JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

COMMUNICATION

pubs.acs.org/JACS

How Does a Drug Molecule Find Its Target Binding Site?

Yibing Shan,[†] Eric T. Kim,[†] Michael P. Eastwood,[†] Ron O. Dror,[†] Markus A. Seeliger,[§] and David E. Shaw*[,†,‡]

[†]D. E. Shaw Research, New York, New York 10036, United States

[‡]Center for Computational Biology and Bioinformatics, Columbia University, New York, New York 10032, United States

[§]Department of Pharmacological Sciences, Stony Brook University, Stony Brook, New York 11794, United States

Supporting Information

ABSTRACT: Although the thermodynamic principles that control the binding of drug molecules to their protein targets are well understood, detailed experimental characterization of the process by which such binding occurs has proven challenging. We conducted relatively long, unguided molecular dynamics simulations in which a ligand (the cancer drug dasatinib or the kinase inhibitor PP1) was initially placed at a random location within a box that also contained a protein (Src kinase) to which that ligand was known to bind. In several of these simulations, the ligand correctly identified its target binding site, forming a complex virtually identical to the crystallographically determined bound structure. The simulated trajectories provide a continuous, atomic-level view of the entire binding process, revealing persistent and noteworthy intermediate conformations and shedding light on the role of water molecules. The technique we employed, which does not assume any prior knowledge of the binding site's location, may prove particularly useful in the development of allosteric inhibitors that target previously undiscovered binding sites.

The vast majority of all drugs act by binding to protein targets, but the physical process by which drugs and other ligands bind to proteins has proven difficult to elucidate, particularly at the single-molecule level. Here we report molecular dynamics (MD) simulations that capture this process, from beginning to end, in full atomic detail. In our simulations, the cancer drug dasatinib (sold under the trade name Sprycel, Bristol-Myers Squibb) and the kinase inhibitor PP1 bound spontaneously to their target, Src kinase, to produce complexes whose structures are virtually identical to those captured by X-ray crystallography[1,2] (Figure 1B). As far as we are aware, these are the first crystallographically accurate simulations of ligand binding that do not rely on prior knowledge of the binding site's location or on electrostatic forces that attract the ligand to its binding site from a distance: our simulations began with ligands positioned at random locations, the ligands were uncharged, and no artificial attractive forces were employed. The simulated ligands circled the target protein extensively before finding the binding site (Figure 1A).

Dasatinib found and bound to its binding site after 2.3 μs in one of four separate simulations, which together totaled 35 μs. In the simulation in which the dasatinib molecule reached the native binding pose, the molecule was mostly in the vicinity of the N-lobe (β-sheet) region; in other simulations, however, dasatinib contacted many other regions of Src kinase. PP1 bound after 15.1, 1.9, and 0.6 μs, respectively, in 3 of 7 separate simulations totaling 115 μs (Figure 2A). Although the accurate calculation of binding on-rates would require substantially more binding events than we have observed, these on-rates may be estimated from our simulations to be $\sim$1.9 s^{-1} μM^{-1} for dasatinib and $\sim$4.3 s^{-1} μM^{-1} for PP1; notably, the estimate for dasatinib is in line with the experimentally measured value ($\sim$5 s^{-1} μM^{-1}).[4]

It has been suggested that protein−ligand binding involves funnel-like free-energy landscapes[5,6] whose global minima correspond to the native conformations of the bound complexes. Our binding simulations—each of which demonstrates a presumably global energetic minimum corresponding to the native binding pose (i.e., that observed in X-ray structures), local energy minima corresponding to the trapping of the ligand in non-native poses, and a funnel-shaped energy landscape that extends to $\sim$10 Å root-mean-square deviation (rmsd) from the native pose (Figure 1C)—are highly consistent with this scenario.

Water molecules are known to play a critical role in protein−ligand binding.[7,8] In addition to correctly identifying the native ligand-binding poses of PP1 and dasatinib, our simulations resolved the locations of water molecules observed in X-ray structures of the Src−PP1 complex (Figure 2B). The simulations also revealed an intermediate state preceding PP1 binding, in which a single shell of water molecules separates the protein and ligand in the binding site (conformation a; Figure 2B). Our simulations suggest that this water shell, which survives for $\sim$0.1 μs before being disrupted upon PP1 binding, gives rise to a kinetic barrier to binding. A similar water shell is also observed immediately before dasatinib adopts its crystallographic binding pose (Figure 2D).

