HEALTH AND MEDICINE

How and when does an anticancer drug leave its binding site?

Pratyush Tiwary,[1] Jagannath Mondal,[2] B. J. Berne[1]*

Obtaining atomistic resolution of drug unbinding from a protein is a much sought-after experimental and computational challenge. We report the unbinding dynamics of the anticancer drug dasatinib from c-Src kinase in full atomistic resolution using enhanced sampling molecular dynamics simulations. We obtain multiple unbinding trajectories and determine a residence time in agreement with experiments. We observe coupled protein-water movement through multiple metastable intermediates. The water molecules form a hydrogen bond bridge, elongating a specific, evolutionarily preserved salt bridge and enabling conformation changes essential to ligand unbinding. This water insertion in the salt bridge acts as a molecular switch that controls unbinding. Our findings provide a mechanistic rationale for why it might be difficult to engineer drugs targeting certain specific c-Src kinase conformations to have longer residence times.

INTRODUCTION

Traditional computational drug design has been based on the optimization of the binding affinity of a drug for a target. However, it is becoming increasingly recognized that the efficacy of a drug is determined not only by its equilibrium binding affinity but also by its residence time in the target protein (1, 2). Thus, effective drug design should not be exclusively based on the thermodynamic binding affinity, and it becomes desirable to extend the traditional design strategies to incorporate calculation of the drug-target lifetime (often expressed through its inverse, the off-rate k_{off}). Although experimental techniques can directly measure the unbinding time, the unbinding dynamics of a drug is much more than just a number. Along with quantifying the residence time, it is important to understand the molecular determinants of unbinding and how they give rise to dissociation pathways and barriers. These determinants include, but are not limited to, key protein-drug and protein-protein interactions, conformational states of the protein, ligand flexibility, and solvation effects (1–4). Armed with an atomistic resolution understanding of the role played by these determinants in the unbinding dynamics, one can then imagine proposing structural modifications to a ligand that would lead to ideally desired longer residence times.

Here, we gain atomic-level insight into the unbinding mechanisms of the anticancer drug dasatinib from the kinase c-Src by performing extensive fully atomistic molecular dynamics (MD) simulations. c-Src is a nonreceptor cytoplasmic tyrosine kinase that is involved in a multitude of cell signaling activities and whose heightened activity has been linked to the progression of a variety of cancers, including those in the colon, liver, lung, and pancreas (5). Dasatinib (sold under the trade name Sprycel by Bristol-Myers Squibb) is a U.S. Food and Drug Administration–approved drug that is used to treat patients with chronic myelogenous leukemia and other related ailments by blocking c-Src as well as Abl kinases. Using fully atomistic simulations to understand the detailed basis of how dasatinib unbinds arguably could have profound implications for designing tailored drugs with desired residence times. However, the reported residence times of dasatinib, and of most successful drugs in general, are well into seconds or slower,

which is far beyond the reach of atomistic simulations without the use of carefully designed enhanced sampling methods (6–9). As such, in this work, we use an enhanced sampling method called "infrequent metadynamics" (see Materials and Methods) that accelerates unbinding to make it computationally tractable yet allows the recovery of unbiased kinetic information, including pathways and time scales. This method has been successfully used to study biophysical dynamics in a variety of systems (6, 10–13) and is ideally suited for the problem at hand.

Here, we obtain 12 independent unbinding trajectories that begin from the bound x-ray structure for dasatinib with c-Src kinase [PDB (Protein Data Bank) 3G5D] and culminate when the ligand is either fully solvated or diffusing on the surface of the protein far from the binding site. Our simulations show that the unbinding of dasatinib from c-Src kinase proceeds through multiple intermediate stable states, facilitated by a fluctuating protein and aided by movement of water molecules that modulate the disruption of shielded protein-protein and protein-drug interactions. We quantify the lifetimes of these intermediates (ranging from fractions of milliseconds to seconds) and identify two broad classes of slow steps in the unbinding process. The first slow step involves an initial hydration of the binding pocket but with the ligand itself relatively stationary. This event is coupled with a temporary disruption of a key evolutionarily preserved Glu-Lys salt bridge interaction (14), which becomes water-mediated and effectively acts like a molecular switch controlling the onset of unbinding. When distorted, this salt bridge, which is considered as a hallmark of an activated kinase when fully formed, creates space for further entry of water molecules into the binding pocket. In the second and even slower step, assisted by the water molecules now present in the binding pocket, the drug breaks its primary interactions with the kinase residues and then finally exits it. Typically, after the drug exits, the aforementioned water molecules also leave the pocket, and the Glu-Lys salt bridge reforms, recovering the x-ray pose. The drug then diffuses in the solvent or on the surface of the protein, moving through a domain of attractive surface spots. Here, we do not explore the time scales of this diffusion but instead focus on the residence time of the drug in the primary binding pocket. It is possible that these attractive spots on the surface that we identified could either be alternate binding locations or have allosteric effects on the binding to the primary pocket. Future work will consider these very relevant questions.

We were able to calculate a residence time that agrees with reported experimental measurements and determine mechanistic

[1]Department of Chemistry, Columbia University, New York, NY 10027, USA. [2]Tata Institute of Fundamental Research, Center for Interdisciplinary Sciences, 21 Brundavan Colony, Narsingi, Hyderabad, India.
*Corresponding author. Email: bb8@columbia.edu

features pertaining to the rate-limiting steps in the unbinding of dasatinib that also agree with experiments. This atomistic insight sheds light on previously unexplained experimental observations (15), such as why the targeting of certain c-Src kinase conformational forms might be futile for engineering desired increased drug residence times (16). Our work also provides direct evidence that water is not just a spectator solvent in drug unbinding but a very active participant (17–19). Our results mark a significant step forward in the ability to calculate drug unbinding pathways and rates with statistical reliability and to obtain a detailed understanding of the complex nature of drug unbinding from an important target for various diseases, the c-Src kinases.

