

Role of water and steric constraints in the kinetics of cavity–ligand unbinding

Pratyush Tiwary^a, Jagannath Mondal^{a,b}, Joseph A. Morrone^c, and B. J. Berne^{a,1}

^aDepartment of Chemistry, Columbia University, New York, NY 10027; ^bTata Institute of Fundamental Research - Centre for Interdisciplinary Sciences, Hyderabad 500075, India; and ^cLaufer Center for Physical and Quantitative Biology, Stony Brook University, Stony Brook, NY 11794

Contributed by Bruce J. Berne, August 20, 2015 (sent for review July 12, 2015; reviewed by Pablo G. Debenedetti and Michael L. Klein)

A key factor influencing a drug's efficacy is its residence time in the binding pocket of the host protein. Using atomistic computer simulation to predict this residence time and the associated dissociation process is a desirable but extremely difficult task due to the long timescales involved. This gets further complicated by the presence of biophysical factors such as steric and solvation effects. In this work, we perform molecular dynamics (MD) simulations of the unbinding of a popular prototypical hydrophobic cavity–ligand system using a metadynamics-based approach that allows direct assessment of kinetic pathways and parameters. When constrained to move in an axial manner, the unbinding time is found to be on the order of 4,000 s. In accordance with previous studies, we find that the cavity must pass through a region of sharp wetting transition manifested by sudden and high fluctuations in solvent density. When we remove the steric constraints on ligand, the unbinding happens predominantly by an alternate pathway, where the unbinding becomes 20 times faster, and the sharp wetting transition instead becomes continuous. We validate the unbinding timescales from metadynamics through a Poisson analysis, and by comparison through detailed balance to binding timescale estimates from unbiased MD. This work demonstrates that enhanced sampling can be used to perform explicit solvent MD studies at timescales previously unattainable, to our knowledge, obtaining direct and reliable pictures of the underlying physiochemical factors including free energies and rate constants.

ligand unbinding | kinetics | enhanced sampling | dewetting transition

The unbinding of ligands from host substrates is a phenomenon widely occurring across biological and chemical sciences. It is of great interest to be able to understand the thermodynamics and kinetics of such processes, especially how they are influenced by solvent and steric effects. An accurate estimate of unbinding kinetics is in fact of crucial importance for drug discovery paradigms (1, 2). However, despite the advent of massively parallel computer resources, it has not been so easy to simulate the dynamics of ligand unbinding and calculate associated rate constants. The complications are mainly twofold. First, as has been seen in studies of model systems (3–10), various proteins (11–13), HIV (14), and actual anticancer drugs (15–17), the solvent often manifests itself at the molecular scale. Whereas coarse-grained models can be fit to explicit solvent molecular dynamics (MD) simulations (3), predictive power can be attained only by performing all-atom MD. Second, performing all-atom MD for unbinding of such systems is however plagued by the timescale problem. MD is restricted to integration timesteps of a few femtoseconds, which can be partially mitigated by multiple timestep MD algorithms (18). However, it is not yet routinely feasible to go into the millisecond regime and beyond for any system with more than a few thousand atoms.

In this paper we consider a popular prototypical cavity–ligand system in explicit water where the attraction between water and the two nanoscale objects, namely a fullerene molecule and a spherical cavity, is weak (3, 4, 6–10). We provide a full dynamical picture of the unbinding process demonstrating the clear role of water. We find that even in this relatively simple system there

exists a rich range of dynamics that changes qualitatively and quantitatively as a function of the cavity–ligand distance and the motional degrees of freedom. Previous pioneering studies (6–8, 10, 19–21) involving explicit all-atom MD, Brownian dynamics, transition path sampling, and other approaches have clearly shown that the association in such systems has a clear signature of solvent fluctuations and a sharp dewetting transition as the nanoscale objects approach each other—if sterically constrained to move along the axis of symmetry. This is a popular setup that has been considered in numerous studies over the years (3, 6–8, 10), and is suggestive of biological systems where steric hindrances in the binding pocket do not allow the ligand to roll or move in a free manner (12, 15). Using unbiased MD and Brownian dynamics tools, it was previously possible to calculate the timescales of association or binding for such systems that explicitly accounted for the dewetting transition (3).

However, apart from one related recent work (12), to the best of our knowledge there is no reported study in which the timescales of the analogous dissociation or unbinding process were calculated through MD simulations. For such cavity–ligand systems due to the very high energy barrier of $\sim 30\text{--}40 k_B T$ (3), the unbinding timescales are simply too slow to be amenable through unbiased MD calculations. As such, in this work we use the popular enhanced sampling technique metadynamics (22–25) along with its recent extension (12, 26) for obtaining unbiased dynamics to calculate free-energy profiles and unbinding rate constants for the cavity–ligand system. Furthermore, we ask and answer the following question: How does the dynamics of cavity–ligand association and dissociation depend on motional degrees of freedom? That is, how would the timescales of binding/unbinding

Significance

The unbinding of ligand–substrate systems in molecular water is a problem of great theoretical and practical interest. To understand the dynamical nature of the unbinding, it is desirable to use atomistic techniques like molecular dynamics (MD). However, the associated timescales are typically too long for MD to be applicable. In this work, we apply a recent metadynamics scheme that allows the use of MD to get information on thermodynamics and kinetics of systems that are virtually impossible to treat with unbiased MD. We calculate unbinding pathways and timescales, and show that solvent molecules play a crucial role in cooperation with steric effects. The approaches used here demonstrate a broadly applicable methodology for studying ligand unbinding.

