

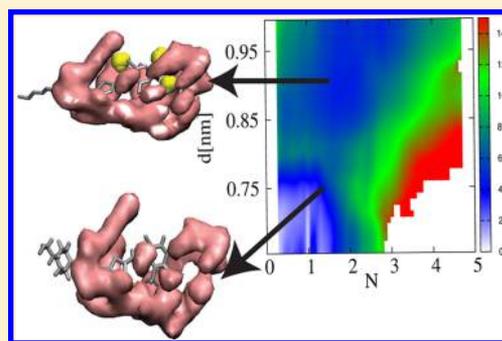
# Role of Desolvation in Thermodynamics and Kinetics of Ligand Binding to a Kinase

Jagannath Mondal, Richard A. Friesner, and B. J. Berne\*

Department of Chemistry, Columbia University, 3000 Broadway, New York, New York 10027, United States

## S Supporting Information

**ABSTRACT:** Computer simulations are used to determine the free energy landscape for the binding of the anticancer drug Dasatinib to its src kinase receptor and show that before settling into a free energy basin the ligand must surmount a free energy barrier. An analysis based on using both the ligand-pocket separation and the pocket-water occupancy as reaction coordinates shows that the free energy barrier is a result of the free energy cost for almost complete desolvation of the binding pocket. The simulations further show that the barrier is not a result of the reorganization free energy of the binding pocket. Although a continuum solvent model gives the location of free energy minima, it is not able to reproduce the intermediate free energy barrier. Finally, it is shown that a kinetic model for the on rate constant in which the ligand diffuses up to a doorway state and then surmounts the desolvation free energy barrier is consistent with published microsecond time-scale simulations of the ligand binding kinetics for this system [Shaw, D. E. et al. *J. Am. Chem. Soc.* **2011**, *133*, 9181–9183].



## 1. INTRODUCTION

Computational drug discovery efforts are generally focused on calculation of equilibrium protein–ligand binding affinities. There is no question that the equilibrium binding affinity is an important predictor of ligand efficacy, and achieving adequate values of this quantity is a necessary (although not sufficient) condition for a viable drug candidate. The calculation of equilibrium binding free energy, while very challenging, can be pursued via both fast approximate methods (empirical scoring function, MM-GBSA,<sup>1</sup> etc.) as well as more expensive, but rigorous, statistically mechanical simulation approaches such as free energy perturbation theory (FEP).<sup>2–4</sup> Recent work in these areas suggests that significant progress is possible, provided sufficient computation time is available and accurate force field representations of the ligand and protein are employed.

However, there is also an increasing belief that, at least for some ligands and pharmaceutical targets, the kinetics of binding can play a significant role in *in vivo* efficacy. It is generally suggested<sup>5</sup> that existence of intermediate free energy barriers can influence the so-called on-rate and off-rate of ligand-binding, which in turn will influence the potency of the drug, by tuning the residence time of the ligand. Hence, the design of drug candidates with increased activation barrier to dissociation is of potential interest as a strategy for finding superior compounds in a drug discovery project.

Computational methods in principle can facilitate the design of ligands with intermediate activation barriers (which must be introduced along with the manifold of other properties required for suitability as a drug). However, the calculation of activation barriers in a complex condensed phase system of this type is even a more formidable task at present than determination of

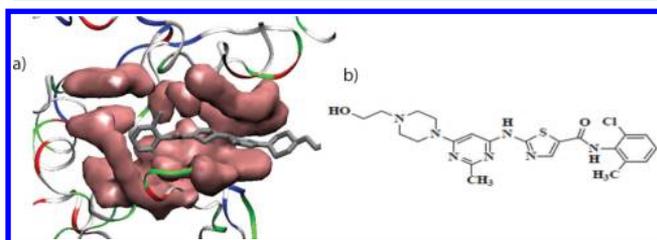
equilibrium binding affinities. Molecular dynamics simulations are one possible way of determining binding kinetics. Many factors contribute to the activation barriers including, for some systems, the movement of water molecules into or out of the binding pocket (desolvation and resolvation). Several studies have evaluated the free energy changes accompanying such processes,<sup>6–9</sup> but there are relatively few simulations focusing on the kinetic barriers arising from desolvation. For example, some recent studies focus on binding kinetics in systems where there is a large scale drying transition before binding.<sup>10,11</sup> Shaw and co-workers have used ANTON, a specialized massively parallel molecular dynamics engine, to generate microsecond to millisecond trajectories of ligands diffusing toward and binding to a protein. Using this approach they have been able to provide a microscopically detailed view of the real-time binding pathways of several beta-blockers to a G-protein coupled receptor<sup>12</sup> as well as the FDA-approved anticancer drugs PP1 and Dasatinib to a kinase.<sup>13</sup> In both of these studies a shell of water molecules is ejected just before the ligand binds to the native pocket. Such calculations provide important insights into the kinetics of the binding process, but they are probably too expensive at present to be employed on a routine basis in evaluating libraries of candidate ligands.

In order to obtain physical insight and useful semi-quantitative information concerning the barriers to ligand binding and unbinding, in this paper we analyze how the free-energy changes as the ligand approaches its binding pose. We explore the binding of a FDA-approved kinase-inhibitor

Received: July 8, 2014

Published: November 13, 2014

Dasatinib<sup>14–16</sup> to src-kinase,<sup>17</sup> the same kinase and ligand studied by Shaw and co-workers. A representative snapshot of Dasatinib in its native binding pose in the pocket of kinase is shown in Figure 1 a. In the present work, we take a different



**Figure 1.** a) Representative snapshot of Dasatinib (silver color) in its native binding pocket (purple color) of kinase. The density map of the native binding pocket has been obtained by averaging over frames of trajectory. Only part of the proteins has been shown (in ribbon representation), and water and ions are not shown for clarity. b) The chemical structure of the ligand Dasatinib used in the work (identical to that provided in ref 16).

approach. In particular we explore the role played by explicit water molecules in the binding process with a focus on the free energy of pocket desolvation. Using molecular dynamics and free energy simulation methods we show that the ligand encounters a free-energy barrier just prior to reaching its binding pose. By treating *both* the water-occupancy and the ligand-pocket separation distance as reaction coordinates, we quantitatively trace the origin of this barrier to the displacement of pocket-waters by the ligand en route to its binding. In addition, application of the WaterMap technique<sup>6</sup> helps to clarify our findings by characterizing the thermodynamic nature of pocket-waters at various ligand-pocket distances. As expected, application of a continuum solvent model (GBSA) fails to find the free energy barrier as implicit solvent models ignore explicit solvation. Finally, it is shown that a kinetic model for the on rate constant, in which the ligand diffuses up to a doorway state and then surmounts the desolvation free energy barrier, is consistent with the Shaw group's very long simulations of the ligand binding kinetics for this system.

