Manuscript Title

Authors

- **John Doe**
  ⓘ [XXXX-XXXX-XXXX-XXXX](#) · ⓖ [johndoe](#) · 🐦 [johndoe](#)
  Department of Something, University of Whatever · Funded by Grant XXXXXXXX

- **Jane Roe**
  ⓘ [XXXX-XXXX-XXXX-XXXX](#) · ⓖ [janeroe](#)
  Department of Something, University of Whatever; Department of Whatever, University of Something

Abstract

Designing ligands that bind their target with high affinity and specificity is a key step in small-molecule drug discovery. Yet accurate predictions of protein-ligand binding free energies are difficult and errors in the calculations can be traced to challenges adequately sampling conformational space, ambiguous protonation states, and other causes. Noncovalent complexes between a cavity-containing host molecule and drug-like guest molecules have emerged as powerful tools for modeling protein-ligand binding. Due to their small size and extensive experimental characterization, calculations of host-guest binding free energies, enthalpies, and entropies offer an opportunity to directly probe, and ultimately optimize, force fields.

The Open Force Field Initiative aims to create a modern, open software infrastructure for automatically generating and validating force fields using high-quality data sets. The first force field to arise out of this effort, named SMIRNOFF99Frosst, has one tenth the number of parameters of a typical general small molecule force field, such as GAFF, yet predicts binding thermodynamics that are on average, at least as accurate. Here, we report the results of free energy calculations on 43 α and β-cyclodextrin host-guest pairs for which experimental data are available. Our calculations were performed using the attach-pull-release method as implemented in the open source package, `pAPRika`. On binding free energies, the root mean square error of the predictions relative to experiment is 0.82 kcal/mol for SMIRNOFF99Frosst and 1.58 kcal/mol for GAFF version 2.1. These results suggest significant room for improvement in force fields, and will help create a transparent and robust method of evaluating future candidate parameter sets from the Open Force Field Group. Improving the performance of force fields for predicting binding affinities will help reduce the timescale and cost required to generate drug candidates.

Introduction

Methods

Chose of host-guest system

In this study, we report the binding thermodynamics of 43 host-guest complexes (Figure [1](#)) computed using three different force fields. The complexes consist of either α- or β-cyclodextrin as host molecule and a series of ammonium, carboxylate, or cyclic alcohol small molecule guests. Cyclodextrins are cyclic polymers consisting of six (αCD) or seven (βCD) glucose monomers in the shape of a truncated cone. The equilibrium constants and standard molar enthalpies of binding for these 43 complexes have been measured using isothermal titration calorimetry [1] and computationally in [2]. As in [2], only a single structural isomer was considered for the 1-methylammonium guests.
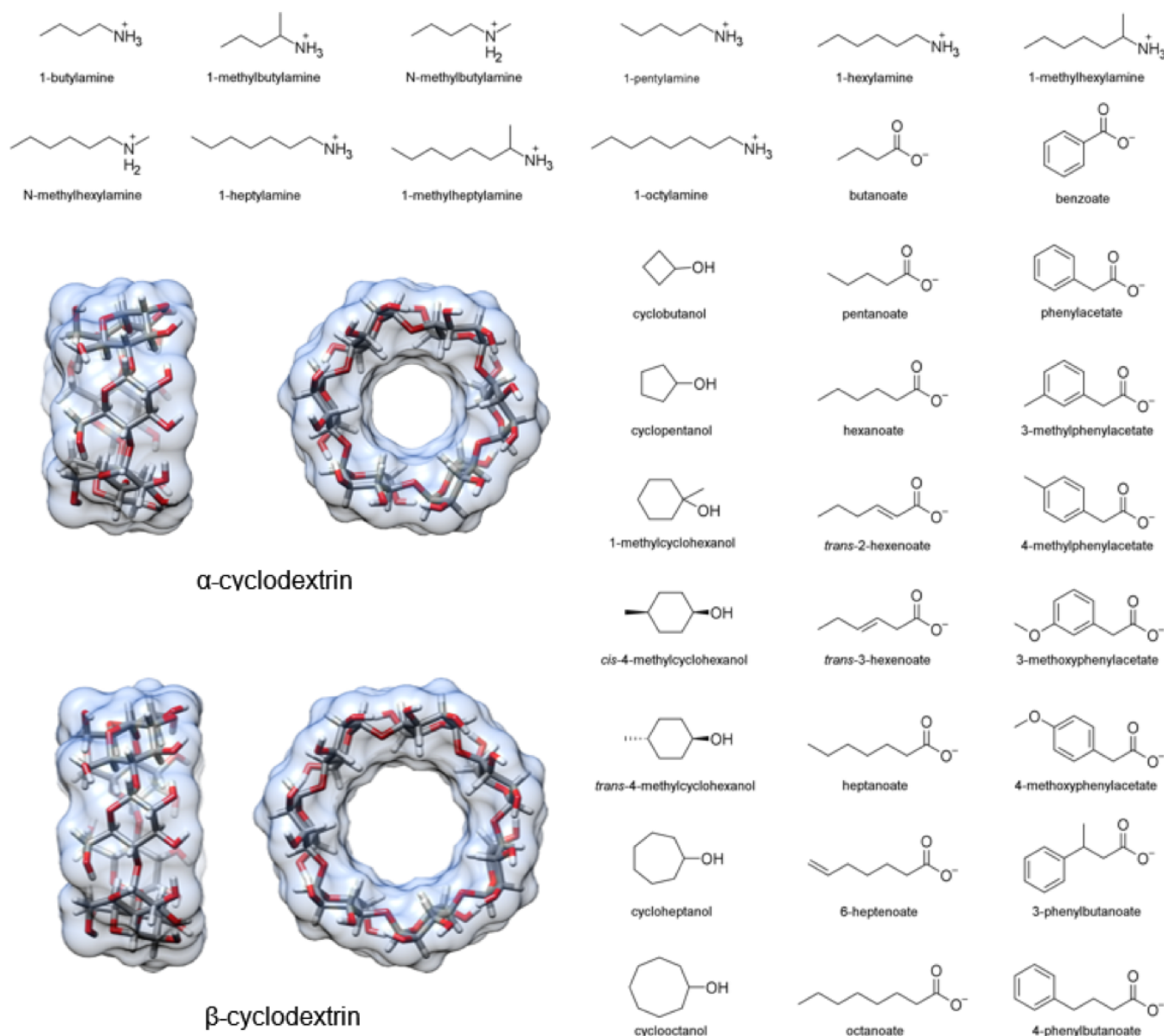
Figure 1: Structures of the two cyclodextrin hosts and 33 guest molecules in this study which together comprise 43 unique host-guest pairs.