Author contributions: P.T., J.M., J.A.M., and B.J.B. designed research; P.T. and J.M. performed research; P.T., J.M., and J.A.M. analyzed data; and P.T., J.M., J.A.M., and B.J.B. wrote the paper.

Reviewers: P.G.D., Princeton University; and M.L.K., Temple University.

The authors declare no conflict of interest.

¹To whom correspondence should be addressed. Email: bb8@columbia.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1516652112/-DCSupplemental.

vary between the cases where the ligand can/cannot undergo free-to-move motion in any direction?

We find that the unbinding, in the case of motion being sterically constrained along the axis of symmetry of the system, proceeds through a sharp wetting transition at a critical ligand–cavity separation, conforming to the picture presented by Mondal et al. through their position-dependent friction calculation (3). The mean unbinding time of this system is found to be around 4,000 s. The transition pathways harvested from our metadynamics-assisted MD runs are in perfect accordance with the previous calculations of Mondal et al. (3). However, when the motion restraint is removed and the ligand is free to move in any direction, it finds an alternate pathway wherein there is no abrupt wetting transition, and the mean unbinding time reduces 20-fold to 200 s. The binding times are also reduced. We validate rigorously all free-energy profiles and rate constants through a combination of umbrella sampling, unbiased MD when feasible, and the principle of detailed balance. The rate constants are further validated also through the Kolmogorov–Smirnov test for Poisson distribution proposed in ref. 27, thus giving further confidence in the estimated dynamics.

This work thus provides useful insight into the phenomenon of hydrophobic interaction in solvated nanoscale systems (28), allowing one to directly simulate the unbinding process in MD despite the very high associated barriers. It also demonstrates that with a careful use of recently developed enhanced sampling techniques, one can perform molecular dynamics studies of unbinding/binding that extend well into the seconds timescale, and provide statistically accurate thermodynamic and kinetic information. The ideas that are used in this work are fairly generic and should be applicable to a large range of studies pertaining especially to ligand unbinding in explicit solvent.

Methods

Throughout this work, we consider a cavity–ligand system (Fig. 1) where the attraction between water and both the nanoscale objects is weak and leaning toward a hydrophobic system (see the *SI Appendix* for detailed potential forms and parameters). The ratio of diameters of the ligand and cavity as well as other interaction parameters are taken to be the same as in ref. 3. In that study, the particular cavity–ligand diameter ratio was chosen to be just large enough to allow water molecules to enter/exit. Whereas the primary method in this paper is metadynamics, we directly and indirectly validate any findings from metadynamics with alternate independent approaches.

Metadynamics for Free-Energy Reconstruction. Metadynamics is a widely used method for exploring complex free-energy surfaces characterized by high free-energy barriers (22–25, 29). One first identifies a small number of slowly changing order parameters, called collective variables (CVs) (30). A memory-dependent biasing potential is constructed through the simulation as a function of these CVs, typically in the form of repulsive Gaussians added wherever the system visits in the CV space. Thus, the system slowly starts to avoid the places where it has already visited. This leads to a gradual enhancement of the fluctuations in the CVs, through which the system is discouraged from getting trapped in the low free-energy basins. At the end of a metadynamics run the probability distribution of any observable, whether biased directly or not, can be computed through a reweighting procedure (24, 31). This easy reweighting functionality is one of the many features of metadynamics that has made it a very popular method for calculation of free-energy surfaces.

For the free-energy reconstruction through metadynamics, in the case when the ligand is sterically constrained to move along the axis of symmetry (z axis in Fig. 1), we perform 1D metadynamics with the z coordinate as the only CV. For the case when the ligand is free to move in any direction, we perform 2D metadynamics with z and the radial distance ($\rho = \sqrt{x^2 + y^2}$) from the axis of symmetry as the two CVs. In either case bias is added every 600 fs. In the *SI Appendix* we report the values of all other relevant parameters for both cases. In either case, we use restraining walls at high ligand–cavity separation to facilitate multiple reentry events (see the *SI Appendix* for details of restraints).

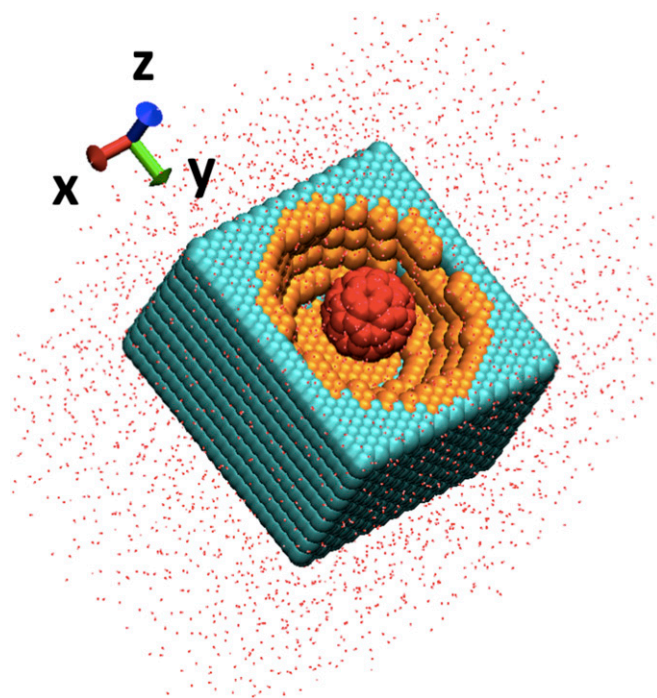


Fig. 1. Cavity–ligand system in explicit water with axes marked. Red: fullerene-shaped ligand atoms. Orange: cavity atoms that interact with the ligand. Green: wall atoms. See the *SI Appendix* for corresponding interaction potentials.

