, the HG-optimized radii yield marked overestimation of the magnitudes of the HFEs (Figure 8B), ΔG_{solv}, increasing the RMSE to 19.1 kcal/mol and decreasing the R^2 to 0.4. Thus, fitting GB cavity radii to host–guest ΔG_{b} improves the accuracy of ABFE calculations, but this improvement comes at the cost of lower accuracy in ΔG_{solv}.

To probe this further, we inspected the small molecules in the outlier region, which we defined as an absolute error in $\Delta G_{\text{solv}} \geq 10\,\text{kcal / mol}$. The chemical structures in Figure S14 reveal that all but one of these outliers contain at least one nitrogen atom, whose GB cavity radius was reduced from 1.55 Å to 0.53 Å during the HG-driven optimization (Table 1). Such a reduction of the GB cavity radius increases the atom's interaction with the solvent, as observed for the HFEs, but also weakens the electrostatic interactions of polar solute atoms. Since the HG-based optimization is mainly driven by the excessively strong electrostatic interactions between the host and guest in the cucurbit[n]uril complexes, ForceBalance decreased the nitrogen radius to a very small value to reduce the large negative ΔG_{b}. This, however, comes at the cost of increasing the hydration of nitrogen atoms, thus giving the excessively favorable ΔG_{solv} observed here.

4 Conclusions and Outlook

The present study demonstrates the systematic optimization of a potential function against host–guest binding free energy data using an automatic and scalable workflow. This advance is made possible by our extension of the OpenFF-Evaluator software framework and curation of 126 host–guest complexes ready to be used in ABFE calculations. Our initial application involves the optimization of GB cavity radii in the OBC2 implicit solvent model, but the present technology can also be used to optimize force field parameters against binding data in the context of an explicit treatment of solvent.

We find that host–guest binding affinities are grossly overestimated by ABFE calculations with the baseline OBC2 model, and that our HG-optimized GB cavity radii dramatically improve the accuracy, with excellent transferability to a test set comprising 90 additional

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

host–guest systems. Importantly, the optimized parameters also transfer well to benchmark dataset of 59 protein–ligand binding free energies, demonstrating the utility of using host–guest systems as a model for binding and data for force field development. However, optimizing the GB radii to ΔG_b resulted in a marked deterioration of the ability of the OBC2 model to reproduce ΔG_{solv}. This results from the fact that the optimization against host–guest binding data drove a marked reduction of selected GB cavity radii, as needed to weaken the excessively attractive host–guest electrostatic interactions, particularly for the cucurbit[*n*]uril cases. However, the artificially small radii lead to excessively favorable HFEs.

From here, one can take different strategies toward further improvement of the implicit solvent model — a goal that we believe is well justified, given the 10x speedup of ABFE calculations that it affords. We suspect that it will be impossible to accurately fit both ABFEs and HFEs with the OBC2 model, but a model optimized for binding and not hydration may be of value, since binding affinities are certainly of more practical importance than the HFEs of small molecules. Our HG-optimized radii work well for binding already, but to improve the model further, one could increase the diversity of the training data to include host–guest complexes with other host molecule types and possibly include protein–ligand binding as well. The former can be done by adding more systems to *Taproom*, while the latter will require extending the Evaluator framework to include a protein–ligand binding affinity workflow. Including some protein–ligand binding in the training data would avoid the risk of overfitting to host–guest systems and might resolve the excessive protein flexibility observed in Figure 7. On the other hand, it may be difficult to arrive at a set of GB cavity radii that, when used with the standard OBC2 model, improves not only RMSE but also R^2 for protein–ligand systems, because changes in hydration are particularly marked for protein–ligand binding

Thus, it would be preferable to derive an implicit solvent model that can reproduce both ABFEs and HFEs, and we anticipate that such a model could yield further improvement in the accuracy of ABFE calculations because hydration certainly contributes to the balance of forces that determine the overall binding affinity. To accomplish this, one may need to modify the functional forms of the GB model. Accordingly, we are currently looking for better alternatives to the traditional GB kernel, f_{GB} in Eq (2) and to the formula used to compute the effective radii R_i in Eq (3). The simplest changes to these functional forms would retain their radially symmetric character, so that the solvation free energy depends explicitly only on interatomic distances, rather than on the angular distributions of atoms around each other. Given the directionality of solute-water hydrogen bonds, it may ultimately be important to use information about the angular distributions of atoms around each other as well.

New functional forms can be implemented through the OpenFF-SMIRNOFF plugin code, which integrates seamlessly with the OpenFF infrastructure. The plugin and the Evaluator framework may enable one to build a next-generation implicit solvent model from the ground up and fit them to physical properties rather than to finite difference PB energies as previously done.[28]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

More broadly, although the work presented here showcases the optimization of an implicit solvent model with the OpenFF-Evaluator framework, the same general procedure can be applied to other potential functions and force field components. For example, it would be of interest to reoptimize Sage's LJ parameters, in the context of an explicit solvent model, against a dataset comprising the original reference data used for Sage and supplemented by host–guest binding data. This will be more computationally demanding than the present study, but should be well within reach of current computing resources. This will add another restraint to the optimization, and the inclusion of host–guest binding affinities to the optimization may improve the ability of future force fields to accurately predict protein–ligand binding affinities.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgement

This work used computational resources from the Triton Shared Computing Cluster at UCSD. The authors acknowledge funding from the National Institute of General Medical Sciences (R01GM061300 and R01GM132386). The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of the NIH. The authors declare the following competing financial interest(s): M.K.G. has an equity interest in and is a cofounder and scientific advisor of VeraChem LLC.

5: Appendix: Combining the Binding Free Energies of Multiple Poses

Equation (9) can be simply derived as follows. Consider the binding of a receptor R with a ligand L, where the bound state is split into two nonoverlapping sectors of conformational space, termed pose 1 and pose 2. The equilibrium constants for the formation of each pose from the unbound state are given by[69]

$$K_1 = \frac{1}{C^\circ} e^{-\beta \Delta G_1^\circ} = \left(\frac{C_1}{C_R C_L} \right)_{eq} \tag{12}$$

$$K_2 = \frac{1}{C^\circ} e^{-\beta \Delta G_2^\circ} = \left(\frac{C_2}{C_R C_L} \right)_{eq} \tag{13}$$

where C° is the standard concentration, C_1 and C_2 are the concentrations of the complex in poses 1 and 2, respectively, C_R and C_L are the concentrations of the unbound receptor and ligand, respectively, and)e_q indicates a quantity at equilibrium. The total concentration of bound complexes is $C_1 + C_2$, so the overall equilibrium constant for binding is

$$K = \left(\frac{C_1 + C_2}{C_R C_L} \right)_{eq} \tag{14}$$

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

$$= K_1 + K_2 \tag{15}$$

Hence, the overall free energy of binding is given by

$$\Delta G^\circ = -RT \ln(C^\circ K) \tag{16}$$

$$= -RT \ln\left(e^{-\beta \Delta G_1^\circ} + e^{-\beta \Delta G_2^\circ}\right) \tag{17}$$

as written in Eq (9).

