Supplemental Information

Large scale affinity calculations of cyclodextrin host-guest Complexes: Understanding the role of reorganization in the molecular recognition process

Lauren Wickstrom, Peng He, Emilio Gallicchio, Ronald M. Levy

BioMaPS Institute for Quantitative Biology and Department of Chemistry and Chemical Biology, Rutgers University, Piscataway, NJ 08854.

Methods

Convergence analysis

In order to correctly access the quality of an effective potential, there should be confidence that the sampling algorithm is generating representative conformational ensembles of the bound and unbound states for that force field as opposed to representing unequilibrated samples due to quasi-ergodic behavior of the simulations. To assess this, we have evaluated the convergence of the free energy calculations using several different measures: the overlap of the binding energy distributions, trajectories in λ- and binding energy space and the binding free energy as a function of the simulation time. As an example, we included convergence analysis for the beta-cyclodextrin+1-butanol host-guest system.

The number of replicas and the λ-assignments influences the convergence of the free energy calculations. For a valid calculation of the binding free energy, the binding energy distributions ranging between λ=0 and λ=1 should, at a minimum, overlap pairwise. Figure S2 shows the binding energy distribution ranging from λ=0.1 to λ=1.0 for 1-butanol for the H-REMD simulations with 16 replicas. For the β-CD+1-butanol BEDAM simulation, the binding energy distributions show an excellent overlap between consecutive λ values. We also show the replica diffusion through λ and binding energy space (Figures S3 and S4). The ideal case is where all replica-walkers explore all possible λ values to guarantee the robustness of the calculation. The replica walkers make several trips through lambda space and explore every lambda value multiple times (Figure S4). In order to investigate conformational transitions, we monitored the time evolution of binding energy for the replicas. Replicas in the simulations of the β-CD+1-butanol complex sample a wide range of binding energies (Figure S4) and do not

appear to have a dominant bound conformation based on the width of the binding energy distribution in Figure S2. Lastly, we examined the time evolution of the binding free energy for a series of windows collected for the BEDAM simulations (Figure S5) for these host-guest systems. The first nanosecond of each host-guest simulation was collected in 0.5 ns windows while the last 4 ns was collected at lower temporal resolution in 1 ns increments which resulted in 5,000 time-contiguous binding energy samples. For the β-CD+1-butanol complex, the time dependent profile shows an initial equilibration period between 0 and 0.5 ns with a binding free energies of approximately -1.5 kcal/mol. After this brief lag phase, the time profile flattens and fluctuates between binding free energies between -1.35 and -1.25 kcal/mol for the duration of the simulation. This behavior indicates that the simulation has converged to a stable binding free energy estimate.

Relationship between PC1/PC2 and compensating/reinforcing components

As illustrated in Figure 6 of the main text, the PC1 and PC2 axes are obtained by principal component analysis of the $\Delta E_{bind}/\Delta G^{\circ}_{reorg}$ data (see Methods). The coordinates of the unitary vectors along the PC1 and PC2 axes are (x_1, x_2) and (y_1, y_2), where $x_1 = 0.801$, $x_2 = -0.598$, and $y_1 = 0.598$, $y_2 = 0.801$, respectively. The set of coordinates (*PC1*,*PC2*), obtained by projecting $(\Delta E_{bind}, \Delta G^{\circ}_{reorg})$ pairs along the PC1 and PC2 axes, constitute an alternative representation of the binding energy/reorganization free energy data in terms of statistically uncorrelated variables (see Methods). A linear relationship exists between these two representations:

$$\Delta E_{bind} = x_1 PC1 + y_1 PC2 \qquad (1)$$

$$\Delta G^{\circ}_{reorg} = x_2 PC1 + y_2 PC2 \qquad (2)$$

where the values of the coefficients are given above. By summing Eqs.1 and 2 we obtain the following decomposition of the standard free energies of binding:

$$\Delta G^{\circ}_{b} = \Delta G_{comp} + \Delta G_{reinf} \quad (3)$$

where $\Delta G_{comp} = (x_1 + x_2)PC1 \approx 0.2PC1$ is the "compensating" component of the binding free energy, while $\Delta G_{reinf} = (y_1 + y_2)PC2 \approx 1.4PC2$ is the corresponding "reinforcement" component.

ANOVA Tests of Significance

We list below the results of Analysis of Variance (ANOVA) tests comparing single variable linear regressions of the experimental binding free energies versus computed binding energies ΔE_{bind} and compensating free energy components ΔG_{comp}, and two-variable models also including reorganization free energies ΔG°_{reorg} and reinforcing free energy components ΔG_{reinf}.

The tests were performed for the complexes with the rigid (number of rotatable bonds < 2) and flexible guests (number of rotatable bonds > 2). For each model we list the residual sum of squares (RSS) and the p-value measuring the statistical significance of the reduction of the RSS due to the addition of the second regression variable. The p-value is computed by means of an F-test.[1] A p-value smaller than 0.05 is considered statistically significant.

Binding energy/reorganization

Model1: $\Delta G_{exp} = c_0 + c_1 \Delta E_{bind}$ (1 variable)

Model2: $\Delta G_{exp} = c_0 + c_1 \Delta E_{bind} + c_2 \Delta G^{\circ}_{reorg}$ (2 variables)

		RSS	p-value
Rigid Guests			
	Model 1	7.44	
	Model 2	6.27	0.09
Flexible Guests			
	Model 1	8.27	
	Model 2	3.17	$8.54\ 10^{-6}$

Compensating/reinforcing

Model 1: $\Delta G_{exp} = c_0 + c_1 \Delta G_{comp}$ (1 variable)

Model 2: $\Delta G_{exp} = c_0 + c_1 \Delta G_{comp} + c_2 \Delta G_{reinf}$ (2 variables)

		RSS	p-value
Rigid Guests			
	Model 1	9.94	
	Model 2	6.27	0.005
Flexible Guests			
	Model 1	11.2	
	Model 2	3.17	$6.66\ 10^{-7}$

Development of hydration sites in β-cyclodextrin

We developed hydration correction parameters for the water sites in the β-CD cavity (Figure S1) using a training set of five β-CD host-guest systems previously studied by Chen et al (Table S6).[2] The training set contained guests with hydrophobic moieties, alcohol, ketone and carboxylic acid functional groups. We evaluated the parameters using the RMS_{error} of the calculated binding free energy relative to experimental binding affinity and the percentage of correct predictions (where the calculated binding free energy is within 2.0 kcal/mol of the

experimental binding affinity). The hydration strength of the water sites was initially set to h_w=+0.4 kcal/mol based on our previous work with host-guest systems.[3] Overall, the binding affinities were less favorable than the experiment (average calculated and experimental binding affinities for this training set was -2.38 kcal/mol and -3.7 kcal/mol respectively) with a RMS_{error} of 1.73 kcal/mol. The most problematic guests were resorcinol and flurbiprofen, whose binding free energies were underestimated by more than 2.0 kcal/mol.