RESULTS

Calculation of k_{off}

The enhanced sampling technique used in this work, metadynamics (20), is a biasing procedure whereby periodically depositing bias as a function of carefully chosen collective variables (CVs), the system is assisted in moving across high-energy barriers that would ordinarily be difficult to escape. Traditionally, the final outcome of this biasing procedure is the unbiased probability distribution of the system, which can be obtained directly by using Eq. 4 (Materials and Methods) or through a reweighting procedure (20, 21). Independent of the free energy estimate, one can also obtain kinetic rate constants through the calculation of an acceleration factor, which directly gives a speed-up in time achieved through metadynamics, assuming that certain important conditions have been met (see Materials and Methods and Eq. 1) (22, 23). One of the key conditions is that the bias should have been deposited infrequently enough to leave the transition states between stable states untouched. This can be done if the underlying dynamical problem is characterized by a few sharp and dominant free energy barriers. The time scales reweighted from metadynamics can then be trusted if they conform to time-homogeneous Poisson statistics (see Materials and Methods). We can expect to use this self-consistency criterion here to verify the accuracy of the accelerated time scales because ligand unbinding has been shown in various previous experimental and numerical studies to conform to the aforementioned energy landscape picture with a small number of prominent barriers and short transition path times (24–28).

Here, we perform 12 independent metadynamics runs, all of which result in unbinding of the ligand. These give a direct measurement of the residence time equaling 21 ± 10 s, which agrees with reported experimental observations lying in the range of 18 s [using the values K_d (dissociation constant) = 0.011 μM and k_{on} = 5 s^{-1} μM^{-1} and the relation $k_{off} = k_{on} K_d$; residence time = $1/k_{off}$ (29, 30)] to 900 s (15) (coming from the fluorescent-labeled analog of dasatinib). We performed the calculations using the OPLS all-atom force field and TIP4P water models (31, 32) and also validated the results using the AMBER all-atom force field (33) with TIP4P water to check for qualitative consistency of the key features (see the Supplementary Materials). Note that our measurement of k_{off} in this work as well as the experimental measurement of Kwarcinski et al. (15) are direct and do not assume the relation $k_{off} = k_{on} K_d$. However, if we take the liberty of using this relation along with the experimental measured value of K_d, our measured k_{off} corresponds to k_{on} = 4.3 s^{-1}μM^{-1}, which is in marked and possibly fortuitous agreement with the experimental measurement (k_{on} = 5 s^{-1}μM^{-1}) (30). Previous computational studies on this system have been semiquantitative at best because of insufficient sampling but have obtained k_{on} in a similar order of magnitude (30, 34).

Here, the metadynamics runs were performed by gradually biasing two CVs, the distance d of the drug to the ligand and the solvation state w of the binding pocket (see Materials and Methods for details). Our particular choice of CVs for this system was motivated by reported experiments and previous computer simulations of protein-ligand systems [including that in this work (30, 34)], which suggest that both the ligand-protein distance and the hydration of the binding pocket play prominent roles in the unbinding process. In the Supplementary Materials, we provide statistical analyses of the unbinding times obtained. The Poisson distribution of the unbinding time scales obtained by biasing the CVs d and w gives us some confidence that, at least for the system in this work, these two CVs do suffice to capture the slow dynamics relevant to the unbinding of dasatinib, and any other slow variables relevant to unbinding are nonorthogonal to these two. In future work on drug unbinding, we intend to use the methods of Tiwary and Berne (35), which was shown to be useful for CV selection with minimal preknowledge of the system. In the study of Tiwary and Berne (36), we also consider the question of how diffusional anisotropy in the CVs can be incorporated into this framework. In the Supplementary Materials, we provide further detailed validation of a master equation built from our simulations, which adds to our confidence that any other slow variables for unbinding are not orthogonal to d and w.

Approximate free energy surface and stable intermediates

In Fig. 1, we provide the approximate free energy surface (FES) (defined in Eq. 4 in Materials and Methods) as a function of d and w, obtained by averaging the various unbinding runs (see the Supplementary Materials for theory regarding relation between bias and equilibrium free energy in metadynamics). To converge FES reported in this work in terms of the relative stability of basins, one would need frequent bias deposition, which is not our aim here, and we use the FES averaged from the various infrequent biasing runs using Eq. 4 only to establish the approximate locations of the various stable states and draw useful information about the existence of metastable intermediates from the topology of this FES. Also note that at no point do we make use of the free energy magnitudes to obtain the time scales or rate constants. This is done through the calculation of the acceleration factor (Eq. 1 in Materials and Methods), which does not need any estimates of the free energy. Because we have already demonstrated that using d and w can capture relevant slow motions in the unbinding, we do not need a more sophisticated full-atom clustering scheme for the purpose of demarcating stable intermediates.