In what follows, we shall focus our attention on calculation of the barrier to protein–ligand association (enabling estimation of the on-rate for ligand binding), as opposed to the reverse process of ligand dissociation. There are several reasons that we make this choice. First, the dissociation barrier can be calculated via the detailed balance relation if the association-barrier and the equilibrium binding constant is known. Calculation of the equilibrium binding constant can in principle be accomplished using free energy perturbation (FEP) methods. Recent work<sup>18</sup> has demonstrated that very reasonable RMS errors (on the order of 1.2 kcal/mol) can be obtained for relative protein–ligand binding affinities in congeneric series, for multiple data sets comprising 200 compounds in all. While additional validation will be required to fully establish the accuracy and robustness of FEP methods (including that of ref 18), these promising results suggest that if a reliable approach to the calculation of the association barrier can be established, it will be possible to compute the barrier to dissociation, exploiting the detailed balance relation, within reasonable error bars. The FEP methods we cite above involve the calculation of relative (compared to a reference molecule, the binding affinity of which is known experimentally), as opposed absolute, free energies of binding. This imposes some

important limitations on the proposed strategy to address dissociation. First, a suitable reference molecule with experimental binding affinity data is required. Second, because it addresses relative rather than absolute effects, the FEP calculation is limited with regard to the physical insight it can provide. For example, as discussed in detail by Pearlstein and co-workers,<sup>19–21</sup> the process of desolvation during ligand dissociation is critical to the magnitude of the dissociation barrier. However, the FEP calculation addresses only relative desolvation effects, a complete picture as to how the filling of the cavity with waters as the ligand is evacuated may not be available. The problem of extracting the information that is available with regard to the dissociation barrier from the FEP calculations will not be discussed further in this paper; we will instead focus entirely on the barrier to association. However, this issue will need to be carefully investigated in future work. Second, the final steps in the binding process, in which the ligand and protein fully adjust to each other to form the equilibrium protein–ligand complex, represent a more challenging computational problem than the early steps and would require a substantially greater amount of computation time, and perhaps an expanded definition of the reaction coordinate, in order to accurately map out the potential of mean force which undergoes a very rapid change as the protein and ligand adjust to full complementarity. In fact, we can estimate the computational effort as comparable to that required for the computation of absolute protein–ligand binding free energies. In contrast, umbrella sampling or FEP methods computing relative binding affinities are much less expensive, as long as there is an acceptable reference molecule with experimentally known binding affinity; this is always the case in a drug discovery project in its lead optimization phase. Finally, direct simulation using explicit time dependent molecular dynamics is only feasible for the association rate-constant; the barrier for the dissociation is too large for optimized ligand binding to enable access to this barrier with current computational hardware. Comparison between the approximate methods that we investigate herein (WaterMap based estimation of the barrier and FEP computation of the potential of mean force) and direct simulation is highly desirable in order to calibrate the validity of the approximations, exclusive of the details of the force field.

## 2. SIMULATION MODEL AND METHODS

**2.1. Simulation Model.** The initial structure of Dasatinib bound to the protein was obtained from protein data bank (PDB Id: 3G5D).<sup>16</sup> All the non-protein molecules (except Dasatinib) were then removed, and protein preparation wizard<sup>22</sup> within the Schrodinger software package<sup>23</sup> was used to post process the structures of proteins and Dasatinib for simulation. Protonation states of the amino acid residues were assigned assuming the systems are at pH 7.0. Any missing hydrogens and any missing atoms of side chains were added, and finally any missing loops in the pdb structures were filled in. The chemical formula of Dasatinib as provided in the RCSB protein data bank is  $C_{22}H_{26}ClN_7O_2S$ . Based on the chemical structure of Dasatinib illustrated in the original article by Getlik et al.<sup>16</sup> which also provided the pdb coordinates (3G5D), Dasatinib is rendered as a charge-neutral ligand (i.e., there is no proton bound to the nitrogen atoms of piperazine rings). To be consistent with the same chemical formula and chemical structure, in our simulation model, Dasatinib is also treated as charge-neutral molecule. The chemical structure of the

Dasatinib as used in the current work, is also shown in Figure 1b).

The OPLSAA all-atom force-field<sup>24</sup> was used to model Dasatinib bound to the protein, and the TIP4P model<sup>25</sup> was used for water. OPLS force field parameters for Dasatinib were obtained from the Schrodinger package by selecting parameters for the analogous atoms of the ligand Dasatinib. All force-field parameters of Dasatinib have been provided near the end of the Supporting Information of this paper (gromacs-compatible parameter files are available upon request). The postprocessed pdb structure of Dasatinib bound to protein was then solvated with TIP4P water. We defined any protein-atom which is within 0.5 nm of the center of mass of the ligand in its native binding pose as belonging to the 'binding pocket'. Counterions were added to ensure electrical neutrality. To be consistent with the work of Shaw and co-workers,<sup>13</sup> sodium chloride (NaCl) salt was added to a concentration of 150 mM to maintain the medium at physiological conditions. Then the entire solvated system was run in the NPT ensemble at a constant pressure of 1 bar and a constant temperature of 300 K, keeping the position of ligand restrained in its native binding pose. Nose Hoover Thermostat<sup>26,27</sup> and Parinello-Rahman barostat<sup>28</sup> were employed to maintain temperature and pressure respectively to their desired average value. The dimension of the equilibrated box was 7 nm × 7 nm × 8 nm. Periodic boundary conditions were applied in all three directions, and the Lennard-Jones (LJ) interactions were smoothly switched off at 1.0 nm. Particle-mesh ewald (PME)<sup>29,30</sup> summation was used for the long-range electrostatic interactions. The PME parameters used a real space cutoff distance of 1.4 nm and an interpolation order of 6, with a maximum fast Fourier transform grid spacing of 0.12 nm. The SETTLE<sup>31</sup> algorithm was used to keep the water molecules rigid. The LINCS<sup>32</sup> algorithm was used to keep all the bonds with hydrogen rigid. The system was evolved with a time step of 2 fs. All atomistic simulations were carried using the software package GROMACS4.5.4.<sup>33</sup>

**2.2. Free Energy Simulation.** One of our aims is to compute the free energy landscape for the approach of the ligand to its native binding pose in the protein. To accomplish this we employ the umbrella sampling technique as implemented in GROMACS4.5.4. The scalar distance  $d$  between the center of masses (COM) of the binding pocket and the ligand is treated as a preliminary reaction coordinate for the purpose of calculating the free energy profile. In order to generate reasonable initial configurations for each individual umbrella-sampling windows, we have employed Steered Molecular Dynamics (SMD) simulations. In this protocol, we start with the configuration of ligand in its native pose in the binding pocket. Then we slowly pull the ligand away from the pocket along  $d$  by attaching a spring of force constant 1000 kJ/mol/nm<sup>2</sup>. We employ a pulling speed of 0.0005 nm/ps which is slow enough to ensure that the ligand spends significant time at different  $d$ . Then from the saved SMD trajectories, the starting configurations for each of the umbrella-sampling windows were chosen based on the respective pocket-ligand distance  $d$  which ranges from 0.1 nm (ligand closest to binding pocket) to 1.5 nm (ligand just out of pocket) at a spacing of 0.05 nm (produced using a higher umbrella-potential force constant (4500 kJ/mol/nm<sup>2</sup>)). Each of the configurations corresponding to individual umbrella-sampling windows was run for 10 ns in the NVT ensemble at 300 K. The average temperature was maintained at the desired value of 300 K by using a Nose

Hoover thermostat.<sup>26,27</sup> Our principal aim is to determine the free energy barriers that might be encountered by the ligand as it moves into the binding pocket, and not to compute the absolute free energy of binding. The latter calculation would require a significantly more extensive exploration of phase space, including a comparison to the unbound state, which is not available from our current simulation protocol. The restricted nature of the present calculation is essential in reducing the required computation time to manageable proportions.

The extent to which pocket reorganization contributes to the free energy surface as the ligand approaches the pocket was investigated by imposing position-restraining harmonic potentials with force constant equal to 1000 kJ/mol applied to each of the X, Y, and Z coordinates of each of the atoms constituting the binding pocket as implemented within Gromacs4.5.4. In the presence of these position restraints, all umbrella sampling simulations were repeated using the same protocol as described previously for an unrestrained pocket.

Continuum solvent models are unable to treat desolvation. Thus, it is of interest to see if such models predict the intermediate free energy barrier observed in the full MD simulations with explicit water. We performed umbrella sampling simulations using the Generalized Born implicit solvent model,<sup>34,35</sup> as implemented within Gromacs4.5.4. The OBC<sup>36</sup> algorithm was used for computing the Born Radii and a simple ACE type approximation due to Schaefer et al.<sup>37</sup> was used for computing the nonpolar part of the solvation free energy. The Wham<sup>38</sup> algorithm (<http://membrane.urmc.rochester.edu/content/wham>) was used to reconstruct the free energy landscape.