In order to improve the agreement between the calculations and experiment, adjustments were made to the h_w parameter in order to make binding more favorable. The h_w parameter was increased to a value of +0.5 for each interior hydration site located in the β-CD cavity. By increasing this hydration site parameter, the unbound state was penalized more for an unoccupied hydration site which should shift binding affinities to more favorable values (a shift of -1.4 kcal/mol for every shift in the h_w parameter by +.1 kcal/mol). Using h_w = +0.5 kcal/mol for each hydration site, the RMS_{error} and % correct predictions shifted to 1.37 kcal/mol and 80%, respectively and the largest outlier in the set was the β-CD+resorcinol host-guest system which deviated approximately 2.0 kcal/mol away from the experimental affinity. Overall, the set deviated from the experimental affinity by 0.85 kcal/mol. Using h_w= +0.6 kcal/mol for each hydration site, the RMS_{error} and % correct predictions shifted to more favorable values of 1.10 kcal/mol and 100% respectively. Furthermore, the average calculated binding affinity deviated by only .2 kcal/mol away from the experimental binding affinity. As a result, the hydration strength of these water sites was set to h_w=+0.6 kcal/mol and applied to the rest of the β-CD host-guest systems included in this study.

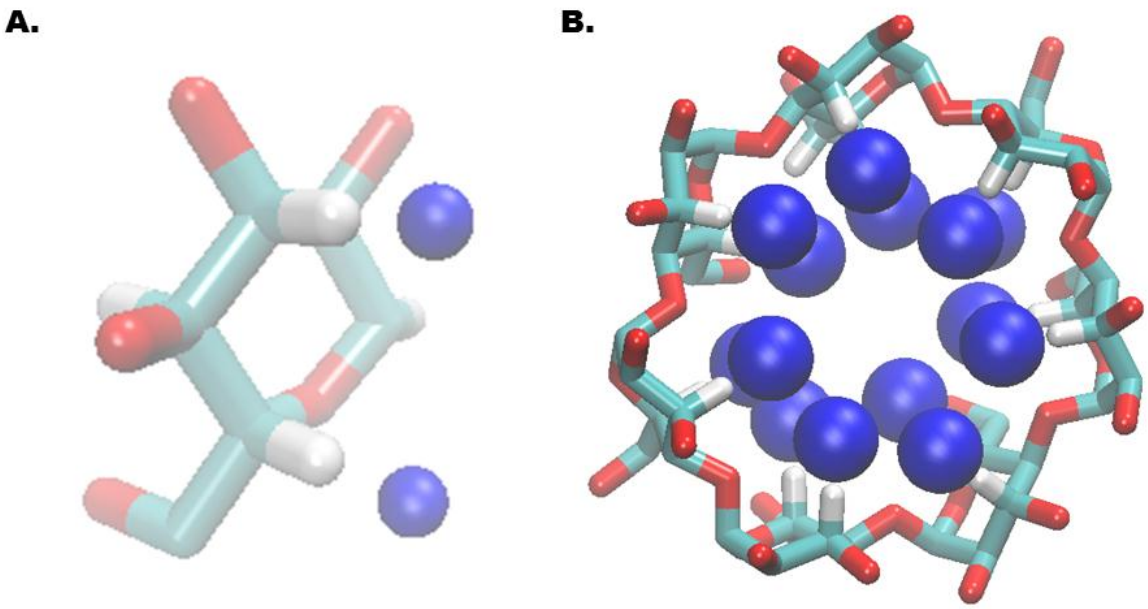


Figure S1.(A) Positioning of the auxillary hydration sites (blue spheres) relative to the glucose monomer and (B) representative conformation of β-cyclodextrin show with the 14 auxiliary interior hydration sites; a guest bound in the interior occupies the hydration sites, mimicking the water expulsion process.

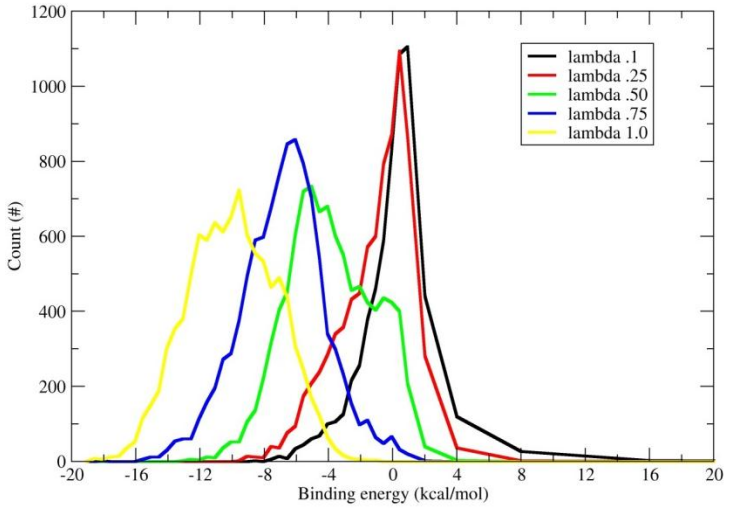


Figure S2. Binding energy distributions for β-CD+1-butanol host-guest system.

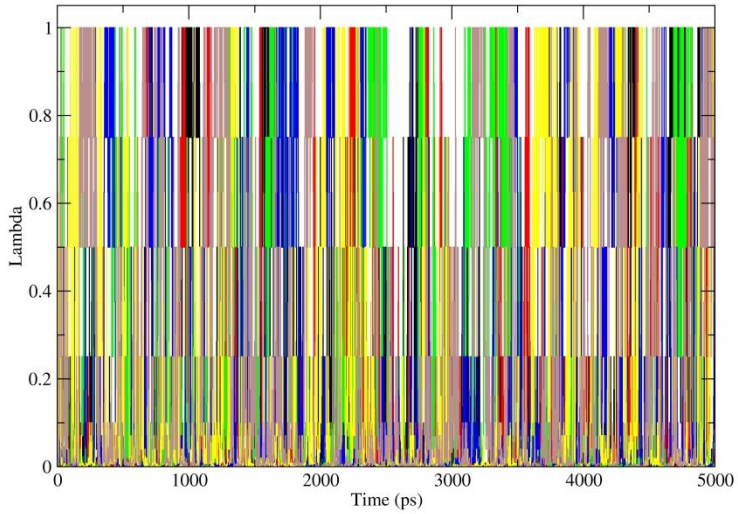


Figure S3. Time trajectories in λ-space of six replicas of the β-CD+1-butanol host-guest system. Each color represents a replica. The degree of ‘‘color mixing’’represents the rate at which replicas diffuse in lambda-space.

