

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/6873527>

# Modeling Competitive Guest Binding to $\beta$ -Cyclodextrin Molecular Printboards

ARTICLE *in* THE JOURNAL OF PHYSICAL CHEMISTRY B · SEPTEMBER 2006

Impact Factor: 3.3 · DOI: 10.1021/jp062553n · Source: PubMed

---

CITATIONS

25

---

READS

133

2 AUTHORS, INCLUDING:



Andreas Larsson

Luleå University of Technology

56 PUBLICATIONS 990 CITATIONS

SEE PROFILE

# Modeling Competitive Guest Binding to $\beta$ -Cyclodextrin Molecular Printboards

D. Thompson\* and J. A. Larsson

Tyndall National Institute, Lee Maltings, Cork, Ireland

Received: April 26, 2006

Anchoring of functionalized guest molecules to self-assembled monolayers (SAMs) is key to the development of molecular printboards for nanopatterning. One very promising system involves guest binding to immobilized  $\beta$ -cyclodextrin ( $\beta$ -CD) hosts, with guest:host recognition facilitated by a hydrophobic interaction between uncharged anchor groups on the guest molecule and  $\beta$ -CD hosts self-assembled at gold surfaces. We use molecular dynamics free energy (MDFE) simulations to describe the specificity of guest: $\beta$ -CD association. We find good agreement with experimental thermodynamic measurements for binding enthalpy differences between three commonly used phenyl guests: benzene, toluene, and *t*-butylbenzene. van der Waals interaction with the inside of the host cavity accounts for almost all of the net stabilization of the larger phenyl guests in  $\beta$ -CD. Partial and full methylation of the secondary rim of  $\beta$ -CD decreases host rigidity and significantly impairs binding of both phenyl and larger adamantane guest molecules. The  $\beta$ -CD cavity is also very intolerant of guest charging, penalizing the oxidized state of ferrocene by at least 7 kcal/mol.  $\beta$ -CD hence expresses moderate specificity toward uncharged organic guest molecules by van der Waals recognition, with a much higher specificity calculated for electrostatic recognition of organometallic guests.

## 1. Introduction

Self-assembly is used extensively by nature to build complex structures via specific noncovalent guest:host interactions. Understanding recognition mechanisms in physical and chemical terms means accessing the information contained in (macro)-molecules, crucial for understanding interactions such as antigen:antibody and ligand:enzyme recognition, and the design of new molecules with tailored recognition properties. In recent years, “molecular printboards”, exploiting the guest complexation properties of cyclodextrin, have been developed and used for nanopatterning applications.<sup>1</sup>  $\beta$ -cyclodextrin ( $\beta$ -CD) molecules can be anchored to gold via long alkanethiol and alkanethioether chains, forming highly ordered self-assembled monolayers (SAMs) with cyclodextrin cavities exposed at the surface. The hydrophobic interior of  $\beta$ -CD allows it to bind uncharged guest molecules in its cavity. This combination of host structural order and specific guest binding provides a molecular printboard. These molecular printboards serve as a first approximation to the much more complex molecular machines used to build and regulate living systems.

The ability to fine-tune guest:host specificity on molecular printboards will allow the use of more complex hosts and many-guest systems, with each host site tailored to bind only one specific guest. Such is the case in living systems, and we use here a method<sup>2–4</sup> which has proved to be a valuable complementary tool to experimental measures of ligand:enzyme interactions, showing the crucial and sometimes interdependent contributions of host conformation and charge states to the overall measured binding energies. The binding free energy difference ( $\Delta\Delta G$ ) for complexation of the host with different guest molecules can be obtained from molecular dynamics free energy (MDFE) simulations<sup>2–4</sup> and the overall  $\Delta\Delta G$  decom-

posed into van der Waals and electrostatic interactions between the guests and individual host sites. The free energy change associated with guest complexation in  $\beta$ -CD is the same both at the SAM and in bulk solution<sup>1g</sup>—the long alkyl chains linking  $\beta$ -CD to the underlying gold surface screen long-range electrostatic effects at the guest:host binding site, and thus, we perform binding simulations to unanchored  $\beta$ -CD. A thermodynamic cycle (see below) is used to obtain binding free energy differences, and a group decomposition of the free energy is used to identify the host sites determining guest binding specificity.

MDFE simulations require a treatment of long-range interactions that is accurate and also efficient, to allow extensive conformational sampling. All-atom simulations are performed for the  $\beta$ -CD:guest complexes, centered on the guest, and immersed in a 40 Å cubic water box. Periodic boundary conditions were applied. van der Waals interactions were truncated at atomic separations of 14 Å. All long-range electrostatic interactions were computed efficiently by the particle mesh Ewald (PME) method. All simulations of guest binding lasted at least 2 ns.