Application of force field parameters

We sought to compare force fields directly, and as such, attempted to minimize additional differences between the simulation sets. In all simulations, we applied AM1-BCC [3,4] partial atomic charges to both the host and guest molecules using the `antechamber` program. The host charges were calculated using a single glucose molecule with methoxy caps on the O1 and O4 alcohols (Figure 2); each glucose monomer in the cyclodextrin polymer has identical charges. We used TIP3P water [5] and Joung-Cheatham monovalent ion parameters for Na^+ and Cl^- [6] in each simulation set.
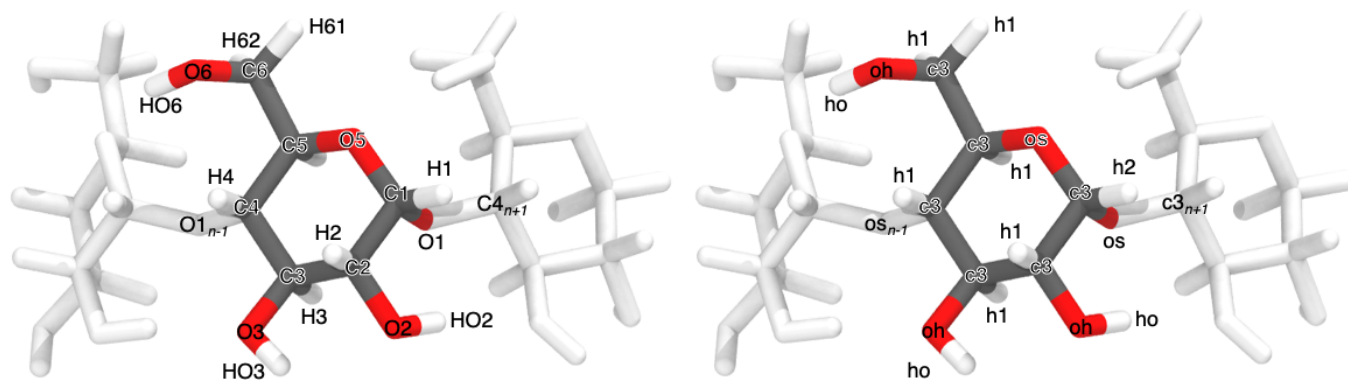
Figure 2: Atom names (left) and GAFF atom types (right) for a glucose monomer in αCD shown with two flanking monomers. The remaining three glucose monomers are hidden for clarity..

GAFF v1.7 bond, angle, torsion, and Lennard-Jones parameters were applied using the `tleap` program distributed with AmberTools16. These simulations were performed as part of Henriksen, et al. [2] and are described in additional detail therein.

GAFF v2.1 parameters were applied in an identical manner to the GAFF v1.7 parameters, using the `tleap` program distributed with AmberTools18 and substituting `leaprc.gaff` for `leaprc.gaff2` in the `tleap` input file. In GAFF v2.1, the bond and angle parameters have been updated to reproduce small molecule geometries obtained from high-level quantum mechanical calculations. The force constants for the bond and angle parameters were tuned to reproduce the vibrational spectra of over 600 molecules. The torsion parameters were optimized to reproduce the rotational potential energy surface of 400 model compounds. Finally, the Lennard-Jones coefficients were redeveloped to reproduce interaction energies and pure liquid properties.

> We might want to cite `gaff2.dat` somehow, because this is where I obtained the above information; we could also cite personal communications with Junmei Wang based on our conversations at the Amber developers' meeting.

To apply SMIRNOFF99Frosst parameters, we followed a multistep process, beginning with the prepared GAFF v1.7 files. The host and guest molecules were parameterized with the Open Force Field Toolkit version 0.0.3, SMIRNOFF version 1.0, and SMIRNOFF99Frosst version 1.0.5. Once parameterized with SMIRNOFF99Frosst, ParmEd [7] was used to combine the host and guest with the solvent and ions, which retained their parameters as TIP3P water and Joung-Cheatham ion parameters, respectively.

> I decided to skip a detailed listing of what I did, and instead link to the Jupyter notebook I used, because these instructions may change with future toolkit versions, and in fact, it is already out of date in some ways. If this is a good way to cite this, I can upload that GitHub repository to Zenodo, get a DOI, and cite the DOI. It may also be feasible to include this as SI, but that can get messy for anything other than trivial code listings.