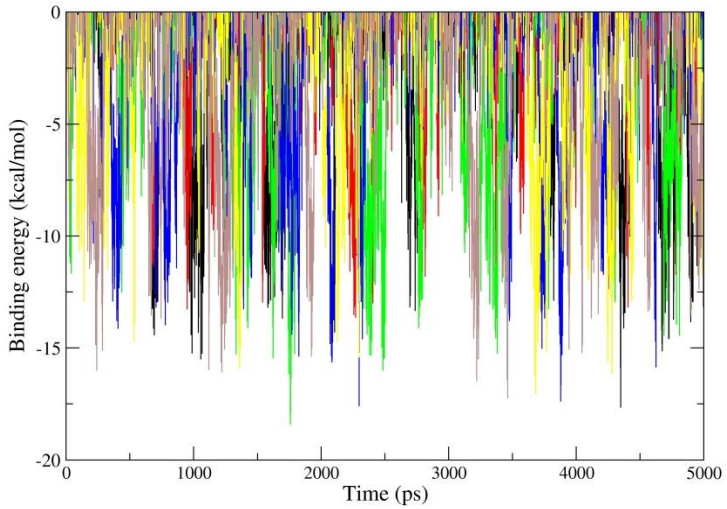


Figure S4. Time trajectories of the binding energies of four representative replicas of the β-CD+1-butanol host-guest system.

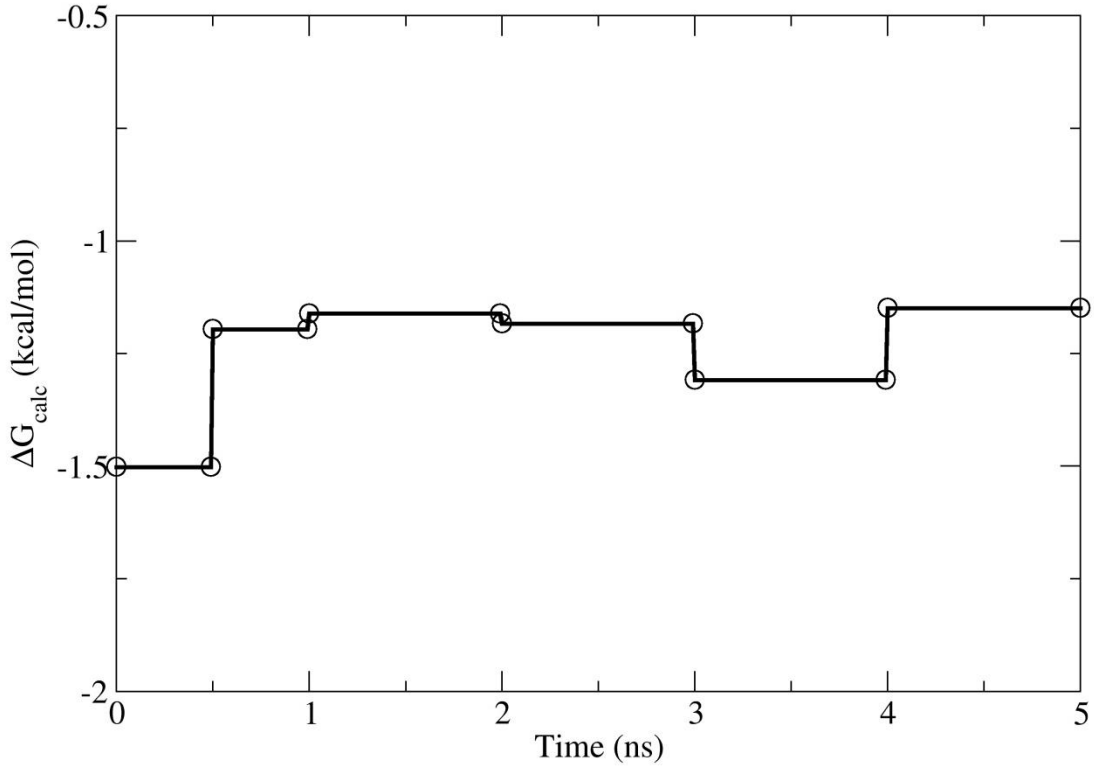
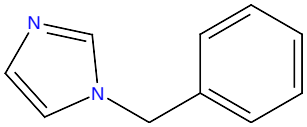
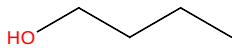


Figure S5. Binding free energy estimate vs simulation time for β-CD+1-butanol host-guest system .

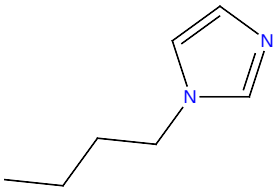
Table S1. 2D chemical structures of the 57 guests.