Guest binding to cyclodextrins serves as a model system for hydrophobic self-assembly and has been the subject of many experimental<sup>5</sup> and theoretical<sup>6–10</sup> studies. Comparison of our approach to that presented in a number of recent theoretical studies of guest binding in  $\beta$ -CD hosts<sup>7–10</sup> shows the power of molecular dynamics simulations to probe guest recognition modes. For example, Brooks et al.<sup>7</sup> calculated binding free energy differences for various monosubstituted benzene guests with  $\beta$ -CD and obtained highly correlated results for two different MDFE schemes,  $\lambda$ -dynamics and free energy perturbation. Kollman et al.<sup>8</sup> used MDFE and also an approach based on just single-state MD trajectories (MM/PBSA) to compute ester versus amide specificity in a  $\beta$ -CD dimer with a disulfide bridge linking the  $\beta$ -CD units. Very long MD simulations of up to 12 ns were used by Varady et al.<sup>9</sup> to model the

\* To whom all correspondence and requests for reprints should be addressed. Phone: +353-21-490-4327. Fax: +353-21-427-0271. E-mail: damien.thompson@tyndall.ie.

complexation of benzyl alcohol to  $\beta$ -CD in solution. The simulations reproduced the experimental complexation state even from totally uncomplexed starting geometries and also identified a number of highly populated conformational states, in addition to the experimental snapshot. Finally, an in-depth treatment of binding of a range of guests, including benzene, resorcinol, flurbiprofen, naproxen, and nabumetone, to  $\alpha$ -,  $\beta$ -, and  $\gamma$ -CD has recently appeared.<sup>10</sup> Gilson and co-workers<sup>10</sup> used a static, but rather sophisticated, method to dock guests. By incorporating estimates of configurational entropy into their free energy calculations, excellent agreement with available experimental binding free energies was obtained.

The following sections describe our computational methods, results, and conclusions. We show that, in agreement with thermodynamic measurements, guest binding affinity to  $\beta$ -CD is *t*-butylbenzene > toluene > benzene. Increased van der Waals stabilization of the larger guest by the inside of the host accounts for most of the observed binding free energy difference, and methylation of the outer rim of  $\beta$ -CD loosens the guest:host steric “fit” by reducing the rigidity of the cavity. Host methylation impedes binding of both phenyl and larger adamantane guest molecules. Charged guests do not bind to native  $\beta$ -CD, with a very large penalty computed for binding oxidized versus neutral ferrocene.

## 2. Methods

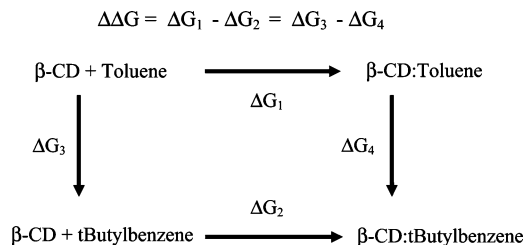
**2.1. Molecular Dynamics Simulations.** Starting structures were generated from a benzyl alcohol: $\beta$ -CD crystal structure, Cambridge Structural Database entry DEBGOG.<sup>11</sup> This structure has the guest polar OH group oriented toward the secondary rim of  $\beta$ -CD. The benzyl alcohol OH group was replaced by an acetamide tail group to give the benzene: $\beta$ -CD complex; benzene: $\beta$ -CD is used as shorthand for *N*-phenyl-acetamide: $\beta$ -CD. Similarly, for the other complexes described below where the *guest* in *guest*: $\beta$ -CD denotes the phenyl headgroup of the guest, the neutral acetamide tail group is present on all guest molecules but is oriented away from the  $\beta$ -CD cavity and not involved in complexation. For example, what we call *t*-butylbenzene: $\beta$ -CD is formally 4-*t*-butylphenylacetanilide: $\beta$ -CD.

The toluene: $\beta$ -CD complex was then created by building in a methyl group on the benzene carbon para to that hosting the acetamide tail group. *t*-Butylbenzene: $\beta$ -CD was generated by simply replacing the toluene CH<sub>3</sub> hydrogens with methyl groups. Toluene and *t*-butylbenzene are flipped relative to benzyl alcohol to account for the higher hydrophobicity of the primary rim of  $\beta$ -CD.

A similar situation exists for the adamantane and ferrocene guest: $\beta$ -CD complexes. Adamantane has an acetamide tail, like the phenyl guests above, and is placed in  $\beta$ -CD by aligning the adamantane headgroup center of mass with the center of the cavity. To probe host functionalization, we measured adamantane, and also toluene, binding to both native and methylated  $\beta$ -CD hosts. We built methylated hosts by replacing either one or both secondary OH groups on each glucose monomer with OMe groups. Finally, we placed a methyl alcohol tail on ferrocene to more closely match the guest molecule used in the experimental binding studies.<sup>18</sup> Ferrocene was placed initially with its central iron atom at the  $\beta$ -CD center of mass.

Hydrogens were constructed with ideal stereochemistry. A 40 Å cubic box of water was overlaid, and waters overlapping the complex were removed. Periodic boundary conditions were assumed; that is, the entire 40 Å box was replicated periodically in all directions.

Each complex was subjected to short minimization prior to MD, to remove any steric clashes. Molecular dynamics



**Figure 1.** Thermodynamic cycle used for computation of the guest binding free energy differences ( $\Delta\Delta G$ ).  $\Delta G_1$  and  $\Delta G_4$  refer to free energy changes for guest mutation in bulk solution and in the  $\beta$ -cyclodextrin ( $\beta$ -CD) host, respectively, while  $\Delta G_2$  and  $\Delta G_3$  are free energy changes for toluene and *t*-butylbenzene binding to host. Similar cycles compare benzene to *t*-butylbenzene and ferrocene to ferrocenium guests.

simulations were performed for 2 ns (for each complex) at constant room temperature and pressure with a Nosé–Hoover algorithm, following 300 ps of thermalization with gradually reducing constraints on the  $\beta$ -CD host heavy atoms. To aid trajectory analysis, the  $\beta$ -CD center of mass was weakly restrained at the origin throughout the dynamics. This is a harmless and efficient way of preventing the complex “wandering” in the large water box; all other degrees of freedom were left unrestrained. Trajectories of 400 ps were also produced for each guest and host molecule in solution, with the uncomplexed guest or host in each case solvated at the center of a box of water molecules.