Thermodynamic calculation

We used the attach-pull-release (APR) method as implemented in the open source package pAPRika version 0.0.3, to calculate absolute binding free energies. A complete description of the APR method has been characterized in the literature [10,11,8,9]. The attachment and release phases consisted of 15 independent windows; the pulling phase consisted of 45 independent windows. During the attachment phase, the force constants on the host and guest are scaled by a λ parameter that goes from $\lambda = 0$, when

all restraints are turned off, to $\lambda = 1$, when all restraints reach their maximum force constant. The λ windows are more densely spaced where the force constant is smaller to improve sampling along highly curved regions of the potential of mean force. Conformational restraints were applied between neighboring glucose units of the cyclodextrin to limit the incursion of monomers into the host cavity. These restraints were applied along the pseudodihedrals $O5_n - C1_n - O1_n - C4_{n+1}$ and $C1_n - O1_n - C4_{n+1} - C5_{n+1}$ to improve convergence and sampling of the bound state (see Figure [2](#) for atom naming conventions). To further improve convergence of the bound state, we applied a hard wall restraint that confined the guest molecule to within a sphere of 12.3 and 13.5 Å of αCD and βCD, respectively.

During the pulling phase, the λ parameter represented a distance restraint and is increased uniformly in 45 increments of 0.4 Å, yielding windows that separate the host and guest by 18 Å. An explicit release calculation is performed to turn off the conformational restraints on the host in the absence of guest molecules. This is performed once for each cyclodextrin. An analytic release calculation is performed to compute the work of moving the guest from the restricted volume enforced by the APR restraints to standard state at 1 M concentration.

> Needs editing

Due to the lack of symmetry of cyclodextrin and the small molecule guests, there are generally two distinct binding poses. We separately compute the binding free energy and enthalpy for each orientation [8] and combine the results to produce a single value for each host-guest combination.

% For simple guest molecules, with one polar functional group, there are generally two possible binding modes in the CD cavity: one with the polar group oriented out of the CD opening with primary alcohols, and one with the polar group oriented out of the wider opening with secondary alcohols (Figure 2). However, experimental data report binding free energy and enthalpy values based on a Boltzmann-weighted ensemble of these two orientations. In order to compare with experiment, we separately compute and appropriately combine the binding free energy and enthalpy for each orientation, as previously described.40

Thermodynamic integration was used to compute the binding free energy. The binding enthalpy was computed as the difference in mean potential energy of the bound state (in the absence of any restraints) and the unbound state (where the guest is held far away from the host, but the conformational restraints on the host are disabled).

% The binding enthalpy was computed by subtracting the average potential energy of the last window of the release phase, in which the guest has been pulled far away from the host, from the mean energy of the first window of the attachment phase, in which restraint force constants are off. The uncertainties for average properties calculated in each window (e.g., restraint coordinate values and potential energies) were estimated with the blocking method42 and propagated into the final reported values using bootstrapping for the thermodynamic integration calculations, or addition in quadrature for the binding enthalpy calculations.

Uncertainties were computed using blocking analysis [12].

Simulations

Simulations were performed with the `pmemd.cuda` module of AMBER 16 (calculations of the GAFF v1.7 force field) and AMBER 18 (calculations of the GAFF v2.1 and SMIRNOFF99Frosst force fields) molecular dynamics software. Each window for each system was independently solvated and simulated.

Solvation consisted of 2000 TIP3P waters for the αCD systems and 2210 waters for the βCD systems in an orthorhombic box. Each simulation contained enough Na^+ or Cl^- ions to neutralize the host-guest complex and an additional 50 mM $NaCl$ to match the experimental conditions in [1]. The host and guest were oriented via non-interacting dummy atoms along the simulation box z axis to minimize the amount of solvent required. Hydrogen mass repartitioning was used to adjust the mass of hydrogen atoms to 3.024 Da by transferring mass from bonded heavy atoms, enabling a simulation timestep of 4 fs. Equilibration consisted of 50,000 steps of energy minimization, 100 ps of heating from 0 to 300 K, followed by 2000 ps of additional NPT simulations. A Langevin thermostat, the Monte Carlo barostat, a nonbonded cutoff of 9 Å and default PME parameters, were used for the NPT simulations.

Production NPT simulations were run for a minimum of 5 ns and maximum of 50 ns per window, except for the endpoint windows used to calculate the enthalpy, which were simulated for 1 μs. In the GAFF v1.7 and GAFF v2.1 simulations, the exact length of each window's simulation was determined by the restraint coordinate uncertainty, estimated via blocking analysis [12]. That is, for restraint energy U in window i, each window was simulated until the variable

$$w(\lambda_i) = \begin{cases} \left\langle \frac{\partial U}{\partial \lambda_i} \right\rangle_{\text{SEM}} \cdot \frac{\lambda_{i+1}}{2} & i = 0 \\ \left\langle \frac{\partial U}{\partial \lambda_i} \right\rangle_{\text{SEM}} \cdot \frac{\lambda_{i+1}-\lambda_{i-1}}{2} & i \in [1, N-1] \\ \left\langle \frac{\partial U}{\partial \lambda_i} \right\rangle_{\text{SEM}} \cdot \frac{1-\lambda_{i-1}}{2} & i = N \end{cases}$$

was below a threshold = 0.02 kcal/mol during the attach phase and 0.1 kcal/mol during the pull phase.

> This is ambiguous because $i = 0$ and $i = N$ during attach and release are run for longer than this.

SMIRNOFF99Frosst simulations calculations were run for 10 ns per window.

Statistics

The overall RMSE and R^2 values for each comparison are reported as the mean with the 95% confidence interval in brackets after 100,000 bootstrapping cycles. The R^2 values for each functional group subset is reported in the bottom right corner.

For each force field, we computed RMSE, MSE, R^2, Kendall τ.

> What else?

