THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

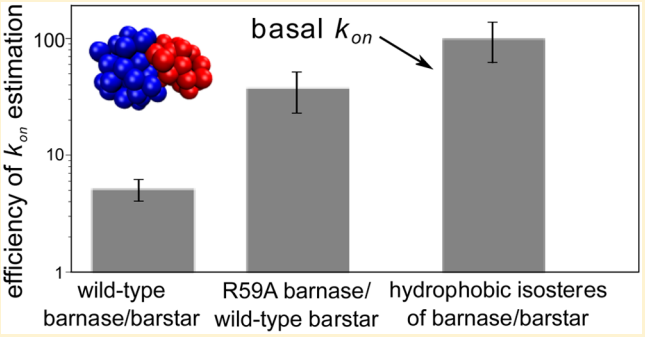
Highly Efficient Computation of the Basal k_{on} using Direct Simulation of Protein−Protein Association with Flexible Molecular Models

Ali S. Saglam and Lillian T. Chong*

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

Supporting Information

ABSTRACT: An essential baseline for determining the extent to which electrostatic interactions enhance the kinetics of protein−protein association is the "basal" k_{on}, which is the rate constant for association in the absence of electrostatic interactions. However, since such association events are beyond the milliseconds time scale, it has not been practical to compute the basal k_{on} by directly simulating the association with flexible models. Here, we computed the basal k_{on} for barnase and barstar, two of the most rapidly associating proteins, using highly efficient, flexible molecular simulations. These simulations involved (a) pseudoatomic protein models that reproduce the molecular shapes, electrostatic, and diffusion properties of all-atom models, and (b) application of the weighted ensemble path sampling strategy, which enhanced the efficiency of generating association events by >130-fold. We also examined the extent to which the computed basal k_{on} is affected by inclusion of intermolecular hydrodynamic interactions in the simulations.



INTRODUCTION

Of fundamental interest to biology is the extent to which electrostatic interactions enhance the rate of protein−protein association. An essential baseline for determining the magnitude of these rate enhancements is the "basal" k_{on}, which is the rate constant for association in the absence of electrostatic interactions.[1] In principle, the basal k_{on} should be measured in the same solvent environment using the hydrophobic isosteres—that is, hypothetical mutants with molecular shapes that are identical to those of the wild-type proteins, but are entirely uncharged. However, due to the difficulty of engineering hydrophobic isosteres, experimental studies have instead estimated the basal k_{on} by measuring the k_{on} for the wild-type proteins at various salt concentrations and extrapolating to the limit of infinite salt concentration where electrostatic interactions would be completely screened.[1]

An alternative approach is to construct the exact hydrophobic isosteres *in silico* by setting all partial charges of the wild-type proteins to zero and directly computing the basal k_{on} by simulating the association of the hydrophobic isosteres. Ideally, such simulations would involve the use of flexible molecular models in order to capture conformational changes during the association process. However, since the weak associations of completely hydrophobic proteins are beyond the milliseconds time scale,[1−7] it has only been feasible to directly compute the basal k_{on} using rigid, models with atomically detailed simulations.[2] Theoretical estimates of the basal k_{on} have also been made using spherical models with orientational constraints[3−6] and applications of transition-rate theory to rigid, atomistic models.[7]

Here, for the first time, we directly computed the basal k_{on} for a protein−protein association process using flexible models with molecular simulations. We focused on barnase and barstar, which are among the most rapidly associating proteins by virtue of their complementary electrostatic surfaces.[2] Our simulations employed flexible, pseudoatomic protein models of barnase and barstar that were designed by Frembgen-Kesner and Elcock[8] to retain the molecular shapes, electrostatic potentials, and diffusion properties of the corresponding atomistic models at the experimental ionic strength (50 mM).[9] The same authors have demonstrated that the use of these models with standard "brute force" simulations can reproduce the experimental k_{on} values of both the wild-type and mutant protein pairs. However, they were unable to carry out such simulations to obtain a statistically robust estimate of the k_{on} for the hydrophobic isosteres (i.e., the basal k_{on}) due to the large computational cost.[8]

A critical feature of our study is the application of the weighted ensemble (WE) strategy[11] to enhance the sampling of rare events, e.g. the slow association of completely hydrophobic, uncharged proteins. Although the WE strategy has been previously applied to protein binding processes using Brownian dynamics (BD) simulations,[11,12] these studies were carried out

Received: November 3, 2015
Revised: December 16, 2015
Published: December 16, 2015