From Metadynamics to Dynamics. Recently, Tiwary and Parrinello extended the scope of metadynamics by showing how to extract unbiased rates from biased ones with minimal extra computational burden (26). For this, they made two key assumptions on the dynamics:

- i) The transition processes are characterized by movements from one stable state to another via dynamical bottlenecks that are crossed only rarely, but when such a transition does happen, the time spent in the bottleneck is small.
- ii) Whereas there is no need to know beforehand the precise nature or location of these bottlenecks, one should have CVs that can distinguish between all stable basins of relevance. Note that this CV does not have to be the true reaction coordinate (26, 32).

Under these two key assumptions, by making the bias deposition slower than the time spent in dynamical bottlenecks, it is possible to keep the transition states (TS) relatively bias-free through the course of metadynamics. This so-called “infrequent metadynamics” approach (12, 26, 27) preserves the unbiased sequence of state-to-state transitions and allows one to access the acceleration of transition rates achieved through biasing by appealing to generalized TS theory (33) and calculating the following simple running average (26, 34, 35):

$$\alpha = \langle e^{\beta V(s,t)} \rangle_t, \quad [1]$$

where s is the collective variable being biased, β is the inverse temperature, $V(s, t)$ is the bias experienced at time t , and the subscript t indicates averaging under the time-dependent potential. The above expression gives the ratio of actual to metadynamics transition times, and is valid even if there are multiple intermediate states and numerous alternative reactive pathways (26, 27).

In a successive work, a way was also proposed to assess the reliability of the two assumptions above (27). This relies on the fact that the escape times from a long-lived metastable state obey a time-homogeneous Poisson statistics (27) with a single rate law. A statistical analysis based on the Kolmogorov–Smirnov test can quantitatively assess how precisely the above assumptions have been met (27). Thus, if (i) significant bias got deposited in the TS region even with infrequent biasing, or (ii) there are hidden unidentified timescales at play that the CV does not resolve, it would lead to failing the test for time-homogeneous Poisson statistics.

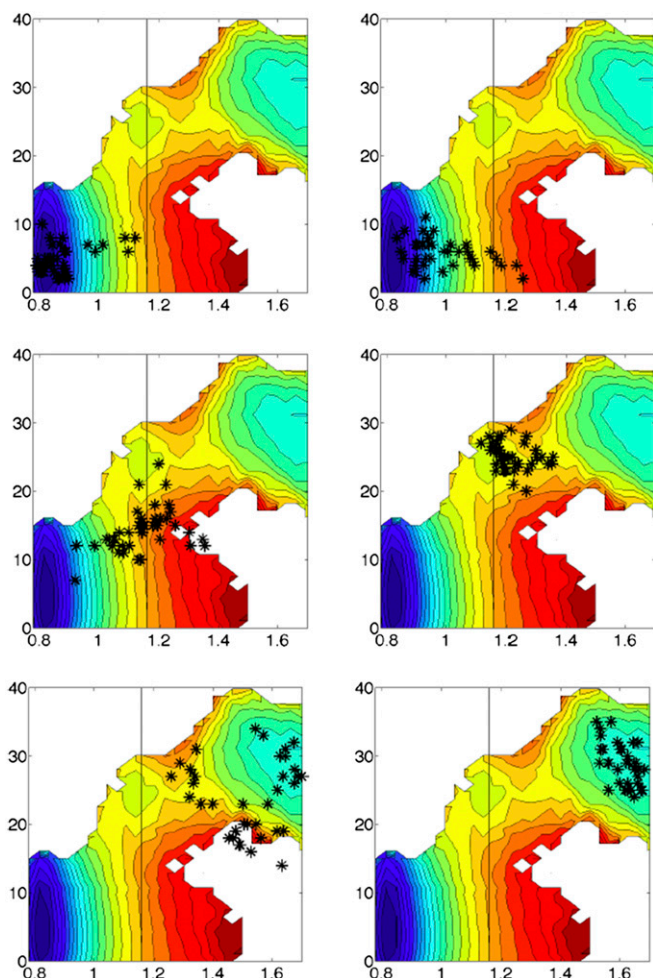


Fig. 3. Typical 60-ps-long reactive trajectory obtained via infrequent metadynamics providing a clear description of the wetting transition as the sterically constrained ligand crosses over from bound to unbound state. Each snapshot proceeding row-by-row from top left to bottom right represents data points from 10-ps-long MD trajectory sections projected on (z, w) space. x axis here is the z distance, and y axis is w , the number of pocket waters. The solid black vertical line marks the critical z value at which ref. 3 had predicted the existence of large solvent fluctuations. The underlying free-energy surface, axes, and the contours are same as in Fig. 2B.

the *SI Appendix* for detailed analysis), demonstrating that (i) the bias deposition was infrequent enough to not gradually corrupt the TSs, and (ii) biasing the CV z was sufficient to ensure the timescale separation needed for the infrequent metadynamics approach to be applicable.

To further validate the unbinding time estimate of $3,863 \pm 1,032$ s, we invoke the principle of detailed balance. We use the 1D free-energy profile (Fig. 2) and the relation unbinding time = binding time $\times e^{-\beta\Delta G}$, where ΔG is the free-energy difference between the solvated and bound ligand. Here the mean binding time means the time for the fully solvated ligand ($z=1.4$) to get bound. Note that for the purpose of checking detailed balance, as is our objective here, any z value would be fine. As per Fig. 2, the estimate of $\Delta G = -72.5$ kJ/mol from metadynamics, giving us a mean binding time of 769 ± 198 ps, which is well within order of magnitude agreement with the value of 476 ± 33 ps through unbiased MD simulations reported by Mondal et al. (3).