The main features of this FES are six basins or stable states such that the barrier between them is at least $3k_B T$. There is nothing special about this particular choice of this parameter for demarcating the stable states. A higher threshold would lead to a model with less number of states and thus less information, whereas a lower number would have more detailed information but also larger errors in the sampling of transition rates between the different states. Any choice in the end has to be examined through self-consistency tests, which we describe here and in the Supplementary Materials. We mark respective boundaries near the transition regions between the states, as was done, for instance, by Swope et al. (37). By counting transition rates between these stable basins, we set up a master equation for the unbinding dynamics, with absorbing boundary conditions in state 6 (unbound state). By solving for the slowest eigenvalues of the transfer matrix (detailed in Fig. 2 and in the Supplementary Materials), we get an estimate of the rate-determining steps in the full unbinding dynamics reported in the previous section. Note that in no way are we implying that this classification is the only one consistent with the unbinding dynamics—the problem of finding
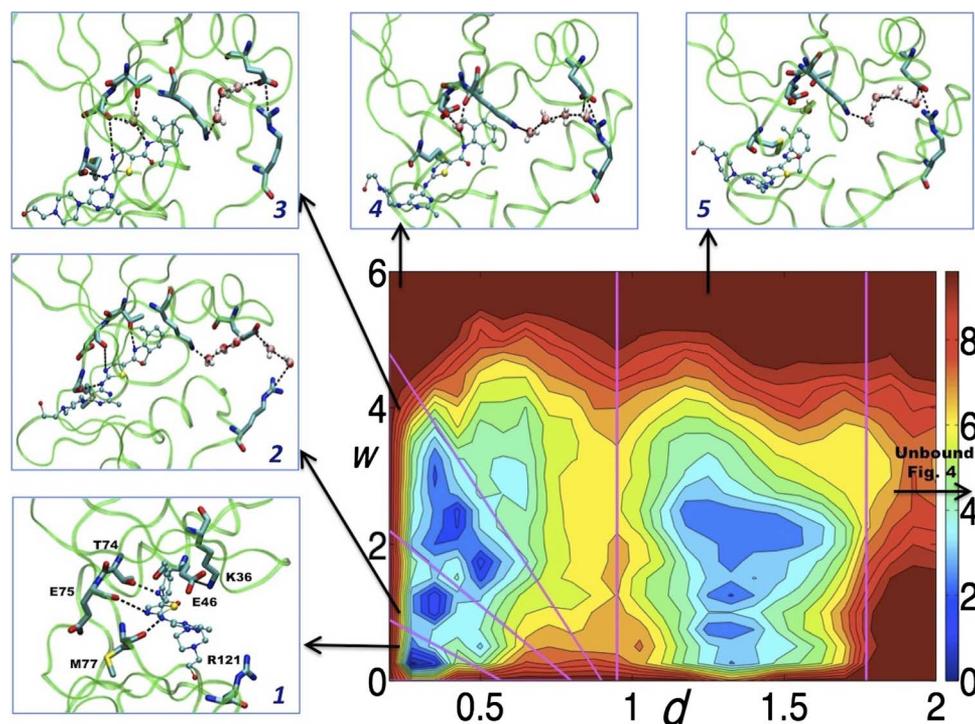
Fig. 1. Metastable intermediates, key residues, and interactions relevant to the unbinding of dasatinib from c-Src kinase, along with the associated FES as a function of ligand-binding pocket distance d (in nanometers) and hydration state of binding pocket w. Energy is in units of kcal/mol, with contours drawn every 0.5 kcal/mol. Various relevant residues along with water molecules that enter the protein have been marked, and dasatinib is shown using the ball and stick model. State 6 (unbound state) is illustrated separately in Fig. 4. See Results for detailed descriptions of the defining interactions in every state.
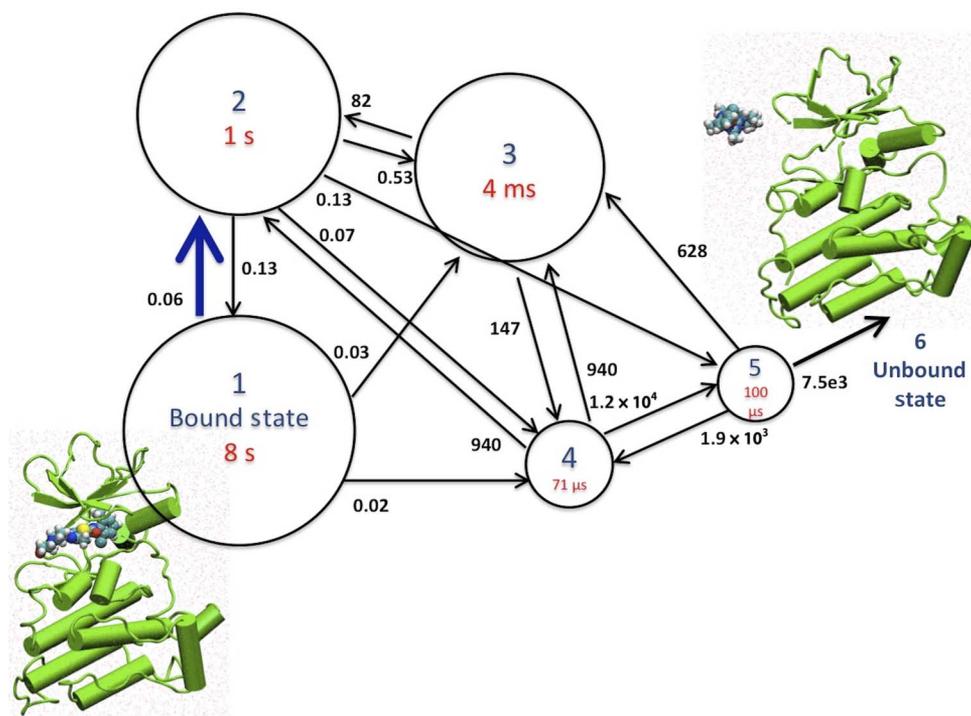
Fig. 2. The complex network of state-to-state transitions for dasatinib as it moves from bound to unbound state, with rate constants in s^{-1} and various residence times. See Fig. 1 for definition of the various states. The radii of circles for various states are approximately logarithmically proportional to their respective residence times. The full rate constant matrix is provided in the Supplementary Materials.

a coarse Markovian model that fits higher-resolution observations has many solutions, and the analysis we present is simply one Markovian classification, consistent with all observations at hand. In the Supplementary Materials, we present a detailed robustness analysis of the master equation with respect to recrossings between stable states.

Molecular characteristics of various stable intermediates

The various stable intermediates involved in the unbinding of dasatinib are illustrated alongside the FES in Fig. 1, and a schematic summarizing the complex network of transitions between these states is provided in Fig. 2. The lifetimes reported here are obtained by summing the times spent in the respective stable states in Fig. 1. Detailed analyses of these along with error estimates are provided in the Supplementary Materials.

State 1 is the x-ray–bound pose, whose lifetime we found to be around 8 s. The important stabilizing interactions in this pose are between the ligand and the residues Thr^{74}, Glu^{75}, and Met^{77} in the protein, as documented in various other published works (4, 29). Also note the salt bridge between Lys^{36} and Glu^{46}, which is evolutionarily preserved across the kinome and is critical for kinase functioning.

The next stable state is state 2, with an average lifetime of around 1 s. The most distinct characteristic of this state is that the Lys^{36}-Glu^{46} direct interaction is broken and is instead water-mediated, with the αC helix rotated outward, relative to state 1. Glu^{46} engages in water-mediated interactions with either of the two arginine residues, Arg^{121} or Arg^{145}, which play a role in the transition of the kinase from active to inactive (38). In these first two states, there is relatively minimal movement of the ligand with respect to the protein, and the long residence times can be explained by the key interactions being preserved.