DOI: 10.1021/acs.jpcb.5b10747
J. Phys. Chem. B 2016, 120, 117−122

without the inclusion of HIs) between, and within, the diffusing proteins. In the absence of explicit solvent, it has been demonstrated that the translational and rotational diffusion coefficients of flexible protein models are drastically underestimated unless intramolecular HIs are included in the simulations.[13] In addition, the neglect of intermolecular HIs in previous BD studies of protein binding processes[2,5,14] is likely to have contributed to their consistent overestimation of association rate constants.[8] Importantly, our simulations were validated by computing the k_{on} values for both wild-type barnase and its R59A mutant, which associates more slowly than wild-type barnase with barstar,[9] and comparing the computed values to experiment.

METHODS

The protein model and energy function. The wild-type and mutant pairs of barnase and barstar were represented using flexible, pseudoatomic models developed by Frembgen-Kesner and Elcock.[8] Full details of these models are provided in ref 5. Briefly, the generation of these models began with all-atom models of the wild-type proteins, which were based on the crystal structure of the barnase−barstar complex (PDB code: 1BRS);[15] the same models were used for both the unbound and bound states. Approximately one pseudoatom was then used to represent every three amino acid residues (33 pseudoatoms for the 110 residues of barnase and 27 pseudoatoms for the 89 residues of barstar). For the wild-type proteins and R59A mutant barnase, the effective charge method[16] was used to derive effective charges for the pseudoatomic models such that the electrostatic potentials of the corresponding all-atom models were reproduced. Electrostatic potentials were obtained by numerically solving the nonlinear Poisson−Boltzmann equation under experimental conditions (pH 8, 25 °C, and ionic strength of 50 mM).[9] Pseudoatoms were then positioned and sized to replicate the electron density envelope of the all-atom model. To generate models of the exact hydrophobic isosteres of barnase and barstar, we started with the pseudoatomic models of the wild-type proteins and set all effective charges to zero.

The energy function consisted of a single intramolecular term involving flexible, harmonic bonds between the pseudoatoms and intermolecular terms for electrostatic and nonelectrostatic interactions. To maintain the molecular shapes of the proteins, three bonds per pseudoatom were formed on average. All intermolecular electrostatic interactions between pseudoatoms were calculated using the Debye−Hückel equation; intramolecular electrostatic interactions were omitted. Nonelectrostatic interactions were calculated using a *very* weak Gō-type potential energy function with a shallow well depth (ε = 0.1 kcal/mol). Thus, native contacts were only slightly rewarded by a weakly attractive Lennard-Jones-like potential and nonnative contacts were penalized by a purely repulsive potential.[17,18] The well-depth was kept at a minimal value in order to avoid implicitly double counting the attractive electrostatic interactions, which are assumed to be a primary driving force for the formation of the barnase-barstar complex.[9] Two pseudoatoms were considered to form a native contact if any non-hydrogen atoms of the residues in the all-atom model were within 5.5 Å of each other in the crystal structure of the native complex, yielding a total of 34 intermolecular native contacts.

Weighted Ensemble (WE) Simulations. All simulations were carried out using the WE path sampling strategy,[11] as implemented in the WESTPA software package (https://westpa.github.io/westpa).[10] In this strategy, a large number of simulations, or trajectory "walkers", are started in parallel from the initial state and iteratively evaluated at fixed time intervals τ for resampling in which walkers are either replicated or combined to maintain a similar number of walkers per bin along a progress coordinate toward the target state. Rigorous management of the statistical weights associated with each walker ensures that no bias is introduced into the dynamics.

In this study, the WE strategy was applied using steady-state simulations within the framework of the Northrup−Allison−McCammon (NAM) method.[19] This framework involves the definition of two concentric spherical surfaces with radii b and q that correspond to center-to-center separation distances for barnase and barstar. The inner sphere, or b surface, represents the initial, unbound state, and the outer sphere, or q surface, is an absorbing surface that is positioned at a much larger separation distance ($q \gg b$) to avoid wasting computational effort sampling the indefinite drifting apart of the proteins. Each WE simulation was started from 24 configurations of the unbound state in which barnase and barstar were randomly oriented at a center-to-center separation distance of b. A walker was continued until the pair of proteins either exceeded a separation distance q or satisfied the criterion for the target state for successful association, i.e. reaching a threshold value, Q_{rxn}, in the fraction of native intermolecular contacts, Q, that reproduces the experimental k_{on} for the wild-type proteins. Consistent with previous brute force simulations,[8] b and q were set to 100 and 500 Å, respectively. Upon reaching the q surface, a walker was "recycled" by starting a new walker from the unbound state with the same statistical weight thereby maintaining a steady state and enforcing a constant effective protein concentration (3.2 μM). Upon reaching a particular Q_{rxn} value, a walker was effectively recycled after completing the WE simulation by removing the walker and its replicas prior to calculating the k_{on}.

