J Comput Aided Mol Des (2014) 28:401–415
DOI 10.1007/s10822-014-9716-4

Converging free energies of binding in cucurbit[7]uril and octa-acid host–guest systems from SAMPL4 using expanded ensemble simulations

Jacob I. Monroe · Michael R. Shirts

Received: 15 November 2013 / Accepted: 21 January 2014 / Published online: 8 March 2014
© Springer International Publishing Switzerland 2014

Abstract Molecular containers such as cucurbit[7]uril (CB7) and the octa-acid (OA) host are ideal simplified model test systems for optimizing and analyzing methods for computing free energies of binding intended for use with biologically relevant protein–ligand complexes. To this end, we have performed initially blind free energy calculations to determine the free energies of binding for ligands of both the CB7 and OA hosts. A subset of the selected guest molecules were those included in the SAMPL4 prediction challenge. Using expanded ensemble simulations in the dimension of coupling host–guest intermolecular interactions, we are able to show that our estimates in most cases can be demonstrated to fully converge and that the errors in our estimates are due almost entirely to the assigned force field parameters and the choice of environmental conditions used to model experiment. We confirm the convergence through the use of alternative simulation methodologies and thermodynamic pathways, analyzing sampled conformations, and directly observing changes of the free energy with respect to simulation time. Our results demonstrate the benefits of enhanced sampling of multiple local free energy minima made possible by the use of expanded ensemble molecular dynamics and may indicate the presence of significant problems with current transferable force fields for organic molecules when used for calculating binding affinities, especially in non-protein chemistries.

Keywords Expanded ensemble · Host–guest · Binding free energy · Molecular dynamics

Introduction

Computing the binding affinity of small molecule host–guest complexes is both an interesting scientific problem and a useful test of free energy calculation methodology. Computed free energies of binding in protein systems have been shown to be highly sensitive to conformational sampling of both the receptor [1–3] and the ligand [3, 4]. Small molecular hosts represent an intermediate step in complexity of free energy calculations between solvation calculations and full protein–ligand binding affinities. Host–guest complexes achieve a range of binding affinities and require sampling over possible binding states and ligand configurations to achieve converged free energy estimates. However, small molecule hosts are relatively rigid compared to proteins and because of their smaller size allow roughly a tenfold reduction in the number of required atoms. Different methods can therefore be tested in a reasonable amount of simulation time. If a method fails to converge for a host–guest system, then it is unlikely to work for a more challenging protein–ligand complex.

The small molecule cucurbit[7]uril (CB7) and octa-acid (OA) hosts bind to a number of small guest molecules with a large range of binding affinities [5, 6]. For cucurbit[*n*]uril systems, guest binding has a strong dependence on both the physical and chemical properties of the host and guest, as well as on solvent conditions such as pH and salt content [5, 7]. In addition to this, CB7 has been shown to bind

Electronic supplementary material The online version of this article (doi:10.1007/s10822-014-9716-4) contains supplementary material, which is available to authorized users.

J. I. Monroe · M. R. Shirts (✉)
Department of Chemical Engineering, University of Virginia, Charlottesville, VA 22904-4741, USA
e-mail: michael.shirts@virginia.edu

J. I. Monroe
e-mail: jim5bm@virginia.edu

ligands with affinities across the range of those observed in biological systems [8, 10]. Accurately predicting binding affinities for these systems using molecular simulation has still proven very challenging despite the simplicity of the systems [11]. In this study, we attempt to compute fully converged binding free energies for these two host–guest systems. Many of the ligand binding affinities computed in this study were blind predictions as part of the SAMPL4 challenge [11].

While convergence of host–guest binding complexes may be accomplished in some cases through standard atomistic molecular dynamics, in this study we speed convergence by enhancing sampling with expanded ensemble [12] simulations along the coupling parameter determining the strength of ligand intermolecular interactions with the rest of the system [13, 14]. By making Monte Carlo moves between intermediate states of ligand-system interaction, expanded ensemble simulations improve sampling, because many configurations creating kinetic barriers to conformational transitions disappear at decoupled intermediate states. To sample all states equally, we must determine the proper weighting function for each state so that the simulation does not spend all of the time in low free energy states. We determine these weights self-consistently using the Wang–Landau algorithm [14, 15]. In this procedure, we continually update the weight of each state according to the frequency that each state is visited, with the update magnitude gradually decreased over time. The weights are adjusted such that the average probability of visiting each state is nearly equal and thus even sampling in state space is achieved. This will occur when the weights are equal to the free energies of each state. Simulations can then easily traverse state space allowing ligands to escape trapped metastable binding states and sample the full binding ensemble. This acceleration of escape from local minima is similar to what can be done with Hamiltonian replica exchange simulations [16], but requires only a single simulation, rather than a set of parallel simulations.

Testing for convergence is vital to validating free energy estimates from simulations [17]. Without such checks, it is impossible to determine whether the difference between simulation and experiment lies in the method itself or in the model chosen to describe the system of interest. Thus, before attempting to improve force field parameters or adjust environmental conditions, we must show that the methods are converging to a single estimate for the model employed. There is no formal methodology for assuring the convergence of free energy estimates from molecular simulation, though various methods for assessing sampling have been described for both thermodynamic and kinetic observables [18–20]. In any complex molecular system, it is impossible to rule out the existence of configurations that are low energy but that have not yet been sampled. However, there are a number of heuristics that can be used to significantly reduce the chance that simulations are not converged. Starting multiple simulations from the same or different initial conformations and observing whether they give the same or different answers is popular, but may still be biased by the choice of starting configurations [17]. We can also investigate whether key degrees of freedom are sampled sufficiently within the simulation time scale, and that the correlation times of each of the relevant observables are sufficiently low [20].

In order to examine convergence in this study, we implemented a range of different tests. We started by checking our expanded ensemble methodology against the well-established use of separate simulations at each intermediate state of ligand interaction. We also obtained free energies of binding from expanded ensemble simulations via an alternative alchemical pathway, disappearing the host instead of the guest. Theoretically, this will yield identical results to those obtained from bringing the guest to a non-interacting state. To examine the completeness of configurational sampling in these systems, we performed cluster analysis of ligand conformations within each host, and monitored the rate of sampling between these structural clusters. Finally, we examined the behavior of the free energy estimate and the values of the Wang–Landau weights with respect to simulation time, monitoring convergence to the final estimate. If we are achieving sufficient conformational sampling by visiting all of the accessible configurational states properly during the initial weight adjustment phase, the estimated weights will agree with the weights (i.e., free energies) determined by a more rigorous free energy method. This last test represents both a test of conformational sampling and of the convergence of the Wang–Landau algorithm, which is known to fail to fully converge in a stochastic manner [21].

Methods

Preparation of systems

For both hosts and all ligands (structures shown in Figures 1 and 2 of Online Resource 1), atom types, bond types, and partial charges were assigned using antechamber run through acpype [22, 23], which assigns General AMBER Force Field [24] parameters with AM1-BCC charges [25]. Protonation states were assigned to match experimental conditions provided in the SAMPL4 challenge, namely a pH of 7.4 for the CB7 guests, and a pH of 8.7 for the OA system. At these pH conditions, all nitrogens on CB7 guests are expected to be protonated, while all carboxylic acid groups on the OA host and guests are expected to be

fully deprotonated. This resulted in a −8 charge being assigned to the OA host and −1 charges to each of its respective guests. Parameter files for both hosts and each ligand are included in Online Resource 2.

Once all molecules were consistently parametrized, each simulated system was built. VMD [26] scripts were employed to move the geometric center of the ligand to the geometric center of the host. GROMACS 4.6 [27, 28] tools `editconf` and `genbox` were then used to center the system in a cubic solvent box of TIP3P [29] water with the sides of the box at least 1.5 nm away from the solute. The GROMACS tool `genion` was then used to add enough ions to bring the net charge of the system to zero. For systems containing CB7 ligands, Cl^- ions were added to balance the charge of the ligand. Na^+ ions were used to balance the charge on the OA host, as well as to balance the −1 charges of the carboxylic acid guests of this system. Following preparation and energy minimization, CB7-complexed systems contained around 2,700 water molecules with an average box size of 84 nm^3. Octa-acid complexed systems contained around 4,100 water molecules with an average box size of 126 nm^3.

This procedure resulted in all anion concentrations approximately half of those used experimentally for CB7 systems [30], and a complete lack of cations that might competitively bind the CB7 system [7]. While salt concentration has been shown to influence CB7 guest binding significantly [7, 30], the complexities of correctly incorporating reliable models for specific ionic components used experimentally was deemed to outweigh the possible increase in accuracy. Further, the experimental conditions modeled should not influence our studies of the convergence of the sampling methodology employed.