Results and discussion

This results section is organized as follows. We first present a comparison of SMIRNOFF99Frosst and two iterations of the General AMBER Force Field (GAFF [13]) on predicting binding free energies (ΔG) and binding enthalpies (ΔH) of small molecule guests to α-cyclodextrin (αCD) and β-cyclodextrin (βCD). We then detail how the behaviors of guest molecules changes between force fields and finally we summarize the parameter differences between SMIRNOFF99Frosst and GAFF along with the effects of the parameter differences.

Binding free energies and enthalpies

SMIRNOFF99Frosst does about as well as, or even better than GAFF v1.7, depsite have far fewer numerical parameters on predicted ΔG and ΔH compared to the values measured with isothermal titration calorimetry. SMIRNOFF99Frosst has an overall agreement of under 1 kcal/mol on binding free energies and under 2 kcal/mol on binding enthalpies across the 43 host-guest systems (Figure 3). Overall, the correlation between SMIRNOFF99Frosst and experiment is mediocre, ranging from an R^2 value of 0.34 to 0.54 for ΔG. The results are more varied for ΔH with R^2 ranging from 0.09 to 0.54. In particular, SMIRNOFF99Frosst underestimates the binding free energy and binding enthalpy of cyclic alcohols and mostly underestimates the binding entropy (Figure 5). Guests containing ammonium or carboxylate functional groups appear to be equally over- and under-estimated for binding free energy.

Notably, SMIRNOFF99Frosst does better on αCD than βCD (Figure **¿fig:S99-vs-experiment-by-cyclodextrin?**).
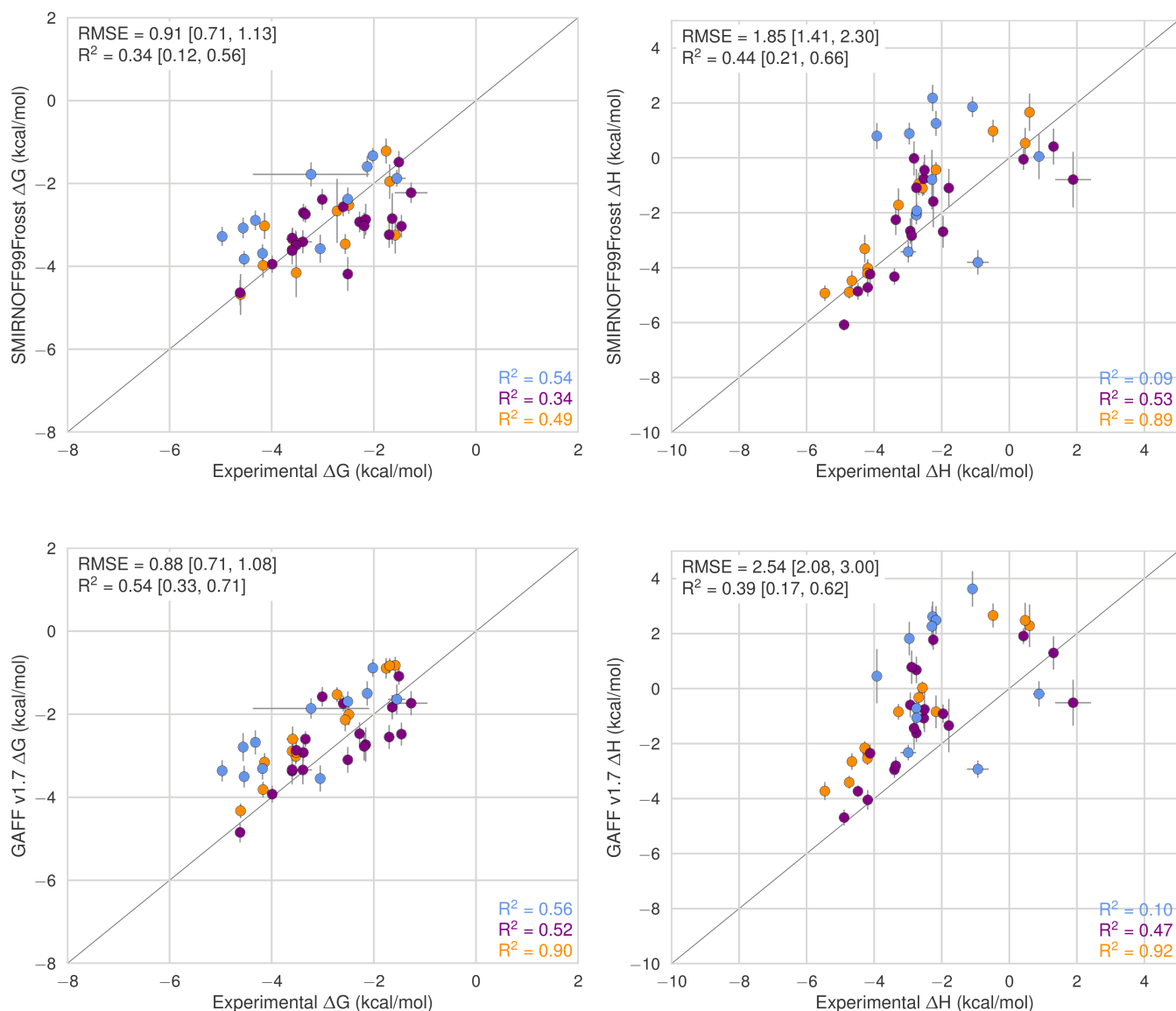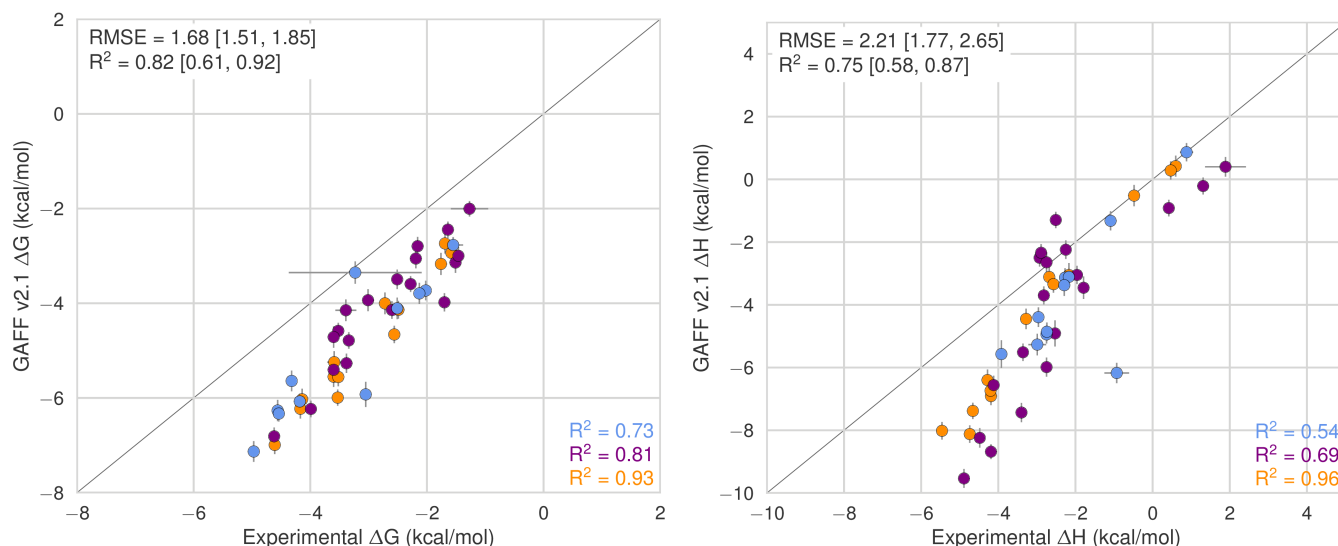
Figure 3: Comparison of calculated absolute binding free energies (ΔG) and binding enthalpies (ΔH) with experiment with SMIRNOFF99Frosst parameters (top), GAFF v1.7 parameters (middle), or GAFF v2.1 parameters (bottom) applied to both host and guest. The orange, blue, and purple coloring distinguish the functional group of the guest as an ammonium, alcohol, or carboxylate, respectively..