For each barnase-barstar pair, five independent WE simulations were performed with different initial random seeds for BD propagation. In each simulation, the configurational space of the protein pairs was divided into 760 bins along a progress coordinate that was intended to capture the slowest protein motions of the association process. We used a progress coordinate that consisted of three zones: (a) a "far" zone involving the distance between barnase and barstar, (b) an "intermediate" zone involving the RMS deviation of barstar from its bound-state position following alignment of barnase, and (c) a "near" zone involving the same RMS deviation metric as in part b and the fraction of native contacts between barnase and barstar. Simulations were evaluated for resampling at fixed time intervals τ (or iterations) of 2 ns to maintain 24 walkers per bin. Each simulation was carried out for 1000 iterations, or a molecular time of 2 μs (defined as $N\tau$ where N is the number of iterations).

Propagation of Dynamics. The dynamics of our WE simulations were propagated using the UIOWA-BD software,[13,20] which is the same BD engine that was used for the brute force simulations by Frembgen-Kesner and Elcock.[8] Consistent with these simulations, our WE simulations were performed at a constant temperature of 25 °C using a standard BD algorithm with the inclusion of hydrodynamic interactions (HIs) via calculation of the diffusion tensor using the equations of Rotne and Prager[21] and Yamakawa;[22] the same values were used for the hydrodynamic radii of the pseudoatoms to

DOI: 10.1021/acs.jpcb.5b10747
J. Phys. Chem. B 2016, 120, 117−122

reproduce the translational diffusion coefficients of the corresponding all-atom protein models by the hydrodynamics program HYDROPRO;[23] and a time step of 0.25 ps was used throughout the simulations.

Calculation of k_{on} values. For each barnase-barstar pair, the k_{on} value was computed from each of five independent WE simulations using conformations that were sampled every 20 ps once a steady state was achieved (Figure S1, Supporting Information). These values were then averaged. All WE simulations were sufficiently long to yield relative percent uncertainties in the average k_{on} of <20% (Figure S2, Supporting Information). Uncertainties in the average k_{on} values were represented by calculating 95% confidence intervals. The k_{on} from each WE simulation was calculated using the NAM method according to the following equation:[19]

$$k_{on} = \frac{k_D(b)\beta}{1 - (1 - \beta)k_D(b)/k_D(q)}$$

where $k(b)$ and $k(q)$ are the diffusion rate constants for achieving separation distances of b and q, respectively, and β is the probability of successful collisions, i.e. that a simulation starting from the unbound state with a separation distance of b (100 Å) reaches the bound state before drifting apart to a separation distance of q (500 Å). Assuming that the motions of the binding partners are isotropic, $k(b)$ and $k(q)$ are given by the Smoluchowski result; $k(r) = 4Dr$, where D is the relative translational diffusion coefficient of the two proteins (i.e., the sum of their corresponding diffusion coefficients). As done for the brute force simulations by Frembgen-Kesner and Elcock,[8] we used the estimate from HYDROPRO[23] for D (2.672×10^2 Å^{2}ps^{-1}). The β value was calculated using the following equation:

$$\beta = \frac{f_{SS}^{bound}}{f_{SS}^{bound} + f_{SS}^{qsurf}}$$

where f_{SS}^{bound} is the steady-state flux into the bound state and f_{SS}^{qsurf} is the steady-state flux into the q surface. As evident in the above equations, the influence of HIs is considered in our calculation of the probability of successful collisions (β), but only approximately on the diffusion of the two proteins by using the sum of their diffusion coefficients (D).[24]