Solvated systems were then energy minimized using GROMACS, first with 5,000 steepest descent steps, then with 5,000 steps (or until the maximum force was lower than 10.0 kJ/mol/nm) using the L-BFGS approach [31]. Sequential temperature and pressure simulations were performed, with 100 ps performed in first the NVT then NPT ensembles. Reference temperatures were set to 300 K and reference pressures were set to 1 bar, with a Berendsen barostat and thermostat used to control each during temperature and pressure equilibration.

Training set ligands for the CB7 system were obtained from experimental studies of Moghaddam et al. [8, 9], Wyman and Macartney [10]. Ligands from Wyman and Macartney [10] were numbered 20 through 30, while those from Moghaddam et al. [8] kept the same numbering as in the original paper (see Figure 1 in Online Resource 1). All ligands except 26, A4, and B11 were used. These three guests had difficulties converging the Wang–Landau weights, requiring more simulation time than could be dedicated to the development of the training set. The training set excluded guests involving coordination with metal ions, but was chosen so as to encompass a broad variety of CB7 ligands, including both weak and strong binders of diverse chemical structure.

Production molecular dynamics

Fully atomistic, expanded ensemble simulations were run for each system using GROMACS with time steps of 0.002 ps. Representative run input files may be found in Online Resource 2. Nośe–Hoover temperature coupling and Martyna–Tuckerman–Tobias–Klein pressure coupling [32] were both used and bonds to hydrogen atoms were constrained using the SHAKE algorithm [33, 34]. Metropolized Gibbs Monte Carlo moves [35] between all intermediate ligand coupling states were performed every 100 time steps based on weights calculated with the Wang–Landau algorithm [15]. The metropolized Gibbs move in state space proposes jumps to all states except the current state, with a rejection step to satisfy detailed balance [35].

For all ligands alone in solution with an absolute value of the charge less than 2, the initial Wang–Landau incrementor was set to 2 k_BT, while for larger or more highly charged systems, an initial incrementor of 10 k_BT was used. Weights were updated at each step, and the increment amount was reduced by a factor of 0.8 each time a flat histogram was observed, meaning that the ratio between the least visited and most visited states since the last change in the weight increment was less than 0.7. The process of updating the weights was halted when the incrementing amount fell below 0.001 k_BT. The weight of the fully coupled state is normalized to zero, meaning that the weight of the uncoupled state corresponds to the free energy of the process.

The last stage of the simulation, during which period WL weights were no longer updated, was termed the "production" stage since it was the only part of the trajectory used to calculate the final free energy change. Once weights are fixed, the expanded ensemble simulation exactly preserves the correct NPT ensemble at each state. Expanded ensemble simulations disappearing the hosts were exceptions to this procedure due to severely uneven coupling state sampling after the weights became fixed (discussed in "Free energy convergence over time in expanded ensemble simulations" section). For these simulations, data from approximately 20–30 ns before the production region up until the beginning of the production region were instead used to compute the free energies since this provided a more even distribution of state sampling to the free energy calculation method. We will show that the error introduced by using the samples just before final equilibration of the weights is essentially zero in this case.

In all simulations, the initial state along our alchemical pathway involved full intermolecular interactions of the

molecule. Twenty-one states with equal spacing of 0.05 were used to first turn off Coulombic interactions. An additional 19 states were used to turn off van der Waals interactions, with spacing chosen so that intermediate states are more closely spaced in regions containing higher variance [36]. This resulted in a total of 40 states of varying levels of ligand intermolecular interactions used in expanded ensemble simulations for all systems studied. Specific intermediate states used are provided in Online Resource 1 and in simulation input files in Online Resource 2. To speed convergence of the simulations of the complexes, the converged Wang–Landau weights from the solute being decoupled from the solution alone were used as the initial weights for the simulation of the same solute being decoupled from the complex. The *pymbar* [37] implementation of MBAR [38] was then used to compute all free energy differences using the potential energy differences between the current state and all other intermediate states stored at 2 ps increments.

For CB7 guests 1, 10 (both SAMPL4 ligands), and 23 (a previously published ligand), for comparison, simulations were also run at fixed lambda states. In this case, only the decoupling of the guest in both solvent and bound to the CB7 host was performed. For these simulations, an even lambda state spacing of 0.1 was used to turn off both Coulombic and van der Waals forces, resulting in 21 total lambda states. Each lambda state was run for 15 ns using identical simulation conditions to those described above, except that Monte Carlo moves between lambda states were not employed. The first 5 ns were excluded as equilibration time when the data were passed to the *pymbar* analysis script. Fewer states are necessary in these fixed state simulations since expanded ensemble moves between states, which require relatively small energy differences to achieve high acceptance probabilities, are not performed. We note that the fixed state simulations used approximately 8,000 cpu hours per guest, while expanded ensemble simulations used approximately 2,500 cpu hours per guest.

Of course, using more than triple the computational expense does not necessarily translate into improved estimates of the binding free energy. Because more time is spent sampling at each intermediate state in the fixed state simulations, the *estimated* statistical uncertainty in the free energy of binding will frequently be lower for fixed lambda state simulations, even after we subsample the sampled data to obtain statistically uncorrelated information. However, we will later show that configurational sampling is kinetically limited for fixed state simulations at longer time scales. In this study, these fixed state simulations will highlight the ability of expanded ensemble to more efficiently sample configuration space, and thus provide more accurate free energy estimates at lower computational expense when long timescale conformational transitions limit sampling.

The usage and benefits of ligand restraints in the uncoupled states in computing free energies of binding have been extensively documented [39–41]. However, in this work, we have not restrained ligand conformations in simulations of the host–guest complex other than to prevent the ligand from drifting away from the host when attractive Coulombic forces are turned off. This absence of orientation restraints is preferred over the use of such additional restraints because the precise bound configurations of these complexes are not known. Thus, simple harmonic restraints between the host and ligand centers of mass were employed. The distance between the host and guest centers of mass in the initial, energy minimized bound state was used as a reference distance for the harmonic potential, which was fully turned on between λ_{coul} values of 0.5–1. For complexed systems in which the ligand or the host was removed, a harmonic restraint between the centers of mass of the guest and the host was applied. Again, we emphasize that this restraining potential was not on in the fully coupled state, and thus does not thermodynamically bias the bound state to any location, but merely serves to keep the ligand from drifting away from the guest when non- or weakly interacting.

The free energy change upon imposition of the restraint in solvent alone is $-k_BT \ln \frac{Q_{res}}{Q_{non}}$, where Q_{res} and Q_{non} are the partition functions of the restrained and non-restrained systems, respectively. The restrained and non-restrained partition functions are the integrals over the entire box volume with and without restraint forces, respectively. The result is $\Delta G_{res} = k_BT \ln\left[\frac{3}{V}\left(\frac{\pi k_BT}{K}\right)^{3/2}\right]$, where K is the spring constant and V is the molecular volume (1.6605 nm^3) corresponding to the 1 mol/L reference concentration. For the choice of parameters in this system, this yields 3.98 kcal/mol with this analytical correction directly added to the computed free energy change from simulation to obtain the physically relevant binding free energy.

Periodic electrostatic effects in small solvent boxes

If the volume of the molecule being disappeared is large relative to the size of the simulation box, non-trivial artifacts in the free energy may be observed [42]. In order to ascertain whether or not differences in box size significantly impacted free energy changes, simulations at fixed lambda states were performed for both small (3.7 nm per side) and large (5.1 nm per side) cubic solvent boxes. These box sizes are representative of the size of a solvent box for a guest alone in solvent, and for the largest of CB7 complexed systems. Since artifacts are most pronounced for highly charged systems [42], CB7 guest 10, which has a +3 charge, was used. Free energies of solvation at both the smallest and largest box sizes employed in our simulations

revealed no statistically significant difference in the predicted free energies. In the small box, the solvation free energy was 370.8 ± 0.10 kcal/mol, which is identical to that within the larger box to within statistical uncertainty. Higher order corrections to the electrostatics due to differences in dielectrics of the solvent and complex systems may also occur [43] but were neglected in this study.

Clustering

Clustering was performed by first using the GROMACS tool `trjconv` to remove periodicity, center the host and guest, and remove translation and rotation of the host from all molecules. We then used `g_cluster` to cluster ligand conformations using the single linkage algorithm implemented in this GROMACS utility. Full trajectories of all CB7 complex simulations and fixed weight regions of these trajectories were subject to separate cluster analysis. The cut-off for cluster differentiation in the full trajectories was set to an RMSD between conformations of 0.1 nm. The cutoff was increased to 0.13 nm for the equilibrated regions of the trajectories. This did not significantly change the representative structures of the predominant clusters, but did reduce the total number of clusters and thus increased the accuracy of calculations for flipping rates between dominant clusters. Since flips were counted as any transition from one cluster to another, a smaller number of clusters reduced the number of flipping events involving very small clusters. Both RMSD cutoffs used resulted in between two and four predominant clusters for all guests at the fully interacting lambda state, while at the decoupled state no significant clusters were observed.