References

(1). Gilson MK; Given JA; Bush BL; McCammon JA The Statistical-Thermodynamic Basis for Computation of Binding Affinities: A Critical Review. Biophysical Journal 1997, 72, 1047–1069. [PubMed: 9138555]

(2). Boresch S; Tettinger F; Leitgeb M; Karplus M Absolute Binding Free Energies: A Quantitative Approach for their Calculation. Journal of Physical Chemistry B 2003, 107, 9535–9551.

(3). Woo HJ; Roux B Calculation of Absolute Protein–Ligand Binding Free Energy from Computer Simulations. Proceedings of the National Academy of Sciences 2005, 102, 6825–6830.

(4). Heinzelmann G; Gilson MK Automation of Absolute Protein–Ligand Binding Free Energy Calculations for Docking Refinement and Compound Evaluation. Scientific Reports 2021, 11, 1–18. [PubMed: 33414495]

(5). Gapsys V; Yildirim A; Aldeghi M; Khalak Y; van der Spoel D; de Groot BL Accurate Absolute Free Energies for Ligand–Protein Binding Based on Non-equilibrium Approaches. Communications Chemistry 2021, 4, 1–13. [PubMed: 36697560]

(6). Alibay I; Magarkar A; Seeliger D; Biggin PC Evaluating the Use of Absolute Binding Free Energy in the Fragment Optimization Process. Nature Communications Chemistry 2022, 5, 1–13.

(7). Muddana HS; Fenley AT; Mobley DL; Gilson MK The SAMPL4 Host–Guest Blind Prediction Challenge: An Overview. Journal of Computer-Aided Molecular Design 2014, 28, 305–317. [PubMed: 24599514]

(8). Yin J; Henriksen NM; Slochower DR; Shirts MR; Chiu MW; Mobley DL; Gilson MK Overview of the SAMPL5 Host–Guest Challenge: Are We Doing Better? Journal of Computer-Aided Molecular Design 2017, 31, 1–19. [PubMed: 27658802]

(9). Rizzi A; Murkli S; McNeill JN; Yao W; Sullivan M; Gilson MK; Chiu MW; Isaacs L; Gibb BC; Mobley DL; Chodera JD Overview of the SAMPL6 Host–Guest Binding Affinity Prediction Challenge. Journal of Computer-Aided Molecular Design 2018, 32, 937–963. [PubMed: 30415285]

(10). Amezcua M; El Khoury L; Mobley DL SAMPL7 Host–Guest Challenge Overview: Assessing the Reliability of Polarizable and Non-Polarizable Methods for Binding Free Energy Calculations. Journal of Computer-Aided Molecular Design 2021, 35, 1–35. [PubMed: 33392951]

(11). Amezcua M; Setiadi J; Ge Y; Mobley DL An Overview of the SAMPL8 Host–Guest Binding Challenge. Journal of Computer-Aided Molecular Design 2022, 36, 707–734. [PubMed: 36229622]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(12). Henriksen NM; Gilson MK Evaluating Force Field Performance in Thermodynamic Calculations of Cyclodextrin Host–Guest Binding: Water Models, Partial Charges, and Host Force Field Parameters. Journal of Chemical Theory and Computation 2017, 13, 4253–4269. [PubMed: 28696692]

(13). Zhang H; Yin C; Yan H; van der Spoel D Evaluation of Generalized Born Models for Large Scale Affinity Prediction of Cyclodextrin Host–Guest Complexes. Journal of Chemical Information and Modeling 2016, 56, 2080–2092. [PubMed: 27626790]

(14). Slochower D; Henriksen N; Wang LP; Chodera JD; Mobley DL; Gilson MK Binding Thermodynamics of Host–Guest Systems with SMIRNOFF99Frosst 1.0.5 from the Open Force Field Initiative. Journal of Chemical Theory and Computation 2019, 15, 6225–6242. [PubMed: 31603667]

(15). Park H; Zhou G; Baek M; Baker D; Dimaio F Force Field Optimization Guided by Small Molecule Crystal Lattice Data Enables Consistent Sub-Angstrom Protein–Ligand Docking. Journal of Chemical Theory and Computation 2021, 17, 2000–2010. [PubMed: 33577321]

(16). Boothroyd S; Wang LP; Mobley DL; Chodera JD; Shirts MR Open Force Field Evaluator: An Automated, Efficient, and Scalable Framework for the Estimation of Physical Properties from Molecular Simulation. Journal of Chemical Theory and Computation 2022, 18, 3566–3576. [PubMed: 35507313]

(17). Frenkel M; Chirico RD; Diky V; Dong Q; Marsh KN; Dymond JH; Wakeham WA; Stein SE; Königsberger E; Goodwin AR XML-based IUPAC Standard for Experimental, Predicted, and Critically Evaluated Thermodynamic Property Data Storage and Capture (ThermoML)(IUPAC Recommendations 2006). Pure and Applied Chemistry 2006, 78, 541–612.

(18). Mobley DL; Guthrie JP FreeSolv: A Database of Experimental and Calculated Hydration Free Energies, with Input Files. Journal of Computer-Aided Molecular Design 2014, 28, 711–720. [PubMed: 24928188]

(19). Mobley DL; Shirts M; Lim N; Chodera J; Beauchamp K; Wang LP Mobleylab/Freesolv: Version 0.52. 2018; https://zenodo.org/record/1161245.