Calculation of WE Efficiency. For each barnase−barstar pair, we determined the efficiency of a single WE simulation relative to brute force simulation in computing the k_{on} for each of five independent WE simulations; these efficiencies were then averaged and uncertainties in the efficiencies were determined by calculating the 95% confidence intervals. The efficiency of each WE simulation was calculated using the following equation:

$$\text{efficiency of WE} = \frac{t_{BF}}{t_{WE}}$$

where t_{BF} and t_{WE} are the wall-clock times required by brute force simulation and the WE simulation, respectively, to generate the same number of independent (uncorrelated) association events using the same computing resource (i.e., 256 CPU cores of 2.3 GHz AMD Interlagos processors). 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. The correlation time was determined by monitoring autocorrelation of the flux into the bound state as a function of the lag time and identifying the first lag time that results in zero autocorrelation (within a 95% confidence interval; see Figure S2, Supporting Information). Since it was not practical to directly obtain t_{BF} for all of the barnase-barstar pairs (i.e., the hydrophobic isosteres), we estimated t_{BF} in a consistent manner for each pair using the following equation:

$$t_{BF} = M_{BF}\left(\frac{0.02 \text{ days/trajectory/CPU core}}{256\text{CPU core}}\right)$$

$$M_{BF} = \frac{\text{number of association events}}{\beta}$$

where M_{BF} is the number of trajectories in a brute force simulation to generate the same number of independent association events observed in a WE simulation−given that the brute force trajectories are terminated when the proteins either associate or reach a separation distance of q according to the NAM method; 0.02 days/trajectory/core is the average wall-clock time that would be required to complete a single brute force trajectory before the proteins reach a separation distance of q; and β (as defined above) is the probability calculated by WE for a single brute force trajectory to generate a successful association event before dissociating to a separation distance of q.

RESULTS AND DISCUSSION

Our general strategy for computing k_{on} values from our simulations was to first identify a criterion for successful association that reproduces the experimental k_{on} for wild-type barnase and barstar. Next, we validated the simulations by using this criterion to calculate the k_{on} for R59A mutant barnase and wild-type barstar, which associates 9-fold more slowly than the wild-type proteins,[9] and comparing the calculated k_{on} to the experimental value. Finally, we used this criterion to estimate the basal k_{on}, i.e., the k_{on} for the hydrophobic isosteres in which all effective charges of the wild-type proteins are set to zero. Following the brute force simulations by Frembgen-Kesner and Elcock,[8] our criterion for successful association was to reach a threshold value, Q_{rxn}, in the fraction of native intermolecular contacts, Q; dynamics were propagated using the same BD engine with the inclusion of intramolecular HIs to achieve realistic diffusive properties of the individual proteins; and k_{on} values were calculated according to the NAM method (see Methods).[19]

Validation of the Simulation Strategy. Figure 1 shows the computed k_{on} as a function of Q_{rxn} for all five independent WE simulations of each barnase−barstar pair. The experimental k_{on} for wild-type barnase and barstar (2.86×10^8 M^{-1} s^{-1})[9] was reproduced when using Q_{rxn} values of 0.27 and 0.56 for simulations with and without intermolecular HIs, respectively. These values differ slightly from those determined by Frembgen-Kesner and Elcock using brute force simulations and the same protein models (0.32 and 0.47, respectively)[8] due to more frequent monitoring of the reaction criterion (every 20 ps instead of 100 ps); thus, our WE simulations are less likely to have missed conformations that satisfy the reaction criterion. Importantly, using the Q_{rxn} values that we have identified, the computed k_{on} values for R59A barnase and wild-type barstar are in excellent agreement with experiment, regardless of whether or not intermolecular HIs were included (Figure 2; see also Table S1, Supporting Information). The reproduction of

DOI: 10.1021/acs.jpcb.5b10747
J. Phys. Chem. B 2016, 120, 117−122

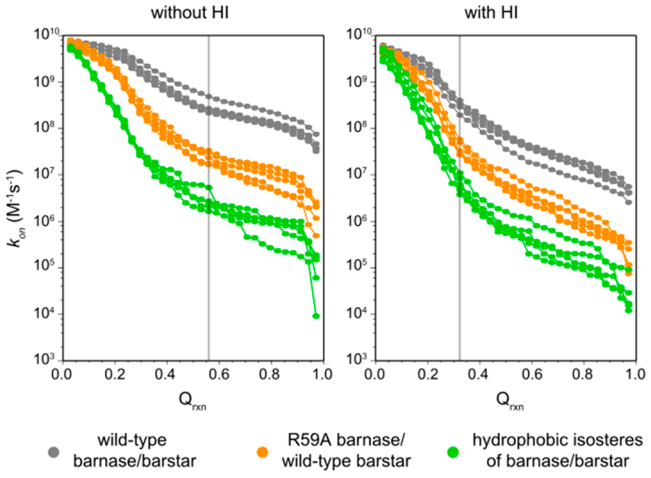


Figure 1. Computed k_{on} values for each barnase-barstar pair from each of five independent WE simulations as a function of the fraction of intermolecular native contacts Q_{rxn}. Results from simulations without and with the inclusion of intermolecular HI are shown in the left and right panels, respectively. The vertical gray line in each panel indicates the Q_{rxn} value that reproduces the experimental k_{on} for the wild-type pair for simulations without and with HI (0.56 and 0.27, respectively) and was used for calculating k_{on} values for the mutant pairs in that panel.

