Highly Efficient Computation of the Basal k_{on} using Direct Simulation of Protein-Protein Association with Flexible Molecular Models

*Ali S. Saglam and Lillian T. Chong**

Department of Chemistry, University of Pittsburgh, Pittsburgh, Pennsylvania, 15260

AUTHOR INFORMATION

Corresponding Author

*To whom correspondence should be addressed. E-mail: ltchong@pitt.edu.

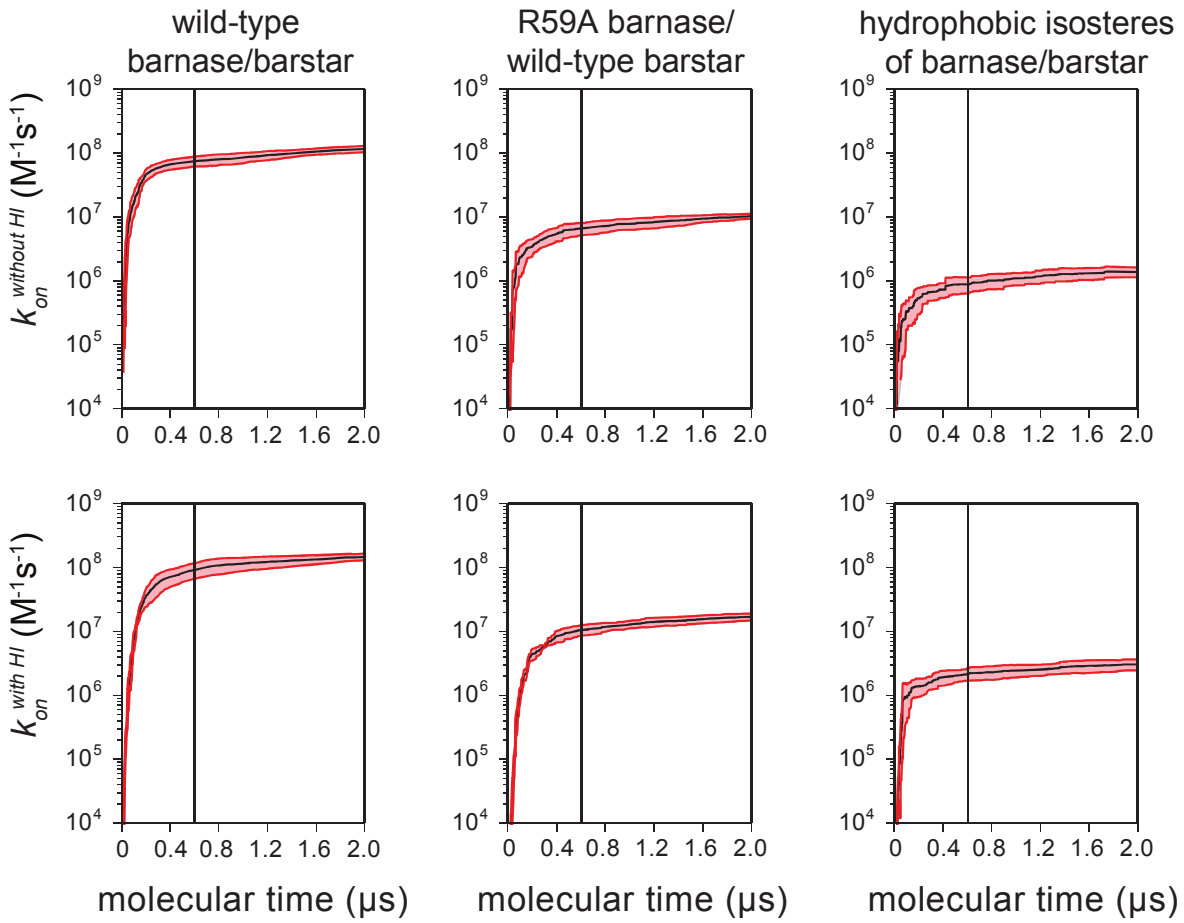

Figure S1. Average calculated k_{on} values for each barnase-barstar pair as a function of the molecular time from five independent WE simulations without and with the inclusion of intermolecular HI (top and bottom panels, respectively). The molecular time is defined as $N\tau$ where N is the number of WE iterations and τ is the fixed time interval of each iteration. Uncertainties (shaded in pink) are 95% confidence intervals. All subsequent analysis of the simulations was performed starting from a molecular time where an approximate steady state has been reached (vertical lines).