In Fig. 3 we provide a set of snapshots of a typical ~ 60 -ps-long reactive trajectory as the system moves successfully from bound to unbound state. Here each snapshot corresponds to

10 ps of MD projected on the 2D (z, w) space, where w is the number of pocket waters (see ref. 3 for precise definition of w). This provides extremely clear dynamical evidence of the dewetting transition in such a hydrophobic and sterically constrained system. The solid black vertical line in Fig. 3 marks the critical z value at which ref. 3 had predicted the existence of large solvent fluctuations. Our MD thus qualitatively and quantitatively validates the prediction of that and previous works. In Table 1, we provide a summary of the timescales for all of the cases obtained through metadynamics, unbiased MD, and through the use of detailed balance.

Ligand Free To Move in Any Direction.

Thermodynamic profile. Fig. 4A shows the 2D free energy as a function of z and $\rho = \sqrt{x^2 + y^2}$ obtained through metadynamics. For all practical purposes this free-energy surface is indistinguishable from the one obtained through umbrella sampling and reported in the *SI Appendix*. In the *SI Appendix* we also provide a comparison of the 1D free energy as a function of z obtained from metadynamics and from umbrella sampling. Note from Fig. 4 that when constraints are lifted, the free-energy minimum is now slightly off-center and deeper than the free energy at the geometric center of the cavity. We find that for this case the system avoids the central dewetting pathway and instead takes an alternate route, rolling in and out along the sides of the cavity. This pathway is favorable compared with the axially symmetric path as hydrophobic contacts between the ligand and the cavity can be maintained until late stages of unbinding. This can be seen from the shape of the free-energy profile in Fig. 4A, which shows how the minimum energy path is along the sides of the pocket, thus maintaining contacts with the pocket wall (white-space region in Fig. 4A). Indeed, even though the free-energy difference between bound and unbound states is larger than for the sterically constrained case, the barrier to unbinding along this alternate pathway is smaller (Fig. 4 and *SI Appendix*, Fig. S1B), and as we show next in *Kinetics*, the unbinding is faster.

Fig. 4B shows the 2D free energy but now as a function of ligand–cavity separation $d = \sqrt{x^2 + y^2 + z^2}$ and the pocket water occupancy w . This is to be contrasted with Fig. 2 for the sterically constrained motion case. This clearly demonstrates that the water

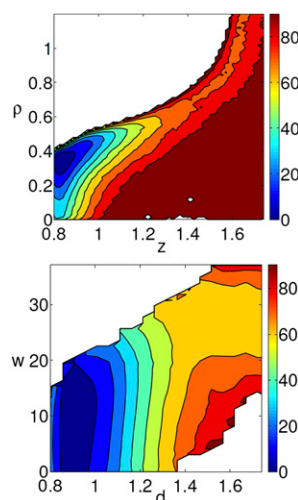


Fig. 4. Various free-energy profiles from metadynamics for the case when the ligand is free to move in any direction. A is a function of (z, ρ) , whereas B is as a function of (d, w) . Note the clear lack of a sharp dewetting transition in B, in contrast with Figs. 2 and 3. All energies are in kJ/mol ($1 k_B T$ at 300 K = 2.5 kJ/mol) and contours are separated by 10 kJ/mol.

occupancy in this case changes gradually without any sharp transitions at specific cavity–ligand separations.

Kinetics. To obtain unbinding kinetics for this case, we perform 14 independent infrequent 1D metadynamics biasing $d = \sqrt{x^2 + y^2 + z^2}$, starting from fully bound ligand and stopping when unbound (see *Methods* and *SI Appendix* for detailed simulation parameters). For the current purpose, we take the unbound ligand as $z = 1.5$, $\rho = 0.6 - 0.8$. The unbinding time now reduces to 200 ± 51 s, with again an excellent Poisson fit demonstrating reliability of the choice of CV and the timescales so generated. Using the free-energy difference of -69.8 kJ (*SI Appendix, Fig. S1B*) between the bound and unbound states and the principle of detailed balance, we obtain a mean binding time of 118 ± 31 ps. This is again in very good agreement with the estimate of 157 ± 5 ps through unbiased MD.

Discussion

In this paper, we have provided detailed thermodynamic and kinetic insights into the binding and unbinding of a popular prototypical ligand-substrate system through fully atomistic metadynamics-assisted MD simulations performed in explicit water. Such systems are very relevant to the understanding of a range of processes across biology, chemistry, and biochemistry. We find that in accordance with previous Brownian-dynamics-based studies (3, 6, 7), the binding of such a system proceeds through a marked dewetting transition, but only when the system is sterically constrained to move along its axis of symmetry. We calculated the unbinding time to be around 4,000 s, with an associated barrier of roughly 80 kJ/mol. When the steric constraint is removed, the dewetting transition becomes continuous, and the unbinding

proceeds through an alternate preferred pathway in which the residence time of the ligand decreases 20-fold to 200 s. These extremely long unbinding timescales were obtained through the use of metadynamics with its recent extension (26), and were validated through alternate simulation techniques and the principle of detailed balance. Clearly the ratio of diameters of cavity and ligand will be a crucial parameter in determining the kinetics and thermodynamics of cavity–ligand systems, adding further richness and complexity to the process, but which can still be treated through the simulation methodologies of this work.