Results and discussion

Comparison to experiment

For the CB7 system, predicted absolute free energies of binding were well off from experiment with an RMSE value of 8.96 ± 0.74 kcal/mol (RMSE uncertainty from 10,000 bootstrapped samples). Absolute predictions for OA guests were much closer to experiment, yielding an RMSE of 1.67 ± 0.38 kcal/mol. This value differs from that reported in the SAMPL4 overview paper since the average error from experiment has not been removed from our values before calculation. In the overview, the procedure of removing the average error was implemented to avoid penalizing submissions of relative free energies of binding for which the prediction for the reference guest was especially poor. In the overview, RMSE for our OA8 submission is reported as 1.3 ± 0.4 kcal/mol, while submission IDs 170 and 526, which employed similar models and methodologies, had RMSE values of 1.0 ± 0.4 and 0.8 ± 0.4 kcal/mol, respectively. Though methodologies differed, the model employed by ID 170 was identical to our own, using the same water model, force field, and partial charges as in this study, while ID 526 only differed in the use of RESP partial charges rather than AM1-BCC. It is thus reassuring that differences in the overview RMSE values to experiment between our own and each of these other submissions is statistically indistinguishable.

However, more direct comparison of our results with submission IDs 170 and 526 is merited considering the similarity of the models employed. Since these other studies computed all free energies of binding relative to OA guest 6, we similarly subtracted our computed absolute free energy of guest 6 from our results. Computing RMSE between our submission adjusted relative to guest 6 and ID 170 using the same procedure as described by Mobley, we obtain 0.50 ± 0.13 kcal/mol, while for ID 526 we obtain 0.57 ± 0.14 kcal/mol. The errors in the RMSE mainly indicate the expected uncertainty in the RMSE for replication of each method for computing binding free energies with different guests. To gain an idea of the magnitude of differences in free energies of binding between the two methods for specific guests, we determine z-scores of the difference in the two predictions, assuming a mean of zero and a standard deviation obtained from the specified errors in the predictions. Excluding guest 6, we find that half of the free energy differences between our results and the other two studies are within one standard deviation of zero, while guests 1, 3, 4, and 5 yield differences around two standard deviations from zero. The differences between our calculations and these calculations are thus statistically distinguishable at some low level.

Although the parameterization schemes were listed as being identical in the papers for submission 170 and 562, we found significant differences in partial charges of the hosts. Changes in the partial charges for the ligand itself were smaller than 0.01 charge units for all ligands. However, our BCC-parameterizd OA host charges were significantly less polarized than the BCC-parameterized OA-host used in the submission 170. In our host, partial charges of the ethers, carbonyl oxygens, and carbonyl carbons were approximately −0.2, −0.6, and +0.35 charge units respectively. For submission 170, these three BCC-parameterized atom types had partial charges of approximately −0.3, −0.85 and +0.9 charge units respectively. Our hosts were parameterized as an entire unit, following the same scheme as the ligand. Details are not provided for the parameterization of the host in the other paper, but it seems likely that whatever the reason, this charge difference may explain the statistically significant differences of 0.5 RMS kcal/mol in binding affinity between two studies with putatively the same parameterization schemes.

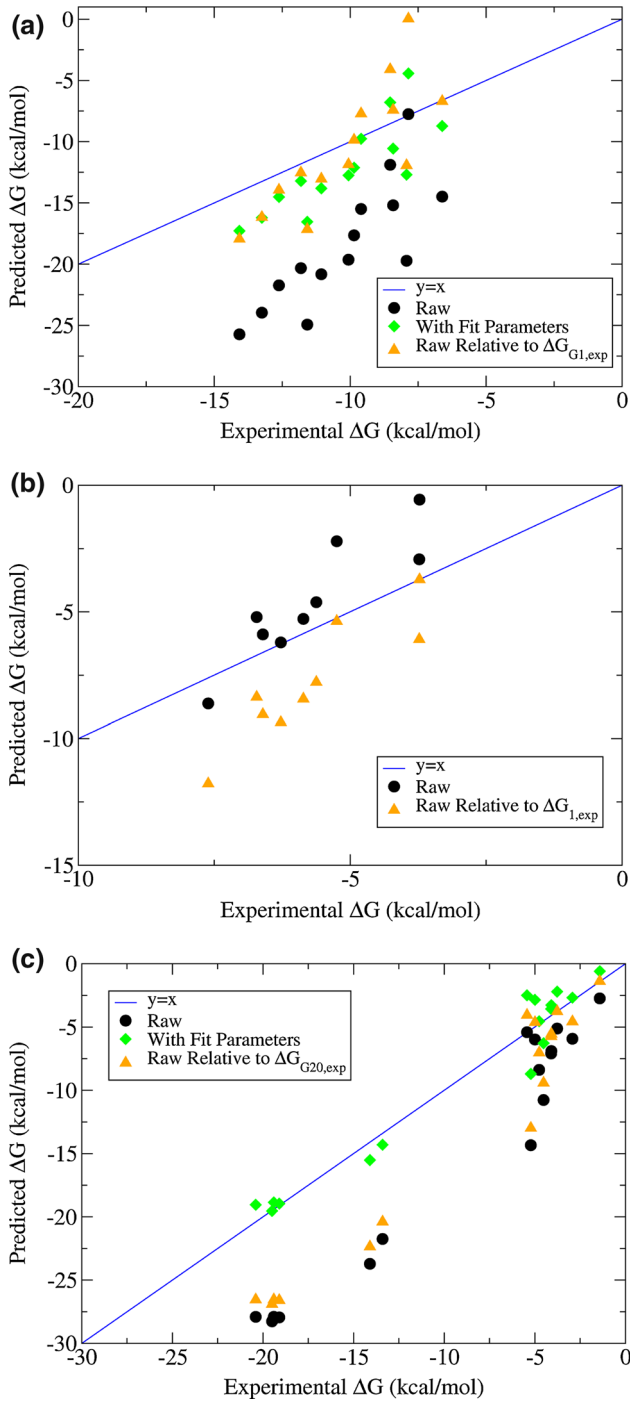

Fig. 1 Computed free energies of binding are plotted against experimental values for SAMPL4 CB7 guests (**a**), SAMPL4 OA guests (**b**), and for the CB7 training set (**c**). Relative values were computed with respect to CB7 guest 1 (**a**), OA guest 1 (**b**), and CB7 guest 20 (**c**). Predicted absolute binding free energies were inconsistent with experiment for both CB7 and OA (**a**, **b**). Corrections from empirical fit parameters to the van der Waals and Coulombic terms reduced the RMSE of the training set from 5.26 ± 0.67 to 1.53 ± 0.27 kcal/mol. For the SAMPL4 ligands, correcting using the parameters derived from fitting the training set only reduced RMSE from 3.58 ± 0.67 to 2.88 ± 0.33 kcal/mol indicating limited transferability of these empirically fit parameters

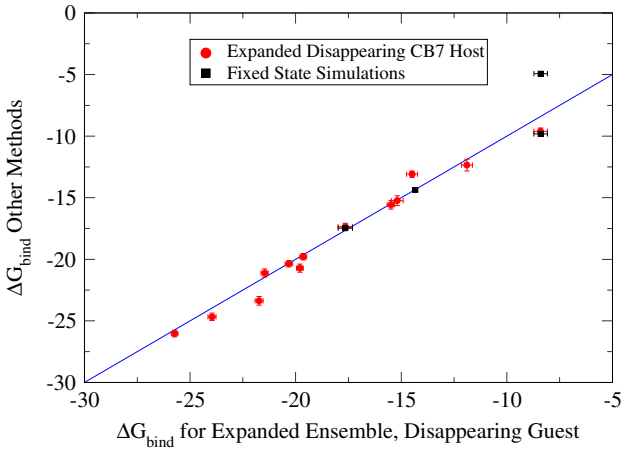

Fig. 2 Free energies of binding from fixed lambda state simulations and expanded ensemble simulations disappearing the host are plotted against those computed from expanded simulations uncoupling the guest. More than half of the points fall within 1 SD of the y = x line, with an RMSE of 0.78 ± 0.15 kcal/mol and a correlation coefficient of 0.99 observed between all expanded ensemble simulations disappearing the host and the guest (*red dots*). This indicates convergence between expanded ensemble simulations employing either alchemical pathway. The two squares for ligand 10 at about −8 kcal/mol for the expanded ensemble free energies represent results from two initial starting configurations for fixed lambda simulations. Uncertainties are standard errors in the calculations

In accordance with SAMPL4 guidelines, raw absolute free energies were submitted for the OA host, while raw and "corrected" free energies (both described below) relative to guest 1 were submitted for the CB7 host. Generally, predicted free energies were too negative for CB7 ligands (mean unsigned absolute error of −8.22 kcal/mol), while, conversely, predictions of affinity were insufficiently negative for OA guests, with an unsigned MAE of 1.10 kcal/mol (Fig. 1). Relative free energies without empirical adjustments still diverged from experiment, at RMSE values of 3.58 ± 0.67 kcal/mol for CB7 and 2.56 ± 0.37 kcal/mol for the OA host. This indicates that errors were system and molecule dependent and cannot be easily corrected with a simple overall linear fit parameter (Fig. 1c). This further implies that errors may stem from different sources in different guests. The larger relative RMSE value observed for the OA system likely stems from the exceptionally poor prediction for OA guest 1. At 3.2 kcal/mol away from the experimental value, this is the largest error in the prediction of any OA guest.

