

UC San Diego

UC San Diego Previously Published Works

Title

On the Role of Dewetting Transitions in Host–Guest Binding Free Energy Calculations

Permalink

<https://escholarship.org/uc/item/2bn3g360>

Journal

Journal of Chemical Theory and Computation, 9(1)

ISSN

1549-9618

Authors

Rogers, Kathleen E

Ortiz-Sánchez, Juan Manuel

Baron, Riccardo

et al.

Publication Date

2013-01-08

DOI

10.1021/ct300515n

Peer reviewed

# On the Role of Dewetting Transitions in Host–Guest Binding Free Energy Calculations

Kathleen E. Rogers,<sup>\*,†,⊥</sup> Juan Manuel Ortiz-Sánchez,<sup>†,‡</sup> Riccardo Baron,<sup>\*,†,‡,‡,‡</sup> Mikolai Fajer,<sup>†,‡</sup> César Augusto F. de Oliveira,<sup>\*,†,‡,§</sup> and J. Andrew McCammon<sup>†,‡,§,||</sup>

<sup>†</sup>Center for Theoretical Biological Physics, University of California San Diego, La Jolla, California 92093-0365, United States

<sup>‡</sup>Department of Chemistry & Biochemistry, University of California San Diego, La Jolla, California 92093-0365, United States

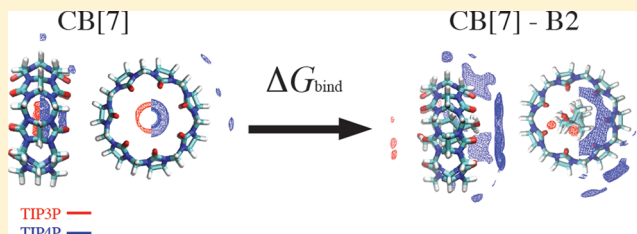
<sup>§</sup>Howard Hughes Medical Institute, University of California San Diego, La Jolla, California 92093-0365, United States

<sup>||</sup>Department of Pharmacology, University of California San Diego, La Jolla, California 92093-0365, United States

<sup>⊥</sup>Biomedical Sciences Graduate Program, University of California San Diego, La Jolla, California 92093, United States

## Supporting Information

**ABSTRACT:** We use thermodynamic integration (TI) and explicit solvent molecular dynamics (MD) simulation to estimate the absolute free energy of host–guest binding. In the unbound state, water molecules visit all of the internally accessible volume of the host, which is fully hydrated on all sides. Upon binding of an apolar guest, the toroidal host cavity is fully dehydrated; thus, during the intermediate  $\lambda$  stages along the integration, the hydration of the host fluctuates between hydrated and dehydrated states. Estimating free energies by TI can be especially challenging when there is a considerable difference in hydration between the two states of interest. We investigate these aspects using the popular TIP3P and TIP4P water models. TI free energy estimates through MD largely depend on water-related interactions, and water dynamics significantly affect the convergence of binding free energy calculations. Our results indicate that wetting/dewetting transitions play a major role in slowing the convergence of free energy estimation. We employ two alternative approaches—one analytical and the other empirically based on actual MD sampling—to correct for the standard state free energy. This correction is sizable (up to 4 kcal/mol), and the two approaches provide corrections that differ by about 1 kcal/mol. For the system considered here, the TIP4P water model combined with an analytical correction for the standard state free energy provides higher overall accuracy. This observation might be transferable to other systems in which water-related contributions dominate the binding process.



## INTRODUCTION

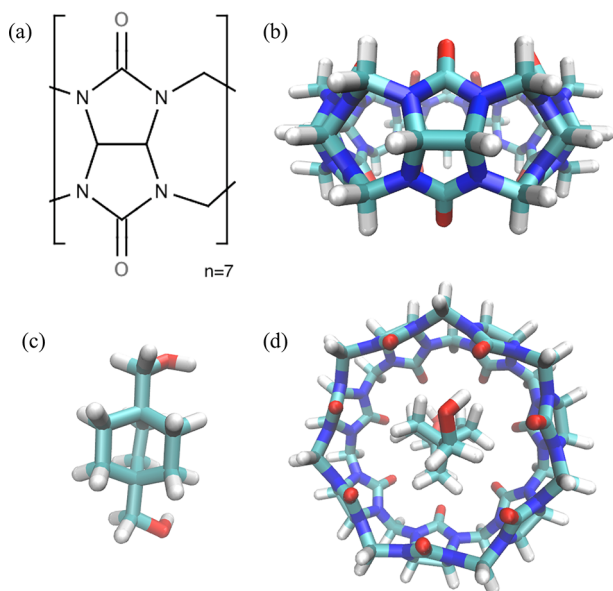
Cucurbit[7]uril (CB[7]; Figure 1) is a synthetic host that is attracting increasing interest for its ability to selectively bind various neutral or positively charged aromatic guests and metal complexes in aqueous solution.<sup>1,2</sup> CB[7] can achieve affinities with some cationic guests that are greater than those typically measured for protein–ligand complexes.<sup>3,4</sup> With their recognition properties and ease of synthesis, members of the CB[n] family have a wide range of potential applications including catalysis, gas purification and waste-stream remediation, crystal engineering, self-assembling and self-sorting systems, molecular machines, supramolecular polymers, self-assembled monolayers, and gene transfection.<sup>5</sup> As a host for metal complexes, CB[7] has also shown promise as a drug carrier for platinum chemotherapeutics such as Oxaliplatin by improving stability and decreasing negative side effects of the drug.<sup>6</sup> This auspicious molecular carrier demonstrates low toxicity, efficient cellular internalization, and delivered drug bioactivity nearly as effective as that of the unbound drug in some cases,<sup>7,8</sup> suggesting that it could be employed as a sophisticated drug delivery system.<sup>9</sup> In order to better utilize CB[7] in these

applications, it is important to understand its behavior within a water environment and its molecular recognition of possible guests. Recent work on CB[7] describes its molecular dynamics<sup>10</sup> and affinity for several synthetic ferrocene and bicyclo[2.2.2]octane based guest molecules (e.g., B2 shown in Figure 1c) both experimentally and computationally.<sup>11–13</sup> Even in seemingly simple, small host–guest systems with relatively rigid and symmetrical structures such as these (see Figure 1d), estimating the binding free energy can be quite challenging. However, to our knowledge no study to date investigated the wetting/dewetting events of the CB[7] system and the relevance of hydration in host–guest binding.