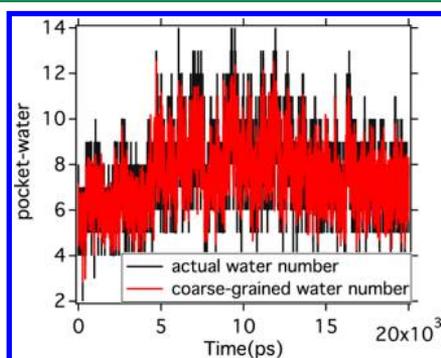
**2.3. Analysis of Pocket-Water Profile.** One of the main focuses of the current work is to determine to what degree desolvation of water molecules in the binding pocket contributes to the free energy barriers encountered by the ligand in approaching its binding pose in the pocket. To accomplish this, we must first decide on a consistent definition of *pocket-water* to be used for all the umbrella-sampling windows. For this purpose, we define any water molecule whose oxygen atom is within 0.5 nm of center of mass of the binding pocket to be a pocket-water molecule (based on the radial distribution function). We then use umbrella sampling techniques to compute the free energy change for bringing the COM of the ligand from bulk solution to a separation  $d$  from the COM of the pocket using umbrella sampling.

Umbrella sampling with  $d$  constrained is straightforward, but constraining the number of pocket-water requires some ingenuity as described in the literature where the number of water molecules  $N$  in the pocket is approximated by a continuous function.<sup>39–41</sup> Based on an analysis of the pair-correlation function of water molecules with respect to the position of the center-of-mass of the pocket, we count all water molecules within an effective cutoff distance  $r_0 = 0.5$  nm of the center of mass of the pocket as "pocket water molecules". We then take the coarse-grained water occupancy number  $N$  in the pocket to be

$$N = \begin{cases} \sum_w \left[ \frac{1 - (r/r_0)^6}{1 - (r/r_0)^{12}} \right] = \sum_w \frac{1}{1 + (r/r_0)^6}, & \text{if } r \leq r_0 \\ 0, & \text{otherwise} \end{cases}$$

where  $r$  is the distance of the oxygen atom of each water molecule from center of mass position,  $r_{com}$  of the binding

pocket, that is  $r = |\mathbf{r}_W - \mathbf{r}_{com}|$ , and the summation runs over all water molecules in the simulation box. As shown in Figure 2,



**Figure 2.** Comparison performance of the coarse-grained representation of water-number to reproduce time-profile of actual number of water.

this coarse-grained representation of the water occupation number reproduces the time profile of ‘atomistic’ water number reasonably well and makes it possible to apply restraints to bias the pocket-water occupancy in an umbrella sampling simulation of the water occupancy number. With this definition, the pocket water occupancy is a continuous variable that adopts noninteger values, and its corresponding probability distribution  $P(N)$  (and free energy  $G(N) = -k_B T \ln P(N)$ ) becomes a continuous function.

We determine  $G_+(N;d)$ , the free energy corresponding to  $N$  water molecules being in the pocket when the COM of the ligand is at a separation  $d$  from the COM of the pocket, by performing umbrella sampling simulations where both the ligand-pocket distance  $d$  and the number of pocket-water  $N$  are constrained. Umbrella sampling with constraints on the two continuous reaction coordinates  $d$  and  $N$  allows us to determine  $G_+(N;d)$ . The special case of  $G_+(N;d = \infty)$ , the free energy as a function of  $N$  in the *ligand-free* pocket, requires umbrella sampling only along the single dimension  $N$ . As Hummer has shown,<sup>41</sup> this is accomplished using harmonic restraints  $U = 0.5K_w(N-n_r)^2$  in which  $n_r$  is the reference water number. We first carry out multiple independent equilibrium simulations to sample the water-number in a ligand-free pocket. Then using the snapshots in the trajectory as initial coordinates, we performed umbrella sampling simulations in each of 16 windows in the range of 0 to 8 at a separation of 0.5 with a spring constant  $K_w = 2.5$  kcal/mol. Each simulation was run for 2 ns. and the last 1.5 ns of each trajectory was used for computing the free energy in a ligand-free pocket as a function of  $N$  using WHAM.

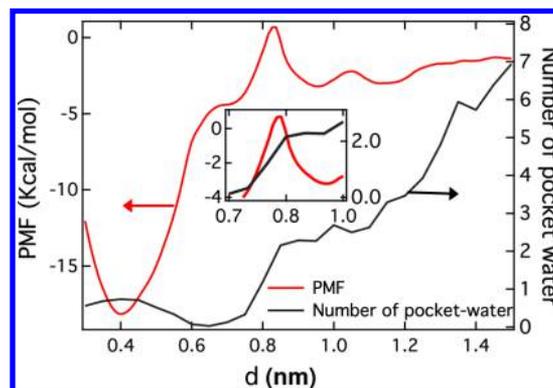
The computation of the free energy  $G_+(N;d)$  for relevant values of  $d$  requires umbrella sampling along both  $d$  and  $N$  resulting in a two-dimensional potential of mean force. To minimize the computational expense of such simulations, we restricted the calculations for  $d$  to three windows at a separation of 0.10 nm in the range  $d = 0.75$  to 0.95 nm, the range of ligand-pocket separations corresponding to the intermediate free energy barrier, and we restricted the calculations for  $N$  to 20 windows of width 0.250 in the range from  $N = 0$  to 4.75, observed to be the range associated with this range of  $d$ . We adopt harmonic restraints  $U = 0.5K_w(N-n_r)^2 + 0.5K_d(d-d_r)^2$  with  $K_d = 478$  kcal/mol/nm<sup>2</sup> and  $K_w = 2.5$  kcal/mol. The values of  $K_d$  and  $K_w$  were chosen such that the distributions of the corresponding reaction coordinates around the desired  $d$  and  $w$

are Gaussian in nature and there is significant overlap among adjacent windows. Thus, in the two-dimensional reaction space there are a total of  $3 \times 20 = 60$  windows. We use the snapshots from the previous umbrella sampling simulation along  $d$  as the initial configurations. Each of 60 windows was run for 2 ns, and the last 1.5 ns of each of trajectory were used for computing the two-dimensional free energy as a function of  $N$  and  $d$  using WHAM. The latest version of PLUMED software<sup>42</sup> was used for this part of the analysis.

**2.4. WaterMap Analysis.** The WaterMap technique<sup>7,9</sup> allows us to thermodynamically characterize the principal hydration sites in the pocket as a function of the ligand-pocket distance,  $d$ . Based on the computation of relative free energies and inhomogeneous solvation theory,<sup>43,44</sup> the WaterMap technique identifies water-clusters of relatively high free energy which when displaced by the ligand significantly contribute to a favorable ligand-binding affinity. We perform independent WaterMap calculations using the final protein configuration from each umbrella-sampling window (equilibrated for 10 ns) when the ligand distance is restrained at  $d = 0.95$  nm (just before the ligand overcomes the barrier and 0.75 nm (just after the ligand overcomes the barrier). The high occupancy hydration sites (of radii 1 nm) in the binding pocket were identified, the waters were clustered, and their associated enthalpies and entropies of displacement are calculated using inhomogeneous solvation theory from MD runs (of 2 ns duration) in the NPT ensemble, with the heavy atoms of the protein harmonically restrained at their respective positions as determined from umbrella simulations in  $d$ -windows.

