

# Supporting Information: Quantitative Ranking of Ligand Binding Kinetics with a Multiscale Milestoning Simulation Approach

Benjamin R. Jagger, Christopher T. Lee, and Rommie E. Amaro\*

*Department of Chemistry and Biochemistry, University of California, San Diego, 9500  
Gilman Drive, La Jolla, California 92093-0340, United States*

E-mail: ramaro@ucsd.edu

# Methods

## System Preparation

GAFF<sup>1,2</sup> forcefield parameters for the seven guest molecule along with both GAFF and Q4MD-CD<sup>3</sup> parameterizations of  $\beta$ -cyclodextrin were obtained from Tang and Chang.<sup>4</sup> For comparison we use identical structure and parameterizations as those used in their study. These initial structures were used by the SEEKR software for preparation of the milestoning simulations. The preparation procedure was the same for each of the seven guest molecules and followed standard SEEKR protocols.<sup>5</sup> All systems were solvated with TIP3P waters.<sup>6</sup>

### Preparation of milestoning simulations with SEEKR

The bound state of the host-guest complex was defined as the center of mass (COM) of the  $\beta$ -cyclodextrin. The guest molecule was considered to be bound when its COM was within 1.5 Å of the bound state coordinates. From this bound state, spherical milestones were defined in increasing 1.5 Å increments from 1.5 Å to 13.5 Å. Furthermore, spherical milestones of radius 7.5 Å and less, were divided into two half-spheres to better capture the asymmetries between the two faces. When the ligand is less than 7.5 Å away from the bound state, it is trapped on a particular face due to the size of the host molecule. For milestone distances greater than 7.5 Å, the ligand was found to freely sample both faces, and therefore a single spherical milestone was sufficient for sampling host-guest interactions. In total, 14 unique milestones were defined. In practice, this was achieved via post-processing the simulations on each face to identify any trajectories that crossed from one face to the other, and modifying the transitions accordingly in the milestoning model. In addition, two simulations were conducted for the outermost milestones (one with the ligand initiated on each face), and these were then combined into a single milestone (with double the sampling) for milestones that were not restricted to a particular face.

The first 13 milestones correspond to the MD region, while the 14th and outermost mile-

stone corresponds to the BD region. The standard SEEKR preparation protocol<sup>5</sup> was then used to generate the coordinate, parameter, and simulation files necessary for a milestoneing calculation. For each of the MD milestones, a copy of the apo  $\beta$ -cyclodextrin structure was generated and the guest molecule was then placed at the appropriate radius from the bound state coordinates. Any water molecules that clashed with the guest molecule were removed. The guest distribution for the BD milestone was constructed by first running a conventional BD simulation where trajectories terminated at the appropriate distance for the milestone surface (13.5 Å).

## Simulation

### MD Simulations

A modified version of NAMD 2.12 was used for all MD simulations.<sup>7</sup> For all 13 milestones in the MD region, the standard SEEKR procedure for minimization, equilibration and simulation was followed. First, 5000 steps of minimization were performed to allow for relaxation, particularly of solvent, around the newly placed guest molecule. Further relaxation of the solvent was achieved by a series of 2 ps heating simulations that gradually increased the temperature from 298 K to 350 K and then cooled back to 298 K. Host and guest atoms were constrained during these heating simulations to ensure that the guest remained on the appropriate milestone surface. To obtain the equilibrium distribution of the guest molecule on each milestone, 200 ns of constant volume simulation was performed. A  $90 \text{ kcal} \cdot \text{mol}^{-1} \cdot \text{\AA}^{-2}$  harmonic restraint was used to hold the COM of the guest molecule at the appropriate distance from the binding site. To minimize any bias of the arbitrary guest starting conformation, the first 40 ns of each simulation were discarded and therefore were not included as part of the equilibrium distribution. From these trajectories, position and velocity configurations were selected every 0.2 ns, resulting in a total of 800 configurations per milestone. To obtain the first hitting point distribution (FHPD) of each milestone, 10 independent and unrestrained simulations were initiated from these equilibrium configurations. Each simulation was prop-

agated backwards in time by reversing its velocity at constant energy and volume (a total of 8000 reversals for each milestone). Only trajectories that struck an adjacent milestone before recrossing the milestone on which they originated were included as part of the FHPD for that milestone. To obtain the transition probabilities and times necessary for the calculation of kinetic parameters, all members of the FHPD were brought back to their starting position and velocity and new unrestrained simulations were then initiated from each configuration. These simulations were propagated forward in time at constant energy and volume. Once a simulation crossed its starting milestone again, it was monitored for crossing of adjacent milestones. When an adjacent milestone was crossed, the simulation was stopped and the transition and incubation time were recorded. Although more of the equilibrium simulation trajectories could likely have been used without biasing the results based on the starting conformation, the 8000 reversal trajectories that resulted from the 160 ns used were more than sufficient to sample the transitions between the milestones, resulting in hundreds of observed transitions between each milestone. In total,  $2.6 \mu s$  of equilibrium sampling (160 ns for 16 milestones) were used and approximately 570 ns of FHPD sampling for a total of  $3.2 \mu s$  of simulation used in the milestones model. The total cost per ligand (including simulation discarded for equilibration) was therefore  $\sim 3.8 \mu s$ .

## BD Simulations

All BD simulations were performed using the BrownDye software package.<sup>8</sup> Electrostatic potentials of the host and guest molecules used as inputs for the BD simulation were calculated with APBS version 1.4.<sup>9</sup> To match experimental conditions, APBS calculations and BD simulations were carried out with a solvent dielectric of 78, a solute dielectric of 2, and zero ionic concentration. An initial series of simulations were initiated from the b-surface, a sphere that encloses the entire host molecule and has sufficient radius for the guest molecule to be situated in bulk solvent such that forces between the host and guest are centrosymmetric.  $10^6$  independent BD simulations were initiated from random points on the b-surface

and and were propagated until the guest either contacted the outermost milestone (13.5 Å) or escaped. Trajectories that successfully contacted the 13.5 Å milestone were used as the FHPD for this milestone. Another series of  $10^6$  BD simulations were initiated from the FHPD and propagated until contacting the second-outermost milestone (12.0 Å) or escaping to the q-surface. This procedure is automated by SEEKR.