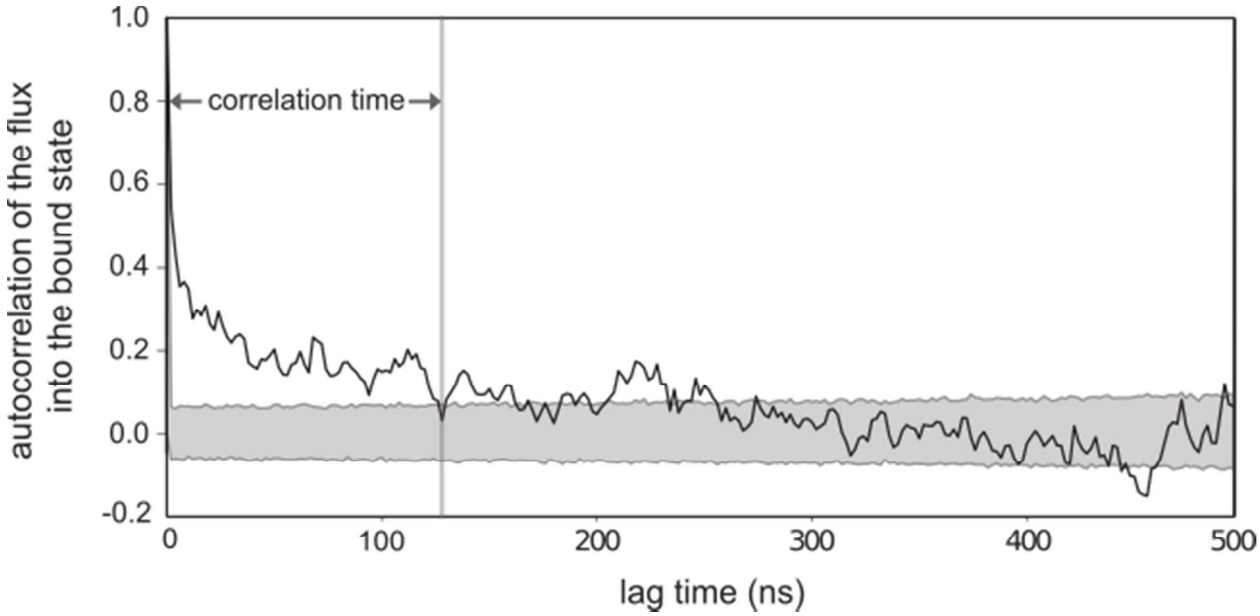

Figure S2. A representative autocorrelation of the flux into the bound state from a single WE simulation as a function of the lag time and its use in determining the number of independent association events. Data shown corresponds to a simulation involving the hydrophobic isosteres with the inclusion of intermolecular HI. The lag time is the frequency with which the flux into the bound state (Q_{rxn} = 0.27) was sampled and the correlation time (vertical gray line) is the time interval required to reach zero autocorrelation in the flux, i.e. the 95% confidence interval centered on zero (gray shaded region). Association events were considered independent if, within the period between the event and one correlation time before the event, their corresponding trajectories did not share a common simulation segment.

Table S1. Average calculated k_{on} values for each barnase-barstar pair from five independent WE simulations (with and without intermolecular HI) vs. experimental values. Uncertainties represent 95% confidence intervals.

	$k_{on}/10^8$($M^{-1}s^{-1}$)		
	simulation		experiment
	without HI	with HI	
wild-type barnase/barstar	2.94 ± 0.96	3.04 ± 0.66	2.86
R59A barnase/ wild-type barstar	0.243 ± 0.067	0.39 ± 0.124	0.31
hydrophobic isosteres of barnase/barstar	0.0285 ± 0.0130	0.0579 ± 0.0017	0.14

Table S2. Average efficiencies of weighted ensemble (WE) vs brute force (BF) simulation in estimating the k_{on} using five independent WE simulations. The efficiency of each WE simulation was estimated using t_{BF}/t_{WE} where t_{WE} is the wall-clock time required of the WE simulation and t_{BF} is the wall-clock time required of BF simulation to generate the same number of independent association events using the same computing resource (256 CPU cores of 2.3GHz AMD Interlagos processors). The latter was estimated using the probability β of capturing a successful association event over the course of the WE simulation (see Methods section). Averages for the free energy barrier to association, number of independent association events, t_{WE}, t_{BF}, and β are also provided. Uncertainties represent 95% confidence intervals.

	wild-type barnase/barstar		R59A barnase/ wild-type barstar		hydrophobic isosteres of barnase/barstar	
	with HI	without HI	with HI	without HI	with HI	without HI
number events	1926 ± 70	2182 ± 168	1329 ± 62	958 ± 23	1096 ± 102	374 ± 18
t_{WE} (days)	2.1 ± 0.3	3.0 ± 0.6	2.0 ± 0.5	2.8 ± 1.0	3.1 ± 0.8	3.4 ± 0.5
t_{BF} (days)	13.6 ± 3.2	15.9 ± 3.0	77.0 ± 23.8	87.8 ± 22.0	386.0 ± 80.6	311.4 ± 103.6
β / 10^{-2}	1.214 ± 0.261	1.182 ± 0.364	0.156 ± 0.049	0.095 ± 0.003	0.025 ± 0.007	0.011 ± 0.005
efficiency of WE	6 ± 1	5 ± 1	46 ± 27	35 ± 13	131 ± 26	92 ± 34