Supporting Information for:

Exploring pH Dependent Host/Guest Binding Affinities

Thomas J. Paul,[†] Jonah Z. Vilseck,[†] Ryan L. Hayes,[†] and Charles L. Brooks III[†‡]

[†]*Department of Chemistry,* [‡]*Department of Biophysics, University of Michigan, Ann Arbor, Michigan 48109, United States*

[*]Corresponding Author. E-mail: brookscl@umich.edu. Phone: 734-647-6682.

Thermodynamic cycles for pK_a and binding free energy determination

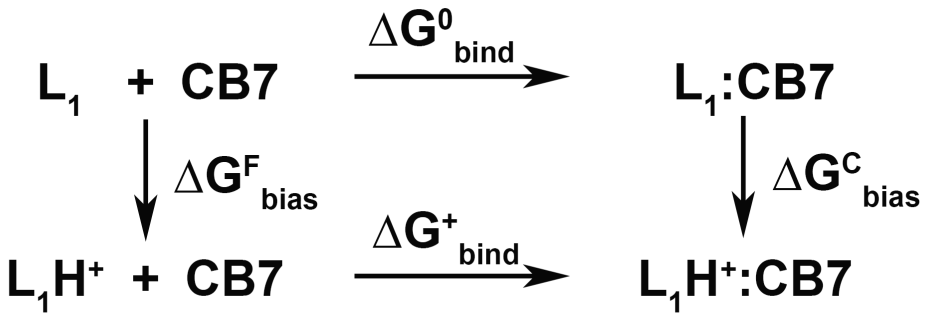


Figure S1: Thermodynamic cycle for the complexation of cucurbit[7]uril (CB7) and a titratable ligand (L_1). L_1 refers to benzimidazole (BZ) and its derivatives ABZ, CBZ, FBZ, and TBZ.

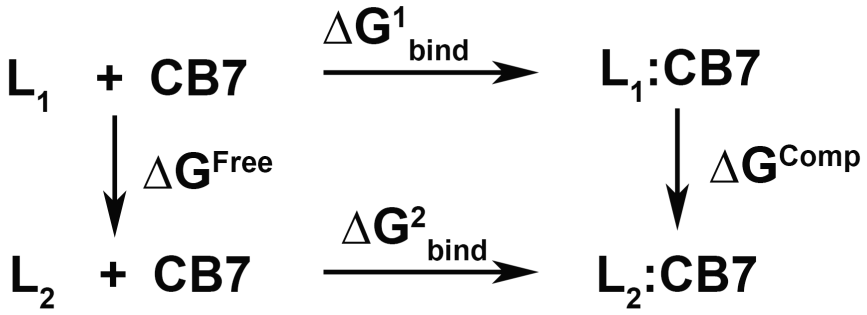


Figure S2: Thermodynamic cycle for the perturbation of one ligand into another bound to a host (CB7), where the vertical arms represent free and bound simulations.

Flattened free energy barriers

Flattened free energy barriers were used as a measure of convergence for our pK_a and reference binding free energy simulations. Where sampling is iteratively optimized by flattening free energy barriers. By reducing the free energies barriers, trapping of individual states is avoided and sampling is more evenly distributed across all unprotonated and protonated states.