One simple possible explanation for the deviation of calculated free energies from experimental values is that there is an overall error in the scaling in either or both of van der Waals or Coulombic energies, but that relative values of these components are correct. To test this hypothesis, we examined an empirical fit with separate parameters for the Coulombic and van der Waals

Table 1 Differences in the fit parameters for the vdW and Coulombic terms using only the training set ligands, only the final test set of SAMPL4 CB7 ligands, or all of the ligands pooled

	Training set	SAMPL4	All CB7 ligands
RMSE absolute	6.18 ± 0.72	8.96 ± 0.74	7.61 ± 0.57
RMSE relative	5.26 ± 0.66	3.58 ± 0.66	3.96 ± 0.44
RMSE absolute with fit from training set	1.53 ± 0.27	2.88 ± 0.33	2.26 ± 0.25
RMSE absolute with fit from same set	1.53 ± 0.27	1.82 ± 0.42	1.97 ± 0.21
vdW parameter (α)	0.81 ± 0.13	0.65 ± 0.14	0.70 ± 0.11
Coulombic parameter (β)	0.70 ± 0.023	0.62 ± 0.028	0.66 ± 0.020

Although in the case of the training set the fit significantly decreases RMSE of predictions from experimental values compared to the same statistic computed for relative free energies of binding, the lack of agreement within error between the best fit parameters demonstrates that such a fitting procedure is not fully transferable. Errors in fit parameters come from the covariance matrix of parameters M determined through the linear least squares formalism, which may be expressed as $M = \frac{S}{m-n} * (X^T X)^{-1}$, where S is the sum of the squared residuals, X is the data matrix, m is the number of samples, and n is the number of degrees of freedom used in the fitting procedure (here $n = 2$). All RMSE are in kcal/mol with errors in the RMSE from 10,000 bootstrapped samples

contributions to the free energy change. We take a simple model $\Delta G_{bind} = \alpha \Delta G_{vdw} + \beta \Delta G_{coul}$, where ΔG_{coul} and ΔG_{vdw} are the differences in free energies of discharging in the presence and absence of the host and removing Lennard–Jones interactions in the presence and absence of the host. We then adjust the parameters α and β to minimize the root mean square difference between the experimental and calculated free energies for the training set.

The resulting fit parameters for the bilinear regression are shown in Table 1, with values less than one indicating that both the Coulombic and vdW terms result in excessively negative binding free energies to the CB7 host. Since a training set was not developed for the OA host, implementation of this empirical fitting approach in the OA system was not investigated. Figure 1c shows that using linear fitting parameters for the training set significantly reduces RMSE from the experimental values over the computed relative binding affinities.

However, when the correlation parameters fit to the training set were used to "correct" the SAMPL4 ligands, the RMSE only improved by 0.7 kcal/mol from that for the relative free energies. This compares unfavorably with the reduction in RMSE of almost 4.0 kcal/mol for the training set. In the initial submission to SAMPL4, these

Table 2 Different simulation methodologies and alchemical pathways yield very close free energies of binding (in kcal/mol)

CB7 guest	$\Delta G_{expanded}$	$\Delta G_{expanded\_host}$	ΔG_{fixed}	$\Delta G_{experiment}$
1	−17.7 ± 0.34	−17.4 ± 0.27	−17.5 ± 0.097	−9.9 ± 0.18
2	−15.5 ± 0.18	−15.6 ± 0.35	–	−9.6 ± 0.18
3	−14.5 ± 0.26	−13.1 ± 0.26	–	−6.6 ± 0.18
4	−15.2 ± 0.28	−15.2 ± 0.40	–	−8.4 ± 0.20
5	−11.9 ± 0.26	−12.4 ± 0.48	–	−8.5 ± 0.17
6	−19.7 ± 0.16	–	–	−7.9 ± 0.18
7	−19.6 ± 0.16	−19.8 ± 0.26	–	−10.1 ± 0.17
8	−20.3 ± 0.18	−20.4 ± 0.26	–	−11.8 ± 0.17
9	−21.7 ± 0.19	−23.4 ± 0.35	–	−12.6 ± 0.17
10	−8.4 ± 0.32	−9.6 ± 0.27	−4.9 ± 0.17, − 9.8 ± 0.16	−7.9 ± 0.19
11a	−19.8 ± 0.16	−20.7 ± 0.33	–	−11.1 ± 0.15
11b	−21.5 ± 0.17	−21.1 ± 0.33	–	−11.1 ± 0.15
12	−24.0 ± 0.19	−24.7 ± 0.30	–	−13.3 ± 0.17
13	−25.7 ± 0.16	−26.0 ± 0.27	–	−14.1 ± 0.19
14	−24.9 ± 0.20	–	–	−11.6 ± 0.19
23	−14.3 ± 0.12	–	−14.4 ± 0.06	−5.2 ± 0.09

The discrepancy observed between fixed state and expanded simulations for guest 10 is likely attributable to trapping in a local free energy minima due to the starting configuration used in fixed intermediate state simulations (see "Analysis of configurational sampling" section). As seen from the second value in the table for fixed lambda simulation of guest 10, if the simulation is started from a different configuration, a free energy of binding that is nearly 5 kcal/mol more negative than the first, much closer to the expanded ensemble results, is obtained. Estimated statistical uncertainties are lower for fixed lambda state simulations due to greater sampling at each lambda state, but as seen for guest 10 can sometimes be large underestimates of error. For guest 11 the experimental value is for a racemic mixture of the two enantiomers R (a) and S (b)

"corrected" free energies actually produced a larger RMSE than the raw, relative submission, however, this was due to a mathematical error corrected after submission. A reduction of RMSE by 0.7 kcal/mol is not a statistically significant improvement over simple calculation of relative binding free energies (Fig. 1a) when considering that the standard deviation of the RMSE for the relative free energies of binding is 0.67 kcal/mol. The parameters derived from this fitting procedure do not appear to be particularly transferable from one set of CB7 guests to another (Table 1). One likely reason is the fact that the guests in the training set were not representative of the SAMPL4 ligand set. While both low and high affinity binders are included, all low affinity guests were small, uncharged molecules. All high affinity training set guests were rigid, polycyclic molecules, with the majority being charged. The SAMPL4 guests were all charged and were more varied in their chemical structure, resulting in a wider range of predicted and experimental affinities. While a more representative training set would help in deriving more appropriate parameters, there appears to be no significant benefit to "correcting" binding free energies of SAMPL4 ligands with the parameters derived from our training set. While such an empirical fit directly applied to the SAMPL4 CB7 ligands themselves does improve the RMSE to 1.82 ± 0.42 kcal/mol (Table 1), the parameters seem to be highly specific to the data being fit and lack consistent transferability.

Checking the convergence of multiple simulation methods

To check the convergence of our CB7 host expanded ensemble simulations for the force field parameters used, we calculated free energies of binding using fixed lambda simulations, as well as through an alternate alchemical pathway where the host was decoupled from its environment rather than the ligand. Calculating the free energies of binding through these different methodologies should result in answers that are the same within statistical uncertainty if they are sufficiently converged. Table 2 shows that expanded ensemble simulations disappearing both the guest and the host result in surprisingly similar free energies of binding given the extremely different processes. Most of these simulations actually agreed well enough that the difference between the predicted free energies was within 1 SD of zero. Exceptions were CB7 guests 3, 9, 10, 11a, and 12, with the maximum z-score