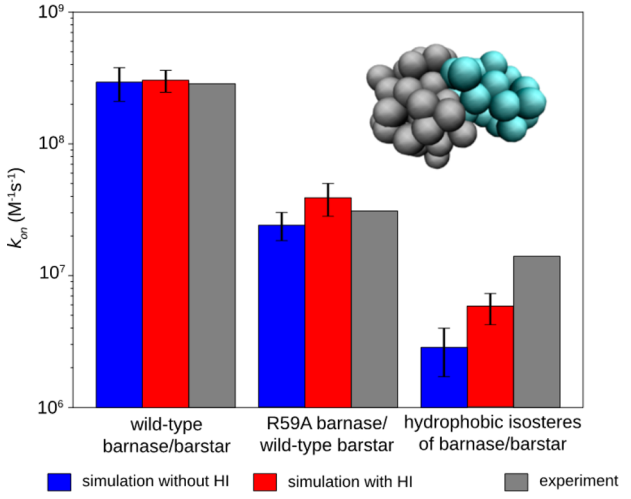


Figure 2. Comparison of computed and experimental[9] k_{on} values on a log scale. Reported values from simulation (with and without intermolecular HI) are averages from five independent WE simulations with the error bars representing 95% confidence intervals. The pseudoatomic protein model of barnase (gray) and barstar (cyan) is shown in the upper right corner.

experimental k_{on} values for both wild-type and mutant pairs of barnase and barstar is consistent with results from brute force simulations,[8] providing validation of our WE simulation protocol. Relative to the basal k_{on}, our computed k_{on} values for wild-type barnase and barstar are 53- and 103-fold faster with and without intermolecular HIs, respectively. These rate enhancements are solely due to the electrostatic interactions between the wild-type proteins given the omission of intramolecular electrostatic interactions in our simulations.

Estimation of the Basal k_{on}. The basal k_{on} computed from our simulations with and without intermolecular HIs are (2.85 ± 1.30) × 10^6 M^{-1} s^{-1} and (5.79 ± 0.17) × 10^6 M^{-1} s^{-1}, respectively. At the effective protein concentration maintained in our simulations (3.2 μM), these rate constants correspond to time scales beyond tens of milliseconds. Our computed basal k_{on} values are similar to those using less computationally intensive strategies; in particular, the use of spherical models with orientational constraints[3−6] has provided estimates in the range of 10^5−10^6 M^{-1} s^{-1} and the use of rigid, atomistic models in either the application of transition-rate theory[7] or direct BD simulation of protein−protein association[2] has yielded estimates of ∼1 × 10^6 M^{-1} s^{-1}. The similarity of our estimates to these previous estimates suggests that flexible models may not be essential for obtaining realistic estimates of k_{on} values for proteins such as barnase and barstar that do not undergo significant conformational changes upon binding (the C_α RMS deviation between the crystal structures of the unbound[25,26] and bound[15] conformations is only 0.5 Å for both barnase and barstar). However, it has not been possible to directly estimate the basal k_{on} with uncertainties of <100% using standard BD simulations with rigid, atomistic models since the association events were much slower in the absence of electrostatic forces.[2] On the other hand, our WE simulations with flexible molecular models enable significantly more precise calculations of the k_{on} (uncertainties of 22−46%) and could therefore be used for even more complicated binding processes, including ones that involve large conformational changes. Notably, our computed k_{on} values are significantly lower than that obtained by experiment from extrapolation to infinite salt concentration (1.4 × 10^7 M^{-1} s^{-1}),[1] suggesting that the favorable electrostatic interactions between the proteins are not completely eliminated at high salt concentrations.