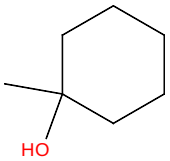
name: 1-benzylimidazole



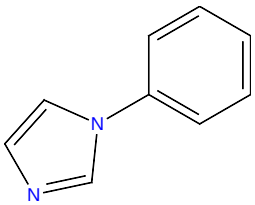
name: 1-butanol



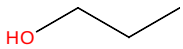
name: 1-butylimidazole



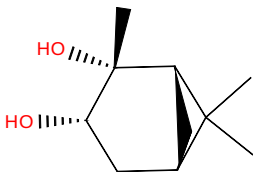
name: 1-methylcyclohexanol



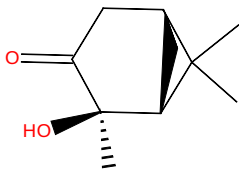
name: 1-phenylimidazole



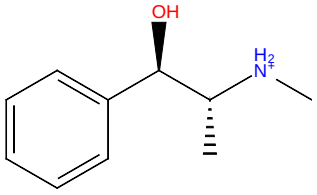
name: 1-propanol



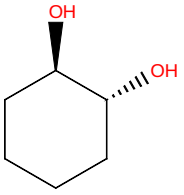
name: 1R,2R,3S,5R-pinanediol



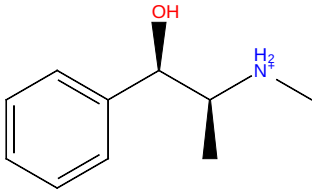
name: 1R,2R,5R-2-hydroxy-3-pinanone



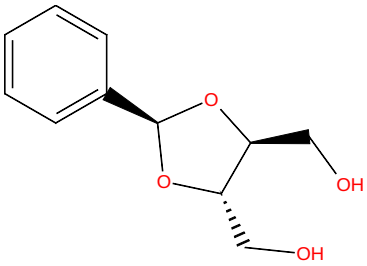
name: 1R,2R-pseudoephedrine



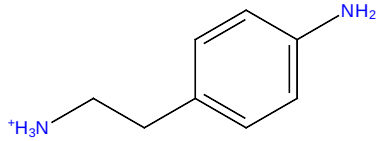
name: 1R,2R-trans-1,2-cyclohexanediol



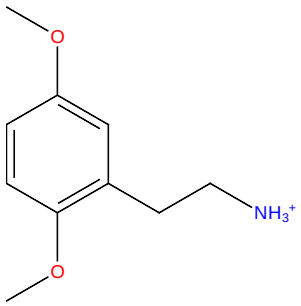
name: 1R,2S-ephedrine



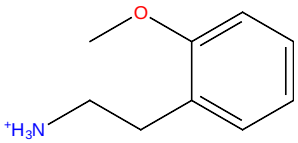
name: 2,3-O-benzylidene-L-threitol



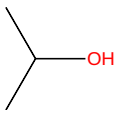
name: 2,4-aminophenylethylammonium



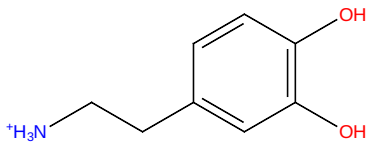
name: 2,5-dimethoxyphenethylammonium



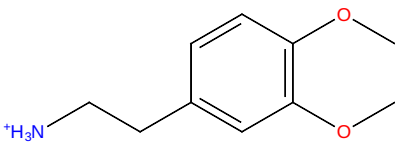
name: 2-methoxyphenethylammonium



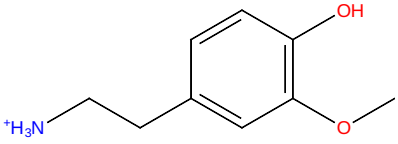
name: 2-propanol



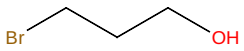
name: 3,4-dihydroxyphenethylammonium



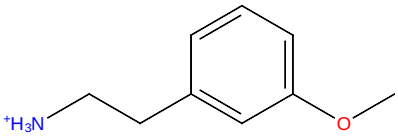
name: 3,4-dimethoxyphenethylammonium



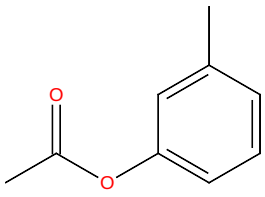
name: 3-O-methyldopamine



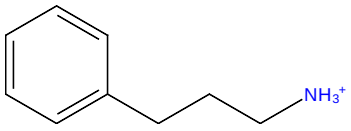
name: 3-bromo-1-propanol



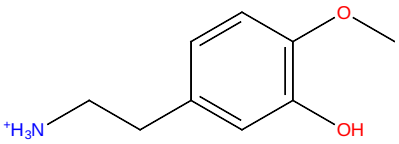
name: 3-methoxyphenethylammonium



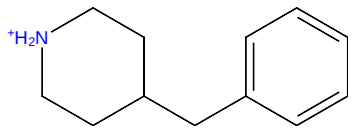
name: 3-methylphenylacetate



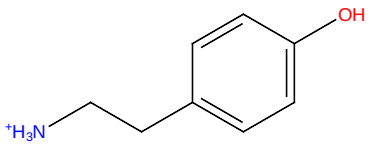
name: 3-phenylpropylammonium



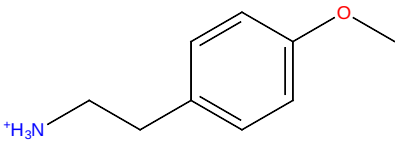
name: 4-O-methyldopamine



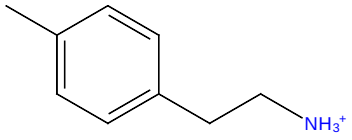
name: 4-benzylpiperidine



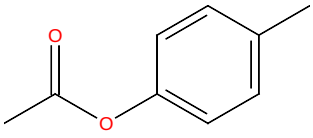
name: 4-hydroxyphenethylammonium



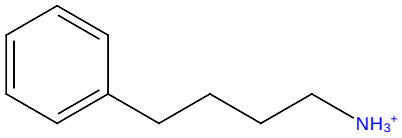
name: 4-methoxyphenethylammonium



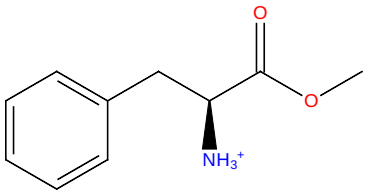
name: 4-methylphenethylammonium



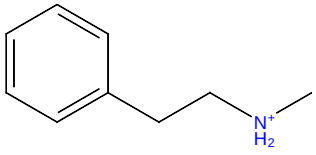
name: 4-methylphenylacetate



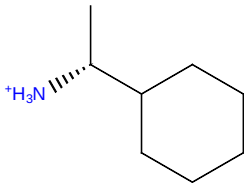
name: 4-phenylbutylammonium



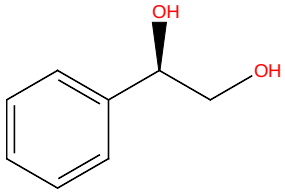
name: L-phenylalaninemethyl-ester



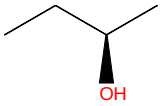
name: N-methylphenethylammonium



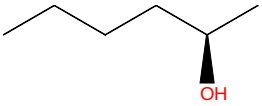
name: R-1-cyclohexylethylamine



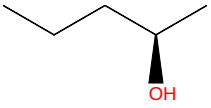
name: R-1-phenyl-1,2-ethanediol



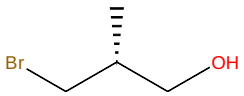
name: R-2-butanol



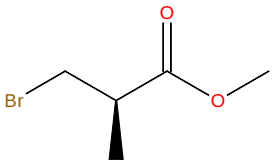
name: R-2-hexanol



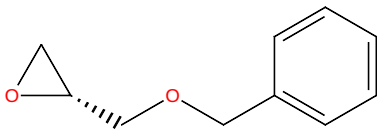
name: R-2-pentanol



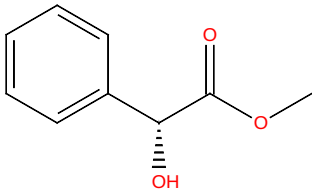
name: R-3-bromo-2-methyl-1-propanol



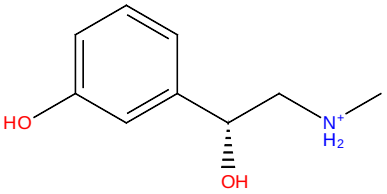
name: R-3-bromo-2-methylpropionic-aci...