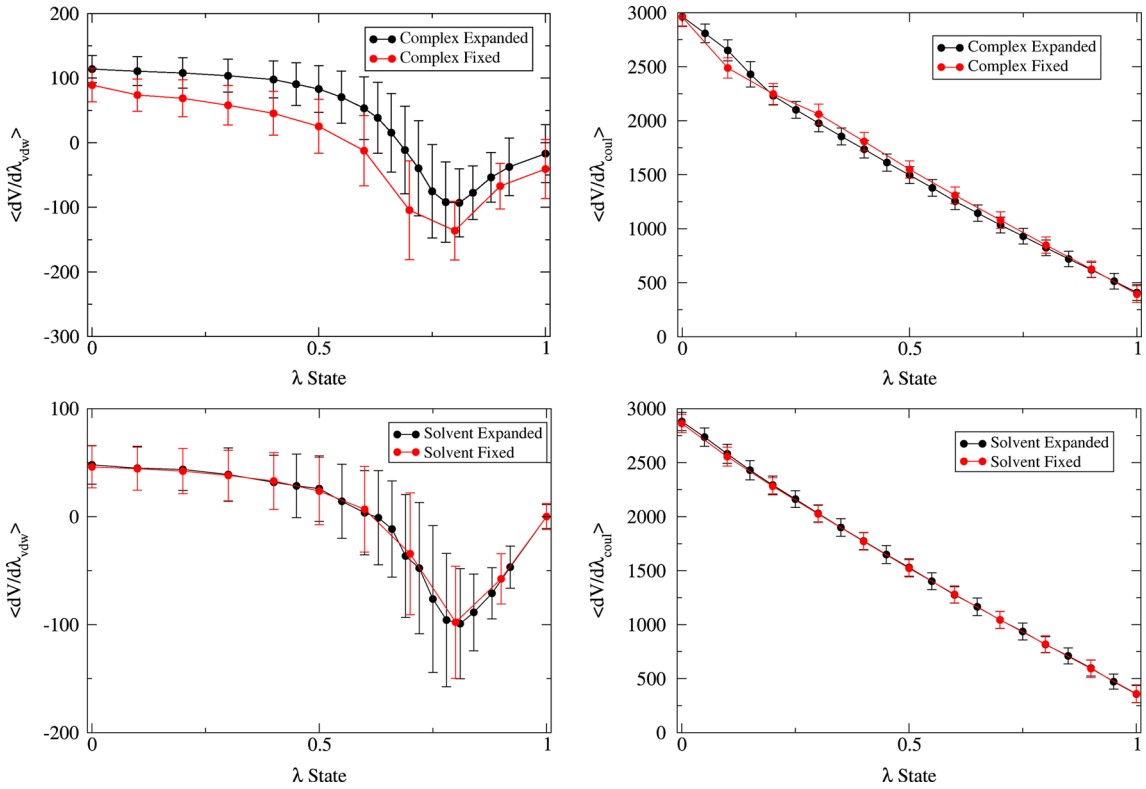

Fig. 3 Plots of the average derivative of the Hamiltonian for both solvation (*bottom*) and complex uncoupling (*top*) of CB7 guest 10 with respect to coupling parameter (lambda) values associated with van der Waals (*left*) and Coulombic (*right*) interactions. Lower derivatives across lambda values for the fixed state simulation of the complex indicate that sampling may not be complete compared to the expanded ensemble simulation

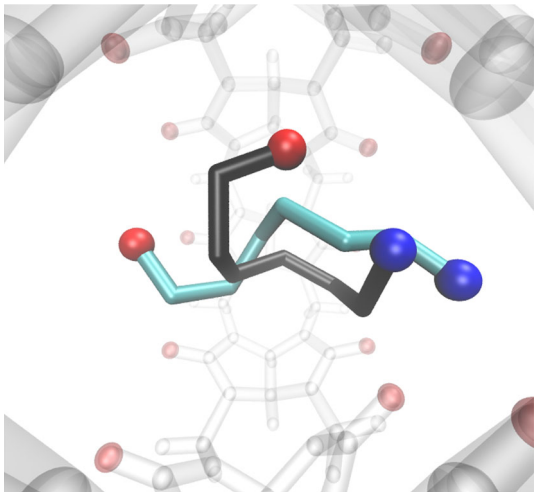

Fig. 4 Visualization of the two predominant configurations observed for CB7 guest 3: a "U" shape (*black sticks*) with the amine group at the portal and the hydroxyl facing in the same direction but slightly more inside the host, and an extended configuration (*cyan sticks*) with the amine and hydroxyl groups residing on opposite faces of the host

being 4.3 where the difference was 1.7 kcal/mol. It is particularly impressive that the systems were able to converge to this extent given the magnitudes of the individual $\Delta G_{complex}$ and $\Delta G_{solvent}$ calculations that are subtracted to give the ΔG_{bind}. For decoupling of the guests, these numbers range from 10 to 180 kcal/mol, depending on the charge or polarity of the ligand, and for decoupling the hosts, the numbers range between 150 and 180 kcal/mol, involving the decoupling of about 80 heavy atoms. Converging the differences in these numbers to within 0.1–0.3 kcal/mol, the statistical uncertainty, in cases without obvious sampling errors demonstrates the power of this expanded ensemble approach (Fig. 2).

Difficulty in converging the weights for guests 3, 9, and 10 (see "Free energy convergence over time in expanded ensemble simulations" section) seems to correlate with larger discrepancies between expanded ensemble simulations disappearing the guest and the host. Simulations at fixed ligand interaction states for guests 1 and 23 also resulted in values statistically equivalent to those produced from expanded ensemble (Table 2), though guest 10 did not. To trace the source of this 3.5 kcal/mol discrepancy in different methods of calculating the binding free energy of guest 10, we plotted $\langle dH/d\lambda \rangle$ for all Coulombic and van der Waals intermediate states for simulations of guest 10 freely solvated and complexed to CB7 for both fixed state and expanded ensemble simulations (Fig. 3). The fixed state simulations under-predict the Hamiltonian derivative at nearly all van der Waals intermediate states for the CB7/guest 10 complex. Since all other curves in Fig. 3 are nearly identical, this appears to be the only source of the error. Configurational clustering for the CB7/guest 10 complex indicate that this may be due to excessive sampling of a local free energy minimum conformation in the fixed state simulation (see "Analysis of configurational sampling" section). This is reinforced for the fixed lambda simulation by a greater average number of water molecules in the binding pocket across nearly all ligand interaction states when compared to expanded simulation averages of bound waters. Such increases in the level of host hydration have been shown by Rogers et al. [44] to correspond to a decreased coupling parameter (lambda) derivative of the Hamiltonian. In the fixed state simulation of the fully coupled state, the maximum observed number of waters in the pocket (7 waters) is no longer observed after the guest flips conformations. The implications of this conformational flipping are described more fully in the following section.

Analysis of configurational sampling

Cluster analysis of sampled configurations for the CB7 and OA systems generally fit with expectations for converged sampling in these simple systems. Due to the symmetry of the CB7 host, one indication of convergence is the even sampling of mirror image guest configurations at both portals. The ratio of sampling between the most and second most frequently sampled of these two configurations is less than 1.5 for all CB7 ligands except guests 10 and 3. During the fixed weight portions of the expanded ensemble simulations of guests 10 and 3, this ratio jumps to 2.2 and 16.7, respectively. This lack of equal sampling of symmetric states occurs in simulations of the same molecules that have greater difficulty in converging their Wang–Landau weights (see "Free energy convergence over time in expanded ensemble simulations" section).

For expanded ensemble simulations of the highly flexible guest 3, both "U-shaped" and extended conformations were observed, in which the hydroxyl group resides either slightly inside of the pocket on the same side as the amine, or on the opposite portal of the host facing into solution (Fig. 4). Since the host is symmetric, we would expect that four predominant clusters would be observed. However this is not the case. While "U" conformations are observed to flip sides of the host throughout the entire simulation, the extended conformation does not, indicating lack of converged sampling. The extended configuration is also the most frequently sampled, with 94 % of the configurations sampled after fixing the weights contained in this cluster. During this same time period in the guest 3 simulation, sampling in intermolecular interaction space was limited to only a very few intermediate states. This indicates that the simulation essentially became a fixed coupling state simulation, preventing conformational flipping and equal sampling of symmetric configurations of guest 3. Failure to sample is thus in large part due to problems with premature

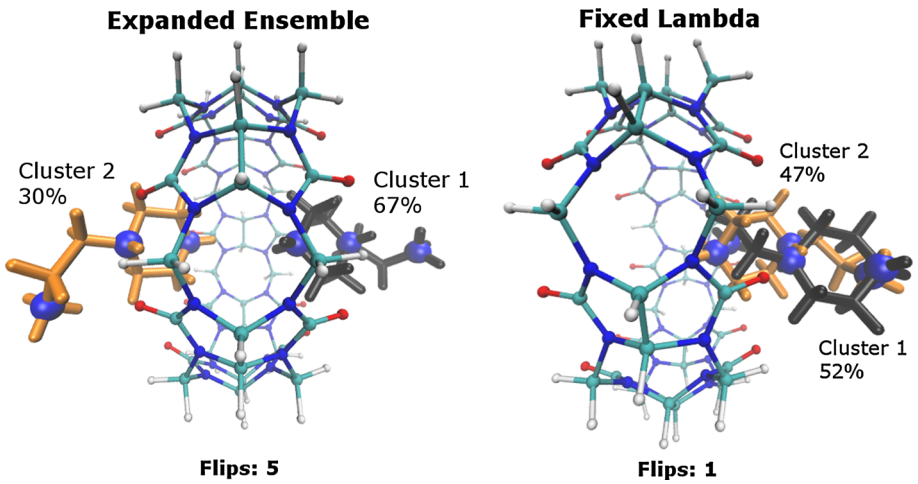

Fig. 5 Expanded ensemble simulation results in enhanced sampling compared with fixed state simulation, which is strongly dependent on the starting structure due to kinetic limitations to configurational sampling. This is evident in the predominant conformations of CB7 guest 10 observed in the fully coupled state for each of the simulation methodologies. The expanded ensemble simulation includes symmetric orientations of the ligand on both sides of the molecule, while simulations at fixed states do not result in this flipping