### 3. RESULTS AND DISCUSSION

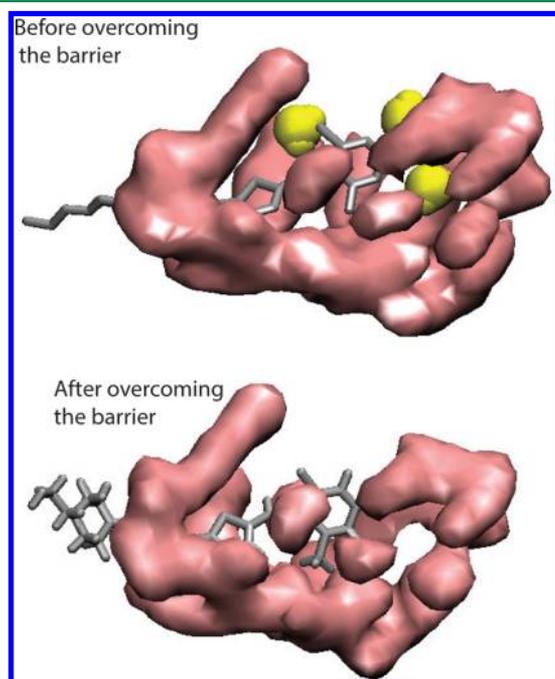
**3.1. Interplay of Free Energy Profile and Pocket-Water.** The free energy profile, or PMF, of the binding complex as a function of the distance,  $d$ , is shown in Figure 3



**Figure 3.** Free energy profile (PMF) of ligand-approach (red curve, left scale) and profile of number of pocket-water (black curve, right scale) as a function of the distance  $d$  between the pocket and ligand centers of mass. Inset: The free energy profile is zoomed in between  $d$  of 0.75 and 0.95 nm to show the free energy barrier due to dewetting.

(left-hand scale). The distance  $d$  is a simple reaction coordinate containing no orientational bias that provides some insight into the binding process yet, as is well-known, the choice of reaction coordinate is critical and the results will be sensitive to its choice. As we shall see, it will be necessary to augment this choice at the very least with the solvent occupation number. Nevertheless, as shown below, the choice of  $d$  generates some interesting observations. We find for example that as the ligand

approaches the pocket from  $d = 1.5$  nm to  $d = 0.95$  nm, the free energy profile remains almost flat. However, as shown in the



**Figure 4.** Representative snapshots of pocket-water at ligand-pocket distances corresponding to  $d = 0.95$  and  $0.75$  nm. The average density map of the binding pocket is shown in pink, the waters are shown in blue color with space-filling representation, and the ligand is shown in silver color with licorice representation.

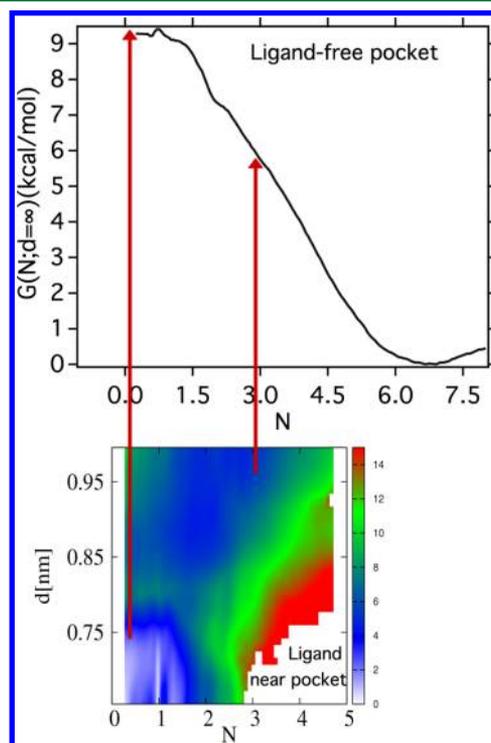
inset of Figure 3, as the ligand passes from  $d = 0.95$  nm to  $d = 0.75$  nm, it has to surmount a free energy barrier of more than 4.0 kcal/mol at an intermediate ligand-pocket distance. We also observe a weak shoulder at  $d = 0.65$  nm, and we find this is mainly steric in origin resulting from the pocket-ligand interaction. At closer separations  $d < 0.75$  nm, a steep decrease in free energy is observed before the ligand settles into a deep global free-energy minimum at  $d = 0.4$  nm. At short distances,  $d < 0.4$  nm, the free energy increases, mainly due to steric repulsions between pocket-wall and ligand. We have checked the reliability of the PMF landscape and the corresponding pocket-water profile, by repeating the same umbrella sampling simulation for a different set of initial configurations. As depicted in Figures S1 and S2 of the Supporting Information, the two different sets of umbrella sampling simulation results are in reasonable mutual agreement. The approximate agreement of two simulations does not of course provide a rigorous estimation of the error, but it does suggest that there are not any immediately obvious major sampling problems. We have also provided the convergence plot of number of pocket-water profile along with the measure of standard deviation from average profile in Figure S6 of the Supporting Information.

The contribution of desolvation to large binding affinities in other complexes has been explored before.<sup>6,41,45</sup> We investigate the desolvation of the binding pocket on ligand approach and determine its contribution to the free energy. Figure 3 (right-hand scale) shows that the number of water molecules present in the kinase binding pocket as a function of  $d$  first decreases from 6, when the ligand is out of the pocket, to 3, when it is at the barrier, to essentially zero, right after it passes the barrier, and then increases slightly as some water molecules reenter the

pocket when the ligand takes up its binding pose. Figure 4 shows representative snapshots (which are also similar to final snapshots of each window) of the distribution of water before and after the ligand overcomes the barrier.

However, at smaller ligand-pocket distances, the number of pocket-water molecules increases slightly until a relatively sharp rehydration takes place at distances close to the position of the free-energy minimum. These water molecules then occupy positions no longer occupied by the invading ligand. Interestingly, Figure 3 (right-hand scale) shows that desolvation gives rise to an intermediate free-energy barrier. In the next subsection, we will analyze the implications of desolvation in the free energy profile.

**3.2. Role of Pocket Dewetting in the Free Energy Profile.** We determine the contribution of desolvation to the free energy barrier that the ligand must overcome in the process of binding to the pocket using the umbrella sampling strategy introduced by Hummer et al.<sup>41</sup> The Gibbs free energy  $G(N; d = \infty)$  shown in the top panel of Figure 5 exhibits a strong monotonic decrease with  $N$  between  $N = 0$  and 8 water molecules. The minimum of  $G(N; d = \infty)$  gives the average number of pocket-water molecules,  $\bar{N}$  in the ligand-free pocket.



**Figure 5.** Top: Free energy  $G(N; d = \infty)$  in kcal/mol as a function of number of pocket water ( $N$ ) (which is effectively similar to the case of a ligand-free pocket). (The free energy  $G(N; d = \infty)$  was obtained from umbrella sampling simulation using water-number as the reaction coordinate as described in the Method section.) Bottom: The dependence of the free energy  $G(N; d)$  on the number of pocket-waters,  $N$ , and on the ligand-pocket separation  $d$  in the range of intermediate free energy barrier. The blue basins correspond to the free energy minima appearing on either side of the intermediate free-energy barrier. The average numbers of water molecules,  $\bar{N}$ , at  $d = 0.75$  nm (after overcoming the barrier) is  $\approx 0$  and at  $\bar{N}$  at  $d = 0.95$  nm (before overcoming the barrier) is  $\approx 3$ . The two red lines provide the changes in desolvation free energies for protein–ligand separations relative to the case when the ligand is outside the pocket just prior to and just after surmounting the intermediate free energy barrier.