Likewise, GAFF v1.7 has very similar properties with an RMSE for ΔG of 0.88 kcal/mol and RMSE for ΔH of 2.54 kcal/mol. Both SMIRNOFF99Frosst and GAFF v1.7 systematically underestimate the binding free energy for cyclic alcohols except for beta-cyclodextrin with cyclopentanol (b-cpe).

Overall, GAFF v2.1 has a larger RMSE than either SMIRNOFF99Frosst or GAFF v1.7 on binding free energies, but strikingly strong correlations with the experimental values across all three functional group classes. Compared to SMIRNOFF99Frosst and GAFF v1.7, GAFF v2.1 uniformly overestimates both the binding free energy and binding enthalpy.

SMIRNOFF99Frosst does a great job predicting the binding free energy of αCD with alcohols, with four of the five points nearly falling on the line of identity.

Show highlights here

Likewise, SMIRNOFF99Frosst does a great job with amines (?) with three of the four points falling nearly on the line of identity.

With βCD, on teh other hand, things are a little different: two of the three carboxylates are overestimated, ammoniums are great, but the alcohols are all over the place (two overestimated, two understimated.)

This difference is likely because… (?) of flexibility? (Is alpha more flexible than beta?)

Guest preferences for primary and secondary binding

There are two orientations for binding: one where the primary functional group is interacting with the more narrow opening primary alcohols and one with the polar group oriented out of the wider, secondary opening.

The experimental data is a Boltzmann weighted ensemble of these two orientations.

Guest preferences for αCD and βCD

Functional group interactions

Differences in force field parameters between SMIRNOFF99Frosst and GAFF

Next, we summarize the parameter differences between SMIRNOFF99Frosst, a decendent of parm@Frosst and GAFF v1.7 (released circa March 2015 according to `gaff.dat` distributed with AMBER16) and GAFF v2.1 (under active development).