State 3 has a lifetime of 4 ms and involves further entry of water molecules into the binding pocket space opened up by the rotation of the αC helix. Most notably, one of these water molecules now helps break the shielded hydrogen bond between the buried Thr^{74} residue and dasatinib (Fig. 1). As per our classification, there is a range of positions of Glu^{46} in this state, where it can continue to interact with the arginine of state 2 directly or through water mediation. This state, which is a necessary and long-lived intermediate for dasatinib unbinding, is very close in root mean square deviation (RMSD) to the αC helix–out structure reported for c-Src kinase (PDB 4YBK) in complex with certain αC helix–out binders (see Fig. 3). The activation loop displayed high flexibility in our simulations while staying unfolded through the entirety of all our runs. Note that the DFG motif (residues 140 to 142) of the kinase stays in the so-called "in" state.

Once the deeply buried hydrogen bond between dasatinib and the protein is broken in state 3, the ligand becomes more mobile, as can be seen from the elongation of basin 3 in the FES in Fig. 1. States 4 and 5 have lifetimes of 71 and 100 μs, respectively. They correspond to the gradual breaking of the interactions between the ligand and the residues Glu^{75} and Met^{77}, again aided by water molecules that entered the binding pocket. After exiting state 5, the drug diffuses either in the solvent or on the surface of the protein by moving through a network of interconnected domains, which we illustrate in Fig. 4. These domains are in good qualitative agreement with the sites observed as the ligand binds in unbiased binding simulations reported by Shan et al. (30). Here, we do not characterize in detail the time scales for surface diffusion, but from our runs, it appeared (at least qualitatively) that the time scales were shorter by several orders of magnitude than the residence time in the primary binding pocket. This is further supported by the observation that even for surface sites located in roughly the same region of the protein, the stabilizing interactions were not unique and

changed from run to run. The small residence times in these surface regions are probably due to the solvent-exposed nature of these regions and their accessibility to water. An accurate characterization of the time scales of the surface diffusion is beyond the scope of the infrequent metadynamics method used in this work but has been attempted in the past using other methods (34). Furthermore, it has also been documented that many force fields display a systematic bias toward increased hydrophobic interactions that could possibly lead to spurious compaction of unfolded proteins and systematic deviations in solvation free energies (39). An accurate study of the surface diffusion would also need to take this complication into account.

What are the slow and rate-limiting steps in the unbinding process?

As summarized in Fig. 2, the unbinding of dasatinib is a fairly concerted process that involves moving through the various stable states described in the previous subsection and is reminiscent of the induced-fit mechanism described, for instance, by Copeland (1, 40). Upon solving for the eigenvalues of the associated master equation, we identify two key slow
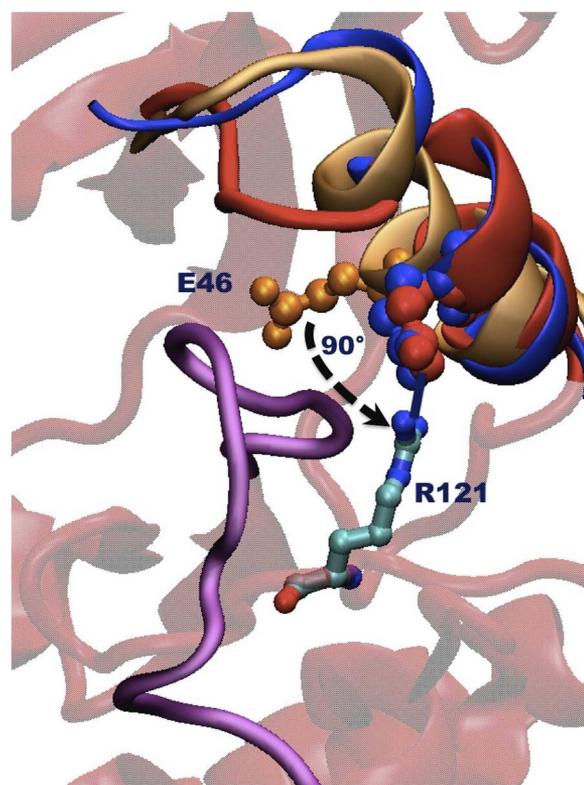
Fig. 3. Various protein structures aligned using the full protein's heavy atoms, showing the rotation of the αC helix and of the glutamic acid residue (Glu⁴⁶ in PDB 3G5D). Orange represents equilibrated αC helix–in structure, red is the αC helix–out structure (or state 3 in Fig. 1) obtained from our runs, and blue is the αC helix–out structure directly from the PDB database (PDB 4YBK) as reported by Kwarcinski et al. (15). Note how the αC helices in the two out structures overlap fairly well and especially how the glutamic residue is displaced almost identically in a roughly orthogonal direction to what it had in the αC helix–in structure. The stabilizing interaction between residues Arg¹²¹ and Glu⁴⁶ is also illustrated, along with the disordered activation loop in purple. The underlying base structure is a representative αC helix–out structure obtained from our runs. The reported structure (PDB 3G5D) in the protein database has missing residues after the end of the helix, and as such, we do not include them in the comparison here.
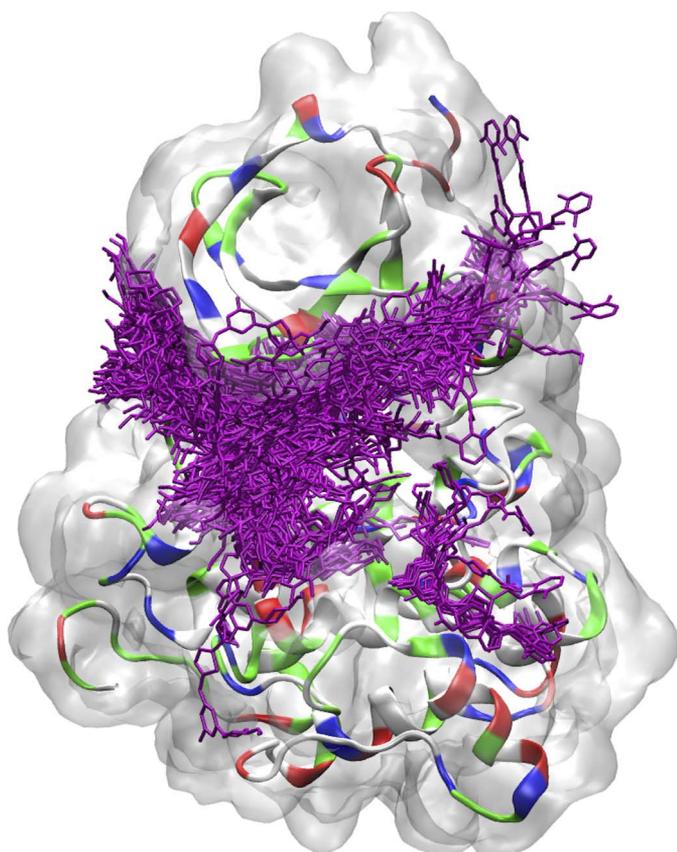
Fig. 4. An interconnected network of domains through which dasatinib moves once unbound from the primary binding pocket (that is, *d* > 1.8 nm; see Fig. 1), clearly demonstrating the presence of surface attractive spots. Here, various dasatinib snapshots from the unbinding trajectory satisfying *d* >1.8 nm are marked in purple on the protein (colored by residue name), which for illustrative purposes is kept in the bound pose, and the surface of the protein is represented in gray.