An emerging challenge in drug discovery concerns the identification of allosteric ligand-binding sites, through which drugs can modulate the effects of ligands that bind at the primary site. Allosteric binding sites may be obscure and underformed in the protein conformations captured by most X-ray structures. While the great majority of kinase inhibitors, including dasatinib and PP1, bind at the ATP-binding site, several allosteric sites have been identified in kinases.[9] The fact that PP1 molecules are repeatedly trapped at these alternative sites (Figure 1A) highlights the potential of such simulations as a tool for discovering allosteric sites. We identified, for example, a previously unreported binding site—a pocket, located

Received: March 25, 2011
Published: May 05, 2011

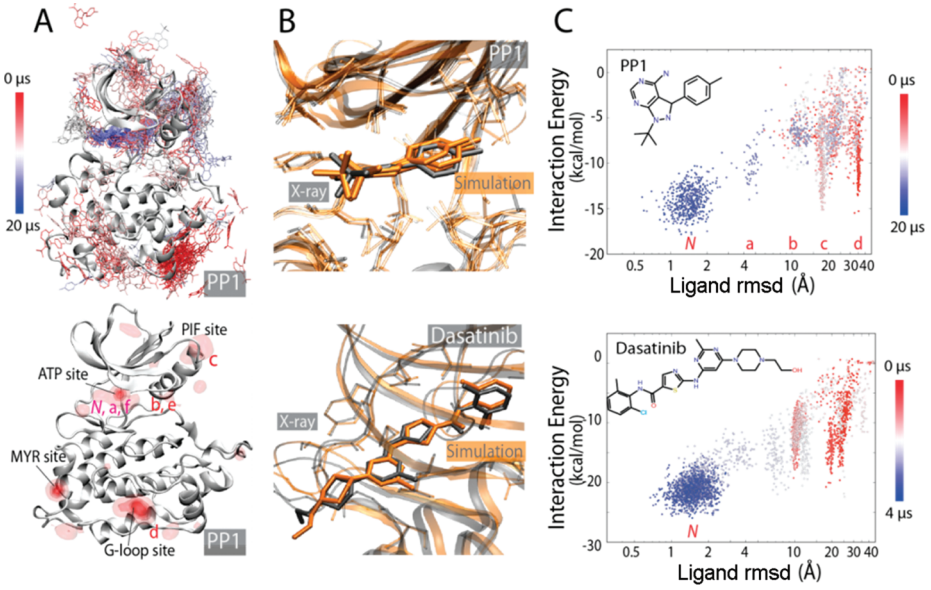

Figure 1. Simulation of the binding of PP1 and dasatinib to Src kinase. (A) Top: traces of a PP1 molecule in one simulation, in which the native pose is reached in 15 μs; for clarity, only the average conformation of the protein is shown. Bottom: density map (depicted by red isocontour surfaces) representing the frequency of PP1 interaction on the protein surface, based on all seven similar simulations we carried out. The high-density "hot spots" largely correspond to known ligand-binding sites (labeled) of protein kinases, but the "G-loop site" has not been identified previously. The locations of the native binding poses (*N*) and non-native poses (a−f) are also labeled. Multiple ligand poses may contribute to a density-map hot spot. (B) Binding poses of PP1 (from three simulations) and dasatinib (from one simulation) superimposed onto the X-ray structures (PDF entries 1QCF and 3G5D, respectively). (C) Binding poses of PP1 and dasatinib adopted in simulations, which are highly consistent with the crystallographic binding poses (within 2.0 Å rmsd), representing the most favorable ligand−protein interaction energies (estimated using a generalized Born model[3]). In principle, in the absence of the X-ray structures, the native binding poses may be identified by limited conformational fluctuation and low interaction energy of the bound ligands. When PP1 is in the native pose, for instance, its conformation fluctuates up to ∼2 Å rmsd with respect to the crystallographic binding pose (by contrast, when PP1 is in the G-loop pocket, its conformation fluctuates as much as 5 Å rmsd from an average conformation); the average interaction energy of PP1 in the native pose was found to be lower than that of PP1 in the G-loop pocket in our simulation by ∼2 kcal/mol. As shown in (A), the non-native poses b, c, and d correspond to PP1 interacting with the P loop, the αC helix region, and the αG helix, respectively; pose a features a single layer of water molecules between the protein and the ligand (see Figure 2).