(20). Wang LP; Head-Gordon T; Ponder JW; Ren P; Chodera JD; Eastman PK; Martinez TJ; Pande VS Systematic Improvement of a Classical Molecular Model of Water. Journal of Physical Chemistry B 2013, 117, 9956–9972. [PubMed: 23750713]

(21). Wang LP; Martinez TJ; Pande VS Building Force Fields: An Automatic, Systematic, and Reproducible Approach. Journal of Physical Chemistry Letters 2014, 5, 1885–1891. [PubMed: 26273869]

(22). Wang LP; McKiernan KA; Gomes J; Beauchamp KA; Head-Gordon T; Rice JE; Swope WC; Martínez TJ; Pande VS Building a More Predictive Protein Force Field: A Systematic and Reproducible Route to AMBER-FB15. Journal of Physical Chemistry B 2017, 121, 4023–4039. [PubMed: 28306259]

(23). Qiu Y; Smith DGA; Boothroyd S; Jang H; Hahn DF; Wagner J; Bannan CC; Gokey T; Lim VT; Stern CD; others, Development and Benchmarking of Open Force Field v1.0.0—the Parsley Small-Molecule Force Field. Journal of Chemical Theory and Computation 2021, 17, 6262–6280. [PubMed: 34551262]

(24). Yin J; Henriksen NM; Muddana HS; Gilson MK Bind3P: Optimization of a Water Model Based on Host–Guest Binding Data. Journal of Chemical Theory and Computation 2018, 14, 3621–3632. [PubMed: 29874074]

(25). Yin J; Fenley AT; Henriksen NM; Gilson MK Toward Improved Force-Field Accuracy through Sensitivity Analysis of Host–Guest Binding Thermodynamics. Journal of Physical Chemistry B 2015, 119, 10145–10155. [PubMed: 26181208]

(26). Onufriev AV; Case DA Generalized Born Implicit Solvent Models for Biomolecules. Annual Review of Biophysics 2019, 48, 275–296.

(27). Wickstrom L; He P; Gallicchio E; Levy RM Large Scale Affinity Calculations of Cyclodextrin Host–Guest Complexes: Understanding the Role of Reorganization in the Molecular Recognition Process. Journal of Chemical Theory and Computation 2013, 9, 3136–3150. [PubMed: 25147485]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(28). Onufriev A; Bashford D; Case DA Exploring Protein Native States and Large-Scale Conformational Changes with a Modified Generalized Born Model. Proteins: Structure, Function and Genetics 2004, 55, 383–394.

(29). Mongan J; Simmerling C; McCammon JA; Case DA; Onufriev A Generalized Born Model with a Simple, Robust Molecular Volume Correction. Journal of Chemical Theory and Computation 2007, 3, 156–169. [PubMed: 21072141]

(30). Nguyen H; Roe DR; Simmerling C Improved Generalized Born Solvent Model Parameters for Protein Simulations. Journal of Chemical Theory and Computation 2013, 9, 2020–2034. [PubMed: 25788871]

(31). Aguilar B; Shadrach R; Onufriev AV Reducing the Secondary Structure Bias in the Generalized Born Model via R6 Effective Radii. Journal of Chemical Theory and Computation 2010, 6, 3613–3630.

(32). Gallicchio E; Levy RM AGBNP: An Analytic Implicit Solvent Model Suitable for Molecular Dynamics Simulations and High-Resolution Modeling. Journal of Computational Chemistry 2004, 25, 479–499. [PubMed: 14735568]

(33). Gallicchio E; Paris K; Levy RM The AGBNP2 Implicit Solvation Model. Journal of Chemical Theory and Computation 2009, 5, 2544–2564. [PubMed: 20419084]

(34). Eastman P; Swails J; Chodera JD; McGibbon RT; Zhao Y; Beauchamp KA; Wang LP; Simmonett AC; Harrigan MP; Stern CD; Wiewiora RP; Brooks BR; Pande VS OpenMM 7: Rapid Development of High Performance Algorithms for Molecular Dynamics. PLOS Computational Biology 2017, 13, e1005659. [PubMed: 28746339]

(35). Hawkins GD; Cramer CJ; Truhlar DG Pairwise Solute Descreening of Solute Charges from a Dielectric Medium. Chemical Physics Letters 1995, 246, 122–129.

(36). Schaefer M; Bartels C; Karplus M Solution Conformations and Thermodynamics of Structured Peptides: Molecular Dynamics Simulation with an Implicit Solvation Model. Journal of Molecular Biology 1998, 284, 835–848. [PubMed: 9826519]

(37). Weiser J; Shenkin PS; Still WC Approximate Atomic Surfaces from Linear Combinations of Pairwise Overlaps (LCPO). Journal of Computational Chemistry 1999, 20, 217–230.

(38). Huang H; Simmerling C Fast Pairwise Approximation of Solvent Accessible Surface Area for Implicit Solvent Simulations of Proteins on CPUs and GPUs. Journal of Chemical Theory and Computation 2018, 14, 5797–5814. [PubMed: 30303377]

(39). Clark Still W; Tempczyk A; Hawley RC; Hendrickson T Semianalytical Treatment of Solvation for Molecular Mechanics and Dynamics. Journal of the American Chemical Society 1990, 112, 6127–6129.

(40). Bondi A van der Waals Volumes and Radii. Journal of Physical Chemistry 1964, 68, 441–451.

(41). Velez-Vega C; Gilson MK Overcoming Dissipation in the Calculation of Standard Binding Free Energies by Ligand Extraction. Journal of Computational Chemistry 2013, 34, 2360–2371. [PubMed: 24038118]

(42). Henriksen NM; Fenley AT; Gilson MK Computational Calorimetry: High-Precision Calculation of Host–Guest Binding Thermodynamics. Journal of Chemical Theory and Computation 2015, 11, 4377–4394. [PubMed: 26523125]

(43). Mobley DL; Chodera JD; Dill KA On the use of Orientational Restraints and Symmetry Corrections in Alchemical Free Energy Calculations. Journal of Chemical Physics 2006, 125.

(44). Aldeghi M; Heifetz A; Bodkin MJ; Knapp S; Biggin PC Accurate Calculation of the Absolute Free Energy of Binding for Drug Molecules. Chemical Science 2016, 7, 207–218. [PubMed: 26798447]

(45). Aldeghi M; Heifetz A; Bodkin MJ; Knapp S; Biggin PC Predictions of Ligand Selectivity from Absolute Binding Free Energy Calculations. Journal of the American Chemical Society 2017, 139, 946–957. [PubMed: 28009512]

(46). Pham TT; Shirts MR Identifying Low Variance Pathways for Free Energy Calculations of Molecular Transformations in Solution Phase. Journal of Chemical Physics 2011, 135.