Effect of Intermolecular HIs on the Kinetics of Association. Although the inclusion of intermolecular HIs has no effect on the ability of the simulation model to reproduce the effects of mutation on the k_{on}, for a fixed value of Q_{rxn}, the inclusion of intermolecular HIs significantly slows down the rate of association for all three pairs of the barnase-barstar system (Figures 2 and 3). Surprisingly, the *extent* to

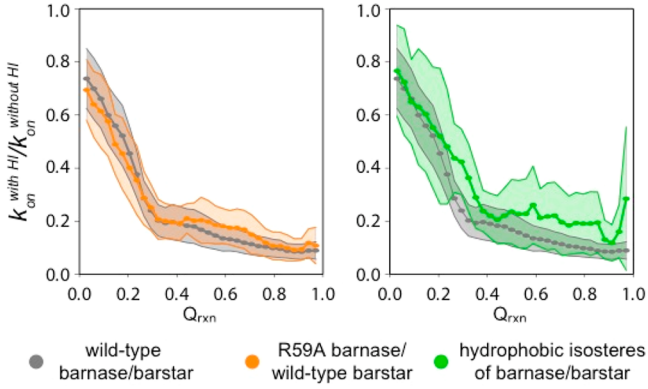


Figure 3. Ratio of association rate constants $k_{on}^{with\ HI}/k_{on}^{without\ HI}$ computed from simulations with and without intermolecular HIs as a function of the fraction of intermolecular contacts Q_{rxn}. The shaded regions represent 95% confidence intervals for averages (filled circles) from five independent WE simulations.

which k_{on} decreases is essentially the same for wild-type and R59A mutant pairs (e.g., by ∼5-fold at Q_{rxn} = 0.27). In contrast, the impact of intermolecular HIs in the brute force simulations by Frembgen-Kesner and Elcock was more pronounced for slower associating mutants of barnase such as R59A in which the electrostatic interactions with barstar are diminished.[8] On

DOI: 10.1021/acs.jpcb.5b10747
J. Phys. Chem. B 2016, 120, 117−122

the basis of these results, it was predicted that the impact would be the most pronounced for the hydrophobic isosteres of barnase and barstar. However, the enhanced sampling provided by the WE strategy reveals no statistical difference between the impact of the intermolecular HIs on the k_{on} for the wild-type and R59A mutant pairs. For the hydrophobic isosteres, our results are inconclusive. Although it was possible to obtain statistically robust estimates of the basal k_{on}—which was the primary goal of this work—our simulations did not reach the level of precision in the *ratio* of the k_{on} values with and without intermolecular HIs that would be required to determine the effect of HIs on the association kinetics relative to the wild-type pair (note the large confidence intervals in Figure 3). For future studies of this effect, significantly greater sampling using a larger number of simulations and/or longer simulations would be required to achieve a sufficient level of precision in the computed k_{on} values, particularly in the absence of intermolecular HIs.

Efficiency of WE Simulation. Finally, it would not have been practical to obtain converged estimates of the basal k_{on} without the use of the WE strategy. In addition, a highly scalable, parallel implementation of this strategy was essential since it would have otherwise required >2 years to carry out the simulations using a serial implementation. To determine the efficiency of parallelized WE vs brute force simulation in estimating the k_{on}, we compared the wall-clock time that would be required of WE vs brute force simulation (both using the NAM framework) to generate the same number of independent (uncorrelated) association events using the same computing resource (256 CPU cores of 2.3 GHz AMD Interlagos processors). Figure 4 shows the efficiencies of WE

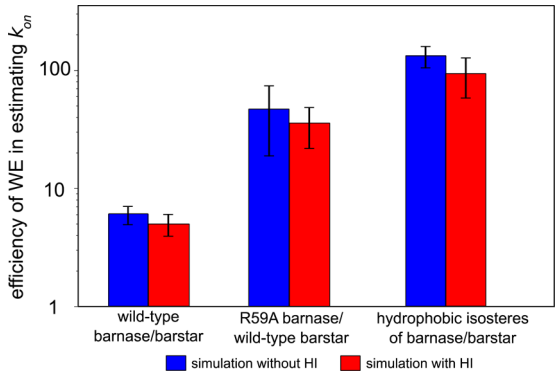


Figure 4. Average efficiencies of WE relative to brute force simulation in computing the k_{on}. Full details for estimating efficiencies are described in Methods. Uncertainties represent 95% confidence intervals for averages from five independent WE simulations.

simulations relative to brute force simulations for each barnase-barstar pair (see also Table S2, Supporting Information). For the wild-type pair, a WE simulation was 6-fold more efficient than brute force simulation with the inclusion of intermolecular HI. This efficiency increased to 46-fold for the R59A mutant pair and ultimately 131-fold for the hydrophobic isosteres. In the latter case, brute force simulation using the same flexible protein models would be highly impractical, requiring 386 days in wall-clock time to generate the same number of association events (>1000) as a single WE simulation, which required only 3 days. The greater efficiency of WE observed for the slower processes (i.e., increasing with the barrier height) is consistent with previous WE studies of other rare events.[12,27−29]