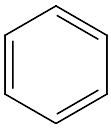
name: R-benzylglycidyl-ether



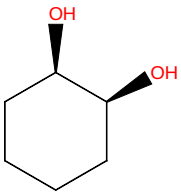
name: R-mandelic-acid-methyl-ester



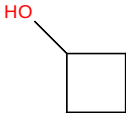
name: R-phenylephrine



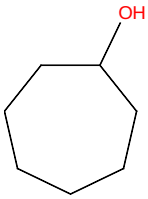
name: benzene



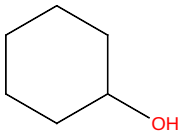
name: cis-1,2-cyclohexanediol



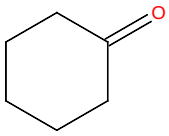
name: cyclobutanol



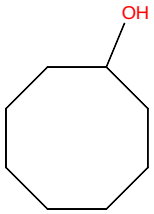
name: cycloheptanol



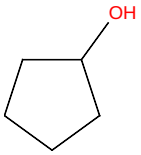
name: cyclohexanol



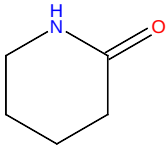
name: cyclohexanone



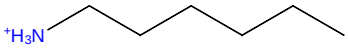
name: cyclooctanol



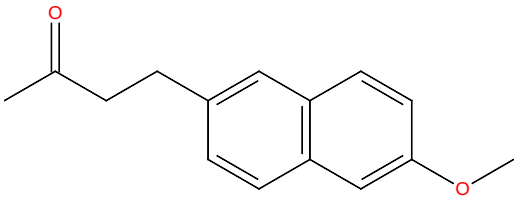
name: cyclopentanol



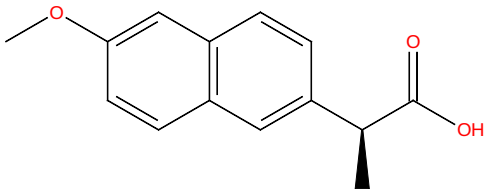
name: delta-valerolactam



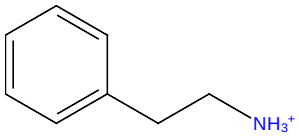
name: hexylammonium



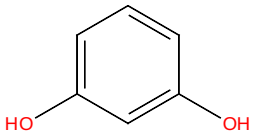
name: nabumetone



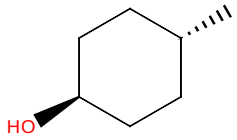
name: naproxen



name: phenethylammonium



name: resorcinol



name: trans-4-methylcyclohexanol

Table S2A. Calculated and experimental binding free energies for the 57 host-guest systems. The table also lists thermodynamic decomposition of the binding free energies into binding energies and reorganization free energy.

Name	# rb	ΔG°_{exp}	ΔG°_{calc}**(error)**	ΔE_{bind}**(error)**	ΔG°_{reorg}**(error)**
2-propanol	0	-0.57	-0.98(0.07)	-8.34(0.18)	7.36(0.25)
3-O-methyldopamine	3	-0.86	0.58(0.09)	-10.44(0.22)	11.02(0.31)
1-propanol	1	-0.88	-0.94(0.05)	-8.43(0.10)	7.49(0.15)
2-methoxyphenethylammonium	3	-1.23	0.28(0.04)	-9.79(0.22)	10.07(0.26)
3-methylphenylacetate	2	-1.46	-2.74(0.07)	-13.45(0.11)	10.71(0.18)
L-phenylalaninemethyl-ester	4	-1.48	1.22(0.09)	-9.15(0.36)	10.37(0.45)
R-2-butanol	1	-1.53	-1.43(0.05)	-9.89(0.17)	8.46(0.22)
Cyclobutanol	0	-1.55	-2.19(0.05)	-10.06(0.10)	7.87(0.15)
3-4-dimethoxyphenethyl-ammonium	4	-1.56	-0.32(0.05)	-11.03(0.17)	10.71(0.22)
1-butanol	2	-1.65	-1.21(0.05)	-9.95(0.11)	8.74(0.16)
Phenethylammonium	2	-1.78	-0.71(0.09)	-9.64(0.21)	8.92(0.30)
N-methylphenethylammonium	3	-1.82	-1.45(0.09)	-12.52(0.24)	11.07(0.33)
3-bromo-1-propanol	2	-1.84	-1.58(0.07)	-10.56(0.14)	8.99(0.21)
delta-valerolactam	0	-1.91	-3.76(0.07)	-13.51(0.10)	9.75(0.17)
1-phenylimidazole	1	-1.91	-0.01(0.07)	-12.24(0.23)	12.23(0.30)
3-4-dihydroxyphenethyl-ammonium	2	-2.02	0.48(0.09)	-11.81(0.29)	12.29(0.38)
R-2-pentanol	2	-2.08	-1.41(0.06)	-11.24(0.22)	9.83(0.28)
2-5-dimethoxyphenethyl-ammonium	4	-2.14	-0.11(0.08)	-9.98(0.11)	9.87(0.19)
R-phenylephrine	3	-2.18	0.54(0.08)	-12.15(0.26)	12.69(0.34)
4-methylphenylacetate	2	-2.19	-2.83(0.07)	-12.58(0.07)	9.75(0.14)
4-O-methyldopamine	3	-2.34	-0.56(0.08)	-11.66(0.08)	11.10(0.16)
2-4-aminophenylethyl-ammonium	2	-2.37	0.01(0.08)	-9.41(0.27)	9.43(0.35)
1R,2S-ephedrine	3	-2.42	-0.76(0.10)	-13.62(0.29)	12.86(0.39)
R-1-phenyl-1-2-ethanediol	2	-2.45	-2.05(0.08)	-14.19(0.23)	12.14(0.31)
4-hydroxyphenethylammonium	2	-2.46	0.07(0.07)	-10.69(0.18)	10.76(0.25)
3-methoxyphenethylammonium	3	-2.49	-0.76(0.09)	-11.26(0.20)	10.51(0.29)
R-mandelic-acid-methyl-ester	3	-2.49	-1.41(0.09)	-12.69(0.10)	11.29(0.19)
Hexylammonium	4	-2.49	-0.56(0.07)	-9.59(0.26)	9.02(0.33)
1R,2R-pseudoephedrine	3	-2.51	-1.07(0.10)	-15.79(0.38)	14.71(0.48)
4-methylphenethylammonium	2	-2.56	-1.46(0.06)	-10.24(0.11)	8.78(0.17)
4-methoxyphenethylammonium	3	-2.58	-1.01(0.06)	-10.39(0.13)	9.38(0.19)
1R-2R-trans-1-2-cyclohexanediol	0	-2.63	-1.91(0.08)	-14.04(0.32)	12.13(0.40)
3-phenylpropylammonium	3	-2.7	-1.73(0.10)	-12.55(0.23)	10.81(0.33)
Resorcinol	0	-2.77	-1.47(0.05)	-11.95(0.16)	10.48(0.21)
Benzene	0	-2.77	-3.69(0.07)	-10.43(0.04)	6.75(0.11)
2-3-O-benzylidene-L-threitol	3	-2.81	-1.74(0.07)	-16.53(0.56)	14.79(0.63)
R-2-hexanol	3	-2.82	-1.90(0.09)	-13.12(0.39)	11.23(0.48)
R-3-bromo-2-methyl-1-propanol	2	-2.94	-2.34(0.09)	-13.24(0.24)	10.90(0.33)