The CHARMM22 force field<sup>12</sup> was used for the phenyl guest molecules and  $\beta$ -CD. Adamantane carbon–carbon parameters were obtained from a web-based X-Plor database.<sup>13</sup> Charges were assigned on the basis of standard CHARMM methodology, that is, +0.09 charge on each hydrogen bonded to carbon with a countercharge on the carbon. Ferrocene and ferrocenium force field parameters were generated from existing electronic structure data.<sup>14,15</sup> A slightly modified TIP3P model was used for the water.<sup>16</sup> Bonds involving hydrogen were constrained to their experimental lengths with the SHAKE algorithm,<sup>17</sup> allowing the use of a 2 fs time step for dynamics, except for adamantane complexes where a 1 fs time step gave more stable trajectories. We used the CHARMM program<sup>18</sup> version c31b2 for all calculations.

**2.2. Free Energy Calculations.** The methods used to compute guest:host binding free energy differences ( $\Delta\Delta G$ ) have been well described previously.<sup>2–4,19–21</sup> To compare binding of phenyl guests with increasingly large headgroups, we use the thermodynamic cycle in Figure 1. The vertical legs in Figure 1 correspond to rigorous, alchemical, molecular dynamics free energy (MDFE) simulations in which the guest molecule is transformed into a competitor molecule, both in bulk solution and in the host. The larger guest is reversibly grown in during a series of simulations; the corresponding work is derived from a thermodynamic integration formula.<sup>2</sup> Binding *t*-butylbenzene is compared to binding both toluene and benzene.

MDFE mutation runs were performed in 40 Å water boxes, with periodic boundary conditions and an Ewald treatment of electrostatics. The guest geometry, atom types, and charges were reversibly changed from that of one guest to the other over a series of 10 simulations, or “windows”, of 100 ps. Dummy atoms of zero charge and zero mass were used to represent atoms not present at the smaller guest endpoint. Runs were performed in both forward and backward directions, growing in and disappearing the *t*-butyl group.

The energy function can be expressed as a linear mixture of terms for the competitive guests. Comparing, for example,

**TABLE 1: MDFE Binding Free Energy Differences ( $\Delta\Delta G$ , kcal/mol) for Phenyl Guest Binding to  $\beta$ -CD<sup>a</sup>**

competitor guest	$\Delta G$ in bulk solution	$\Delta G$ in $\beta$ -CD	$\Delta\Delta G$
toluene/ <i>t</i> -butylbenzene	−8.5, −8.3; −9.0, −9.0	−6.6, −7.2, −7.6, −7.4; −5.7, −6.3, −6.6, −5.3	2.1 ± 0.8
benzene/ <i>t</i> -butylbenzene	5.8, 6.6; 6.9, 5.6	8.6, 8.3, 9.0, 8.3; 9.5, 11.5, 10.1, 10.1	3.2 ± 1.2

<sup>a</sup> MDFE  $\Delta\Delta G$  is computed from the vertical, alchemical legs in Figure 1:  $\Delta\Delta G = \Delta G_1 - \Delta G_4$ . Uncertainty in  $\Delta\Delta G$  is given as the standard deviation in the individual pairs of forward and backward runs, two each in solution and four each in  $\beta$ -CD. A positive  $\Delta\Delta G$  value indicates preferential *t*-butylbenzene binding.

toluene and *t*-butylbenzene, the energy function may be written as<sup>2</sup>

$$U(\lambda) = U_0 + (1 - \lambda)U_{t\text{-butylbenzene}} + \lambda U_{\text{toluene}}$$

where  $\lambda$  is a weight, or “coupling parameter”, and  $U_0$  represents interactions between parts of the system other than the hybrid guest. We gradually mutated toluene into *t*-butylbenzene by changing  $\lambda$  from 1 to zero. The successive weights were  $\lambda = 0.99, 0.95, 0.9, 0.8, 0.6, 0.4, 0.2, 0.1, 0.05$ , and  $0.01$  for toluene, and  $1 - \lambda$  for *t*-butylbenzene. The derivative of the free energy ( $G$ ) can be written as

$$dG/d\lambda = \langle U_{\text{toluene}} - U_{t\text{-butylbenzene}} \rangle_\lambda$$

where the brackets represent a time average over an MD trajectory performed with the energy function  $U(\lambda)$ . The free energy derivatives were computed at each  $\lambda$  value from a 100 ps MD simulation. Each run thus corresponds to 1.0 ns in all. To obtain the free energy change, trapezoidal integration was then used for all but near-endpoint ( $\lambda = 0-0.03$  and  $\lambda = 0.97-1$ ) van der Waals contributions, where the steepness of the derivative makes analytical integration more suitable.<sup>22</sup>

### 3. Results

**3.1. Structure, Dynamics, and Solvation of the Phenyl Guest Binding Site in  $\beta$ -CD.** To prevent the benzene guest from flipping (every 100 ps or so) in the cavity and orienting its tail into the  $\beta$ -CD cavity, it was necessary to increase the length of the tail, replacing the methyl termination with a propyl group. A longer tail was also used for phenyl guests with small anchor groups in experimental studies of guest binding to  $\beta$ -CD.<sup>18</sup> The toluene and *t*-butylbenzene guests did not flip during the 2 ns of dynamics. For all three phenyl guest: $\beta$ -CD complexes, the average root-mean-square deviation (rmsd) of  $\beta$ -CD heavy atoms from the starting energy-minimized geometry was 0.9–1.0 Å, similar to that of uncomplexed  $\beta$ -CD (0.9 Å).