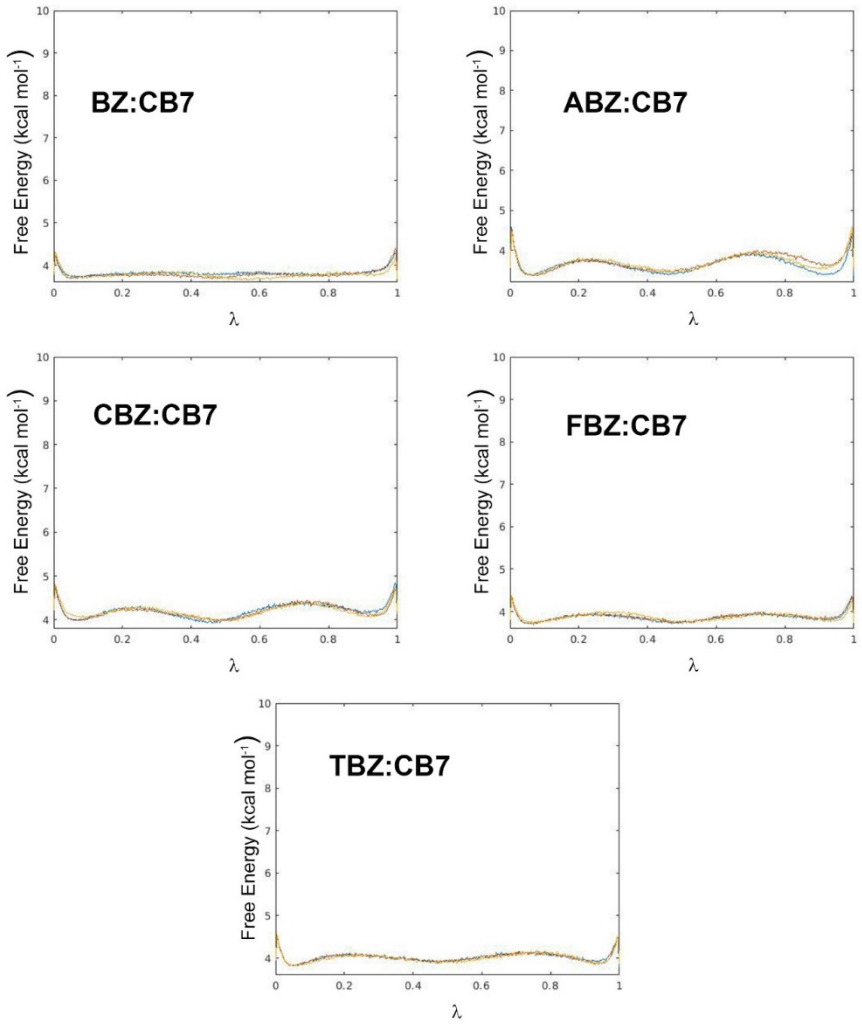


Figure S3. Flattened free energy barriers for *BZ:CB7utlizing a three-state model

pK_a determination utilizing RESP partial charges

Simulations were also performed using RESP[1-3] partial charges and CGenFF[4] parameters (Figure S4 and Table S5) which yielded poor results with negative Pearson and Spearman coefficients indicating that both the correlation to experiment and predicted relative ranking were significantly off. Small positive shifts were observed for ligands ABZ, CBZ, FBZ, and TBZ, however, a negative shift in pK_a (-1.4 pK unit) was predicted for BZ. This prediction is consistent with a previous theoretical study performed by B. Brooks and coworkers that showed predicted pK_a values were very sensitive to the partial charge scheme used.[5] They observed a similar drop in pK_a (-1.36 pK units) for BZ upon complexation with CB7 utilizing RESP[1-3] partial charges and an enveloping distribution sampling method paired with Hamiltonian replica exchange.[5]