The importance of wetting/dewetting transitions in receptor or model host cavities has recently received much attention already.<sup>14–18</sup> One of the most important consequences of these studies was the demonstration that noncovalent binding can largely depend on water-related contributions, not—as sometimes assumed—primarily on direct host–guest interaction. In

Received: June 20, 2012

Published: October 26, 2012



**Figure 1.** Host and guest molecules studied. (a) Chemical representation of glycoluril repeat unit constituting the host. (b) Side view of the [ $n = 7$ ] nonpolar host cucurbit[7]uril (CB[7]) structure. (c) B2 nonpolar guest molecule. (d) Top view of CB[7]-B2 host-guest complex in the structure used to initialize free energy calculations.

particular, while the dehydration of the ligand molecules was shown to be the driving factor for binding, removal of solvent fluctuations can result in entropic penalties that are sizable compared with the binding free energy.<sup>14,16</sup> However, our understanding of solvation effects and their thermodynamic role in molecular recognition of more realistic systems remains quite poor.<sup>15,18–20</sup>

Here, we study CB[7] hydration dynamics and its role upon binding of B2, a recently designed bicyclo[2.2.2]octane apolar guest.<sup>21</sup> We investigate the dependence of hydration properties using the popular TIP3P and TIP4P water models<sup>22</sup> with thermodynamic integration (TI) and explicit solvent molecular dynamics (MD) simulation. Overall, the closest match of estimated host-guest absolute binding free energy against experimental data (within 0.3 kcal/mol) is achieved using the TIP4P water model. However, both TIP3P and TIP4P water models lead to qualitatively similar host hydration in the bound and unbound states, and along the unphysical  $\lambda$  state intermediates. For the unphysical  $\lambda$  points at which significant host wetting/dewetting transitions occur, we observe remarkably large  $(\partial V)/(\partial \lambda)$  fluctuations along the simulation time, ranging up to 140 kcal/mol and 130 kcal/mol for TIP3P and TIP4P water models, with standard deviations up to 22 kcal/mol and 19 kcal/mol, respectively. Furthermore, these fluctuations directly correlate with changes in host hydration, indicating that wetting/dewetting effects play a major role in the host-guest binding processes and, as a result, in determining free energy estimates. The work presented herein underscores the importance of water-related contributions and their convergence in free energy calculations of noncovalent binding.

## COMPUTATIONAL DETAILS

**Molecular Models.** Initial coordinates of the host-guest complexes<sup>21</sup> were previously generated with the Vdock<sup>23</sup>

program from the CB[7] crystal structure.<sup>6,24</sup> The general Amber force field (GAFF) was used to describe all intramolecular energy contributions.<sup>25,26</sup> The partial charges of CB[7] and the bicyclo[2.2.2]octane (B2) were calculated using the RESP program.<sup>27,28</sup> To ensure symmetry, charge equivalence was enforced on each one of the seven units of the cucurbituril host (CB[7]). The molecular electrostatic potential was calculated at the HF/6-31G\* level. Lennard-Jones parameters for B2 and CB[7] were assigned as previously described by Moghaddam et al.<sup>12</sup> These host-guest complexes were solvated in cubic boxes with a buffer region of 12 Å.

**Molecular Dynamics Simulations.** All simulations were performed using a modified version of AMBER<sup>29,30</sup> that enables the calculation of absolute binding free energies with restraining and soft-core potentials. For both B2 and B2-CB[7], decoupling of the guest partial charges (ele) occurred first at 11  $\lambda$  values (0.0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0). Decoupling of the van der Waals (vdW) energy terms utilized 19 different  $\lambda$  values (0.0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.65, 0.7, 0.75, 0.775, 0.8, 0.825, 0.85, 0.875, 0.9, 0.925, 0.95, 1.0). All simulations were performed using rectangular periodic boundary conditions with isotropic position scaling in the isothermic-isobaric (NPT) ensemble, with a pressure reference of 1 atm and a relaxation time of 0.5 ps. The system was kept at the reference temperature of 300 K using the Langevin thermostat<sup>31</sup> (collision frequency of 20 ps<sup>-1</sup>), and Newton's equations of motion were integrated using the leapfrog algorithm<sup>32</sup> with a MD time step of 0.001 ps. Using the *sander* Amber module, long-range electrostatics interactions were handled with the particle-mesh Ewald (PME) procedure and long-range van der Waals interactions approximated by a continuum model.<sup>29</sup> We refer the reader to ref 29 for additional computational details.

**Thermodynamic Integration.** The change in free energy between two states, A and B, can be estimated using standard thermodynamic integration (TI)<sup>33</sup> as

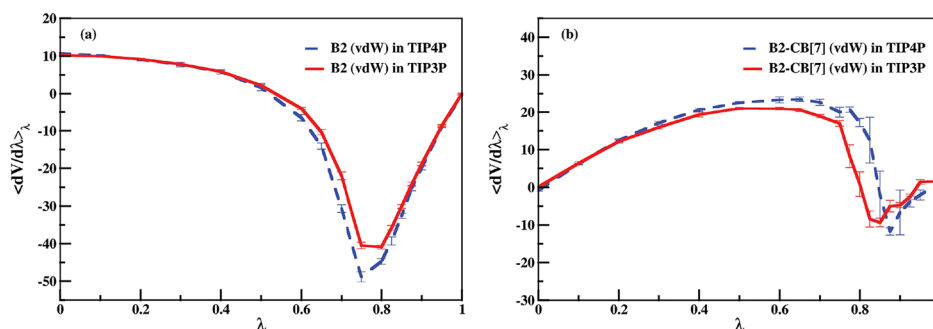
$$\Delta G_{A \rightarrow B} = \int_{\lambda_A}^{\lambda_B} d\lambda \left\langle \frac{\partial V(\lambda)}{\partial \lambda} \right\rangle \quad (1)$$

with  $V(\lambda)$  representing the potential energy from a single trajectory of the system as a function of the coupling parameter  $\lambda$ , and with  $\langle \dots \rangle$  denoting the cumulative backward average at the given  $\lambda$  point. The thermodynamic perturbation in this work went from state A ( $\lambda = 0$ ), in which the guest (G) experiences full interactions with the host (H) and the water environment, to state B ( $\lambda = 1$ ), in which the guest-related interaction energy vanishes. In order to improve phase-space sampling and avoid free energy singularities, the soft-core potential of Zacharias et al.<sup>30,34</sup> was employed for all guest atoms ( $scalpha = 0.5$ ).<sup>29,35</sup> Equation 1 was integrated numerically using the trapezoidal rule.