1-butylimidazole	3	-2.99	-2.68(0.09)	-13.77(0.15)	11.08(0.24)
Cyclopentanol	0	-3.05	-2.87(0.04)	-11.91(0.12)	9.03(0.16)
R-benzylglycidyl-ether	4	-3.23	-2.84(0.10)	-13.73(0.16)	10.88(0.26)
R-3-bromo-2-methylpropionic-acid-methyl-ester	3	-3.3	-1.75(0.05)	-11.41(0.10)	9.67(0.15)
cis-1-2-cyclohexanediol	0	-3.32	-2.27(0.09)	-14.13(0.21)	11.86(0.30)
R-1-cyclohexylethylamine	1	-3.44	-0.70(0.15)	-13.07(0.20)	12.37(0.35)
4-phenylbutylammonium	4	-3.56	-2.23(0.09)	-12.08(0.22)	9.86(0.31)
1-benzylimidazole	2	-3.57	-3.10(0.08)	-14.69(0.24)	11.59(0.32)
Cyclohexanone	0	-3.71	-3.15(0.09)	-12.36(0.07)	9.22(0.16)
Cyclohexanol	0	-3.88	-3.13(0.08)	-13.78(0.18)	10.65(0.26)
1-methylcyclohexanol	0	-4.18	-3.18(0.10)	-14.52(0.29)	11.35(0.39)
Naproxen	3	-4.33	-4.85(0.13)	-20.02(0.28)	15.17(0.41)
4-benzylpiperidine	2	-4.5	-2.97(0.18)	-17.12(0.41)	14.15(0.59)
trans-4-methylcyclohexanol	0	-4.54	-3.06(0.10)	-14.41(0.23)	11.35(0.33)
Cycloheptanol	0	-4.56	-3.29(0.06)	-15.53(0.31)	12.24(0.37)
Nabumetone	4	-4.59	-4.31(0.08)	-16.54(0.16)	12.23(0.24)
1R,2R,5R-2-hydroxy-3-pinanone	0	-4.62	-2.24(0.09)	-16.26(0.21)	14.02(0.30)
Cyclooctanol	0	-4.97	-3.29(0.09)	-16.78(0.33)	13.48(0.42)
1R,2R,3S,5R-pinanediol	0	-5.2	-2.60(0.16)	-19.35(0.46)	16.75(0.62)

$\Delta \boldsymbol{G}^{\circ}_{\boldsymbol{exp}}$, the experimental binding affinity; $\Delta \boldsymbol{G}^{\circ}_{\boldsymbol{calc}}$, the calculated binding free energy; ΔE_{bind}, the binding energies; $\Delta \boldsymbol{G}^{\circ}_{\boldsymbol{reorg}}$, the reorganization free energy; and #rb, number of rotatable bonds. All values are expressed in kcal/mol.

Table S2b. Calculated and experimental binding free energies for the 18 small guests (aromatics, alkanols, ethers and imidazoles) from Table S1A. See section B of main text.

Name	ΔG°_{exp}	ΔG°_{calc} **(error)**
2-propanol	-0.57	-0.98(0.07)
1-propanol	-0.88	-0.94(0.05)
R-2-butanol	-1.53	-1.43(0.05)
Cyclobutanol	-1.55	-2.19(0.05)
1-butanol	-1.65	-1.21(0.05)
3-bromo-1-propanol	-1.84	-1.58(0.07)
R-2-pentanol	-2.08	-1.41(0.06)
R-1-phenyl-1,2-ethanediol	-2.45	-2.05(0.08)
1R,2R-trans-1,2-cyclohexanediol	-2.63	-1.91(0.08)
Benzene	-2.77	-3.69(0.07)
R-2-hexanol	-2.82	-1.90(0.09)
R-3-bromo-2-methyl-1-propanol	-2.94	-2.34(0.09)
1-butylimidazole	-2.99	-2.68(0.09)
Cyclopentanol	-3.05	-2.87(0.04)
R-benzylglycidyl-ether	-3.23	-2.84(0.10)
cis-1,2-cyclohexanediol	-3.32	-2.27(0.09)
1-benzylimidazole	-3.57	-3.10(0.08)
Cyclohexanol	-3.88	-3.13(0.08)
RMSD		0.74
Spearman rho		0.84

ΔG°_{exp}, the experimental binding free energy; ΔG°_{calc}, the calculated binding free energy; and RMSD, root mean squared deviation between the calculated and experimental binding free energies. All values are expressed in kcal/mol.

Table S3. Populations of the binding modes for the set of β-CD host-guest systems.

Guests	DS	H-bond DS	US	H-bond US
2-propanol	0.41	0.34	0.59	0.35
3-O-methyldopamine	0.11	0.10	0.89	0.01
1-propanol	0.34	0.26	0.66	0.45
2-methoxyphenethylammonium	0.05	0.05	0.95	0.04
3-methylphenylacetate	0.18	0.13	0.82	0.72
L-phenylalaninemethyl-ester	0.55	0.53	0.45	0.06
R-2-butanol	0.41	0.32	0.59	0.36
Cyclobutanol	0.42	0.35	0.58	0.38
3,4-dimethoxyphenethylammonium	0.11	0.08	0.89	0.01
1-butanol	0.36	0.33	0.64	0.53
Phenethylammonium	0.47	0.41	0.53	0.01
N-methylphenethylammonium	0.82	0.78	0.18	0.01
3-bromo-1-propanol	0.57	0.52	0.43	0.32
1-phenylimidazole	0.09	0.07	0.91	0.86
delta-valerolactam	0.12	0.09	0.88	0.78
3,4-dihydroxyphenethylammonium	0.32	0.26	0.68	0.02
R-2-pentanol	0.56	0.51	0.44	0.32
2,5-dimethoxyphenethylammonium	0.06	0.04	0.94	0.02
R-phenylephrine	0.62	0.46	0.38	0.01
4-methylphenylacetate	0.18	0.12	0.82	0.68
4-O-methyldopamine	0.23	0.20	0.77	0.00
2,4-aminophenylethylammonium	0.42	0.36	0.58	0.01
1R,2S-ephedrine	0.93	0.63	0.07	0.00
R-1-phenyl-1,2-ethanediol	0.45	0.41	0.55	0.28
4-hydroxyphenethylammonium	0.69	0.61	0.31	0.01
Hexylammonium	0.75	0.65	0.25	0.01
3-methoxyphenethylammonium	0.44	0.36	0.57	0.01
R-mandelic-acid-methyl-ester	0.10	0.09	0.90	0.86
1R,2R-pseudoephedrine	0.87	0.71	0.13	0.00
4-methylphenethylammonium	0.28	0.20	0.72	0.01
4-methoxyphenethylammonium	0.41	0.27	0.59	0.00
1R-2R-trans-1-2-cyclohexanediol	0.54	0.51	0.46	0.32
3-phenylpropylammonium	0.77	0.72	0.23	0.01
2-3-O-benzylidene-L-threitol	0.03	0.03	0.97	0.81
R-2-hexanol	0.67	0.64	0.33	0.25
R-3-bromo-2-methyl-1-propanol	0.73	0.68	0.27	0.20
1-butylimidazole	0.21	0.18	0.79	0.67

Cyclopentanol	0.30	0.25	0.70	0.24
R-benzylglycidyl-ether	0.24	0.14	0.76	0.33
R-3-bromo-2-methylpropionic-acid-methyl-ester	0.32	0.26	0.68	0.52
cis-1-2-cyclohexanediol	0.57	0.51	0.43	0.27
R-1-cyclohexylethylamine	0.88	0.87	0.12	0.01
4-phenylbutylammonium	0.51	0.33	0.49	0.01
1-benzylimidazole	0.69	0.49	0.31	0.24
Cyclohexanone	0.22	0.16	0.78	0.59
Cyclohexanol	0.48	0.44	0.52	0.36
1-methylcyclohexanol	0.41	0.35	0.59	0.40
Naproxen	0.84	0.79	0.16	0.08
4-benzylpiperidine	0.89	0.85	0.12	0.00
trans-4-methylcyclohexanol	0.68	0.66	0.32	0.19
Cycloheptanol	0.35	0.33	0.65	0.23
Nabumetone	0.18	0.11	0.82	0.01
1R,2R,5R-2-hydroxy-3-pinanone	0.25	0.24	0.75	0.61
Cyclooctanol	0.53	0.51	0.47	0.34
1R,2R,3S,5R-pinanediol	0.74	0.72	0.27	0.20

Binding modes are defined as down-state (DS) and up-state binding mode (US). In the down-state mode, the guest has its polar functional group pointed toward the primary alcohols. In the up-state binding mode, the polar group points towards the secondary alcohols. The h-bond DS is the down-state mode where a hydrogen bond is formed between the primary alcohols and the polar functional group on the guest. The h-bond US is the up-state mode where a hydrogen bond is formed between the secondary alcohols and the polar functional group on the guest. A hydrogen bond is considered formed when the distance between the donor and acceptor heavy atoms is less than 4.0 Å.