Three to four water molecules were displaced from the cavity to allow for guest binding, with the uncomplexed  $\beta$ -CD cavity holding on average seven water molecules.<sup>23</sup> No correlation was observed between guest size and the number of excluded waters. No direct or water-mediated guest:host H-bonding was observed.

**3.2. Toluene/*t*-Butylbenzene Binding to  $\beta$ -CD.** To obtain the toluene/*t*-butylbenzene binding free energy difference, we use MDFE alchemical simulations. The data are summarized in Table 1. Each MDFE  $\Delta\Delta G$  value is computed from four simulations in solution and eight simulations in  $\beta$ -CD (10 windows of 100 ps each; see Methods). In solution, two runs were performed in each direction: mutating toluene to *t*-butylbenzene or the reverse, yielding  $\Delta G_1$  in Figure 1. Solution MDFE runs were initiated from MD snapshots taken 200 and 400 ps out on the toluene:solution and *t*-butylbenzene:solution trajectories. In the  $\beta$ -CD host, four runs were performed in each direction, yielding  $\Delta G_4$  in Figure 1. These host MDFE calculations were started from snapshots taken 0.5, 1.0, 1.5, and 2.0 ns out on the toluene:host and *t*-butylbenzene:host native trajectories. In all cases, the first 40 ps of each window was

used for equilibration at the new  $\lambda$  value and derivatives were computed from the last 60 ps. As shown in Table 1,  $\Delta G$  in solution is −8.7 kcal/mol pro-toluene, while the host system has  $\Delta G = -6.6$  kcal/mol pro-toluene. Subtracting the host and solution values (Figure 1) gives an overall value of  $\Delta\Delta G = +2.1$  kcal/mol in favor of *t*-butylbenzene. The uncertainty, taken as the standard deviation over all solution:host pairs of runs, 32 in all, is  $\pm 0.8$  kcal/mol. An overall free energy difference of 3.1 kcal/mol in favor of the *t*-butylbenzene guest was measured using isothermal titration calorimetry.<sup>18</sup> Subtracting the small entropic contribution gives 2.5 kcal/mol,<sup>18</sup> well within the range of our predicted  $\Delta\Delta G$  value of  $2.1 \pm 0.8$  kcal/mol.

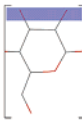
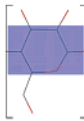
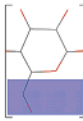
MDFE calculations allow a decomposition of the overall  $\Delta\Delta G$  value into van der Waals and electrostatic components. As shown in Table 2, van der Waals effects account for most of the *t*-butylbenzene stabilization, as expected for these uncharged guests. The lack of electrostatic guest:host interaction is confirmed by Poisson–Boltzmann (PB)<sup>21</sup> electrostatic interaction energy calculations (not shown); PB  $\Delta G$  values are always less than 1 kcal/mol for all three guests. To identify the source of binding specificity, we project the MDFE van der Waals  $\Delta\Delta G$  value onto  $\beta$ -CD and water. We do this by simply switching off all atoms except those in the guests and in the host sites of interest. Free energy components in Table 2 show that van der Waals interaction with the  $\beta$ -CD host stabilizes *t*-butylbenzene, offset slightly by a small water penalty for the larger guest. The inner cavity of the host, rather than the peripheral hydroxyl groups, confers specificity. The  $\beta$ -CD cavity can get a tighter “grip” on the larger *t*-butylbenzene guest.

**3.3. Benzene/*t*-Butylbenzene Binding to  $\beta$ -CD.** To complete our analysis of the binding of the phenyl guests to  $\beta$ -CD, we also alchemically mutated benzene into *t*-butylbenzene and calculated the binding free energy difference between the two guests. As explained above, benzene requires a longer tail to prevent flipping in the cavity. Thus, in morphing benzene into *t*-butylbenzene, we simultaneously grow out this longer tail while growing in the *t*-butylbenzene anchor group. The computed  $\Delta\Delta G$  value is  $+3.2 \pm 1.2$  kcal/mol in favor of *t*-butylbenzene (Table 1). The measured binding enthalpy difference is  $+2.9$  kcal/mol,<sup>18</sup> very close to our computed value, and we also get the right rank ordering of benzene, toluene, and *t*-butylbenzene binding strengths. The van der Waals components in Table 2 show again how the inner cavity of  $\beta$ -CD provides most of the specificity for binding the larger *t*-butylbenzene guest. The larger the difference in guest size, the more cyclodextrin stabilizes the large *t*-butylbenzene guest.