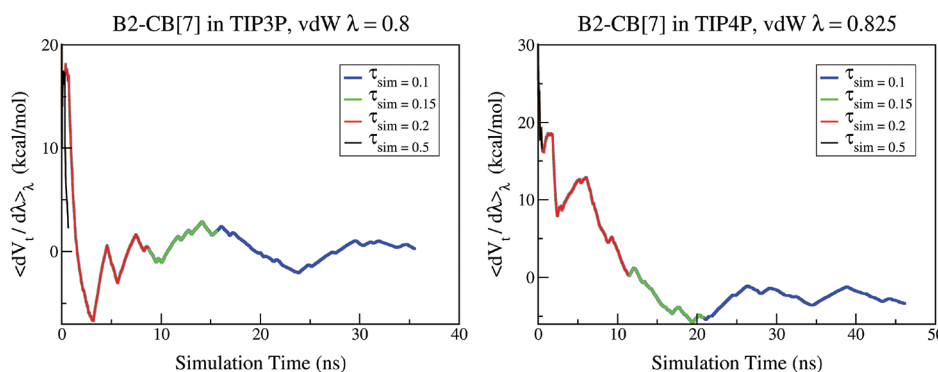
**Statistical Analysis of Uncertainties.** A simulation standard error  $\sigma_{\text{sim}}(t)$  of the time-varying potential energy derivative at a given  $\lambda$  can be calculated as

$$\sigma_{\text{sim}}(t) = \sqrt{\frac{1}{T-1} \sum_{t=1}^T \left[ \left( \frac{\partial V_t(\lambda)}{\partial \lambda} \right)_\lambda - \left\langle \frac{\partial V_t(\lambda)}{\partial \lambda} \right\rangle_\lambda \right]^2} / \sqrt{T} \quad (2)$$

with  $T$  being the total number of block averages<sup>36</sup> throughout the single  $i$ th trajectory or all  $N$  concatenated independent trajectories.  $((\partial V_t(\lambda))/(\partial \lambda))_\lambda$  denotes the potential energy



**Figure 2.** Thermodynamic integration along the  $\lambda$  coupling parameter. vdW-decoupling TI curves of (a) guest and (b) host–guest system simulations in TIP3P (red) and TIP4P (blue).  $\langle \partial V(\lambda) / \partial \lambda \rangle_{\lambda}$  values and corresponding uncertainties,  $\sigma_{\text{sim}}(t)$ , are shown at each  $\lambda$  sampled.



**Figure 3.** Convergence monitoring of  $\langle \partial V(\lambda) / \partial \lambda \rangle_{\lambda}$ , the potential energy derivative backward average, across time with different criteria,  $\tau_{\text{sim}}$ , during B2-CB[7] vdW decoupling simulations where the highest  $\partial V / \partial \lambda$  fluctuations occur using either TIP3P (left) or TIP4P (right) water models ( $\lambda = 0.8$  and  $\lambda = 0.825$ , respectively).

**Table 1. Absolute Host–Guest Binding Free Energy Estimates<sup>a</sup>**

water model	system	$\Delta G_{\text{ele}}$ [kcal/mol]	$\Delta G_{\text{vdW}}$ [kcal/mol]	correction term for potential bias from host–guest restraint		$\Delta G_{\text{calcd}}^{\text{c}}$ [kcal/mol] <sup>e</sup>		$\Delta G_{\text{exptl}}^{\text{d12}}$ [kcal/mol]	
				empirical <sup>b</sup> (avg) <sup>c</sup>	analytical <sup>d</sup>	empirical	analytical		
TIP3P	B2	$-21.6 \pm 0.22$	$-4.87 \pm 0.44$	Volume [ $\text{\AA}^3$ ]	$-6.4 (0.076)$	n/a	$-12.0 \pm 0.9$	$-11.2 \pm 0.9$	$-13.4 \pm 0.1$
	B2-CB[7]	$-22.4 \pm 0.23$	$11.2 \pm 0.77$	correction [kcal/mol]	$-3.3 (-6.0)$	$-4.1$			
TIP4P	B2	$-20.8 \pm 0.33$	$-6.61 \pm 0.60$	Volume [ $\text{\AA}^3$ ]	$-6.2 (0.071)$	n/a	$-14.4 \pm 1.0$	$-13.7 \pm 1.0$	
	B2-CB[7]	$-22.0 \pm 0.34$	$12.3 \pm 0.82$	correction [kcal/mol]	$-3.3 (-6.0)$	$-4.1$			

<sup>a</sup>Uncertainties were calculated as in eq 3. See also Figure.SI-1 and Figure 2 for raw  $\Delta G_{\text{ele}}$  and  $\Delta G_{\text{vdW}}$  data, respectively. <sup>b</sup>From eq 5, using maximum  $r_{\text{G}}-r_{\text{H}}$  center of mass (COM)-COM distance (n/a for B2 alone). <sup>c</sup>From eq 5 using average  $r_{\text{G}}-r_{\text{H}}$  center of mass (COM)-COM distance (n/a for B2 alone). <sup>d</sup>From eq 6; no restraint used for B2. <sup>e</sup>Absolute binding free energy calculation derived from thermodynamic integration cycle (decoupled partial charges and decoupled van der Waals terms) with correction for standard state:  $(\Delta G_{\text{B2\_ele}} + \Delta G_{\text{B2\_vdW}}) - (\Delta G_{\text{B2-CB[7]_ele}} + \Delta G_{\text{B2-CB[7]_vdW}} + \text{B2-CB[7] correction term for potential bias from restraint})$ .

derivative, block-averaged at time  $t$ , and  $\langle (\partial V_T(\lambda)) / (\partial \lambda) \rangle_{\lambda}$  is the ensemble average over the entire simulation time at a given  $\lambda$ . As an example,  $\sigma_{\text{sim}}(t)$  uncertainties are reported as error bars for  $\langle (\partial V_T(\lambda)) / (\partial \lambda) \rangle_{\lambda}$  vs  $\lambda$  in Figure 2 (vertical bars). Then, a corresponding free-energy uncertainty can be obtained as

$$\sigma = \left( \int_{\lambda_A}^{\lambda_B} \sigma_{\text{sim}}^2(t) d\lambda \right)^{1/2} \quad (3)$$

We define a criterion,  $\tau_{\text{sim}}$ , for convergence monitoring of all TI simulations and automated decisions for termination of individual  $\lambda$  runs. This criterion is based on the backward cumulative averages of the  $\langle (\partial V_t(\lambda)) / (\partial \lambda) \rangle_{\lambda}$  terms. This reads

$$\tau_{\text{sim}}(T-t) = \frac{1}{\sqrt{T-t}} \sqrt{\frac{1}{T-t-1} \sum_{t=T}^{T-t} \left[ \left( \frac{\partial V'_T(\lambda)}{\partial \lambda} \right)_{\lambda} - \left\langle \frac{\partial V'_{T-1}(\lambda)}{\partial \lambda} \right\rangle_{\lambda} \right]^2} \quad (4)$$

Examples of the variation of this quantity backward in simulation time are given in Figure 3 and Table SI-1. A criterion of  $\tau_{\text{sim}} < 0.1$  was employed to enforce proper convergence and automatically decide when to terminate TI runs. The term proposed in eq 4 is a well-behaved quantity that tends to zero for increasingly long trajectories.