Lys^{36}-Glu^{46} salt bridge functions as a molecular switch that must be "on" for the drug unbinding to progress. Designing drugs that target kinase conformations with this switch already on might at best lead to similar, if not shorter, residence times. Our findings are well supported by recent experimental measurements (*15*) where this was found to be the case: Dasatinib had a very similar residence time in c-Src kinase compared to a ligand that bound to c-Src kinase in its αC helix–out configuration. In Fig. 3, we provide an overlay of the latter configuration with our state 3 to demonstrate that they are very similar. It will be interesting to see whether this same reasoning applies to other kinases, which adapt both the αC helix–in and αC helix–out conformations.

Note that the molecular switch mechanism described only where the Glu-Lys salt bridge distorts and where water enters the binding pocket and does not indicate which is the cause and which is the effect: that is, whether the salt bridge distortion leads to the entry of water molecules or the water entry causes the salt bridge to distort. One way to settle this question would be to start independent MD runs from a configuration where water molecules had entered the pocket but the Lys^{36}-Glu^{46} direct interaction was still preserved, applying the static bias until this point and adding no further bias after this point. However, given the high–free energy barrier of the event, it might be very tricky to draw statistically reliable information from this procedure. A more quantitatively robust approach to settling this question, which will be the subject of future work, would be through the use of the algorithm spectral gap optimization of order parameters (*35*), which allows distinguishing between driven versus driving variables.

DISCUSSION

Here, we have studied in detail the unbinding dynamics of the anticancer drug dasatinib from c-Src kinase. By collecting statistics over multiple unbinding trajectories, we obtain a residence time in marked agreement with the experimental measurements. It is possible that our residence time estimates are slightly underestimated because of the neglect of diffusive parts of the trajectory once the ligand is on the surface of the protein. Apart from a quantitative agreement in residence times, we also obtain a descriptive understanding of the pathways of unbinding, where coupled protein-water movement controls and facilitates the unbinding of the drug, and an evolutionarily preserved salt bridge acts like a molecular switch. This mechanism has been previously proposed for slow unbinding systems, such as biotin-streptavidin (*24, 25, 28*), where it was found that water entry coupled with rupture of a salt bridge is a necessary but insufficient step for unbinding. In our case, this salt bridge comes from the αC helix, which is not entirely surprising given that this helix is flexible across a variety of kinases (*14, 38*). The errors in classical force fields could have significant effects on any such findings. For instance, the OPLS force field, as used in this work, overestimates the strength of salt bridges (*41*), such as the one we discuss here. However, the fact that we obtain the same qualitative mechanism with another force field gives us confidence in our findings. Furthermore, the ability to do these detailed kinetics calculations with mechanisms and rate constants comparable to other experiments should also serve as a tool in the endeavor to improve force fields.

Our work has at least two obvious pharmacological implications. First, although hydration of the binding pocket and of the ligand is important in unbinding, our work shows exactly when and where on the unbinding pathway the hydration happens and what roles it plays. Although, strictly speaking, the hydration of the binding pocket is a

steps in the network of steps (fig. S1B). The rate-limiting step involves moving from states 3, 4, and 5 to 6, that is, the movement of the drug accompanied with the breaking of the interactions between the drug and the Glu^{75} and Met^{77} residues in a configuration where the αC helix is already rotated. The second slow step, which is around seven times faster than the true rate-limiting step but always happens before the second step, is the rotation of the αC helix–out configuration from state 1 to 2, through coupled protein-water movements but with the drug itself staying stationary. In this step, because the Lys^{36}-Glu^{46} direct interaction is broken (Fig. 3), the outward movement of Glu^{46} creates space for water molecules to freely move in and out of the protein. Because we obtained the same qualitative unbinding mechanism using a different classical force field, our confidence in this proposed mechanism is strengthened (see the Supplementary Materials). However, erroneous stabilization of salt bridge strength could still be a more generic force field artifact and will be investigated in future work.

The above observation is a crucial finding because it demonstrates that a drug that binds to the αC helix–out conformation will not necessarily have a longer residence time than a drug that binds to the true αC helix–in configuration because even the latter has to move through the αC helix–out configuration; moreover, the rate-limiting step comes into play only after this conformation has been attained. That is, the

slightly faster step than the movement of dasatinib, we find that the hydration step is the first slow step, and the associated barrier must be surmounted before anything else can happen. Thus, one way to engineer drugs with longer residence times that do not depend on specific ligand-pocket interactions would be to focus on controlling water access to the binding pocket.