CONCLUSIONS

In conclusion, we have directly computed the basal k_{on} for a protein−protein association process for the first time using flexible models with molecular simulations. In particular, we computed the basal k_{on} for the barnase-barstar system using highly efficient WE simulations. Our computed basal k_{on} is significantly lower than that obtained by experiment from extrapolation to infinite salt concentration, suggesting that the electrostatic interactions are not completely eliminated at high salt concentrations. This result underscores the importance of directly computing the basal k_{on} using the true hydrophobic isosteres of the proteins under regular salt concentrations—a goal that can only be achieved by molecular simulation. Relative to our basal k_{on}, the electrostatic interactions of the wild-type proteins enhance the rate of association by >130-fold. As demonstrated by Frembgen-Kesner and Elcock using brute force simulations,[8] the inclusion of intermolecular HIs significantly decreases the computed k_{on} values for both wild-type and mutant pairs. However, the extensive sampling provided by our WE simulations has revealed that the extent by which the k_{on} is reduced is the same for both the wild-type and R59A mutant pairs. For the hydrophobic isosteres, the relative extent to which the k_{on} was affected by the intermolecular HIs was inconclusive due to insufficient precision in the ratio of the k_{on} with and without intermolecular HI. Finally, our results demonstrate that WE simulations are orders of magnitude more efficient than brute force simulation in providing converged estimates of rate constants for the slow associations of proteins in the complete absence of electrostatic interactions. The computation of such rate constants is otherwise impractical when using flexible protein models—even when these models are coarse-grained. Given its high efficiency, the simulation strategy used in this study would be useful for even more complicated systems, including those that undergo large conformational changes upon binding.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcb.5b10747.

Figures S1 (average calculated k_{on} values as a function of time) and S2 (representative autocorrelation of the flux into the bound state) and Tables S1 (average calculated k_{on} values) and S2 (average efficiencies of weighted ensemble vs brute force) (PDF)

AUTHOR INFORMATION

Corresponding Author

*(L.T.C.) E-mail: ltchong@pitt.edu.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Adrian Elcock for valuable discussions and making the UIOWA-BD software available. We also thank Alex DeGrave and Adam Pratt for critical reading of the manuscript. Financial support was provided by NSF CAREER MCB-0846216, NIH 1R01GM115805-01, and a DAAD graduate research grant. Computational resources were provided by NSF CNS-1229064 and the University of Pittsburgh's Center for Simulation and Modeling.