Structural consequences of the force field parameter differences

In both SMIRNOFF99Frosst and GAFF v1.7, the average RMSD of βCD is between 2 and 2.5 Å over 43 μs of simulation. GAFF v2.1 is significantly more rigid, with an average RMSD less than 1.0 Å from the initial structure (Figure 4).
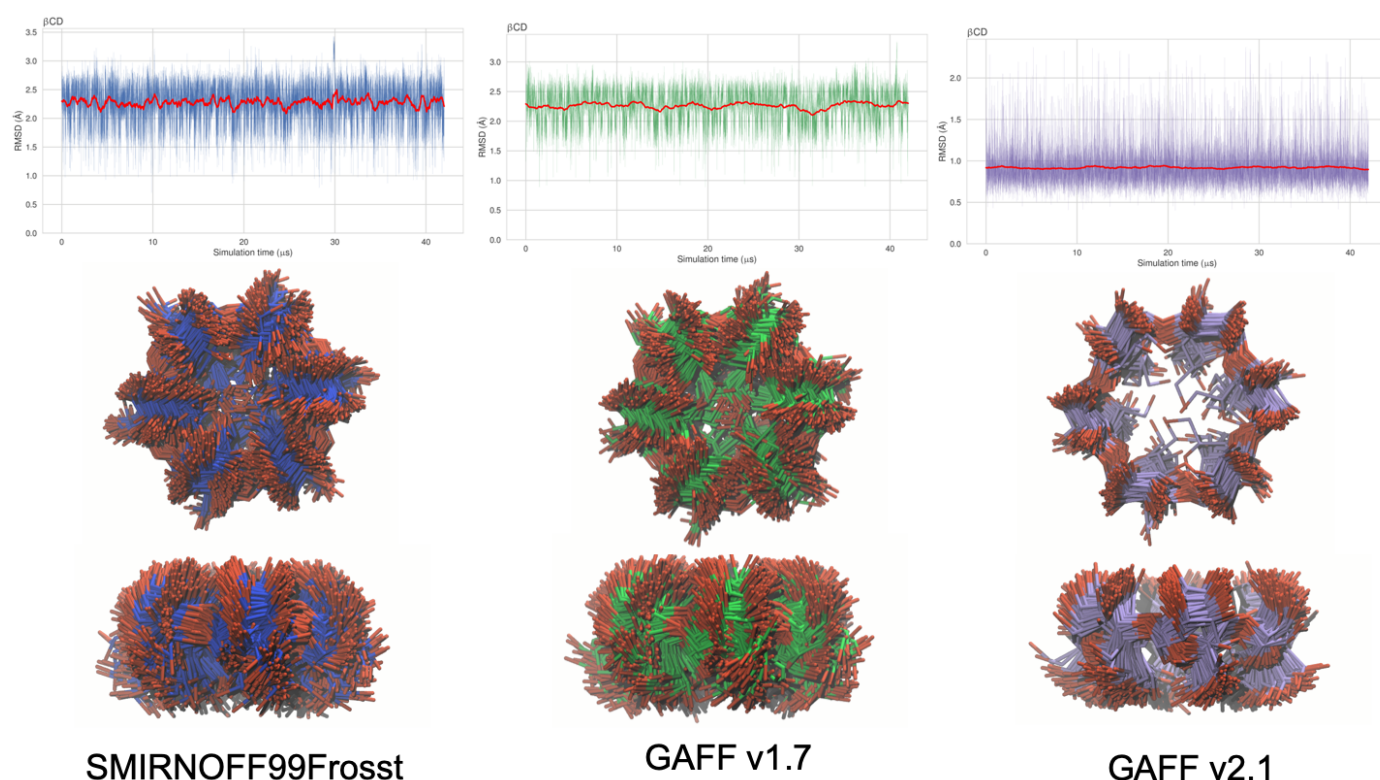
Figure 4: Top: Root mean square deviation (RMSD) of free βCD in the three force fields, all relative to the same initial structure. A 1000 frame moving average is plotted in red. Middle: to-view of the open cavity of βCD with no guest (200 snapshots over 1 μ s). Bottom: side-view of the open cavity. The carbons are colored blue in SMIRNOFF99Frosst, green in GAFF v1.7, and purple in GAFF v2.1. Hydrogen atoms have been hidden for clarity.