Second, it explains and quantifies the role played by the αC helix and the evolutionarily preserved Lys-Glu salt bridge as a molecular switch in unbinding, helping us to understand the recent observation that drugs binding to αC helix–out kinase conformations have similar residence times as drugs binding to αC helix–in conformations. This understanding has obvious pharmacological implications and, if verified for other kinases, could be of considerable use in drug design by indicating that slow off rates are not an inherent feature of αC helix–out unbinders, as also noted in the experiments of Kwarcinski et al. (15). For instance, extensive efforts have been made to design conformationally selective kinase inhibitors, following the huge success of Gleevec (imatinib, Novartis) (42), which binds only to a very specific conformation (43). Our work provides preliminary hints that chasing after specific kinase conformations might be rather futile from the perspective of tailoring residence times. However, given that ligand unbinding involves passing through the αC helix–out state, a better strategy to engineer drugs with long residence times would be to block the accessibility of this state.

Given our identification of the slow steps in unbinding, we can make forward-looking suggestions that would be directly verifiable in experiments and, if validated, could serve as general guiding principles for engineering related drugs with long residence times. For instance, it should be possible to introduce mutations in the αC helix itself or in parts of the kinase domain that engage with the αC helix that restrict the movement of the αC helix but do not disturb the key stabilizing interactions in the bound dasatinib configuration. We expect this intervention to have minimal effect on the K_d but an order of magnitude effect on k_{off} and possibly k_{on}. Differences between how k_{off} and k_{on} are affected would also serve as evidence for the asymmetrical nature of binding/unbinding. Similar slowing down of the unbinding should be expected by substituting D_2O for H_2O in the solvent.

We realize that the residence time of seconds we have achieved in this work, although it represents a step forward for atomistic simulations, can still be short of the desirable time scales for many practical applications, when the target residence time exceeds the pharmacokinetic half-life of the drug in systemic circulation (a time scale of tens of minutes, hours, or even days) (1). That said, our current work marks an ambitious and potential step toward mapping the unbinding mechanisms of various drugs from kinases and other proteins with all-atom insight. Many successful kinase inhibitors on the market, which are often only one of a kind, are distinguished among other things by their long residence times in selected kinases, for example, Iclusig (ponatinib, Ariad Pharmaceuticals), which has long residence times even in the difficult case of tyrosine kinases with mutated gatekeeper residues (44), quite possibly a reason why it is one of the only available drugs to treat certain rare blood and bone marrow diseases. Although the calculations performed in this work are still computationally too expensive to be applied to a series of ligands against a series of mutated proteins, the continued improvements in computing power and force fields, coupled with the development of reliable enhanced sampling methodologies such as but not limited to those presented by Tiwary and Berne (35), makes us optimistic that we will soon be able to speed up these calculations, enough to achieve the design of drugs with tailored residence times in specific proteins, directly guided by atomistic simulations.

MATERIALS AND METHODS

System specifications

The initial structure of dasatinib bound to kinase was obtained from PDB (PDB 3G5D). OPLS all-atom force field and TIP4P water models were used. The full system, which consisted of 50,411 atoms, including protein, solvent, and charge-compensating counterions, was prepared and equilibrated using the exact same specifications as in the work of Mondal et al. (34). We refer to this work in full detail. All our simulations were performed using constant temperature and volume ensemble at temperature T = 300 K using the software GROMACS 4.6.7 patched with the plugin PLUMED 2.1.

We also repeated some of the calculations of this work using a complete different force field and ligand charge parametrization to ascertain the robustness of the results in a qualitative sense. For this, we used the AMBER ff99SB*-ILDN force field to repeat three independent unbinding runs with a frequent biasing protocol (bias every 0.5 ps; all other biasing parameters are identical to other reports). The ligand charges were parametrized using the general amber force field and AM1-BCC charges.

Infrequent metadynamics

Here, we used the infrequent metadynamics approach (22, 23) to obtaining kinetic rate constants. This involves periodic biasing of a few (typically one to three) CVs or order parameters to increase the probability of escape from metastable states where the system would ordinarily be trapped (20). If these CVs demarcate all relevant stable states of interest and if the time interval between biasing events is much longer than the transition path time spent in the ephemeral transition state regions, then the likelihood of not adding bias in the transition state (TS) regions increases. This preserves the sequence of transitions between stable states that the unbiased trajectory would have taken (22). Furthermore, the acceleration α of transition rates through biasing, which directly yields the true unbiased rates, can be calculated through generalized transition state theory (22, 45)

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

Here, s is the CV being biased, $\beta = 1/k_B T$ 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. One can verify a posteriori whether the requirements in the work of Tiwary and Parnello (22) and Salvalaglio (23) regarding quality of CVs and infrequency of bias addition were satisfied by checking whether the cumulative distribution function for the transition times from each stable state is Poissonian (23).

Two CVs were used in this work: the distance d of the drug to the ligand and the solvation state w of the binding pocket. The distance variable d was taken as the center-of-mass distance between two of the hydrogen bond formers in the bound pose, namely, between the groups Met^{77}:O,Met^{77}:N and ligand:N,ligand:N1. Because the CVs need to be smoothly differentiable to be biased, the pocket solvation variable w was implemented through the function

$$w = \sum_{i \in A} \sum_{j \in B} f(r_{i,j}) \tag{2}$$

Here, set A denotes the protein atoms Thr^{74}:Oγ1, Glu^{75}:O, and Asp^{140}:Cα that comprise the core of the buried interactions that the protein makes with the drug in the bound pose. Set B comprises all the

oxygen atoms of various water molecules. $r_{i,j}$ is the distance in nanometers between atoms i and j coming from these two sets, respectively. The function

$$f(r) \equiv \frac{(1 - (r/0.6)^6)}{(1 - (r/0.6)^{10})} \tag{3}$$

is a switching function that makes w smoothly differentiable.

The bias was added once every 8 ps as a bivariate Gaussian function of the above-defined variables d and w, $V(d, w) = h_0 e^{-\frac{V(d_t, w_t)}{k_B \Delta T}} e^{-(d - d_t)^2 / 2\sigma_d^2 - (w - w_t)^2 / 2\sigma_w^2}$.

Here, d_t and w_t are the values of the CVs at a certain time t, h_0 is the initial Gaussian amplitude, which we kept at 1.8 kJ/mol. ΔT, which we acquired to be $\Delta T = 14T$, is the so-called tempering parameter, which smoothly modulates the amplitude of the Gaussian each time a certain point in the CV space is revisited. The Gaussian widths σ_d and σ_w were 0.02 nm and 0.02 U, respectively. The free energy as function of these two variables, defined below, can be estimated from the bias as [within an additive constant (21)]

$$F(d, w) = -k_B T \log P(d, w) = -\frac{T + \Delta T}{\Delta T} V(d, w) \tag{4}$$

where $P(d, w)$ is the stationary probability density as a function of d and w.