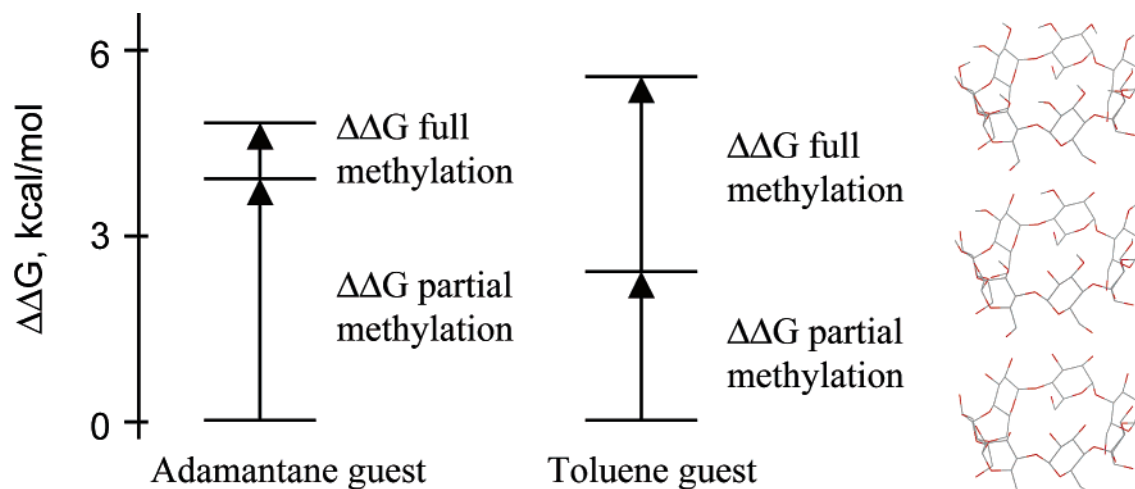
**3.4. Adamantane Binding to Native and Methylated  $\beta$ -CD.** To further explore host stabilization of large guest molecules, we performed additional MDFE calculations for adamantane binding, to both native and methylated  $\beta$ -CD. Host methylation involved partial and full methylation of secondary OH groups, replacing either 7 or all 14 of the secondary OH groups with OCH<sub>3</sub> (see the structures in Figure 2). As expected,<sup>6</sup> replacement of some or all of the polar secondary OH groups with OMe leads to loss of the  $\beta$ -CD truncated cone shape. The guest binding site warps due to the loss of hydrogen bonding,



TABLE 2: MDFF Binding Free Energy Components (kcal/mol) for Phenyl Guest Binding to  $\beta$ -CD<sup>a</sup>

Competitor guests	$\Delta\Delta G_{\text{electrostatic}}$	$\Delta\Delta G_{\text{van der Waals}}$	van der Waals components		
					
Toluene / tButylbenzene	0.3	1.8	0.4	2.2	0.2
Benzene / tButylbenzene	0.3	2.9	0.4	3.0	0.3

<sup>a</sup> A positive  $\Delta\Delta G$  value indicates preferential *t*-butylbenzene binding.  $\beta$ -CD van der Waals components sum to 2.8 and 3.7 kcal/mol for toluene/*t*-butylbenzene and benzene/*t*-butylbenzene, respectively. The remaining  $-1.0$  and  $-0.8$  kcal/mol, respectively, is penalization of the larger *t*-butylbenzene guest by the water molecules remaining in the  $\beta$ -CD cavity.



**Figure 2.** Guest binding free energy changes upon methylation, for adamantane and toluene binding to native, partially methylated and fully methylated  $\beta$ -CD. Binding free energy changes are computed from guest:host interaction energies, averaged over 2 ns. Shown on the right are native, partially methylated and fully methylated  $\beta$ -CD structures.

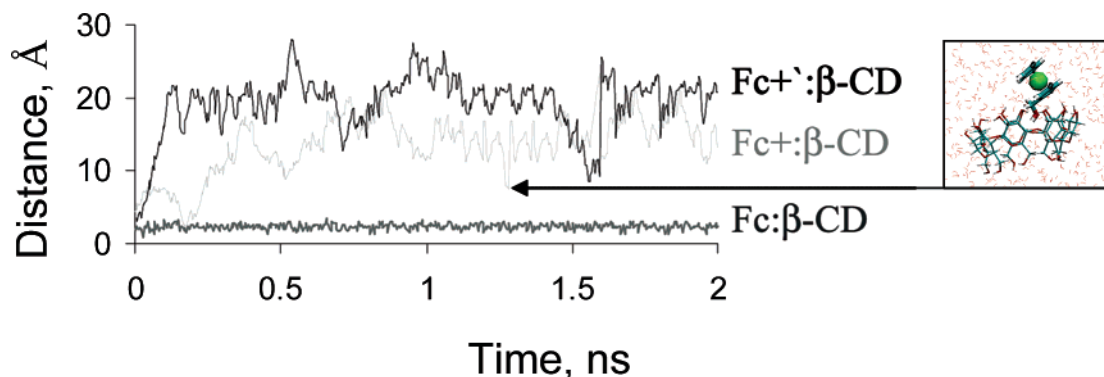
becoming less spherical. We see in our simulations that the parent glucose fragments become tilted with respect to each other, most noticeably in the fully methylated structure, with two to three of the seven sugar units rotating by up to  $90^\circ$ . This allows the  $\text{CH}_2\text{OH}$  groups on the primary rim of  $\beta$ -CD to poke up into the bottom of the cavity.