**Separation of Thermodynamic States.** Corrections for the standard state free energy<sup>37,38</sup> and separation of the

thermodynamic states of interest were obtained using a harmonic restraining potential between the guest center of mass,  $r_G$ , and the host center of mass,  $r_H$ , to confine the guest to the region within the host cavity  $V_{\text{cavity}}$ , with a force constant  $k_h$  of 2.5 kcal/mol. This  $V_{\text{cavity}}$  was either estimated empirically from a 10 ns MD trajectory, according to

$$V_{\text{cavity}} = \frac{4}{3}\pi(r_G - r_H)^3 \quad (5)$$

or defined analytically as<sup>39</sup>

$$V_{\text{cavity}} = \left( \frac{2\pi RT}{k_h} \right)^{3/2} \quad (6)$$

for comparison. Equation 5 or 6 was then utilized to find the standard-state free energy

$$\Delta G_{\text{Bind}}^{\circ} = \int_{\lambda_A}^{\lambda_B} d\lambda \left\langle \frac{\partial V(\lambda)}{\partial \lambda} \right\rangle + RT \ln \left( \frac{V_{\text{cavity}}}{V^{\circ}} \right) \quad (7)$$

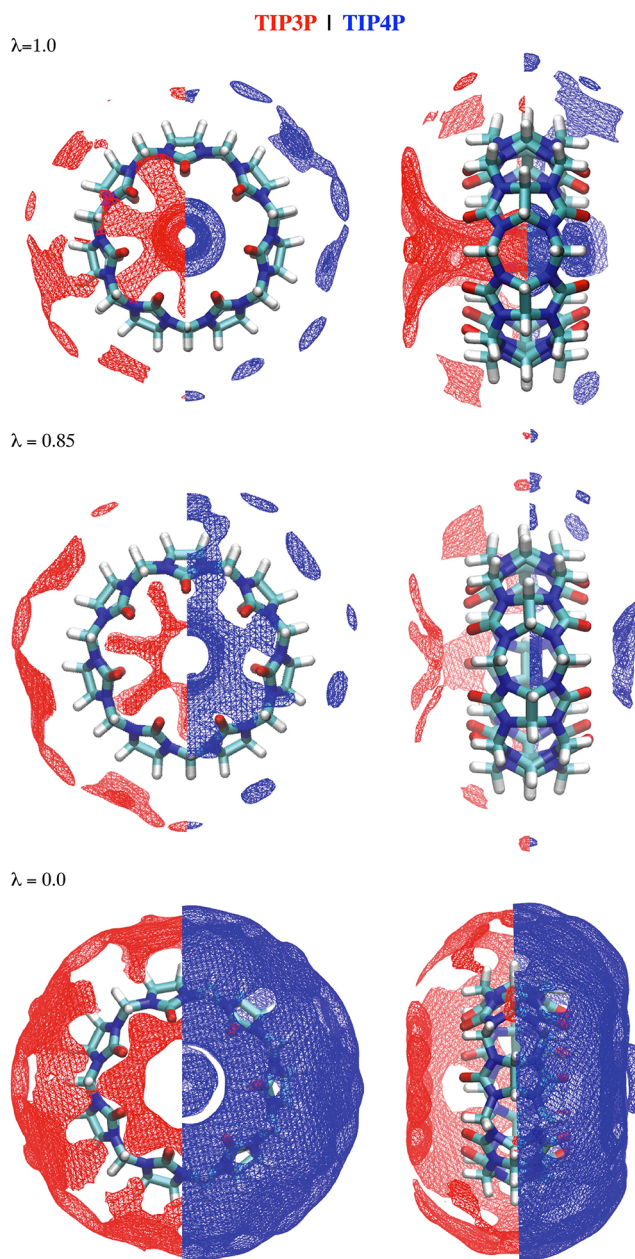
for guest transfer from this restricted  $V_{\text{cavity}}$  volume to the bulk volume  $V^{\circ}$  (as described in ref 40).

## RESULTS AND DISCUSSION

In Figure 2, we include the following  $\langle (\partial V_T(\lambda))/(\partial \lambda) \rangle_{\lambda}$  vs  $\lambda$  TI curves for the van der Waals decoupling steps: (i) solvated guest vdW decoupled along  $\lambda$  in the unbound state (Figure 2a) and (ii) the solvated guest vdW decoupled along  $\lambda$  in the bound state (Figure 2b). Although  $\langle (\partial V)/(\partial \lambda) \rangle$  at most  $\lambda$  values shows very good convergence (see also Supporting Information Figure SI-1), we observe a region of  $\lambda$  space from 0.7 to 0.875 with noticeably higher uncertainty and where convergence is considerably hindered within the host–guest vdW decoupling simulations. This narrow region is also where the  $\langle (\partial V_T(\lambda))/(\partial \lambda) \rangle_{\lambda}$  vs  $\lambda$  TI curve changes most dramatically in the B2-CB[7] vdW transformation (Figure 2b).

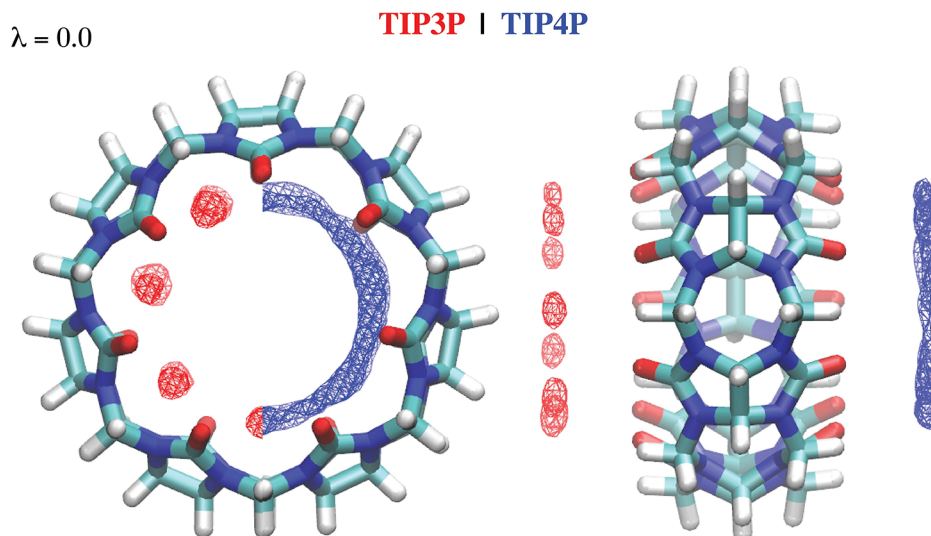
In order to observe the dependence of  $\langle (\partial V_T(\lambda))/(\partial \lambda) \rangle$  on total simulation time, we use a convergence criterion based on backward block-averaging. Figure 3 depicts the simulation convergence monitoring used in this work (see Computational Details) at the  $\lambda$  points with the largest statistical uncertainty for either TIP3P or TIP4P. These results further confirm that the convergence of  $\langle (\partial V_T(\lambda))/(\partial \lambda) \rangle_{\lambda}$  is considerably impeded at  $\lambda$  values within this region of higher uncertainty (e.g.  $\lambda = 0.8$  for TIP3P and  $\lambda = 0.825$  for TIP4P). Interestingly, the shape of the TI curves for the vdW decoupling simulations display remarkable dependence on the water model employed (Figure 2). The large dip in  $\langle (\partial V_T(\lambda))/(\partial \lambda) \rangle_{\lambda}$  is noticeably dissimilar in both width and location for the B2-CB[7] trajectories (Figure 2b).