DOI: 10.1021/acs.jpcb.5b10747
J. Phys. Chem. B 2016, 120, 117−122

REFERENCES

(1) Schreiber, G.; Haran, G.; Zhou, H. X. Fundamental aspects of protein-protein association kinetcs. *Chem. Rev.* **2009**, *109*, 839−860.
(2) Gabdoulline, R. R.; Wade, R. C. Simulation of diffusional association of barnase and barstar. *Biophys. J.* **1997**, *72*, 1917−1929.
(3) Northrup, S. H.; Erickson, H. P. Kinetics of protein-protein association explained by Brownian dynamics computer simulation. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89*, 3338−3342.
(4) Zhou, H. X. Enhancement of protein-protein association rate by interaction potential: accuracy of prediction based on local Boltzmann factor. *Biophys. J.* **1997**, *73*, 2441−2445.
(5) Camacho, C. J.; Kimura, S. R.; Delisi, C.; Vajda, S. Kinetics of desolvation-mediated protein-protein binding. *Biophys. J.* **2000**, *78*, 1094−1105.
(6) Schlosshauer, M.; Baker, D. Realistic protein-protein association rates from a simple diffusional model neglecting long-range interactions, free energy barriers, and landscape ruggedness. *Protein Sci.* **2004**, *13*, 1660−1669.
(7) Alsallaq, R.; Zhou, H. X. Prediction of protein-protein association rates from a transition-state theory. *Structure* **2007**, *15*, 215−224.
(8) Frembgen-Kesner, T.; Elcock, A. H. Absolute protein-protein association rate constants from flexible coarse-grained Brownian dynamics simulations: The role of intermolecular hydrodynamic interactions in barnase-barstar association. *Biophys. J.* **2010**, *99*, L75−L77.
(9) Schreiber, G.; Fersht, A. R. Rapid, electrostatically assisted association of proteins. *Nat. Struct. Biol.* **1996**, *3*, 427−431.
(10) Zwier, M. C.; Adelman, J. L.; Kaus, J. W.; Pratt, A. J.; Wong, K. F.; Rego, N. B.; Suarez, E.; Lettieri, S.; Wang, D. W.; Grabe, M.; et al. WESTPA: An interoperable, highly scalable software package for weighted ensemble simulation and analysis. *J. Chem. Theory Comput.* **2015**, *11*, 800−809.
(11) Huber, G. A.; Kim, S. Weighted-Ensemble Brownian dynamics simulations of protein association reactions. *Biophys. J.* **1996**, *70*, 97−110.
(12) Rojnuckarin, A.; Livesay, D. R.; Subramaniam, S. Bimolecular reaction simulation using Weighted Ensemble Brownian dynamics and the University of Houston Brownian Dynamics program. *Biophys. J.* **2000**, *79*, 686−693.
(13) Frembgen-Kesner, T.; Elcock, A. H. Striking effects of hydrodynamic interactions on the simulated diffusion and folding of proteins. *J. Chem. Theory Comput.* **2009**, *5*, 242−256.
(14) Gabdoulline, R. R.; Wade, R. C. Protein-protein association: investigation of factors influencing association rates by brownian dynamics simulations. *J. Mol. Biol.* **2001**, *306*, 1139−1155.
(15) Buckle, A. M.; Schreiber, G.; Fersht, A. R. Protein-protein recognition: crystal structural analysis of a barnase-barstar complex at 2.0-A resolution. *Biochemistry* **1994**, *33*, 8878−8889.
(16) Gabdoulline, R. R.; Wade, R. C. Effective charges for macromolecules in solvent. *J. Phys. Chem.* **1996**, *100*, 3868−3878.
(17) Go, N. Theoretical studies of protein folding. *Annu. Rev. Biophys. Bioeng.* **1983**, *12*, 183−210.
(18) Takada, S. Gō-ing for the prediction of protein folding mechanisms. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 11698−11700.
(19) Northrup, S. H.; Allison, S. A.; McCammon, J. A. Brownian dynamics simulation of diffusion-influenced bimolecular reactions. *J. Chem. Phys.* **1984**, *80*, 1517−1524.
(20) Elcock, A. H. Molecular simulations of cotranslational protein folding: Fragment stabilities, folding cooperativity, and trapping in the ribosome. *PLoS Comput. Biol.* **2006**, *2*, e98.
(21) Rotne, J.; Prager, S. Variational treatment of hydrodynamic interaction in polymers. *J. Chem. Phys.* **1969**, *50*, 4831−4837.
(22) Yamakawa, H. Transport properties of polymer chains in dilute solution - hydrodynamic interaction. *J. Chem. Phys.* **1970**, *53*, 436−443.
(23) Garcia de la Torre, J.; Huertas, M. L.; Carrasco, B. Calculation of hydrodynamic properties of globular proteins from their atomic-level structure. *Biophys. J.* **2000**, *78*, 719−730.
(24) Dlugosz, M.; Antosiewicz, J. M.; Zielinski, P.; Trylska, J. Contributions of far-field hydrodynamic interactions to the kinetics of electrostatically driven molecular association. *J. Phys. Chem. B* **2012**, *116*, 5437−5447.
(25) Martin, C.; Richard, V.; Salem, M.; Hartley, R.; Mauguen, Y. Refinement and structural analysis of barnase at 1.5 Å resolution. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1999**, *55*, 386−398.
(26) Ratnaparkhi, G. S.; Ramachandran, S.; Udgaonkar, J. B.; Varadarajan, R. Discrepancies between the NMR and X-ray structures of uncomplexed barstar: analysis suggests that packing densities of protein structures determined by NMR are unreliable. *Biochemistry* **1998**, *37*, 6958−6966.
(27) Zhang, B. W.; Jasnow, D.; Zuckerman, D. M. Efficient and verified simulation of a path ensemble for conformational change in a united-residue model of calmodulin. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 18043−18048.
(28) Adelman, J. L.; Grabe, M. Simulating rare events using a weighted ensemble-based string method. *J. Chem. Phys.* **2013**, *138*, 044105.
(29) Zwier, M. C.; Kaus, J. W.; Chong, L. T. Efficient explicit-solvent molecular dynamics simulations of molecular association kinetics: Methane-methane, Na+/Cl-, methane/benzene, and K+/18-crown-6 ether. *J. Chem. Theory Comput.* **2011**, *7*, 1189−1197.

DOI: 10.1021/acs.jpcb.5b10747
J. Phys. Chem. B 2016, 120, 117−122