★ Update font size in RMSD plots.
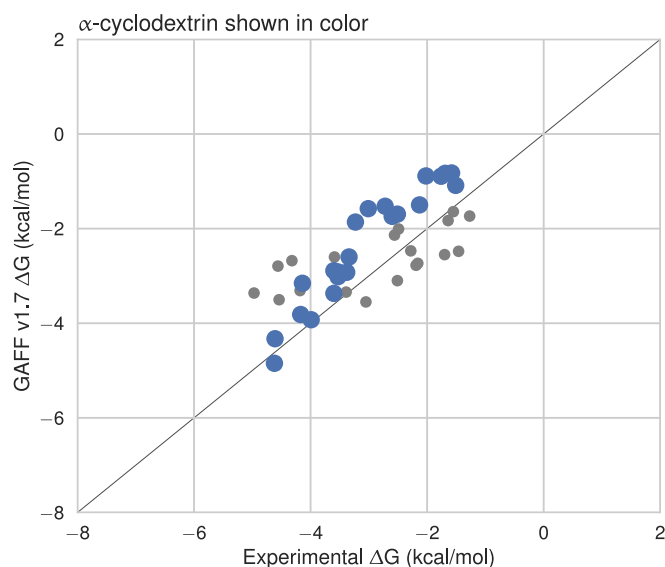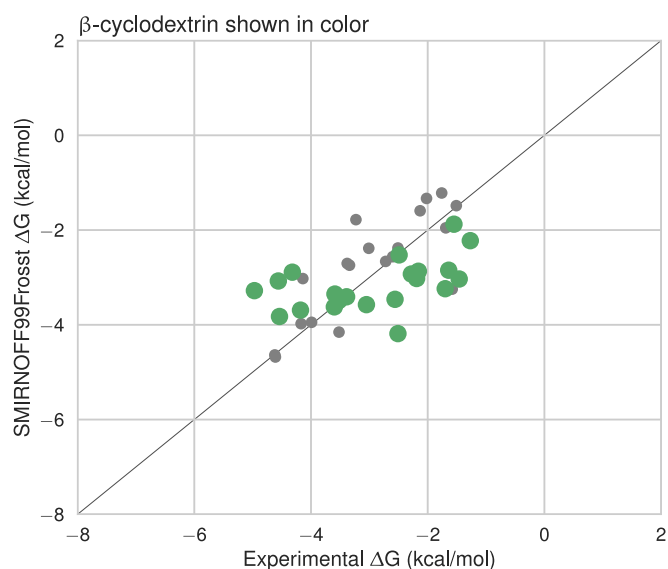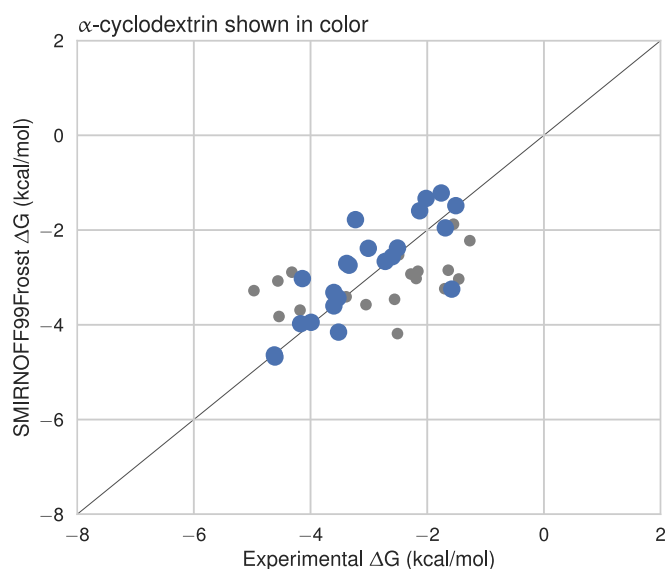
Supporting Information
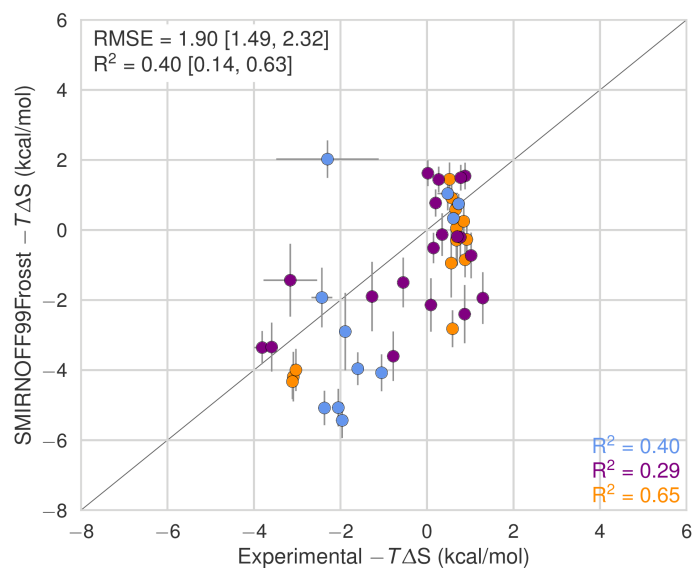
Figure 5: Comparison of calculated absolute binding enthalpy with experiment with SMIRNOFF99Frosst parameters applied to both host and guest. The orange, blue, and purple coloring distinguish the functional group of the guest as an ammonium, alcohol, or carboxylate, respectively..
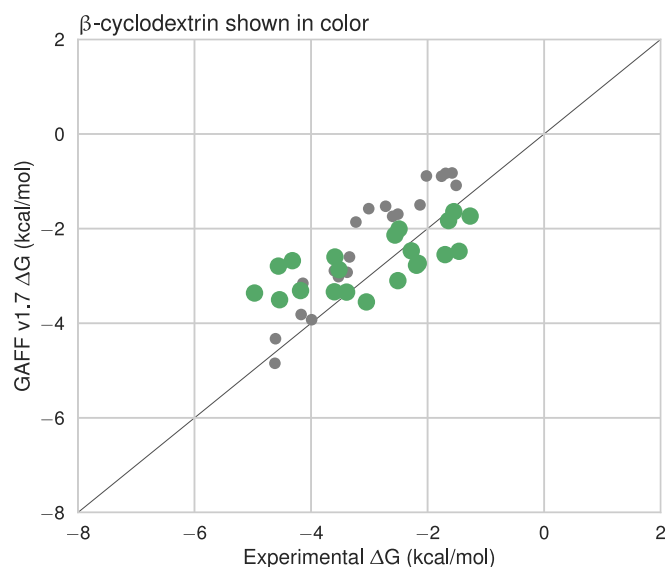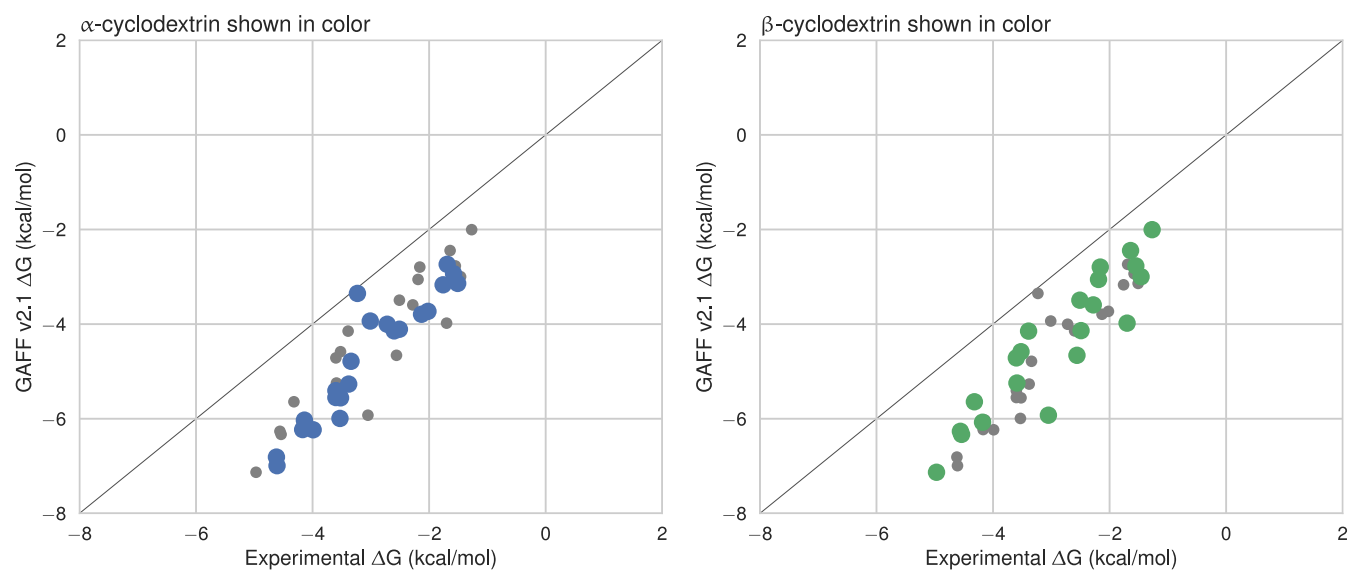
Figure 6: Binding free energies (ΔG) replotted from Figure [3], with αCD points colored in blue and βCD points in grey (left) or αCD points in grey with βCD points colored in green (right)..