Even though the systems considered here are far simpler than an actual complex drug-protein system, we believe this work has multiple useful implications. Firstly, this is one of the first times, to our knowledge, that such a quantitatively insightful study has been carried out on realistic ligand-substrate systems with very slow unbinding kinetics, and validated through a range of simulation techniques. The relative simplicity of the system allowed us to validate the timescales through detailed balance by performing thorough unbiased MD simulations of the binding, thereby giving confidence in the use of metadynamics type technique for getting unbinding kinetics in more complex protein-ligand systems (12). Secondly, our work shows how the presence of simple steric constraints can heavily influence the role played by molecular solvent. We hope that our work will serve as a useful addition to the understanding of hydrophobic interactions in solvated systems in biology and chemistry.

ACKNOWLEDGMENTS. This work was supported by grants from the National Institutes of Health (NIH-GM4330) and the Extreme Science and Engineering Discovery Environment (TG-MCA08X002).

3. Copeland RA, Pompliano DL, Meek TD (2006) Drug-target residence time and its implications for lead optimization. *Nat Rev Drug Discov* 5(9):730–739.
4. Pan AC, Borhani DW, Dror RO, Shaw DE (2013) Molecular determinants of drug-receptor binding kinetics. *Drug Discov Today* 18(13-14):667–673.
5. Mondal J, Morrone JA, Berne BJ (2013) How hydrophobic drying forces impact the kinetics of molecular recognition. *Proc Natl Acad Sci USA* 110(33):13277–13282.
6. Setny P, Baron R, Michael Kekenes-Huskey P, McCammon JA, Dzubiel J (2013) Solvent fluctuations in hydrophobic cavity-ligand binding kinetics. *Proc Natl Acad Sci USA* 110(4):1197–1202.
7. Remsing RC, et al. (2015) Pathways to dewetting in hydrophobic confinement. *Proc Natl Acad Sci USA* 112(27):8181–8186.
8. Morrone JA, Li J, Berne BJ (2012) Interplay between hydrodynamics and the free energy surface in the assembly of nanoscale hydrophobes. *J Phys Chem B* 116(11):378–389.
9. Li J, Morrone JA, Berne BJ (2012) Are hydrodynamic interactions important in the kinetics of hydrophobic collapse? *J Phys Chem B* 116(37):11537–11544.
10. Sharma S, Debenedetti PG (2012) Evaporation rate of water in hydrophobic confinement. *Proc Natl Acad Sci USA* 109(12):4365–4370.
11. Giovambattista N, Rossky PJ, Debenedetti PG (2006) Effect of pressure on the phase behavior and structure of water confined between nanoscale hydrophobic and hydrophilic plates. *Phys Rev E Stat Nonlin Soft Matter Phys* 73(4 Pt 1):041604.
12. Baron R, Setny P, McCammon JA (2010) Water in cavity-ligand recognition. *J Am Chem Soc* 132(34):12091–12097.
13. Young T, Abel R, Kim B, Berne BJ, Friesner RA (2007) Motifs for molecular recognition exploiting hydrophobic enclosure in protein-ligand binding. *Proc Natl Acad Sci USA* 104(3):808–813.
14. Tiwary P, Limongelli V, Salvalaglio M, Parrinello M (2015) Kinetics of protein-ligand unbinding: Predicting pathways, rates, and rate-limiting steps. *Proc Natl Acad Sci USA* 112(5):E386–E391.
15. Rasaiah JC, Garde S, Hummer G (2008) Water in nonpolar confinement: From nanotubes to proteins and beyond. *Annu Rev Phys Chem* 59:713–740.
16. Braaten D, Ansari H, Luban J (1997) The hydrophobic pocket of cyclophilin is the binding site for the human immunodeficiency virus type 1 Gag polyprotein. *J Virol* 71(3):2107–2113.
17. Shan Y, et al. (2011) How does a drug molecule find its target binding site? *J Am Chem Soc* 133(24):9181–9183.
18. Mondal J, Friesner RA, Berne BJ (2014) Role of desolvation in thermodynamics and kinetics of ligand binding to a kinase. *J Chem Theory Comput* 10(12):5696–5705.
19. Ladbury JE (1996) Just add water! The effect of water on the specificity of protein-ligand binding sites and its potential application to drug design. *Chem Biol* 3(12):973–980.
20. Tuckerman M, Berne BJ, Martyna GJ (1992) Reversible multiple time scale molecular dynamics. *J Chem Phys* 97(3):1990–2001.
21. Willard AP, Chandler D (2008) The role of solvent fluctuations in hydrophobic assembly. *J Phys Chem B* 112(19):6187–6192.
22. Bolhuis PG, Chandler D (2000) Transition path sampling of cavitation between molecular scale solvophobic surfaces. *J Chem Phys* 113(18):8154–8160.
23. Hummer G, Garde S (1998) Cavity expulsion and weak dewetting of hydrophobic solutes in water. *Phys Rev Lett* 80(19):4193–4196.
24. Laio A, Parrinello M (2002) Escaping free-energy minima. *Proc Natl Acad Sci USA* 99(20):12562–12566.
25. Barducci A, Bussi G, Parrinello M (2008) Well-tempered metadynamics: A smoothly converging and tunable free-energy method. *Phys Rev Lett* 100(2):020603–020606.
26. Tiwary P, Parrinello M (2015) A time-independent free energy estimator for metadynamics. *J Phys Chem B* 119(3):736–742.
27. Dama JF, Parrinello M, Voth GA (2014) Well-tempered metadynamics converges asymptotically. *Phys Rev Lett* 112(24):240602–240605.
28. Tiwary P, Parrinello M (2013) From metadynamics to dynamics. *Phys Rev Lett* 111(23):230602.
29. Salvalaglio M, Tiwary P, Parrinello M (2014) Assessing the reliability of the dynamics reconstructed from metadynamics. *J Chem Theory Comput* 10(4):1420–1425.
30. Berne BJ, Weeks JD, Zhou R (2009) Dewetting and hydrophobic interaction in physical and biological systems. *Annu Rev Phys Chem* 60:85–103.
31. Bonomi M, et al. (2009) Plumed: A portable plugin for free-energy calculations with molecular dynamics. *Comput Phys Commun* 180(10):1961–1972.
32. Barducci A, Bonomi M, Parrinello M (2011) Metadynamics. *Wiley Interdiscip Rev Comput Mol Sci* 1(5):826–843.
33. Bonomi M, Barducci A, Parrinello M (2009) Reconstructing the equilibrium Boltzmann distribution from well-tempered metadynamics. *J Comput Chem* 30(11):1615–1621.
34. Best RB, Hummer G (2005) Reaction coordinates and rates from transition paths. *Proc Natl Acad Sci USA* 102(19):6732–6737.
35. Berne BJ, Borkovec M, Straub JE (1988) Classical and modern methods in reaction rate theory. *J Phys Chem* 92(13):3711–3725.
36. Voter AF (1997) Hyperdynamics: Accelerated molecular dynamics of infrequent events. *Phys Rev Lett* 78:3908–3911.
37. Grubmüller H (1995) Predicting slow structural transitions in macromolecular systems: Conformational flooding. *Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics* 52(3):2893–2906.
38. Bolhuis PG, Chandler D, Dellago C, Geissler PL (2002) Transition path sampling: Throwing ropes over rough mountain passes, in the dark. *Annu Rev Phys Chem* 53(1):291–318.