Nevertheless, integration of each CB[7]-B2 bound vdW decoupling curve produces comparable differences in free energy contribution for TIP3P and TIP4P ( $\Delta G_{\text{vdW}}$  of 11.2 and 12.3 kcal/mol, respectively). Similar results were also obtained for the unbound B2 guest, with a  $\Delta G_{\text{vdW}}$  of  $-4.87$  kcal/mol and  $-6.61$  kcal/mol for TIP3P and TIP4P, respectively (Table 1). Figure SI-1 displays the TI curves calculated from the electrostatic decoupling steps. Not surprisingly, the electrostatic contribution of free energy change between the two water models is even smaller ( $\Delta G_{\text{ele}}$  estimates for the unbound state in TIP3P and TIP4P differ by 0.8 kcal/mol and those for bound state by only 0.4 kcal/mol).



**Figure 4.** Average hydration surfaces of 50 ns vdW host–guest simulations for unbound ( $\lambda = 1.0$ ; top) and bound ( $\lambda = 0.0$ ; bottom) thermodynamic states. Corresponding maps are also shown for the  $\lambda = 0.85$  intermediate unphysical state (middle). Hydration maps for TIP3P (red wireframe, left of vertical slice) and TIP4P (blue wireframe, right of vertical slice) water models were generated using the average density of the water oxygen atom on a grid (0.3 Å resolution) after superimposing MD snapshots using all atoms of the host (shown as licorice representation here). The surface isovalues represent water densities that are  $\sim 1.1$  times TIP4P bulk water density and  $\sim 0.95$  times TIP3P bulk water density.

To calculate the final free energy values according to eq 7, either an empirical or an analytical correction was used to take into account the standard state for  $\Delta G$ . The empirical volumes were estimated from the total spherical space of radius  $(r_G - r_H)$  sampled by the guest in the bound state simulations. By using the maximum distance between the guest center of mass (COM) and the host COM  $(r_G - r_H)$  in eq 5, the empirical correction energy was estimated to be  $-3.3$  kcal/mol for both



**Figure 5.** Average hydration surfaces (as in Figure 4) for bound host–guest vdW simulation ( $\lambda = 0.0$ ) thermodynamic state with TIP3P (red wireframe, left side) and TIP4P (blue wireframe, right side) water oxygen atom densities shown in and around the host structure (licorice representation) from topview (left) and sideview (right) for each at  $\sim 1.5$  times TIP4P bulk water density and  $\sim 1.3$  times TIP3P bulk water density.

TIP3P and TIP4P (Table 1). If instead we use the average  $r_G - r_H$  during sampling, we estimate a correction of  $-6.0$  kcal/mol using either water model (Table 1). Using eq 6, the calculated analytical restraining bias correction term for both water models was  $-4.1$  kcal/mol. We hope that the average  $r_G - r_H$  correction term would converge (with further sampling in the host–guest MD simulation) to the same value as the analytical correction term, but we believe that the difference in free energy values here might be attributable to underestimating the volume sampled by the guest using this average value and overestimating it using the maximum  $r_G - r_H$  sampled.

The absolute binding free energy estimate derived from simulations in TIP4P (best estimate:  $-13.7$  kcal/mol) is slightly closer to the experimental value ( $-13.4$  kcal/mol) than from those in TIP3P (best estimate:  $-12.0$  kcal/mol). Taken as a whole, our calculations with TIP4P in combination with either correction term and the parameters used for the CB[7]-B2 host–guest system provide a more accurate estimate of absolute binding free energy. Our results do suggest that the TIP4P water model may describe the binding thermodynamics more accurately compared with the TIP3P model by bringing our calculated value significantly closer to the experimental value, but further studies would be necessary to confirm this.

These results are in good agreement with a previous study that showed that the TIP4P model describes water properties significantly more accurately compared with the TIP3P water model.<sup>41</sup> Due to the important role of the solvation/desolvation of the CB[7] and B2 during the binding event, we decided to further investigate the relationship between slow simulation convergence and host hydration with both water models.

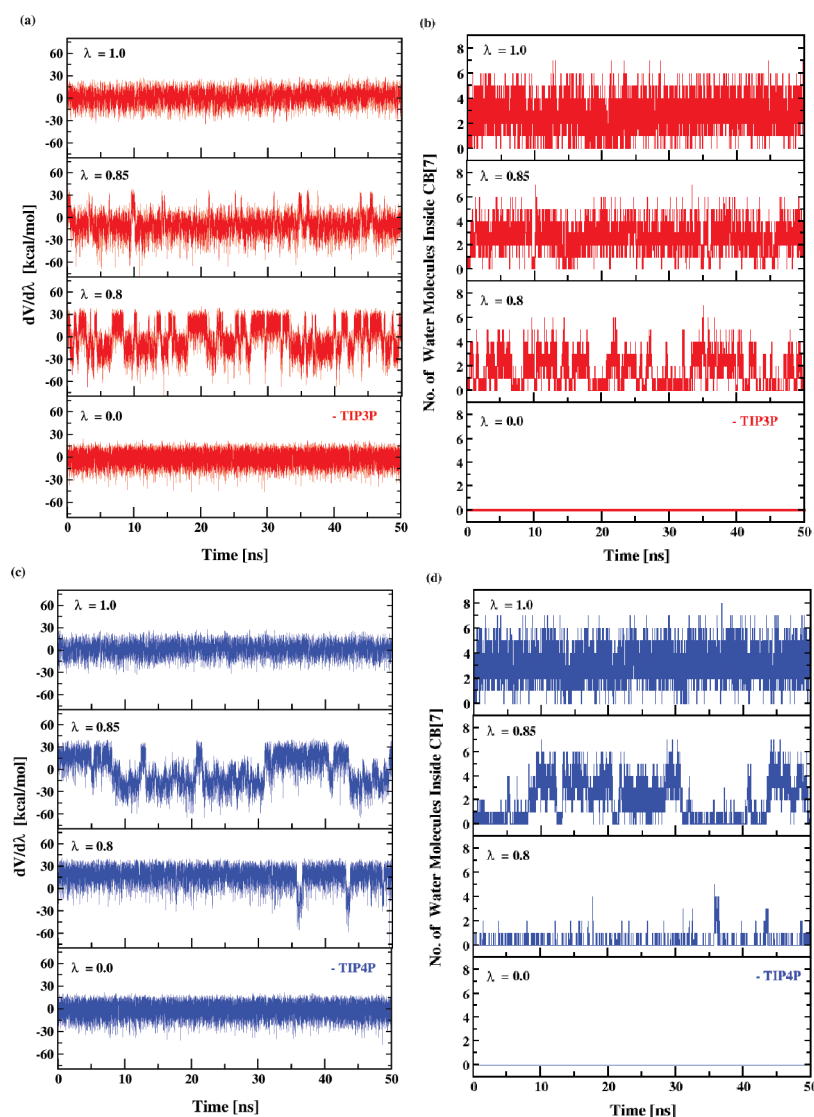
In Figures 4 and 5, we examine the average water density surfaces throughout selected trajectories with TIP3P and TIP4P. Because both host–guest water model systems reach their  $\langle (\partial V_T(\lambda)) / (\partial \lambda) \rangle_\lambda$  minimum on the vdW TI curve at  $\lambda = 0.85$  (Figure 2b), in Figure 4 we include the average water density surfaces at this point in addition to those for the