The increased structural disorder upon methylation is reflected in the increased time-averaged root-mean-square deviations (rmsd's) of host heavy atoms from the starting geometry. The rmsd values are  $0.9 \text{ \AA}$  in native  $\beta$ -CD,  $1.4 \text{ \AA}$  in partially methylated  $\beta$ -CD, and  $1.7 \text{ \AA}$  in fully methylated  $\beta$ -CD. "Growing in" seven methoxyl groups, in the same way that the guest molecules were alchemically mutated above, always carries a penalty of  $+127 \pm 2 \text{ kcal/mol}$  (data not shown), for both native  $\rightarrow$  partially methylated and partially methylated  $\rightarrow$  fully methylated  $\beta$ -CD, both alone in solution and when complexed to adamantane or toluene. The penalty for methylation is thus additive, and more importantly, guest binding does not affect the transition from native to partially or fully methylated  $\beta$ -CD. That is, the host is skewed to the same extent, whether alone in solution or with a bound guest. This constant formation energy, irrespective of guest binding, together with visual inspection of the MD structures, indicates that the new methoxy groups do not dip into the cavity.

Adamantane has a stronger interaction with  $\beta$ -CD than toluene, in agreement with experiment.<sup>1g</sup> The time-averaged interaction between adamantane and native  $\beta$ -CD is around 4

kcal/mol lower than that for toluene: $\beta$ -CD. Adamantane also has the lowest rms deviation and fluctuation of all the guests in this study, showing how well adamantane "fits" the  $\beta$ -CD cavity. Computed guest:host interaction energies show however that methylation significantly impedes guest binding. These losses in guest binding energy due to methylation are plotted in Figure 2. A partial loss in host rigidity has less of an effect on toluene binding than it does on adamantane binding. This is because toluene is smaller than adamantane, and so not as strongly affected by partial methylation of the host. The excellent steric fit between adamantane and  $\beta$ -CD is upset by partial methylation, reducing the adamantane: $\beta$ -CD binding strength by 4 kcal/mol (Figure 2). Upon full methylation, the host becomes extremely distorted and thus binding of both adamantane and toluene is severely impaired, by approximately 5 kcal/mol relative to binding in native  $\beta$ -CD. Thus, it is clear that the native host, with its strong intramolecular H-bonding network, is best for guest binding. Though the new methyl groups do not directly contact the guest, warping of the methylated host means it partially loses its grip on the guest and so guests have a lower binding strength in methylated cyclodextrin.

**3.5. Ferrocene/Ferrocenium Binding to  $\beta$ -CD.** In a preliminary investigation of charge effects in  $\beta$ -CD hosts, we compared binding of neutral and oxidized ferrocene guests to  $\beta$ -CD. Unbinding of guest molecules was observed experimentally when the neutral ferrocene anchor was oxidized to the ferrocenium cation.<sup>1e,24</sup> The rather exotic electronic structure



**Figure 3.** Timeline of guest:host separations for neutral ferrocene (Fc, dark gray) and cationic ferrocenium (Fc+<sup>+</sup> in black, Fc+ in light gray) binding to  $\beta$ -CD. Fc+ is ferrocenium with the extra +1 charge on iron, Fc+<sup>+</sup> is ferrocenium with an extra +0.5 on iron and the remaining +0.5 spread over the cyclopentadienyl carbons. The distance plotted is between the  $\beta$ -CD center of mass and the central iron atom of ferrocene. The structure shown on the right-hand side is a typical snapshot of partially bound ferrocenium; Fe and the cyclopentadienyl groups are shown in sphere and thick stick format, respectively, while  $\beta$ -CD is shown as thin sticks.

**TABLE 3: MDFE and PBFE Binding Free Energy Differences ( $\Delta\Delta G$ , kcal/mol) for Neutral Ferrocenium (Fc) and the Ferrocenium Cation in  $\beta$ -CD<sup>a</sup>**

competitor guest	$\Delta G$ in bulk solution	$\Delta G$ in $\beta$ -CD	MDFE $\Delta\Delta G$	PBFE $\Delta\Delta G$	
				$\epsilon = 2$	$\epsilon = 1$
Fc/Fc+	-25.3; -25.6	-18.4; -18.2	7.2	6.5	13.1
Fc/Fc+ <sup>+</sup>	-12.4; -12.1	-2.1; -4.5	9.0	6.9	13.8

<sup>a</sup> Fc+ and Fc+<sup>+</sup> denote the cation with two different charge smearing schemes, as described in the text. MDFE  $\Delta\Delta G$  is averaged over forward and backward runs both in solution and in the host. PBFE  $\Delta\Delta G$  is computed from 2 ns trajectories, sampling every 10 ps, where  $\epsilon$  is the dielectric constant used for the host. A positive  $\Delta\Delta G$  value indicates preferential binding of neutral ferrocene.

of ferrocene and its cation<sup>15</sup> means that computationally intensive quantum mechanical calculations are required to obtain accurate charge distributions and geometrical force constants. Here, we use existing ferrocene force field parameters<sup>14</sup> and, given the observed minor changes in geometry upon oxidation,<sup>15</sup> place an overall positive charge on the guest by simply either adding +1.0 to Fe or, alternatively, adding +0.5 to Fe and then smoothing the remaining +0.5 over the cyclopentadienyl groups. Obviously, rigorous *ab initio* calculations are required for a more flexible and polarizable representation of the ferrocenium cation. Nevertheless, with both charge sets, the ferrocenium guest unbinds from  $\beta$ -CD within 0.2 ns (Figure 3). Both the anchor and tail of the charged guest is solvated by bulk water for most of the remaining 1.8 ns, though the ferrocenium guest occasionally samples a partially bound conformation whereby its methanol tail is oriented toward the secondary hydroxyls at the entrance of the  $\beta$ -CD cavity (Figure 3 and below).