Table S4. Principal component analysis of the binding free energies for the 57 host-guest systems. The table lists the calculated binding free energy, reinforcing and compensating free energy. See section D of main text.

Name	ΔG°_{calc} **(error)**	ΔG_{comp}**(error)**	ΔG_{reinf}**(error)**
2-propanol	-0.98(0.07)	-2.22(0.29)	1.26(0.31)
3-O-methyldopamine	0.58(0.09)	-2.99(0.36)	3.62(0.38)
1-propanol	-0.94(0.05)	-2.25(0.17)	1.34(0.18)
2-methoxyphenethylammonium	0.28(0.04)	-2.77(0.33)	3.09(0.34)
3-methylphenylacetate	-2.74(0.07)	-3.44(0.20)	0.74(0.21)
L-phenylalaninemethyl-ester	1.22(0.09)	-2.71(0.56)	3.96(0.58)
R-2-butanol	-1.43(0.05)	-2.60(0.27)	1.20(0.28)
Cyclobutanol	-2.19(0.05)	-2.55(0.17)	0.40(0.18)
3-4-dimethoxyphenethyl-ammonium	-0.32(0.05)	-3.05(0.27)	2.77(0.28)
1-butanol	-1.21(0.05)	-2.64(0.18)	1.47(0.19)
Phenethylammonium	-0.71(0.09)	-2.61(0.35)	1.94(0.37)
N-methylphenethylammonium	-1.45(0.09)	-3.33(0.39)	1.93(0.41)
3-bromo-1-propanol	-1.58(0.07)	-2.77(0.24)	1.23(0.25)
delta-valerolactam	-3.76(0.07)	-3.33(0.18)	-0.38(0.20)
1-phenylimidazole	-0.01(0.07)	-3.42(0.36)	3.46(0.38)
3-4-dihydroxyphenethyl-ammonium	0.48(0.09)	-3.36(0.46)	3.90(0.48)
R-2-pentanol	-1.41(0.06)	-2.98(0.34)	1.61(0.36)
2-5-dimethoxyphenethyl-ammonium	-0.11(0.08)	-2.78(0.20)	2.71(0.22)
R-phenylephrine	0.54(0.08)	-3.47(0.41)	4.05(0.43)
4-methylphenylacetate	-2.83(0.07)	-3.18(0.14)	0.40(0.15)
4-O-methyldopamine	-0.56(0.08)	-3.20(0.16)	2.68(0.18)
2-4-aminophenylethyl-ammonium	0.01(0.08)	-2.64(0.43)	2.69(0.44)
1R,2S-ephedrine	-0.76(0.10)	-3.72(0.47)	3.02(0.49)
R-1-phenyl-1-2-ethanediol	-2.05(0.08)	-3.73(0.37)	1.73(0.39)
4-hydroxyphenethylammonium	0.07(0.07)	-3.00(0.29)	3.11(0.31)
3-methoxyphenethylammonium	-0.76(0.09)	-3.06(0.33)	2.35(0.35)
R-mandelic-acid-methyl-ester	-1.41(0.09)	-3.38(0.19)	2.03(0.21)
Hexylammonium	-0.56(0.07)	-2.62(0.41)	2.09(0.42)
1R,2R-pseudoephedrine	-1.07(0.10)	-4.29(0.59)	3.28(0.61)
4-methylphenethylammonium	-1.46(0.06)	-2.69(0.19)	1.27(0.20)
4-methoxyphenethylammonium	-1.01(0.06)	-2.79(0.22)	1.82(0.23)
1R-2R-trans-1-2-cyclohexanediol	-1.91(0.08)	-3.70(0.50)	1.85(0.51)
3-phenylpropylammonium	-1.73(0.10)	-3.30(0.38)	1.62(0.40)
Resorcinol	-1.47(0.05)	-3.17(0.25)	1.74(0.26)
Benzene	-3.69(0.07)	-2.48(0.10)	-1.17(0.11)
2-3-O-benzylidene-L-threitol	-1.74(0.07)	-4.42(0.83)	2.75(0.84)
R-2-hexanol	-1.90(0.09)	-3.45(0.60)	1.60(0.62)
R-3-bromo-2-methyl-1-propanol	-2.34(0.09)	-3.43(0.39)	1.14(0.41)

1-butylimidazole	-2.68(0.09)	-3.53(0.26)	0.90(0.28)
Cyclopentanol	-2.87(0.04)	-2.99(0.19)	0.16(0.20)
R-benzylglycidyl-ether	-2.84(0.10)	-3.50(0.28)	0.71(0.30)
R-3-bromo-2-methylpropionic-acid-methyl-ester	-1.75(0.05)	-2.99(0.17)	1.28(0.18)
cis-1-2-cyclohexanediol	-2.27(0.09)	-3.68(0.35)	1.46(0.37)
R-1-cyclohexylethylamine	-0.70(0.15)	-3.57(0.37)	2.93(0.40)
4-phenylbutylammonium	-2.23(0.09)	-3.12(0.36)	0.94(0.38)
1-benzylimidazole	-3.10(0.08)	-3.74(0.38)	0.70(0.40)
Cyclohexanone	-3.15(0.09)	-3.08(0.15)	-0.02(0.17)
Cyclohexanol	-3.13(0.08)	-3.48(0.30)	0.40(0.32)
1-methylcyclohexanol	-3.18(0.10)	-3.69(0.47)	0.56(0.49)
Naproxen	-4.85(0.13)	-5.02(0.47)	0.24(0.50)
4-benzylpiperidine	-2.97(0.18)	-4.44(0.68)	1.53(0.72)
trans-4-methylcyclohexanol	-3.06(0.10)	-3.67(0.38)	0.66(0.40)
Cycloheptanol	-3.29(0.06)	-3.95(0.47)	0.72(0.48)
Nabumetone	-4.31(0.08)	-4.12(0.27)	-0.13(0.29)
1R,2R,5R-2-hydroxy-3-pinanone	-2.24(0.09)	-4.28(0.35)	2.10(0.37)
Cyclooctanol	-3.29(0.09)	-4.30(0.52)	1.07(0.53)
1R,2R,3S,5R-pinanediol	-2.60(0.16)	-5.11(0.74)	2.58(0.77)

Calculated values shown: $\Delta \boldsymbol{G}^{\circ}_{\boldsymbol{calc}}$, the calculated binding affinity; ΔG_{comp}, the compensating component of the binding free energy; and ΔG_{reinf}, the reinforcing component of the binding free energy. All values are expressed in kcal/mol.