unbound ( $\lambda = 1.0$ ) and bound ( $\lambda = 0.0$ ) cases, with densities shown at  $\sim 0.95$  times bulk TIP3P and  $\sim 1.1$  times bulk TIP4P. In Figure 5, we display higher density water surfaces of the B2-bound case, with densities shown at  $\sim 1.3$  times TIP3P bulk water density and  $\sim 1.5$  times TIP4P bulk water density.

At  $\lambda = 1$ , the average water density from our 50 ns simulation was highest inside of the host. The average volume surfaces of the TIP4P water oxygen atoms are shown here as a few small clusters, around and between the nitrogen atoms of the host, and three toroidal shapes with the largest ring in the center of the host and two others near the top and bottom (Figure 4), which is consistent with other models of the host alone.<sup>42</sup> Interestingly, at this same density value, TIP3P waters display average surfaces that appear as more of a cylinder of ordered water oxygen atoms inside and more of a seven-petal flower shape outside the host but show some similar density clusters near the host nitrogen atoms.

Both water models still show similar ordering near the host nitrogen atoms at  $\lambda = 0.85$ , but we do observe some different volume density patterns (Figure 4). TIP4P waters still form ring-like average density surfaces, but the center toroid is smaller and those on the outside of the host become more concave than in the unbound ( $\lambda = 1$ ) case. Here, we also observe an additional concave flower-shaped surface of TIP3P waters at this density (as at  $\lambda = 1$ ) appearing on both ends of the host. However, while TIP3P water molecules do occupy the host at this  $\lambda$ , the surfaces do not appear inside the host at this high-density value, unlike the TIP4P and unbound TIP3P cases.

In the B2-bound host case ( $\lambda = 0.0$ ), again we see unique surface patterns, even between the two water models. The TIP4P surface almost completely covers the outside of the host, but TIP3P surfaces only appear as large clusters on the outside of the host and as fuller flower shapes near the host oxygen atoms (Figure 4). As expected though, no water densities are seen within the host cavity because the guest vdW parameters remain fully coupled in this case. Higher TIP4P density surfaces



**Figure 6.** Fluctuations of (a,c)  $dV/d\lambda$  values and (b,d) number of water molecules in the CB[7] host system along time for the TIP3P (red, a,b) and TIP4P (blue, c,d) water models during 50 ns host-guest vdW simulations. The physical process of host-guest binding goes from the unbound state ( $\lambda = 1$ ) to the bound state ( $\lambda = 0$ ), while absolute binding free energies were effectively calculated conveniently in the opposite direction by decoupling of the ligand interactions (from  $\lambda = 0$  to  $\lambda = 1$ ).

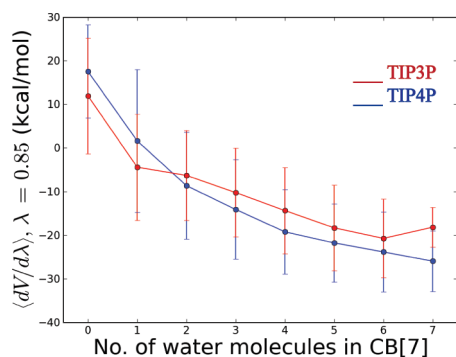
**Table 2. Host Hydration During the End Point and Unphysical States Employed for Thermodynamic Integration<sup>a</sup>**

host hydration features	water model	$\lambda = 0.0$	$\lambda = 0.775$	$\lambda = 0.8$	$\lambda = 0.825$	$\lambda = 0.85$	$\lambda = 0.875$	$\lambda = 1.0$
time hydrated (%)	TIP3P	0	29.5	59.0	87.8	93.8	96.7	97.0
	TIP4P	0	8.02	18.3	62.1	86.2	88.8	97.7
time to first max. hydration event (ns)	TIP3P	n/a	6.47	35.0	45.7	10.2	6.88	12.7
	TIP4P	n/a	21.0	3.40	2.26	2.02	2.00	1.44
mean no. of water molecules inside the host	TIP3P	0	$0.6 \pm 1$	$1.4 \pm 1$	$2.4 \pm 1$	$2.6 \pm 1$	$2.9 \pm 1$	$2.7 \pm 1$
	TIP4P	0	$0.1 \pm 0.5$	$0.4 \pm 1$	$1.8 \pm 2$	$3.0 \pm 2$	$3.3 \pm 2$	$3.0 \pm 1$
maximum no. of water molecules inside the host	TIP3P	0	6	7	7	7	7	7
	TIP4P	0	6	6	7	7	7	7
minimum no. of water molecules inside the host	TIP3P	0	0	0	0	0	0	0
	TIP4P	0	0	0	0	0	0	0

<sup>a</sup>Results are summarized for the  $\lambda$  states at which largest wetting/dewetting transitions occur and compared to end point states.

appear just as small clusters circling the host and as two toroids, one at each end of the host, while TIP3P surfaces form only small circular clusters along these outside rings but do not form solid toroids there yet (Figure 5). As expected here, where the guest occupies the host, still no water densities appear within its

cavity. The unique hydration maps observed for TIP4P and TIP3P models at the same, specific density value are expected, owing to both the different mixed interactions of these water models with the same host molecule as well as the well-known structural disparities between the water models themselves.



**Figure 7.** Correlation between  $\langle \partial V / \partial \lambda \rangle$  values and number of waters within host for  $\lambda = 0.85$  vdW simulation with TIP3P and TIP4P water models. Vertical error bars are the standard deviation values.