(47). Rizzi A Improving Efficiency and Scalability of Free Energy Calculations through Automatic Protocol Optimization. Ph.D. thesis, Weill Medical College of Cornell University, United States – New York, 2020.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(48). Shirts MR; Chodera JD Statistically Optimal Analysis of Samples From Multiple Equilibrium States. Journal of Chemical Physics 2008, 129, 124105. [PubMed: 19045004]

(49). Rekharsky MV; Mayhew MP; Goldberg RN; Ross PD; Yamashoji Y; Inoue Y Thermodynamic and Nuclear Magnetic Resonance Study of the Reactions of α- and β-Cyclodextrin with Acids, Aliphatic Amines, and Cyclic Alcohols. Journal of Physical Chemistry B 1997, 101, 87–100.

(50). Kellett K; Kantonen SA; Duggan BM; Gilson MK Toward Expanded Diversity of Host–Guest Interactions via Synthesis and Characterization of Cyclodextrin Derivatives. Journal of Solution Chemistry 2018, 47, 1597–1608.

(51). Muddana HS; Varnado CD; Bielawski CW; Urbach AR; Isaacs L; Geballe MT; Gilson MK Blind Prediction of Host–Guest Binding Affinities: A New SAMPL3 Challenge. Journal of Computer-Aided Molecular Design 2012, 26, 475–487. [PubMed: 22366955]

(52). Moghaddam S; Yang C; Rekharsky M; Ko YH; Kim K; Inoue Y; Gilson MK New Ultrahigh Affinity Host–Guest Complexes of Cucurbit[7]uril with Bicyclo[2.2.2]octane and Adamantane Guests: Thermodynamic Analysis and Evaluation of M2 Affinity Calculations. Journal of the American Chemical Society 2011, 133, 3570–3581. [PubMed: 21341773]

(53). Kim HJ; Jeon WS; Ko YH; Kim K Inclusion of Methylviologen in Cucurbit[7]uril. Proceedings of the National Academy of Sciences of the United States of America 2002, 99, 5007–5011. [PubMed: 11917115]

(54). Gibb CL; Gibb BC Well-Defined, Organic Nanoenvironments in Water: The Hydrophobic Effect Drives a Capsular Assembly. Journal of the American Chemical Society 2004, 126, 11408–11409. [PubMed: 15366865]

(55). Gan H; Benjamin CJ; Gibb BC Nonmonotonic Assembly of a Deep-Cavity Cavitand. Journal of the American Chemical Society 2011, 133, 4770–4773. [PubMed: 21401093]

(56). Böttcher J; Dilworth D; Reiser U; Neumüller RA; Schleicher M; Petronczki M; Zeeb M; Mischerikow N; Allali-Hassani A; Szewczyk MM; Li F; Kennedy S; Vedadi M; Barsyte-Lovejoy D; Brown PJ; Huber KV; Rogers CM; Wells CI; Fedorov O; Rumpel K; Zoephel A; Mayer M; Wunberg T; Böse D; Zahn S; Arnhof H; Berger H; Reiser C; Hörmann A; Krammer T; Corcokovic M; Sharps B; Winkler S; Häring D; Cockcroft XL; Fuchs JE; Müllauer B; Weiss-Puxbaum A; Gerstberger T; Boehmelt G; Vakoc CR; Arrowsmith CH; Pearson M; McConnell DB Fragment-Based Discovery of a Chemical Probe for the PWWP1 Domain of NSD3. Nature Chemical Biology 2019, 15, 822–829. [PubMed: 31285596]

(57). Murray CW; Carr MG; Callaghan O; Chessari G; Congreve M; Cowan S; Coyle JE; Downham R; Figueroa E; Frederickson M; Graham B; McMenamin R; O'Brien MA; Patel S; Phillips TR; Williams G; Woodhead AJ; Woolford AJ Fragment-Based Drug Discovery Applied to Hsp90. Discovery of Two Lead Series with High Ligand Efficiency. Journal of Medicinal Chemistry 2010, 53, 5942–5955. [PubMed: 20718493]

(58). Grädler U; Schwarz D; Blaesse M; Leuthner B; Johnson TL; Bernard F; Jiang X; Marx A; Gilardone M; Lemoine H; Roche D; Jorand-Lebrun C Discovery of Novel Cyclophilin D Inhibitors Starting from Three Dimensional Fragments with Millimolar Potencies. Bioorganic and Medicinal Chemistry Letters 2019, 29, 126717. [PubMed: 31635932]

(59). Friberg A; Vigil D; Zhao B; Daniels RN; Burke JP; Garcia-Barrantes PM; Camper D; Chauder BA; Lee T; Olejniczak ET; Fesik SW Discovery of Potent Myeloid Cell Leukemia 1 (Mcl-1) Inhibitors Using Fragment-Based Methods and Structure-Based Design. Journal of Medicinal Chemistry 2013, 56, 15–30. [PubMed: 23244564]

(60). Duarte Ramos Matos G; Kyu DY; Loeffler HH; Chodera JD; Shirts MR; Mobley DL Approaches for Calculating Solvation Free Energies and Enthalpies Demonstrated with an Update of the FreeSolv Database. Journal of Chemical and Engineering Data 2017, 62, 1559–1569. [PubMed: 29056756]

(61). Qiu Y; Smith DGA; Boothroyd S; Jang H; Wagner J; Bannan CC; Gokey T; Lim VT; Stern CD; Rizzi A; Lucas X; Tjanaka B; Shirts MR; Gilson MK; Chodera JD; Bayly CI; Mobley DL; Wang LP Development and Benchmarking of Open Force Field 2.0.0: The Sage Small Molecule Force Field. Journal of Chemical Theory and Computation 2023, 19, 3251–3275. [PubMed: 37167319]

(62). Jakalian A; Bush BL; Jack DB; Bayly CI Fast, Efficient Generation of High-Quality Atomic Charges. AM1-BCC Model: I. Method. Journal of Computational Chemistry 2000, 21, 132–146.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(63). Jakalian A; Jack DB; Bayly CI Fast, Efficient Generation of High-Quality Atomic Charges. AM1-BCC Model: II. Parameterization and Validation. Journal of Computational Chemistry 2002, 23, 1623–1641. [PubMed: 12395429]

(64). Hornak V; Abel R; Okur A; Strockbine B; Roitberg A; Simmerling C Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. Proteins: Structure, Function and Bioinformatics 2006, 65, 712–725.

(65). Lindorff-Larsen K; Piana S; Palmo K; Maragakis P; Klepeis JL; Dror RO; Shaw DE Improved Side-Chain Torsion Potentials for the Amber ff99SB Protein Force Field. Proteins: Structure, Function and Bioinformatics 2010, 78, 1950–1958.