We defined the unbound state as states with $d > 1.8$ nm, beyond which we find that the ligand diffuses either on the surface of the protein or in the bulk. All 12 independent simulations were run until they gave full unbinding. This took between 150 ns of unscaled MD time for the fastest exit event and as much as 750 ns for the slowest event. Through each of the simulations, we found no deterioration of the protein structure, as measured by the RMSD value of the heavy atoms, which stayed within 3.5 Å of the native state.

SUPPLEMENTARY MATERIALS

Reliability of time scales
Master equation
Effect of force field
fig. S1. Empirical (dashed line) and fitted cumulative distribution functions (solid line).
fig. S2. Residence time, eigenvector, and eigenvalue analysis.
fig. S3. FES as a function of ligand-binding pocket distance d (in nanometers) and hydration state of binding pocket w calculated using the second set of force fields.
table S1. Residence times and P values for various states
table S2. Matrix K of state-to-state transition rates.

REFERENCES AND NOTES

1. R. A. Copeland, The drug–target residence time model: A 10-year retrospective. *Nat. Rev. Drug Discov.* **15**, 87–95 (2016).
2. C. S. Tautermann, Impact, determination and prediction of drug–receptor residence times for GPCRs. *Curr. Opin. Pharm.* **30**, 22–26 (2016).
3. A. C. Pan, D. W. Borhani, R. O. Dror, D. E. Shaw, Molecular determinants of drug–receptor binding kinetics. *Drug Discov. Today* **18**, 667–673 (2013).
4. J. Mondal, P. Tiwary, B. J. Berne, How a kinase inhibitor withstands gatekeeper residue mutations. *J. Am. Chem. Soc.* **138**, 4608–4615 (2016).
5. S. M. Dehm, K. Bonham, *SRC* gene expression in human cancer: The role of transcriptional activation. *Biochem. Cell Biol.* **82**, 263–274 (2004).
6. P. Tiwary, V. Limongelli, M. Salvalaglio, M. Parrinello, Kinetics of protein–ligand unbinding: Predicting pathways, rates, and rate-limiting steps. *Proc. Natl. Acad. Sci. U.S.A.* **112**, E386–E391 (2015).
7. I. Teo, C. G. Mayne, K. Schulten, T. Lelièvre, Adaptive multilevel splitting method for molecular dynamics calculation of benzamidine-trypsin dissociation time. *J. Chem. Theory Comput.* **12**, 2983–2989 (2016).
8. H. Wu, F. Paul, C. Wehmeyer, F. Noé, Multiensemble Markov models of molecular thermodynamics and kinetics. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E3221–E3230 (2016).
9. H.-J. Li, C.-T. Lai, P. Pan, W. Yu, N. Liu, G. R. Bommineni, M. Garcia-Diaz, C. Simmerling, P. J. Tonge, A structural and energetic model for the slow-onset inhibition of the *Mycobacterium tuberculosis* enoyl-ACP reductase InhA. *ACS Chem. Biol.* **9**, 986–993 (2014).
10. Y. Wang, E. Papaleo, K. Lindorff-Larsen, Mapping transiently formed and sparsely populated conformations on a complex energy landscape. *eLife* **5**, e17505 (2016).
11. J. I. Stuckey, B. M. Dickson, N. Cheng, Y. Liu, J. L. Norris, S. H. Cholensky, W. Tempel, S. Qin, K. G. Huber, C. Sagum, K. Black, F. Li, X.-P. Huang, B. L. Roth, B. M. Baughman, G. Senisterra, S. G. Pattenden, M. Vedadi, P. J. Brown, M.T. Bedford, J. Min, C. H. Arrowsmith, L. I. James, S. V. Frye, A cellular chemical probe targeting the chromodomains of Polycomb repressive complex 1. *Nat. Chem. Biol.* **12**, 180–187 (2016).
12. P. Tiwary, J. Mondal, J. A. Morrone, B. J. Berne, Role of water and steric constraints in the kinetics of cavity–ligand unbinding. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 12015–12019 (2015).
13. P. Tiwary, B. J. Berne, How wet should be the reaction coordinate for ligand unbinding? *J. Chem. Phys.* **145**, 054113 (2016).
14. S. S. Taylor, A. P. Kornev, Protein kinases: Evolution of dynamic regulatory proteins. *Trends Biochem. Sci.* **36**, 65–77 (2011).
15. F. E. Kwarcinski, K. R. Brandvold, S. Phadke, O. M. Beleh, T. K. Johnson, J. L. Meagher, M. A. Seeliger, J. A. Stuckey, M. B. Soellner, Conformation-selective analogues of dasatinib reveal insight into kinase inhibitor binding and selectivity. *ACS Chem. Biol.* **11**, 1296–1304 (2016).
16. R. V. Agafonov, C. Wilson, R. Otten, V. Buosi, D. Kern, Energetic dissection of Gleevec's selectivity toward human tyrosine kinases. *Nat. Struct. Mol. Biol.* **21**, 848–853 (2014).
17. J. E. Ladbury, Just add water! The effect of water on the specificity of protein-ligand binding sites and its potential application to drug design. *Cell Chem. Biol.* **3**, 973–980 (1996).
18. P. Schmidtke, F. J. Luque, J. B. Murray, X. Barril, Shielded hydrogen bonds as structural determinants of binding kinetics: Application in drug design. *J. Am. Chem. Soc.* **133**, 18903–18910 (2011).
19. T. Young, R. Abel, B. Kim, B. J. Berne, R. A. Friesner, Motifs for molecular recognition exploiting hydrophobic enclosure in protein–ligand binding. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 808–813 (2007).
20. O. Valsson, P. Tiwary, M. Parrinello, Enhancing important fluctuations: Rare events and metadynamics from a conceptual viewpoint. *Annu. Rev. Phys. Chem.* **67**, 159–184 (2016).
21. P. Tiwary, M. Parrinello, A time-independent free energy estimator for metadynamics. *J. Phys. Chem. B* **119**, 736–742 (2014).
22. P. Tiwary, M. Parrinello, From metadynamics to dynamics. *Phys. Rev. Lett.* **111**, 230602 (2013).
23. M. Salvalaglio, P. Tiwary, M. Parrinello, Assessing the reliability of the dynamics reconstructed from metadynamics. *J. Chem. Theory Comput.* **10**, 1420–1425 (2014).
24. D. E. Hyre, L. M. Amon, J. E. Penzotti, I. Le Trong, R. E. Stenkamp, T. P. Lybrand, P. S. Stayton, Early mechanistic events in biotin dissociation from streptavidin. *Nat. Struct. Mol. Biol.* **9**, 582–585 (2002).
25. S. Freitag, V. Chu, J. E. Penzotti, L. A. Klumb, R. To, D. Hyre, I. L. Trong, T. P. Lybrand, R. E. Stenkamp, P. S. Stayton, A structural snapshot of an intermediate in the streptavidin–biotin dissociation pathway. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 8384–8389 (1999).
26. F. Schwesinger, R. Ros, T. Strunz, D. Anselmetti, H.-J. Güntherodt, A. Honegger, L. Jermutus, L. Tiefenauer, A. Plückthun, Unbinding forces of single antibody-antigen complexes correlate with their thermal dissociation rates. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 9972–9977 (2000).
27. C.-K. Lee, Y.-M. Wang, L.-S. Huang, S. Lin, Atomic force microscopy: Determination of unbinding force, off rate and energy barrier for protein–ligand interaction. *Micron* **38**, 446–461 (2007).
28. D. Huang, A. Caflisch, The free energy landscape of small molecule unbinding. *PLOS Comput. Biol.* **7**, e1002002 (2011).
29. M. Getlik, C. Grütter, J. R. Simard, S. Klüter, M. Rabiller, H. B. Rode, A. Robubi, D. Rauh, Hybrid compound design to overcome the gatekeeper T338M mutation in cSrc. *J. Med. Chem.* **52**, 3915–3926 (2009).
30. Y. Shan, E. T. Kim, M. P. Eastwood, R. O. Dror, M. A. Seeliger, D. E. Shaw, How does a drug molecule find its target binding site? *J. Am. Chem. Soc.* **133**, 9181–9183 (2011).
31. G. A. Kaminski, R. A. Friesner, J. Tirado-Rives, W. L. Jorgensen, Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *J. Phys. Chem. B* **105**, 6474–6487 (2001).
32. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926 (1983).
33. K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror, D. E. Shaw, Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* **78**, 1950–1958 (2010).