Figure 6 displays both the  $\langle \partial V / \partial \lambda \rangle$  values (Figure 6a,c) and the number of waters inside the host (Figure 6 b,d) along simulation time for several  $\lambda$  points. All waters within 4 Å of both the very top and bottom atoms of the B2 core were measured at each frame of the 50 ns vdW host–guest complex trajectories and compared to the  $\langle \partial V / \partial \lambda \rangle$  measured at that time point. Clearly, two distinctive energy extremes are represented by hydrated and dehydrated states (Figure 6). Frequent fluctuations can occur between the two hydration states when the guest vdW decoupling is sufficiently high to allow for increased host hydration, but still low enough that the host cannot remain stably hydrated throughout the rest of the simulation. The  $\langle \partial V / \partial \lambda \rangle$  of the host–guest system at several of the higher  $\lambda$ 's near 0.825 fluctuate between two very different hydration extremes with periods of several nanoseconds during vdW TI simulations (Figure 6). The frequency of wetting/dewetting transitions varies depending on the  $\lambda$  point considered, with highest frequencies of up to 10–15 events occurring in under 10 ns of simulation. This behavior is the major factor slowing free energy convergence. The slowest simulations to converge were at  $\lambda = 0.8$  for TIP3P and at  $\lambda = 0.825$  in TIP4P as these experienced the most frequent and extreme host hydration fluctuations, and therefore energy fluctuations, throughout their trajectories (Figure 6). We note that apparent convergence would be misleadingly attributed to the simulations considering MD periods shorter than the wetting/dewetting transition time scales (see Figure 3 and Supporting Information Table SI-1 for an example).

Because host hydration plays such a major role in the B2-CB[7] vdW decoupling simulations and free energy estimates, we include Table 2, which reveals some additional key differences in the TIP3P and TIP4P host wetting events during our simulations at  $\lambda$ 's where fluctuations occurred most. There is a striking yet conceivable difference in both the percent of time that the host is hydrated as well as the mean number of waters inside the host from  $\lambda = 0.775$  to 0.825 between the simulations using the two different water models, with more TIP3P molecules occupying the host cavity than TIP4P during these simulations. However, at the unbound ( $\lambda = 1.0$ ) and bound ( $\lambda = 0.0$ ) states, the two models behave quite similarly with the host stably hydrated or dehydrated throughout most if not all of the simulation, respectively. Additionally, at  $\lambda = 0.8$  our TIP3P simulation shows up to seven waters inside the host, while that of TIP4P fits only a maximum of six inside. The time to reach the first maximum hydration event and the mean number of waters inside the host

are also included for each case. Overall, this analysis does confirm the opposite behavior at the two  $\lambda$  end points though, with a hydrated cavity at  $\lambda = 1$  where the guest is fully decoupled (Figures 4 and 6b), representing the unbound state, and a dehydrated host cavity at  $\lambda = 0$  where the guest is fully coupled, representing the bound state (Figures 4, 5, and 6b).

Figure 7 further confirms the correlation of system free energy and host hydration seen with both water models (Figure 6, see also Supporting Information). Here, we look at the  $\langle \partial V / \partial \lambda \rangle$  vs the number of TIP4P or TIP3P waters in the host cavity at  $\lambda = 0.85$ , a point where the host is usually hydrated with up to seven waters but on average holds about three waters (Table 2). Although the standard deviations of  $\langle \partial V / \partial \lambda \rangle$  are large, there is a clear trend toward lower energy with an increasing number of water molecules using both models, which is consistent across all  $\lambda$ 's where the host cavity is capable of hydration (see Figure SI-2). At the  $\lambda$  end point states representing the unbound and bound host, we observe very different, stable behavior and no such correlation (Figure 6 and Figure SI-2). The range of  $\langle \partial V / \partial \lambda \rangle$  values is smaller ( $\sim 20$  kcal/mol) at both end points, with the unbound case showing a steadier  $\langle \partial V / \partial \lambda \rangle$  at or near 1.2 kcal/mol, regardless of the number of waters inside the host, and the bound case (where no waters ever fit inside the host) averaging around 0.2 kcal/mol (Figure SI-2). Clearly, this system presents a fascinating yet challenging behavior for TI absolute binding free energy calculations using explicit water and demonstrates how important the role of wetting and dewetting events is in the accuracy and convergence of free energy calculations.

## CONCLUSIONS

Thermodynamic integration was used to calculate the absolute binding free energy of a high affinity host–guest system in explicit solvent. Using both TIP3P and TIP4P water models, our estimations correspond remarkably to experimental calculations. Although both models proved to be adequate for our studies of this host–guest system, our simulations in TIP4P did lead us to an absolute binding free energy estimate gratifyingly closer to the experimental value. We observe a toroidal average water density surface first appearing inside the host in the unbound state ( $\lambda = 1.0$ ) or ordered around the top and bottom host oxygen molecules in the guest-bound state ( $\lambda = 0.0$ ), with fascinating patterns of intermediate water ordering seen both inside and out of the host at varying  $\lambda$  values in between. Significant changes in  $\langle \partial V / \partial \lambda \rangle$  occur along molecular dynamics simulations as the guest van der Waals interactions are decoupled along  $\lambda$  space using both water models. The dynamic behavior of host wetting and dewetting events is directly correlated with these large energy fluctuations of the system. For both water models, we see a decrease in  $\langle \partial V / \partial \lambda \rangle$  with increasing water molecules inside the host during van der Waals simulations with sufficient guest decoupling. The work described here emphasizes the importance of wetting/dewetting transitions and their influence on free energy estimation for the host–guest system considered in this study. We believe that the wetting transitions observed are relevant in more complex biological scenarios as well, including protein–ligand binding.<sup>18</sup> However, it is also expected that a protein cavity, which is accessible from the bulk from one side only, might have quite largely different hydration properties than a host guest system, such as CB[7], as the latter is open to bulk from two symmetric openings. Our



study is, to our knowledge, the first that addressed these slow water transitions in a host–guest system.

## ■ ASSOCIATED CONTENT

### ■ Supporting Information

TI curves of partial charge (ele) decoupling steps and correlation graphs of  $\langle \partial V / \partial \lambda \rangle$  vs the number of waters inside the host at additional  $\lambda$ 's are included. Examples of the variation of the convergence monitoring criterion,  $\tau_{\text{sim}}$ , are also provided to show changes in statistical uncertainty with increasing simulation time. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## ■ AUTHOR INFORMATION

### Corresponding Author

\*E-mail: [krogers@ucsd.edu](mailto:krogers@ucsd.edu); [r.baron@utah.edu](mailto:r.baron@utah.edu); [cesar@mccammon.ucsd.edu](mailto:cesar@mccammon.ucsd.edu).

### Present Address

<sup>#</sup>Department of Medicinal Chemistry, College of Pharmacy, and The Henry Eyring Center for Theoretical Chemistry, The University of Utah, Salt Lake City, Utah, United States.

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

This work was supported by NSF, NIH, HHMI, the National Biomedical Computation Resource, and the Center for Theoretical Biological Physics. The authors thank Michael Gilson and his group for providing original system coordinates and helpful discussions. R.B. acknowledges the University of Utah, Department of Medicinal Chemistry for startup funding to support the last stage of this project. J.M.O.-S. acknowledges the Fulbright Commission/Generalitat de Catalunya Program and the Generalitat de Catalunya for a Fulbright grant and a Beatriu de Pinos postdoctoral grant, respectively.