## Milestoning Calculations

Statistics from all milestones in the MD and BD regions were extracted using SEEKR and combined to construct a transition kernel as well as an incubation time vector. These are the two key quantities for the calculation of kinetic parameters in milestoning theory.<sup>10</sup> As described previously, a post-simulation analysis was performed to account for ligand transitions between the two faces of the  $\beta$ -cyclodextrin ring for milestones of radius 7.5 Å or less. The analysis portion of SEEKR was then used to compute the desired kinetic quantities,  $k_{\text{on}}$  and  $k_{\text{off}}$ , as well as the free energy of binding  $\Delta G_{\text{bind}}$ .

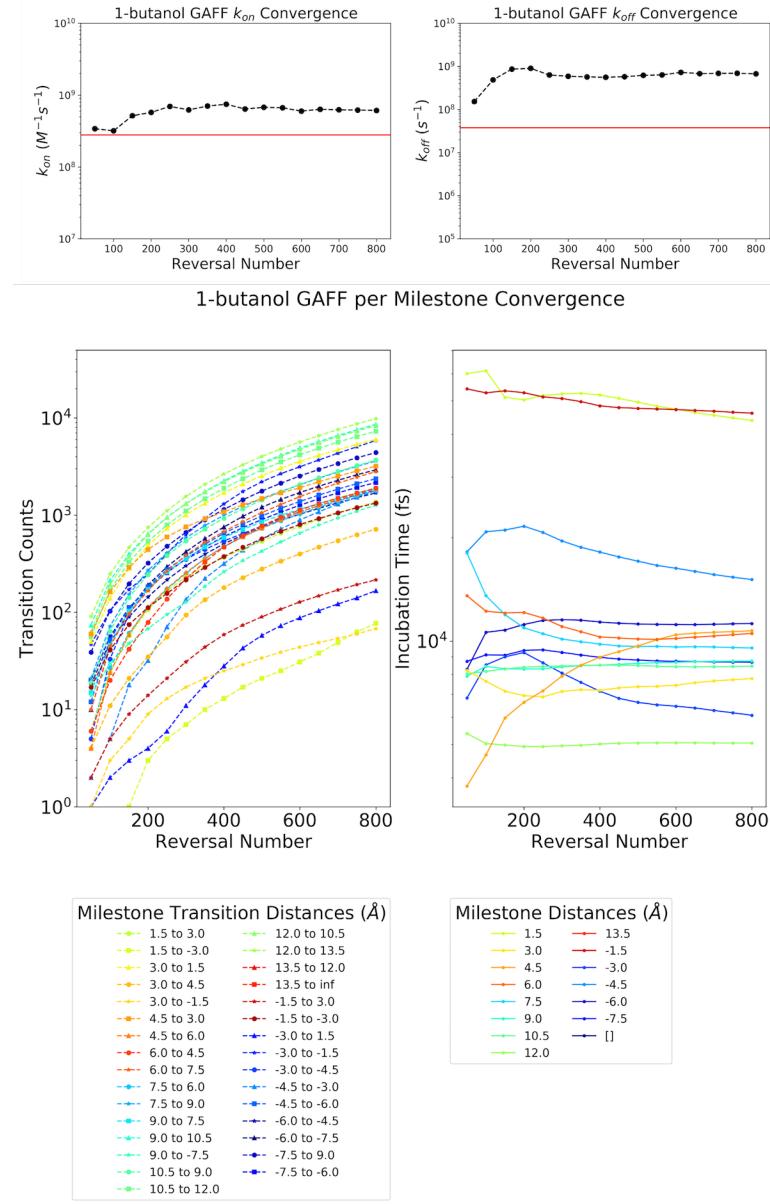


Figure 1: 1-butanol GAFF per milestone convergence plot

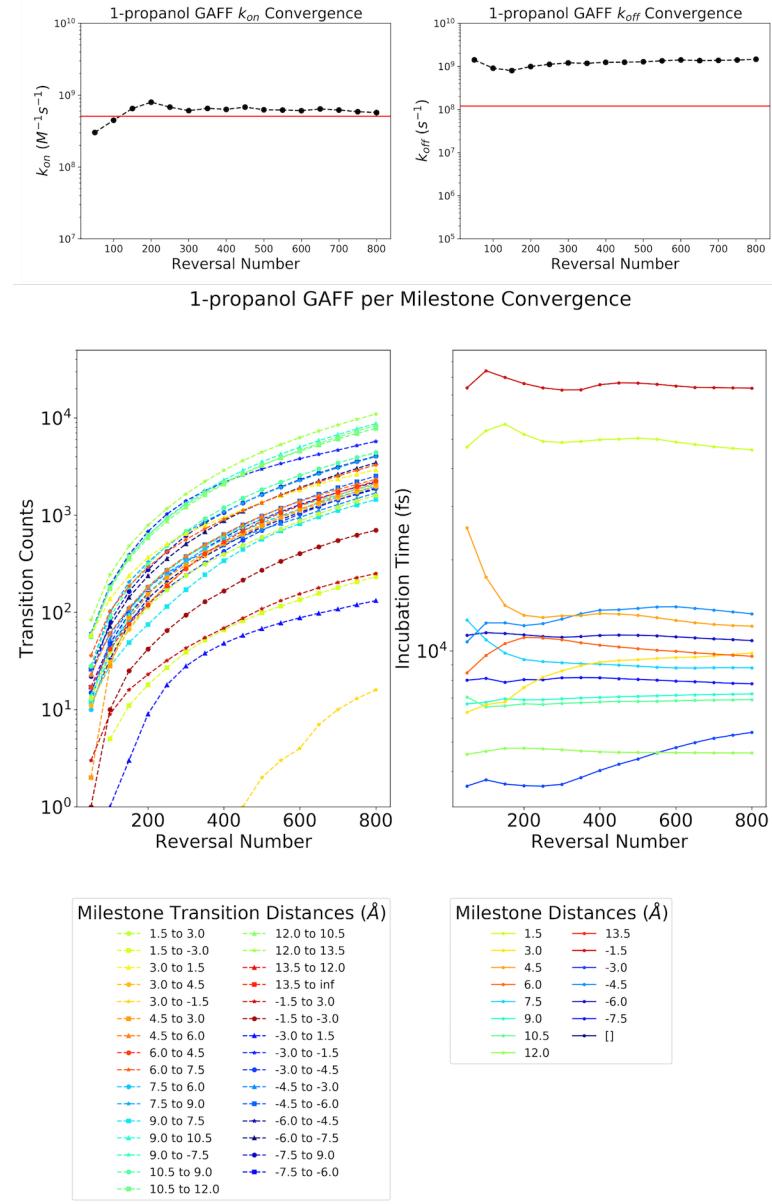


Figure 2: 1-propanol GAFF per milestone convergence plot

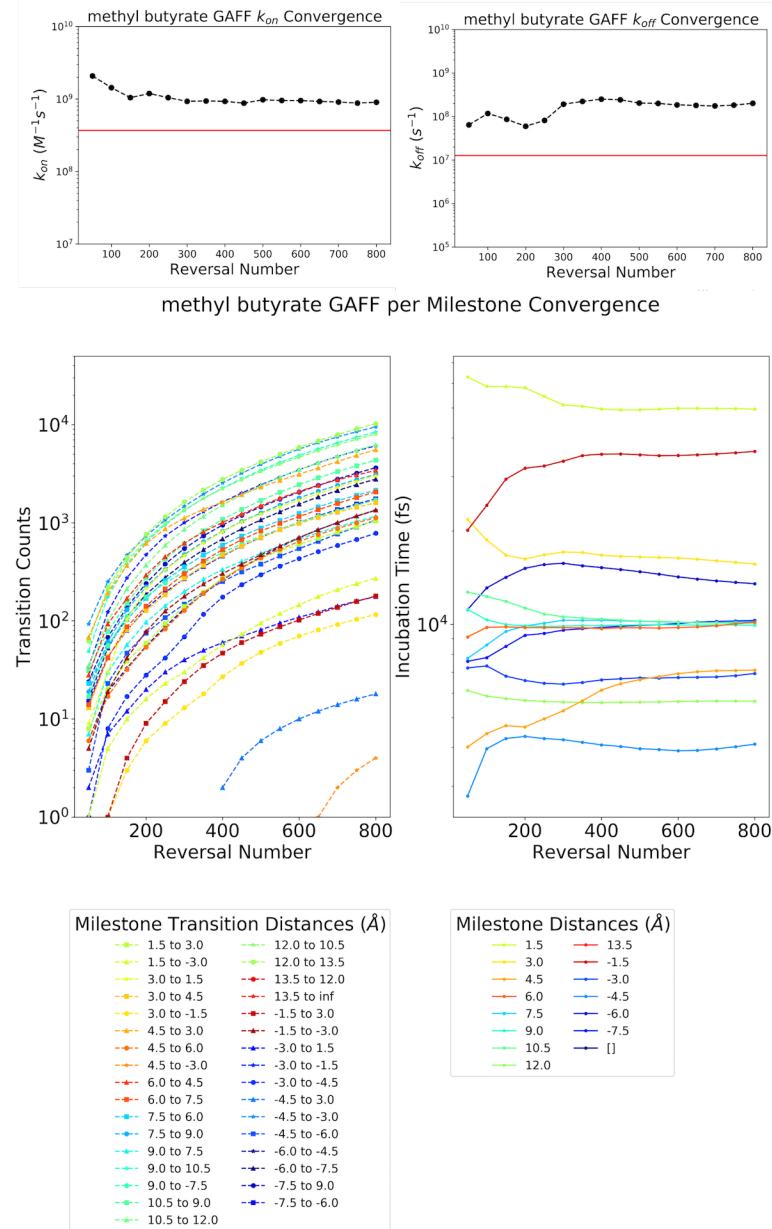


Figure 3: methyl butyrate GAFF per milestone convergence plot

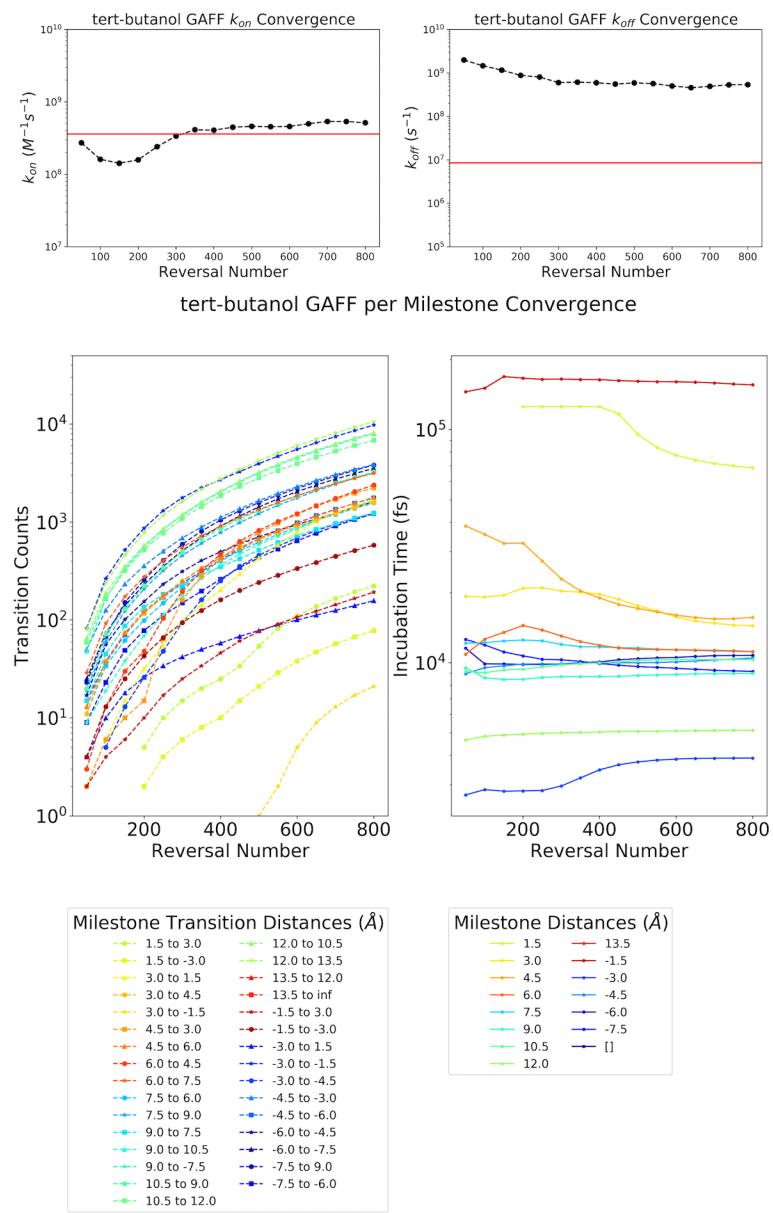