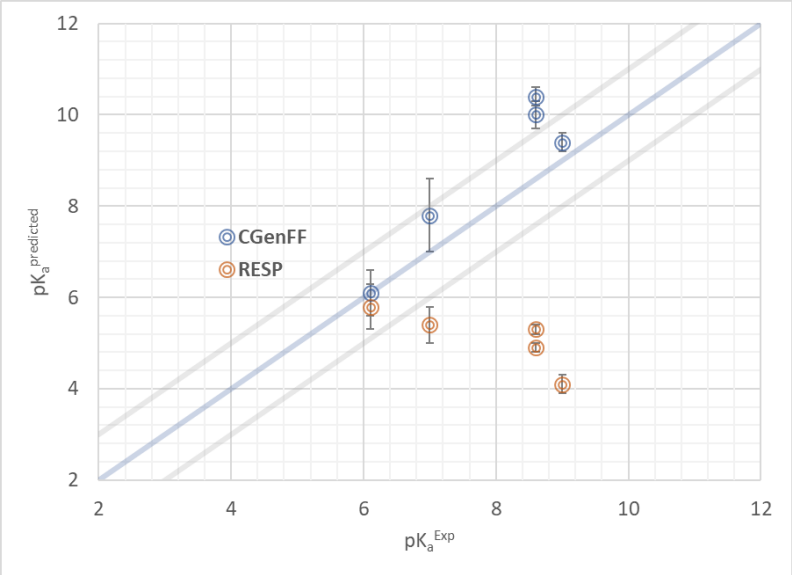


Figure S4. A comparison between the experimental and computed pK_a. The grey lines indicate ±1 pK unit.

This is in contrast to a large shift in pK_a (≈+7.5 pK units) when they changed the partial charge scheme to CGenFF[4]. We in fact saw a very similar trend, albeit a smaller shift in pK_a for BZ (+3.9 pK units), using a completely different method. Sampling and overstabilization of the protonated form of BZ was cited as a possible cause of the significantly elevated pK_a shift by B. Brooks and coworkers.[5] In contrast, our approach does not suffer from sampling or overstabilization issues by design due to the employment of λ-dynamics[6] with enhanced

sampling (ALF)[7] simulations, where sampling is iteratively optimized by flattening free energy barriers which reduces trapping of individual states. This allows determination of free energies where trapping is avoided, and sampling is evenly distributed across all unprotonated and protonated states.

Ligand	pK_a^{Exp}	pK_a^{RESP}
BZ	9.0	4.1±0.2
ABZ	6.1	5.8±0.5
CBZ	7.0	5.4±0.4
FBZ	8.6	5.3±0.1
TBZ	8.6	4.9±0.1
	MUE	**2.5**
	Pearson (R)	**-0.8**
	Spearman	**-0.9**

Table S1. A comparison between the experimental pK_a and those computed using the RESP charge scheme

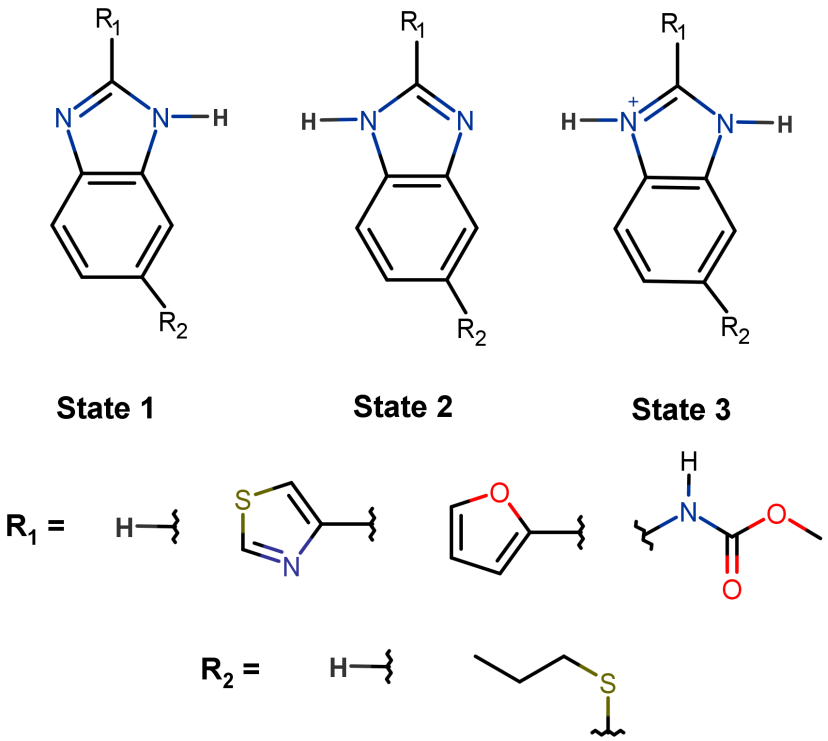
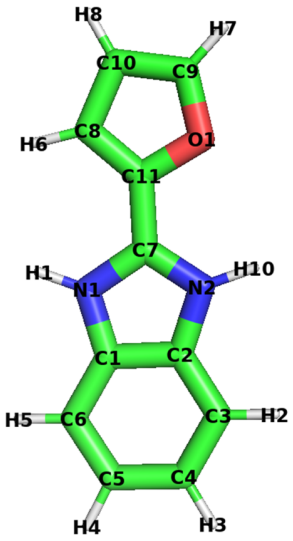
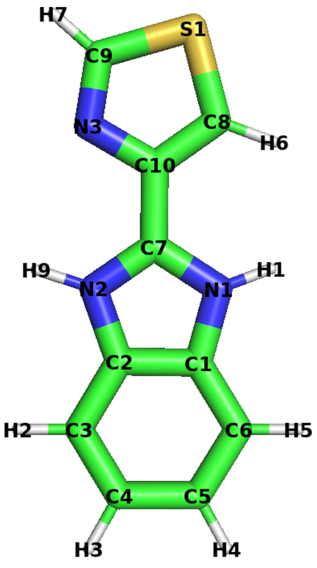