From this we can compute  $\Delta G$  for the transition from  $\bar{N}$  to  $N$ , corresponding to partial desolvation of the *ligand-free pocket*. The bottom plate of Figure 5 shows the free energy profile  $G_+(N;d)$  as a function of  $N$  and  $d$ , in the neighborhood of the intermediate free energy barrier. We clearly see the presence of two basins before and after the intermediate free energy maximum. At  $d = 0.95$  nm (just before overcoming the barrier) the average water occupation number is  $\bar{N} \approx 3$ , and at  $d = 0.75$  nm (just after overcoming the barrier) it is  $\bar{N} \approx 0$ . Thus, for a given ligand-pocket distance, we can read from the graph in the bottom plate the average number of water molecules in the pocket with the ligand present, and then from the top plate the corresponding value of  $\Delta G$  for partial desolvation of the ligand-free pocket to that same number of water molecules. This allows us to estimate the desolvation free energy required for the ligand to replace the water molecules in a ligand-free pocket upon binding. From Figure 3, we estimate the desolvation free energy change accompanying the change in ligand-pocket distance from  $d = 0.95 \rightarrow d = 0.75$  nm (which are the position before and after the intermediate free-energy barrier respectively) to be  $\approx 3.7$  kcal/mol, a result almost equal to the total free energy cost (4 kcal/mol) for the ligand to surmount the barrier on going from  $d = 0.95 \rightarrow d = 0.75$  nm. This suggests that the intermediate free energy barrier encountered by the ligand is mainly due to pocket-desolvation, and that pocket-desolvation is thus a key step in the kinetics of ligand binding. Both these one-dimensional and two-dimensional free energetics of pocket-water are in reasonable agreement with an independently performed pocket-water free energy analysis, using different initial configuration (Figures S2 and S3 in the Supporting Information). We have also found that extending the simulation length of each window from 2 to 5 ns does not change the result significantly.

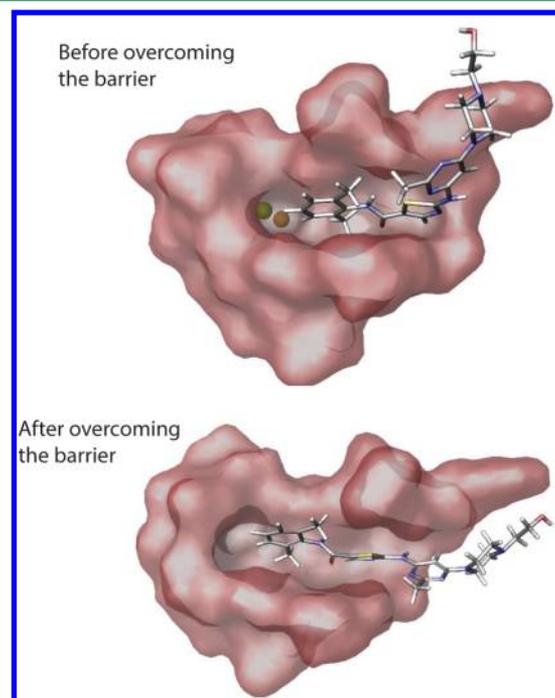
The above-mentioned three pocket-waters corresponding to ligand-pocket separation of  $d = 0.95$  nm are mainly stabilized in the pocket by hydrogen bonding interactions. A hydrogen bonding analysis involving these three pocket waters (using a distance cutoff of 0.3 nm and angle cutoff of 20 degrees) reveals that the contribution comes in two forms: water–water hydrogen bonding interaction among these three closely seated waters and hydrogen bonding between water and certain polar amino residues constituting the binding pocket. A study of the hydrogen bonding occupancy over the simulation trajectory corresponding to the  $d = 0.95$  nm identifies three amino acid residues which form considerably stable hydrogen bonds with these water molecules through their polar side chain groups located at the proximity of the water molecules. As illustrated in Figure S7 in the Supporting Information, these three amino acid residues participating in protein–water hydrogen bonds are LYS36 through its side chain amino group (acting as hydrogen bond donor), GLU46 through its side chain carboxylate group (acting as hydrogen bond donor), and THR74 through its side chain hydroxyl group (acting as hydrogen bond acceptor). Even though there are other transient hydrogen bonds between pocket amino residues and three waters, analysis of hydrogen bonding occupancies shows that these three particular hydrogen bonds between pocket and water make most significant contributions in stabilizing those water molecules.

**3.3. WaterMap Analysis of Pocket-Waters and Its Relation to Free Energy Barriers.** The use of WaterMap to understand ligand association and dissociation barriers has been pioneered in a series of papers by Pearlstein et al.<sup>19–21</sup> over the

past several years. These papers correlate water displacement upon ligand binding (or resolution upon unbinding) with the on and off rates of ligand binding, respectively, in a number of different types of receptor systems. In the present paper, we build on ideas in this work but utilize descriptors based on water displacement free energies, calculated from WaterMap, that are different in their details from those proposed in those references. Specifically, we evaluate water displacement as a function of a proposed reaction coordinate and identify the association barrier as the free energy cost of water displacement at an approximate transition state, as opposed to utilizing water displacement enthalpies from the bound ligand pose, as is done in the work by Pearlstein et al. As was discussed above, we do not investigate the relationship of water resolution to the off-rate, as our approach to binding kinetics exploits calculation of the off-rate from the on-rate and the equilibrium binding constant.

We explore the possibility of using the recently developed WaterMap method<sup>6</sup> to determine the position and approximate magnitude of the free energy barrier, by analyzing the hydration of the pocket at two different pocket-ligand distances  $d$ , right before ( $d = 0.95$  nm) and right after ( $d = 0.75$  nm) the ligand overcomes the barrier. If successful, this method would provide a relatively inexpensive way to predict the position and magnitude of the barrier arising from desolvation. The WaterMap method predicts regions of high-water density (referred to as 'hydration sites') and rank-orders them based on the relative free-energies of the water-clusters. The two WaterMap snapshots depicted in Figure 6 represent the principal hydration sites before and after the ligand overcomes the free energy barrier, respectively.

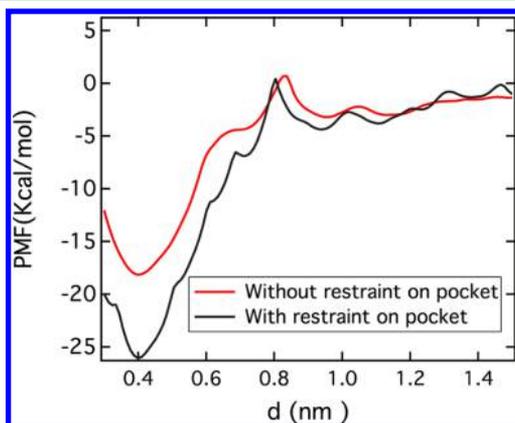
Based on this WaterMap analysis, we find that before the ligand passes over the intermediate barrier from the right, the



**Figure 6.** Snapshot from WaterMap analysis showing principal hydration sites near the binding pocket before overcoming the free energy barrier (corresponding to  $d = 0.95$  nm) and after overcoming the barrier (corresponding to  $d = 0.75$  nm).

water hydration sites have a total free energy of 5.5 kcal/mol, but immediately after passing over the barrier no hydration sites are observed in the binding pocket (at  $d = 0.75$  nm), independently confirming the complete desolvation observed in the umbrella sampling simulation. The water map analysis suggests there is a desolvation free energy cost of 5.5 kcal/mol, in fairly good agreement with the 4 kcal/mol estimate obtained from free energy simulation. WaterMap thus seems to provide an inexpensive tool for semiquantitatively predicting the existence of free energy barriers encountered by the ligand as it approaches or leaves the binding pocket.