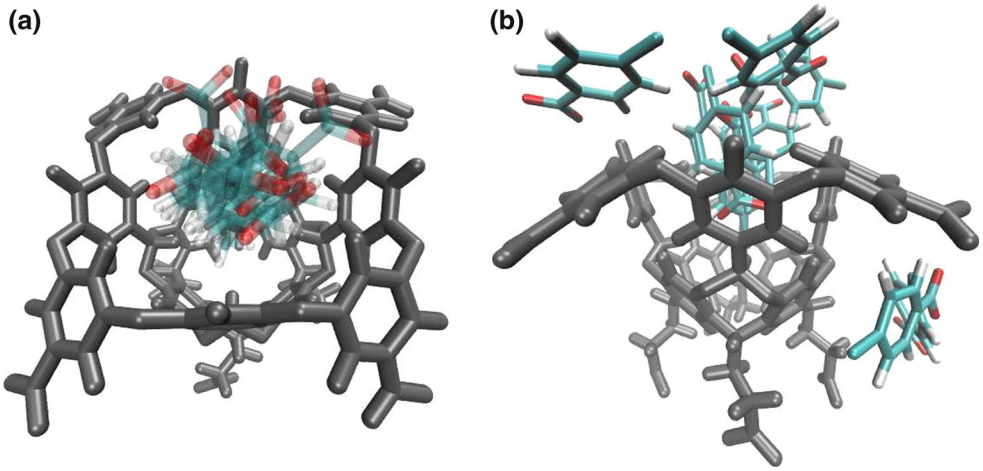

Fig. 6 Guests of the OA host all orient themselves with their carboxyl group pointed out into solution in the predominant clustered conformation (**a**). Guest 4, which binds moderately weakly, experiences rarely occurring conformations in which the carboxyl group flips towards the host cavity, and in which the molecule leaves the host through the cavity wall during simulation at decoupled lambda states (**b**); such configurations would likely not be observed if the ligand was strictly restrained in the binding site, or if fixed intermediate state simulations were used

convergence in the Wang–Landau algorithm, rather than the inability of the expanded ensemble simulations to visit all configurations. Uneven sampling after self-consistent determination of weights is a symptom of poorly converged weights by the Wang–Landau process, and is discussed further in "Free energy convergence over time in expanded ensemble simulations" section.

The ability of expanded ensemble simulations to escape kinetic traps, enhancing sampling and thus improve statistical convergence is most clearly observed between expanded and fixed state simulations of CB7 guest 10. During the fixed state simulation of the fully coupled state, guest 10 started in and subsequently remained in a conformation with the single terminal amine facing the oxygen portal of CB7 while both charged amines in the ring faced into solution (Fig. 5, right). The latter conformation allows a larger number of water molecules to occupy the pocket. Thus, we expect that this conformation is less favorable due to lost electrostatic interactions with the oxygen portal, but also less favorable due to positive free energies of solvation for the host cavity [44].

While this unfavorable conformation was observed in the fully coupled state during expanded ensemble simulation, it is not sampled nearly as frequently as the opposite orientation. Since all fixed ligand interaction state simulations were started from the same structure, many states in

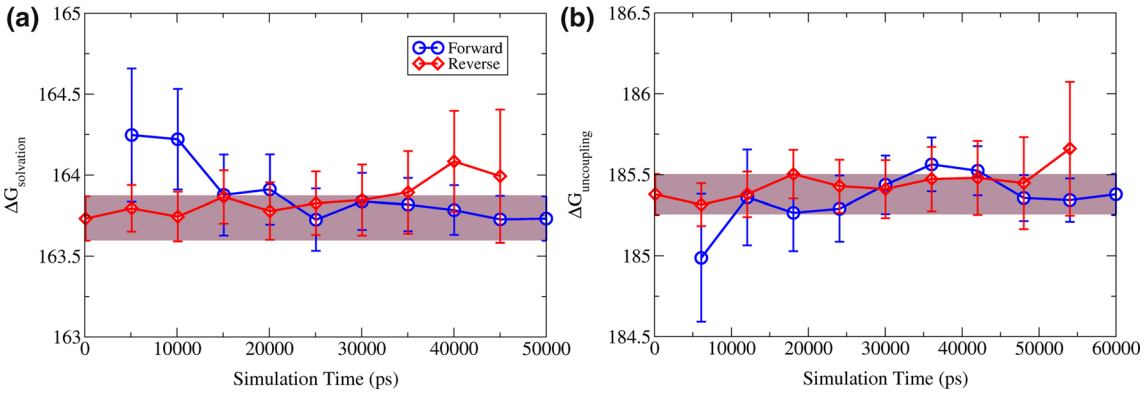

Fig. 7 Cumulatively calculated free energy changes converge when computed in both the forward (*blue*) and reverse (*red*) directions. Shown are the free energy of solvation (**a**) and free energy of uncoupling in complex (**b**) for CB7 guest 1. The *tinted region* indicates values within 1 SD of the final value of the cumulative free energy

addition to the fully coupled state were also biased. If configurations are clustered for various intermediate states along the alchemical pathway, we observe that clustering differences between fixed and expanded simulations are maintained similar to Fig. 5 until highly decoupled states, at which point no predominant clusters are observed for either type of simulation. This is a likely cause of the differences between $\langle dH/d\lambda \rangle$ curves shown in Fig. 3. Biasing of the fixed state simulations by the starting configuration is further confirmed by a more negative free energy of -9.8 ± 0.16 kcal/mol obtained by repeating these simulations with a starting configuration in which the terminal amine is facing into solution. This demonstrates the presence of kinetic limitations on fixed lambda state simulations for some systems. Such kinetic trapping is further demonstrated by a guest 10 fixed state simulation flipping rate between predominant clusters of 0.07 flips/ns, while expanded simulations of CB7 complexes averaged 3.8 flips/ns, a more than 50-fold increase.

In the OA system, opposite faces of the host are not symmetric, resulting in only one entrance for ligands and only one predominant cluster containing a large number of members for the fully coupled state of each ligand. In every case, this cluster was oriented such that the hydrophobic ring remained facing the inside of the cavity while the carboxyl group of the guest protruded out into the solution (Fig. 6a), as expected. For nearly half of the guests, at least one conformation is observed in which the ligand flips its carboxylic acid group towards the host. This is observed for guest 4 in addition to configurations in which the ligand has fully left the host cavity (Fig. 6b).

At fully or almost fully decoupled states, the ligand is able to enter the host and actually pass through the cavity wall. Once completely outside the host and free of steric clashes with the cavity wall, the ligand is able to sample the fully coupled state, resulting in the observed clusters. The ratio of sampling in any fully coupled unbound state of guest 4 to the predominant configuration at full interaction is 1:1,167, which represents a free energy difference of −4.2 kcal/mol. This is close to the predicted free energy of binding of −5.2 kcal/mol, so such configurations are consistent with the binding free energy. Although such rarely observed configurations will contribute negligibly to the free energy, the fact that they are observed shows the sampling power of expanded ensemble simulations compared to standard molecular dynamics simulations.

Free energy convergence over time in expanded ensemble simulations

We can also monitor convergence by observing the behavior over the entire simulation time of the cumulative free energy estimate and the Wang–Landau weights. Figure 7 shows the behavior of the cumulative free energy estimate over simulation time computed using MBAR for both solvation and uncoupling while complexed for CB7 guest 1. As the simulation progresses, and weights adjust via the Wang–Landau algorithm to provide even sampling, the free energy appears to equilibrate to a steady value. However, cumulative averaging in the forward direction is inherently biased by the portions of the simulation spent equilibrating [45]. Extensive sampling of the true binding free energy will then be necessary in order to overcome biases left over from the non-equilibrated period of sampling. To overcome this problem, reverse cumulative averaging may be employed. This technique computes the free energy estimate cumulatively from data starting at the end of the simulation, as demonstrated in Fig. 7.