Table S5. Decomposition of the reorganization free energy into strain energy and conformational entropy.

Name	ΔE_{strain}**(R)**	ΔE_{strain}**(L)**	$-T\Delta S^{\circ}_{conf}$	ΔG°_{reorg}
2-propanol	-0.10	0.18	7.29	7.36
3-O-methyldopamine	-0.83	-0.02	11.86	11.02
1-propanol	-0.80	0.01	8.28	7.49
2-methoxyphenethylammonium	-0.52	-0.05	10.63	10.07
3-methylphenylacetate	-0.71	0.33	11.09	10.71
L-phenylalaninemethyl-ester	-1.62	-0.12	12.11	10.37
R-2-butanol	-0.18	0.10	8.54	8.46
Cyclobutanol	-0.80	0.04	8.63	7.87
3-4-dimethoxyphenethylammonium	-0.75	0.38	11.07	10.71
1-butanol	-0.37	0.38	8.74	8.74
Phenethylammonium	-1.14	-0.22	10.28	8.92
N-methylphenethylammonium	-0.90	-0.08	12.05	11.07
3-bromo-1-propanol	-0.05	0.26	8.78	8.99
1-phenylimidazole	-0.35	0.04	12.54	12.23
delta-valerolactam	-0.29	0.32	9.73	9.75
3-4-dihydroxyphenethylammonium	-0.75	0.20	12.84	12.29
R-2-pentanol	0.20	0.26	9.37	9.83
2,5-dimethoxyphenethylammonium	-0.89	-0.09	10.84	9.87
R-phenylephrine	-0.97	0.01	13.65	12.69
4-methylphenylacetate	-1.15	-0.02	10.92	9.75
4-O-methyldopamine	-1.04	0.06	12.08	11.10
2,4-aminophenylethylammonium	-0.42	-0.14	10.01	9.43
1R-2S-ephedrine	-1.11	0.11	13.86	12.86
R-1-phenyl-1,2-ethanediol	0.00	0.15	11.99	12.14
4-hydroxyphenethylammonium	-0.68	0.43	11.00	10.76
3-methoxyphenethylammonium	-1.33	-0.08	11.91	10.51
Hexylammonium	-1.01	0.11	9.93	9.02
R-mandelic-acid-methyl-ester	-0.70	-0.28	12.27	11.29
1R,2R-pseudoephedrine	-1.05	0.56	15.21	14.71
4-methylphenethylammonium	-0.95	-0.04	9.77	8.78
4-methoxyphenethylammonium	-0.97	-0.04	10.39	9.38
1R,2R-trans-1,2-cyclohexanediol	0.34	0.26	11.53	12.13
3-phenylpropylammonium	-1.66	0.14	12.33	10.81
Benzene	-0.53	0.07	7.21	6.75
Resorcinol	1.41	0.37	8.71	10.48
2,3-O-benzylidene-L-threitol	-0.65	0.61	14.83	14.79

R-2-hexanol	-0.55	-0.23	12.01	11.23
R-3-bromo-2-methyl-1-propanol	-0.05	-0.09	11.03	10.90
1-butylimidazole	-1.45	0.13	12.40	11.08
Cyclopentanol	-0.33	0.07	9.28	9.03
R-benzylglycidyl-ether	-0.31	0.03	11.16	10.88
R-3-bromo-2-methylpropionic-acid-methyl-ester	-0.56	0.04	10.19	9.67
cis-1-2-cyclohexanediol	0.04	-0.12	11.95	11.86
R-1-cyclohexylethylamine	-1.63	0.08	13.92	12.37
4-phenylbutylammonium	-1.15	0.11	10.89	9.86
1-benzylimidazole	-0.28	0.23	11.63	11.59
Cyclohexanone	-0.53	0.35	9.40	9.22
Cyclohexanol	-0.04	0.12	10.57	10.65
1-methylcyclohexanol	-0.34	-0.09	11.78	11.35
Naproxen	-0.33	-0.01	15.50	15.17
4-benzylpiperidine	-1.95	-0.01	16.11	14.15
trans-4-methylcyclohexanol	-0.34	-0.06	11.75	11.35
Cycloheptanol	0.14	0.06	12.04	12.24
Nabumetone	-2.23	-0.39	14.85	12.23
1R-2R-5R-2-hydroxy-3-pinanone	0.11	0.04	13.87	14.02
Cyclooctanol	0.33	-0.10	13.25	13.48
1R,2R,3S,5R-pinanediol	1.41	0.27	15.08	16.75

Calculated values shown are $\Delta \boldsymbol{G}^{\circ}_{\boldsymbol{reorg}}$, the reorganization free energy; $\Delta \boldsymbol{E}_{\boldsymbol{strain}}$, the strain energy; and $-\boldsymbol{T}\Delta \boldsymbol{S}^{\circ}_{\boldsymbol{conf}}$, conformational entropy differences between the bound and unbound states. All values are expressed in kcal/mol.

Table S6. Comparison of calculated and experimental binding affinities for a training set of β-cyclodextrin host-guest systems using different hydration correction energies for water sites located inside the host.

h_w		**0.40**	**0.50**	**0.60**
Guests	ΔG°_{exp}	ΔG°_{calc}	ΔG°_{calc}	ΔG°_{calc}
Benzene	-2.77	-2.44	-3.07	-3.69
Resorcinol	-2.77	-0.28	-0.78	-1.47
Flurbiprofen	-4.97	-2.56	-2.92	-3.19
Naproxen	-4.33	-3.42	-3.79	-4.85
Nabumetone	-4.59	-3.18	-3.70	-4.31
RMSD		1.73	1.37	1.10
% correct		60	80	100
Max	-2.77	-0.28	-0.78	-1.47
Min	-4.97	-3.42	-3.79	-4.85
Average	-3.71	-2.38	-2.85	-3.50

h_w , hydration correction parameter; ΔG°_{exp}, the experimental binding affinity; ΔG°_{calc} , the calculated binding free energy; and RMSD, root mean squared deviation between the calculated and experimental binding free energies. These values are expressed in kcal/mol. % correct, the percentage of correct predictions (where the calculated binding free energy is within 2.0 kcal/mol of the experimental binding affinity).

References

1. Dalgaard, P., *Introductory Statistics with R* Springer: New York, 2008.

2. Chen, W.; Chang, C. E.; Gilson, M. K., Calculation of cyclodextrin binding affinities: energy, entropy, and implications for drug design. *Biophys. J.* **2004,** *87*, 3035-3049.

3. Gallicchio, E.; Levy, R. M., Prediction of SAMPL3 host-guest affinities with the binding energy distribution analysis method (BEDAM). *J. Comput.-Aided Mol. Des.* **2012,** *26*, 505-516.