References

1. **Thermodynamic and Nuclear Magnetic Resonance Study of the Reactions of α- and β-Cyclodextrin with Acids, Aliphatic Amines, and Cyclic Alcohols**
Mikhail V. Rekharsky, Martin P. Mayhew, Robert N. Goldberg, Philip D. Ross, Yuko Yamashoji, Yoshihisa Inoue
The Journal of Physical Chemistry B (1997-01) https://doi.org/cmwfwn
DOI: 10.1021/jp962715n

2. **Evaluating Force Field Performance in Thermodynamic Calculations of Cyclodextrin Host–Guest Binding: Water Models, Partial Charges, and Host Force Field Parameters**
Niel M. Henriksen, Michael K. Gilson
Journal of Chemical Theory and Computation (2017-07-11) https://doi.org/gd2z2t
DOI: 10.1021/acs.jctc.7b00359 · PMID: 28696692 · PMCID: PMC5606194

3. **Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation**
Araz Jakalian, David B. Jack, Christopher I. Bayly
Journal of Computational Chemistry (2002-10-18) https://doi.org/cktk6g
DOI: 10.1002/jcc.10128 · PMID: 12395429

4. **Fast, efficient generation of high-quality atomic charges. AM1-BCC model: I. Method**
Araz Jakalian, Bruce L. Bush, David B. Jack, Christopher I. Bayly
Journal of Computational Chemistry (2000-01-30) https://doi.org/cvvpkv
DOI: 10.1002/(sici)1096-987x(20000130)21:2<132::aid-jcc5>3.0.co;2-p

5. **Comparison of simple potential functions for simulating liquid water**
William L. Jorgensen, Jayaraman Chandrasekhar, Jeffry D. Madura, Roger W. Impey, Michael L. Klein
The Journal of Chemical Physics (1983-07-15) https://doi.org/dg9sq8
DOI: 10.1063/1.445869

6. **Determination of Alkali and Halide Monovalent Ion Parameters for Use in Explicitly Solvated Biomolecular Simulations**
In Suk Joung, Thomas E. Cheatham III
The Journal of Physical Chemistry B (2008-07) https://doi.org/cgwnj7
DOI: 10.1021/jp8001614 · PMID: 18593145 · PMCID: PMC2652252

7. **Lessons learned from comparing molecular dynamics engines on the SAMPL5 dataset**
Michael R. Shirts, Christoph Klein, Jason M. Swails, Jian Yin, Michael K. Gilson, David L. Mobley, David A. Case, Ellen D. Zhong
Journal of Computer-Aided Molecular Design (2016-10-27) https://doi.org/f9m3wn
DOI: 10.1007/s10822-016-9977-1 · PMID: 27787702 · PMCID: PMC5581938

8. **Computational Calorimetry: High-Precision Calculation of Host–Guest Binding Thermodynamics**
Niel M. Henriksen, Andrew T. Fenley, Michael K. Gilson
Journal of Chemical Theory and Computation (2015-08-26) https://doi.org/f7q3mj
DOI: 10.1021/acs.jctc.5b00405 · PMID: 26523125 · PMCID: PMC4614838

9. **Overcoming dissipation in the calculation of standard binding free energies by ligand extraction**
Camilo Velez-Vega, Michael K. Gilson
Journal of Computational Chemistry (2013-08) https://doi.org/gbdv9w
DOI: 10.1002/jcc.23398 · PMID: 24038118 · PMCID: PMC3932244

10. **The SAMPL4 host–guest blind prediction challenge: an overview**

Hari S. Muddana, Andrew T. Fenley, David L. Mobley, Michael K. Gilson

Journal of Computer-Aided Molecular Design (2014-03-06) https://doi.org/f5585s

DOI: 10.1007/s10822-014-9735-1 · PMID: 24599514 · PMCID: PMC4053502

11. **Bridging Calorimetry and Simulation through Precise Calculations of Cucurbituril–Guest Binding Enthalpies**

Andrew T. Fenley, Niel M. Henriksen, Hari S. Muddana, Michael K. Gilson

Journal of Chemical Theory and Computation (2014-08) https://doi.org/f6hv4x

DOI: 10.1021/ct5004109 · PMID: 25221445 · PMCID: PMC4159218

12. **Error estimates on averages of correlated data**

H. Flyvbjerg, H. G. Petersen

The Journal of Chemical Physics (1989-07) https://doi.org/bjpm5j

DOI: 10.1063/1.457480

13. **Development and testing of a general amber force field**

Junmei Wang, Romain M. Wolf, James W. Caldwell, Peter A. Kollman, David A. Case

Journal of Computational Chemistry (2004) https://doi.org/cdmcnb

DOI: 10.1002/jcc.20035 · PMID: 15116359