To obtain the free energy change upon guest oxidation and calculate the penalty associated with guest charging, we placed a 1–3 kcal/mol per Å<sup>2</sup> center-of-mass restraint on ferrocenium in further simulations, forcing it to remain in the cavity. We then alchemically mutated ferrocene into ferrocenium, both in solution and in the host (Table 3), and obtained a binding free energy difference ( $\Delta\Delta G$ ) of at least 7 kcal/mol in favor of neutral ferrocene. This computed ( $\Delta\Delta G$ ) value is higher than that estimated from binding experiments<sup>24</sup> and may be due to the restraint term preventing sampling of partially unbound ferrocenium: $\beta$ -CD states, some of which were observed in the unrestrained ferrocenium: $\beta$ -CD trajectories (inset of Figure 3). These weak ferrocenium: $\beta$ -CD complexes (Figure 3) always involved the guest methanol tail orienting toward the OH groups at the entrance to the  $\beta$ -CD cavity. Also, in binding experiments, ferrocenium is generally present with a short polar tail, making

it difficult to separate out guest anchor and tail contributions to measured binding strengths.<sup>24</sup> Thus, our calculated high  $\Delta\Delta G$  value may be more representative of the specificity in ferrocene/ferrocenium binding as, for example, anchor groups on atomic force microscopy (AFM) probes, dendrimers, and nanoparticles<sup>1</sup>, where only the ferrocene/ferrocenium anchor can bind in the cavity. The high electrostatic specificity calculated with MDFE, similar in magnitude to that expressed in highly tuned biological processes,<sup>25</sup> is supported by Poisson–Boltzmann estimates of  $\Delta\Delta G$  (PBFE).<sup>21</sup> PBFE calculations estimate guest binding strengths using the horizontal legs in the thermodynamic cycle in Figure 1, and using a host dielectric constant ( $\epsilon$ ) between 1 and 2,<sup>10</sup> we find good agreement with MDFE  $\Delta\Delta G$  values (Table 3).

#### 4. Conclusions

Thermal calorimetry and surface plasmon resonance provide a wealth of information on guest binding thermodynamics, allowing calculation of overall binding free energies as well as their enthalpic and entropic constituents. Theoretical work provides a valuable aid, allowing insight into the structure and dynamics of guest:host complexation. The host sites contributing strongly to guest binding specificity may also be identified. With a sufficiently rigorous treatment of binding, it is possible in some cases to accurately compute binding enthalpies and even binding free energies, though entropy estimates remain difficult.<sup>26</sup> Through the use of thermodynamic cycles, one may avoid some of the pitfalls of binding free energy estimates, principally the huge sampling space often required for movement of the guest from a completely solvated to completely bound environment. By alchemically mutating from one state to the other, it is possible to rigorously compute the binding free energy difference between similarly sized guests with similar binding modes. Indeed, we show here the power of MDFE calculations to reproduce experimental specificities based on both steric and electrostatic recognition in  $\beta$ -CD. Also, the steric binding energy penalty associated with secondary rim methylation is calculated and rationalized from the MD structures.

The ability to functionalize both guest and host structures, and predict accurately the resulting binding energies, will drive the next generation of nanopatterning technologies. More precise and simultaneous control over multiple specific noncovalent interactions will facilitate the creation of inhomogeneous large-scale assemblies, such as those found in living systems.

**Acknowledgment.** This work was partially funded by the EC NaPa project (Contract No. NMP4-CT-2003-500120). We

thank Jim Greer, Jurriaan Huskens, and Bart Jan Ravoo for discussions. The authors wish to acknowledge the SFI/HEA Irish Centre for High-End Computing (ICHEC) for the provision of some computational facilities and support.