While quantitative techniques to identify the equilibration time have been proposed by Yang et al. [45], we focus on a simple and quick visual analysis. For plots such as shown in Fig. 7, we define the time for equilibration as the time for both the forward and reverse cumulative free energy values to converge within 1 SD of the final free

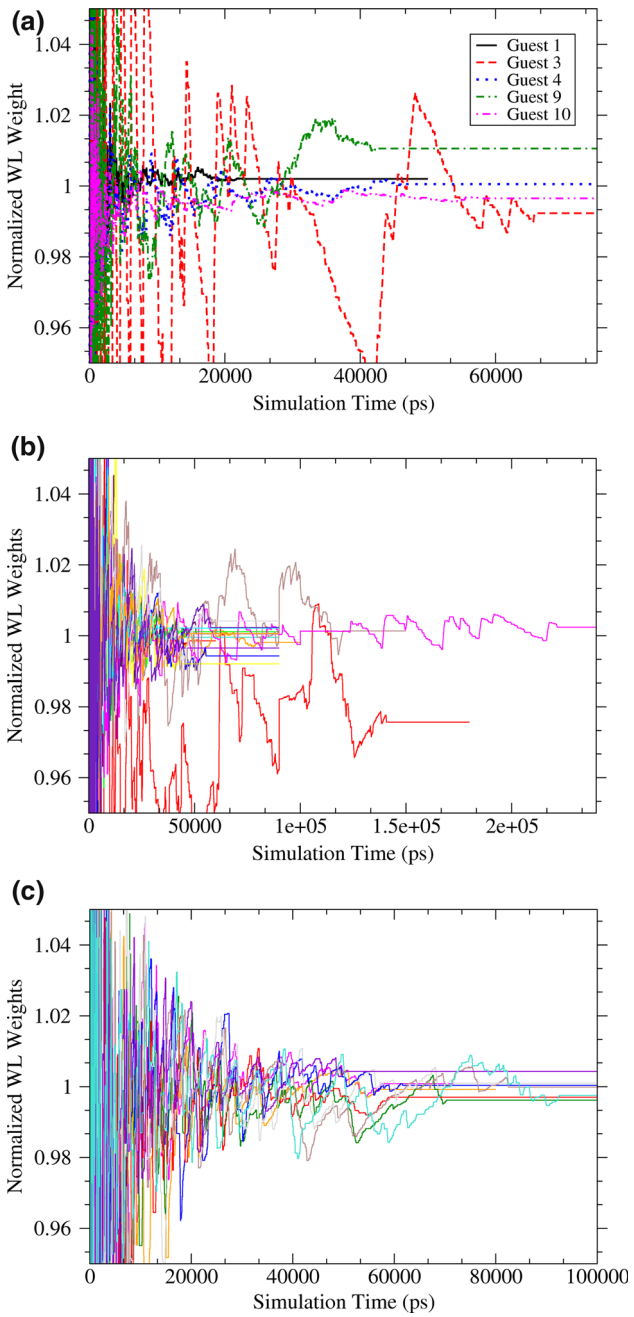

Fig. 8 Near convergence is generally observed for weights during the Wang–Landau process (WL weights) of the fully uncoupled state for each ligand normalized to free energy changes computed with MBAR. Shown are normalized weights for solvation of selected CB7 guests (**a**), CB7 guest uncoupling in complex (**b**), and OA guest uncoupling in complex (**c**). Plots for all SAMPL4 guests are shown for guest uncoupling in the CB7 and OA complexes, though no legend is shown for space reasons. Solvation of OA guests generated well converged curves similar to those for uncoupling in the OA complex. CB7 guest 3, which is shown in red in **a** and **b**, was difficult to converge, possibly due to its highly flexible nature and multiple binding modalities to CB7

energy estimate. CB7 guest 1 (Fig. 7) demonstrates the behavior typically seen in our simulations in that convergence is even observed before the weights are fixed at the start of the production portion of the simulation. While this is not a precise determination of convergence, it is useful for identifying simulations which may require further simulation time. Since nearly all free energies appear to reach steady values in both forward and reverse directions before the production period (Figures 3–8 in Online Resource 1), additional simulation time was only suggested for the CB7 guest 3 system using this approach. Thus, while useful as a quick visual check, this method may not identify difficulties in convergence such as observed for guests 9 and 10 (see Table 1, Fig. 8).

An alternative method of monitoring convergence is to observe the behavior of the weights over time during the Wang–Landau process. Because we use the fully coupled state as the reference state in setting the weights, the weight corresponding to the uncoupled state is an estimate of the free energy change of the entire process and is expected to converge to a value very close to that computed by MBAR using data from the period after which the Wang–Landau weights are no longer adjusted.

For nearly all guests of both the CB7 and OA hosts, the self-consistently determined weight of the uncoupled state and the free energy computed with MBAR agree within 0.8 k_BT. This is not generally within the statistical uncertainty of MBAR, nor is it supposed to be because of the approximations in the Wang–Landau process, but is still in general agreement, and will lead to near even sampling of states during the production stage of the simulations. The exceptions to the agreement between the Wang–Landau derived weights and the MBAR analysis are for guest 3 in complex and guest 10 in both solvent and complex (see Fig. 8). The differences between final Wang–Landau weights and the MBAR analysis for guest 10 are 1.8 and 2.2 k_BT different for complex and solvent, which represents only around a 0.5 % error due to the large magnitude of the free energy differences associated with the high charge (+3) on guest 10.

In contrast, disappearing guest 3 in complex results in a 4 % difference between the weight determined from the Wang–Landau process and that calculated from MBAR. This likely stems from a significant change in intermediate state and configurational (see "Analysis of configurational sampling" section) sampling during the production period, which is the fixed weight portion of the simulation that is passed to MBAR. Highly uneven sampling along the coupling coordinate is also observed in nearly all expanded simulations disappearing the host, and in some cases

resulted in a single state being sampled 900 more frequently than the least sampled state. Although the host is rigid, removing the large complex from the water is a slower process than removing the much smaller guest. That this difficulty is observed in the expanded ensemble simulation of the removal of guest 3 and not with the other ligands appears to be due to the guest's high flexibility and multiple binding modalities (see "Analysis of configurational sampling" section). The high number of guest 3 degrees of freedom is also the likely cause of the high variability and slow convergence of the free energy of solvation for guest 3 (Fig. 8). In the case of disappearing the host, its large size likely contributes to sampling difficulties similarly to the flexibility of guest 3.

Sampling biased to a small number of intermediate states likely stems from saturation of the error in the WL weights, leading to sampling issues after active adjustment of their values to achieve even sampling is halted. Belardinelli et al. [21] demonstrates that this may be avoided in most cases by reducing the Wang–Landau incrementor by a factor of the Monte Carlo time instead of reduction of the weights by a constant multiplicative factor, as we have done in this work. An alternative methodology for dealing with this issue would be to compute new WL weights using MBAR, which is highly tolerant of uneven sampling. The new weights could then be used to run further fixed-weight simulations. While this process could be automated, it might still require a substantial number of iterations of this process, and thus extensive simulation time, before observing WL weight convergence. Regardless of the appropriate solution, our methods for checking convergence of WL weights over time provide a simple and quick visual test that seems to correlate well to inhomogeneities in configuration and coupling space sampling.

Conclusions

In general, our predicted absolute free energies of binding were too high for the SAMPL4 CB7 guests and too low for the OA guests. For the CB7 guests, relative free energies of binding were significantly closer to experiment, though still not particularly good, going from 8.96 to 3.58 kcal/mol with an improvement in RMSE of 5.38 kcal/mol for CB7, whereas the OA results got worse, going from 1.67 to 2.56 kcal/mol, because of a particularly bad prediction of the ligand designated as the reference ligand. Empirical fit parameters using a training set of known binding affinities to linearly correct the van der Waals and Coulombic free energies separately for CB7 ligands were found to be relatively non-transferable, suggesting high chemical heterogeneity in key interactions with this host. This would also imply strong dependence of the free energies of binding on the accuracy of the force field used and the solvent conditions modeled. These inaccuracies may contribute to lack of key cation interactions with the oxygen portal of the host, which could lead to our over-prediction of CB7 guest affinities [7, 30].

We further demonstrate that differences in our estimates from experiment mainly lie in the choice of model, including force field parametrization and environmental conditions such as salt concentration and pH, by examining the extent of convergence in our simulations with a variety of methods. Close to identical free energies of binding were obtained through alternative free energy simulation methodologies and through an alternative alchemical pathway in which the host was disappeared instead of the ligand. Conformational clustering reveals that in most cases we obtain fairly good convergence in conformational space for expanded ensemble simulations. However, when large kinetic barriers exist between relevant configurations, simulations at fixed ligand interaction states struggle to overcome kinetic barriers, as seen with CB7 guest 10. Such problems are significantly less prominent with expanded ensemble simulations. In many of the simple cases examined here, this would likely not have been an issue. However a strong dependence, in some cases up to 5 kcal/mol, on the starting configuration suggests substantial care must be taken when setting up systems for fixed state simulation.