Figure 4: tert butanol GAFF per milestone convergence plot

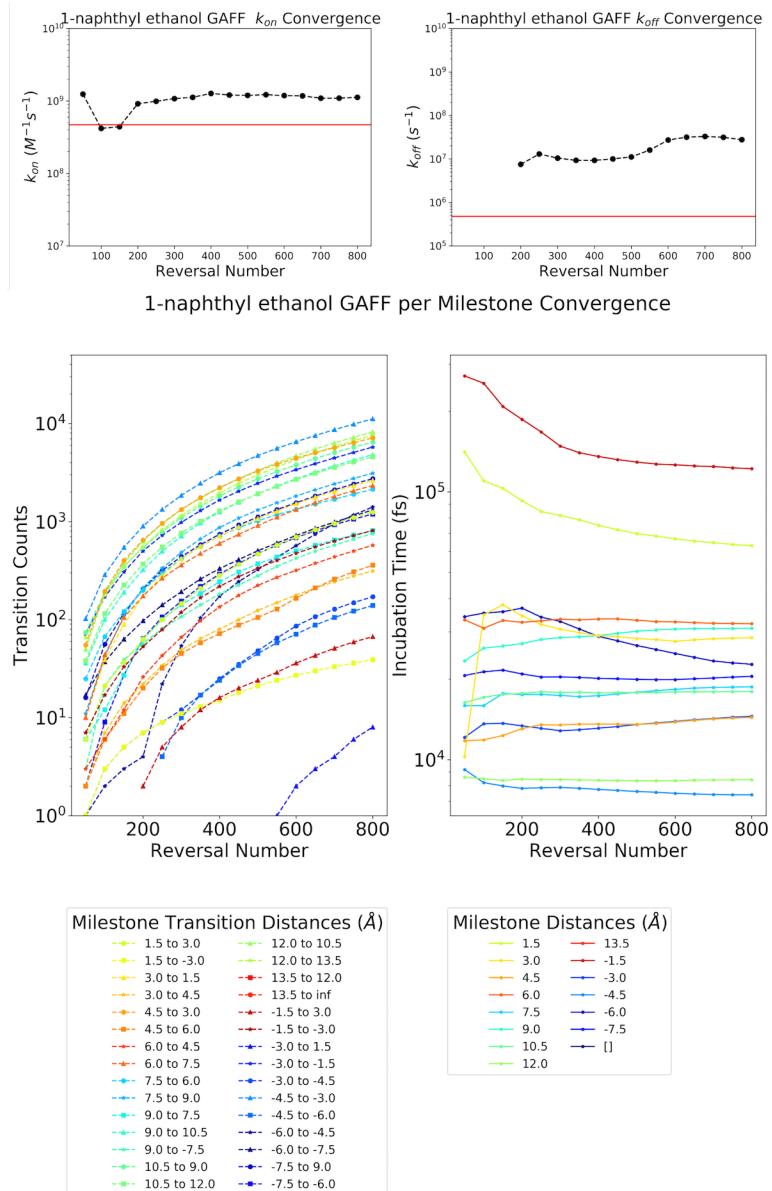


Figure 5: 1-naphthyl ethanol GAFF per milestone convergence plot

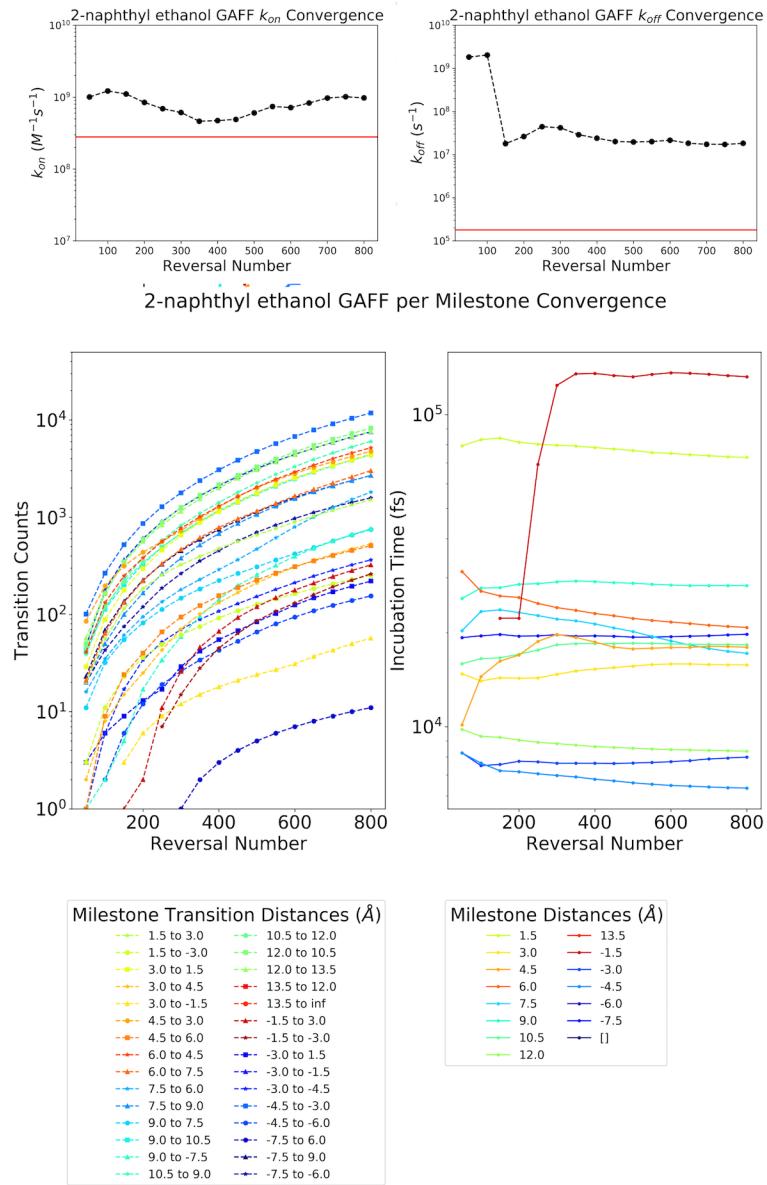


Figure 6: 2-naphthyl ethanol GAFF per milestone convergence plot

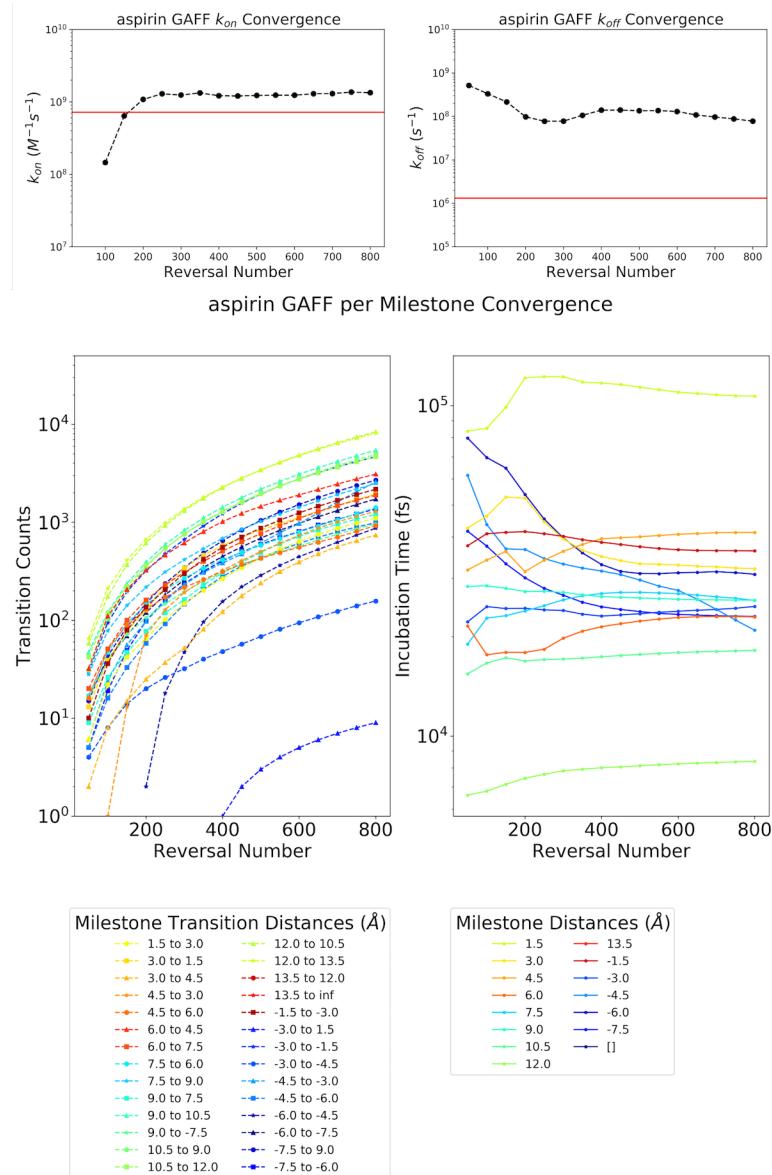


Figure 7: aspirin GAFF per milestone convergence plot

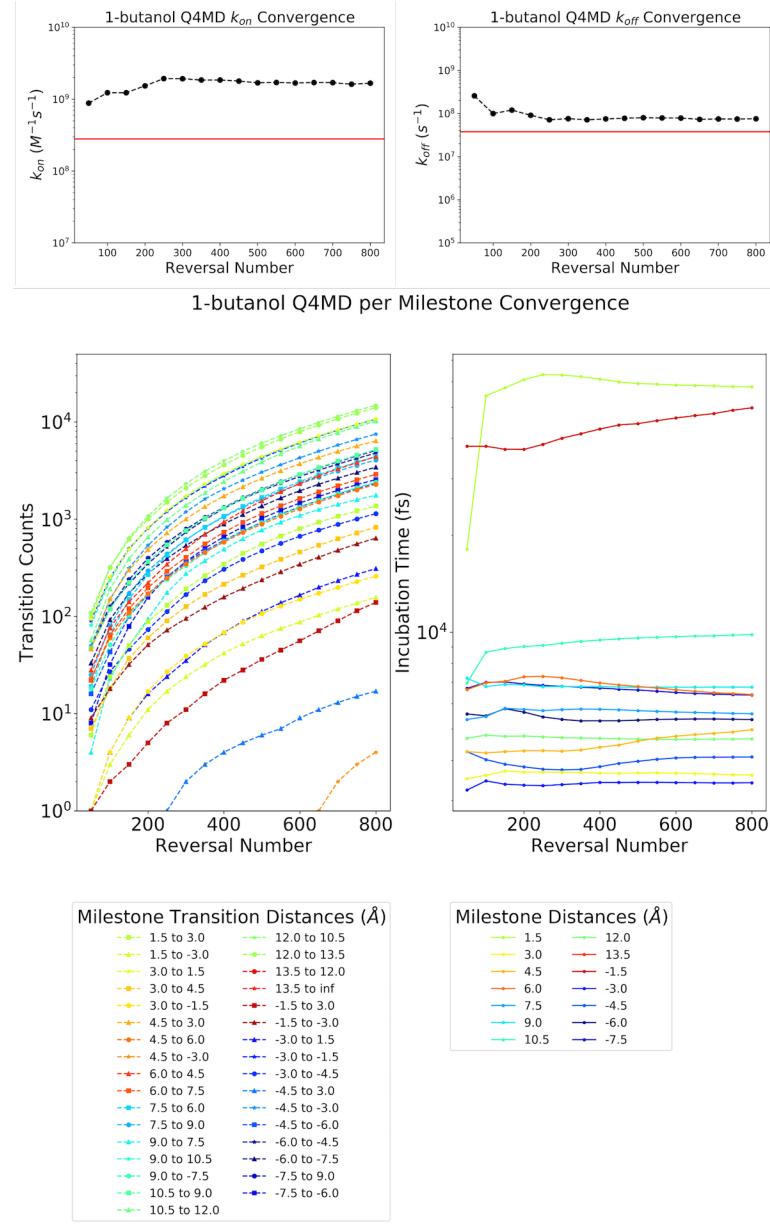


Figure 8: 1-butanol Q4MD per milestone convergence plot

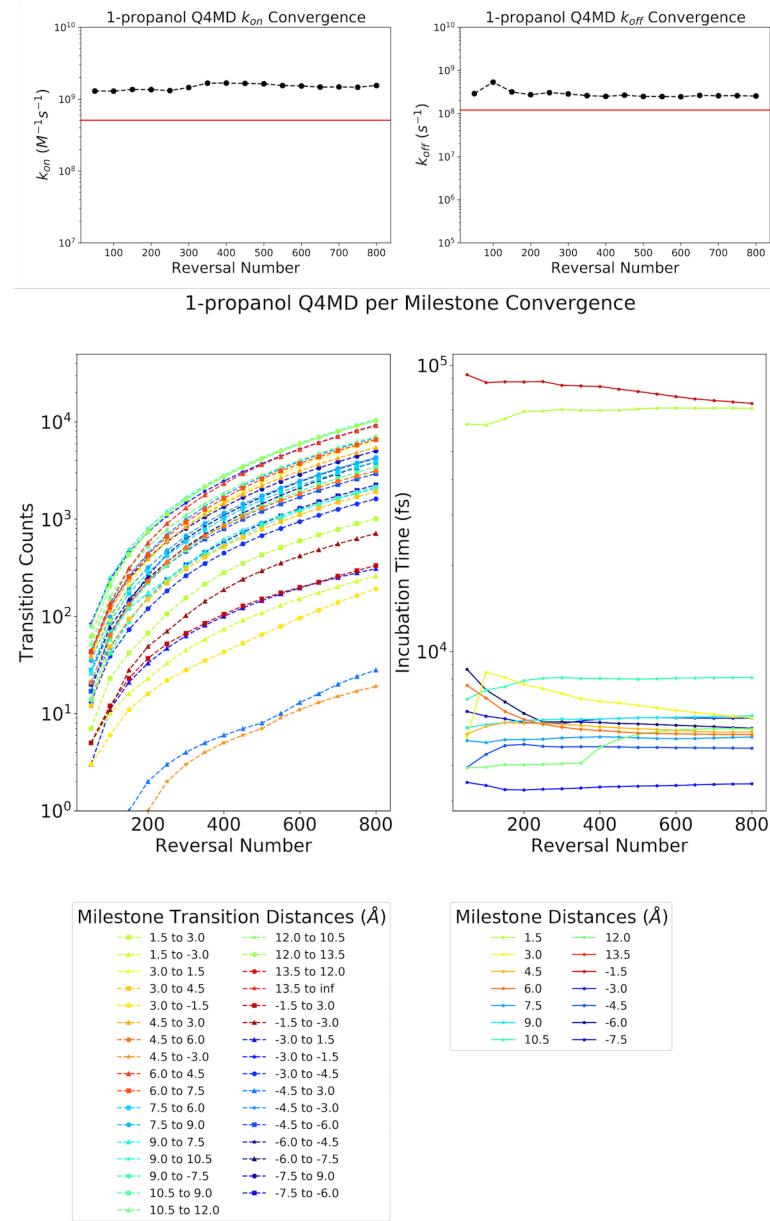