## References and Notes

- (1) (a) Schonherr, H.; Beulen, M. W. J.; Bugler, J.; Huskens, J.; van Veggel, F. C. J. M.; Reinhoudt, D. N.; Vancso, G. J. *J. Am. Chem. Soc.* **2000**, *122*, 4963–4967. (b) de Jong, M. R.; Huskens, J.; Reinhoudt, D. N. *Chem.—Eur. J.* **2001**, *7*, 4164–4170. (c) Huskens, J.; Deij, M. A.; Reinhoudt, D. N. *Angew. Chem., Int. Ed.* **2002**, *41*, 4467–4471. (d) Zapotoczny, S.; Auletta, T.; de Jong, M. R.; Schonherr, H.; Huskens, J.; van Veggel, F. C. J. M.; Reinhoudt, D. N.; Vancso, G. J. *Langmuir* **2002**, *18*, 6988–6994. (e) Auletta, T.; Dordi, B.; Mulder, A.; Sartori, A.; Onclin, S.; Bruinink, C. M.; Nijhuis, C. A.; Beijleveld, H.; Péter, M.; Schönher, H.; Vancso, G. J.; Casnati, A.; Ungaro, R.; Ravoo, B. J.; Huskens, J.; Reinhoudt, D. N. *Angew. Chem., Int. Ed.* **2004**, *43*, 369–373. (f) Onclin, S.; Mulder, A.; Huskens, J.; Ravoo, B. J.; Reinhoudt, D. N. *Langmuir* **2004**, *20*, 5460–5466. (g) Auletta, T.; de Jong, M. R.; Mulder, A.; van Veggel, F. C. J. M.; Huskens, J.; Reinhoudt, D. N.; Zou, S.; Zapotoczny, S.; Schönher, H.; Vancso, G. J.; Kuipers, L. *J. Am. Chem. Soc.* **2004**, *126*, 1577–1584. (h) Mulder, A.; Auletta, T.; Sartori, A.; Del Ciotto, S.; Casnati, A.; Ungaro, R.; Huskens, J.; Reinhoudt, D. N. *J. Am. Chem. Soc.* **2004**, *126*, 6627–6636. (i) Huskens, J.; Mulder, A.; Auletta, T.; Nijhuis, C. A.; Ludden, M. J. W.; Reinhoudt, D. N. *J. Am. Chem. Soc.* **2004**, *126*, 6784–6797. (j) Nijhuis, C. A.; Huskens, J.; Reinhoudt, D. N. *J. Am. Chem. Soc.* **2004**, *126*, 12266–12267. (k) Bruinink, C. M.; Nijhuis, C. A.; Péter, M.; Dordi, B.; Crespo-Biel, O.; Auletta, T.; Mulder, A.; Schönher, H.; Vancso, G. J.; Huskens, J.; Reinhoudt, D. N. *Chem.—Eur. J.* **2005**, *11*, 3988–3996. (l) Crespo-Biel, O.; Péter, M.; Bruinink, C. M.; Ravoo, B. J.; Reinhoudt, D. N.; Huskens, J. *Chem.—Eur. J.* **2005**, *11*, 2426–2432. (m) Mulder, A.; Onclin, S.; Péter, M.; Hoogenboom, J. P.; Beijleveld, H.; ter Maat, J.; García-Parajó, M. F.; Ravoo, B. J.; Huskens, J.; van Hulst, N. F.; Reinhoudt, D. N. *Small* **2005**, *1*, 242–253. (n) Nijhuis, C. A.; Yu, F.; Knoll, W.; Huskens, J.; Reinhoudt, D. N. *Langmuir* **2005**, *21*, 7866–7876.
- (2) Simonson, T. In *Computational Biochemistry and Biophysics*; Becker, O. M.; MacKerell, A. D.; Roux, B.; Watanabe, M., Eds.; Marcel Dekker Inc.: New York, 2001.
- (3) Simonson, T.; Archontis, G.; Karplus, M. *Acc. Chem. Res.* **2002**, *35*, 430–437.
- (4) Thompson, D.; Plateau, P.; Simonson, T. *ChemBioChem* **2006**, *7*, 337–344.
- (5) See, for example, the papers in ref 1 and references therein.
- (6) Lipkowitz, K. B. *Chem Rev.* **1998**, *98*, 1829–1874.
- (7) Damodaran, K. V.; Banba, S.; Brooks, C. L. *J. Phys. Chem. B* **2001**, *105*, 9316–9322.
- (8) Bea, I.; Cervello, E.; Kollman, P. A.; Jaime, C. *Comb. Chem. High Throughput Screening* **2001**, *4*, 605–611.
- (9) Varady, J.; Wu, X.; Wang, S. *J. Phys. Chem. B* **2002**, *106*, 4863–4872.
- (10) Chen, W.; Chang, C.; Gilson, M. K. *Biophys. J.* **2004**, *87*, 3035–3049.
- (11) Hirata K.; Uekama K.; Otagiri M.; Hirayama. F.; Ohtani Y. *Bull. Chem. Soc. Jpn.* **1985**, *58*, 1234–1238.
- (12) MacKerrell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, D. L.; Evanseck, J. D.; Field, M. J. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
- (13) Kleywegt, G. (Uppsala University); Adams, P. (Yale University). Online X-PLOR Topology and Parameter Library for Hetero Compounds. <http://atb.csb.yale.edu/xplor/hetero/hetero.html>.
- (14) Thompson, N. D.; Landis, C. R.; Bosnich, B. *J. Am. Chem. Soc.* **1992**, *114*, 7264–7272.
- (15) Baik, M.-H.; Crystal, J. B.; Friesner, R. A. *Inorg. Chem.* **2002**, *41*, 5926–5927.
- (16) Jorgensen, W.; Chandrasekar, J.; Madura, J.; Impey, R.; Klein, M. *J. Chem. Phys.* **1983**, *79*, 926–935.
- (17) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327–341.
- (18) Brooks, B. R.; Brucoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187–217.
- (19) Archontis, G.; Simonson, T.; Moras, D.; Karplus, M. *J. Mol. Biol.* **1998**, *275*, 823–846.
- (20) Archontis, G.; Simonson, T. *J. Am. Chem. Soc.* **2001**, *123*, 11047–11056.
- (21) Archontis, G.; Simonson, T.; Karplus, M. *J. Mol. Biol.* **2001**, *306*, 307–327.
- (22) Simonson, T. *Mol. Phys.* **1993**, *80*, 441–447.
- (23) Linder, K.; Saenger, W. *Carbohydr. Res.* **1982**, *99*, 103–115.
- (24) (a) Matsue, T.; Evans, D. H.; Osa, T.; Kobayashi, N. *J. Am. Chem. Soc.* **1985**, *107*, 3411–3417. (b) Kaifer, A. E. *Acc. Chem. Res.* **1999**, *32*, 62–71. (c) Osella, D.; Carretta, A.; Nervi, C.; Ravera, M.; Gobetto, R. *Organometallics* **2000**, *19*, 2791–2797.
- (25) Fersht, A. R. *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding*; W. H. Freeman and Company: New York, 1999.
- (26) Chang, C.-E.; Gilson, M. K. *J. Am. Chem. Soc.* **2004**, *126*, 13156–13164.