The convergence of cumulative free energy estimates from MBAR over the course of the simulation indicates that free energies generally converged before the production period. Similarly, weights of the uncoupled state after the Wang–Landau convergence process are found in nearly all guests to closely approach final MBAR free energies computed over simulations with fixed states. We found free energies of binding for guest 3 to be the most difficult to converge for both solvation and decoupling from the complex, likely due to the high number of degrees of configurational freedom that guest 3 contains. Examining the convergence of simulations via a range of methods provides high confidence in the predicted free energies for the chosen set of force field parameters and environmental conditions. The fact that all assessments of simulation quality tended to agree indicates that simple comparison of the weights with the free energies calculated from MBAR, or even inspection of the rate of convergence of weights, may be useful tools for quickly checking free energy convergence in expanded ensemble free energy calculations. Such checks will prove useful in future free energy studies, and will also allow for confidence in differences observed while experimenting with the use of various force fields.

Acknowledgments The authors would like to thank the NanoSTAR Institute at the University of Virginia for an undergraduate research

grant, David Mobley (UC-Irvine) for his patience and strong leadership in the SAMPL4 competition, OpenEye for sponsorship of the SAMPL4 competition, and Lyle Isaacs (University of Maryland-College Park) and Mike Gilson (UC-San Diego) for their work in preparing the host–guest systems for SAMPL4.

References

1. Mobley DL, Dill KA (2007) Confine-and-release method: obtaining correct binding free energies in the presence of protein conformational change. J Chem Theory Comput 3:1231–1235
2. Gallicchio E, Lapelosa M, Levy RM (2010) Binding energy distribution analysis method (BEDAM) for estimation of protein–ligand binding affinities. J Chem Theory Comput 6:2961–2977
3. Boyce SE, Mobley DL, Rocklin GJ, Graves AP, Dill Ka, Shoichet BK (2009) Predicting ligand binding affinity with alchemical free energy methods in a polar model binding site. J Mol Biol 394(4):747–63
4. Jayachandran G, Shirts MR, Park S, Pande VS (2006) Parallelized-over-parts computation of absolute binding free energy with docking and molecular dynamics. J Chem Phys 125(8):084,901
5. Isaacs L (2009) Cucurbit[n]urils: from mechanism to structure and function. Chem Commun (6):619–29
6. Sun H, Gibb CLD, Gibb BC (2008) Calorimetric analysis of the 1:1 complexes formed between a water-soluble deep-cavity cavitand, and cyclic and acyclic carboxylic acids. Supramol Chem 20(1–2):141–147
7. Ong W, Kaifer AE (2004) Salt effects on the apparent stability of the cucurbit[7]uril-methyl viologen inclusion complex. J Org Chem 69(4):1383–5
8. Moghaddam S, Yang C, Rekharsky M, Ko YH, Kim K, Inoue Y, Gilson MK (2011) 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. J Am Chem Soc 133(10):3570–81
9. Moghaddam S, Inoue Y, Gilson MK (2009) Host–guest complexes with protein–ligand-like affinities: computational analysis and design. J Am Chem Soc 131(11):4012–21
10. Wyman IW, Macartney DH (2008) Cucurbit[7]uril host–guest complexes with small polar organic guests in aqueous solution. Org Biomol Chem 6(10):1796–801
11. Muddana HS, Fenley AT, Mobley DL, Gilson MK (2014) Blind prediction of the host–guest binding affinities from the SAMPL4 challenge. J Comput Aided Mol Des (in press)
12. Lyubartsev AP, Martsinovski AA, Shevkunov SV, Vorontsov-Velyaminov PN (1992) New approach to Monte Carlo calculation of the free energy: method of expanded ensembles. J Chem Phys 96(3):1776
13. Escobedo Fa, Martínez-Veracoechea FJ (2007) Optimized expanded ensembles for simulations involving molecular insertions and deletions. I. Closed systems. J Chem Phys 127(17):174103
14. Desgranges C, Delhommelle J (2012) Evaluation of the grand-canonical partition function using expanded Wang–Landau simulations. I. Thermodynamic properties in the bulk and at the liquid-vapor phase boundary. J Chem Phys 136(18):184107
15. Wang F, Landau D (2001) Efficient, multiple-range random walk algorithm to calculate the density of states. Phys Rev Lett 86(10):2050–2053
16. Wang K, Yang Y, Chodera JD, Shirts MR (2013) Identifying ligand binding sites and poses using GPU-accelerated Hamiltonian replica exchange molecular dynamics. J Comput Aid Mol Des 12(27):989–1007
17. Mobley DL (2012) Let's get honest about sampling. J Comput Aided Mol Des 26(1):93–5
18. Flyvbjerg H, Petersen HG (1989) Error estimates on averages of correlated data. J Chem Phys 91(1):461
19. Hess B (2002) Convergence of sampling in protein simulations. Phys Rev E 65(3):1–10
20. Grossfield A, Zuckerman DM (2009) Quantifying uncertainty and sampling quality in biomolecular simulations. Ann Rep Comput Chem 5:23–48
21. Belardinelli RE, Manzi S, Pereyra VD (2008) Analysis of the convergence of the 1/t and Wang–Landau algorithms in the calculation of multidimensional integrals. Phys Rev E 78:067701
22. da Silva AWS, Vranken WF (2012) Acpype—antechamber python parser interface. BMC Res Notes 5:367
23. Wang J, Wang W, Kollman PA, Case DA (2006) Automatic atom type and bond type perception in molecular mechanical calculations. J Mol Graph Model 25:247–260
24. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA (2004) Development and testing of a general amber force field. J Comput Chem 25:1157–1174
25. Jakalian A, Jack DB, Bayly CI (2002) Fast, effcient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. J Comput Chem 23(16):1623–1641
26. Humphrey W, Dalke A, Schulten K (1996) VMD—visual molecular dynamics. J Mol Graph 14:33–38
27. Hess B, Kutzner C, Spoel DVD, Lindahl E (2008) Gromacs 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. J Chem Theory Comput 4:435–447
28. Pronk S, Páll S, Schulz R, Larsson P, Bjelkmar P, Apostolov R, Shirts MR, Smith JC, Kasson PM, van der Spoel D, Hess B, Lindahl E (2013) GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. Bioinformatics 29(7):845–54
29. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. J Chem Phys 79:926–935
30. Cao L, Isaacs L (2013) Absolute and relative binding affinity of cucurbit[7]uril towards a series of cationic guests. Supramol Chem. doi:10.1080/10610278.2013.852674
31. Liu DCL, Nocedal J (1989) On the limited memory method for large scale optimization. Math Program B 45(3):503–528
32. Martyna GJ, Tuckerman ME, Tobias DJ, Klein ML (1996) Explicit reversible integrators for extended systems dynamics. Mol Phys 87:1117–1157
33. Ryckaert JP, Ciccotti G, Berendsen HJ (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J Comput Phys 23(3):327–341
34. Andersen C (1983) RATTLE: a "velocity" version of the SHAKE algorithm for molecular dynamics calculations. J Comput Phys 52:24–34
35. Chodera JD, Shirts MR (2011) Replica exchange and expanded ensemble simulations as Gibbs sampling: simple improvements for enhanced mixing. J Chem Phys 135(19):194110
36. Paliwal H, Shirts MR (2011) A benchmark test set for alchemical free energy transformations and its use to quantify error in common free energy methods. J Chem Theory Comput 7(12):4115–4134
37. Chodera JD, Shirts MR (2009) A python implementation of the multistate Bennet acceptance ratio (MBAR). https://simtk.org/home/pymbar
38. Shirts MR, Chodera JD (2008) Statistically optimal analysis of samples from multiple equilibrium states. J Chem Phys 129(12):124105
39. Boresch S, Tettinger F, Leitgeb M, Karplus M (2003) Absolute binding free energies: a quantitative approach for their calculation. J Phys Chem B 107(35):9535–9551
40. Deng Y, Roux B (2006) Calculation of standard binding free energies: aromatic molecules in the T4 lysozyme L99A mutant. J Chem Theory Comput 2(5):1255–1273

41. Wang J, Deng Y, Roux B (2006) Absolute binding free energy calculations using molecular dynamics simulations with restraining potentials. Biophys J 91(8):2798–814
42. Hunenberger PH, McCammon JA (1999) Ewald artifacts in computer simulations of ionic solvation and ionion interaction: a continuum electrostatics study. J Chem Phys 110(4):1856
43. Rocklin GJ, Mobley DL, Dill Ka, Hünenberger PH (2013) Calculating the binding free energies of charged species based on explicit-solvent simulations employing lattice-sum methods: an accurate correction scheme for electrostatic finite-size effects. J Chem Phys 139(18):184103
44. Rogers KE, Ortiz-Sánchez JM, Baron R, Fajer M, de Oliveira CAF, McCammon JA (2013) On the role of dewetting transitions in host–guest binding free energy calculations. J Chem Theory Comput 9(1):46–53
45. Yang W, Bitetti-Putzer R, Karplus M (2004) Free energy simulations: use of reverse cumulative averaging to determine the equilibrated region and the time required for convergence. J Chem Phys 120(6):2618–28