(66). Maier JA; Martinez C; Kasavajhala K; Wickstrom L; Hauser KE; Simmerling C ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. Journal of Chemical Theory and Computation 2015, 11, 3696–3713. [PubMed: 26574453]

(67). Case DA; Aktulga HM; Belfon K; Ben-Shalom IY; Berryman JT; Brozell SR; Cerutti DS; Cheatham TE; Cisneros GA; Cruzeiro VWD; Darden TA; Duke RE; Giambasu G; Gilson MK; Gohlke H; Goetz AW; Harris R; Izadi S; Izmailov SA; Kasavajhala K; Kaymak MC; King E; Kovalenko A; Kurtzman T; Lee TS; LeGrand S; Li P; Lin C; Liu J; Luchko T; Luo R; Machado M; Man V; Manathunga M; Merz KM; Miao Y; Mikhailovskii O; Monard G; Nguyen H; O'Hearn KA; Onufriev A; Pan F; Pantano S; Qi R; Rahnamoun A; Roe DR; Roitberg A; Sagui C; Schott-Verdugo S; Shajan A; Shen J; Simmerling CL; Skrynnikov NR; Smith J; Swails J; Walker RC; Wang J; Wang J; Wei H; Wolf RM; Wu Y; Xiong Y; Xue Y; York DM; Zhao S; Kollman P Amber 2022, University of California, San Francisco. 2022.

(68). Baumann HM; Gapsys V; De Groot BL; Mobley DL Challenges Encountered Applying Equilibrium and Nonequilibrium Binding Free Energy Calculations. Journal of Physical Chemistry B 2021, 125, 4241–4261. [PubMed: 33905257]

(69). Zhou HX; Gilson MK Theory of Free Energy and Entropy in Noncovalent Binding. Chemical Reviews 2009, 109, 4092–4107. [PubMed: 19588959]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

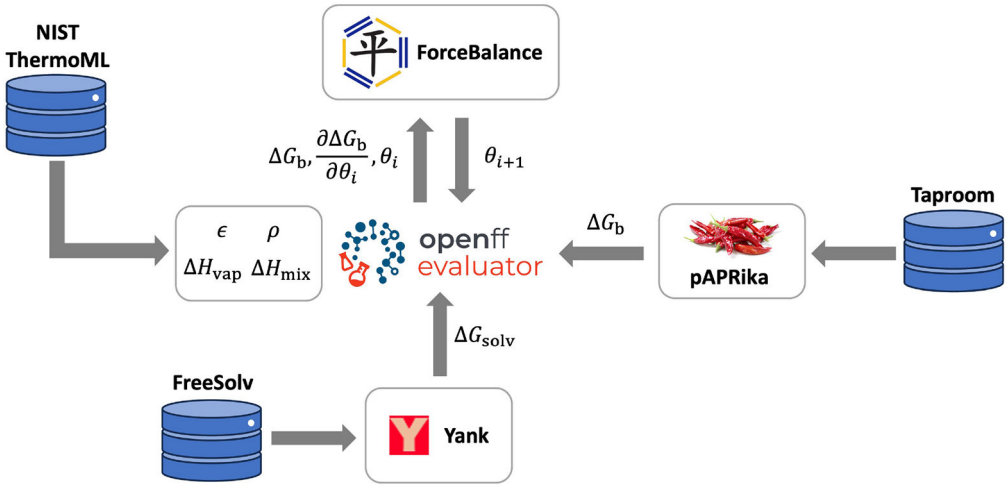

Figure 1:

Workflow of the ForceBalance-Evaluator framework for iterative optimization of force field parameters, θ, against physical properties. The baseline OpenFF-Evaluator can estimate and optimize force fields to experimental liquid-state properties and HFEs. ForceBalance can also interface with quantum mechanical software for the inclusion of quantum chemical reference data in the optimization (omitted from the diagram for simplicity). We have extended OpenFF-Evaluator to include the host–guest binding affinities, ΔG_b, part of the workflow. The host–guest system definitions and experimental binding data are stored in *Taproom*, and our pAPRika binding free energy tool computes ΔG_b and the gradient with respect to a set of force field parameters θ_i (i.e. $\frac{\partial \Delta G_b}{\partial \theta_i}$). OpenFF-Evaluator sends these quantities to ForceBalance, which returns updated parameters θ_{i+1} for a new iteration. In the diagram above, we only show OpenFF-Evaluator passing ΔG_b and $\frac{\partial \Delta G_b}{\partial \theta_i}$ to ForceBalance, though this is generalizable to the other properties as well.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