Figure S5. The three-state model used for the pK_a calculation with pendant groups labeled. States 1 and 2 are in a neutral charge state while state 3, the protonated form, is in a +1 charge state.



Atom	Charge
C1	0.297
C2	0.193
C3	-0.241
H2	0.196
C4	-0.218
H3	0.196
C5	-0.218
H4	0.196
C6	-0.34
H5	0.196
C7	0.57
N1	-0.632
H1	0.487
N2	-0.632
H10	0.487
C8	-0.217
H6	0.179
C9	0.151
H7	0.196
C10	-0.16
H8	0.258
C11	0.33
O1	-0.274

Figure S6. Modified CGenFF charges used for FBZ+



Atom	Charge
C1	0.297
C2	0.193
C3	-0.241
H2	0.196
C4	-0.218
H3	0.196
C5	-0.218
H4	0.196
C6	-0.341
H5	0.196
C7	0.57
N1	-0.631
H1	0.487
N2	-0.631
H9	0.487
C8	-0.152
H6	0.236
C9	0.303
H7	0.206
C10	0.352
N3	-0.561
S1	0.078

Figure S7. Modified CGenFF charges used for TBZ+

References

1. Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. A Well-behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. *J. Phys. Chem.* **1993,** *97*, 10269-10280.
2. Cieplak, P.; Cornell, W. D.; Bayly, C.; Kollman, P. A. Application of The Multimolecule and Multiconformational RESP Methodology to Biopolymers: Charge Derivation for DNA, RNA, and Proteins. *J. Comput. Chem.* **1995,** *16*, 1357-1377.
3. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Kollman, P. A. Application of RESP Charges to Calculate Conformational Energies, Hydrogen Bond Energies, and Free Energies of Solvation. *J. Am. Chem. Soc.* **1993,** *115*, 9620-9631.
4. Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P., et al. CHARMM General Force Field: A Force Field for Drug-like Molecules Compatible with The CHARMM All-atom Additive Biological Force Fields. *J. Comput. Chem.* **2010,** *31*, 671-690.
5. Lee, J.; Miller, B. T.; Brooks, B. R. Computational Scheme for pH-dependent Binding Free Energy Calculation with Explicit Solvent. *Protein Sci.* **2016,** *25* (1), 231-243.
6. Knight, J. L.; Brooks, C. L., III. λ-dynamics Free Energy Simulation Methods. *J. Comput. Chem.* **2009,** *30*, 1692-1700.
7. Hayes, R. L.; Armacost, K. A.; Vilseck, J. Z.; Brooks, C. L., III. Adaptive Landscape Flattening Accelerates Sampling of Alchemical Space in Multisite λ Dynamics. *J. Phys. Chem. B.* **2017,** *121*, 3626-3635.