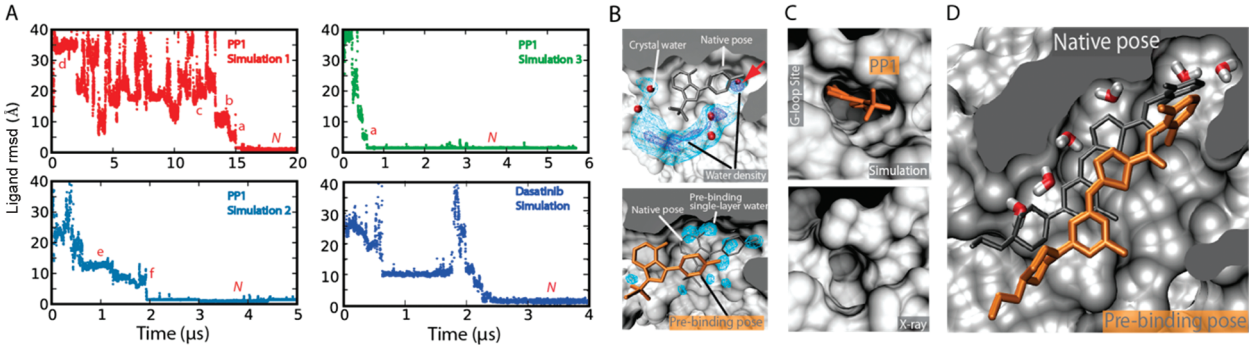

Figure 2. Water molecules in the simulated ligand binding and a putative allosteric binding site. (A) Ligand rmsd with respect to the native pose as a function of time for each binding simulation; each conformation is superimposed on the X-ray structures of the protein. Poses a−f refer to non-native PP1 poses, while *N* refers to the native poses of PP1 and dasatinib. Specifically, b−e refer to poses in which PP1 remains for 1 μs or more, while a and f refer to poses immediately preceding the native pose. In Simulation 2, PP1 binds via an alternative pathway that circumvents pose a. Poses e and f are observed only in Simulation 2, one in the ATP-binding site and the other in the extended ATP-binding site adjacent to the αC helix. (B) Water density (blue surfaces) in the native pose and in intermediate pose a, from a PP1 binding simulation. The simulated water density for the native pose, particularly that of the enclosed water (indicated by the red arrow), mirrors that of the X-ray structure. In pose a, which immediately precedes binding, a single layer of water molecules separates the ligand from the binding pocket surface. (C) A close-up of the G-loop site occupied by PP1, which is under-formed and too small to accommodate PP1 in the X-ray structure from which the simulation was initiated (PDB entry 1Y57). (D) In an intermediate state preceding dasatinib binding by 130 ns, a layer of water molecules separates dasatinib from the protein. In both (B) and (C), the protein is clipped to show the otherwise enclosed region.

dx.doi.org/10.1021/ja202726y |*J. Am. Chem. Soc.* 2011, 133, 9181–9183

between the αFG and αGH loops and the αF and αG helices, to which PP1 (and, in later simulations, dasatinib) bound—which is too shallow to accommodate PP1 or dasatinib in the X-ray structure (Figures 1A and 2C). An ongoing ligand-design effort targeting this putative allosteric binding site is currently underway in our laboratory.

More generally, an important limitation of traditional virtual drug screening is that it must start with a well-defined binding site, despite promising recent developments.[10,11] By allowing the identification of previously unknown binding sites, MD simulations of protein−ligand binding, such as those presented here, may substantially broaden the applicability of computational techniques to drug development.

■ ASSOCIATED CONTENT

Supporting Information. Methods of simulation and analysis; movie of dasatinib binding. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
David.Shaw@DEShawResearch.com

■ ACKNOWLEDGMENT

We thank Prof. Bruce Berne of Columbia University and Prof. Jin Wang of Stony Brook University for helpful discussions.

■ REFERENCES

(1) Sicheri, F.; Moarefi, I.; Kuriyan, J. *Nature* **1997**, *385*, 602.
(2) Getlik, M.; Grütter, C.; Simard, J. R.; Klüter, S.; Rabiller, M.; Rode, H. B.; Robubi, A.; Rauh, D. *J. Med. Chem.* **2009**, *52*, 3915.
(3) Onufriev, A.; Bashford, D.; Case, D. A. *J. Phys. Chem. B* **2000**, *104*, 3712.
(4) Shan, Y.; Seeliger, M. A.; Eastwood, M. P.; Frank, F.; Xu, H.; Jensen, M. Ø.; Dror, R. O; Kuriyan, J.; Shaw, D. E. *Proc. Natl. Acad. Sci. U. S.A.* **2009**, *106*, 139.
(5) Bryngelson, J. D.; Onuchic, J. N.; Socci, N. D.; Wolynes, P. G. *Proteins: Struct., Funct., Bioinf.* **1995**, *21*, 167.
(6) Wang, J.; Verkhivker, G. M. *Phys. Rev. Lett.* **2003**, *90*, 188101–1.
(7) Ladbury, J. E. *Chem. Biol.* **1996**, *3*, 973.
(8) Young, T.; Abel, R.; Kim, B.; Berne, B. J.; Friesner, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 808.
(9) Taylor, S. S.; Kornev, A. P. *Trends Biomed. Sci.* **2011**, *36*, 65.
(10) Hetényi, C.; van der Spoel, D. *FEBS Lett.* **2006**, *580*, 1447.
(11) Davis, I. W.; Raha, K.; Head, M. S.; Baker, D. *Protein Sci.* **2009**, *18*, 1998.