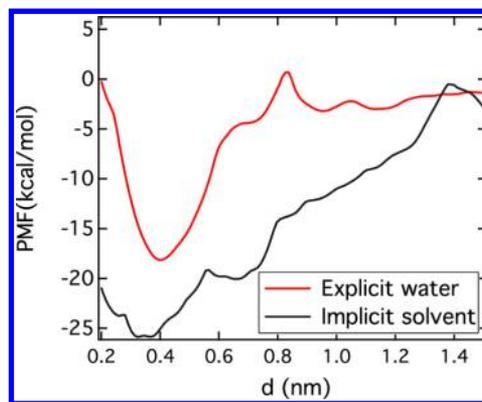
**3.4. Does Pocket-Reorganization Contribute to the PMF?** One might wonder how much the reorganization of the pocket contributes to the intermediate barrier seen in the PMF. To investigate this, we have repeated the computation of the PMF using position restraints on each of the atoms in the binding pocket in order to prevent the pocket from reorganizing. In Figure 7, a similar intermediate free energy



**Figure 7.** Comparison of free energy profile of ligand approach to the binding pocket of kinase in the absence of restraint on the pocket (original result) with that in the presence of restraint on the pocket.

barrier appears even when reorganization of the pocket is restrained suggesting that this intermediate barrier does not result from pocket reorganization. Interestingly, the free energy of the binding pose is lower for restrained pocket than for the unrestrained pocket probably because there is less steric hindrance to the ligand. In addition we also have found that the radius of gyration of the binding pocket increases at most by a few angstroms when the ligand is moved from bulk solution to the position of the intermediate free energy barrier, a change that is not significant. Therefore, we believe that the major contribution to the intermediate free energy barrier comes from the free energy of desolvation and not from reorganization of the pocket.

**3.5. A Continuum Solvent Model Does Not Give the Intermediate Free Energy Barrier.** We suspect that continuum solvation models will not be able to account for the intermediate free energy barrier if this barrier is due, as we claim, to desolvation, because such models cannot account for desolvation accurately when an atomic level description of individual water molecules, sensitive to the size of the particles, is important. Many popular existing drug-binding scoring functions like MM/GBSA or MM/PBSA are based on implicit solvent models. We repeated the computation of the PMF using a popular continuum solvent model GBSA.<sup>34,35</sup> Figure 8 compares the PMFs for GBSA with that of explicit water. We

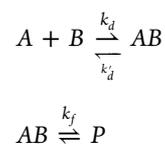


**Figure 8.** Comparison of free energy profile of ligand approach to the binding pocket of kinase in implicit solvent with that in explicit water.

find that although the position of the well corresponding to the binding pose is essentially the same for GBSA and explicit water, GBSA does not predict the intermediate free energy barrier, as we suspected, and also overstabilizes the overall ligand binding. The result is reminiscent of a previous study by Mondal et al.<sup>46</sup> where the implicit solvent description of free energy of association of a pair of beta-peptides was found to be overstabilized compared to that of explicit solvent simulation, even though there was agreement in the location of free-energy minima. In an earlier study, Berne and co-workers<sup>47</sup> also found that the continuum solvent description of a free energy landscape of the hairpin part of G-protein overweights many non-native configurations. The free energy profile in implicit solvent has been verified against independent runs (Figure S5 of the Supporting Information).

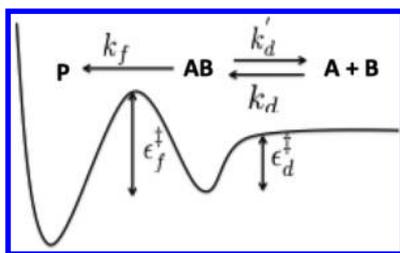
**3.6. Estimates of the Second-Order Rate Constant for Ligand Binding.** The kinetics of ligand binding involves the diffusion of the ligand up to a doorway configuration followed by transitions over one or more free energy barriers. In the foregoing we showed that the final barrier encountered before the ligand binds is to a large extent due to desolvation of the pocket. We were able to estimate the activation free energy for this transition assuming that the distance between the center-of-mass of the pocket and the center-of-mass of the ligand is a good reaction coordinate. We also showed that this distance is strongly correlated with the number of pocket waters so that strictly speaking the number of pocket water molecules should be included in the reaction coordinate. It is of interest to see if we can make a reasonable estimate of the second-order rate constant for binding based on diffusion up to and activation over this desolvation barrier. What follows is a very rough analysis of the kinetics to provide guidance for future theoretical models.

Consider the following simple kinetic scheme for this binding process



and its associated schematic free energy diagram presented in Figure 9.

Here  $A$  is the ligand,  $B$  is the protein,  $P$  is the protein–ligand complex in its binding pose,  $k_d \approx 4\pi\kappa(D_A + D_B)\sigma$  is the diffusion controlled rate constant for the ligand to arrive at the



**Figure 9.** Schematic of potential energy surface corresponding to this kinetic scheme.  $k_f$  and  $e_f^\ddagger$  are the rate constant and corresponding activation barrier for the transition from the doorway state to the final binding pose.  $k'_d$  and  $e_d^\ddagger$  are the rate constant and activation barrier for the dissociation of the doorway complex into free ligand and protein.  $k_d$  is the diffusion controlled rate constant for the ligand and protein to diffuse to the doorway state.

doorway state located at a separation  $\sigma$  with relative diffusion constant  $(D_A + D_B)$ ,  $\kappa$  is a steric factor taking into account the probability that the ligand and receptor will be correctly aligned,  $k'_d$  is the rate constant for the doorway complex to fall apart, and  $k_f$  is the rate constant for the doorway complex to transition to the final binding pose of the protein–ligand bound state. Treating the doorway complex in the steady-state approximation yields the effective second-order rate constant for binding as

$$k_r = \frac{k_f k_d}{k_f + k'_d} = \frac{(k_f/k'_d)}{1 + (k_f/k'_d)} k_d$$

In this expression the rate constants  $k_b$ ,  $k'_b$ , and  $k_f$  are respectively second-order, first-order, and first-order rate constants.

The rate constant,  $k_b$  for the final binding step would then be described by an activation free energy  $e_b^\ddagger$  and frequency factor  $A_f$  and the rate constant for the dissociation of the doorway complex,  $k'_b$  would be described by an activation free energy  $e'_b^\ddagger$  and frequency factor  $A'_b$ . Because these two rate constants correspond to leaving the same well by two different paths, we assume that  $A'_b \approx A_f$ . This would be true for a simple one-dimensional reaction coordinate in the transition state approximation, and we assume this here. When these approximations are substituted in the preceding equation we find

$$k_r \approx \frac{e^{-\beta(e_f^\ddagger - e'_d^\ddagger)}}{[e^{-\beta(e_f^\ddagger - e'_d^\ddagger)} + 1]} k_d$$

As might have been expected from the free energy diagram, the overall rate constant for binding depends on the difference in activation energies  $(e_f^\ddagger - e'_d^\ddagger)$  and not only on the activation energy  $e_f^\ddagger$  to go from the doorway state to the bound state.

The second-order rate constant for the binding of Dasatinib was estimated by Shaw and co-workers.<sup>13</sup> They simulated an aqueous solution of one kinase molecule and four ligand molecules (at a molar concentration of 0.007 M) and were able to determine the time it took for a ligand to completely bind and from this made the following rough estimate of the second-order rate constant for binding,  $k_2 \approx 2 \times 10^6 \text{ s}^{-1} \text{ M}^{-1}$ . Since this estimate is based on only one successful trajectory, it contains significant uncertainties, and certainly experimental data is needed to validate this estimate. Nevertheless, we shall regard it as a reference for comparison purposes and assume that  $k_r = 2 \times 10^6 \text{ s}^{-1} \text{ M}^{-1}$ . In order to proceed we need to estimate the

diffusion controlled rate constant  $k_d$ . The expression  $4\pi(D_A + D_B)\sigma$  without the steric factor  $\kappa$  is for two spheres freely diffusing until they touch, but this will considerably overestimate the rate constant for the molecular system where the molecules must be correctly aligned before they can react. The result for two spheres will thus provide an approximate upper bound for  $k_r$ . To determine this upper bound we compute the mean square displacement of Dasatinib in water from MD simulations, and thus find that its translational diffusion coefficient is  $D_A = 5 \times 10^{-10} \text{ m}^2/\text{s}$ . Assuming that the diffusion coefficient of the kinase is smaller, and taking  $\sigma = 10^{-9} \text{ m}$ , we estimate that  $k_d = 3.8 \times 10^9 \text{ s}^{-1} \text{ M}^{-1}$ . Because the geometries of most reactant encounters will not lead to successful binding events such steric effects will reduce this estimate considerably even by as much as 2 orders of magnitude. With estimates of  $k_d$  (omitting the steric factor) and  $k_r$ , we can invert the above equation for  $k_r$ , and thus find that  $(e_f^\ddagger - e'_d^\ddagger) \approx 4.5 \text{ kcal/mol}$ . However, from our PMF we estimate that  $(e_f^\ddagger - e'_d^\ddagger) \approx 2.2 \text{ kcal/mol}$ , a result consistent with the rate constant being  $\approx 10^8 \text{ M}^{-1}$ . Taking  $\kappa = 2 \times 10^{-2}$  would give  $k_r = 2 \times 10^6 \text{ s}^{-1} \text{ M}^{-1}$ , the result of Shaw et al. This rough estimate shows that the desolvation barrier determined in this paper is consistent with full scale MD simulations.