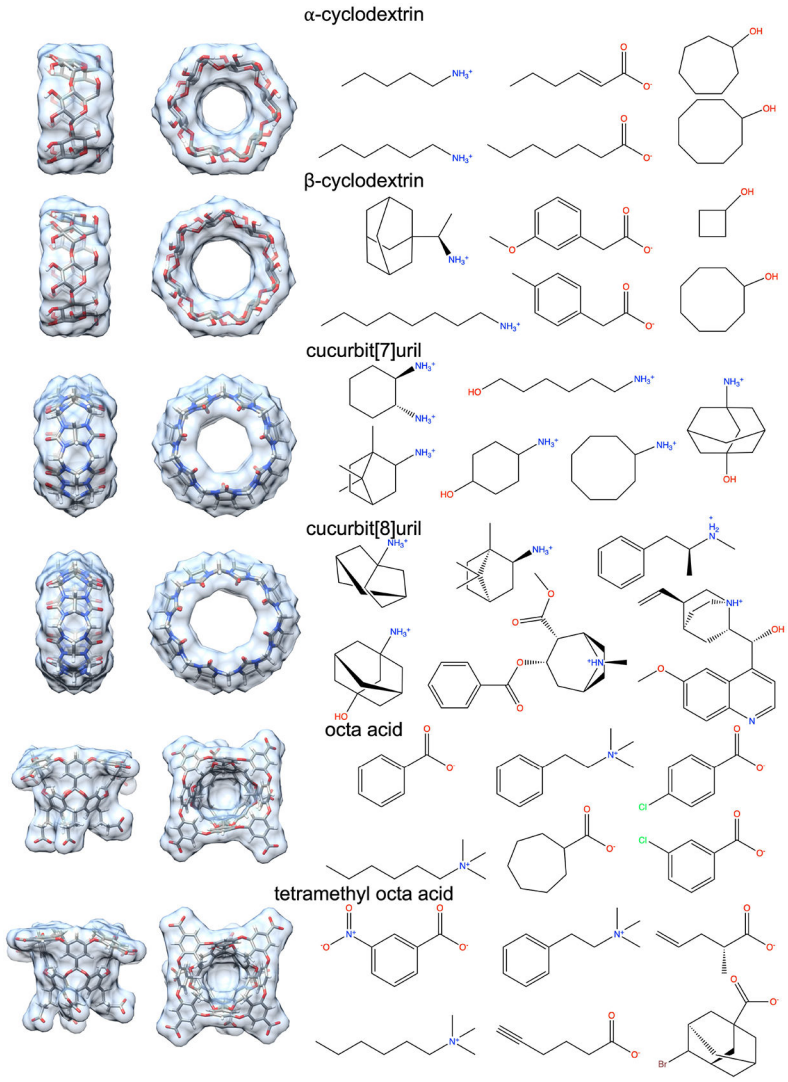

Figure 2:
Hosts and guests used in the training set. The *Taproom* names and experimental values are summarized in Table S1.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

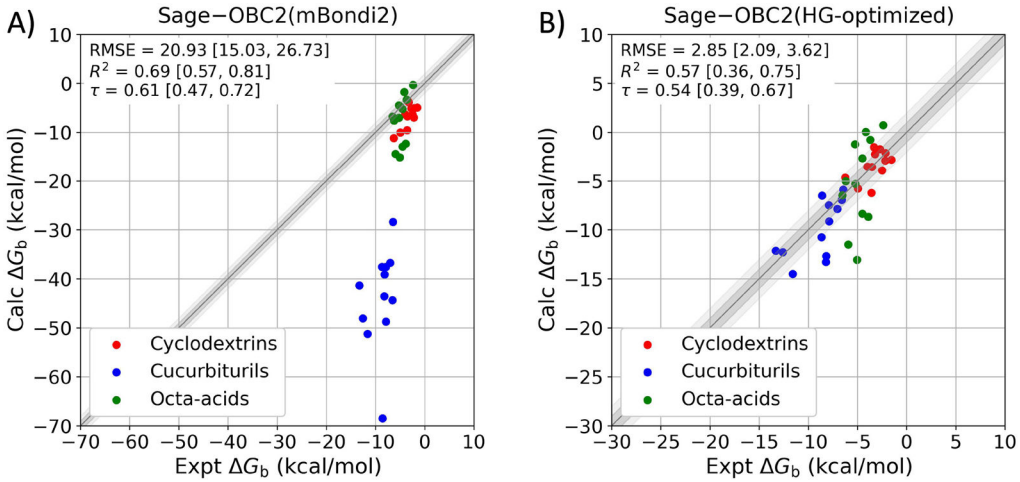

Figure 3:

Calculated versus experimental absolute binding free energies of the host–guest training set. A: Calculations with the original mBondi2 radii. B: Calculations with the HG-optimized radii. The dark and light gray shaded areas represent 1 kcal/mol and 2 kcal/mol deviations from the unity line, respectively. Values in square brackets are 95% confidence intervals from bootstrapping over the whole data set. The full data and error statistics are available in Tables S1 and S4.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

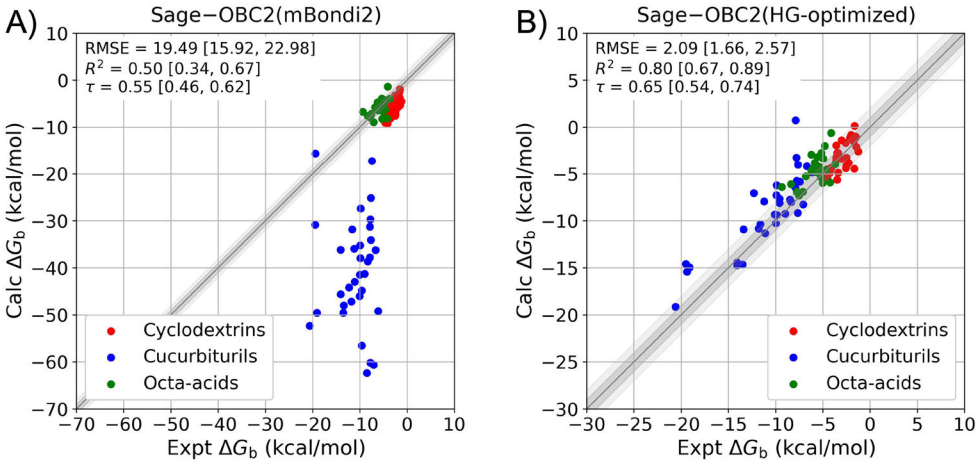

Figure 4:

Computed versus experimental binding free energies of the host–guest test set. A: Calculations with the original mBondi2 radii. B: Calculations with the HG-optimized radii. The dark and light gray shaded areas represent 1 kcal/mol and 2 kcal/mol deviations from the unity line, respectively. Values in square brackets are 95% confidence intervals from bootstrapping over the whole data set. The full data and error statistics are available in Tables S3 and S5.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