Figure 9: 1-propanol Q4MD per milestone convergence plot

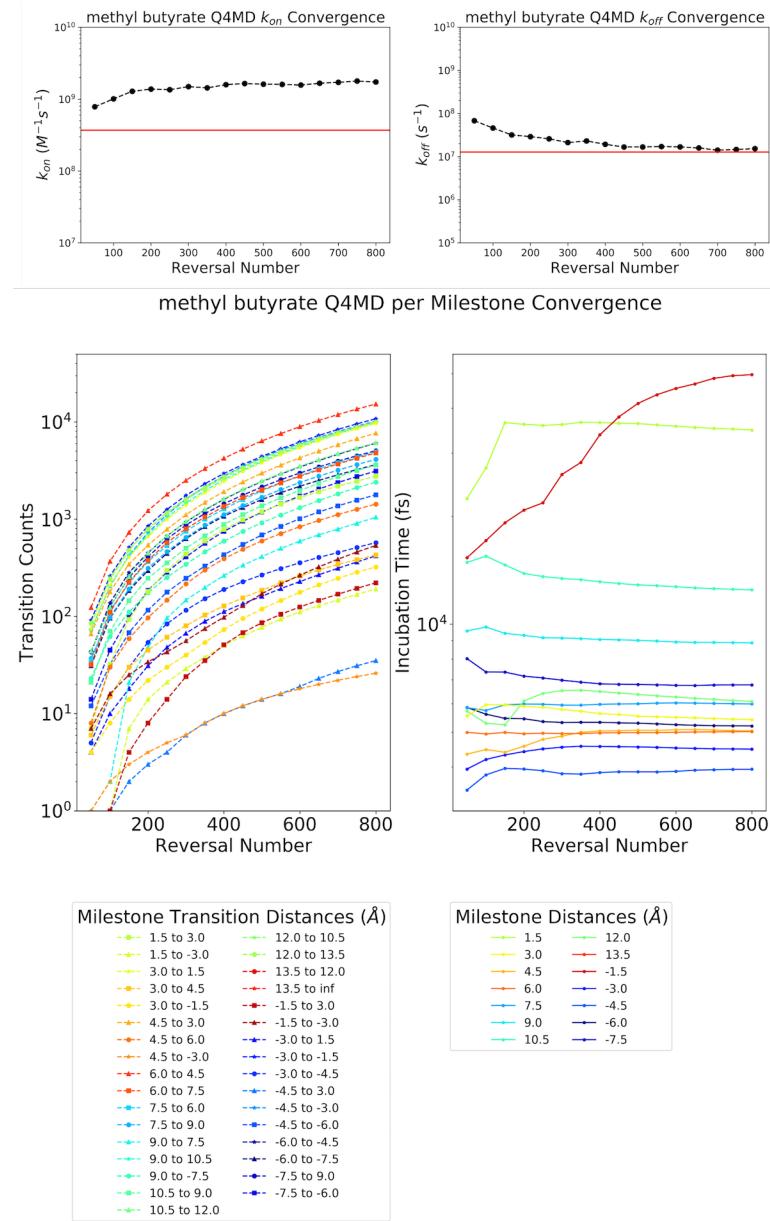


Figure 10: methyl butyrate Q4MD per milestone convergence plot

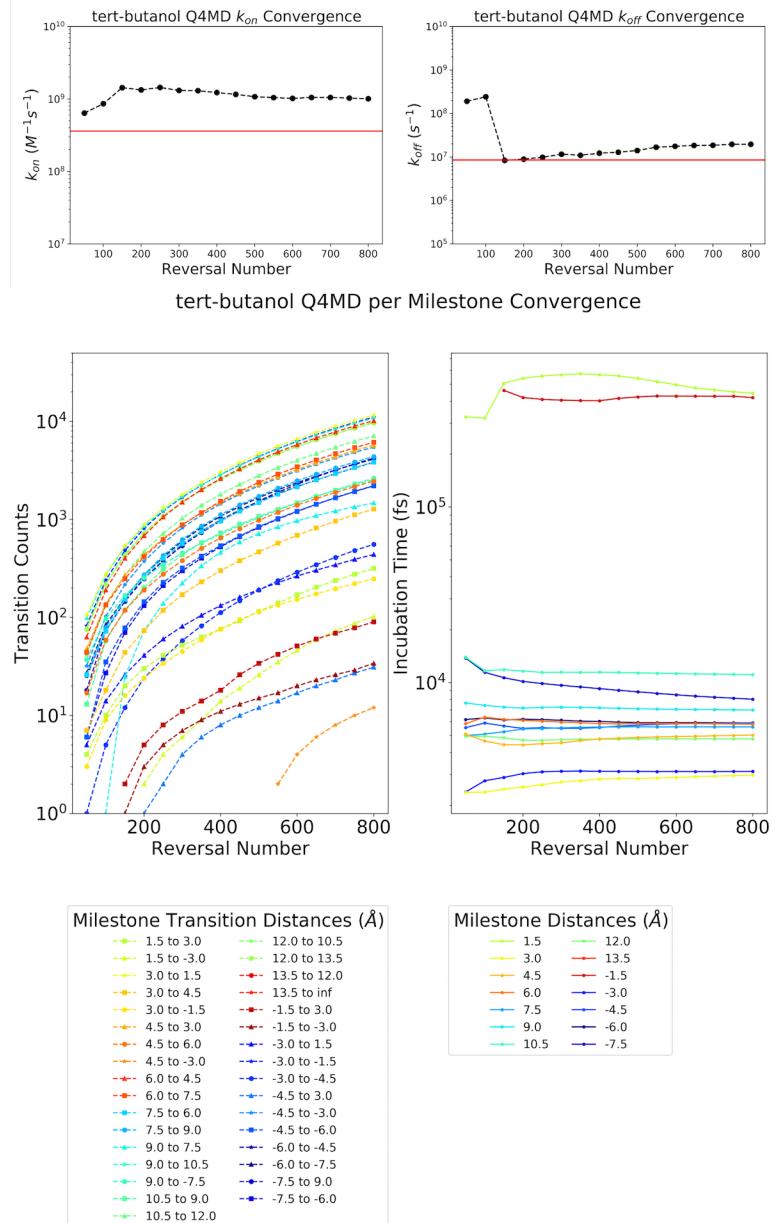


Figure 11: tert butanol Q4MD per milestone convergence plot

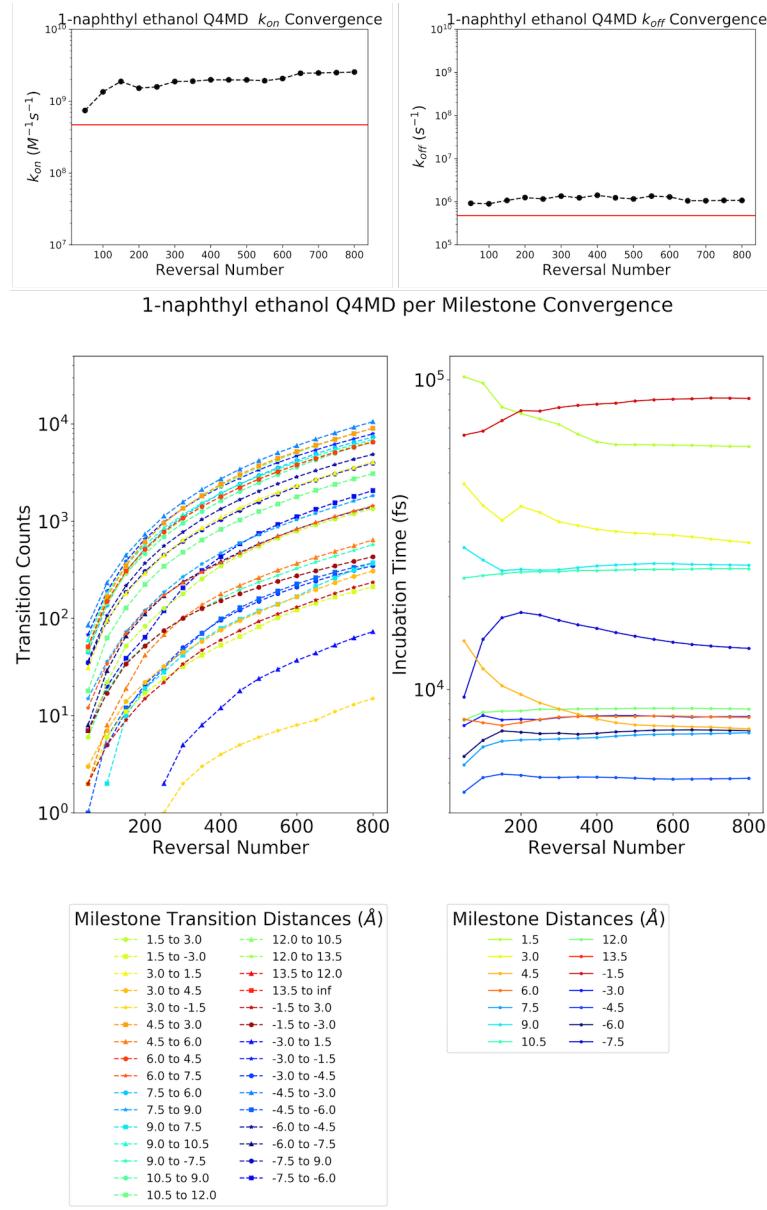


Figure 12: 1-naphthyl ethanol Q4MD per milestone convergence plot

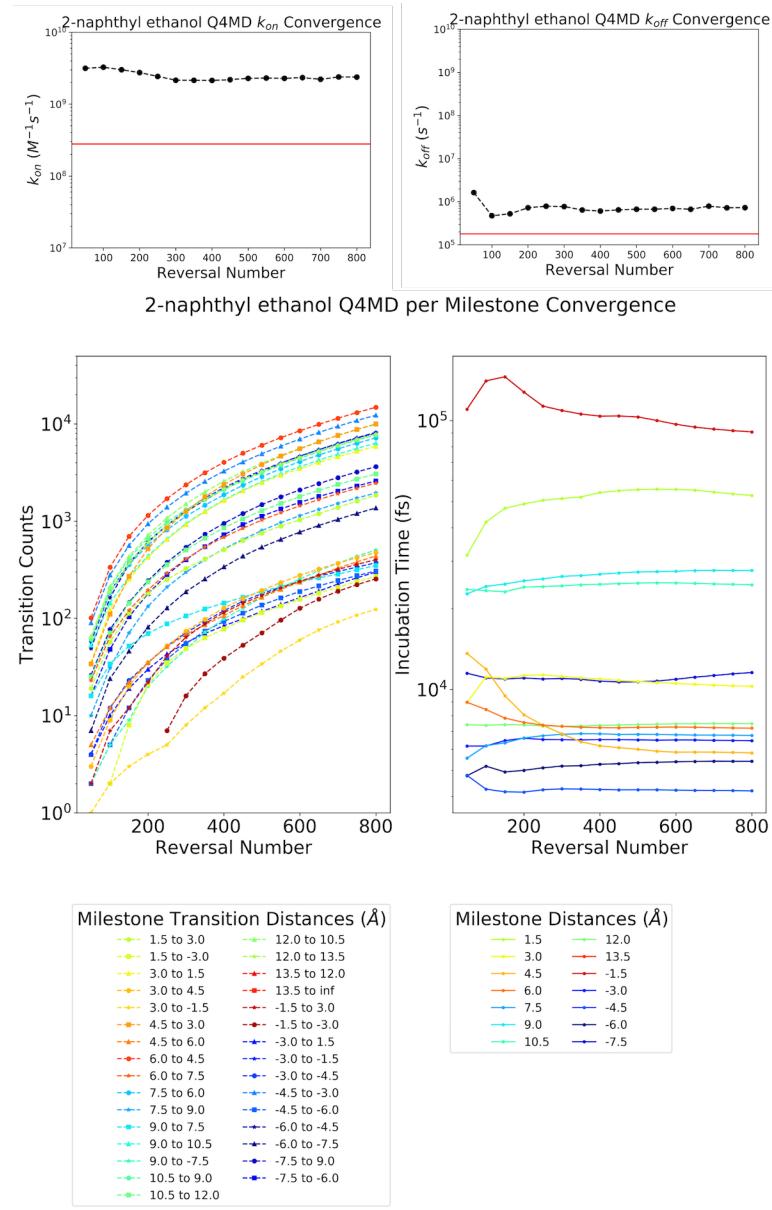


Figure 13: 2-naphthyl ethanol Q4MD per milestone convergence plot

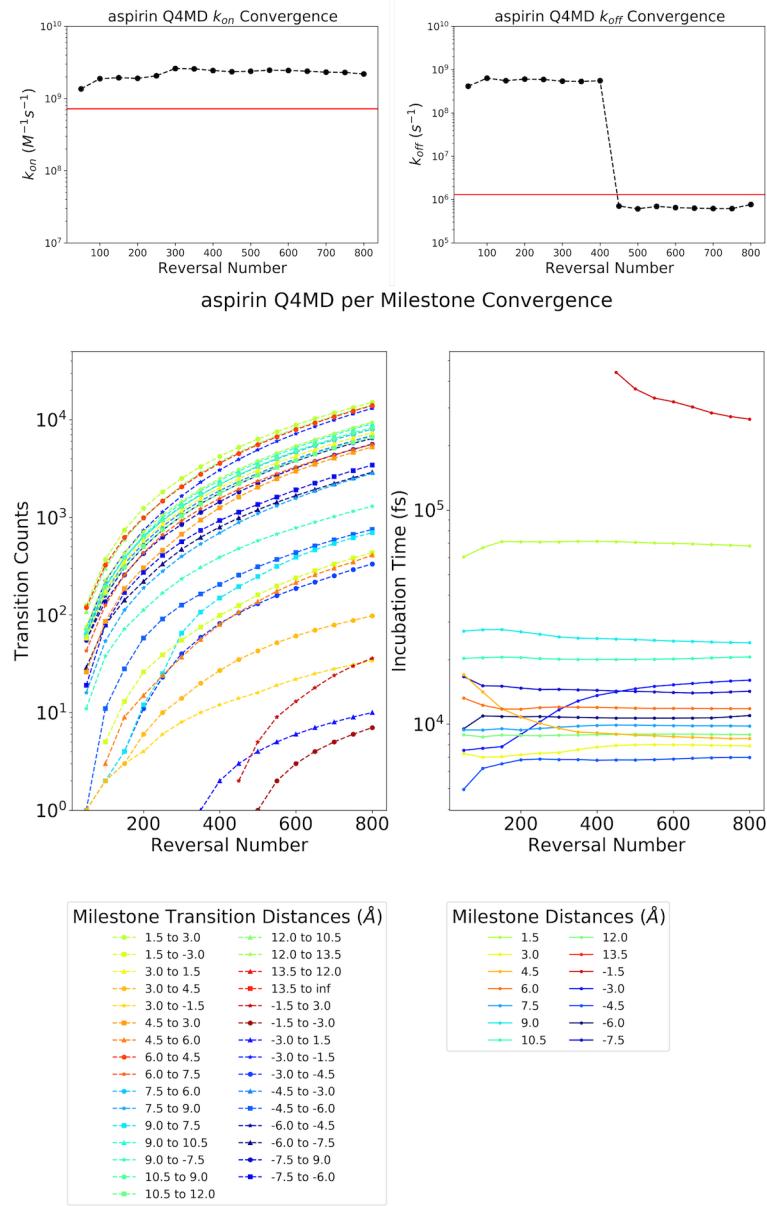