#### 4. CONCLUSIONS

In this article, we have determined, in atomistic detail, the existence of an intermediate free energy barrier encountered by the kinase-inhibitor Dasatinib, a potent anticancer drug, as it approaches the receptor src-kinase. This was accomplished by introducing two reaction coordinates, namely the ligand-pocket separation and the water occupation number and by performing free energy umbrella sampling simulations to determine the free energy surface as a function of these two reaction coordinates, we showed that the intermediate free energy barrier is mainly a result of desolvation of the binding pocket on the approach of the ligand to its binding pose. This barrier is close to the position of the ligand's native-binding pose. Desolvation thus seems to play an important role in the thermodynamics and kinetics of binding. Shaw and co-workers provided a pictorial view of the entire binding process of Dasatinib to its native binding pocket and also observed the presence of an intermediate shell of water molecules prior to binding to the pocket. Our study provides a free energy analysis of the role of the water encountered by the ligand in approaching its binding pose and shows that the free energy cost for the ligand to move from bulk to the position of the intermediate free energy barrier is comparable to the free energy cost for complete desolvation of the ligand-free pocket. Therefore, water molecules in the pocket act as a significant barrier to ligand binding. We find that the intermediate free energy barrier is the same for the pocket-restrained system as it is for the restraint-free system, a result that suggests that this barrier is not due to pocket reorganization but due to desolvation of the pocket. A continuum model like GBSA correctly locates the position of the free energy minimum corresponding to the binding pose but predicts an over-stabilized binding pose.

This paper also highlights the underlying contribution of the desolvation free energy barrier to the kinetics of the kinase-ligand binding. Surprisingly, a simple kinetic model for the on rate constant in which the ligand diffuses up to a doorway state and then surmounts the desolvation free energy barrier was found to be consistent with microsecond long simulations of

the ligand binding kinetics for this system [see ref 13], despite the fact that a rigorous treatment of the kinetics will require a multidimensional reaction coordinate in which the state of the cavity water is included. Nevertheless, the simple model suggests that it may be possible to use data from free energy analysis together with a more extensive diffusion based doorway state model to predict on and off rate constants for ligand binding in at least this class of protein–ligand systems. This would greatly increase the throughput of drug evaluation protocols and would allow the bypassing of very long computer simulations.

## ■ ASSOCIATED CONTENT

### ■ Supporting Information

Effect of different independent initial configurations on free energy profile as a function of ligand-pocket separation in explicit water, in implicit solvent, pocket-water profile, one-dimensional PMF as a function of water number in a ligand free pocket, two-dimensional PMF as a function of water number and ligand pocket separation, the convergence plot of pocket water profile and the illustration of hydrogen bonding interaction of key water molecules with the proximal amino acid residues. Also provided the detailed parameter tables for ligand Dasatinib as used in the paper. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## ■ AUTHOR INFORMATION

### Corresponding Author

\*Phone: 212-854-2186. E-mail: [bb8@columbia.edu](mailto:bb8@columbia.edu).

### Notes

The authors declare the following competing financial interest(s): B.J.B. is a consultant to Schrodinger, Inc. and is on their Scientific Advisory Board. R.A.F. has a significant financial stake in, is a consultant for, and is on the Scientific Advisory Board of Schrodinger, Inc.

## ■ ACKNOWLEDGMENTS

This work was supported by grants from the National Institutes of Health [NIH-GM4330 (to B.J.B.)] and by the National Science Foundation through [via Grant No. NSF-CHE-0910943]. We gratefully acknowledge the computational support of the Computational Center for Nanotechnology Innovations (CCNI) at Rensselaer Polytechnic Institute (RPI). This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number OCI-1053575. B.J.B. is a consultant to Schrodinger, Inc. and is on their Scientific Advisory Board. R.A.F. has a significant financial stake in, is a consultant for, and is on the Scientific Advisory Board of Schrodinger, Inc.

## ■ REFERENCES

- (1) Massova, I.; Kollman, P. A. Combined Molecular Mechanical and Continuum Solvent Approach (MM-PBSA/GBSA) to Predict Ligand Binding. *Perspect. Drug Discovery Des.* **2000**, *18*, 113–135.
- (2) Moble, D.; Dill, K. A. Binding of Small-molecule Ligands to Proteins: "What You See" is Not Always "What you Get". *Structure* **2009**, *17*, 489–498.
- (3) Gilson, M. K.; Zhou, H. X. Calculation of Protein-Ligand Binding Affinities. *Annu. Rev. Biophys. Biomol. Struct.* **2007**, *36*, 21–42.
- (4) Guvench, O.; MacKerell, A. D. J. Computational Evaluation of Protein-small molecule Binding. *Curr. Opin. Struct. Biol.* **2009**, *19*, 56–61.

- (5) Copeland, R. A.; Pompliano, D. L.; Meek, T. A. Drug-target Residence Time and its Implications for Lead Optimization. *Nat. Rev. Drug Discovery* **2006**, *5*, 730–739.

- (6) Young, T.; Abel, R.; Kim, B.; Berne, B. J.; Friesner, R. A. Motifs for molecular recognition exploiting hydrophobic enclosure in protein-ligand binding. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 808–813.

- (7) Abel, R.; Young, T.; Farid, R.; Berne, B. J.; Friesner, R. A. Role of The Active-site Solvent in The Thermodynamics of Factor Xa ligand binding. *J. Am. Chem. Soc.* **2008**, *130*, 2817–2831.

- (8) Baron, R.; Setny, P.; McCammon, J. A. Water in Cavity-Ligand Recognition. *J. Am. Chem. Soc.* **2010**, *132*, 12091–12097.

- (9) Wang, L.; Berne, B. J.; Friesner, R. A. Ligand Binding to Protein-binding Pockets with Wet and Dry Regions. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 1326–1330.

- (10) Mondal, J.; Morrone, J. A.; Berne, B. J. How Hydrophobic Drying Forces Impact the Kinetics of Molecular Recognition. *Proc. Nat. Acad. Sci. U.S.A.* **2013**, *110*, 13277–13282.

- (11) Setny, P.; Baron, R.; Kekenes-Huskey, P. M.; McCammon, J. A.; Dzubiella, J. Solvent Fluctuations in Hydrophobic Cavity-ligand Binding Kinetics. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 1197–1202.

- (12) Dror, R. O.; Pan, A. C.; Arlow, D. H.; Borhani, D. W.; Maragakis, P.; Shan, Y. Pathway and Mechanism of Drug Binding to G-protein-coupled Receptors. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 13118–13123.

- (13) Shan, Y.; Kim, E. T.; Eastwood, M. P.; Dror, R. O.; Seeliger, M. A.; Shaw, D. E. How Does a Drug Molecule Find its Target Binding Site? *J. Am. Chem. Soc.* **2011**, *133*, 9181–9183.

- (14) Zhang, J.; Yang, P.; Ullrich, A. Targeting Cancer with Small Molecule Kinase Inhibitors. *Nat. Rev. Cancer* **2009**, *9*, 28–39.