## ■ REFERENCES

- Freeman, W. A.; Mock, W. L.; Shih, N. Y. *J. Am. Chem. Soc.* **1981**, *103* (24), 7367–7368.
- Kim, J.; Jung, I. S.; Kim, S. Y.; Lee, E.; Kang, J. K.; Sakamoto, S.; Yamaguchi, K.; Kim, K. *J. Am. Chem. Soc.* **2000**, *122* (3), 540–541.
- Liu, S.; Ruspic, C.; Mukhopadhyay, P.; Chakrabarti, S.; Zavalij, P. Y.; Isaacs, L. *J. Am. Chem. Soc.* **2005**, *127* (45), 15959–15967.
- Rekharsky, M. V.; Mori, T.; Yang, C.; Ko, Y. H.; Selvapalam, N.; Kim, H.; Sobransingh, D.; Kaifer, A. E.; Liu, S.; Isaacs, L.; Chen, W.; Moghaddam, S.; Gilson, M. K.; Kim, K.; Inoue, Y. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104* (52), 20737–20742.
- Lagona, J.; Mukhopadhyay, P.; Chakrabarti, S.; Isaacs, L. *Angew. Chem., Int. Ed. Engl.* **2005**, *44* (31), 4844–4870.
- Jeon, Y. J.; Kim, S. Y.; Ko, Y. H.; Sakamoto, S.; Yamaguchi, K.; Kim, K. *Org. Biomol. Chem.* **2005**, *3* (11), 2122–2125.
- Zhao, Y. J.; Buck, D. P.; Morris, D. L.; Pourgholami, M. H.; Day, A. L.; Collins, J. G. *Org. Biomol. Chem.* **2008**, *6* (24), 4509–4515.
- Ma, D.; Hettiarachchi, G.; Nguyen, D.; Zhang, B.; Wittenberg, J. B.; Zavalij, P. Y.; Briken, V.; Isaacs, L. *Nat. Chem.* **2012**, *4* (6), 503–510.
- Hettiarachchi, G.; Nguyen, D.; Wu, J.; Lucas, D.; Ma, D.; Isaacs, L.; Briken, V. *PLoS One* **2010**, *5* (5), e10514.
- El-Barghouthi, M. I.; Assaf, K. I.; Rawashdeh, A. M. M. *J. Chem. Theory Comput.* **2010**, *6* (4), 984–992.
- Chen, W.; Chang, C. E.; Gilson, M. K. *Biophys. J.* **2004**, *87* (5), 3035–3049.
- Moghaddam, S.; Yang, C.; Rekharsky, M.; Ko, Y. H.; Kim, K.; Inoue, Y.; Gilson, M. K. *J. Am. Chem. Soc.* **2011**, *133* (10), 3570–3581.
- Muddana, H. S.; Gilson, M. K. *J. Chem. Theory Comput.* **2012**, *8* (6), 2023–2033.
- Baron, R.; Setny, P.; McCammon, J. A. *J. Am. Chem. Soc.* **2010**, *132* (34), 12091–12097.
- Hummer, G. *Nat. Chem.* **2010**, *2* (11), 906–907.
- Setny, P.; Baron, R.; McCammon, J. A. *J. Chem. Theory Comput.* **2010**, *6* (9), 2866–2871.
- Wang, L.; Friesner, R. A.; Berne, B. J. *Faraday Discuss.* **2010**, *146*, 247–262; discussion pp 283–298, 395–401.
- Wang, L.; Berne, B. J.; Friesner, R. A. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108* (4), 1326–1330.
- Young, T.; Hua, L.; Huang, X.; Abel, R.; Friesner, R.; Berne, B. J. *Proteins* **2010**, *78* (8), 1856–1869.
- Baron, R.; Setny, P.; McCammon, J. A., Hydrophobic Association and Vol.-Confined Water Molecules. In *Methods & Principles in Medicinal Chemistry - Protein-Ligand Interactions*; Gohlke, H., Ed.; Wiley-VCH: Weinheim, Germany, 2012; pp 145–170.
- Moghaddam, S.; Inoue, Y.; Gilson, M. K. *J. Am. Chem. Soc.* **2009**, *131* (11), 4012–4021.
- Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79* (2), 926–935.
- Kairys, V.; Gilson, M. K. *J. Comput. Chem.* **2002**, *23* (16), 1656–1670.
- Jeon, W. S.; Moon, K.; Park, S. H.; Chun, H.; Ko, Y. H.; Lee, J. Y.; Lee, E. S.; Samal, S.; Selvapalam, N.; Rekharsky, M. V.; Sindelar, V.; Sobransingh, D.; Inoue, Y.; Kaifer, A. E.; Kim, K. *J. Am. Chem. Soc.* **2005**, *127* (37), 12984–12989.
- Wang, J. M.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25* (9), 1157–1174.
- Wang, J. M.; Wang, W.; Kollman, P. A.; Case, D. A. *J. Mol. Graphics Modell.* **2006**, *25* (2), 247–260.
- Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97* (40), 10269–10280.
- Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Kollman, P. A. *J. Am. Chem. Soc.* **1993**, *115* (21), 9620–9631.
- Case, D. A.; Cheatham, T. E., III; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *J. Comput. Chem.* **2005**, *26* (16), 1668–1688.
- Fajer, M.; Baron, R.; de Oliveira, C. A. F.; Rogers, K.; Ortiz-Sanchez, J. M.; Wereszczynski, J.; Steinbrecher, T.; Walker, R. C.; McCammon, J. A. Manuscript in preparation.
- Grest, G. S.; Kremer, K. *Phys. Rev. A* **1986**, *33* (5), 3628–3631.
- Pastor, R. W.; Brooks, B. R.; Szabo, A. *Mol. Phys.* **1988**, *65* (6), 1409–1419.
- Kirkwood, J. G. *J. Chem. Phys.* **1935**, *3*, 300–313.
- Zacharias, M.; Straatsma, T. P.; McCammon, J. A. *J. Chem. Phys.* **1994**, *100* (12), 9025–9031.
- Steinbrecher, T.; Mobley, D. L.; Case, D. A. *J. Chem. Phys.* **2007**, *127* (21), 214108.
- Allen, M. P.; Tildesley, D. J. *Computer Simulations of Liquids*; Oxford University Press: New York, 1989.
- General, I. J. *J. Chem. Theory Comput.* **2010**, *6* (8), 2520–2524.
- Gilson, M. K.; Given, J. A.; Bush, B. L.; McCammon, J. A. *Biophys. J.* **1997**, *72* (3), 1047–1069.
- Hamelberg, D.; McCammon, J. A. *J. Am. Chem. Soc.* **2004**, *126* (24), 7683–7689.
- Lawrenz, M.; Baron, R.; Wang, Y.; McCammon, J. A. *J. Chem. Theory Comput.* **2011**, *7* (7), 2224–2232.
- Vega, C.; Abascal, J. L. *Phys. Chem. Chem. Phys.* **2011**, *13* (44), 19663–19688.
- Nguyen, C. N.; Kurtzman Young, T.; Gilson, M. K. *J. Chem. Phys.* **2012**, *137* (4), 044101.