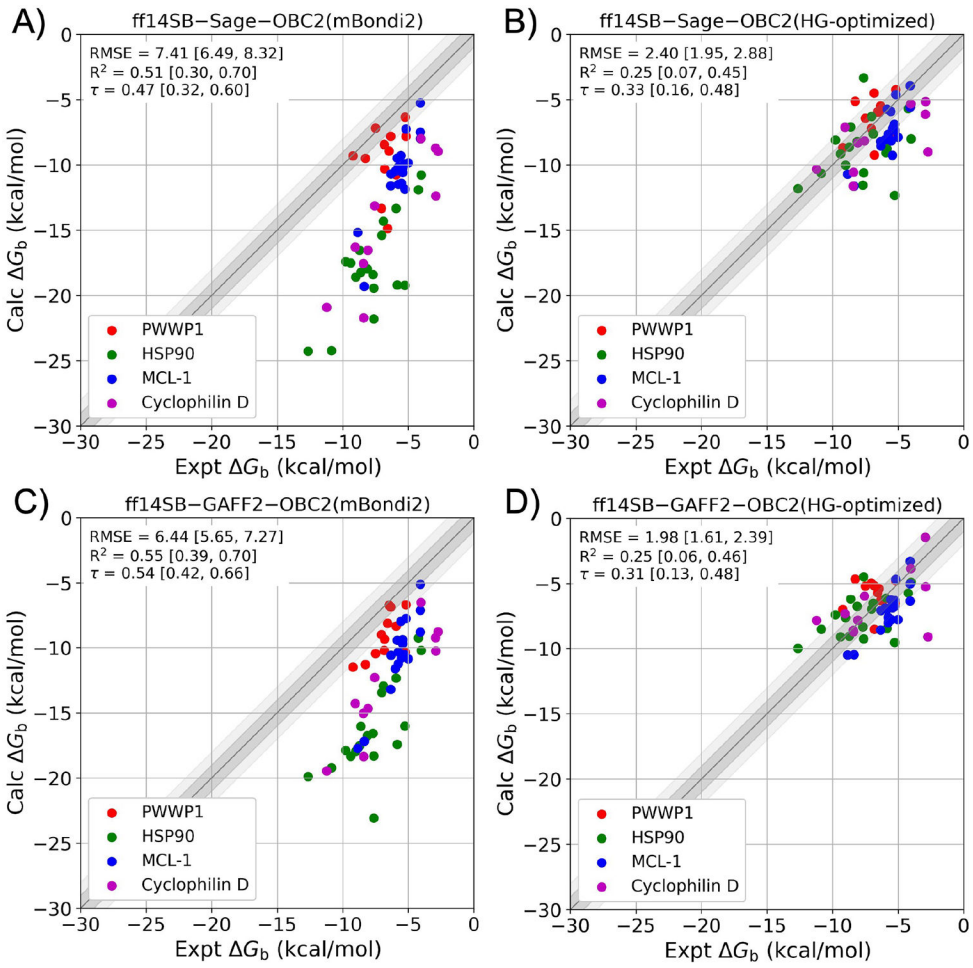

Figure 5:

Protein–ligand absolute binding free energy benchmark with the ff14SB protein force field. Top row: Sage small molecule force field with the (A) mBondi2 and (B) HG-optimized radii set. Bottom row: GAFF2 small molecule force field with the (C) mBondi2 and (D) HG-optimized radii set. The statistics on the top left of each graph are for all four proteins. The dark and light gray shaded areas represent 1 kcal/mol and 2 kcal/mol deviations from the unity line, respectively. Values in square brackets give the 95% confidence intervals from bootstrapping over the whole data set. The full data are available in Tables S6 and S7, and the error statistics are summarized in Table S10.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

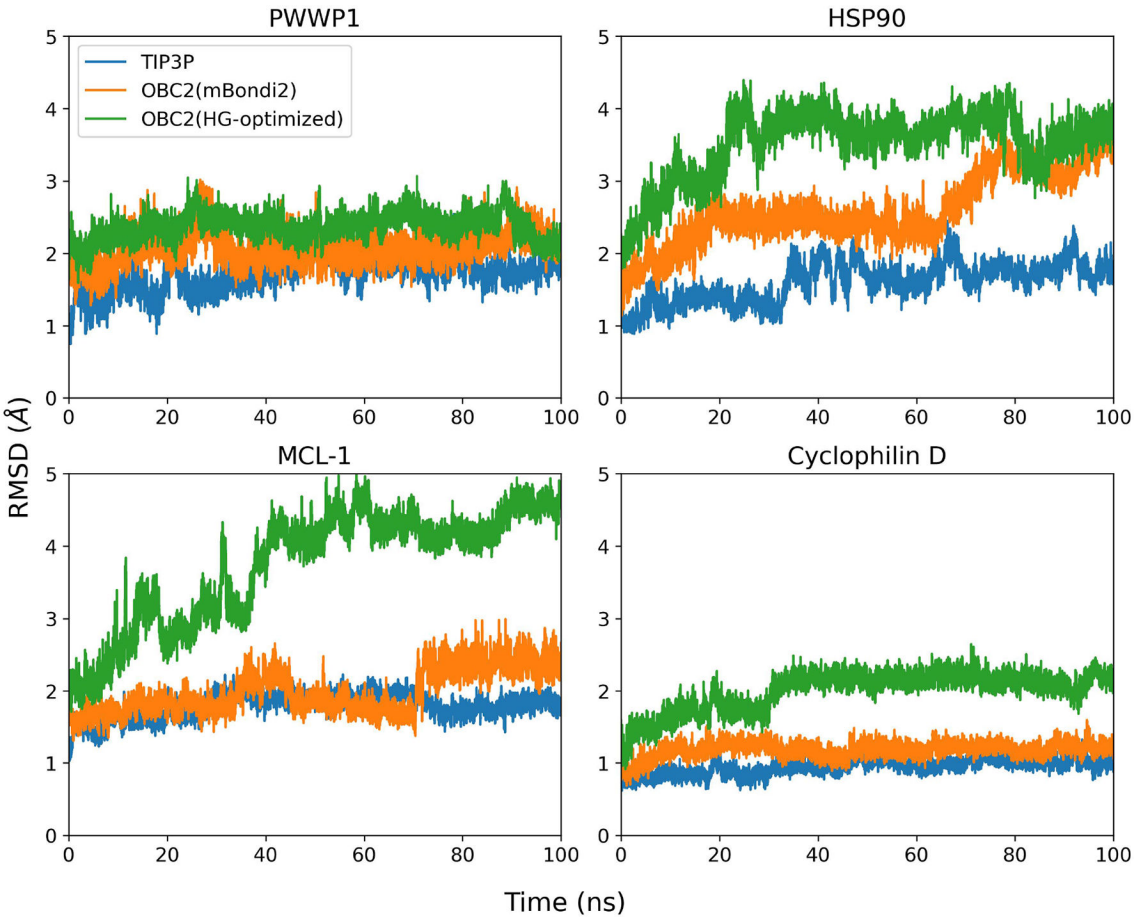

Figure 6:
RMSD of the four proteins (PWWP1, HSP90, MCL-1, and Cyclophilin D) over 100 ns of unrestrained MD simulation without any ligand bound. The proteins are simulated in TIP3P (blue), and OBC2 implicit solvent with the mBondi2 (orange) and HG-optimized (green) radii sets.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