Supporting Information

Tiwary, Mondal, Morrone and Berne

September 3, 2015

1 Simulation set-up for MD, metadynamics and umbrella sampling calculations

The model ligand used in this work is a C_{60} fullerene and the pocket is an ellipsoidal hole carved from a hydrophobic slab, all interacting via Lennard-Jones potentials and enclosed by a periodic box with explicit water and cubic edge length 5.96nm. This system was introduced previously in Ref.[1]. The pocket sites are fixed and interact with the model ligand with a Lennard-Jones potential having $\sigma = 0.4152$ nm, kept same for all interactions. The pocket itself comprises 2 types of atomic species, interacting with the ligand atoms (color red in Fig. 1 in main text) as described below. The system comprised a total of 34296 atoms, with the total number of ligand, cavity and solvent atoms equaling 60, 9020 and 25216 respectively.

1. cavity atoms (color orange in Fig. 1 in main text), with Lennard-Jones $\epsilon=0.008$ kJ/mol.
2. wall atoms (color green in Fig. 1 in main text), with Lennard-Jones $\epsilon=0.0024$ kJ/mol.

The solute-solvent interactions are represented by the geometric mean of the respective water and solute parameters, in accordance with the OPLS formalism[2]. All simulations are performed in explicit TIP4P water [3] with the GROMACS 4.5.4 MD package[4], patched with the PLUMED plugin[5]. During the equilibration stage, temperature and pressure are controlled with the stochastic velocity rescaling thermostat [6] and Berendsen barostat[7]. The production runs were NVT (constant number, volume, temperature) with a temperature of 300K.

The PLUMED plugin [5] was used to carry out metadynamics and umbrella sampling calculations. An integration time-step of 2fs was used for all unbiased MD as well as metadynamics runs.

2 Simulation details for umbrella sampling

We employed umbrella sampling simulation technique for independent test of the free energy surfaces obtained through metadynamics. The umbrella sampling protocol followed for the case when the ligand is sterically constrained, is identical to that reported in Ref. [1] and we refer the reader to that reference.

Here we report the protocol followed to get the two-dimensional free energy surfaces in the scenario when the ligand is free to roll in any direction. Specifically, we discretize along z direction in a range of $z = 0.8$ to 2.0 nm at intervals of 0.1 nm and along $\rho = \sqrt{x^2 + y^2}$ direction in the range $\rho = 0$ to 1.2 nm at intervals of 0.1 nm. Here a value of $\rho = 0$ implies that the ligand is in centrosymmetric orientation with the cavity and any positive value of ρ implies a deviation of the centrosymmetric arrangement of the ligand relative to pocket. Thus in the 2-dimensional reaction space there are a total of $13 \times 13=169$ windows. We generate the initial configurations corresponding to each umbrella window by performing a steered MD (SMD) simulations along ρ for a given z window starting with a centrosymmetrical ($\rho = 0$) orientation and subsequently collecting corresponding configurations from the resulting SMD trajectories for all desired ρ and z combinations. For each of the 169 windows, we used a two-dimensional harmonic restraining potential. We adopt harmonic restraints $U=0.5K_\rho(\rho - \rho_0)^2 + 0.5K_z(z - z_0)^2$ with $K_\rho=4500$ kJ/mol/nm² and $K_z=4500$ kJ/mol/nm². The values of K_ρ and K_z were chosen such that the distributions of the corresponding reaction coordinates around the desired ρ_0 and z_0 are Gaussian in nature and there is significant overlap among adjacent windows.

Each of the 169 windows was run on 8 processors simultaneously. Each of the windows took 14.5 hours for completion and hence total wall clock time was 102 days. However, availability of 15 machines reduced the total required time to 102/15 or around 1 week.