- (15) Gschwind, A.; Fisher, O.; Gray, N. The Discovery of Receptor Tyrosine Kinases: Targets for Cancer Therapy. *Nat. Rev. Cancer* **2004**, *4*, 361–370.

- (16) Getlik, M.; Grütter, C.; Simard, J. R.; Klüter, S.; Rabiller, M.; Rode, H. B.; Robubi, A.; Rauh, D. Hybrid Compound Design to Overcome The Gatekeeper T338M Mutation in cSrc. *J. Med. Chem.* **2009**, *52*, 3915–3926.

- (17) Cohen, P. Protein kinases—The Major drug Targets of the Twenty-first Century? *Nature Rev.* **2002**, *1*, 309–315.

- (18) Wang, L.; Wu, Y.; Deng, Y.; Kim, B.; Pierce, L.; Krilov, G.; Lupyán, D.; Robinson, S.; Dahlgren, M.; Greenwood, J.; Romero, D. L.; Masse, C.; Knight, J. L.; Steinbrecher, T.; Damm, W.; Harder, E.; Sherman, W.; Brewer, M.; Wester, R.; Murcko, M.; Frye, L.; Farid, R.; Lin, T.; Mobley, D.; Jorgensen, W. L.; Berne, B. J.; Friesner, R. A.; Abel, R. Accurate and Reliable Prediction of Relative Protein-Ligand Binding Affinities via Free Energy Calculations: Validation in Prospective Drug Discovery. 2014; Manuscript In Preparation.

- (19) Pearlstein, R. A.; Hu, Q.-Y.; Zhou, J.; Yowe, D.; Levell, J.; Dale, B.; Kaushik, V. K.; Daniels, D.; Hanrahan, S.; Sherman, W.; Abel, R. New Hypotheses About the Structure function of Proprotein Convertase subtilisin/kexin type 9: Analysis of The Epidermal Growth Factor-like Repeat A Docking Site Using WaterMap. *Proteins: Struct., Funct., Bioinf.* **2010**, *78*, 2571–2586.

- (20) Tran, Q.-T.; Williams, S.; Farid, R.; Erdemli, G.; Pearlstein, R. The translocation kinetics of antibiotics through porin OmpC: Insights From Structure-based Solvation Mapping Using WaterMap. *Proteins: Struct., Funct., Bioinf.* **2013**, *81*, 291–299.

- (21) Pearlstein, R. A.; Sherman, W.; Abel, R. Contributions of Water Transfer Energy to Protein-Ligand Association and Dissociation Barriers: Watermap Analysis of a Series of p38 aMAP Kinase Inhibitors. *Proteins: Struct., Funct., Bioinf.* **2013**, *81*, 1509–1526.

- (22) Madhavi Sastry, G.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W. Protein and Ligand Preparation: Parameters, Protocols, and Influence on Virtual Screening Enrichments. *J. Comput.-Aided Mol. Des.* **2013**, *27*, 221–234.

- (23) Schrödinger suite 2012 Protein Preparation Wizard; Schrödinger: New York. <http://www.schrodinger.com> (accessed 2012-09-30).

- (24) Kaminski, G.; Friesner, R.; Tirado-Rives, J.; Jorgensen, W. Evaluation and Reparametrization of the OPLS-AA Force Field for

Proteins via Comparison with Accurate Quantum Chemical Calculations on Peptides. *J. Phys. Chem. B* **2001**, *105*, 6474–6487.

(25) Jorgensen, W.; Chandrasekhar, J.; Madura, J.; Impey, R.; Klein, M. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–935.

(26) Nosé, S. A Molecular Dynamics Method for Simulations in The Canonical Ensemble. *Mol. Phys.* **1984**, *52*, 255–268.

(27) Hoover, W. Canonical Dynamics: Equilibrium Phase-Space Distributions. *Phys. Rev. A* **1985**, *31*, 1695–1697.

(28) Parrinello, M.; Rahman, A. Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method. *J. Appl. Phys.* **1981**, *52*, 7182–7190.

(29) Darden, T.; York, D.; Pederson, L. G. Particle mesh Ewald: An Nlog(N) Method for Ewald Sums in Large System. *J. Chem. Phys.* **1993**, *98*, 10089–10092.

(30) Essman, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pederson, L. G. A Smooth Particle Mesh Ewald Potential. *J. Chem. Phys.* **1995**, *103*, 8577–8592.

(31) Miyamoto, S.; Kollman, P. Settle: An Analytical Version of the SHAKE and RATTLE Algorithm for Rigid Water Models. *J. Comput. Chem.* **1992**, *13*, 952–962.

(32) Hess, B.; Bekker, H.; Berendsen, H.; Fraaije, J. G. E. LINCS: A Linear Constraint Solver for Molecular Simulations. *J. Comput. Chem.* **1997**, *18*, 1463–1472.

(33) Hess, B.; Kutzner, C.; Van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435–447.

(34) Qiu, D.; Shenkin, P. S.; Hollinger, F. P.; Still, W. C. The GB/SA Continuum Model for Solvation. A Fast Analytical Method for the Calculation of Approximate Born Radii. *J. Phys. Chem. A* **1997**, *101*, 3005–3014.

(35) Larsson, P.; Lindahl, E. A High-Performance Parallel-Generalized Born Implementation Enabled by Tabulated Interaction Rescaling. *J. Comput. Chem.* **2010**, *31*, 2593–2600.

(36) Onufriev, A.; Bashford, D.; Case, D. Exploring Protein Native States and Large-scale Conformational Changes with a Modified Generalized Born Model. *Proteins: Struct., Funct., Genet.* **2004**, *55*, 383–394.

(37) Schaefer, M.; Bartels, C.; Karplus, M. Solution Conformations and Thermodynamics of Structured Peptides: Molecular Dynamics Simulation with an Implicit Solvation Model. *J. Mol. Biol.* **1998**, *284*, 835–848.

(38) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. The Weighted Histogram Analysis Method for Free-energy Calculations on Biomolecules. I. The method. *J. Comput. Chem.* **1992**, *13*, 1011–1021.

(39) Patel, A. J.; Varilly, P.; Chandler, D. Fluctuations of Water Near Extended Hydrophobic and Hydrophilic Surfaces. *J. Phys. Chem. B* **2010**, *114*, 1632–1637.

(40) Andreev, S.; Reichman, D.; Hummer, G. Effect of Flexibility on Hydrophobic Behavior of Nanotube Water Channels. *J. Chem. Phys.* **2005**, *123*, 194502.

(41) Zhu, F.; Hummer, G. Drying Transition in the Hydrophobic Gate of the GLIC Channel Blocks Ion Conduction. *Biophys. J.* **2012**, *103*, 219–227.

(42) Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. PLUMED 2: New Feathers For an Old Bird. *Comput. Phys. Commun.* **2014**, *185*, 604–613.

(43) Lazaridis, T. Inhomogeneous Fluid Approach to Solvation Thermodynamics. 1. Theory. *J. Phys. Chem. B* **1998**, *102*, 3531–3541.

(44) Gallicchio, E.; Kubo, M. M.; Levy, R. M. Enthalpy, Entropy and Cavity Decomposition of Alkane Hydration Free Energies: Numerical Results and Implications for Theories of Hydrophobic Solvation. *J. Phys. Chem. B* **2000**, *104*, 6271–6285.

(45) Zhu, F.; Hummer, G. Pore Opening and Closing of a Pentameric Ligand-gated Ion Channel. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 19814–19819.

(46) Mondal, J.; Zhou, X.; Cui, Q.; Yethiraj, A. Self-Assembly of  $\beta$ -Peptides: Insight from the Pair and Many-Body Free Energy of Association. *J. Phys. Chem. B* **2010**, *114*, 13551–13556.

(47) Zhou, R.; Berne, B. J. Can a Continuum Solvent Model Reproduce The Free Energy Landscape of a  $\beta$ -hairpin Folding in Water? *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 12777–12782.