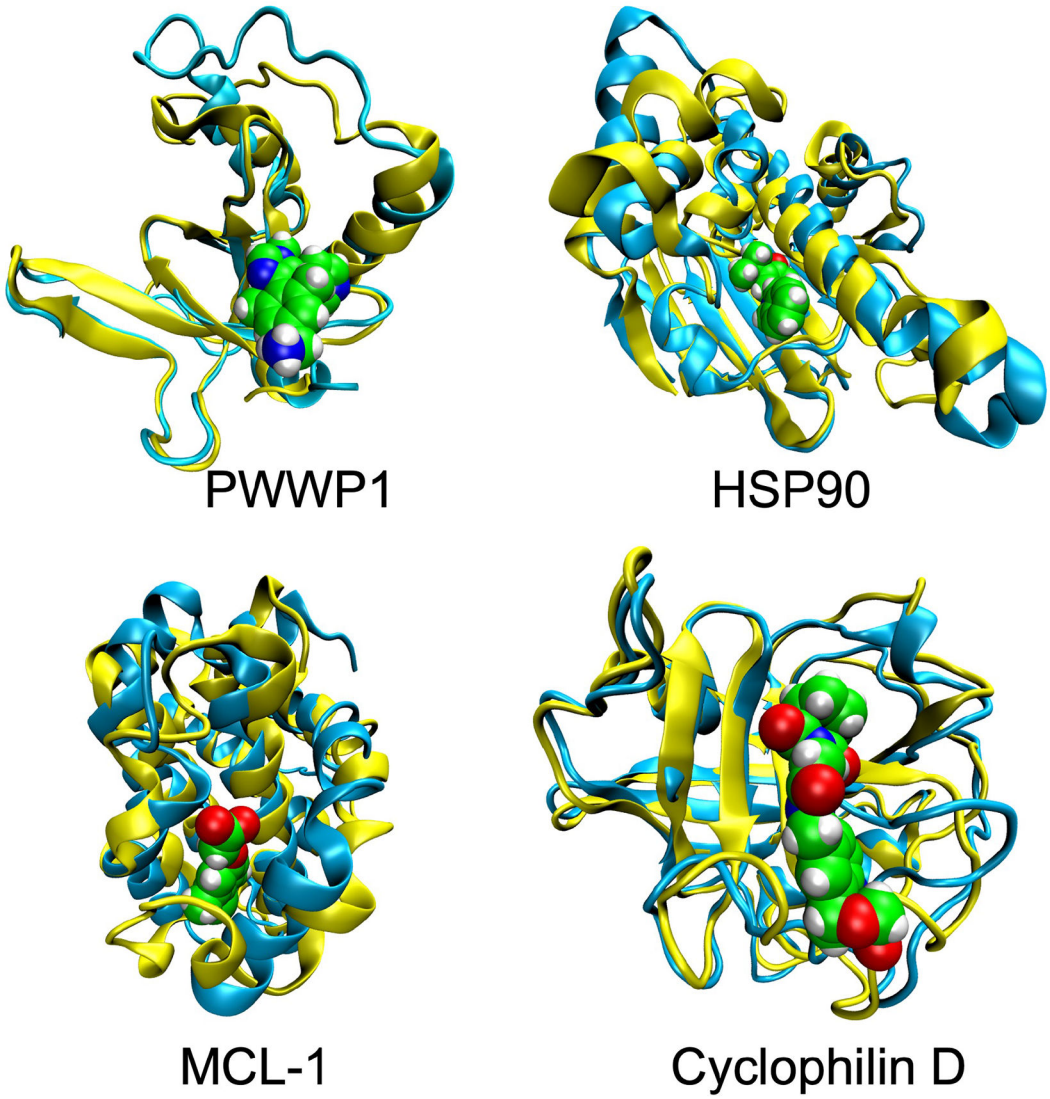

Figure 7:
Sample structures of each of the four proteins from our simulations with TIP3P water (cyan) and OBC2 implicit solvent with our HG-optimized radii (yellow). Each structure was aligned to the respective protein's initial structure from Alibay et al.[6] These 100 ns simulations were run without bound ligands, and the sample structure in each case is that last frame. We added the ligand molecule from the reference Alibay conformation as a reference to show the binding pocket.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

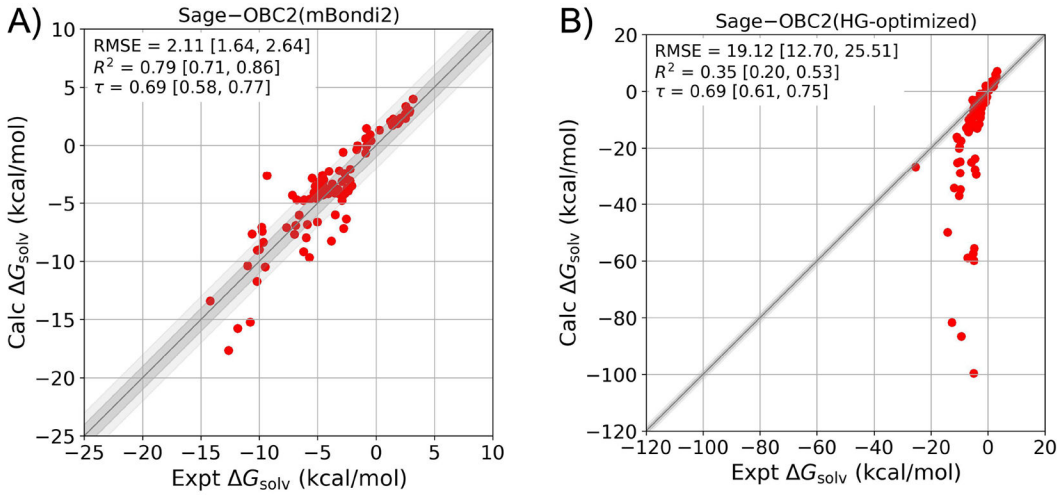

Figure 8:

Small molecule hydration free energy benchmark with the Sage force field. A: Calculations with the original mBondi2 radii. B: Calculations with the HG-optimized radii. The dark and light gray shaded areas represent 1 kcal/mol and 2 kcal/mol deviations from the unity line, respectively. Values in square brackets give the 95% confidence intervals from bootstrapping over the whole data set. The full data and error statistics are summarized in Tables S12 and S13.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1:

GB cavity radii before (mBondi2) and after (HG-optimized) optimization against host–guest binding free energies.

Atom Type	SMIRKS	Description	mBondi2 (Å)	HG-optimized (Å)
H	[#1:1]	hydrogen	1.20	0.707
H-N	[#1:1]~[#7]	hydrogen bound to nitrogen	1.30	1.218
C	[#6:1]	carbon	1.70	1.663
N	[#7:1]	nitrogen	1.55	0.533
O	[#8:1]	oxygen	1.50	1.632