In Fig. 1a we provide the 2-d free energy obtained through umbrella sampling for the case when the ligand is free to move in any direction. This is virtually indistinguishable from the free energy obtained through metadynamics (see Fig. 4a in main text). A similar comparison for the case when the ligand is sterically constrained has already been provided in the main text (Fig. 2a). In Fig. 1b we give the 1-d free energy as a function of z obtained from umbrella sampling and metadynamics. The barrier can be seen to be lower than the case reported in main text where the fullerene is sterically constrained (Fig 2a in main text).

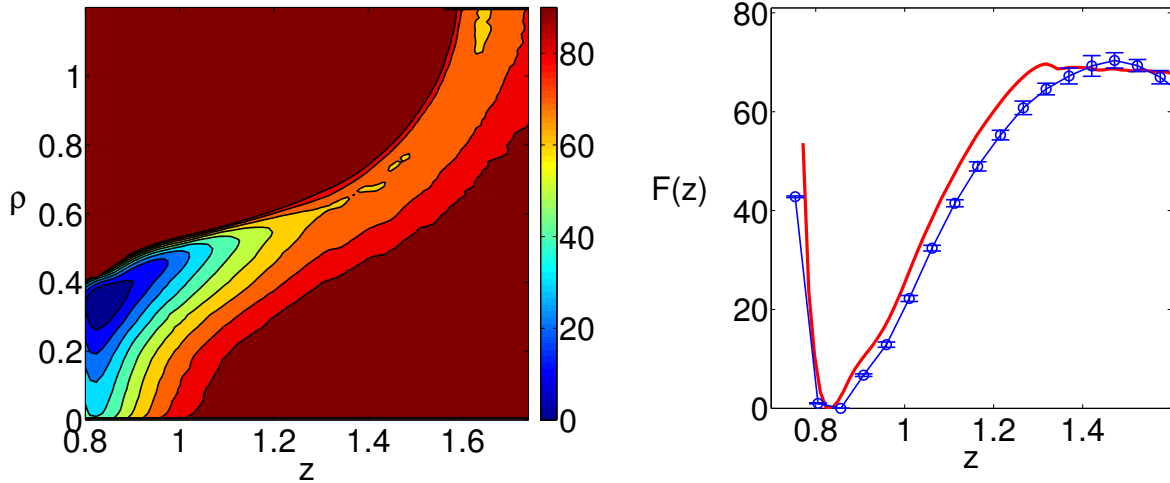


Figure 1: (a) Free energy profile from umbrella sampling for the case when the ligand is free to move in any direction, as a function of (z, ρ) . This is to be compared with Fig. 4(a) in the main text, which was obtained through metadynamics. (b) A comparison of the 1-d free energy as a function of z obtained from umbrella sampling (red line) and metadynamics (blue line with error bars). All energies are in kJ/mol and contours are separated by 10 kJ/mol.

3 Simulation details for metadynamics

In this section we provide details of CV, biasing kernel, biasing frequency and well-tempered metadynamics bias factor for various cases. The metadynamics run for free energy construction took around 5 days of wall-clock time on 32 processors, while the relatively slower infrequent metadynamics runs consumed a total of nearly 2 weeks of wall-clock time on 120 processors, with 8 processors for each of the 14 independent runs. Note that for the latter category, this number represents an averaged value given the Poisson statistics of unbinding time and thus of the required computer time per event.

3.1 Fullerene sterically constrained

3.1.1 Metadynamics for free energy construction

For these calculations, bias was added on z -coordinate as the only CV using gaussian width of 0.02 nm. Gaussians were deposited every 0.6 ps, with a starting height of 2 kJ/mol and gradually decreased on the basis of well-tempered metadynamics biasing factor $\gamma = 15$ [8]. Quartic restraining wall was used at $z = 1.7$.

3.1.2 Metadynamics for rate constants

For these calculations, bias was added on z -coordinate as the only CV using gaussian width of 0.01 nm. Gaussians were deposited every 10 ps, with a starting height of 0.25 kJ/mol and gradually decreased on the basis of well-tempered metadynamics biasing factor $\gamma = 10$ [8]. Since these calculations were stopped after first escape, no restraining walls were used.

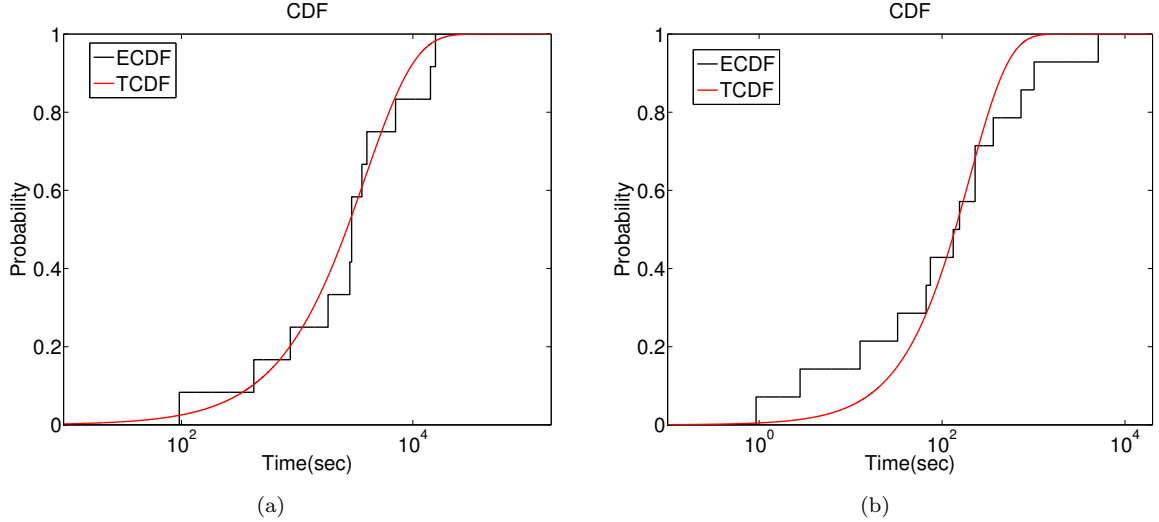


Figure 2: Poisson fit analysis for ligand (a) sterically constrained and (b) free to move in any direction. ECDF (black line) and TCDF (red line) denote empirical and theoretical (i.e. best fit) cumulative distribution functions respectively. The respective p-values for (a) and (b) are 0.72 and 0.64, which are well above the statistical threshold of 0.05.