34. J. Mondal, R. A. Friesner, B. J. Berne, Role of desolvation in thermodynamics and kinetics of ligand binding to a kinase. *J. Chem. Theory Comput.* **10**, 5696–5705 (2014).

35. P. Tiwary, B. J. Berne, Spectral gap optimization of order parameters for sampling complex molecular systems. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 2839–2844 (2016).

36. P. Tiwary, B. J. Berne, https://arxiv.org/abs/1704.03912 (2017).

37. W. C. Swope, J. W. Pitera, F. Suits, M. Pitman, M. Eleftheriou, B. G. Fitch, R. S. Germain, A. Rayshubski, T. J. C. Ward, Y. Zhestkov, R. Zhou, Describing protein folding kinetics by molecular dynamics simulations. 2. Example applications to alanine dipeptide and a β-hairpin peptide. *J. Phys. Chem. B* **108**, 6582–6594 (2004).

38. D. Shukla, Y. Meng, B. Roux, V. S. Pande, Activation pathway of Src kinase reveals intermediate states as targets for drug design. *Nat. Commun.* **5**, 3397 (2014).

39. R. B. Best, J. Mittal, Free-energy landscape of the GB1 hairpin in all-atom explicit solvent simulations with different force fields: Similarities and differences. *Proteins* **79**, 1318–1328 (2011).

40. R. A. Copeland, Conformational adaptation in drug–target interactions and residence time. *Future Med. Chem.* **3**, 1491–1501 (2011).

41. K. T. Debiec, A. M. Gronenborn, L. T. Chong, Evaluating the strength of salt bridges: A comparison of current biomolecular force fields. *J. Phys. Chem. B* **118**, 6561–6569 (2014).

42. T. Schindler, T. Schindler, W. Bornmann, P. Pellicena, W. T. Miller, B. Clarkson, J. Kuriyan, Structural mechanism for STI-571 inhibition of abelson tyrosine kinase. *Science* **289**, 1938–1942 (2000).

43. M. Tong, M. A. Seeliger, Targeting conformational plasticity of protein kinases. *ACS Chem. Biol.* **10**, 190–200 (2014).

44. J. M. Gozgit, Y. Song, T. Baker, F. Gonvalvez, X. Zhu, S. G. Zech, M. T. Greenfield, T. Zhou, P. M. Taslimi, D. C. Dalgarno, V. M. Rivera, Comparative analysis of BCR-ABL and VEGFR2 inhibitory activities of ponatinib, axitinib, and PF-114. *Haematologica* **100**, 427 (2015).

45. B. J. Berne, M. Borkovec, J. E. Straub, Classical and modern methods in reaction rate theory. *J. Phys. Chem.* **92**, 3711–3725 (1988).

Citation: P. Tiwary, J. Mondal, B. J. Berne, How and when does an anticancer drug leave its binding site? *Sci. Adv.* **3**, e1700014 (2017).

Science Advances

How and when does an anticancer drug leave its binding site?

Pratyush Tiwary, Jagannath Mondal and B. J. Berne

ARTICLE TOOLS	http://advances.sciencemag.org/content/3/5/e1700014
SUPPLEMENTARY MATERIALS	http://advances.sciencemag.org/content/suppl/2017/05/26/3.5.e1700014.DC1
REFERENCES	This article cites 44 articles, 8 of which you can access for free http://advances.sciencemag.org/content/3/5/e1700014#BIBL
PERMISSIONS	http://www.sciencemag.org/help/reprints-and-permissions

Use of this article is subject to the Terms of Service