Figure 14: aspirin Q4MD per milestone convergence plot

## References

- (1) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25*, 1157–1174.
- (2) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graph. Model.* **2006**, *25*, 247–260.
- (3) Cézard, C.; Trivelli, X.; Aubry, F.; Djedaiñi-Pilard, F.; Dupradeau, F.-Y. Molecular dynamics studies of native and substituted cyclodextrins in different media: 1. Charge derivation and force field performances. *Phys. Chem. Chem. Phys.* **2011**, *13*, 15103.
- (4) Tang, Z.; Chang, C.-e. A. Binding Thermodynamics and Kinetics Calculations Using Chemical Host and Guest: A Comprehensive Picture of Molecular Recognition. *J. Chem. Theory Comput.* **2018**, *14*, 303–318.
- (5) Votapka, L. W.; Jagger, B. R.; Heyneman, A. L.; Amaro, R. E. SEEKR: Simulation Enabled Estimation of Kinetic Rates, A Computational Tool to Estimate Molecular Kinetics and Its Application to Trypsin–Benzamidine Binding. *J. Phys. Chem. B* **2017**, *121*, 3597–3606.
- (6) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935.
- (7) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781–1802.
- (8) Huber, G. A.; McCammon, J. A. Browndye: A Software Package for Brownian Dynamics. *Comput. Phys. Commun.* **2010**, *181*, 1896–1905.

- (9) Baker, N. A.; Sept, D.; Joseph, S.; Holst, M. J.; McCammon, J. A. Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proc. Natl. Acad. Sci.* **2001**, *98*, 10037–10041.
- (10) Faradjian, A. K.; Elber, R. Computing time scales from reaction coordinates by milestoning. *J. Chem. Phys.* **2004**, *120*, 10880–10889.