3.2 Fullerene free to move in any direction

3.2.1 Metadynamics for free energy construction

For these calculations, bias was added on z - and the radial distance ($\rho = \sqrt{x^2 + y^2}$) from the axis of symmetry as the two CVs, using gaussian widths of 0.03 nm and 0.01 nm respectively. Gaussians were deposited every 0.6 ps, with a starting height of 2 kJ/mol and gradually decreased on the basis of well-tempered metadynamics biasing factor $\gamma = 15$ [8]. Quartic restraining walls were used at $z = 2.1$ and $\rho = 1.2$.

3.2.2 Metadynamics for rate constants

For these calculations, bias was added on $d = \sqrt{x^2 + y^2 + z^2}$ -coordinate as the only CV, using gaussian width of 0.03 nm. Gaussians were deposited every 10 ps, with a starting height of 1.2 kJ/mol and gradually decreased on the basis of well-tempered metadynamics biasing factor $\gamma = 15$ [8]. Since these calculations were stopped after first escape, no restraining walls were used [8].

4 Kolmogorov-Smirnoff tests

The unbinding of a ligand from a binding pocket is a typical example of a rare event; the distribution of the associated transitions times is thus expected to be exponential, characteristic of a homogeneous Poisson process[9]. Comparing the empirical distribution of times obtained from metadynamics simulation using Eq. (1) in main text, with a theoretical exponential distribution provides an *a posteriori* assessment[9] of the fulfillment of requirements of the approach detailed in Ref. [10]. To compare the two distributions we proceed constructing the empirical cumulative distribution function (ECDF) of transition times obtained from simulations. Then the characteristic time of the corresponding theoretical Poisson process τ , which is the timescale reported in the text, is obtained through a least squares fitting of the ECDF with the theoretical expression of a cumulative distribution function in the case of a homogeneous Poisson process (TCDF):

$$\text{TCDF} = 1 - \exp\left(-\frac{t}{\tau}\right) \quad (1)$$

This allow to obtain an estimate of the characteristic time of the Poisson process associated with the transition times obtained from simulation. To compare the theoretical (TCDF) and empirical (ECDF) distributions we carry out a Kolmogorov-Smirnoff (KS) test, in which we check the null hypothesis

that the sample of transition times extracted from metadynamics and a large sample of times randomly generated according to an exponential probability distribution reflect the same underlying distribution.

As a quantitative measure of the similarity between the empirical and theoretical distributions we use the p -value associated to the KS statistic. The p -value represents the probability that the distribution of times extracted from metadynamics is obtained from the theoretical exponential distribution. In Ref. [9] this analysis is validated and discussed at length.

In Fig. 2, we provide the empirical and fitted cumulative distribution functions for both the cases considered in this work: (a) when the ligand is sterically constrained to roll along the axis of symmetry, (b) when the the ligand is free to move in any direction. The respective p-values respectively for the two cases corresponding to the fitted timescales reported in the text, are 0.72 and 0.64, well above the threshold of 0.05[9, 11].

Bibliography

- [1] Mondal J, Morrone JA, Berne BJ (2013) How hydrophobic drying forces impact the kinetics of molecular recognition. *Proc. Natl. Acad. Sci.* 110(33):13277–13282.
- [2] Jorgensen WL, Maxwell DS, Tirado-Rives J (1996) Development and testing of the opls all-atom force field on conformational energetics and properties of organic liquids. *J. Amer. Chem. Soc.* 118(45):11225–11236.
- [3] Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79(2):926–935.
- [4] Hess B, Kutzner C, Van Der Spoel D, Lindahl E (2008) Gromacs 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theor. Comp.* 4(3):435–447.
- [5] Bonomi M et al. (2009) Plumed: A portable plugin for free-energy calculations with molecular dynamics. *Comp. Phys. Comm.* 180(10):1961–1972.
- [6] Bussi G, Donadio D, Parrinello M (2007) Canonical sampling through velocity rescaling. *J. Chem. Phys.* 126(1):014101–014107.
- [7] Berendsen HJ, Postma JPM, van Gunsteren WF, DiNola A, Haak J (1984) Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* 81(8):3684–3690.
- [8] Barducci A, Bussi G, Parrinello M (2008) Well-tempered metadynamics: A smoothly converging and tunable free-energy method. *Phys. Rev. Lett.* 100(2):020603–020606.
- [9] Salvalaglio M, Tiwary P, Parrinello M (2014) Assessing the reliability of the dynamics reconstructed from metadynamics. *J. Chem. Theor. Comp.* 10(4):1420–1425.
- [10] Tiwary P, Parrinello M (2013) From metadynamics to dynamics. *Phys. Rev. Lett.* 111(23):230602–230606.
- [11] Tiwary P, Limongelli V, Salvalaglio M, Parrinello M (2015) Kinetics of protein–ligand unbinding: Predicting pathways, rates, and rate-limiting steps. *Proc. Natl. Acad. Sci.* 112(5):E386–E391.